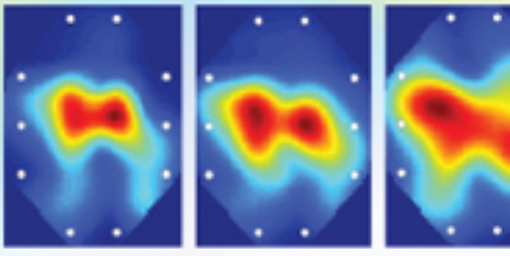
EDITOR C H CHEN



ULTRASONIC AND ADVANCED METHODS FOR NONDESTRUCTIVE TESTING AND MATERIAL CHARACTERIZATION

ULTRASONIC AND ADVANCED METHODS
FOR NONDESTRUCTIVE TESTING AND MATERIAL CHARACTERIZATION

This page intentionally left blank

EDITOR C H CHEN
University of Massachusetts Dartmouth, USA

ULTRASONIC AND ADVANCED METHODS FOR NONDESTRUCTIVE TESTING AND MATERIAL CHARACTERIZATION

World Scientific
NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

ULTRASONIC AND ADVANCED METHODS FOR NONDESTRUCTIVE TESTING AND MATERIAL CHARACTERIZATION

ISBN 978-981-270-409-2
ISBN 981-270-409-4

Printed in Singapore.

PREFACE

For over sixty years, the nondestructive testing (NDT) or nondestructive evaluation (NDE) of materials has been an area of continued growth. The need for NDT has increased dramatically in recent years for various reasons such as product safety, in-line diagnostics, quality control, health monitoring, and security testing, etc. Besides the practical demands the progress in NDT has a lot to do with its interdisciplinary nature. NDT is an area closely linked to aerospace engineering, civil engineering, electrical engineering, material science and engineering, mechanical engineering, nuclear engineering, petroleum engineering and physics among others. There are at least two dozen NDT methods in use. In fact any sensor that can examine the inside of material nondestructively is useful for NDT. However the ultrasonic methods are still most popular because of its capability, flexibility, and relative cost effectiveness. For this reason this book places a heavy emphasis on the ultrasonic methods which are covered in Part I of the book. Many other or advanced methods of NDT have also been well developed in recent years, such as the use of terahertz waveforms, the thermography method, etc. and they are covered in Part II of the book. The application of signal and image processing to NDT problems is another major development in the last twenty years and the subject is presented in Part III. The NDT area is so broad that the representative chapters as presented in the book can cover only certain aspects of the progress and development. However the book is intended to present the essential aspects of the NDT with the chapters authored by leading experts in the field.

We like to emphasize that the book provides a balance between the basics and the applications of NDT. Chapter 1 of Part I on “Fundamental models and measurements for ultrasonic nondestructive evaluation systems by Prof. Schmerr has the most clear and concise presentation of the mathematical principles of characterizing an ultrasonic nondestructive flaw measurement system, which are useful for both medical and industrial evaluations. Combining the ultrasonic beam models, flaw scattering models and measurements in a complete ultrasonic measurement model gives us the capability to make absolute predictions of the measured A-scan signals seen in ultrasonic flaw measurements. Chapter 2 on

"Integrated and flexible high temperature piezoelectric ultrasonic transducers" by Dr. Jen, et al. presents a technology to fabricate flexible transducers consisting of a thin substrate, a piezoelectric film and electrodes. The integrated transducer can operate at high temperature. The flexibility is realized owing to the porosity of piezoelectric film and the thinness of substrate and electrodes. Chapter 3 on "Real-time and in-line ultrasonic diagnostics of polymer process" by Dr. Jen, et al. shows ultrasonic diagnostic capability in polymer extrusion melt quality monitoring application. Also for injection molding diagnostics filling monitoring is illustrated. Laser-ultrasound is a major development in ultrasonic NDE in recent years. We are fortunate to have Dr. Jean-Pierre Monchalin to write Chapter 4 on "Laser ultrasonics: principles and industrial applications", which provides a broad overview of the field of laser-ultrasonics encompassing basics and applications. This chapter is a "must" reading for students and researchers in the laser ultrasound area. Another major new development in ultrasonic NDT is the methodology of guided waves as described in Chapter 5, "Guided wave based nondestructive testing: a reference-free pattern recognition approach", by Dr. Sohn, et al. Defects such as crack and/or corrosion can be detected without requiring pattern comparison with previously stored baseline data. The robustness of the proposed technique against undesirable variations in the system, such as temperature and external loading, can be attractive for long-term continuous monitoring.

A unique feature of this book is that it deals with both industrial and medical ultrasound. The latter is covered in the next four chapters. Both Chapter 6 and Chapter 7 by Dr. Zhang and Dr. Yan et al. respectively are focused on the tissue harmonic imaging based on the second harmonic ultrasound propagation. Chapter 8 on "Applications of high-frequency ultrasound to diagnostic and regenerative medicine" by Dr. Walters starts with a discussion of propagation and scattering properties of soft tissue and then presents a high frequency ultrasound system using 50 MHz transducer for the study of pulmonary arterial hypertension. This study can be important to the diagnosis and treatment of coronary arterial disease. An important advantage of the ultrasonic system is that it can be quantitative and noninvasive or minimally invasive. Chapter 9 is concerned with the receiver operating characteristics (ROC) analysis for computer aided detection. Both threshold specific performance measure such as sensitivity and specificity, and threshold independent performance measure based on the area under ROC curve are examined. The use of ROC curve is not limited to medical ultrasound problems however. Recently there has been greatly increased interest in ROC analysis for NDT problems in general. Finally the last

chapter of Part I, "Waveguides in acoustic sensor systems" by Larry Lynnworth is one of the few chapters of the book dealing with ultrasonic NDT devices.

The development of new sensors or emerging methods of nondestructive testing has contributed greatly to the advances of NDT. Part II deals with advanced methods other than ultrasound for NDT. It begins with Chapter 11 on "Electromagnetic NDE techniques for materials characterization" by Dr. Jayakumar et al. which provides an excellent overview of eddy current, magnetic, micro-magnetic, alternating current field measurement (ACFM), potential drop and microwave techniques for material characterization. Terahertz waveforms represent a new method for NDT. Chapter 12 "Terahwertz NDE for aerospace applications" by Bob Anastasi, et al. presents the pulsed terahertz technology which is an excellent example of emerging techniques that provide nondestructive testing of difficult-to-inspect material such as the Sprayed-on Foam insulation of the Space Shuttle External Tank. Terahertz NDE is useful also for metallic surface roughness evaluation and characterization of corrosion under paint and corrosion under shuttle tile. Chapter 13 is "Large area time domain terahertz (T-Ray) imaging non-destructive evaluation for security and aerospace" by Dr. Zimdars, et al. T-ray imaging employs single cycle sub-picosecond electromagnetic impulses with hyper wide band spectral content spanning the milimeter-wave to terahertz region of the spectrum. T-Ray imaging can offer up to 10 to 50 times greater resolution than microwave imaging at 40–90 GHz. There are other advantages offered by this new technology as presented in this chapter.

The next three chapters deal with the thermography technique in NDT. Chapter 14 on: "Active infrared thermography techniques for the nondestructive testing of materials", by Dr. Clemente Ibarra-Castanedo et al. provides an overview of three active thermography techniques including lock-in thermography, pulsed theomography and vibrothermography, which have various capability in speed and depth of penetration. The next chapter on "Quantitative evaluation of emerging infrared thermography techniques for aerospace applications" by Dr. DiMambro is emphasized on the sonic infrared probability of detection study using titanium specimens. Sonic Infrared Imaging is an emerging inspection technology that employs short pulse of ultrasound, typically less than a second, via a 20–40 kHz ultrasonic transducer. The increase in temperature in the vicinity of crack in aircraft components being tested can then be detected with an infrared camera. Further to the discussion of this subject is the next chapter on "Sonic infrared imaging: a novel NDE technology for detection of crack/delaminations/disbands in materials and structures" by Prof.

Han, which provides further physical insights and capability of the sonic infrared technology.

Chapter 17, "A data fusion framework for interpreting multiple NDT images" by Dr. Liu, et al. represents another major research in interpreting NDT data. As more than one sensor is likely to be used in NDT nowadays, the basic question is how to best integrate the information from different sensors. In this chapter an image fusion framework for interpreting NDI images based on Dempster-Shafer (DS) theory is presented. The use of DS theory to fuse the inspection results of aircraft lap joints with conventional eddy current and pulsed eddy current is demonstrated. The list of 30 references will help readers to gain more insight of the NDT data fusion approaches in general.

Although most NDT work has an objective of health monitoring or structural health monitoring, the next three chapters deal specifically with the issue of structural health monitoring. Chapter 18 on "SMART structures approach: achieving global structural health monitoring through the application of distributed, embedded micro-sensor systems", by Dr. Roach presents a clear introduction to the needs and approaches in structural health monitoring with special reference to the aging infrastructure. The application of distributed sensor systems can reduce the costs involved by allowing condition-based maintenance practice to be substituted for the current time-based maintenance approach. It may be possible in the near future to quickly, routinely, and remotely monitor the integrity of a structure in service. The next chapter, "Nondestructive testing for bearing condition monitoring and health diagnosis" by Prof. Gao, et al. is concerned with diagnosing the health condition of rolling bearings. Commonly used NDT techniques based on mechanical and acoustic sensing are reviewed. Vibration has been shown to be an informative and convenient variable for revealing bearing conditions. Bearing can be measured by using displacement, velocity, and acceleration sensors. The strength and weakness of the sensing techniques are examined. The authors also mentioned of the recent advancement of wireless sensor networks for NDT that allows for distributed sensing and improved decision making. This idea of integrating sensors is clearly shared by the following chapter, "Advanced sensors for health monitoring systems" by Dr. Mrad. This chapter presents non-mathematically the key sensor technologies which are integrated within an aircraft diagnostics, prognostics and health Management (DPHM) system making use of the fiber-optic based sensor networks.

Part III of the book is on the use of signal and image processing in the NDE problems. This is an area that has made progress for over twenty years and its importance is gaining attention gradually especially since the publication of my

edited book, "Signal Processing and Pattern Recognition in Nondestructive Evaluation of Materials", Springer-Verlag 1987. Clearly the area has many different approaches that may not be well correlated. Chapter 21, "Computational intelligence methodologies for non destructive testing/evaluation applications", by Dean Morabito, et al. describes how the advanced signal processing techniques, such as wavelet transform and independent component analysis, broadly listed as the computational intelligence methods, can be used in solving NDT problems such as feature extraction, denoising, identification of defects, etc. Perhaps the most influential signal processing development for NDT is the split spectrum processing (SSP). The following chapter, "Spit spectrum processing and automatic flaw detection" by Prof. Donohue has the most extensive but concise treatment of the subject from mathematical principles to the examples demonstrating the application of SSP in an automatic flaw detection scheme for multiple flaws embedded in nonstationary grain noise. The nonlinearities of the SSP algorithm effectively change the flaw and grain echo distribution to enhance the separation of their amplitudes beyond that of simple envelope detection. The ultrasonic nondestructive characterization of thin layered composite materials or structures can be difficult because the reflected signals are highly overlapped. The classical signal processing approach would have difficulty to separate the layers and determine the thickness and velocity of each layer. Chapter 22, "An application of the EM algorithm to thickness estimation of thin layered materials" by myself and X. Wu presents the use of expectation-maximization algorithm that is robust and effective to estimate layer thicknesses. It is noted that like in any signal processing problems, the quality of available data has a lot to do with the final signal processing performance. However signal processing allows us to extract information not easily available from the NDT measurements and thus essentially extend the resolution of the measurement beyond what is offered by the physical system. The actual signal processing effect may depend on the individual problem under consideration. The next chapter, "Homomorphic devonvolution of ultrasonic images" by Dr. Jirik et al. addresses the basic 1-D and 2-D deconvolution issues with examples of medical ultrasonic images. They are able to demonstrate that the images of kidney and aorta are sharper and more precise after homomorphic deconvolution.

Signal and image processing can help the human users to get a better visualization of the damage as clearly illustrated in Chapter 25 on "Frequency-wavenumber domain filtering for improved damage visualization" by Prof. Ruzene. The chapter presents a filtering technique aimed at improving the damage visualization capabilities of full wavefield measurements. The technique operates in the frequency/wavenumber domain, where the presence of reflections

associated to damage can be easily observed and separated from the main signal generated for inspection. Chapter 26 on "Recent advanced signal processing methods for waveform-based NDE methods" by Dr. Roth, et al., focuses on developing broad-based signal and image processing software for terahertz waveforms as used in the nondestructive evaluation of foam-based thermal protection system of NASA's Space Shuttle external tank. Some novel ideas of waveform feature extraction for image generation are presented. For example, the centroid of the power spectral density for the main (substrate) reflection off a metallic substrate under the foam shows promise versus conventional parameters of the signal used to form images. One of unique aspects of the software is the inclusion of a merit assessment algorithm to "grade" various signal processing methods. Finally in Chapter 27, "A Robust algorithm for nondestructive testing of weld seams", by Dr. Mery, et al. provides a detailed survey of available image segmentation algorithms suitable for NDT of weld seams. The steps of a proposed algorithm are presented and a quantitative performance measure of the algorithm is shown with the use of ROC curves. Both Chapters 26 and 27 show details of experimental results with the use of signal and image processing algorithms.

The nature of an edited book like this with a collection of articles from leading experts is not the same as journal or conference publications. NDT journals and conference proceedings may not have the details as presented in this volume. Topics of the book have been organized in such a manner that the readers can easily seek for additional and relevant information from a vast number of references cited. I hope the book can contribute in a small way toward the rapid progress of the ultrasonic and advanced (other) methods of nondestructive evaluation. Though the book is an outgrowth of the workshop of the same title held in June 2006 at our campus, and sponsored by our Ultrasonic NDE Lab., the detailed content is quite different. Readers may obtain a copy of the proceedings of the June workshop from the July 2006 issue of NDT e-Journal available from www.NDT.net.

I like to take this opportunity to thank the advisory board members of our Lab. for their valuable advices to the Lab. in the past few years. They are listed as follows.

Dr. Cheng-Kuei Jen, Principal Research Officer, Institute of Material Research, National Research Council, Boucherville, Quebec, Canada

Dr. Baldev Raj, Director, Indira Gandhi Centre for Atomic Research, Kalpakkam, India

Mr. Robert Anastasi, Manager, US Army Research Laboratory, Vehicle Technology Directorate, Hampton, VA, USA

Finally my special thanks go to all authors of the book whose work as presented in this volume, I firmly believe, can add significantly to the continued growth in NDT.

C. H. Chen
Ultrasonic NDE Lab.
University of Massachusetts Dartmouth
January 25, 2007

This page intentionally left blank

CONTENTS

Part 1

Ultrasonic Methods

This page intentionally left blank

CHAPTER 1

FUNDAMENTAL MODELS AND MEASUREMENTS FOR ULTRASONIC NONDESTRUCTIVE EVALUATION SYSTEMS

Lester W. Schmerr Jr.

Center for Nondestructive Evaluation and the Dept. of Aerospace Engineering
Iowa State University Ames, Iowa 50011
lschmerr@cnde.iastate.edu

It will be shown that it is possible to completely characterize an ultrasonic nondestructive flaw measurement system and simulate the measured flaw signals. This characterization includes the pulser/receiver, cabling, and transducers present as well as the propagating and scattered acoustic/elastic wave fields. All the elements of this comprehensive ultrasonic measurement model will be discussed as well as practical methods for obtaining those elements through a combination of models and measurements.

1. Introduction

Ultrasound is widely used for both medical evaluations and industrial inspections. In medical applications the ultrasound is typically propagating in tissue and the interactions of the sound with objects such as tumors are normally displayed in the form of ultrasonic images for evaluation by a physician. In industrial nondestructive evaluation (NDE) ultrasonic tests the ultrasonic waves are traveling in structural materials or components and the focus is often on evaluating those components for flaws such as cracks. While the results of the interaction of the sound with flaws can also be displayed as an ultrasonic image, a more common type of display used in industrial inspections is a received voltage versus time trace, $\tilde{V}_R(t)$, on an oscilloscope, called an *A-scan*, as shown in Fig. 1 for a typical immersion ultrasonic NDE flaw inspection system. Since images are generated by processing combinations of such A-scans, A-scan data provides the fundamental basis for describing the output of both medical and industrial ultrasonic systems.

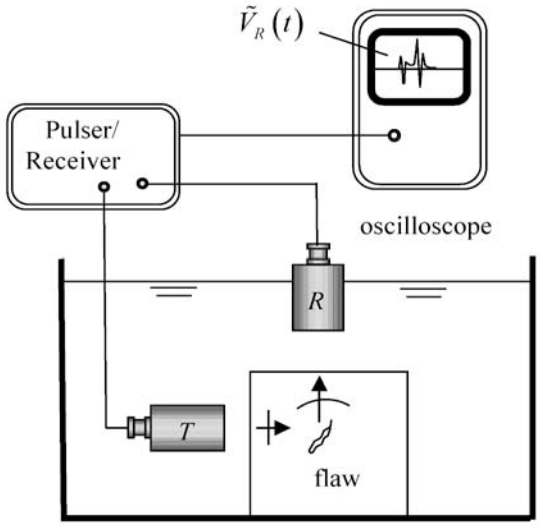


Figure 1. An ultrasonic pitch-catch flaw measurement system.

The basic components of the ultrasonic inspection system of Fig. 1 are the pulser/receiver, the cabling, the transducers, and the acoustic/elastic wave propagation and scattering processes present. The pulser section of the pulser/receiver generates short electrical pulses which travel through the cabling to the transmitting transducer. The transducer converts these electrical pulses into acoustic pulses at its acoustic output port. In the immersion setup of Fig. 1 these acoustic pulses first travel as a beam of sound in the fluid. This sound beam is transmitted into the solid component being inspected and the beam interacts with any flaws that are present. The flaw generates scattered wave pulses traveling in many directions, but some of these pulses reach the receiving transducer which then converts them into electrical pulses. These electrical pulses travel again through cabling to the receiver section of the pulser/receiver, where they are amplified and displayed as a received A-scan voltage, $\tilde{V}_R(t)$, as a function of the time *t*.

If we assume the components of the ultrasonic system shown in Fig. 1 can be represented as *linear time-shift invariant* (LTI) *systems* [1], then the output voltage pulse, $\tilde{V}_R(t)$, can be written as a convolution integral of the form

$$\tilde{V}_R(t) = \int_{-\infty}^{+\infty} \tilde{s}(\tau)\tilde{t}_A(t-\tau)\,d\tau, \tag{1.1}$$

where $\tilde{s}(t)$ is a function that characterizes all the electrical and electrical-to-acoustic and acoustic- to-electrical elements present (the pulser/receiver, cabling,

and transducers) and $\tilde{t}_A(t)$ is an function that contains the acoustic/elastic wave propagation and scattering processes occurring in the fluids and solids. We will find it convenient not to deal directly with these time dependent functions. Instead, we will model the ultrasonic system of Fig. 1 in the frequency domain by applying the Fourier transform to whatever pulses are present. In this chapter the Fourier transform and its inverse are defined as [2]:

$$\begin{aligned} g(f) &= \int_{-\infty}^{+\infty} \tilde{g}(t)\exp(2\pi i f t)\,dt \\ \tilde{g}(t) &= \int_{-\infty}^{+\infty} g(f)\exp(-2\pi i f t)\,df \end{aligned} \tag{1.2}$$

for a time-domain function, $\tilde{g}(t)$, and its frequency domain spectrum, $g(f)$, where t is time (usually measured in μsec) and f is frequency (usually measured in megaHertz (MHZ)). In Eq. (1.2) $i = \sqrt{-1}$. For an LTI system, Eq. (1.1) becomes in the frequency domain simply a product relationship, *i.e.* [1]:

$$V_R(f) = s(f)t_A(f), \tag{1.3}$$

where $V_R(f)$ are the frequency components(spectrum) of $V_R(t)$, $s(f)$ is the *system function*, and $t_A(f)$ is the *acoustic/elastic transfer function*.

Even though Eq. (1.3) is very simple in form, it also is extremely useful. We will show that it is possible to measure the system function directly in a reference calibration experiment and to model the contributions of the complex propagating and scattered waves present in the acoustic/elastic transfer function. This means that through a combination of models and measurements it is possible to synthesize the frequency components of the measured signals for many ultrasonic NDE setups. An inverse Fourier transform (done with a Fast Fourier Transform (FFT)) then yields a prediction of the measured A-scan signals. To make this approach quantitative we need to explicitly define the acoustic/elastic transfer function and the system function. The physical meaning of both of these functions will be described in detail in the following sections.

2. The Acoustic-Elastic Transfer Function

Dang *et al.* [3, 4] have used mechanical reciprocity principles to develop an expression for the acoustic/elastic transfer function, $t_A(f)$, for the immersion setup of Fig. 1. They found

$$t_A(f) = \frac{1}{Z_r^{T;a} v_T^{(1)} v_R^{(2)}} \int_{S_f} \left(\boldsymbol{\tau}^{(1)} \cdot \mathbf{v}^{(2)} - \boldsymbol{\tau}^{(2)} \cdot \mathbf{v}^{(1)} \right) dS. \tag{1.4}$$

In Eq. (1.4) two solutions are present. The first solution (labeled as state (1)) is when the transmitting transducer T is radiating into the fluid with an average velocity, $v_T^{(1)}(f)$, on its face. In this solution the waves arriving at the flaw from the transducer and those scattered from the flaw generate the stress vector, $\boldsymbol{\tau}^{(1)}$, and velocity, $\mathbf{v}^{(1)}$, on the surface of the flaw, S_f. The second solution (labeled as state (2)) is for the case when the receiving transducer R is acting as a transmitter with an average velocity, $v_R^{(2)}(f)$, on its face and the flaw is absent, so that $\boldsymbol{\tau}^{(2)}, \mathbf{v}^{(2)}$ are the stress vector and velocity, respectively, on S_f due the arriving waves only. The quantity $Z_r^{T;a}$ is the *acoustic radiation impedance* of the transmitting transducer, T. If we model the transmitting transducer as a piston probe where the output velocity is spatially uniform over the acoustic output port of the transducer, then at the high frequencies normally found in NDE tests it can be shown that $Z_r^{T;a} = \rho_1 c_{p1} S_T$, where ρ_1, c_{p1} are the density and wave speed of the fluid the transducer is radiating into and S_T is the area of the active face of the transducer [3]. Commercial ultrasonic NDE transducers can often be very successfully modeled as piston probes so that in this case the acoustic radiation impedance is a simple, known quantity. Equation (1.4) relies primarily on assumptions of linearity and reciprocity of the ultrasonic measurement system so that although it was explicitly developed for the immersion setup of Fig.1 it is also applicable to most ultrasonic bulk wave measurement systems, including contact cases and those using angle beam shear waves. Since the stress vector and velocity fields on the surface of the flaw for states (1) and (2) in Eq. (1.4) are divided by the average velocities acting on the faces of the transmitting transducers in those two states, we only need to know "normalized" stress vector and velocity fields at the flaw for states (1) and (2), *i.e.* when the transmitting transducers have unit average velocity on their faces. These normalized stress vectors and velocity vectors in states (1) and (2) on the surface of the flaw can, therefore, be obtained directly from models if one has a sufficiently general transducer beam model and flaw scattering model for the ultrasonic flaw measurement setup under consideration and one does not need to know the actual velocities generated on the transducer faces to obtain this transfer function.

If we place Eq. (1.4) into Eq. (1.3) and assume that the system function is known or can be measured then we have a general *ultrasonic measurement model* given by:

$$V_R(f) = \frac{s(f)}{Z_r^{T;a} v_T^{(1)} v_R^{(2)}} \int_{S_f} \left(\boldsymbol{\tau}^{(1)} \cdot \mathbf{v}^{(2)} - \boldsymbol{\tau}^{(2)} \cdot \mathbf{v}^{(1)} \right) dS. \tag{1.5}$$

A form similar to Eq. (1.5) was first derived by Bert Auld in 1979 [5] and has been used world-wide by many researchers in conjunction with analytical or numerical wave propagation and scattering models that can evaluate the $(\boldsymbol{\tau}, \mathbf{v})$ fields for states (1) and (2). Many of these applications of Eq. (1.5) can be found in past volumes of the Review of Progress in Quantitative Nondestructive Evaluation [6].

Now, assume that the incident waves in both these states in the vicinity of the flaw can be written in quasi-plane wave form, *i.e.*

$$\begin{aligned} \mathbf{v}^{(1);inc} &= v_T^{(1)}(f) \hat{V}^{(1)}(\mathbf{x}, f) \mathbf{d}^{(1)} \exp\left[ik_{\beta 2} \mathbf{e}^{(1)} \cdot \mathbf{x} \right] \\ \mathbf{v}^{(2)} &= v_R^{(2)}(f) \hat{V}^{(2)}(\mathbf{x}, f) \mathbf{d}^{(2)} \exp\left[ik_{\alpha 2} \mathbf{e}^{(2)} \cdot \mathbf{x} \right], \end{aligned} \tag{1.6}$$

where $\hat{V}^{(m)}$ (m = 1,2) are complex-valued "amplitudes" that represent the incident velocity fields normalized by the driving velocities $v_T^{(m)}$ (m = 1,2) and $\mathbf{d}^{(m)}$ are the polarization vectors for these incident velocity fields in both states (m = 1,2). The parameters $k_{\alpha 2} = 2\pi f / c_{\alpha 2}$ and $k_{\beta 2} = 2\pi f / c_{\beta 2}$ are the wave numbers in the solid surrounding the flaw for the incident waves in states (1) and (2), respectively, where α and β denote the incident bulk wave type (p or s) present and $(c_{\alpha 2}, c_{\beta 2})$ are their wave speeds. Note that in state (2) the incident velocity field is the entire field present since the flaw is absent while in state (1) this field is only part of the total field, so that this difference has been indicated explicitly in Eq. (1.6) by labeling the incident field in state (1) with the superscript "*inc*". If we place the expressions of Eq. (1.6) into Eq. (1.5) and also assume the amplitude terms in Eq. (1.6) do not vary significantly over the surface of the flaw, then Eq. (1.5) reduces to [1]:

$$V_R(f) = s(f) \hat{V}_0^{(1)}(f) \hat{V}_0^{(2)}(f) A(f) \left[\frac{4\pi \rho_2 c_{\alpha 2}}{-ik_{\alpha 2} Z_r^{T;a}} \right], \tag{1.7}$$

where $\hat{V}_0^{(m)}(f) \equiv \hat{V}^{(m)}(\mathbf{x}_0, f)$ (m = 1,2) are the beam "amplitudes" evaluated at some fixed point, $\mathbf{x}_0$, in the vicinity of the flaw (usually taken as at the flaw "center"), ρ_2 is the density of the solid medium surrounding the flaw, and $A(f)$ is a scalar component of the vector plane wave far-field scattering amplitude of the flaw, a term we will discuss in more detail in section 1.4. A form similar to Eq. (1.7) was first derived by Thompson and Gray in 1983 [7] and is referred to as the *Thompson-Gray measurement model*. Although Eq. (1.7)

is less general than Eq. (1.5), it explicitly identifies all the contributions to the measured voltage response in a modular form which makes it an extremely useful tool for analyzing ultrasonic NDE flaw inspection systems. The flaw response, for example, is contained entirely in $A(f)$. The effects of the instrumentation and transducers are contained in the system function, $s(f)$, and the ultrasonic beams incident on the flaw are given by the terms $\hat{V}_0^{(m)}(f)$. The remaining term in Eq. (1.7) is just a combination of simple, known parameters. This modularity of Eq. (1.7) has been exploited in numerous quantitative flaw studies, many of which can also be found in past volumes of the Review of Progress in Quantitative Nondestructive Evaluation [6].

In both Eq. (1.5) and Eq. (1.7) the acoustic/elastic transfer function is defined implicitly in terms of the fields present in a flaw measurement. It is also possible to obtain more explicit analytical expressions for the acoustic/elastic transfer function in some simple calibration setups. Such setups are used for determining the system function, $s(f)$, as well as finding transducer and other system parameters. Figure 2 shows two examples of setups where the acoustic/elastic transfer function is known. In Fig. 2(a), a circular, planar piston P-wave transducer of radius a generates and receives the waves reflected from a flat fluid-solid interface. For this pulse-echo immersion setup the acoustic/elastic transfer function, $t_A^{planar}(f)$ is [1]

$$t_A^{planar}(f) = R_{12}\exp\left(2ik_{p1}D\right)\exp\left[-2\alpha_{p1}(f)D\right]\tilde{D}_p\left(k_{p1}a^2/2D\right) \tag{1.8}$$

with

$$\tilde{D}_p(u) = 2\left[1-\exp(iu)\left\{J_0(u)-iJ_1(u)\right\}\right]. \tag{1.9}$$

Here R_{12} is the plane wave reflection coefficient for the interface, based on the ratio of the reflected pressure to that of the incident pressure given by

$$R_{12} = \frac{\rho_2 c_{p2} - \rho_1 c_{p1}}{\rho_2 c_{p2} + \rho_1 c_{p1}}, \tag{1.10}$$

where ρ_1, c_{p1} and ρ_2, c_{p2} and are the density and compressional wave speed of the fluid and solid, respectively, $k_{p1} = 2\pi f / c_{p1}$ is the wave number for waves traveling in the fluid, a is the radius of the transducer, and D is the distance from the transducer to the fluid-solid interface. The functions $J_0(u), J_1(u)$ are Bessel functions of the first kind of order zero and one, respectively, and $\alpha_{p1}(f)$ is the frequency dependent attenuation of P-waves traveling in the fluid. For degassed

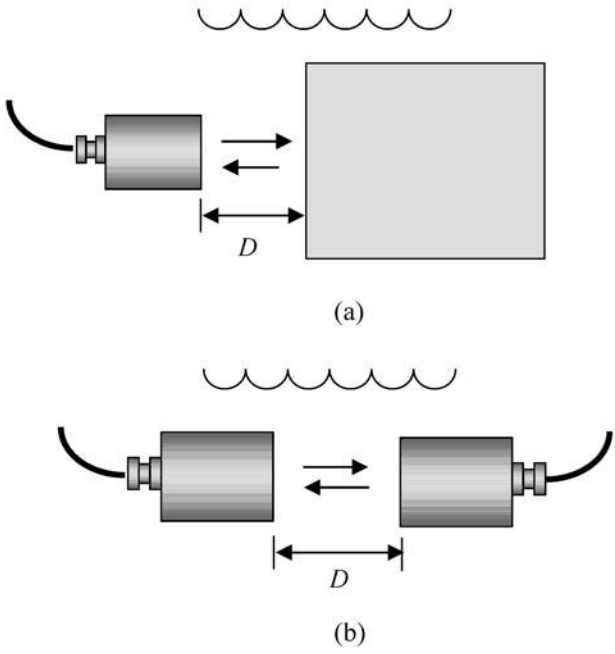


Figure 2. (a) A pulse-echo reference ultrasonic setup and (b) a pitch-catch ultrasonic reference setup.

water, this attenuation function has been experimentally determined [1]. In Eq. (1.8) the first three terms on the right hand side of that equation represent a plane wave that has been reflected from the interface and has returned to the transducer. Thus the $\tilde{D}_p$ function is a *diffraction correction* that accounts for the fact that the transducer generates a reflected beam of sound, not just a plane wave. Although Eq. (1.8) is for a planar transducer, it can also be modified to determine the acoustic/elastic transfer function, $t_A^{sph}(f)$, for a circular, spherically focused piston transducer with a geometrical focal length, R_0, in the setup of Fig. 2(a) where the distance $D = R_0$. In that case, one finds [8]

$$t_A^{sph}(f) = R_{12}\exp\left(2ik_{p1}R_0\right)\exp\left[-2\alpha_{p1}(f)R_0\right]\left[\tilde{D}_p\left(k_{p1}a^2/2R_0\right)\right],^* \quad (1.11)$$

where $[\ ]^*$ denotes the complex-conjugate. In a similar fashion, for a rectangular piston transducer in the pulse-echo immersion setup of Fig. 2(a) the acoustic/elastic transfer function, $t_A^{rect}(f)$, is [9]:

$$t_A^{rect}(f) = R_{12}\exp\left(2ik_{p1}D\right)\exp\left[-2\alpha_{p1}(f)D\right]\tilde{D}_p^{rect}\left(k_{p1}a^2/2D\right) \quad (1.12)$$

with

$$\tilde{D}_p^{rect} = \frac{4}{i}\left\{F\left(\sqrt{2k_{p1}a^2/\pi D}\right) + \frac{i}{\pi\sqrt{2k_{p1}a^2/\pi D}}\left[\exp\left(ik_{p1}a^2/D\right)-1\right]\right\} \cdot\left\{F\left(\sqrt{2k_{p1}b^2/\pi D}\right) + \frac{i}{\pi\sqrt{2k_{p1}b^2/\pi D}}\left[\exp\left(ik_{p1}b^2/D\right)-1\right]\right\}, \tag{1.13}$$

where $F(x)$ is the Fresnel integral

$$F(x) = \int_0^x \exp\left(i\pi t^2/2\right)dt \tag{1.14}$$

and the lengths of the two sides of the transducer are $(2a, 2b)$, respectively. Similarly, for a cylindrically focused rectangular piston transducer with geometrical focal length, *R*, along the side of the transducer whose length is 2*b*, and where the distance $D = R$ we can also easily modify Eq. (1.12) to obtain the acoustic/elastic transfer function, $t_A^{cyl}(f)$, for this cylindrically focused transducer as [9]

$$t_A^{cyl}(f) = R_{12}\exp\left(2ik_{p1}R\right)\exp\left[-2\alpha_{p1}(f)R\right] \cdot\frac{4}{i}\left\{F\left(\sqrt{2k_{p1}a^2/\pi R}\right) + \frac{i}{\pi\sqrt{2k_{p1}a^2/\pi R}}\left[\exp\left(ik_{p1}a^2/R\right)-1\right]\right\} \cdot\left\{F\left(\sqrt{2k_{p1}b^2/\pi R}\right) + \frac{i}{\pi\sqrt{2k_{p1}b^2/\pi R}}\left[\exp\left(ik_{p1}b^2/R\right)-1\right]\right\}^*. \tag{1.15}$$

All of the examples so far have been for pulse-echo immersion setups. For the pitch-catch setup of Fig. 2(b), the acoustic/elastic transfer function for a pair of circular, planar piston transducers, both having the same radius, *a*, is given simply by [1]

$$t_A(f) = \exp\left(ik_{p1}D\right)\exp\left[-\alpha_{p1}(f)D\right]\tilde{D}_p\left(k_{p1}a^2/D\right), \tag{1.16}$$

involving the same diffraction correction function, $\tilde{D}_p$, as before.

We have given these examples to show a few important cases where the acoustic/elastic transfer function can be obtained but there are many other calibration setups where this transfer function is also known [10]. For more complex calibration configurations (such as found in contact and angle beam transducer setups) where analytical expressions for this transfer function cannot

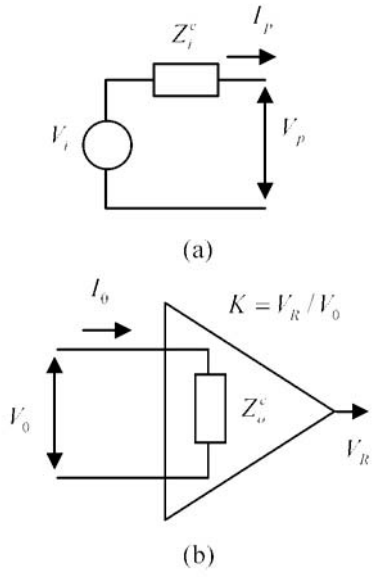


Figure 3. (a) A model of the pulser section of a pulser/receiver as a voltage source and an electrical impedance, and (b) a model of the receiver section as an electrical impedance and a gain factor.

be obtained, it is possible to numerically determine the acoustic/elastic transfer function by use of a transducer beam model [11].

3. The System Function

It is necessary to model the acoustic/elastic transfer function in some form since it involves the propagation and scattering of wave fields in fluids and solids, fields that are often not able to be measured directly. The system function, however, can be measured in reference setups (such as those just discussed) where the acoustic/elastic transfer function is known. For any such reference setup, if the acoustic elastic transfer function is $t_A^{ref}(f)$ and the frequency spectrum, $V_R^{ref}(f)$, of the received signal is measured, then by Eq. (1.3) we have

$$s(f) = \frac{V_R^{ref}(f)}{t_A^{ref}(f)}, \tag{1.17}$$

i.e. we can obtain the system function by deconvolution. In practice the division in Eq. (1.17) is carried out with the aid of a *Wiener filter* to desensitize the deconvolution process to noise. The equivalent Wiener filter form is [12]

$$s(f)=\frac{V_R^{ref}(f)\left[t_A^{ref}(f)\right]^*}{\left|t_A^{ref}(f)\right|^2+\varepsilon^2\max\left\{\left|t_A^{ref}(f)\right|^2\right\}}, \tag{1.18}$$

where ε is a constant that is chosen to characterize the noise level present in the measurement. As $\varepsilon \to 0$ Eq. (1.18) reduces to Eq. (1.17).

The system function characterizes all the electrical and electromagnetic components of the ultrasonic measurement system (pulser/receiver, cabling, transducers) so that it does not depend on the particular reference acoustic/elastic transfer function used to determine it. Schmerr *et al.* have demonstrated this fact for a number of different reference setups [10]. This means that the same system function can be used in conjunction with different acoustic/elastic transfer functions to determine the output voltage as long as the system settings (gain, *etc*.) are not changed and the same components (transducers, cabling, *etc*.) are present. For example, the system function found with the pitch-catch setup of Fig. 2(b) could also be used in Eq. (1.5) or Eq. (1.7) for the flaw measurement setup of Fig. 1.

It is also possible to determine the system function by characterizing all the elements that it contains. For example, the pulser part of the pulser/receiver can be modeled as a Thévenin equivalent voltage source, $V_i(f)$, and electrical impedance, $Z_i^e(f)$ as shown in Fig. 3(a) [3]. Similarly the receiver can be modeled as an electrical impedance, $Z_0^e(f)$, and a gain factor, $K(f)$, as shown in Fig. 3(b), where $K(f)=V_R(f)/V_0(f)$ [3]. All of these parameters can be determined experimentally through a series of voltage and current measurements so that both the pulser and receiver can be modeled explicitly in this fashion, assuming they act as linear devices.

The cabling used for both sound generation and reception can be modeled as linear two port systems as shown in Figs. 4(a), (b) [3]. On the sound generation side the voltage and current $\left(V_p(f), I_p(f)\right)$ at the output port of the pulser can be related to the voltage and current $\left(V_{in}(f), I_{in}(f)\right)$ at the electrical input port of the transmitting transducer through a 2x2 transfer matrix, $[\mathbf{T}]$, *i.e.*

$$\begin{Bmatrix} V_p \\ I_p \end{Bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{Bmatrix} V_{in} \\ I_{in} \end{Bmatrix}. \tag{1.19}$$

On the reception side, the voltage and current $\left(V_t(f), I_t(f)\right)$ at the electrical output port of the receiving transducer can be related to the voltage and current $\left(V_0(f), I_0(f)\right)$ at the receiver input by another 2x2 transfer matrix, $[\mathbf{R}]$, *i.e.*

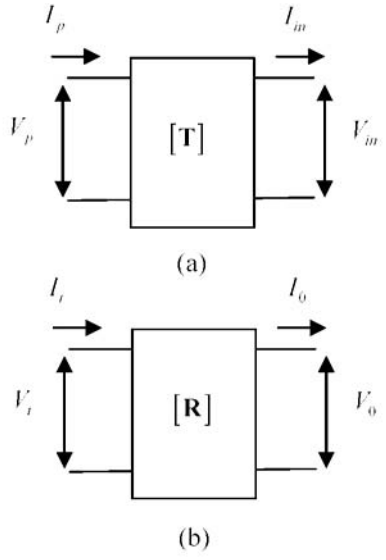


Figure 4. (a) A two port model of the cabling connecting the pulser and the transmitting transducer, and (b) a two port model of the cabling between the receiving transducer and the receiver.

$$\begin{Bmatrix} V_t \\ I_t \end{Bmatrix} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{Bmatrix} V_0 \\ I_0 \end{Bmatrix}. \tag{1.20}$$

At the MHz frequencies found in NDE tests, unless the cables are very short (typically much less than a meter), the cables do not act as a pure pass-through devices, leaving the signals unchanged. Thus, typically the components of these transfer matrices are frequency dependent. Since a cable is a simple passive electrical device, these transfer matrix components can be easily obtained by making a series of voltage and current measurements of the cabling under different electrical termination conditions [4]. If we assume the cabling to act as linear, reciprocal devices we must have det[**T**] = 1, det[**R**] = 1 which are conditions on the transfer matrix components that can be used to check the measurements. Note that in some setups the cabling may consist of flexible cables and cabling contained in transducer support fixtures. All of these components can be characterized together *in situ*, using Eqs. (1.19) and (1.20) to model their transfer matrix.

Ultrasonic transducers are more difficult devices to characterize in an ultrasonic setup. In principle they could also be characterized as linear two port systems that transform electrical signals (voltage and current) into acoustic signals (force and velocity) and vice-versa. However, to date there has not been a

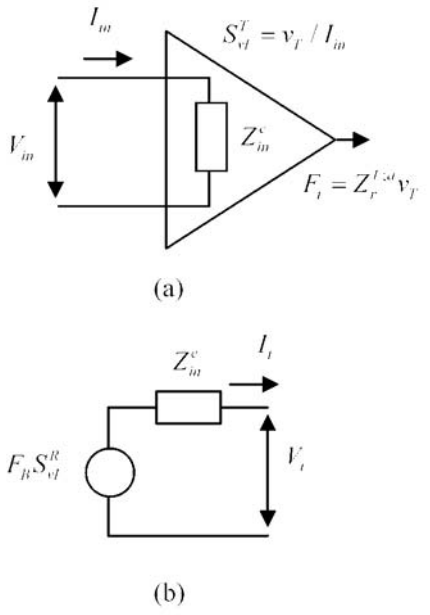


Figure 5. (a) A model of the transmitting transducer as an electrical impedance, $Z_{in}^{T;e}$, and a sensitivity, S_{vI}^{T}, and (b) the model of a receiving transducer and the acoustic sources that drive it as a voltage source, $F_B S_{vI}^{R}$, and an impedance, $Z_{in}^{R;e}$.

practical way to determine the components of the transducer transfer matrix in a fashion similar to what is done with the cabling. It is difficult, for example, to enforce different acoustic termination conditions at the transducer acoustic port and also to make the needed measurements of force and velocity. Fortunately, there is an alternate model one can use for the transmitting and receiving transducers that does not require knowledge of their transfer matrices, as shown in Figs. 5(a), (b). In Fig. 5(a) the transmitting transducer, *T*, is modeled as an electrical input impedance, $Z_{in}^{T;e}(f)$, and a sensitivity, $S_{vI}^{T}(f)$. The input impedance is just the ratio of the driving voltage and current at the transducer electrical input port, *i.e.* $Z_{in}^{T;e} = V_{in} / I_{in}$, and the sensitivity we will use here is defined as the ratio of the average output velocity, $v_T(f)$, at the acoustic output port of the transducer to the input current, *i.e.* $S_{vI}^{R} = v_T / I_{in}$. The output compressive force, $F_t(f)$, of the transducer is linearly related to the average output velocity through the acoustic radiation impedance of the transducer, $Z_r^{T;a}$, *i.e.* $F_t = Z_r^{T;a} v_T$. As discussed previously for piston transducers at high frequencies this acoustic radiation impedance is just the same value as that of a plane wave (although the transducer itself does not generate purely plane waves) so that if $v_T(f)$ is known the output force is also known. The reason why it is

useful to characterize the transducer in terms of its impedance and sensitivity is that both of these parameters can be obtained experimentally using purely electrical measurements. For example, consider a circular planar piston transducer in the pulse-echo configuration of Fig. 2(a). If one measures the frequency components of the voltage and current $\left(V_{in}(f), I_{in}(f)\right)$ at the electrical input port of the transducer when it is radiating waves into the fluid but before any waves have been reflected from the fluid-solid interface, then from the definition of the transducer's impedance we have [13]

$$Z_{in}^{T;e}(f) = \frac{V_{in}(f)}{I_{in}(f)}. \tag{1.21}$$

In the same fashion, if one measures the frequency components of the voltage and current $\left(V_{out}(f), I_{out}(f)\right)$ at the electrical output port of the transducer when it is receiving the beam of sound reflected from the fluid-solid interface, it can be shown that the sensitivity is given by [13]

$$S_{vI}^{T}(f) = \sqrt{\frac{V_{in}(f)I_{out}(f) + V_{out}(f)I_{in}(f)}{t_A(f)Z_r^{A;a}\left[I_{in}(f)\right]^2}}, \tag{1.22}$$

where $t_A(f)$ is the acoustic/elastic transfer function given by Eq. (1.8). Thus, with these two sets of electrical voltage and current measurements it is possible to obtain both the transducer's electrical impedance and its sensitivity. Knowledge of the transducer's electrical impedance defines its role as an electrical element in the generation of sound and knowledge of its sensitivity defines how the electrical inputs are transformed to acoustic outputs, so these two parameters (impedance, sensitivity) are all that is needed to fully characterize the role of the transducer as an electrical component and as a electrical-to-acoustic conversion device in the sound generation process.

When the ultrasonic transducer acts as a receiver, it converts the received acoustic waves into electrical signals. Remarkably, the same transducer impedance and sensitivity which we have defined for the case when the transducer is acting as a transmitter are also all that is needed to model the role of the transducer in the reception process. As illustrated in Fig. 5(b), it can be shown that a receiving transducer *R* and the acoustic waves that drive it can be modeled as a voltage source of strength $F_B S_{vI}^{R}$ in series with the electrical impedance, $Z_{in}^{R;e}$, where $F_B(f)$ is the *blocked force* acting on the receiving transducer by all the waves present [3]. By definition, this blocked force is just the compressive force exerted on the transducer face when that face is held rigid (motionless). This particular force appears naturally when modeling the sound

reception process and also appears in the definition of the acoustic/elastic transfer function as will be shown shortly.

Since we have shown that its is possible to characterize all the elements of both the sound generation and sound reception processes with simple models whose parameters can be obtained experimentally with a number of electrical measurements, one can combine all these models into explicit expressions for transfer functions that define the complete sound generation and reception processes. For example, if we define the sound generation transfer function, $t_G(f)$, as the ratio of the output compressive force, $F_t(f)$, on the face of the transmitting transducer to the Thévenin equivalent voltage source of the pulser, $V_i(f)$, it can be shown that [3] (see Fig. 6)

$$t_G(f) \equiv \frac{F_t(f)}{V_i(f)} = \frac{Z_r^{T;a} S_{vI}^T}{\left(Z_{in}^{T;e} T_{11} + T_{12}\right) + \left(Z_{in}^{T;e} T_{21} + T_{22}\right) Z_i^e}. \tag{1.23}$$

Similarly, if we define a sound reception transfer function, $t_R(f)$, as the ratio of the output voltage of the pulser, $V_R(f)$, to the blocked force, $F_B(f)$, we have (see Fig. 6) [3]

$$t_R(f) \equiv \frac{V_R(f)}{F_B(f)} = \frac{K Z_o^e S_{vI}^R}{\left(Z_{in}^{R;e} R_{11} + R_{12}\right) + \left(Z_{in}^{R;e} R_{21} + R_{22}\right) Z_o^e}. \tag{1.24}$$

The system function, $s(f)$, is then defined in terms of these transfer functions and the pulser source voltage as

$$\begin{aligned} s(f) &\equiv t_G(f) t_R(f) V_i(f) \\ &= \frac{Z_r^{T;a} S_{vI}^T}{\left(Z_{in}^{T;e} T_{11} + T_{12}\right) + \left(Z_{in}^{T;e} T_{21} + T_{22}\right) Z_i^e} \\ &\quad \cdot \frac{K Z_o^e S_{vI}^R}{\left(Z_{in}^{R;e} R_{11} + R_{12}\right) + \left(Z_{in}^{R;e} R_{21} + R_{22}\right) Z_o^e} V_i. \end{aligned} \tag{1.25}$$

By measuring all the elements contained in Eq. (1.25), it is possible to directly determine the system function. Of course, this result should agree with the direct measurement of the system function in a calibration setup, as discussed previously. Calculation of the system function in both of these ways has shown that they do indeed agree with each other [9, 14]. Equation (1.25) shows the very important role that the transducer sensitivity plays in the system function. In a pulse-echo setup where both the transmitting and receiving transducers are identical ($T = R$) the system function is proportional to the square of the

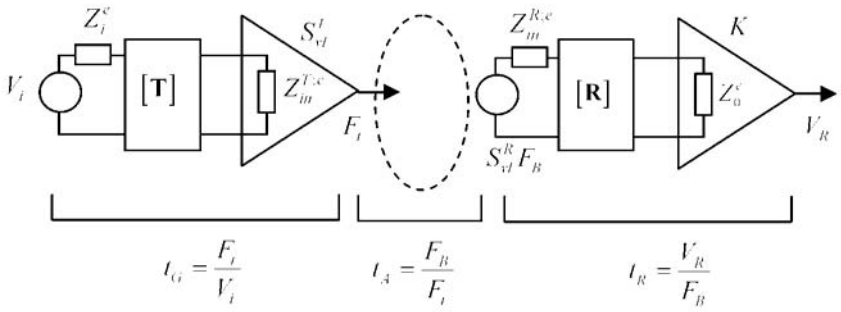


Figure 6. A complete model of an ultrasonic measurement system where the sound generation process can be described by a transfer function, t_G, the sound reception process by a transfer function, t_R, and all the acoustic/elastic propagation and scattering processes by the transfer function, t_A.

transducer sensitivity. Equation (1.25) also shows explicitly how the cabling properties and other electrical properties of the ultrasonic measurement system combine to generate the system function.

Having two methods to determine the system function gives us a set of very useful options. For example, in flaw measurement systems where we are primarily interested in the acoustic/elastic flaw signals, the direct determination of the system function by deconvolution (Eq. (1.18)) in conjunction with ultrasonic beam models and flaw scattering models gives us the ability to predict the measured flaw signals in an absolute manner. This approach allows us to efficiently study how the ultrasonic beam and flaw characteristics affect the measured response. In contrast, by synthesizing the system function directly from the measurement of its elements (Eq. (1.25)) one can study the effect that changes of system components and system settings have on the measured signals and optimize inspection setups, choice of transducers, *etc*. In the entire ultrasonic measurement system, the output voltage, $V_R(f)$, can be written as

$$\begin{aligned} V_R(f) &= \frac{V_R(f)}{F_B(f)}\frac{F_B(f)}{F_t(f)}\frac{F_t(f)}{V_i(f)}V_i(f) \\ &= t_R(f)\frac{F_B(f)}{F_t(f)}t_G(f)V_i(f) \\ &= s(f)\frac{F_B(f)}{F_t(f)}, \end{aligned} \tag{1.26}$$

which, when comparing with Eq. (1.3), shows that the acoustic/elastic transfer function, $t_A(f)$, is given by

$$t_A(f) = \frac{F_B(f)}{F_t(f)}, \tag{1.27}$$

i.e. this transfer function is just the ratio of the blocked compressive force acting on the face of the receiving transducer to the compressive output force generated on the face of the transmitting transducer (Fig. 6). This is the definition used in obtaining all the acoustic/elastic transfer functions discussed previously. If we assume that the interactions of the incident waves on the receiving transducer with its receiving face can be treated as plane wave interactions, then the blocked force is given by [3]

$$F_B(f) = 2F_{inc}(f), \tag{1.28}$$

where $F_{inc}(f)$ is the compressive force acting on the receiving transducer generated by only the incident waves, *i.e.* $F_{inc}(f)$ is the compressive force acting on the area of the transducer if the transducer is absent. Since this plane wave interaction assumption is likely a good approximation in many cases, Eq. (1.28) provides a convenient wave to obtain the acoustic/elastic transfer function. For example, in calculating the acoustic/elastic transfer function for the pulse-echo reference setup of Fig. 2(a) the normalized average pressure, $p_{ave}/\rho_1 c_{p1} v_0$, generated at the transducer face by the sound beam reflected from the fluid/solid interface was modeled. This pressure can be used to determine the acoustic elastic transfer function since from Eq. (1.28) we have

$$t_A(f) = \frac{2p_{ave}S_T}{F_t} = \frac{2p_{ave}S_T}{Z_r^{T;a}v_0} = \frac{2p_{ave}}{\rho_1 c_{p1} v_0}, \tag{1.29}$$

where S_T is the area of the face of the transducer T and v_0 is the constant velocity acting on its face , assuming the transducer acts as a piston source. The factor of two appearing in Eq. (1.29) also appears in the diffraction correction term for this transfer function (see Eq. (1.9)) as well in the other acoustic/elastic transfer function expressions given previously for different reference setups because of Eq. (1.28) and the use of the blocked force in the definition of the acoustic/elastic transfer function.

Figure 6 gives a complete model of an ultrasonic measurement system. While Eq. (1.23) and Eq. (1.24) define explicitly the sound generation and reception transfer function in this complete model, calculation of the acoustic/elastic transfer function in a flaw measurement setup requires knowledge of both the propagating waves present in the measurement system and the waves scattered

from the flaw. In the next two sections we will discuss ultrasonic beam models and flaw scattering models that can give these wave fields.

4. Transducer Beam Models

In this section we will concentrate on the role that transducer beam models play in determining the acoustic/elastic transfer function. In the Thompson-Gray measurement model (Eq. (1.7)) acoustic/elastic beam models are needed to calculate the normalized velocity fields $\hat{V}_0^{(m)}(\mathbf{x}_0, f)$ (m = 1, 2) incident on a flaw in states (1) and (2) at a fixed point, $\mathbf{x}_0$, that appear in that model. These same fields evaluated at an arbitrary point, $\mathbf{x}$, on the flaw surface, $\hat{V}^{(m)}(\mathbf{x}, f)$, are also implicitly present in the Auld form of the ultrasonic measurement model (Eq. (1.5)) since they are needed to determine the stress vectors and velocities present in that equation.

Modeling the velocity fields $\hat{V}^{(m)}(\mathbf{x}, f)$ is a difficult task since in many ultrasonic setups the sound beam generated by the transducer may interact with one or more interfaces and surfaces before it reaches a flaw. Here we will discuss beam models for the immersion setup shown in Figure 7 where the sound beam is generated by a planar piston transducer in a fluid and the sound must pass through a plane fluid-solid interface. This is a simple setup but one that will allow us to describe many of the main issues associated with beam modeling. In all the cases to be discussed we will treat the materials as perfect and neglect the attenuation of the ultrasound but attenuation can easily be introduced into the beam models as long as it is not too severe [1].

Most ultrasonic beam models express the transducer sound beam as a superposition of some simpler wave fields that act as set of fundamental basis functions. One choice of basis function is to express the waves generated by the transducer in the fluid by a superposition of plane waves and inhomogeneous waves traveling in different directions. This type of beam model is called an *angular plane wave spectrum beam model*. It is a convenient model to use when the fluid-solid interface is planar since it is possible to describe analytically how plane waves are transmitted or reflected at plane interfaces. For example, for the setup of Fig. 7 it can be shown that [1]

$$\hat{V}^{(m)}(\mathbf{x}, f) = \frac{f}{2\pi c_1} \iint_{S_T} \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{T_{12}(k_{1x}, k_{1y})}{k_{1z}} \cdot \exp\left[i\left(\mathbf{k}_2 \cdot (\mathbf{x} - \mathbf{x}_I) + \mathbf{k}_1 \cdot (\mathbf{x}_I - \mathbf{x}_T)\right)\right] \, dk_{1x} dk_{1y} \right] dS(\mathbf{x}_T), \tag{1.30}$$

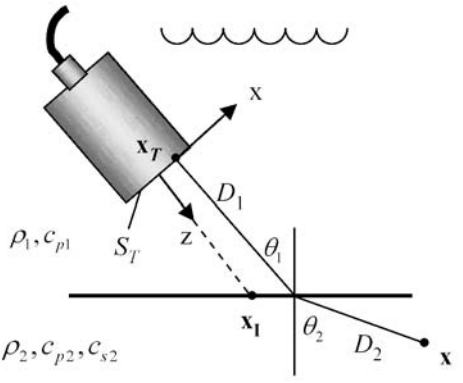


Figure 7. The radiation of a sound beam from a planar transducer through a plane fluid-solid interface.

where $(\mathbf{k}_1, \mathbf{k}_2)$ are the wave number vectors in the fluid and solid, respectively, T_{12} is the plane wave transmission coefficient for the fluid-solid interface (based on a velocity ratio). Point $\mathbf{x}$ is an arbitrary point in the solid, point $\mathbf{x}_T$ is an arbitrary point on the surface, S_T, of the transducer, and point $\mathbf{x}_I$ is the point on the interface where a ray from the origin O along the central axis of the transducer face intersects the interface. The quantities (k_{1x}, k_{1y}, k_{1z}) are the components of the wave number vector in the fluid along the *x*- ,*y*- and *z*-axes, respectively, (see Fig. 7 but note the *y*-axis is not shown explicitly in that figure) where k_{1z} is given by

$$k_{1z} = \begin{cases} \sqrt{k_1^2 - k_{1x}^2 - k_{1y}^2} & k_{1x}^2 + k_{1y}^2 < k_1^2 \\ i\sqrt{k_{1x}^2 + k_{1y}^2 - k_{p1}^2} & k_{1x}^2 + k_{1y}^2 \geq k_1^2 \end{cases}. \tag{1.31}$$

[Note that strictly speaking we should identify the modes associated with quantities such as wave speeds and wave numbers in these wave field expressions since in the solid we can describe sound beams for both P-waves and S-waves. However, for economy of notation we will only designate these and other variables in this section by the material in which the waves are traveling. Thus, (c_1, c_2) will represent the wave speeds of waves traveling in media one and two, respectively, and (k_1, k_2) their corresponding wave numbers for whatever mode we are considering.] If the surface integral in Eq. (1.30) is defined as

$$C(k_{1x}, k_{1y}) = \iint_{S_T} \exp[-i\mathbf{k}_1 \cdot \mathbf{x}_T] \, dS(\mathbf{x}_T) \tag{1.32}$$

then for circular, elliptical, and rectangular transducers this integral can be done explicitly, leaving only the two infinite k-space integrals to be done

numerically [1]. In most cases the integrand values for only real values of k_{1z} are retained in Eq. (1.30) which corresponds to neglecting the inhomogeneous waves and reduces Eq. (1.30) to two finite integrals. Even with this assumption one must superimpose a large number of plane wave components so this beam model is numerically intensive.

An alternative approach is to make a high frequency approximation and evaluate the *k*-space integrals in Eq. (1.30) by the method of stationary phase. In this case, Eq. (1.30) reduces to a *Rayleigh-Sommerfeld* type of integral [1]

$$\hat{V}^{(m)}(\mathbf{x}, f) = \frac{if}{c_{p1}} \iint_{S_T} T_{12} \frac{\exp\left[ik_1 D_1 + ik_2 D_2\right]}{\sqrt{D_1 + \frac{c_2}{c_1} D_2}\sqrt{D_1 + \frac{c_2 \cos^2 \theta_1}{c_1 \cos^2 \theta_2} D_2}} dS, \tag{1.33}$$

where the distances (D_1, D_2) and angles (θ_1, θ_2) are calculated from a point $\mathbf{x}_T$ on the face of the transducer to the point $\mathbf{x}$ in the solid along a stationary phase (ray) path that satisfies Snell's law and the plane wave transmission coefficient, T_{12}, is evaluated for this same path (see Fig. 7). The ordinary Rayleigh-Sommerfeld integral for a single fluid medium is a superposition of spherical waves arising from point sources points distributed over the face of the transducer [1]. Similarly, Eq. (1.33) represents the wave field in the solid as a superposition of waves from point sources on the transducer face but the elastic waves in the solid in this case have elliptical wave fronts due to the interaction of the spherical waves in the fluid with the plane interface at oblique incidence. Because one must take many source points on the transducer surface to evaluate Eq. (1.33) accurately, like the angular plane wave spectrum beam model, this beam model is not very computationally efficient.

One way to develop a fast beam model is to introduce the paraxial approximation. Loosely speaking, this approximation assumes that the ultrasonic wave field can be described as a well-collimated beam where there is a single predominant propagation direction. In the first (fluid) medium this direction is normal to the face of the transducer. In the *k*-space integrations of Eq. (1.30) the paraxial approximation corresponds to assuming that $k_{1x}, k_{1y} << k_1$. If one introduces this paraxial approximation into Eq. (1.30) and one also uses Stokes' theorem to replace the surface integral over S_T by an explicit term and a line integral around the edge, C_T, of the transducer, one obtains a *boundary diffraction wave beam model* of the transducer wave field in the quasi-plane wave form [1]

$$\hat{V}^{(m)}(\mathbf{x}, f) = T_{12}^0 \exp\left(ik_1 D_{10} + ik_2 D_{20}\right) \mathbb{C}(\mathbf{x}, f) \tag{1.34}$$

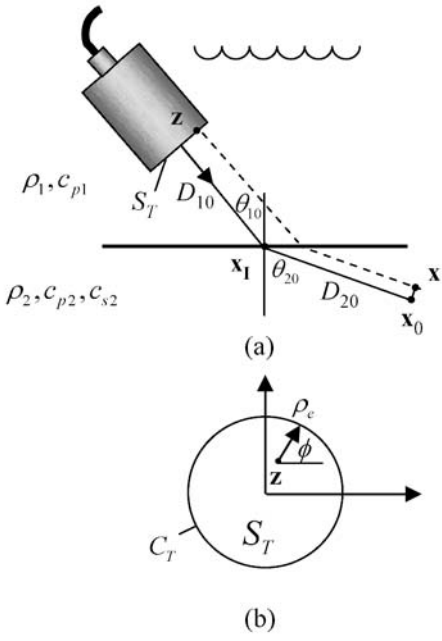


Figure 8. (a) Geometrical parameters that define the propagation of the waves from the transducer into the solid in the paraxial approximation, and (b) those parameters defined in the plane of the transducer.

where the diffraction correction, $\mathbb{C}(\mathbf{x}, f)$, is given by

$$\mathbb{C}(\mathbf{x}, f) = \frac{1}{\sqrt{\Delta_{x0}}\sqrt{\Delta_{yo}}} \oint_{C_T} \frac{\left\{ \Theta(\mathbf{x}) - \exp\left[\frac{ik_1 \rho_e^2(\mathbf{x}, \phi)}{2} \left(\frac{\cos^2\phi}{\Delta_{x0}} + \frac{\sin^2\phi}{\Delta_{y0}} \right) \right] \right\} d\phi}{\frac{\cos^2\phi}{\Delta_{x0}} + \frac{\sin^2\phi}{\Delta_{yo}}} \tag{1.35}$$

and where

$$\Theta(\mathbf{x}) = \begin{cases} 1 & \mathbf{z} \text{ inside } S_T \\ 1/2 & \mathbf{z} \text{ on } C_T \\ 0 & \mathbf{z} \text{ outside } S_T \end{cases} \tag{1.36}$$

with

$$\Delta_{x0} = D_{10} + \frac{c_2}{c_1} D_{20}$$
$$\Delta_{y0} = D_{10} + \frac{c_2 \cos^2 \theta_{10}}{c_1 \cos^2 \theta_{20}}. \tag{1.37}$$

The distance $\rho_e(\mathbf{x}, \phi)$ is shown in Fig. 8(b). This distance is measured from a point, **z**, on the plane of the face of the transducer to an arbitrary point on the edge of the transducer, where **z** is the point obtained by tracing a ray (the dashed line in Fig. 8(a)) from **x** to the plane of the transducer face along a ray that is parallel to the fixed ray that lies along the central axis of the transducer in the fluid and satisfies Snell's law (the solid line in Fig. 8(a)). The distances (D_{10}, D_{20}) and the plane wave transmission coefficient, T_{12}^0, are calculated along the central axis ray. This boundary diffraction wave beam model is typically several orders of magnitude faster to compute than either the angular plane wave spectrum model or the Rayleigh-Sommerfeld model so that it can be used effectively as a beam model whenever plane interfaces are present. For curved interfaces, however, all of these beam models have difficulty in dealing with cases where focal points or caustics are present in the solid [1].

A beam model that is both fast and can handle curved interfaces is the multi-Gaussian beam model [15]. In this model, the basis functions used are Gaussian beams which have an advantage over plane waves or spherical waves in that they remain well-behaved even after transmission through curved surfaces and the laws that define their propagation and reflection /refraction can be obtained analytically. Also, by superimposing only ten Gaussian beams one can accurately represent the wave fields of circular piston transducers that are either planar or focused so there is very little computation required for a Gaussian beam model. Gaussian beam models do rely on the paraxial approximation so they can degrade if the conditions of that approximation are not satisfied (as with tightly focused probes, near critical and grazing angles, *etc*.).

For the interface problem of Fig 7 the transducer wave field of a circular planar piston transducer is given by the multi-Gaussian beam model in the same quasi-plane form of Eq. (1.34), where the diffraction correction is expressed as [15]

$$\mathbb{C}(\mathbf{x}, f) = \sum_{r=1}^{10} \frac{\sqrt{\det\left[\mathbf{M}_2(D_{20})\right]_r}}{\sqrt{\det\left[\mathbf{M}_2(0)\right]_r}} \frac{\sqrt{\det\left[\mathbf{M}_1(D_{10})\right]_r}}{\sqrt{\det\left[\mathbf{M}_1(0)\right]_r}}$$
$$\cdot \left[V_1(0)\right]_r \exp\left[i \frac{k_1}{2} \mathbf{y}^T \left[c_1 \mathbf{M}_2(D_{20})\right]_r \mathbf{y} \right] \tag{1.38}$$

and the ten values of $\left[V_1(0)\right]_r$ and $\left[\mathbf{M}_1(0)\right]_r$ (r = 1, 2, ..10) can be related to a set of ten complex coefficients (A_r, B_r) given by

$$\left[V_1(0)\right]_r = A_r$$
$$\left[\mathbf{M}_1(0)\right]_r = \begin{bmatrix} \dfrac{iB_r}{c_1\tilde{D}_R} & 0 \\ 0 & \dfrac{iB_r}{c_1\tilde{D}_R} \end{bmatrix} \tag{1.39}$$

where the Rayleigh distance $\tilde{D}_R = k_1 a^2/2$ in terms of the radius, a, of the transducer and $\mathbf{y} = \mathbf{x} - \mathbf{x}_0 = (y_1, y_2)$ is a vector in a plane perpendicular to the central ray of the transducer in the solid medium from point $\mathbf{x}_0$ on that central ray to point $\mathbf{x}$ (see Fig. 8(a)). Wen and Breazeale [16] obtained the set of ten coefficients (A_r, B_r) appearing in Eq. (1.38). Those coefficients approximate the piston profile on the face of the transducer and generate through Eq. (1.38) a very accurate model of the transducer wave field except in the very near field. The propagation of the Gaussian beam in the fluid is given by the propagation law [15]

$$\left[\mathbf{M}_1(D_{20})\right]_r = \begin{bmatrix} \dfrac{\left[\left\{\mathbf{M}_1(0)\right\}_{11}\right]_r}{1 + D_{10}c_1\left[\left\{\mathbf{M}_1(0)\right\}_{11}\right]_r} & 0 \\ 0 & \dfrac{\left[\left\{\mathbf{M}_1(0)\right\}_{22}\right]_r}{1 + D_{10}c_1\left[\left\{\mathbf{M}_1(0)\right\}_{22}\right]_r} \end{bmatrix}. \tag{1.40}$$

If the transducer is radiating through a curved interface where a principal axis of the surface is aligned with the plane of incidence of the transducer (the plane that contains the central transducer axis and the normal to the interface) the transmission law across the interface is [15]

$$\left[\mathbf{M}_2(0)\right]_r = \begin{bmatrix} \dfrac{\left(\cos^2\theta_1 + Kh_{11}\right)/\cos^2\theta_2}{c_1\left(D_{10} - i\tilde{D}_R/B_r\right)} & 0 \\ 0 & \dfrac{1 + Kh_{22}}{c_1\left(D_{10} - i\tilde{D}_R/B_r\right)} \end{bmatrix}, \tag{1.41}$$

where

$$K = \left(D_{10} - i\tilde{D}_R / B_r \right) \left(\cos\theta_1 - \frac{c_1}{c_2} \cos\theta_2 \right) \tag{1.42}$$

and (h_{11}, h_{22}) are the principal curvatures in the plane of incidence and perpendicular to that plane, respectively.

Finally, the propagation law in the solid is [15]

$$\left[\mathbf{M}_2 \left(D_{20} \right) \right]_r = \begin{bmatrix} \dfrac{\left[\left\{ \mathbf{M}_2 (0) \right\}_{11} \right]_r}{1 + D_{20} c_2 \left[\left\{ \mathbf{M}_2 (0) \right\}_{11} \right]_r} & 0 \\ 0 & \dfrac{\left[\left\{ \mathbf{M}_2 (0) \right\}_{22} \right]_r}{1 + D_{20} c_2 \left[\left\{ \mathbf{M}_2 (0) \right\}_{22} \right]_r} \end{bmatrix}. \tag{1.43}$$

Thus, with the (A_r, B_r) coefficients and these propagation/transmission laws, all the terms in Eq. (1.38) are explicitly known.

In exactly the same fashion a multi-Gaussian beam model can be constructed for generally oriented curved interfaces and multiple solid media, including anisotropic elastic solids. Rectangular transducers and spherically focused and cylindrically focused transducers can also be easily simulated with a multi-Gaussian beam model. Thus, a multi-Gaussian beam model is often the beam model of choice to simulate ultrasonic transducer wave fields.

The plane waves, spherical waves, and Gaussian beams discussed above are the basis functions most commonly used to generate ultrasonic beam models. Other more direct numerical methods such as finite elements [17], boundary elements [18], finite differences [19], and the finite elastodynamic integration (EFIT) technique [20] can also be applied but in general those methods are computationally intensive which severely limits their use in conducting significant parametric studies.

5. Flaw Scattering Models

In the general ultrasonic model of Eq. (1.5), if we assume that the incident waves in states (1) and (2) can be described by the quasi-plane wave forms of Eq. (1.6) but if we do not assume that the variations of these incident waves over the surface of the flaw are negligible, that measurement model becomes

$$V_R(f) = s(f)\left[\frac{4\pi\rho_2 c_{\alpha 2}}{-ik_{\alpha 2} Z_r^{T;a}}\right]\int_{S_f} \hat{V}^{(1)}(\mathbf{x}, f)\hat{V}^{(2)}(\mathbf{x}, f)\mathcal{A}(\mathbf{x}, f)\exp\left[ik_{\alpha 2}\mathbf{e}^{(2)}\cdot\mathbf{x}\right]dS(\mathbf{x}) \tag{1.44}$$

with

$$\mathcal{A}(\mathbf{x}, f) = \frac{1}{4\pi\rho_2 c_{\alpha 2}^2}\left[\tilde{\tau}_{ji}^{(1)} d_i^{(2)} + C_{ijkl} d_k^{(2)}\left(e_l^{(2)}/c_{\alpha 2}\right)\tilde{v}_i^{(1)}\right]n_j, \tag{1.45}$$

where C_{ijkl} is the elastic constant tensor for the material surrounding the flaw and n_j are the components of the outward unit normal on the flaw surface. The quantities $\left(\tau_{ij}^{(1)}, v_i^{(1)}\right)$ are the stresses and velocity components on the surface of the flaw due to both the incident and scattered waves and the normalized quantities $\left(\tilde{\tau}_{ij}^{(1)}, \tilde{v}_i^{(1)}\right)$ appearing in Eq. (1.45) are defined as

$$\begin{aligned}\tilde{\tau}_{ij}^{(1)} &= \frac{-2\pi i f \tau_{ij}^{(1)}}{v_T^{(1)}\hat{V}^{(1)}} \\ \tilde{v}_j^{(1)} &= \frac{-2\pi i f v_j^{(1)}}{v_T^{(1)}\hat{V}^{(1)}}.\end{aligned} \tag{1.46}$$

Physically, these normalized fields are the stresses and velocity components due to an incident quasi-plane wave of unit amplitude incident on the flaw.

Writing the measurement model in this manner is revealing as, like the Thompson-Gray measurement model, we see the ultrasonic beam model terms appear explicitly in Eq. (1.44). The quantity $\mathcal{A}(\mathbf{x}, f)$ thus contains the flaw scattering response. However, unlike the Thompson-Gray measurement model we see that the beam model terms and flaw scattering response cannot be treated separately in Eq. (1.44) as they must be integrated together over the flaw surface. Comparing Eq. (1.7) and Eq. (1.44) we see that the far-field scattering amplitude component, $A(f)$, appearing in the Thompson-Gray measurement model is given by

$$A(f) = \int_S \mathcal{A}(\mathbf{x}, f)\exp\left(ik_{\alpha 2}\mathbf{e}^{(2)}\cdot\mathbf{x}\right)dS(\mathbf{x}). \tag{1.47}$$

In an elastic solid, the incident waves that strike a flaw can be either P- or S-waves. In the far-field of the flaw the scattered waves are both spherical P- and S-waves. For an incident plane wave of type β $(\beta = p, s)$ and scattered waves of type α $(\alpha = p, s)$ the scattered displacement vector, **u**, in the far-field is given by [1]

$$u(\mathbf{x},f)=\mathbf{A}^{p;\beta}U_0\frac{\exp(ik_{p2}r)}{r}+\mathbf{A}^{s;\beta}U_0\frac{\exp(ik_{s2}r)}{r}, \tag{1.48}$$

where U_0 is the displacement amplitude of the incident wave, (k_{p2},k_{s2}) are the wave numbers for scattered P-waves and S-waves, respectively, and r is the radial distance from a fixed point on the flaw, usually taken as the flaw "center" to point **x**. The quantities $\mathbf{A}^{\alpha;\beta}$ are the *far-field vector scattering amplitudes* of the flaw. These vector scattering amplitudes can be expressed as integrals of the stress and velocity fields over the surface of the flaw from which it is easy to show that for a given set of incident and scattered waves (i.e. for specified α and β wave types) we have [1]

$$A(f)=\int_S \mathcal{A}(\mathbf{x},f)\exp\left(ik_{\alpha 2}\mathbf{e}^{(2)}\cdot\mathbf{x}\right)dS(\mathbf{x})=\mathbf{A}^{\alpha;\beta}\cdot\left(-\mathbf{d}^{(2)}\right), \tag{1.49}$$

where recall $\mathbf{d}^{(2)}$ is the polarization of the wave incident on the flaw from the receiving transducer (acting as a transmitter) in state (2). It is only in the Thompson-Gray model that this specific component of the vector scattering amplitude appears explicitly, but even in the more general case, as Eq. (1.49) shows, the quantity $\mathcal{A}(\mathbf{x},f)$ is also closely related to this same component.

Obtaining either $\mathcal{A}(\mathbf{x},f)$ or $A(f)$ is a challenging task since real flaw morphologies are often not simple and to find the stresses and velocities on the flaw surface rigorously it is necessary in general to solve a complex boundary value problem. Numerical methods such as boundary elements or finite elements may often be required to obtain these wave fields but one can obtain some useful flaw scattering with approximations. Here we will describe some results that have been obtained with the *Kirchhoff approximation*.

The Kirchhoff approximation assumes that on that part of the flaw surface, called the lit surface, S_{lit}, where the incident wave (assumed to be planar) can strike the surface directly, the scattered fields are given by the reflected and transmitted plane waves calculated at a plane interface whose normal locally coincides locally with the normal to the surface of the flaw at every point on the lit surface. On the remaining part of the flaw surface it is assumed that the total fields are identically zero. Since the interactions of a plane wave with a plane interface is a problem that can be solved analytically, using the Kirchhoff approximation leads to an explicit expression for $\mathcal{A}(\mathbf{x},f)$ and hence an integral expression for $A(f)$ [1].Recently it has been shown that for the pulse-echo response of an arbitrary stress-free flaw in an elastic solid, where the scattered wave direction is directly opposite to the incident wave direction, $\mathbf{e}^{(1)}$, in state (1),

the Kirchhoff approximation for $\mathcal{A}(\mathbf{x},f)$ is identical to the same pulse-echo response for a void in a fluid [21], namely

$$\mathcal{A}(\mathbf{x},f)=\frac{-ik_{\beta 2}}{2\pi}\left(\mathbf{e}^{(1)}\cdot\mathbf{n}\right)\exp\left[ik_{\beta 2}\left(\mathbf{e}^{(1)}\cdot\mathbf{x}\right)\right], \tag{1.50}$$

where in pulse-echo we have set the incident wave and scattered wave modes equal $(\alpha=\beta)$ and $k_{\beta 2}$ $(\beta=p,s)$ are the wave numbers for the host material surrounding the flaw for P-waves or S-waves, respectively. Thus, the pulse-echo scattering amplitude component, $A(f)$, for an arbitrary stress-free flaw in the Kirchhoff approximation is given for either P-waves or S-waves as

$$A(f)=\frac{-ik_{\beta 2}}{2\pi}\iint_{S_{lit}}\left(\mathbf{e}^{(1)}\cdot\mathbf{n}\right)\exp\left[2ik_{\beta 2}\left(\mathbf{e}^{(1)}\cdot\mathbf{x}\right)\right]dS(\mathbf{x}). \tag{1.51}$$

There are a number of canonical stress-free geometries where Eq. (1.51) can be evaluated explicitly. For a spherical void (pore) of radius b we find [22]

$$A(f)=\frac{-b}{2}\exp\left(-ik_{\beta 2}b\right)\left[\exp\left(-ik_{\beta 2}b\right)-\frac{\sin\left(k_{\beta 2}b\right)}{k_{\beta 2}b}\right]. \tag{1.52}$$

For an elliptical flat crack with semi-major axes (b_1,b_2) along the unit vector directions $(\mathbf{u}_1,\mathbf{u}_2)$, respectively, [22]

$$A(f)=\frac{-ib_1b_2\left(\mathbf{e}^{(1)}\cdot\mathbf{n}\right)}{2r_e}J_1\left(2k_{\beta 2}r_e\right), \tag{1.53}$$

where

$$r_e=\sqrt{b_1^2\left(\mathbf{e}^{(1)}\cdot\mathbf{u}_1\right)^2+b_2^2\left(\mathbf{e}^{(1)}\cdot\mathbf{u}_2\right)^2}. \tag{1.54}$$

Finally, for a side-drilled hole of radius b and length L, which is a reference reflector commonly used in NDE, when the incident wave direction, $\mathbf{e}^{(1)}$, lies in a plane perpendicular to the axis of the hole we have [22]

$$A(f)=\frac{\left(k_{\beta 2}b\right)L}{2}\left[J_1\left(2k_{\beta 2}b\right)-iS_1\left(2k_{\beta 2}b\right)\right]+\frac{i\left(k_{\beta 2}b\right)L}{\pi}, \tag{1.55}$$

where S_1 is the Struve function. The beam variations over the length of a side-drilled hole in an ultrasonic measurement cannot normally be neglected, so that Eq. (1.55) cannot be used directly in the Thompson-Gray measurement model.

However, Schmerr and Sedov [23] have derived an equivalent measurement model that assumes that the beam variations are negligible only over the cross-section of the hole. In that case the general reciprocity-based measurement model reduces to

$$V_R(f) = s(f)\left[\int_L \hat{V}_0^{(1)}(z,f)\hat{V}_0^{(2)}(z,f)\,dz\right]\left[\frac{A(f)}{L}\right]\left[\frac{4\pi\rho_2 c_{\alpha 2}}{-ik_{\alpha 2}Z_r^{T;a}}\right], \qquad (1.56)$$

where $\hat{V}_0^{(m)}(z,f)$ are the normalized incident wave fields evaluated along the central z-axis of the side-drilled hole. For holes whose length exceeds the incident beam width the z-integrations in Eq. (1.56) are simply truncated at locations where the beam amplitude has become sufficiently small. Equation (1.55) can be placed into this measurement model to simulate the pulse-echo response of the side-drilled hole.

The Kirchhoff approximation can also be used to model the response of flaws with complex shapes. In those cases one must perform the integrations of the known fields over the lit surface numerically. However, if one evaluates the surface integral approximately at high frequencies with the method of stationary phase, an explicit expression can be derived for what is called the *leading edge response* of a volumetric (i.e. non crack-like) flaw. In a pulse-echo setup, for example, the leading edge response of a convex inclusion, A_{le}, for either P-waves or S-waves is given by [1]

$$A_{le}(f) = \frac{R_{12}\sqrt{R_1 R_2}}{2}\exp\left(-2ik_{\beta 2}d\right), \qquad (1.57)$$

where R_{12} is the plane wave reflection coefficient at normal incidence to a planar interface between the host and flaw, (R_1, R_2) are the principal radii of curvature of the flaw surface at the stationary phase point, and d is the distance from a fixed point (usually taken as the "center" of the flaw) to where the constant phase plane of a propagating plane wave traveling in the incident wave direction first touches the flaw surface. Similar results can also be obtained for a more general pitch-catch setup. For very complex shaped non-convex flaws there may be multiple stationary phase points and hence multiple leading edge response contributions that must be obtained.

Since the magnitude of the leading edge response is a constant, this contribution to the total scattered wave response is present at all frequencies and is relatively insensitive to the limited bandwidth present in ultrasonic systems. In many cases, therefore, this leading edge response dominates the scattering response of even complex-shaped volumetric flaws.

We have given only a few examples of flaw scattering responses in this section, concentrating on cases where significant analytical results are possible with the Kirchhoff approximation. Other approximate methods such as the Born approximation [24], the geometrical theory of diffraction [25], and low-frequency approximations [26] also can yield useful results without resorting to brute force numerical methods.

6. Summary

As shown in this chapter we can model and/or measure all the elements of a typical bulk wave ultrasonic NDE system. The system function can be obtained by making a single output voltage measurement in a reference setup or by making a detailed set of electrical measurements of the electrical and electromechanical elements of the measurement system. The acoustic/elastic transfer function can be found with the use of ultrasonic beam models and flaw scattering models. Combining these models and measurements in a complete ultrasonic measurement model gives us the capability to make absolute predictions of the measured A-scan (voltage versus time) signals seen in ultrasonic flaw measurements. These developments make it practical to design, engineer, and optimize ultrasonic NDE systems at many levels. This model-based approach also provides the basis for making quantitative flaw evaluations, which in many cases is the end goal of ultrasonic NDE inspection efforts. Significant uses have already been made of this comprehensive ultrasonic measurement model and many more applications are possible in the future. By necessity, the discussion in this chapter has only outlined the main features of the underlying models and measurements. Much more detail can be found in [9].

References

1. L. W. Schmerr, *Fundamentals of Ultrasonic Nondestructive Evaluation – A Modeling Approach*, (Plenum, New York, N.Y., 1998).
2. J. S. Walker, *Fast Fourier Transforms, 2nd Ed.*, (CRC Press, New York, NY, 1996).
3. C. J. Dang, L. W. Schmerr and A. Sedov, *Research in Nondestructive Evaluation*, **14**, 141 (2002).
4. C. J. Dang, L. W. Schmerr and A. Sedov, *Research in Nondestructive Evaluation*, **14**, 177 (2002).
5. B. A. Auld, *Wave Motion,* **1**, 3 9 (1979).
6. D. O. Thompson and D. E. Chimenti, Eds., *Review of Progress in Quantitative Nondestructive Evaluation*, (Vols. 1-18, Plenum Press, N.Y. 1981-1999), (Vols. 19 – 24, (American Institute of Physics, Melville, N.Y., 2000-2006).
7. R. B. Thompson and T. A. Gray, *J. Acoust. Soc. Am.*, **74**, 140 (1983).
8. X. Chen and K. Q. Schwarz, *J. Acoust. Soc. Am.*, **95**, 3049 (1994).

9. L. W. Schmerr and S. J. Song, *Ultrasonic Nondestructive Measurement Systems – Models and Measurements*, (Springer, Cambridge, MA, 2007)
10. L. W. Schmerr, S. J. Song and H. Zhang, in *Nondestructive Characterization of Materials, VI*, Eds. R.E. Green Jr., K.J. Kozaczek and C. O. Ruud, (Plenum Press, New York, NY, 1994, p. 111.
11. H. J. Kim, J. S. Park, S. J. Song and L. W. Schmerr, *J. Nondestr. Eval.,* **23**, 81 (2004).
12. A. Papoulis, *Signal Analysis*, (McGraw-Hill, New York, NY., 1977).
13. A. Lopez-Sanchez and L. W. Schmerr, *IEEE Trans. Ultrasonics, Ferroelectrics and Frequency Control*, **53**, 2101 (2006)
14. A. L. Lopez-Sanchez and L. W. Schmerr, *Res. in Nondestructive Evaluation*, (2006, to appear).
15. R. J. Huang, L. W. Schmerr and A. Sedov, *Res. in Nondestructive Evaluation*, **16**, 143 (2005).
16. J. J. Wen and M. A. Breazeale, *J. Acoust. Soc. Am.,* **83**, 1752 (1988).
17. W. Lord, R. L. Ludwig and Z. You, *J. Nondestr. Eval*., **9**, 155 (1990).
18. Q. C. Guo and J. D. Achenbach, *Ultrasonics*, **33**, 449 (1995).
19. H. Yim and E. Baek, "Two-dimensional numerical modeling and simulation of ultrasonic testing," *J. Korean Soc. N.D.T*, **22**, 649 (2002).
20. P. Fellinger, R. Marklein, K. J. Langenberg and S. Klaholz, *Wave Motion*, **21**, 47 (1995).
21. L. W. Schmerr and A. Sedov, in *Review of Progress in Quantitative Nondestructive Evaluation*, Eds. D.O. Thompson and D.E. Chimenti, (American Institute of Physics, Melville, NY, 22B, 2003) p. 1776.
22. A. Lopez-Sanchez, H. J. Kim, L. W. Schmerr and T. Gray, *Res. in Nondestructive Evaluation,* **17**, 49 (2006)
23. A. L. Lopez, H. J. Kim, L. W. Schmerr and A. Sedov, J. *Nondestr. Eval.,* **24**, 83 (2005).
24. J. E. Gubernatis, E. Domany, J. A. Krumhansl and M. Hubermann, *J. Appl. Phys.,* **48**, 2812 (1977).
25. J. D. Achenbach, A. K. Gautesen and H. McMaken, *Ray Methods for Waves in Elastic Solids*, (Pitman Books Ltd., Boston, 1982).
26. J. H. Rose, *Ultrasonics*, **25**, 141 (1987).

This page intentionally left blank

CHAPTER 2

INTEGRATED AND FLEXIBLE HIGH TEMPERATURE PIEZOELECTRIC ULTRASONIC TRANSDUCERS

Cheng-Kuei Jen and Makiko Kobayashi

Industrial Materials Institute, National Research Council Canada
75 de Mortagne Blvd., Boucherville, QC, J4B 6Y4, Canada
cheng-kuei.jen@crnc-nrc.gc.ca

Thick (> 40 μm) ceramic films as piezoelectric and ultrasonic transducers (UTs) have been successfully deposited on metallic and non-metallic substrates by a spray technique. In the film fabrication a composite consisting of piezoelectric powders well mixed with solution of high dielectric constant is directly sprayed onto the substrate. It is then dried, fired or annealed by heat. Multiple coating is used to achieve preferred thicknesses. A corona poling is utilized to achieve the piezoelectricity of the film. The top electrode is accomplished by a painting method. All fabrication processes may be carried out on-site and by handheld devices. Integrated ultrasonic longitudinal, shear, surface and plate wave transducers have been made. The same technology has been also used to fabricate flexible transducers consisting of a thin substrate, a piezoelectric ceramic film and electrodes. The flexibility is realized owing to the porosity of piezoelectric film and the thinness of the piezoelectric film, substrate and electrodes. All transducers have been tested at least up to 150°C. Their applications for non-destructive testing of different materials are demonstrated.

1. Introduction

Non-destructive testing (NDT) of materials are commonly performed to identify, characterize, assess voids, defects and damage in metals, metal alloys, composites and other materials [1,2]. Furthermore, the increasing demand to improve the performance, reduce downtime, increase reliability and extend the life of transportation vehicles, structures and engineering systems, requires the use of systems that have integrated capabilities with built-in sensors that perceive and process in-service information and take actions to accomplish desired

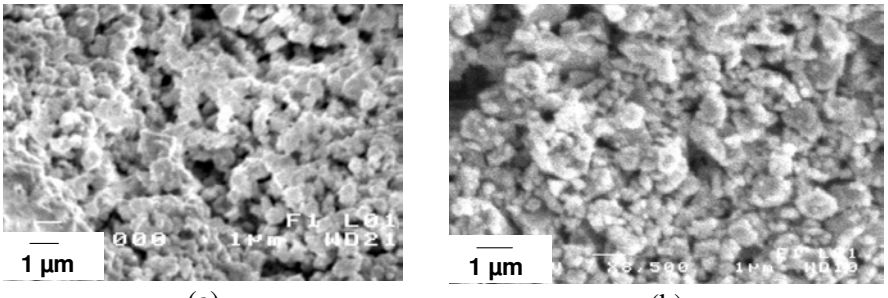
operations and tasks [3-6]. It is established that ultrasonic methods are widely used for real-time, in-situ or off-line NDT and evaluation of large metallic and polymeric composite structures including airplanes, ships, automobiles, pressure vessels and pipelines [5,6]. Because of their subsurface inspection capability, elastic property characterization ability, fast inspection speed, simplicity and cost-effectiveness, there has been considerable interest in ultrasonic NDT [5-15]. In many applications, ultrasonic inspections may need to be applied to curved surfaces or complex geometries [1,2,5,6]. It is known that conventional planar ultrasonic transducers (UTs) show poor inspection performance on curved surfaces. Also applications may often be subjected to high temperature (HT) environments [7-15]. Common limitations of current piezoelectric UTs are (a) lack of suitability for use on curved surfaces and (b) the difficulty for use at temperatures higher than 100°C. Therefore the objective of this investigation is to develop integrated longitudinal (L), shear (S), Rayleigh surface (RAW) and Lamb plate acoustic wave (PAW) transducers and flexible HTUTs operated at least up to 150°C. NDT of different materials employing these HTUTs will be demonstrated.

2. Fabrication and Characterization

In this study thick (> 40 μm) piezoelectric ceramic films as UTs are of interest. These films can be made by the technologies of jet printing [16,17], screen printing [18,19], tape casting [20,21], dipping [22,23], hydrothermal method [24,25], etc. Here an alternative sol-gel spray technique is used [14,15,26]. The ball-milled sub-micron piezoelectric lead-zirconate-titanate (PZT) or bismuth titanate (BIT) powders were dispersed into PZT sol-gel solution. The PZT and BIT powders were chosen because of their high piezoelectric constant and high Curie temperature (675°C), respectively. The final PZT/PZT or BIT/PZT mixture (paint) was then sprayed directly onto selected metallic substrates, such as stainless steel (SS) and aluminum (Al) alloys through an airbrush. With this sol-gel spray technique, the films can be produced at desired locations using a paper shadow mask. After spraying the coating, thermal treatments such as drying, firing and/or annealing were normally carried out using a heat gun. In special cases in this study a furnace would be used and mentioned specifically. Multiple coatings were made in order to reach desired film thicknesses. In this study the film thickness is between 40 and 200 μm. Piezoelectric films were then electrically poled using a corona discharging technique. The corona poling method was chosen because it could pole the piezoelectric film over a large area

with complex geometries. Finally, silver paste, platinum paste or silver paint spray method was used to form the top electrodes at room temperature.

The measured relative dielectric constant of the PZT/PZT film and BIT/PZT film was about 320 and 80, respectively. The d_{33} measured by an optical interferometer is 30 (10^{-12} m/V) for PZT/PZT and 10 (10^{-12} m/V) for BIT/PZT. The thickness mode electromechanical coupling constant measured was 0.2 for PZT/PZT and that for BIT/PZT was weaker. Figures 1(a) and 1(b) show the SEM images of the PZT/PZT and BIT/PZT film, respectively. They indicate that the grain size is less than 1 μm and the film is not dense. It is suspected that the porosity contributes to the low values of the dielectric constant, d_{33} and the thickness mode electromechanical coupling constant.



(a) (b)

Figure 1. SEM images of (a) PZT/PZT film and (b) BIT/PZT film.

3. Integrated Longitudinal and Shear Wave Transducers

It is known that the broadband MHz frequency L UT is suitable for defect or void detection. Here we would like to demonstrate that integrated thick film UTs at 440°C could perform these defects detection. Figure 2(a) shows a 200 μm thick BIT/PZT film was deposited onto a 12.7 mm thick steel plate. In this case the heating was performed using a furnace instead of a heat gun. The ultrasonic signals reflected from a bottom drilled flat hole of 1.0 mm diameter are shown in Fig. 2(b). The four traces from top to bottom show that the length of this hole measured from the bottom of the plate are 0, 1 mm, 2 mm and 3 mm, respectively. The large signal at the end of the traces in Fig. 2(b) was the reflected echo from the bottom of the steel substrate. Figure 2(b) demonstrates that ultrasonic monitoring of the extension of the defect length can be performed at 440°C, and the time delay of the signal reflected from the defect can be used to know the position of the tip of this bottom drilled flat hole.

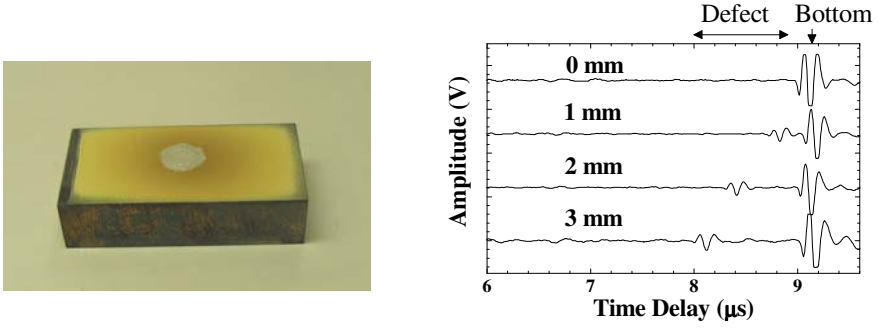


Figure 2. (a) A 200 μm thick BIT/PZT film deposited onto a planar steel substrate and (b) the ultrasonic monitoring of the extension of the artificial vertical defect at 440°C.

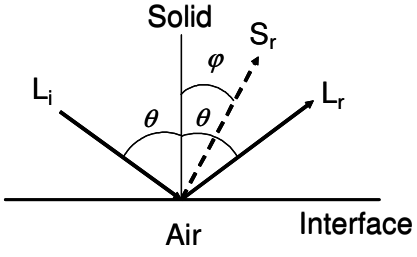


Figure 3. Reflection and mode conversion with an incident longitudinal wave at a solid-air interface.

Various efforts have been devoted to the development of piezoelectric HTUTs of large bandwidth and high efficiency [7-15] and they may be supplied by several companies. However, it is understood that S waves may be advantageous over L waves for NDT and characterization of materials because liquid and gas medium do not support S waves. In addition, for the evaluation of material properties, sometimes it is important to measure shear modulus and viscoelastic properties in which S wave properties are a requisite. Furthermore, a HTUT setup to generate and receive both L and S waves at the same sensor location would be also of interest.

The mode conversion from L to S wave due to reflection at a solid-air interface was reported [27,28]. It means that the L wave UT together with L-S mode conversion caused by the reflection at a solid-air interface can be effectively used as a S wave probe as shown in Fig. 3. In Fig. 3, L_i waves generated by an L wave UT reach a solid-air interface and reflected as L_r and S_r waves. The equations governing the reflection and mode conversion with respect

to the L wave incident angle θ can be given in Eqs.(1)-(3) [29], where V_l and V_s are L and S wave velocities in the solid, respectively, and R_{ll} and R_{sl} are energy reflection coefficients of the L and S waves, respectively.

$$\frac{V_l}{\sin\theta} = \frac{V_s}{\sin\varphi} \tag{1}$$

$$R_{ll} = \left[\frac{\cos^2 2\varphi - \left(V_s/V_l\right)^2 \cdot \sin 2\theta \cdot \sin 2\varphi}{\cos^2 2\varphi + \left(V_s/V_l\right)^2 \cdot \sin 2\varphi \cdot \sin 2\theta} \right]^2 \tag{2}$$

$$R_{sl} = \frac{4\left(V_s/V_l\right)^2 \cdot \cos^2 2\varphi \cdot \sin 2\theta \cdot \sin 2\varphi}{\left[\cos^2 2\varphi + \left(V_s/V_l\right)^2 \cdot \sin 2\varphi \cdot \sin 2\theta\right]^2} \tag{3}$$

In this study, a mild steel with the L wave velocity V_l = 5900 m/s and S wave velocity V_s = 3200 m/s at room temperature was used as the substrate. Figure 4 shows the calculated energy reflection coefficient based on Eqs. (2) and (3) for the mild steel substrate. It indicates that the maximum energy conversion rate from the L_i wave to the S_r wave is 97.5% at θ = 67.2°, and the reduction of the energy conversion rate is within 1% in the θ range between 60.8° and 72.9°.

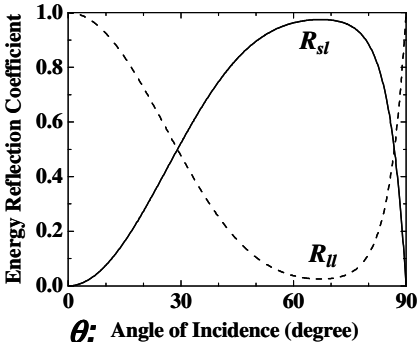
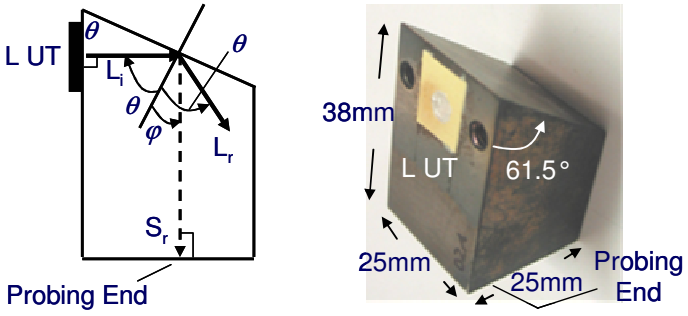


Figure 4. Energy reflection coefficient vs. θ (incident angle).

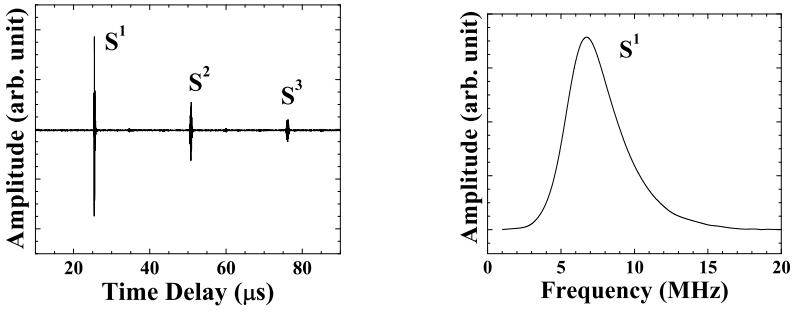
In order to achieve S wave HTUTs firstly we fabricated the 100 μm thick BIT/PZT film L wave UT using the sol-gel spray technique described in Section 2. Let this L UT be in a plane parallel to the mode converted S wave direction as shown in Fig. 5(a). This approach could reduce the machining time of the substrate and thick UT film fabrication difficulty. The top electrode was made by platinum paste which can sustain the temperature up to more than 450°C. By considering this criterion, $\theta + \varphi$ is required to be 90°. From Eq. (1), which is the Snell's law, we can obtain θ = 61.5°. At this angle, the conversion

rate is 96.7% that is only 0.8% smaller than the maximum conversion rate at 67.2°, based on the result in Fig. 4. Therefore, Figs. 5(a) and 5(b) show the design schematic and actual device developed for this study, respectively. Figures 6(a) and 6(b) show the ultrasonic signal in time and frequency domain, respectively, of the received S_r wave in the pulse-echo mode at 350°C. The S^n represents nth round trip of the S wave echoes traversing back and forth between the L UT and the probing end in Fig. 5. The center frequency of the S^1 echo was 6.7 MHz and the 6 dB bandwidth was 3.8 MHz. The signal-to-noise ratio (SNR) of S^1 echo was about 30 dB. The SNR is defined as the ratio of the amplitude of the S^1 echo over that of the undesired signals between the S^n echoes in Fig. 6(a). The signal strength of the S^1 echo at 350°C was 5 dB smaller than that at room temperature. It can be seen that the received L wave is not visible due to the fact that the dimension of the substrate has been chosen so that the reflected L wave from the probing end does not enter into the aperture of the L UT.



(a) (b)

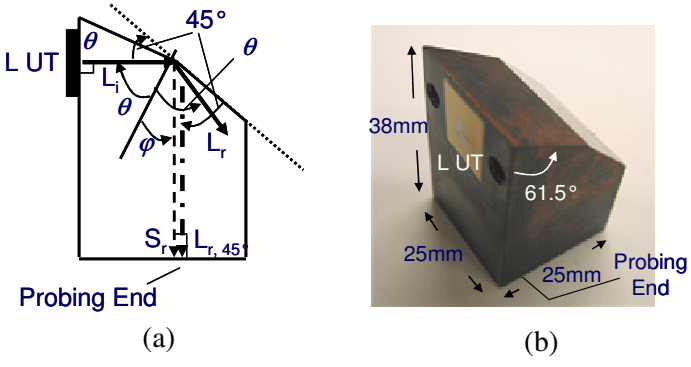
Figure 5. (a) Schematic diagram and (b) actual device of an integrated S wave UT probe with the L wave UT is located in a plane parallel to the direction of mode converted S wave where $\theta = 61.5°$.



(a) (b)

Figure 6. Ultrasonic signal in (a) time and (b) frequency domain of the S wave UT probe shown in Fig. 5 at 350°C.

If one would like to generate and receive both L and S waves at the same time, then the S wave probe shown in Fig. 5 can be modified to achieve such a purpose. In fact, it simply makes a slanted surface with an angle 45° from the intersection of the slanted plane and the line from the center of the L UT as shown in Fig. 7(a). The 45° angle plane will reflect the energy of the L_i wave into the $L_{r,45°}$ wave normal to the probing end as shown in Fig. 7(a). Therefore, in principle, the upper part of the L wave, generated from L UT, can be used to produce the S_r wave and the lower part to produce the $L_{r,45°}$ wave. Figure 7(b) shows an actual device developed. Figure 8 shows ultrasonic signal in time domain in the pulse-echo mode at 350°C, in which the S_r (S^1) and L_r (L^1) waves are observed simultaneously. The L^1 represents the first round trip L wave echo traversing between the L UT and the probing end. The center frequencies of the S^1 and L^1 echoes were 7.0 MHz and 7.0 MHz and the 6 dB bandwidths were 3.0 MHz and 3.8 MHz, respectively. During the top electrode fabrication for the device shown in Fig. 7(b), the area of the top electrode was adjusted so that the amplitude of the reflected S_r and $L_{r,45°}$ waves were nearly the same. The SNR of the L^1 and S^1 was about 20 dB.



(a) (b)

Figure 7. (a) Schematic diagram and (b) actual device of an integrated L and S wave probe with the L wave UT located in a plane parallel to the direction of S_r wave.

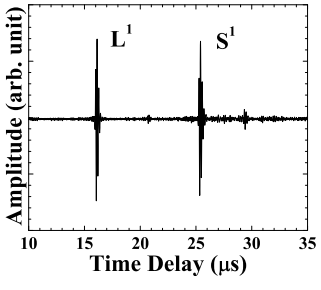


Figure 8. Ultrasonic signal in time domain of the L and S wave UT probe shown in Fig. 5 at 350°C.

4. Integrated RAW and PAW Transducer

For NDT applications there is a critical need for integrated in-situ sensors for local and global damage detection and assessment [3,4]. In the past the local aspect of the damage monitoring on metals and graphite/epoxy composites has been reported [14,15]. It is known that RAW and PAW transducers [27,30] can be used for NDT of metals such as SS and Al alloys in the range of several centimeters or meters depending on the attenuation characteristics of the substrates. In the common practice the L or S wave UTs and a wedge are used to generate and receive the desired RAW and PAW with the proper mode conversion inside and through the wedge [1,2]. However, there is a requirement of an ultrasonic couplant between the wedge and the sample under test. It is difficult to apply these UTs, wedges and couplants on curved surfaces and at HT.

In this section, the purpose is to develop techniques for on-site fabrication of RAW and PAW transducers directly onto desired SS and Al alloy substrates for NDT applications at 150°C. Since pulse-echo modes are of interest for NDT applications, most of our measurement data will be shown for this mode although measurement data in transmission modes will be demonstrated as well. One goal is to use these structurally integrated sensors, to inspect, for example, in-flight aircraft critical components, thus increasing platform availability and reducing associated maintenance costs.

The fabrication processes of RAW and PAW transducers are the same as those for the fabrication of L UT except that the top electrode is made in a shape of interdigital transducer (IDT). Since in this study the desired RAW and PAW operation frequency range is between 0.5 and 2.0 MHz, one mask used in this study is designed, fabricated by the electrical discharge machining (EDM) method and shown in Fig. 9. The top and bottom connection electrodes parallel to the wave propagation direction are called bus-bars, the other thin electrodes perpendicular to the wave propagation direction are called fingers. The finger widths of the IDT are 0.5 mm for the IDT. The separation among the fingers is also 0.5 mm wide. The mask is made of a 0.57 mm thick SS plate. The thickness is chosen so that the mask is flat, has negligible shadow effect during the colloidal silver spray and is reusable. Such finger width is convenient for the colloidal silver spray method. This IDT fabrication technique makes the selection of finger size and sensor size simple and convenient.

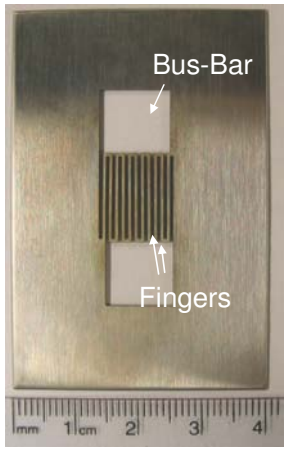


Figure 9. Mask of an IDT pattern.

4.1. *Integrated Ultrasonic Transducer on Al Alloy Substrates for RAW and PAW Measurements*

Al alloys are common materials for aircraft structures and other transportation systems, such as automobiles. A 25 mm thick Al alloy plate is used here for RAW experiments. Two IDTs were made on top of the 86 μm thick PZT/PZT film by the colloidal sliver paint spray. The thickness was about several microns. Figure 10 shows the integrated PZT/PZT composite film transducer with the IDT near edge "A" operated in the pulse-echo mode at 150°C. The limitation of 150°C came from the consideration of the 350°C Curie temperature of PZT/PZT composite films and sprayed colloidal silver paint IDT electrodes.

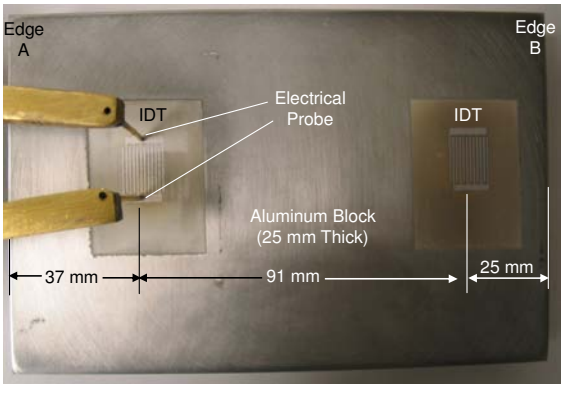


Figure 10. A 25 mm thick Al alloy plate with an IDT RAW transducer operated in pulse-echo mode. The 86 μm thick piezoelectric film was made of a PZT/PZT composite.

The measured RAW signals in time domain with a band pass filter between 0.5 MHz and 2.0 MHz is given in Fig. 11. In Fig. 11 R_A, R_B and R_{A+B} are the reflected echoes either from the edge "A" or the edge "B" through the corresponding Rayleigh wave travel paths (distances) of 2A, 2B, 2(A+B), respectively. The longest travel distance in this figure was 306 mm (for R_{A+B}). The two edges can be considered as large deep defects (cracks) in practical NDT applications. It means that this integrated RAW transducer can be regarded as a good NDT tool at 150°C for a sensing distance of 306 mm (e.g. R_{A+B}). The noises in Fig. 11 came from the scattered bulk waves from the lower corners of the two edges of the Al alloy substrate. It is noted that when both IDTs shown in Fig. 10; one as the RAW transmitter and another as the receiver, are used, transmitted RAW can be obtained. The IDT near edge "B" can be also operated in the pulse-echo mode.

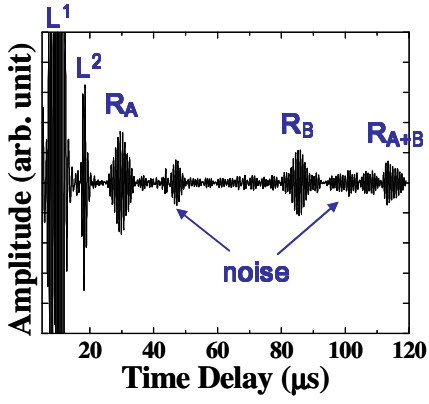


Figure 11. Ultrasonic performance of an IDT RAW transducer shown in Fig. 10 and operated in pulse-echo mode at 150°C in time domain with a band pass filter between 0.5 MHz and 2.0 MHz.

At room temperature the measured L wave velocity V_l and S wave velocity V_S of this Al alloy substrate are 6343 m/s and 3044 m/s, respectively. Using these data and Eq. (4) in [27,30]

$$\left(\frac{V_R}{V_S}\right)^6 - 8\left(\frac{V_R}{V_S}\right)^4 + 8\left\{3 - 2\left(\frac{V_S}{V_l}\right)^2\right\}\left(\frac{V_R}{V_S}\right)^2 - 16\left\{1 - \left(\frac{V_S}{V_l}\right)^2\right\} = 0 \qquad (4)$$

The calculated RAW group and phase velocity V_R is 2846 m/s. It is noted that for isotropic substrates the phase velocity is equal to the group velocity [27,30]. At room temperature the measured V_R is 2840 m/s which agrees well with the theoretically calculated value. At 150°C, as shown in Fig. 11, the measured RAW

velocity is 2700 m/s and the signal strength at 150°C is 13 dB weaker than that at room temperature. The reason of this high attenuation at 150°C is being investigated.

Figure 11 also shows the existence of L waves represented by L^1 and L^2 through the thickness direction of the Al alloy substrate. They are generated and detected by the same IDT transducer in Fig. 10. The L^1 and L^2 are the first and second round-trip echoes through the thickness of the 25 mm thick Al alloy substrate. These films are excellent bulk wave UT as illustrated in [14,15]. In Fig. 11 the L^1 signal was saturated. The strength of the L wave can be adjusted by the electrode size of the upper and lower bus-bars and finger width of the IDT. By comparing the IDT shown in Fig. 9 with that shown in Fig. 10 we can see that the 2 mm width of the bus-bars of the IDT pattern in Fig. 10 is narrower than 9.5 mm shown in Fig. 9. The narrower is the width of these bus-bars, the weaker is the L wave and the stronger the strength of RAW. This indicates that NDT can be carried out not only by the RAW along the surface of the Al alloy substrate but also by the L wave along the thickness direction.

Since the center frequency of the bulk L wave with this PZT/PZT composite film is around 7 MHz, if the pass band is extended to higher frequency than 2 MHz used for RAW, the L wave signal strength will become stronger. The pass band of the band pass filtering can be also adjusted and carried out by software in real-time. Because the RAW transducer shown in Fig. 10 is of layer structure which consists of the PZT/PZT composite film and Al alloy substrate, the excitation efficiency [31] of the RAW device with respect to the PZT/PZT film thickness and the finger width of the IDT which affect operating frequency will be further investigated in order to strengthen the RAW signals. It is noted that the measured L wave velocity in PZT/PZT composite film is 2200 m/s which is slower than that 6343 m/s in Al alloy, the V_R along the surface of PZT/PZT composite film on Al alloy shown in Fig. 10 should be slower than that on the surface of Al alloy surface without the film [27,30,31].

4.2. *Integrated PAW on a SS Plate*

For PAW experiments a 111 μm thick BIT/PZT composite film [14,15] was deposited on a 0.702 mm thick SS plate as shown in Fig. 12. Then two IDT electrodes were made on top of the film by the colloidal silver spray using the IDT mask shown in Fig. 9. Their thickness was about several microns. Figure 13 shows the measured PAW signals in time domain using the integrated PAW transducer shown in Fig. 12 near edge "A" operated in the pulse-echo mode at 150°C. The pass band of the band pass filter was also between 0.5 MHz and

2.0 MHz. In Fig. 13 P_A, P_B , P_{A+B}, P_{2A+B} and P_{A+2B} are the reflected echoes either from the edge "A" or the edge "B" through the corresponding PAW travel paths (distance) of 2A, 2B, 2(A+B), 2(2A+B) and 2(A+2B), respectively. The longest travel distance in this figure was 594 mm (for P_{A+2B}). The echoes of P_A and P_B are weaker than P_{A+B} because they travel uni-directionally along A or B direction, respectively, but echo P_{A+B} travel both along "A" and "B" direction and summed at the generating and receiving IDT near the edge "A". Along longer propagation distance echo strength will be weaker due to higher loss in the path. However, the echoes in Fig. 13 show good SNRs. The two edges can be considered as large deep defects (cracks) in practical NDT applications. It means that this integrated PAW transducer can be regarded as excellent NDT tool at 150°C for a sensing distance of 594 mm (e.g. P_{A+2B}).

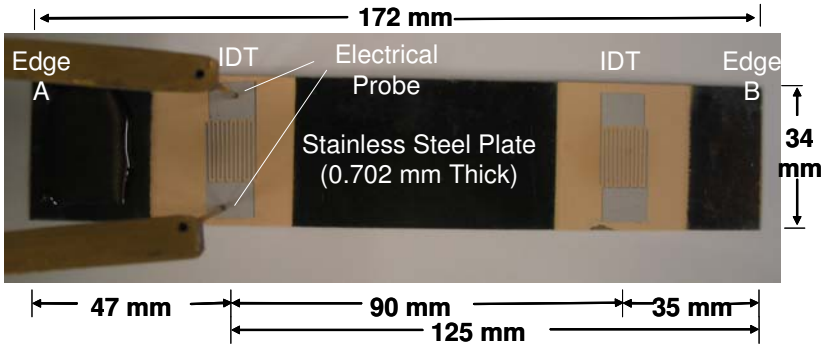


Figure 12. A 0.702 mm thick SS plate with a PAW transducer with an IDT operated in pulse-echo mode. The 111 µm piezoelectric film was made of a BIT/PZT composite.

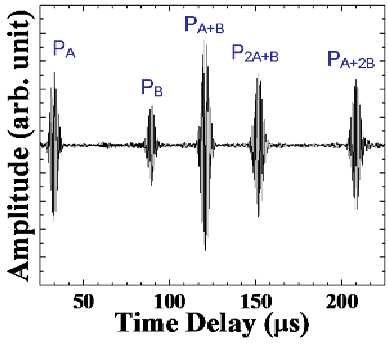


Figure 13. Ultrasonic performance of PAW transducer shown in Fig. 10 with an IDT operated in pulse-echo mode at 150°C in time domain. The SS substrate is 0.702 mm thick.

When both IDTs shown in Fig. 12; one as the PAW transmitter and another as the receiver, are used, the measured PAW signal, P, in time domain and in transmission mode is shown in Fig. 14. The IDT near edge "B" can be also individually operated in the pulse-echo mode. BIT/PZT films were shown to be able to sustain operation temperature more than 440°C [14,15].

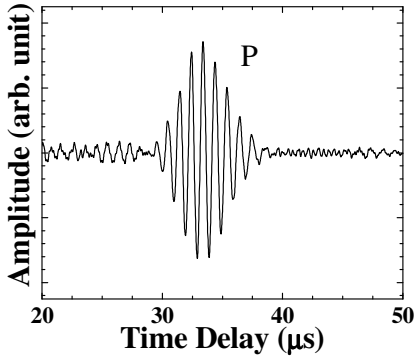


Figure 14. Ultrasonic performance of PAW transducers shown in Fig. 12 with IDT operated in transmission mode at 150°C in time domain.

At room temperature the measured PAW group velocity of the device shown in Fig. 12 is 2936 m/s. The measured V_l and V_s of the SS plate at room temperature are 5828 m/s and 3151 m/s, respectively. Let $k_l = \omega/V_l$, $k_S = \omega/V_S$, $k = \omega/V$, $\omega = 2\pi f$ and $q = (k^2 - k_l^2)^{1/2}$ and $s = (k^2 - k_s^2)^{1/2}$ where f is the frequency, then the equations governing the anti-symmetric and symmetric modes Lamb waves are given in Eqs. (5) and (6), respectively [30].

$$\left(k^2 + s^2\right)^2 \coth\left(qh/2\right) - 4k^2 qs \coth\left(qh/2\right) = 0 \tag{5}$$

$$\left(k^2 + s^2\right)^2 \tanh\left(qh/2\right) - 4k^2 qs \tanh\left(qh/2\right) = 0 \tag{6}$$

Using Eqs. (5) and (6) the calculated PAW phase V (in solid lines) and group velocities V_g (in dashed lines), and RAW velocity V_R (2919 m/s) are shown in Fig. 15. It is found that at the plate thickness h = 0.702 mm and f = 1 MHz the theoretically calculated V_g for the first anti-symmetric mode a_0 is 3030 m/s. The measured V_g was 2936 m/s which agrees well with the calculated value of 3030 m/s. Note that the calculated result does not include the effect of the BIT/PZT composite film loading. Therefore the PAW shown in Fig. 13 is the first anti-symmetrical plate wave mode a_0. Because the BIT/PZT composite film has slower L wave velocity, 2450 m/s, than that, 5828 m/s, of SS plate, its

111 μm thickness will slow down the PAW in the plate region coated with this composite film [27,30,31].

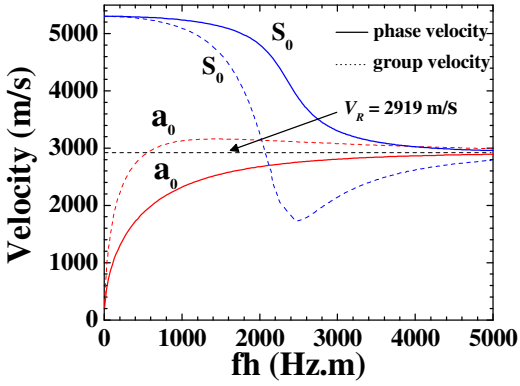


Figure 15. The calculated PAW and SAW (Rayleigh wave) dispersion curve for the 0.702 mm thick SS plate shown in Fig. 12.

The geometry of the PAW transducer, such as thickness of the BIT/PZT film and finger widths of the IDT which affect operating frequency will be investigate in order to excite PAW waves efficiently. By comparing the measurement data in Fig. 13 with those at room temperature it is found that the signal strength was 2 dB weaker at 150°C. The measured PAW (a_0) at 150°C was 2860 m/s.

5. Flexible Ultrasonic Transducers

For NDT of objects having curved surfaces flexible UTs are more suitable than classical UTs with flat probing ends. This capability insures the self-alignment of the UTs to the curved and complex geometry so that the transmitted ultrasonic energy can be maximized for improved diagnoses. Thus, the research interest here is to develop flexible UTs for potential implementation at room temperature and HT applications on samples of simple and complex geometries. In commercially available flexible transducers, piezoelectric polymers such as polyvinylidene fluoride (PVDF) [32,33] and piezoelectric ceramic/polymer composites [34-36] are mainly used as piezoelectric materials. Both materials include polymer which prevents the use of such flexible transducers at elevated temperature. For instance, PVDF shows significant piezoelectric deterioration above 65°C. Several copolymers have superior temperature stability compared to PVDF, however, operation temperature is limited to around 90-100°C [32]. In addition, piezoelectric polymers have low electromechanical coupling

coefficients [33,37]. Compared to bulk piezoelectric ceramics or polymers, piezoelectric ceramic/polymer composites may have superior electromechanical coupling properties in addition to flexibility and low dielectric losses. However, at HT good ultrasonic performances were reported only up to 80°C in particular under thermal cycling environment [38] due to the soft state of the resin as it reaches its glass transition temperature of 150°C. Other flexible HTUTs have been also reported [39]. Because of single crystal films used, to provide the flexibility, the thickness of the piezoelectric film ranged from 0.2 to 10 µm. These films provide an operating frequency normally higher than 30 MHz that may not be suitable for NDT of thick and highly attenuating materials.

5.1. *Flexible UT Using SS Foil as Substrates*

Since SS substrate can sustain operation temperature more than several hundred degrees Celsius (> 700°C), it may be brazed or welded onto metallic structures for HT operation. It may be also glued onto metallic, polymer or graphite/epoxy composite materials for NDT applications. The techniques described in Section II were used to fabricate the PZT/PZT composite films. The top electrodes were made by silver paste. Figure 16 illustrates two views of a PZT/PZT 120 µm thick composite film five UTs array directly fabricated onto a 75 µm thick SS foil with this particular merit. Thinner SS foils may also be used as reported in [40]. The entire transducer array structure was sandwiched by polyimide films excluding the probing side of the membrane (the side opposite to thick piezoelectric film) so that it may be protected from the moisture in the environment and can operate at temperatures up to 150°C which is currently limited by the operating temperature of PZT/PZT composite. Copper strips were used for electrical connections.

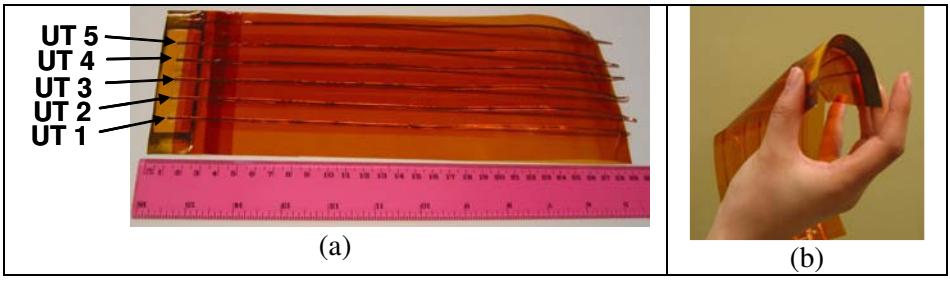


Figure 16. A flexible PZT/PZT 120 µm thick composite film five UT array using a 75 µm thick SS membrane.

When this flexible UT array was pressed onto a steel plate of 13.8 mm together with HT coupling oil and fixed by a mechanical clamp, the measured ultrasonic response of UT3, shown in Fig. 16, in time and frequency domains at 150°C, is obtained and presented in Fig. 17. L_2, L_4 and L_6 represent the first, second and third round trip echo through thickness of the steel plate. The center frequency, 6 dB bandwidth and SNR of the L_2 echo are determined to be 2.6 MHz, 2.6 MHz and 22 dB, respectively. We have also used steel, titanium, nickel and copper membranes for the fabrication of flexible UTs [15]. This type of flexible UTs has been also employed to inspect metallic tubes, graphite/epoxy composite and polymer materials [40].

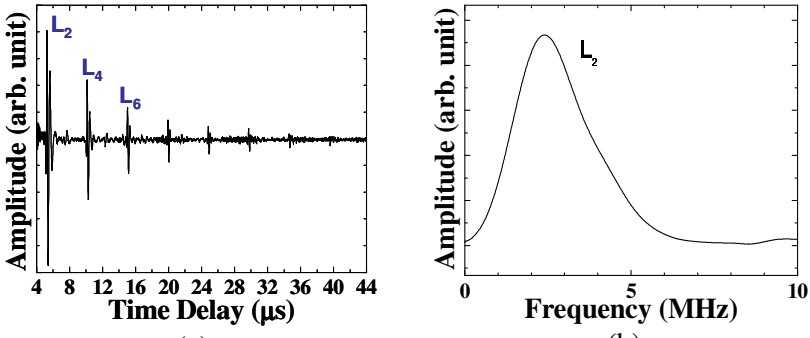


Figure 17. Ultrasonic performance of the flexible ultrasonic transducer (UT3) in (a) time and (b) frequency domains for NDT of a 13.8 mm thick steel plate at 150°C.

In order to ensure the flexibility, the thickness of the SS foil must be reasonably thin. For example, 150 μm thick SS sheet has a poor fitting and restoring ability with a curvature of 60 mm radius. However, foil thickness has influences on the ultrasonic performance of the flexible UT such as frequency and coupling with the test sample. Therefore, experiments with different foil thicknesses were performed. First, PZT/PZT films about 70 μm thick were deposited onto 38 and 75 μm thick SS foils, respectively. Both UTs were fabricated under the same process. Each transducer had a silver paste top electrode with a diameter of 6 mm. The test object was a 19.0 mm thick Al plate. The results at room temperature of the 1st round-trip echo L_2 in time and frequency domains are shown in Figs. 18(a) and (b), respectively. It is found that the flexible UT coated onto a 38 μm thick SS foil has higher signal strength by 5 dB, a broader bandwidth, and a higher center frequency than the UT that coated onto a 75 μm thick foil.

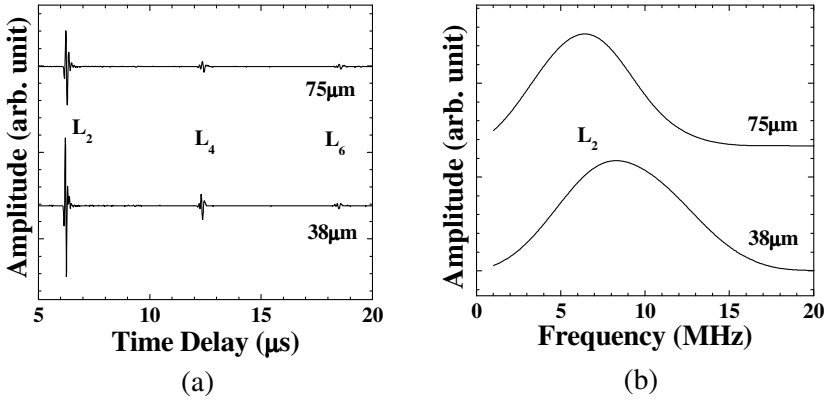


Figure 18. Experimental results of PZT/PZT flexible UTs with 38 and 75 μm thick SS foils onto a 19 mm Al plate in the (a) time domain and (b) frequency domain at room temperature.

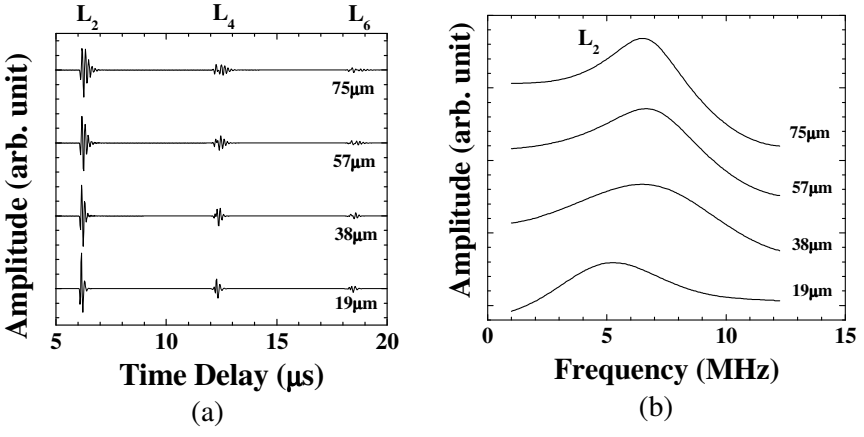


Figure 19. Calculated results of PZT/PZT flexible UTs with 19, 38, 57 and 75 μm thick SS foils onto a 19 mm Al plate in the (a) time domain and (b) frequency domain at room temperature.

In order to verify this tendency, theoretical calculations were carried out for comparison to experimental results. The transfer matrix formalism [41] was used to evaluate the ultrasonic wave interaction with a multilayered structure composed of a silver paste top electrode, a PZT/PZT piezoelectric film, a SS foil, a silicone oil layer as couplant, and a 19.0 mm thick Al plate. Information on each layer, such as thickness, density, longitudinal wave velocity, was required for the calculations. The data for PZT/PZT was taken from Ref. [42]. During the simulation, the only variable factor was the thickness of SS foil. In addition to the thicknesses of 38 and 75 μm, which were used in the experiment, the thicknesses of 19 and 57 μm were also considered. The results of the first echo L^1 in the time

domain and the frequency domain are shown in Figs. 19(a) and (b), respectively. A fairly good agreement of the tendency, i.e. relative signal strength, bandwidth and center frequency for the different SS foil thicknesses, is found between the experimental and calculated results. The calculations also show that even thinner foil can be used while not disturbing the ultrasonic performance of the flexible UT. With the thinner substrate, the robust properties would be of concern.

5.2. *Flexible UT Using Polyimide as Substrates*

In order to increase the flexibility of the flexible UT polyimide films are used for this investigation because polyimide films can sustain 350°C and the sol-gel PZT/PZT multilayer fabrication process. Its acoustic impedance is also different than that of the SS substrate. Because the polyimide film is an insulator, a colloidal silver spray and nickel alloy electroless plating methods were developed to coat a conductive layer onto the piezoelectric composite film side of the film as the bottom electrode before the coating of the piezoelectric composite thick films. However, in this particular sample, colloidal silver spray method was used and 2 μm thick silver paint was deposited between the top polyimide film and the PZT/PZT composite. Care has been taken to strengthen the adhesion between this electrode layer and the polyimide. Figure 20 shows the flexible UT with the 60 μm polyimide film substrate with a 60 μm thick PZT/PZT composite film. This PZT/PZT film thickness is obtained using ten layers of coating process. Such fabrication process implies that the polyimide film has sustained ten thermal cycles during the drying and firing of the PZT/PZT composite film and also the corona poling heating process. As shown in Fig. 20, the current process produces less than perfect flat UT surfaces (presence of wrinkles) due to the applied processing heat. Efforts are being devoted to reduce or eliminate such wrinkles and imperfections.

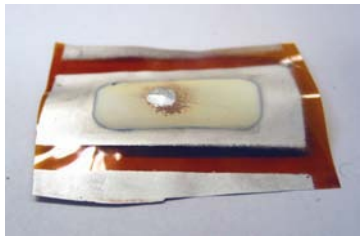


Figure 20. A flexible UT using polyimide film as the substrate.

In order to demonstrate the performance of the flexible UT using polyimide film as the substrate at elevated temperature, the transducer was pressed onto an

Al plate heated to 150°C. HT oil couplant was placed between the probing side of the polyimide film and Al plate. Figures 21(a) and (b) show the transducer response, in time and frequency domains respectively, in pulse-echo mode at room temperature. L_2 and L_4 are the first and second round trip echoes through the thickness of the Al plate. The center frequency, the 6 dB bandwidth and SNR of the L_2 echo are determined to be 15.4 MHz, 10.8 MHz and 25 dB, respectively. The ultrasonic response of the transducer at 150°C is shown in Fig. 22. The center frequency, the 6 dB bandwidth and SNR of the L_2 echo at this operating temperature is 13.4 MHz, 6.7 MHz and 20 dB, respectively. It is observed that the signal strength of the L_2 echo is decreased by about 3 dB and the frequency bandwidth reduced by 4.1 MHz. At 150°C, the time delay of the L_2 echo traveling in Al substrate has 0.2 μs more delay than that measured at room temperature.

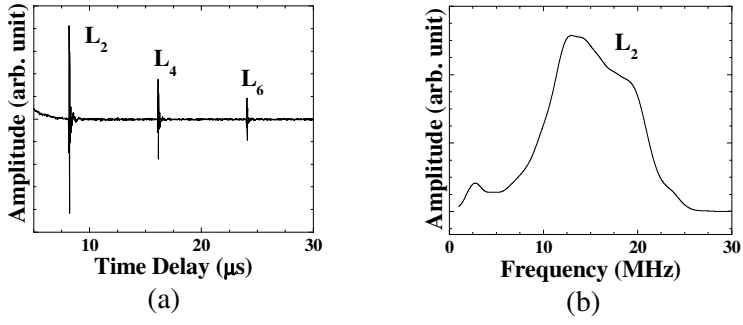


Figure 21. Room temperature flexible UT response in (a) time and (b) frequency domain for a 25.2 mm thick Al plate.

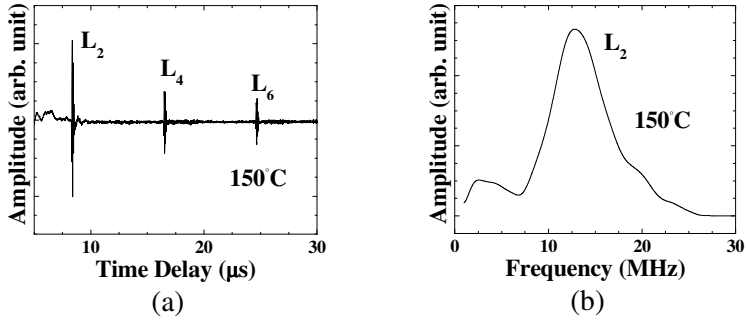


Figure 22. HT (150°C) flexible UT response in (a) time and (b) frequency domain for a 25.2 mm thick Al plate.

6. Conclusions

Thick (> 40 μm) ceramic films as piezoelectric UTs were successfully deposited on metallic and non-metallic substrates by a spray technique. In the film fabrication a composite consisting of piezoelectric BIT or PZT powders well mixed with PZT solution was directly sprayed onto different substrates. It is then dried, fired or annealed by heat. Multiple coating is used to achieve preferred thicknesses. A corona poling is used to achieve the piezoelectricity of the film. The top electrode is accomplished by a painting method with silver paste, platinum paste or silver paint. All fabrication processes may be carried out on-site and by handheld devices. Films have been characterized by SEM images, dielectric constants, d_{33} and thickness mode electromechanical coupling constants. Integrated ultrasonic longitudinal, shear, surface and plate wave transducers have been made. The same technology was also used to fabricate flexible transducers consisting of a thin substrate, a piezoelectric ceramic film and electrodes.

Ultrasonic L wave probe was first integrated onto a planar steel substrate. The ultrasonic monitoring of the extension of artificial defect at 440°C has been demonstrated. Similar L wave sensors have been used for real-time, in-line, noninvasive and non-destructive ultrasonic diagnostics of barrel and screw wear during polymer extrusion [43], filling conditions and solidification of the part during polymer injection molding [44], and polymer degradation and filling incompleteness during micromolding [45].

Integrated ultrasonic S wave probes were fabricated onto steel substrates with the use of mode conversion from L to S waves. The theoretical calculation indicates that the maximum energy conversion rate from the L to S wave is 97.5% at the L wave incident angle θ = 67.2° as shown in Fig. 4, and the reduction of the energy conversion rate is within 1% in the θ range between 60.8° and 72.9°. Let the L UT be made in a plane parallel to the propagation direction of the mode converted S waves as shown in Fig. 5 and the θ be 61.5°, clear S waves were generated. The reduction of energy conversion rate at θ equal to 61.5° is only 0.8% smaller than the maximum conversion rate at θ = 67.2°. A probe that can simultaneously generate and receive both L and S waves was also demonstrated. All the above mentioned probes have been made and operated up to 350°C with a center frequency of 5–7 MHz, 6 dB bandwidth of 4–6 MHz, and SNR of more than 20 dB.

The fabrication and ultrasonic performance of integrated RAW and PAW transducers directly onto Al alloy and SS substrates for NDT applications at

150°C have been demonstrated. The IDT mask was made by an EDM method. A colloidal silver spray method was used to form the IDT electrodes at room temperature through the IDT mask. Experimental results show RAW propagation along a 25 mm thick Al alloy substrate and PAW propagation in the first anti-symmetrical mode along a 0.702 mm thick SS plate. The pass band of the filters for RAW or PAW is between 0.5 MHz and 2.0 MHz. The measured RAW and PAW results agree well with the theoretically calculated values. If the edges of the substrates can be considered as large defects or cracks, measured signals with good SNR demonstrated the NDT capability in a distance of several centimeters or tens of centimeters of these integrated RAW or PAW transducers operated at 150°C. If the substrate is thick, both SAW and bulk L wave can be generated and detected simultaneously for NDT applications along the surface or thickness direction, respectively of the substrate. Arrays of these transducers can be readily achieved as well.

Flexible UTs consisting of a 75 μm thick SS foil or a 60 μm thick polyimide film, a piezoelectric PZT/PZT composite film and electrodes were developed. The flexibility was realized owing to the porosity of piezoelectric film and the thinness of substrate and electrodes. PZT/PZT composite was chosen as piezoelectric film because of its high piezoelectric strength. Thicknesses of 40 to 120 μm were obtained by the sol-gel spray technique. Electrode materials are silver paste, platinum paste or sprayed silver paint. SS foil itself serves as the bottom electrode; however, polyimide film required a coating of a conducting layer before the deposition of PZT/PZT composite film. The UT array can be conveniently configured by making several top electrodes. The flexible UT array with the SS foil has been successfully tested at 150°C for NDT of a 13.8 mm thick steel. In order to investigate the influence of the metal foil thickness on the ultrasonic performance of the flexible UTs, 70 μm PZT/PZT films were fabricated onto 38 and 75 μm thick SS foils and their ultrasonic performances were compared with the simulation results. The flexible UT with 38 μm SS foil showed a higher signal strength, broader bandwidth, and higher center frequency than that with 75 μm SS foil and calculation results are in agreement with this tendency. The flexible UT with the polyimide film was also used for NDT of a 25.2 mm thick Al plate at room temperature and 150°C. Such flexible UTs are expected to be applicable for health monitoring of humans and structures with complex shapes and geometries.

Acknowledgments

Authors would like to thank Y. Ono, J.-F. Moisan, H. Hebert, D. Lévesque, K.-T. Wu, Q. Liu, L. Song, J. Tatibouet and N. Mrad for their technical assistance. Financial support from Natural Sciences and Engineering Research Council of Canada is acknowledged.

References

1. J. Krautkrämer and H. Krautkrämer, *Ultrasonic Testing of Materials* (Springer-Verlag, Berlin, 1990).
2. A. S. Birks, R. E. Green, Jr. and P. McIntire, Eds., *Nondestructive Testing Handbook*, 2nd Ed., Vol. 7: Ultrasonic Testing (ASNT, 1991).
3. M. V. Gandhi and B. S. Thompson, *Smart Materials and Structures* (London; New York, Chapman & Hall, 1992).
4. J.-B. Ihn and F.-K. Chang in Ultrasonic Nondestructive Evaluation Engineering and Biological Material Characterization, Ed. T. Kundu (CRC Press, New York, 2004).
5. L. C. Lynnworth, *Ultrasonic Measurements for Process Control* (Academic Press, New York, 1989).
6. Kundu, T. Ed, Ultrasonic Nondestructive Evaluation: Engineering and Biological Material Characterization (CRC Press, New York, 2004).
7. J. R. Fothergill, P. Willis, and S. Waywell, *British Journal of NDT* **31**, 259 (1989).
8. N. D. Patel, S. X. Fulford, and P. S. Nicholson, *Review of Progress in QNDE* **9**, 823 (1990).
9. T. Arakawa, K. Yoshikawa, S. Chiba, K. Muto, and Y. Atsuta, *Nondestr. Test. Eval*. **7**, 263 (1992).
10. T. N. Nguyen, M. Lethiecq, B. Karlsson, and F. Patat, *Acta Acustica* **3**, 331 (1995).
11. H. Mrasek, D. Gohlke, K. Matthies, and E. Neumann, *NDTnet* **1**, no. 9 (1996).
12. H. Karasawa, M. Izumi, T. Suzuki, S. Nagai, M. Tamura, and S. Fujimori, *J. of Nuclear Science and Technology* **37**, 769 (2000).
13. R. Kazys, A. Voleisis, R. Sliteris, B. Voleisiene, L. Mazeika, P. H. Kupschus, and H. A. Abderrahim, *IEEE Sensor J.* **6**, 1134 (2006).
14. M. Kobayashi and C.-K. Jen, *Smart Materials and Struct.* **13**, 951 (2004).
15. M. Kobayashi, C.-K. Jen, Y. Ono and J.-F. Moisan, *CINDE* **26** 5 (2005).
16. H. Adachi, Y. Kuroda, T. Imahashi, and K. Yanagisawa, *Jpn. J. Appl. Phys*. **36**, 1159 (1997).
17. X. Zhao, J. R. G. Evans, and M. J. Edrisinghe, *J. Am. Ceram. Soc.* **85**, 2113 (2002).
18. R. A. Dorey, R. W. Whatmore, S. P. Beeby, R. N. Torah and N. M. White, *Integrated Ferroelectrics* **54**, 651 (2003).
19. P. Marechal, F. Levassort, J. Holc, L.-P. Tran-Huu-Hue, M. Kosec and M. Lethiecq, *IEEE Trans. UFFC* **53**, 1524 (2006).
20. C. Galassi, E. Roncari, C. Capiani, and P. Pinasco, *J. Eur. Ceram. Soc*. **17**, 367 (1997).
21. F. Levassort, T. Bove, E. Ringgaard, L. P. Tran-Huu-Hue, J. Hole and M. Lethiecq, *Proc. IEEE Ultrason. Symp.* (2003).
22. K. L. Gentry, J. M. Zara, S.-D. Bu, C.-B. Eom and S. W. Smith, *Proc. IEEE Ultrason. Symp.* **2** (2000).

23. X.-Y. He, A.-L. Ding, X.-S. Zheng, P.-S. Qiu and W.-G. Luo, *Microelectronic Eng.* **66**, 865 (2003).
24. M. Shimomura, T. Tsurumi, Y. Ohba and M. Daimon, *Jpn. J. Appl. Phys.* **30**, 2174 (1991).
25. N. Katsura, M. Ishikawa, T. Sato, M. Takeuchi, N. Kawashima, M. Kurosawa and S. Takeuchi, *Proc. IEEE Ultrasonics Symp.*, 1300 (2003).
26. D. A. Barrow, T. E. Petroff, R. P. Tandon and M. Sayer, *J. Appl. Phys.* **81**, 876 (1997).
27. B. A. Auld, *Acoustic fields and waves in solids*, Vol.2 (John Wiley & Sons, New York, 1973), p. 30-38.
28. M. O. Si-Chaib, H. Djelouah, and M. Bocquet, *NDT&E Int'l* **33**, 91 (2000).
29. W. G. Mayer, *Ultrasonics* **3**, 62 (1965).
30. I. A. Viktorov, *Rayleigh and Lamb waves* (Plenum, New York, 1967).
31. G. S. Kino and R. S. Wagers, *J. Appl. Phys.* **44**, 1480 (1973).
32. L. F. Brown and A. M. Fowler, *Proc. IEEE Ultrason. Symp.*, 607 (1998).
33. J.-M. Park, J.-W. Kong, D.-S. Kim, and D.-J. Yoon, *Composites Science and Technology* **65**, 241 (2005).
34. T. F. McNulty, V. F. Janas, A. Safari, R. L. Loh and R. B. Cass, *J. Am. Ceram. Soc.* **78**, 2913 (1995).
35. H. Karasawa, M. Izumi, T. Suzuki, S. Nagai, M. Tamura, and S. Fujimori, *J. of Nuclear Science and Technology* **37**, 769 (2000).
36. L. J. Bowen, R.L. Gentilman, H.T. Pham, D.F. Fiore, and K.W. French, *Proc. IEEE Ultrason. Symp.*, 499 (1993).
37. A. Gachagan, P. Reynolds, G. Hayward, and A. Mcnab, *Proc. IEEE Ultrason. Symp.*, 853 (1996).
38. C. Devallencourt, S. Michau, C. Bantignies, and N. Felix, *Proc. IEEE Ultrason. Symp.*, 1294 (2004).
39. M. Akiyama, T. Kamohara, K. Nishikubo, N. Ueno, H. Nagai, and T. Okutani, *Appl. Phys. Lett.* **86**, 022106 (2005).
40. M. Kobayashi, C.-K. Jen, and D. Lévesque, *IEEE Trans. UFFC* **53**, 1478 (2006).
41. D. Lévesque and L. Piché, *J. Acoust. Soc. Am* **92**, 452 (1992).
42. M. Sayer, M. Lukas, T. Olding, G. Pang, L. Zou, and Y. Chen, *Proc. Mat. Res. Soc. Symp.* **541**, 599 (1999).
43. C.-K. Jen, Z. Sun and M. Kobayashi C.-K. Jen, Z. Sun and M. Kobayashi, *Measurement Science and Technology* **16** 842 (2005).
44. M. Kobayashi, Y. Ono, C.-K. Jen and C.-C. Cheng, *IEEE Sensors Journal* **6**, 55 (2006).
45. B. R. Whiteside, E. C. Brown, Y. Ono, C.-K. Jen and P.D. Coates, *Plastics, Rubbers and Composites* **34**, 387 (2005).

This page intentionally left blank

CHAPTER 3

REAL-TIME AND IN-LINE ULTRASONIC DIAGNOSTICS OF POLYMER PROCESSES

Cheng-Kuei Jen, Zhigang Sun, Jacques Tatibouët, and Yuu Ono[†]

Industrial Materials Institute, National Research Council Canada
75 de Mortagne Blvd., Boucherville, QC, J4B 6Y4, Canada
cheng-kuei.jen@cnrc-nrc.gc.ca

[†]*Department of Systems and Computer Engineering, Carleton University*
1125 Colonel By Drive, Ottawa, Ontario, K1S 5B6, Canada

Real-time and in-line ultrasonic diagnostics of polymer extrusion and injection molding processes together with the high temperature ultrasonic probes used are presented. For polymer extrusion, melt quality, filler concentration, and dispersion monitorings are the examples. Clad ultrasonic rod probes consisting of a core and a cladding fabricated by a thermal spray technique are used. For injection molding, diagnostics of melt flow front, average flow speed, filling completion, and solidification are illustrated. Sprayed high temperature ultrasonic thick film transducers integrated onto the mold inserts are employed.

1. Introduction

Presently, world market and international competition demands high quality products manufactured by efficient and cost-effective technologies. The technical specifications of products are increasingly advanced and precise. New materials and mass produced items, even those made from conventional materials, need continuous and automated quality control to ensure production efficiency and product superiority. Such quality control requires advanced instruments that diagnose the manufacturing process and provide precise feedback on process parameters and materials properties during production. In this study real-time, in-line, non-invasive, and non-destructive diagnostics methods was introduced to monitor the polymer extrusion [1,2] and injection molding (IM) [3,4] processes

which are two popular mass production technologies for plastic products and of interest here.

Ultrasound, owing to its capability to probe inside materials nondestructively, is an attractive diagnostic approach to satisfy these requirements. Recently, in-line or in-situ ultrasonic monitoring of polymer extrusion [5-11], foam extrusion [12,13] and IM processes [14-19] has been demonstrated with the potential of process optimization. Since these polymer production processes are operated at elevated temperatures, e.g. 100 to 350°C, the ultrasonic sensors should operate at such elevated temperatures. Here, piezoelectric ultrasonic transducers (UTs) are considered because of their simplicity, high sensitivity, and cost-effectiveness. In addition to high temperature (HT) operation, these UTs are also required large bandwidth and high signal-to-noise ratio (SNR).

In the process monitoring and diagnostics, ultrasonic parameters such as ultrasonic velocity and attenuation in the polymer melts, scattered ultrasonic signals, for instance, by fillers in the polymer melts, reflected ultrasonic signal from the probe-polymer melt interface, etc. can be used to characterize melt properties such as viscosity [5,8], filler concentration [5,6,10], filler dispersion [5,6,10], melt and mixing quality [5], residence time distribution [7,11], filling completion [15,18,19], flow front arrival [16,17], solidification [17], polymer degradation [6,19], etc. and process parameter such as temperature [9,14,17]. All the experiments presented in this study were conducted in ultrasonic transmission mode for polymer extrusion and in pulse-echo mode for IM. The ultrasonic data acquisition system was composed of two pulser-receivers, two 12-bit dual-channel digitizing boards with a sampling rate of 50 MHz for each channel and one personal computer.

2. Diagnostics of Polymer Extrusion Processes

Polymer extrusion [1,2] is one of the most used industrial mass production processes of plastic products. In the process, polymer pellets are fed into the feed hopper, transported, melted, mixed and blended by screws in the heated barrel section under elevated temperatures (normally about 150°C ~ 220°C), and then converted into the melt at the extruder exit die that shapes the product. The shapes can be rods, tubes, films, sheets, filaments and pipes. In the process additives such as colorants and fillers, etc, which are also in pellet form, may be added. At the feed hopper polymer pellets enter into the extruder barrel and come into contact with the screw. The rotating screw or screws force the pellets forward into the barrel which is heated to the desired melt temperature. In most processes, a heating profile is set for the barrel in which three or more

independently controlled heaters gradually increase the temperature of the barrel from the feed hopper to the die exit. This allows pellets to melt gradually as they are pushed through the barrel and lowers the risk of overheating which may cause degradation in the polymer. Extra heat is contributed by the intense pressure and friction taking place inside the barrel.

At present, single screw extruders dominate the extrusion field. However, twin-screw and multi-screw extruders can achieve improved dispersion and mixing [1,2]. Polymer extrusion is also a main part of IM, blow molding, micromolding, etc. Therefore it is highly desired to have real-time in-line diagnostic capability to evaluate the performance of each segment of the extruder and all extrudates. In this section ultrasonic probes, which can be installed from the feed hopper to die exit [11,20], and their application in monitoring melt quality as well as filler concentration and dispersion will be presented.

2.1. *Ultrasonic Probes for Polymer Extrusion Monitoring*

For the experiments the ultrasonic probes used were composed of a steel clad buffer rod [20] and a longitudinal (L) wave piezoelectric UT attached to the UT end of the probe. The probing end of the probes had the same external diameter and threads as conventional Dynisco pressure and temperature probes as shown in Fig. 1 and could be installed at any ports designed for hosting Dynisco probes from feed hopper to the die exit. The probe was made of a 130.9 mm long clad buffer rod consisting of a steel core and a 1mm thick stainless steel cladding. The probing end was of 7.7 mm diameter. In this approach, the probing end of this rod was in contact with molten polymers whereas its UT end could be cooled by air or water, or coated with HTUT without any cooling. If cooling is used, commercially available ambient temperature piezoelectric UTs and couplants can be readily used, and the temperature of the probing end can be up to more than 600°C. Therefore such probe is suitable for diagnostics of all polymers during processing.

Figure 2a shows a schematic diagram of a setup with two ultrasonic buffer rod probes, which can be the one shown in Fig.1, operating in the transmission mode at an extrusion die. The probing end of the probes was flush with the inner surface of the exit die. Let the length of the probe be d and the gap of the exit die be h, which is also the thickness of the polymer melt during extrusion at the die. The ultrasonic wave was transmitted from the transmitting UT, propagating through the buffer rod with a length d, the thickness h of the melt and another buffer rod with a length d, and then reached the receiving UT. This signal is designated as L_A. L_B and L_C were the signals traveled through the melt thickness

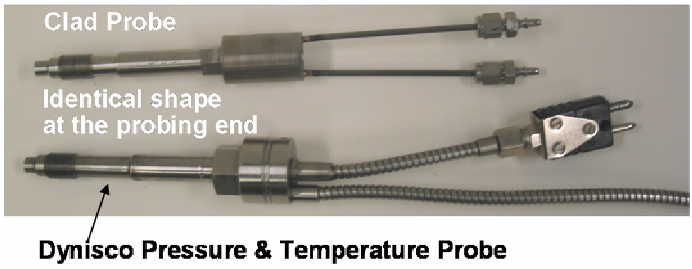


Figure 1. A clad ultrasonic buffer rod probe having the same external probing end diameter and thread as conventional Dynisco pressure and temperature probe. The probing end is of 7.7 mm diameter.

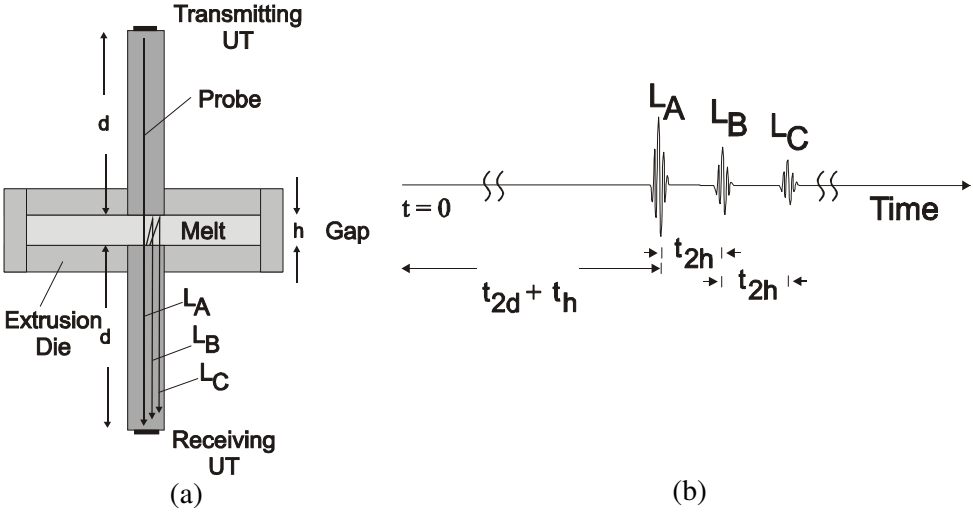


Figure 2. (a) Schematic diagram of a setup with two ultrasonic buffer rod probes operating in the transmission mode at an extrusion die and (b) time delays of transmitted ultrasonic signals.

two and four more times, respectively than L_A. The time delays of these three signals are illustrated in Fig. 2b. The ultrasonic L wave velocity in the melt can be obtained as $V_{melt} = 2h/t_{2h}$. The amplitude reduction of the signals which is associated with the ultrasonic attenuation in the melt can be expressed as $\alpha_{melt} = 10 \{\log (|L_A|/|L_B|)\}/h$, where $|L_A|$ and $|L_B|$ are the amplitude of the L_A and L_B signals, respectively.

Figure 3 shows reflected L wave signals (echoes) at an ultrasonic frequency of 5 MHz from the probing end of the probe shown in Fig. 1 and operated in the pulse-echo mode. L^1 and L^2 were the 1st and 2nd round trip echoes. The SNR of

the reflected L^1 signal was above 35 dB. It is known that the SNR of the signals obtained in the transmission mode are higher than those obtained in the pulse-echo mode. This is an advantage of the transmission mode [20]. The drawback of the transmission mode is the requirement of two probes with proper alignment. The dimensions of the probe shown in Fig. 1 were chosen to achieve high SNR and small size. The higher is the SNR, the higher is the sensitivity to monitor the fluctuations of the V_{melt} and α_{melt}. In this investigation only L wave signals were used. The ability of ultrasound to diagnose the molten polymers is realized if enough ultrasonic energy propagated through the polymer melts is received by the receiving UT. In the pulse-echo mode the transmitting UT is also used as the receiving UT.

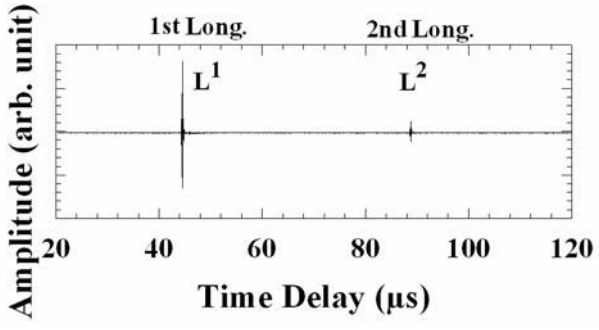


Figure 3. Reflected 5 MHz L wave echoes from the probing end of the probe shown in Fig. 1.

Since the absorption in melts is normally proportional to the square of the operating ultrasonic frequency, for thicker polymer melts and those having higher absorption, the operating ultrasonic frequency needs to be lower, and this necessitates a clad rod having a larger diameter [20]. Figure 4 shows a clad buffer rod having a 33 mm diameter probing end which can probe thicker polymer melts than that shown in Fig. 1.

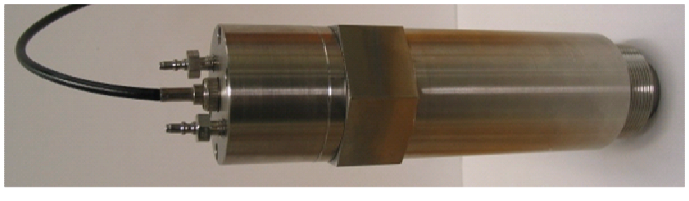
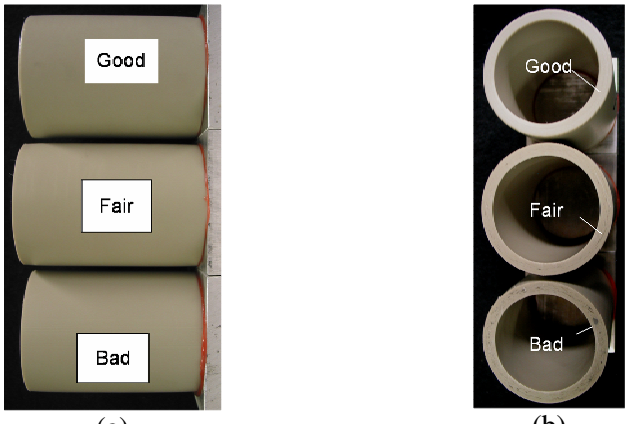


Figure 4. A clad ultrasonic buffer rod probe having a probing end of 33-mm diameter.

2.2. *Melt Quality Monitoring*

Because each manufacturer desires to enhance the productivity of polymers in a given extruder by maximizing the extrusion rate (tons/hr), and then a part of pellets may not be properly melted due to excessive production speed. Also, the extruders produced by different machine builders perform differently with regard to various polymers or same polymers supplied by different suppliers at different extrusion rate. Therefore there is a need to have an instrument which can help maximize the production rate and maintain or enhance the quality of extruded polymers. Below a demonstration will be made to illustrate how and why ultrasonic technology can be used for this purpose.

Figures 5a and 5b show the outer and inner looks, respectively, of three extruded polymer pipe samples having an outer diameter of 57.15 mm and a wall thickness of 6.35 mm obtained with low, medium and high extrusion speed. In Fig. 5b there are no visible un-melted pellets for the sample obtained with low extrusion speed and this sample is designated as “good”. However, visible un-melted pellets can be seen from the samples obtained with medium and high speed designated as “fair” and “bad” samples, respectively, in Fig. 5b. There were more and large un-melted pellets in “bad’ than those in “fair” sample. Clearly Fig. 5a shows that the optical surface examination technique was not sensitive to monitor the melt quality of the samples presented here.



(a) (b)

Figure 5. (a) Outer and (b) inner look, respectively of extruded polymer pipe with low (good), medium (fair) and high extrusion speed (bad). The outer diameter of the pipe was 57.15 mm and wall thickness was 6.35 mm.

To evaluate the effectiveness of the ultrasonic monitoring, two clad ultrasonic probes having the Dynisco shape as shown in Fig. 1 were used for the real-time and in-line diagnostics of polymer melt quality at the die exit here. These two probes shown in Fig. 6 were operated in the transmission mode. In addition, temperature and pressure probes were also installed at the same location but in different angular positions.

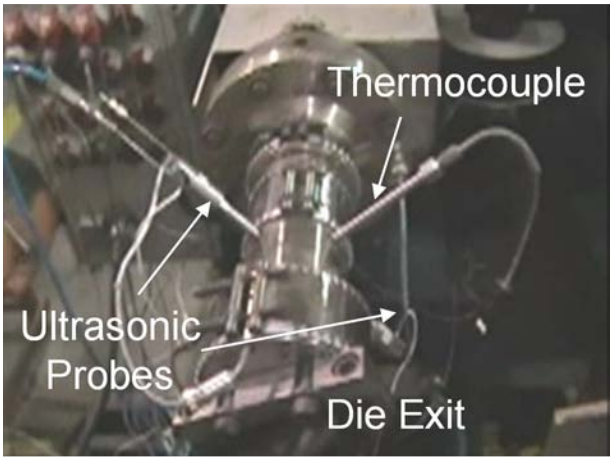


Figure 6. Two ultrasonic clad probes shown in Fig. 1 are used for the polymer melt quality monitoring at the die exit in a transmission mode.

Figure 7 shows a demonstration that during the extrusion of a polypropylene (PP) with the addition of several nylon pellets of higher melting temperature than PP, the ultrasonic probe could sense the un-melted nylon pellets in real time due to the instantaneous reduction of the transmitted ultrasonic signal as shown in Fig. 7. The reduction was caused by the acoustic impedance mismatch between the hard nylon pellet and the PP melt where acoustic impedance is defined by the product of the material density and ultrasonic L wave velocity in the material. Figure 7 also indicates that pressure and temperature probes are not sensitive to detect the un-melted pellets.

At an industrial 120 mm single screw extruder the ultrasonic sensors have been set up as shown in Fig. 6. The extruded materials were PP. The thickness of the melt between two probing ends of the two aligned clad ultrasonic probes was 33 mm which was also the gap size of the die exit. The extrusion throughput was nearly 400 kg/hr. The upper center plot of Fig. 8 shows the reduction of the transmitted ultrasonic signal strengths (amplitude) which correlate to the melt quality index according to the experience of the extrusion. The optical images of the samples extruded at the 29%, 36%, 43%, 50%, 57% and 64% of the

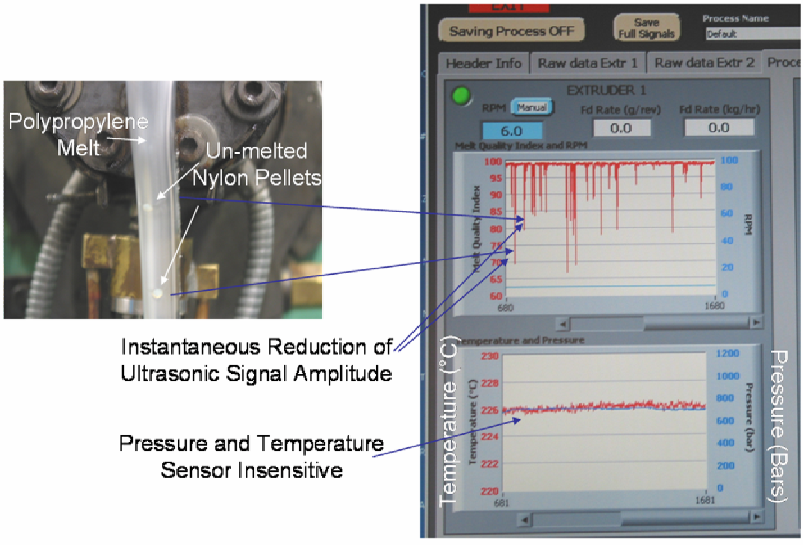


Figure 7. Sensitivity of ultrasonic, pressure and temperature probes for real-time and in-line detection of the un-melted nylon pellets in PP melt.

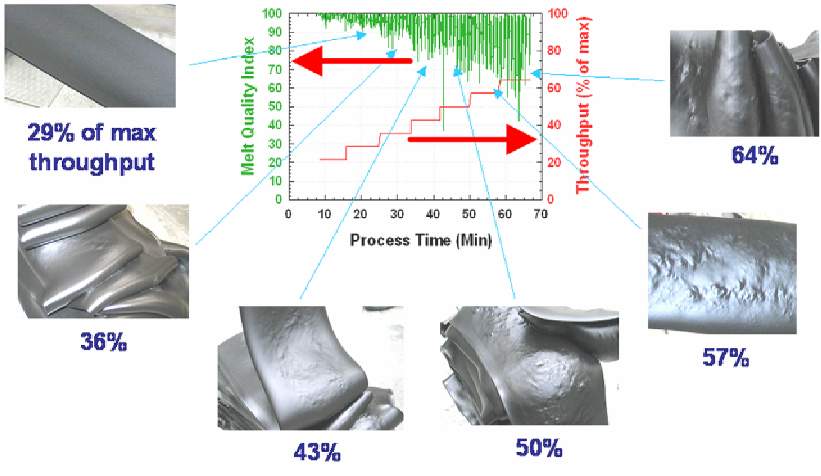


Figure 8. Real-time and in-line ultrasonic monitoring of melt quality index with corresponding optical images of the extruded PP with different amount of un-melted pellets due to different extrusion throughput.

maximum throughput of this extruder are also given in Fig.8. It can be seen that when the throughout increases, the amount of un-melted pellets increases as well, and the ultrasonic signals obtained in real-time and in-line accurately and timely indicate such property. It is noted that if the extrudates in Fig. 8 were in pipe

shape, the un-melted pellets may not be seen as indicated in Fig. 5. In Fig. 8 the small variations of melt quality index versus throughput may come from the fluctuations in size distribution and dispersion of the un-melted pellets. It is noted that this extruder and/or the extrusion conditions were not effective to extrude PP because of the indication that the melt quality of PP was good only when the extrusion throughput was nearly 30% of the maximum rate. This also means that the ultrasonic diagnostic technology can help evaluate and improve the extruder including screw design for extrusion of particular polymers.

2.3. *Filler Concentration and Dispersion Monitoring*

PP resins are used for numerous applications of plastic products. One common way to achieve different mechanical properties is to add calcium carbonate ($CaCO_3$) as fillers with various quantities into PP. Monotonic dependency has been obtained for both ultrasonic velocity and attenuation in $CaCO_3$ added PP with respect to the concentrations of $CaCO_3$, which can then be used as calibration curve for the determination of filler concentration, for a given attenuation or ultrasonic velocity value. Since ultrasonic attenuation is less sensitive to the fluctuation of pressure and temperature during extrusion [6], it is used in this investigation to monitor the $CaCO_3$ concentration and dispersion in PP.

A single screw extruder with an internal barrel diameter of 63.5 mm was used for the experiments. Two ultrasonic probes were installed in an instrumentation die in a similar fashion as that shown in Fig. 6. The pellets of PP with different $CaCO_3$ concentrations were first prepared on a twin-screw extruder, then extruded and monitored on the single screw extruder. During sample preparation, step changes in the feeding rate of the $CaCO_3$ filler were made while maintaining constant feeding rate of PP in order to obtain several concentration levels allowing one to derive a calibration curve for the ultrasonic characteristics as a function of $CaCO_3$ concentration. In all experiments, the average level of attenuation was determined for each filled PP of a given $CaCO_3$ concentration. For each concentration level, the extrudate was sampled and actual concentration was obtained from pyrolysis (ashing). The results are shown in Fig. 9.

To illustrate specifically the impact of $CaCO_3$ particle dimension on ultrasonic attenuation, two $CaCO_3$ samples were used with two mean particle diameters: 3.0 μm (Camel-Wite) and 0.7 μm (Camel-Cal). The curves for these two particle dimensions are clearly distinct, the smaller diameter leading to a lower attenuation at a given concentration. Figure 9 also shows results for stearate-coated grades known to impart improved particle dispersion for both

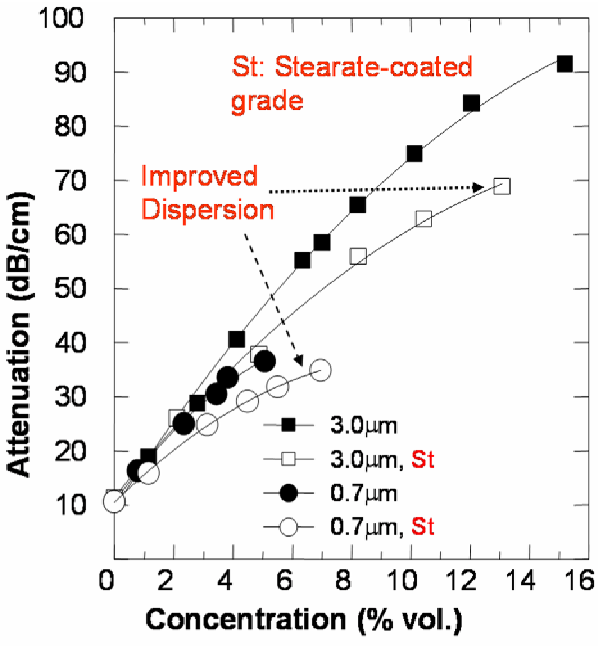


Figure 9. Variation of ultrasonic attenuation in PP as a function of $CaCO_3$ concentration.

3.0 μm and 0.7 μm mean diameter $CaCO_3$ in PP. This stearate-coated grade gave lower attenuation than non-treated one, indicating that the apparent diameter in the latter case was larger, due to agglomeration of particles. It thus implies that a specific calibration curve of ultrasonic attenuation versus $CaCO_3$ concentration could be obtained for each different grade of $CaCO_3$, with respect to particle size and dispersion.

3. Diagnostics of Polymer Injection Molding (IM) Process

Conventional polymer IM [3,4] is a process of forcing melted polymer into a mold cavity having a unique shape for a designed production parts. Each IM machine is equipped with a polymer extrusion section which will convey the polymer pellets, melt and mix them and collect them into the nozzle. Then the typical IM cycle can be described as: (i) the polymer melt is injected into the cavity of a mold through a gate from the nozzle; (ii) the cavity is completely filled with the material and the additional melt is forced into the cavity under high pressure until the gate is frozen in order to compensate for the shrinkage due to the continuous cooling. Then, the part is further cooled until it is sufficiently solidified; (iii) the mold is opened and the part is detached from the immobile

mold; and (iv) the part is ejected from the cavity of the mobile mold by the ejection pin. Then the mold is closed and the entire cycle, (i)-(iv), is repeated as shown in Fig. 10.

Due to the requirement of high speed, high efficiency and cost-effective production, in-line diagnostics technology is desired. Although the pressure, temperature and positioning sensors are equipped for the machine control effectively but since the material properties of the part are not known, process control was not achieved by the machine manufacturers. Ultrasonic diagnostics technology is introduced here to illustrate the feasibility of process control because part properties may be obtained from the ultrasonic parameters such as ultrasonic velocity and attenuation in the part, reflection and transmission coefficient at the mold-part interface and ultrasonic scattered signal in the part, etc.

In order to provide varieties of the products, co-IM [21], gas [22] and water assisted IM [22], powder IM [23] and micromolding [19] have been developed. In this study only ultrasonic diagnostics of convention IM is given as an example.

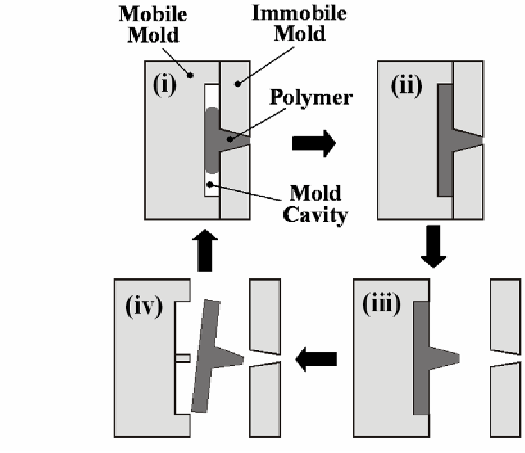


Figure 10. Cycle of the IM process: (i) filling; (ii) packing, holding and cooling; (iii) mold open; and (iv) part ejection.

3.1. *Integrated Ultrasonic Sensor Inserts for IM Diagnostics*

In order to monitor the behavior of the polymer inside the mold cavity during all types and sizes of IM processes, integrated ultrasonic sensor inserts approach was adopted. Thick film HTUT sensors were fabricated onto small size steel inserts by a sol-gel spray technique [18,24] and then embedded into the mold insert of proper size. Multiple mold inserts of various sizes are commonly used by

injection molders, in particular, for multi-cavity molding processes. Figure 11 shows the HTUT sensor inserts with (right) and without (left) an electrical connection. The HTUT sensors can operate up to more than 400°C [24] which can cover all polymer IM processes. These sensors do not require ultrasonic couplant and can be coated onto curved sensor inserts. Teflon insulated coaxial cables were used for electrical connection between the top electrode and the pulser-receiver used. The maximum operating temperature of this coaxial cable was 200°C. For applications at temperatures higher than 200°C, a ceramic insulated coaxial cable can be used as an electrical cable.

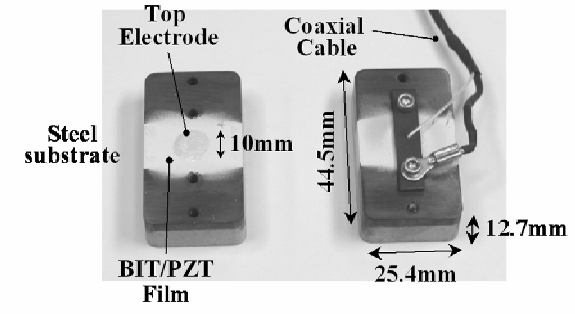


Figure 11. HTUT sensor inserts with (right) and without (left) an electrical connection used for monitoring of injection molding process.

Four HTUT sensor inserts (UT1-4) embedded into the mold insert are shown in Fig. 12, which also demonstrates that sensor array configuration is feasible. The mold, mold insert and sensor inserts were made of steel. The mold insert had dimensions of 76 mm-width, 165 mm-length and 21 mm-thickness, which was 1 mm thinner than the mobile mold shown in Figs. 12a and 12b so that a rectangular part with dimensions of 76 mm-width, 165 mm-length and 1 mm-thickness could be molded for the study. A hole at the center of the mold insert was for a part ejection pin. By replacing the mold insert, the shape and dimensions of the molded part can be easily modified to meet the customer's demands. The bottom surfaces of the integrated HTUT sensor inserts were flush with the mold insert surface (cavity surface), as seen in Fig. 12a. A distance between the center of the UT1 (UT3) and UT2 (UT4) was 34.9 mm and that of the UT2 and UT3 was 44.5 mm as shown in Fig. 12b. All the HTUTs had almost the same ultrasonic performance, which indicated that fabrication of the HTUTs was consistent. The center frequency and the 6 dB band width were 9-11 MHz and 6-8 MHz, respectively.

Figure 13 presents a cross-sectional view of the mold (mobile and immobile), mold insert and molded part (polymer) with four HTUT sensor inserts (UT1-4).

Polymer melt was injected into the cavity of the mold through the gate at the center of the immobile mold. For comparison purpose with ultrasonic data, a Kistler temperature and cavity pressure sensor, whose sensing end had circular shape with a diameter of 4mm and was flush with the internal surface of the immobile mold, was attached to the immobile mold. This Kistler sensor probing end was facing to the UT1 as shown in Fig. 13. L^n (n = 1,2,3…) represents n-th round trip echoes propagating in the HTUT insert and reflected at the insert-polymer interface, and L_{2n} (n = 1,2,3…) represents those in the polymer and reflected at the polymer-immobile mold interface.

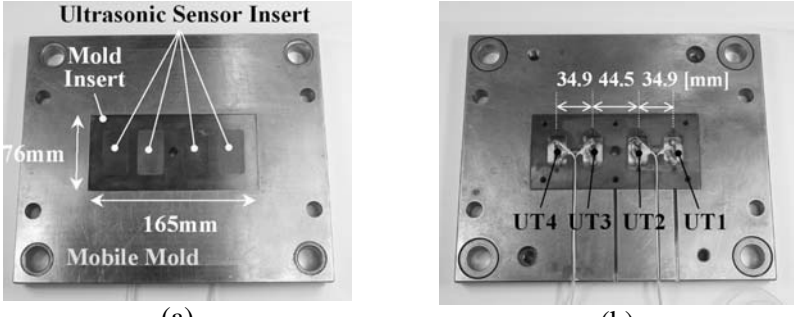


Figure 12. Mobile mold with four HTUT sensor inserts embedded in the mold insert.

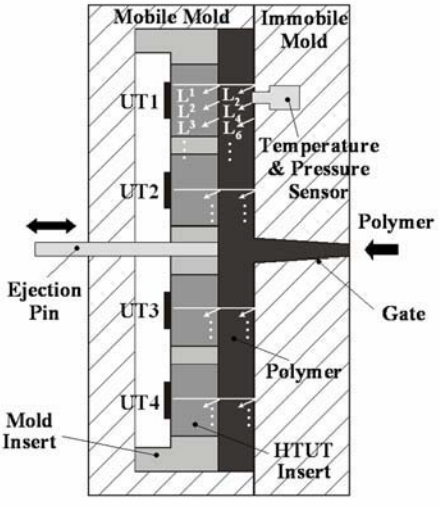


Figure 13. Cross-sectional view of the mold (mobile and immobile), mold insert and molded part (polymer) with four HTUT sensor inserts (UT1-4). L^n and L_{2n} (n = 1,2,3…) represent n-th round trip echoes propagating in the HTUT insert and those in the polymer, respectively.

The experiments were carried out using a 150-ton injection molding machine. A material employed was an injection grade polycarbonate, which is amorphous thermoplastic. The molding conditions were set as follows: melt temperature was 320°C; mold temperature was 120°C; and injection (plunger) speed was 110 mm/sec. The total cycle time was about 30 sec. The data acquisition rate was 10 Hz unless particularly mentioned.

3.2. *Ultrasonic Measurement Results*

Figure 14 shows a typical result of acquired signals with the UT1 during the injection process. Although the signals were acquired during whole cycle (30 sec) in time delay range from 4 μsec to 24 μsec covering from the L^1 to L^5 echoes, only the signals at the beginning (from 4 sec to 9 sec) and the end (from 22 sec to 27 sec) of the process in the time delay range from 4 μsec to 9.6 μsec are shown in Fig. 14. One can see the L^1 and L^2 echoes reflected at the insert-polymer or -air interface depending if the polymer existed at the UT location or not. S^1 represents the first round trip shear wave echo reflected at the insert-air or -polymer interface. When the polymer melt arrived at the UT1 location at the process time of 5.8 sec, the L_2 and L_4 echoes propagating in the polymer started to appear. At this moment, the L_6 and L_8 echoes were out of the time delay range shown in Fig. 14. At the process time of 25.5 sec, the L_2, L_4, L_6 and L_8 echoes vanished, which indicates that the molded part was detached from the UT sensor insert.

In order to investigate further the correlation between ultrasonic signals observed and the process cycle, the amplitude values of the L^4 and L_2 echoes with respect to the process time were obtained using the signals measured with the UT1. Here the L^4 echo was chosen instead of the L^1 echo since, in principle; higher-order round trip echoes of L^n can lead to higher sensitivity to the UT insert-polymer interface condition because of more reflections at this interface [2]. The results are presented in Fig. 15a. Temperature and pressure values measured by the Kistler sensor are also presented in Fig. 15b. It is noted that the ultrasonic data shown in Fig. 15a were acquired with a different cycle from that for the data shown in Fig.14. Hence the process time at which the L_2 echo appeared or disappeared in Fig.15a was different from that shown in Fig. 14. At the process time A, the polymer melt arrived at the UT1 and the Kistler sensor locations simultaneously. The amplitude of the L^4 echo suddenly decreased and the L_2 echo started to appear as seen in Fig. 15a due to the fact that a part of the ultrasonic energy was transmitted into the polymer through the UT1 insert-polymer interface and then reflected back at the polymer-immobile mold interface.

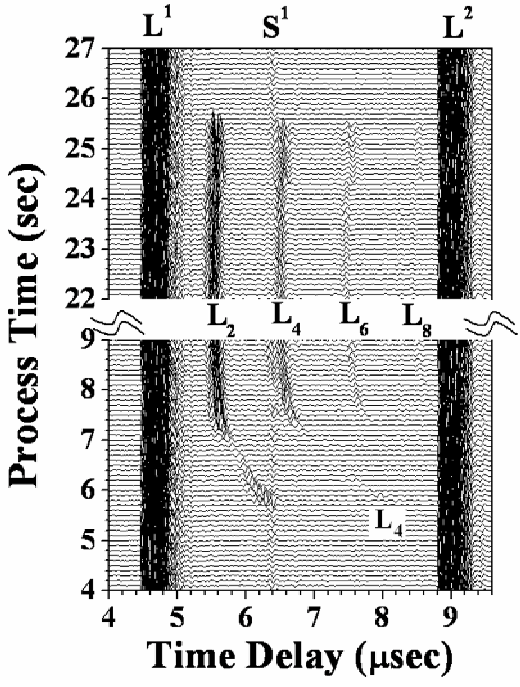
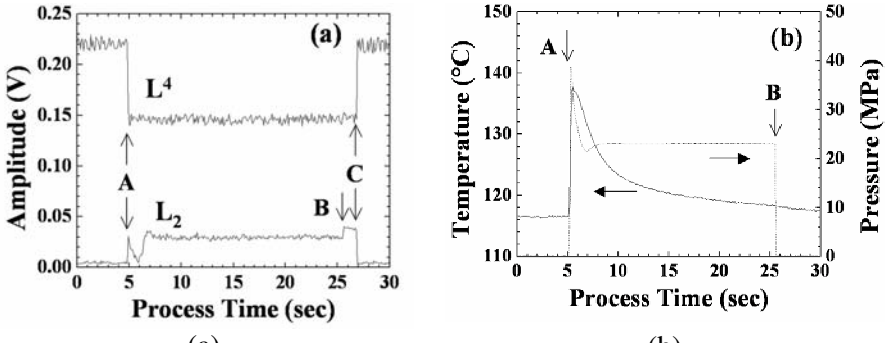


Figure 14. Typical signals measured with the UT1 during injection process.



(a) (b)

Figure 15. (a) Amplitude variations of L^4 and L_2 echoes measured with UT1, and (b) temperature (solid line) and pressure (dotted line) variation measured with Kistler sensor during injection process. Arrows A, B and C indicate the time for flow front arrival at UT1 location, mold open and part ejection, respectively.

At the time A, the pressure and temperature raised up sharply from zero to 40 MPa and from 116°C to137°C, respectively, as shown in Fig. 15b. It should be noted that the pressure sensor has a threshold in its response, resulting in a slight time delay comparing with the ultrasonic data (~0.1 sec with our

experimental conditions). In addition, the measured temperature was smaller than the melt temperature of 320°C since it was not the melt temperature but "contact (or surface)" temperature of the part, which is significantly affected by the interface condition between the sensor and the polymer.

Just after the time A, the amplitude of the L_2 echoes decreased and increased during a few seconds in the time range between 5 sec and 7 sec, which might reflect the variation of ultrasonic attenuation in the polymer because of its solidification. This will be discussed further later. At the process time B, the mold was opened and the molded part was detached from the immobile mold as shown in Fig. 10(iii). Consequently, the amplitude of the L_2 echo increased due to the almost total reflection at the polymer-air interface, and the pressure dropped to zero. In addition, the phase change of 180° was observed on the L_2 echo at the time B since acoustic impedance of the polymer is smaller than that of the steel mold but larger than that of air, resulting in the change of sign of reflection coefficient at this interface from plus to minus. Such phase reverse of the L_2 echo is also a good indication of the part detachment from the mmobile mold. At the time C, the amplitude of the L^4 echo recovered to its initial value and the L_2 echo disappeared, indicating that the part was detached from the UT1 insert due to the part ejection as shown in Fig. 10(iv). Hence, the times of the flow front arrival, mold open and part detachment were clearly observed on the ultrasonic echoes as indicated by the arrows A, B and C, respectively, in Fig. 15a.

3.3. *Discussion*

3.3.1. *Flow Front Arrival and Flow Speed Monitoring*

During the filling stage, as shown in Fig. 10(i), the flow front advancement and flow front velocity are critical information. The flow front position can be used to control the plunger speed so as to allow smooth transition from the filling to the packing/holding stages, as shown in Fig. 10(ii), to avoid part flashing and mold damage due to high impact. Figure 16 shows the L^4 echoes measured with the UT3 and UT4. In this experiment, the signals were acquired every 3 msec (333 Hz) for higher time resolution. As mentioned previously, the amplitude of the L^4 echoes steeply decreased at the time A for UT3 and A' for UT4, indicating flow front arrival at each UT location. Therefore, the average flow front velocity, V_f, between the UT3 and UT4 was calculate to be 306 mm/sec by using $V_f = L/\Delta t_f$, where L (= 34.9 mm) was the distance between UT3 and UT4, and Δt_f (= 114 msec) was the time difference between A and A'.

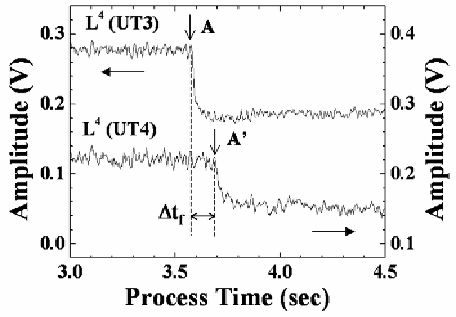


Figure 16. Amplitude variation of L^4 echoes measured with UT3 and UT4 showing the time difference (A-A') of flow front arrival at each UT location.

3.3.2. *Filling Completion Monitoring*

Filling completion of the mold with the materials is the most critical requirement for the molding process since the incomplete part must be rejected. Figure 17 shows one incomplete (#1) and three complete (#2-4) parts molded successively under the same molding conditions. Accidentally the part #1 had defects on both bottom edges as indicated by the arrows in Fig. 17. The volume filling rate for the part #1 was calculated to be 99%. Figure 18 presents the amplitude of the L_2 echoes, obtained during the cycles for the parts #1-4 shown in Fig. 17, with the UT1-4, whose locations are indicated in Fig. 17. Even though the part was filled at UT3 and UT4 locations for the part #1, one can see that the L_2 echoes measured with the UT4 appeared for a few seconds only at the beginning of the cycle. In addition, the amplitude of the L_2 echoes measured with the UT3 gradually decreased to noise level before the mold opened at time B. These are due to the fact that the part detachment at the UT3 and UT4 locations occurred before the mold opened because of shrinkage of the part caused by the lack of enough filling pressure. The L_2 echoes measured with the UT2 for the part #1 appeared in the entire time range between A and B, however the amplitude was a little smaller comparing with those for the part #2-4. It is concluded that the presented ultrasonic method has the ability to monitor the incomplete filling of the part even with filling rate of 99%.

It is noted that the part ejection time C was not observed on the L_2 echoes measured with UT2 and UT3 for the part #2-4, as shown in Fig. 18. This is due to that the part detachment at UT2 and UT3 locations had already occurred at time B because the part in the gate area was weakly pulled towards the immobile mold when the mold was opened, resulting in the slight bending of the part. The part in the gate area was cut from the molded parts after the ejection, hence it is not seen on the parts shown in Fig. 17.

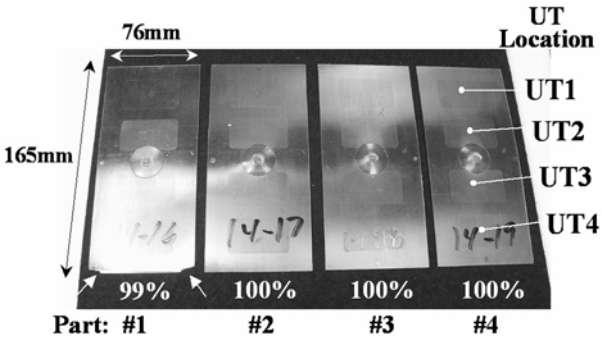


Figure 17. One incomplete (#1) and three complete (#2-4) parts molded successively under the same molding conditions. Filling rate of part #1 was 99%.

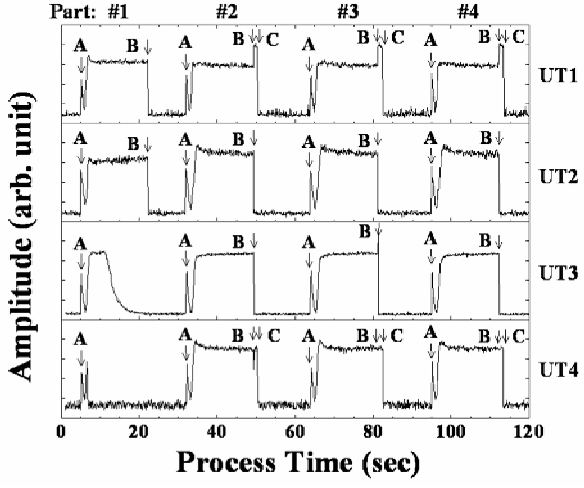


Figure 18. Amplitude variation of L_2 echoes measured during injection cycles for the parts #1-4 shown in Fig. 17.

3.3.3. *Solidification Monitoring*

Ultrasonic velocity and attenuation inside the molded parts are strongly related to the physical properties of the molded part. Therefore, solidification of the molded part may be monitored with the velocity and attenuation during the process. They can be determined using the time delay and amplitude variations of the L_{2n} echoes propagating in the polymer, as shown in Fig. 14, provided that the echoes have sufficient SNR. The ultrasonic velocity, V, is determined by $V = 2l / (t_4 - t_2)$, where l is the thickness of the mold cavity measured at the UT location by a micrometer, and t_2 and t_4 are the time delay of the L_2 and L_4 echoes, respectively. It is difficult to estimate the acoustic properties such as acoustic impedance of the

UT insert and the mold during the process due to their temperature change, which is necessary to calculate the ultrasonic attenuation in the molded part. Hence the attenuation, α, of the amplitude of the L_{2n} echo, which is associated with the ultrasonic attenuation in the molded part, is calculated by $\alpha = 10\ \{\log(|L_2|/|L_4|)\}/l$, where $|L_2|$ and $|L_4|$ are the amplitude of the L_2 and L_4 echoes, respectively.

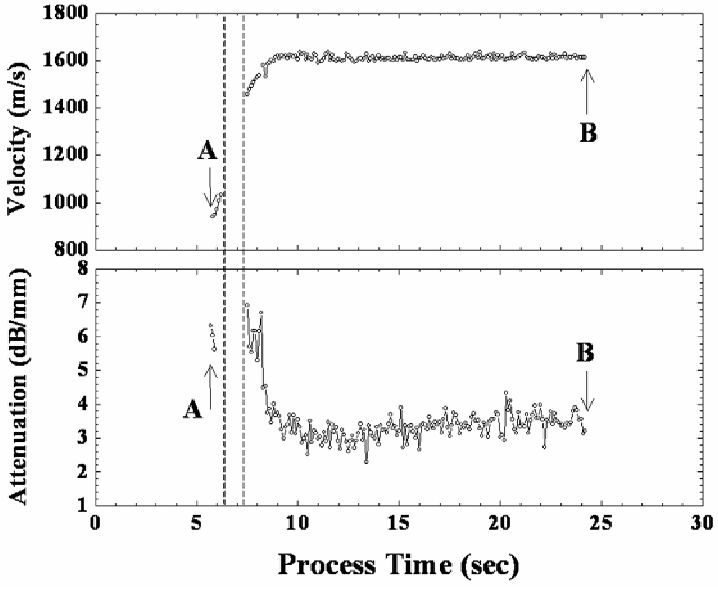


Figure 19. Ultrasonic velocity in the molded part and attenuation of L_{2n} echo obtained using L_2 and L_4 echoes measured with UT2.

Figure 19 presents the ultrasonic velocity and the attenuation of the L_{2n} echo obtained at the UT2 location. At the process time between 6.3 sec and 7.2 sec, indicated by the dotted lines in Fig. 19, the data are missed, since the ultrasonic attenuation in the polymer was so high due to the solidification [16] that the L_4 echoes didn't have sufficient SNR to determine its time delay and amplitude. The measured velocity inside the polymer varied from 950 m/s to 1600 m/s and the attenuation from 6 dB/mm to 3 dB/mm due to the variation of the elastic properties of the part, such as elastic constants, viscosity and density, indicating a transformation of the polymer from the molten state to solid state [5].

4. Conclusions

Real-time and in-line ultrasonic diagnostics of polymer extrusion and injection molding processes together with the high temperature ultrasonic probes used

were presented. For polymer extrusion, clad ultrasonic rod probes consisting of a core and a cladding fabricated by a thermal spray technique were used [20]. This type of probe has the advantage of high SNR which is necessary for diagnostics to achieve high sensitivity. For the demonstration two ultrasonic probes operated in the transmission mode and installed at the exit die were used to monitor the melt quality of extruded PP and $CaCO_3$ filler concentration and dispersion in PP. For the melt quality the fluctuations of the transmitted ultrasonic signals were used to quantify the amount of un-melted pellets. The monitored results showed that the ultrasonic diagnostics system could be used to maximize the production throughput and in the meantime obtain the accepted extrudate quality. In addition, ultrasonic attenuation variation was also demonstrated to be able to monitor the concentration amount and dispersion of $CaCO_3$ in PP.

Thick film HTUTs were successfully deposited on steel substrates by a spray technique [18,24] to develop integrated ultrasonic sensor inserts for polymer IM process diagnostics. The center frequency of ultrasonic signals generated by these HTUTs was around 10 MHz and a SNR was more than 30 dB in the pulse-echo mode at 500°C. For IM process diagnostics, the polymer behavior in the mold cavity was monitored using the multiple HTUTs sensor inserts embedded in the mold insert of proper size at different locations. Multiple mold inserts of various sizes are commonly used by injection molders, in particular, for multi-cavity molding processes. Utilization of sensor inserts in the mold insert allows us to monitor the injection process with molds of all sizes. The material employed was injection grade polycarbonate and the molded part had a rectangular shape. Flow front arrival and average flow front velocity of the polymer inside the mold cavity was determined using the amplitude variation of the echo reflected at the UT sensor insert-air or -polymer interface. This echo was also used to detect the time of the opening of the mold. The amplitude and/or phase variations of the echoes propagating in the molded part were used to detect the part ejection and to monitor the filling completion of the mold cavity with the polymer. The amplitude of the echoes was sensitive to the part detachment from the mold so that uncompleted part with filling rate of 99% could be detected. The solidification monitoring of the molded part was performed by measuring the ultrasonic velocity and attenuation in the molded part using the multiple echoes propagating in the part. The variation of ultrasonic velocity and attenuation were reflecting the elastic properties of the part, such as elastic constants, viscosity and density, during the solidification.

Acknowledgments

Authors would like to thank M. Kobayashi, C.-C. Cheng, Y. Simard, H. Hebert, R. Gendron, M.M. Dumoulin, L. Piché, A. Hamel, A. Derdouri and J.-F. Moisan for their technical assistance.

References

1. C. Rauwendaal, *Polymer Extrusion, 4th Ed.*, (Hanser Gardner, New York, 2001).
2. C.I. Chung, *Extrusion of Polymers: Theory and Practice*, (Hanser Gardner, Cincinnati, Ohio, 2000).
3. D.V. Rosato, D.V. Rosato and G. Marlene, Eds., *Injection Molding Handbook (3rd Ed.)*, (Sparinger-Verlag, Berlin, 2000).
4. T.A. Osswald, L.-S. Turnq and P.J. Gramann, Eds., *Injection Molding Handbook*, (Hanser Gardner, Cincinnati, Ohio, 2001).
5. L. Piché, A. Hamel, R. Gendron, M.M. Dumoulin and J. Tatibouët, US Patent no. 5,433,112, July 18, 1995.
6. R. Gendron, M.M. Dumoulin, J. Tatibouët. L. Piche and A. Hamel, Proc. SPE ANTEC, 2256 (1993).
7. R. Gendron, L.E. Daigneault, J. Tatibouët and M.M.Dumoulin, *Advances in Polymer Technology*, **15**, 111 (1996).
8. E.C. Brown, P. Olley; and P.D. Coates, Proc. SPE ANTEC, 1042 (1997).
9. E.C. Brown, P. Olley, P.D. Coates, *Plastics, Rubber and Composites*, **29**, 3 (2000).
10. J. Tatibouët and M.A. Huneault, *Inter. Polymer Processing*, **XVII**, 49 (2002).
11. Z. Sun, C.-K. Jen, C.-K. Shih and D. Denelsbeck, *Polymer Eng. and Science*, **43**, 102 (2003).
12. A. Sahnoune, R. Gendron, M.A. Huneault, J. Tatibouët and L. Piché, *J. Cellular Plastics*, **37**, 429 (2001).
13. J.Tatibouët and R.Gendron, *Journal of Cellular Plastics*, **41**, 57 (2005).
14. M. Konno, A. Cui, N. Nishiwaki, S. Hori, Proc. SPE ANTEC, 2798 (1993).
15. C.L. Thomas, A.A. Tseng, A.J. Bur, J.L. Rose, *Advances in Poly. Tech*., **15**, 151 (1996).
16. H. Wang, B. Cao, C.-K. Jen, K.T. Nguyen and M. Viens, *Polym. Eng. Sci*., **37**, 363 (1997).
17. S.-S.L. Wen, C.-K. Jen and K.-T. Nguyen, *Int'l Polymer Processing*, **XIV**, 175 (1999).
18. M. Kobayashi, Y. Ono, C.-K, Jen, and C.-C. Cheng, *IEEE Sensors J*., **6,** 55 (2006).
19. B.R. Whiteside, E.C. Brown, Y. Ono, C.-K. Jen and P.D. Coates, *Plastics, Rubbers and Composites* **34**, 387 (2005).
20. C.-K. Jen, J.-G. Legoux and L. Parent, *NDT & E* Int'l, **33**, 145 (2000).
21. Y. Ono, M. Kobayashi, C.-K. Jen, C.-C. Cheng, A. Derdouri, and Y. Simard, Proc. SPE ANTEC, 556 (2004).
22. L. Mulvaney-Johnson, C.-C. Cheng, Y. Ono, E.C. Brown, C.-K. Jen and P.D. Coates, Private communication, 2007.
23. C.-C. Cheng, Y. Ono, B.R. Whiteside, E.C. Brown, C.-K. Jen and P.D. Coates, *Int'l Polymer Processing*, **XXII**, 2007 (in press).
24. M. Kobayashi and C.-K. Jen, *Smart Materials and Struct.* **13**, 951 (2004).

This page intentionally left blank

CHAPTER 4

LASER-ULTRASONICS: PRINCIPLES AND INDUSTRIAL APPLICATIONS

Jean-Pierre Monchalin

Industrial Materials Institute, National Research Council of Canada
75, de Mortagne Blvd, Boucherville, Québec, J4B 6Y4, Canada
jean-pierre.monchalin@cnrc-nrc.gc.ca

A broad overview of the field of laser-ultrasonics is presented. This overview draws from developments at the Industrial Materials Institute of the National Research Council of Canada as well as elsewhere. The principles of generation and detection are presented, stressing a few key characteristics of laser-ultrasonics: the material is actually the emitting transducer and transduction is made by light, thus eliminating any contact. These features carry both advantages and limitations that are explained. Another feature, which has been an impediment, is actually the complexity of the "laser-ultrasonic transducer", but in spite of this complexity, it can be made very reliable for use in severe industrial environments. It also can be very cost effective for a number of applications. Four applications that are now used in industry are presented: the inspection of polymer matrix composites used in aerospace, the thickness gauging of hot steel tubing in production and the measurement and characterization of thin layers in microelectronics by two different approaches. Technological aspects, such as interferometer design, detection lasers and others are also discussed. As an overall conclusion, laser-ultrasonics that was for a long time a laboratory curiosity has definitely now made its transition to industry. Nevertheless, developments should continue to improve performance, to make it well adapted to specific inspection or characterization tasks and more affordable.

1. Introduction

Ultrasonics is a widely used technique in industry for thickness gauging, flaw detection and to a limited but increasing extent for material characterization. Ultrasonics is relatively simple to apply, the key element being a piezoelectric

transducer often used for emission and reception. Ultrasonics is also an active area of research with efforts pursued in many different directions. However, ultrasonics as it is usually applied, has several constraints or limitations, the main one being the need of contact or fluid medium for ultrasound coupling to the tested part. Others are the limited bandwidth of piezoelectric transducers and the sensitivity to the phase of the ultrasonic wavefront across the transducer element, which makes the inspection of parts with acute curvature difficult. As recognized by researchers for many years, these limitations could be eliminated by using lasers and light for generating and detecting ultrasound, so called laser-ultrasound, laser-based ultrasound or laser-ultrasonics. Progress to take this concept out of the laboratory into industry has been quite slow, but during the last few years, we have seen the start of real industrial use, while observing at the same time increasing research activities.

The generation of ultrasound with lasers actually nearly goes back to the invention of the first laser, the ruby laser[1], which was used to generate ultrasound or shock waves in materials[2,3]. Optical detection of ultrasound has been on the other hand developing quite slowly, the type of interferometer typically used by researchers being not applicable to industrial conditions.

We are presenting here a broad overview of the field of laser-ultrasonics encompassing basics and applications. This overview draws from accomplishments and results obtained at the Industrial Materials Institute of the National Research Council of Canada (IMI/NRC) as well as elsewhere. The principles of generation and detection are first outlined, stressing a few key characteristics of laser-ultrasonics: transduction is made by light, thus eliminating any contact and the material is actually the emitting transducer. These features carry both advantages and limitations that are explained. Another feature, which has been actually an impediment, is the complexity of the "laser-ultrasonic transducer", which is sketched in Fig. 1. As shown in Fig. 1, such a transducer has typically three basic elements, a generation laser, a detection laser and an interferometer, followed by data acquisition and processing electronic hardware. In spite of this complexity, such a transducer can be made very reliable for use in severe industrial environments, as described below. It also can be made very cost effective for a number of applications. Four applications that have now become sufficiently mature for use in industry are described: the inspection of polymer matrix composites used in the aerospace industry, the wall thickness gauging of hot steel tubing in production and the characterization of very thin layers used in the microelectronic industry by two different approaches. Other potential applications that are now being studied in the laboratory or have been demonstrated in industrial conditions are also indicated.

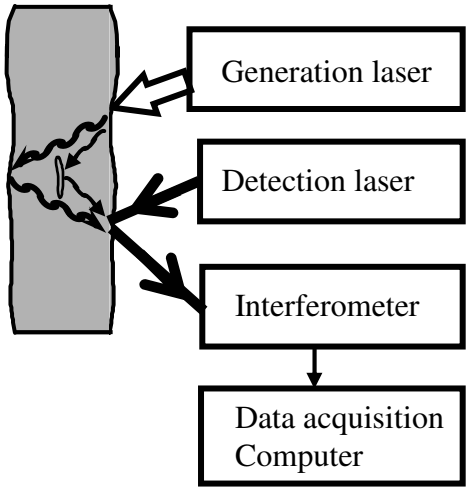


Figure 1. Sketch of the principle of laser-ultrasonics or of a "laser-ultrasonic transducer".

2. Laser Generation of Ultrasound

There are essentially two kinds of mechanisms for generating ultrasound, the first one is perfectly nondestructive and is based on a thermoelastic mechanism (see Fig. 1) while the second one is invasive and is based on the ablation of the sample or on the vaporization of some surface layer. Independently of the mechanism used, it should be recognized that the material is actually the emitting transducer of ultrasound. This distinguishes laser-ultrasonics from piezoelectric-based conventional ultrasonics in which the source of ultrasound is a piezoelectric element separated from the tested material. This feature has important consequences, some beneficial and others detrimental.

2.1. *Thermoelastic Generation*

The principle is the following: laser light is absorbed to some depth inside the material releasing heat locally; the heated region then expands producing a strain and a corresponding stress that is the source of waves propagating in the material or at its surface. Each heated point inside the material is a center of expansion to which 3 orthogonal force dipoles can be associated[4,5]. When the surface is free and when light penetration is very small compared to the acoustic wavelength (which is often the case in metals), the dipole normal to the surface cancels leaving only the in-plane dipoles. This situation has been extensively studied theoretically and experimentally[4–8]. In the near field, the displacement signal

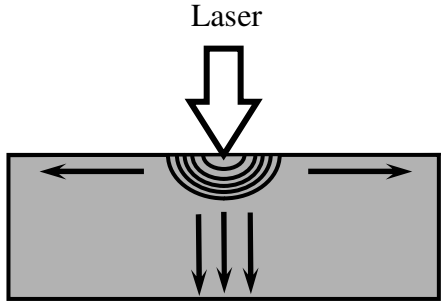


Figure 2. Thermoelastic generation.

appearing on the opposite surface of a plate sample consists of a first depression step corresponding to the longitudinal wave arrival followed by a second one that corresponds to the shear wave arrival. In the far field, the emission pattern for the longitudinal and the shear wave presents symmetrical lobes inclined strongly off the normal (the order of 30° for shear waves to more than 60° for longitudinal waves) and there is no emission along the normal[9]. Note that it is possible to modify the directivity of emission in the far field by having an ensemble of point sources phased in such a way that the various contributions add up at a particular location or along a particular direction. This phased array source can be implemented by sending the generation beam through several optical fibers of lengths chosen to give proper time delays[10]. This method is not very flexible, so the use of separate lasers triggered at proper times has been explored[11], although such a system appears too complex and too costly to be used in practice. Another scheme worth noting is the use of a ring illumination source (practically realized by focusing with a conical lens or axicon), which has the effect of concentrating energy along the axis normal to the surface and passing through the center of the ring (i.e. in other words to focus emitted ultrasound along this axis)[12].

In order to have emission along the normal to the surface in the far field, which is desirable for many applications, there should be some light penetration resulting in a thermoelastic source that is buried underneath the surface. The effect of light penetration can easily be understood by considering the laser heated zone as an ensemble of slices, as shown in Fig. 3a. As shown in Fig. 3b, each slice gives at an observation point inside the medium a displacement compression wavefront first, followed by a displacement rarefaction wavefront after reflection by the free surface and delayed by $2d/V_L$, where d is the slice depth and V_L the longitudinal velocity. These two displacements have opposite polarity but the same magnitude and rise time (equal to the laser pulse duration τ). The actual displacement pulse observed in the medium results from the sum

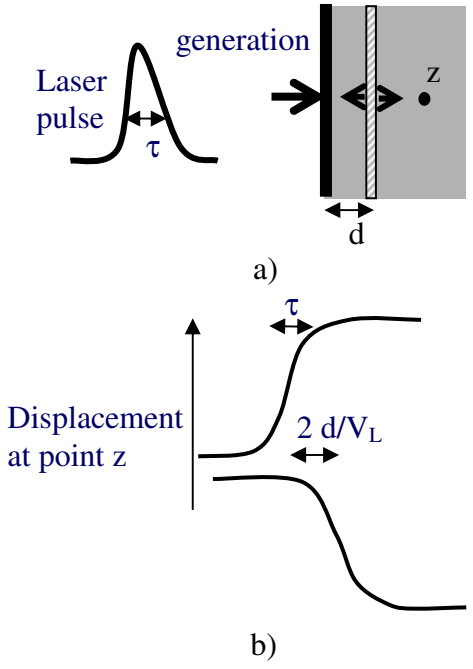


Figure 3. Thermoelastic generation with penetrating light. Fig. 3a shows schematically a laser pulse of duration τ and a layer at depth d heated by this laser pulse in a given medium. An observation point is shown in this medium at depth z. Fig. 3b shows schematically the two displacement signals observed at point z, one being associated to a compression and arriving directly at z and the other one being associated to a rarefaction and reaching z after reflection by the free surface. The actual displacement at z results from the sum of these two signals.

from these two contributions. Therefore, when penetration is very small the displacement pulse is nearly zero, which is the case for metals at frequencies in the range of 1 MHz to about 1 GHz (but not at higher frequencies or for very short pulses, when the typical penetration of 5 to 10 nm for metals becomes comparable to the acoustic wavelength). When penetration increases, the displacement pulse starts to increase and its shape is found to follow proportionally that of the laser pulse. With further increase of light penetration, when the penetration depth becomes about τ V_L, the displacement pulse amplitude saturates and its width increases.

The qualitative explanation above can be made easily quantitative and an analytical expression can be readily obtained if the laser pulse shape can be described analytically[13–15]. Further to this unidimentional model, 3-D cases have been treated assuming axisymetry[16] and an othotropic medium[17]. Light penetration has been experimentally studied in detail with colored glasses

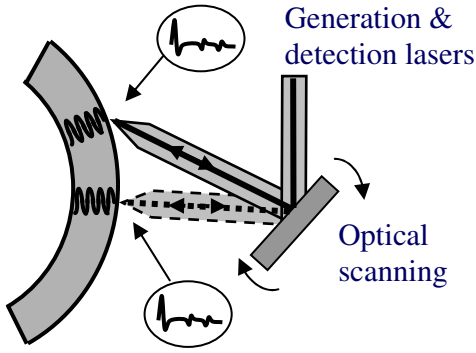


Figure 4. Effect of penetrating laser light: generation occurs perpendicularly as if a conventional piezoelectric transducer was moved over the surface.

giving a wide range of penetration depths[15,18]. Light penetration is very important in practice since its gives a piston source at the surface of the material emitting normally propagating longitudinal waves, independently of the surface curvature and of the orientation of the laser beam. This is at the basis of the inspection of polymer-matrix composites, which is one of the industrial applications of laser-ultrasonics (see Fig. 4).

Enhancement of generation along the normal also occurs when the absorbing material is covered by a transparent layer[19]. In this case, as explained above for the penetrating light source, there are also two delayed and opposite polarity displacement wavefronts that do not cancel if the layer is sufficiently thick. This occurs even if the material has practically no light penetration (metals).

Surface waves and plate waves can also be generated efficiently and in a very versatile manner. When the laser beam is focused to a small circular spot, a surface wave with a cylindrical symmetry is emitted from this spot. Its amplitude has a maximum when the laser pulse duration is about D/V_R where D is the spot diameter and V_R the Rayleigh velocity of the material[20]. Detailed theoretical analysis has been performed and verified experimentally[21]. Good directivity is obtained by focusing the beam with a cylindrical lens to get a line source[22]. More complicated patterns can even be used, such as an array of lines, giving narrower band emission but having the advantage to distribute the laser energy over a broader area, so ablation and surface damage can be avoided. A chirp array has also been used in combination with a matched pattern used for detection[23], as well as a pattern obtained by using a computer-generated hologram or a spatial light modulator[24]. A converging circular Rayleigh surface wave giving very strong displacement at the center of convergence can be readily obtained by using, in addition to the conventional spherical lens, an axicon (conical lens)[25].

Of course when the thickness of the sample becomes of the order of the Rayleigh wavelength or less than it, dispersion effects of plate waves should be taken into account. When a line source is used, emission can be strongly enhanced by sweeping the laser line over the surface of the sample (for example by reflecting the laser beam off a fast rotating mirror)[26]: enhancement occurs when the sweeping velocity of the laser beam is equal to the Rayleigh velocity V_R. With a grating source, strong enhancement is observed when the grating is made to move with a velocity equal also to the Rayleigh velocity[26]. Such a moving grating can be obtained by making two beams issued from the same generation laser to intersect at the surface of the sample, one being frequency shifted by an offset f (for example by sending it through an acousto-optic shifter). By reference to Fig. 5, the enhancement condition is: $\lambda/2\sin\theta = V_R/f$, where λ is the optical wavelength and θ the angle between the laser beams and the normal (note that $(f\lambda)/(2\sin\theta)$ is the grating velocity).

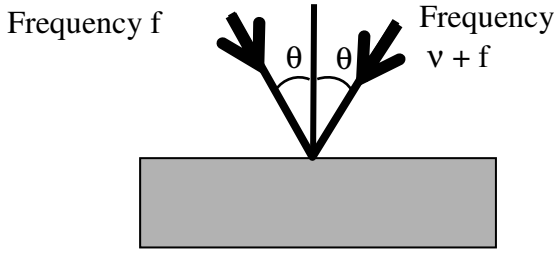


Figure 5. Rayleigh surface wave enhancement by moving line grating.

2.2. *Generation by Ablation or Vaporization*

If one increases the energy density, particularly for small light penetration (metals), one reaches the threshold where the surface starts to melt and then to get vaporized. At this point, matter is ejected from the surface and through various physical processes this vapor and the surrounding air is ionized, thus producing a plasma plume that expands away from the laser spot on the surface (sketched in Fig. 6). On the opposite surface of a plate-like sample, one first observes a surface elevation spike of duration of the order of the laser pulse duration, which is associated with the recoil effect produced by the matter blown off the surface. This displacement then continues all the time the plasma applies a pressure on the surface and diminishes with plasma expansion and cooling. Therefore, depending upon the energy density, the phenomena taking place goes from solely a vaporization effect for which the ultrasonic pulse has duration of

the order of the laser pulse to a strong plasma regime with a much longer duration[4–6]. A similar vaporization effect occurs also when the material is covered by a thin absorbing layer, which is blown off, leaving the substrate underneath substantially unaffected if the energy density is below some threshold. In the strong plasma regime on the other hand, a crater mark is left on the surface (as sketched in Fig. 6). It should also be noted that in addition of vaporization and plasma contributions, there is always some thermoelastic contribution, which is clearly or barely noticeable depending upon the regime.

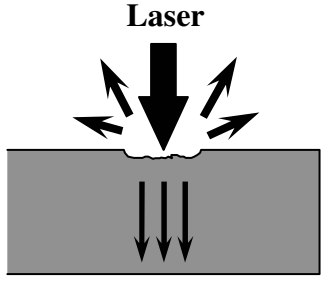


Figure 6. Generation by ablation or vaporization. In the case of strong ablation, a crater is left on the surface.

The generation by ablation has not been modeled as well as thermoelastic generation, for which precise computer codes have been written. Some modeling has however been performed[27,28,29.] The vaporization regime or ablation in vacuum can be modeled relatively easily[30]. An approximate description has been obtained by assuming that the effect is represented by a surface force pointing inwards with appropriate duration[4,5]. When this surface force is combined with a thermoelastic dipole, all the generation conditions can actually be described. This approach also allows predicting the far field patterns[4,5] that have been experimentally observed.

3. Laser Detection of Ultrasound

To detect ultrasound, the surface is illuminated by a laser beam, continuous or of pulse duration sufficiently long to capture all the ultrasonic signal of interest. Scattered or reflected light is then collected by an optical receiver that could be either based on an interferometer or on a non-interferometric device. Practically useful non-interferometric schemes are based on the detection of the change of reflectivity produced by the ultrasonic strain or the knife-edge technique[31]. The detection of the change of reflectivity, which is of the order of the ultrasonic

strain and small (in the 10^{-4} to 10^{-6} range) is particularly useful at high frequencies since strain increases proportionally to frequency for a given displacement. The knife-edge technique on the other hand is based on monitoring with a knife edge positioned in front of a detector or with a position sensitive detector the tilt produced by the ultrasonic surface ripple. This very simple to implement and quite sensitive technique, is however limited to polished surfaces and laboratory experimentation.

Regarding interferometric techniques, it is first useful to recall the effect of the ultrasonic surface motion on the scattered or reflected light. A surface displacement $\delta(t)$ produces a shift $4\pi\delta(t)/\lambda$ on the phase $2\pi\nu t$ of the backscattered or reflected light field, where ν and λ are the optical frequency and wavelength, respectively. This phase can be readily rewritten to show that the instantaneous frequency is changed to $\nu(1 - 2\ v(t)/c)$, which expresses the well known Doppler effect and where $v(t)$ is the ultrasonic surface velocity ($v = d\delta(t)/dt$). Further, assuming surface motion at a frequency f, simple trigonometric manipulation shows that the backscattered or reflected field includes 3 terms: a central carrier at the laser frequency ν plus two optical sidebands at frequencies $\nu + f$ and $\nu - f$. The ratio of the field amplitudes between the sidebands and the central carrier is $2\pi U/\lambda$, where U is the ultrasonic surface displacement amplitude. When the surface motion is not a single frequency, the sidebands are broadened into two symmetric lobes on both sides of the carrier optical frequency ν. These three descriptions of the effect of the surface motion are equivalent and any one could be used to interpret any interferometric detection scheme.

Interferometric detection is based on the conversion by the interferometer of the phase or frequency modulation produced by the surface motion into an intensity modulation, which is detected by the optical detector, as schematically represented in Fig. 7[31,32]. Regarding sensitivity of interferometric detection, it should be noted that there is a fundamental limitation associated with the nature of light as an ensemble of discrete particles. The minimum displacement that can be detected is a function of the number of collected photons and is given by the following formula: $\delta_{lim} = (\lambda/4\pi)\ (B\ h\nu/\ 2\eta P_0)^{1/2}$, where h is the Plank's constant, η the quantum efficiency of the detector, B the electronic bandwidth and P_0 the collected power[33]. This formula assumes optimum conditions: the optical field scattered by the surface coherently and uniformly interferes with a much more intense reference wave (without phase or frequency modulation or optical sidebands) directly derived from the detection laser, i.e. there should be no speckle effect and perfectly matched signal and reference wavefronts. This formula also assumes a quadrature condition between signal and reference. For

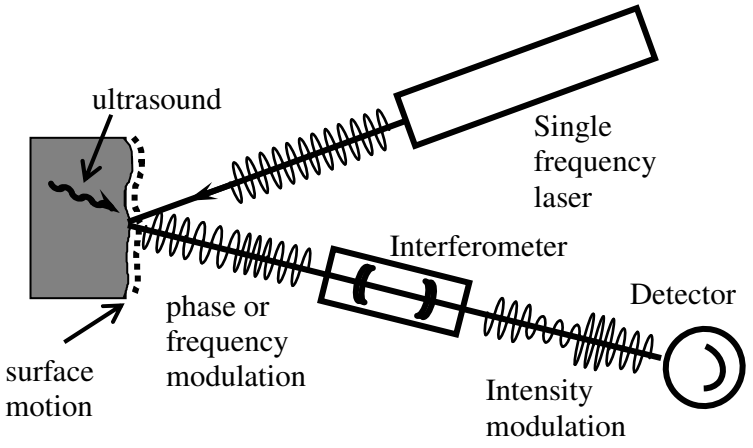


Figure 7. Principle of optical detection of ultrasound with an interferometer: the interfrometer converts the phase or frequency modulation produced by the surface motion in an intensity modulation.

example, for $\lambda = 1.06$ μm and $\eta = 0.9$, the ultimate sensitivity is 3×10^{-8} nm $(W/Hz)^{1/2}$ or 0.003 nm for 1 mW received and 10 MHz bandwidth. It should be noted that this sensitivity is the ultimate that can be obtained with the classical light fields produced by all existing lasers, but it can be improved in principle by using squeezed states or other non classical states of light[34]. When working with a speckled signal beam, the sensitivity is reduced well below the ultimate limit unless the interferometer also acts as a signal-to-reference wavefront adapter or one speckle is collected.

A simple Michelson interferometer, sketched in Fig. 8, has been used for many years for laboratory experimentation. Its drawback for use in industry (except in conditions where the surfaces are polished or smooth, such as in the microelectronic industry) is its sensitivity to optical speckle. Any optical device has some capacity to collect scattered light, called its etendue or throughput, which is defined as the product of the surface of the illuminated spot by the solid angle of the input aperture seen from the surface. In the case of the simple Michelson interferometer of Fig. 8, this parameter is small and of the order of the square of the optical wavelength (about 10^{-6} mm^2.sr), which in practice requires to focus the beam onto the surface to collect essentially one speckle[31]. One will note that the surface is actually in this scheme part of the interferometer as one of its mirrors.

The limitation of one speckle detection can be circumvented by using a matrix of detectors that detects several speckles and by combining the various

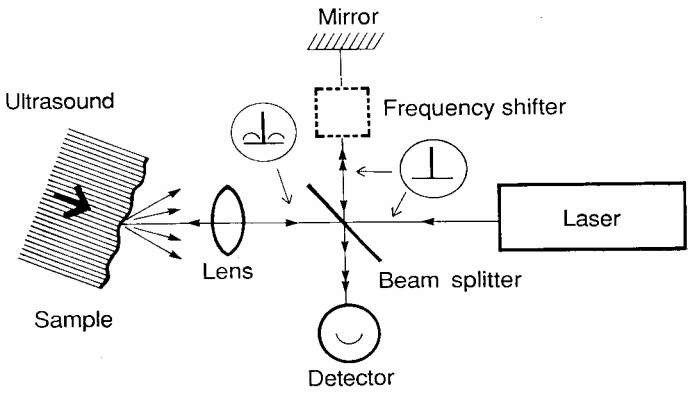


Figure 8. Michelson interferometer: the inserts indicate schematically the optical spectra along the reference and beam paths.

signals from the detectors. A recently reported system has several channels each being based on a quadrature Michelson interferometer[31]. Each channel provides either the square of the ultrasonic displacement or the displacement itself, depending upon the number of detectors used per channel and signal processing[35, 36].

Still, a more practical scheme is to use the Michelson interferometer as a light filter working as a frequency analyzer of the scattered light[37]. Such a filtering action is obtained by giving very different pathlengths to the two arms of the interferometer. The two interfering waves in the interferometer are then given a time-delay (time-delay interferometry), which should be of the order of half the ultrasonic period for optimum sensitivity. For the frequency range of 1 to 20 MHz, this results into a very bulky system, particularly if immersion in a liquid bath is added to increase etendue. A much more practical implementation of time-delay interferometry and of demodulation by an optical filter is based on the use of multiple beam interference in the confocal Fabry-Perot interferometer.

3.1. *Confocal Fabry-Perot Interferometer*

This is a simple optical resonator made of two concave identical mirrors separated by a distance equal to their radius of curvature L_R[38]. As shown in Fig. 9, any incident ray gives 4 rays, two on the transmission side and two in reflection. In excellent approximation each ray retraces its path after multiple reflections in the resonator and the multiple beam interference phenomenon occurs on the 4 output ports. Since this takes place independently of the

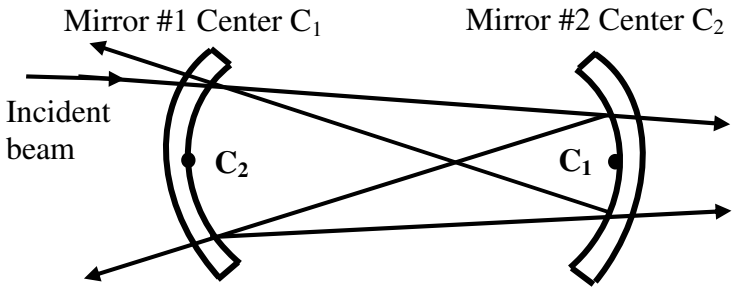


Figure 9. Schematic of the confocal Fabry-Perot interferometer.

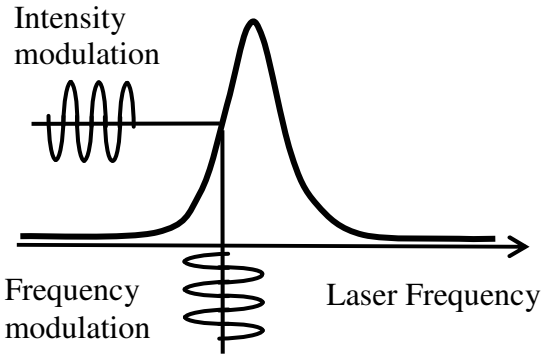


Figure 10. Principle of demodulation with a confocal Fabry-Perot.

direction and height of the incident ray, the system has a large etendue. For example, a meter long cavity with mirrors with 85% reflectivity provides an etendue exceeding that of a collecting optical fiber with core diameter of 1 mm and numerical aperture of 0.35 (etendue equal to 0.4 mm^2.sr).

This system is a resonator that has resonance peaks of width $\Delta\nu_{FP}$, which decreases with increased mirror reflectivity R ($\Delta\nu_{FP} = (c/4L_R)/F$, where c is the speed of light and F the finesse, $F = \pi R/(1 - R^2)$). As shown in Fig. 10, demodulation is realized by locating the detection laser frequency on the slope of one peak[31]. Note that this interpretation, which shows that the system responds as a velocimeter (the frequency modulation of the scattered light is proportional to the surface velocity, as mentioned above) is valid only at low frequencies (frequencies much below $\Delta\nu_{FP}$). When the frequency increases to about $\Delta\nu_{FP}$, the response levels off and has to be calculated[39,40]. Figures 11 and 12 show the calculated responsivities and their experimental verification for two confocal

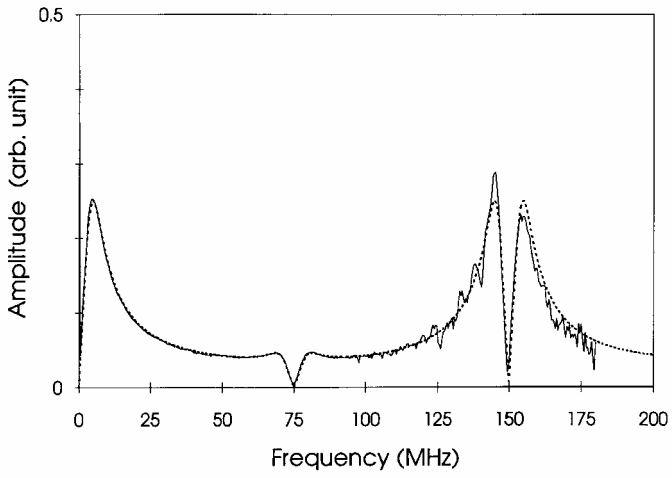


Figure 11. Frequency response of a confocal Fabry-Perot 1 m long used in transmission; full line: experimental data, dash line: calculated data.

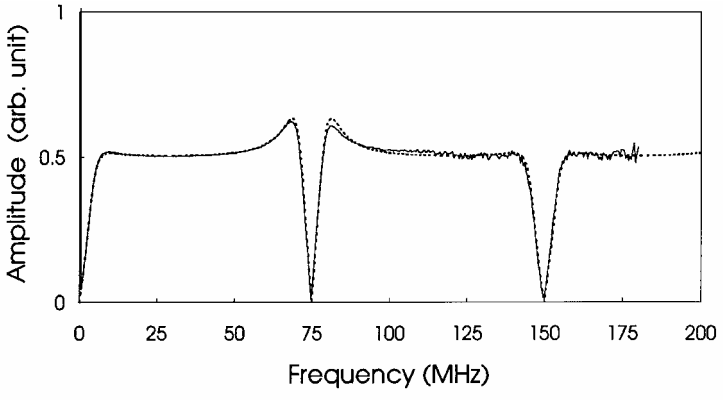


Figure 12. Frequency response of a confocal Fabry-Perot 1 m long used in reflection; full line experimental data, dash line: calculated data.

Fabry-Perot configurations. In one configuration, the system is used in transmission (detector on the transmission side) and in the other one it is used in reflection (detector on the reflection side, also called sidebands stripping scheme)[41]. The reflection scheme has nearly flat frequency response above $\Delta\nu_{FP}$ (except for the periodic drops every $c/4L_R$) and will be used preferably for detection in this high frequency range. It should be noted that the responsivity is practically zero at very low frequencies, which means that this system is

intrinsically insensitive to vibrations, a key advantage for use in industrial environments.

Regarding sensitivity, the minimum detectable displacement is about 4 δ_{lim} at maximum response for the transmission and reflection configurations. When high sensitivity is only required at high frequencies, the back mirror could be made totally reflecting. This leads to improved sensitivity by about a factor two. This can be explained by the fact that the four output beams exiting a confocal cavity, two on the transmission side and another two on the reflection side, giving four independent output ports with multiple interference on each port, are now reduced to two, contributing to an increase of the intensity of interfering terms and consequently of sensitivity. A further improvement by $\sqrt{2}$ can be obtained with a configuration with only a single output port on the reflection side. This can be realized with a totally reflecting back mirror and a front mirror that is totally reflecting over half of its surface, as in the original confocal Fabry-Perot design[33,38]. It should also be noted that the transmission scheme could operate with unpolarized light whereas the use in reflection requires polarizing optics for optimum operation. Therefore often in practice, especially if a large core multimode fiber is used to transmit light to the Fabry-Perot, the transmission configuration gains a sensitivity factor of about $\sqrt{2}$ with respect to the others. If the range of frequencies of interest is between 1 and 15 MHz, which is often the case in nondestructive testing, this configuration will be the one usually selected with a proper choice of mirror reflectivity R and cavity length L_R to give adequate etendue and frequency response.

The main weakness of the Fabry-Perot demodulators is their lack of sensitivity at low ultrasonic frequencies (below 2 MHz), which is circumvented by the devices based on two-wave mixing in photorefractive materials described below. Regarding the reduction of the effect of laser noise, a differential scheme which eliminates or at least minimizes laser intensity noise has been reported[42]. More recently another one, more powerful which eliminates in addition phase noise has been described[43].

3.2. *Photorefractive Two-wave Mixing Interferometer*

In the two-wave mixing approach, wavefront adaptation is performed actively, by opposition to the confocal Fabry-Perot in which adaptation is performed by passive or linear optical components; the technique used is also known as real-time holography. This active wavefront adaptation eliminates the need of an external stabilization device against thermal drift or ambient vibrations, as

required for the confocal Fabry-Perot. The basic setup of the two-wave mixing interferometer is sketched in Fig. 13. A signal beam which acquires phase shift and speckle after reflection on a surface in ultrasonic motion, is mixed in a photorefractive crystal with a pump plane wave to produce a speckle adapted reference wave that propagates in the same direction as the transmitted signal wave and interferes with it. The quadrature condition is provided by passive optical components after the crystal[44,45].

The etendue, determined by the size of the crystal and the angle between the signal and pump beams, is easily made larger than the etendue of the fiber mentioned above. The response of such a device is flat from a low cut-off frequency f_c, which depends on crystal properties and pump intensity, up to the detector cut-off frequency. There are no periodic sensitivity drops like in the confocal Fabry-Perot. With an InP crystal with proper iron doping, operating at 1.06 μm and with the application of an electric field, the sensitivity is about the same as the maximum sensitivity of the confocal Fabry-Perot used in transmission, i.e. the detection limit is about 4 δ_{lim}[46]. Better sensitivity has been demonstrated with a CdTe crystal with Vanadium doping[47], but CdTe photorefractive crystals need further development before becoming a reliable source for such a device. At 1.06 μm, GaAs crystals are also used without an electric field giving a sensitivity reduced by about 2.5 compared to the maximum sensitivity of the confocal Fabry-Perot[45,48].

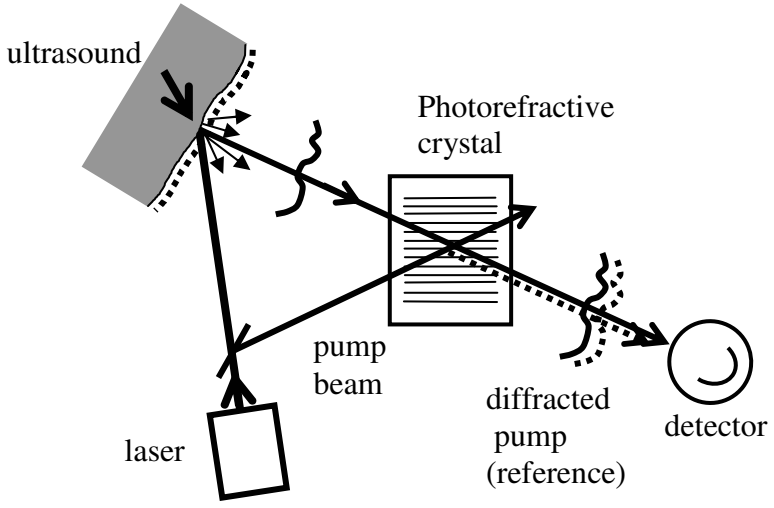


Figure 13. Demodulation with a two-wave mixing photorefractive interferometer.

One important advantage of the photorefractive demodulator with respect to the confocal Fabry-Perot is its better sensitivity at low ultrasonic frequencies (below 1 MHz), thus allowing probing more easily materials with strong ultrasonic attenuation (materials with coarse microstructure, porous or mushy materials…). The system has also the advantage to be easily combined with a differential or balanced scheme (two detectors giving responses to phase modulation of opposite sign), so the noise coming from the laser intensity fluctuations can be eliminated to a large extent[49]. Further, by making the pathlengths from the laser to the crystal along the signal and the pump beams to be sensibly equal, the effect of laser phase noise is eliminated.

Its weakness, in spite of the use of semiconductor photorefractive crystals with high photoconductivity (GaAs, InP, CdTe), is its rather slow response time, i.e. the time needed for the photorefractive grating to be built up or to be erased. This affects the ability of the system to adapt to motions of the probed object that cause a change of the speckle pattern of the scattered light or a change of its frequency by the Doppler effect. This time τ_c is related to the low cut-off frequency f_c by the relation $\tau_c = 1/(2\pi f_c)$. The response time also scales essentially as 1/pumping intensity. Responses times as short as 400 ns are obtained at 1.06 μm with an intensity of about 3 kW/cm^2 in a GaAs crystal without an applied electric field. The application of an electric field, which increases sensitivity, has however the drawback to lengthen the response time. With an InP:Fe crystal under a field of 4.9 kV/cm and a power density of about 100 W/cm^2, a time constant of about 2 μs is obtained[46]. These values are longer than those typically obtained with a confocal Fabry-Perot, which can be estimated as $(4 L_R/c)F$. For a one-meter long Fabry-Perot with a finesse of 10 or a 50 cm long with a finesse of 20, this gives $\tau_c \approx 130$ ns. This means that this system cannot be used as easily as the confocal Fabry-Perot on moving objects. This limitation is illustrated in Fig. 14, which shows the drop of response with the object velocity or Doppler shift (Doppler shift $\Delta\nu = 2V/\lambda$, where V is the projected velocity along the line-of-sight). To be unaffected by a velocity of 1 m/s would require a response time τ_c much less than 1 μs.

It should be noted that the photorefractive system is much more tolerant to motions transverse to the line-of-sight that cause only a change of the speckle pattern without a change of frequency. This is illustrated in Fig. 15. Although the effect of sensitivity reduction by the object motion along the line-of-sight or transverse has not been thoroughly investigated, it can be understood that a displacement of $\lambda/2$ or less during the time τ_c causes grating washout[46], whereas the displacement should be much larger (the size of a speckle) to produce speckle

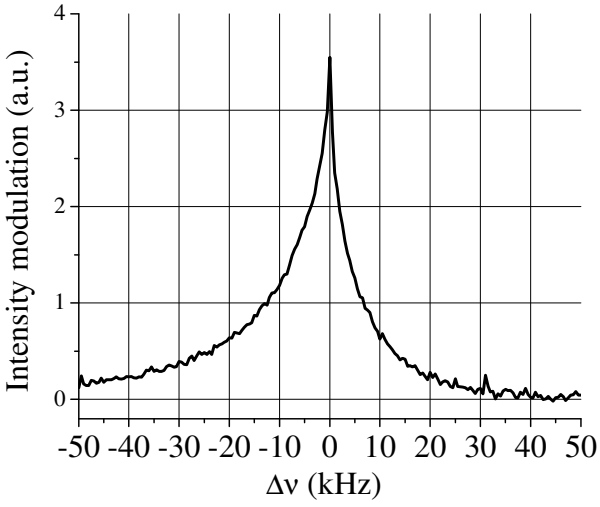


Figure 14. Response of a GaAs photorefractive demodulator to object motion causing a Doppler shift $\Delta\nu$ ($\Delta\nu = 2\ V/\lambda$, where V is the velocity of the object along the line-of-sight and λ is the optical wavelength). The crystal is pumped with 350 mW and the response time is about 35 μs. The beam is phase modulated at 2 MHz to simulate ultrasonic displacement.

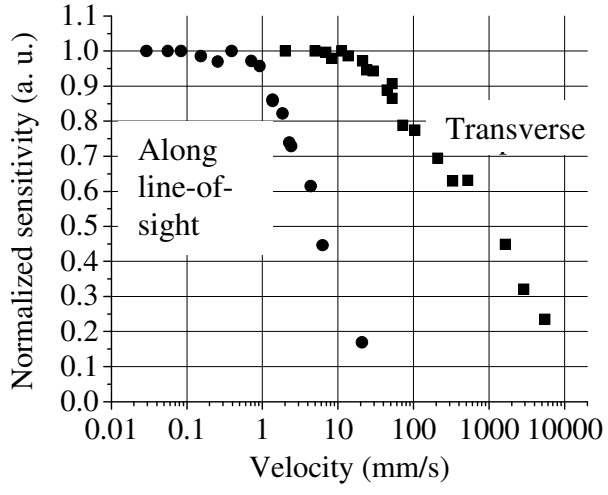


Figure 15. Response of the same GaAs phorefractive demodulator to motion along the line-of-sight and transverse to it.

decorrelation and a corresponding reduction of sensitivity if there is only transverse displacement.

To compensate for the Doppler shift, the frequency of one of the interfering beams can be changed (for example by sending the beam through an acoustooptic

shifter) in such a way that the photorefractive grating becomes stationary in spite of the object motion. An automatic system that performs this compensation has been demonstrated and is based on the tracking of the error signal that appears at the output of the balanced receiver or differential amplifier after the crystal[50].

3.3. *Imaging and Multiplexing with a Photorefractive Two-wave Mixing Interferometer*

Any optical receiver that has a sizable etendue can demodulate over several spots in parallel or over an image. Although this can be demonstrated with the confocal Fabry-Perot (using the Connes type to avoid spurious superposition of images), this is more easily implemented with a photorefractive two-wave mixing interferometer. In the multiplex scheme, as shown in Fig. 16, several spots are projected onto the surface of the tested objects according to an arrangement that fits the application (line, grid array, circle...). Scattered light from these spots is then projected onto the photorefractive crystal, either at distinct locations or at the same location from different directions and is finally received by separate detector elements[51,52]. In this last case, it can be shown that there is essentially no cross talk between the various channels when the pump intensity is much larger than the intensity of any signal beam. Such a configuration can obviously be used to increase the speed or throughput of an inspection task by the number of spots but can also be applied for realizing a focusing detector. For example, if the spots

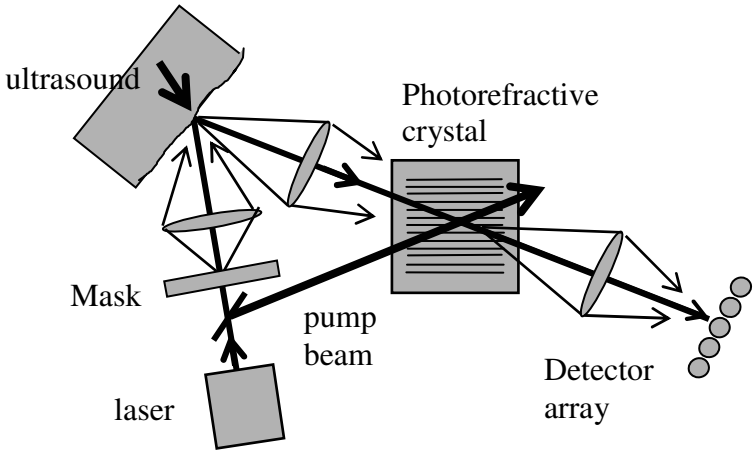


Figure 16. Sketch of an optical multiplexing scheme associated with a photrefractive demodulator. The mask could be replaced by other optical elements such as an holographic grating or a fiber bundle array to project a light spot distribution onto the surface of the object.

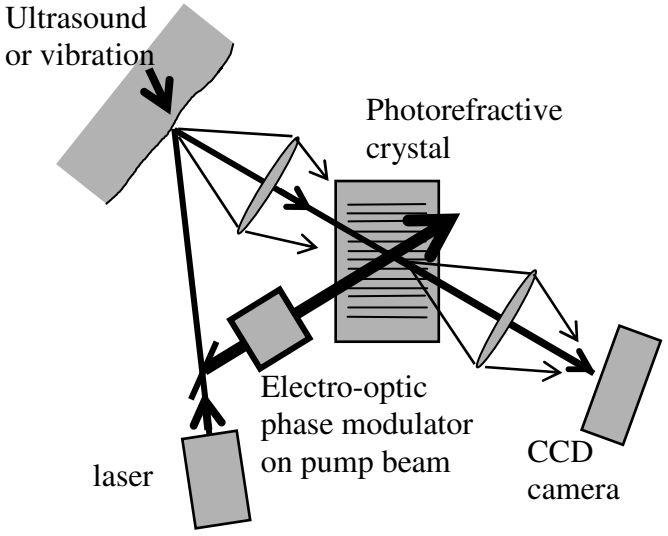


Figure 17. Sketch of full field imaging of vibration or ultrasonic motion with a photorefractive demodulator.

are distributed along a line, the system has a peaked response for a surface wave incident perpendicularly to the line. If the spots are on the other hand located on a circle, the system detects essentially a surface wave originating from the center of the circle.

Full field imaging of vibration or ultrasonic motion at the surface of an object can be obtained with the setup sketched in Fig. 17. The pump beam is strongly phase modulated at a frequency close to that of the surface disturbance to be detected (the difference of frequencies should be less than the photorefractive crystal low cut-off frequency f_c). This modulated pump has essentially the effect to produce inside the crystal a photorefractive grating that is nearly stationary and diffracts a beam that interferes with the transmitted signal beam. Then, at each point across the crystal or across an image of the surface, a signal representative of the surface displacement at this point is obtained and could be seen on a CCD camera[53,54].

3.4. *Detection Laser*

One key element of the detection scheme is the detection laser. It should be high power, since sensitivity increases with power, and should not contribute to any noise in addition to the fundamental photon or shot noise. High power is

particularly needed when the surface is absorbing and detection is at a large distance giving a small collection solid angle. The pulse duration should be sufficiently long to capture all the signal of interest, which means for many applications a duration between 10 and 100 μs. Nd-YAG technology at 1.06 μm, which is known to provide high amplification gain, is particularly suited for realizing such a laser by amplifying a small and very stable Nd-YAG laser oscillator[55]. A suitable oscillator is commercially available with power from 100 mW to about 2 W and is based on a small monolithic cavity pumped by a laser diode[56]. Depending upon the repetition rate, the amplifier could be flashlamp pumped (up to 100 Hz) or diode pumped (above 100 Hz)[57]. Faraday isolators are typically used after the small laser oscillator and between amplification stages to prevent feedback and spurious oscillations in such a high gain system. Peak powers of 1 kW and more are typically obtained with pulse duration of the order of 50 μs.

The effect of the additional noise in laser amplifiers that originates from the Amplified Spontaneous Emission (ASE) has been studied for two possible configurations: the amplifier is located ahead of the probed object (pre-amplification scheme) or after it (post-amplification scheme). Pre-amplification is generally preferable because ASE is attenuated by surface absorption and collection losses in the same proportion as the signal beam, making this additional noise negligible in comparison to the shot noise[58].

More recently, since such a master oscillator-pulsed amplifier system is complex and costly, IMI/NRC has been working on the development of a pulsed oscillator without seeding by a low power oscillator. In this system, the laser pulse presents strong relaxation oscillations followed by a plateau of about 100 μs with weak residual oscillations. Power of more than ten watts on the plateau was obtained on a single transverse and longitudinal mode with either flashlamp or diode pumping. Such a laser is used with a photorefractive demodulator equipped with a balanced receiver so the residual oscillations are not detected[59]. With one stage flashlamp pumped amplification, output power approaching the kilowatt range was also obtained.

4. Digital Signal Processing

Laser-ultrasonics, like any other ultrasonic technique could benefit advantageously of digital signal processing to increase sensitivity. Operations such as averaging, adaptive filtering could be used as well as more advanced methods such as split-spectrum processing or wavelet denoising.

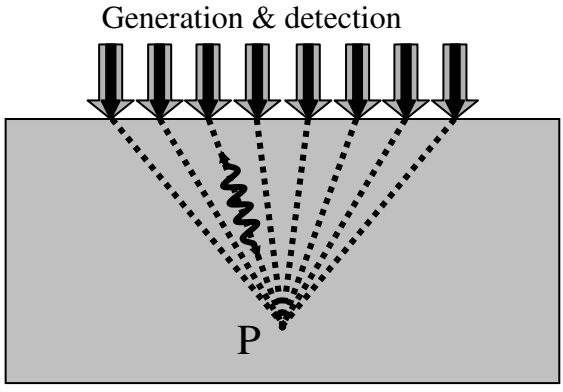


Figure 18. Principle of laser-ultrasonic SAFT processing.

Regarding imaging, a numerical imaging approach such as the one based on the Synthetic Aperture Focusing Technique (SAFT) can be combined with laser-ultrasonics to obtain high quality images[60,61,62]. The principle of SAFT is explained by reference to Fig. 18 as follows. An array of signals is obtained by generating and detecting ultrasound over a grid of points. A point P in the volume of the object is interrogated by summing in a computer all the signals with a proper delay. The delay applied corresponds to the propagation time between the point P and the generation/detection point from the grid at the origin of the signal. If, after summation a result above a certain noise threshold is obtained, this is indicative of a flaw at the point P. Otherwise there is no flaw at P. Since such processing performed in the time domain can be very computation intensive and fairly long, methods that operate in Fourier space and make use of Fast Fourier Transform algorithms have been developed. These methods, significantly different from the conventional time SAFT, although sometimes called F-SAFT, are based on a plane wave decomposition of the acoustic field for each frequency combined with a back-propagation algorithm[63,64,65]. This processing technique has been applied in particular to the imaging of stress corrosion cracks in steel. One example of the results obtained is shown in Fig. 19 where the crack opening image of stress corrosion cracks in a stainless steel test sample is compared with the image obtained conventionally by liquid penetrants[66]. The laser-ultrasonic image was obtained by scanning the opposite surface and applying an F-SAFT reconstruction algorithm using the laser generated shear waves. It should also be noted that SAFT processing could be applied to laser generated surface and plate waves[67].

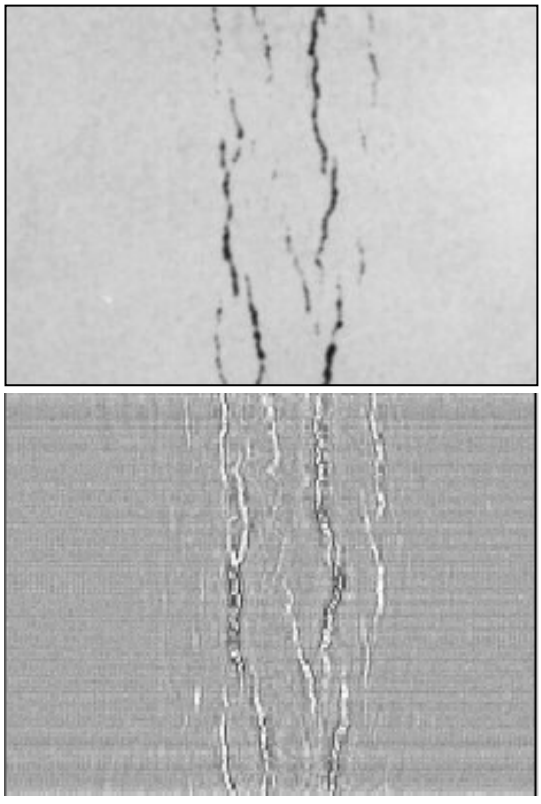


Figure 19. Comparison of crack opening images obtained by liquid penetrants (top) and laser-ultrasonic F-SAFT processing using shear waves and an annular aperture (bottom). The grey scale intensity in the bottom image is indicative of the crak depth.

5. Industrial Applications

We have seen that laser-ultrasonics has been the object of continuous efforts by the scientific and research community to better understand the phenomena taking place during generation and detection of ultrasound and to devise ways to make it usable in an industrial environment. These ways have aimed to push sensitivity, increase bandwidth, insure robustness and reliable operation in spite of ambient disturbances (such as vibrations). However, implementation of the technique is complex, particularly in comparison with conventional piezoelectric-based ultrasonics and also costly and these factors have slowed its industrial use. Nevertheless, we have seen it transitioned during the recent years to the industrial floor for four applications, namely the inspection of aircraft structures made of polymer-matrix composites, the on-line wall thickness gauging of seamless tubes

and the thickness determination and characterization of microelectronic thin layers by two approaches.

5.1. *Inspection of Aircraft Structures Made of Polymer-matrix Composites*

The development of this application has been actively pursued for many years by IMI/NRC, UltraOptec, Inc. a licensee of IMI/NRC and General Dynamics of Forth Worth, Texas (now Lockheed Martin Aeronautics Company)[55,68,69]. Lockheed-Martin has at one point continued independently its own development, which has led to systems that are now routinely used for testing composite parts of the F-22 fighter fabricated by the company[70,71]. Laser-ultrasonics is particularly interesting for inspecting complex parts and for such parts has advantages regarding ease of operation (no additional tooling, no previous detailed knowledge of the part shape, no part precise orientation) and inspection throughput.

This application is first based on the generation of a longitudinal wave normal to the surface independently of the laser beam incident direction, as shown in Fig. 4. As explained above this characteristic is based on the penetration of laser light below the surface for efficient generation and the use of a sufficiently large spot size for minimum ultrasonic beam spreading (5mm typical). For adequate absorption and penetration in these materials, a TEA CO_2 laser operating at 10.6 μm was found quite appropriate and is used in all systems built and presently in use. Penetration is about 20 μm in epoxy and of the same order of magnitude in many paints or peel-plies[13]. Since the system should inspect large areas, a standoff distance of more than one meter is required. This requirement coupled with a scattering surface, which could be also highly absorbing, makes the use of a high power detection laser mandatory. In all the systems developed and presently used, this detection laser is based on Nd-YAG technology and has been described above. A confocal Fabry-Perot is also used as demodulator. Unlike the CO_2 generation beam, which is directly coupled to the inspected part, the beam from the detection laser and the collected beam scattered by the surface are transmitted by optical fibers. In addition to the use at Lockheed Martin, such a technology has been also used at validation and production stages by aerospace companies in Europe using a system assembled by UltraOptec[72].

IMI/NRC has also demonstrated that this technology could be used not only for inspecting fabricated parts but also for inspecting an aircraft during a maintenance operation. Figure 20 shows the laser-ultrasonic C-scan image of the horizontal stabilizer of a CF-18 airplane. Inspection was performed with no surface preparation in a maintenance hangar on a plane in flying conditions[73,74].

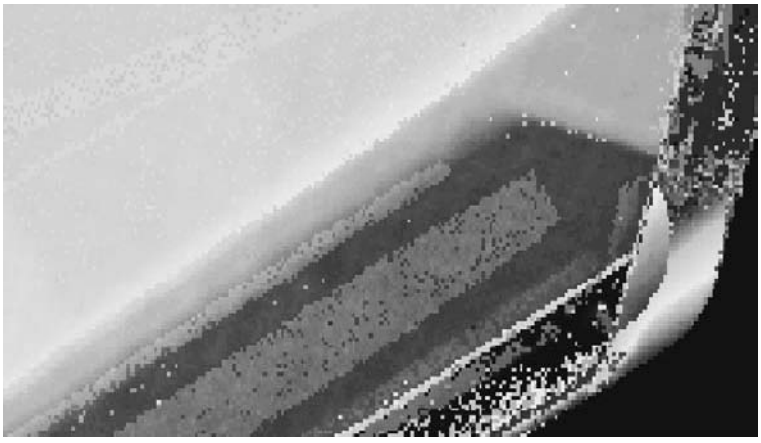


Figure 20. Laser-ultrasonic C-scan image of part of the horizontal stabilizer of a CF-18 airplane in undismantled and ready for take-off conditions. One will notice that, unlike conventional water jet ultrasonics, laser-ultrasonic allows scanning to the very edge of the part.

If much larger areas have to be inspected or for complete flexibility of access all around a given stationary part, the generation unit that houses essentially the generation laser could be mounted at the end of the arm of a gantry robot (with 3 translation axis and one rotation axis in addition of the optical scanning mirror). Such a system has been implemented by UltraOptec for the US Air Force McClellan base and its configuration is sketched in Fig. 21[75]. This system was thoroughly tested and its capability to inspect many parts of various shapes, materials and surface conditions was validated[76]. The results obtained with the system were confronted with those obtained by conventional ultrasonics. The conclusion reached was a definite advantage of this laser-ultrasonic system over conventional ultrasonics for the inspection of contoured parts, but a lower throughput for parts essentially flat, which can be rapidly inspected by multiple conventional transducers in parallel. Also, as expected, shiny surfaces could not be easily inspected over a wide area because of the strong variation of reflectivity with incident angle. Sensitivity for very thick composite parts was found insufficient in many cases although such a limitation could probably be eliminated or diminished by using a photorefractive demodulator, a technology that was not available when this system was built. The systems presently in use at Lockheed Martin while having several distinguishing features with respect to the system installed for the US Air Force make however use of the same concept of a generation unit mounted on a gantry robot[71].

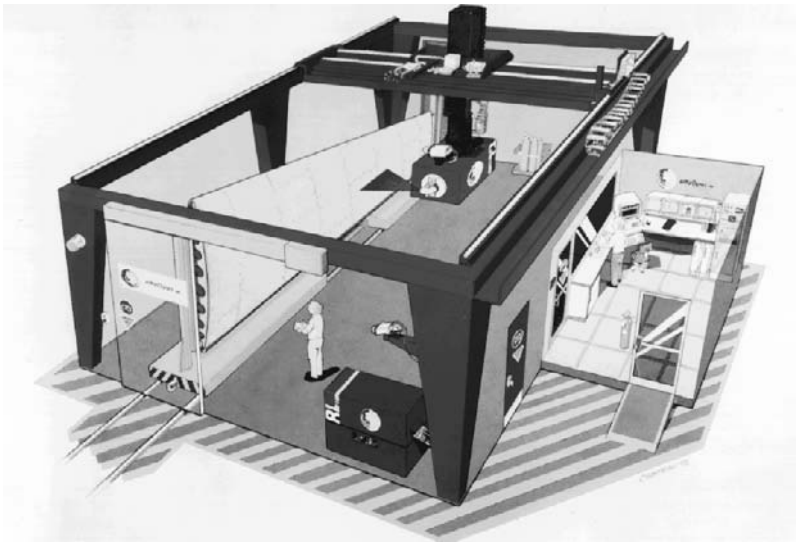


Figure 21. Sketch of the system implemented for the McClellan US Air Force base.

5.2. *On-line Wall Thickness Gauging of Seamless Tubes*

The development of this application follows from a collaboration between IMI/NRC and the Timken Company, which received the support of the US Department of Energy for the implantation on a Timken production line of a laser-ultrasonic wall thickness gauge. This gauge has now measured reliably since its deployment more than one million tubes[77]. IMI/NRC had demonstrated the feasibility of such an application many years before on a seamless tube production line of Algoma Steel[55,73]. This was also followed by similar demonstrations elsewhere[78].

The use of laser-ultrasonics in this case follows from the need of a sensor for measuring at elevated temperature right on the production line the wall thickness and eccentricity. These tubes are fabricated by hot piercing and are used in particular to make hollow round parts with added value, such as pressure cylinders, races of ball bearings… The fabrication process, consisting essentially of forcing a mandrel through a hot billet in rotation, results in relatively large thickness variations and eccentricity, which are undesirable since these variations lead to added fabrication costs for the final products (e.g. more machining time, increased tool wear). Therefore better control of the fabrication of the tubes was needed and has been made possible by the laser-ultrasonic thickness gauge installed on-line and providing in real time information on tube characteristics.

This system allows in particular obtaining very quickly and reliably better mill setups, thus reducing out-of-tolerance products (less scrap and rework) and troubleshooting time. It has resulted in significant productivity increase.

Since the system has to operate in a severe industrial environment, it was made of essentially two units, an environment controlled cabin and a measuring head on top of the line linked by an umbilical cord containing optical fibers for transmission of the generation and detection laser beams and for bringing the scattered light to the interferometer. The cabin houses the lasers, the confocal Fabry-Perot interferometer, control electronics, processing and display data computers. This system also includes a fiber-coupled pyrometer to measure tube temperature and a fiber-coupled coordinate measuring system to determine the measuring locations on the passing tube in rotation and to provide full thickness mapping. A picture of the measuring head on top of the line and above a passing tube is shown in Fig. 23. Note that the implemented system provides eye-safe operation all the time, convenient and fast removal of the measuring head for line service or modification and adequate laser servicing (e.g. periodic flashlamp change) by the location of the lasers in a dust free clean environment. It was also found that the system could provide more than thickness information such as a measurement of the austenitic grain size by proper analysis of ultrasonic attenuation[79]. This application has been commercialized by Tecnar Automation Inc.[80].

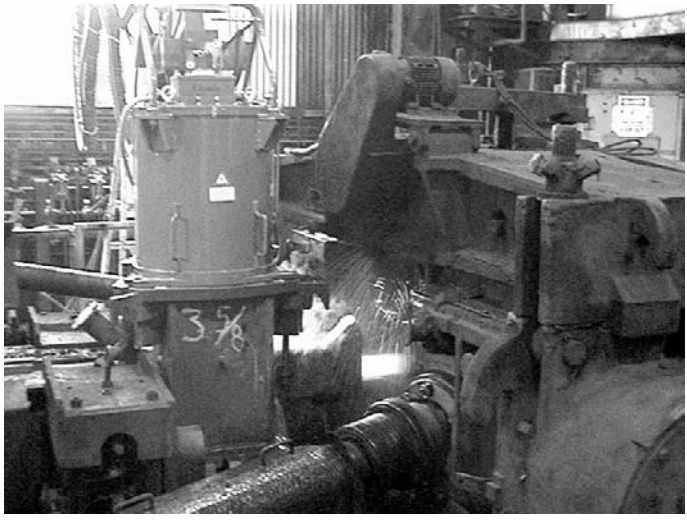


Figure 23. View of the inspection head measuring on-line a tube being processed.

5.3. *Thickness Determination of Microelectronic Thin Layers*

This industrial application follows from developments performed at Brown University on the generation and detection of very high frequency ultrasound with very short pulse lasers (typically femtosecond Ti:Sapphire lasers with pulse duration of about 100 fs)[81,82]. As shown in Fig. 22, the short pulse laser is directed onto the tested sample, where it generates normally propagating longitudinal waves. We have seen above that such a generation is efficient because the laser pulse is very short, even if light penetration is very small, such as in metals (5 to 10 nm). For such a penetration, the propagation delay though the heated layer is even larger than the pulse duration. Therefore, in practice to resolve very thin layers that are not strongly absorbing, a thin metallic layer has to be added as a transducer layer to insure a sufficiently short pulse. This limitation is not important in practice since the technique is applied essentially to opaque thin films, transparent films being well measured by ellipsometry. The stress pulse is detected by monitoring the change of reflectivity using a stroboscopic technique: as shown in Fig. 22, part of the laser pulse is sent to a delay line and is used to probe the reflectivity change after a given delay. By varying the length of the delay line, the whole ultrasonic signal coming from propagation through the sample, usually a stack of thin layers, can be retrieved. Using suitable modeling of the propagation through a multi-layer system, the thicknesses of all the layers of the stack can be determined[83]. This technique has been commercialized by Rudolph Technologies Inc.[84]

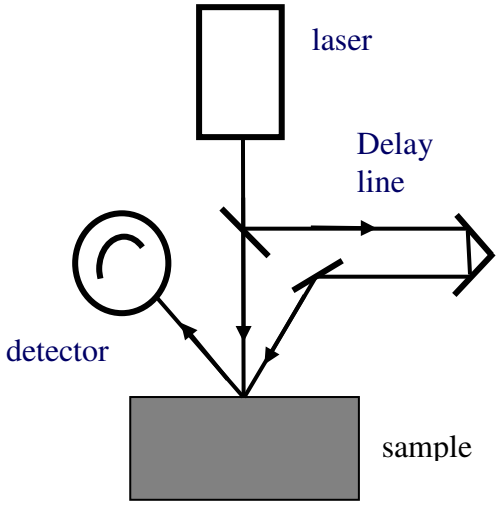


Figure 22. Sketch of a pump-probe setup based on the change of reflectivity for detection.

5.4. *Thickness Determination and Elastic Properties Determination of Microelectronic Thin Layers*

This industrial application follows from developments performed essentially at the Massachusetts Institute of Technology and is based on the generation of ultrasonic surface waves at the surface of the coated material[85,86]. As shown in Fig. 23, two laser beams (the pump beams) interfere on the surface and produce a light grating that in turn, by thermoelastic effect, produce a surface wave with a well defined wavelength. Actually, counter-propagating waves are produced, thus giving a standing surface wave. The time variation of this wave is then monitored by using a probe beam and detecting the beam diffracted by the surface ripple. By varying the angle between the pump beams, the wavelength or k-vector of the surface wave can be changed and its dispersion curve can be measured. Then using suitable modeling, the thickness of the layer and its elastic properties can be determined. This technique has found applications for characterizing films deposited on silicon wafers, such as the thickness measurement of copper interconnections and the determination of the mechanical properties of low permittivity films. It has been commercialized by Philips Advanced Metrology Systems[87].

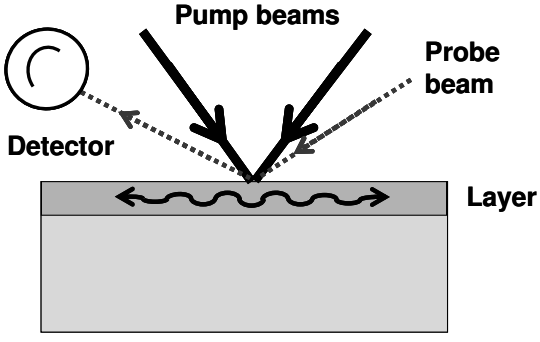


Figure 23. Sketch of a pump-probe setup based on the transient grating approach.

6. Other Potential Industrial Applications

We have seen that four applications of laser-ultrasonics have made the transition from the laboratory to industry. Many other applications are potential for industrial use and are at various stages of development. Among those, there is the application to the characterization of hot metals during processing. As

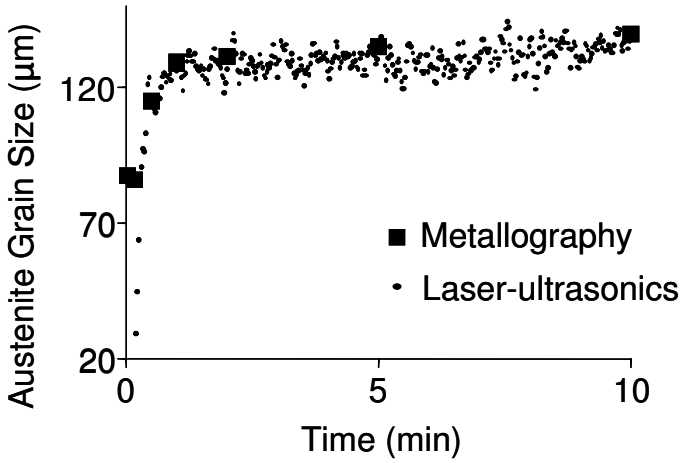


Figure 24. Laser-ultrasonic "metallography"

noted above with respect to the laser-ultrasonic tube thickness gauge, laser-ultrasonics could be used to measure the grain size of steel in the austenitic phase[88]. Figure 24 shows an example of this "laser-ultrasonics metallography", in which grain size determined by analysis of ultrasonic attenuation and grain size measured by conventional metallography after quenching are compared. This data was obtained by heating the steel sample in a Gleeble™ thermal simulator. Grain size determination by laser-ultrasonics, either in the laboratory or on-line, have the potential for commercialization. It should also be noted that laser-ultrasonics could provide on a metal processing line, at the hot stage (hot rolling) or at lower temperature near finishing (cold rolling, annealing) further information such as texture (or anisotropy), distribution between various phases, degree of recrystallisation and information on the mechanical properties[89,90]. Laser-ultrasonics has been in particular tested on a finishing line by IMI/NRC under support of the American Iron and Steel Institute material with the purpose of measuring these properties.

Other potential industrial applications of laser-ultrasonics include the detection of degradation of metallic structures, such as fatigue cracking[91], stress corrosion cracking[66] and corrosion thinning in airframes[92]. There is also definite potential for measuring on-line the mechanical properties of a paper web and its tension[93,94,95]. The measurement of the thickness of an oil spill from an airplane is also another but challenging application which has been demonstrated[96,97]. The detection of defects at a smaller scale (e.g. on chips or electronic boards) has also been explored and is an application well adapted to laser-ultrasonics by the

capability of the technique to provide high frequency ultrasonic testing without water coupling[98].

7. Conclusion

We have presented a broad overview of the basics and of the various technological aspects of laser-ultrasonics. We have outlined the various principles and discussed the phenomena involved in laser generation and detection. We have also presented the advantages and drawbacks of this technique, particularly in comparison with conventional piezoelectric-based ultrasonics. Several of those are linked to the basic characteristic of laser-ultrasonics, which is the use of light as a means for ultrasound emission and detection. Light allows ultrasonic testing without contact and at a distance, thus making possible a wide range of high temperature applications, but sensitivity is on the other hand dependant upon the number of detected photons, which in turn could require a special high power laser for detection (unless the surface has good reflection properties, which is usually the case for microelectronic applications). Many advantages and drawbacks of laser-ultrasonics are also linked to the fact that the material is actually the emitting transducer. This distinguishing feature allows probing more easily parts with complex shapes, but may lead to several limitations depending upon the material, such as very weak emission or material damage. In this regard, we have outlined ways that have been devised to increase the emitted ultrasonic wave amplitude for a specific configuration or task. Detection interferometers well adapted to industrial applications have also been described, including the confocal Fabry-Perot and the two-wave mixing photorefractive interferometer. Since this interferometer has many intrinsic advantages and a remaining drawback, the sensitivity to the Doppler effect associated with object motion having been overcome, it is expected that it will find a broader use, displacing the confocal Fabry-Perot.

A third feature of laser-ultrasonics that has been noted is the complexity of the technique that includes usually two lasers and a detection interferometer. This makes it generally a high cost solution, but in spite of that, it turned out to be cost effective for several applications. To our knowledge, four have really made the transition from the laboratory to the industrial floor: the inspection of polymer-matrix composites, the wall thickness measurement of hot steel tubes during processing and the thickness measurement and characterization of thin layers in microelectronics by two different approaches. Other potential applications that are at various stages of advancement have also been outlined. In spite of the successes, which show that this technology has really a practical impact and is

not simply a laboratory curiosity, efforts should continue to improve its performances and to find ways to diminish its implementation costs.

References

1. Maiman, T. H., *Nature*, **187**, 493 (1960).
2. White, R. M., "Generation of elastic waves by transient surface heating", *J. Appl. Phys.*, **34**, 3559 (1963).
3. Askar'yan, A., Prokhorov, A. M., Chanturya, G. F. and Shipulo, M. P., "The effect of a laser beam in a liquid", *Sov. Phys. JETP*, **17**, 1463 (1963).
4. Scruby, C. B. and Drain, L. E., *Laser-Ultrasonics: Techniques and Applications*, Adam Hilger, Bristol, UK (1990).
5. Hutchins, D. A., "Ultrasonic generation by pulsed lasers", in *Physical Acoustics*, vol. 28, eds. W. P. Mason and R. N. Thurston, Academic Press, New-York, 1988, p. 21.
6. Dewhurst, R. J., Hutchins, D. A., Palmer, S. B. and Scruby, C. B., "Quantitative measurements of laser-generated acoustic waveforms", *J. Appl. Phys.* **53**, 4064 (1982).
7. Rose, L. R. F., "Point-source representation for laser-generated ultrasound", *J. Acoust. Soc. Am.*, **75**, 723 (1984).
8. Doyle, P. A., "On epicentral waveforms for laser-generated ultrasound", *J. Phys.D: Appl. Phys.* **19**, 1613 (1986).
9. Scruby, C. B., Dewhurst, R. J., Hutchins, D. A. and Palmer, S. B. "Laser generation of ultrasound in metals", in *Research techniques in Nondestructive testing* , vol. 5, chapter 8, edited by R. S. Sharpe, Academic press, New-York, p. 281.
10. Vogel, J. A. and Bruinsma, A. J. A., *Non-destructive testing*, vol. 4, eds. J. M. Farley and R. W. Nicols, Pergamon, Oxford, 1988, p. 2267.
11. Noroy, M.-H., Royer, D. and Fink, M., "Transient elastic wave generation by an array of thermoelastic sources", *Appl. Phys. Lett.*, **63**, 3276 (1993).
12. Wang, X., Littman, M. G., McManus, J. B., Tadim M., Kimm Y. S., Askarm A. and Rabitzm H., "Focused bulk ultrasonic waves generated by ring-shaped laser illumination and application to flaw detection", *J. Appl. Phys.,* **80**, 4274 (1996).
13. Dewhurst, R. J., Hutchins, D. A., Palmer, S. B. and Scruby, C. B., "Quantitative measurements of laser-generated acoustic waveforms", *J. Appl. Phys.* **53**, 4064 (1982).
14. Rose, L. R. F., "Point-source representation for laser-generated ultrasound", *J. Acoust. Soc. Am.*, **75**, 723 (1984).
15. Doyle, P. A., "On epicentral waveforms for laser-generated ultrasound", *J. Phys.D: Appl. Phys.* **19**, 1613 (1986).
16. McDonald, F. A., "Practical quantitative theory of photoacoustic pulse generation", *Appl. Phys. Lett.*, **54**, 1504 (1989).
17. Dubois, M., Enguehard, F., Bertrand, L., Choquet, M. and Monchalin, J.-P., "Modeling of laser thermoelastic generation of ultrasound in an orthotropic medium", *Appl. Phys. Lett.*, **64**, 554 (1994).
18. Dubois, M., Enguehard, F., Bertrand, L., Choquet, M. and Monchalin, J.-P., "Numerical and experimental studies of the generation of ultrasound by laser", 8th International Topical Meeting on Photoacoustic and Photothermal Phenomena, Pointe-à-Pitre, Guadeloupe, Jan. 22-25 1994, *J. Physique IV*, Colloque C7, vol. 4, p. 689 (1994).

19. Hutchins, D. A., Dewhurst, R. J. and Palmer, S. B., "Laser generated ultrasound at modified metal surfaces", *Ultrasonics*, **19**, 103 (1981).
20. Arnold, W., Betz, B. and Hoffman, B., "Efficient generation of surface acoustic waves by thermoelesticity", *Appl. Phys. Lett.*, **47**, 672 (1985).
21. Scala, C. M. and Doyle, P. A., "Time- and frequency-domain characteristics of laser-generated ultrasonic surface waves", *J. Acoust. Soc. Am.*, **85**, 1569 (1989).
22. Aindow, A. M., Dewhurst, R. J. and Palmer S. B., "Laser generation of directional surface acoustic wave pulses in metals", *Optics Com.*, **42**, 116 (1982).
23. Murray, T. W., Baldwin, K. C. and Wagner, J. W., "Laser ultrasonic chirp sources for low damage and high delectability without loss of temporal resolution", *J. Acoust. Soc. Am.*, **102**, 2742 (1997).
24. Sharples, S. D., Clark, M. and Somekh, M. G., "All-optical adaptive scanning acoustic microscope", *Ultrasonics*, **41**, 295 (2003).
25. Maldague, X., Cielo, P. and Jen, C. K., "NDT applications of laser-generated focused ultrasonic waves", *Materials Evaluation*, **44**, 1120 (1986).
26. Yamanaka, K., Kolosov, O. V., Nagata, Y., Koda, T., Nishino, H. and Tsukahara, Y., "Analysis of excitation and coherent amplitude enhancement of surface acoustic waves by phase velocity scanning method", *J. Appl. Phys.*, **74**, 6511 (1993).
27. Hoffman, A. and Arnold, W., "Calculation and measurement of the ultrasonic signals generated by ablating material with a Q-switched pulse laser", *Appl. Surface Science*, **96**, 71 (1996).
28. Hoffman, A. and Arnold, W., "Modeling of the ablation source in laser-ultrasonics", in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 19A, eds. D. O. Thompson and D. E. Chimenti , Plenum, New-York, 2000, p. 279.
29. Hébert, H, Vidal, F., Martin, Kieffer, J.-C., Nadeau, A., Johnston, T.W., Blouin, A., Moreau, A. and Monchalin, J.-P. "Ultrasound generated by a femtosecond and a picosecond laser pulse near the ablation threshold", *J. Appl. Phys.*, **98**, 033104 (2005).
30. Murray, T. W. and Wagner, J. W., "Laser generation of acoustic waves in the ablative regime", *J. Appl. Phys.*, **85**, 2031 (1999).
31. Monchalin, J.-P., "Optical detection of ultrasound", *IEEE Trans. Sonics, Ultrasonics*, Freq. Control, **UFFC-33**, 485 (1986).
32. Dewhurst, R. J. and Shan, Q., "Optical remote measurement of ultrasound", *Meas. Sci. Technol.*, **10**, R139 (1999).
33. Blouin, A., Delaye, P., Drolet, D., de Montmorillon, L.-A., Launay, J.-C., Roosen, G. and Monchalin, J.-P., "Optical detection of ultrasound using two-wave mixing in semiconductor photorefractive crystals and comparison with the Fabry-Perot", in *Nondestructive Characterization of Materials VIII*, Proceedings of the conference, Boulder, Colorado, June 15-20, 1997, eds. R. E. Green Jr. and C. O. Ruud, Plenum Press, NY, 1998, p. 13.
34. Caves, C. M., "Quantum-mechanical noise in an interferometer", *Phys. Rev D*, **23**, 1693 (1981).
35. Pouet, B., Breugnot S. and Clemenceau, P., "An Innovative Interferometer for Industrial laser Ultrasonic Inspection", in *Review of Progress in QNDE*, vol. 24A, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2005, p. 273.

36. Pouet, B., Breugnot S. and Clemenceau, P., "Recent Progress in Multi-Channel Quadrature Interferometer: Demonstration of a Compact Fiberized Architecture", in *Review of Progress in QNDE*, vol. 26, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2007, in press.
37. Kaule, W., in Proceedings of 8th World NDT conference, Cannes, France, paper 3J5, 1976.
38. Connes, P., " L'étalon de Fabry-Pérot sphérique", *J. Physique et le Radium*, **19**, 262 (1958).
39. Monchalin, J.-P., and Héon, R., "Laser ultrasonic generation and optical detection with a confocal Fabry-Perot interferometer", *Materials Evaluation*, **44**, 1231 (1986).
40. Shan, Q., Bradford, A. S. and Dewhurst, R. J., "New field formulas for the Fabry-Perot interferometer and their application to ultrasound detection", *Meas. Sci. Technol.,* **9**, 24 (1998).
41. Monchalin, J.-P., Héon, R., Bouchard, P. and Padioleau, C., "Broadband optical detection of ultrasound by optical sideband stripping with a confocal Fabry-Perot", *Appl. Phys. Lett.*, **55**, 1612 (1989).
42. Monchalin J.-P. and Héon, R., "Laser optical ultrasound detection using two interferometer systems", US patent No 5,080,491 (14 Jan. 1992).
43. Blouin, A., Padioleau, C., Néron, C., Lévesque, D. and Monchalin J.-P., "Differential Confocal Fabry-Perot for the Optical Detection of Ultrasound", in *Review of Progress in QNDE*, vol. 26, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2007, in press.
44. Ing, R. K. and Monchalin, J.-P., "Broadband optical detection of ultrasound by two-wave mixing in a photorefractive crystal", *Appl. Phys. Lett.* **59**, 3233 (1991).
45. Blouin, A. and Monchalin, J.-P., "Detection of ultrasonic motion of a scattering surface by two-wave mixing in a photorefractive GaAs crystal", *Appl. Phys. Lett.*, **65**, 932 (1994).
46. Delaye, P., Blouin, A., Drolet, D., de Montmorillon, L.-A., Roosen, G. and Monchalin, J.-P., "Detection of an ultrasonic motion of a scattering surface by photorefarctive InP:Fe under an applied DC field", *J. Opt. Soc. Am.*, **14**, 1723 (1997).
47. de Montmorillon, L.-A., Biaggio, I., Delaye, P., Launay, J.-C. and Roosen, G., "Eye-safe large field-of-view homodyne detection using a photorefractive CdTe:V crystal", *Opt. Comm.* **129**, 293 (1996).
48. Drolet, D., Néron, C., Blouin, A. and Monchalin, J.-P., "Specifications of an ultrasonic receiver based on two-wave mixing in photorefractive gallium arsenide implemented in a laser-ultrasonic system", in *Review of Progress in QNDE*, vol. 15, eds. D.O. Thompson and D.E. Chimenti, Plenum, New-York, 1996, p. 637.
49. P. Delaye, A. Blouin, L.-A. de Montmorillon, D. Drolet, J.-P. Monchalin and G. Roosen, "Photorefractive detection of ultrasound", *Proc. SPIE*, **3137**, 171 (1997).
50. Campagne, B., Néron, C., Blouin, A. and Monchalin, J.-P., "Doppler frequency-shift compensated photorefractive interferometer for ultrasound detection on objects in motion", in *Review of Progress in QNDE*, vol. 22A, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2003, p. 273.
51. Murray, T. W. and Krishnaswamy, S., "Multiplexed interferometer for ultrasonic imaging", *Opt. Eng.*, **40**, 1321 (2001).
52. Grappin, F., Delaye, P. and Roosen, G., "Parallel processing of ultrasonic signals in a single photorefractive crystal", Proceedings of the *9th conference on Photorefractive Effects,*

Materials and Devices, La colle sur Loup , France, June 17-21, 2003, *Trends-in-Optics-and-Photonics*, Optical Society of America, Washington DC, vol. 87, p. 463.

53. Hale, T. C., Telschow, K. L. and Deason, V. A., "Photorefractive optical lock-in vibration spectral measurement", *Appl. Optics*, **36**, 8248 (1997).
54. Telschow, K. L., Deason, V. A., Schley, R. S. and Watson, S. M., "Direct imaging of traveling Lamb waves in plates using photorefractive dynamic holography", *J. acoust. Soc. Am.*, **106**, 2578 (1999).
55. Monchalin, J.-P., "Progress towards the application of laser-ultrasonics in industry", in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 12, edited by D. O. Thompson and D. E. Chimenti, Plenum, New-York, 1993, p. 495.
56. Kane, T. J. and Byer, R. L., "Monolithic unidirectional single-mode Nd: YAG ring laser", *Opt. Lett.*, **10**, 65 (1985).
57. Drake, T., Osterkamp, M., Yawn, K. and Do, T., "Improved laser technology for Industrial Laser Ultrasonics", presented at *Review of Progress in QNDE*, Golden, Colorado, July 25-30, 2004, unpublished.
58. Blouin, A. and Monchalin, J.-P., "Optical amplification of the laser ultrasonic signal", in *Review of Progress in QNDE*, vol. 23, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2004, p. 270.
59. Carrion, L., Blouin, A., Padioleau, C. and Monchalin J.-P., "Single-frequency pulsed laser oscillator and system for laser-ultrasonics", *Meas. Sci. Technol*., **15**, 1939-1946 (2004).
60. Seydel, J. A., "Ultrasonic synthetic aperture focusing techniques in NDT", in *Research Techniques in Nondestructive Testing*, vol. 6, ed. R. S. Sharpe, Academic, New-York, 1983.
61. Blouin, A., Lévesque, D., Néron, C., Drolet, D. and Monchalin, J.-P., "Improved resolution and signal-to-noise ratio in laser-ultrasonics by synthetic aperture focusing technique (SAFT) processing", *Optics Express* **2**, 531 (1998).
62. Lorraine, P. W., Hewes R. A. and Drolet, D., "High resolution laser ultrasound detection of metal defects", in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 16, eds. D. O. Thompson and D. E. Chimenti , Plenum Press, New-York, 1997, p. 555.
63. Mayer, K., Marklein, R., Langenberg, K. J. and Kreutter, T., "Three-dimensional imaging system based on Fourier transform synthetic aperture focusing technique", *Ultrasonics*, **28**, 241 (1990).
64. Busse, L. J., "Three-dimensional imaging using a frequency-domain synthetic aperture focusing technique", IEEE Transactions **UFFC 39**, 174 (1992).
65. Lévesque, D., Blouin, A., Néron, C. and Monchalin, J.-P., "Performance of laser-ultrasonic F-SAFT imaging", *Ultrasonics*, **40**, 1057 (2002).
66. Ochiai, M., Lévesque, D., Talbot, R., Blouin, A., Fukumoto, A. and Monchalin, J.-P., "Visualization of surface-breaking tight cracks by laser-ultrasonic F-SAFT", in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 22A, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2003, p. 1497.
67. Lorraine, P. W., "Laser ultrasonic imaging of Lamb waves in thin plates", in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 17, eds. D. O. Thompson and D. E. Chimenti , Plenum Press, New-York, 1998, p. 603.
68. Padioleau, C., Bouchard, P., Héon, R., Monchalin, J.-P., Chang, F. H., Drake, T. E. and McRae, K. I., "Laser ultrasonic inspection of graphite epoxy laminates", in *Review of*

Progress in Quantitative Nondestructive Evaluation, vol., 12, eds. D. O. Thompson and D. E. Chimenti, Plenum, New-York, 1993), p. 1345.

69. Chang, F. H., Drake, T. E., Osterkamp, M. A., Prowant, R. S., Monchalin, J.-P., Héon, R., Bouchard. P., Padioleau, C., Froom, D. A., Frazier, W. and Barton, J. P., "Laser ultrasonic inspection of honeycomb aircraft structures" in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 12, eds. D. O. Thompson and D. E. Chimenti, Plenum, New-York, 1993, p. 611.
70. Drake Jr, T. E., "Review of LaserUT™ Production Implementation", presented at *Review of Progress in Quantitative Nondestructive Evaluation*, Green Bay, Wisconsin, July 27-August 1, 2003, unpublished.
71. Drake Jr, T. E., Dubois, M., Osterkamp, M., Yawn, K., Tho Do, Kaiser, D., Maestas, J. and Thomas, M., "Review of Progress of the LaserUT™ Technology at Lockheed Martin Aeronautics Company", presented at *Review of Progress in Quantitative Nondestructive Evaluation*, Brusnwick, Maine, July 31-August 5, 2005, unpublished.
72. Pétillon, O., Dupuis, J.-P, Voillaume, H. and Trétout, H., "Applications of Laser Based Ultrasonics to Aerospace Industry", *Proceedings of the 7th European Conference on Non-Destructive Testing*, Copenhagen, 26-29 May 1998, p.27.
73. Monchalin, J.-P., Néron, C., Bussière, J.F., Bouchard, P., Padioleau, C., Héon, R., Choquet, M., Aussel, J.-D., Carnois, C., Roy, P., Durou, G. and Nilson, J.A., "Laser-ultrasonics: from the laboratory to the shop floor", *Advanced Performance Materials*, **5**, 7 (1998).
74. Choquet, M., Héon R., Padioleau C., Bouchard P., Néron C. and Monchalin, J.-P., "Laser ultrasonic inspection of the composite structure of an aircraft in a maintenance hangar", *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 14, eds. D. O. Thompson and D. E. Chimenti, Plenum, New-York, 1995, p. 545.
75. Fiedler, C. J., Ducharme, T. and Kwan, J., "The laser-ultrasonic inspection system (LUIS) at The Sacramento Air Logistics Center", *Review of Progress in Quantitative Nondestructive Evaluation*, vol.16, eds. Thompson D. O. and Chimenti D. E., Plenum, New-York, 1997, vol. 16, p. 515.
76. Nelson, J. M., Nerenberg, R. L., Rempt, R. D., Shepherd, W. B., Sorsher, S. M., Woodmansee, W. E., Choquet, M. and Monchalin, J.-P., "Enhanced laser generated ultrasound", Contract F33615-96-C-5268, Final report to the US Air Force, D950-10322-1, The Boeing Information, Space and Defense Systems, Seattle, 8 July 1998.
77. Monchalin, J.-P., Choquet, M., Padioleau, C., Néron, C., Lévesque, D., Blouin, A., Corbeil, C., Talbot, R., Bendada, A., Lamontagne, M., Kolarik II, R. V., Jeskey, G. V., Dominik, E. D., Duly, L. J., Samblanet, K. J., Agger, S. E., Roush, K. J. and Mester, M. L., "Laser ultrasonic system for on-line steel tube gauging", in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 22A, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2003, p. 264.
78. Deppe, G. J., "Wall thickness measurement by laser UT on hot tubes in a rolling mill", *International Tube Association Conference*, Bilbao, Spain, Oct. 24-28, 2001; available from NDT.net (http://www.ndt.net), March 2002, vol. 7, No. 03.
79. Kruger, S., Lamouche, G., Monchalin, J.-P., Kolarik II, R. V., Jeskey, G. and Choquet, M., "On-line Monitoring of Wall thickness and Austenitic Grain size of a Seamless l Tubing Production Line at The Timken Co.", *Iron & Steel technology*, Oct 2005, pp. 25-31.

80. See web site: http://www.tecnar-automation.com.
81. Thomsen, C., Grahn, H. T., Maris, H. J. and Tauc, J., "Surface generation and detection of phonons by picosecond light pulses", *Phys. Rev. B.,* **34**, 4129 (1986).
82. Maris, H., "Picosecond Ultrasonics", *Scientific American*, p. 86, Jan. 1998.
83. Morath, C. J., Collins, G. J., Wolf, R. G. and Stoner, R. J., "Ultrasonic multiplayer metal film metrology", *Solid State Technology*, p. 85, June 1997.
84. See web site: http://www.rudolphtech.com.
85. Rogers, J. A., Fuchs, M., Banet, M. J., Hanselman, J. B., Logan, R. and Nelson, K. A., "Optical system for rapid materials characterization with transient grating technique: Application to nondestructive evaluation of thin films used in microelectronics", *Appl. Phys. Lett.*, **71**, 225 (1997).
86. Rogers, J. A., Maznev, A. A., Banet, M. J. and Nelson, K. A., "Optical Generation and Characterization of Acoustic Waves in Thin Films: Fundamentals and Applications", *Ann. Rev. Mater. Sci.*, **30**, 117 (2000).
87. See web site: http://www.advancedmetrologysystems.com.
88. Dubois, M., Moreau, A., Militzer M. and Bussière, J. F., "A New Technique for the Quantitative Real-Time Monitoring of Austenite Grain Growth in Steel", *Scripta Materialia* , **42**, 867 (2000).
89. Lamouche, G., Kruger, S. E., Gille, G., Giguère, N., Bolognini S. and Moreau, A., "Laser-Ultrasonic Characterization of the Annealing Process of Low-Carbon Steel", in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 22B, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2003, p. 1681.
90. Kruger,S. E., Moreau, A., Militzer M. and Biggs, T., "In-Situ, Laser-Ultrasonic Monitoring of the Recrystallization of Aluminum Alloys" *Thermec 2003 International Conference on Processing & Manufacturing of Advanced Materials*, Part 1. Eds. T. Chandra, J. M. Torralba and T. Sakai., Trans Tech Publications Ltd., Uetikon-Zurich, 2003, p. 483.
91. Fomitchov, P. A., Kromine, A. K., Krishnaswamy, S. and Achenbach, J. D., "Ultrasonic imaging of small surface-breaking defects using scanning laser source technique", in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 21A, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2002, p. 356.
92. Choquet, M., Lévesque, D., Massabki, M., Néron, C., Bellinger, N. C., Forsyth, D., Chapman, C. E., Gould, R., Komorowski, J. P. and Monchalin, J.-P., "Laser-ultrasonic detection of hidden corrosion in aircraft lap joints: results from corroded samples", in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 20A, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2001, p. 300.
93. Blouin, A., Reid, B. and Monchalin, J.-P., "Laser-ultrasonic Measurements of the Elastic Properties of a Static and a Moving Paper Web and of the Web Tension", in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 20A, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2001, p. 271.
94. Ridgway, P. L., Hunt, A. J., Quinby-Hunt, M. and Russo, R. E., "Laser-ultrasonics on moving paper", *Ultrasonics*, **37**, 395 (1999).
95. Ridgway, P. L., Russo, R. E., Lafond, E. F., Jackson, T. G., Baum G. A. and Zhang, X., "Laser Ultrasonic Measurement of Elastic Properties of Moving Paper: Mill Demonstration"

in *Review of Progress in QNDE*, vol. 25, eds. D. O. Thompson and D. E. Chimenti, AIP Conference Proceedings, New-York, 2006, ?.

96. Choquet, M., Héon, R., Vaudreuil, G., Monchalin, J.-P., Padioleau, C. and Goodman, R. H., "Remote Thickness Measurement of Oil Slicks on Water by Laser-Ultrasonics", Proceedings of the *1993 International Oil Spill Conference*, Tampa, Florida, March 29-April 1, 1993, American Petroleum Institute, Washington DC, publication # 4580, p. 531.
97. Néron, C., Padioleau, C., Monchalin, J.-P., Choquet, M., Brown, C. E. and Fingas, M. F., "Laser Ultrasonic Measurement of the Thickness of Oil Spills from an Airplane", presented at the *Review of Progress in Quantitative Nondestructive Evaluation* , Brunswick, Maine, July 31- August 5, 2005, unpublished.
98. Howard, T., Erdahl, D., Ume, I. C., Gamalski, J. and Achari, A., "Development of an advanced system for inspection of flip chip and chip scale package interconnects using laser ultrasound and interferometric techniques", 2002 International Conference on Advanced Packaging and Systems, *Proceedings of the SPIE*, vol. 4828, 2002, pp. 136-142.

This page intentionally left blank

CHAPTER 5

GUIDED WAVE BASED NONDESTRUCTIVE TESTING: A REFERENCE-FREE PATTERN RECOGNITION APPROACH

Hoon Sohn, Seung Bum Kim and Seung Dae Kim

Carnegie Mellon University
Pittsburgh, PA 15213, USA
hsohn@cmu.edu

A new methodology of guided wave based nondestructive testing (NDT) is developed to detect crack/corrosion damage in metallic structures without using prior baseline data. In conventional guided wave based techniques, damage is often identified by comparing the "current" data obtained from a potentially damaged condition of a structure with the "past" baseline data collected from the pristine condition of the structure. However, it has been reported that this type of pattern comparison with the baseline data can lead to increased false alarms due to its susceptibility to varying operational and environmental conditions of the structure. To develop a more robust damage diagnosis technique, a new concept of NDT is conceived so that defects such as crack and/or corrosion can be detected without direct comparison with previously obtained baseline data. The proposed NDT technique utilizes the polarization characteristics of the piezoelectric wafer transducers attached on the both sides of the thin metallic structure. Crack/corrosion formation creates Lamb wave mode conversion due to a sudden change in the thickness of the structure. Then, the proposed technique instantly detects the appearance of the defects by extracting this mode conversion from the measured Lamb waves even at the presence of changing operational and environmental conditions. Numerical and experimental results are presented to demonstrate the applicability of the proposed technique to crack/corrosion detection.

1. Introduction

There has been an increasing demand in using Structural Health Monitoring (SHM) and Nondestructive Testing (NDT) techniques for continuous monitoring of aircraft, civil infrastructure and mechanical system to prevent catastropic failure and to reduce costs in maintenace. For SHM/NDT, guided waves have

received a great deal of attention and have been a topic of considerable interest, because they can propagate over considerable distances with little attenuation. Conventional guided wave studies have focused on schemes where baseline signals are measured so that changes from the baseline can be detected. However, there are significant technical challenges to realizing this pattern comparison. For instance, structural defects typically take place long after the initial baseline are collected, and other operational and environmental variations of the system can produce significant changes in the measured response, masking any potential signal changes due to structural defects.

As an alternative that can overcome the drawbacks of the conventional NDT methods, a new concept of NDT technique, which does not rely on previously obtained baseline data, is proposed for crack/corrosion detection. In a thin elastic medium such as a metal plate, the formation of a crack/corrosion causes the conversion of the propagation waves to other modes. In this cahpter, a technique that can isolate this mode conversion is developed using the poling directions of piezoelectric materials such as Lead Zirconate Titanate (PZT). The uniqueness of the proposed crack damage detection technique is that this mode conversion due to a crack/corrision formation is instantly identified without using prior baseline data. By removing the dependency on the prior baseline data, the proposed damage detection system becomes less vulnerable to operational and environmental variations that might occur throughtout the life span of the structures being monitored.

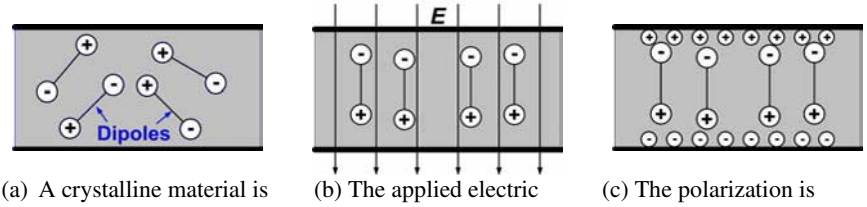
This chapter is organized as follows. First, the polarization process of crystalline materials is briefly described. Then, the effect of a PZT polarization direction on Lamb wave measurement is investigated, and the proposed reference-free diagnosis technique is developed based on the PZT poling directions. Furthermore, a thresholding technique is developed to make the proposed NDT technique less sensitive to variations in PZT size, bonding condition and alignment. Finally, experimental tests as well as numerical simulations are executed to investigate the applicability of the proposed NDT technique to crack/corrosion detection.

2. Theoretical Development

2.1. *Piezoelectric Material and its Polarization Characteristics*

Piezoelectric materials are natural or artificially polarized ceramics which have piezoelectricity[1]. These materials develop an electrical charge or voltage when a mechanical pressure is applied, which is the simplest description of

piezoelectricity. Conversely, piezoelectric materials produce deformation (strain) when exposed to an applied electric field. Due to this unique nature of the piezoelectric materials, they are commonly used as both sensors and actuators in many applications[2]. For instance, wafer-type piezoelectric materials such as PZT are commonly used for exciting and measuring guided waves for SHM and NDT applications[3]. In some natural ceramic materials such as quartz, crystal cells that behave similarly to electric dipoles are oriented along the crystal axes. However, other crystalline materials often have random orientation of the dipoles at the initial state, and they are polarized during a thermal poling process[1]. In the first stage, a crystalline material with randomly oriented dipoles is slightly warmed up below its Curie temperature [Figure 1 (a)]. After a strong electric field E is applied to the crystalline material, the dipoles in the material align along the field lines [Figure 1 (b)]. Finally, the material is cooled down, and the electric field is removed [Figure 1 (c)]. The polarization of the material is permanently maintained as long as the poled material stays below its Curie temperature. The overall behavior of a piezoelectric material as well as its electrical characteristics is governed by the poling direction of the material[2]. In the next section, the influence of the poling direction on Lamb waves is discussed.



(a) A crystalline material is heated up near its Curie temperature.

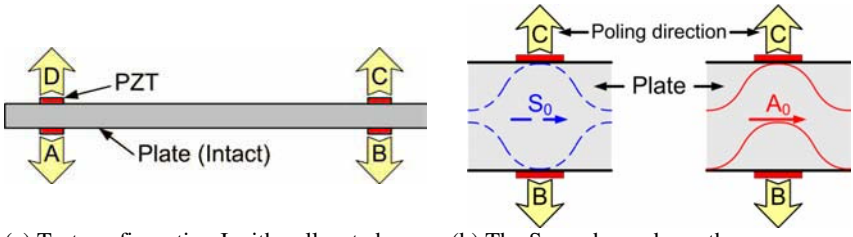
(b) The applied electric field aligns the dipoles along the field lines.

(c) The polarization is permanently maintained after cooling.

Figure 1. A poling process of an artificially polarized material[1].

2.2. *The Effect of PZT Poling Direction on Lamb Wave Propagation*

In this section, it is investigated how the phase of a Lamb wave mode changes depending on (1) the poling directions of exciting and sensing PZT wafer transducers and (2) whether a wafer transducer is attached either on the top or bottom surface of a plate. For illustration, it is assumed four identical PZT wafer transducers, labeled as “A”, “B”, “C, and “D”, are attached to the plate as shown in Figure 2 (a). The arrows indicate the positive poling directions of each PZT transducers. PZTs A and D are placed exactly at the same position but on the other side of the plate. PZTs B and C are positioned in a similar fashion. Furthermore, it is assumed that a narrow-band tone burst is applied as an input



(a) Test configuration I with collocated PZTs with the opposite poling directions.

(b) The S_0 mode produces the same bending for PZTs B and C while the A_0 mode results in the opposite bending.

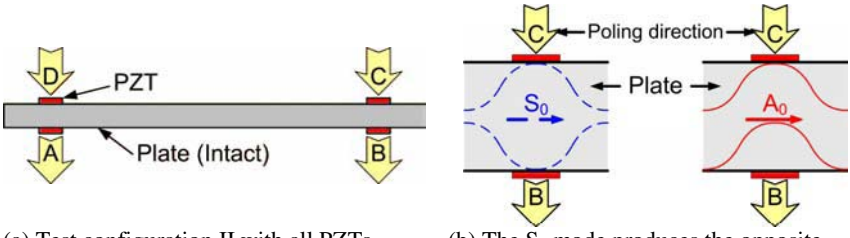
Figure 2. The effect of the PZT poling directions on the phases of the S_0 and A_0 modes (Configuration I).

signal, and the driving frequency (150 kHz) is chosen such that only the fundamental symmetric (S_0) and anti-symmetric (A_0) modes are generated. Although only the S_0 and A_0 modes are shown here for simplicity, the proposed concept is applicable even with multiple higher modes.

When PZT A is excited, the S_0 and A_0 modes are generated and measured at PZTs B and C[4]. In an ideal condition, the amplitude and arrival time of the S_0 mode measured at PZTs B and C should be identical. In addition, both PZTs B and C should be subjected to positive bending because of the symmetric nature of the S_0 mode [See the figure on the left in Figure 2 (b)]. (Hereafter, the term of "positive bending" is used when the positively polarized side of the PZT is subjected to tensile strain. On the other hand, the PZT is subjected to negative bending when the negatively polarized size of the PZT is subjected to tensile strain. The positive bending produces a "positive" output voltage while the negative bending results in a "negative" output voltage value.) Because both PZTs B and C are subjected to the positive bending, the phase of the S_0 mode measured at these PZTs are identical as well as the amplitude and arrival time [See the S_0 mode in Figure 4 (a)]. As far as the A_0 mode is concerned, PZT B is subjected to the negative bending although PZT C still undergoes the positive bending [See the figure on the right in Figure 2 (b)]. Therefore, the A_0 modes measured at PZTs B and C are out-of-phase [See the A_0 mode in Figure 4 (a)]. However, when the poling direction of the PZT C is switched [Figure 3 (a)], PZTs B and C will produce out-of-phase S_0 modes and in-phase A_0 modes [Figure 3 (b) and Figure 4 (b)].

This idea of using the PZT poling directionality in Lamb wave propagation is not a completely new idea. However, the majority of the past work has focused on selective generation of S_0 and A_0 modes[5-8]. For instance, by exciting PZTs A and D shown in Figure 2 (a) in-phase, only the S_0 mode can be excited. In this

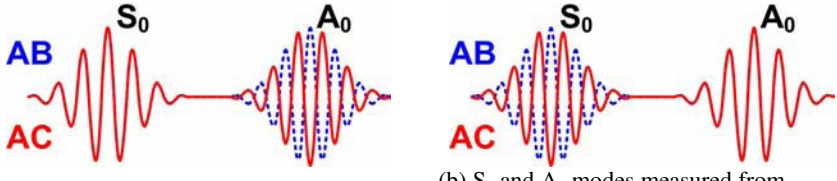
study, the polarization characteristic of the PZT is utilized not only for selective generations of Lamb wave modes but also for selective measurements. In the following section, this concept is further advanced so that the mode conversion due to crack formation can be extracted from the measured Lamb wave signals.



(a) Test configuration II with all PZTs with the same poling directions.

(b) The S_0 mode produces the opposite bending for PZTs B and C while the A_0 mode results in the same bending.

Figure 3. The effect of the PZT poling directions on the phases of the S_0 and A_0 modes (Configuration II).



(a) S_0 and A_0 modes measured from configuration I in Figure 2 (a): S_0 modes in-phase & A_0 modes out-of-phase.

(b) S_0 and A_0 modes measured from configuration II in Figure 3 (a): S_0 modes out-of-phase & A_0 modes in-phase.

Figure 4. A schematic comparison of the S_0 and A_0 modes measured from Configurations I and II shown in Figure 2 (a) and Figure 3 (a), respectively: AB (a dash line) and AC (a solid line) denote the response signals measured at PZTs B and C when a tone burst input is applied at PZT A.

2.3. *Extracting Mode Conversed Signals due to Defect Formation*

In this subsection, the PZT polarization characteristic is further advanced so that the mode conversion due to crack/corrosion formation can be detected without using any prior baseline data. First, the effect of a crack on Lamb wave modes is described. If Lamb waves propagating along a thin plate encounter a discontinuity, some portion of the waves are reflected at the discontinuity point and others are transmitted through it. When a S_0 mode arrives at the notch as shown in Figure 5, it is separated into S_0 and A_0 modes (denoted as S_0/S_0 and A_0/S_0, respectively). In a similar manner, an A_0 mode is also divided into S_0 and A_0 modes (S_0/A_0, and A_0/A_0). This phenomenon is called mode conversion[9].

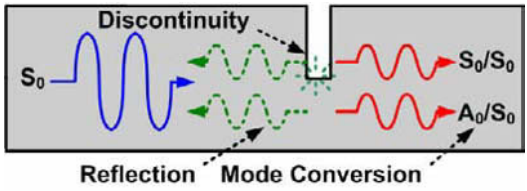


Figure 5. A schematic diagram of mode conversion and reflection due to a discontinuity on a plate

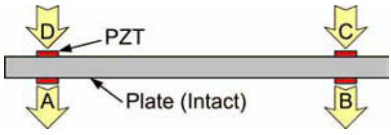
When the plate is in a pristine condition and four identical PZTs are instrumented as shown in Figure 6 (a), it can be shown that signal AB becomes identical to signal CD as illustrated in Figure 6 (b)[10]. Here, signal AB denotes the response signal measured at PZT B when the excitation is applied at PZT A, and signal CD is defined in a similar fashion. However, signal AB is no longer identical to signal CD [Figure 6 (c) and (d)] when there is a crack between PZTs A and B (or PZTs C and D). As for signal AB, the S_0/A_0 mode arrives at PZT B earlier than the A_0/S_0 mode when the notch is located closer to PZT A than PZT B (assuming that the S_0 mode travels faster than the A_0 mode). Conversely, the S_0/A_0 mode arrives at PZT D after the A_0/S_0 mode in the case of signal CD. In Figure 6 (d), signals AB and CD are drawn considering not only the arrival time of each mode but also the poling directions of the PZTs.

Note that, while the S_0 and A_0 modes in Figure 6 (d) are in-phase, the S_0/A_0 and A_0/S_0 modes in signals AB and CD are fully out-of-phase. Therefore, the additional modes generated by a notch can be extracted simply by subtracting signal AB from signal CD as shown in Figure 6 (d). Because this approach relies only on comparison of two signals obtained at the current state of the system rather than comparison with previously recorded reference data, it is expected that this approach reduces false alarms of defect due to changing operational and environmental variations of the system. For instance, it can be readily shown that temperature change of the system does not affect this approach.

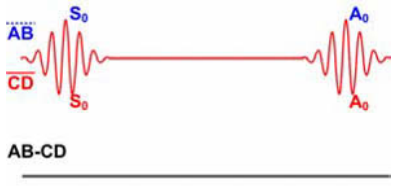
A similar observation can be made with corrosion formation. Alleyne et al.[11] reported that corrosion introduces mode conversion and corrosion in steel pipeline is detected using Lamb waves. Sicard et al.[12] applied Lamb waves to corrosion detection in airframe structures. Experimental results with accelerated corrosion tests are presented later.

2.4. *Development of a Filtering Technique to Address Variation in PZT Size, Alignment and Bonding Condition*

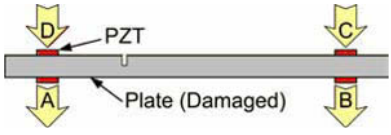
In the previous section, it is shown that signals AB and CD are indistinguishable when there is no crack/corrosion [Figure 6 (b)]. This is based on the assumption



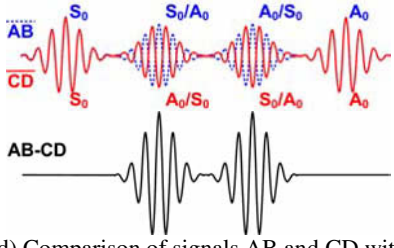
(a) An intact plate with the PZT configuration II shown in Figure 3 (a).



(b) Comparison of signals AB and CD without a notch: the S_0 & A_0 modes are Identical.



(c) A damaged plate with the PZT configuration II shown in Figure 3 (a).



(d) Comparison of signals AB and CD with a notch: the S_0 & A_0 modes are identical, but the S_0/A_0& A_0/S_0 modes are out-of-phase.

Figure 6: Extraction of the additional Lamb wave modes generated by a notch using the poling directionality of the PZT transducers (A_0/S_0 mode denotes an A_0 converted from S_0 when it passes through a crack. S_0/A_0 is defined similarly.)

that all PZT transducers are identical and PZTs A and D (or PZTs B and C) are perfectly collocated. In practice, these assumptions can not be fully satisfied because of variations in PZT size, alignment and bonding condition[10]. Therefore, residual differences would remain after subtracting the two signals even at the absence of a crack, and this could be a source of positive false alarms. To tackle this practical implementation issue, a thresholding technique is developed here.

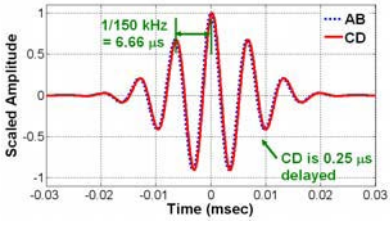
The development of this signal processing technique is based on the observation that the additional modes generated by a crack are out-of-phase while the other modes are in-phase when signals AB and CD are compared [Figure 6 (d)]. Because th S_0 modes and A_0 modes in signals AB and CD are in-phase according to Figure 6 (d), the point-by-point product (PPP) values between the S_0 modes or the A_0 modes are always positive. On the other hand, the PPP values between the S_0/A_0 mode of signal AC and the A_0/S_0 mode of signal BD (or between the A_0/S_0 mode of signal AC and the S_0/A_0 mode of signal BD) always become negative. Therefore, the existence of a crack can be detected by observing the negative PPP values between signals AB and CD. Then, it is investigated how the PPP values are affected by variations in PZT size,

alignment and bonding condition. The main effects of the non-ideal PZTs on signals AB and CD can be summarized as the amplitude change of the response signals and the time shifting of one signal with respect to the other. For instance, the debonding and/or cracking of PZT A can reduce the coupling area between PZT A and the substrate and consequently decrease the amplitude of signal AB with respect to signal CD. It can be readily shown that this pure amplitude change does not alter the sign of the PPP values. However, the variations in PZT size, alignment and bonding condition can also cause phase shifting, and this requires a special treatment.

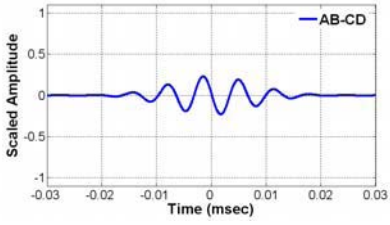
In Figure 7, an example of the phase shift caused by PZT transducer misalignment is illustrated. The results shown in Figure 7 are obtained from the same configuration shown in Figure 3 (a) except that PZT C is shifted 0.76 mm to the right with respect to PZT B. That is, the distance between PZTs A and B becomes 0.76 mm shorter than the distance between PZTs C and D. Due to this misalignment, the A_0 mode in signal AB arrives about 0.25 μs earlier than the A_0 mode in signal CD (0.76 mm / 3.055 mm/μs = 0.25 μs) [Figure 7 (a)]. Because the A_0 mode travels slower than the S_0 mode at 150 kHz (V_S = 5.088 mm/μs, V_A = 3.055 mm/μs), the A_0 mode is more severely affected by the misalignment than the S_0 mode. Therefore, the discussion here focuses on the delay of A_0 modes. Furthermore, note that 0.76 mm misalignment is equivalent to 8% of the 10 mm $\times$ 10 mm PZT wafer transducer size. Although 0.25 μs may be considered a small time delay, Figure 7 (b) shows that this time delay produces a substantial difference between signals AB and CD when they are subtracted from each other.

On the other hand, Figure 7 (c) shows that the PPP values between signals AB and CD are mostly positive except at a few negative points. The negative PPP values are the results of the misalignment and amplified in Figure 7 (d). It can be analytically shown that the PPP between the two signals at 150 kHz produces a 300 kHz frequency component as shown in Figure 7 (c). Therefore, the time period of the PPP values becomes 3.33 μs (= 1/300 kHz). The observation of Figure 7 (d) further reveals that the duration of the negative PPP values is shorter than that of the positive PPP values. For instance, 3 mm misalignment, which is equivalent to 30% misalignment in a 10 mm$\times$ 10 mm PZT wafer transducer, produces negative PPP values with a duration of 0.98 μs (= 3 mm / 3.055 mm/μs). Note that the mode conversion also produces negative PPP values. However, the duration of the negative PPP values caused by the mode conversion becomes 2.35 μs (= 3.33 μs – 0.98 μs) in this configuration. Therefore, the duration of the negative PPP values due to the mode conversion remains longer than the duration caused by PZT misalignment as long as the PZT misalignment is controlled under certain precision. In this paper, it is assumed

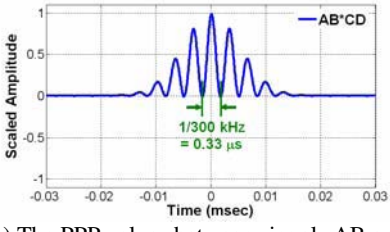
that the PZT misalignment can be controlled under 30% of the 10 mm × 10 mm PZT transducer and this misalignment produces the negative PPT values with a duration less than or equal to 0.98 μs. Then, the effect of the misalignment is eliminated by removing the negative PPP values, whose duration is less than 0.98 μs [Figure 7 (e)]. The effectiveness of this thresholding technique is further investigated in the following numerical studies.



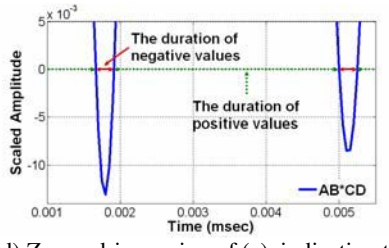
(a) Signals AB &CD when PZT C is shifted 0.76 mm to the right with respect to PZT B.



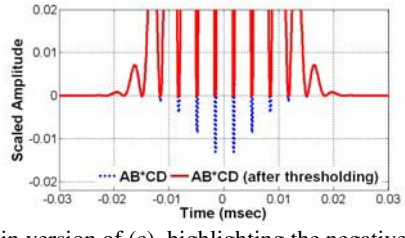
(b) Difference between signals AB and CD.



(c) The PPP values between signals AB and CD when PZT C is shifted 0.76 mm to the right with respect to PZT B.



(d) Zoomed-in version of (c), indicating the durations of positive values and negative values due to misalignment.



(e) Zoomed-in version of (c), highlighting the negative values of the PPP: a dash line – before thresholding, and a solid line – after thresholding.

Figure 7. Compensation of sensor misalignment using the proposed thresholding process (AB*CD denote the point-by-point product (PPP) between signals AB and CD).

3. Numerical Simulation

The idea of using the PZT polarity for crack detection was first validated by numerical simulation. Using COMSOL Multiphysics software (www.comsol.com),

Lamb wave propagation in a two dimensional aluminum plate was simulated using the combination of plain strain, piezo plain strain, and electrostatics modules in COMSOL software. The length of the plate was 70 cm, and its thickness was 6 mm. Four identical PZTs with a size of 10 mm × 10 mm × 0.508 mm were attached to the plate model as shown in Figure 8. Note that PZTs A and D were collocated but on the other side of the plate with the same poling direction. PZTs B and C were placed in a similar fashion. The parameter values used in the numerical simulation are listed in Table 1. A narrowband tone-burst signal at 150 kHz was used as an input signal. In the simulation, Rayleigh damping coefficients were set to 10^{-4} for a mass damping coefficient and 0 for a stiffness damping coefficient, respectively. The simulation results were obtained from a time dependent solver, and a time step was set to 0.25 μs, which is equivalent to 4 M samples/sec. To control the error in each integration step, relative tolerance and absolute tolerance for the solution were chosen to be 10^{-3} and 10^{-10}, respectively. The maximum backward differentiation formula (BDF) order for setting the degree of the interpolating polynomials in the time-stepping method was set to order 2. Finally, the model was meshed using a mapped mesh option, and the size of each mesh was limited to 1 mm × 1 mm[13].

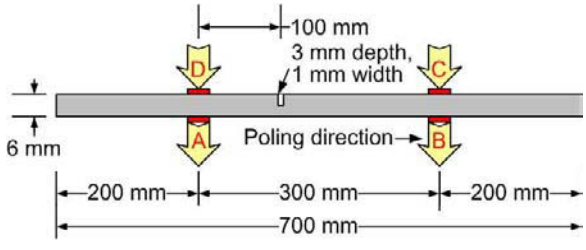
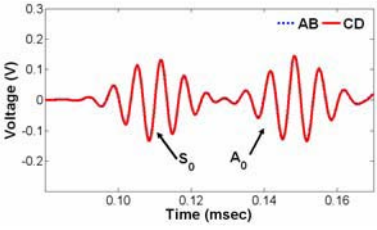


Figure 8. Dimension of an aluminum plate used in numerical simulation.

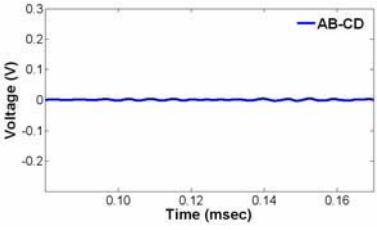
Table 1. Parameters used in numerical simulation.

Exciting frequency	150 kHz
α (Mass damping coefficient)	10^{-4}
β (Stiffness damping coefficient)	0
Sampling rate	4 Ms/s
Relative tolerance	10^{-3}
Absolute tolerance	10^{-10}
Maximum BDF order	2
Mesh size (mapped mesh)	1 mm × 1 mm maximum

Figure 9 illustrates that signals AB and CD were almost identical and this well corresponds to the theoretical expectation. Once a notch of 3 mm depth and 1 mm width was introduced 100 mm away from PZT D toward PZT C, signal AB became different from signal CD as a result of the mode conversion induced by the crack [Figure 10 (a)]. Then, the mode conversion due to crack formation was extracted simply by subtracting signal CD from signal AB [Figure 10 (b)].

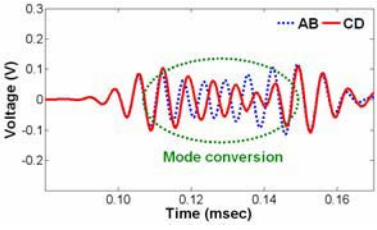


(a) Signals AB and CD without a notch.

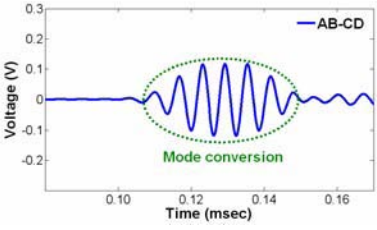


(b) Difference between signals AB and CD.

Figure 9. Simulated Lamb wave signals without a notch.

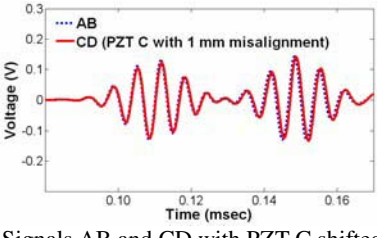


(a) Signals AB and CD with a notch (3 mm depth, 1 mm width).

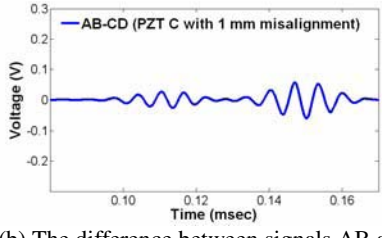


(b) Difference between signals AB and CD with a notch (3 mm depth, 1 mm width).

Figure 10. Simulated Lamb wave signals with a notch of 3 mm depth and 1 mm width.

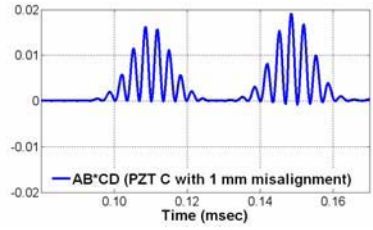


(a) Signals AB and CD with PZT C shifted 1 mm to the right wrt. PZT B.

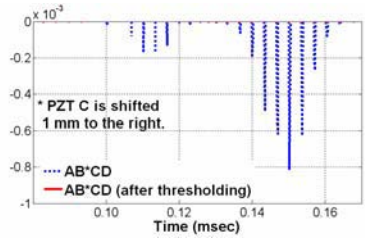


(b) The difference between signals AB and CD with PZT C shifted 1 mm to the right wrt. PZT B.

Figure 11. The effect of PZT misalignment on the difference between signals AB and CD (PZT C shifted 1 mm to the right with respect to PZT B).

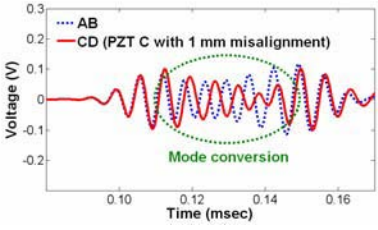


(a) The PPP values between signals AB and CD when PZT C is shifted 1 mm to the right wrt. to PZT B.

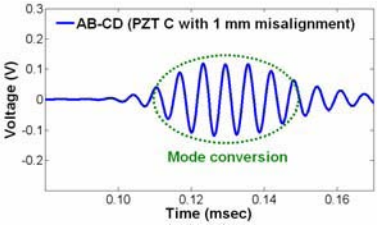


(b) Negative PPP values between signals AB and CD before and after applying the proposed thresholding.

Figure 12. Compensation of the PZT transducer misalignment using the proposed thresholding technique (PZT C: shifted 1 mm to the right with respect to PZT B, without a notch).

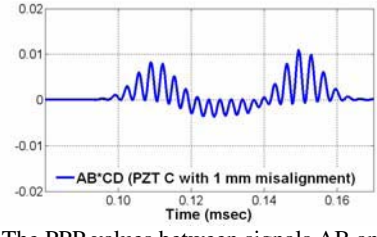


(a) Signals AB and CD with PZT C shifted 1 mm to the right and a 3 mm depth notch.

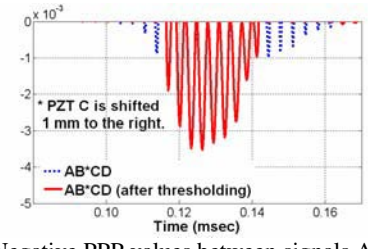


(b) Difference between signals AB and CD with PZT C shifted 1mm to the right and a 3 mm depth notch.

Figure 13. Comparison of signals AB and CD with PZT transducer misalignment and a 3 mm depth notch.



(a) The PPP values between signals AB and CD with PZT C shifted 1 mm to the right and a 3 mm depth notch.



(b) Negative PPP values between signals AB and CD before and after applying the proposed thresholding.

Figure 14. The extraction of the mode conversion using the proposed thresholding technique at the presence of the PZT transducer misalignment (PZT C: shifted 1 mm to the right with respect to PZT B and a 3 mm depth notch).

So far it has been assumed that all PZTs are identical, and two PZTs on the both sides of the beam are precisely collocated. In practice, the size, bonding condition and electrical impedance of PZT transducers will vary from one device

to another, and there can be PZT misalignment. To investigate the effectiveness of the proposed thresholding techniques for addressing these practical implementation issues, the previous numerical simulation was repeated after introducing misalignment between PZTs B and C: PZT C was shifted 1 mm to the right with respect to PZT B. Although the shapes of signals AB and CD were almost identical as shown in Figure 11 (a), their subtraction produced residuals [Figure 11 (b)], whose amplitudes were much higher compared to the one in Figure 9 (b). Figure 12 (a) shows the PPP values between signals AB and CD, and the negative PPP values were magnified in Figure 12 (b). By determining the duration of the negative PPP values and removing negative values whose duration is less than the threshold value (0.98 μs defined in Section 2.4.), the effect of the PZT misalignment was removed in Figure 12 (b).

In Figure 13 (a), the numerical simulation was repeated again with a crack introduced at 100 mm away from PZT A toward PZT B as well as the PZT misalignment. Subtracting signal AB from signal CD produced the residual signal shown in Figure 13 (b). Without relying on prior baseline data, it is challenging to determine whether this remaining signal appears due to mode conversion or sensor misalignment. However, by calculating the PPP between signals AB and CD and applying the proposed thresholding technique, the mode conversion due to a crack was extracted as shown in [Figure 14 (a) and (b)]. This numerical example demonstrates that crack formation can be identified even when there is PZT transducer misalignment. This numerical finding is further substantiated in the following experimental study.

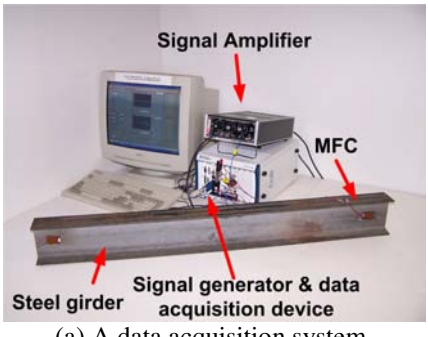
4. Experimental Results

4.1. *Description of Experimental Setup*

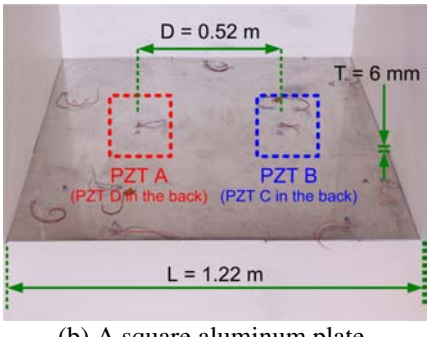
To further examine the proposed reference-free NDT technique, experimental tests have been conducted on an aluminum plate. The overall test configuration of the experiment and the test specimen are shown in Figure 15. The data acquisition system was composed of an arbitrary waveform generator (AWG), a high-speed signal digitizer (DIG), a low noise preamplifier (LNP) and a multiplexer. The dimension of the plate was 122 cm $\times$ 122 cm $\times$ 0.6 cm, and four PSI-5A4E type PZT wafer transducers (1.0 cm $\times$ 1.0 cm $\times$ 0.0508 cm) were mounted in the middle of the plate. PZTs A and D were collocated and attached on the different side of the plate, and PZTs B and C were mounted in a similar fashion. The PZTs were attached so that their poling directions were identical to the configuration shown in Figure 3 (a). PZTs A and B (or PZTs C and D) were

0.52 m apart each other. In this experiment, the PZT transducers were attached to either the top or the bottom surface of the plate with a commercial cyanoacrylate adhesive.

Using the 14-bit AWG, a tone-burst signal with a ±10 peak-to-peak voltage and a driving frequency of 150 kHz was generated and applied. First, PZT A in Figure 15 (b) was excited using the tone-burst waveform. Then, PZT A generated elastic waves, and the response was measured at PZT B. When the waves arrived at PZT B, the voltage output from PZT B was amplified by the LNP with a gain of fifty and measured by the DIG. The sampling rate and resolution of the DIG were 20 MS/sec and 16 bits, respectively. In order to increase a signal-to-noise ratio, the forwarding signals were measured twenty times and averaged. After the forwarding signal from PZT A to PZT B (signal AB) was measured, the same process was repeated by exciting PZT C and measuring response at PZT D (signal CD). Finally, the PPP values between signals AB and CD were calculated, and the negative PPP values due to sensor misalignment were selectively removed using the proposed thresholding technique. Detailed crack and corrosion test results are described in Sections 4.2 and 4.3, respectively.



(a) A data acquisition system.



(b) A square aluminum plate.

Figure 15. An aluminum plate testing configuration for crack detection.

4.2. *Crack Detection*

In Figure 16 (a), Lamb wave signals obtained from an intact plate are presented. Using the test setup described in Section 4.1, signals AB and CD were measured. Although signal AB was supposed to be identical to signal CD, a residual signal was observed in Figure 16 (b) due to variations in PZT alignment, size and bonding condition.

Then, a 1 mm (width) × 60 mm (length) notch was introduced between PZTs A and B (150 mm away from PZT A toward PZT B), and its depth was

gradually increased up to 3 mm. Due to crack formation, two additional modes appeared between the existing S_0 and A_0 modes as shown in Figure 16 (d), (f) and (h). The comparison of Figure 16 (b) with Figure 16 (f) and (h) clearly shows the effect of the mode conversion on the difference between signals AB and CD. However, the existence of the mode conversion may not be conclusive for the 1 mm depth notch shown in Figure 16 (d) unless it is compared with the signal in Figure 16 (b).

Next, the PPT values between signals AB and CD were computed and the proposed thresholding technique was applied. Using the proposed thresholding technique, the negative values associated with the PZT variations were selectively removed while the negative PPP values resulted from the mode conversion were preserved. As shown in Figure 17 (a), the negative PPP values corresponding to the intact condition became negligible after applying the thresholding technique. Once the notch was introduced, the effect of the two additional modes clearly appeared as shown in Figure 17 (b), (c), and (d). Note that even the existence of the incipient notch (1 mm) was clearly detected [Figure 17 (b)], and the arrival times of these two additional modes in Figure 17 (b), (c), and (d) well matched with the theoretical calculation.

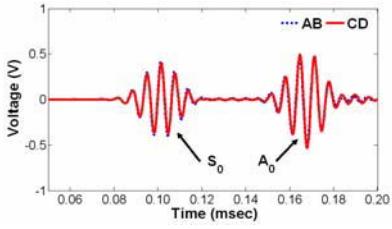
In addition, an attempt was made to estimate the location of the notch by measuring the arrival times of the two converted modes in signals AB and CD. Based on signal AB in Figure 16 (a), the group velocities of the S_0 and A_0 modes were estimated to be 5.128 m/ms and 3.099 m/ms (theoretically, V_S = 5.088 m/ms and V_A = 3.055 m/ms), respectively. Assuming that the notch is closer to the PZT A, the arrival time of the first converted mode, t_a, was estimated using Eq. (1):

$$t_a = \frac{s}{V_A} + \frac{l-s}{V_S} \tag{1}$$

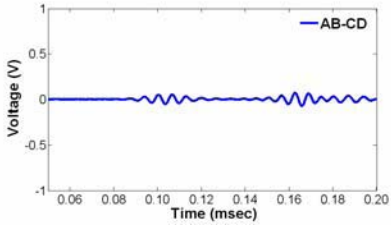
where s and l denote the distances between the notch and PZT A and between PZT A and PZT B, respectively. V_A and V_S, denote the group velocities of the A_0 and S_0 modes. By measuring the arrival time of the first converted mode (0.1206 ms) and using Eq. (1), the two possible locations of the notch were estimated to be at 15.034 cm away from PZT A or PZT B. This estimated distance was close to the actual notch location (0.2% error).

Note that, when a notch is closer to PZT A than PZT B, the S_0/A_0 mode arrives at PZT B earlier than the A_0/S_0 mode. Conversely, the A_0/S_0 arrives first when the notch is closer to PZT B. However, it could not be determined whether the first arrived mode on each signal was the S_0/A_0 mode or the A_0/S_0 mode, because these modes were not distinguishable [Figure 6 (d)]. (When signals AB

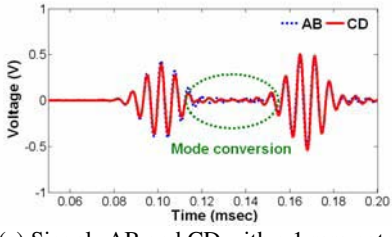
and CD are compared, both modes are out-of-phase) Therefore, it was only possible to narrow down the crack position into two alternatives. In this example, the crack location was determined assuming that the crack was closer to PZT A. However, the S_0/A_0 and A_0/S_0 modes could have been distinguishable if signals AB and AC had been compared instead of signals AB and CD. When signals AB and AC are compared, the S_0/A_0 and A_0/S_0 modes become out-of-phase and in-phase, respectively. Therefore, each mode can be selectively identified. Once each mode is identified, the crack location along the wave propagation path can be uniquely determined.



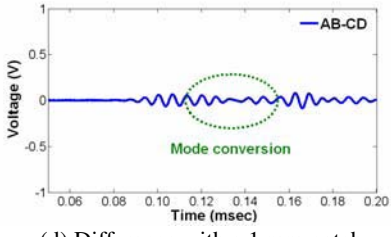
(a) Signals AB and CD without a notch.



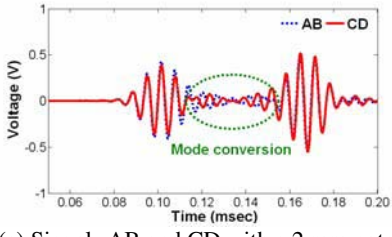
(b) Difference without a notch.



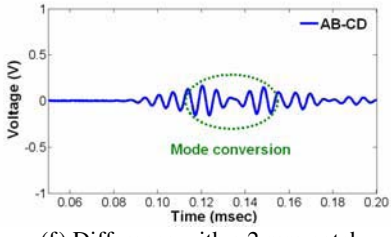
(c) Signals AB and CD with a 1 mm notch.



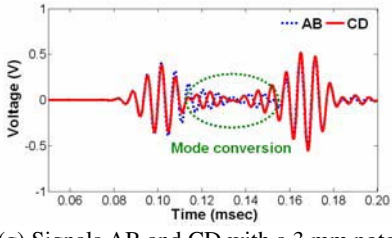
(d) Difference with a 1 mm notch.



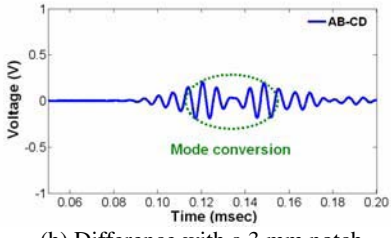
(e) Signals AB and CD with a 2 mm notch.



(f) Difference with a 2 mm notch.

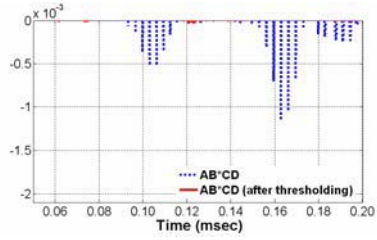


(g) Signals AB and CD with a 3 mm notch.

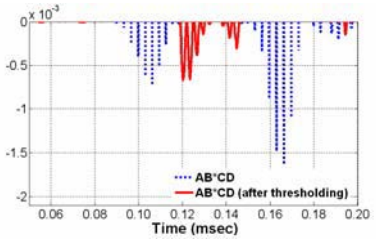


(h) Difference with a 3 mm notch.

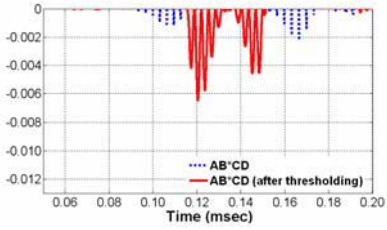
Figure 16. Comparison of signals AB and CD by increasing the depth of a notch.



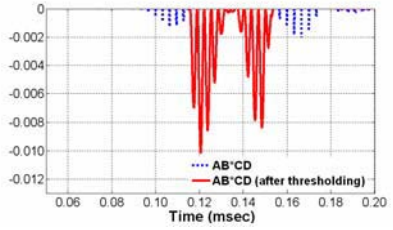
(a) Negative PPP values before and after applying the proposed thresholding (without a notch).



(b) Negative PPP values before and after applying the proposed thresholding (with a 1 mm depth notch).



(c) Negative PPP values before and after applying the proposed thresholding (with a 2 mm depth notch).



(d) Negative PPP values before and after applying the proposed thresholding (with a 3 mm depth notch).

Figure 17. The extraction of the mode conversion produced by a notch using the proposed thresholding technique at the presence of the PZT transducer misalignment (a dashed line before thresholding and a solid line after thresholding).

4.3. *Corrosion Detection*

In this section, the same principle is applied to corrosion detection in an aluminum beam shown in Figure 18. The dimension of the beam was 91 cm × 2 cm × 0.124 cm. Four PZTs (PZTs A to D) of an identical size (2 cm × 1.5 cm) were placed, and the distance between two PZTs was set to 250 cm. The driving frequency value and the sampling rate of the digitizer were set to 50 kHz and 5 MHz, respectively. In order to find the optimal excitation frequency that showed the best separation between A0 and S0 mode, a frequency sweep test was performed from 10 kHz to 150 kHz with a 5 kHz incremental value. During the data collection, 50,000 time data points, equal to 0.01 msec, were collected. The rest of the test configuration was similar to those described in Section 4.1.

The corrosion tests examined are summarized in Table 2. Initially, the beam was tested without any corrosion. Then, corrosion (1 cm × 2 cm) was gradually introduced 5 cm apart from PZT A using a Hydrochloric Acid (HCl). The thickness of the target area decreased from 1.24 mm to 1.17 mm, 1.04 mm and 0.91 mm. They correspond to 6.1%, 16.3% and 26.5% thickness reduction, respectively.

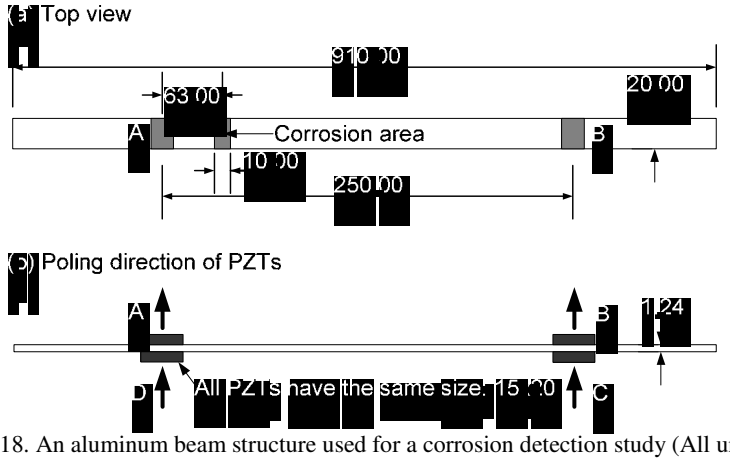
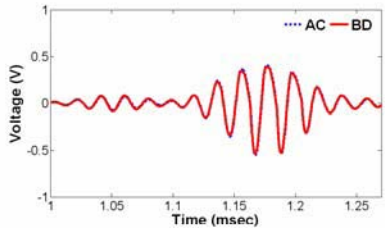


Figure 18. An aluminum beam structure used for a corrosion detection study (All units are in mm).

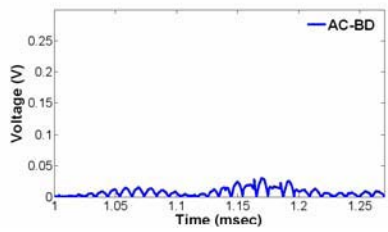
Table 2. Gradual reduction of the specimen's thickness through accelerated corrosion test.

Test	Thickness (mm)	Loss (mm)	Loss rate (%)
Initial	1.24	0.00	0.0
#1	1.17	0.08	6.1
#2	1.04	0.20	16.3
#3	0.91	0.33	26.5

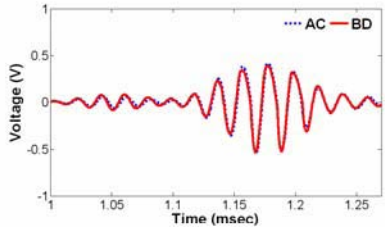
The signals AB and CD and their differences for all test cases are shown in Figure 19. Note that signals AB and CD were normalized so that they do have the unit standard deviation and could be easily compared. Figure 19 clearly illustrates that the difference between signals AB and CD gradually increased as the thinning of the aluminum article continued. In Figure 20, the negative PPP values corresponding to all test cases are depicted. The dashed line shows the negative PPT values before applying the proposed thresholding technique, and the solid line displays the values after applying the thresholding. The negative PPP values were negligible for the initial condition and increased as the corrosion progressed. When the duration of the negative PPP values was less than the threshold value, 2.0 μs, defined in Section 2.4, the corresponding negative values were eliminated during the proposed thresholding process. The thresholding technique might have removed some of the negative PPP values that are associated with PZT variations. It is believed that the negative PPP values after thresholding are mainly associated with mode conversion created by corrosion.



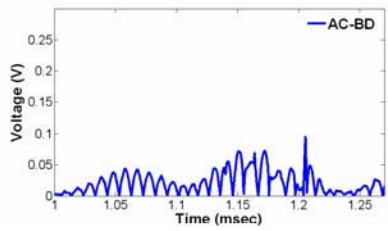
(a) Signals AB and CD without corrosion.



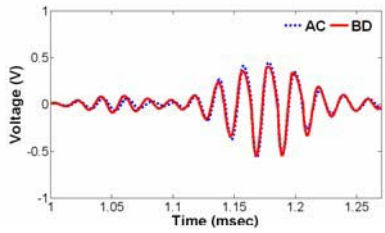
(b) Difference between signals AB and CD without corrosion.



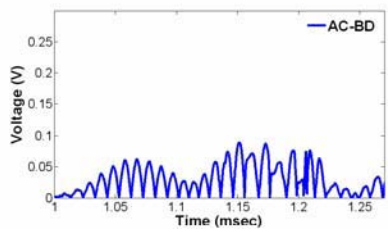
(c) Signals AB and CD with 6.1% thickness reduction.



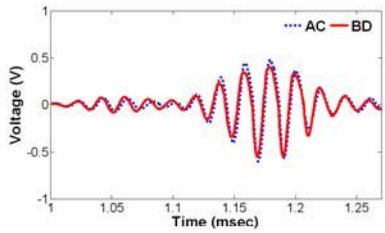
(d) Difference between signals AB and CD with 6.1% thickness reduction.



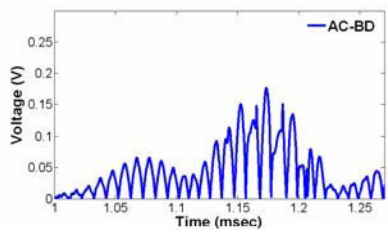
(e) Signals AB and CD with 16.3% thickness reduction.



(f) Difference between signals AB and CD with 16.3% thickness reduction.

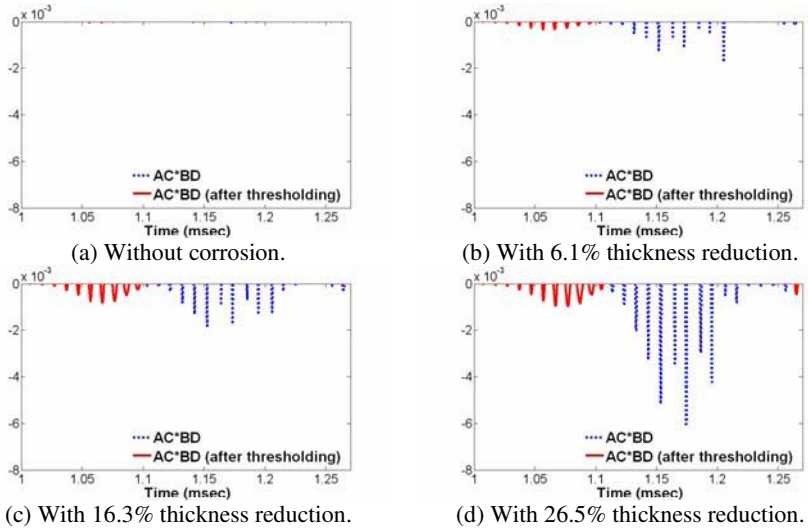


(g) Signals AB and CD with 26.5% thickness reduction.



(h) Difference between signals AB and CD with 26.5% thickness reduction.

Figure 19. Comparison of signals AB and CD when corrosion gradually reduced the thickness of the aluminum specimen.



(a) Without corrosion. (b) With 6.1% thickness reduction.

(c) With 16.3% thickness reduction. (d) With 26.5% thickness reduction.

Figure 20: Negative PPP values between signals AB and CD before and after applying the proposed thresholding (a dashed line before thresholding and a solid line after thresholding).

5. Conclusion

A new concept of nondestructive testing is developed in this study so that crack/corrosion formation in a thin metallic structure can be instantaneously detected without referencing to previously stored baseline data. This reference-free technique for nondestructive testing is developed based on Lamb wave propagations and PZT polarization characteristics. When Lamb waves propagate through a crack or corrosion, the Lamb wave modes are converted to other modes, and this phenomenon is called mode conversion. The appearance of this mode conversion is identified by strategically placed PZT wafer transducers on the both side of the test article. Numerical simulations and experimental tests conducted in this study substantiate the effectiveness of the proposed reference-free technique for damage detection. Practical instrumentation issues such as variations in PZT size, bonding condition as well as alignment are addressed by developing a thresholding technique. Because this reference-free technique does not rely on previously obtained baseline data for crack detection, it is expected that this approach minimize false alarms of damage due to changing operational and environmental variations experienced by many in-service structures. The robustness of the proposed technique against undesirable variations in the system, such as temperature and external loading, can make it attractive for long-term continuous monitoring. Further investigation is underway to extend the

proposed concept to more complex structures such as steel girders and welded connections.

Acknowledgements

This work was supported by the National Science Foundation under grants CMS-0529208 and the Pennsylvania Infrastructure Technology Alliance (PITA), a partnership of Carnegie Mellon University, Lehigh University and the Commonwealth of Pennsylvania Department of Community and Economic Development. The authors like to acknowledge Dr. S. Liu for the NSF support and Prof. Gary K. Fedder for PITA support.

The work was performed while the authors were with Carnegie Mellon University. Dr. Hoon Sohn is now with Korean Advanced Institute of Science and Technology in its Civil and Environmental Engineering Department. Seung Dae Kim is now with the University of California, San Diego.

References

1. Fraden, J., Handbook of Modern Sensors, 2nd Edition, American Institute of Physics, 2001, 556pp.
2. Buchanan, R.C., Ceramic Materials for Electronics, 3rd Edition, Marcel Dekker, New York, 2004, 676pp.
3. Giurgiutiu, V., and Zagrai, A.N., "Characterization of Piezoelectric Wafer Active Sensors", Journal of Intelligent Material Systems and Structures, Vol. 11, No. 12, 2000, pp. 959-976.
4. Viktorov, I. Rayleigh and Lamb Waves, Plenum Press, New York, 1967, 154pp.
5. Su, Z., and Ye, L., "Selective generation of Lamb wave modes and their propagation characteristics in defective composite laminates", Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications, Vol. 218, No. 2, 2004, pp. 95-110.
6. Yamanaka, K., Nagata, Y., and Koda, T., "Selective excitation of single-mode acoustic waves by phase velocity scanning of a laser beam," *Applied Physics Letters*, 58(15), 1591-1593, 1991.
7. Wilcox, P.D., Lowe, M.J.S., and Cawley, P., "Mode and Transducer Selection for Long Range Lamb Wave Inspection," *Journal of Intelligent Material Systems and Structures*, 12(8), 553-565, 2001.
8. Giurgiutiu, V., "Tuned Lamb Wave Excitation and Detection with Piezoelectric Wafer Active Sensors for Structural Health Monitoring," *Journal of Intelligent Material Systems and Structures*, 16(4), 291-305, 2005.
9. Cho, Y., "Estimation of Ultrasonic Guided Wave Mode Conversion in a Plate with Thickness Variation", IEEE transactions on ultrasonics, ferroelectrics, and frequency control, Vol. 47, No. 3, 2000, pp. 591-603.

10. Park, H.W., Kim, S.B., and Sohn, H., "Understanding a Time Reversal Process in Lamb Wave Propagations", Proceedings on Structural Control and Health Monitoring, 2006.
11. Alleyne, D.D., Pavlakovic, B., Lowe, M.J.S., and Cawley, P., "Rapid, Long Range Inpsection of Chemical Plant Pipework using Guided Waves," Review of Progress in QNDE, Vol. 20., pp. 180-187, 2001.
12. Sicard, R., Chahbaz, A., and Goyette, J., "Cprrosion Monitoring of Airframe Structures using Ultrasonic Arrays and Guided Waves," AIP Conference Proceedings, No. 657A, pp. 806-813, 2003.
13. COMSOL AB., "COMSOL Multiphysics User's Guide", Version 3.2, 2005.

CHAPTER 6

ACOUSTIC NONLINEAR IMAGING AND ITS APPLICATION IN TISSUE CHARACTERIZATION

Dong Zhang and Xiu-fen Gong

Institute of Acoustics, Lab. of Modern Acoustics
Nanjing University, Nanjing 210093, China
dzhang@nju.edu.cn

Acoustic nonlinear imaging has brought about significant improvements in image quality by taking advantage of the nonlinear components. This article reviews works of acoustic nonlinear imaging for biological tissues in Nanjing University, including acoustic nonlinearity parameter B/A imaging, tissue harmonic imaging by using the multi-phase-coded-pulse technique and super harmonic imaging. Theoretical analysis and experimental imaging of biological tissues by using these methods are presented.

1. Introduction

Ultrasound imaging has been widely accepted for clinical diagnosis because of its capability to provide important information on the diseased state of the tissues in a human body non-invasively and non-destructively. In recent years, ultrasound image quality has been greatly improved due to advances in technology and introduction of new techniques, such as tissue harmonic imaging [1-3]. The acoustic nonlinear effects in biological tissues at biomedical frequencies and intensities were first theoretically predicted and experimentally demonstrated [4-5]. The second and higher harmonic components are generated as ultrasound propagates through biological tissues due to the phenomenon of nonlinear sound propagation. Tissue harmonic imaging uses ultrasound at twice the transmitted frequency to form the image, unlike the fundamental imaging that uses ultrasound at the transmitted frequency to form the image. Although it is superior to fundamental imaging mode due to improvement of the spatial resolution and suppression of side lobe levels, two typical shortcomings are associated with this technique. One is that the low signal-to-noise ratio (SNR)

limits the penetration depth in ultrasonic diagnosis due to the much lower energy at the second harmonic frequency than that at the fundamental frequency. Another is related to the narrow bandwidth signals in the transmitting system for reducing the side-lobe levels and the harmonic leakage, which results in degraded axial resolution. Much attention has been focused on the study of the nonlinear effect in biological tissues [6-13], and on the development of the acoustic nonlinear imaging techniques [14-25]. In this chapter, works on acoustic nonlinear imaging in our Lab are reviewed. (1) As a measure of the nonlinearity in the pressure-density relation for a medium, the nonlinearity parameter B/A plays a significant role in nonlinear acoustics [26], and has been found some dependence on the compositions and structure features of biological tissues [8,10-12]. Thereby this parameter has been proposed to be a new parameter for tissue characterization, and could have some potential applications in ultrasonic diagnosis. We present here the B/A imaging via the measurement of second harmonics. (2) We present a multiple-phase-coded-pulse technique to provide the desired enhancement of SNR for high order harmonic ultrasonic imaging. In this technique, N phase-coded pulses are transmitted in sequence, the Nth order harmonic component in the echo signal is extracted with a gain of $20\log_{10} N$ dB compared with that in the single pulse harmonic imaging mode, whereas the fundamental and other order harmonic components are efficiently suppressed to reduce image confusion. Consequently, SNR is increased by $10\log_{10} N$ dB. (3) A super-harmonic component is defined as a summation of the third-, fourth- and fifth-order harmonic components [25]. We study the super-harmonic generation and imaging in biological tissues.

2. Acoustic Nonlinearity Parameter B/A Imaging

2.1. *Definition of B/A*

B/A arises from the Taylor series expansion expressing the variations in pressure with density in a medium. It represents the fact that density does not follow linear ly the changes in applied pressure, and may be put in the form [26]:

$$p - p_0 = A\left(\frac{\rho-\rho_0}{\rho_0}\right) + \frac{B}{2}\left(\frac{\rho-\rho_0}{\rho_0}\right)^2 + \frac{C}{6}\left(\frac{\rho-\rho_0}{\rho_0}\right)^3 + \cdots, \qquad (1)$$

where p is the pressure in the medium and ρ is the density, with p_0, ρ_0 is the static value of the same quantities. Coefficients $A = \rho_0(\partial p/\partial\rho)_{s,0} = \rho_0 c_0^2$, $B = \rho_0^2(\partial^2 p/\partial\rho^2)_{s,0}$, $C = \rho_0^3(\partial^3 p/\partial\rho^3)_{s,0}$ and c_0 is the sound speed of small signal. All partial derivations are evaluated at equilibrium configuration and at constant

entropy, which are denoted by subscripts "0", and "s". The third term on the right-hand side of Eq. (1) is orders of magnitude smaller than the second. The quantity of B/A is the ratio of the quadratic to linear terms in the Taylor series,

$$\frac{B}{A} = 2\rho_0 c_0 \left(\frac{\partial c}{\partial p}\right)_{s,0}, \tag{2}$$

Suppose that an initially sinusoidal wave travels in a lossless medium, a progressive change in shape will be found as shown in Fig. 1. The explanation for this change lies in the variation in phase speed v(x) at a position x, which is described as:

$$v(x) = c_0 + (1 + B/2A)u(x) \tag{3}$$

where u(x) is the particle velocity at the position x. Eq. (3) shows that the phase velocity at a point on the wave travel depends on the wave velocity, as well as the particle velocity and B/A. The phase speed is highest in the high-density compressions, and lowest in the low-density rarefactions. As a result of the changed waveform, the monochromatic wave progressively becomes richer in harmonic frequencies. By using Burgers' equation in lossless medium, and boundary condition x = 0, $u = u_0 \sin\omega t$, here u_0 is the amplitude of particle velocity at the source, ω is the angular frequency, the sound pressure at distance x is given as: [26]

$$p = \sum_n \frac{2p_0}{n\sigma} J_n(n\sigma)\sin n(\omega t - kx) \tag{4}$$

where p_0 is the sound pressure amplitude at the source, J_n is the Bessel function at the order n, $\sigma = x/\ell$, $\ell = 1/(\beta M k)$ is the discontinuity distance, $\beta = 1 + B/2A$ is the acoustic nonlinear coefficient, $M = u_0/c_0$ is the acoustic Mach number. When n = 2 and $\sigma << 1$, the sound pressure amplitude of second harmonics is

$$p_2 = \frac{\omega x p_0^2}{4\rho_0 c_0^3}(B/A + 2), \tag{5}$$

Considering the nonlinearity of a medium, the high harmonics are generated during the propagation of sound wave with fundamental frequency f as shown in Eqs. (4) and (5). According to the law of energy conservation, the energy of high harmonics comes from that of the fundamental wave. Therefore, the fundamental wave will have an extra attenuation, and the effective propagation distance of ultrasound wave, will be shortened. The existence of these nonlinear phenomena will obviously influence the propagation of ultrasound in biological media, and it also will have influences on the accuracy of diagnosis and the efficiency of the therapy in biomedicine.

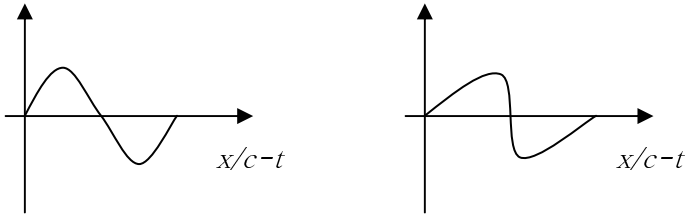


Figure 1. Progressive change in waveform for an initially sinusoidal wave in water. (B/A = 5.2, c_0 = 1500 m/s, p_0 = 1 atm, x= 10 cm, f = 20 MHz).

2.2. *Determination Of B/A for Biological Tissues*

A thermodynamic method [6-7] and a finite amplitude method [8] are well known as the traditional measurement methods for the determination of B/A for fluids and biological tissues. The former is generally considered to be capable of the greater accuracy, which requires the measurement of sound velocity as a function of hydrostatic pressure and temperature. B/A is derived from consideration of the expression in Eq. (2). The latter requires the measurement of the second harmonic amplitude, and B/A is derived from Eq. (5). Corrections are necessary for attenuation and diffraction for biological tissues [6].

In our laboratory, a finite amplitude insert substitution method was proposed [7], in which B/A of a sample is determined by comparing the amplitude of harmonics in the sample with that in a reference liquid with known B/A value (such as water). The advantages of the method are that the absolute value of the sound pressure for the second harmonics is unnecessary, and the effect of diffraction could be effectively reduced. In the measurement, a sample box with thickness of d, is inserted into water between the transmitter and the receiver. The distance between two transducers is L, and the value of B/A is $(B/A)_0$. By measuring the relative amplitudes of the sound pressure for the second harmonics with and without insertion of the sample, respectively, then the B/A value of the sample is obtained as:

$$\left(\frac{B}{A}\right)_x = \left[\frac{p_{2x}}{p_{20}}\frac{L}{d}\frac{1}{I_1 I_2} - \left(\frac{L}{d}-1\right)\frac{I_2}{I_1}D'D''\right]\frac{(\rho c^3)_x[(B/A)_0+2]}{(\rho c^3)_0 D'^2 D''} - 2, \qquad (6)$$

where the subscripts “0” and “x” stand for values in the water and the sample, respectively. $I_1 = \exp(-\alpha_1 d)$, $I_2 = \exp(-\alpha_2 d/2)$, Transmission coefficients $D' = 2(\rho c)x/[(\rho c)_0 + (\rho c)_x]$, $D'' = 2(\rho c)x/[(\rho c)_0 + (\rho c)_x]$.

An improved thermodynamic method was also developed in our laboratory [7], in which the change of the sound velocity with respect to the hydrostatic

pressure is replaced by measuring the phase change with respect to the hydrostatic pressure.

$$\frac{B}{A}=2\rho_0 c_0\left(\frac{\partial c}{\partial p}\right)_s=-2\frac{\rho_0 c_0^3}{L}\left(\frac{\Delta t}{\Delta p}\right)_s=-2\frac{\rho_0 c_0^3}{\omega L}\left(\frac{\Delta\phi}{\Delta p}\right)_s, \tag{7}$$

where L is the distance between the transmitter and the receiver, $\Delta\phi$ is the change of the phase.

2.3. *Tissue Characterization Using B/A*

The B/A parameter has been studied in various biological media, and its relationship with compositions and structure features of media has been investigated. It has been found that (1) the values of B/A in biological media range from 5-12. (2) Fatty soft tissues have higher values of B/A than other kinds of normal tissues. [13] Sehgal et al. obtained the following equation fitting data at 37^0c [8]:

$$\frac{10^{15}}{\rho c^3}(B/A)=1.378+0.106x_w+4.037x_f, \tag{8}$$

where x_w and x_f are the volume fractions of water and fat content. (3) B/A is dependent on the ratio of free to bound water. [8,10-11] (4) B/A may also depend on structure features at the tissue level, cellular lever, and the macromolecular level [12]. Table I gives our results of B/A values for normal porcine liver and various pathological porcine livers.

Table I. Values of B/A for normal porcine liver and various abnormal porcine livers.

Samples	velocity (m/s)	Density (g/cm3)	Attenuation(dB/cm)		B/A
			2 MHz	4 MHz	
Normal liver	1607	1.05	1.10	2.16	6.9
Stasis Liver	1720	1.01	1.28	2.23	7.9
Fatty liver	1601	1.05	1.27	2.20	8.0
Hepatocirrhosis	1580	1.06	1.05	2.24	8.5

2.4. *Acoustic Nonlinearity Parameter Imaging*

Based on the finite amplitude theory, we have investigated the transmission B/A tomography via the second harmonics, the parametric array, as well as B/A tomography in reflection mode [20-22]. Here, the B/A imaging via the measurement of the second harmonics is reviewed.

When a sample is inserted into the reference medium (water) between the transmitter and the receiver, the sound pressure amplitude for the second harmonics at distance L is:

$$p_{2x}(L)=\frac{\omega}{2}p_0^2\int_0^L\beta_i(x)\exp[\int_0^x-2\alpha_1(x)dx-\int_x^L\alpha_2(x)dx]dx, \tag{9}$$

where p_0 is the source amplitude of sound pressure, $\beta_i(x) = \beta(x)/[\rho(x)c^3(x)]$, $\beta(x)=1+\frac{1}{2}\left(\frac{B}{A}\right)(x)$. $\alpha_1(x)$, $\alpha_2(x)$ are sound attenuation coefficients for fundamental and second harmonic waves. When the sample is absent, we have

$$p_{20}(L)=\frac{\omega L}{2}p_0^2\beta_{i0}, \tag{10}$$

where β_{i0} is the value of β_i for water. Using Eqs. (9) and (10), we obtain

$$\frac{p_{2x}(L)}{p_{20}(L)}=\frac{1}{\beta_{i0}L}\int_0^L\beta_i(x)\exp\left[\int_0^x-2\alpha_1(x)dx-\int_x^L\alpha_2(x)dx\right]. \tag{11}$$

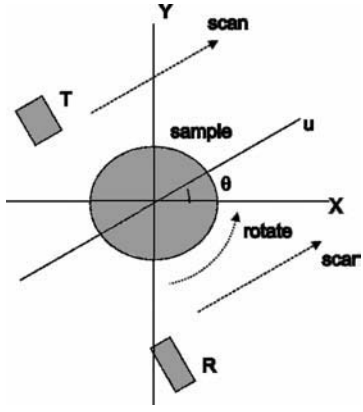


Figure 2. The conventional CT scanning.

By considering $p_{2x}(L)/p_{20}(L)$ as projection data $P(u,\theta)$ in a conventional CT scanning technique (Fig. 2), we reconstruct the tomography of $\beta_i(x,y)$ corrected by an attenuation matrix C(x,y) defined by

$$C(x,y)=\frac{1}{2\pi}\int_0^{2\pi}\exp[\int_{-V_1}^{v}-2\alpha_1(x,y)dv-\int_v^{V_1}\alpha_2(x,y)dv]d\theta, \tag{12}$$

where $V_1 = L/2$. Figure 3 shows a B/A tomography for a sample combined with normal porcine liver and a kind of pathological porcine liver (fatty liver). From the result, we find that the value of B/A for the fatty liver increases, and the corresponding part in the image exhibits higher gray.

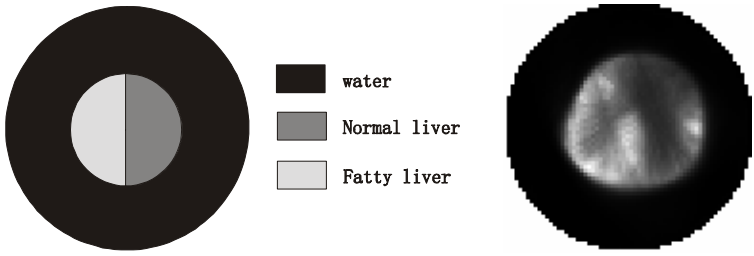


Figure 3. (a) Cross section of sample model, (b) B/A image of the sample.

3. Multi-Phase-Coded-Pulse Technique

A burst signal with angular frequency ω and initial phase φ_0 is radiated in a medium, and the sound pressure at distance x = 0 is $p_0 \sin(\omega t + \varphi_0)$. From the Fubini series solution of the Burgers' equation [26], the sound pressure at distance x is:

$$p(x) = \sum_{m=1}^{\infty} \frac{2p_0}{m\sigma} J_m(m\sigma) \sin(m\omega\tau + m\varphi_0), \tag{13}$$

where m is the order of the harmonics. $\sigma = x / x_k$ is the axial normalized distance, $x_k = (\beta M k)^{-1}$ is the distance of shock formation, $k = \omega / c_0$ is the wave number, β is the nonlinearity coefficient, $M = v_0 / c_0$ is the acoustic Mach number, where v_0 and c_0 are the particle velocity and isentropic sound speed. t is the transmitting time and $\tau = t - x / c_0$ is the retarded time.

Suppose that N pulse signals with an equidistant phase φ_n are used to excite the transducer in sequence. If the initial phase of the first pulse signal is φ_0, the phase of the nth pulse signal is set to be $\varphi_n = \varphi_0 + 2\pi n / N$, where $n = 0, 1, K, N-1$. The sound pressure of the n[th] transmitted pulse signal at the source is expressed as

$$p_n(0) = p_0 \sin(\omega t + \varphi_n) = p_0 \sin(\omega t + \varphi_0 + 2\pi n / N) \qquad n = 0 \quad \text{to} \quad N-1. \tag{14}$$

According to Eqs. (13) and (14), the sound pressure at distance x for the n[th] transmitted pulse signal is:

$$p_n(x) = \sum_{m=1}^{\infty} \frac{2p_0}{m\sigma} J_m(m\sigma) \sin[m\omega\tau + m(\varphi_0 + 2\pi n / N)]. \tag{15}$$

The summation of the received signals of the N phase-coded transmitted signals at distance x is:

$$p_s(x) = \sum_{n=0}^{N-1} \sum_{m=1}^{\infty} \frac{2p_0}{m\sigma} J_m(m\sigma) \sin[m\omega\tau + m(\varphi_0 + 2\pi n / N)]. \tag{16}$$

From Eq. (16), we can find that the sound pressure of the m[th] order harmonic component in the summation is:

$$p_{sm}(x) = \sum_{n=1}^{N-1} \frac{2p_0}{m\sigma} J_m(m\sigma) \sin[m\omega\tau + m(\varphi_0 + 2\pi n / N)] . \quad (17)$$

From Eq. (17), the sound pressure of the summation is:

$$p_{sm}(x) = \begin{cases} \dfrac{2p_0}{m\sigma} J_m(m\sigma) \sum\limits_{n=1}^{N-1} \sin[m\omega\tau + m(\varphi_0 + 2\pi n / N)] = 0 & m \neq N \\ N \dfrac{2p_0}{m\sigma} J_m(m\sigma) \sin(m\omega\tau + m\varphi_0) & m = N \end{cases} . \quad (18)$$

Comparing Eq. (18) with Eq. (13), the sound pressure of the N[th] order harmonic component using N phase-coded pulses is N times that of the single pulse, meanwhile, the other harmonic components including the fundamental frequency are fully eliminated. In addition, the KN[th] (k is integer) order harmonics are also can be in creased by N times. The phase inversion technique is the special case of N = 2 [24]. It is concluded that the selected harmonic component can be efficiently extracted with harmonic enhancement of $20\log_{10} N$ dB in high order harmonic imaging by the use of N phase-coded pulses, and SNR still can be improved by $10\log_{10} N$ dB considering the noise energy incorporation.

In this chapter, the second and higher harmonic imaging with multiple phase-coded pulses are carried out for biological tissues. As shown in Fig. 4, the sample is a piece of porcine fatty tissue with two holes (diameter 3 mm and 2 mm). In the reconstructed images in Fig. 5, the different gray scale corresponds to different biological tissues due to the difference in the amplitudes of the harmonic components. By means of multiple phase-coded pulses (N = 2, 3, 4 and 5 respectively), the processed acoustic nonlinear images at the second to fifth order harmonic frequencies are presented in Fig. 5(b) to (e). For comparison, the fundamental frequency images are also given in Fig. 5(a). The two holes in Fig. 5(a) are vague and the shapes are not clearly displayed. In addition, there are lots of speckle noises in the fatty tissue for the ultrasound diffraction and side lobe level of the fundamental frequency. By means of the phase-coded pulse technique, the amplitudes of the processed high order harmonics are enhanced, which yield the improved the brightness and contrast of the images as shown in Fig. 5(b) to (e). Along with the increase of the harmonic order, the resolution and the clarity of the images become better.

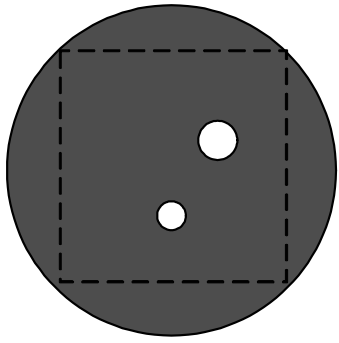


Figure 4. Cross section of sample model.

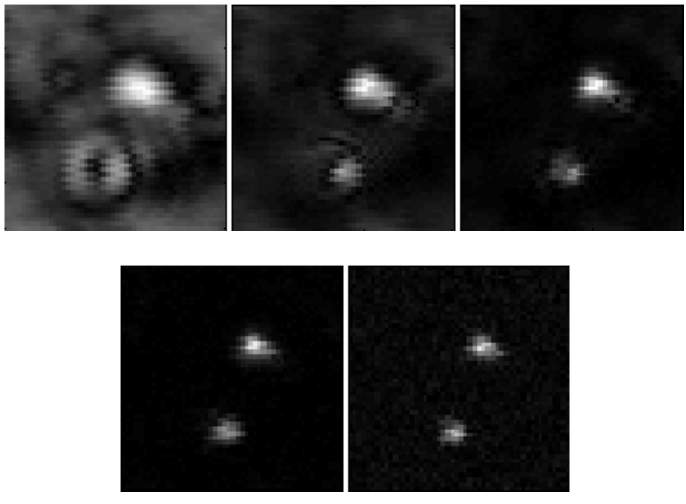


Figure 5. Reconstructed images of sample model by the multi-phase-coded-pulse technique: (a) 1st image, (b) 2nd harmonic image, (c) 3rd image, (d) 4th image and (e) 5th image.

4. Super Harmonic Imaging

When a finite amplitude wave propagates in a medium, the nonlinear effects occur and the second and the higher harmonics are generated. The super-harmonic component is obtained by the linear summation of the absolute amplitudes of 3rd, 4th and 5th harmonics [25].

The widely accepted KZK model [27-28] is used to calculate the amplitudes of the fundamental, second, third, fourth and fifth harmonics during the nonlinear sound propagation, Numerical simulations are performed for a planar piston source with a diameter of 8 mm operating at 2.5 MHz. Distilled water is considered as the medium for ultrasound propagation. The sound velocity is

1492 m/s at 23°C and the attenuation coefficient is $25\times10^{-15}\,\mathrm{Npm^{-1}Hz^{-2}}$ with a frequency dependency of f^2 [23]. The coefficient of nonlinearity β is set to 3.5 [23]. The non-dimensional axial and lateral grid steps of $\Delta\sigma$ and Δr are equal to 10^{-4} and 5×10^{-3} respectively. The calculated steps of the axial and lateral spatial spaces are set to 10000 and 2000; hence the maximum values for axial and lateral directions are 1 and 10, respectively. The maximum number of the calculated harmonic order is 50. Figure 6 shows a comparison of the axial sound pressure distributions between the super-harmonic and second harmonic components at a source pressure of 0.5 MPa. Figure 7 gives the same comparison at a source pressure of 1 MPa. In these two figures, the dash and solid lines represent respectively the calculated super-harmonic and the second harmonic components, which are both normalized by the maximum sound pressure of the second harmonic. Figure 8 displays the normalized lateral beam patterns of the super-harmonic and the second harmonic components at a normalized axial distance of 0.6.

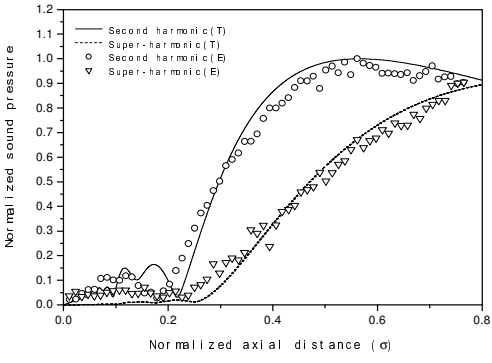


Figure 6. Comparison of the normalized axial sound pressures between the superharmonic and the second harmonic components at $p_0 = 0.5$ MPa (T-Theory, E- Experiment).

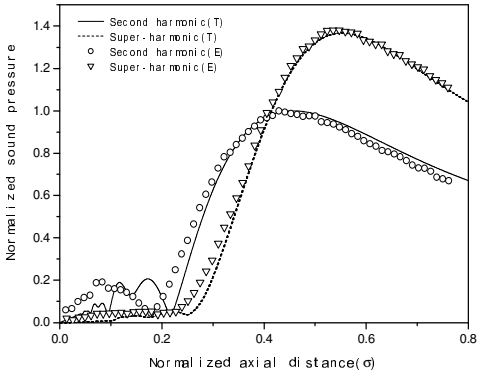


Figure 7. Comparison of the normalized axial sound pressures between the superharmonic and the second harmonic components at $p_0 = 1$ MPa (T – Theory, E – Experiment).

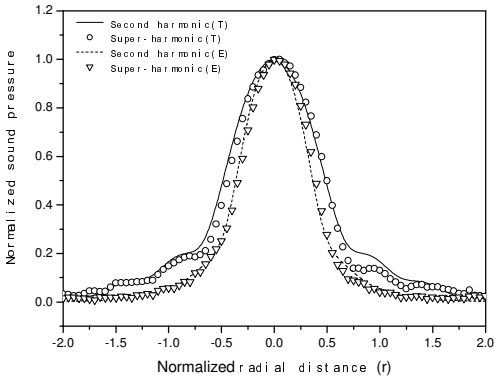


Figure 8. Comparison of the normalized lateral beam patterns between the superharmonic and the second harmonic components at $\sigma = 0.6$ (T – Theory, E – Experiment).

As shown in Fig. 6, at a source sound pressure of 0.5 MPa, the amplitude of the super-harmonic component is lower than that of the second harmonic for the normalized axial distance $\sigma \leq 0.8$. As shown in Fig. 7, when the source sound pressure is equal to 1 MPa, the amplitude of the super-harmonic component begins to exceed that of the second harmonic for $\sigma \geq 0.4$, reaches a maximum value at $\sigma = 0.54$, and then decreases due to the high attenuation, but it is still higher than that of the second harmonic. Therefore the source pressure amplitude and the sound propagation distance are two important variables in obtaining a super-harmonic component with higher energy over the second harmonic.

Figure 8 compares the lateral beam patterns of the super-harmonic and the second harmonic components, which are normalized by their corresponding maximum amplitudes. It is observed that the super-harmonic component has a narrower main lobe and its –3dB beam width is about 77% of that of the second harmonic. The narrower beam width of the super-harmonic component results in an improved lateral resolution. Compared with those of the second harmonic, the side lobe levels of the super-harmonic component are reduced by about 10 dB because the energy is concentrated in the main lobe. This will help reducing artifacts and aberrations in medical imaging.[19]

Figure 10 shows the arrangement of the experimental set-up for studying the super-harmonic sound propagation and imaging for biological tissues. A programmable function generator (Agilent 33250, Loveland, CO, USA) is controlled by a computer via GPIB interface and generates a tone burst signal (2.5 MHz, 8 cycles, pulse repetition frequency 1 KHz). This signal is first amplified using a broadband 55 dB power amplifier (ENI A150, Rochester, NY, USA) and then sent to a PZT transducer (center frequency 2.5 MHz, radius

4 mm). The received acoustic pressure is detected by a broadband needle hydrophone with 0.6 mm in diameter (NP1000, 0.1 MHz to 20 MHz, NTR Systems. Inc. Seattle, WA, USA). Both the transmitter and the receiver are coaxially fixed and immersed in distilled water. A sound permeable sample container with a thickness of 20 mm is placed between the hydrophone and the transmitter and its rear surface is about 1 mm from the receiver to minimize the influence of sound diffraction. After passing through a 30 dB pre-amplifier (PFS017A, NTR Systems. Inc. Seattle, WA, USA), the received signals are acquired using a digital oscilloscope (Agilent 54810, Colorado Springs, CO, USA). The digitized signals are further processed using a FFT software. The super-harmonic component amplitude is thus obtained by the linear summation of the 3rd, 4th and 5th harmonic components.

A mechanical scanning system (Newport Motion Controller MM3000, Irvine, CA, USA) is used in two-dimensional motion for the measurement of the super-harmonic sound distribution and the super-harmonic imaging. The axial and lateral resolutions are 1 μm.

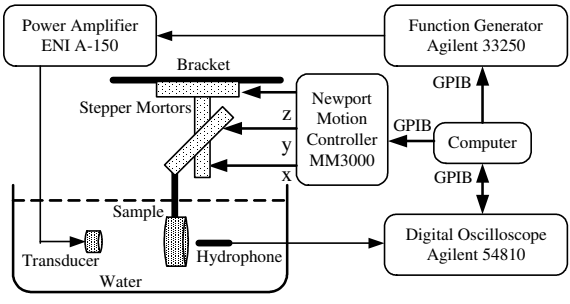


Figure 9. The diagram of experimental set-up.

Figure 10 shows a three-layer model B with a porcine fatty tissue located at both the external ring and the internal circle, and a porcine liver tissue located at the middle ring. The corresponding fundamental, second harmonic and super-harmonic images are presented in Fig. 11, in which the different gray scales correspond to different biological tissues due to the differences in the amplitudes of selected component of the received signal. Because the beam pattern at the super-harmonic frequency is improved (–3 dB beam width at the fundamental frequency is 1.3 times greater than that at the second order harmonic component and 1.7 times greater than that at the super-harmonic frequency), the super-harmonic image exhibits better spatial resolution than the fundamental or the second harmonic images. In the fundamental frequency image Fig. 11(a), it is

observed that the edges of different tissues are vague and there exists an intergradation of about 2 mm from the poor spatial resolution at the fundamental frequency. With an improvement in spatial resolution, the second harmonic images demonstrate recognizable tissue combinations as shown in Fig. 11(b). The edges of the liver and fatty tissues are confirmed, but the lower amplitude of the second harmonic results in less contrast of the image. In the super-harmonic image as shown in Fig. 11(c), both the dimensions and positions of tissues are confirmed by the sharp borders and the contrast of the images is improved. In addition, the speckle noise in the fatty tissue is reduced.

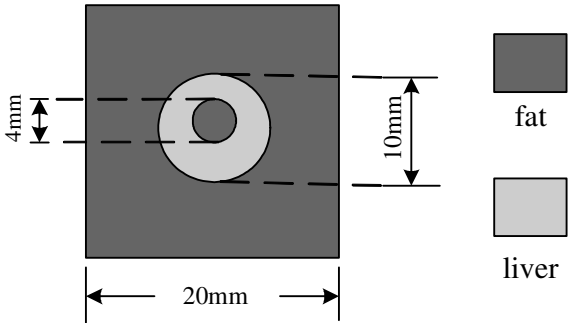


Figure 10. Cross section of sample model.

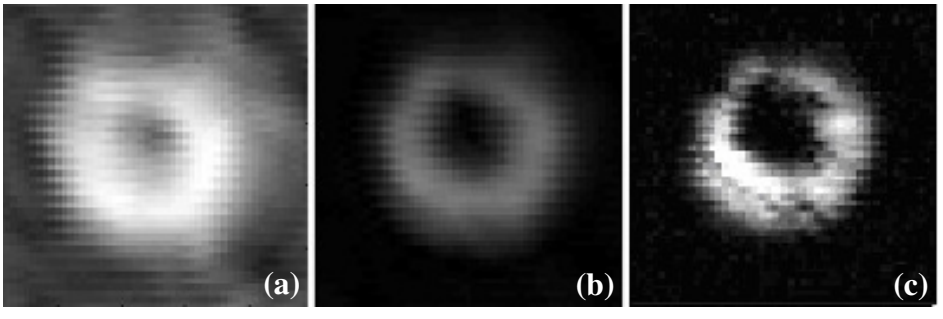


Figure 11. Images of sample model: (a) 1st imaging, (b) 2nd image, (c) super-harmonic image.

5. Conclusion

This chapter reviewed some researches on the acoustic nonlinear imaging in our laboratory. Further assessment of those techniques in clinical application is needed.

References

1. T. Christopher, Finite amplitude distortion-based inhomogeneous pulse echo ultrasonic imaging, IEEE UFFC, 44, 125-139 (1997).
2. C. C. Shen and P. C. Li, Harmonic leakage and image quality degradation in tissue harmonic imaging, IEEE UFFC, 48, 728-736 (2001).
3. P. Jiang, Z. Mao, and J. C. Lazenby, A new tissue harmonic imaging scheme with better fundamental frequency cancellation and higher signal-to-noise ratio, IEEE Ultrasonics Symposium, 1589-1594 (1998).
4. T. G. Muir, and E. L. Carstensen, Prediction of nonlinear acoustic effects at biomedical frequencies and intensities, Ultrasound in Med. & Biol., 4, 345-357 (1980).
5. W. K. Carstensen, W. Law, N. D. Mckay, and T. G. Muir, Demonstration of nonlinear acoustical effects at biological frequencies and intensities, Ultrasound in Med. & Biol., 6, 359-368 (1980).
6. W. K. Cobb, Finite amplitude method for the determination of the acoustic nonlinearity parameter B/A, J. Acoust. Soc. Am. 73, 1525-1531 (1983).
7. X. F. Gong, Z. M. Zhu, T. Shi and J. H. Huang, Determination of acoustic nonlinearity parameter in biological media using FAIS and ITD methods, J. Acoust. Soc. Am. 86, 1-5 (1989).
8. C. M. Seghal, R. C. Bahn, and J. F. Greenleaf, Measurement of the acoustic nonlinearity parameter B/A in human tissues by thermodynamic method, J. Acoust. Soc. Am., 76, 1023-1029 (1984).
9. X. F. Gong, R. Feng, C. Y. Zhu, and T. Shi, Ultrasonic investigation of the nonlinearity parameter B/A in biological media, J. Acoust. Soc. Am., 76, 949-950 (1984)
10. J. Zhang, M. S. Kuhlenschmidt, and F. R Dunn. Influences of structural factors of biological media on the acoustic nonlinearity parameter B/A. J. Acoust. Soc. Am. 1991, 99: 2397-2402.
11. J. Zhang, and F. Dunn, In vivo B/A determination in a mammalian organ, J. Acoust. Soc. Am., 81, 1635-1637 (1987)
12. C. M. Sehgal , G. M. Brown, R. C. Bahn, and J. F Greenleaf, Measurement and use of acoustic nonlinearity and sound speed to estimate composition of excised livers. Ultrasound in Med. & Biol. 12: 865-874. (1986).
13. R. L. Errabolu, C. M. Sehgal, R. C. Bahn, and J. F. Greenleaf, Ultrasound in Med. & Biol. 14 (1988) 137.
14. K. Yoshizumi, T. Sato, and N. Ichida, J. Acoust. Soc. Am. 82 (1987) 302.
15. N. Ichida, T. Sato, and M. Linzer, Imaging the nonlinear parameter of a medium. Ultrason. Imag. 5: 295-299 (1983).
16. N. Ichida, S. Takuso, H. Miwa, and K. Murakami, IEEE trans. Son. Ultrason. SU-31 (1984) 635.
17. C. A. Cain, Ultrasonic reflection mode imaging of nonlinear parameter B/A: theoretical basis, J. Acoust. Soc. Am., 80, 28-32 (1986).
18. F. Tranquart, N. Grenier, V. Eder, and L. Pourcelot, Clinical use of ultrasound tissue harmonic imaging, Ultrasound in Med. & Biol., 6, 889-894 (1999).
19. S. Tanaka, O. Oshikawa, T. Sasaki, T. Ioka, and H. Tsukuma, Evaluation of tissue harmonic imaging for the diagnosis of focal liver lesions, Ultrasound in Med. & Biol., 26, 183-187 (2000).

20. D. Zhang, and X .F. Gong, Experimental investigation of acoustic nonlinearity parameter tomography for excised pathological biological tissues, Ultrasound in Med. & Biol. 25, 593-599 (1999)
21. D. Zhang, X. F. Gong, and S. G. Ye, Acoustic nonlinearity tomography for biological specimens via measurements of the second harmonic wave, J. Acoust. Soc. Am. 99, 2397-2402 (1996).
22. D. Zhang, X. F. Gong, and X. Chen, Experimental imaging of acoustic nonlinearity parameter B/A for biological tissues via a parametric array. Ultrasound in Med. & Biol.; 27: 1359-1365 (2001).
23. Q. Y. Ma, D. Zhang, X. F. Gong, and Y. Ma, J. Acoust. Soc. Am. 119 (2006) 2518.
24. Q. Y. Ma, Y. Ma, and X. F. Gong, et al. Improvement of tissue harmonic imaging using pulse inversion technique. Ultrasound in Med. & Biol. 2005, 31: 889-894.
25. A. Bouakaz, and N. de Jong, "Native tissue imaging at superharmonic frequencies," IEEE Trans. Ultrason., Ferroelect., Freq. Contr., 44, 125-139 (2003)
26. R. T. Beyer, Nonlinear Acoustics, Naval Ship System Command, Dept. of the Navy, Washington DC, WA, USA, (1974) 91.
27. V. Kuznetsov. Equation of nonlinear acoustics. Sov. Phys. Acoust. 1970, 16: 467-470
28. V. A. Khokhlova, R. Souchon, and J. Tavakkoli, et al. Numerical modeling of finite-amplitude sound beams: Shock formation in the near field of a cw plane piston source. J. Acoust. Soc. Am. 2001, 110: 95-108

This page intentionally left blank

CHAPTER 7

ANGULAR SPECTRUM DECOMPOSITION ANALYSIS OF SECOND HARMONIC ULTRASOUND PROPAGATION AND ITS RELATION TO TISSUE HARMONIC IMAGING

Xiang Yan and Mark F. Hamilton

Department of Mechanical Engineering, The University of Texas at Austin
Austin, Texas 78712-1063
xiang.yan@olympusndt.com

An angular spectrum decomposition method (ASM) for calculating second-harmonic propagation in fluid-like media is discussed. This method is based on a perturbation solution of Westervelt's equation, which can be augmented to take into account inhomogeneity of the medium. ASM can be applied to the analysis of tissue harmonic imaging (THI), a diagnostic method that utilizes second harmonics for imaging. One of the advantages of THI is the second harmonic's ability to improve image quality by reducing aberration caused by inhomogeneity in the body wall. The body wall aberration can be modeled as a localized phase screen in the wave propagation path. The results show that the interactions of angular spectrum components of the fundamental field generate a second-harmonic field that is relatively immune to the aberration. An experiment was performed to demonstrate this phenomenon by using a phase screen created from low density Polyethylene material. The measurements were in close agreement with calculations using ASM.

1. Introduction

When ultrasound propagates in fluid-like medium, due to nonlinearity inherent in the equations of motion and state, the waveform distorts progressively and new frequency components are generated. These nonlinear phenomena magnify as the amplitude of the ultrasound wave increases. The mechanism of nonlinear generation has been widely investigated and numerous model equations have been proposed under different conditions.[1] Westervelt's equation is chosen for our analysis for its inclusion of three dimensional diffraction, nonlinearity and

absorption.[1] The solution of the Westervelt equation is solved under quasilinear conditions, in which the second-harmonic field level is one order smaller than the field level at the source frequency. When the ultrasound is transmitted at a single frequency f, the Westervelt equation can be written in the frequency domain, where the fundamental component propagates according to a homogeneous Helmholtz equation, and the second harmonic propagates according to an inhomogeneous Helmholtz equation with its forcing function determined by the fundamental field. These two equations can be solved separately by using a 2D spatial Fourier transform in the plane perpendicular to the propagation axis. This transform is called the angular spectrum of the wave field. The angular spectrum decomposition method (ASM) was first introduced in nonlinear acoustics by Alais and Hennion to analyze sum- and difference frequency generation by a nonlinear parametric array.[2] Subsequently it has been used to analyze reflection and refraction of the second-harmonic beam at a fluid-solid boundary.[3] In this chapter, the method is used to analyze second-harmonic propagation when inhomogeneity is present in the medium.

A recent and very successful application of nonlinear acoustics is use of the second harmonics in diagnostic ultrasound, a modality frequently referred to as tissue harmonic imaging (THI). In THI, the source transmits an ultrasound pulse at frequency f, and the receiver registers the signal at $2f$. Although the second-harmonic amplitude is usually an order of magnitude smaller than that of the fundamental component, it is sufficiently high to be measured under conditions encountered in clinical cases.[4,5,6] Advances in ultrasound source design, detection instruments, and signal processing capability enable THI to be feasible.

THI is capable of improving image resolution due to its narrow main lobe and reduced side lobe reverberations.[5,6,8,9] A rule of thumb is that the main lobe directivity of the $2f$ beam can be approximated as the square of the main lobe directivity of the f beam. Second harmonics are also capable of reducing artifacts and haze due to their relative immunity to phase aberrations coming from intervening tissue inhomogeneity.[5,8] In conventional imaging (linear acoustics), the inhomogeneity introduces phase aberration in the f beam, which degrades the main lobe directivity and increases the side-lobe levels. But the self-interaction of the f beam, from which the second-harmonic ($2f$) is generated, is capable of filtering out these aberrations.[5] The generation of the second harmonic is strongest in regions where the fundamental field is strongest. The low side-lobe amplitudes are not sufficient to generate a second harmonic very efficiently, and therefore side lobe suppression is usually higher for the second harmonic beam than in the fundamental beam. Many numerical and experimental investigations have been conducted to assess quantitatively the nonlinear beam quality with

aberrations in the medium, corresponding either to real inhomogeneous tissues[6,8] or constructed from tissue phantoms.[9,10,11]

In the present chapter, we analyze second-harmonic generation and propagation using ASM taking into account inhomogeneity (phase aberrations) between the source and receiver. Measurements and simulations are focused on the distorted beam patterns introduced by the phase aberration. The aberration screen was constructed to provide phase variations having random distributions over the plane perpendicular to the propagation axis. The phase screen possesses statistics that are similar to those provided by Hinkleman based on measurements of human abdominal wall.[12] The simulated beam patterns obtained through ASM are compared with measurements. Finally, a quantitative measure called focusing quality, originally proposed by Tabei[13] is used to characterize the fundamental and second harmonic beam patterns.

2. Angular Spectrum Decomposition Method

In THI, the transmitted ultrasound is a short pulse instead of a single frequency. The frequency of the pulse is tuned properly to cover sufficient bandwidth in order to provide good axial resolution. On the other hand, the bandwidth should be sufficiently narrow to avoid overlapping of the fundamental and second-harmonic spectra.[5] To compare beam pattern distortion at two discrete frequencies *f* and 2*f*, a frequency domain analysis is efficient. Here we will assume that the source transmits a single frequency *f*, which greatly simplifies solution of the Westervelt equation.

The Westervelt equation, which describes the nonlinear propagation of ultrasound waves in an absorbing medium, is written as[1]

$$\nabla^2 p - \frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2} + \frac{\delta}{c_0^4}\frac{\partial^3 p}{\partial t^3} = -\frac{\beta}{\rho_0 c_0^4}\frac{\partial^2 p^2}{\partial t^2} \tag{1}$$

where $\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ and $p, \rho_0, c_0, \delta, \beta$ represent acoustic pressure, medium density, sound speed, sound diffusivity, and coefficient of nonlinearity, respectively. The right-hand side of Eqn. (1) accounts for the harmonic generation and is considered to be a small correction to the linear wave equation. We let $p = p_1 + p_2$, where $|p_1| >> |p_2|$, p_1 is the sound pressure at the fundamental frequency *f*, and is the sound pressure for the second harmonic, at frequency *2f*. The higher harmonics are neglected due to their much smaller amplitudes and higher attenuations. Equation (1) is expanded according to the smallness of p_1 and p_2, and they are, individually, described by Helmholtz equations. Introducing the following notation,

$$p_n(\mathbf{r},z,t)=\frac{1}{2j}P_n(\mathbf{r},z)e^{jn\omega t}+\text{c.c.},\quad \text{n}=1,2.\quad \omega=2\pi f \tag{2}$$

where $\mathbf{r}=(x,y)$, we arrive two equations describing the complex pressure $P_n(\mathbf{r},z)$:

$$\nabla^2 P_1+k_1^2 P_1=0 \tag{3}$$

$$\nabla^2 P_2+k_2^2 P_2=-\frac{j2\beta k^2}{\rho_0 c_0^2}P_1^2 \tag{4}$$

where $k_1^2=k^2-j\delta k^3/c_0$, $k_2^2=4k^2-j8\delta k^3/c_0$, and $k=\omega/c_0$. Equations (3) and (4) are solved using ASM. The following 2D spatial Fourier transform pair is introduced:

$$\hat{P}(\boldsymbol{\kappa},z)=F\{P(\mathbf{r},z)\}=\int P(\mathbf{r},z)e^{j\boldsymbol{\kappa}\bullet\mathbf{r}}d\mathbf{r} \tag{5}$$

$$P(\mathbf{r},z)=F^{-1}\{\hat{P}(\boldsymbol{\kappa},z)\}=\int \hat{P}(\boldsymbol{\kappa},z)e^{-j\boldsymbol{\kappa}\bullet\mathbf{r}}\frac{d\boldsymbol{\kappa}}{(2\pi)^2} \tag{6}$$

where $\boldsymbol{\kappa}=(k_x,k_y)$ is wave number vector (see Figure 1). The Fourier transform pair shows that any pressure field $P(\mathbf{r},z)$ can be viewed as a superposition of planes waves that propagate at angles determined by $\boldsymbol{\kappa}$. The amplitude of each plane wave is determined by the angular spectrum $\hat{P}(\boldsymbol{\kappa},z)$. Talking the Fourier transforms of Eqns. (3) and (4), we arrive at the solution for P_1:

$$P_1(\mathbf{r},z)=F^{-1}\left\{\hat{P}_1(\boldsymbol{\kappa},z_0)e^{-jk_{z1}(z-z_0)}\right\} \tag{7}$$

with

$$k_{z1}=\sqrt{k_1^2-|\boldsymbol{\kappa}|^2}=\sqrt{k^2-|\boldsymbol{\kappa}|^2-j\delta k^3/c_0} \tag{8}$$

where $\hat{P}_1(\boldsymbol{\kappa},z_0)$ is the known angular spectrum distribution at a given propagation distance , and is at the target plane z, with $z>z_0$. The solution for P_2 is composed of two parts, P_2^h, which accounts for linear propagation of $P_2(\mathbf{r},z_0)$ from z_0 to z, and P_2^p, which is the nonlinear generation from $P_1(\mathbf{r},z)$ from z_0 to z. The first part, P_2^h, has a form similar to Eq. (7), and is obtained by changing k to k. The solution for P_2^p comes from the nonlinear interactions of plane waves represented by $\hat{P}_1(\boldsymbol{\kappa},z_0)$, which is found to be[3,15]

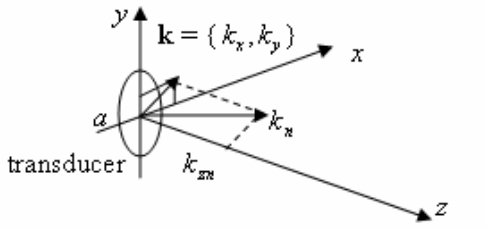


Figure 1. Geometry of ultrasound beam propagation, transducer is located in plane, propagation is along axis.

$$P_2^p(\mathbf{r},z)=F^{-1}\left\{\hat{P}_2^p(\boldsymbol{\kappa},z)e^{-jk_{z2}(z-z_0)}\right\} \tag{9}$$

$$\hat{P}_2^p(\boldsymbol{\kappa},z)=\frac{j\beta k^2}{2\pi^2\rho_0 c_0^2}\int\frac{\hat{P}_1(\boldsymbol{\kappa}',z_0)\hat{P}_1(\boldsymbol{\kappa}-\boldsymbol{\kappa}',z_0)}{(k_a+k_b)^2-k_{z2}^2}\left[e^{-j(k_a+k_b-k_{z2})(z-z_0)}-1\right]d\boldsymbol{\kappa}', \tag{10}$$

with

$$k_a=\sqrt{k^2-|\boldsymbol{\kappa}'|^2-j\delta k^3/c_0}, \quad k_b=\sqrt{k^2-|\boldsymbol{\kappa}-\boldsymbol{\kappa}'|^2-j\delta k^3/c_0} \tag{11}$$

and

$$k_{z2}=\sqrt{4k^2-|\boldsymbol{\kappa}|^2-j8\delta k^3/c_0} \tag{12}$$

The total solution of $P_2(\mathbf{r},z)$ is the sum of $P_2^h(\mathbf{r},z)$ and $P_2^p(\mathbf{r},z)$, i.e.,

$$P_2(\mathbf{r},z)=F^{-1}\left\{\left(\hat{P}_2^h(\boldsymbol{\kappa},z_0)+\hat{P}_2^p(\boldsymbol{\kappa},z)\right)e^{-jk_{z2}(z-z_0)}\right\} \tag{13}$$

From Eqns. (10) ~ (12) one can see that the angular spectrum for the second harmonic $\hat{P}_2(\boldsymbol{\kappa},z)$ broadens due to the doubling of frequency and convolution-like filtering with $\hat{P}_1(\boldsymbol{\kappa},z_0)$. A broader spectrum indicates a narrower beam width. Equations (12) and (13) also indicate that the second harmonic spectrum at distance z depends only on the fundamental spectrum at distance z_0, assuming the intervening medium is homogeneous. This facilitates the calculation of second harmonics since there are no intermediate calculations are needed. Although the second harmonic is generated cumulatively from z_0 to z, the perturbation assumption and frequency domain calculation eliminate any intermediate steps.

It is also possible to include inhomogeneity along the propagation path, provided these regions are localized and lend themselves to modeling with phase screens. In THI, significant inhomogeneity comes from the scattering from septa in the subcutaneous fat in the body wall.[16] This localized aberration causes a phase change that can be modeled as phase screen,[12] which makes ASM suitable

for analyzing THI. Suppose the phase screen is situated at z_s and introduces a phase modulation to $P_n(\mathbf{r}, z_s)$. The ASM computation then separates into two parts. First, the pressures $P_n(\mathbf{r}, z)$ are calculated from $z = 0$ to $z = z_s$ using Eqs. (7) and (10), with $P_2^h(\mathbf{r},0) = 0$. Then, at z_s, phase aberrations are applied to $P_n(\mathbf{r}, z)$ and $P_n(\mathbf{r}, z)$, which then propagate from z_s to z individually. Notice that the phase change doubles for the second harmonic due to the wavelength being half that of the fundamental. If medium inhomogeneity is distributed along the propagation path, multiple phase screens can be applied and calculations can be done from screen to screen.

Early models for THI assume that the tissue is homogenous.[6] Later models taking inhomogeneity into account involved numerical simulations with nonlinear propagation algorithms that march the wave field forward incrementally step by step. Christopher[8] modeled the body wall inhomogeneity with phase screens and propagated the wave field incrementally thereafter. A time-domain algorithm for solving the KZK algorithm has also been used to model inhomogeneity with terms included that account for continuous variations of the sound speed and density.[11] All results show that even in the presence of aberrations, the second-harmonic field still provides better beam quality than the fundamental field. The second-harmonic field exhibits a tighter main lobe and lower side lobes than the fundamental beam.

3. Measurement

In this chapter, the second-harmonic field is calculated using ASM with inhomogeneity taken into account by phase screens. Experiments were performed and compared with calculations. The experimental setup is shown in Figure 2. A focused circular piston source (Panametrics Acuscan series) with center frequency 2.25 MHz, diameter 38 mm and focal length 152 mm, was placed in a water tank in front of a phase screen that was created by machining a uniform 3 mm-thick low-density polyethylene plate to varying thickness over one of its surfaces. The focal beam patterns were measured using a PVDF membrane hydrophone. The hydrophone was used to scan over the plane perpendicular to the propagation axis. The source, hydrophone and plate were all immersed in a water tank containing de-ionized, degassed fresh water.

Random indentations were engraved on the plate to introduce phase aberration. They were small circular cuts that were randomly distributed on the surface according to a desired statistical distribution. All the circles had the same 1 mm indentation. This indentation introduced a 2.7 rad phase shift at 2.25 MHz. Two patterns were created with different circle diameters and different spatial

distributions, as shown in Figure 3. Screen A had 200 small circles with diameter of 3.18 mm distributed over an area of 100 mm x 100 mm. Screen B had 40 circles with diameter 12.7 mm distributed over the same area. Their inhomogeneity length scales are described by the characteristic lengths of the 2D autocorrelation functions (correlation lengths). This is the interval between the half maximum amplitude points of the autocorrelation function. Screen A had an inhomogeneity length scale of about 3 mm, and screen B of about 11.7 mm, which are close to the diameters of the circles. These two values correspond to the measured minimum and maximum inhomogeneity length scales of the human abdominal wall.[17]

In addition to introducing phase changes, the phase screen also introduces amplitude aberrations due to the impedance mismatch and absorption. It is estimated that the amplitude variation across the phase screen is within 10%. This amplitude perturbation has a substantially smaller effect on the wave field than the phase changes, and it can be neglected without introducing significant error.[15] The insertion loss introduced by the plate can be compensated for by increasing the drive level. The finite thickness of the phase screen (maximum 3 mm) is about 3 wavelengths of the fundamental field. The refraction introduced by the plate can slightly reduce the focal distance, but in the case of a beam at normal incidence, the effect is found to be negligible. Under conditions of the experiment, the plate is was as a phase screen with zero thickness, introducing uniform amplitude variation and random phase in the calculations.

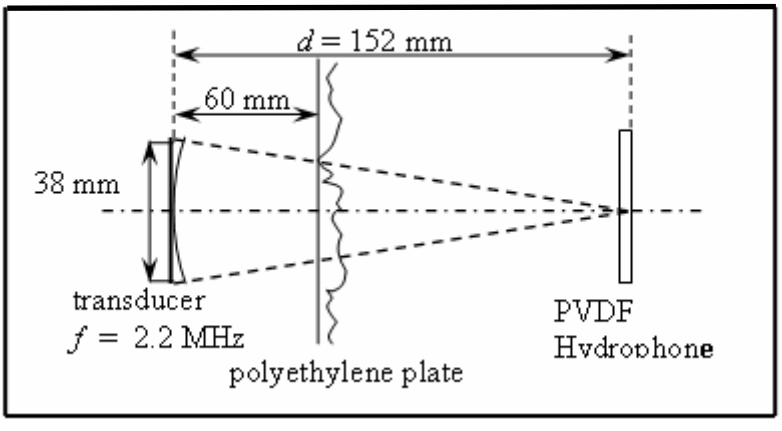


Figure 2. Experiment set-up. A phase screen made from low-density polyethylene plate is placed in front of a focused transmitter to simulate the aberration caused by body wall.

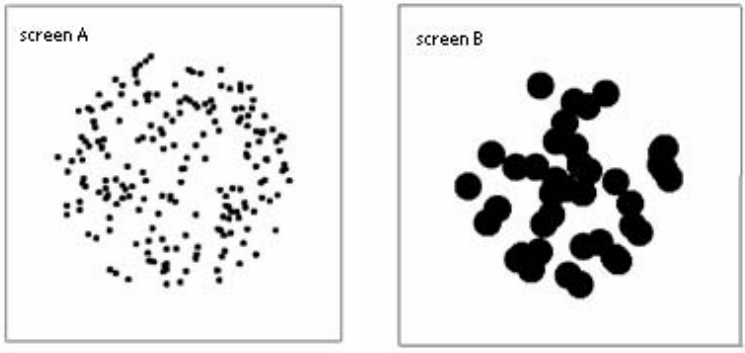


Figure 3. Geometry of circle indentations in the polyethylene plates used as the phase screens.

In the measurements, the transmitted signals are monofrequency pulses (around 15 cycles) at 2.25 MHz generated by a HP 3314A function generator. They were amplified by 50 dB before being sent to the transducer. The signal received by the hydrophone was amplified by 20 dB before digitization. The digitized waveform was finally sent to the computer to be analyzed. A time domain FFT was used to extract the spectrum information from the waveforms.

4. Results

The focal beam patterns for the fundamental and second harmonic fields were measured first without the phase screen. The results are shown in Figure 4. They are compared with the ASM calculation using Eqns. (7), (9) and (10), and it is seen that good agreement is achieved. The beam patterns exhibit good axial symmetry. The measured beam patterns do not show deep nulls as compared to the simulations due to the deviation of the source condition from an ideal piston and the finite frequency bandwidth of the CW pulse. The calculated angular spectra of the beam patterns are shown in Figure 5. The spatial frequencies are normalized to the wavelength, i.e. λf_x (at 2.25MHz), where $f_x = k_x / 2\pi$. Figure 5 shows that the angular spectrum of the fundamental (Figure 5(a)) has isolated ring-type components, whereas that of the second harmonic (Figure 5(b)) is more continuously distributed and spreads out over a larger domain, which explains the narrow main lobe for the second harmonics. Continuous spectral distributions of the second harmonics result from the convolution-like interactions of fundamental spectrum as shown in Eqn. (10), which acts like a smooth filter. The resulting pressure field, after applying the inverse FFT to the angular spectrum, possesses reduced side lobes.

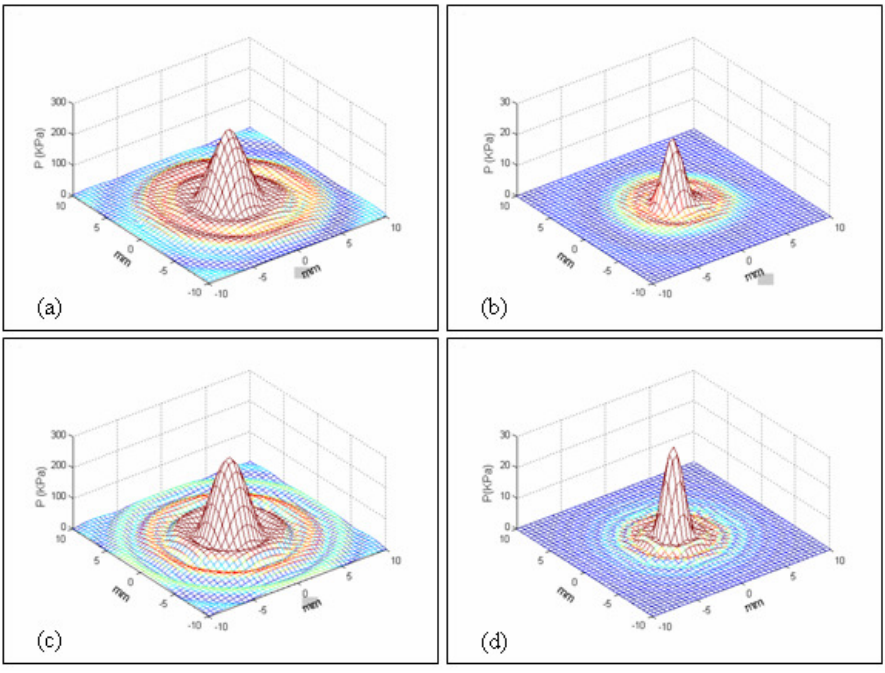


Figure 4. Measured fundamental (a) and second harmonic (b) beam patterns compared to calculated fundamental (c) and second harmonic (d) beam patterns using angular spectrum method in the absence of phase screen.

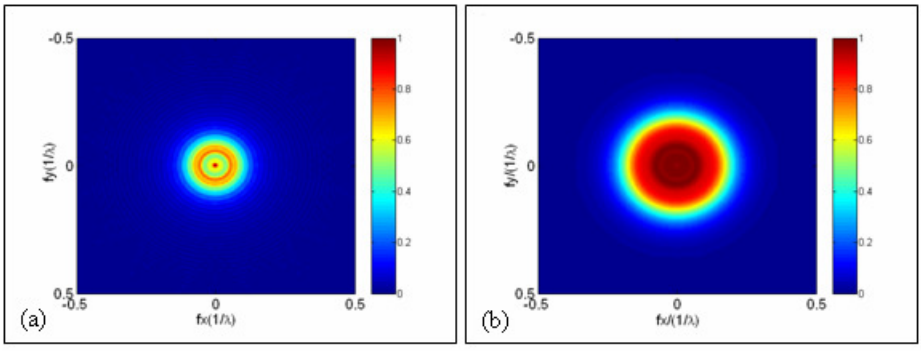


Figure 5. Calculated normalized angular spectrum distribution for fundamental (a) and second harmonic (b) beam pattern with no phase aberration.

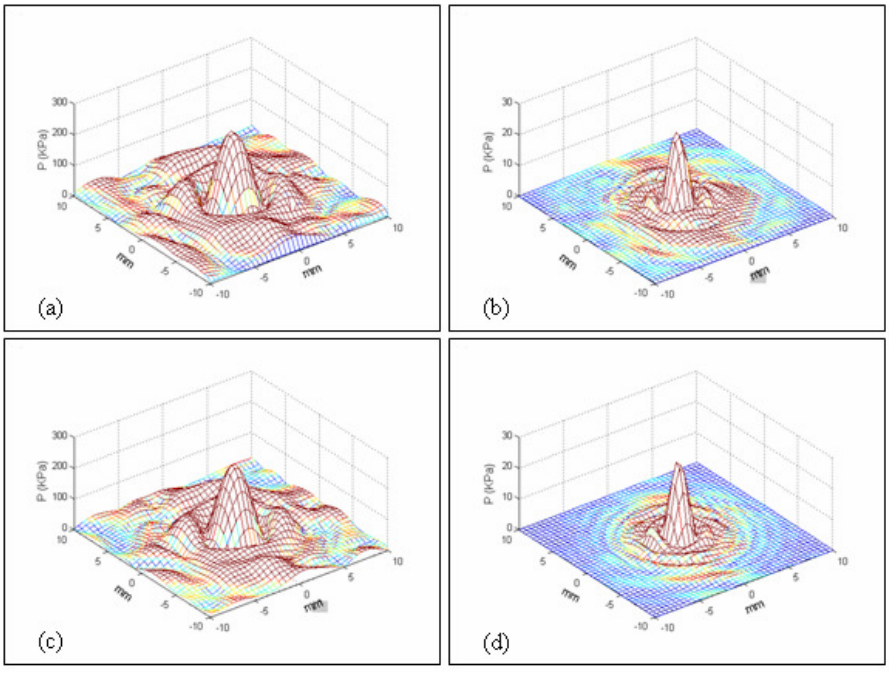


Figure 6. Measured fundamental (a) and second harmonic (b) beam patterns compared to calculated fundamental (c) and second harmonic (d) beam with phase screen A in the propagation path.

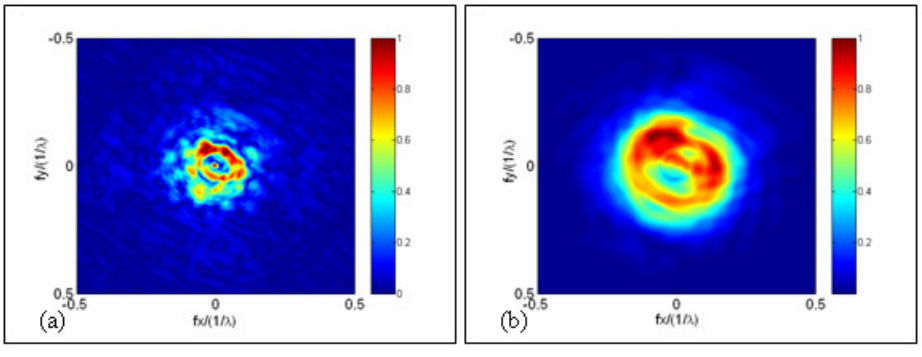


Figure 7. Calculated normalized angular spectrum distribution for fundamental (a) and second harmonic (b) beam pattern with phase screen B.

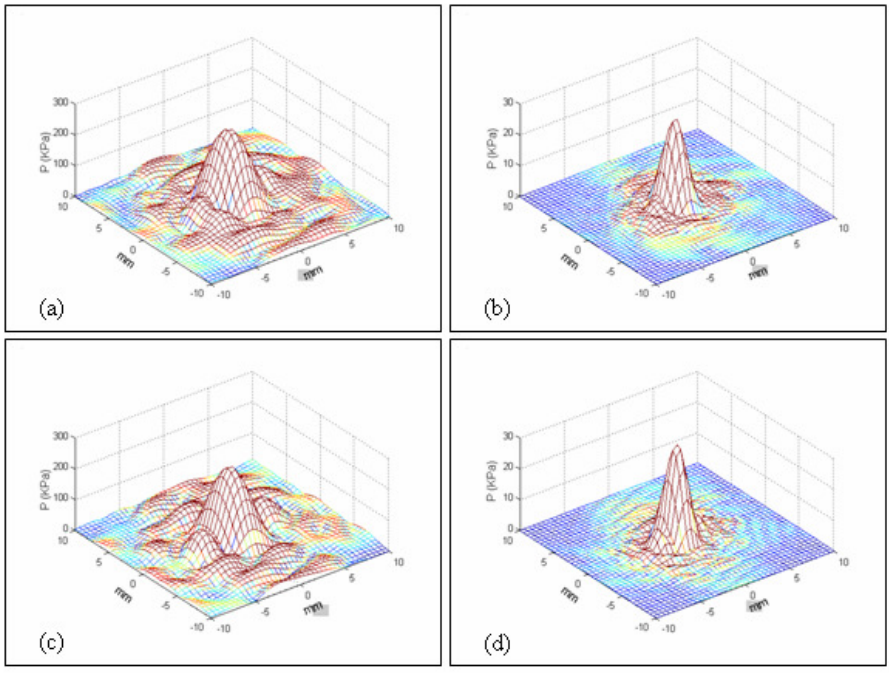


Figure 8. Measured fundamental (a) and second harmonic (c) beam patterns compared to calculated fundamental (b) and second harmonic (d) beam with phase screen B in the propagation path.

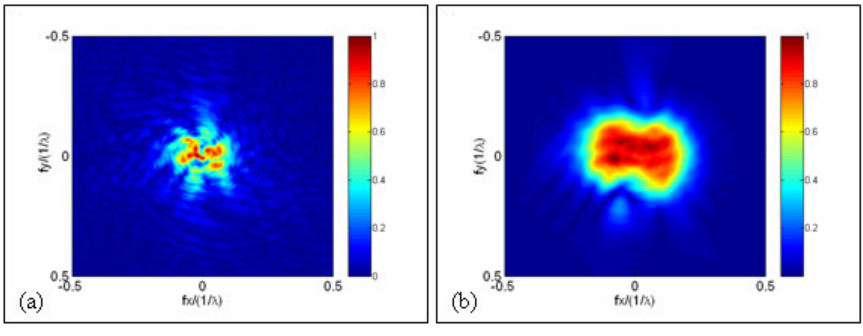


Figure 9. Calculated normalized angular spectrum distribution for fundamental (a) and second harmonic (b) beam pattern with phase screen B.

For the case in which phase screen A is placed between the transmitter and receiver, as shown in Figure 2, the measured and simulated focal beam patterns are as shown in Figure 6. One sees readily that the scattered side lobe levels are greatly increased, while the main lobes are less affected. The ASM simulated beam patterns predict these results and agree well with the measurements. The angular spectra are shown in Figure 7. When comparing Figure 7(a) with Figure 5(a), one can see there is a significant amount of energy scattered to the higher angular frequencies. The frequency distribution has isolated patches, with a small amount of energy concentrated around the dc component, and a ring structure is observed at high spatial frequencies. The interference of the high frequency spectrum with the low frequency spectrum increases side lobe levels in the beam pattern, and this in turn leads to increased reverberation. The spectrum for the second harmonic, Figure 7(b), exhibits similar distortion, also leading to increased side lobe levels. But compared to Figure 7(a), the second harmonic spectrum is more continuously distributed and the corresponding side lobe levels are obviously lower than fundamental beam.

The focal beam patterns resulting from use of screen B are shown in Figure 8, and their angular spectra are shown in Figure 9. Screen B has a larger inhomogeneity length scale that scatters less energy into the high spatial frequencies in comparison to screen A, and their spectra clearly illustrate this point. The second harmonic spectrum, Figure 9(b), almost maintains the same shape and continuity as in Figure 5(b).

The measured width of the main lobe (3-dB beam width) and maximum side lobe levels for fundamental and second harmonics are presented in Table 1.

Table 1. 3-dB main lobe beam width and maximum side lobe levels for the measured fundamental and second harmonic beam patterns.

	3dB main lobe width		Maximum side lobe level	
	fundamental	2nd harmonic	fundamental	2nd harmonic
No screen	2.9 mm	1.7 mm	-17.6 dB	-26.2 dB
Screen A	2.8 mm	1.5 mm	-8.2 dB	-12.2 dB
Screen B	3.7 mm	2.0 mm	-10.0 dB	-15.0 dB

Another quantitative measure for a focused beam pattern, referred to as the focusing quality, was proposed by Tabei[13] to characterize the degradation of the focus caused by aberration. The focusing quality is defined as the reciprocal of the Peripheral Energy Ratio (PER), i.e., the ratio of acoustic energy in the peripheral area to the energy in the focal area,

$$Q = E_{\text{foc}} / E_{\text{peri}} \tag{14}$$

The focal area is defined as the region where amplitude is within half of the peak value in the beam, and the peripheral area is the rest of the measurement area outside this focal area. The focusing quality Q for the measured beam patterns are shown in Table 2. One can see that Q is decreased in the presence of aberration. The value of Q for second harmonic beam is always larger than the value for the fundamental. The discrepancy in their values of Q is even larger in the case of the phase screen. This means that the phase screen causes more degradation of the fundamental field than of the second harmonic field.

From the above discussion, one can see that second harmonic generation and propagation is inherently filtered in such a way that its angular spectrum is smoothed and maintains its continuity, even when components in spectrum of the fundamental are scattered into isolated bands (in the spatial frequency domain) due to inhomogeneity. The filtering process arises from the nonlinearity of the hydrodynamics and the state equations that characterize the pressure in the propagation medium as indicated by Eqn. (4) in space domain. The continuity of its angular spectrum ensures that the second harmonic side lobes are at a low level. At the same time, the width of the main lobe is reduced due to doubling of the frequency. The main difference between the nonlinearly generated second harmonic and the linear second harmonic is the side lobe level difference. The linear wave at the second harmonic ($2f$) has a narrower main lobe and more side lobes, and the inhomogeneity creates twice the phase variation at $2f$, which drives the side lobes to higher levels than in the fundamental beam.[15] Therefore the enhanced scan resolution of the linear $2f$ beam is at the cost of increased artifacts and haze. The nonlinear second harmonics thus provide enhanced scan resolution and reduced artifacts. The cost is the lower amplitude of the signal, and therefore lower signal-to-noise ratio, leading to the requirement of a more sensitive receiver and lower noise amplifier.

Table 2. Focusing quality Q for the measured fundamental and second harmonic beam patterns

	Focusing Quality Q	
	Fundamental	2nd harmonic
No screen	2.0	2.4
Screen A	0.7	1.3
Screen B	1.4	2.3

5. Conclusions

In this chapter we investigated the angular spectrum of the second harmonic arising from the nonlinear propagation of an ultrasonic wave in fluid-like medium. The angular spectrum in the observation plane can be described as an integral over the angular spectrum of fundamental component in the source plane, provided the intervening medium is homogeneous. This method is adapted to account for inhomogeneity of the medium encountered during THI. The inhomogeneity arising from the body wall is modeled as a phase screen. Measurements and corresponding simulations have been presented for the scattered beam patterns radiated from a focused piston source through an artificial random phase-screen. The statistics of the phase screen correspond to the measured human abdominal wall statistics. The simulation employing ASM agrees well with the measurements, and the reduced side lobe reverberation of the second harmonics is explained from the angular spectrum point of view. The focusing quality factor quantifies the beam degradation and shows improvement of second harmonic beam over fundamental beam. This provides a theoretical basis to explain the reduced artifacts and haze in the image quality in THI.

References

1. M. F. Hamilton, in *Nonlinear Acoustics*, Ed. M. F. Hamilton and D. T. Blackstock (Academic Press, Boston, 1998).
2. P. Alais and P. Y. Hennion, *Acustica*, 1 (1979).
3. B. J. Landsberger and M. F. Hamilton, *J. Acoust. Soc. Am.*, 488 (2001).
4. M. A. Averkiou, *IEEE Ultrasonics Symposium*, 1563 (2000).
5. M. A. Averkiou, *J. Acoust. Soc. Am.*, 2435 (2003).
6. M. A. Averkiou, D. N. Roundhill, and J. E. Powers, *IEEE Ultrasonics Symposium*, 1561 (1997).
7. B. Ward, A. C. Baker and V. F. Humphrey, *J. Acoust. Soc. Am.*, 143 (1997).
8. P. T. Christopher, *IEEE Trans. Ultrason. Ferroelctr. Freq. Control*, 125 (1997).
9. G. Wojcik, J. Mould, S. Ayter, and L. Carcione, *IEEE Ultrasonics Symposium*, 1583 (1998).
10. P. T. Christopher, *IEEE Trans. Ultrason. Ferroelctr. Freq. Control*, 158 (1998).
11. Y. Jing, X. Yang and R. Cleveland, *J. Acoust. Soc. Am.*, 2324 (2003).
12. L. M. Hinkelman, D.-L. Liu, L. A. Metlay, and R. C. Waag, *J. Acoust. Soc. Am.*, 530 (1994).
13. M. Tabei, T. D. Mast, and R. C. Waag, *J. Acoust. Soc. Am.*, 1166 (2003).
14. M. F. Hamilton, in *Nonlinear Acoustics*, Ed. M. F. Hamilton and D. T. Blackstock (Academic Press, Boston, 1998).
15. X. Yan, *Ph.D. dissertation, The University of Texas at Austin*, (2004).
16. T. D. Mast, L. M. Hinkelman, M. J. Orr, V. W. Sparrow and R. C. Waag, *J. Acoust. Soc. Am.*, 1177 (1997).

CHAPTER 8

APPLICATIONS OF HIGH-FREQUENCY ULTRASOUND TO DIAGNOSTIC AND REGENERATIVE MEDICINE

Kendall R. Waters

Volcano Corp, Advanced Technology Laboratory
Cleveland, Ohio, USA
kwaters@volcanocorp.com

High-frequency ultrasound (HFUS) measurement techniques have been developed for application to diagnostic and regenerative medicine. The ultrasonic propagation and scattering properties of several biological materials have been measured over frequency ranges between 15 and 65 MHz. *Ex vivo* measurements of murine pulmonary arterial walls show the potential role of HFUS in monitoring pulmonary arterial hypertension. Similarly, *ex vivo* measurements of bovine cartilage and poly(ethylene glycol) based hydrogels indicate that HFUS techniques may permit real-time, online monitoring of functional tissue-engineered cartilage. Lastly, *ex vivo* measurements of human coronary arteries with atherosclerotic plaque have been performed to develop an *in vivo* tool (VH™ IVUS) for management of coronary artery disease. Additional discussion concerning standards and technologies for HFUS as well as opportunities to improve understanding of the interaction of HFUS and tissue is provided.

1. Introduction

Quantitative ultrasonic characterization of biological materials involves the application of physical acoustics measurement techniques to investigate how the biomechanical environment affects the health or quality of a biological material. Ultrasound-based measurement techniques enable the probing of biological materials at multiple length scales, including the tissue, matrix, and cellular levels. These techniques enable determination of the physical properties of biological materials and their response to physical stimuli such as a mechanical vibration or strain. The ability to monitor the physical properties of biological materials permits investigation of disease mechanisms in natural tissue as well as the bioresponse of synthetic tissues to the mechanical environment.

A goal of diagnostic medicine is to understand the structure and function of normal tissue and how it remodels under disease conditions. Regenerative medicine (or tissue engineering) is a multidisciplinary field that exploits living cells for therapeutic application with the goal of enhancing or restoring tissue function.[1] The spatial and temporal resolutions of state-of-the-art ultrasound imaging technologies are capable of accurately reconstructing anatomical morphology that can provide diagnostically relevant information. In addition to the gross morphology of tissue, structure and composition plays a critical role in its function. Consequently, the ability to evaluate tissue based on structure and composition can potentially provide additional information of clinical relevance.

The majority of medical ultrasound applications utilize frequencies that are below 10 MHz (λ > 150 um). If the goal is to provide microscale resolution (< 100 μm) that can be correlated with histologic findings, it is necessary to have an imaging and measurement device that is capable of such spatial resolutions. However, the attenuation of ultrasound in soft tissue increases exponentially with frequency and, correspondingly, limits the depth of penetration. Consequently, the application of high-frequency ultrasound (HFUS) is naturally limited to applications where propagation lengths and attenuation properties are sufficiently small such that imaging and measurement are possible.

For purposes of this chapter, we will consider HFUS to be >10 MHz (λ < 150 μm). Fundamental principles of ultrasonic propagation and scattering in biological materials are discussed, including the application of ultrasound in order to elucidate of the relationship of structure and function in such materials. Basic and clinical research investigations are described to illustrate the role of quantitative ultrasonic characterization of biological materials in diagnostic and regenerative medicine. These investigations are motivated by the impact on health of pulmonary arterial hypertension, osteoarthritis, and coronary artery disease.

2. Interaction of Ultrasound and Tissue

The structure and composition of tissue is in general anisotropic and heterogeneous and can be described as having multiple length scales of importance. Tissue exhibits both longitudinal mode and shear mode ultrasonic properties, where the longitudinal mode is particularly dominant in soft tissues. The conventional longitudinal mode properties that are considered include linear propagation (speed of sound and attenuation) and scattering properties. The attenuation of ultrasound includes both absorption and scattering, where absorption accounts for the majority of attenuation in soft tissue. Tissue also

exhibits nonlinear propagation properties in which ultrasonic energy at a fundamental frequency can be promoted to harmonic frequencies.[2] These ultrasonic propagation properties can also be related to elastic moduli, such as the bulk and shear moduli, assuming an appropriate structural model of the tissue is available. In short, there are a range of mechanical properties of tissue that are available for investigation of tissue. These properties may be affected during remodeling due to disease processes or by therapeutic treatment and, consequently, are potentially useful for characterization.

The uses of ultrasound for biomedical purposes utilize frequencies from below 1 MHz to over 1 GHz with corresponding wavelengths of millimeters down to micrometers. Figure 1 illustrates this range of frequencies and wavelengths with corresponding tissue length scales and medical ultrasound applications. Of present interest are those frequencies from approximately 10 MHz to 100 MHz (150 μm down to 15 μm wavelength). This corresponds to the tissue and so-called matrix length scales, and it provides an idea of what we ultimately are able to interrogate. The ability to image the morphological features of tissue is certainly of interest and clinical utility, but it is the ability to determine the intrinsic ultrasonic properties of the tissue in a quantitative manner that is of present interest. For purposes of this chapter, we will limit our discussion to linear, longitudinal mode propagation and scattering properties of soft tissues.

2.1. *Propagation Properties of Soft Tissue*

The small-signal propagation of ultrasound through soft tissue can be described by a complex, frequency-dependent wave number

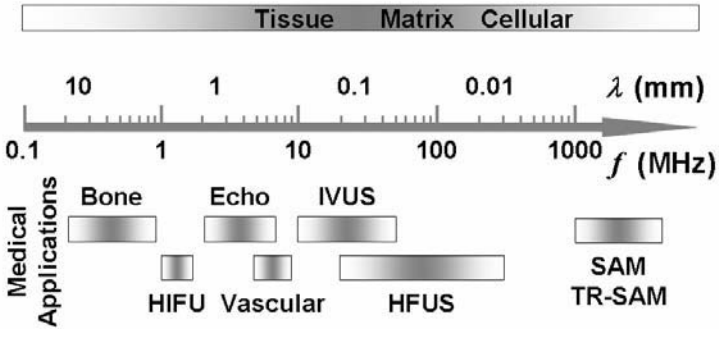


Figure 1. Medical ultrasound applications and techniques and their corresponding frequencies and wavelengths compared to tissue length scales: bone, high-intensity focused ultrasound (HIFU), echocardiography (Echo), non-invasive vascular, intravascular ultrasound (IVUS), high-frequency ultrasound (HFUS), scanning acoustic microscopy (SAM), and time-resolved SAM (TR-SAM).

$$\kappa(\omega) = \omega / c(\omega) + i\alpha(\omega), \tag{1}$$

where ω is angular frequency, c is phase velocity, and α is the (amplitude) attenuation coefficient. The complex wave number κ is spatially dependent, but in practice is generally treated as an effective wave number. An ultrasonic field will propagate through soft tissue with a characteristic phase velocity (or speed). A phase velocity that is only weakly dependent upon frequency

$$c(\omega) \approx c(\omega_0) = c_0 \tag{2}$$

is conventionally referred to as a speed of sound. There is generally not a large amount of variation of the speeds of sound between soft tissues in either normal or diseased states.

An ultrasonic field loses energy as it propagates through soft tissue due to absorptive and scattering processes

$$\alpha(\omega) = \alpha_{abs}(\omega) + \alpha_{scat}(\omega), \tag{3}$$

and the loss due to absorption is generally much larger than that due to scattering $(\alpha_{abs} \gg \alpha_{scat})$. A considerable amount of literature has examined the propagation properties of soft tissue.[3-5]

Two mechanisms that have been proposed for absorption of ultrasound in tissue are viscosity and relaxation in light of the similarities between soft tissues and fluids and polymers. Viscous loss is thought of as a fluid friction and exhibits an absorption per wavelength that is proportional to frequency

$$\alpha_{\text{visc}} \lambda \sim \omega. \tag{4}$$

Relaxation losses exhibit the characteristic form associated with a relaxation frequency ω_r

$$\alpha_{\text{relax}} \lambda \sim \omega \left[1 + (\omega / \omega_r)^2 \right]^{-1}. \tag{5}$$

For low frequencies ($\omega \ll \omega_r$), the absorption due to relaxation is approximately proportional to ω^2. However, soft tissues do not exhibit either strictly viscous or relaxation losses. It has been empirically found that over modest bandwidths (particularly <10 MHz) that the attenuation increases approximately linearly with frequency.[6] Assuming that losses due to scattering are negligible, then the attenuation (or absorption) can be written as

$$\alpha(\omega) = \beta\omega, \tag{6}$$

where β is the conventional slope of attenuation and has been found to be a useful parameter for tissue characterization. This phenomenological description of absorption in soft tissue has been interpreted as a distribution of viscous and

relaxation losses.[7] Alternatively, an hysteretic loss model[8] has been proposed to describe this form of absorption where the loss of energy per cycle is constant and, consequently, the absorption grows linearly with frequency. Nevertheless, it is the slope of attenuation β that is typically reported.

2.2. *Scattering Properties of Soft Tissue*

The scattering properties of soft tissue are of considerable interest[9] due to the fact that there are a limited number of acoustic windows into the human body that permit interrogation of tissue by ultrasound. Specifically, it is the backscatter properties of tissue that are relevant for most clinical applications. Scattering arises due to spatial variations in the acoustic impedance of the tissue. The scattered field from a volume of interest can be written as being proportional to a scattering amplitude function f ,

$$p_{scat}\left(\vec{r},\omega\right) \sim f\left(\hat{k}_{inc},\hat{k}_{scat},\omega\right), \tag{7}$$

where $\hat{k}_{inc}$ is the direction of the incident field, $\hat{k}_{scat}$ is the direction of the scattered field, and $\vec{r}$ is the position where the scattered field is observed. (See Fig. 2.) Because the backscatter properties are of particular importance, one often considers the backscatter coefficient[10] (or differential backscatter cross section)

$$\eta\left(\omega\right) = \left|f\left(0,\pi,\omega\right)\right|^2, \tag{8}$$

which in general depends in part on frequency, the bulk moduli, and the mass densities. Scattering in soft tissue is often described as occurring in the Rayleigh regime where the wavelength of the ultrasound is much larger than the size of the scatterers. In such cases, the so-called frequency dependence of the backscatter coefficient is found to be 3 for ensembles of cylindrical scatterers and 4 for ensembles of spherical scatterers.[11] This behavior is not necessarily expected to be observed for heterogeneous tissues.

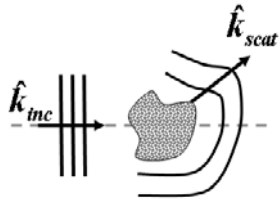


Figure 2. Scattering of an incident ultrasonic field by a volume of tissue. The angle of incidence is $\hat{k}_{inc}$, and the direction of scattering is $\hat{k}_{scat}$.

3. Basic Research Applications

We consider the use of HFUS for characterization of biological materials that may potentially play a role in diagnostic and regenerative medicine. The first study includes *ex vivo* measurements of pulmonary arterial wall from a murine (rat) model in order to evaluate the role of hypoxia (low oxygen content) in pulmonary arterial hypertension (PAH) and its corresponding effects on ultrasonic properties of tissue. The second study demonstrates the utility of HFUS to monitor the development of tissue-engineered cartilage that is based on a biodegradable hydrogel scaffold. Both studies employ a 50 MHz HFUS research system.

3.1. *Pulmonary Arterial Hypertension*

Pulmonary arterial hypertension[12] includes multiple diseases that lead to hypertension in the pulmonary circulation. Hypertension (or elevated blood pressure) of the pulmonary circulation is defined as a mean pulmonary arterial pressure >25 mmHg[13] at rest, which is considerably lower than pressures of the systemic and coronary circulations. Chronic hypertension if not diagnosed early or if left untreated leads to progressive increases in pulmonary vascular resistance and eventually to right-side heart failure.[14]

The diagnosis of PAH can sometimes be difficult because its symptoms are similar to those of other diseases. Cardiac catheterization utilizes intravascular ultrasound to determine pulmonary vascular resistance and is considered a reliable diagnostic test.[13,14] However, catheterization is an invasive procedure that can preclude its use in some patients, and consequently a noninvasive alternative is of interest, particularly for screening tests.[15]

In contrast to arterial tissue of the systemic[16,17] and coronary circulations,[18-21] little has been reported on the ultrasonic propagation and scattering properties of the arterial tissue of the pulmonary circulation. The knowledge of such properties has the potential to improve tools for the diagnosis of PAH.

3.1.1. *Tissue Specimens and Preparation*

Long-Evans murine models were raised at the University of Colorado Health Sciences Center in Denver (UCHSCD). The murine models were sacrificed at 12 – 13 weeks of age, and the extrapulmonary system (main trunk and left and right branches) were excised. The pulmonary arteries (PAs) were kept on ice until measurements were performed at the facilities at the National Institute of Standards and Technologies (NIST) in Boulder, Colorado. All measurements

were performed within 24 hours of sacrifice. Two populations of murine models were investigated: normoxic Long-Evans (N_N = 8) and hypoxic Long-Evans (N_H = 7). The hypoxic rat models were placed for three weeks in a hypobaric chamber that simulated an altitude of 5200 m. This altitude has roughly half the oxygen content of air at sea level such that the hypoxia presumably induces PAH.

The extrapulmonary system was further prepared prior to ultrasonic measurements. Each arterial section (left branch, right branch, main trunk) was segmented, and the PA wall was cut in the longitudinal direction. The arterial wall was opened as a membrane, and the adventitia was removed under an optical microscope by means of tissue scissors. Measurements on fresh tissue were performed in the radial direction (through the thickness of the arterial wall). Additional details regarding the tissue specimens have been previously described.[22]

3.1.2. *Data Acquisition*

Measurements were performed using a HFUS system in double-transmission and pulse-echo modes to respectively measure the propagation and backscatter properties of the PA specimens. (See Figure 3.) The 50 MHz transducer had the following characteristics: peak frequency of 40 MHz, -6 dB bandwidth of 40 MHz (25 MHz to 65 MHz), -14 dB pulse duration of 80 ns, beam diameter of 90 μm, and a depth of field of 1.2 mm. This transducer permitted measurements over length scales of approximately 60 μm in the axial direction. The transducer was electrically excited by a 150 MHz square wave pulser/receiver. The ultrasonic pressure field traveled through a nutritive solution (also known as media) that was maintained at 37°C before interacting with the tissue. Detected ultrasonic signals were received and sampled (1 GSamples/s with 8-bit resolution). Pulmonary arterial tissue specimens were mounted on a custom tissue fixture that was made of stainless steel. Reference measurements were performed with the tissue removed from the propagation path.

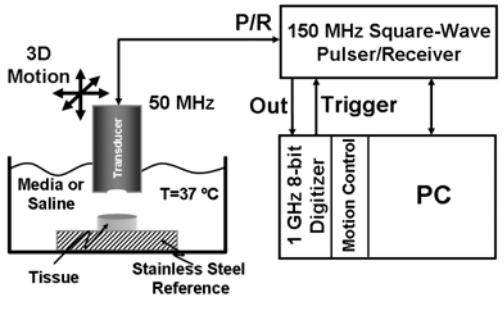


Figure 3. Schematic diagram of high-frequency ultrasound measurement system.

3.1.3. *Data Reduction*

Standard data reduction methods were implemented to calculate the ultrasonic propagation and backscatter properties, including speed of sound,[23] attenuation coefficient,[24] and backscatter coefficient.[10] The speed of sound was determined by monitoring differences in times-of-flight of the ultrasonic signal between reference and specimen measurements

$$v = v_r \left[1 + v_r \Delta t / (2h)\right]^{-1}, \tag{9}$$

where v_r is the speed of sound in the reference medium (nutrative solution), Δt is the difference in times-of-flight, and h is the thickness of the tissue specimen (also determined ultrasonically). The speed of sound of the nutrative solution was approximated by a polynomial expression for the temperature-dependent speed of sound in water.[25]

The attenuation coefficient is determined using a log-spectral subtraction technique

$$\alpha(\omega) = \frac{1}{2h}\left[20\log_{10}\left(\frac{|F_r(\omega)|}{|F_t(\omega)|}\right) - 20\log_{10}\left(\frac{T_{rt}R_{ts}T_{tr}}{R_{rs}}\right)\right], \tag{10}$$

where F_x is the Fourier transform of a received (pressure) signal, T_{xy} is an amplitude transmission coefficient, and R_{xy} is an amplitude reflection coefficient, and the subscripts r, t, and s respectively represent the reference medium, the tissue, and the stainless steel tissue fixture.

The backscatter coefficient

$$\eta(\omega) = \left\{\frac{|F_t(\omega)|^2}{|F_r(\omega)|^2}\right\}\left\{\frac{r_0^2}{\pi a^2 l E_\infty}|D_r(\omega)|^2\right\}\left\{\frac{1}{A(\alpha(\omega), l, x_0)}\right\}\left\{R_{rs}^2\frac{T_r^2}{T_t^2}\right\} \tag{11}$$

is determined by calculation of the apparent backscatter transfer function[26] with compensation for diffraction effects, attenuation in the intervening path of tissue, and interface losses. The Rochester method[10] has been used for to account for diffraction effects (second term) and is referred to for additional details regarding Eq. (11). The attenuation term (third term) is equivalent to that originally described by Sigelmann and Reid.[27]

Statistical analysis was performed using single-factor analysis of variance with a confidence interval of 0.05. Results are reported as the mean plus or minus one standard deviation.

3.1.4. *Results*

Figure 4 summarizes the ultrasonic properties of the normoxic and hypoxic murine populations. Hypoxia was found to increase the speed of sound (Fig. 4a), the slope of attenuation (Fig. 4b), and the mean backscatter coefficient (Fig. 4c) in the media of the extrapulmonary arterial walls. The differences were found to be statistically significant in the right main branch for all three properties, the left main branch for the slope of attenuation and mean backscatter, and the trunk for only the slope of attenuation. In addition, the frequency dependence of the backscatter coefficient was found to be statistically larger for the hypoxic trunk than the normoxic trunk.

The increased speed of sound is consistent with a stiffening of the arterial walls related to hypertension.[28] The increased scattering found in the hypoxic specimens is consistent with the increase in attenuation. The values for backscatter of the media from normoxic specimens are similar to that of other elastic arteries, such as the carotid artery.[16]

This evidence supports the hypothesis that hypoxia leads to remodeling of the pulmonary arterial wall. It is still an open question as to what remodeling causes these changes in the ultrasonic properties (e.g., increased elastin content). Figure 5 shows a histologic image of a section of a normoxic pulmonary trunk that has been stained with Movat pentachrome. Elastin is stained black, collagen is yellow/orange, smooth muscle cells are pinkish, and ground substance is blue.

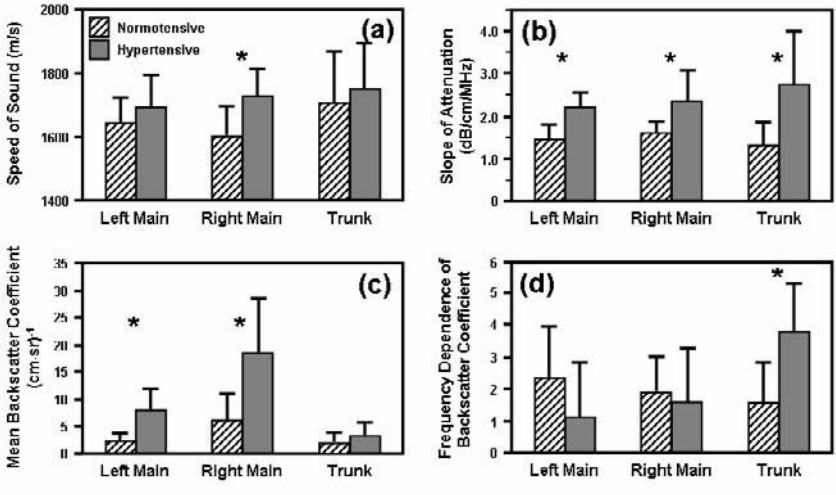
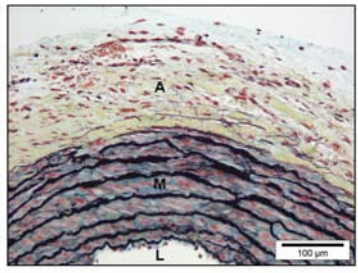


Figure 4. Propagation and backscatter properties of the extrapulmonary system for control (normoxic) and experimental (hypoxic) groups of Long-Evans murine (rat) models, including (a) speed of sound, (b) slope of attenuation, (c) mean backscatter coefficient, and (d) frequency dependence of backscatter coefficient. * indicates statistical difference between groups at the 0.05 level.



Figue 5. Optical micrograph of a histologic slice of the pulmonary arterial wall from a hypoxic murine model stained with Movat pentachrome. The lumen (L), media (M), and adventitia (A) are noted. A 100 µm scale bar is included.

Hypoxia is known to increase the number of medial layers (medial thickening) (not shown). Medial thickness was found to be larger for the hypoxic specimens, particularly the right main branch which was on average 64% thicker (approx. 280 µm). However, a definitive study of the correlation of the ultrasonic properties and remodeling due to hypoxia requires a comprehensive histologic study which was not available for this preliminary study.

3.1.5. *Perspectives*

The present results represent one of the few reports on the ultrasonic propagation and scattering properties of the media of the pulmonary artery. The paucity of such data may reflect the larger historical interest in diseases associated with the systemic and coronary circulations rather than the pulmonary circulation. However, there may be increased interest based on recent scientific meetings such as the National Institutes of Health meeting on the Evolution of Pulmonary Hypertension and the establishment of Specialized Centers of Clinically Oriented Research for Pulmonary Vascular Disease (D. Dunbar Ivy, UCHSD, personal communication, October 2006).

It is the progressive increase in pulmonary vascular resistance that is potentially fatal in patients that suffer from PAH. Consequently, it is the hemodynamics in the pulmonary circulation that is of particular interest for diagnosis and therapy of PAH. Knowledge of the ultrasonic propagation and scattering properties of the pulmonary arterial wall can benefit the design of ultrasound-based screening and monitoring approaches. However, the development of a robust non-invasive estimation of pulmonary vascular resistance based on flow may be most attractive to clinicians.[29]

3.2. *Osteoarthritis and Functional Tissue Engineering*

Osteoarthritis (OA) is a degenerative joint disease that affects over 20 million Americans with annual costs to society of more than US$60 billion.[30] OA leads to a degraded quality of life for the patient due to the loss of articular cartilage that serves to cushion and lubricate joints.[31] At present there is no known cure for OA. Furthermore, the avascular nature of cartilage limits its ability for self repair. Consequently, the availability of a regenerative or reparative therapy would improve the quality of life for patients that suffer from OA and reduce the economic impact of this disease.

Functional tissue engineering of articular cartilage aims to provide an engineered cartilage that exhibits structure and function of sufficient quality to replace or repair native articular cartilage. Principles of functional tissue engineering have been identified that can significantly impact the quality of engineered cartilage.[32] Briefly, the degradation profiles of hydrogel carriers have important effects on many aspects of neotissue formed in these systems. In addition, knowledge of the complete mechanical properties of native cartilage can be exploited to establish required target properties.

An ultrasound measurement technique is attractive because it can be quantitative, noninvasive or only minimally invasive, and produce minimal detrimental biological effects. The use of ultrasound for monitoring mechanical properties of native or engineered cartilage has until recently received only little attention.[3-5,33] Ultrasonic attenuation of human and bovine cartilage was measured as early as the 1950s.[34] However, more comprehensive measurements of the anisotropic and heterogeneous properties of cartilage were not performed until the early 1990s following a developed interest in the potential role of ultrasound for monitoring effects of OA on cartilage. High-frequency ultrasound measurements of the speed of sound and the ultrasonic attenuation of bovine cartilage were correlated with extracellular matrix (ECM) components.[35,36] Studies of the effects of OA on the ultrasonic propagation and scattering properties of cartilage have continued to be performed at several research hospitals and university laboratories.[37-42] In addition, the ultrasonic properties of engineered cartilage following transplantation have recently been investigated.[43] The use of standard ultrasonic measurement methods would further elucidate the structure-function relationship in cartilage and facilitate the use of ultrasound as a tool for online monitoring of the mechanical properties of engineered cartilage.

3.2.1. *Tissue Specimens and Preparation*

Three groups of three tissue specimens each were investigated. The first group consisted of cartilage specimens that were harvested from the femoropatellar joint of immature (2 weeks old) bovine models and had a total collagen content of (58.3 ± 13.2) mg/g wet weight and a glycosoaminoglycan (GAG) content of (19.2 ± 3.8) mg/g wet weight. The second group consisted of cartilage specimens that were from the same bovine animal models and were additionally treated with a collagenase solution (0.05% weight) for 60 min. The collagenase treatment removed a portion of the crosslinked collagen network of the cartilage, and these specimens served as models for engineered cartilage with a developing ECM. The collagenase-treated cartilage specimens had total collagen content of (30.0 ± 8.3) mg/g wet weight and GAG content of (5.7 ± 2.2) mg/g wet weight values. The third group consisted of a poly(ethylene glycol) (PEG) based hydrogel scaffold having triggerable degrading pathways of dimethacrylated poly(caprolactone)-b-PEG-b-poly(caprolactone) (PEG-CAP-DM) that was seeded with chondrocytes (Day 0) and was assumed to have zero total collagen content and zero GAG content. Total collagen content was determined by a hydroxyproline assay, and total GAG content was measured using a dimethylmethylene blue assay.

3.2.2. *Data Acquisition and Reduction*

The experimental setup and acquisition protocol were similar to that of the PAH study, with the exception of the use of a phosphate-buffered saline solution as the reference medium. The data reduction techniques were also similar to that of the PAH study. Speed of sound and slope of attenuation were calculated. Integrated backscatter,[44]

$$IBS = \frac{1}{\Delta\omega}\int_{\omega_l}^{\omega_h} 20\log_{10}\left(\frac{\left|F_t(\omega)\right|}{\left|F_r(\omega)\right|}\right)d\omega, \tag{12}$$

was calculated and can be thought of as a poor man's backscatter coefficient where compensation for diffraction, attenuation, and insertion losses are ignored. Here ω_l and ω_h are respectively the low and high ends of the useful bandwidth $\Delta\omega$. The use of an integrated backscatter parameter rather than backscatter coefficient was motivated by the fact that such a parameter could be more easily implemented in an online monitoring system.

3.2.3. *Results*

The ultrasonic properties of the bovine cartilage specimens were considered as target properties for engineered cartilage. It was assumed that as the engineered cartilage developed that the collagen and GAG content would increase. The speed of sound and slope of attenuation increased in a quasi-linear manner with respect to total collagen content. (See Fig. 6(a,b).) The integrated backscatter parameter (relative to a near-perfect reflector) was also sensitive to total collagen content (not shown) and the ratio of total collagen content to GAG content (Fig. 6c). The large uncertainty in the ratio (for the integrated backscatter value of approximately -30 dB) is due to the large relative uncertainty in the measurement of the GAG content for the collagenase-treated cartilage. A multi-parametric approach to quantify the quality of the engineered cartilage may be found to be useful for online monitoring.

We hypothesize that the efficiency of a region of natural or engineered cartilage to scatter sound in the backwards direction is higher for a less homogeneous specimen. Briefly, a tissue-engineered construct that is considered to initially be relatively homogeneous (i.e., mostly fluid) becomes less homogenous as the ECM develops. Beyond some threshold, the construct would once again become more homogeneous (i.e., mostly ECM). The ratio of total collagen content to GAG content could be considered as a measure of this homogeneity. Consequently, the integrated backscatter would be expected to increase for a period during development and then be followed by a decrease. However, it should be noted that the conventional definition of integrated backscatter does not account for effects due to diffraction or attenuation of the intervening path. These effects were minimized by selecting regions-of-interest that were at shallow depths within the tissue specimens, but explicit compensation for these effects may be necessary for a detailed characterization.

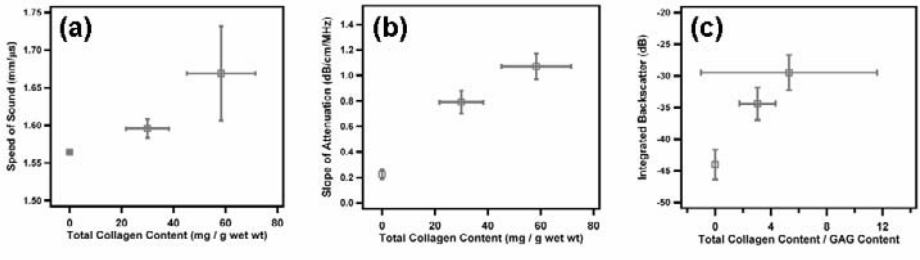


Figure 6. (a) Speed of sound and (b) slope of attenuation as a function of total collagen content and (c) integrated backscatter as a function of the ratio of total collagen content to glycosoaminoglycan (GAG) content in a PEG-CAP-DM hydrogel, collagenase-treated bovine cartilage, and untreated bovine cartilage.

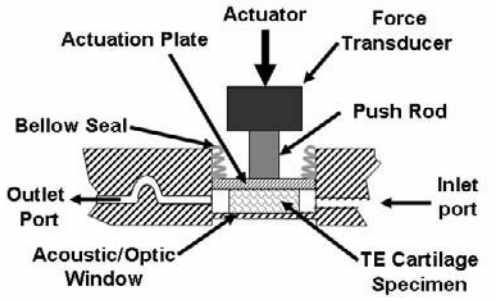


Figure 7. Schematic of a tissue-engineered (TE) specimen well for a compressive-stimulation bioreactor that permits online monitoring. Courtesy of Timothy P. Quinn of the Materials Reliability Division at the National Institute of Standards and Technology in Boulder, Colorado.

3.2.4. *Perspectives*

These preliminary measurements are indicative of the potential role that ultrasound may play in monitoring the development of engineered cartilage. The online monitoring of engineered cartilage requires expertise of biological and chemical engineering for the tissue engineering, mechanical engineering for the biomechanical stimulation of the tissue, and ultrasound physics for the monitoring. Figure 7 provides an example of an incubation cell that is part of a bioreactor and is currently under development at NIST. This effort combines the expertise of these three fields to enable a noninvasive real-time approach to online monitoring of engineered cartilage.

4. Clinical Application: Coronary Artery Disease

Coronary artery disease (CAD) is a type of cardiovascular disease that is known to affect over 13 million American adults, and CAD is the single largest killer of males and females. Each year 700,000 Americans are estimated to have a new coronary attack and approximately 500,000 have a recurrent attack.[45] Coronary artery disease is defined as cardiac events or symptoms related to myocardial ischemia and/or injury due primarily to atherosclerosis. Such events include unstable angina (chest pain), myocardial infarction (heart attack), and sudden death due to ischemic heart disease. While atherosclerosis and ischemia need not necessarily lead to a cardiac event, it is rare that a cardiac event is not preceded by atherosclerosis or ischemia. The challenge is to not only detect conditions of atherosclerosis or ischemia, but also predict what conditions will lead to a cardiac event.[46] It is striking that myocardial infarction or sudden cardiac death is the *initial* presentation of CAD in more than 50% of patients. This reflects a

so-called "detection gap," namely a difference between CAD cases currently detected and the total burden of the disease among the population.[47]

Atherosclerosis can be described as a disease of the intima, and the progression of atherosclerotic plaque has been previously described in detail.[48] There are a range of plaque types that can be present in a lesion. Four plaque types that are of present interest include fibrotic tissue, fibro-fatty tissue, necrotic core (with microcalcifications), and dense calcium.[49] The absolute and relative proportions of these plaque types in a given lesion are indicative of the so-called vulnerability of the lesion to rupture.

Some early plaques are characterized by a thickening of the intima. As plaques progress they develop a well-defined lipid or necrotic core that is covered by a fibrous cap. Such plaques are referred to as fibroatheromas, and it is the case of a thin-cap fibroatheroma that is prone to rupture. A rupture event can lead to thrombosis which may in turn lead to occlusion of the artery and myocardial infarction. A thin-cap fibroatheroma is an example of a so-called vulnerable plaque,[50] and the ability to detect vulnerable plaques can arguably facilitate prevention of acute coronary events.

There are multiple modalities that are used for the detection of vulnerable plaques.[51] Of present interest is the use of spectral-based ultrasound approaches for the estimation of plaque composition in order to assess the vulnerability of a plaque to rupture.

4.1. *Diagnosis and Treatment of Coronary Arterial Disease*

The standard clinical means for diagnosing CAD is minimally invasive coronary (x-ray) angiography.[52] This involves cardiac catheterization and the use of radio-opaque contrast dye. It provides the interventional cardiologist with a "shadow" of the vascular lumen (lumenogram) where the dye was able to flow and, consequently, permits estimation of the level of stenosis. While the degree of stenosis is an important factor in treatment planning, it is now realized that plaque composition is also an important factor.[53] Intravascular ultrasound (IVUS) is a complementary tool to angiography and provides tomographic views of the arterial wall and plaque lesions. Consequently, an improved description of plaque morphology as well as determination of plaque composition is possible.

Current clinical IVUS systems can be differentiated by the type of catheter that is used. The catheter is a flexible tube which encases a side-looking ultrasound probe and is guided from the femoral artery, where it is introduced into the body, to the coronary arteries for imaging of the plaque. The ultrasound probe can be one of two types: a single-element mechanically rotating transducer

or a solid-state array. Both types of IVUS catheters have their advantages and disadvantages. Briefly, mechanically rotating devices can achieve higher nominal center frequencies (up to 45 MHz) compared to solid-state arrays (up to 20 MHz) and, consequently, provide finer spatial resolution at the cost of reduced depth of penetration. Solid-state arrays permit focusing via synthetic aperture techniques.[54] Image artifacts are possible for both types of catheters. Grating lobes may be present for solid-state arrays when echogenic regions (e.g., dense calcium) are present in the field of view. Non-uniform rotational distortion can occur for mechanically rotating catheters if the transducer does not rotate at a uniform rate. Additional details regarding IVUS hardware and catheters have been described elsewhere.[49]

For the remaining discussion regarding characterization of coronary arterial plaque, discussion will be limited to the use of a solid-state array. The catheter is approximately 1 mm in diameter. The array consists of 64 elements and has a nominal center frequency of 20 MHz. The spatial resolution is approximately 140 μm in the axial direction and several hundred microns in the azimuthal and elevational (out-of-plane) directions.

4.2. *VH™ IVUS: Real-Time Lesion Assessment*

Several approaches for the ultrasonic evaluation of atherosclerotic plaque composition have been explored. Qualitative methods have been proposed to visually evaluate plaque morphology, brightness, and heterogeneity.[55] Quantitative approaches involving computer-assisted estimation of image texture[56] or mean grayscale have also been proposed.[57] Although such systematic approaches to image evaluation provide additional information related to plaque content, the assessment of the image can be strongly dependent on system settings such as image compression techniques, time-gain-compensation levels, and beam focusing parameters.

Quantitative approaches that are based upon the analysis of RF-based signals aim to reduce system-dependent effects on the signal backscattered from the atherosclerotic plaque. The application of appropriate signal analysis techniques can provide estimates of the ultrasound propagation and scattering properties[58,59] and, consequently, insight into tissue remodeling due to atherosclerosis.[60,61]

VH IVUS is a spectral-based approach to characterization of coronary atherosclerotic plaque that was developed at the Cleveland Clinic Foundation and licensed to Volcano Corp. It provides automated generation of lumen and vessel border contours as well as colorized tissue maps of plaque composition. Plaque is classified into one of the four aforementioned plaque types, namely fibrotic

tissue, fibro-fatty tissue, necrotic core (with microcalcifications), and dense calcium.

The VH IVUS algorithm has been described in some detail in previous publications.[20,49,62-65] Briefly, spectral analysis is performed for localized regions of plaque using a sliding window technique. System-dependent effects can be estimated by use of blind deconvolution techniques.[66-68] Localized regions of plaque can then be classified by use of a classification tree, a type of statistical learning method.[69] The VH IVUS classification tree is trained using an *ex vivo* measurement database of fresh human coronary arteries obtained at autopsy. The measurement database contains regions of interest of known plaque type (as assessed by histology) and corresponding spectral parameters. Spectral parameters are based on the estimated backscatter transfer function[70] and include in part the integrated backscatter[44] and so-called linear regression parameters.[71]

4.3. *Results*

Figure 8a provides an example of an IVUS grayscale image of a section of a coronary artery with plaque that was measured *ex vivo*. The catheter is centered in the image, and the central bright ring corresponds to catheter ringdown. The lumen (red) and external elastic lamina (blue) contours are identified. The saline-perfused lumen is echolucent. In addition, the media of the (muscular) coronary artery is typically echolucent in contrast to the adventitia which is echogenic, and this transition is indicative of the location of the external elastic lamina. The backscatter signal between the contours is from the plaque and media. The signal from the plaque is analyzed according to the VH IVUS algorithm, and a color image is constructed that depicts the plaque composition as shown Fig. 8b. The VH IVUS image indicates that this plaque section contains a calcified necrotic core that is covered by a fibrous cap (at 6 o'clock). The corresponding histology is shown in Fig. 8c and indicates the presence of a calcified necrotic core with a fibrous cap with microcalcifications. A limitation of such comparisons is that the thickness of a histology slice is 4 μm whereas the ultrasonic image upon which VH is based is derived from the backscatter from a volume with a thickness of several hundred microns. Comparison of IVUS and histology images is necessarily imprecise because of this discrepancy in slice thickness.

It is the thin-cap fibroatheroma which is of particular concern in the context of the vulnerable plaque. Fibrous caps that are less than 65 μm thick are prone to rupture.[72] However, the current axial resolution in VH IVUS images is approximately 250 μm which precludes the explicit detection of thin-cap fibroatheromas. The identification of so-called "VH thin" cap fibroatheromas has

been suggested as a compromise in which the fibroatheroma may not appear to have any fibrous cap in the VH IVUS image.

Figure 9 is a screen shot of VH IVUS as implemented on a clinical imaging system. The VH IVUS image overlays the IVUS grayscale image in the tomographic view at left. Catheter ringdown is masked out. Estimated plaque composition is shown in a pie chart form and indicates that this section of the lesion is fibrocalcific. The right side of the figure shows a slice of the longitudinal view that is available when a catheter is pulled back through the vessel (from distal to proximal) and indicates the length of a lesion. Use of the longitudinal view facilitates the planning and deployment of stent devices (mesh tubes) that reduce the level of stenosis.[73,74]

4.4. *Perspectives*

Vulnerable plaque detection is an active area of research. All imaging modalities attempt to exploit some *contrast* mechanism that can indicate the presence of a vulnerable plaque. VH IVUS exploits the contrast in the backscatter properties of tissue to determine plaque composition. There are additional ultrasound-based techniques that can provide complementary information.

Palpography is an elastographic-based technology developed at Erasmus Medical Center (Rotterdam) and licensed to Volcano Corp. It exploits the elastic properties of tissue in order to identify regions of high strain such as the shoulders of thin-cap fibroatheromas which are indicative of vulnerability to rupture. Figure 10 provides an example of a palpogram that overlays a grayscale IVUS image of a coronary atherosclerotic plaque. Schaar et al.[75] have shown that such a palpogram represents a thin-cap fibroatheroma.

In addition to exploitation of the scattering and elastic properties by respectively VH IVUS and Palpography, other experimental techniques aim to take advantage of still other contrast mechanisms. Photoacoustics[76] is a promising approach that combines the tremendous optical contrast properties of tissue with the attractive depth of penetration and spatial resolution of ultrasound. Harmonic IVUS offers the potential of image improvements similar to that of harmonic imaging for echocardiographic applications. The appeal of harmonic IVUS increases when this modality is combined with the use of contrast agents such as encapsulated gas microbubbles or site-targeted agents. Harmonic IVUS can exploit the nonlinear scattering properties of microbubbles to detect the vasa vasorum which may facilitate clinical assessment of atherosclerosis.[77] Site-targeted contrast agents could potentially bind to specific plaque components that could aid detection of the vulnerable plaque.[78]

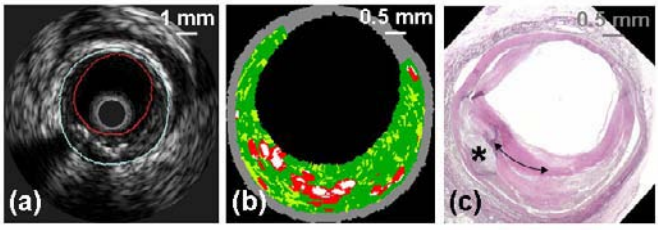


Figure 8. (a) Grayscale, (b) VH IVUS, and (c) histology images of a human coronary arterial plaque. Grayscale and VH IVUS images were acquired *ex vivo* with a Volcano s5 system with an Eagle Eye catheter. The histologic specimen was stained with hematoxylin and eosin. * indicates presence of a necrotic core, and the arrow indicates a fibrous cap with microcalcifications. (VH Colors: Green = Fibrotic Tissue, Yellow-Green = Fibro-Fatty Tissue, Red = Necrotic Core with microcalcifications, White = Dense Calcium, Gray = Media)

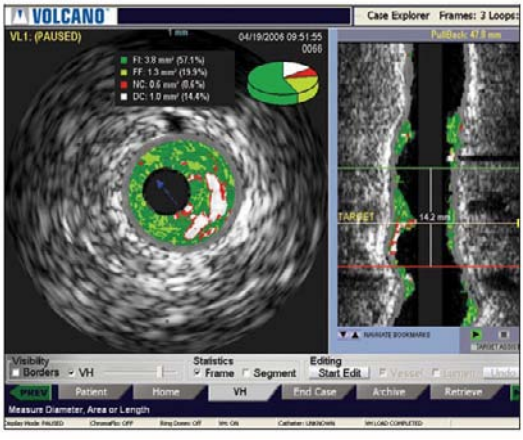


Figure 9. Screen shot of a Volcano s5 system using the Eagle Eye catheter. A VH IVUS image overlays a grayscale image of a coronary arterial plaque at left. A longitudinal view of the vessel is shown at right.

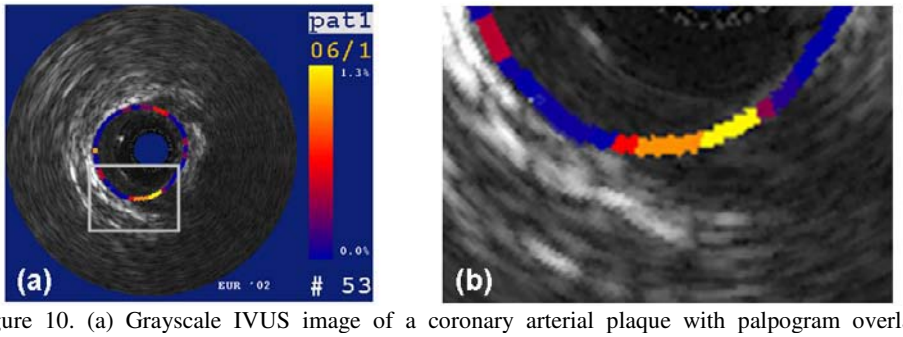


Figure 10. (a) Grayscale IVUS image of a coronary arterial plaque with palpogram overlay. (b) Zoom of palpogram shown in boxed region in (a) indicating that shoulder regions of eccentric plaque are areas of high strain. Blue corresponds to regions of low strain, red and yellow corresponds to regions of higher strain. Courtesy of Antonius F. W. van der Steen of the Erasmus Medical Center, Thorax Centre, Rotterdam, The Netherlands.

The diagnostic utility of any of these ultrasound-based approaches to the management of coronary artery disease must ultimately be verified through large-scale clinical trials. There are several clinical trials which are currently underway in which VH IVUS and Palpography are primary endpoints.[79,80] Furthermore, a global multi-center registry has been established that will correlate VH IVUS data with demographics and clinical presentation.[81]

5. Discussion

The role of HFUS in diagnostic and regenerative medicine has been investigated for two basic and one clinical research topics. The propagation and backscatter properties of the extrapulmonary arterial walls have been measured in order to observe the effects of PAH. Similarly, the propagation and backscatter properties of natural cartilage have been measured to establish target properties for tissue-engineered cartilage. In addition, the use of these same properties has been proposed to permit noninvasive, online monitoring of the engineered cartilage during its development. Lastly, the backscatter properties of atherosclerotic plaque have been exploited in an effort to improve the management of coronary artery disease.

The use of HFUS has been investigated for over 50 years,[82] and the range of potential medical applications continues to grow. However, there appears to be a significantly larger body of literature[3-5] available for the properties of biological materials at low frequency ultrasound (<10 MHz) due in part that such frequencies have historically had wider applicability to medical problems. This indicates opportunity to not only develop new applications for HFUS (>10 MHz) but also validate the conventional wisdom concerning the interaction of (low frequency) ultrasound and tissue that has developed over decades.

There are several rules of thumb that exist concerning the propagation and scattering of biological materials. Dispersion in soft tissue is negligible such that a frequency-independent speed of sound (1540 m/s) can be assumed. Attenuation due to scattering is negligible compared to that due to absorption, and that absorption increases linearly with frequency with a given slope of approximately 0.5 dB/cm/MHz. Each of these points is worth confirming or updating for the case of high frequency ultrasound. For the case of attenuation, it has been shown that the slope of attenuation in soft tissue is larger than 0.5 dB/cm/MHz which is consistent with the description that the increasing attenuation with respect to frequency is better described as a power law than as linear with frequency.[83] Furthermore, the knowledge of such high-frequency properties can be utilized for the construction of tissue-mimicking phantoms[84] which have applicability to

imaging system evaluation as well as algorithm development for tissue characterization purposes.

There is also a need to accurately characterize high-frequency devices.[85] State-of-the-art commercial hydrophones have active elements with effective diameters on the order of 100 μm. This may be too large for some high-frequency transducers such that phase cancellation at the face of the hydrophone may occur.[86] Furthermore, there is the additional issue concerning the availability and accuracy of calibrations of hydrophones for frequencies above 60 MHz (S. Howard, Onda Corp, personal communication, October 2006). The ability to accurately measure exposure levels may be of particular interest for the case of using HFUS to monitor neotissue development.

These examples motivate the complementary roles that industrial, academic, and governmental organizations can have. The US Food and Drug Administration regulates the amount of acoustic energy that can be deposited by a medical device. Estimations of this level for a particular device will depend upon the availability of reliable calibration measurements (i.e., hydrophones) and accurate measurement of the propagation properties of the tissue of interest. Government organizations such as NIST can facilitate the development of technologies for reliable calibration devices and protocols. Funding organizations such as the National Institutes of Health can establish priorities and provide support for the appropriate research and development in which both academia and industry can participate.

6. Conclusions

The application of high-frequency ultrasound to diagnostic and regenerative medicine has been described for three cases that concern pulmonary arterial hypertension, osteoarthritis, and coronary artery disease. Quantitative ultrasonic characterization techniques based on spectral analysis have been applied and indicate the role that such approaches may have in monitoring the progression of disease and the development of tissue-engineered materials. Additional opportunities have been described to determine the relative importance of the propagation and scattering properties in biological materials for purposes of high-frequency ultrasound tissue characterization as well as to develop standards and technologies for calibration measurements.

Acknowledgments

Parts of this research were performed while the author held a National Research Council Research Associateship Award at NIST. The author thanks Elizabeth S. Drexler, Andrew J. Slifka, and Chris N. McCowan of NIST for discussions and support related to the PAH study and D. Dunbar Ivy and Kelley Colvin at UCHSCD for supply and preparation of the Long-Evans rats. The author thanks Mark A. Rice and Kristi S. Anseth of the University of Colorado at Boulder for supply, preparation, and biochemical analysis of the bovine cartilage and hydrogel specimens. The author also thanks D. Geoffrey Vince and Anuja Nair of Volcano Corp for many discussions related to VH IVUS.

References

1. G. C. Gurtner, M. J. Callaghan, and M. T. Longaker, Progress and Potential for Regenerative Medicine, *Annu Rev Med* (2006).
2. M. F. Hamilton and D. T. Blackstock, *Nonlinear acoustics* (Academic Press, San Diego, CA, 1998).
3. F. A. Duck, *Physical properties of tissue: a comprehensive reference book* (Academic Press, London ; San Diego, 1990).
4. S. A. Goss, R. L. Johnston, and F. Dunn, Comprehensive compilation of empirical ultrasonic properties of mammalian tissues, *J Acoust Soc Am* **64**, 423 (1978).
5. S. A. Goss, R. L. Johnston, and F. Dunn, Compilation of empirical ultrasonic properties of mammalian tissues. II, *J Acoust Soc Am* **68**, 93 (1980).
6. J. F. Greenleaf, *Tissue characterization with ultrasound* (CRC Press, Boca Raton, Fla., 1986).
7. J. C. Bamber, Attenuation and Absorption in *Physical Principles of Medical Ultrasonics*, edited by C. R. Hill (Halsted Press, New York, 1986), p. 119.
8. R. L. Johnston and F. Dunn, Ultrasonic hysteresis in biological media, *Radiat Environ Biophys* **19**, 137 (1981).
9. K. K. Shung, Introduction in *Ultrasonic scattering in biological tissues*, edited by K. K. Shung and G. A. Thieme (CRC Press, Boca Raton, 1993), p. 1.
10. X. Chen, D. Phillips, K. Q. Schwarz, et al., The measurement of backscatter coefficient from a broadband pulse-echo system: a new formulation, *IEEE Trans Ultrason Ferroelec Freq Contr* **44**, 515 (1997).
11. J. J. Faran, Sound Scattering by Solid Cylinders and Spheres, *J Acoust Soc Am* **23**, 405 (1951).
12. G. Simonneau, N. Galie, L. J. Rubin, et al., Clinical classification of pulmonary hypertension, *J Am Coll Cardiol* **43**, 5S (2004).
13. D. D. Ivy, Diagnosis and treatment of severe pediatric pulmonary hypertension, *Cardiol Rev* **9**, 227 (2001).
14. T. D. Nauser and S. W. Stites, Diagnosis and treatment of pulmonary hypertension, *Am Fam Physician* **63**, 1789 (2001).

15. R. Shandas, C. Weinberg, D. D. Ivy, et al., Development of a noninvasive ultrasound color M-mode means of estimating pulmonary vascular resistance in pediatric pulmonary hypertension: mathematical analysis, in vitro validation, and preliminary clinical studies, *Circ* **104**, 908 (2001).
16. G. R. Lockwood, L. K. Ryan, J. W. Hunt, et al., Measurement of the ultrasonic properties of vascular tissues and blood from 35-65 MHz, *Ultrasound Med Biol* **17**, 653 (1991).
17. K. R. Waters, S. L. Bridal, C. Cohen-Bacrie, et al., Parametric analysis of carotid plaque using a clinical ultrasound imaging system, *Ultrasound Med Biol* **29**, 1521 (2003).
18. M. Kawasaki, H. Takatsu, T. Noda, et al., In vivo quantitative tissue characterization of human coronary arterial plaques by use of integrated backscatter intravascular ultrasound and comparison with angioscopic findings, *Circ* **105**, 2487 (2002).
19. J. C. Machado, F. S. Foster, and A. I. Gotlieb, Measurement of the ultrasonic properties of human coronary arteries in vitro with a 50-MHz acoustic microscope, *Braz J Med Biol Res* **35**, 895 (2002).
20. A. Nair, B. D. Kuban, N. Obuchowski, et al., Assessing spectral algorithms to predict atherosclerotic plaque composition with normalized and raw intravascular ultrasound data, *Ultrasound Med Biol* **27**, 1319 (2001).
21. F. R. Pereira, J. C. Machado, and F. S. Foster, Ultrasound characterization of coronary artery wall in vitro using temperature-dependent wave speed, *IEEE Trans Ultrason Ferroelec Freq Contr* **50**, 1474 (2003).
22. E. S. Drexler, T. P. Quinn, A. J. Slifka, et al., Comparison of mechanical behavior among the extrapulmonary arteries from rats, *J Biomech* (2006).
23. D. K. Hsu and M. S. Hughes, Simultaneous ultrasonic velocity and sample thickness measurement and application in composites, *J Acoust Soc Am* **92**, 669 (1992).
24. A. C. Kak and K. A. Dines, Signal processing of broadband pulsed ultrasound: measurement of attenuation of soft biological tissues, *IEEE Trans Biomed Eng* **25**, 321 (1978).
25. V. A. Del Grosso and C. W. Mader, Speed of sound in pure water, *J Acoust Soc Am* **52**, 1442 (1972).
26. S. A. Wickline, J. E. Pérez, and J. G. Miller, Chapter 10: Cardiovascular tissue characterization *in vivo* in *Ultrasonic Scattering in Biological Tissues* (CRC Press, Inc., 1993), p. 313.
27. R. A. Sigelmann and J. M. Reid, Analysis and measurement of ultrasound backscattering from an ensemble of scatterers excited by sine-wave bursts, *J Acoust Soc Am* **53**, 1351 (1973).
28. A. V. Chobanian, G. L. Bakris, H. R. Black, et al., Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure, *Hypertension* **42**, 1206 (2003).
29. K. Dyer, C. Lanning, B. Das, et al., Noninvasive Doppler tissue measurement of pulmonary artery compliance in children with pulmonary hypertension, *J Am Soc Echocardiogr* **19**, 403 (2006).
30. J. A. Buckwalter, C. Saltzman, T. Brown, et al., The impact of osteoarthritis: implications for research, *Clin Orthop Relat Res*, S6 (2004).
31. C. C. Wang, J. M. Deng, G. A. Ateshian, et al., An automated approach for direct measurement of two-dimensional strain distributions within articular cartilage under unconfined compression, *J Biomech Eng* **124**, 557 (2002).

32. D. L. Butler, S. A. Goldstein, and F. Guilak, Functional tissue engineering: the role of biomechanics, *J Biomech Eng* **122**, 570 (2000).
33. S. Tepic, T. Macirowski, and R. W. Mann, Mechanical properties of articular cartilage elucidated by osmotic loading and ultrasound, *Proc Natl Acad Sci* **80**, 3331 (1983).
34. K. T. Dussik, D. J. Fritch, M. Kyriazidou, et al., Measurements of articular tissues with ultrasound, *Am J Phys Med* **37**, 160 (1958).
35. D. H. Agemura, W. D. O'Brien, Jr., J. E. Olerud, et al., Ultrasonic propagation properties of articular cartilage at 100 MHz, *J Acoust Soc Am* **87**, 1786 (1990).
36. D. A. Senzig, F. K. Forster, and J. E. Olerud, Ultrasonic attenuation in articular cartilage, *J Acoust Soc Am* **92**, 676 (1992).
37. E. H. Chiang, T. J. Laing, C. R. Meyer, et al., Ultrasonic characterization of in vitro osteoarthritic articular cartilage with validation by confocal microscopy, *Ultrasound Med Biol* **23**, 205 (1997).
38. M. Fortin, M. D. Buschmann, M. J. Bertrand, et al., Dynamic measurement of internal solid displacement in articular cartilage using ultrasound backscatter, *J Biomech* **36**, 443 (2003).
39. H. Kuroki, Y. Nakagawa, K. Mori, et al., Mechanical effects of autogenous osteochondral surgical grafting procedures and instrumentation on grafts of articular cartilage, *Am J Sports Med* **32**, 612 (2004).
40. S. G. Patil, Y. P. Zheng, J. Y. Wu, et al., Measurement of depth-dependence and anisotropy of ultrasound speed of bovine articular cartilage in vitro, *Ultrasound Med Biol* **30**, 953 (2004).
41. S. Saarakkala, R. K. Korhonen, M. S. Laasanen, et al., Mechano-acoustic determination of Young's modulus of articular cartilage, *Biorheology* **41**, 167 (2004).
42. A. Saied and P. Laugier, High-resolution ultrasonography for analysis of age- and disease-related cartilage changes, *Methods Mol Med* **101**, 249 (2004).
43. K. Hattori, Y. Takakura, H. Ohgushi, et al., Which cartilage is regenerated, hyaline cartilage or fibrocartilage? Non-invasive ultrasonic evaluation of tissue-engineered cartilage, *Rheumatology* **43**, 1106 (2004).
44. M. O'Donnell, J. W. Mimbs, and J. G. Miller, The relationship between collagen and ultrasonic attenuation in myocardial tissue, *J Acoust Soc Am* **65**, 512 (1979).
45. T. Thom, N. Haase, W. Rosamond, et al., Heart disease and stroke statistics--2006 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee, *Circ* **113**, e85 (2006).
46. A. J. Taylor, C. N. Merz, and J. E. Udelson, 34th Bethesda Conference: Executive summary--can atherosclerosis imaging techniques improve the detection of patients at risk for ischemic heart disease?, *J Am Coll Cardiol* **41**, 1860 (2003).
47. R. C. Pasternak, J. Abrams, P. Greenland, et al., 34th Bethesda Conference: Task force #1--Identification of coronary heart disease risk: is there a detection gap?, *J Am Coll Cardiol* **41**, 1863 (2003).
48. H. C. Stary, D. H. Blankenhorn, A. B. Chandler, et al., A definition of the intima of human arteries and of its atherosclerosis-prone regions. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association, *Circ* **85**, 391 (1992).
49. A. Nair, M. P. Margolis, S. C. Davies, et al., Intracoronary Ultrasound in *Mastering Endovascular Techniques: A Guide to Excellence*, edited by P. Lanzer (Lippincott Williams & Wilkins, Philadelphia, 2007), p. 77.

50. W. C. Little, Angiographic assessment of the culprit coronary artery lesion before acute myocardial infarction, *Am J Cardiol* **66**, 44G (1990).
51. Detection of Vulnerable Plaques (Vol. 47, Issue 8 Suppl C): *Supplement of the Journal of the American College of Cardiology* (Elsevier, Philadelphia, 2006), p. 1.
52. E. Escolar, G. Weigold, A. Fuisz, et al., New imaging techniques for diagnosing coronary artery disease, *CMAJ* **174**, 487 (2006).
53. P. Schoenhagen, R. D. White, S. E. Nissen, et al., Coronary imaging: angiography shows the stenosis, but IVUS, CT, and MRI show the plaque, *Cleve Clin J Med* **70**, 713 (2003).
54. M. O'Donnell and L. J. Thomas, Efficient synthetic aperture imaging from a circular aperture with possible application to catheter-based imaging, *IEEE Trans Ultrason Ferroelec Freq Contr* **39**, 366 (1992).
55. M. L. Gronholdt, B. M. Wiebe, H. Laursen, et al., Lipid-rich carotid artery plaques appear echolucent on ultrasound B-mode images and may be associated with intraplaque haemorrhage, *Eur J Vasc Endovasc Surg* **14**, 439 (1997).
56. J. E. Wilhjelm, M. L. Gronholdt, B. Wiebe, et al., Quantitative analysis of ultrasound B-mode images of carotid atherosclerotic plaque: correlation with visual classification and histological examination, *IEEE Trans Med Imaging* **17**, 910 (1998).
57. M. P. Urbani, E. Picano, G. Parenti, et al., In vivo radiofrequency-based ultrasonic tissue characterization of the atherosclerotic plaque, *Stroke* **24**, 1507 (1993).
58. B. Barzilai, J. E. Saffitz, J. G. Miller, et al., Quantitative ultrasonic characterization of the nature of atherosclerotic plaques in human aorta, *Circ Res* **60**, 459 (1987).
59. E. Picano, L. Landini, F. Lattanzi, et al., Ultrasonic tissue characterization of atherosclerosis: state of the art 1988, *J Nucl Med Allied Sci* **32**, 174 (1988).
60. J. A. Campbell and R. C. Waag, Normalization of ultrasonic scattering measurements to obtain average differential scattering cross sections for tissues, *J Acoust Soc Am* **74**, 393 (1983).
61. F. L. Lizzi, M. Greenebaum, E. J. Feleppa, et al., Theoretical framework for spectrum analysis in ultrasonic tissue characterization, *J Acoust Soc Am* **73**, 1366 (1983).
62. J. D. Klingensmith, R. Shekhar, and D. G. Vince, Evaluation of three-dimensional segmentation algorithms for the identification of luminal and medial-adventitial borders in intravascular ultrasound images, *IEEE Trans Med Imaging* **19**, 996 (2000).
63. A. Nair, D. Calvetti, and D. G. Vince, Regularized autoregressive analysis of intravascular ultrasound backscatter: improvement in spatial accuracy of tissue maps, *IEEE Trans Ultrason Ferroelec Freq Contr* **51**, 420 (2004).
64. A. Nair, J. D. Klingensmith, and D. G. Vince, Real-time plaque characterization and visualization with spectral analysis of intravascular ultrasound data, *Stud Health Technol Inform* **113**, 300 (2005).
65. A. Nair, B. D. Kuban, E. M. Tuzcu, et al., Coronary plaque classification with intravascular ultrasound radiofrequency data analysis, *Circ* **106**, 2200 (2002).
66. K. F. Kaaresen and E. Bolviken, Blind deconvolution of ultrasonic traces accounting for pulse variance, *IEEE Trans Ultrason Ferroelec Freq Contr* **46**, 564 (1999).
67. R. Jirik and T. Taxt, High-resolution ultrasonic imaging using two-dimensional homomorphic filtering, *IEEE Trans Ultrason Ferroelec Freq Contr* **53**, 1440 (2006).
68. O. V. Michailovich and A. Tannenbaum, Despeckling of medical ultrasound images, *IEEE Trans Ultrason Ferroelec Freq Contr* **53**, 64 (2006).

69. T. Hastie, R. Tibshirani, and J. H. Friedman, *The elements of statistical learning : data mining, inference, and prediction* (Springer, New York, 2001).
70. S. A. Wickline, J. E. Pérez, and J. G. Miller, Cardiovascular Tissue Characterization In Vivo in *Ultrasonic Scattering in Biological Tissue*, edited by K. K. Shung and G. A. Thieme (CRC Press, Inc., Boca Raton, 1993), p. 313.
71. F. L. Lizzi, M. Astor, E. J. Feleppa, et al., Statistical framework for ultrasonic spectral parameter imaging, *Ultrasound Med Biol* **23**, 1371 (1997).
72. A. P. Burke, R. Virmani, Z. Galis, et al., 34th Bethesda Conference: Task force #2--What is the pathologic basis for new atherosclerosis imaging techniques?, *J Am Coll Cardiol* **41**, 1874 (2003).
73. A. L. McLeod, D. B. Northridge, and N. G. Uren, Ultrasound guided stenting, *Heart* **85**, 605 (2001).
74. G. S. Mintz, S. E. Nissen, W. D. Anderson, et al., American College of Cardiology Clinical Expert Consensus Document on Standards for Acquisition, Measurement and Reporting of Intravascular Ultrasound Studies (IVUS). A report of the American College of Cardiology Task Force on Clinical Expert Consensus Documents, *J Am Coll Cardiol* **37**, 1478 (2001).
75. J. A. Schaar, E. Regar, F. Mastik, et al., Incidence of High-Strain Patterns in Human Coronary Arteries, *Circ* **109**, 2716 (2004).
76. S. Emelianov, S. R. Aglyamov, A. B. Karpiouk, et al., Synergy and Applications of Ultrasound, Elasticity, and Photoacoustic Imaging, 2006 IEEE Ultrasonics Symposium (in press).
77. D. E. Goertz, M. E. Frijlink, D. Tempel, et al., Contrast harmonic intravascular ultrasound: a feasibility study for vasa vasorum imaging, *Invest Radiol* **41**, 631 (2006).
78. F. Villanueva, Development of Targeted Ultrasound Contrast Agents, 2006 IEEE Ultrasonics Symposium (in press).
79. P. W. Serruys, Multimodality Imaging and Biomarker Evaluation of Vulnerable Plaque: Integrated Insights From IBIS-1 and IBIS-2, and Future Predictions, Transcatheter Cardiovascular Therapeutics (2006).
80. G. W. Stone, Virtual Histology I: First Presentation of the Baseline Clinical Features and Plaque Characteristics From the PROSPECT Trial, Transcatheter Cardiovascular Therapeutics (2006).
81. G. S. Mintz, Virtual Histology II: Technique Evolution and Insights From the Global VH Registry, Transcatheter Cardiovascular Therapeutics (2006).
82. F. Dunn and J. E. Breyer, Generation and detection of ultra-high-frequency sound in liquids, *J Acoust Soc Am* **34**, 775 (1962).
83. T. L. Szabo, Causal theories and data for acoustic attenuation obeying a frequency power law, *J Acoust Soc Am* **97**, 14 (1995).
84. E. L. Madsen, T. A. Stiles, K. R. Waters, et al., Tissue-Mimicking Materials for Use in Intravascular Ultrasound Phantoms, 2006 IEEE Ultrasonics Symposium (in press).
85. V. Wilkens and W. Molkenstruck, Construction and Characterization of a Broadband Reference PVDF Membrane Hydrophone for Comparisons of Hydrophone Calibration Methods up to 140 MHz, 2006 IEEE Ultrasonics Symposium (in press).
86. P. W. Marcus and E. L. Carstensen, Problems with absorption measurements of inhomogeneous solids, *J Acoust Soc Am* **58**, 1334 (1975).

CHAPTER 9

A ROC FRAMEWORK FOR COMPUTER-AIDED DETECTION SYSTEMS

Ulrich Scheipers

Resonant Medical Inc.
2050 Bleury, Montreal, Quebec, H3A 2J5, Canada
ulrich@scheipers.org

The application of receiver operating characteristic (ROC) methods for computer-aided detection systems is reviewed. A statistical framework for ROC analysis is presented and different methods of evaluating the performance of computer-aided detection systems are derived.

Most detection systems that are used today are dependent on a separation threshold, which, in many cases, can be chosen by the operator without restrictions. The separation threshold separates the range of output values of the detection system into different target groups, thus, conducting the actual detection process.

In the first part of this chapter, threshold specific performance measures, e.g. sensitivity and specificity, are presented. In the second part, a threshold-independent performance measure, the area under the ROC curve, is reviewed. Only the use of separation threshold-independent performance measures provides detection results which are overall representative for computer-aided detection systems.

This chapter provides the interested scientist with all information needed to conduct ROC analysis and to integrate algorithms performing ROC analysis into detection systems while understanding the basic principles of detection.

1. Introduction

Many works on computer-aided detection have been published during the last years. However, no single measure is used to express the detection performances of the systems proposed but many different measures are used instead. This multitude of different measures makes comparing different systems with each other rather complicated and sometimes impossible.

Furthermore, the lack of a representative performance measure is striking in many works. As shown in the following text, not every performance measure that can be found in current publications is fully representative for the underlying detection system. Some performance measures are representative for one operating point of the detection system only. However, performance measures that are representative for all of the systems operating points are needed for an efficient performance assessment. The intention of this chapter is not to evaluate different methods of detection and classifier design, but to discuss the different modalities of measurement that can be applied to judge the performance of classifiers, detection processes and computer-aided detectors.

Although several software packages dealing with the numerical evaluation of detector performance and with measuring detection results are available today [1-4], basic knowledge about performance measures is important when new detection systems or computer-aided detectors have to be designed and developed. The development of computer-aided detection systems always involves numerous iterative steps to achieve the final and hopefully satisfying detection results, which finally allow application of the system in an industrial, clinical or experimental environment. Additional software might be a tremendous help in visualizing intermediate and final results, but at the same time may never replace inline code that is used during the iterative search for optimal detection results. This chapter is intended to provide the reader with the knowledge to include reasonable performance measures directly into the training routines of a computer-aided detection system.

This review was motivated by the work on ultrasound tissue characterization by the same author. During the last years, a system for ultrasonic tissue characterization based on a multiparameter approach and fuzzy inference systems has been developed at Ruhr-University Bochum, Germany [5-10]. Fuzzy inference systems are applied to combine up to eight different tissue parameters and to model their underlying distributions. The system has successfully been applied to various problems in ultrasonic tissue characterization, e.g. the differentiation between plaque, vessel wall and blood in intravascular ultrasound [11], the characterization and staging of deep venous thrombosis [12-13], the monitoring of thermal ablation therapy of the liver [14], the exploration of metabolism effects of the liver [15] and the differentiation between different types of parotid tumors [16].

Throughout this chapter, an illustrative numerical detection example is presented and discussed. The most important steps of ROC analysis are reviewed and performed using the simulated data of this example. The example and all figures shown in this chapter have been calculated and visualized using the software package published by the author [4].

Historically, the term 'receiver operating characteristic' originates from the evaluation of systems and techniques for military reconnaissance. The ROC measure is used to describe the dependence of one's own losses due to friendly fire in contrast to hit enemy targets. Furthermore, the term sometimes is expressed as 'receiver-operator characteristic', because it describes the relationship between the receiver hardware (e.g. sonar or radar system) and the operator interpreting the display or audible signals of the receiver. Today, the terminology 'relative operating characteristic' is sometimes used in order to break with the past. The problem of reconnaissance can easily be compared to computer-aided detection systems. Both approaches seldom achieve 100% accuracy, thus an ideal operating point depending on actual requirements has to be found. In both cases, only the ROC curve, and accordingly, the area under the ROC curve provide an estimate of the overall performance of the system.

2. The Detection Problem

Most detection systems provide a quasi-continuous output value k for each single case of the population. The population is the entire set of data that is available and processed by the computer-aided detection system. To name but a few, in the case of computer-aided diagnostic systems in medical environments, the population may consist of parameter vectors with entries belonging to one patient at a time, but it may also consist of parameter vectors that contain several values that have been extracted from each patient as a set. In the special case of medical diagnostic imaging, these parameter vectors may have been extracted from numerous regions of interest (ROI) of diagnostic images (e.g. ultrasound, MRI, CT, etc.) of several patients or even of a single patient. In the case of industrial detection systems, the population may consist of parameter vectors with entries belonging to an analyzed product or sample at a time. In the case of optical character recognition systems, the population may consist of parameter vectors that contain information about a letter to be recognized.

For example in ultrasonic tissue characterization, parameter vectors typically contain tissue-describing features, such as spectral or textural parameter estimations, but they may also contain morphological parameters or clinical variables in addition to ultrasonic features [5-10, 17-31]. In an industrial environment, the parameter vectors may contain diverse measures, like image features originating from a camera or an x-ray system used for flaw control. In a detection system used for reconnaissance, the parameter vectors may contain features extracted from satellite or radar images. In speech recognition experiments, the

parameter vectors may contain features or cues that might lead to the detection and recognition of words.

In many approaches, the detector must differentiate between two target groups or target classes: Ω_P and Ω_N, where all positive cases belong to the class Ω_P and all negative cases belong to the class Ω_N, respectively. A two-class problem is called a dichotomous problem [32].

The true membership of the cases to the target classes is defined by the so-called 'gold standard', which is the best estimation of class membership available. For example, in medical applications, histological results are typically used as the gold standard, but even diagnostic modalities may also be used as a gold standard, if no histological results are available. In the example of speech recognition experiments, the gold-standard will be the actual spoken or transmitted word as known by the operator of the experiment.

When the gold standard is known, the probabilities $p_P(k)$ and $p_N(k)$ of the output value k of the detection system can be determined by simply comparing the output values of the detection system to the results of the gold standard. The probabilities $p_P(k)$ and $p_N(k)$ stand for the probability of membership between an output value k and a certain class or target group Ω_P or Ω_N, respectively.

The probabilities $p(\Omega_P)$ and $p(\Omega_N)$, which are independent of the actual output value k, are the *a priori* probabilities of the two target groups. The *a priori* probabilities simply are the probabilities of incidence of the single target groups within the whole population. The probabilities $p(k|\Omega_P)$ and $p(k|\Omega_N)$ are the conditional probabilities of the output value k of the detection system belonging to a certain target group without regarding the *a priori* probabilities of the target groups themselves. The symbol '|' in the definitions initiates the condition. The definition 'conditional probability' means that the probability of the output value k is provided under the condition that the case belongs to the certain target group Ω_P or Ω_N given in the definition:

$$p_P(k) = p(\Omega_P) \cdot p(k|\Omega_P), \tag{1}$$

$$p_N(k) = p(\Omega_N) \cdot p(k|\Omega_N). \tag{2}$$

The *a priori* probabilities $p(\Omega_P)$ and $p(\Omega_N)$ of the two target groups Ω_P and Ω_N are usually estimated from the number of cases in the gold standard:

$$p(\Omega_P) = \frac{N_{GS,P}}{N_{GS,P} + N_{GS,N}}, \tag{3}$$

$$p(\Omega_N) = \frac{N_{GS,N}}{N_{GS,P} + N_{GS,N}}, \tag{4}$$

where $N_{GS,P}$ stands for the overall amount of positive cases within the population according to the gold standard, and $N_{GS,N}$ stands for the overall amount of negative cases according to the gold standard, respectively.

The sum of the integral values of both probability functions is unity:

$$\int_{k=-\infty}^{\infty} p_P(k) + p_N(k)\, dk = 1. \tag{5}$$

The integral equation above can be replaced by its corresponding sum equation for discrete variables, which typically is the case in computer-aided detection systems:

$$\sum_{k=0}^{K} \hat{p}_P(k) + \hat{p}_N(k) = 1. \tag{6}$$

The integral values of both conditional probability functions are unity as $p(\Omega_P) + p(\Omega_N) = 1$ with $0 \le p(\Omega_P) \le 1$ and $0 \le p(\Omega_N) \le 1$:

$$\int_{k=-\infty}^{\infty} p(k|\Omega_P)\, dk = 1, \tag{7}$$

$$\int_{k=-\infty}^{\infty} p(k|\Omega_N)\, dk = 1. \tag{8}$$

Again, the above integral equations can be replaced by its corresponding sum equations when dealing with discrete variables:

$$\sum_{k=0}^{K} \hat{p}(k|\Omega_P) = 1, \tag{9}$$

$$\sum_{k=0}^{K} \hat{p}(k|\Omega_N) = 1. \tag{10}$$

An important fact about the estimation of the probabilities that has to be considered is the following: not only the threshold-dependent probabilities of a group may vary in different populations, but also the estimations of the *a priori* probabilities may tend to vary, thus, making a comparison between different populations difficult.

In many cases, the disposal of *a priori* probabilities is preferred when using computer-aided detection systems [7]. In the case of parametrical approaches, this

is done by normalizing the threshold-dependent probability functions by known *a priori* probabilities. Training the system with the same amount of data of all groups also leads to equalized probability functions. Unequalized training affects the separation thresholds of the system and often leads to the preference of the group with the highest amount of training data involved. In most cases, the preference of a certain target group is not desired, but the system is designed to provide unbiased separation thresholds. The final choice of the separation threshold has to be made by the user during the evaluation of the detection results. Unequalized training may also lead to biased ROC estimations.

As a demonstrative example, simulated detection results were chosen to illustrate the methods and definitions mentioned throughout this chapter. A detection system was applied on data originating from 30 cases. The detection system presented the following output vector $\mathbf{k}$ containing $N = 30$ output values k in the range of 0 to 9:

$$\mathbf{k} = [3\ 5\ 5\ 3\ 0\ 2\ 6\ 3\ 7\ 4\ 8\ 5\ 6\ 5\ 8\ 3\ 1\ 3\ 7\ 1\ 9\ 4\ 2\ 6\ 4\ 2\ 6\ 5\ 4\ 1]. \quad (11)$$

The gold standard k_{GS} was known for all 30 simulated cases and was given as the following binary vector $\mathbf{k}_{GS}$:

$$\mathbf{k}_{GS} = [1\ 0\ 1\ 1\ 1\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 1\ 1]. \quad (12)$$

As a dichotomous problem, e.g. the differentiation between positive cases and negative cases, was proposed, ‘1’ stands for positive cases and ‘0’ stands for negative cases, respectively.

Several separation thresholds k_C were applied to analyze the systems behavior at numerous operating points. The applied separation thresholds are given as the following vector $\mathbf{k}_C$:

$$\mathbf{k}_C = [-0.5\ 0.5\ 1.5\ 2.5\ 3.5\ 4.5\ 5.5\ 6.5\ 7.5\ 8.5\ 9.5]. \quad (13)$$

In this example, the separation thresholds k_C were manually chosen to lie between the values of the output vector $\mathbf{k}$ in order to keep the example as concise and simple as possible. In addition, one separation threshold was chosen to lie below the minimum entry and one separation threshold was chosen to lie above the maximum entry of the output vector. In real life applications, the vector of separation thresholds has to be chosen to be resolved sufficiently fine to allow precise evaluation of the underlying detection system.

The probability curves for the simulated dichotomous detection problem are shown in Fig. 1. As already mentioned above, dichotomous detection problems are typical for many computer-aided detection systems. In many cases, higher order detection problems are reduced to dichotomous problems in order to simplify the detection task.

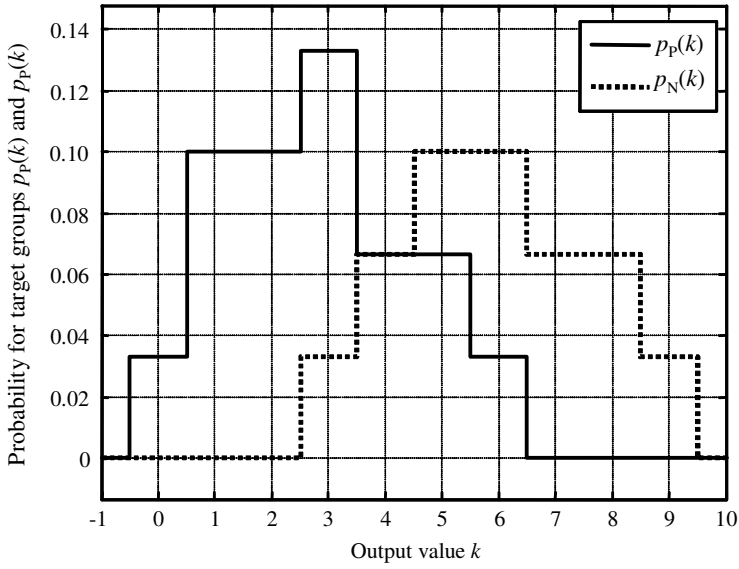


Figure 1. Probability curves $p_P(k)$ and $p_N(k)$ constructed from the output values of a detection system for a simulated but typical dichotomous problem. Because of the limited size of the underlying data, the typical Gaussian shape is only rudimentary apparent. In this example, the two target groups cannot be separated completely. Both curves show a considerable overlap, which is typical for real life detection problems.

When the size of the population is large enough to support sufficient stability of the system, the shapes of the probability curves, in most cases, approximate Gaussian distributions according to the central limit theorem [33-35]. In the ideal case, which is the complete separation of the classes, the overlap of the curves is zero. Under Monte Carlo conditions, i.e. the special case in which both classes are distributed equally, the curves overlap completely, thus, making a separation impossible [34].

A detection system is designed to separate the two groups while keeping the overlap as low as possible. When the probability curves cannot be separated completely, which typically is the case in real life detection systems, a separation threshold has to be defined to separate the groups and to determine which output values have to be considered to belong to which target group.

Often the presentation of the detection results as a single value is preferred, in order to compare different detectors with each other or to optimize the computer-aided detection system [32, 35-37]. The performance of a detection system can be provided using threshold specific measures, which provide a performance measure of the system for one fixed operating point only. These measures are described in the following section 3. Another method to measure the performance

of a detection system is using threshold-independent approaches, which yield a measure that describes the performance of the system for all possible operating points. These threshold-independent measures are described in section 4.

3. Threshold Specific Performance Measures

When considering threshold specific performance measures, any single decision of a detection system yields a value that can be characterized as true positive $TP(k_C)$, true negative $TN(k_C)$, false positive $FP(k_C)$, or false negative $FN(k_C)$ in relation to the gold standard and depending on the separation threshold or cut-off point k_C, which separates the output range into different target classes. $TP(k_C)$ is defined as the number of correctly detected positive cases while $TN(k_C)$ is defined as the number of correctly detected negative cases, respectively. Whereas $FP(k_C)$ is the number of incorrectly as positive detected but negative cases and $FN(k_C)$ is the number of incorrectly as negative detected but positive cases, respectively.

Using these four values, the whole population of cases, which is $TP(k_C) + TN(k_C) + FP(k_C) + FN(k_C)$, can be characterized according to Fig. 2 and the following definitions. The diagram shown in Fig. 2 is known as the statistical nomination matrix [38].

Gold standard / Detector	$N_{GS,P}$ Positive	$N_{GS,N}$ Negative
$N_{D,P}(k_C)$ Positive	$TP(k_C)$ True positive	$FP(k_C)$ False positive
$N_{D,N}(k_C)$ Negative	$FN(k_C)$ False negative	$TN(k_C)$ True negative

Figue 2. Statistical nomination matrix showing the denominations of threshold specific measures of detection results. This table is also referred to as contingency table or confusion matrix. The results of a detection system are compared to the results of the so-called gold standard, which, in many cases, is a detection system itself.

In the following definitions, $N_{D,P}(k_C)$ stands for the number of cases detected as positive by the detection system, while $N_{D,N}(k_C)$ stands for the number of cases detected as negative by the detection system, respectively. Whereas $N_{GS,P}$ stands for the actual number of positive cases within the population according to the gold standard, and $N_{GS,N}$ stands for the actual number of negative cases according

to the gold standard, respectively. The symbol 'I' in the following definitions initiates the condition:

$$TP(k_C) = \left\{ N_{D,P}(k_C) \middle| N_{D,P}(k_C) \in \Omega_P \right\}, \text{ with } TP(k_C) = \left\{ 0...N_{GS,P} \right\}, \quad (14)$$

$$TN(k_C) = \left\{ N_{D,N}(k_C) \middle| N_{D,N}(k_C) \in \Omega_N \right\}, \text{ with } TN(k_C) = \left\{ 0...N_{GS,N} \right\}, \quad (15)$$

$$FP(k_C) = \left\{ N_{D,P}(k_C) \middle| N_{D,P}(k_C) \in \Omega_N \right\}, \text{ with } FP(k_C) = \left\{ 0...N_{GS,N} \right\}, \quad (16)$$

$$FN(k_C) = \left\{ N_{D,N}(k_C) \middle| N_{D,N}(k_C) \in \Omega_P \right\}, \text{ with } FN(k_C) = \left\{ 0...N_{GS,P} \right\}. \quad (17)$$

Eight typical threshold-specific performance measures are used to describe the performance of computerized detection systems and detection systems in general. The two most commonly used measures are sensitivity $SE(k_C)$ and specificity $SP(k_C)$. Sensitivity is also known as the probability of detection $PD(k_C)$ or true-positive fraction. Specificity is also known as the true-negative fraction. The values for both measures are always in the range of 0 to 1:

$$SE(k_C) = \frac{TP(k_C)}{TP(k_C) + FN(k_C)}, \quad (18)$$

$$SP(k_C) = \frac{TN(k_C)}{TN(k_C) + FP(k_C)}. \quad (19)$$

Neither sensitivity nor specificity is affected by possible prevalence of one of the underlying classes. This property makes these two measures the measures of choice for most detection problems.

The probability of false alarm $PFA(k_C)$ or the false-positive fraction is one minus specificity. This measure is often used to plot the ROC curve, which is described in the next section. The probability of false clearance $PFC(k_C)$ or false-negative fraction, which is one minus sensitivity, is rarely used but mentioned here for completeness:

$$PFA(k_C) = 1 - SP(k_C) = \frac{FP(k_C)}{TN(k_C) + FP(k_C)}, \quad (20)$$

$$PFC(k_C) = 1 - SE(k_C) = \frac{FN(k_C)}{TP(k_C) + FN(k_C)}. \quad (21)$$

Like sensitivity and specificity, neither the probability of false alarm nor the probability of false clearance is affected by a possible prevalence of one of the underlying classes.

In contrast to the first four measures, the positive predictive value $PPV(k_C)$ and the negative predictive value $NPV(k_C)$, its opposite, are affected by a possible prevalence of one group or class within the population and, thus, have to be used with care because they easily can lead to false conclusions when comparing different detection results originating from different datasets with each other:

$$PPV\left(k_C\right) = \frac{TP\left(k_C\right)}{TP\left(k_C\right) + FP\left(k_C\right)}, \tag{22}$$

$$NPV\left(k_C\right) = \frac{TN\left(k_C\right)}{TN\left(k_C\right) + FN\left(k_C\right)}. \tag{23}$$

From an analytical point of view, the sensitivity $SE(k_C)$ at threshold k_C is calculated by integrating the probability curve $p(k|\Omega_P)$ of the output values belonging to the positive target group from $-\infty$ to k_C. The specificity $SP(k_C)$ at threshold k_C is calculated by integrating the probability curve $p(k|\Omega_N)$ of the output values belonging to the negative target group from k_C to $+\infty$, as shown in the following definitions:

$$SE\left(k_C\right) = \int_{-\infty}^{k_C} p\left(k\left|\Omega_P\right.\right) dk, \quad \text{for } \mu_P < \mu_N, \tag{24}$$

$$SP\left(k_C\right) = \int_{k_C}^{\infty} p\left(k\left|\Omega_N\right.\right) dk, \quad \text{for } \mu_P < \mu_N. \tag{25}$$

In the equations stated above, the direction of integration is shown for the case that the positive density curve is placed to the 'left', i.e. towards lower values, of the negative density curve. This assumption is consistent with the numerical example given in this chapter and shown in Fig. 1. Assuming a Gaussian distribution, this means $\mu_{P<}\mu_N$, where μ_P stands for the center of the positive density and μ_N for the center of the negative density, respectively. In other cases, the positions of the densities may be inverted, thus, the directions of integration have to be inverted as well in order to achieve consistent results.

For the discrete case, the above integral equations can be replaced by sum equations. However, a special case has to be considered: in the following definitions, the probability at exactly the separation threshold k_C is considered *per definitionem* to belong to the sensitivity measure:

$$SE_k\left(k_C\right) = \sum_{k=0}^{k_C} \hat{p}\left(k\left|\Omega_P\right.\right), \quad \text{for } \mu_P < \mu_N, \tag{26}$$

$$SP_k(k_C) = \sum_{k=k_C+1}^{K} \hat{p}(k|\Omega_N), \quad \text{for } \mu_P < \mu_N. \tag{27}$$

Like sensitivity and specificity, the probability of false alarm and the probability of false clearance can be expressed as integrals over the probability curves $p(k|\Omega_N)$ and $p(k|\Omega_P)$ as follows:

$$PFA(k_C) = \int_{-\infty}^{k_C} p(k|\Omega_N)\,dk, \quad \text{for } \mu_P < \mu_N, \tag{28}$$

$$PFC(k_C) = \int_{k_C}^{\infty} p(k|\Omega_P)\,dk, \quad \text{for } \mu_P < \mu_N. \tag{29}$$

Again, the integral equations above can be replaced by their corresponding sum equations for the case of discrete variables. In the discrete case, the probability at exactly the separation threshold k_C is considered *per definitionem* to belong to the probability of false clearance:

$$PFA_k(k_C) = \sum_{k=0}^{k_C-1} \hat{p}(k|\Omega_N), \quad \text{for } \mu_P < \mu_N, \tag{30}$$

$$PFC_k(k_C) = \sum_{k=k_C}^{K} \hat{p}(k|\Omega_P), \quad \text{for } \mu_P < \mu_N. \tag{31}$$

The accuracy $ACC(k_C)$, which is often referred to in the literature, and the probability of error $PE(k_C)$ also have the disadvantage of being dependent on the *a priori* probabilities or prevalence of the target groups [39].

These measures are only representative measures if the class distributions on the evaluation dataset are meaningful [39]. If the *a priori* probabilities of the target groups are not equal, the accuracy curve will reflect the ratio of the sizes of the positive and the negative target group. To overcome this disadvantage, the accuracy measure is sometimes defined as the product or the weighted sum of sensitivity and specificity [40]. However, the accuracy measure is presented here using the conventional and widely used definition:

$$ACC(k_C) = \frac{TP(k_C) + TN(k_C)}{TN(k_C) + TP(k_C) + FN(k_C) + FP(k_C)}, \tag{32}$$

$$PE(k_C) = 1 - ACC(k_C) = \frac{FP(k_C) + FN(k_C)}{TN(k_C) + TP(k_C) + FN(k_C) + FP(k_C)}. \tag{33}$$

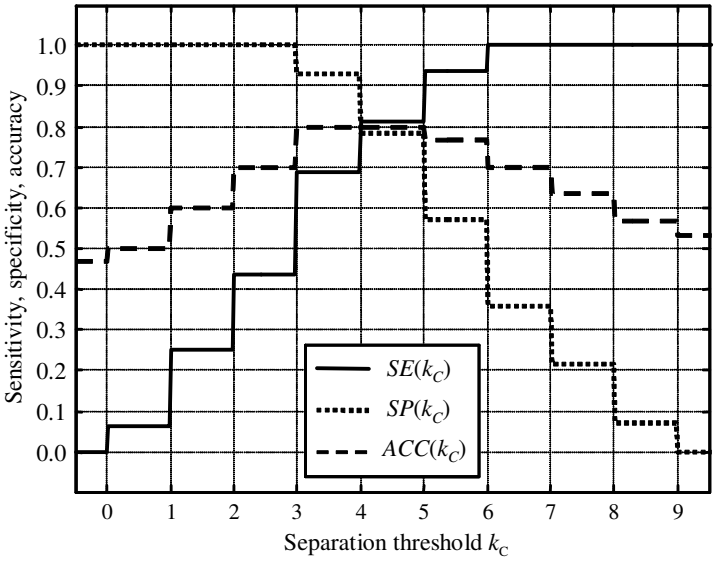


Figure 3. Typical curves for sensitivity $SE(k_C)$, specificity $SP(k_C)$ and accuracy $ACC(k_C)$ under dependence of the separation threshold k_C. The dependence of the accuracy of the individual sizes of the target groups, i.e. the *a priori* probabilities of the target groups, can be seen as the factor underlying the accuracy curve. Sensitivity and specificity are independent of any prevalence in the target groups.

In the equations above, k_C is the separation threshold or cut-off point. To visualize the trade-off between sensitivity and specificity, these two measures are often plotted against the threshold k_C. Sometimes, the accuracy $ACC(k_C)$ is plotted additionally into the same figure, as can bee seen in Fig. 3.

Sensitivity $SE(k_C)$ and specificity $SP(k_C)$ of the exemplary detection system discussed in this chapter can be estimated from the simulated detection results $\mathbf{k}$ and from the gold standard $\mathbf{k}_{GS}$ for all separation thresholds $\mathbf{k}_C$ following the definitions given in equations (18) and (26) and in equations (19) and (27), respectively. The accuracy $ACC(k_C)$ is estimated additionally for illustrative reasons following definition (32), although only sensitivity and specificity are needed for the estimation of the area under the ROC curve, which will be the next step. The results are shown in Fig. 3 and given in the following vectors:

$$\mathbf{SE}(\mathbf{k}_C) = [0\ 0.06\ 0.25\ 0.44\ 0.69\ 0.81\ 0.94\ 1\ 1\ 1\ 1], \tag{34}$$

$$\mathbf{SP}(\mathbf{k}_C) = [1\ 1\ 1\ 1\ 0.93\ 0.79\ 0.57\ 0.36\ 0.21\ 0.07\ 0], \tag{35}$$

$$\mathbf{ACC}(\mathbf{k}_C) = [0.47\ 0.5\ 0.6\ 0.7\ 0.8\ 0.8\ 0.77\ 0.7\ 0.63\ 0.57\ 0.53]. \tag{36}$$

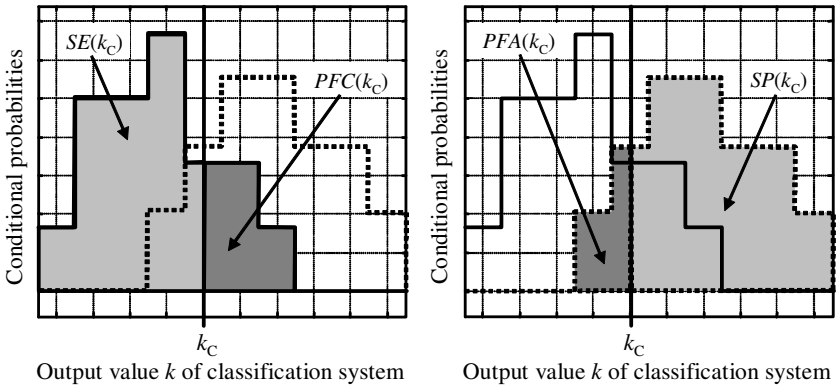


Figure 4. Sensitivity $SE(k_C)$, probability of false clearance $PFC(k_C)$, probability of false alarm $PFA(k_C)$ and specificity $SP(k_C)$ as calculated as areas under the conditional probability curves for a fixed separation threshold k_C. Sensitivity and specificity stand for the amount of cases that are correctly classified as positive or negative, respectively. The probability of false clearance and the probability of false alarm express the fraction of cases that are erroneously classified as negative or positive, respectively.

An illustrative visualization of the four most important threshold-dependent performance measures, sensitivity, specificity, probability of false clearance and probability of false alarm is given in Fig. 4.

There is one important point about the use of threshold specific performance measures: most computer-aided detection systems do not provide a binary output value. At the same time, designing a system that only yields binary output values is not desired in most cases. Most detectors yield a quasi-continuous output. Therefore, providing sensitivity and specificity only at a single threshold or cut-off point k_C does not provide a representative indication of the overall performance of the system. A system ideally is evaluated at all operating points k, to provide an objective judgment about the overall performance of the system. Even when the final detection results are presented as a binary decision, e.g. yes/no or positive/negative, the overall performance of a parameter or a classifier is required for comparison, and thus, must be computed during the performance evaluations. The easiest way to provide a single value measure for a detection system is to calculate the area, A_{ROC}, under the receiver operating characteristic curve, as described in section 4.

4. Threshold-independent Performance Measures

The equal detection rate *EDR* or equal error rate *EER*, which is the sensitivity and the specificity at the operating point $SE(k_C) = SP(k_C)$, is often used as the only performance measure of a detection system [41]:

$$EER = SE\left(k_C\right)\Big|_{SE(k_C)=SP(k_C)} = SP\left(k_C\right)\Big|_{SP(k_C)=SE(k_C)} . \tag{37}$$

However, this measure may be misleading, because the performance of the detection system is provided for only one single operating point k_C [36]. The overall performance of a detection system that is not dependent of a separation threshold or a cut-off point can easily be provided using the area A_{ROC} under the receiver operating characteristic curve [1-4, 33, 38, 39, 42-54]. The ROC curve is best constructed by plotting threshold-dependent sensitivity $SE(k_C)$ against specificity $SP(k_C)$ for all possible separation thresholds k_C:

$$SE\left(k_C\right) = f\left(SP\left(k_C\right)\right). \tag{38}$$

However, the definition above is considered 'non-standard' by several authors. Another way to construct the ROC curve that is often used in the literature plots the threshold-dependent sensitivity $SE(k_C)$ against $1 - SP(k_C)$, which actually is the probability of false alarm $PFA(k_C)$. This method provides the same value for the area under the ROC curve, but simply flips the ROC curve from left to right [33, 42].

In the present article, ROC curves are constructed by plotting the prevalence-independent sensitivity and specificity measures against each other, which best displays the tradeoff between these two variables and is considered more meaningful than the commonly used method of plotting sensitivity against the probability of false alarm.

According to definition (38), the area under the ROC curve A_{ROC} is estimated by the following integral equation:

$$A_{ROC} = \int_0^1 SE\left(SP\left(k_C\right)\right) dSP\left(k_C\right) = \int_0^1 SP\left(SE\left(k_C\right)\right) dSE\left(k_C\right). \tag{39}$$

Considering discrete variables $SE_k(k_C)$ and $SP_k(k_C)$, the area under the ROC curve is best estimated using the following equation, which takes into account the staircase characteristics of the underlying measures:

$$\begin{aligned} A_{ROC} = {} & \tfrac{1}{2}\left[SE_k\left(0,1,...,K-1\right) + SE_k\left(1,2,...,K\right)\right] \\ & \cdot\left[SP_k\left(0,1,...,K-1\right) - SP_k\left(1,2,...,K\right)\right]^T . \end{aligned} \tag{40}$$

The area under the ROC curve is usually provided as a value between 0 and 1, where 1 stands for ideal detection results. Areas less than 1 occur when actually positive cases are classified as negative and actually negative cases are classified as positive. Areas of exactly 0.5 stand for Monte Carlo results, which theoretically mean a worst-case detection with completely overlapping

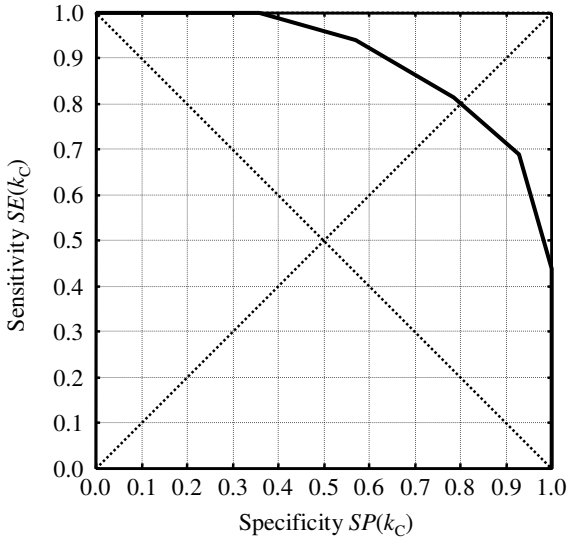


Figure 5. Typical ROC curve. Sensitivity $SE(k_C)$ is plotted against specificity $SP(k_C)$. Sometimes sensitivity is plotted against one minus specificity (i.e. probability of false alarm), which results in a left/right flipped ROC curve. The area under the ROC curve is usually provided as a value between 0 and 1, where 1 stands for ideal detection results and 0.5 stand for Monte Carlo results, which theoretically mean a worst-case detection with completely overlapping probability densities. The area under the ROC curve can also take values below 0.5. In these cases, the probability curves have 'changed their places', which may occur when using cross-validation methods. Two dashed lines are added to the diagram. The first, going down from the upper left, is called the line of zero discrimination. The second line marks possible positions of equal detection rate, and, thus, is called the line of equal detection.

probability densities. The area under the ROC curve can also take values below 0.5. In these cases, the probability curves have 'changed their places', which may occur when using cross-validation methods. Sometimes the area under the ROC curve is given in percentage between 0 and 100 instead of values between 0 and 1.

The ROC curve based on the simulated data of the example presented in this chapter and constructed according to definition (38) is shown in Fig. 5. The area under the ROC curve of the example is estimated as $A_{ROC} = 0.9$ following the definition given in (40).

Often, two diagonal lines are drawn into the same figure. The line going from [0, 0] to [1, 1] crosses the point of the ROC curve where sensitivity $SE(k_C)$ equals specificity $SP(k_C)$ and, thus, may help reading the ROC curve and evaluating the equal error rate or equal detection rate. The other diagonal line highlights the shape of the ROC curve for Monte Carlo conditions, i.e. $A_{ROC} = 0.5$.

Compared to performance measures that are dependent on a certain separation threshold k_C, such as sensitivity $SE(k_C)$ or specificity $SP(k_C)$, the area A_{ROC} under the ROC curve provides an estimate of the overall performance of a detection system for all operating points and, thus, is the ultimate measure for all detection tasks.

For most detectors or classifiers, including neural networks and fuzzy inference systems, the separation threshold k_C can be varied continuously, thus, permitting the estimation of sensitivity and specificity for all operating points of the system [51, 54]. This approach allows the construction of ROC curves representative for the whole system. For special cases, where only partial data are available, the ROC curve can be constructed from available data only. The so-called partial area index can be estimated from incomplete ROC curves.

In many cases, detectors that depend on a fixed separation threshold or even locate a single threshold that is optimal from a particular point of view are used. For example, the conventional Bayes classifier includes *a priori* knowledge about the target groups in order to work on an optimal fixed separation threshold. When ROC methods have to be applied for a detection method using a fixed separation threshold, recurrent calculation of sensitivity $SE(k_C)$ and specificity $SP(k_C)$ for different sizes of the target groups $N_{T,P}$ and $N_{T,N}$ within the training-data set can be performed. Using this method, the formerly fixed separation threshold is varied according to the manually introduced *a priori* probabilities, which are proportional to $N_{T,P}$ and $N_{T,N}$, respectively. The separation threshold k_C can then be defined as:

$$k_C = \frac{N_{T,P}}{N_{T,N}}. \tag{41}$$

Cross-validation experiments are usually applied to measure the performance of detection systems. When using cross-validation, different systems are repeatedly trained using different training datasets and repeatedly evaluated using different test datasets. The test datasets naturally have to be independent of the data included in the associated training datasets. Usually applied cross-validation methods are five-fold cross-validation or leave-one-out cross-validation, also known as total cross-validation [55]. The area under the ROC curve A_{ROC} is then stated as the mean area over cross-validation experiments together with its cross-validation standard deviation σ_{ROC} or its confidence intervals. The usually applied confidence interval is 95 %.

In a fundamental paper on ROC methods, Hanley *et al.* introduced a method to calculate the standard error E_{ROC} of an estimate of the area under the ROC

curve [49]. The method is based on nonparametric Wilcoxon or Mann-Whitney statistics, and is defined as follows:

$$E_{\mathrm{ROC}} = \sqrt{\frac{A_{\mathrm{ROC}}\left(1 - A_{\mathrm{ROC}}\right) + \left(N_{\mathrm{GS,P}} - 1\right)\left(Q_1 - A_{\mathrm{ROC}}{}^2\right) + \left(N_{\mathrm{GS,N}} - 1\right)\left(Q_2 - A_{\mathrm{ROC}}{}^2\right)}{N_{\mathrm{GS,P}} N_{\mathrm{GS,N}}}}, \quad (42)$$

with the two probability estimations Q_1 and Q_2 defined as:

$$Q_1 = \frac{A_{\mathrm{ROC}}}{\left(2 - A_{\mathrm{ROC}}\right)}, \quad (43)$$

and

$$Q_2 = \frac{2A_{\mathrm{ROC}}{}^2}{\left(1 + A_{\mathrm{ROC}}\right)}. \quad (44)$$

These definitions above will lead to reliable results, if the numbers of cases depicted by $N_{\mathrm{GS,P}}$ and $N_{\mathrm{GS,N}}$ represent independent cases, e.g. individual samples. However, interdependence or correlation between the cases might lead to biased estimations of E_{ROC} [16, 56].

According to the gold standard of the simulated example presented in this chapter, the number of negative cases is $N_{\mathrm{GS,N}} = 14$ and the number of positive cases is $N_{\mathrm{GS,P}} = 16$, respectively. Thus, the standard error of the estimate of the area under the ROC curve yields $E_{\mathrm{ROC}} = 0.06$ for the numerical example of this chapter and following definitions (42), (43) and (44).

Special care has to be taken when comparing different ROC curves intersecting each other. Methods for comparing different ROC curves are provided in [35, 39, 43, 44, 57].

Sometimes, only the equal error rate *EER* or equal detection rate *EDR* is reported in publications on detection systems, although the area under the ROC curve would be interesting to know in order to compare the results of the study with the results of other studies. In addition, sometimes only the area under the ROC curve is reported while the equal error rate is missing. In Fig. 6, the equal error rate *EER* is given as a function of the area under the ROC curve A_{ROC} in the range of 0.5 to 1.0 for the special case of normally distributed target groups.

No closed form for the relationship can be given, because the error function is involved in estimating the equal error rate from a given estimate of the area under the ROC curve and vice versa. For distributions not satisfying a Gaussian shape, the relationship given in Fig. 6 can only be used as a first approximation.

The following polynomial approximates the equal error rate *EER* as a function of the area under the ROC curve A_{ROC} in the range of 0.5 to 1.0. The root mean square error of the approximation is smaller than 0.005:

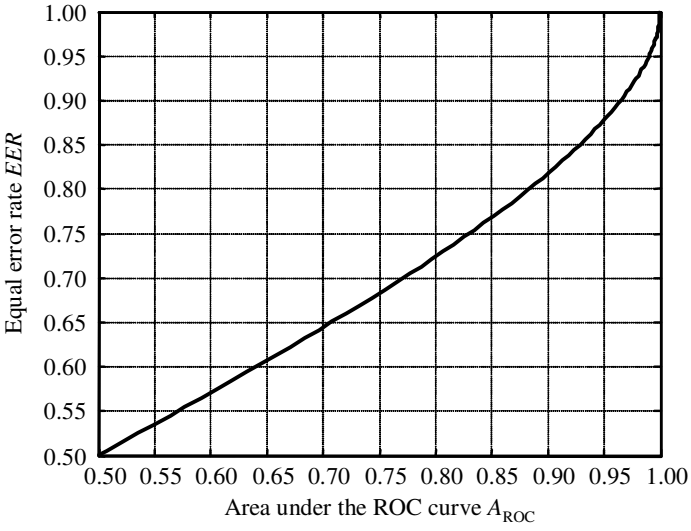


Figure 6. Equal error rate *EER* as a function of the area under the ROC curve A_{ROC} for the special case of normally distributed target groups. No closed form of the relationship can be given because the error function is involved in estimating the equal error rate from a given estimate of the area under the ROC curve. For distributions not satisfying a Gaussian shape, the relationship given in the figure can only be used as a first approximation.

$$\begin{aligned} EER = & -604.28 + 6052.85\, A_{ROC} - 25731.35\, A_{ROC}^{\ 2} + 60219.39\, A_{ROC}^{\ 3} \\ & - 83792.43\, A_{ROC}^{\ 4} + 69332.59\, A_{ROC}^{\ 5} - 31593.55\, A_{ROC}^{\ 6} + 6117.78\, A_{ROC}^{\ 7}. \end{aligned} \tag{45}$$

Vice versa, the following polynomial approximates the area under the ROC curve A_{ROC} as a function of the equal error rate *EER* in the range of 0.5 to 1.0. The root mean square error of the approximation is smaller than 0.0006:

$$A_{ROC} = 0.199 - 0.703\, EER + 3.727\, EER^2 - 2.223\, EER^3 . \tag{46}$$

For the simulated example presented in this chapter and yielding an area under the ROC curve of $A_{ROC} = 0.9$, an equal error rate of $EER = 0.82$ is calculated according to definition (45). A satisfying agreement can be found comparing this outcome with the results of the example shown in Fig. 3 and Fig. 5 where an equal error rate of $EER = 0.80$ is estimated.

Special attention has to be paid when applying the above approximations to real life data, as they are only applicable to normally distributed data. If the data are not normally distributed, a bias will be introduced into the results. The simulated data of the finite example presented within this chapter can only be considered to approximate a normal distribution. Therefore, the small deviation of the true equal error rate could have been expected.

5. Conclusion

Several measures of detection performance of computer-aided detection systems have been reviewed in this article. Disadvantages and problems of the different performance measures have been pointed out. In addition to separation threshold-specific performance measures, a separation threshold-independent performance measure, the area under the ROC curve, has been discussed.

When performance measures, such as sensitivity and specificity, are based on a single separation threshold, the measures only express the performance of the system at the operating point that corresponds to a single threshold. This pitfall may easily misdirect the scientist, who is evaluating or designing a complex detection system. Only a performance measure, such as the area under the ROC curve, that takes the entire operating range of the underlying detection system into account, can provide the user with a meaningful and effective means of selecting, designing, and optimizing a detection system. Nevertheless, the practical use of ROC analysis is greatly overlooked in the literature on computer-aided detection systems, which often focuses on ideal conditions of the underlying data.

The statistics of ROC methods have been summed up and put into a theoretical framework. Only performance measures that are independent of the separation threshold can fully represent the overall performance of a computer-aided detection system. The application of the area under the ROC curve as a performance measure for detection systems in general, including computer-aided methods, is highly recommended for making meaningful performance comparisons of different detection techniques and systems.

Acknowledgments

The author would like to thank Dr. Helmut Ermert and his team at the Institute of High Frequency Engineering at Ruhr-University Bochum, Germany, and all colleagues at the Center of Excellence for Medical Technology (Kompetenzzentrum Medizintechnik Ruhr - KMR), Bochum, Germany. The major part of this work was developed at these two institutes. Furthermore, many thanks go out to Dr. Tony Falco and his research team at Resonant Medical Inc., Montreal, Canada.

References

1. Stephan, C., Wesseling, S., Schink, T., Jung, K., Comparison of eight computer programs for receiver-operator characteristic analysis, *Clin Chem 49(3)*, 433-439 (2003).
2. Zweig, M. H., Campbell, G., Receiver-operator characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine, *Clin Chem 39(4)*, 561-577 (1993).
3. Scheipers, U., Perrey, C., Siebers, S., Hansen, C., Ermert, H., A tutorial on the use of ROC analysis for computer-aided diagnostic systems**,** *Ultrason Imaging 27*, 181-198 (2005)
4. Software for estimation of ROC curves 'SCHROC' of Ruhr-Univ. Bochum, Germany: www.scheipers.org.
5. Scheipers, U., Lorenz, A., Pesavento, A., et al., Ultrasonic multifeature tissue characterization for the early detection of prostate cancer, *Proceedings IEEE Ultrasonics Symposium*, 1265-1268 (2001).
6. Scheipers, U., Ermert, H., Lorenz, A., et al., Neuro-fuzzy inference system for ultrasonic multifeature tissue characterization for prostate diagnostics, *Proceedings IEEE Ultrasonics Symposium*, 1347-1350 (2002).
7. Scheipers, U., Ermert, H., Sommerfeld, H. J., et al., Ultrasonic multifeature tissue characterization for prostate diagnostics, *Ultrasound Med Biol 29(8)*, 1137-1149 (2003).
8. Scheipers, U., König, K., Sommerfeld, H. J., et al., Ultrasonic tissue characterization for the classification of prostate tissue, *Proceedings WCU*, 637-640 (2003).
9. Scheipers, U., Ermert, H., Sommerfeld, H. J., et al., Ultrasonic tissue characterization for prostate diagnostics: spectral parameters vs. texture parameters, *Biomed Tech 48(5)*, 122-129 (2003).
10. Scheipers, U., König, K., Sommerfeld, H. J., et al., Diagnostics of prostate cancer based on ultrasonic multifeature tissue characterization, *Proceedings IEEE Ultrasonics Symposium*, 2153-2156 (2004).
11. Perrey, C., Scheipers, U., Bojara, W., Holt, S., Lindstead, M., Ermert, H., Computerized segmentation of blood and luminal borders in intravascular ultrasound. *Proceedings IEEE Ultrasonics Symposium*, 1122-1125 (2004).
12. Siebers, S., Geier, B., Scheipers, U., Vogt, M., Mumme, A., Ermert, H., Staging of deep venous thrombosis using ultrasound. *Biomed Tech 49*, 846-847 (2004).
13. Siebers, S., Geier, B., Vogt, M., Scheipers, U., Mumme, A., Ermert, H., Classification of venous thrombosis combining ultrasound elastography and tissue characterization. *Proceedings IEEE Ultrasonics Symposium*, 1761-1764 (2004).
14. Siebers, S., Schwabe, M., Scheipers, U., Welp, C., Werner, J., Ermert, H., Evaluation of ultrasonic texture and spectral parameters for coagulated tissue characterization. *Proceedings IEEE Ultrasonics Symposium*, 1804-1807 (2004).
15. Hansen, C., Scheipers, U., Hüttebräuker, N., Gebel, M., Ermert, H., Trends in time series of parameters from ultrasonic images due to metabolic activities of the Human Liver. *In: Meinzer, H.-P., et al.: Bildverarbeitung für die Medizin*, 445-449 (2005).
16. Scheipers, U., Siebers, S., Gottwald, F., Ashfaq, M., Bozzato, A., Zenk, J., Iro, H., Ermert, H., Sonohistology for the computerized differentiation of parotid-gland tumors, *Ultrasound Med Biol 31(10)*, 1287-1296 (2005).
17. Basset, O., Sun, Z., Mestas, J. L., et al., Texture Analysis of ultrasonic images of the prostate by means of co-occurrence matrices, *Ultrason Imaging 15*, 218-237 (1993).
18. Feleppa, E. J., Porter, C. R., Ketterling J., et al., Recent developments in tissue-type imaging (TTI) for planning and monitoring treatment of prostate cancer, *Ultrason Imaging 26*, 163-172 (2004).

19. Hartman, P. C., Oosterveld, B. J., Thijssen, J. M. et al., Detection and differentiation of diffuse liver disease by quantitative echography. A retrospective assessment, *Invest Radiol 28*, 1-6 (1993).
20. Jenderka, K. V., Gärtner, T., Cobet, et al., Tissue characterization by imaging the local frequency dependent relative backscatter coefficient, *Ultrasonic Imaging and Signal Processing, Proceedings of SPIE 3982*, 270-277 (2000).
21. Lang, M., Ermert, H., Heuser, L., In vivo study of on-line liver tissue classification based on envelope power spectrum analysis, *Ultrason Imaging 16*, 77-86 (1994).
22. Lizzi, F. L., Feleppa, E. J., Astor, M., et al., Statistics of ultrasonic spectral parameters and liver examinations, *IEEE Trans Ultrason Ferroelec Freq Control 44*, 935-942 (1997).
23. Nair, A., Kuban, B. D., Obuchowski, N., et al., Assessing spectral algorithms to predict atherosclerotic plaque composition with normalized and raw intravascular ultrasound data, *Ultrasound Med Biol 27(10)*, 1319-1331 (2001).
24. O'Brien, S. P., Sigel, B., Justin, J., et al., Carotid plaque spaces relate to symptoms and ultrasound scattering, *Ultrasound Med Biol 30(5)*, 611-616 (2004).
25. Oelze, M. L., O'Brien Jr., W. D., Blue, J. P., et al., Differentiation and characterization of rat mammary fibroadenomas and 4T1 mouse carcinomas using quantitative ultrasound imaging, *IEEE Trans Med Imaging 23(6)*, 764-771 (2004).
26. Oosterveld, B. J., Thijssen, J. M., Hartman, P. C., et al., Ultrasound attenuation and texture analysis of diffuse liver disease: methods and preliminary results, *Phys Med Bio 36*, 1039-1064 (1991).
27. Schmitz, G., Ermert, H., Senge, T., Tissue characterization and imaging of the prostate using radio frequency ultrasonic signals, *IEEE Trans Ultrason Ferroelec Freq Control 46*, 126-138 (1999).
28. Sigel, B., Feleppa, E. J., Swami, V., et al., Ultrasonic tissue characterization of blood clots, *Surg Clin North Am. 70(1)*, 13-29 (1990).
29. Siebers, S., Welp, C., Werner, J., et al., Ultrasound-based imaging modalities for thermal therapy monitoring, *Biomed Tech 47*, 438-440 (2002).
30. Thijssen, J. M., Oosterveld, B. J., Hartman, P. C., et al., Correlations between acoustic and texture parameters from RF and B-mode liver echograms. *Ultrasound Med Biol 19*, 13-20 (1993).
31. Valckx, F. M. J., Thijssen, J. M., van Geemen, A. J., et al., Calibrated parametric medical ultrasound imaging, *Ultrason Imaging 22*, 57-75 (2000).
32. Sprent, P., and Smeeton, N. C., Applied Nonparametric statistical methods, (Chapman & Hall, London/United Kingdom, 2000).
33. Kroschel, K., Statistische Nachrichtentechnik, (Springer-Verlag, Berlin/Germany, 1996).
34. Böhme, J. F., Stochastische Signale, (B. G. Teubner, Stuttgart/Germany, 1993).
35. Hanley, J. A., and McNeil, B. J., A method for comparing the areas under receiver operating characteristic curves derived from the same cases, *Radiology 148(3)*, 839-843 (1983).
36. Verlinde, P., Chollet, G., and Acheroy, M., Multi-modal identity verification using expert fusion, *Information Fusion 1(1)*, 17-33 (2000).
37. Zhou, X. H., and Gatsonis, C. A., A simple method for comparing correlated ROC curves using incomplete data, *Statistics in Medicine 15(15)*, 1687-1693 (1996).
38. Webb, S., The physics of medical imaging, (Adam Hilger Publishing, Bristol/United Kingdom, 1988).
39. Metz, C. E., Basis principles of ROC analysis, *Semin Nucl Med 48(4)*, 283-298 (1978).
40. Anastasio, M. A., Kupinski, M. A., and Nishikawa, R. M., Optimization of FROC analysis of rule-based schemes using a multiobjective approach, *IEEE Trans Med Imaging 17(6)*, 1089-1093 (1998).

41. Oglesby, J., What's in a number? Moving beyond the equal error rate, *Speech Communication 17(2)*, 193-208 (1995).
42. Obuchowski, N. A., Nonparametric analysis of clustered ROC curve data, *Biometrics 53,* 567-578 (1997).
43. DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L., Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach", *Biometrics 44,* 837-845 (1988).
44. Metz, C. E., Wang, P.-L., Kronman, H. B., A new approach for testing the significance of differences between ROC curves measured from correlated data, *in F. Deconinck: Information processing in medical imaging, The Hague, Nijhoff,* 432-445 (1984).
45. Bamber, D., The area above the ordinal dominance graph and the area below the receiver operating characteristic graph", *Journal Mathematical Psychology 12*, 387-415 (1975).
46. Metz, C. E., and Kronman, H. B., Statistical significance tests for binormal ROC curves, *Journal Mathematical Psychology 22*, 218-243 (1980).
47. Egan, J. P., Greenberg, G. Z., and Schulman, A. I., Operating characteristics, signal detectability, and the method of free response, *Journal Acoustical Society America 33(8)*, 993-1007(1961).
48. Hanley, J. A., and McNeil, B. J., The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology 143(1)*, 29-36 (1982).
49. Grey, D. R., Some aspects of ROC curve-fitting: normal and logistic models, *Journal Mathematical Psychology 9*, 128-139 (1972).
50. Fukunaga, K., Introduction to statistical pattern recognition, (Academic Press, Boston, MA, 1990).
51. Hastie, T., Tibshirani, R., Friedman, J., The elements of statistical learning: data mining, inference, and prediction, (Springer Verlag, Series in Statistics, 2001).
52. Wagner, R. F., Beiden, S. V., Campbell, G., Metz, C. E., Sacks, W. M., Assessment of medical imaging and computer-assist systems: lessons from recent experience, *Acad Radiol 9*, 1264-1277 (2002).
53. Beiden, S. V., Maloof, M., Wagner, R. F., A general model for finite-sample effects in training and testing of competing classifiers, *IEEE Trans Patt Anal Machine Intell 25*, 1561-1569 (2003).
54. Bishop, C. M., Neural networks for pattern recognition, (Oxford University Press, Oxford/United Kingdom, 1995).
55. Hanley, J. A., McNeil, B. J., The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology 143*, 29-36 (1982).
56. Scheipers, U., Siebers, S., Ashfaq, M., Gottwald, F., Bozzato, A., Zenk, J., Iro, H., Ermert, H., Ultrasonic tissue characterization for the differentiation of parotid gland tumors, *Proceedings IEEE Ultrasonics Symposium*, 828-830 (2005).
57. Rockette, H. E., Obuchowski, N., Metz, C. E., Gur, D., Statistical issues in ROC curve analysis, *SPIE 1234, Medical Imaging IV: PACS System Design and Evaluation*, 111-119 (1990).

CHAPTER 10

WAVEGUIDES IN ACOUSTIC SENSOR SYSTEMS

Lawrence C. Lynnworth

Lynnworth Technical Services
77 Graymore Road, Waltham Massachusetts 02451-2201 USA
www.lynnworthtechnicalservices.com

Waveguide sensors and buffers have been applied to measuring three of the four principal measurands of interest in industrial process control, namely, liquid level, flow, and temperature. Buffers are used when the measurand is at high temperature, which sometimes occurs in NDE/NDT situations. Waveguide sensors can also sense liquid density and viscosity. Problems include isolating the sought measurand from interfering variables, and eliminating acoustic or ultrasonic noise conveyed by a pipe or other structural element that acts as an unwanted waveguide, channeling interference to a transducer and reducing the signal to noise ratio at the transducer.

1. Introduction

This chapter is mainly concerned with practical aspects of using guided waves as part of acoustic sensing systems. Emphasis is on industrial process control applications. Readers seeking theoretical foundation for the various waves to be discussed below are referred to texts such as [61] and the ultrasonic delay line literature. Numerous pre-1967 innovative solutions to buffering a transducer from high temperature are illustrated in [13]. Additional discussions of these thermal buffers and pre-1989 techniques and applications appeared in [38]. In a sense, as the Curie point or maximum use temperature of transducers increases above today's limits, the need for some of these thermal buffers diminishes. Where laser or other noncontact (emat) means evolve for launching or receiving acoustic/ultrasonic waves in hot inhospitable environments, thermal buffers may become obsolete [53].

My personal introduction to ultrasonic waveguides began in 1962 when I accepted, along with employment at Parametrics, Inc. (later, renamed

Panametrics, Inc.) the task of measuring elastic moduli of solids at elevated temperature. In that work, the waveguide was merely a thermal buffer that allowed bulk wave ultrasound to be conveyed from a transducer at ordinary temperature (T) to laboratory oven-heated specimens that were red-hot or hotter [12]. While there was some novelty involved in designing apparatus to *pressure-couple momentarily* to metal and fused silica specimens up to ~1000°C in air and later, >2000°C in vacuum or inert atmospheres, it was apparent that the basic design of buffers had been solved by other investigators in the fields of ultrasonic delay lines, NDT, or elsewhere. Examples: Frederick [19] in 1948 and Levitt and Martin [33] in 1960 for bulk waves; Bell [4] in 1957 for extensional and later, [4(b, c)] for torsional waves. Hollow buffers were already known in 1962 or earlier [64, 1] but clad buffers [26], bundles of thin rods packed in a rigid pipe [36], spiraled sheet and thin-plate buffers conveying shear waves [41] came later. Many of these are tabulated and diagrammed on page 805 in [40]. A number of these and other pre-1967 high-temperature ultrasonic measuring arrangements and techniques are illustrated in [13].

In this chapter, elastic moduli examples will be used as an introduction to a more general discussion of various uses of acoustic waveguides as industrial sensors, particularly for sensing or at least contributing to the sensing of three industrial process control measurands: flow, liquid level, and temperature. Waveguides can also be used as part of an analytical instrument, responding to the density and/or viscosity of a fluid adjacent the waveguide. Waveguides have also been used to monitor polymerization or degree of cure of epoxy systems in which they are immersed or embedded. Pre-1989 examples are mentioned on pages 93-94 in [38]; a 2003 example is [69].

Consider a pipe as a waveguide. Besides yielding useful signals that provide information about the flow or other characteristics of a fluid within or outside the pipe, a pipe may also act as a waveguide in an undesired manner, yielding noise rather than or in addition to a desired flow signal. Typically the noise includes coherent crosstalk transmitted from transmitter to receiver through the pipe wall. Remedies include electronic and acoustical measures, e.g. quadrature demodulation [62(b)], roughening the surface, adding blocks or damping material to the pipe between transducers. Examples are shown on page 815 in [40]. A differential transducer arrangement due to Jacobson [62(d)] combined electronic and acoustical aspects. Other remedies include damping within a rearward-directed waveguide [54], and improving the isolation between a transducer and its supports [51].

2. Waveguides, Waves, Frequencies, and Measurands

We shall limit most of the discussion to *solid* and *hollow* waveguides, longitudinal and shear *bulk* waves, extensional and torsional waves guided by acoustically-slender elastic waveguides, and flexural waves in cylindrical shells. Rayleigh examples are also included. Rayleigh waves have enjoyed considerable success when implemented in high-frequency SAW (surface acoustic wave) devices. In this chapter, the selected Rayleigh waveguide examples illustrate concepts that are not yet available commercially, to this writer's knowledge. (Lack of commercialization sometimes can be traced to insufficient optimization or no published demonstration relative to an important problem. Concepts not yet commercialized might thereby provide an interesting challenge or opportunity as a student project.)

Guiding of the *polarization* of a shear wave was investigated in [24]. It was found that polarization rotated, following the twist of a twisted rectilinear waveguide to at least 90 degrees. But polarization did not follow the twist for a cylindrical bar (circular cross section). In other words, the twisted crystal lattice does *not* appear to influence polarization. These remarks are based on an experimental procedure having a limit of angular resolution of ±5 degrees. Details: http://www.asnt.org/publications/Materialseval/solution/jan04solution/jan04sol.htm.

Because ultrasound has been used so extensively in flow measurements, we include pipes as examples of a hollow waveguide (cylindrical shell). Historically, the use of a hollow waveguide as part of an acoustic measuring system dates back at least to Colladon, 1826. Colladon used a tube as a buffer to avoid having to immerse his head in Lake Geneva when he made his classic measurement of the speed of sound in water. This experiment is reported in Lindsay's translation, pages 194-201 in [34]. Other early applications of waveguides, besides natural ones such as in our ears or voice passageways, are recognized in musical instruments; animal horns; since 1816 in stethoscopes (pages 166-171 in [34]; see also, http://www.antiquemed.com/monaural_stethoscope.htm). Nearly fifty years after Colladon's 1826 sound speed measurements in Lake Geneva, i.e., *ca.* 1873, organ pipes were used in early demonstrations of acoustical thermometry by Mayer [50]. In 1962, Apfel [1] used a tubular lead-in and He-filled sensor cavity to measure T at high temperature. A number of early (pre-1967) high-temperature ultrasonic applications are illustrated in a chapter by Panametrics' late co-founder Edmund H. Carnevale [13]. At Avco, pre-1960, Carnevale himself used fused silica buffers in a momentary contact arrangement to measure plasma temperatures to around 10^4 kelvins [11]. At Panametrics, starting in 1962,

momentary contact was extended to measuring elastic moduli of solids at elevated temperature using *metal* buffers (funded in part by ONR) [12], and by the mid-1960s to measuring transport properties of gases and plasmas, again using fused silica buffers (funded in part by NASA) [14].

In this chapter, frequencies generally shall be limited to 0.5 to 5 MHz for bulk waves and 50 to 500 kHz for extensional and torsional waves. Our selection of frequencies puts many interesting SAW or plate wave sensors and nanoscale devices beyond the chapter's scope. Flexural waves are considered, with frequencies down to the *audible* range. Measurands shall be those mentioned in the introduction. Fluids whose density, flow velocity or other characteristic is sought may surround a sensor waveguide or be surrounded by a tubular sensor or conduit. When the medium acts as its own sensor, the waveguides convey the interrogating acoustic wave to and from the medium. In passive arrangements [20, 69(c)] the waveguide may convey energy from the sensed region to one or more nearby or remote transducers. Flowmeter transducers may be outside the pressure boundary, e.g. clamped on a pipe. In a downhole flow tool the transducers may be *inside* the pressure boundary.

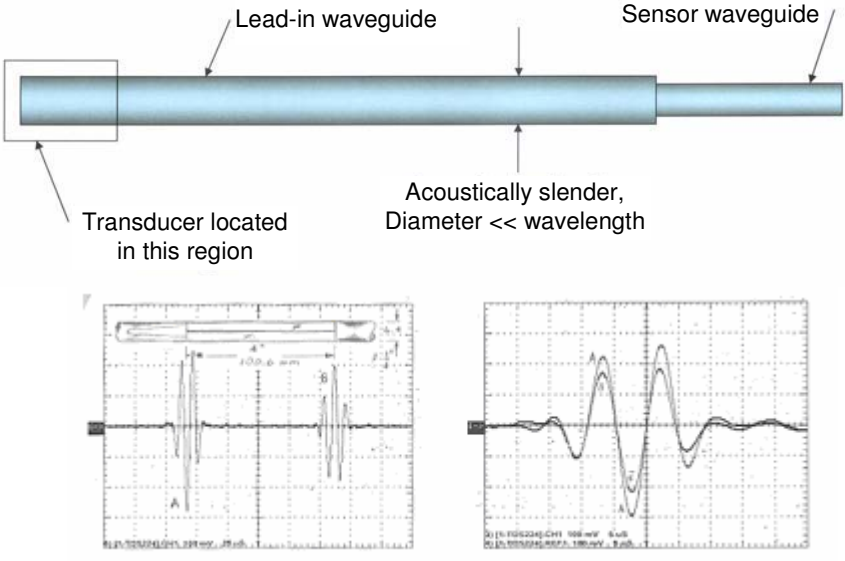


Figure 1. *Top:* Pulse-echo waveguide system for measuring elastic moduli in acoustically-slender specimens. The specimen comprises the sensor waveguide. © 2006 Lynnworth Technical Services, reproduced with permission. *Bottom:* impedance-matching by a taper [44] between round lead-in waveguide and diamond sensor that is terminated by a round lead-out waveguide. Torsional echoes are of similar shape but of opposite polarity. Oscillogram at *bottom right* shows the similarity of echo shapes after echo B was reversed in polarity and superimposed on echo A [55(c)]. © 2004 IEEE, reproduced with permission.

2.1. *Waveguide is the Sensor – Elastic Moduli Example Leads to Other Measurands*

Consider a waveguide system comprised of a long lead-in wire attached to a much shorter segment of specimen wire of circular cross section and of density ρ. Suppose all the diameters to be small compared to wavelength, and the specimen to be isotropic. Assume a pulse-echo transducer at the left, and a pulser/receiver/intervalometer system connected to it (Figure 1). The wave impedance of the sensor is sufficiently different from that of the lead-in so echoes are generated at their interface.

If the lead-in and sensor have the same *characteristic* acoustic impedance, then absent attenuation, echoes of equal amplitude for extensional waves are generated at the interface and free end if the larger-to-smaller diameter ratio is 2:1. Echoes of equal amplitude are generated for torsional waves if the larger-to-smaller diameter ratio is $\sqrt{2}$:1. If the sensor is between circular segments, matching between lead-in and sensor can be accomplished by introducing a gradual transition such as a cone or, as shown in Figure 1 (*bottom left*), a taper similar to that of a bladed screwdriver. The result illustrated by the oscillograms is that echoes A and B are of nearly equal *magnitude*; their polarities are reversed because of the sequence of impedance steps [44, 55(c)].

Let the length and round-trip transit time in a specimen of circular cross section be measured. Then for the extensional and torsional modes, the corresponding elastic moduli are, respectively, $E = \rho(c_{EXT})^2$ and $G = \rho(c_{TORS})^2$ where the two sound speeds are determined using one or multiple echoes, and one specimen length or different lengths if the specimen can be cut shorter and shorter. The equation for E is simpler than if bulk waves were used. There is no Poisson's ratio term because the waveguide is thin enough to not be stiffened by surrounding elastic material. The equation for G is the same as for bulk shear waves. However, if the specimen is not round in cross section, then a shape factor $K_{SH} < 1$ must be included in the equation for G. Poisson's ratio σ is determined by the usual formula, if the specimen is isotropic: $\sigma = (E/2G) - 1$.

In 1957 Bell [4] employed magnetostriction to generate and detect extensional waves in heated specimens. As applications emerged in thermometry, the specimen was made of a refractory metal or nonmetal, becoming a temperature sensor calibrated by heating in appropriate atmospheres to >2000°C. The magnetostrictive transducer method was retained for over four decades, although some investigators preferred piezoelectric transducers butted against the end of a thin rod to generate a compressional wave and detect the echoes. By 2004, however, a piezoelectric method was demonstrated wherein NDT shear

transducers were dry pressure coupled to the cylindrical surface at one end of the lead-in, leading first to torsional waves and later, to extensional and flexural waves too [55].

In the pioneering hands of Bell, Thorne and their colleagues in the UK, by 1960 the technology had migrated within the UK across technical fields, from metallurgical studies to thermometry. In July 1964, Bell's instrument technology began its migration across the Atlantic to the USA on board the *Queen Elizabeth*. After redesign for *automatic* measurements of time, and introduction of echo select and blanking delay windows, the Panatherm® 5010 intervalometer was commercialized by Panametrics in the latter half of the 1960s and sold in the USA, Europe and Asia for moduli and thermometry applications. It was also used in NASA, ONR and AEC high-temperature R&D programs at Panametrics.

In contrast to more traditional AE (acoustic emission) studies e.g. as reported in [23], at Panametrics using waveguides, AE applications [17] took two forms: transmission and reception. In the transmission form, the waveguide's sensor shown in Figure 1 was replaced with a hard needle and became a point source of reproducible 100-kHz impulses that could be pressure coupled to any accessible point on a structure. These impulses simulated AE signals for calibrating AE triangulation equipment. In the reception form, the waveguide was pressure coupled or, for a more permanent installation, metallurgically bonded to a structure. The structure in at least one application was at high temperature. The waveguide conveyed actual AE signals to a remote receiving transducer. (Apart from AE *waveguide* efforts, Fowler, who headed the NDT Division of Panametrics for many years, also developed a 100-kHz piezoelectric AE transducer which by ~1980 was impedance-matched to air and sealed for use with hazardous or aggressive gases. In various forms, these sealed, impedance-matched "AE" transducers continue to be used in ultrasonic flowmeters for flare gas [49(b)], natural gas, etc. [46]. (For these transducers, gas flowmeter applications since ~1984 number in the thousands.) This illustrates a successful ultrasonic technology transfer paradigm, NDT to process control. However, there were temperature limits to this AE transducer. By ~1995 these were solved using a different transducer combined with a buffer waveguide made of a bundle of thin wires, to be described subsequently (Figure 5, *top left*). Elsewhere, air flow containing particulates was measured with a probe in the flow. The sensing portion hit by the particles guided "AE" impact noise to an external pickup.

Returning to the individual wire waveguide, a typical thermometry application in the 1960s was measuring the temperature (T) profile along a fuel pin centerline, which was accomplished in several laboratories but did not

materialize into a commercial product satisfying a large market. Timing between 100-kHz broadband echoes to ±100 ns was achieved using a centerline method developed by Bernard M. Gordon and colleagues around 1967 and was used for approximately twenty years in specialized Panatherm applications. An early Panatherm 5010 instrument was replaced in the late 1980s at Rice University by a modified Panametrics flowmeter employing correlation for timing [62(c)], leading to several high-temperature moduli publications, e.g. [62(e)].

Returning to temperature: Figure 2 shows a multizone waveguide sensor with echoes generated at discontinuities. Researchers at Sandia National Laboratory [10] made thoriated tungsten waveguides having five zones of length ~10 mm each, and used them in thoria sheaths to measure T to about 2750°C. In 1979 researchers at B&W (summarized in Ch. 5 in [38]) used blackened Ti ribbon, ¼-mm thick, for fast-response T sensing as part of a CHF (critical heat flux) experiment. This thin rectangular cross section Ti sensor had a precursor: ten years earlier, Fowler [18] measured Young's modulus as a function of direction in rectangular cross section Ti specimens cut from a rolled sheet.

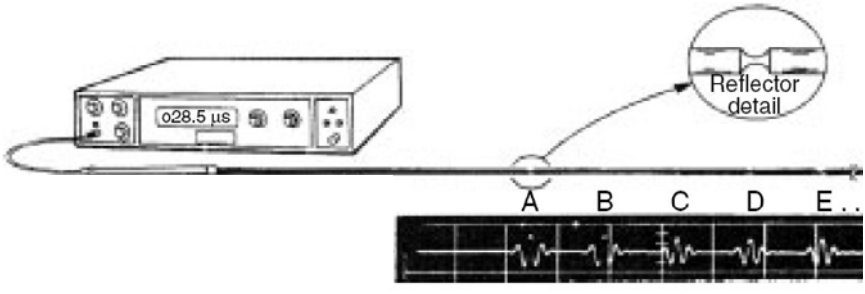


Figure 2 (a). Echoes A, B, C, D, E generated at notches. At each notch the magnitude of the sound pressure reflection coefficient |R| << 1. Panatherm® intervalometer measured time between centerlines of 100-kHz broadband echoes to 100 ns based on a timing method due to B. Gordon and colleagues *ca*. 1967. Multizone T-profiling waveguide & instrumentation details: [10, 46(d), 56(b)].

Figure 2 (b). Torsional T sensor design from a 2001 undergraduate student project at Tufts University. Students [9] started with a 4.76-mm (3/16″) diameter cold rolled steel rod, milled one end to create a semicircular cross section of length 25.4 mm, and measured the torsional transit time between the shoulder echo and the end echo. (A larger-diameter notched bar was used in 1947 for bulk guided wave determinations of elastic properties [19].) In [9], torsional calibration was conducted from room temperature to 650°C. The sensor was first placed in a SS (stainless steel) sheath. The sheath was heated with a propane torch. The reference T was a thermocouple in contact with the sensor. Among results obtained in [9]: (1) $K_{SH} \approx 0.77$; (2) From ~140 to 440°C, torsional transit time increased linearly as T increased, ~5 ns/°C for one round trip in the sensor.

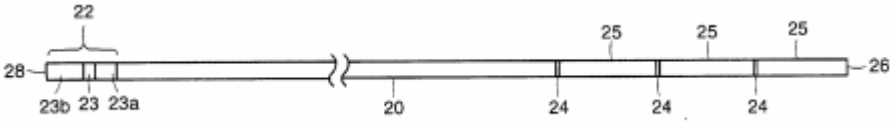


Figure 2 (c). Grain-stabilized refractory waveguide e.g. polycrystalline alumina, zones defined by circumferential grooves, waves generated by a piezoelectric or magnetostrictive transducer assembly, in 2003 thermometer design [21].

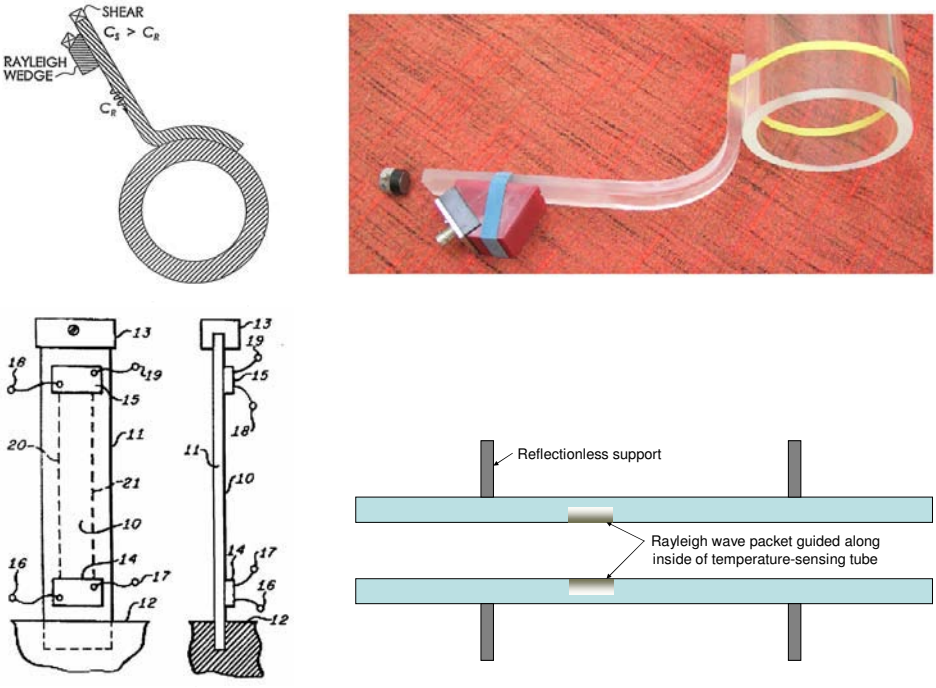


Figure 2 (d). Rayleigh waveguides. *Top left*, diagram from a 2000 patent proposes Rayleigh wave clamp-on buffer for pipes [41(b)]. *Top right*, photo of mock-up illustrates concept similar to that diagrammed at *top left*. The waveguide, of rectangular cross section and about 9 mm thick, has a short straight portion (the "shoe") and a longer straight portion (the "standoff") connected by a circular arc. In this mock-up the waveguide is positioned against transparent pipe of diameter ~400-mm. Assuming a fused silica waveguide, an NDT-style angle beam transducer could launch Rayleigh wave along one surface of that waveguide at the velocity c_R = 3410 m/s, taking the c_R value for fused quartz in http://www.kayelaby.npl.co.uk/general_physics/2_4/2_4_1.html. Smaller shear wave contact transducer is represented near end of waveguide, corresponding to diagram, *top left*. Elastic bands represent mechanical clamping. *Bottom left,* two views of Rayleigh sensor for measuring strain, displacement or acceleration [68]. *Bottom right*: Tubular Rayleigh sensor concept uses Rayleigh waves *inside* a tube of wall thickness $w \geq \lambda_R$ where λ_R is the Rayleigh wavelength. Thick wall avoids generating echoes from external supports, avoiding, in principle, one of the problems of introducing extensional or torsional waves through a pressure boundary, especially if a welded penetration seal is required. Potential application: thermometry.

Noncircular cross sections make it awkward to determine G using torsional waves because one needs to calculate a shape factor K_{SH}. But by 1977 a rectangular shape was found effective for determining with torsional waves the density ρ_L of a liquid surrounding the sensor, and also, the liquid level H. The cross section was optimized for ρ_L by Kim in 1989. He found a diamond (♦) or cusped diamond to be preferred shapes if the objective is measuring ρ_L or H. Kim also found a way to determine liquid viscosity η using a circular cross section sensor in series with the diamond sensor. It turns out that the diamond or cusped diamond shape is potentially useful for T too because, for aspect ratio = 3, $K_{SH} \approx 1/2$. This means, torsional transit time is twice that of a circular cross section of the same sensor material and length, implying twice the T sensitivity. By 2004 it was demonstrated at GE Infrastructure Sensing that the diamond shape can also be used as a shedder of vortices in flowing air or water. If the flow velocity V is bidirectional, a symmetrical shedder that looks the same to flow coming from either end of the pipe offers a bidirectionality advantage over the more usual asymmetric bluff body shapes. However, in most V applications anticipated at this time, where flow is not bidirectional, a more conventional bluff shedder shape (e.g. bluff polygon, Figure 7(a) or 11) will probably be preferred [45].

This brief history of slender waveguides excited in extensional or torsional mode has taken us on a journey starting with E, G and σ to T, H, V, ρ_L and η. In subsequent sections dealing with some of these topics, more technical details are included, e.g. wave shapes for specific waveguides; elaboration of "flow" to include velocity V or mass flowrate M_F.

Dr. Sidney Lees of the Forsyth Dental Center, Boston, remarked some years ago that a sensor that is good for everything is good for nothing. Among the designer's challenges is the task of designing and instrumenting the waveguide sensor to respond primarily to one measurand and not be influenced by interfering variables. Examples of analogous problems in designing quartz resonators are given in [32]. An example in waveguide densitometry: to reduce the influence of T over a specified T range, one solution is to make the waveguide of an alloy that can be heat treated to achieve a controllable near-zero thermoelastic coefficient, e.g. NI-SPAN-C® alloy 902, available from Alloys International, Special Metals Corporation.

2.2. *Intrusive vs Nonintrusive vs Noninvasive Sensor*

Ultrasonic equipment manufacturers often point to the noninvasive aspect of their products. Does a waveguide have to be invasive? No, it can be the wall of a vessel (such as an aboveground fuel storage tank). It can be a portion of a pipe.

We may take the term noninvasive to strictly mean, no penetration whatsoever of the pressure boundary of the vessel, conduit or structure bounding or containing a fluid. In practice, wall surfaces are often rubbed clean or cleared of scale or paint to accommodate coupling. In some cases a curved surface will be ground flat so that a flat-faced member can couple to it efficiently.

Invasive implies a significant penetration, e.g. boring a hole through a wall so the ultrasonic transducer or buffer attached to it can radiate directly into the fluid. Intrusive means invasive plus further penetration into the fluid, e.g. into or all the way across a flowing fluid. In order to obtain a representative measure of T, a T sensor may have to be installed a distance of 300 mm into a large-diameter pipe's interior. The previously-mentioned vortex shedder strut is typically positioned across a diameter. However, insertion flowmeters are also installed only part way into a flowing stream.

An example of a waveguide whose last inch (25.4-mm) has the cusped diamond cross section and intended for use intrusively, is shown in Figure 3.

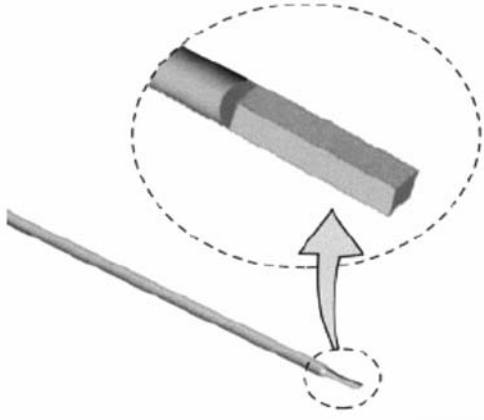


Figure 3. Lead-in waveguide of length ~300 mm and diameter 6.4 mm conveys torsional waves at a frequency nominally 100 kHz towards the short sensor at the end opposite the transducers (not shown). The sensor is of cusped diamond cross section, aspect ratio three, and has been fabricated of metals such as Ti or 316SS.

2.3. *Waveguide Contributes to the Sensing as a Thermal Buffer*

Here we consider using a waveguide not primarily as a sensor but rather as a tool to convey ultrasound to and from a sensor that is not necessarily a waveguide. A noninvasive example would be the use of threaded rods as used by Araki and Matsunaga [2] to measure flow (Figure 4).

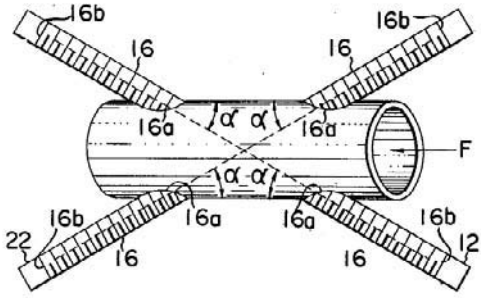


Figure 4. Use of threaded rods as thermal buffers in a flowmeter [2].

Another example is the use of relatively thin but rigid plate shaped like a small hockey stick, and also used in clamp-on flowmeters to achieve a noninvasive interrogation (Figure 5, *top right*). This buffer, often referred to as the "OKS" waveguide, includes reference reflectors which can be used to sense T in a shoe portion of the buffer. If that part of the system is well insulated, then to the extent the shoe of the buffer, which is in thermal contact with the hot pipe, comes to thermal equilibrium with the pipe, T in the shoe is a measure of T in the pipe.

Coupling between the shoe and the pipe is usually accomplished by a viscous liquid, gel or grease at room temperature, and by high pressure aided by soft oxidation-resistant metal foil at cryogenic or high temperature, e.g. sandwich comprised of zinc foil wrapped with gold. Another coupling alternative is welding. Welding provides permanent coupling. Welding of the OKS buffer to a steel shell has been used in AECL (Atomic Energy of Canada Limited) downcomer studies in nuclear power plants [25], where the steam generator outer steel shell, nominally 75-mm thick, was much thicker than most pipe walls. Where welding is permitted, the buffer facilitates permanent coupling. But welding of the buffer to a conduit would not be permitted in, for example, chemical or nuclear plant situations with unresolved safety concerns regarding one or more of: pipe wall thickness, metallurgical changes during welding, or combustible fluid inside or outside the pipe.

AECL's ultrasonic downcomer contrapropagation flow technology, utilizing OKS buffers and including welding procedures (provided by B&W), and also dealing with steam carryunder (a two-phase flow situation) has been applied to measuring flow velocities in steam generators at Ontario Power Generation's Darlington and Bruce Power's Bruce A and B Nuclear Generating Stations. At Bruce B, equipment was recently installed on steam generators at three reactors, in response to long-term boiler degradation due to Flow-Accelerated Corrosion

(FAC) of internal components. Any associated loss of thermal-hydraulic performance is expected to affect the downcomer flow rates. It is intended that measured velocities will also be used as part of an information database required for reactor uprating (Figure 5, *bottom right*) and to provide a general long-term steam-generator "health monitor." During a recent two-year period, 16 steam generators at the Bruce B Nuclear Generating Station have been instrumented in this manner, according to [25].

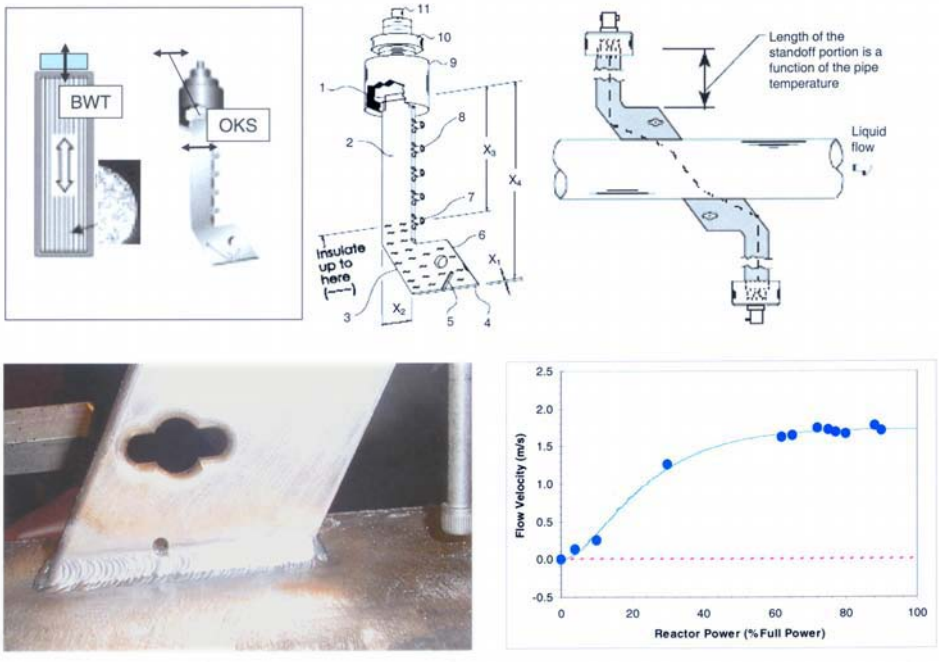


Figure 5. *Top box*: *Left,* Fiberacoustic bundle packed inside a welded metal enclosure. Main use: wetted buffer to measure flow of hot gases. *Right:* Shear buffer. Double-headed arrows indicate principal particle displacements. *Top middle & right*: Hockey-stick-shaped thin blade "OKS" buffer waveguide conveys transverse shear waves at 0.5 to 2 MHz towards the wall of a hot pipe. Main use: flow measurement of hot liquids, up to 400°C [36, 41, 43]. *Bottom*: AECL (Atomic Energy of Canada Limited) results [25], courtesy of V. P. Janzen. *Bottom left*: OKS welded to downcomer boiler, for use to 270°C. *Bottom right*: Example of the measured downcomer water flow velocity (m/s) as a function of percent reactor full power for a CANDU steam generator.

Next, consider invasive buffers such as a bundle of hundreds of thin rods encapsulated within a pipe which may be 15 to 30 mm diameter and 300 mm long, welded closed at both ends (Figure 5, *top left*) [36, 43]. Thin rods are free of dispersion, even when tightly packed. While noninvasive solutions are preferred if they yield the required accuracy, when the fluid is a gas and accuracy

better than 1% is required, it is difficult to meet this requirements with external i.e. clamp-on transducers. Accordingly, in many cases, a wetted transducer (actually a wetted buffer) is acceptable and may be preferred or specified by the user. One can think of the bundle buffer solution as superposition in parallel of hundreds of acoustically-slender extensional-wave lead-in rods of the type depicted in Figure 1. The piezoelectric transducer is encapsulated and the transducer assembly, an explosion-proof design, is removably coupled to the end of the bundle.

The last example in this section illustrates the use of the wall of an aboveground storage tank as the sensor to determine, from flexural wave velocity in the wall, the height H of liquid hydrocarbon fuel height inside the tank. Fuels were gasoline, as well as oil. The transducers used in a 1993 study [35] were magnetostrictive, in fact quarterwave waveguides, pressure-coupled normal to the tank wall to generate and detect asymmetric (A_o) plate waves. In some respects the 1993 magnetostrictive Ni transducers resembled those used thirty-six years earlier by Bell [4] in his metallurgical studies. In [35] by making the frequency sufficiently low, in the audio range, the phase velocity of the A_o plate wave was made slower than the longitudinal wave in the adjacent liquid. The low-frequency flexural wave therefore does not radiate or leak into the liquid. This keeps attenuation small. Liquid presence is detected from transit time. Transit time increases because of mass loading. A more rigorous explanation is given in [15].

By experimenting with the transducers on different waveguides, it was found in [35(b)] that the phase velocity in plastic pipe increased or decreased as a function of the flow direction relative to propagation direction, while the average phase velocity was a function of the density of the fluid inside the plastic pipe. This offers a noninvasive flexural wave method of determining mass flowrate M_F. Multimode analyses due to Sato et al. appeared in 2006 [35(d)].

3. Industrial Process Control Measurands (T, H, V) – Additional Details

The following sections elaborate on problems and solutions related to acoustic waveguides as part of sensing systems, particularly for three of the principal measurands in industrial process control: T, H and V. The fourth important measurand, pressure (P), receives minor attention in this chapter as its measurement apparently has not yet successfully utilized acoustic *waveguide* technology to any significant extent. Quartz resonators, however, are widely used as P sensors, and one may argue that they are indeed waveguides, even if short.

3.1. *Temperature T*

As indicated above, the elastic moduli E and G, which depend on T, are proportional to the square of sound speed of extensional and torsional waves, respectively. This means a database of moduli vs T can be used to estimate the T-dependence of extensional and torsional waves even if one neglects the small change in waveguide density with T. This simple basis for waveguides used as T sensors is partly offset by complications stemming from other aspects of propagation such as attenuation (which tends to increase rapidly above the recrystallization temperature, about half the absolute melting point for a number of refractory polycrystalline metals [56]). One remedy is to use a ceramic crystal with high melting point, e.g. sapphire, or a polycrystalline ceramic such as grain stabilized alumina [21]. Considerable T sensitivity is sacrificed (compared to a metal sensor) but lower attenuation even for torsional (↻) waves allows use of higher frequency, e.g. ≥1 MHz, and the sub-ns timing possible at frequencies ≥1 MHz supports this choice. Brittleness, however, is more of a problem with ceramic sensors than with metal. Torsional waves allow use of smaller notches as fiduciary marks on the waveguide, compared to extensional waves. If a metal waveguide is used, another potential interference is inadvertent welding of the waveguide to its sheath if T is close to the melting point (diffusion bonding). With oscillatory rotary motion, that source of bonding between sensor and sheath might be eliminated. Even with nonmetal (ThO_2) sheaths, problems can occur at high temperature [10]. Other sources of attenuation are reflection losses at seals and supports. These and other problems motivated solutions as detailed in pre-1989 references on waveguide thermometers summarized in Chapter 5 in [38].

3.2. *Liquid Level H*

There are numerous waveguide approaches to measuring liquid level H. In [38] these were categorized according to access from the top, side or bottom of the vessel. Noninvasive waveguide solutions use the vessel wall as the sensor. An example using SV vertically-polarized shear waves circumferentially guided by the wall and leaking off into the adjacent liquid when liquid is present at their level is illustrated on page 380 in [42]. Invasive probes that guide shear waves vertically downward or upward can measure liquid presence by arranging for some of the guided energy to be reflected normal to the axis of one probe, across a gap, towards the other probe, shown on page 37 of [38]. Sometimes the vessel wall or another structure can reflect the interrogating energy back into the first probe, in which case one probe suffices.

A hollow waveguide with multiple reference reflectors can be installed from the top of a vessel to measure H according to [67], Figure 6 *left.* A variant of this idea [60] was installed from the top of buried underground gasoline storage tanks but the transducer radiated upwards through gasoline to sense the top liquid level (gasoline-air interface). The transducer also radiated downward to detect the gasoline-water interface, sensing condensed water lying at the bottom.

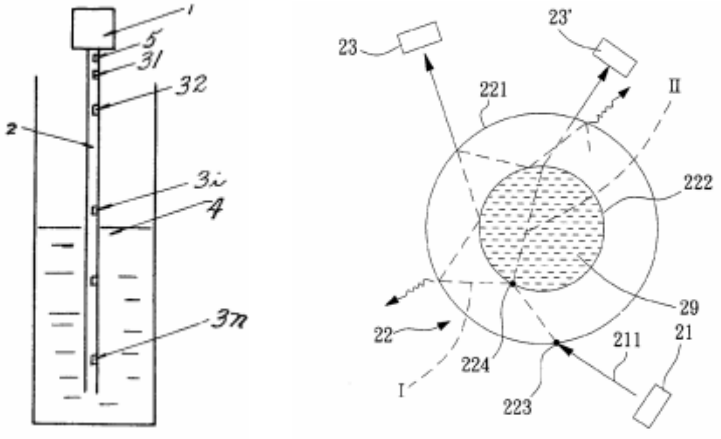


Figure 6. *Left*: Hollow waveguide (2) in a 1968 design contains reference reflectors to generate timing echoes [67]. *Right*: Liquid (29) presence and level detection utilizes waveguiding property of a cylindrical shell (221) in 2006 NDT leakage application [58]. See also, pages 379-380 in [42].

An invasive solid torsional waveguide of rectangular cross section, developed initially to measure liquid density ρ_L, was found to respond by a mass loading effect to the product of ρ_L times the length of the immersed sensor. This type of H sensor can be T-compensated using extensional waves in the same sensor, demonstrated in [52]. The same cross section that is optimum for sensing ρ_L is also optimum for sensing H, in terms of sensitivity. Therefore, in the early 1990s, corrosion-resistant material drawn into the diamond shape was used in several high-temperature H-sensing programs [29], although the rectangular cross section still found use [63]. For a given material, for aspect ratio 3, the torsional sensitivity to density using the diamond is three times that of the rectangle.

To illustrate the sensitivity S, where S is the fractional increase in torsional transit time on immersing the sensor from air into water, compare the transit times shown in the oscillograms of Figure 8(a, b). These oscillograms show that for a 316SS "flat" diamond sensor, aspect ratio 3, S = 15%. In similar tests using a cusped diamond, S = 18%. For a similar SS waveguide but of *rectangular* cross

section, S was found to be 5% (page 49 in [38]. S depends on sensor density ρ_S and details of the cross section.

The main problem experienced in the early 1990s with the SS (stainless steel) diamond waveguide of aspect ratio 3 but probably not limited to that shape, was a tendency to become fouled by microbubbles and bubbles when used in ordinary water and especially water being heated that was not previously de-aerated [29]. Polishing and/or coating with Teflon did not prevent nucleation of bubbles. Accordingly, work on torsional waveguide sensing of H in superheated water apparently stopped for a dozen years, until recently resumed e.g. at ORNL (Oak Ridge National Laboratory) for application within integral primary system reactor (IPSR) in-vessel environments. See Figure 7 (b-d) [31].

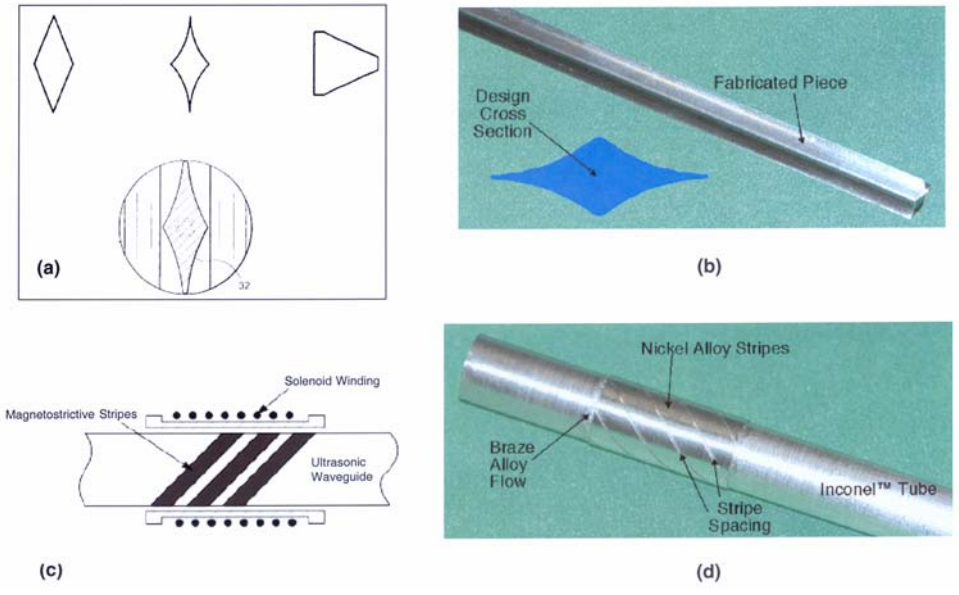


Figure 7. Evolution of diamond and polygonal cross sections for torsional sensing of liquid density ρ or liquid level H, or V and potentially M_F when used as a vortex shedder, 1989-2006. (a) "Flat" diamond and cusped diamond, after Jin O. Kim [28]. When made of 316SS, density $\rho_{S.316SS} \approx 8$ g/cm^3, aspect ratio = 3, these provide sensitivity S = 15% and 18%, respectively [45, 55(c)]. *Top right*: conventional bluff polygon, estimated S = 4% in 316SS [45]. *Bottom*: cusped diamond with two tips radiused to maintain cross sectional shape in flowing fluids [44]. (b) Cusped diamond waveguide sensor currently under development at ORNL. Material: Inconel™ 600, aspect ratio 2.7, density $\rho_{S.Inc600} \approx 8.5$ g/cm^3 [31]; (c) 2005 concept for oriented magnetostrictive generation of torsional waves [30, 57], redrawn for clarity by D. E. Holcomb [31]; (d) 2006 implementation at ORNL [31].

In a mass flowmeter project that seemingly was unrelated to measuring H, it was found in 2004 that in *flowing* water, bubbles did not nucleate on the waveguide to any measurable degree, implying that if adequate flow were maintained, the bubbles, if they form at all, would be swept away [45].

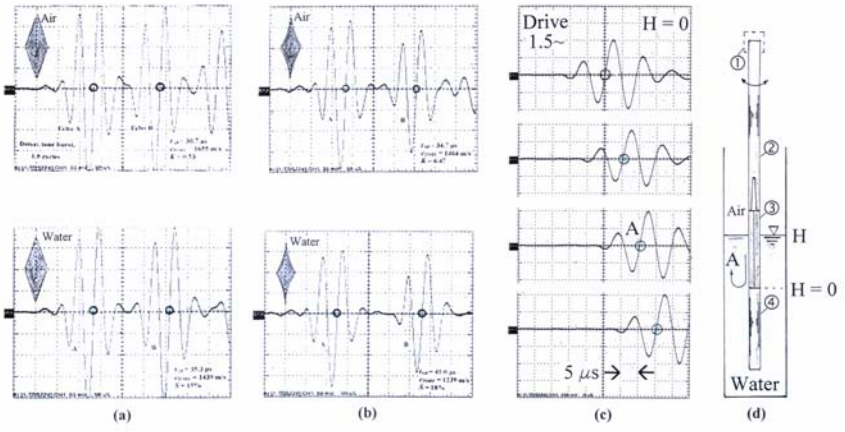


Figure 8. Torsional transit time in flat and cusped diamond sensors, 316SS, in air and in water [55(c)]. Ratio of transit times (water to air) yields sensitivity S, 15% in (a), 18% in (b). Sensor is represented in Figure 3. Each progressive ~5μs time-shift in (c) corresponds to ~34-mm increased depth of immersion H in (d). Note preservation of 100-kHz broadband waveforms. These sensors operate primarily on a mass loading principle [28, 29, 63], but that principle might not fully account for the sensor's performance, according to [15]. Legend: ① Torsional transducer [55]. ② Lead-in, 316SS, ϕ9.5mm × ~300-mm long. ③ Sensor, 101.6-mm long. ④ Lead-out, ~100-mm long. Portions © 2004 IEEE, reproduced with permission.

We conclude this section with an illustration (Figure 9) of a noninvasive solution corresponding to a flexural guided wave high level alarm system installed at over ninety points of liquid level measurement in the early 1990s [35]. The requirement was to measure from outside the tank, the liquid level, to assure that it did not exceed levels which in the trade are termed high and high-high. By using a low-frequency A_o wave, the flexural transit time was found to increase by about 10 μs when the liquid level reached the centerline of the path between a pair of horizontally-spaced transducers. The distance between transducers was about 250 mm.

In later applications, problems were encountered where the tank was internally coated with residue. Simulation experiments by Peter J. Stein and colleagues at SSI (Scientific Solutions, Inc., Nashua NH) with peanut butter of

varying thickness showed that frequencies even below 1 kHz might be needed to sense beyond the residue and respond to actual changes in the liquid level of interest [65].

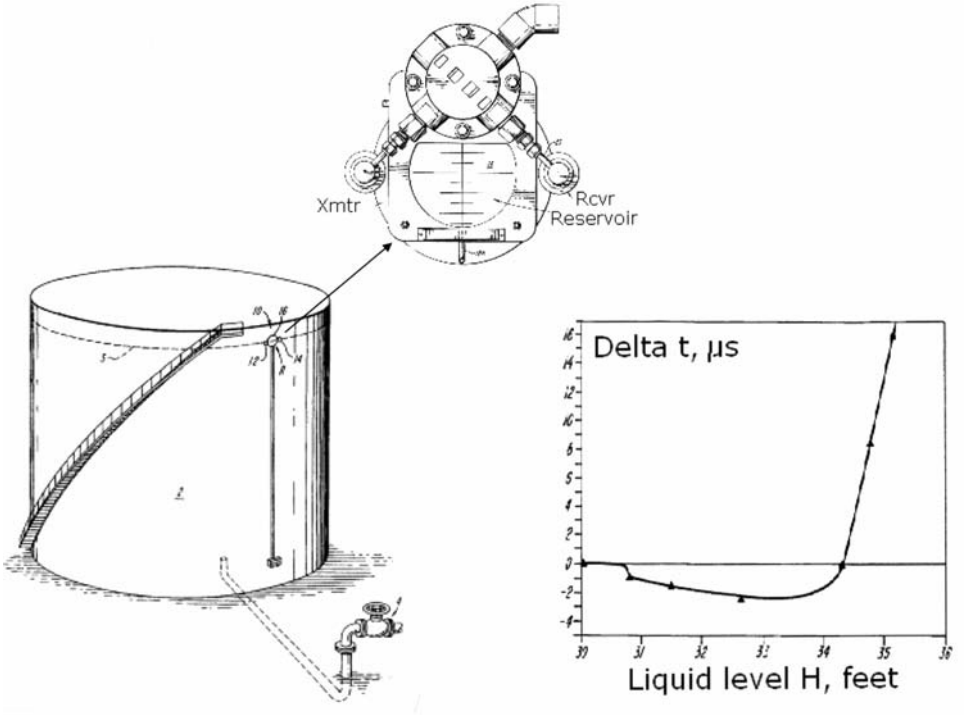


Figure 9. Noninvasive high level alarm system measures H. System uses flexural wave transit time to sense the mass loading of liquid inside the tank. Liquid, if present, slows down the velocity of flexural waves. These waves are guided by the steel wall. They propagate across a ~250-mm-long horizontal path between transducers that are coupled or attached externally to that wall. Graph shows that at a liquid level H ≈ 10.7 m (~35 ft), resolving transit time to 1 µs resolves H to about 25 mm. System can be checked by temporarily filling an external reservoir between the two transducers with water or other allowable liquid which mass-loads the *outside* of the vessel wall as closely as possible to the way the actual liquid inside the vessel would load the wall if the liquid reached the same level.

3.3. *Flow Velocity V and Related Parameters*

Noninvasive measurements of V or terms derived from V such as volumetric flowrate Q and mass flowrate M_F in principle could be achieved using a low-frequency method described in [59] in which the conduit and its contents

comprise a waveguiding system. Another noninvasive waveguide solution is described in [35] based on flexural waves. As far as is known to the author, however, neither of these waveguide flowmeter methods has been commercialized. Figure 10 illustrates these and other flowmeters that utilize the conduit and its contents as a waveguide in which flow is measured using an interrogating wave transmitted unidirectionally as in (a) or bidirectionally (b-d). In the early 2000s method (d) found applications in high-purity semiconductor liquid flowmetering. A passive noninvasive audible-frequency method from CiDRA uses acoustic spatial array signal processing to obtain V from the difference in speed of noise propagating with vs against the flow. Also, the sound speed is analyzed as a function of frequency to reveal density, void fraction or other fluid characteristics. The noise is sensed as dynamic (ac) pressure by means of external pressure sensors. These may be strain gages [20].

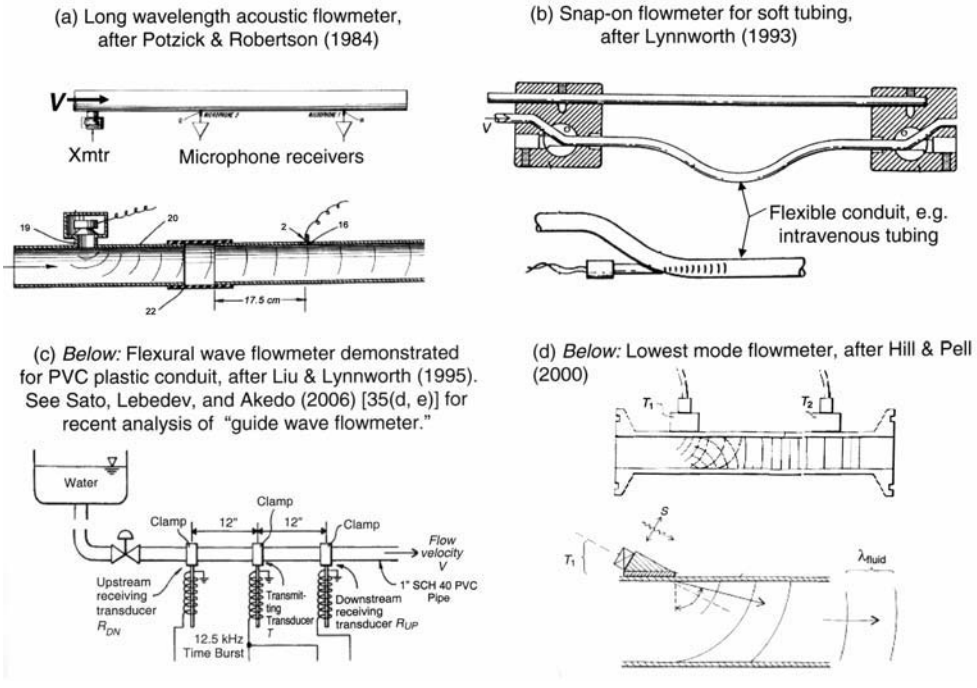


Figure 10. Noninvasive axial-path flowmeter inventions in which the conduit and its fluid contents comprise the waveguide. Related diagrams: http://www.uspto.gov | Patents | Search | Patent Number Search | type the US patent number.

Noninvasive buffer waveguides have been used in many high temperature flow applications to convey ultrasound from a transducer at a relatively comfortable

temperature to the surface of a hot pipe, or in some cases, to a pipe at a cryogenic temperature. The shear wave buffer shaped like a small hockey stick ("OKS") was mentioned in the introduction. The bundle, also mentioned in the introduction for wetted applications (BWT® bundle waveguide technology transducer) could also be applied as a clamp-on and may eventually find noninvasive applications in tag flowmetering [62]; page 537 in [7].

A portion of a tag flowmeter (GE Infrastructure Sensing, Model CTF868) was combined with a waveguide used as a vortex shedding strut in laboratory experiments reported in [45]. The frequency at which vortices are shed is proportional to V. In principle, a torsional waveguide of noncircular cross section could serve as a ρ_L liquid density sensor as well as a shedder [44] in which case a mass flowmeter might be achieved in a compact arrangement. So far, only the V-sensing portion of this potential M_F meter has been demonstrated under flow. The experimental arrangement is represented in Figure 11. The shedder is intrusive. The sensing transducers are external to the pressure boundary, clamped onto the pipe. The pipe was PVC plastic for air tests (data, Figure 12). When the fluid was water, the pipe was steel.

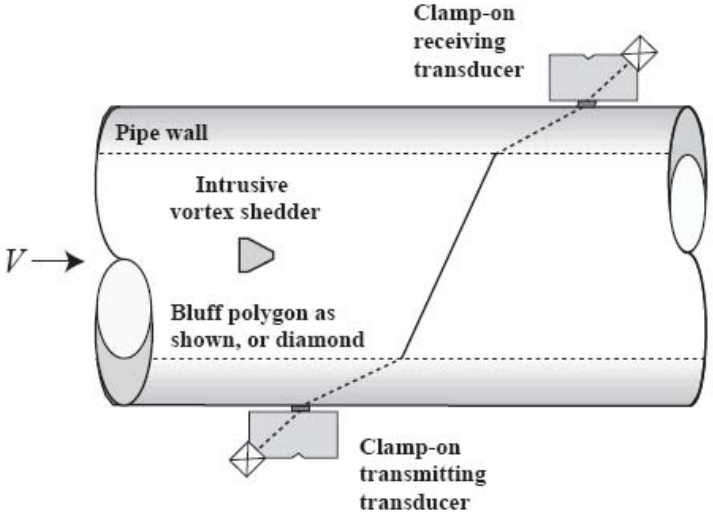


Figure 11. Schematic of lab experiment using as the intrusive shedder, a bluff body shaped as a polygon (as shown) or as a diamond [shown in Figures 3 and 7(a)]. Modulation by the shed vortices of the cw 500-kHz beam transmitted unidirectionally across the shedder's wake to the receiving transducer is sensed by one channel of a commercial tag flowmeter (GE Model CTF868), not shown. The modulation frequency is proportional to V [45].

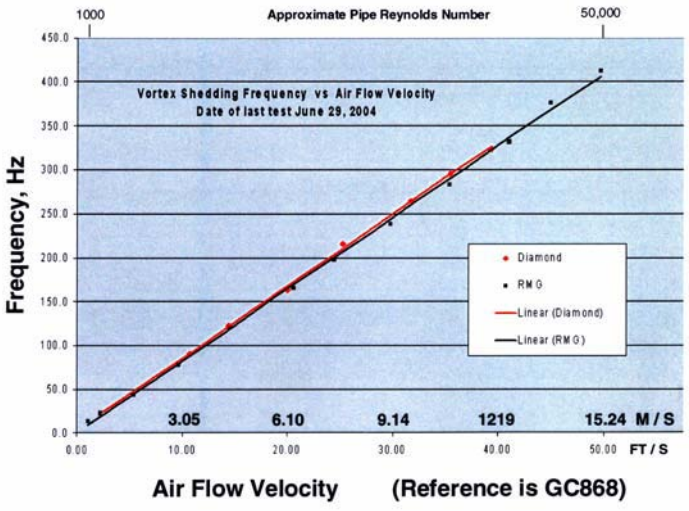


Figure 12. Vortex shedding frequency vs air flow velocity, for pipe *Re* (Reynolds number) ≈ 1000 to 50,000. The shedders were either a diamond or a conventional bluff polygon. Data were obtained in the lab experiment depicted in Figure 11. Air was drawn through a 50-mm ID plastic pipe [45]. © 2006 IEEE, reproduced with permission.

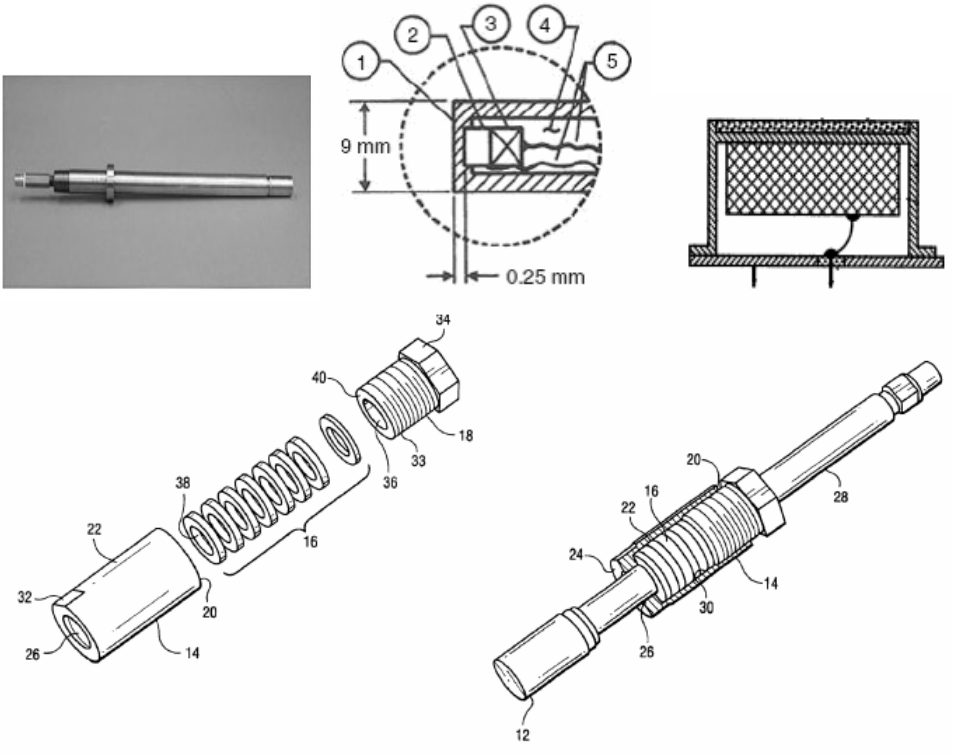
Next, we show a waveguide comprising the rearward extension of a transducer, containing means to abate noise, due to Nguyen [54]. Noise is a problem in virtually all measurement systems where high accuracy is required. Many patents show inventors' attempts to mitigate acoustic noise traveling unintentionally around the pipe. Some are shown on page 815 in [40]. The pipe itself may be viewed as one of the most troubling examples of an acoustic waveguide used in an acoustic sensing system, especially if the fluid is attenuating and/or of low acoustic impedance.

The transducer in Figure 13 (except *top right*) typically is a metal-enclosed impedance-matched configuration that evolved around 1980 [46].

3.4. *Pressure P*

Pressure P has a small effect on elastic moduli and in turn has a small effect on the speed of a guided wave. It is difficult to measure P accurately using the "baro-elastic" effect because the moduli generally change much more in response to a small T increment or decrement (thermoelastic effect), than in response to a small P change. Nevertheless, T-compensated P sensors based on quartz resonators manufactured, for example, by Paroscientific [8] and Quartzdyne [16] measure P (or load, when used in load cells). Quartz crystal cuts were developed

by HP [5] to measure T to high accuracy; HP formerly manufactured T and P sensors using cuts of quartz crystals whose resonant frequency yielded either T or P with little cross-sensitivity. A recent thickness-shear design is analyzed in [70].



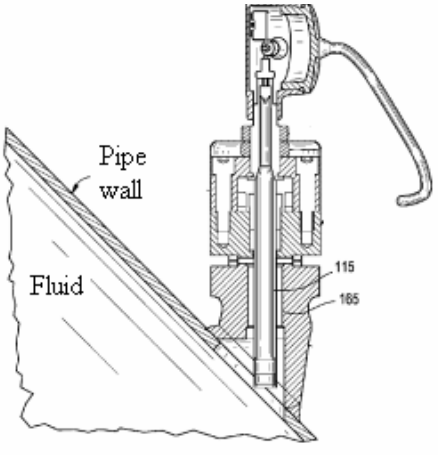


Figure 13. *Top row, Left:* GE's T11 transducer used in gas flowmetering. *Middle*: Impedance-matched, metallurgically-sealed transducer [46] widely used in gas flowmetering since the 1980s. ①Front face; ② Quarterwave Z matcher; ③ Piezoelectric element; ④ Attenuating backing; ⑤ Electrical lead wires. *Right:* 2006 design of an impedance-matched transducer using as the Z-matcher, a dry gel of an inorganic oxide or an organic polymer, the solid skeletal part of the gel having been rendered hydrophobic, achieving low density and homogeneity ("nanofoam") [66]. *Middle row:* Exploded and assembly views show details of Nguyen's 2005 method [54] of using a stack of compressed damping rings adjacent to and abutting the shaft portion of a transducer holder, for attenuating noise traveling rearward through that holder. The holder thereby becomes a damped waveguide. *Bottom:* Isolation built into mounting system, for T11 transducer [51].

Barkhoudarian [3] in 1991 proposed determining P noninvasively based on measuring a T-compensated transit time along a circumferential path 30 in the wall of the pipe. We may think of the wall as a waveguide sensor which stretches as P increases, Figure 14. Axial path 28 was intended to provide T compensation.

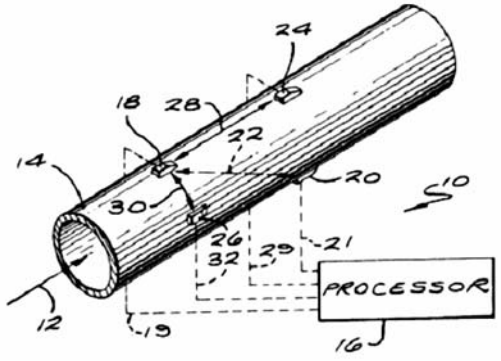


Figure 14. Barkhoudarian's 1991 design [3] utilizes the pipe wall as a waveguide to measure P and T noninvasively. The external transducers 18, 20 measure V by contrapropagation. Invention's intention: sense T along axial path 28 and P along circumferential path 30 after T-compensating.

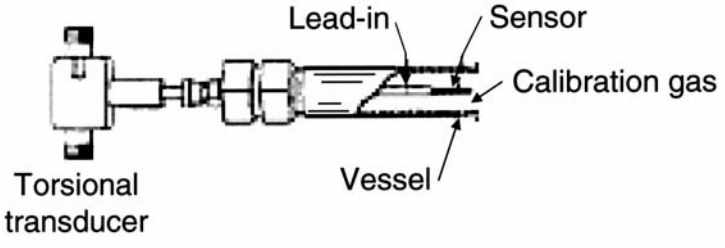


Figure 15. In this diagram, adapted from [44], assume isothermal conditions. Imagine the vessel to be filled with gas such as Ar or He, at atmospheric pressure. By changing the gas from He to Ar, the molecular weights of which are very nearly 4 and 40, respectively, the mass loading increases by essentially one order of magnitude. This simulates a tenfold pressure increase. The fractional torsional $(\Delta t)/t$ is still small, reaching (for a 316SS diamond sensor of aspect ratio three) 15% upon immersion in water. Densities of He, air and Ar are approximately 0.2, 1.3 and 2 kg/m^3 at standard conditions; for water, ρ_L = 1000 kg/m^3. Assume the torsional system responds linearly to fluid density. Consider that a 316SS sensor of diamond cross section, aspect ratio 3, length 53 mm, in vacuum yields a pulse-echo torsional transit time t = 67 μs. Then the increments for immersion in He, air, Ar and water would be ~2 ns, 13 ns, 20 ns, and 10 μs, respectively. Immersion in air at P =10 bar would yield Δt = 130 ns. Measuring Δt to 1.3 ns corresponds to resolving P to 1% at 10 bar, degrading to 10% near atmospheric pressure. These approximations assume perfect T compensation. In [45] the T coefficient of torsional sound speed c_{TORS} for a 316SS diamond sensor was measured in an environmental chamber. Result: $\triangle c_{TORS}/c_{TORS} \approx$ -0.026% per °C between 0 and 100°C.

References

1. Apfel, J. H. (1962). Acoustic thermometry. *Review of Scientific Instruments*, 33 (4), 428–430.
2. Araki, H. and Matsunaga, Y. (1977). *Ultrasonic flow meter*, U.S. pat 4,014,211.
3. Barkhouradian, S. (1991), *Nonintrusive flow sensing system*, US pat 5,040,415.
4. Bell, J. F. W. (1957). The velocity of sound in metals at high temperatures. *Philosophical Magazine*, 2, 1113–1120. (b) Bell, J. F. W. and Doyle, B. P. (1964). Sound pulses in thin non-uniform rods, *Ultrasonics* 2 (1) pp. 39–42. (c) Bell, J. F. W., Doyle, B. P., and Smith, B. S. (1966). An instrument for the measurement of acoustic pulse velocity and attenuation in a solid probe, *J Sci Instrum* 43 (1) pp. 28–31.
5. Benjaminson, A. and Rowland, F. (1972). The development of the quartz resonator as a digital temperature sensor with a precision of 1×10^{-4}, pp. 701–708 in Plumb, H. H. (Ed.-in-chief) *Temperature-its measurement and control in science and industry,* 4, ISA.
6. Brazhnikov, N. I., (1978). *Method of and device for controlling gas-liquid or liquid-liquid interface in monolayer reservoirs*, US pat 4,118,983.
7. Brown, A. E. and Lynnworth, L. (2001). Ultrasonic flowmeters, Chap. 20, pp. 515–573 in Spitzer, D. W. (ed.), *Flow Measurement*, 2nd Ed., ISA.
8. Busse, D. W. (1987). Quartz transducers for precision under pressure, *Mechanical Engineering,* 109 (5) pp. 52–56. See also: http://www.paroscientific.com/qtechnology.htm .
9. Cannistraro, M., Fried, J., and Schultz, J. (2002). Slow torsional wave temperature sensors for electronically-remote temperature sensing, Final Report, ME 43 Senior Design Project, Tufts University (presented in part, Dec. 2001).
10. Carlson, G. A., Sullivan, W. H., and Plein, H. G. (1977). Application of ultrasonic thermometry in LMFBR safety research, *Proc Ultrason Symp*, IEEE, pp. 24–28.
11. Carnevale, E. H., Poss, J. L., and Yos, J. M. (1962). Ultrasonic temperature determinations in a plasma, pp. 959–967 in Herzfeld, C. M. (Ed.), *Temperature-Its measurement and control in science and industry,* Reinhold.
12. Carnevale, E. H., Lynnworth, L. C., and Larson, G. S. (1964). Ultrasonic measurement of elastic moduli at elevated temperatures, using momentary contact, *J Acous Soc Am* 36 (9) pp. 1678–1684.
13. Carnevale, E. H. (1967). Acoustic techniques and applications at elevated temperatures, in Mannella, G. G. (Ed.) *Aerospace Measurement Techniques,* NASA SP-132, pp. 77–103.
14. Carnevale, E. H., Lynnworth, L. C., and Larson. G. S. (1967). Ultrasonic determination of transport properties of monatomic gases at high temperatures, *J Chem Phys* 46 (8) pp. 3040–3047.
15. Craster, R. V. (1997). The light fluid loading limit for fluid/solid interactions, *J Appl Math* 7, pp. 485–506.
16. Eer Nisse, E. P. (1985). *Resonator pressure transducer*, US patent 4,550,610; http://www.quartzdyne.com/.
17. Fowler, K. A. (1971). Acoustic emission simulation test set, *MTRSA* 11 (3) pp. 35–36; with Papadakis, E. P. (1972), Observation and analysis of simulated ultrasonic acoustic emission waves in plates and complex structures, pp. 222–237 in *Acoustic Emission,* ASTM Special Techn Publ 505.
18. Fowler, K. A. (1969). A new way to measure elastic moduli, *Metal Progress* 95 (6) p. 21.

19. Frederick, J. R. (1948). Ultrasonic measurement of the elastic properties of polycrystalline materials at high and low temperatures (Abstr.). *J Acous Soc Am*, 20, 586. See also: (1947). *A study of the elastic properties of various solids by means of ultrasonic pulse techniques*, Ph.D. thesis, U of Michigan.
20. Gysling, D. L., Kersey, A. D., and Paduano, J. D. (2002). *Fluid parameter measurement in pipes using acoustic pressures*, US pat 6,354,147; Gysling, D. L. and Loose, D. H. (2006). US pat 7,032,432; Gysling, D. L. (2006). US pat 7,096,719.
21. Herb, G. T. and Fendrock, C. (2003). *Ultrasonic thermometer system*, US pat 6,517,240. (b) Fendrock, C. and Varela, D. W. (1995). Ultrasonic thermometry for industrial furnace temperature measurement, *Sensors* 12 (1) pp. 23–25, 27, and 37.
22. Hill, J. A. and Pell, J. P. (2000). *Flow measurement system with guided signal launched in lowest mode*, US pat 6,065,350.
23. Huang, M., Jiang, L., Liaw, P. K., Brooks, C. R., Seeley, R., and Klarstrom, D. L. (1998). Using acoustic emission in fatigue and fracture materials research, http://www.tms.org/pubs/journals/JOM/9811/Huang/Huang-9811.html#ToC12 .
24. Hurd, J. C., Nguyen, T. H., and Lynnworth, L. C. (2004). Shear wave polarization follows twist of rectangular steel bar, *Materials Eval* 62 (1), pp. 37–42. http://www.asnt.org/publications/Materialseval/solution/jan04solution/jan04sol.htm .
25. Janzen, V. P., Luloff, B. V., Sedman, K., and Galan, G. (2005). An improved ultrasonic downcomer flow-measurement system for CANDU steam generators, *Proc. 7th CNS International Conference on CANDU Maintenance,* Toronto; (b) Janzen, V. P. (2006), priv. comm.
26. Jen, C. K. and Legoux, J. G. (1998). *Clad ultrasonic waveguides with reduced trailing echoes*, US pat 5,828,274.
27. Jen, C.-K., and Legoux, J. G. (1996). High performance clad metallic buffer rods, *Proc. IEEE Ultrasonics Symp*. 771–776. (b) ------, Nguyen, K. T., Ihara, I., and Hébert, H. (1998). Novel clad ultrasonic buffer rods for the monitoring of industrial materials processing, *Proc. 1st Pan American Conference for Nondestructive Testing*, paper 16, http://www.ndt.net/article/pacndt98/16/16.htm. (c) Legros. N., Jen, C.-K., and Ihara, I. (1999). Ultrasonic evaluation and application of oriented polymer rods, *Ultrasonics* 37 291–297. (d) Jen, C. K., Dai, J. D., Yu, J. C. H., Wang, Z., and Cheeke, J. D. N. (1991). An analysis of thin-rod flexural acoustic wave gravimetric sensors immersed in liquid, IEEE *TUFFC* 38 (3) 312–314. (e) Cheeke, J. D. N., *Fundamentals and Applications of Ultrasonic Waves*, CRC (2002).
28. Kim, J. O. (1989). The Interaction between stress waves transmitted in solid waveguides and adjacent media, Ph. D. dissertation, University of Pennsylvania.
29. Kim, J. O., Bau, H. H., Liu, Yi, Lynnworth, L. C., Lynnworth, S. A., Hall, K. A., Jacobson, S. A., Korba, J. M., Murphy, R. J., Strauch, M. A., and King, K. G., (1993). Torsional sensor applications in two-phase fluids, *TUFFC*, 40 (5) pp. 563–576; (b) Bau, H. H., Kim, J. O., Lynnworth, L. C., and Nguyen, T. H. (1990). *Torsional wave fluid sensor and system*, US pat 4,893,496.
30. Kim, Y. Y., Park, C. I., Cho, S. H., and Han, S. W. (2005). Torsional wave experiments with a new magnetostrictive transducer configuration*, J Acous Soc Am* 117 (6), pp. 3459–3466.
31. Kisner, R. A., Jamarillo, R. A., Freudenberg, K. D., and Holcomb, D. E. (2006). Private communication.

32. Kusters, J. A. and Vig, J. R. (1991). Hysteresis in quartz resonators-a review, IEEE *TUFFC* 39 (3) http://www.ieee-uffc.org/freqcontrol/Hysteresis/Hysteresis_final.html; see also: http://www.oscilent.com/esupport/TechSupport/ReviewPapers/IntroQuartz/vigrefs.html
33. Levitt, A. P. and Martin, A. G. (1960). Ultrasonic determination of elastic constants of metals at elevated temperatures, *Nondestructive Testing* 18 (10) pp. 333–336.
34. Lindsay, R. B., (1972). *Acoustics-Historical and Philosophical Development*, Dowden, Hutchinson & Ross. See also, regarding Colladon's reputation as the "father of light guiding": http://www.rpi.edu/~schubert/Educational%20resources/1800s%20Colladon%20-%20the%20guiding%20of%20light.pdf
35. Liu, Y. and Lynnworth, L. C. (1993). Flexural wave sidewall sensor for noninvasive measurement of discrete liquid levels in large storage tanks, pp. 385–390 in IEEE *Proc.Ultrasonics Symposium*. (b) Liu, Y. and Lynnworth, L. C. (1995). *Elastic wave sensing system*, US pat 5,456,114. (c) See [27]. (d) Sato, H., Lebedev, M., and Akedo, J. (2006). Theoretical and experimental investigation of propagation of guide waves in cylindrical pipe filled with fluid, *Jpn J Appl Phys* 45 (5B) pp. 4573–4576.
36. Liu, Y., Lynnworth, L. C. and Zimmerman, M. A. (1998). Buffer waveguides for flow measurement in hot fluids. *Ultrasonics*, 36(1–5), 305–315. (b) Liu, Y. and Lynnworth, L. C. (1999) *Ultrasonic path bundle and system*, US pat 5,962,790; (2002) US pat 6,343,511.
37. Lynnworth, L. C. (1988). *Method and apparatus for measuring fluid characteristics using surface generated volumetric interrogation signals*, US pat 4,735,097.
38. Lynnworth, L. C. (1989). *Ultrasonic Measurements for Process Control,* Academic Press.
39. Lynnworth, L. C. (1993). *Snap-on flow measurement system*, US pat 5,179,862.
40. Lynnworth, L. C. (2005). Article 119, Ultrasonic instrumentation principles, pp. 802-809; Article 120, Ultrasonic instrumentation design, pp. 810–821 in Sydenham, P. H. and Thorn, R. (Eds.), *Handbook of Measuring System Design*, Wiley. (b) Lynnworth, L. C. and Liu, Yi (2006). Ultrasonic flowmeters: Half-century progress report, 1955–2005, *Ultrasonics* Volume 44, Supplement 1, 22 December 2006, e1371-e1378, *Proc. Ultrasonics International (UI'05) and World Congress on Ultrasonics (WCU)*, http://dx.doi.org/10.1016/j.ultras.2006.05.046.
41. Lynnworth, L. C., Jossinet, G., and Chérifi, E. (1996). 300°C Clamp-on ultrasonic transducers for measuring water flow and level, pp. 407–412, in IEEE *Proc. Ultrasonics Symposium*. (b) Lynnworth, L. C. (2000). *Ultrasonic buffer/waveguide*, US pat 6,047,602. (c) http://www.gesensing.com/products/resources/whitepapers/ur267b.pdf.
42. Lynnworth, L. C. and Mágori, V. (1999). Industrial process control sensors and systems, Chap. 4, pp. 275–470 in Papadakis, E. P. (Guest Ed.), *Physical Acoustics* XXIII, Academic Press.
43. Lynnworth, L. C., Liu, Y., and Umina, J. (2005). Extensional bundle waveguide techniques for measuring flow of hot fluids, *IEEE Trans UFFC* 52 (4) pp. 538–544; http://www.gesensing.com/products/resources/brochures/bwt.pdf http://www.gesensing.com/products/resources/whitepapers/ur257.pdf; http://www.gesensing.com/products/resources/whitepapers/ur262.pdf .
44. Lynnworth, L. C., Smart, C., and Nicholson, T. (2005*). Mass flow sensor and methods of determining mass flow of a fluid,* US pat 6,912,918.
45. Lynnworth, L. C., Cohen, R., Rose, J. L., Kim, J. O., and Furlong, E. R. (2006). Vortex shedder fluid flow sensor, *IEEE Sensors J* 6 (6), pp. 1488–1496.

46. Lynnworth, L. C., Fowler, K. A. and Patch, D. R. (1981). *Sealed, matched piezoelectric transducer*, U.S. Patent 4,297,607. (b) Lynnworth, L. C., Patch, D. R. and Mellish, W. C. (1984). Impedance-matched metallurgically-sealed transducers. *IEEE Trans Sonics and Ultras*, SU-31(2), 101–104. (c) See also: [66]. (d) Lynnworth, L. C., Patch, D. R., and Carnevale, E. H. (1972). *Ultrasonic profile measuring apparatus*, US pat 3,636,754.
47. Lynnworth, L. C., Papadakis, E. P. and Fowler, K. A. (1977). Ultrasound propagation measurements and applications, pp. 77–115 in *International Advances in Nondestructive Testing*, Vol. 5 (ed. W.J. McGonnagle), Gordon & Breach.
48. Matson, J. E., Nguyen, T. H. and Lynnworth, L. C. (1987). Ultrasonic measurement of liquid flow using clamp-on Rayleigh wave transducers, *IEEE Transactions I & M Conference Proceedings,* pp.197–206. (b) Matson, J., Mariano, C., Khrakovsky, O. and Lynnworth, L. (2002). http://www.gesensing.com/products/resources/whitepapers/ur240.pdf
49. Matson, J. and McCormick, J. (2006). Recent tests of crossed diameter paths 10D downstream of a multihole flow conditioner, *Proc. AGA Operations Conf.* Boston. (b) Matson, J. and Smith, R.G. (2007). The future of flare gas measurement, *Flow Control* XIII (1) 14–20.
50. Mayer, A. M. (1873). On an acoustic pyrometer, *Phil Mag* Ser 4, 45, pp. 18–22.
51. McCarey, R. A. Jr. and Frail, C. (2006). *Transducer holder and nozzle,* US pat 7,086,286.
52. Miller, G. N., Anderson, R. L., Rogers, S. C., Lynnworth, L. C., Studley, W. B. and Wade, W. R. (1980). High temperature, high pressure water level sensor, *Proc Ultras Symp*, pp. 877–881, IEEE.
53. Monchalin, J.-P. (2007). Laser ultrasonics, Chapter 4 in this volume.
54. Nguyen, T. H. (2005). *Short-circuit noise abatement device and method for a gas ultrasonic transducer*, US pat 6,876,128.
55. Nguyen, T. H., Smart, C. D., and Lynnworth, L. C. (2004). Transverse shear to torsion mode converter and applications, *Materials Evaluation* 62 (6) 690-698; (b) Nguyen, T. H. and Lynnworth, L. C. (2005), *Acoustic waveguide system*, US pat 6,889,552; (c) Lynnworth, L. C., Cohen, R., and Nguyen, T. H. (2004). Clamp-on shear transducers simplify torsional and extensional investigations, *Proc. IEEE Ultras Symp*, pp. 1603–1607.
56. Papadakis, E. P., Fowler, K. A., Lynnworth, L. C., Robinson, A., and Zysk, E. D. (1974). Ultrasonic measurements of Young's modulus and extensional wave attenuation in refractory metal wires at elevated temperatures with application to ultrasonic thermometry, *J. Appl Phys* 45 (6) pp. 2409-2420. (b) Papadakis, E. P., Lynnworth, L. C., Patch, D. R., and Carnevale, E. H. (1972). *Ultrasonic thermometry for LMFBR systems*, Final report, NYO 3906-13.
57. Park, C. I., Kim, W., Cho, S. H., and Kim, Y. Y. (2005). Surface-detached V-shaped yoke of obliquely bonded magnetostrictive strips for high transduction of ultrasonic torsional waves, *Applied Physics Letters* 87, 224105.
58. Perng, K.-N. and Liu, C.-S. (2006). *Nondestructive method for inspecting cladding tubes*, US pat 7,104,127.
59. Potzick, J. E. and Robertson, B. (1984). *Long wavelength acoustic flowmeter*, US pat 4,445,389.
60. Rountree, S. P. and Berjaoui, S. W. (1996). *Method and apparatus for use in ultrasonic ranging*, US pat 5,568,449.
61. Royer, D. and Dieulesaint, E. (1999). *Elastic Waves in Solids I - Free and Guided Propagation*, Springer-Verlag NY; (b) Rose, J. L. *Ultrasonic Waves in Solid Media* (1999). Cambridge University Press.

62. Shen, C. and Jacobson, S. A. (2001). *Coherent multi-path flow measurement system*, US pat 6,293,156; (b) Jacobson, S. A., Denbigh, P. N., and Naudé (1985). A new method for the demodulation of ultrasonic signals for cross-correlation flowmeters, *Ultrasonics* 23 (3) 128–132; (c) Jacobson, S. A. , Lynnworth, L. C., and Korba, J. M. (1988). *Differential correlation analyzer*, US pat 4,787,252; (d) Jacobson, S. A. (1995). *Flow measurement system including ultrasonic transducers*, US pat 5,460,047; (e) Collard, S. M. and McLellan, R. B. (1990). Modulus defect in the noble metals, *Scripta Metallurgica et Materialia* 24, pp. 623–627.
63. Shepard, C. L., Burghard, B. J., Friesel, M. A., Hildebrand, B. P., Moua, X., Diaz, A. A. and Enderlin, C. W. (1999). Measurements of density and viscosity of one- and two-phase fluids with torsional waveguides, *IEEE TUFFC* 46 (3) pp. 536–548.
64. Sherratt, G. G. and Griffiths, E. (1934). The determination of the specific heat of gases at high temperatures by the sound velocity method, I-carbon monoxide. *Proceedings of the Royal Society (London), Series A*, 147, 292–308.
65. Stein, P. J. and Euerle, S. E. (2001). *Non-invasive low frequency elastic wave sensing system for sludge laden environments*, US pat 6,192,751.
66. Suzuki, M., Hashida, T., and Hashimoto, M. (2006). *Ultrasonic transducer*, US pat 7,087,264. (b) See also: Nagahara, H., Suginouchi, T., and Hashimoto, M. (2006). Acoustic properties of nanofoam and its applied air-borne ultrasonic transducers, Paper P1M-8, pp. 1541–1544, *Proc. Ultrason. Symp.* Hashimoto, M., Nagahara, H., and Suzuki, M. (2007). *Ultrasonic sensor*, US Patent 7,162,930.
67. Tomioka, G. (1968). *Apparatus for measuring liquid level,* US pat 3,394,589.
68. Van de Vaart, H., Matthews, H. B., and Worley, J. C. (1975). *Acoustic delay surface wave motion transducers*, US pat 3,863,497.
69. Vogt, T., Lowe, M., and Cawley, P. (2003). Cure monitoring using ultrasonic guided waves in wires, *J Acous Soc Am* 114 (3) pp. 1303–1313. (b) See also, waveguide studies listed in http://www.imperial.ac.uk/ndt/public/publications/2005.htm. (c) Matsiev, L. (2006). Measurements of liquid density and viscosity with flexural resonators using noise as an excitation source, Paper 3I-2, pp. 880–883, *Proc. Ultrason. Symp.*
70. Y. Hu, J. Yang, Y. Zeng, and Q. Jiang (2006). A high-sensitivity, dual-plate, thickness-shear mode pressure sensor, *TUFFC* 53 (11) pp. 2193–2197.

Part 2

Advanced (Other) Methods

This page intentionally left blank

CHAPTER 11

ELECTROMAGNETIC NDE TECHNIQUES FOR MATERIALS CHARACTERIZATION

B. P. C. Rao, T. Jayakumar and Baldev Raj

Non-destructive Evaluation Division
Metallurgy and Materials Group
Indira Gandhi Center for Atomic Research
Kalpakkam – 603 102, TN, India
bpcrao@igcar.gov.in

In this chapter, various electromagnetic NDE techniques such as eddy current, magnetic, micro-magnetic, potential drop and microwave techniques for materials characterization are covered. For each technique, principle, instrumentation along with sensors, capabilities, applications, and recent advances towards quantitative characterization of defects, microstructures and residual stresses in engineering materials are given. Special emphasis is also given to numerical modelling, sensors and signal processing for automated NDE.

1. Introduction

Engineering components and structures in power, chemical, petrochemical, fertiliser, transport and other industry are usually subjected to complex service environment viz. exposure to elevated temperature, hostile media, loading, transients etc. causing formation and growth of defects, accumulation or redistribution of residual stresses and degradation due to microstructural changes, corrosion, creep and fatigue damage [1, 2]. Likewise, discontinuities may be present in the raw material stage or might have been introduced during machining, fabrication, heat treatment and assembling. Defects are discontinuities that influence structural integrity of a component and these include cracks, inclusions, voids, corrosion wall loss, lack of fusion and penetration in welds, tensile residual stresses, embrittlement etc.

According to fracture mechanics, defects present in materials lead to failure by growing to a critical and self-propagating size. The fracture mechanics concepts allow one to calculate the critical sizes of defects as a function of defect dimensions, active stress system, stress intensity and material properties such as elastic modulus, yield strength and fracture toughness. By knowing the dimensions of defects present in a component, it is possible to estimate both extent of degradation and remaining life of the component and defects' growth characteristics under operating conditions. Early detection and characterisation of defects, microstructural variations and residual stresses, is essential to avoid damage and loss of structural integrity of components [3]. In the industrial scenario, with the incorporation of strategies such as total quality management, structural integrity assessment and residual life assessment, the role of NDE is becoming more vital.

The two important aspects in materials characterisation are i) reliable detection of discontinuities i.e. defects, microstructural variations, stresses and ii) characterization of discontinuities i.e. length, width, depth, location, orientation, shape, distribution etc. It is a pre-requisite for the later aspect that the NDE response of different discontinuities is different so that a detected discontinuity is characterized by utilising the difference in the NDE response.

NDE techniques that use some form of electromagnetic excitation are termed as electromagnetic techniques and some of these include eddy current (EC), magnetic particle (MP), magnetic flux leakage (MFL), magnetic Barkhausen emission (MBE), potential drop and microwave techniques [3-7]. In these techniques, material under investigation is excited electromagnetically and the manifestation of electromagnetic fields due to material discontinuities affecting electrical conductivity or magnetic permeability or dielectric permittivity are measured using a sensor, with the exception of MP testing in which sensor is not used. With the field-discontinuity interaction being different, the detection sensitivity and hence the capability, applicability, and versatility of various electromagnetic techniques differ and as a result selection of NDE technique is important. Choice of NDE technique depends on material, geometry, defect type and location, applicability, accessibility and suitability. It may be sometimes necessary to use a combination of two or more techniques, in the best complementary way, to carry out reliable inspection utilising data fusion concepts which allow combining information of different techniques or sensors to provide comprehensive information of the inspected regions. In such situations, detailed mock-up studies are essential. Reference defects or heat treatments are used for calibrating the instruments and also for quantitative characterisation of discontinuities. Natural cracks or set of artificial defects made

by electric-discharge machining (EDM) or specially produced fatigue cracks or heat treated microstructures are used for this purpose.

Many developments are taking place in the existing electromagnetic NDE techniques incorporating the rapid progress in the fields of microelectronics, instrumentation, sensors, computers, signal processing, and robotics. The way electromagnetic NDE is practised now is different from it was two decades ago [3]. These concurrent advances in other fields have enhanced the capabilities of the electromagnetic NDE techniques enabling detection and sizing of incipient as well as buried defects, changes in microstructures and accumulated deformation, stress or damage e.g. prior to crack formation. Such possibilities allow efficient preventive actions to be taken avoiding catastrophic failure of materials. Digital signal processing, image processing and artificial intelligence are being increasingly applied for high sensitivity detection and characterisation of defects even in the presence of disturbing variables and noise [7]. Automated NDE is widely practised to inspect large areas, eliminate human fatigue and uncertainty, avoid access constraints for manual inspection, increase inspection speed and to carry out on-line inspection. This chapter attempts to explain how materials characterization is performed using electromagnetic NDE techniques utilising the advances in related fields of science and technology.

2. Eddy Current Testing

Eddy current (EC) testing is the most popular and widely used electromagnetic NDE technique. In industrial scenario, among other electromagnetic NDE techniques, this technique finds large number of applications and used in majority of the engineering industries. In EC testing, a coil also called probe is excited with sinusoidal alternating current usually in the frequency range of 1 kHz-10 MHz and change in coil impedance that arises due to distortion of eddy currents by discontinuities in electrical conductivity, magnetic permeability and geometry is measured [4]. The impedance change in the probe coil is also affected by test frequency and spacing between the coil and the material (also called lift-off or fill factor). EC coils possess directional properties i.e. regions of high and low sensitivity (impedance change) [8, 9]. Discontinuities that cause maximum perturbation to eddy currents produce large change in impedance that is detected with high sensitivity. The locus of impedance change during the movement of EC probe over the test material is called the EC signal. While the amplitude of the EC signal provides information about the severity of the discontinuity, the phase angle provides information about its location or depth. Due to the *skin effect*, one can readily detect the surface-breaking defects as

compared to the sub-surface defects or buried defects. Selection of test frequency is very important in the EC tests and in general, it is chosen such that a maximum amplitude signal is produced for desirable discontinuities or defects and with a good phase separation from lift-off [10]. Besides detection of defects, EC technique is widely used for material sorting, determination of hardness, heat treatment inadequacies, coating thickness measurements, and detection of defects in tubes, rods, bars, multi-layer structures, discs, welds, blades and other regular and irregular geometries [11].

Usual test procedure for characterisation of discontinuities involves first calibration. Artificial defects such as saw cuts, flat bottom holes and EDM notches, produced in a material with similar chemical composition and geometry as that of the actual component, are used. Well-characterised natural defects such as service induced fatigue cracks and stress corrosion cracks are preferred, if available [12]. The test frequency, instrument gain and other instrument functions are optimised such that all the specified artificial defects are detected applying a threshold to EC signal parameters, e.g. signal peak-to-peak amplitude, and phase angle etc. During actual testing, any region that produces signals greater than the threshold is recorded defective. For detailed characterisation, master calibration graphs, e.g. between EC signal parameters and defect sizes are obtained.

Detection and characterisation of defects in the presence of disturbing variables is difficult using calibration graph based method. While analogue filters are useful in some situations, use of digital signal processing and artificial intelligence methods is very attractive [13]. Signal processing methods produce noise-free signals from which a variety of signal parameters are determined for input to artificial intelligence (AI) methods for decision making. The input parameters can be from time-domain, frequency-domain, cepstral domain, wavelet-domain etc. and the AI module output can be classification based or quantitatively modelled, as in multiple-regression. The AI module can be off-line or on-line. For example, a multi-layer perceptron (MLP) based artificial neural network (ANN), can classify heat exchanger tube defects off-line into shallow, medium, deep and through wall, based on the wall loss. Likewise, a radial-basis function (RBF) based ANN which performs on-line multi-frequency multi-parameter EC testing allows detection and characterisation of defects in stainless steel cladding tubes (outer diameter 5.1 mm, wall thickness 0.38 mm) with periodic wall thickness variations formed during manufacture and stainless steel plates and welds with variations in surface roughness, lift-of and material properties [14]. The ANN method was applied to two practical situations; a) quantification of depth of stellite coating on carbon steel plates and b) quantification of depth of NiAl coating on Inconel 718.

Eddy current sensors or probes are induction based and are made up of a few turns of copper wire usually wrapped around a ferrite core with or without shielding. Every probe has an operating frequency range and impedance value matching the instrument bridge circuit. It is desirable to operate the probe within that range and if to be used beyond, it is essential to avoid operating near resonance frequency. EC probes can be absolute or differential (comparing) and their design is dependent on the object geometry viz. tube, plate, bar etc. and the expected type and location of discontinuity. While pancake or surface probes are used for testing plates and regular geometries, encircling or bobbin type probes are employed for testing tubes, rods, and other cylindrical objects [10]. The encircling probes exhibit reduced sensitivity for shallow and localised defects and for such applications, motorised rotating probe coil (MRPC), phased-array, plus-point etc. are used. The inspection data from these probes can be displayed as images and allows identifying circumferential location of defects. For inspection of irregular and inaccessible regions, flexible and conformal sensors are employed. The pancake type probes show reduced sensitivity for sub-surface and buried defects and for such needs, integrated probes with coils for excitation and giant magneto resistance (GMR) sensors for reception are very attractive.

GMR sensors consist of magnetic metallic multilayers such as Fe/Cr, Co/Cu, and NiCo/Cu in which ferromagnetic layers are separated by non-magnetic layers of a few nm thick. The application of a magnetic field to the metallic multilayer results in a significant reduction of the electrical resistance of the multilayer due to the spin dependent scattering of electrons at interfaces. The GMR sensors are characterized by high sensitivity, high signal-to-noise ratio, high spatial resolution and frequency independent sensitivity and as a result, they are very attractive for electromagnetic NDE sensing applications, especially for inspection of rivets and multi-layer structures in aircrafts by integrated EC-GMR probes. EC-GMR probes can reliably detect defects located 9mm below the surface in carbon steel and stainless steel plates. Similar to GMR, giant magneto-impedance (GMI) effect in which there is a very large impedance change under an applied magnetic field, in amorphous wire, thin films, and ribbons is finding many applications in NDE. GMI sensors show high sensitivity and give quick response under low magnetic fields and can be used with conventional EC instruments. Other advances in EC sensors include chip-based sensors, wireless sensors and networking of sensors.

2.1. *Non-ferromagnetic Tube Testing*

For periodic monitoring of corrosion of tubes in heat exchangers, steam generators, condensers and boilers in petrochemicals, fertiliser, power and other industries, EC technique is a natural choice. This is because of its ease of operation, sensitivity, versatility, speed (up to 10 m/s) and reliability. EC testing can detect wall thinning, cracks, pitting, stress corrosion cracking, hydrogen embrittlement, carburisation, denting and crud deposits etc. Conventional single frequency EC testing is inadequate when interferences arise from disturbing sources such as support plates, baffle plates and probe wobble [4-8]. Many a time, it is under the support plates corrosion damage takes place. To eliminate signals from support plates, multi-frequency technique is employed. This technique involves simultaneous excitation of two or more frequencies in an EC probe and processing of the corresponding signals to suppress the contributions from disturbing sources [15]. Examination of non-ferromagnetic tubes is performed according to ASME Section V, Article 8, Appendix 1.

2.2. *Remote Field Technique for Non-ferromagnetic Tube Testing*

Examination of ferromagnetic tubes is difficult using conventional testing procedures due to magnetic permeability which restricts penetration of eddy currents and produces disturbing signals. This disturbance can be eliminated by employing bias or saturation current which saturates the material magnetically and makes the tube material behave non-ferromagnetic, thus, allows conventional EC testing. However, in the case of installed tubes of smaller diameters, saturation units can not be accommodated due to limited access. For mildly magnetic materials, partial saturation using high-energy permanent magnets is a possibility. However, the most widely used technique for inspection of ferromagnetic tubes is remote field eddy current (RFEC) technique. This technique uses low frequency excitation and a separate receiver coil kept at two to three tube diameters away from exciter coil. The phase lag of induced voltage in the receiver coil with respect to the exciter is measured using lock-in amplifiers and correlated with wall loss. The advantages of RFEC technique include

- ability to test tubes with equal sensitivity to internal and external wall loss
- linear relationship between wall loss and measured phase lag and
- higher lift-off tolerance and absence of lift-off related problems
- ease of numerical modelling

RFEC technique can achieve inspection speeds of 1 m/s and is used for inspection of carbon steel and other ferromagnetic tubes in process industry. For verification of RFEC results and for taking decision on plugging tubes with reported indications, ultrasonic internal rotary inspection system (IRIS) that gives wall thickness map is employed. One recent application of RFEC technique is discussed. In-service inspection of steam generators (SG) of sodium cooled fast breeder reactors with 23 m long modified 9Cr-1Mo ferromagnetic steel tubes is challenging in view of smaller diameter, expansion bends, support plates, one-side access and sodium deposits. A comprehensive RFEC technology comprising of instrumentation, probes, robotic device has been developed [16]. The transition and remote field zones of RFEC probes have been studied using finite element model to optimise the excitation frequency and the spacing between exciter and receiver coils. In order to negotiate expansion bend regions, flexible RFEC probe has been developed. Wavelet transform based signal processing method and duel-frequency method have been developed to suppress the signals due to bend regions and support plates. The influence of sodium deposits on the outside surface of SG tubes as well as within any defects has been studied concerning the defect detection sensitivity. Change in shape of signal from defects in SG tubes has been observed as shown in Fig. 1 after exposure to sodium. However, it has been established that the peak-to-peak amplitude remains same and does not affect the calibration graphs for sizing. The RFEC technology has been successfully field tested in a 1:1 SG test facility. A compact prototype spider robot with closed loop servomotor controlled device has been developed for automated in-service inspection of SG. The robot moves over the tube sheet and positions winch controlled RFEC probe with an accuracy of ±0.5 mm.

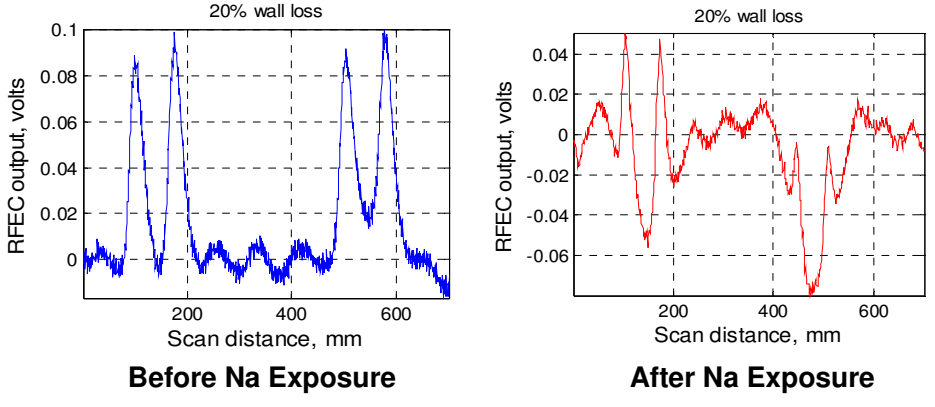


Figure 1. The RFEC signals obtained from modified 9Cr-1Mo ferritic steel before and after exposure to sodium.

2.3. *Eddy Current Imaging*

The widely used single-scan based EC methods fail to give spatial information of defects. However, there is a clear benefit if raster-scan imaging is performed with one EC sensor or line-scan using arrays and the resulting EC data is represented in the form of an image [17]. The image format provides comfortable interpretation enabling discern irregularities in an image by comparing different regions. Further, the processed image and the defect details provide an objective and documentable information that could, for example, be used to monitor the growth of defects with time. More importantly, process automation is possible since computers perform the raster-scan imaging and data processing. Eddy current imaging is used for detection of corrosion and cracks in multi-layer structures of aircrafts, dovetail regions of aeroengines, heat exchanger tubes, automobile components, pressure vessels, and other critical engineering components. One interesting application of eddy current imaging is location of centreline of stainless steel welds in reactor vessel of fast reactors with an accuracy of ±0.5 mm to scan ultrasonic transducers at desired skip distances for inspection of welds and heat-affected zones. Presence of intentionally formed delta ferrite (magnetic phase), to prevent hot cracking of austenitic stainless steel welds, produces large impedance change enabling reliable detection of weld centreline at 200°C and with 10 mm lift-off.

Processing of eddy current images is essential. Image processing methods applicable to eddy current images are different, as compared to ultrasonic or optical images because of larger pixel size and image blur due to convolution of broad point spread function (*foot print*) of probe [17]. Image restoration is possible by deconvolution and Weiner filter based inversion methods. Several image processing methods (time-based, frequency-based and wavelet-based) are developed for image enhancement, image compression, image fusion and delineation of desired information in images and for determination of input features to artificial intelligence schemes for off-line or on-line testing.

Imaging by mechanical scanning of EC probe over the material surface is slow and influenced by material variations and lift-off. For quick and automated EC testing of SS plates and welds, an intelligent imaging scheme, which synergistically combines artificial neural network and image processing methods, has been developed [18]. In this scheme, as shown in Fig. 2, first a coarse scan imaging is made to quickly identify defective regions and then a fine scan imaging of the defective regions is performed to obtain three-dimensional images with accurate length, width, depth, and orientation information of defects [19].

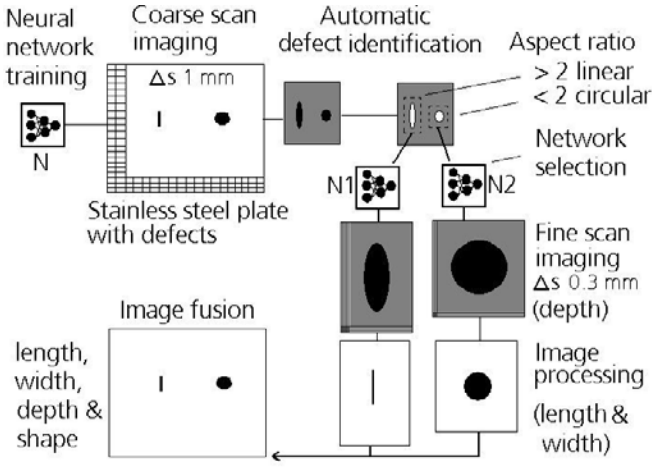


Figure 2. Intelligent imaging scheme for automated eddy current testing.

2.4. *Magneto-optic Imaging*

Magneto-optic imaging (MOI) is a relatively new electromagnetic NDE technique that provides real time images of surface and subsurface defects in metallic materials e.g. aircraft structures [20]. In this technique, flux leakage maps associated with applied currents are recorded and monitored to locate defects. The device combines planar eddy current and magneto-optic imaging, and is applicable to real time imaging of aircraft structures through paint and other surface coverings. Using this technology, magnetic fields associated with the disruption of induced eddy currents produced by intervening rivets, cracks, corrosion, and other defects are detected by a sensor that exhibits the Faraday magneto-optic effect. The operating frequency of MOI applications typically ranges from 100 Hz to 100 kHz. The presence of rivets or defects diverts the eddy current from its uniform flow, and hence generates a normal magnetic field component, which can be measured using Faraday's magneto-optic effect. A magneto-optic sensor, placed parallel to the eddy current induction foil, is excited by a linearly polarized light. The plane of polarization undergoes rotation by an angle that depends on the local magnetic field intensity, and the specific Faraday rotation of sensor material. Perturbations in the magnetic field are monitored by analyzing rotation in light polarization. Images appear directly at the sensor and can be viewed by eye, or as an option the output can be displayed on or computer or TV screen for viewing, monitoring and recording [21]. It is possible to numerically model the MOI technique to study the detection sensitivity in a test

situation. This method is widely used by airlines for inspection of fuselage, wing and control surfaces of aircraft. This technique has proved to be highly reliable and accurate in detection of fatigue cracks below the rivets and also for characterization of corrosion and other subsurface defects. The chief advantages of MOI technique include increased inspection speed, more intuitive, easily interpreted inspection information, and the ability to easily document the inspection images.

2.5. *Detection of Intergrannular Corrosion and Stress Corrosion Cracking*

Stainless steels are prone to intergranular corrosion (IGC) due to formation of chromium carbides at grain boundaries and depletion of chromium level within the grain to less than 12% when the material is exposed to 873-1173 K. Feasibility of detection and accurate characterisation of IGC in AISI type 316L stainless steel using EC technique has been established [22]. Specimens subjected to ageing treatment at 973 K for 1, 20, 50, and 200 hours duration and exposed to Strauss test have been examined. The EC amplitude has been noted to change by a small amount of 0.25 volts with ageing for different durations. On the contrary, the amplitude has been observed to increase monotonously to about 3 volts, after the Strauss test. This increase in the EC voltage has been attributed to the changes that have taken place during the Strauss test i.e. dissolution of Cr depleted regions and formulation of grain boundary grooving. In order to bring out the relative changes clearly, ratio of EC amplitudes of Strauss test exposed to unexposed conditions has been evaluated for each specimen as shown in Fig. 3.

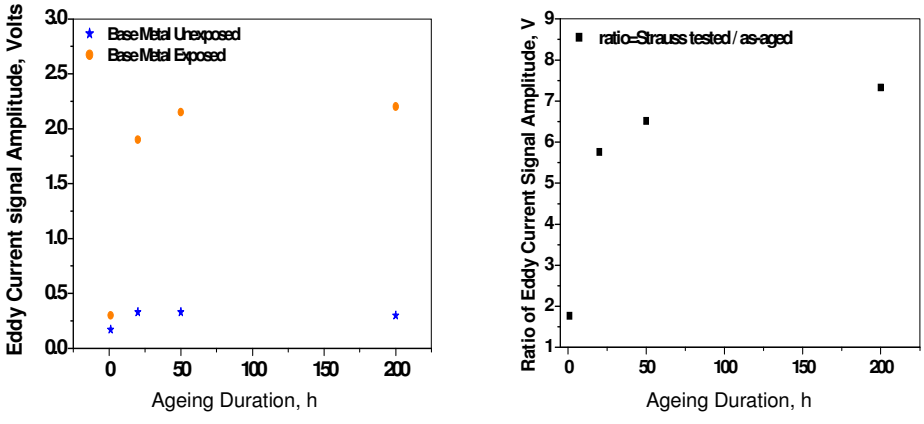


Figure 3. EC response on raw and Strauss test exposed stainless steel 316L specimens.

As can be observed, when the ageing duration is more than 20 h, the ratio of EC amplitude becomes more than 5 and is well delineated from the specimens with ageing duration of 1 h and without ageing. Bend tests on Strauss test exposed specimens have confirmed that sensitisation has taken place in specimens aged for duration of 20 hours and above. The study has been recently extended to SS 316 and it has been found that sensitization of specimens that leads to failure by Strauss test can be detected based on the EC amplitude, even without the Strauss test [22]. As can be observed, when the ageing duration is more than 20 h, the ratio of EC

2.6. *Characterisation of Microstructures*

During manufacturing and service life of components, it is essential to ensure that there is no undesirable degradation in microstructures. Electromagnetic NDE techniques exploit measurement of changes in electrical conductivity and magnetic permeability to characterize microstructural changes in metallic materials. EC testing is a natural choice as eddy currents are influenced by changes in microstructure, precipitates, cold work, deformation, dislocation pile-up etc. and the coil impedance or induced voltage in pick-up coil changes accordingly. The magnitude and phase of induced voltage or impedance change are used to correlate with the microstructures and to estimate the volume fraction of various phases present in the material [1].

Specimens subjected to heat treatments are used to simulate the service exposed conditions and the expected microstructures. Usually, EC measurements for microstructure characterisation are location-based. In general, absolute probes are used and analysis is based on impedance plane signal interpretation. Reference standards with known conductivity and permeability are used for quantitative characterisation of microstructures besides, specimens heat treated to different ageing conditions.

Cold worked and annealed condition, e.g. in stainless steel 316 or 304 effects electrical conductivity in opposite directions and strain-induced martensite, being a magnetic phase, increases the magnetic permeability. This phase can be detected using EC technique [1]. In these steels, delta ferrite that is intentionally formed in welds to arrest hot cracking can be reliably detected, even with 10 mm lift-off, due to its magnetic permeability as discussed Section 2.3. Likewise, in stainless steel welds subjected to fatigue cycles, transformation of magnetic phases into non-magnetic phases reduces the impedance change.

EC technique can characterise microstructures in titanium alloy (VT 14 alloy Ti-4.5Al-3Mo-1V) subjected to a series of heat treatments consisting of

solutionizing for 1 h at selected temperatures in the range of 923-1303 K at an interval of 50K, followed by water quenching. This treatment produces variety of microstructures due to controlled α-β transformation and formation of various phases. Comparison of impedance magnitude and phase angle with ultrasonic velocity and hardness measurements has established that EC technique can be implemented in production line to quickly assess the adequacy of heat treatment [23]. EC technique is also used to characterise precipitation hardening in Duralumin alloy and ageing effects in heat treated M250 Maraging steels that are very attractive for aerospace industry in view of their high strength, fracture toughness and corrosion resistance. In order to study the microstructural changes occurring during heat treatment, M250 Maraging steel specimens thermally aged at 755 K for 0.25, 1, 3, 10, 30, 40, 70 and 100 h of duration have been examined using EC technique. It has been established, for the first time, that after 0.25 h there is a slight increase in permeability and reduction in resistivity due to annihilation of dislocations [24]. The probe response from other specimens indicated that the permeability decreases with ageing duration, as shown in Fig. 4 due to formation of precipitates (Ni_3Ti) and Fe_2Mo and reversion of martensite to austenite phase. EC technique can also be used to characterise cast irons [25].

2.7. *Pulsed Eddy Current Technique*

In contrast to conventional eddy current techniques, in pulsed eddy current (PEC) technique, exciter is driven with a repetitive broadband pulse, such as a square wave which induces transient eddy currents associated with highly attenuated magnetic pulses propagating through the material. The signals reflected from defects are picked up. At each probe location, a series of voltage-time data pairs are produced as the induced field decays, analogous to ultrasonic A-scan data. Defects close to the surface will affect the eddy current response earlier in time than deep defects [26]. The modes of presentation of PEC data are analogous to ultrasonic methods viz. A-, B- and C-scans that enable detailed characterisation of defects. Eddy currents have advantage of penetrating into subsurface layers whether or not the layers are mechanically bonded. Making use of this, PEC technique is used for detection of hidden corrosion in layered structures such as aircraft lap-splices, corrosion under insulation in insulated components [27]. PEC technique has capability to detect and quantify metal loss more than 30% of one plate thickness in case of corrosion in multi-layered structures. Imaging can be carried out with PEC and the data can be sliced in depth direction for detailed characterisation.

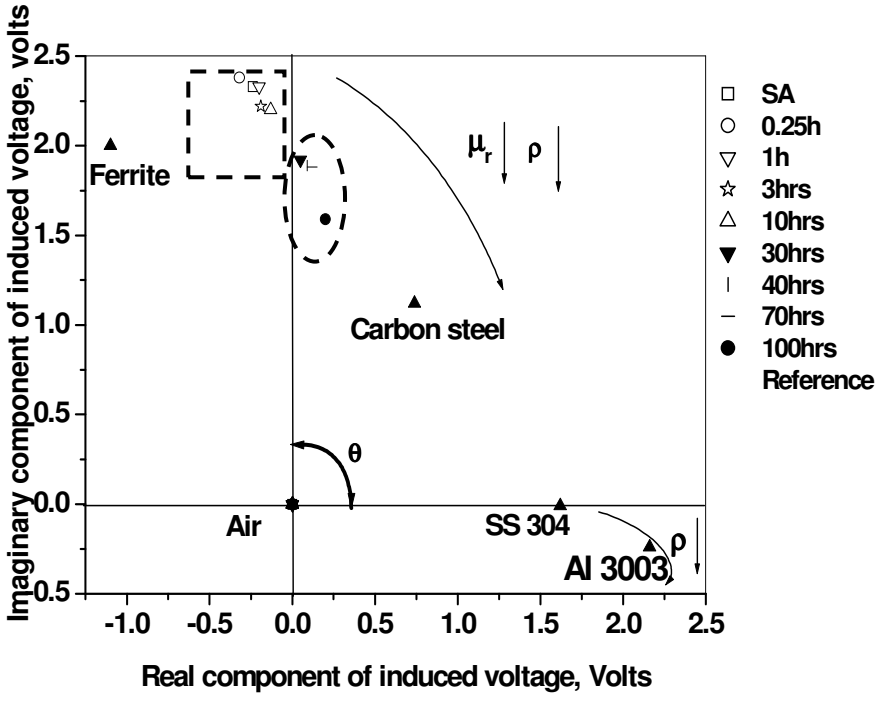


Figure 4. Variation of induced voltage in T/R probe with aging time.

2.8. *Low Frequency Eddy Current Technique*

For inspection of covered components to detect corrosion without removing insulation, low frequency EC (LFEC) technique is an attractive option [28]. The technique is also sensitive to localized damage such as cracks and pits, and can detect severe wall thinning. LFEC technique is a time-base technique that uses very low frequency (less than 50 Hz) and separate exciter and receiver coils. In this technique, current through a transmitter coil is pulsed to induce magnetic fields and eddy currents which vanish when current is switched off. These eddy currents diffuse into the wall, decay with time at a certain rate and are picked up by the receiver coil. At places of wall loss, the time of arrival will be earlier than places with no wall loss and from this time of arrival, average wall thickness can be determined. The method can also be used to detect cracks in multi-layered structures. The most important application of LFEC technique is detection of corrosion on the hidden side of aircraft skin panels and detection of cracks beneath external repair doublers. Further, use of GMR sensors enhances the detection sensitivity of the technique [29]. LFEC technique allows rapid scanning,

in comparison to PEC technique, for detection of corrosion under insulation and corrosion areas less than 25 mm can be detected with thinner insulation.

3. Magnetic Particle Testing

Magnetic particle (MP) testing is simple and highly sensitive technique that provides clear indications of surface and sub-surface physical discontinuities in magnetic materials [30]. When a ferromagnetic component is magnetised, the magnetic lines of force (the magnetic flux) are predominantly inside the material. However when there is a surface breaking or a sub-surface discontinuity, the field is distorted, causing local magnetic flux leakage fields. These leakage fields are detected by sprinkling finely divided ferromagnetic powders. Because of the added advantage of MP testing over liquid penetrant testing in detecting subsurface defects, it is usually specified for inspection of all ferromagnetic components. MP testing can be used for detection of cracks, blowholes, laps, non-metallic inclusions, segregation etc. The sensitivity of MP testing depends on the magnetisation method and on the electromagnetic properties of the material tested as well as on the size, shape and orientation of the defect. MPT is widely used for inspection of cracks in transportation industries like crankshafts, fly wheels, crank hooks, shafts, steam turbine blades etc. [31] Automatic magnetic inspection is used for ball and roller bearings, bearing races and rings, small castings and forgings and steel-mill billets [32]. Scope does not exist in MP tests to apply the signal processing methods for enhanced detection and the inversion methods for accurate characterisation of defects as practised in ultrasonic, eddy current and other NDE methods. In light of this, magnetisation methods, amperage, powders, carrier fluids, sprinkling methods, viewing conditions, recording methods etc. are carefully tailored. After the inspection by magnetic methods, it is necessary to subject the component to demagnetisation.

MP testing consists of magnetisation of the component, applying magnetic powder and examination of powder patterns. The indications of defects can be preserved by photography or video recording or by the use of peel-off transparent adhesive films. Evaluation of length, width and lateral shape of indications or defects is carried out using visual aids and other measuring techniques and more recently, using advance image processing methods. Dry powder methods are employed if large discontinuities (>1 mm), especially the sub-surface ones are expected. Red coloured powders are preferred on dark surfaces and black coated powders on hot objects (up to 400° C). On the contrary, to detect small and shallow surface discontinuities such as tight fatigue cracks, wet fluorescent methods with black light illumination are employed [12]. Figure 5 shows typical

fluorescent MP testing of a steam turbine rotor using cable-wrap technique and typical indication of a tight fatigue crack in the 'fir-tree' root region of a turbine blade. For the inspection of relatively inaccessible regions e.g. bolt holes and internal threads, magnetic rubber inspection methods are adopted. For under water inspections, magnetic ink, putty, or tapes are used.

One major requirement of MP testing is the necessity for application of magnetic field in two different directions, viz. circular and longitudinal. There are a wide variety of methods to magnetise a component such as prod, head shot, central conductor, yoke and cable. The method selected depends on such factors as the geometry of the component, desired direction of magnetic field, the expected orientation of defects and whether or not the defects are subsurface. Both AC and DC methods are used for magnetisation. Because of *skin-effect*, AC methods are restricted to detection of near-surface defects and therefore, are not very effective for detection of subsurface defects. MP tests are usually performed as per ASME-Section V-Article 7, ASTM E1444-05 and ASTM E 709-05.

Figure 5. Fluorescent MP testing of a steam turbine rotor and typical indication from a tight fatigue crack in the 'fir-tree' root region of a low-pressure blade.

To inspect large components, extremely large magnetizing currents are required and if the surface of the component is covered by a film or other nonmagnetic layer, the sensitivity is reduced. For near-surface defects, the detection sensitivity decreases drastically with the location from the surface. A fall out of these facts is the uncertainty that arises from a large near-surface defect and a shallow surface defect. To classify an indication into surface and near-surface, other NDE methods such as visual testing are necessary. Automatic MP testing systems that use video cameras and image processing methods have been developed, e.g. for inspection of turbine blade roots and rotor grooves and automobile parts [32].

It is commonly thought that MP testing is relatively a simple technique and training is usually overlooked. The consequences of such an assumption are missing of harmful defects due to improper magnetisation/demagnetisation, inaccurate calibration of equipment, inadequate illumination, inaccurate particle concentration, and misinterpretation. It is all the more essential to use Gauss meters for measurement of magnetic fields, quality indicators (shims) for controlling the field strength and verifying field direction and more importantly, the Ketos ring for establishing the detection sensitivity. Standard test blocks with one-side and two-side cracks and other test blocks are used to assess the quality MP systems [33].

4. Magnetic Flux Leakage Testing

In MFL technique, the localised magnetic leakage fields are detected using sensors such as pick up coils, Hall sensors, magnetodiodes, anisotropic magnetoresistance (AMR) sensor and GMR sensor [12]. Use of sensors in MFL testing enables automatic testing and evaluation without human inspectors. One of the primary application areas for practical MFL systems is that of axi-symmetric components i.e. tubes, pipes, rods, bars, rounds and billets. Components of these geometries are particularly amenable to fully automated and relatively high speed scanning in a production line during their manufacture. MFL technique is also widely practised for in service inspection of underground pipes in petrochemical industries, inspection of oil storage tank floors and pipes (internal/external) (ASTM E-570-97), steel wire ropes (ASTM-E-1571-01), under water structures and highly irregular components such as helicopter rotor blade D-spars, gear teeth and artillery projectiles. MFL technique can detect local corrosion within a depth of 12 mm. Unlike in the MP tests, the magnetisation levels are usually low and high strength rare earth magnets or compact electromagnets are commonly used for magnetisation in MFL technique. Since magnetisation is local, demagnetisation is usually not required [34].

In general, the MFL unit comprising of magnets and sensors is scanned at uniform speed and the sensor output is recorded continuously. The sensor output depends on the size and orientation of the defects as well as on the level of magnetisation and the inspection speed. ASTM-E-570-97 defines flux leakage examination of ferromagnetic steel tubular products. The usual procedure for characterisation involves generation of a master calibration graph between the sensor output and the size of the calibration standard defects identical to the expected defects. Advances in sensor technology and data acquisition and analysis have resulted in realisation of fully automated MFL test stations that can

achieve very high inspection speeds. MFL units can be portable, battery-powered, and compact. For detection and evaluation of corrosion damage in long oil pipelines, which run into a few hundreds of kilometres, pipe inspection gauges (PIGs) housing the MFL units are widely employed. PIGs consist of MFL unit, stand-alone battery supply, electronics and computer for data acquisition, processing and transmission to a remote log station, where evaluation is carried out [35].

In MFL technique, magnetisation can be achieved using electro-magnets or permanent magnets or by passing electric current through the component. One advantage of using an electro-magnet is that the magnetising current can be increased to cater to a wider range of component thicknesses that can not be achieved with a given permanent magnet. However, this will be at the expense of size, weight, and the use of an independent (battery) power supply. Further, when strong magnetic fields close to saturation are used, reduced sensitivity due to residual magnetization is likely. This can happen during detection of far-side corrosion. For a given magnet system, the flux density achieved in the component will depend on the thickness and permeability of the material.

Selection of sensor is important as it decides the success of MFL testing. While coils measure the rate of change in magnetic fields, Hall, GMR and AMR sensors measure actual magnetic field. Hall sensor consists of indium arsenide, indium antimony, gallium arsenide, silicon or other semiconductor materials and its main features include high linearity, small size, low power consumption and inexpensive. Micro-Hall sensors with 2-dimensional electron gas (2-DEG) GaAs with small active area and made by photolithographic technique show an order of magnitude better sensitivity and can be used to detection of buried defects [36]. The detection sensitivity of various electromagnetic sensors is given in Table 1. Though Hall sensors are undeniably more sensitive than inductive coils for measurement of leakage fields, they are too sensitive to surface conditions and hence, result in an unreliable inspection and the generation of significant false calls. Hence, for the inspection of tubes, the preferred sensor is the traditional coil due to stability and reliability.

Table 1. Sensitivity of electromagnetic field detection sensors.

Sensor	H range (Tesla)	Sensitivity (V/Tesla)	Response Time	Power Consumption	Sensor size
GMI	10^{-11}-10^{-4}	4600	1 MHz	10 mW	1-2 mm
GMR	10^{-12}-10^{-2}	120	1 MHz	10 mW	10-100 μm
Hall	10^{-6}-10^{-2}	0.65	1 MHz	10 mW	10-100 μm
SQUID	10^{-14}-10^{-6}	10^{-14}	1 MHz	10 mW	10-100 μm

Superconducting Quantum Interface Device (SQUIDs) are the world's most sensitive (~ fT) detectors of magnetic fields. The basic phenomena governing the operation of SQUID devices are flux quantization in superconducting loops and the Josephson Effect. Its application ranges from detection of magnetic fields from buried defects, material degradations such as fatigue, thermal aging, and biomedical applications. SQUIDs has been used to characterise stain induced magnetic martensite phase in cold worked stainless steels [37]. SQUIDs can be used to detect and characterise high temperature fatigue damage in stainless steel 304 welds as typically shown in Fig. 6.

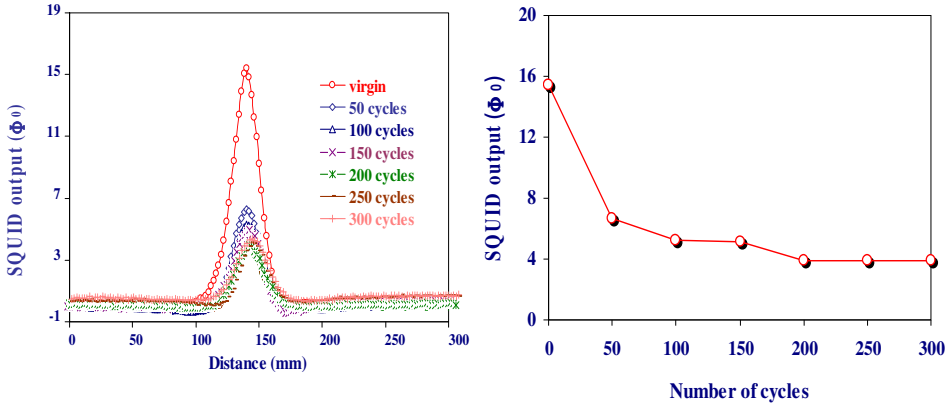


Figure 6. SQUID response from stainless steel 304 weld specimens subjected to fatigue damage.

When Hall sensors are used, normal and tangential components of the leakage field can be detected by aligning the sensing element parallel and perpendicular to the scanning surface, respectively. Typical Hall sensor response for notches located at different depths in a 12 mm thick carbon steel plate is shown in Fig. 7. For MFL applications, GMRs are very attractive. They can pick-up weak magnetic fields from far-side defects or corrosion [38].

In MFL testing, defects exhibiting various combinations of volume loss and through wall dimension can give the same amplitude signal. Therefore, truly quantitative depth information can not be obtained by this technique. Variations in inspection speed and remnant magnetisation produce eddy current noise and similarly, permeability variations and surface roughness produce another kind of noise. By employing high-pass and low-pass filters with suitable cut-off frequencies, noises can be eliminated to a maximum extent. It is interesting to note that the amplitude and shape of MFL signal from a sensor-side defect is significantly different from an equivalent other-side defect. Similarly, presence of ferromagnetic debris may produce spurious MFL signals. For such situations,

employing signal processing and pattern recognition methods such as artificial neural networks and wavelet networks ensure fast and reliable detection and characterisation of defects. Numerical modelling of MFL testing helps simulation of complex component and defect geometries, non-linear magnetic aspects and optimisation of MFL test parameters.

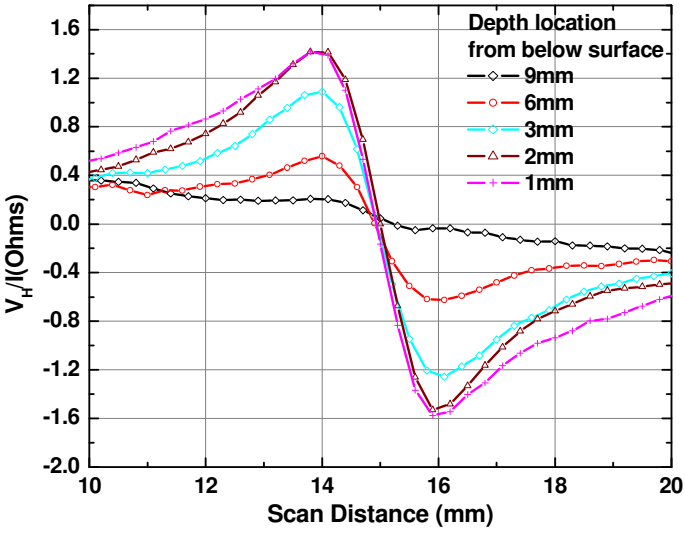


Figure 7. Typical voltage signal (tangential component of magnetic field) of Hall sensor for sub-surface notches located at different depths in a 12 mm thick carbon steel plate.

The use of MFL, visual, and ultrasonic techniques have prompted development of data fusion methods that, on one hand, greatly enhance the probability of detection (POD) and on the other hand, enable complete characterisation of discontinuities [39]. In view of this, the widely followed philosophy is employing variable threshold in MFL testing so that the most severe indications are chosen for subsequent visual or ultrasonic testing for detailed characterisation.

The electromagnetic technique suitable for inspection of very thick ferromagnetic or non-ferromagnetic components is saturation low frequency eddy current (SLOFEC) technique. In this technique, eddy current sensors are used to detect flux variations inside the component [39]. SLOFEC can also be used for inspection of thinner walls covered with thick non-metallic layers such as glass fibre reinforced epoxy coatings on floor of (oil) storage tanks. This technique can differentiate near-surface and far-side defects.

5. Micro-Magnetic Techniques

B-H loop and Barkhausen emission techniques are very useful for NDE of microstructures and residual stresses in ferromagnetic materials. Using micro-magnetic techniques, it is possible to reliably assess adequacy of heat treatment during fabrication, initial microstructures and in-service degradation in microstructures, if any, within 5 mm depth of components [40]. At room temperature, the magnetic domains of a ferromagnetic material are randomly oriented. Upon application of external magnetic field, the domains orient themselves along the magnetic field direction, which leads to an irreversible magnetization process causing hysterisis. The D.C. hysterisis characteristics are a material attribute which provides information of the magnetization of the material and hence, the microstructure [41]. The shape of D.C. hysterisis or B-H curve is sensitive to the microstructure and stress. The ratio of Flux density (B) to the applied field (H) is known as magnetic permeability (μ). Various B-H loop parameters that are used for material characterization include: Coercivity (H at B=0), Retentivity (B at H=0), area of B-H curve (hysteresis loss), harmonic characteristics, initial permeability (μ at H≈0), differential permeability (dB/dH), maximum permeability (μ_{max}) and permeability at saturation (μ at B_{max}). The hysterisis parameters can be used to characterize microstructures and to identify materials subjected to different thermo-mechanical treatments and also to identify improper material processes such as rolling, drawing, swaging and extrusion [42]. Retentivity and coercivity are affected by microstructures and precipitation behaviour substantially affects coercivity. Dislocation density generally affects the retentivity.

More detailed information regarding microstructures can be obtained by using parameters derived from the application of a varying magnetic field to the material. When a cyclic magnetization is applied on a sample, the magnetic domains align in one direction slowly and realign in the opposite direction. During this process, the random movement of the domain walls produces “Barkhausen” jumps. The smooth movement of the walls is affected, due to the pinning of the domain walls by the microscopic defects such as dislocations, grain boundaries, precipitates and voids. During the magnetization process, the slope of the B-H curve and the slope of the tangential magnetic field vary due to the variations in the microstructure. Magnetic Barkhausen Emissions (MBE) are detected using coils on the surface or sometimes encircling the component and the characteristics of MBE signal can be related to a number of different microstructural features [43-45]. Many common surface treatments such as

grinding, shot-peening, carburizing and induction hardening involve some modification of both stress and microstructure and can be readily characterised using this technique. Using MBE, it is possible to assess different microstructural features similar to X-Ray diffraction or electron microscopy. Various dynamic processes such as embrittlement, creep and fatigue similarly involve changes in stress and microstructure and can also be monitored by measuring the MBE. MBE can assess effectiveness of post weld heat treatment in welded tubes/ components given to remove the residual stresses [44].

MBE technique has been successfully used for evaluation of carburization depth in ferritic steel tubes in a petrochemical industry and also to assess fatigue damage in 9Cr-1Mo steel. MBE measurements were carried out on service exposed 0.5Cr-0.5Mo ferritic steel tubes at different depths (cross section) from the carburized ID surface, to simulate the variation in carbon concentration gradient within the case depth on MBE with increasing time of exposure of the tubes to carburization. It has been observed that the MBE level increases with increasing depth of measurement, as shown in Fig. 8. An inverse relation between MBE level and carbon content and hence, hardness has been observed. This study suggests that the MBE measurements on the carburised surface can be correlated with the concentration gradient within the case depth of the MBN which would help in predicting the appropriate depth of the carburized layer with proper prior calibration. The depth of carburized layer is a deciding factor for continued use of these tubes.

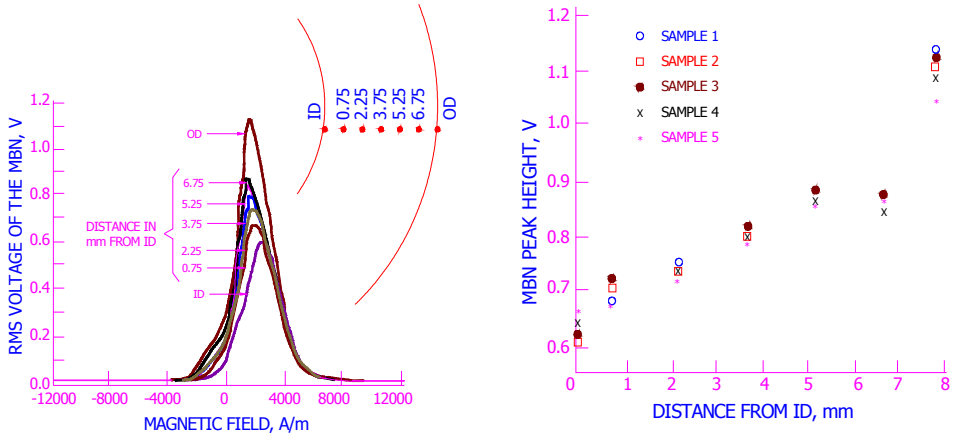


Figure 8. Application of MBE technique for detection and evaluation of carburisation in service exposed 0.5Cr-0.5Mo ferritic steel.

MBE technique can be used to identify different stages of progressive deformation. Since the magnetization process directly interacts with the dislocations, any change in the number density or their regrouping into cell formations will affect the MBE. The onset of micro yielding on ferromagnetic materials can be very easily detected. This technique can locate the area of necking even before any visible deformation occurs.

MBE technique can be used for evaluation of residual stresses. When microstructure is unchanged, the MBE output increases with tensile stress and decreases with compressive stress. A linear relationship between them exists strictly with in the elastic region. The MBE can be used to differentiate the different stages of low cycle fatigue (LCF) damage in ferromagnetic materials. When a material is subjected to progressive low cycle fatigue damage, the material undergoes different stages of deformation prior to microcrack formation. The dislocations produced by the process initially increase in number, rearrange into cell structures and this process continues till the cell saturates. Further, cycling produces micro cracking and then macro cracking. These distinct features can easily be identified by MBE.

6. Alternating Current Field Measurement (ACFM) Technique

ACFM technique is a non-contact electromagnetic technique for detection and sizing of surface cracks in metallic materials. This technique involves inducing local uniform current into a sample and measuring magnetic field above the sample surface using a sensor [46]. The conventional ACFM probe contains a field inducer and one pair of magnetic field sensors measuring the magnetic filed in two directions viz. B_x and B_z. The measure of B_x along the length of the crack is attributed to the depth of a crack and B_z gives a dipole signal produced by circulation of surface current flow and strongest at the crack ends and is an indicative of length of a crack. In order to cover large area inspection in a single scan, a linear array sensor is mounted on a probe head. Since a single field inspection is limited to a particular orientation of defects, incorporating one more field inducer will provide information about other orientations. It will be highly useful when the crack orientation is unknown. ACFM system with high scanning speed and energizing array sensor from 5 kHz to 50 kHz are commercially available.

The special feature of this technique is that it is insensitive to speed of probe and hence it is operator independent and it can operate up to 500^0C. It has an ability to inspect through coatings of several mm thick and an ability to obtain

depth information on deep crack (up to 25 mm). It is used in oil and gas industry to inspect painted, welded structures for detecting fatigue and stress corrosion cracking. This technique is applied on boilers, boiler tube sheet, pipeline and cranes in refinery. This technique has been recently extended for detecting stress corrosion cracking in stainless steel tanks used for nuclear waste storage and in gas pipelines [47]. A variant of ACFM technique is AC stress measurement technique. In this technique, the small change in the strength and direction of magnetic field is measured and related to change in stress state and dynamic load cycles [48]. It is possible to model the ACFM technique for optimization and enhanced performance [49].

7. Potential Drop Techniques

This technique is based on measurement of perturbations in electrical potential across a crack plane in an electrically conducting specimen caused by the flow of current at the crack geometry due to an applied constant direct current (DC) or alternating current (AC) current. Presence of crack in the specimen gives rise to an increase in resistance because of the reduction in the cross sectional area and hence, a potential drop [1]. Precise measurements of the output voltage are related to the crack size. Both DC and AC potential drop techniques are equally sensitive for determination of crack size and are routinely employed for dynamically measuring the crack lengths during fatigue and fracture loading. For a given current in the range of 5-50 A, the output voltage is in the range of 0.1 to 50 mV depending on the specimen dimensions and the electrical conductivity.

Analytical relationship can be derived between potential difference and crack size. Such relationships are usually expressed in terms of normalised voltage (V/V_r) and some reference crack size (a_r) where V is the measured voltage drop and V_r is the voltage drop for a reference crack. The resolution of potential drop systems i.e. the smallest change in crack size which can be distinguished in actual test operation, depends on the electrical conductivity of the specimen and it increases with decreasing conductivity. For the same applied current, the resolution is smaller in higher conducting materials and higher in lower conducing materials. Guidelines for electric potential difference based determination of crack size are provided in ASTM E 647 A6.

ACPD system is very much similar to DCPD system except the operating frequency, which is usually less than 100 Hz. At higher frequencies the skin effect plays an important role and this result in non-uniform current distribution and is advantageous in at least one sense that it provides better resolution for

surface cracks [50]. The high frequency ACPD technique is well established for crack growth measurement during dynamic loading of specimen. The potential difference measured across the crack area increases as the crack propagates or grows. High frequency ACPD technique is used for fatigue and fracture crack growth measurements of structural components such as steel tubular joints of offshore jacket structures, piping components of pressure vessels, automobile components and steel plates. The measurements are useful for arriving at the rate of fatigue crack growth, assessing the fatigue life of components, and checking for leak-before-break criteria for pressure vessels components. Analysis of the fatigue crack growth data compared with beach marks would give the information of the crack profile development.

DCPD systems are much easier to understand and use, and provide reasonable measurement sensitivity and precision. They are sensitive to thermally induced emf which can introduce significant errors into the crack length measurement [51]. On the other hand, ACPD systems provide comparable sensitivity and resolution at greatly reduced operating current as compared to DCPD systems.

Another technique based on electric potential is field signature mapping (FSM) technique. This is a non-intrusive and rugged method and can be used for monitoring localised corrosion, erosion or abrasion and cracking in steel and metal structures, pipelines, pipe bends and vessels [52]. This technique is based on feeding an electric direct current through the selected regions of the structures to be monitored, and sensing the pattern of the electrical field by measuring small potential differences generated on the surface of the monitored structure due to defects. Potential areas for corrosion are identified and the FSM sensing electrodes are mounted and the potential difference is monitored periodically / continuously. Typical defects that can be monitored and located by this technique include: general corrosion, CO_2 corrosion, localised corrosion or pitting; bacterial corrosion; cracking in welds/HAZ and erosion. FSM is attractive for continuous monitoring of corrosion in potential areas. Potential areas for corrosion are identified and the FSM sensing electrodes are mounted. The potential difference generated on the surface of the structure is monitored periodically / continuously. Selected area is fitted with a number of sensing electrodes or pins distributed in a matrix with variable spacing. Typical distance between electrodes is 2-3 times wall thickness. By proper interpretation of the changes in the potential differences, information pertaining to wall thinning of the component can be obtained reliably. In the case of installed components, the measured potentials are compared with those initially measured.

8. Microwave Techniques

Microwave NDE is based on transmitting electromagnetic waves (usually a few tens of GHz) into a dielectric structure such as concrete, and mapping the magnitude or phase of the reflected waves using a receiver to create two- or three-dimensional images of the structure under investigation [53, 54]. As the dielectric properties of the structures highly influence reflectivity of microwaves, the properties of the reflected waves would provide information about the structure under investigation. Success of microwave NDE technique is largely due to its ability to penetrate thick structures, providing good contrast in imaging of composite structures and its non-radioactive nature [55]. Two types of microwave techniques are in use namely, near field and far field techniques. In the near-field, the size of the source is significant but in the far-field, the source is small compared with exploration depth. Optimal design of antennas and arrays is essential. The depth of penetration of microwaves is inversely related to the frequency. Lower frequency would result in more depth of penetration. However, reduction in frequency results in loss of resolution and the presence of eddy current would also come into existence. The resolution achieved is also dependent on the frequency of operation. Hence, proper optimisation of depth and frequency must be obtained prior to the application of this technique. Microwave techniques are increasingly applied for non-destructive assessment of the integrity of reinforced concrete structures [56, 57].

The term ground penetrating radar (GPR) is often used to describe microwave inspection systems for locating utility lines below ground and mild steel rebar in concrete decks/pavements [58]. GPR is a pulse-echo method and uses radio waves. An antenna emits short duration pulses (of the order of nanoseconds) of electromagnetic waves. The pulses travel through the underlying material, and when the pulse encounters an interface between dissimilar materials (materials with different dielectric constants), some of the energy is reflected back towards the antenna. The antenna receives the reflected portion of the pulse and generates an output signal. By measuring the time from the start of the pulse until the reception of the echo, the depth of the interface can be determined if the propagation speed through the material is known. The amplitude of reflection of electromagnetic pulse at an interface depends on the difference between the relative dielectric constants. At a concrete/metal interface such as concrete and steel reinforcement, there is complete reflection of the incident pulse. The arrival time and strength of these echoes can also be used to calculate concrete layer thickness and other properties like moisture content. The dielectric constant of

concrete increases with increase in moisture content and salt concentration. GPR systems have been successfully employed to locate and map large voids, spalling asphalt, and delaminations in concrete decks/slabs and precast concrete elements. Microwave technique has been successfully used for detection of debonds, impact damage and core cracking in fibre reinforced plastic structures used for maritime applications. The modern commercial microwave instruments have the capability for three dimensional tomographic reconstruction of defect images. Moisture content in concrete, sand and wood can be measured within an uncertainty of 2% volume using microwave technique.

9. Electromagnetic Modelling

One advantageous feature of electromagnetic NDE techniques is the possibility to model which allows one to study complex situations, to optimise sensors and to predict signals from defects etc. without a need to carry out experiments [59]. Although, it is possible to find an analytical solution to partial differential equations (PDE) describing the electromagnetic NDE techniques, obtaining solution using numerical modelling is popular. Among other methods, finite differences (FD), finite elements (FE), volume integral (VI), boundary element (BE) and hybrid methods are widely used. Finite element method has gained popularity in MFL, EC, ACPD and microwave (high frequency or wave propagation) techniques. With the advancement in meshing algorithms, several generic commercial FEM packages such as FEMLAB, ANSYS and MAGSOFT are available. The typical vector potential formulation of the static or quasi static differential equation derived from the Maxwell's equations is

$$\nabla \times \left(\frac{1}{\mu} \nabla \times A \right) = J_o + \sigma \frac{\partial A}{\partial t}$$

where A is the vector potential and σ is the conductivity μ is permeability and J_0 is the applied current density. The time derivative will not be present for static case and it becomes jω in case of time harmonic excitation. In case of wave propagation or high frequency modelling, the wave equation involving the vector potential is considered. The above equation is solved in eddy current and magnetic flux leakage problems. In FE method, the solution domain is discretised into several elements e.g. triangles or rectangles in 2D, hexahedral or tetrahedral in 3D and solution is obtained at nodes using e.g. functional minimisation

techniques. In order to validate the models and to allow comparison between different models, benchmarks problems in EC and MFL techniques are defined by the world federation of NDE centres [60, 61]. FE modelling has been used in EC testing to optimise probes and to predict signals for inverse problems. To predict probe velocity effects in MFL testing and to study the influence of motion induced effects, FE modelling has been used [62]. Further, it is also possible to predict flux leakage patterns from pits using FE models [63].

The BE method is based on the integral formulations of the PDE under consideration. The potential advantage of this method is that only boundary (or surface) of the domain of the PDE requires meshing as compared to the whole domain in FE method [64]. The field interactions due to sources enclosed by a domain can be modelled efficiently using this method; hence this method may not be applicable to all kinds of situations when sources are present outside or very near to the solution domain. Although, hybrid methods that use FE and BE methods are attractive [65], volume integral based models e.g. CIVA, developed at CEA, France appears more efficient, accurate and fast and are being used for a variety of eddy current and MFL applications [66, 67].

Conclusion

Electromagnetic NDE techniques find a lot of applications in engineering industry. Rapid advances are taking place in these techniques making use of the progress in allied fields of microelectronics, sensors, signal processing, image processing, artificial intelligence, computers and robotics. Portable and integrated systems having modules of several NDE techniques are becoming popular for field applications. Such systems will enable data fusion for enhanced detection of defects. Numerical modelling has enhanced the usefulness of the techniques by way of better understanding of field-defect interactions and optimisation of sensors as well as instrument parameters. Availability of miniature, smart and flexible solid state sensors is enabling inspection of complex geometries and inaccessible regions. Grater emphasis is being placed on development of array sensors and automated testing for quick and reliable detection and sizing of shallow as well as buried defects. Design and development of microwave antennas holds the key to rapid imaging of concrete structures. Eddy current and micro-magnetic techniques have established capability for quantitative characterisation of microstructures, accumulated deformation and stress, prior to crack formation. Such possibilities allow efficient preventive actions to be taken avoiding catastrophic failures of materials and components.

Acknowledgements

Authors sincerely thank Mr. S. Thirunavukkarasu, Mrs. B. Sasi, Mr. W. Sharatchandra Singh, Mr. S. Vaidyanathan and many other colleagues of NDE Division, India Gandhi Center for Atomic Research (IGCAR), Kalpakkam, India for their help during the preparation of this chapter.

References

1. Baldev Raj, T. Jayakumar and B.P.C. Rao, "Review of NDT techniques for structural integrity", Sadhana, Academy Proc. in Engineering Sciences, Vol. 20, 1995, pp. 5-38.
2. Grandt, A.F. Jr., Fundamentals of Structural Integrity: Damage Tolerant Design and Non-destructive Evaluation, John Wiley & Sons, Inc. Hoboken, NJ, 2004.
3. Baldev Raj, Nondestructive evaluation – An overview, Encyclopedia of Materials: Science & Technology, Elsevier Science Publishers, 2001.
4. H.L. Libby, "Introduction to Electromagnetic Non-destructive Test Methods", Wiley-Interscience, New York, 1971.
5. J. Blitz, "Electrical and Magnetic Methods for Non-destructive Testing", Adam Hilger, Bristol, 1991.
6. W. Lord, "Electromagnetic methods of Non-destructive Testing", Gordon and Breach, New York, 1985.
7. Moore Patrik, Udpa S.S, Non-destructive testing handbook. 3rd edition. Vol. 5: Electromagnetic testing, ASNT, Columbus, OH, USA, 2004.
8. V.S. Cecco V.S, G. Van Drunnen and F.L. Sharp, "Eddy current manual: test method", Vol. 1, AECL-7523, Chalk River, Ontario, Nov., 1983.
9. R.C. McMaster, P. McIntire and M.L. Mester (Eds.), "Non-destructive Testing Handbook: Electromagnetic testing", Vol. 4, ASNT, Columbus, OH, USA, 1986.
10. B.P.C. Rao, Introduction to Eddy Current Testing, Narosa Publishing, New Delhi, India, February, 2007.
11. D.J. Hagemaier, "Fundamentals of Eddy Current Testing", ASNT, Columbus, OH, USA, 1990.
12. B.P.C. Rao and T. Jayakumar, Discontinuity characterisation using electromagnetic methods, J of Non-Destructive Testing & Evaluation, Vol. 2, No. 2, 2002, pp. 23-29.
13. T. Stepinski, "Digital Processing of Eddy Current Signals and Images", Proc. of 6th European Conf. on NDT, Nice, October 1994, pp. 51-55.
14. B.P.C. Rao, Baldev Raj, T. Jayakumar and P. Kalyanasundaram, "Using artificial neural networks to quantify discontinuities in eddy current testing", Materials Evaluation, Vol. 60, 2002, pp. 84-88.
15. ASTM-STP-722 Eddy Current Characterization of Materials and Structures, ASTM Committee E-7, ASNT, and National Bureau of Standards, 1981.
16. S. Vaidyanathan, B.P.C. Rao, T. Jayakumar and Baldev Raj Development of remote field eddy current testing, Part III of Development in instrumentation and automation for NDE applications: in-house experience in the Insight, Vol. 45, No.1 , January 2003, pp. 73-76.

17. M. Gramz and T. Stepinski, "Eddy Current Imaging, Array Sensors and Flaw Reconstruction", Research in Nondestructive Evaluation, Vol. 5, No. 3, 1994, pp. 157-174.
18. B.P.C. Rao, Baldev Raj, T. Jayakumar and P. Kalyanasundaram, "An intelligent imaging scheme for automated eddy current testing", Non-destructive Testing & Evaluation, Vol. 17, 2001, pp. 41-57.
19. B.P.C. Rao, Baldev Raj, T. Jayakumar, P. Kalyanasundaram and W.Arnold, "A New Approach for Restoration of Eddy Current Images", Journal of Non-destructive Evaluation, Vol. 20, 2001, pp. 61-72.
20. Sandra Simmons, Magneto-optic imaging, Materials Evaluation, Vol. 51, 1993, pp 529-534.
21. I. Elshafiey and L Udpa, Sensitivity Analysis of Magneto-Optic Imaging in Non-destructive Evaluation of Pipelines, Proc. 3rd Middle East NDT Conference and Exhibition, Nov 2005 - Bahrain, Manama.
22. H. Shaik, N. Sivaibharasi, B. Sasi, T. Anita, R. Amirthalingam, B.P.C. Rao, T. Jayakumar, H.S. Khatak and Baldev Raj, "Use of eddy current testing method in detection and evaluation of sensitisation and intergranular corrosion in austenitic stainless steels", Corrosion Science, Vol. 48, 2006, pp.1462-1482.
23. B.P.C. Rao, T. Jayakumar, B. Sasi, Anish Kumar, K. Rajkumar and Baldev Raj, "Characterisation of microstructures in heat treated Titanium alloys and Maraging steels using eddy currents", Proc. of Rev. of Prog. in QNDE Conf., USA, 2006 (in press).
24. K.V. Rajkumar, B.PC. Rao, B. Sasi, Anish Kumar, T. Jayakumar, Baldev Raj and K.K. Ray, "Characterization of aging behaviour in M250 grade Maraging steel using eddy current non-destructive methodology". Journal of Material Science Engineering, 2007 (in press).
25. T. Uchimoto, T. Takagi, S. Konoplyuk, T. Abe, H. Huang and M. Kurosawa, Eddy current evaluation of cast irons for material characterization, Journal of Magnetism and Magnetic Materials, Volumes 258-259 , March 2003, pp. 493-496.
26. B.A. Lepine, B.P. Wallace, D.S. Forsyth and A. Wyglinshi, Pulsed Eddy Current Method, Proc. of 1st PACNDT Conf., Toronto, Sept. 1998 (NDT.Net, Jan. 1999, Vol. 4, No.1).
27. J.C. Moulder, M.W. Kubovich, E. Uzal and J.H. Rose, Pulsed Eddy Current Measurements of Corrosion-Induced Metal Loss: Theory and Experiment, Rev. of Prog. in QNDE, Vol. 14, eds. D.O. Thompson and D.E. Chimenti, Plenum Press, New York, 1995, pp. 2065-2072.
28. Gary W. Carriveau and R. Austin, Nondestructive Evaluation of Corrosion Through Insulation, Review of Progress in Quantitative NDE, Eds. D.O. Thomson and D.E. Chimenti, Plenum Press, New York, 1996, Vol. 15B, pp. 1741-1746.
29. Stevens, D.M. and Whaley, H.L., Non Destructive Evaluation (ASME), Vol. 54, 1993, pp. 43-47.
30. D.J. Lovejoy, "Magnetic Particle Inspection: A practical guide", Chapman & Hall, 1993.
31. Betz, C.E., Principles of Magnetic Particle Testing, Magnaflux, Hardwood Heights, Illinois, 1997.
32. J.S. Borucki, "Development of Automated Magnetic Particle Testing Systems", Materials Evaluation, Vol. 49, 1991, pp 324-329.
33. V. Schuster, V. Deutsch, W.A.K. Deutsch, F. Bartholomai, P. Müller, Fluxa-Control: an Automated Monitoring of Fluorescent Particles in Magnetic Particle Testing, Proc. of ECNDT, 2006, paper We.2.6.1.
34. Dobmann, G., Walle, G., and Höller, P., "Magnetic leakage flux testing with probes: physical principles and restrictions for application," NDT International, Vol. 20, No. 2, 1987.

35. L. Clapham and D.L. Atherton, "Magnetic flux leakage inspection of oil and gas pipelines", J. of Non-destructive Evaluation (ISNT), Vol. 20, 2000, pp. 40-45.
36. Ch. Ravi Kumar, T.S. Abhilash, G. Rajaram, B.P.C. Rao, S. Rita, K.A. Gopal, T. Jayakumar, M.P. Janwadkar, P.V. Kumar and Baldev Raj, Magnetic Flux Leakage Measurement using 2DEG based Micro-Hall Sensors, 50th DAE Solid State Physics Symposium, BARC, Mumbai, Dec. 2005.
37. B.P.C. Rao, T. Jayakumar, Baldev Raj, and W. Arnold, Non-destructive characterisation of martensite In AISI Type 304 stainless steel using SQUID and MBN methods, J of Non-destructive Evaluation, Vol. 21, No.1, March 2001, pp 38-43.
38. L. Chen, P.-W. Que and T. Jin, A Giant-Magnetoresistance Sensor for Magnetic-Flux-Leakage Non-destructive Testing of a Pipeline, Russian Journal of Non destructive Testing, Vol. 41, No. 7, 2005, pp. 462-465.
39. X.E.Gross, "NDT Data Fusion", Arnold, London, 1997.
40. M.J. Cohn, J.A. de Raad, Non intrusive inspection for flow accelerated corrosion detection, Proc. of ASME Pressure Vessels and Piping Conference, July 1997, Orlando, Florida.
41. V. Moorthy, S. Vaidyanathan, Baldev Raj, T. Jayakumar, B.P. Kashyap, Insight into the Microstructural Characterisation of Ferritic Steels using Micro-Magnetic Parameters, Metall. Mater. Trans., Vol. 31A, 2000, pp. 1053-1065.
42. D.C. Jiles, "Review of magnetic methods for nondestructive evaluation (Part 2)," NDT International, Vol. 23, No. 2, 1990.
43. D.C. Jiles, "Introduction to magnetism and magnetic materials," Chapman and Hall Publishers, London and New York, 1998.
44. Baldev Raj, T. Jayakumar, V. Moorthy and S. Vaidyanathan, Characterisation of microstructures, deformation, and fatigue damage in different steels using Magnetic Barkhausen Emission technique, Russian journal of nondestructive testing, Vol. 37, No. 11, 2001, pp. 789-798.
45. Lo, C.C.H.; Lee, S.J.; Kerdus, L.C.; Jiles, D.C, "Examination of the relationship between the parameters of Barkhausen effect model and microstructure of magnetic materials", Journal of Applied Physics, Vol. 91, Issue 10, 2002, pp. 7651-7653.
46. M. Lewis, D.H. Michael, M.C. Lugg and R. Collins, "Thin-skin electromagnetic fields around surface-breaking crack in metals", Journal Applied Physics, Vol. 64, No. 8, 1988, pp. 3777-3784.
47. David Topp, "Quantitative in-service inspection using the alternating current field measurement (ACFM) method", Proc. of International conference on NDT contribution to infrastructure safety systems, Nov-22-26, 1999, Torres.
48. Martin Lugg and David Topp, Recent Developments and Applications of the ACFM Inspection Method and ACSM Stress Measurement Method, Proc. 9th European Conf. on NDT, Sept. 2006, Berlin, Germany, pp. Tu.3.6.5.
49. J. Zhou, M.C. Lugg and R. Collins, "A nonuniform model for alternating current field measurement of fatigue cracks in metals", Int. J. Appl. Electromag. Mech., Vol. 10, 1999, pp. 221-235.
50. R. Collins, W.D. Dover, and D.H. Michael, Potential drop techniques, Nondestructive Testing, edited by R.S. Sharpe, Academic, New York, 1985, Chap. 5.
51. W.D. Dover, R. Collins, and D.H. Michael, Use of potential drop technique, British J. NDT. Vol. 33, 1991, p. 121.

52. Roe Strommen et al., FSM (Field Signature Method); The New Technology for Internal Corrosion Monitoring of Pipelines, Vessels and Pressure Equipment, Proc. of ASME Energy Sources Technology Conference, Houston Texas, Feb. 2-4, 1998, XP00105536.
53. A.J. Bahr, Microwave Nondestructive Testing Methods, Vol. 1, Gordon and Breach, New York, 1982, Chap. 4.
54. A.J. Bahr, Experimental Techniques in Microwave NDE, Review of Prog. in QNDE, Volume: 14A, 1995, pp 593-600.
55. Zoughi, R., Microwave Non-Destructive Testing and Evaluation, Kluwer Academic Publishers, 2000.
56. A. Kavitha, M. Vikram, R.Edward, Udpa L, and Udpa Satish S, Microwave NDE for Reinforced Concrete, QNDE, AIP Conference Proceedings, Voln 820, 2006, pp. 455-460.
57. P. Shaw and A. Xu, High Energy Radiography and Radar Applications to Concrete Inspection, Proc. NDT in Civil Engineering Conference, Liverpool, 1997.
58. D. Schaurich, K. Tulla, H. Wiggnehauser, "Moisture measurements in builing materials with microwave", H. Kaariainen, M. Rudolph, NDT&E international, Vol. 34, 2001, pp. 389-394.
59. N. Ida, "Numerical modelling for electromagnetic non-destructive evaluation", Chapman & Hall, London, 1995.
60. Y. Tian, Y. Li, L. Udpa, and S.S. Udpa, "Simulation of the world federation's Second eddy current benchmark problem", Review of progress in QNDE Vol. 22. AIP Conference Proceedings, Vol. 657, 2003, pp. 1816-1823.
61. Zeng Zhiwei, Tian Yong, Udpa S and Udpa L, Finite Element Modeling of the World Federation's Second MFL Benchmark Problem, Review of progress in QNDE Vol. 22. AIP Conference Proceedings, Vol.700, 2004, pp. 1553-1559.
62. Y.K. Shin and W. Lord, "Numerical Modeling of Probe Velocity Effects for Eddy Current and Flux Leakage NDE", Rev. of Prog. in QNDE, Vol. 10, 1991, p. 921.
63. Atherton, D.L., "Finite Element Calculations and Computer Measurements of Magnetic Flux Leakage Patterns for Pits," British Journal of Non-destructive Testing, Vol. 30, No. 3, 1988, pp. 159-162.
64. N. Nakagawa and B.P.C. Rao, A study of boundary element eddy current model validation, Review of Prog in Quantitative NDE, Vol. 24, ed. by D.O. Thompson and D.E. Chimenti, American Institute of Physics, New York, Vol. 24A, 2004, pp 432-439.
65. Yann Le Bihan, Jozsef Pavo, Claude Marchand, "Study and experimental validation of the calculation of the ECT signal induced by a minute crack using a FEM–BIM combination", NDT&E International Vol. 39, 2006, pp. 476-486.
66. G. Pichenot, F. Buvat, D. Premel, and T. Sollier, Eddy Current Modelling for NDT, Proc. 16th World Conf. on NDT, Montreal, paper 338 (http://www.ultrasonic.de/abstract/wcndt2004/338.htm)
67. Fabrice Foucher, Cedrat SA, Laurent Le Ber, Steve Mahaut, "Simulation Software of NDT Techniques", NDT.net, Vol. 11, No. 6, June 2006.

This page intentionally left blank

CHAPTER 12

TERAHERTZ NDE FOR AEROSPACE APPLICATIONS

Robert F. Anastasi[1], Eric I. Madaras[2], Jeffrey P. Seebo[3], and William P. Winfree[2]

[1]*U.S. Army Research Laboratory, Vehicle Technology Directorate*
AMSRD-ARL-VT-SM, Nondestructive Evaluation Sciences Branch
NASA Langley Research Center, MS 231, Hampton, VA 23681
[2]*NASA Langley Research Center, Nondestructive Evaluation Sciences Branch*
MS 231, Hampton, VA 23681
[3]*Lockheed Martin, NASA Langley Research Center, MS 231, Hampton, VA 23681*

Pulsed terahertz nondestructive evaluation is an emerging technology with significant potential for inspecting nonconductive aerospace materials for hidden flaws and damage. This new technology is potentially useful for inspections that have previously been identified as either extremely difficult or impossible with current inspection technologies. Characteristic of a difficult-to-inspect material is the Sprayed-On Foam Insulation of the Space Shuttle External Tank. Recent measurements demonstrate that a terahertz nondestructive evaluation system is able to detect voids and unbonds in this material. Other viable applications of pulsed terahertz technology include examination of metallic surface roughness, measurement of paint thickness, and corrosion detection under Space Shuttle tiles. This chapter discusses the development of a measurement system for performing these inspections and the results of measurements on a variety of materials and structures.

1. Introduction

Aerospace vehicles often involve the use of novel materials and structures in high performance applications. The ensured integrity of these materials and structures is often critical for safe and reliable mission performance. The uniqueness of the materials and structures often requires advanced nondestructive techniques to verify the absence of critical flaws. As a result, emerging technologies are often considered for inspection of aerospace materials and structures.

One emerging technology that has demonstrated applicability for aerospace materials and structures is pulse terahertz electromagnetic waves. The term

terahertz (THz) applies to the electromagnetic wave energy that falls between the frequency ranges of 300 GHz to 3 THz with a free space wavelength range of 1000 to 100 μm [1]. This is a region of the electromagnetic spectrum between the microwave and infrared bands. Until recently short pulse sources and detectors were nonexistent in this range, limiting their application as a viable nondestructive evaluation (NDE) technique.

The technology that enables the NDE applications to be discussed here possible was pioneered in the late 1980's by groups at AT&T Bell Laboratories and IBM's T.J. Watson Research Center. They generated THz pulses using photoconductive antennas illuminated with femtosecond optical pulses and detected the THz when the antenna was gated with the optical excitation [2-6]. Some THz uses have focused on spectroscopic applications in astronomy [2], security application for chemical and biological and weapons detection [7-11], medical applications [12, 13] and more recently NDE applications.

NASA's initial application of pulsed THz NDE was detection of flaws in the thermal protection system (TPS) of the shuttle external tank. The Columbia Accident Investigation Board recommended NASA initiate a program to eliminate all External Tank Thermal Protection System debris shedding at the source and place a particular emphasis on the region where the bipod struts attach to the tank [14]. This recommendation resulted in numerous Space Shuttle Return-to-Flight initiatives to understand and eliminate the debris-shedding problem [15]. One part of this effort included development of NDE techniques for detection of critical flaws in the External Tank's Sprayed-on Foam Insulation (SOFI). A survey of advanced NDE technologies was performed, and pulsed THz was selected for implementation on the tank inspections [16]. In this chapter, pulsed THz inspection applications for the aerospace structures are reviewed. These applications include inspection of the SOFI of the Space Shuttle External Tank, examination of metallic surface roughness, measurement of paint thickness, and feasibility of inspecting for corrosion under Space Shuttle tiles.

2. THz System

The pulsed THz system used in this work has been described in the detail elsewhere [17] and was initially designed to inspect up to 30 cm of Sprayed-On Foam Insulation. This system was integrated with a spatial scanner, data acquisition, and control system that translated the THz transceiver along two dimensions, and acquired and stored the temporal waveforms at incremental scan points. The system, illustrated in Figure 1, consists of an 800 nm femtosecond laser source with a pulse rate of 75 MHz, a THz control unit with a mechanical

scanning optical delay line, a transceiver head with transmitter, receiver, and lens for focusing the THz beam, and a computer. Fiber optics connected the laser source to the transceiver and allowed for a laser-light contained system with a freely movable transmitter and receiver.

In this system, a short laser pulse is used in combination with a specially grown GaAs crystal for both THz generation and detection. When the laser pulse impinges upon the crystal, electrons are rapidly excited into conduction bands and a short burst of THz radiation is generated. The radiation is focused and transmitted into a test object being inspected. The radiation reflected from the test object is refocused to strike a single crystal GaAs receiver. For this configuration, part of the generating laser pulse is directed to the receiving crystal. The laser pulse gates on the receiver producing a voltage potential that is proportional to the THz radiation that is present at the receiver during the laser pulse. By controlling the path of the laser pulse with a scanning optical delay line, the THz radiation at the receiver is sampled to create a voltage-time record of the reflected THz signal.

The initial configuration of the transmitter and receiver for this system was the pitch-catch mode illustrated in Figure 1. This was reconfigured to the confocal arrangement shown in Figure 2. This confocal arrangement used a 15 cm diameter, 285 mm focal length lens to concentrate the THz energy on the sample to a spot diameter of approximately 3 mm. A typical THz signal reflected from an aluminum plate and its spectrum is shown in Figure 3. The digitized time-domain signal is composed of 2048 points and has a duration of 320 ps, which corresponds to a propagation path length in air of 96 mm.

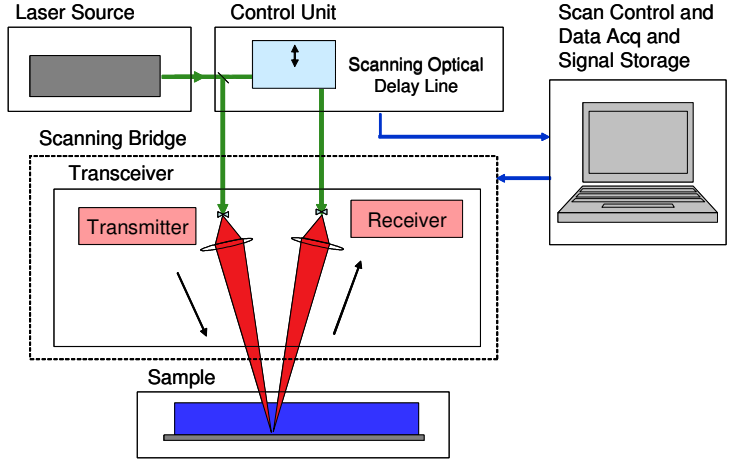
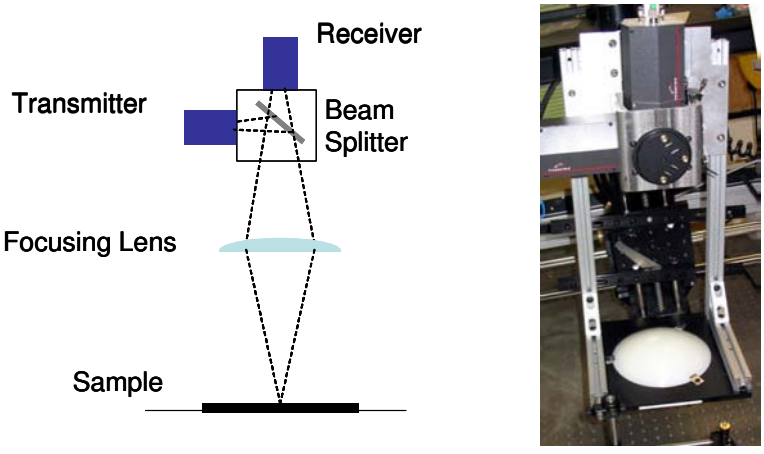
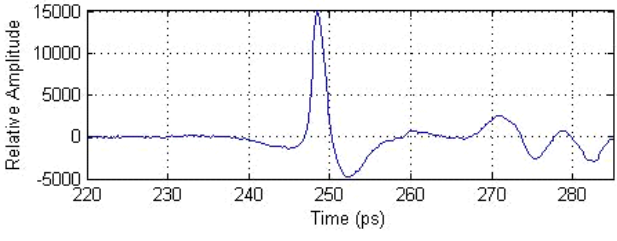


Figure 1. System schematic.

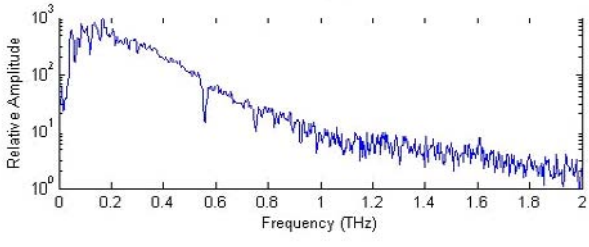


(a) (b)

Figure 2. Transceiver confocal arrangement (a) schematic and (b) photograph.



(a)



(b)

Figure 3. A typical THz (a) signal reflected from an aluminum plate and its (b) spectrum is shown.

The beam spread from the transmitter was measured by placing a variable aperture between the THz-transmitter and focusing lens. The aperture opening and distance from the source was varied until maximum signal output was obtained. Using the source distance and aperture diameter a full angle beam

spread of 27^{o} was calculated. The beam diameter at the focusing lens was approximately 9.5 cm, thus the system f/number was approximately three and the Rayleigh limit of resolution was approximately 1 mm at a frequency of 1 THz.

3. Sprayed-On Foam Insulation (SOFI)

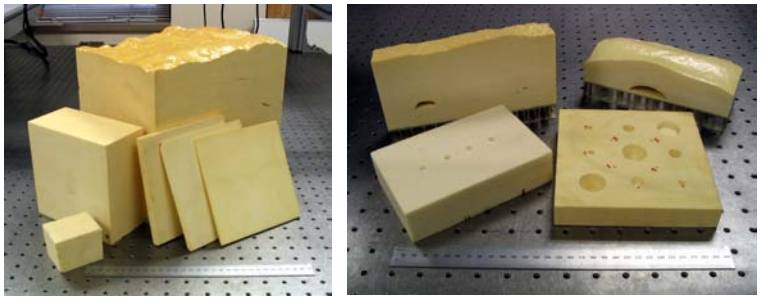
The foam insulation on the Space Shuttle External Tank is the primary thermal protection system for the external tank. Foam is automatically sprayed on the tank on the '*acreage*' areas, the large areas with no obstructions and is '*hand-sprayed*' in various areas where structures or fittings are attached. The spraying is done in layers to build up to the desired thickness. The foam is a type of polyurethane composed of five primary ingredients: polymeric isocyanate or isocyanurate, a flame retardant, a surfactant, a blowing agent, and a catalyst. The blowing agent creates the foam's cellular structure by making millions of tiny bubbles or foam cells [18, 19]. The closed cell structure required for the foam's low thermal conductivity results in a material that is difficult to inspect with conventional one-sided NDE methods such as ultrasonics and thermography. In this section, an overview of THz inspection of foam is presented.

3.1. *Foam Specimens*

To determine the capabilities of THz inspection, foam specimens with and without defects were manufactured and examined. The foam component of these specimens was similar to the primary Shuttle External Tank foam, fabricated in blocks of various sizes. Thicknesses ranged from approximately a centimeter to twenty centimeters and width and depths ranging from tens to hundreds of centimeters. Some of the configurations are shown in Figures 4 and 5. Figure 4 shows samples of various size and with flat-bottom holes and manufactured voids and Figure 5 shows a model of the Inter-tank Region from the External Tank.

3.2. *Foam Inspection Results and Images*

Initial inspections were performed with a THz transceiver in a pitch-catch arrangement and subsequent inspections performed with the previously described confocal arrangement. Initial inspections examined the foam index of refraction, air-gap detection, and unbond and void detection. Later inspections examined samples with flat-bottom holes, samples with manufactured defects, and a modeled External Tank sample.

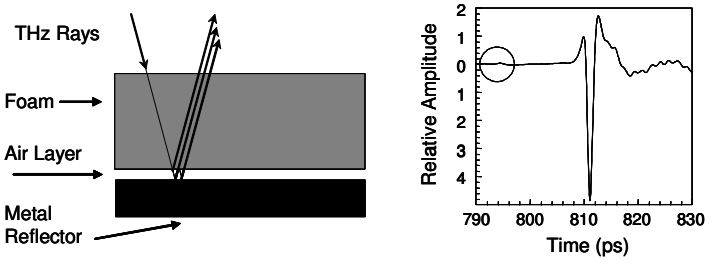


(a) (b)

Figure 4. Foam samples (a) of various size and (b) machined flat-bottom holes and manufactured voids.

(a) (b)

Figure 5. Model of the Inter-tank region of the External Tank showing (a) the substrate structure with stringers and inserts and (b) foam covered sample.



(a) (b)

Figure 6. Air-gap measurements (a) schematic outlining the reflections from the foam-air and the air-metal interfaces and (b) THz signal from a foam block situated approximately 5.7 mm above a metal reflector. The circled part of the signal is from the foam-air interface. The large signal is the air-metal interface reflection.

The index of refraction of the foam calculated from time-of-flight measurements varied between 1.02 and 1.05. This value is close to the index of refraction of air. To characterize the response from an air gap, a foam sample was held above a metallic reflector as illustrated in Figure 6a. This figure also indicates the first few reflections at the air-foam interface and the air-metal interface. A THz signal for a gap of 5.7 mm is shown in Figure 6b where the circled region represents the location of the foam-air interface. By expanding this region of interest, the air-foam interface can be clearly seen as shown in Figure 7. This illustrates the temporal resolving power of the equipment.

The initial specimens for imitating unbonded regions were fabricated with Teflon inserts at the foam-to-metal interface. THz images in Figure 8 are from two foam samples each 20 cm thick. One sample has a 5 by 5 cm insert and the other a 10 by 10 cm insert. Figure 8a shows the 5 cm insert and Figure 8b shows a corner of the 10 cm insert. These measurement and others indicated the Teflon inserts were being detected rather than air gaps at the interface. To produce more realistic unbonds with air gaps, a thin square piece of cured foam was bonded at its corners, with small dots of glue, to the metal substrate before more sprayed-on foam was applied to the specimen. The bonding helped hold the thin foam in place while the sample was built up with more sprayed-on foam that would adhere to the substrate except were the inserts was attached thus producing a more realist unbond. The THz response for these flaws are shown in the Figure 10, where the unbonds are the upper row of indications in Figures 10a and 10b. There are two significant results illustrated in the THz images of these fabricated unbonds. The first result indicates how the sides of the unbond are detected. This detection results from the interference of THz waves propagating through the foam only and through the unbond or air gap region. The second result shows the response of the glue dots at the corners of the thin foam insert. The glue used was not transparent to THz. These responses demonstrate the difficulty of creating realistic flaws in foam specimens.

Another flaw of interest is large voids in the foam. Voids were manufactured in samples using a cylindrical foam inserts with the centers hollowed out. Figure 9, shows THz images of a cylindrical void with a radius of 1.25 cm and height of 1.25 cm. Figure 9a shows a time-of-flight image of the THz response due to the presence of the void. The light region in the images is indicative of the shorter time of flight for the air as compared to the surrounding foam. Figure 9b shows a peak detected amplitude image of the THz response due to the presence of the void, with the edge of the void outlined by the interference effect described above.

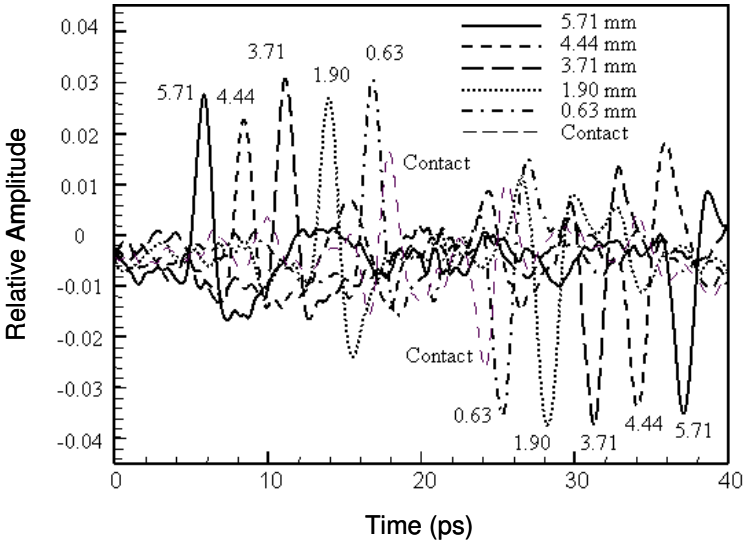
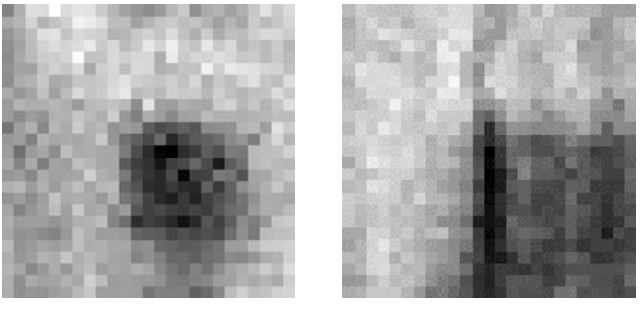
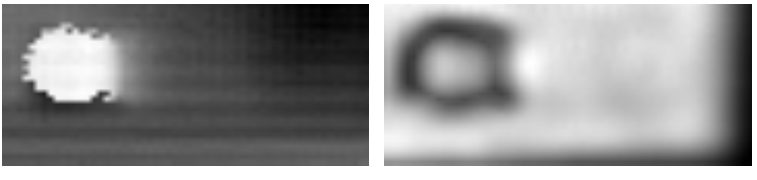


Figure 7. First group of reflections is from the foam-air interface with the metal reflection removed. The second group of reflections is a multiple reverberation from the air-foam interface and has the opposite polarity.



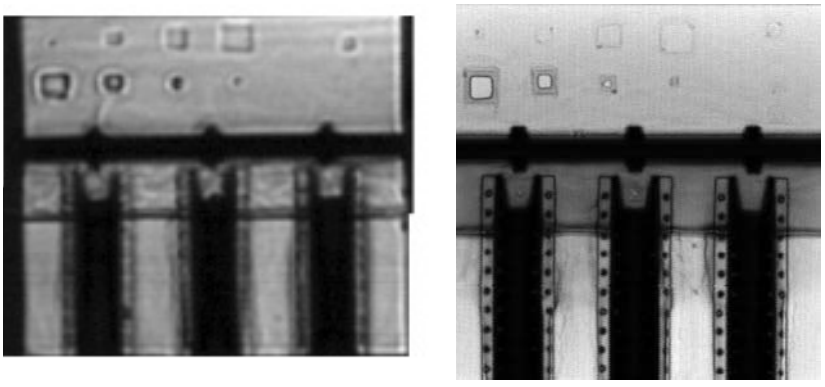
(a) (b)

Figure 8. THz amplitude images of Teflon inserts made through 20 cm of foam. The images are 12.7 by 12.7 cm (a) image of the 5 by 5 cm insert and (b) image of the corner of the 10 by 10 cm insert.



(a) (b)

Figure 9. Terahertz images of a void (a) time-of-flight image and (b) amplitude image.



(a) (b)

Figure 10. THz images of the model of the Inter-tank region (a) an early image and (b) a later image using the confocal transceiver with improved detail of manufactured defects and subsurface structure.

A specimen representing the External Tank Inter-tank region was manufactured with void and unbond-inserts of various size and placements in the sample. A photograph of the sample was shown in Figure 5. Figure 10a shows a THz amplitude image acquired with a collinear head with a relatively small aperture. Figure 10b shows a THz amplitude image acquired with the larger aperture confocal transceiver. As can be seen in these figures the larger aperture confocal transceiver provided improved imaging of manufactured defects. The larger aperture produces an improved focus and thus spatial resolution.

4. Metallic Surface Roughness Evaluation

Metallic surface roughness in a nominally smooth surface is a potential indication of material degradation or damage. When the metallic surface is coated or covered with an opaque dielectric material, such as paint or insulation, then visual inspection for changes in surface roughness as a result of mechanisms such as corrosion becomes impossible. THz is able to penetrate the dielectric coating and inspect the metallic surface. In this section, variation in the THz response as a function of surface roughness is examined.

4.1. *Surface Roughness*

Roughness is a measure of the relative elevations of surface irregularities that result from various sources such as machining processes, surface damage, and

corrosion. In machining applications, a machinist will compare his work piece to a representative sample or standard by running his finger over the work piece and sample. The drawback of this method is that it is qualitative with different machinists arriving at different conclusions for the same work piece. A contact surface-tracing machine is a quantitative method used to measure roughness by pulling a stylus across the surface. Non-contact optical methods measure the scattering of radiation by the surface to characterize surface roughness and yield statistical surface parameters [20, 21]. One instrument for non-contact surface roughness measurement is a Total Integrated Scattering (TIS) instrument. This instrument measures the specular and diffusely scattered laser light from a surface and relates these measurements to the root-mean-squared (rms) surface roughness. This general relationship is [21]

$$TIS = \frac{R_s}{R_0} = e^{-(4\pi\delta/\lambda)^2} \tag{1}$$

where R_s is the specular reflectance, R_0 is the total of the specular reflection and the diffuse reflectance, λ is the wavelength of light, and δ is the rms surface roughness. In this relation, it is assumed that the light is incident normal to the surface and monochromatic.

4.2. *Surface Roughness Measurement Specimens*

The test specimens used for this work were commercially available comparator plates made of electroformed nickel that serve as standards for a machinist to evaluate his work. Two different comparator plates were measured: a 'Planing and Shaping' plate and a 'Casting' plate. The overall dimension of each plate was 5.5 cm x 12.5 cm and each roughness area was approximately 16 mm x 25 mm. These plates covered a range of average roughness (Ra) from 0.04 μm to 100 μm, where the average roughness is defined as the area between the roughness profile and its mean line, or the integral of the absolute value of the roughness profile height over the evaluation length [22]. The plates are shown in Figure 11. The 'Planing and Shaping' plate, Figure 11a, has eight separate areas of roughness. The smoothest area has an average roughness of 0.8 μm and the roughest is 100 μm. The roughness areas on this plate appear as parallel rulings. For the 100 μm, 50 μm, 25 μm, and 12.5 μm areas the ruling spacing is 3.2 mm, 1.6 mm, 0.8 mm and 0.4 mm respectively. The 'Casting' plate, Figure 11b, has seven separate areas of roughness that have a range of average roughness from

50 μm to 0.8 μm. These roughness areas appear as randomly rough surfaces. The average roughness parameter (Ra) has an approximate relation to the rms roughness (δ) used in equation (1), [20].

$$Ra \cong 0.8\,\delta \tag{2}$$

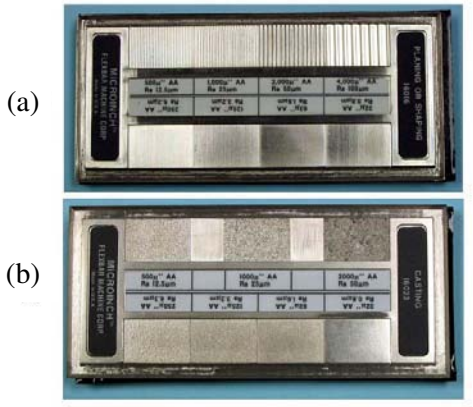


Figure 11. The surface roughness comparator plates (a) 'Planing or Shaping' plate and (b) 'Casting' plate.

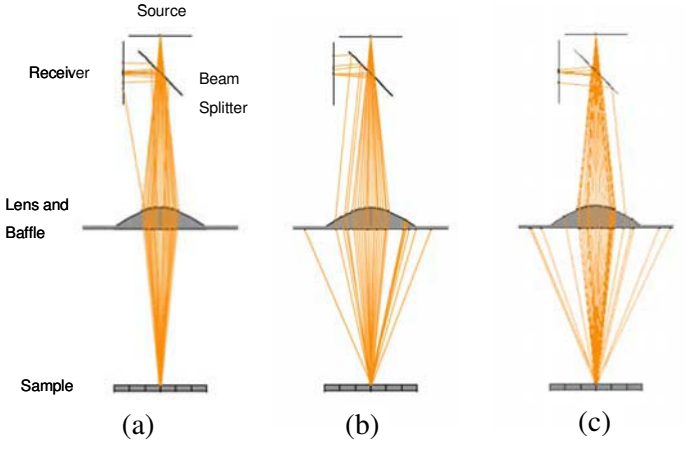


Figure 12. Ray tracing model examples with estimated surface roughness of (a) 0 μm, (b) 96 μm, and (c) 149 μm.

4.3. *Optical Ray Tracing Model*

To gain a better understanding of how the THz system's optical configuration might influence roughness measurement, an optical ray-tracing model was

developed and its results were compared to the experimental measurements. In this simulation, the sample was modeled as a mirror with its surface modified using a function that adds surface roughness. Figure 12 shows three ray tracing results for estimated rms values of 0 μm, 96 μm, and 149 μm. This figure shows that as roughness increases, ray scattering increases. However, Figure 12a shows some rays deviate from the predicted path. This appeared to be a function of the aspherical lens surface defined in the ray-tracing program because ray deviation does not occur when a convex-planar lens was used in place of the aspherical lens. To compensate for this effect, a large number of rays were used to average out any ray tracing errors. To quantify the modeling results the number of rays at the receiver was used as a measure of received specular signal intensity.

4.4. *Measurements and Results*

The samples were inspected with the THz system and the scan data was stored for post processing. To visualize the data, amplitude and frequency images were generated. The amplitude image was generated by plotting the maximum amplitude of each signal as a function of location and the frequency image was generated by calculating the power spectrum of each signal in the THz scan data and then taking the amplitude of the spectrum at 0.4 THz. The image at this frequency exhibited the best contrast between the different roughness areas. Figure 13 shows a gray scale representation of the data acquired on the specimens. For surface roughness below 12.5 μm Ra, variations in surface roughness became indistinguishable.

To quantify the scan data, maximum signal amplitudes for each roughness area on the comparator plates were examined. TIS measurements were obtained by taking the ratio between the specular and total reflection. The specular reflection for each roughness area was calculated by averaging all the signals in each roughness area and then taking its maximum amplitude as the total specular reflection. For the total reflection, the maximum amplitude of the smoothest roughness area, that resembled a polished surface, was used. In practice, a separate polished standard could be used. The standard deviation of this average per roughness area was used to estimate a TIS error. A plot of this TIS value as a function of roughness for each of the comparator plates is shown in Figure 14 along with calculated TIS values from equation (1) and TIS values calculated from the ray tracing model simulations. In the model simulations and calculations, a frequency of 0.4 THz was used.

The measurements, calculations, and modeling results all indicate that as roughness increases the TIS value decreases. However, the values for TIS

obtained from the different methods diverge as roughness increases. The divergence may in part be due to geometry differences of the samples, where the 'Planing and Shaping' plate has a linear roughness and the 'Casting' plate has a random roughness. In addition, errors may be introduced because a TIS instrument is typically used to measures the diffuse reflectance at a single wavelength over a hemisphere while the THz system measures the specular reflectance over a limited angle.

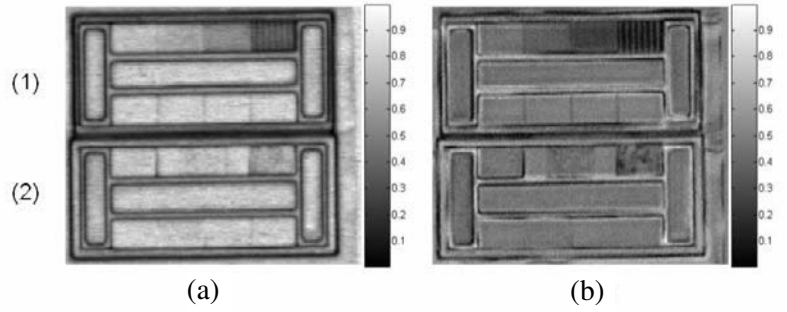


Figure 13. The (a) amplitude and (b) frequency images for Plate 1: 'Planing and Shaping' plate and Plate 2: 'Casting' plate.

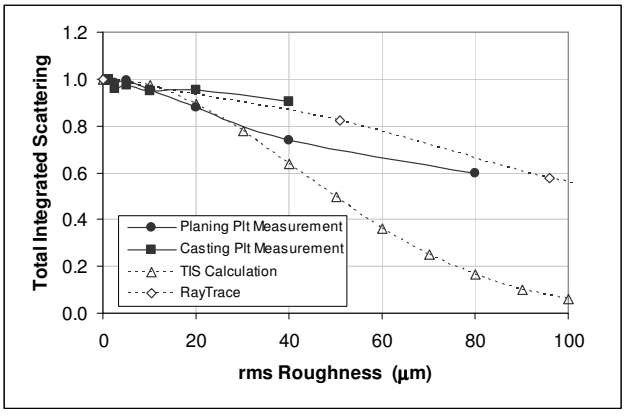


Figure 14. Measurement results as a function of rms roughness compared to calculated total integrated scattering values and ray tracing model results at a frequency of 0.4 THz.

5. Corrosion

Corrosion is a dissolving and wearing away of metal caused by a chemical reaction, mainly oxidation. In this electrochemical process, electrons in a system move from the anode to the cathode. The electron loss at the anode is called

oxidation and metal loss occurs. At the cathode, electrons are captured and damage such as hydrogen embitterment can occur [23]. The corrosion process generally affects a metallic surface that may exhibit corrosion pits, cracking, or thinning and in general changes a nominally smooth surface to an uneven and irregular surface that maybe measured. In this section, a THz system is used to examine the feasibility of inspecting for corrosion effects by examining variations in THz response due to changes in paint thickness and surface variation. Samples inspected include; a paint thickness sample, a helicopter fuselage component with paint condition variations and assumed corrosion, and Shuttle Tiles with manufactured corrosion spots on the metallic substrate.

5.1. *Corrosion Under Paint*

Conventional nondestructive evaluation technologies (infrared, eddy current, ultrasonic, and radiography) have been used to inspect for corrosion under paint with the most promising being thermographic testing [24]. Near-field microwave techniques have also been used to detect for the presence of corrosion on an aluminum substrate under paint and primer [25, 26]. In the microwave measurements, dielectric properties of the paint, primer, and corrosion by-products were examined in the S-, G-, J-, X-, and Ku-Bands. The measurements showed that the dielectric properties of paint, primer, and aluminum oxide are very similar which reduce the detection of aluminum oxide directly. In addition, these experiments demonstrated that higher frequencies with a standoff distance of a few mm were optimal for corrosion detection. Therefore pulse THz, with its higher frequency content, should have improved detection capabilities. This initial effort focused on measuring the thickness of the dielectric paint layer, as a potential way of identifying underlying corrosion and defects.

5.1.1. *Paint Thickness Sample and Measurements*

A specimen for THz measurements was fabricated with varying thickness layers of paint on an aluminum substrate. After sanding and cleaning the sample surface, it was coated with a thin layer of commercially available spray paint primer and then eight topcoat layers of an enamel (green) spray paint, to create a step thickness specimen with each step approximately 25 mm wide.

The paint step thicknesses, mechanically measured with a micrometer, ranged in thickness from 0.03 mm to 0.33 mm. Figure 15 shows the sample and THz signals for three different steps. Figure 15b shows the signal for step 9, paint thickness 0.335 ± 0.0039 mm. In this signal the paint surface echo and aluminum

surface echo are resolved. Figure 15c shows the signal for an intermediate paint thickness, step 6, paint thickness 0.183 ± 0.0039 mm, the individual echoes are not fully resolved in this signal and Figure 15d, show a signal from the bare aluminum surface. Using the time between the paint surface echo and aluminum echo a wave velocity in the paint was found to be 0.190 ± 0.013 mm/ps. This corresponds to a refractive index of 1.58 ± 0.11 and a dielectric constant of 2.50 ± 0.36.

A THz amplitude image in Figure 16 was generated from the maximum signal amplitude at each point in the scan. The image indicates the signal amplitude correlates to paint thickness. The lower amplitude corresponds to thicker paint and higher amplitude to thinner paint. Consideration of this data indicates paint thickness can be determined from the attenuation in the paint to a thickness of ~ 0.050 mm when individual paint surface and aluminum substrate echoes are not temporally resolvable. The variation in attenuation due to difference in the color or paint pigment was not investigated.

5.1.2. *Corrosion Sample and Measurements*

A sample with naturally occurring corrosion was inspected. This sample was a Fuselage Step Door from an out of use Chinook Helicopter and was obtained for use in this study from Ft. Eustis, United States Army Aviation Logistics School, USAALS, Newport News, VA. The sample, shown in Figure 17, is approximately 20 cm by 14 cm and has areas of chipped, flaking and cracked paint. Figure 18 shows a THz peak-to-peak amplitude image, which highlights areas where paint is chipped. The THz frequency images are shown in Figure 19. Figure 19a shows the spectral amplitude between 0.50 and 0.56 THz and Figure 19b shows the spectral amplitude between 0.56 and 0.62 THz. Examining small bandwidths in the frequency domain eliminates the effect of large amplitudes that obscure detail. Due to the complexity of the surface condition, a quantitative assessment of the extent of corrosion is not currently possible. However based on the paint-thickness sample and previous microwave efforts, areas of concern with thicker dielectric coatings can be identified as regions that maybe a combination of paint and corrosion products.

5.2. *Corrosion Under Shuttle Tile*

The black high-temperature reusable surface insulation (HRSI) tiles are mainly seen on the underside of the Shuttle. They are made of a low-density, high-purity silica amorphous fiber insulation that is made rigid by ceramic bonding. After the

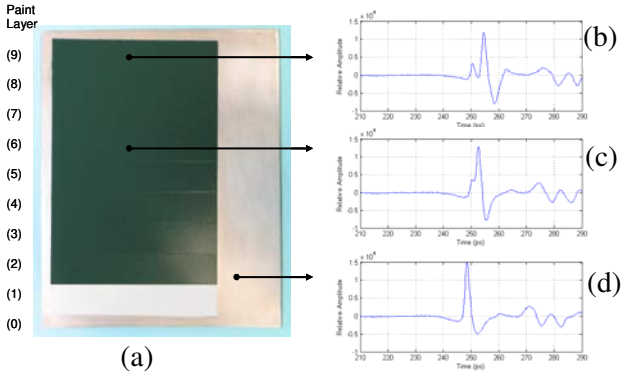


Figure 15. Paint sample (a) and (b) signal through layer-9, (c) through layer-6, and (d) on bare metal.

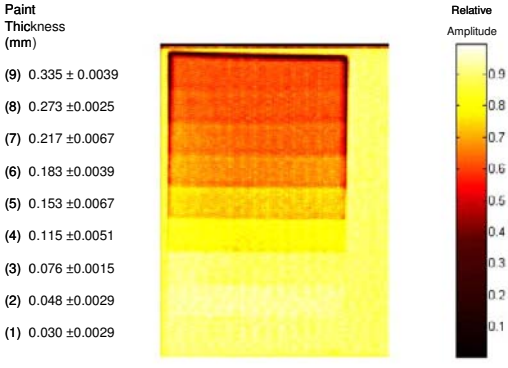


Figure 16. THz Amplitude Image (maximum signal amplitude).

Figure 17. Chinook Fuselage Step Door is shown with flaked, chipped and cracked paint.

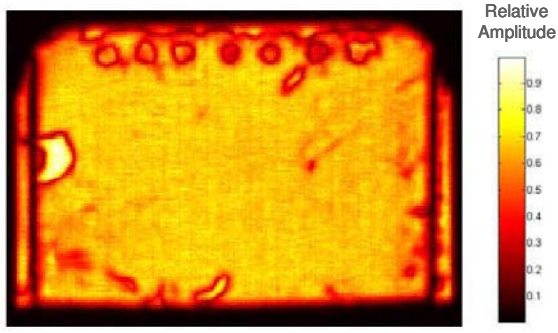


Figure 18. THz images of the Chinook Fuselage Step Door maximum amplitude.

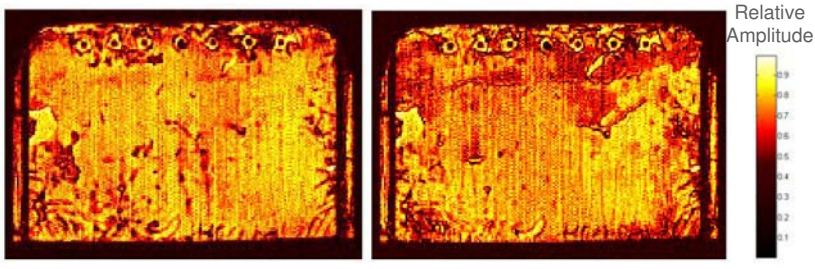


Figure 19. THz frequency images with mean amplitude between (a) 0.50 and 0.56 THz and (b) 0.56 and 0.62 THz.

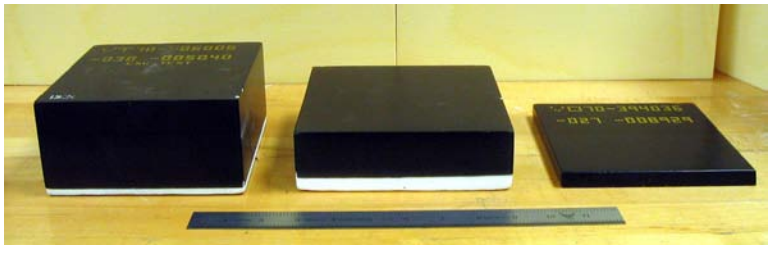


Figure 20. Shuttle tiles of various thicknesses: (a) 76 mm, (b) 48 mm, and (c) 12 mm.

tiles are shaped, they are sintered to produce a rigid block of material that is machined to precise dimensions required for their ultimate location on the Shuttle surface. A tile is 90% air, weighs 0.144 gr/cm^3, and provides protection against temperatures up to 1260° C. Tiles vary in thickness from 25.4 to 127 mm, which

is determined by the heat load encountered during re-entry. Generally, the tiles are thicker at the forward areas of the orbiter and thinner toward the aft end and are nominally 152 by 152 mm square. The HRSI tiles are coated on the top and sides with a mixture of powdered tetrasilicide and borosilicate glass with a liquid carrier and then heated in an oven to form a black, waterproof glossy coating. Tiles cannot withstand airframe load deformation and therefore stress isolation between the tiles and the Shuttle structure is necessary. This isolation is provided by a strain isolation pad (SIP) which isolates the tiles from the Shuttle's structural deflections. The SIP layer consists of a felt material that is bonded to the tile and then bonded to the Shuttle structure with an RTV, a room temperature curing rubber adhesive. Around the SIP layer is an adjacent layer, called a filler bar, used to fill in the space between tiles [27, 28]. The tiles, SIP, and filler bar materials are very porous, so that if they are exposed to moisture, they may hold the water next to the aluminum skin. If the paint and adhesive that covers the skin fails or is missing, the moisture may activate visually undetectable corrosion. THz waves can penetrate the tile, enabling a single sided inspection for corrosion. In this section, a preliminary study of the capabilities of THz for inspecting for corrosion through a Shuttle Tile is presented.

5.2.1. *Tile Response to THz*

To characterize the THz properties of the tiles, reflections from a metal substrate with and without an intermediary thick, medium, and thin tile were measured. These tiles are shown in Figure 20. The thin tile in Figure 20c was made from an originally thicker tile that was cut. From the signals, the calculated attenuation was 3.06 dB/cm and the average wave velocity was $2.78 \pm 0.04 \times 10^8$ m/s. The refractive index for the material, calculated from the average wave velocity, was 1.08 ± 0.02.

5.2.2. *Corrosion Samples and THz Images*

To assess the ability of the THz system for corrosion detection under Shuttle tiles, tiles were placed on a metal substrate with manufactured corrosion defects. In one set of measurements, the corrosion was covered with detachable tiles to allow for measurements with and without the tiles. In other cases, blind tests were performed on tile samples with defects that were covered by a permanently attached tile. The corrosion specimens were part of test panels that were originally test articles for an impact study. These panels consisted of an array of tiles 24.5 mm thick, attached to an aluminum substrate. Each tile numbered was used to reference the THz scans.

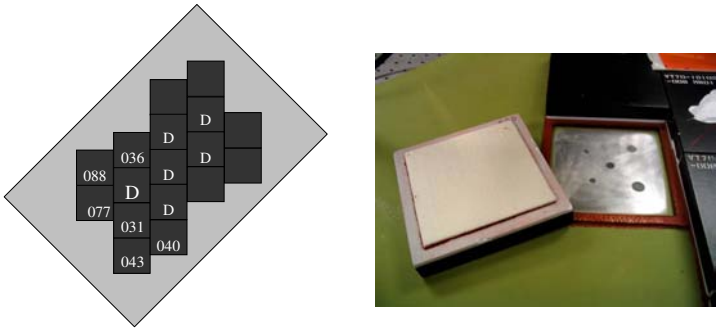


Figure 21. First test panel (a) schematic and (b) close-up of tile '077' inverted and induced corrosion areas on metallic substrate. The numbered tiles were inspected and 'D' indicates tiles that were damaged in previous tests.

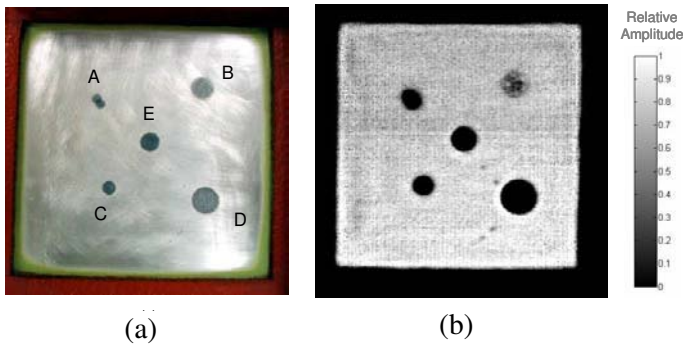


Figure 22. Induced corrosion sites A-E under tile '077' (a) photograph and (b) THz amplitude image.

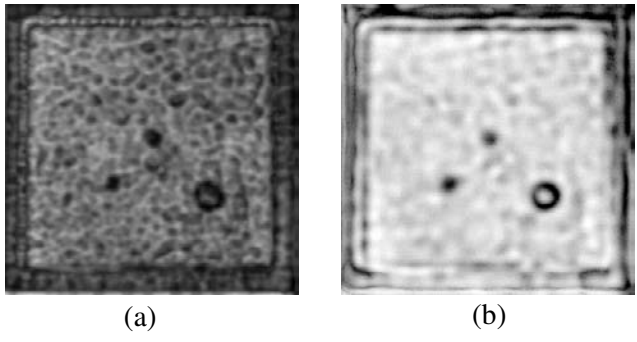


Figure 23. THz images of induced corrosion covered with tile '077' (a) Amplitude image and (b) Frequency image.

The first test panel is depicted in Figure 21a. The number corresponds to the last three digits of the tile identification number and the 'D' indicates tiles that were damaged in previous tests. The numbered tiles were inspected. Two of the tiles were removed and five areas of corrosion were induced. The shape of the corrosion areas was similar for each tile. One tile was re-bonded to the aluminum substrate and the other tile was not re-bonded. This allowed for inspection of the corrosion damage with and without a tile. Figure 21b shows a close-up of the removable tile and its placement position. This tile is numbered '077' in Figure 21a and is shown inverted with the felt pad bonded to it. The location of the re-bonded tile was unknown to the personnel conducting the THz study.

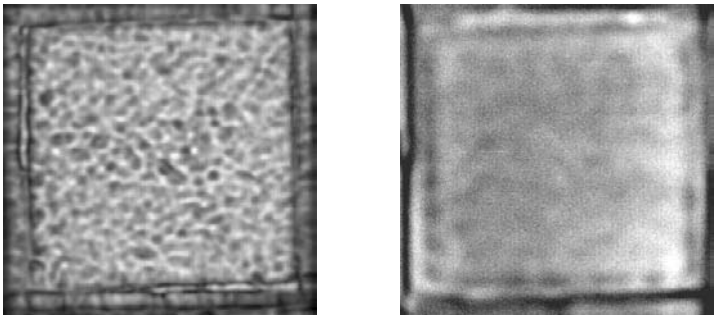
The extent of induced corrosion spots under tile '077' labeled A-E in Figure 22a was measured using a surface replication process. From the replication, estimates of the depth of corrosion were made. Depths were approximately (A) ~0.06 mm, (B) less then 0.06 mm, (C) 0.95 mm, (D) 0.75 mm, and (E) 0.30 mm. Diameters ranged from 2.54 mm to 15.24 mm.

Figure 22b shows a THz scan of the bare aluminum surface. The amplitude image is a grey scale representation of the peak-to-peak value at each scan point. The five corrosion spots and some small surface scratches can be seen in the image. The tile was placed over the corrosion spots and rescanned. The resulting amplitude image and frequency image at 75 GHz are shown in Figure 23. Figure 23a shows some signal variations caused by signal attenuation across the tile and a few darker spot that stand out from the background. The contrast between the corrosion spots and the background improves in the frequency image, Figure 23b. Band-pass filtering eliminates some of the background variations that obscure details in the data. The dark spots in Figure 23b correspond to spots C, D and E or spots with corrosion depth equal to or greater than 0.30 mm.

Other attached or bonded tiles on this first panel were inspected with the THz system. Images of tile '031' shown in Figure 24 are representative of the other tiles. Figure 24a shows the amplitude image and Figure 24b is the frequency image. The distinctive spots like those for tile '077' are not seen in these images and nor in the other tiles except tile '043'. The images for tile '043' are shown in Figure 25. Theses images show the distinctive spots similar to those for tile '077'. According to personnel that manufactured the panel, this tile was re-bonded over the induced corrosion spots similar to those of tile '077'.

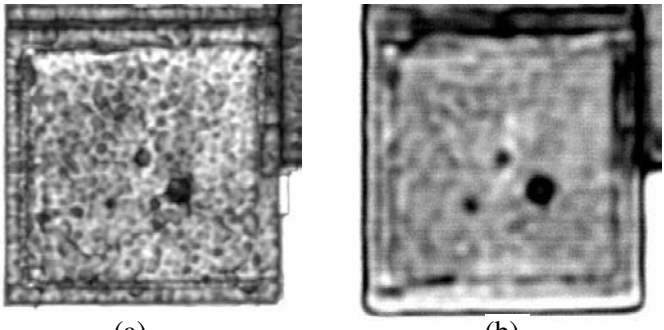
The second panel inspected had 17 tiles and a layout similar to the first panel. Eleven of the 17 tiles were inspected using the THz system. The remaining six

tiles were damaged in previous tests. This inspection was a blind test, the personnel inspecting the panel did not know which tile had been removed and replaced after corrosion was induced. THz images of all but one tile were similar to those in Figure 24. Tile '745' was the exception. The images of this tile showed distinctive spots that stood out from the background. This image along with a picture of the 25 induced corrosion spots (obtained after the inspection process) is shown in Figure 26. Figure 26a shows the 25 induced corrosion spots and depth estimates made during manufacturing. Figure 26b shows the THz image with eight large obvious corrosion spots and another eight that are smaller and not as evident. From the depth measurements and THz image, THz NDE could detect corrosion spot with depths greater than approximately 0.13 mm.



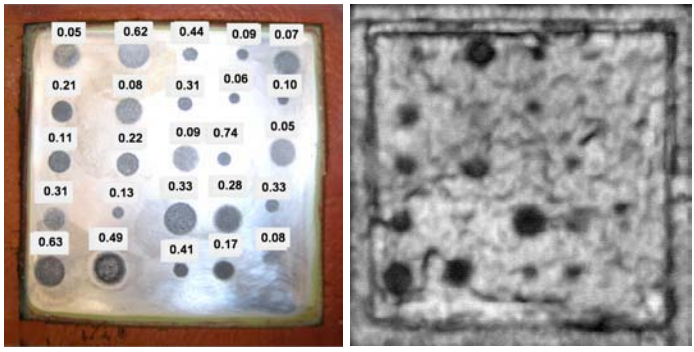
(a) (b)

Figure 24. THz images of tile '031' (a) Amplitude image and (b) Frequency image.



(a) (b)

Figure 25. THz images of tile '043' (a) Amplitude image and (b) Frequency image.



(a) (b)

Figure 26. Tile '745' (a) photograph of induced corrosion spots with estimated depth (mm) at each location and (b) THz amplitude image.

6. Summary

This chapter has reviewed some applications of THz NDE. It has shown how THz is able to detect and image defects in foam and inspect a metallic substrate for roughness and corrosion. The foam investigations indicate the measurement system's temporal resolution and signal-to-noise ratio enables the detection of air-gaps in the foam. Changing from the pitch-catch transceiver to the confocal transceiver improved the focusing and thus images from the system.

The system was able to measure surface roughness to ~12.5 μm Ra. Also shown was that the technique of Total Integrated Scattering are plausible to adopt with the noted difference between the systems being a diffuse measuring system versus a specular measuring system, and a single versus a band of wavelengths. The best correspondence between measurements, formula, and modeling appeared to be in the rms range from ~10 μm to ~ 40 μm. Above this range, the results diverge, which may be due to instrument, modeling, and sample geometry differences. The THz system has the advantage that it could be used for surface roughness measurements under dielectric coatings.

THz response to paint thickness showed that as the paint became thin, an individual paint surface echo could not be resolved from substrate echoes, but signal amplitude continued to be indicative of paint thickness. Minimum thickness resolved was ~ 0.05 mm. The inspection of the painted Fuselage Step Door proved too complex to characterize without studying small sections, but signal amplitude was able to specify regions of concern.

Electromagnetic properties of Shuttle Tiles were examined in the THz regime and found to have a refractive index of 1.08, which was similar to the SOFI

material. Inspection of the corrosion panels showed that corrosion on the bare aluminum could easily be detected while corrosion under the tile was more difficult to detect. This was primarily due to attenuation of high frequency components of the THz wave as the THz pulse propagates through a tile. Measurements indicate that corrosion depths of greater than ~ 0.13 mm could be detected

The applications reviewed here have demonstrated the capability of THz NDE and have shown promising results inspecting dielectric layers and substrate material.

Acknowledgement

The authors would like to acknowledge the contributions and support for this work provided by William W. Ussery of Michoud Assembly Facility, James L. Walker of NASA Marshall Space Flight Center, Richard W. Russell of NASA Johnson Space Center, Janice K. Lomness and Catherine C. Kammerer of NASA Kennedy Space Center, and Stephen W. Smith of NASA Langley Research Center.

References

1. P. H. Siegel, *IEEE Transactions on Microwave Theory and Techniques*, Vol. 50, No. 3, pp 910-928, (2002)
2. D. L. Woolard, E.R. Brown, M. Pepper, and M. Kemp, *Proceedings of IEEE*, Vol. 93, No. 10, pp 1722-1743, (2005)
3. P. R. Smith, D.H. Auston, and M.C. Nuss, *IEEE Journal of Quantium Electronics*, Vol. 24, No. 2, pp 255-260, (1988)
4. C. Fattinger and D. Grischkowsky, *Appl. Phys. Lett*., Vol. 54, pp 490-492, (1989)
5. B. B. Hu and M. C. Nuss, *Opt. Lett*., Vol. 20, No. 16, pp 1716-1718, (1995)
6. D. M. Mittleman, R. H. Jacobson, and M. C. Nuss, *IEEE J. Sel. Topics Quantum Electron*., Vol. 2, No. 3, pp 679-692, (1996)
7. M. J. Fitch, C. Dodson, D. S. Ziomek, and R. Osiander, *Chemical and Biological Standoff Detection II*, edited by James O. Jensen, Jean-Marc Thériault, *Proceedings of SPIE*, Vol. 5584, pp 16-22, (2004)
8. K. Kawase, Y. Ogawa, and Y. Watanabe, *Optics Express*, Vol. 11, No. 20, pp 2549-2554, (2003)
9. M. C. Kemp, P. F. Taday, B. E. Cole, A. J. Fitzgerald, and W. R. Tribe, *Terahertz for Military and Security Applications*, R. J. Hwu, and D. L. Woolard editors, *Proceedings of SPIE* Vol. 5070, pp 44-52, (2003)
10. N. Anscombe, *IEE Review*, pp 26-30, (2005)
11. D. Zimdars, J. S. White, G. Stuk, A. Chemovsky, G. Fichter, and S. Williamson, *THz Technology, Ultrafast Measurements, and Imaging, IEEE*, pp 5-6, (2005)

12. K. Humphreys, J. P. Loughran, M. Gradziel, W. Lanigan, T. Ward, J. A. Murphy, C. O'Sullivan, *Proceedings of the 26th Annual International Conference of the IEEE EMBS*, pp 1302-1305, (2004)
13. E. Pickwell, B. E. Cole, A. J. Fitzgerald, M. Pepper, and V. P. Wallace, *Phys. Med. Biol*., Vol. 49, pp 1595-1607, (2004)
14. Report of the Columbia Accident Investigation Board, Government Printing Office, Washington, D.C., August (2003)
15. C. E. Cockrell, Jr., R. S. Barnes, H.L. Belvin, J. Allmen, and A. Otero, *41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*, Tucson, Arizona, pp 1-11, (2005)
16. "NASA's Implementation Plan for Space Shuttle Return to Flight and Beyond," Vol. 1, Ninth Edition, <http://www.nasa.gov/pdf/111156main_RTF_imp_9th_changes.pdf>, (2005)
17. D. Zimdars, J. A. Valdmanis, J. S. White, G. Stuk, G., S. Williamson, W. P. Winfree, and E. I. Madaras, *Review of Progress in Quantitative Nondestructive Evaluation*, Golden, Colorado, (2004)
18. E. Weiser, T. St. Clair, and M. Nemeth, *NASA Langley Research Center Technical Report,* NASA/TM-2004-21328, (2004)
19. NASA Facts, Thermal Protection System, FS-2004-08-97-MSFC, http://www.nasa.gov/pdf/63758main_TPS_FACT_SHEET.pdf, August (2004)
20. J. A. Ogilvy, "Theory of Wave Scattering from Random Rough Surfaces," Institute of Physics Publishing, Philadelphia, PA (1992)
21. J. M. Bennett, and L. Mattsson, "*Introduction to Surface Roughness and Scattering*," Op Soc of America, Washington, D.C., (1999)
22. Precision Devices, Inc, http://www.predev.com/smg/parameters. htm#Ra%20-%20Average%20Roughness, November (2005)
23. G. A. Matzkanin, and J. L. Easter, *Nondestructive Testing and Information Analysis Center*, Publication NTIAC-SR-98-03, (1998)
24. J. S. Cargill, J. A. Pecina, S. M. Shepard, J. D. Weir, R. Zoughi, and A. V. Bray, *Materials Evaluation,* Vol. 63, No. 2, pp102-109, (2005)
25. D. Hughes, N. Wang, T. Case, K. Donnell, R. Zoughi, R. Austin, M. Novack, *Review of Progress in Quantitative Nondestructive Evaluation*, Vol. 20, ed by D. O. Thompson and D. E. Chimenti, pp 460-466, (2001)
26. D. Hughes, N. Wang, T. Case, K. Donnell, and R. Zoughi, *Surface Sensing Technologies and Applications*, Vol. 2, No. 4, pp 435-451, (2001)
27. NASA Facts, "Obitor Thermal Protection System," KSC Release No. 11-89, http://www-pao.ksc.nasa.gov/nasafact/tps.htm, February (1998)
28. "Space Shuttle Orbiter Systems – Thermal Protection System," http://science.ksc.nasa.gov/shuttle/technology/sts-newsref/sts_sys.html, Accessed August 4, (2006)

CHAPTER 13

LARGE AREA TIME DOMAIN TERAHERTZ (T-RAY) IMAGING NON-DESTRUCTIVE EVALUATION FOR SECURITY AND AEROSPACE

David Zimdars, Jeffrey S. White, G. Stuk, G. Fichter,
G. Sucha, and S. Williamson

Picometrix, LLC., 2925 Boardwalk Dr., Ann Arbor, Michigan 48108, USA
dzimdars@picometrix.com

The method, technology, and examples of high speed time domain terahertz (a.k.a. T-Ray) imaging non-destructive examination (NDE) for security, aerospace, and building structure are discussed. T-Ray imaging can be utilized for non-contact transmission and/or monostatic reflection inspection of non-conductive materials such as plastics, foam, composites, ceramics, paper, wood and glass. The method can be used to generate 2 and 3 dimensional sub-surface structural images. Spectral content analysis enables material characterization. Example subsurface homeland security images of concealed items in baggage and on personnel are shown. Voids and disbonds present in space shuttle external tank sprayed on foam insulation are shown. We tabulate attenuation and penetration characteristics through a selection of building materials, and demonstrate the ability of T-ray instrumentation to sub-surface image building structures such as wall framing and interior wiring and conduits.

1. Introduction

Time domain terahertz (T-Ray) imaging is being adopted for non-destructive evaluation (NDE) applications in aerospace and other government and industrial settings [1-3]. T-Ray imaging employs single cycle sub-picosecond electromagnetic impulses with hyper wide band spectral content spanning the mm-wave (<50 GHz) to terahertz (>2 THz) region of the spectrum. It can be employed in transmission or reflection, and is analogous to a non-contact electromagnetic raster scanned “ultrasound” method but with spectral information. T-Ray waveforms can be analyzed either in the time domain for structural information, or in the frequency domain for spectral features.

NASA is currently employing T-Ray reflection NDE to examine the space shuttle external tank sprayed on foam insulation (SOFI) for voids and disbonds. Homeland security applications such as the inspection of personnel [2], the detection of concealed explosives [2], biological agents, chemical weapons, flammables, metallic and non-metallic weapons, and other potentially dangerous items are the subject of active investigation. Advancement of many of these application beyond small table top experimentation had been limited by slow imaging speed (tens of minutes or hours), small scan areas (<10 square cm) and in many cases the requirement that the sample itself be mechanically raster scanned.

There are several imaging techniques available today for 2-D inspection of surfaces and interfaces within optically opaque media. A few of the more common ones are: thermal imaging which is sensitive to contrast in temperature, microwave, ultrasound, x-ray backscatter, shearography, and infrared imaging (both coherent and incoherent). Each of these approaches draws of the uniqueness of the type radiation that is being used to construct the highest resolution, highest contrast image possible.

The pulsed terahertz (T-Ray) imaging technique is the newest addition to the list of subsurface imaging tools. The T-Ray imaging method system taps the region of the electro-magnetic spectrum that resides between the millimeter-wave region on its low-frequency end and the far-infrared region on its high-frequency end. In a few respects it is akin to microwave imaging. In other respects, it more similar to pulsed ultrasound imaging. However, there are several very important distinctions in comparison to each of these techniques. T-Ray imaging can offer up to 10 to 50 times greater resolution than microwave imaging at 40-90 GHz. Secondly, short pulses of time-domain terahertz imaging eliminate range ambiguity and standing wave patterns which can plague continuous wave microwave techniques. Thirdly, the reflected T-Ray waveform is similar to an ultrasound “B-Scan” with pulses reflected from each interface, allowing 3 dimensional layer structure to be reconstructed. However, it is non-contact, and can penetrate vacuum. Lastly, the hyper-wide bandwidth can allows the determination of spectroscopic information, such as explosive resonances.

The T-Ray ulsed-laser-based technique overcomes past impediments to allow terahertz radiation to both be generated and detected. Because the pulses that are produced are extremely brief, on the order of 1 picosecond (10^{-12} seconds), it is possible with this technique to “time-gate” only that region of the object field that is important. By blocking out all the scattered terahertz radiation emanating from the foreground and background sandwiching the plane being imaged, we

greatly enhance our signal-to-noise. Time-gating also eliminates range ambiguity that result in ghost images.

T-Ray radiation generally penetrates most dielectric materials such as plastic, foam, ceramics, paper, clothing, cardboard, gypsum board and dry wood. Water is highly absoprtive. Metal is reflective.

A non-exhaustive list of aerospace materials which may be imaged include: thermal protection systems (space shuttle silica tiles), sprayed on foam insulation, cork layers used in the boaster rockets, and other composite structures. These materials are in systems in which the examination of new construction for flaws (voids, disbonds, inclusions, improper geometry and dimensions, and incomplete curing) may be critical. Likewise, as spacecraft and aircraft age it will be critical to periodically inspect some systems for damage or chemical degradation. In other areas, T-Ray imaging could assess early signs of corrosion under paint and fatigue at bonded interfaces. T-Ray imaging will find applications in the inspection of automobile dashboards, imaging behind walls and flooring surfaces in construction, inspection for delamination of printed circuit boards and tires as well as with manufactured parts such as tiling and paper products.

2. T-Ray Imaging

To assist in the explanation of T-Ray reflection imaging, Figure 1 shows a hypothetical example object. For heuristic purposes, the object to be imaged is a two-layer laminated object. The object consists of two glued together THz transmissive layers (for example, resin composite) with a reflective backing. There is a small rectangular gap in the glue layer. The glue layer is buried in the opaque resin layers and is not visible.

The imaging scheme is to transmit T-Ray pulses through the resin layers, reflect off of the backing layer, transmit back through the resin layer, and collect the reflected pulse with the receiver. The pulses are generated by a THz transmitter modules and are detected by a receiver module. The transmitter-receiver assembly is raster-scanned above the object. The T-Ray waveform is collected for each location. The waveform consists of reflection pulses from each interface with a change in refractive index, or metallic barrier. The waveform is somewhat analogous to a “B-Scan” in ultrasound. A “C-Scan” can be constructed from the matrix of raster scanned waveforms by choosing a analysis function that reduces the waveform to a scalar intensity. For example, the root-mean-square value of the T-Ray waveform reflected from the backing layer can be used to construct the image. The waveform is windowed around the reflection from the

backing layer. Note that the final reflection from the backing layer is reduced by all of the absorption present from each layer above.

For this example, the glue layer is presumed to be somewhat reflective, so only a portion of the T-Ray pulse propagates through the glue. In this case, the relative amplitude of a pulse traveling through the glue should be less than the amplitude of a pulse traveling through the gap in the glue. To generate the image, the fiber-optic coupled modules are raster scanned above the object. The waveforms are collected at a constant rate, and each waveform represents a pixel in the image. The bottom of Figure 1 shows a two-dimensional image of the object reconstructed from the RMS value of the waveform. The darker glue layer is essentially a "shadow" and the gap in the glue layer is essentially a lighter gap in that shadow.

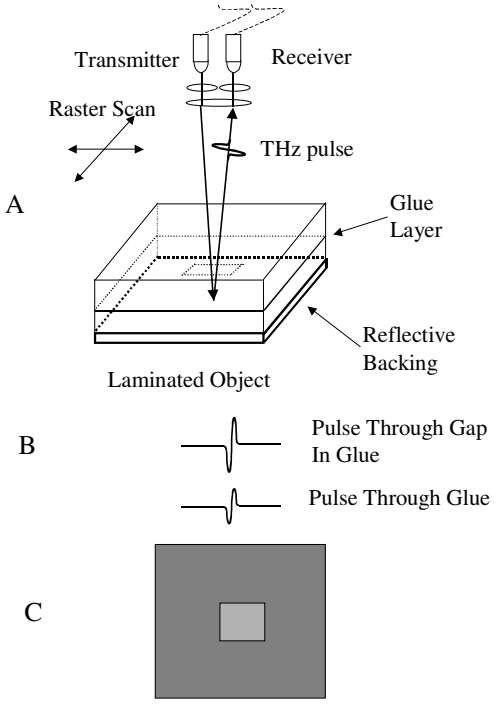


Figure 1. Heuristic diagram of 2D terahertz reflection imaging of glue under cardboard. A: T-Ray pulse beam path through object cross section. B: Attenuation of pulse in glue layer. C: 2 D image from RMS of waveform.

3. T-Ray Instrumentation

The potential of terahertz imaging for non-destructive evaluation was evident since the first femtosecond laser based demonstration of by Nuss and co-workers

[3-4]. However, the cumbersome nature of free space optical bench systems prevented the application of terahertz imaging outside of terahertz development laboratories until the advent of commercial fiber optic coupled turn-key instrumentation. In order to home-build a free space T-Ray system, skills in ultrafast lasers and optics, semiconductor physics, and sub-millimeter waves optics are required. Once assembled from individual components, these homebuilt systems produced images by moving the object under test through a stationary beam, preventing the system from scanning large immovable objects. This is because moving the terahertz transmitter or receiver chips would require re-aligning the free space femtosecond laser beam on the chip with submicron precision. For an application such as scanning the space shuttle external tank sprayed on foam insulation, it is only practical to scan a movable reflection transceiver over the stationary tank.

Recently high speed large area T-Ray non destructive evaluation imaging systems with freely positionable fiber optic coupled components have been developed. The Picometrix QA-1000 terahertz imaging setup, as utilized to generate the images shown in this chapter, is shown in Figure 2. Transmission geometry is shown in Figure 3. The system consists of a rack mounted control unit containing a femtosecond laser, fiber launch, a 320 picosecond window -- 100 Hz repetition rate optical delay, signal processing, motion control and image analysis computer. A 5 meter fiber-optic and electrical umbilical cord attaches the terahertz transmitters and receivers mounted on an overhead gantry with 1.0 m by 1.0 m scanning range. The fiber optic coupling enables the scanning of the transmitter and receivers, allowing for a stationary object. In reflection, the co-linear transceiver has a 30 cm working distance, bringing the THz pulse to a 2 mm focus. In transmission, a "C" shaped yoke holds the transmitter and receiver above and below the object, respectively. Up to 30 cm thick objects may be imaged in transmission, with the focusing lenses achieving a similar 2 mm spot. Both the reflection transceiver or transmission "C" yoke can be moved vertically up to 30 cm to achieve the best focus.

The T-Ray waveform captured at each location is analogous to an "A-scan" in ultrasound. As the gantry motions control raster scans, T-Ray waveforms are captured for each location and are analyzed to generate an image similar to an ultrasound "C-scan."

The terahertz transmitter and receiver modules utilize the method of photoconductive generation and detection in low temperature grown gallium arsenide (LT-GaAs). Femtosecond light pulses are used to trigger the conduction of current in the biased transmitter and to detect the incoming terahertz field inducing current in the antenna time coincident with the optical gate (Figure 4).

The femtosecond light pulses are used to trigger terahertz transients because the optical femtosecond pulses are several orders of magnitude faster than what can be achieved by purely electronics methods. The photoconductive elements are powered by a 100 fs pulse duration laser with 80 MHz repetition rate and an emission wavelength at 800 nm. Typically 20 mW of optical power is coupled into the fiber.

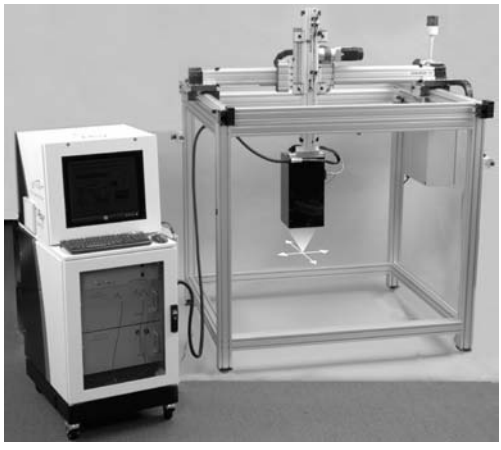


Figure 2. Large area T-Ray (time domain terahertz) non destructive imager (QA1000) with rack mounted laser, high speed optical delay, and control unit configured for co-linear (monostatic) reflection imaging.

Figure 3. QA1000 scanner configured in transmission configuration with overhead X-Y raster scanner and "yoke" mounted transmitter (TX) and receiver (RX).

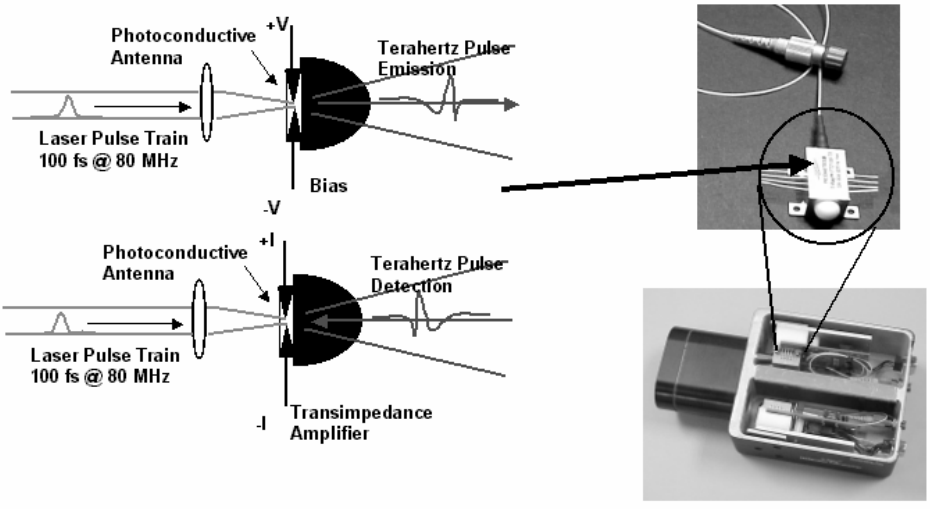


Figure 4. Generation and detection of T-Ray pulses within fiber pigtailed modules. Upper Left: Emission of T-Ray pulse from biased antenna on LT-GaAs semiconductor stimulated by femtosecond laser pulse. Lower Left: Detection of T-Ray pulse current induced through an antenna on LT-GaAs semiconductor gated by femtosecond laser pulse. Upper Right: Miniature packaged T-Ray receiver module with fiber pigtail and internal current preamplifier. Lower Right: Installation of modules within T-Ray reflection transceiver.

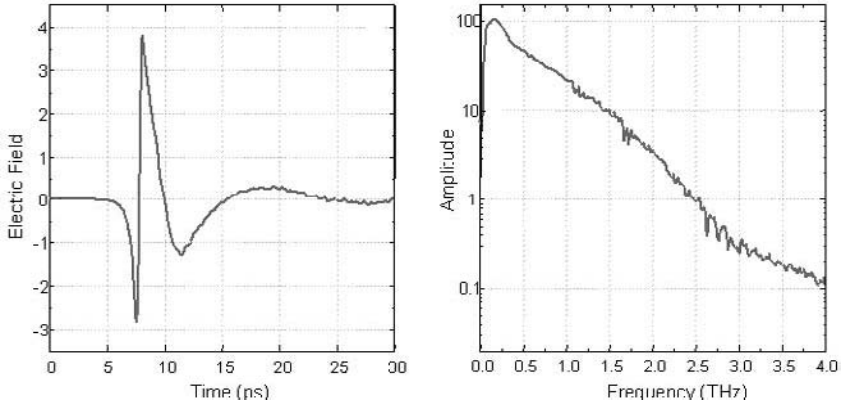


Figure 5. T-Ray waveform. Left: Time domain waveform in 30 picosecond window. Right: Fourier transform of time domain waveform. Spectral content extends from 50 GHz to > 2 THz. Dips in waveform above 1 THz are water vapor absorption lines.

The terahertz waveform is generated by scanning the receiver optical gate in time with respect to the transmitter optical trigger pulse. The QA-1000 scans a waveform window of 80 ps at 20 Hz or optionally a waveform of 320 ps at

100 Hz. Figure 5 shows a 30 ps portion of a terahertz transmission waveform [1]. The signal to noise shown is typical of a well optimized single scan at 20 Hz repetition rate. The 10% bandwidth extends from 0.03 to 1.5 THz with usable bandwidth to 2 THz.

4. T-Ray Transmission Imaging

T-Ray imaging can been performed in either transmission or reflection geometries. The T-Ray pulses emitted from the terahertz transmitter are brought to a focus within the object to be imaged. In transmission imaging, the terahertz pulses pass through the object to be imaged and are re-focused onto the terahertz receiver. In two dimensional imaging, a scalar value is calculated from the T-Ray waveform (A-scan) to assign a brightness value at that pixel point. Common analyses are peak to peak amplitude, delay, spectral power within a range. In a reflection image, the T-Ray pulse reflects from surfaces within the object. T-Ray reflection imaging can be used to determine three dimensional structure within the sample analogous to “B-scans” in ultrasound [3].

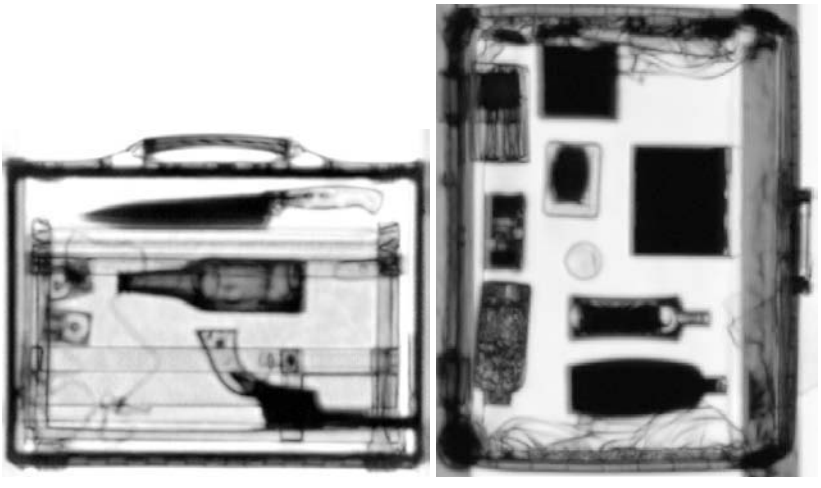


Figure 6. Transmission T-Ray images. Left: Attaché Case. Right: Suitcase, approx. 30 in. by 20 in, by 13 in.

Two sample T-Ray transmission images are shown in Figure 6. The pixel intensity is logarithmically proportional to the power integrated from 0.2 to 2.0 THz, derived from the Fourier transform of the 320 ps time domain waveform. The image shows the attenuation of pulses within the object. Both images have 1.5 mm pixel sizes, but are not shown to equivalent scale.

Pixel acquisition for both images was 66 pixels per second. Raster scan speed was 0.1 m/s in continuous motion. At this rate and resolution, it takes approximately 100 minutes to scan the entire 1 sq. m area of the gantry.

5. Co-Linear (Monostatic) T-Ray Reflection Imaging

An example of a T-Ray reflection image employing the co-linear transceiver for a homeland security application is shown in Figure 7 and Figure 8.

The images in Figure 7 shows an example of time domain THz reflection subsurface images revealing threat objects concealed on a person. A chef's knife beneath a jacket, and a toy metal pistol beneath a jacket are shown. These items had been placed on the chest of a subject who lay beneath the QA1000, imaging time taking approximately 10 minutes. The image was formed by plotting pixel intensity logarithmically proportional to the total power return without time gating. Image resolution was approximately 1.5 mm.

Figure 8 shows THz reflection images of an attaché case containing a knife, a pistol, and a block of boron carbide based explosive simulant. The images demonstrate the ability of time domain THz to display separate layer within an object. The top most image is a visible photograph of the exterior of the case. A single data set of T-Ray waveforms reflecting from the attaché case was used to generate the images. Time gating of reflections in the T-Ray return waveforms is used to isolate specific layers. By time gating one is generating "C scan" at a certain depth within the object.

The surface terahertz image of the attaché case in Figure 8 was generated by time gating only the reflected pulse from the surface and plotting pixel intensity proportional to the power of this top reflection. The top pockets image excludes

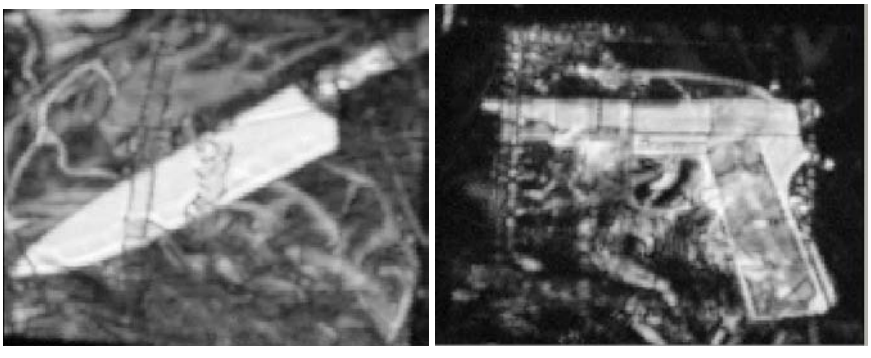
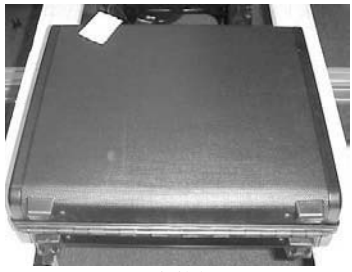
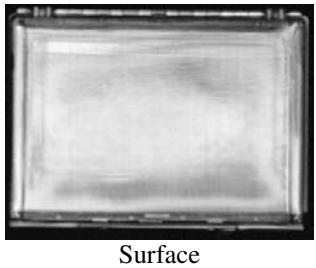


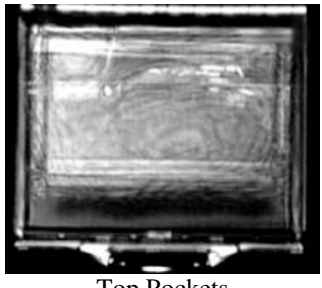
Figure 7. Reflection terahertz images of personnel. Left: knife under jacket; Right: pistol under jacket.



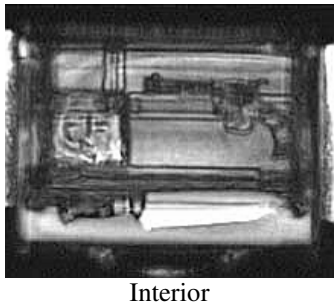
Visible



Surface



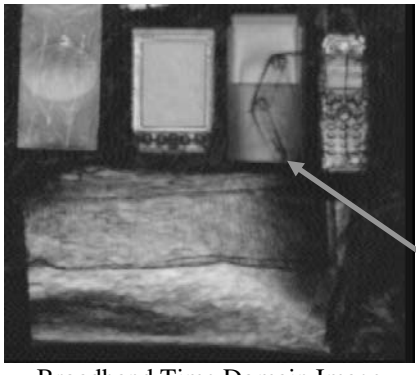
Top Pockets



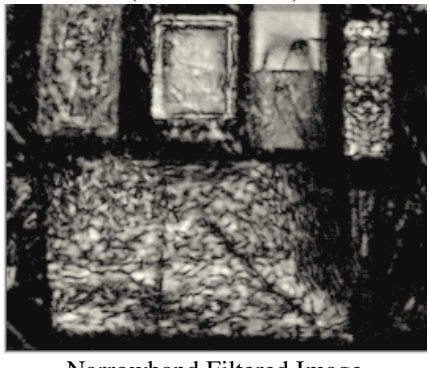
Interior

Figure 8. Reflection T-Ray images of Attaché Case showing multiple layers analyzed from single data set. From Top: Visible image; Return from surface of case attaché case; return excluding surface reflection showing top pockets; and return from interior of case at 2 inch depth showing knife and pistol.

Visible



Broadband Time Domain Image
(0.2 to 1.0 THz)



Narrowband Filtered Image
(0.43 THz +/- 0.003 THz)

Figure 9. Reflection T-Ray images of common personal items underneath cotton knit shirt. Top: Visible image (from top left) cigar pack, PDA, rubber sheet, plastic sheet, cell phone, and quarter fold newspaper; Middle: Image reconstructed from power between 0.2 to 1.0 THz; Bottom: Simulated continuous wave image (narrow band) reconstructed from power at 0.43 THz +/- 0.003 THz.

the topmost reflection, and shows only the power reflecting from the pockets that line the case interior. The interior image time gates a region later in the waveform corresponding to about 2 inches into the case, and shows the hidden interior contents. Image resolution was approximately 1.5 mm.

The reflection from each of these layers is attenuated by the previous layers. The time gating is very crucial to generating a high contrast image of the interior contents, in that if the image was formed from the entire waveform return the surface reflection would dominate, and intermediate layers would obscure the interior contents. This is a very important advantage over continuous wave illumination with total reflected power detection (for example by a mm-wave harmonic mixer detected by a rectifying diode) since there is no inherent way for the continuous wave system to exclude the unwanted reflections which may dominate the image.

The T-Ray reflection images of common benign personal objects beneath clothing (a cotton knit shirt) is shown in Figure 9. The image contains a cigar pack, PDA, rubber sheet, plastic sheet, cell phone, and quarter fold newspaper The center wide band time domain THz image is free of artifactual clutter and the items are recognizable and easily distinguished from the gun and the knife in Figure 7. The image was formed by plotting the pixel intensity proportional to the logarithm of the THz reflected power between 0.2 to 1.0 THz with no time gating.

Cloth has typical thicknesses between 0.5 mm to 2 mm and layer to layer spacing of 3 mm to 10-20 mm. This creates the possibility that single cloth layers and multiple stacks of cloth will generate Fabry-Perot resonances which can create artifactual interference patterns which clutter the image. These resonances will be a 5-20 GHz wide centered unpredictably up to about 1 THz. This interference clutter has been simulated by narrow band filtering the THz waveforms and generating an image at 0.43 THz +/- 0.003 THz (bottom of Figure 9). In comparison to the wide band time domain image, the objects in the narrow band image are highly obscured by the interference pattern. This “oil on water” appearance of narrow band THz images has at times been attributed to “standing wave artifacts” which posits interference between the imaging transmitter, receiver, and lenses and the object imaged. In cases where the illuminating coherence length is very long (10’s of centimeters) this may (and probably does) occur. However, in this case we believe the interfering structures are entirely within the object imaged (clothing).

The standard wide band time domain THz image does not show appreciable interference effects because the resonance is narrow compared to the total range of frequencies integrated, and as the center frequency shifts with the structure distance, the average total power is still approximately constant. An equivalent description would be to note that the time domain pulse has sub-millimeter coherence length, which is shorter than the cloth layer structure. This yields distinguishable pulses in the time domain that simply shift in time (but not power) as the structure distance changes. We speculate that ideally the illumination bandwidth should be >200 GHz to minimize these interference artifacts. This is easy to achieve with the time domain THz source, even with pulses centered below 500 GHz, to maximize penetration through the clothing. However, this bandwidth requirement will challenge many continuous wave sources, which naturally have much narrower emission bandwidths, and may subject to continuous wave architectures to interference artifacts.

6. T-Ray Images of Space Shuttle External Tank Sprayed on Foam Insulation

The detection of flaws within low-density dielectric materials such as sprayed on foam insulation (SOFI) presents a challenge to conventional non-destructive evaluation (NDE) techniques. The extreme attenuation of sound within foam can severely limit the effective depth of ultrasound, for example. Time domain terahertz reflection imaging has been shown to yield high resolution images promising an exciting new quantitative NDE method for materials such as foam.

We have determined that T-Ray pulses can penetrate as much of 20 cm of foam with a maximum frequency of up to 300 GHz (>1 THz for thinner foam). This high frequency, in comparison to microwaves, allows a lateral resolution of less than 2 mm and a depth resolution of <100 microns. T-Ray pulses are shown to scatter from voids with the foam and reflect from disbonds at the foam metal interface.

Disbonds are indicated by a “pre-pulse” reflecting from the air-foam interface above the metal backing, Figure 10. The pre-pulses from a piece of SOFI placed a variable distance (in microns) above a metal backing is shown below. The surface (or fully bonded position) is shown as an estimate because the 75 micron distance represents the foam fully clamped to the plate still exhibiting a pre-pulse, whereas bonded foam exhibits no pre-pulse.

Panels designed to simulate segments of the space shuttle external tank have been constructed to incorporate known voids and disbonds. These panels exhibit

a complex three-dimensional surface geometry with a bolt flange (running the circumference of the tank) and structural "stringer" running along the length of the tank. The foam on this structurally complex region of the taken is sprayed on by hand. The operator sprays on the foam, lets it harden, and then mills the surface to the required external shape. The flaws were incorporated by milling voids or disbonds of known sizes in small foam blocks and then spraying on the overlaying standard foam structure. The figure inset shows typical location placements of voids and disbonds.

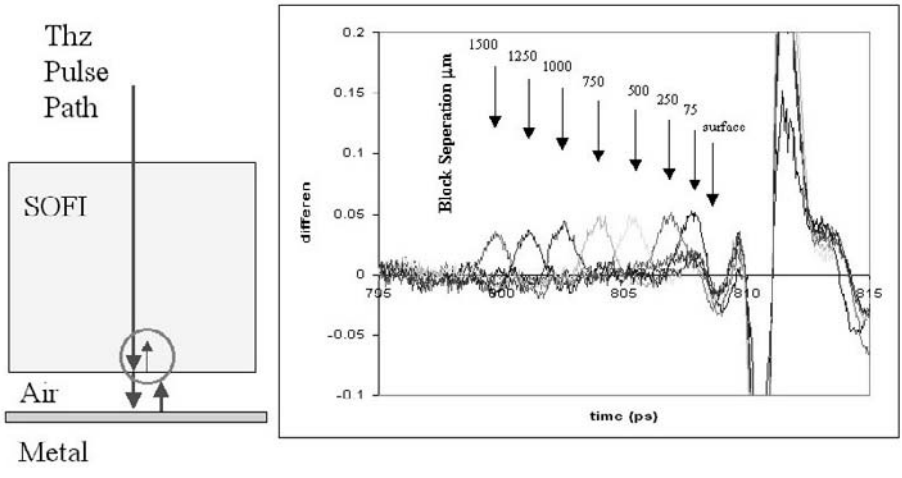


Figure 10. T-Ray pulse interaction with space shuttle external tank sprayed on foam insulation. Terahertz pulses are attenuated at higher frequencies by travel through the foam. Traveling through a void will decrease the amount of scattering due to absence of the foam. A portion of the pulse will be reflected from the surfaces of a void or disbond. Left: Diagram illustrating pulse path through the foam and the reflection from simulated disbond. Right: Measured "Pre-pulse" reflecting from air-foam interface as a function of air-metal separation.

The T-Ray reflection images of panels with synthetic known flaws are shown in below. Voids appear as bright bubbles with dark rings (small voids are dark), and can be sized as small as 0.25 in.

The T-Ray image of a panel intentional incorporating many "natural" defects (typically known as "rollover voids") is shown in Figure 12. These rollover voids may occur when the foam spraying technician moves the nozzle too quickly, and the foam liquid "rolls over" an edge, leaving a gap. Rollover voids are present in the image running horizontally parallel to the six stringers, and appear as dark narrow indications. The two parallel vertical lines running from top to bottom are

"conethane" lines, showing where the foam was milled and covered with glue. The white circles indicate regions where the milled surface broke through voids.

We have also imaged a test panel of the space shuttle thermal protection tiles, Figure 13. The beam was focused on the back-side of the panel, reflecting from the aluminum backing simulating the skin of the orbiter. The image shows the seams of the felt filler bars and bond-pads, as well as machined areas, repairs, and disbonds in the RTV adhesive.

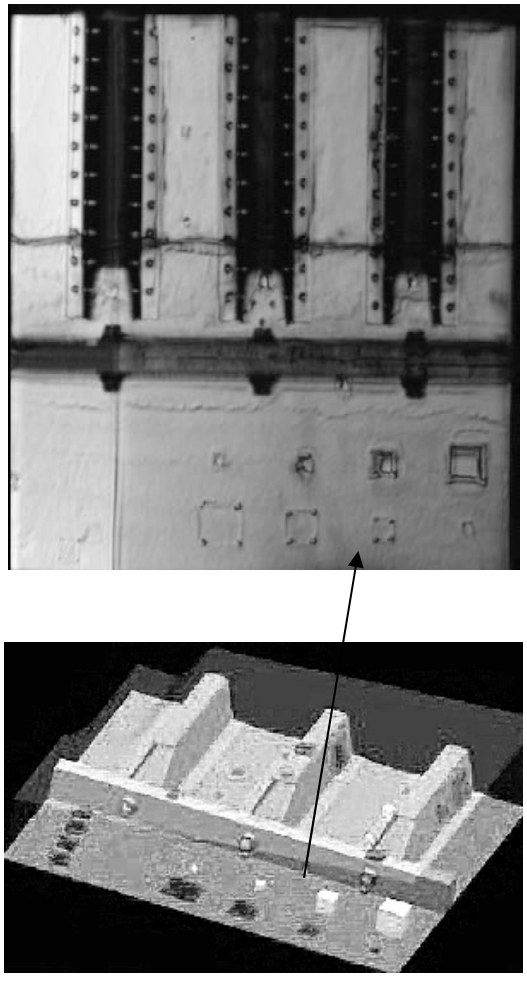


Figure 11. T-Ray images of space shuttle external tank mock-up panel with programmed synthetic voids and disbonds. Top: High resolution image of synthetic voids and disbonds. Bottom: 3D view of programmed location of voids on panel. The 3 structures at top are known as "stringers", the structure from left to right with the 3 bolts is the "flange".

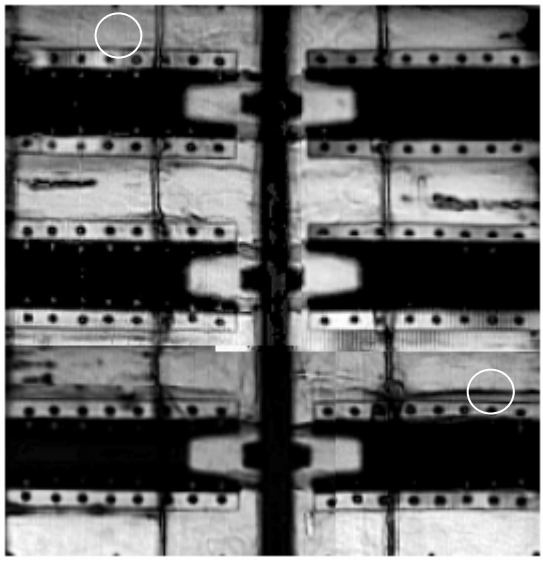


Figure 12. T-Ray image of space shuttle sprayed on insulation test panel with natural void inclusions. Image is a composite from 4 quadrants. Image was taken with high-resolution co-linear transceiver. The circles indicate regions where voids crack through the milled surface of the panel.

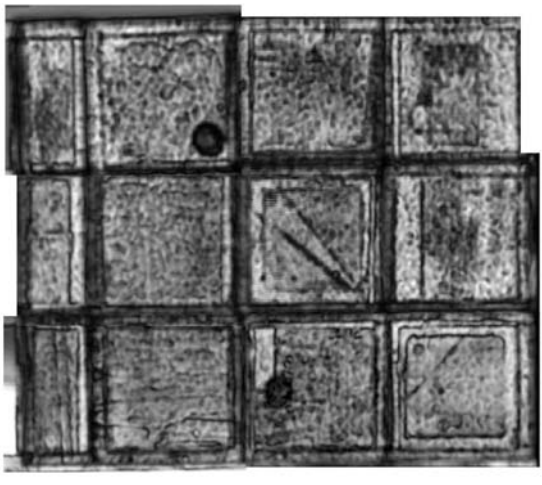


Figure 13. High resolution T-Ray reflection image of space shuttle thermal protection system tiles (0.2 to 0.6 THz). Image shows filler bars, bond pads, machined areas, backfills and disbonds.

7. Terahertz Transmission Properties of Building Materials

The attenuation as a function of frequency for a selection of building materials common to interior construction has been measured with the Picometrix T-Ray 2000 terahertz spectrometer in transmission geometry. The attenuation loss in dB (power) per centimeter of material is shown in Table 1. It is has been assumed that the materials follow Beer's law. Blank entries in table exceeded the dynamic range of the instrument. Those entries marked with an (*) showed an orientation dependence attributed to the polarization of the T-Ray pulse. Generally, the attenuation increases super-linearly with frequency, regardless of material. For some materials the trend is quasi-quadratic at the lower frequencies. The T-Ray 2000 utilized an antenna set which biased higher T-Ray power to frequencies below 500 GHz. The maximal frequency measurable for each material should be considered the measurement to have the greatest uncertainty. Measurements were made on unaltered samples of typical dimensions used in construction. Frequencies higher than 1.5 THz could penetrate these materials, of course, if the thickness were reduced, but this was deemed unrealistic for building inspection.

Each of these materials allows the transmission of a significant portion of the T-Ray pulse. They effectively exhibit a high frequency cut-off which depends on the type and thickness of the material. For the purposes of building structure inspection, the decade of spectra between 50 GHz and 500 GHz matches the transmission well, but the portion of the T-Ray pulse above 500 GHz will generally not be utilized (this is only specific to this construction material example where thicknesses are several centimeters). Loss of frequency may serve to lengthen the pulse in time and wavelength, reducing spatial and depth resolution from sub-millimeter to several millimeters.

As an example to estimate the amount of material to impose a 100 GHz cutoff for gypsum board drywall for a practical instrument with 80 dB dynamic range is calculated. At 100 GHz, loss at this frequency is ~1 dB/cm. Given 80 dB dynamic range / 1 dB/cm loss = 80 cm drywall.

A beneficial implication of the T-Ray method is that it will provide the maximal sub-millimeter lateral and depth resolution allowed by the optical configuration and material. Only the T-Ray image *resolution* is (gracefully) degraded as the material cut-off goes to lower frequencies -- there will generally be a detectable pulse from which to form an image of useable signal to noise rather than a complete blocking of the radiation. In contrast, a single frequency instrument (for example only at 800 GHz) may provide similar high resolution images, but only as long as the operational frequency is below the material cut-off. Additional lower frequency sources would be required to cover all materials.

Because the T-Ray impulse is hyper wide band, only a single instrument is required to image the broadest range of materials. In addition, unlike narrow band frequency methods employed in this region of the spectra, the method is not prone to standing wave artifacts or interference effects.

Table 1. Attenuation in dB per cm of selection of building materials at 0.1, 0.3 and 0.6 THz, 0.9 and 1.2 THz taken by Picometrix T-Ray 2000 THz spectrometer. Blank entries exceeded dynamic range of instrument. * Marked with asterisk shows polarization dependence.

Material	0.1 THz	0.3 THz	0.6 THz	0.9 THz	1.2 THz
0.34 in. plywood	13	31	46		
0.75 in. plywood	12	25	26		
5.2 mm Lauan	5.8	23	52	71	
Antistatic Carpet	9.4	36	76		
Nylon Carpet	6.3	21	58	76	
Cedar, Across Grain*	8.3	19	27	27	
Cedar, With Grain*	4.0	9.7	21	22	
Ceiling Tile	0.7	9.3	24	24	
Ceramic Floor Tile	8.3	36	45		
Clay Floor Tile	13	30	34		
Gypsum Drywall	0.8	12	27		
Hickory Floor Slat	12	24	26		
Marble Floor Tile	3.7	9.3	23	34	
Oak Floor Slat	9.8	25	24		
Pine Board	3.8	9.7	20		
Pine, Across Grain*	22	34	29		
Pine, With Grain*	13	26	26		
0.25 in. Fiberboard	7.9	27	55		
0.5 in. MDFB	4.7	21	29		
Melamine	12	31	32		
Asphalt Shingle	32	64	96	100	140
Slate	2.8	10	28		
Vinyl Drywall	1.6	19	27		
Vinyl Floor	-	8.7	26	52	87
Wood Laminate	11	24	59	63	

8. T-Ray Images of Interior Construction Structure

T-Ray subsurface images of building structure test articles are shown in Figure 14 (reflection) and Figure 15 (transmission). The T-Ray images are shown with corresponding visual images of the surface and interior. The articles were made to mimic typical interior wall construction. The surface of both test structures is 0.5 in gypsum board drywall. Note that the reflection image is of a metal stud structure, whereas the transmission image is of a pine 2x4 stud structure. Metal is reflecting and not transmissive.

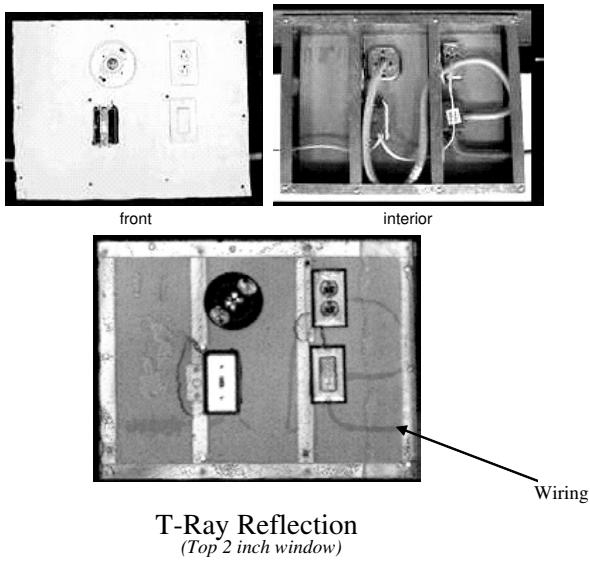


Figure 14. T-Ray subsurface images of building structure in reflection (left) and transmission (right). Visible images of the test articles are shown at the top.

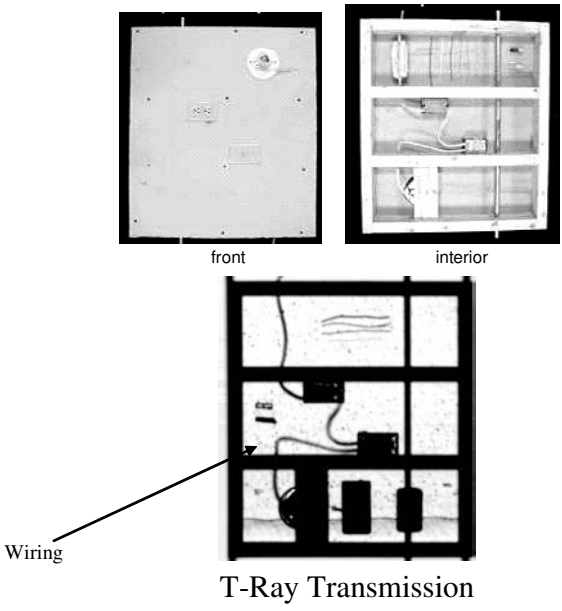


Figure 15. T-Ray subsurface images of building structure in transmission. Visible images of the front and interior of the test articles are shown at the top.

The reflection image shows the location of the studs and plastic wire conduits. Note not all wiring is visible in the reflection image. The QA1000 320 ps window only captures about 2 inches of return and the window was set to capture the front surface reflection and as much of the interior as possible. However the 2 in. depth did not encompass all the wiring or the back side drywall.

Transmission imaging is shown for comparison. In practice, reflection imaging would likely be the preferred mode since access to the far side of the wall is generally not available (except perhaps at a factory). The reflection image shows the direct reflection of the wiring, which can be used to three dimensionally locate it in depth ("B-scan"). An alternative method of collection which can provide much higher contrast for the intermediate structure, is to generate the image from the reflection from the back side drywall. Intermediate structures are then shadowed over the backside reflection, and the contrast is similar to the transmission image.

T-Ray imaging of wall structure and interior objects such as wiring and plumbing shows good feasibility. The T-Ray method has several advantages over alternative methods. Resolution is approximately 2 mm, superior to microwave of mm-wave. Unlike single frequency microwave imaging, T-Ray easily provides depth and layer structure. Front surface reflections which could dominate frequency domain images, can be easily removed by time gating. While similar to ultrasound "C scans", the T-Ray method can couple through air (or vacuum) with no loss resolution. Thus high resolution images can be made even when air separates one or multiple layers in the structure. Unlike back-scatter X-Ray imaging, the T-Ray method does not employ ionizing radiation, and there are no safety concerns for the operator.

Acknowledgements

The authors would like to thank Warren Ussery of the Lockheed-Martin Michoud Assembly Facility (MAF) for providing foam test articles.

References

1. J. V. Rudd, D. Zimdars, and M. Warmuth, "Compact fiber-pigtailed terahertz imaging system," Proceedings of SPIE, Commercial and Biomedical Applications of Ultrafast Lasers II; Joseph Neev, Murray K. Reed; Eds. , San Jose, CA, **3934**, (2000) p. 27-35.
2. D. Zimdars, "Fiber-pigtailed terahertz time domain spectroscopy instrumentation for package inspection and security imaging," Proceedings of SPIE, Terahertz Military and Security Applications, J. Hwu and D.L. Woolard, Ed., Orlando, Fla., **5070**, (2003) p. 108.; M.C. Kemp, P.F. Taday, B.E. Cole, J.A. Cluff, A.J. Fitzgerald, and W.R. Tribe, "Security Applications of

Terahertz Technology," Proceedings of SPIE, Terahertz Military and Security Applications, J. Hwu and D.L. Woolard, Ed., Orlando, Fla., **5070**, p. 44, 2003

3. D. M. Mittleman, M. Gupta, R. Neelamani, R. G. Baraniuk, J. V. Rudd, and M. Koch, "Recent advances in terahertz imaging," *Appl. Phys. B,* **68**, (1999) p. 1085.
4. D. M. Mittleman, R. H. Jacobsen, and M. C. Nuss, "T-Ray Imaging," *IEEE J. Quantum Electron.* **20**, (1996) p. 679.

This page intentionally left blank

CHAPTER 14

ACTIVE INFRARED THERMOGRAPHY TECHNIQUES FOR THE NONDESTRUCTIVE TESTING OF MATERIALS

Clemente Ibarra-Castanedo[1], Marc Genest[2], Jean-Marc Piau[1], Stéphane Guibert[1], Abdelhakim Bendada[1], and Xavier P. V. Maldague[1]

[1]*Computer Vision and Systems Laboratory, Laval University, Quebec City (Quebec) G1K 7P4, Canada*
{IbarraC, JMPiau, Guibert, Bendada, MaldaguX}@gel.ulaval.ca
[2]*Institute for Aerospace Research (IAR), National Research Council Canada (NRC) 1200 Montreal Road, Bldg. M-14, Room 130, Ottawa, ON, Canada, K1A 0R6*
Marc.Genest@nrc-cnrc.gc.ca

Active infrared thermography refers to the group of methods employed to inspect the integrity of materials or systems through the use of an external energy source and an infrared detector. The external stimulus can be of many forms such as warm or cold air, heat pulses, periodic thermal waves, or mechanical oscillations, *e.g.* ultrasounds. The way data is captured and processed, as well as the typical applications differ according to the excitation source. This chapter presents a review of three of the most common active techniques in the field of thermography: lock-in thermography, pulsed thermography and vibrothermography.

1. Introduction

Infrared (IR) thermography is a nondestructive testing (NDT) technique, *i.e.* an inspection method for the examination of a part, material or system without impairing its future usefulness [1]. When compared with other classical NDT techniques such as C-Scan ultrasonics or X-Rays, data acquisition by infrared thermography is safe, nonintrusive and noncontact, allowing the detection of relatively shallow subsurface defects (a few millimeters in depth) under large surfaces (typically 30x30 cm^2 at once, although inspection of larger surfaces is possible) and in a fast manner (from a fraction of a second to a few minutes depending in the configuration, see below). Any object above the absolute zero temperature will emit IR radiation. Infrared radiation corresponds to the band of

the electromagnetic spectrum between 0.74 and 1000 μm. Hence, an excellent way to measure thermal variations is to use an IR radiometer, usually a focal plane array (FPA) camera capable of detecting radiation in the mid (3 to 5 μm) or long (8 to 14 μm) wave infrared bands, denoted as MWIR and LWIR, respectively, corresponding to two of the high transmittance atmospheric windows as shown in Figure 1.

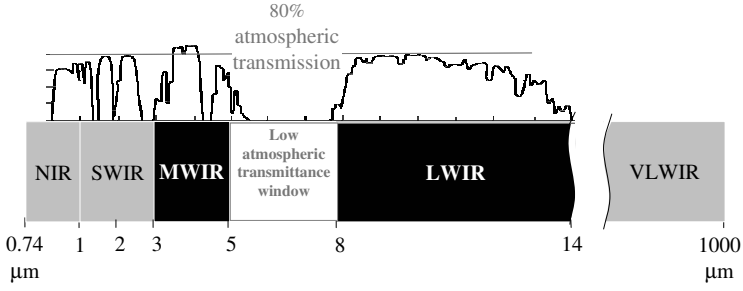


Figure 1. The infrared bands in the electromagnetic spectrum.

There are two approaches in infrared thermography as seen in Figure 2: (1) *passive*, in which the features of interest are naturally at a higher or lower temperature than the background, *e.g.* surveillance of people on a scene; and (2) *active*, in which an energy source is required to produce a thermal contrast between the feature of interest and the background, *e.g.* a specimen with internal flaws.

The active approach is adopted in many cases given that the inspected parts are usually in equilibrium with the surroundings.

A wide variety of energy sources can be used to induce a thermal contrast between defective and non-defective zones that can be divided in *external*, if the energy is delivered to the surface and then propagated through the material until it encounters a flaw; or *internal*, if the energy is injected into the specimen in order to stimulate exclusively the defects. Typically, external excitation is performed with optical devices such as photographic flashes (for heat pulsed stimulation) or halogen lamps (for periodic heating), whereas internal excitation can be achieved by means of mechanical oscillations, with a sonic or ultrasonic transducer for both burst and amplitude modulated stimulation.

As depicted in Figure 2, there are three classical active thermographic techniques based on these two excitation modes: lock-in thermography and pulsed thermography, which are optical techniques applied externally; and vibrothermography, which uses ultrasonic waves (amplitude modulated or pulses) to excite internal features. Lock-in thermography is presented first.

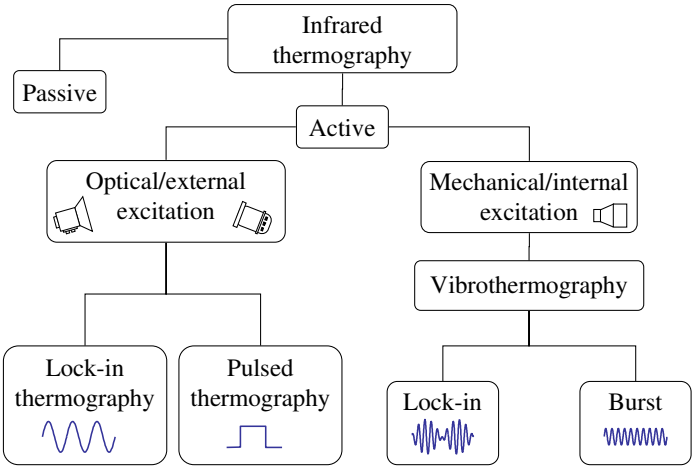


Figure 2. Infrared thermography approaches.

2. Lock-in Thermography

2.1. *Basic Theory for Lock-in Thermography*

Lock-in thermography (LT) also known as *modulated thermography* [2] is a technique derived from *photothermal radiometry* [3], in which, a small surface spot is periodically illuminated by an intensity modulated laser beam to inject thermal waves into the specimen. The thermal response is recorded at the same time using an infrared detector and decomposed by a lock-in amplifier to extract the amplitude and phase of the modulation [4]. Photothermal radiometry was a raster point-by point technique that required long acquisition times (especially in the case of deep defects involving very low modulation frequencies, see below). Furthermore, extra hardware, *i.e.* lock-in amplifier, is needed in order to retrieve the amplitude and phase of the response.

Fortunately, it is possible to dramatically simplify and speed up the acquisition process for NDT applications by replacing: (1) the laser beam with one or several modulated heating sources, *e.g.* halogen lamps, that cover the entire specimen surface instead of only a point; (2) the infrared detector with an infrared camera capable of monitoring the whole (or a large part of the) surface (typically in a 320x256 or 640x512 pixel matrix configuration); and (3) the lock-in hardware with a software capable of recovering mathematically the amplitude and phase of the response. This is what is called lock-in thermography [5].

Sinusoidal waves are typically used in LT, although other periodic waveforms are possible. Using sinusoids as input has the advantage that the frequency and shape of the response are preserved; only the amplitude and phase delay of the wave may change (*i.e.* sinusoidal fidelity). The periodic wave propagates by radiation through the air until it touches the specimen surface where heat is produced and propagates through the material. Internal defects act as barrier for heat propagation, which produces changes in amplitude and phase of the response signal at the surface.

Heat diffusion through a solid is a complex 3D problem that can be described by the Fourier's law of heat diffusion (or the heat equation) [6]:

$$\nabla^2 T - \frac{1}{\alpha} \cdot \frac{\partial T}{\partial t} = 0 \tag{1}$$

where $\alpha = k/\rho c_P$ [m^2/s] is the thermal diffusivity of the material being inspected, k [W/mK] its thermal conductivity, ρ [kg/m^3] its density and c_P [J/kgK] its specific heat at constant pressure.

The Fourier's law 1D solution for a periodic thermal wave propagating through a semi-infinite homogeneous material may be expressed as [7]:

$$T(z,t) = T_o \exp(-\frac{z}{\mu}) \cos\left(\frac{2\pi \cdot z}{\lambda} - \omega t\right) \tag{2}$$

where T_0 [°C] is the initial change in temperature produced by the heat source, $\omega = 2\pi f$ [rad/s] is the modulation frequency, f [Hz] is the frequency, $\lambda \equiv 2\pi\mu$ [m] is the thermal wavelength; and μ [m] is the thermal diffusion length, which determines the rate of decay of the thermal wave as it penetrates through the material and is defined by [7]:

$$\mu \equiv \sqrt{\frac{2 \cdot \alpha}{\omega}} = \sqrt{\frac{\alpha}{\pi \cdot f}} \tag{3}$$

Figure 3 depicts a thermal wave as it travels from the surface through a solid. As can be seen, after the waveform travels only a distance equal to a thermal diffusion length μ, it has been already reduced to 1/535 its initial intensity. From Eq. (3), in order to cover a distance equal to a thermal wavelength λ, a thermal wave at f=10 Hz will be able to travel less than 0.4 mm through plastic, slightly more than 11 mm through aluminum (see Figure 3) and almost 16 mm through air. Hence, thermal waves propagate deeper in more diffusive materials. On the other hand, information about deeper features is available when lower frequencies are used.

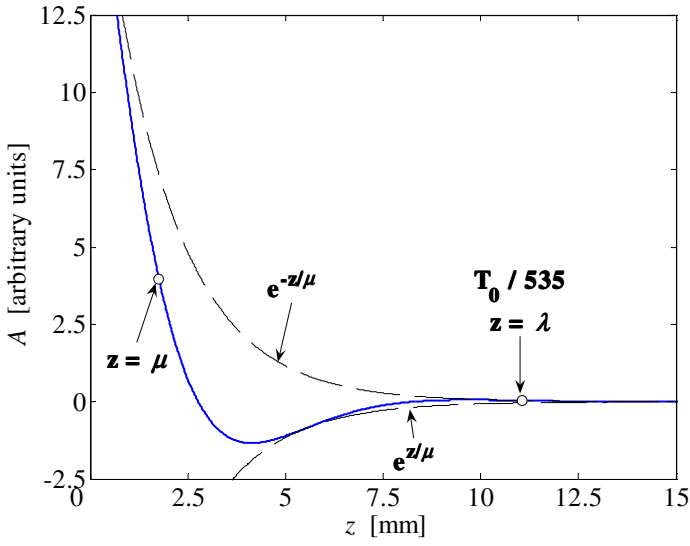


Figure 3. Thermal wave propagation by conduction through an aluminum semi-infinite plate at 10 Hz.

These two aspects are important to know when planning the inspection of a part in order to correctly select the working frequency and to determine the depth of internal defects as will be pointed out in section 2.3. The next paragraph describes the experimental setup and the acquisition system in more detail.

2.2. *Experimental Setup and Data Acquisition for Lock-in Experiments*

Figure 4 depicts a lock-in thermography experiment. Two lamps are shown although it is possible to use several lamps mounted on a frame to reduce the non-uniform heating and/or to increase the amount of energy delivered to the surface. The lamps send periodic waves (*e.g.* sinusoids) at a given modulation frequency ω, for at least one cycle, ideally until a steady state is achieved, which depends on the specimen's thermal properties and the defect depth, see Eq. (3) above. In practice however, only a few cycles are needed to adequately retrieve phase and amplitude data, much before attaining steady state conditions.

Figure 5 shows an example of the raw output signal for a sinusoidal input at two different points. As can be seen, noise is omnipresent and processing is required not only to extract the amplitude and/or phase information but also to de-noise the signal.

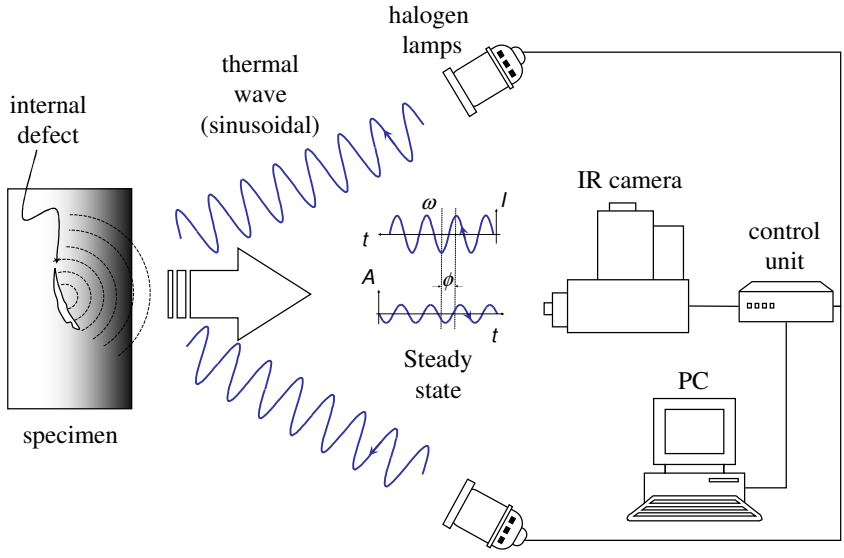


Figure 4. Experimental set-up for lock-in thermography.

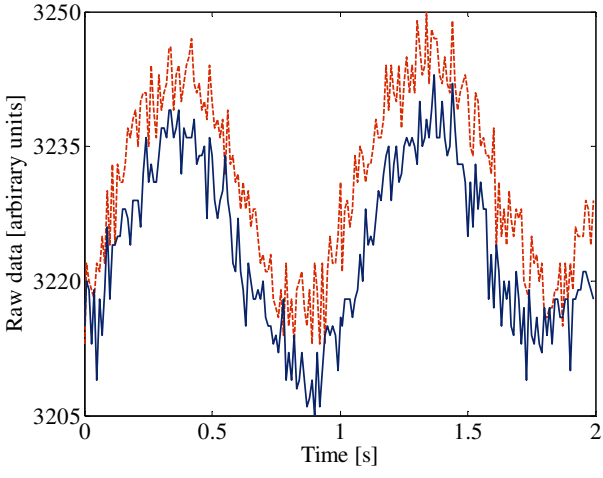


Figure 5. Raw data output signal at f=1 Hz for 4 different points.

In the following paragraph, we present two processing techniques for LT.

2.3. *Processing Lock-in Thermography Data*

A four point methodology for sinusoidal stimulation is available [5], [8] and illustrated in Figure 6.

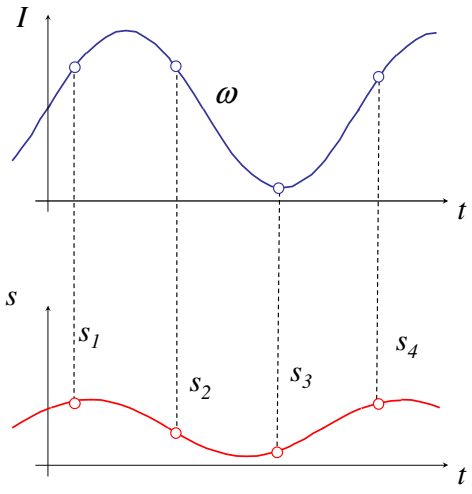


Figure 6. Four point methodology for amplitude and phase delay estimation by lock-in thermography.

The sinusoidal input signal *I*, is represented on top of Figure 6, the response signal *S* is depicted at the bottom. As mentioned before, input and output have the same shape when sinusoids are used, there is only a change in amplitude and phase that can be calculated as follows [8]:

$$A = \sqrt{(S_1 - S_3)^2 + (S_2 - S_4)^2} \tag{4}$$

$$\phi = \arctan\left(\frac{S_1 - S_3}{S_2 - S_4}\right) \tag{5}$$

The 4-point method is fast but it is valid only for sinusoidal stimulation and is affected by noise. The signal can be de-noised in part by averaging of several points instead of a single one and/or by increasing the number of cycles. Another possibility is to fit the experimental data using least squares regression [9] and to use this synthetic data to calculate the amplitude and the phase. This two alternatives contribute to slow down the calculations. Alternatively, the discrete Fourier transform (DFT) can be used to extract amplitude and phase information from LT data. Initially proposed for pulsed thermography data [10] (see pulsed phase thermography below), the DFT can be written as [11]:

$$F_n = \Delta t \sum_{k=0}^{N-1} T(k\Delta t) \exp^{(-j2\pi nk/N)} = \mathrm{Re}_n + \mathrm{Im}_n \tag{6}$$

where j is the imaginary number (j^2=-1), n designates the frequency increment (n=0,1,…N); Δt is the sampling interval; and Re and Im are the real and the imaginary parts of the transform, respectively.

In this case, real and imaginary parts of the complex transform are used to estimate the amplitude and the phase [10]:

$$A_n = \sqrt{\mathrm{Re}_n^2 + \mathrm{Im}_n^2} \tag{7}$$

$$\phi_n = \tan^{-1}\left(\frac{\mathrm{Im}_n}{\mathrm{Re}_n}\right) \tag{8}$$

The DFT can be use with *any* waveform (even transient signals as in pulsed phase thermography and burst phase vibrothermography, see below) and has the advantage of de-noising the signal. Although very useful, Eq. (6) is slow. Fortunately, the fast Fourier transform (FFT) algorithm is available [12] to be implemented or can be found (integrally or simplified) in common software packages.

One of the most interesting characteristics of LT testing is the possibility of performing quantitative operations in a straightforward manner. As discussed at the end of section 2.1, a direct relationship exists between the depth of a defect and the thermal diffusion length, μ [13], *i.e.* Eq. (3). For instance, empirical expressions have been proposed [14], [15] to extract the depth of a given defect from LT data using a relationship of the form:

$$z = C_1 \cdot \mu = C_1 \cdot \sqrt{\frac{\alpha}{\pi \cdot f_b}} \tag{9}$$

where f_b is known as the blind frequency [16], *i.e.* the frequency at which a given defect have enough (phase or amplitude) contrast to be detected (at frequencies higher than f_b is not possible to detect it), C_1 is an empirical constant.

It has been observed that, C_1=1 when using amplitude data [13], whilst reported values when working with the phase are in the range $1.5<C_1<2$ [13]-[17], with C_1=1.82 typically adopted [14], [15] following the research work by Thomas et al., 1980 [18]. The phase is therefore of more interest in NDT than the amplitude, since it provides deeper probing capabilities. Moreover, the ratio involved in Eqs. (5) and (8), cancels part of the artifacts related to non-uniform heating, environmental reflections, emissivity variations and surface geometry variations.

2.4. *Advantages, Disadvantages and Applications*

Given that LT requires to perform an experiment for each and every inspected depth and there is a stabilization time before reaching a permanent regime, inspection by lock-in thermography is in general slower than other approaches such as pulsed thermography, which will be described in the following section. A complete LT experiment is carried out by inspecting the specimen at several frequencies, covering a wide range from low to high frequencies, and then a fitting function can be used to complete the amplitude or phase profiles for each point (*i.e.* each pixel). Nevertheless, there exists a direct relationship between depth and the inspection frequency that allows depth estimations to be performed from amplitude or phase data without further processing. Furthermore, the energy required to perform an LT experiment is generally less than in other active techniques, which might be interesting if a low power source is to be used or if special care has to be given to the inspected part, *e.g.* cultural heritage pieces, works of art, frescoes, etc. [19].

Some typical applications include determination of coatings thickness, detection of delaminations, determination of local fiber orientation, corrosion detection and inspection of cultural properties. [4], [19].

3. Pulsed Thermography

3.1. *Basic Theory for Pulsed Thermography*

In pulsed thermography (PT), the specimen surface is submitted to a heat pulse using a high power source such as photographic flashes. A heat pulse can be though as the combination several periodic waves at different frequencies and amplitudes. After the thermal front come into contact with the specimen's surface, a thermal front travels from the surface through the specimen. As time elapses, the surface temperature will decrease *uniformly* for a piece without internal flaws. On the contrary, subsurface discontinuities (*e.g.* porosity, delaminations, disbonds, fiber breakage, inclusions, etc.), can be though as resistances to heat flow that produce abnormal temperature patterns at the surface, which can be detected with an IR camera.

The 1D solution of the Fourier equation for the propagation of a *Dirac* heat pulse, *i.e.* an ideal waveform defined as an intense unit-area pulse of so brief duration that no measuring equipment is capable of distinguishing it from even shorter pulses [11], in a semi-infinite isotropic solid by conduction has the form [6]:

$$T(z,t) = T_0 + \frac{Q}{\sqrt{k\rho c_P \pi t}} \exp(-\frac{z^2}{4\alpha t}) \tag{10}$$

where Q [J/m^2] is the energy absorbed by the surface and T_0 [K] is the initial temperature.

A Dirac heat pulse is composed of periodic waves at *all* frequencies and amplitudes. Although it is not possible to reproduce such a waveform in practice, a heat pulse provided by a powerful source such as a photographic flash having approximately a square shape, can be used. In this case, the signal is composed by periodic waves at *several* (but not all) frequencies. The shorter the pulse, the broader the range of frequencies. This is very useful for understanding the link between lock-in and pulse thermography, which give rise to pulsed phase thermography that will be discussed in paragraph 3.3.

From Eq. (10), the thermal profiles for an aluminum specimen subjected to a uniform heat pulse follows the behavior portrayed in Figure 7.

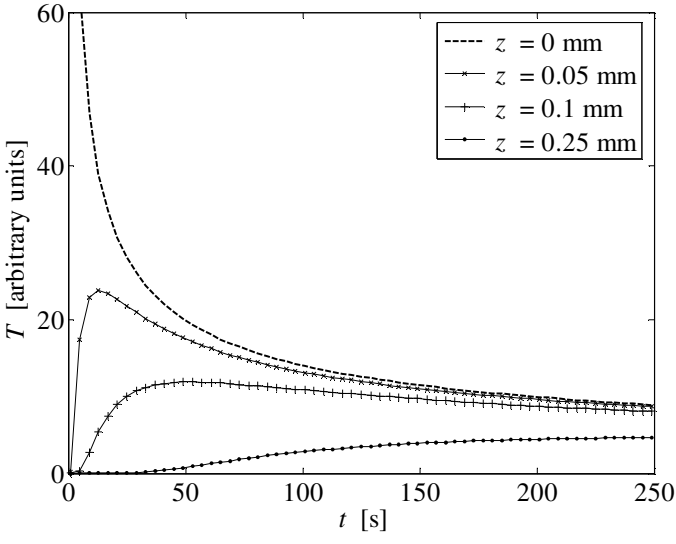


Figure 7. Thermal profiles at four depths (z=0, 0.05, 0.1 and 0.25 mm) for an aluminum plate subjected to a uniform heat pulse.

The profiles correspond to 4 depths: z=0, 0.05, 0.1 and 0.25 mm. Thermal evolution for deep defects starts at zero and reaches peak intensity at a given time (longer for deeper defects) and then slowly decays approximately as the square root of time. Shallower defects show higher peak intensities earlier. At the surface the behavior is somehow different starting at a high temperature and

monotonically decaying following approximately the square root of time. At the surface (z=0 mm), Eq. (10) can be rewritten as follows:

$$T(0,t) = T_0 + \frac{Q}{e\sqrt{\pi t}} \tag{11}$$

where $e \equiv (k\rho c_P)^{1/2}$ [m] is the effusivity, which is a thermal property that measures the material ability to exchange heat with its surroundings.

Although Eq. (11) is only an approximation of the complex 3D diffusion problem described by Fourier's law, *i.e.* Eq. (1), many of the processing techniques have been based on this simplification to perform qualitative and quantitative analysis, as will be discussed in paragraph 3.3. The experimental setup is discussed first.

3.2. *Experimental Setup and Data Acquisition for Pulsed Thermography*

When compared to LT, data acquisition in PT is fast and straightforward as illustrated in Figure 8.

Two photographic flashes are used to heat up the specimen's surface, after what, the thermal changes are recorded with an infrared camera. A synchronization unit is needed to control the time between the launch of the thermal pulse and the recording with the IR camera. Data is stored as a 3D matrix (see Figure 9a) where x and y are the spatial coordinates, and t is the time. Temperature decreases approximately as $t^{1/2}$ (at least at early times), as predicted by Eq. (11), except for the defective areas, where the cooling rate is different (see Figure 9b).

Although heat diffusion is a complex problem, the relationship between defect depth and time, simplified through Eq. (11), has been exploited by many researchers to develop qualitative and quantitative techniques. Next section presents three techniques that are used to analyze PT data. These techniques were considered to be some of the most promising ones among many others and they are currently subjected to extensive investigation.

3.3. *Processing Pulsed Thermography Data*

PT is probably the most extensively investigated approach because of its ease of deployment. Raw PT data however, is difficult to handle and analyze. There are a great variety of processing techniques that have been developed to enhance the subtle IR signatures [10], [20], [21]. Space being limited, it is only possible to

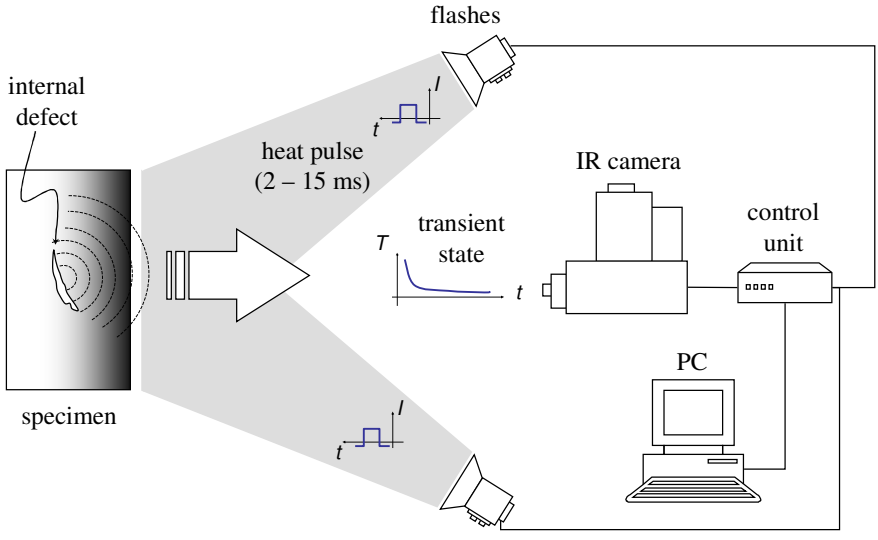


Figure 8. Experimental set-up for pulsed thermography.

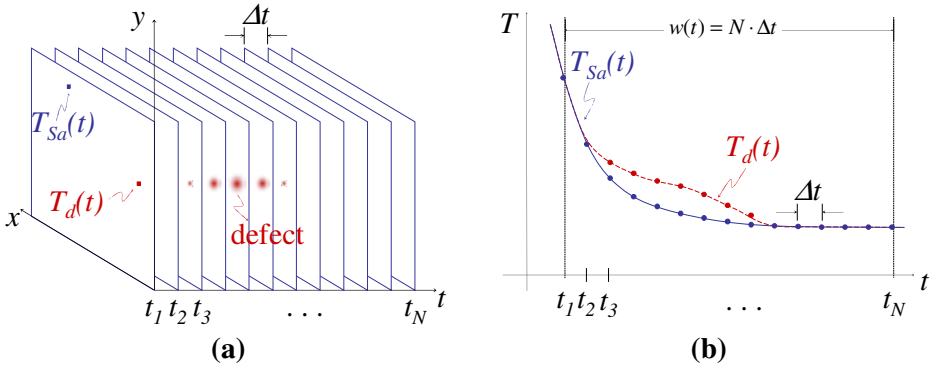


Figure 9. Temperature evolution: (a) data 3D matrix, and (b) temperature profile for a defective (dotted line) and non-defective (continuous line) pixels.

discuss a few of them. We selected three techniques that have shown very promising results for most common applications. Although only a brief discussion is provided; interested readers may consult the references provided below.

3.3.1. *Thermal Contrast Based Techniques*

Thermal contrast is a basic operation that despite its simplicity is at the origin of most of the PT analysis. Various thermal contrast definitions exist [22] p. 198,

but they all share the need to specify a sound area S_a, *i.e.* a non-defective region within the field of view. For instance, the *absolute* thermal contrast $\Delta T(t)$ is defined as [22]:

$$\Delta T(t) = T_d(t) - T_{S_a}(t) \tag{12}$$

with $T(t)$ being the temperature at time t, $T_d(t)$ the temperature of a pixel or the average value of a group of pixels, and $T_{Sa}(t)$ the temperature at time t for the S_a. No defect can be detected at a particular t if $\Delta T(t)$=0.

The main drawback of classical thermal contrast is establishing S_a, especially if automated analysis is needed. Even when S_a definition is straightforward, considerable variations on the results are observed when changing the location of S_a as is well-known [23].

In the *differential absolute contrast* (DAC) method [24], instead of looking for a non-defective area, an ideal S_a temperature at time t is computed locally assuming that on the first few images (at least one image at time t' in particular, see below) this local point behaves as a S_a in accordance to Eq. (11), *i.e.* there is no defect visible. The first step is to define t' as a given time value between the instant when the pulse has been launched t_0, and the precise moment when the first defective spot appears on the thermogram sequence, *i.e.* when there is enough contrast for the defect to be detected, t_1. At t', there is no indication of the existence of a defective zone yet, therefore the local temperature for a S_a is exactly the same as for a defective area [25]:

$$T_{S_a}(t') = T(t') = \frac{Q}{e\sqrt{\pi t'}} \quad \Rightarrow \quad \frac{Q}{e} = \sqrt{\pi t'} \cdot T(t') \tag{13}$$

From this result, T_{Sa} can be computed for every pixel at time t. Substituting Eq. (13) into the absolute contrast definition, *i.e.* Eq. (12), it follows that [25]:

$$\Delta T_{dac} = T(t) - \sqrt{\frac{t'}{t}} \cdot T(t') \tag{14}$$

Actual measurements diverge from the solution provided by Eq. (14) as time elapses and also as plate thickness enlarges with respect to the non-semi-infinite case. Nevertheless, it has proven to be very efficient in reducing artifacts from non-uniform heating and surface geometry [20], and to provide a good approximation even for the case of anisotropic materials [20] at early times. Originally, proper selection of t' required an iterative graphical procedure, for which a graphical user interface was developed [26]. An automated algorithm is now available [27]. Furthermore, a modified DAC technique has been developed

as well [28]. It is based on a finite plate model, which includes the plate thickness L explicitly in the solution, extending in this way the validity of the DAC algorithm to later times.

3.3.2. *Pulsed Phase Thermography*

Pulsed phase thermography (PPT) [10], [29] is another interesting technique, in which data is transformed from the time domain to the frequency spectra using the one-dimensional discrete Fourier transform (DFT), *i.e.* Eq. (6). The use of the DFT, or more precisely the FFT on thermographic data was first proposed by Maldague and Marinetti in 1996 [10]. Since then, it has been applied to other thermographic data, such as lock-in, described in section 2 and for vibrothermography data as will be described in section 4. As for the case of LT, the amplitude and the phase can be computed from Eqs. (7) and (8). The frequency components can be derived from the time spectra as follows:

$$f = \frac{n}{N \cdot \Delta t} \tag{15}$$

where n designates the frequency increment (n=0,1,…N), Δt is the time step and N is the total number of frames in the sequence.

Phase ϕ is of particular interest in NDE given that it is less affected than raw thermal data by environmental reflections, emissivity variations, non-uniform heating, surface geometry and orientation. These phase characteristics are very attractive not only for qualitative inspections but also for quantitative characterization of materials. For instance, a depth inversion technique using the phase from PPT and Eq. (9) is available [30].

The FFT is typically used to extract amplitude and phase information. Nevertheless, it is also possible to used different transformation algorithms such as the wavelet transform [31]. The later has the additional advantages of preserving the temporal information after the transformation and to use *wavelets* as the basis function instead of sinusoids. Wavelets are periodic waves of short duration that allow to better reproduce a transient signal and to used different scales or resolutions [32]. These advantages of the wavelet transform are currently under investigation.

3.3.3. *Thermographic Signal Reconstruction*

Thermographic signal reconstruction (TSR) [33] is an attractive technique that allows increasing spatial and temporal resolution of a sequence, reducing at the same time the amount of data to be manipulated. TSR is based on the assumption

that, temperature profiles for non-defective pixels should follow the decay curve given by the one-dimensional solution of the Fourier Equation, *i.e.* Eq. (2), which may be rewritten in the logarithmic form as:

$$\ln(\Delta T) = \ln\left(\frac{Q}{e}\right) - \frac{1}{2}\ln(\pi t) \tag{16}$$

Next, an *n*-degree polynomial is fitted for each pixel [33]:

$$\ln(\Delta T) = a_0 + a_1 \ln(t) + a_2 \ln^2(t) + ... + a_p \ln^p(t) \tag{17}$$

Typically, *n* is set to 4 or 5 to avoid "ringing" and to insure a good correspondence between data and fitted values. Synthetic data processing brings interesting advantages such as: significant noise reduction, possibility for analytical computations, considerably less storage is required since the whole data set is reduced from *N* to *p*+1 images (one per polynomial coefficient), and calculation of first and second time derivatives from the synthetic coefficients is straightforward.

3.4. *Advantages, Disadvantages and Applications*

Pulsed thermography is fast (from a few seconds for high conductivity materials to a few minutes for low conductivity materials) and easy to deploy. There are numerous processing techniques available although many of them are complex when compared to lock-in. Thermal based techniques are affected by non-uniform heating, emissivity variations, environmental reflections and surface geometry. These problems however, are dramatically reduced using advanced processing algorithms, *e.g.* PPT, DAC and TSR. For instance PPT allows the recovering of amplitude and phase data as in LT with the advantage that, since a heat pulse can be seen as a set of several periodic thermal waves launched at once, several data points (amplitude or phase) can be extracted from a single experiment. Nevertheless, as described in [34], special care must be given to the selection of the sampling and truncation parameters in order to perform quantitative calculations by PPT.

4. Vibrothermography

4.1. *Basic Theory for Vibrothermography*

Vibrothermography (VT), also known as *ultrasound thermography* [35] or *thermosonics* [36], utilizes mechanical waves to directly stimulate internal

defects and without heating the surface as in optical methods (*e.g* LT and PT). If we though of optical lock-in thermography to be the successor of photothermal radiometry as mentioned before, vibrothermography can be seen as the successor of *optoacoustics* or *photoacoustics* [13], [37], [38], in which microphones or piezoceramics in contact with the specimen and a lock-in amplifier were used to detect the thermal wave signature from a defect. This was performed in a point-by-point manner and lack of practical interest at the time. The theory behind however, was the base for the later development of VT.

In the classical ultrasound C-scan NDT inspection, a transducer is placed in contact with the sample with the help of a coupling media. The ultrasonic waves travel through the specimen and are transmitted back to the surface where the transducer pick-up the reflected signal (pulsed-echo technique), or they are collected on the opposite side (transmission). In any case, the principle of defect detection is based on the differences in specific acoustic impedances (Z) between materials. In VT, ultrasonic waves will travel freely through a homogeneous material, whereas an internal defect will produce a complex combination of absorption, scattering, beam spreading and dispersion of the waves, whose principal manifestation will be in the form of heat. Heat then will travel by conduction in all directions, an IR camera can be directed to one of the surfaces of the specimen to capture the defect signature. Ultrasonic waves are ideal for NDT since they are not audible to humans (although some low frequencies harmonics are present), defect detection is independent from of its orientation inside the specimen, and both internal and open surface defects can be detected. Hence, VT is very useful for the detection of cracks and delaminations.

Sonic waves, audible for humans, vibrate between 20 Hz and 20 kHz. The range for ultrasonic waves, no audible to humans, is between 20 kHz and 1 MHz, although most transducers operate between 15 and 50 kHz. Unlike electromagnetic waves, mechanical elastic waves such as sonic and ultrasonic waves do not propagate in a vacuum; on the contrary, they require a medium to travel. They travel faster in solids and liquids than through the air. This brings about an important aspect of VT: although contactless (through air) ultrasonics is currently under intense investigation in many areas [39], the common approach in VT is to use a coupling media such as a piece of fabric, water-based gels or aluminum, between the transducer and the specimen to reduce losses.

The next paragraph describes the experimental setup for a VT experiment.

4.4. *Experimental Setup and Data Acquisition for Vibrothermography*

There are basically two configurations for VT that can be sought as analog to optical methods described above. The first one is lock-in vibrothermography (or

amplitude modulated VT), analog to the LT approach; and the second technique is burst vibrothermography, which is analog to PT. These two approaches are illustrated in Figure 10 and Figure 11.

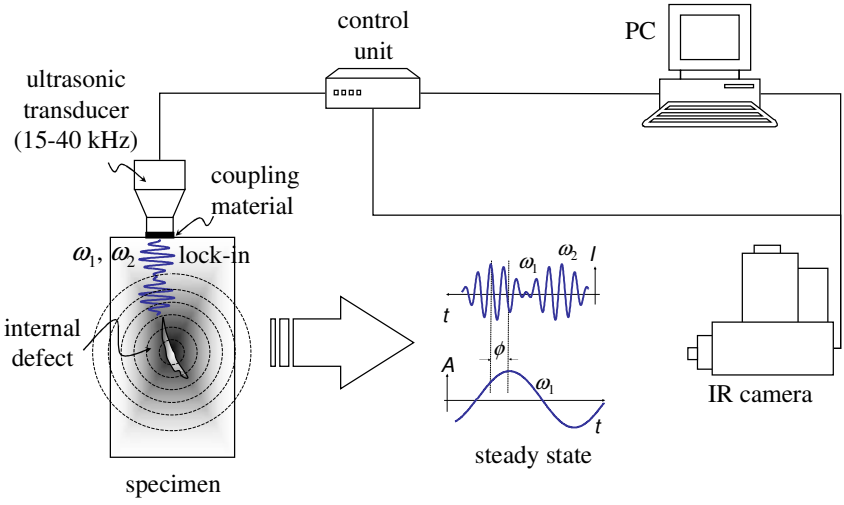


Figure 10. Configuration for lock-in vibrothermography.

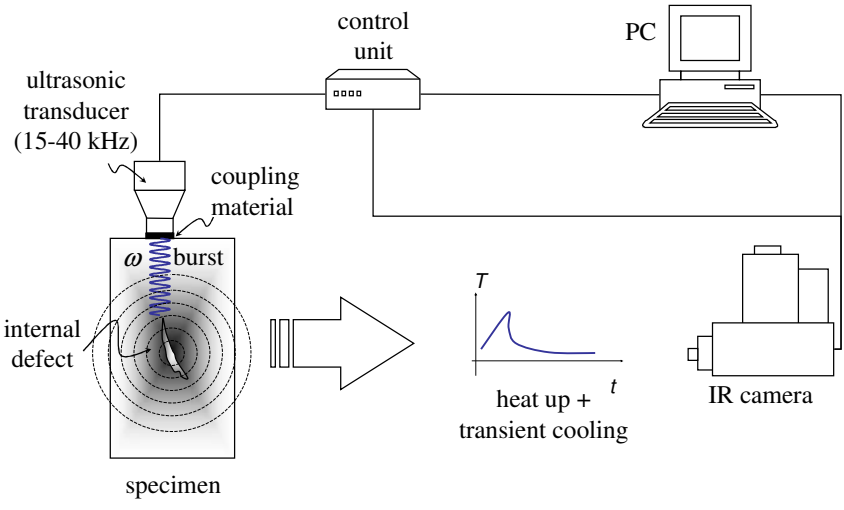


Figure 11. Configuration for burst vibrothermography.

It should be mentioned that is also possible to modulate the frequency either in lock-in or burst VT [40]. This approach is sometimes called *wobbulation*. The idea is to cover a range of ultrasonic frequencies, instead of only one, since it is

not always possible to predict the right frequency for a particular application. Ultrasonic wobbulation can be compared to a heat pulse, which is composed of thermal waves at many frequencies. Wobbulation is useful as well to prevent the appearance of *standing waves*, which are produced when working at the natural harmonics resonance frequency of the material. In practice however, it is sometimes preferable to repeat the acquisition at a different frequency since the commercial transducers commonly used are not suitable for frequency modulations.

The ultrasound wave is produced by a transducer made of a stack of piezo elements and concentrated in a titanium horn that acts like a hammer. Hence, the part being inspected should be firmly immobilized (but without damaging) to avoid cantilever effects, clapping and sliding of the transducer. The transducer horn should be pressed against the sample as well to improve the coupling transmission of the ultrasound into the specimen. Insertion of a material between the transducer and the sample is strongly recommended not only as a coupling medium, but also to avoid damage of the sample and correct misalignment. A bad coupling implies a poor ultrasound transmission but more seriously it creates unwanted heat in the vicinity of the ultrasound injection point.

After the elastic waves are injected to the specimen, they travel through the material and dissipate their energy mostly at the defects so heat is locally released. The thermal waves then travel by conduction to the surface, where they can be detected with an IR camera.

When compared to optical/external techniques, the thermal wave travels half the distance in a VT experiment since heat propagation is performed from the defect to the surface, whereas for optical techniques heat travels from the surface to the defects and back to the surface. Hence, VT is very fast, even faster than PT. A typical experiment last from a fraction of a second to several seconds. In addition, the longer the transducer operates at the surface; the most heat is released at the contact surface, increasing the probability of damaging the area. Furthermore, the pressure applied between the horn and the specimen has a great impact on the thermal response [41].

4.3. *Processing Vibrothermography Data*

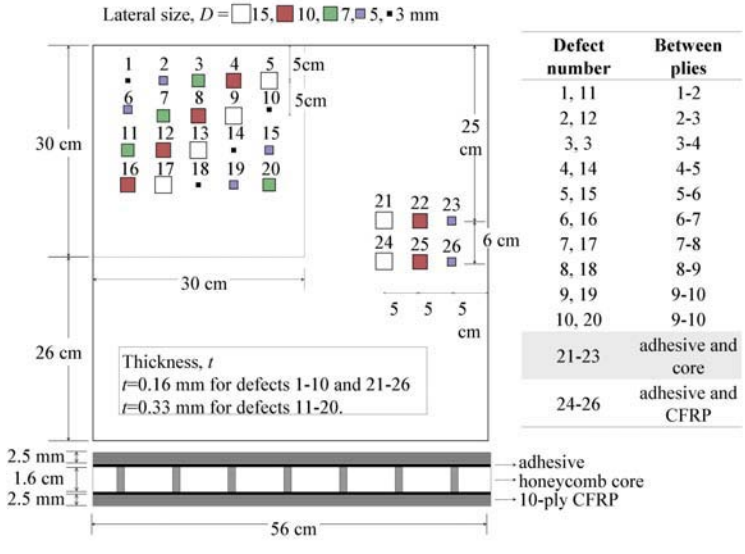
Although raw thermograms present sometimes good enough contrast to detect defects, some processing is required most of the time. As for the case of LT and PT, the FFT algorithm can used and amplitude and phase images are recovered through Eqs. (7) and (8).

4.4. *Advantages, Disadvantages and Applications*

In either lock-in or burst configuration, VT is extremely fast, although it is necessary to relocate the transducer (and to immobilize the specimen again) to cover a large area for inspection. Hence, VT is more suitable for relatively small objects. It is the most appropriate technique to inspect some types of defects, *e.g.* micro cracks. On the contrary, it does not perform very well in some other cases in which application of optical techniques are straightforward, *e.g.* water detection. But probably the most inconvenient aspect of VT is the need of a coupling media between the sample and the transducer, and the need of holding the specimen. On the other hand, there is only minimal heating of the inspected specimen since energy is usually dissipated mostly at the defective areas, although there is some localized heating at the coupling and clamping points.

5. Comparative Results

A honeycomb sandwich panel, shown Figure 12, was inspected by the tree active techniques described above. The panel consists of a honeycomb core of 1.6 cm between two 10-ply CFRP. Several Teflon® inserts of different dimensions and thickness as specified in the figure are placed at different locations and depths. The group of defects at the top (defects 1 to 20) are analyzed first. It is possible to detect almost all of the defects by PPT as can be seen from Figure 13a and b.



Defect number	Between plies
1, 11	1-2
2, 12	2-3
3, 3	3-4
4, 14	4-5
5, 15	5-6
6, 16	6-7
7, 17	7-8
8, 18	8-9
9, 19	9-10
10, 20	9-10
21-23	adhesive and core
24-26	adhesive and CFRP

Figure 12. The honeycomb sandwich panel.

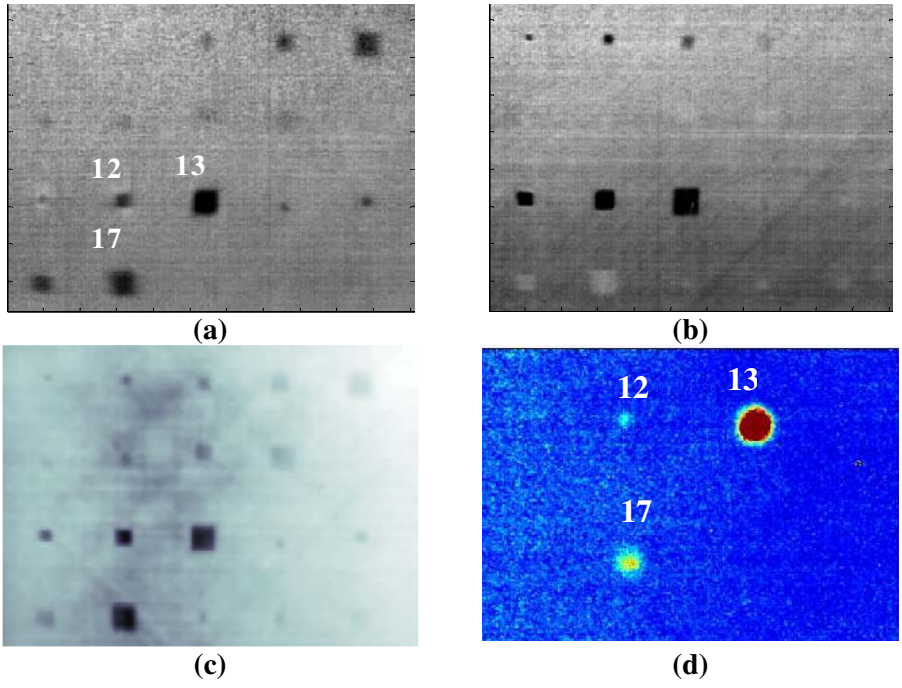


Figure 13. Comparative results for the honeycomb sandwich panel, shallowest group of defects. Synthetic PPT phase from a 7th degree polynomial at f =(a) 0.082 and (b) 0.74 Hz, (c) LT phase at f =0.001 Hz, and (d) VT result.

Contrast is better for shallower defects, although better results could be obtained for deeper defects using longer acquisition times using the same time resolution. The LT phasegram at f=0.001 Hz is presented in Figure 13c. Compared to PPT results, contrast is better for most of the defects and worse for a few others, but with the exception of defect number 10, all defects can be seen in a single phasegram. On the other hand, only the thicker (L=0.33 mm) inserts could be detected by VT. Defect number 12 is 0.5 mm deep and 10 mm large, defect 13 is 0.75 mm deep and 15 mm large, whilst defect 17 is 1.75 mm deep and 15 mm large.

Results for the deepest set of inserts (defects 21 to 26) are seen in Figure 14. While only assumptions of the location of these defect can be made from PT and LT results (Figure 14a and b, respectively), amplitude and phase images by VT (Figure 14c and d, respectively) clearly show the presence of the two largest defects (defects number 21 and 24 in Figure 12).

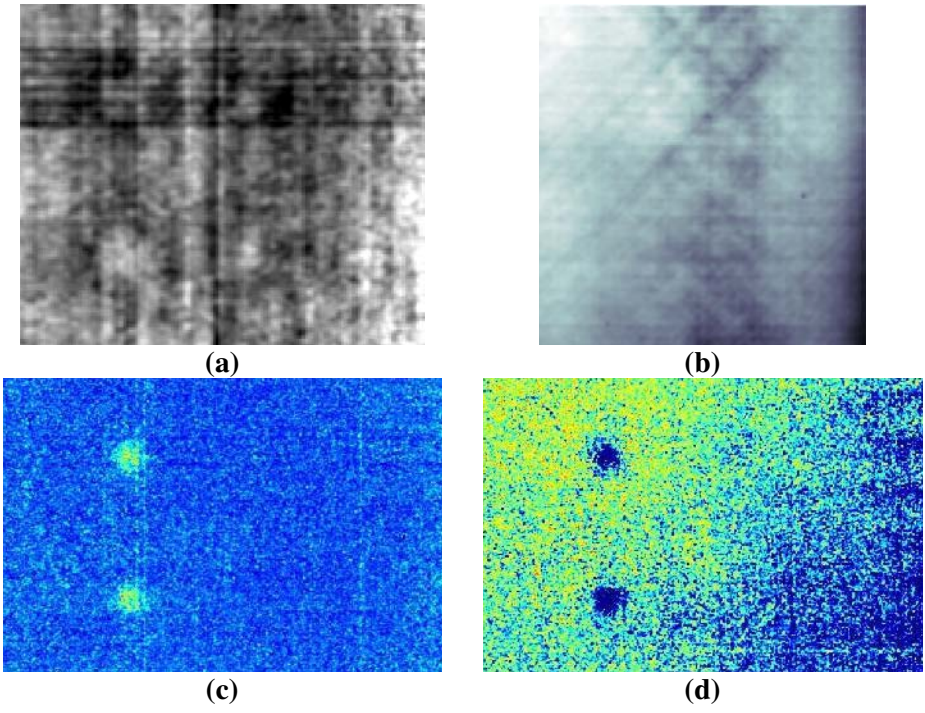


Figure 14. Comparative results for the honeycomb sandwich panel, deepest group of inserts. (a) PPT phase at f = 0.041 Hz, (b) LT phase at f =0.001 Hz, and second frequency components for the (c) amplitude and (d) phase data by VT.

Only the two largest defects can be detected by VT. This is probably due to the way the inserts where positioned during the fabrication process. The transition from the carbon-epoxy matrix to Teflon® is much less dramatic than the one from carbon-epoxy to air. Hence, probably in the case of the largest defects, there was more air trapped between the inserts and the glue promoting the dissipation of heat in the area. For the VT experiment, the IR camera was recording from the side with the inserts and the transducer was placed on the same side. Thermal waves only had to travel the distance from the location of the defects to the surface, while for the case of LT and PT, the thermal excitation came from the same side where data was collected, *i.e.* the thermal waves had to travel twice the distance when compared to VT.

6. Conclusions

The three main active thermography techniques can be used in the NDT assessment of materials and structures in many areas. Lock-in thermography

allows having a better control of the energy that is to be deposited on a surface. This is of critical importance for the inspection of delicate specimens such as cultural properties and works of art. In addition, defect characterization (depth retrieval, determination of thermal properties) is straightforward from the diffusion length equation. Lock-in thermography however, requires performing several experiments to completely cover the range of depths.

Pulsed thermography is faster and easier to deploy, although experiments are performed having less control over the injected energy. Raw data is commonly affected by problems such as uneven heating, emissivity variations, reflections from the environment and surface geometry sharp variations. Nevertheless, a wide variety of processing techniques are available, which allows not only carrying out qualitative but also quantitative inspections.

The vibrothermography approach is relatively new on the active thermography scene. It had showed however to be very attractive for the inspection of some types of defects. For instance, it is possible to detect deeper defects by vibrothermogrpahy than with optical means. This is because of the way heat is generated and transformed into heat. In VT, the thermal waves only have to travel half the distance (from the defect to the surface) than with optical methods (from the surface to the defect and back). Nevertheless, VT is still lacking of quantitative studies, and very often optimal inspection parameters must be found experimentally.

In summary, a group of active techniques is available for a wide variety of applications. The selection of the most adequate approach depends on the particular application and the available experimental and expertise resources.

References

[1] Maldague X. P. V., Streckert H. H and Trimm M. W. “Introduction to infrared and thermal testing: Part 1. Nondestructive testing,” in Nondestructive Handbook, Infrared and Thermal Testing, Volume 3, X. Maldague technical ed., P. O. Moore ed., 3rd edition, Columbus, Ohio, ASNT Press, 2001, 718 p.

[2] Giorleo G. and Meola C. “Comparison between pulsed and modulated thermography in glass-epoxy laminates”, *NDT&E International*, **35**:287–292, 2002.

[3] Nordal P. E. and Kanstand S. O. “Photothermal radiometry,” *Physica Scripta*, **20**:659–662, 1979.

[4] Busse G. “Techniques of infrared thermography: Part 4. Lockin thermography,” in Nondestructive Handbook, Infrared and Thermal Testing, Volume 3, X. Maldague technical ed., P. O. Moore ed., 3rd edition, Columbus, Ohio, ASNT Press, 2001, 718 p.

[5] Wu D. and Busse G. “Lock-in Thermography for NonDestructive Evaluation of Materials,” *Rev. Gén. Therm.*, **37**:693–703, 1998.

[6] Carslaw H. S. and Jaeger J. C. *Conduction of Heat in Solids*, 2nd edition, Clarendon Press, Oxford, 1986.

[7] Favro L. D. and Han X. "Thermal Wave Materials Characterization and Thermal Wave Imaging," in Birnbaum G., Auld B. A. (eds.): *Sensing for Materials Characterization, Processing and Manufacuring*, ASNT TONES, **1**:399–415, 1998.

[8] Busse G., Wu D. and Karpen W. "Thermal Wave Imaging with Phase Sensitive Modulated Thermography," *J. Appl. Phys.*, **71**(8):3962–3965, 1992.

[9] Krapez J. C. "Compared performances of four algorithms used for modulation thermography", *Proc. 4th Conference on Quantitative InfraRed Thermography - QIRT*, D. Balageas, G. Busse, C. Carlomagno (eds.), Eurotherm Seminar 60, Lodz, Pologne, September 7-10, 1998, 148–153.

[10] Maldague X. P. and Marinetti S. "Pulse Phase Infrared Thermography," *J. Appl. Phys.*, **79**(5):2694–2698, 1996.

[11] Bracewell R. The Fourier Transform and its Applications, McGraw-Hill, USA, 1965.

[12] Cooley J. W. and Tukey J. W. "An Algorithm for the Machine Calculation of Complex Fourier Series," *Mathematics of Computation*, **19**(90):297–301, 1965.

[13] Busse G. and Rosencwaig A. "Subsurface Imaging with Photoacoustics," *Appl. Phys. Lett.*, **36**(10):815-816, 1980.

[14] Meola C., Carlomagno G. M. and Giorleo G. "The Use of Infrared Thermography for Materials Characterization", *J. Mater. Process. Technol.*, **155–156**:1132–1137, 2004.

[15] Meola C. and Carlomagno G. M. "Recent Advances in the Use of Infrared Thermography", *Meas. Sci. Technol.*, **15**:27–58, 2004.

[16] Bai W. and Wong B. S. "Evaluation of Defects in Composite Plates under Convective Environments using Lock-In Thermography," *Meas. Sci. Technol.*, **12**:142–150, 2001.

[17] Busse G. "Nondestructive Evaluation of Polymer Materials," *NDT&E Int.*, **27**(5):253–262, 1994.

[18] Thomas R. L., Pouch J. J., Wong Y. H., Favro L. D. and Kuo P. K. "Subsurface Flaw Detection in Metals by Photoacustic Microscopy", *J. Appl. Phys.*, **51**(2): 1152–1156, 1980.

[19] Carlomagno G. M. and Meola C. "Comparison between thermographic techniques for frescoes NDT", *NDT&E International*, **35**:559–565, 2002.

[20] Ibarra-Castanedo C., Bendada A. and Maldague X. "Image and signal processing techniques in pulsed thermography," *GESTS Int'l Trans. Computer Science and Engr.,* **22**(1): 89-100, November 2005.

[21] Ibarra-Castanedo C., González D., Klein M. Pilla M., Vallerand S. and Maldague X. "Infrared Image Processing and Data Analysis," *Infrared Phys. Technol.*, **46**(1–2):75–83, 2004.

[22] Maldague X. P. V., Theory and Practice of Infrared Technology for NonDestructive Testing, John Wiley-Interscience, 684 p., 2001.

[23] Martin R. E., Gyekenyesi A. L., Shepard S. M., "Interpreting the Results of Pulsed Thermography Data," Materials Evaluation, **61**(5):611–616, 2003.

[24] Pilla M., Klein M., Maldague X. and Salerno A., "New Absolute Contrast for Pulsed Thermography," QIRT 2002, D. Balageas, G. Busse, G. Carlomagno (eds.), *Proc. of QIRT*, pp. 53–58, 2002.

[25] Pilla M. "A Novel Contrast Method for Pulse Thermography Data", *Ph.D. Thesis*, Politecnico di Milano, 2002.

[26] Klein M., Pilla M. and Maldague X. "IR_View: A graphical user interface to process infrared images with MATLAB," application and documentation available online [http://irview.m-klein.com].

[27] González D. A., Ibarra-Castanedo C., Pilla M., Klein M., López-Higuera J. M. and Maldague X., "Automatic Interpolated Differentiated Absolute Contrast Algorithm for the Analysis of Pulsed Thermographic Sequence," *Proc. 7th Conference on Quantitative InfraRed Thermography (QIRT)*, Rhode Saint Genèse, Belgium, July 5-8, 2004, H.16.1–H.16.6.

[28] Benítez H. D., Ibarra-Castanedo C., Bendada A., Maldague X. and Loaiza H. "Modified Differential Absolute Contrast Using Thermal Quadrupoles for the Nondestructive Testing of Finite Thickness Specimens by Infrared Thermography," *Proc. CCECE 2006 - Canadian Conference on Electrical and Computer Engineering,* Ottawa (Ontario) Canada, May 7–10, 2006, Paper No. **398**.

[29] Ibarra-Castanedo C. and Maldague X. "Pulsed Phase Thermography Reviewed," *QIRT J.*, **1**(1):47–70, 2004.

[30] Ibarra-Castanedo C. "Quantitative subsurface defect evaluation by pulsed phase thermography: depth retrieval with the phase," *Ph. D. thesis*, Laval University, 2005. [accessible online: http://www.theses.ulaval.ca/2005/23016/23016.pdf].

[31] Galmiche F., Vallerand S. and Maldague X. "Pulsed phase thermography with the wavelet transform," in Review of progress in Quantitative NDE, D. O. Thompson and D.E. Chimenti (eds.), American Institute of Physics, **19A**:609–615, 2000.

[32] Ibarra-Castanedo C., González D. A., Galmiche F., Bendada A. and Maldague X. "Chapter 5 : On signal transforms applied to pulsed thermography," *Recent Res. Devel. Applied Phys.*, **9**(2006):101-127, 2006.

[33] Shepard S. M. "Advances in Pulsed Thermography", *Proc. SPIE - The International Society for Optical Engineering, Thermosense XXVIII*, Orlando, FL, 2001*,* Eds. A. E. Rozlosnik and R. B. Dinwiddie, **4360**:511–515, 2001.

[34] Ibarra-Castanedo C. and Maldague X. "Interactive methodology for optimized defect characterization by quantitative pulsed phase thermography," *Research in Nondestructive Evaluation,* **16**(4):1–19, 2005.

[35] Dillenz A., Zweschper T. and Busse G. "Progress in ultrasound phase thermography," *Proc. SPIE - The International Society for Optical Engineering, Thermosense XXVIII*, Orlando, FL, 2001*,* Eds. A. E. Rozlosnik and R. B. Dinwiddie, **4360**:574–579.

[36] Favro L. D., Han X., Ouyang Z., Sun G., Sui H. and Thomas R. L. "Infrared imaging of defects heated by a sonic pulse," *Rev. Sci. Instr.*, 2000.

[37] Busse G. "Optoacoustic and photothermal material inspection techniques," *Applied Optics*, **21**(1):107–110, 1982.

[38] Thomas R. L. "Thermal NDE techniques – from photoacoustics to thermosonics," *Review of Quantitative Nondestructive Evaluation*, Eds. D. O. Thompson and D. E. Chimenti, **21**:3–13, 2002.

[39] Sinha S. K., Iyer S. R. and Bhardwaj M. C. "Non-contact ultrasonic sensor and state-of-the-art camera for automated pipe inspection," *Proceedings of IEEE Sensors*, **2**(1):493–498, 2003.

[40] Zweschper T., Riegert G., Dillenz A. and Busse G. "Frequency modulated elastic wave thermography," *Proc. SPIE - The International Society for Optical Engineering, Thermosense XXV*, Orlando, FL, 2003*,* Eds. K. E. Cramer and X. P. Maldague, **5073**:386–391.

[41] Shepard S. M., Ahmed T. and Lhota J. R. "Experimental considerations in vibrothermography," *Proc. SPIE - The International Society for Optical Engineering, Thermosense XXVI*, Orlando, FL, 2004*,* Eds. D. D. Burleigh, K. E. Cramer and G. R. Peacock, **5405**:332–335.

CHAPTER 15

QUANTITATIVE EVALUATION OF EMERGING INFRARED THERMOGRAPHY TECHNOLOGIES FOR AEROSPACE APPLICATIONS

Joseph DiMambro

Sandia National Laboratories
PO Box 5800, MS-0615
Albuquerque, NM 87185
jdimamb@sandia.gov

Sandia National Laboratories Airworthiness Assurance Nondestructive Inspection Validation Center (AANC) provides independent and quantitative evaluations of new and enhanced inspection technologies, to developers, users, and regulators of aircraft. In this chapter, we present the results from two recent thermography probability of detection (POD) experiments. The first POD study involves sonic infrared (IR) imaging for detecting fatigue cracks in Inconel® and titanium. The second POD study involves pulsed thermography using a low cost uncooled infrared camera for detecting flaws in composite honeycomb structure.

1.1. Sonic IR Probability of Detection Study

Sonic infrared imaging [1-2] is an emerging inspection technology that employs a short pulse of ultrasound, typically less than a second, via a 20–40 kHz ultrasonic transducer. This causes fatigue crack faces to rub or clap, thus inducing frictional heating. The increase in temperature in the vicinity of the crack can then be detected with an IR camera and appears as a bright IR source within a dark background field.

Several encouraging demonstrations have been performed on a number of different aircraft components (e.g. turbine blades, wheels, brake manifolds, bolts, etc.) having known cracks which were detected with fluorescent penetrant inspection (FPI) in the field [3]. Commercial airlines have expressed interest in replacing some FPI inspections currently performed in the aviation field with

Sonic IR. Benefits include reduced part preparation requirements (e.g. no paint stripping, no specialized cleaning), inspector dependency, processing and evaluation time, floor space requirements, and environmental issues (e.g. elimination of chemicals). However, before Sonic IR can be successfully transferred into the field several questions must be answered such as: How reliably can the technology detect defects? Is the technology truly nondestructive? This effort explores these questions.

1.2. Reliability Experiment Development

To ascertain, in a quantitative manner, the ability of Sonic IR to reliably find cracks, we have developed two probability of detection (POD) experiments. The first experiment consists of 144 titanium 6-4 specimens, and the second of 92 Inconel®-718 specimens. Thirty-six of the titanium specimens contained 37 known cracks, and twenty four of the Inconel® specimens contained 28 known crack detection opportunities. Known fatigue cracks vary in length from 0.016-0.182 inches and 0.022-0.422 inches for titanium and Inconel® specimens respectively. Titanium and Inconel® are common materials used for aircraft engine components such as blades and discs, which are typically inspected for cracks with FPI in the field. Figure 1.1 is a photograph of two test specimens. Each specimen is 6 inches long, 1 inch wide and 0.25 or 0.50 inches thick. Fatigue crack(s) are located on the top surface (6 inch x 1 inch plane) of the specimen that has a surface roughness average (RA) of 5-15 micro inches.

Figure 1.1. Titanium 6-4 and Inconel®-718 Specimens.

Specimens containing fatigue cracks were generated by a three-point bending process and were previously used for FPI research at Iowa State University and at the AANC. Nonflawed specimens were fabricated and incorporated into the experiment to test for false calls. Fluorescent penetrant, eddy current, ultrasonic linear array, and optical inspections were performed to characterize the

specimens. The inspections confirmed the presence of fatigue cracks and the optical inspection was used to measure the length of each fatigue crack and ensure that the flaw length distribution for each experiment was statistically relevant.

1.3. Reliability Experiment Implementation

Wayne State University (WSU) participated in the blind Sonic IR inspections. WSU followed detailed and firm AANC experiment protocols. For instance, prior to the blind test, WSU was given an objective to detect crack lengths 0.050 inches and greater. Cracked titanium and Inconel® reference specimens having cracks 0.050 inch long were given to WSU so they could develop and optimize an inspection setup and procedure. The Sonic IR inspection setup is shown in Figure 1.2 and inspection parameters are shown in Table 1.1. Only the pressure between the ultrasonic transducer and the specimen was changed between the experiments, while all other parameters remained constant.

The protocol limited WSU to only one inspection per specimen unless the sonic IR equipment malfunctioned, which is uncharacteristic. If the inspection resulted in a possible defect call, one additional inspection was allowed with the IR camera positioned closer to the inspection article. However, this occurred infrequently. The inspector team consisted of two graduate students who we considered Sonic IR experts. One student placed the test specimen in a fixture and the second operated the inspection system and made defect and no defect calls. The inspector was asked to make a defect or no defect call based on what he considered a relevant indication. AANC experiment monitors recorded successful defect calls (hits), defects that were missed (misses), and false calls. Figure 1.3 is an infrared inspection image of a test specimen with a crack indication (bright white line) in the center of the specimen.

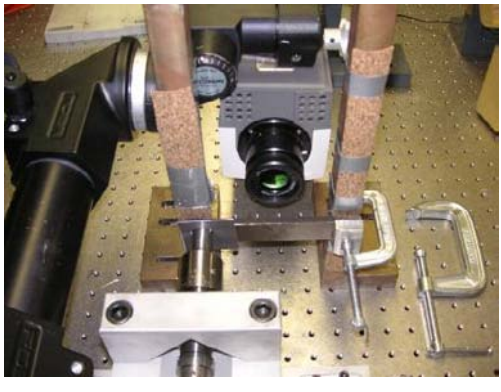


Figure 1.2. Photograph of Sonic IR Inspection Setup.

Table 1.1. Sonic IR POD Experiment Setup Parameters

Parameter	Value
Camera used	Phoenix-MWIR 9803
Resolution of IR camera	640 X 256 pixels
Frame rate	120Hz
Lens used	25mm
Field of view	4.5 x 1.875 inches
Distance from the lens to the sample	6.625 inches
Integration Time	2.0 ms
Image acquisition time	1 sec
Power supply	Branson 2000ea (40kHz)
Size of ultrasonic gun tip	0.75 inches
Coupling material used	Duct tape
Ultrasonic pulse duration	800ms
Strain gage reading Inconel®	2800(initial)-1800(final)
Strain gage reading (0.25" thin) Ti	2100(initial)-1450(final)
Strain gage reading (0.50" thick) Ti	2300(initial) -1600(final)

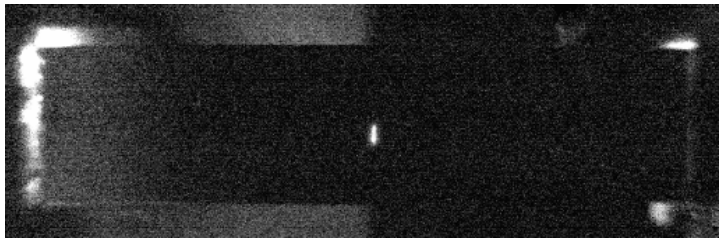


Figure 1.3. Sonic IR Inspection Image with a Crack Indication (bright white line).

The blind experiment was repeated with a non-expert AANC inspector who was given an hour of training using the WSU Sonic IR system.

1.4. FPI Reliability Experiment Implementation

Blind fluorescent penetrant inspections were performed using a Level IV (ultra high) sensitivity penetrant and a Method D: post emulsifiable, hydrophilic removal process. This penetrant process was specifically requested for this comparison study by industry because it is commonly used to inspect critical aircraft engine components (e.g. blades and discs) in the field. Each inspector followed their respective corporate processing procedures/standards for cleaning, dwell times, drying times, etc. AANC protocols defined the inspection area to be a 4" x 1" area in the center of each specimen which is comparable with the Sonic IR inspection field of view. The protocol further stipulated that each inspector

work independently (e.g. perform all processing and inspection steps individually) and call what they consider a relevant FPI indication.

1.5. Reliability Experiment Results

Table 1.2 contains a summary of the Sonic IR and FPI inspection outcomes. The first Sonic IR inspection (Sonic-IR1) was performed by WSU and the second (Sonic-IR2) was accomplished by a less experienced, recently trained AANC inspector. Seven FPI inspections were performed on the titanium specimens and included three inspectors from a community college and four inspectors from an aircraft manufacturer. Two of the college inspectors were ASNT level 1 inspectors; one of the OEM inspectors was a level 2; and the remaining four were level 3. The four lab inspectors also inspected the Inconel® specimens. The penetrant inspections show the ASNT level of the inspector as the first number in the inspection column label in Table 1.2.

Table 1.2. Summary of Inspection Results

	Titanium POD Experiment				Inconel® POD Experiment			
Inspection	Calls	#Detects	#Misses	Detect Rate	Calls	#Detects	#Misses	Detect Rate
Sonic-IR1	33	33	4	0.89	20	20	8	0.71
Sonic-IR2	30	30	7	0.81	17	17	11	0.61
Penetrant 1-1	45*	26	11	0.70	-	-	-	-
Penetrant 1-2	34*	24	13	0.65	-	-	-	-
Penetrant 3-3	39	31	6	0.84	-	-	-	-
Penetrant 2-4	105	22	15	0.59	37	25	3	0.89
Penetrant 3-5	235	22	15	0.59	63	24	4	0.86
Penetrant 3-6	146	27	10	0.73	60	24	4	0.86
Penetrant 3-7	55	21	16	0.57	30	24	4	0.86

*Inspectors only inspected 118 of the Titanium specimens, but all cracked specimens were inspected.

The summary of the inspection results given in Table 1.2 indicate that for the specimens used in the testing the Sonic-IR technique detected a higher total proportion of the titanium cracks but a lower proportion of the Inconel® cracks than did the penetrant inspections. However, the overall summary given in Table 1.2 gives no indication of how the probability of detection might depend on the flaw size. To gain an understanding on how the detection rates may depend on flaw size, we fit to the data the probability of detection model given by:

$$POD(a) = \Phi\big(c + d \cdot \ln(a)\big), \tag{1.1}$$

where $\Phi(\cdot)$ is the cumulative Normal distribution function and a is the flaw size. The parameters c and d are estimated by maximizing the likelihood to provide the best fit to the data. This is the model often used for probability of detection, although it is often parameterized in terms of a "mean" and "standard deviation" to emphasize the connection to distribution functions. We prefer the parameterization given in equation (1), noting that a negative estimate for d implies that the probability of detection decreases with flaw size, and the case of $d = 0$ corresponds to a constant probability of detection with no dependence on flaw size.

Figure 1.4 displays the estimates of the probability of detection (POD) curves for each of the inspections of the titanium sample set using the model given by equation (1). Similarly, Figure 1.5 shows the fits for the Inconel® specimens. In none of the penetrant inspections was the "d" parameter of equation (1) significantly different from 0 at a 90% confidence level. The POD curves are shown using the maximum likelihood estimate of "d" when that estimate was positive and using $d = 0$ when the estimate was negative. In both figures a vertical line indicates the value at the minimum crack size represented in the specimen set.

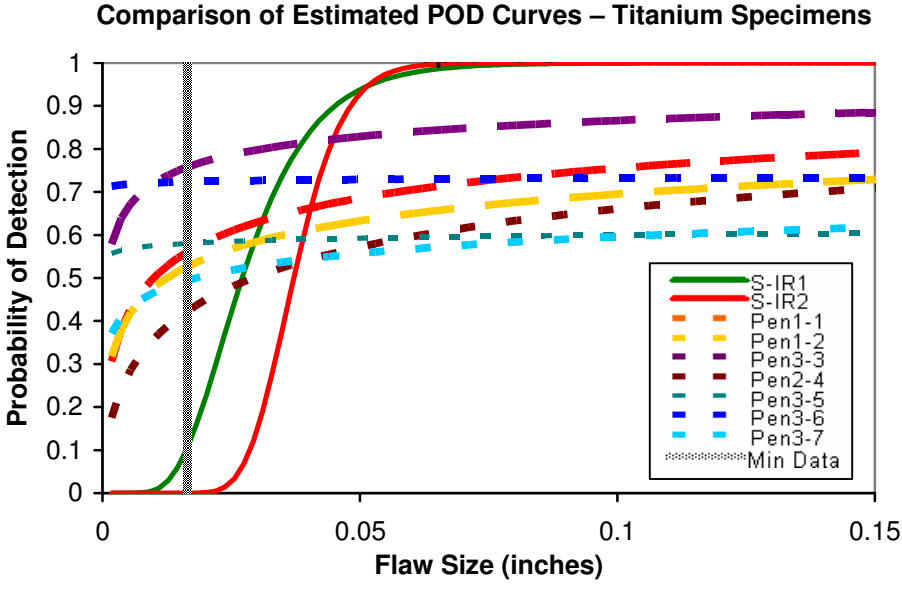


Figure 1.4. Binary Regression POD Curve Estimates for Titanium Specimens.

The relative flatness of the penetrant POD curves in both Figure 1.4 and Figure 1.5 arises because there is a lack of flaw size dependence in the inspection results. That is, the flaws that were missed did not tend to be small flaws, but were fairly uniformly spread over the range of flaw sizes.

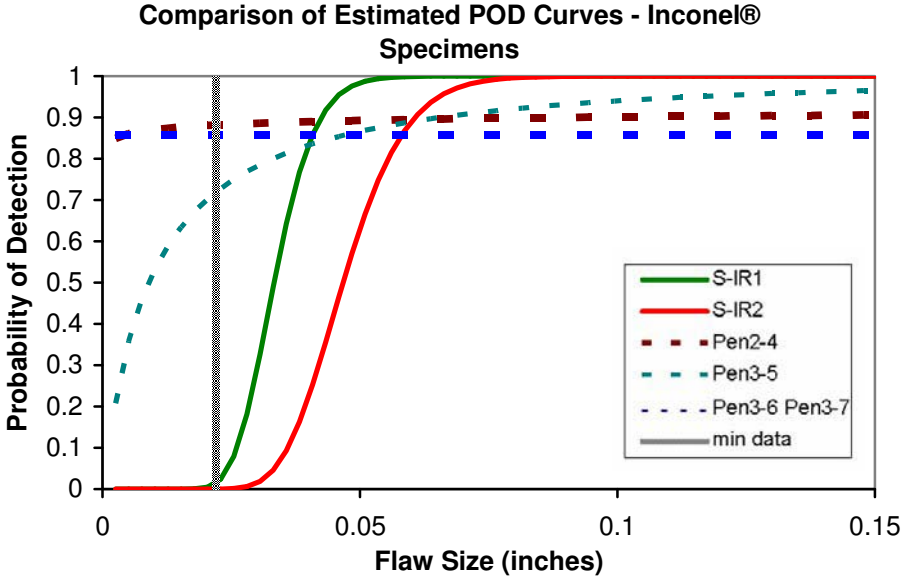


Figure 1.5. Binary Regression POD Curve Estimates for Inconel® Specimens.

It is customary to summarize the outcome of a POD experiment not only with the best estimate of the POD curve, but also with a reflection of the degree of uncertainty implied by the data. The degree of uncertainty is expressed in terms of confidence bands placed about the curves. Suggested methods for calculating confidence bounds are reviewed in references [4] and [5]. Here, we employ the profile likelihood methodology based on likelihood ratio statistics [5] to compute 2-sided confidence bounds at the 90% confidence level[1]. Specifically, we compare the bounds derived from the two Sonic IR inspections to the "best" penetrant inspection for each of the two specimen types in Figures 1.6 and 1.7.

[1] The choice of a two-sided bound is made because the desire is to compare the different POD curves arising from the various sets of inspection data. The usual choice of characterizing POD uncertainty by a lower 95% curve is justified by the desire to take a conservative position in engineering applications, where the POD at specific flaw sizes are to be compared to structural risk requirements. The lower bound from the 90% two sided bound is nominally comparably to a 95% one-sided bound, although the one sided coverage level is not specifically controlled to be 95%.

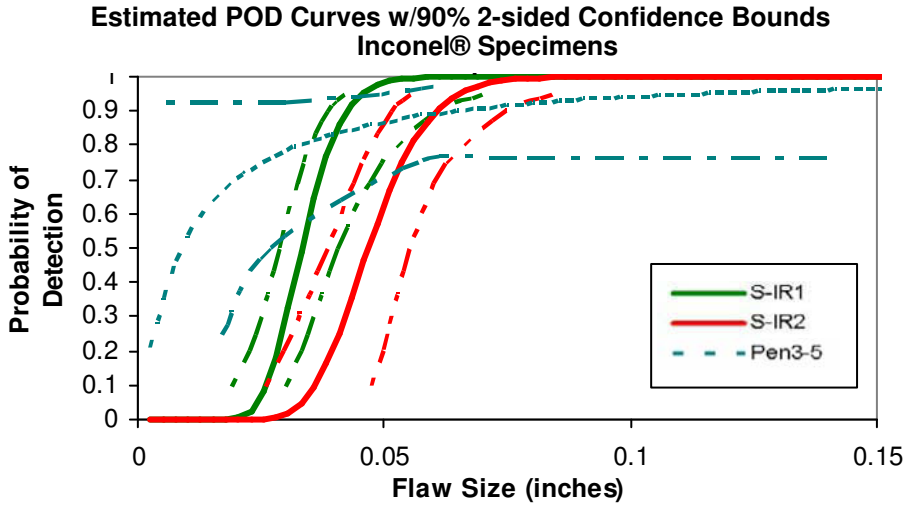


Figure 1.6. Sonic IR and "best" FPI Inspections with 90% two-sided confidence bounds Titanium Specimens.

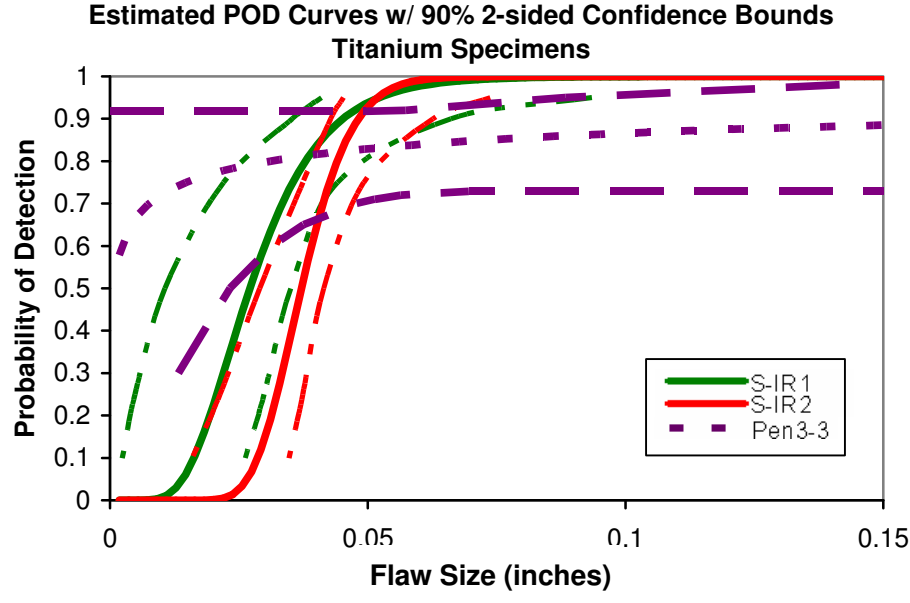


Figure 1.7. Sonic IR and "best" FPI Inspections with 90% two-sided confidence bounds Inconel® Specimens.

It should be clear that if there is little statistical evidence that the probability of detection depends on flaw size for the FP inspections, then the calculation of bounds for a specific level of detection is problematic. For example, in the FP inspections the best estimates of overall detection rates (that is, assuming that, in fact, $d = 0$) are less than 0.90 for all the inspections and for all crack sizes. Therefore, the upper bound on the flaw size that would be detected 90% of the time would be ∞ (infinity).

By moving the IR camera closer to the inspection article it is believed that the increased resolution would enable an inspector to better discern and identify the heat generated from smaller cracks from background noise during inspection. Therefore, the Sonic IR probability of detection curves shown in Figures 1.4–1.7 would shift to the left to some unknown threshold as the IR camera distance to inspection article is reduced. To test this hypothesis we reduced the objective crack length from 0.050" to 0.030" and asked WSU to optimize their inspection setup using a titanium and Inconel reference sample containing a fatigue crack approximately 0.030" in length. The distance between the IR camera and the reference sample was reduced from 6.625" to 2.25" and the titanium and Inconel experiments were both repeated. The tradeoff in moving the IR camera closer to the inspection article is that the inspection field of view was reduced from 4.5" x 1.875" to 2" x 1.6". Figure 1.8 illustrates the improvement in POD for the titanium specimens.

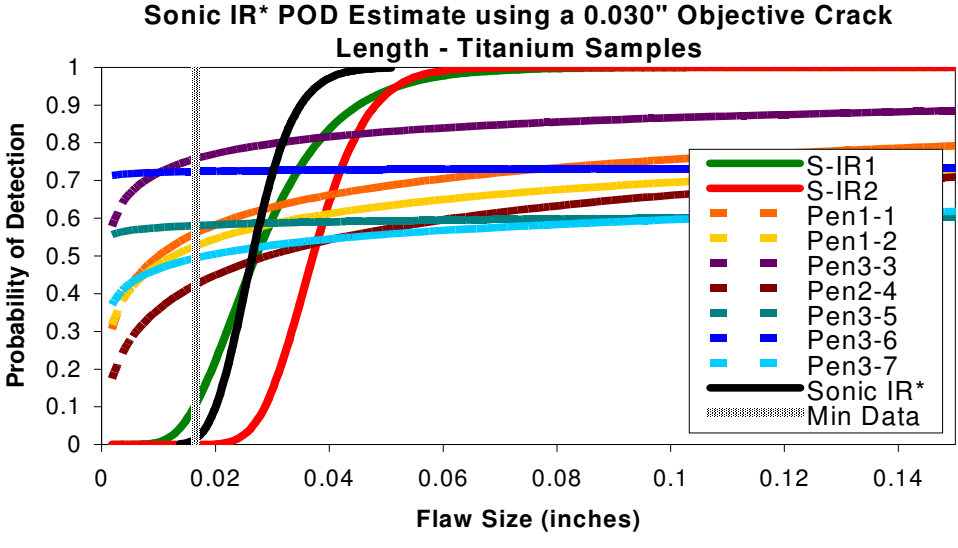


Figure 1.8. Binary Regression POD Curve Estimates for Titanium Specimens.

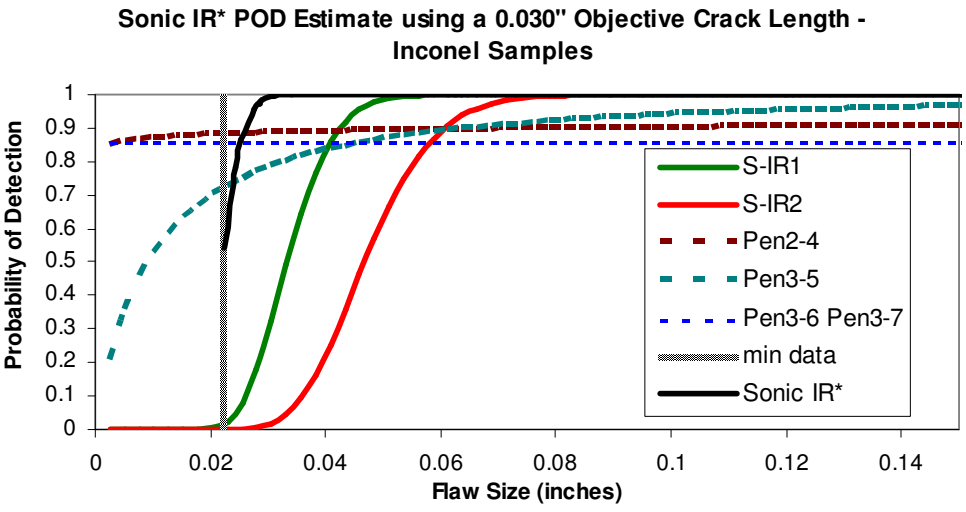


Figure 1.9. Binary Regression POD Curve Estimates for Inconel Specimens.

1.6. Conclusions

Although the objective crack length for Sonic IR was 0.050" in length, the inspection achieved higher POD values than FPI for fatigue cracks 0.040" in length and greater. The upper bound on flaw size that is detectable at a rate of 0.9 is in the 60 to 70 mil range for the Sonic IR inspections. Flaws less than the estimated 0.040" are more likely to be detected by FPI in this experiment, but at the expense of a high false call rate. No false calls were observed in any of the Sonic IR inspections. The false call rate for FPI is reflective of the fact that the FPI is not specifically targeted to deep cracks, but also is responsive to other surface conditions (e.g. porosity, machine marks, etc) and the inspector's interpretation of "relevant indication." The Sonic IR inspection data imply that the probability of detection is dependent upon crack size while there is little evidence that POD for FPI depends on flaw size. The FPI data suggest that a constant probability of detection independent of flaw size adequately characterizes the inspection process for both types of specimens, and the upper bound for the 0.90 detectable crack is infinity. Test specimen material and inspector experience did not affect Sonic IR POD, while the FPI POD indicated greater inspector and material variability. When the objective crack length was reduced from 0.050" to 0.030" and the experiment was repeated, Sonic IR achieved higher POD values than FPI for fatigue cracks 0.030" in length and greater. Finally, several titanium and Inconel® reference standards were

measured before and after repeated Sonic IR inspections, some up to 40-50 times, and no crack growth was observed.

2.1. Pulsed Thermography Probability of Detection Study

Pulsed thermography is a nondestructive inspection method that uses thermal gradients to analyze the physical characteristics of a structure such as internal defects. This is done by converting a thermal gradient into a visible image by using a thermally sensitive detector such as an infrared (IR) camera. By the judicious application of external heat sources, common aircraft defects can be detected by an appropriate infrared survey. The heat source, such as flash lamps, is used to raise the surface temperature of the structure. The subsequent heat transfer into the material is affected by any defects that may be present. As the heat diffuses through the structure and cools, the surface temperature is monitored for a period of time by an infrared camera. Areas that appear hotter or cooler than surrounding areas may indicate the presence of a flaw beneath the surface that is preventing or assisting heat diffusion into deeper layers. By using a computer to analyze the infrared data captured over time, subtle variations can be enhanced in the image.

The infrared cameras used in modern pulsed thermography systems for NDE are typically based on focal plane array photon detectors (e.g. InSb, QWIP, or HgCdTe), which must be cooled to LN2 temperatures during operation. Despite excellent performance of thermographic NDE systems based on these cameras for many aircraft applications, their relative high cost, compared to established inspection technologies, has posed an obstacle to widespread implementation by airlines and their maintenance contractors. A newer generation of microbolometer infrared cameras, based on vanadium oxide (VOx) or barium strontium titanate (BST) detectors, do not require cryogenic cooling (although they do require temperature stabilization, usually provided by a thermoelectric cooler). The microbolometer cameras are generally cheaper, smaller, consume less power, and require less maintenance than their cooled counterparts [6]. As a result, they have become extremely popular for many applications in preventative maintenance. However, uncooled cameras have not been widely adapted for NDE, as they require frequent recalibration that can interrupt an acquisition sequence that is longer than a few seconds, and the noise performance degrades over the course of a typical acquisition cycle. An opportunity for widespread implementation by the airlines can be realized if the resolution and sensitivity levels are sufficient to detect flaws with high confidence for some applications.

2.2. Composite Honeycomb POD Experiment Implementation

The AANC composite honeycomb POD experiment was chosen because there is increasing use and interest in composite materials by the aviation community. A number of NDT technologies have performed the experiment, and a wealth of POD data is currently available for comparison [7] [8]. AANC implemented a blind composite honeycomb flaw detection experiment at Thermal Wave Imaging (TWI) Inc. Again, the goal of the experiment was to quantify how well pulsed thermography with an uncooled camera can reliably find flaws in composite honeycomb structure. A variety of 3-, 6-, and 9-ply fiberglass and carbon graphite laminate and Nomex® core samples were inspected by TWI using a *Flir* A40 microbolometer uncooled IR camera integrated with an Echotherm™ pulsed thermography system. Before the experiment commenced, 3-, 6-, and 9-ply fiberglass and carbon reference samples were inspected to develop and optimize an inspection procedure and setup. The inspection setup is shown in Figure 2.1.

Figure 2.1. Pulsed Thermography Inspection Setup

The blind experiment consisted of 44 specimens having an 18inches x 18inches inspection area. The specimens contain a statistically relevant flaw distribution with defect sizes ranging from 0.5-3 inches in diameter. The engineered flaw types include inter-ply delaminations, skin-to-core air gap disbonds, skin-to-core “kissing” disbonds, and impact damage. Note that TWI previously participated in this experiment in 2003 using a Flir Merlin Mid cooled IR camera. Table 2.1 compares the characteristics of each IR camera used.

Table 2.1. Comparison of IR Cameras Used for the Composite Honeycomb Experiment

Parameter	Merlin Mid	A40
Detector Material	Indium Antimonide (InSb)	Vanadium Oxide (VOx)
Detector Cooling	Integral Stirling or LN2	Uncooled Microbolometer
Spectral Range	3-5 micron	7.5-13 micron
Thermal Sensitivity	0.025 °C	0.08 °C
Focal Plane Array	320 x 256	320 x 240
Frame Rate	60 Hz	60 Hz
Weight	9 lbs	3.1 lbs
Size	9.8" x 5.5" x 5.0"	8.2" x 4.3" x 3.6"
Cost	$43,000	$18,000

2.3. Composite Honeycomb POD Experiment Results

AANC collected hit/miss data from the blind inspection performed at TWI. We then fit to the data the probability of detection model given by equation 1.1. POD curves for pulsed thermography using the A40 uncooled IR camera and Merlin Mid cooled IR camera are shown for the 3-, 6-, and 9-ply specimens in Figures 4–9. POD curves are also depicted in Figures 410 to demonstrate how pulsed thermography compares to NDE equipment which is currently used in the field. These techniques include tap testers, mechanical impedance analysis, and low frequency bond testing and the blind inspections were performed by inspectors from commercial airlines. Pulsed thermography using the Merlin Mid cooled IR camera achieved a detection rate of 100% and 97% for 3-ply fiberglass

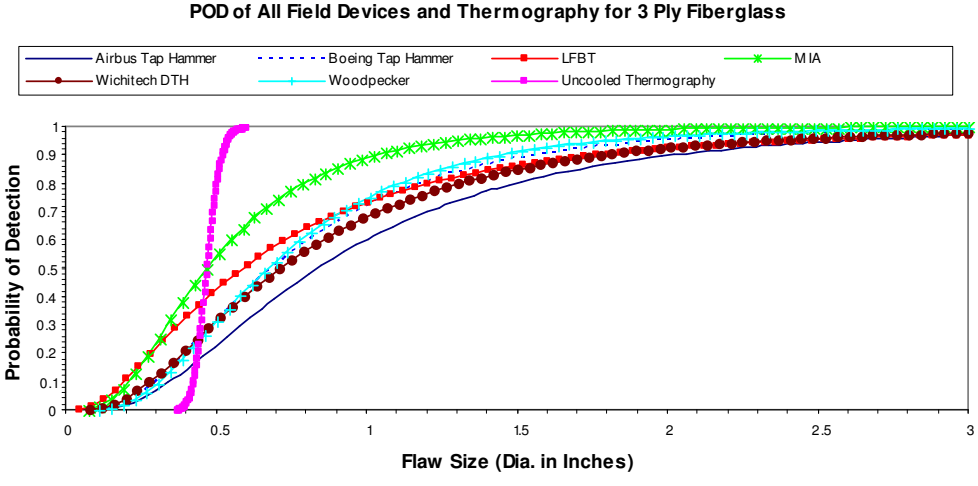


Figure 4. POD of all Field Devices and Thermography for 3-Ply Fiberglass.

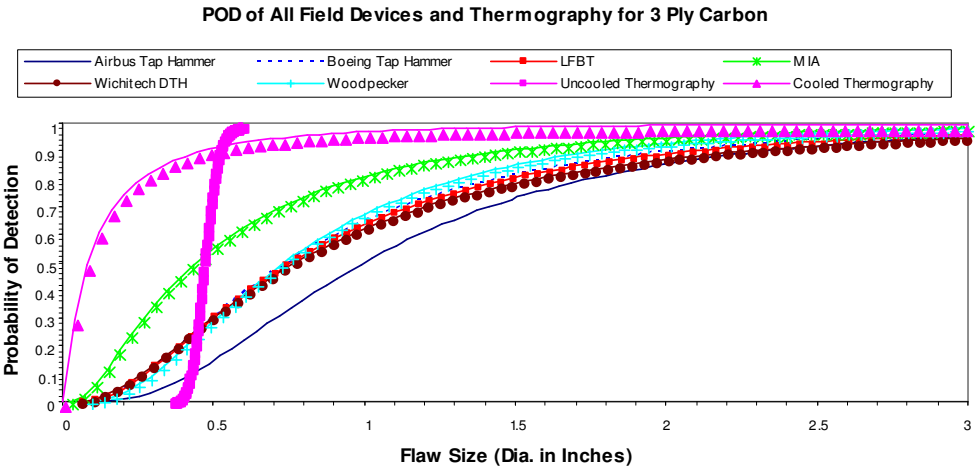


Figure 5. POD of all Field Devices and Thermography for 3-Ply Carbon.

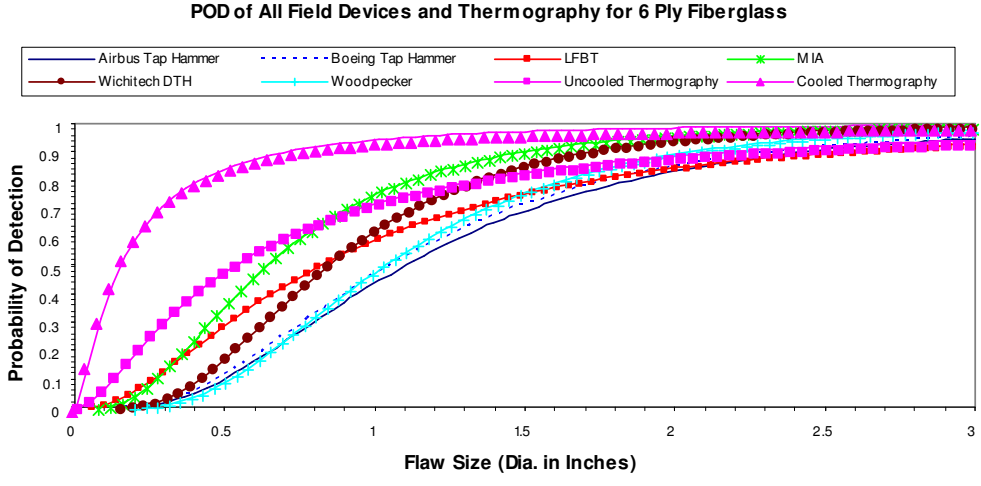


Figure 6. POD of all Field Devices and Thermography for 6-Ply Fiberglass.

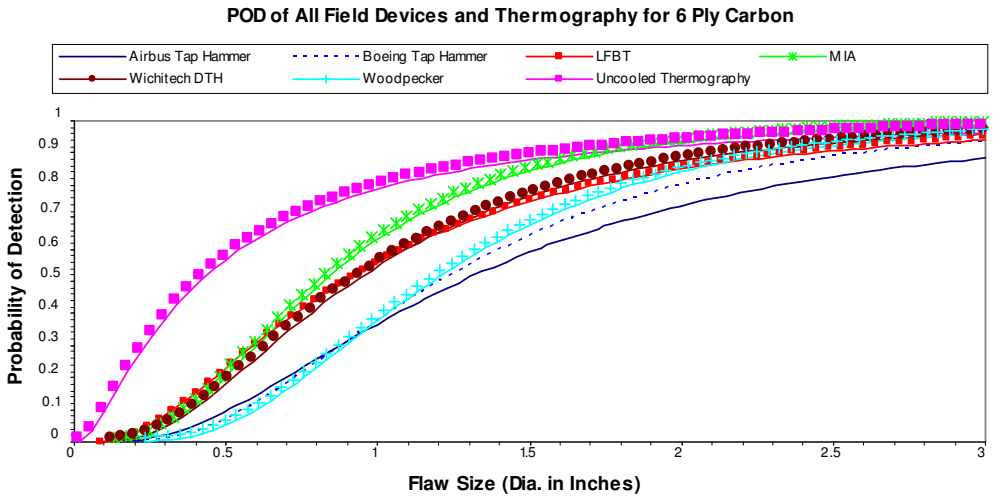


Figure 7. POD of all Field Devices and Thermography for 6-Ply Carbon.

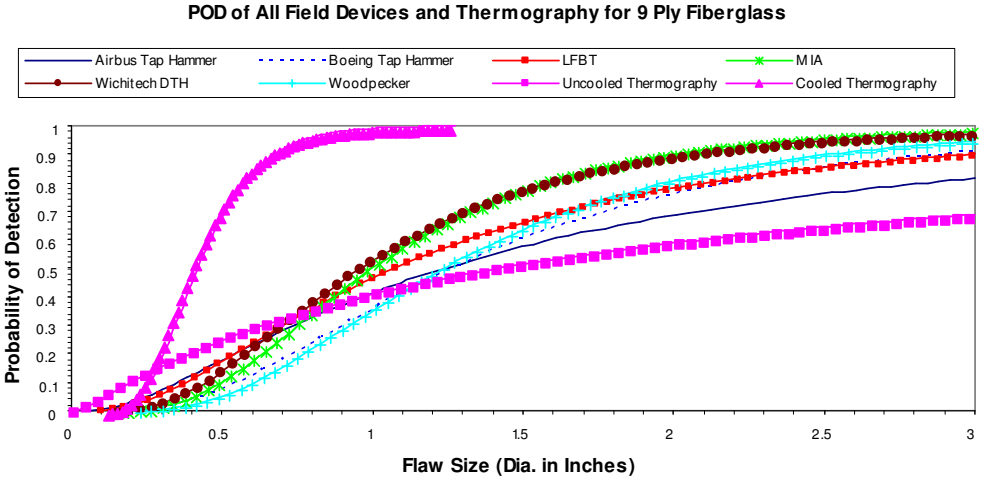


Figure 8. POD of all Field Devices and Thermography for 9-Ply Fiberglass.

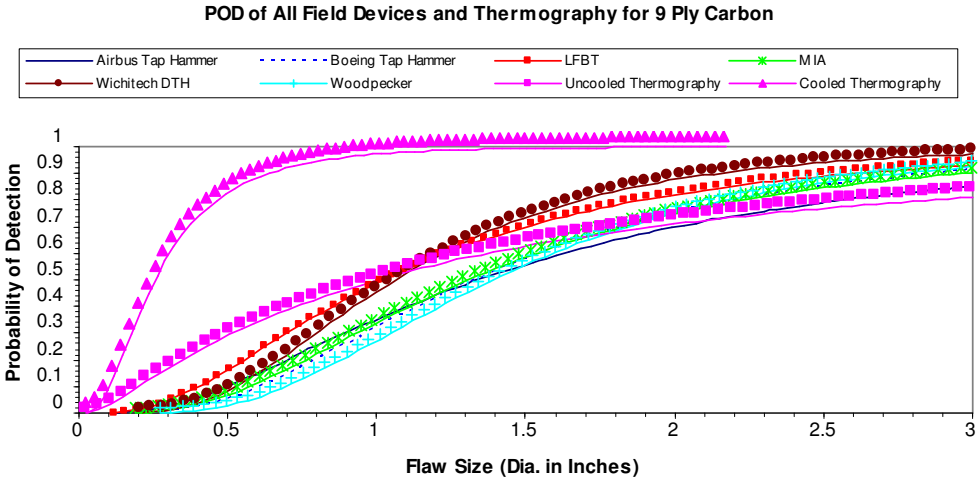


Figure 9. POD of all Field Devices and Thermography for 9-Ply Carbon.

and 6-ply carbon samples respectively. As a result, POD curves could not be generated for these experiments and therefore they do not appear in Figures 4 and 7.

Table 1.2 summarizes and compares the pulsed thermography results based on IR camera that was used. The overall flaw detection rate is calculated by dividing the number of flaws detected from the blind inspection by the number of known flaws in the specimen set. Note that the minimum flaw size present in the POD specimens is 0.5 inches in diameter.

Table 2.2. Cooled versus Uncooled Thermography POD Comparison

Cooled Pulsed Thermography (Merlin Mid) General Inspection Results			
Specimen Type	Overall Flaw Detection Rate	Flaw Size at 90% POD	Number of False Calls
3 Ply Carbon	96%	0.5	2
3 Ply Fiberglass	100%	<0.5	0
6 Ply Carbon	97%	0.6	2
6 Ply Fiberglass	93%	0.7	3
9 Ply Carbon	95%	0.6	0
9 Ply Fiberglass	93%	0.7	2
Cooled Pulsed Thermography (Merlin Mid) General Inspection Results			
3 Ply Carbon	96%	0.5	0
3 Ply Fiberglass	96%	0.5	0
6 Ply Carbon	78%	1.7	2
6 Ply Fiberglass	71%	2.2	2
9 Ply Carbon	53%	>3.0	0
9 Ply Fiberglass	42%	>3.0	4

In addition to hit/miss data, the number of false calls for each inspection technique was recorded and is shown in Table 2.3. Since multiple field inspectors performed the experiment using the Airbus Tap Hammer for instance, the false calls were averaged which produced the fractional false call numbers shown in Table 2.3.

Table 2.3. Cooled versus Uncooled Thermography POD Comparison

	Number of False Calls					
NDI Technique	3F	3C	6F	6C	9F	9C
Cooled Thermography	0	2	3	2	2	0
Uncooled Thermography	0	0	2	2	4	0
Airbus Tap Hammer	0.3	0.8	2.9	3.5	7.1	6.8
Boeing Tap Hammer	1.8	1.0	4.7	4.7	6.9	5.4
Low Freq Bond Tester	0.9	2.3	3.3	6.5	2.8	3.8
MIA	0.5	0.6	1.9	1.3	6.8	3.1
Wichitech DTH	0.9	0.6	1.6	0.9	4.9	4.4
Woodpecker	0.3	0.0	0.1	0.7	6.7	2.8
	X Denotes Near 100% Detection Rate					

Advanced NDE techniques such as the computer aided tap testing (CATT), laser shearography, surface acoustic mapping (SAM), and MAUS resonance and MIA were also tested. The minimum flaw size that each technique achieved with a 90% probability of detection is shown in Table 2.4 below.

Table 2.4. Advanced NDE Techniques POD Comparison.

	Flaw Size Diameter (inches) at 90% POD					
Advanced Technique	3F	3C	6F	6C	9F	9C
Cooled Thermography	X	<0.5	0.7	X	0.7	0.6
Uncooled Thermography	0.5	0.5	2.2	1.7	>3	>3
CATT	1.1	1.2	1.1	1.9	2.2	2.4
Shearography	0.6	X	X	1.0	0.8	2.0
MAUS Resonance	0.7	X	0.5	X	0.6	1.3
SAM	0.7	1.3	0.9	1.6	1.0	>3
MAUS MIA	0.7	0.9	1.5	1.4	>3	>3
	X Denotes Near 100% Detection Rate					

Additional techniques have since been tested, such as air coupled ultrasound, laminography, and Laser UT. These results have not been complied and, therefore, the results in Table 2.3 are preliminary.

Some considerations when comparing the advanced techniques is that the CATT is typically used to inspect composite materials only. Various eddy current and ultrasound inspections can be deployed with the Boeing MAUS scanner

which is appealing. Honeywell ceased all SAM research and system development efforts; therefore, a system is currently not available for purchase. Air coupled ultrasound cannot be deployed from one side so access to the backside of the part is required. Laminography is an X-ray based technology and, therefore, radiation precautions apply. The laser ultrasound system is currently not field deployable. Pulsed thermography can also be used to detect corrosion and disbonds (metallic structure), fluid ingress, lightning strikes, and obtain thickness measurements.

2.4. Conclusions

Pulsed thermography using the *Flir* A40 microbolometer IR (uncooled) camera detected flaws with higher probability (POD) than NDI techniques currently used in the field as well as most advanced NDI techniques tested for 3-ply fiberglass and 3-ply carbon laminate Nomex® honeycomb core samples. Flaws were detected competitively with the field NDI techniques for the 6 ply samples. It was observed that the detection performance degraded as the laminate thickness increased. It is important to note that the A40 camera tested in this study achieved comparable POD for the 3-ply samples than the *Flir* Merlin Mid IR (cooled) camera and is $25,000 less expensive. Although more expensive, the Merlin Mid IR camera used for pulsed thermography detected flaws with higher probability (POD) than the field NDI techniques and most advanced NDI techniques tested for 3-, 6-, and 9-ply fiberglass or carbon laminate and Nomex® honeycomb core samples.

Acknowledgements

This research was sponsored by the Federal Aviation Administration (FAA) William J. Hughes Technical Center, under contract number DTFA0300X90019-Z.

References

1. R.B. Mignogna, R.E. Green, Jr., J.C. Duke, Jr., E.G. Henneke II, and K.L. Reifsnider, Ultrasonics 19, 159 (1981).
2. R.L. Thomas, L.D. Favro, X. Han, Z. Ouyang, G. Sun, and H. Sui, U.S. Patent No. 6,236,049 (22 May 2001).
3. DiMambro, J. et al, "The Potential of Sonic IR to Inspect Aircraft Components Traditionally Inspected with Fluorescent Penetrant and or Magnetic Particle Inspection in the Field," *Review of Progress in Quantitative Nondestructive Evaluation Quantitative, Vol. 25A*, D.O. Thompson and D.E. Chimenti, eds, AIP, 2005, pp 536-543.

4. Harding C. A., and Hugo G. R., "Statistical Analysis of Probability of Detection Hit/Miss Data for Small Data Sets," *Review of Progress in Quantitative Nondestructive Evaluation, Vol. 22B*, D.O. Thompson and D.E. Chimenti, eds, AIP, 2003, pp 1838-1845.
5. Spencer F. W., "The Calculation and Use of Confidence Bounds in POD Models," 2006 QNDE, Portland, Oregon, July 31 - August 4, 2006.
6. Shepard S., Lhota J., Chaudhry B., Hou Y., "Flash Thermography Inspection Using Miniature Uncooled Cameras" ATA NDT Forum, October, 2004.
7. DiMambro, J., Roach, D., Rackow, K., Shepard S., Ahmed, T., "A Composite Honeycomb Probability of Detection (POD) Experiment using an Uncooled Infrared Camera for Pulsed Thermography" ATA NDT Forum, October, 2006.
8. Roach, D., Rackow, K., "Improving In-Service Inspection of Composite Structures: It's a Game of CATT and MAUS", FAA/NASA/DOD Aging Aircraft Conference, September 2003.

This page intentionally left blank

CHAPTER 16

SONIC INFRARED IMAGING: A NOVEL NDE TECHNOLOGY FOR DETECTION OF CRACKS/DELAMINATIONS/DISBONDS IN MATERIALS AND STRUCTURES

Xiaoyan Han

Department of Electrical and Computer Engineering
5050 Anthony Wayne Dr., #3140, Detroit, MI, USA 48202
xhan@eng.wayne.edu

Sonic Infrared Imaging is a novel NDE technique that combines ultrasonic excitation and infrared imaging to detect defects, such as cracks, delaminations and disbonds, in materials and structures. The ultrasonic excitation is typically a fraction-of-a-second-long pulse, which causes friction heating in the defects. The temperature changes in the target due to the heating are simultaneously being imaged by an infrared camera. Thus, defects in a wide variety of materials can be identified. This technique provides a sensitive, fast, and wide-area NDE method

1. Introduction

Sonic Infrared (IR) Imaging is a novel NDE technique that combines ultrasonic excitation and IR imaging to detect defects, surface or subsurface, such as cracks, delaminations, and disbonds, in materials and structures.[1-8] The ultrasonic excitation is typically a fraction-of-a-second-long pulse applied at some convenient point on the surface of the object under inspection. The sound vibrations produce localized frictional heating at the defect, and an IR camera images the surface heating resulting from the effects of friction or other irreversible internal surface interactions in the vicinity of cracks, or other defects with surfaces that are in contact. These effects result from the fact that the two surfaces of internal defects do not move in unison when sound propagates in the object. Thus, for instance, the faying surfaces of a closed crack appear as a planar heat source. If the crack intersects the surface, the heat source first appears as a line in the IR image. This line subsequently blurs and broadens into a diffusely

heated region surrounding the original line. When the sound pulse is turned off, the resulting temperature pattern decays through the process of thermal diffusion. The whole imaging period, including the storage of the data, is about a second or so, depending on the thermal properties of the object under evaluation. Thus, Sonic IR imaging is a sensitive, super-fast, and wide area imaging NDE technique. The (reversible) thermoelastic heating and cooling caused by the vibrations in the object, which are typically in the tens of kilohertz range, are averaged out during the IR sensor's integration time, which is typically around a millisecond. Thus, only the irreversible heating is imaged by the IR camera. This is different from Stress Pattern Analysis by Thermal Emission (SPATE) in which the detection of periodic thermoelastic temperature variations is synchronized with the sinusoidal stress-induced heating and cooling.

2. Experimental Setup

In a typical Sonic IR imaging arrangement, as shown in the schematic drawing of Figure 1, a single frequency ultrasonic excitation source is applied to the sample to cause friction heating in the defects, and an IR camera images the resultant heating of the crack. The camera can be placed in any location that is appropriate to view the region of interest in the sample, and the acoustic excitation transducer can be placed in some convenient location.

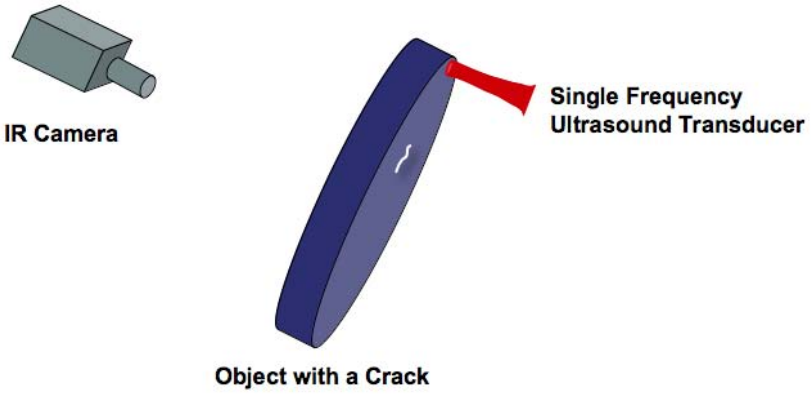


Figure 1. Schematic drawing of a typical Sonic IR imaging arrangement. A single frequency ultrasonic excitation source is applied to the sample to cause friction heating in the defects, and an IR camera images the resultant heating of the crack.

In our laboratory protocol for imaging cracks, the IR video camera begins to acquire images a short time before the sound pulse is injected into the part, and

continues for a short time after it ends. The reason for acquiring images before the sound pulse starts is to allow for the possibility of subtracting one or more of those images from later images to produce true dark-field images in which only the defects are visible. This subtraction eliminates background artifacts and clutter, and greatly increases the sensitivity for detection of the defects. By thresholding such subtracted images, one could design a system to automate the rejection of defective parts. Alternatively, one can leave a small fraction of the background image unsubtracted to produce a "ghost" image of the part to aid in localizing the defects by visual inspection.

The effect of friction heating in Sonic IR imaging is illustrated in Figure 2. In Figure 2a we show an optical micrograph of a fatigue crack in an aluminum test bar. The crack is tight and is only 0.67mm long. In Figure 2b, we show one frame taken from a video sequence of IR images grabbed during the sonic excitation. The IR signal around the crack is almost overwhelming. It is obvious that this technique could be used to detect even smaller cracks. The saw cut itself does not light up since there is no rubbing in the slot. The bright spot in the saw cut is a particle of a foreign material.

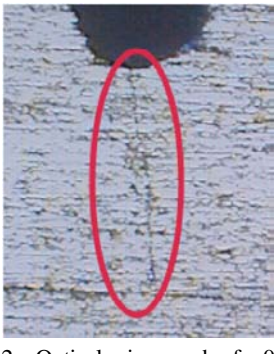


Figure 2a. Optical micrograph of a 0.67mm long crack in an aluminum bar. The crack was initiated at a saw cut, the tip of which can be seen at the top of the image.

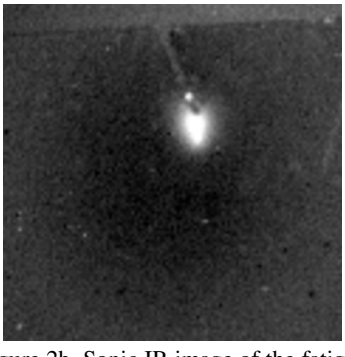


Figure 2b. Sonic IR image of the fatigue crack shown in Fig. 2a during a sound pulse. The bright spot in the saw cut is a particle of some foreign material.

3. Sonic IR Imaging for NDE: Examples of Defect Detection

Although Sonic IR imaging is a new member in the NDE family of techniques, it has demonstrated that it is a powerful technique for detecting cracks (surface or subsurface) and other defects such as delaminations, disbonds, and corrosion in materials including metals and alloys, composites, ceramics, and coatings.[5-10] In the following we show some examples.

We show, in Figure 3, the application of Sonic IR for detecting cracks in a butt weld in a metal rod intended for a specialized automotive application. The diameter of the rod is about 8mm. The left part of this rod is stainless steel, and the right part is carbon steel. The two are friction-welded together. Cracks at the interface of the two metals are typically very tight, often under the surface, and thus not readily visible. It is important to have a suitable NDE technique to identify defective rods out of the production line before they are assembled into the final products. Figures 3a, 3b, 3c, and 3d show a sequence of IR images taken using the Sonic IR imaging technique: just before the sound pulse was turned on, at about 16ms after the sound was turned on, in the middle of the pulse (after 50ms), and 8ms after the sound was turned off, respectively. The IR signal is so strong that an automated system could easily have detected the crack in a small fraction of a second. It is this speed that makes Sonic IR imaging attractive for production environments.

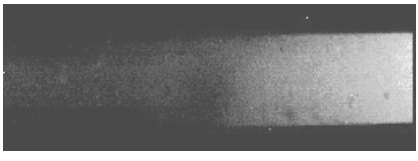


Figure 3a. An IR image taken just before the sound pulse was turned on.

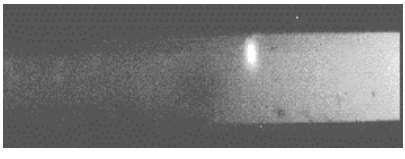


Figure 3b. An IR image taken at about 16ms after the sound was turned on.

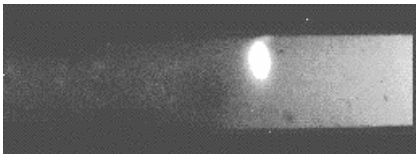


Figure 3c. An IR image taken in the middle of the pulse (after 50ms).

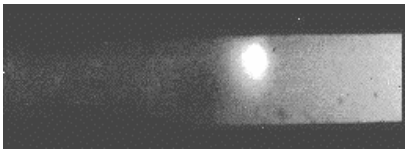


Figure 3d. An IR image taken 8ms after the sound was turned off.

To explore the application of Sonic IR to a different welding problem, a bad weld was purposely made in a notoriously poor welding material: pure copper. A photo of such a bad weld in copper is shown in Figure 4a, which shows an open crack in the middle part of the weld. However, the crack also extends under the surface on both ends well beyond the visible open section. This is shown in the Sonic IR image in Figure 4b. The IR image shows not only the open crack as a line of bright spots due to contacting asperities, together with its bright subsurface extensions, but also the lack of a good bond between a lump of copper and the surface on the left side of the weld.

Figure 4a. A photo of a copper welding sample which shows an open crack just to the right and above a lump of copper on the surface.

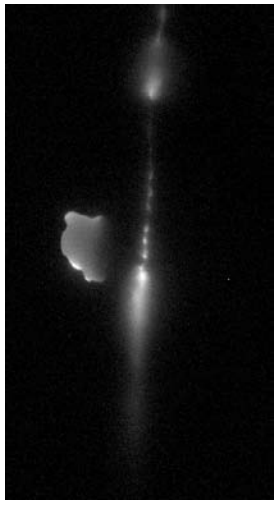


Figure 4b. IR image of the sample which not only shows the open crack (bright dots), but also the sub-surface extensions which can not be seen optically (left).

Since Sonic IR imaging detects the heat generated by rubbing crack faces, it can distinguish closed cracks from EDM notches that are in common use as calibration standards for some other NDE techniques. In Figure 5a, we show a photo of a portion of an experimental bevel gear from a helicopter. This gear has two EDM notches, one with a fatigue crack underneath, and one with no crack. A Sonic IR image is shown in Figure 5b, where only the fatigue crack lights up as marked by the solid-line oval. The EDM notch, which does not have a crack underneath, gives no indication of heating (oval with dashed line). This illustrates the fact that faying surfaces are necessary for Sonic IR imaging.

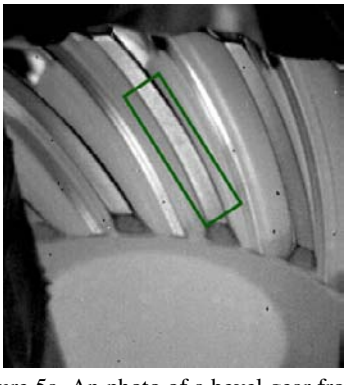


Figure 5a. An photo of a bevel gear from a helicopter which has two EDM notches in the boxed slot, one with a fatigue crack underneath, and one with no crack.

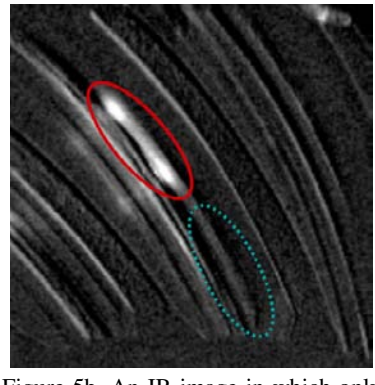


Figure 5b. An IR image in which only the notch with a crack lights up. This illustrates the fact that faying surfaces are necessary for Sonic IR heating.

The ultrasonic excitation source employed in Sonic IR imaging is in a relatively low frequency range, typically 20 - 40kHz. Such low-frequency acoustic waves infuse even a large object with multiply reflected waves within a millisecond. Therefore, this technique can be applied equally well to quite large and irregularly shaped objects, as to small delicate objects using the same apparatus. To illustrate the application of Sonic IR imaging to massive objects, we show, in Figure 6, an 84kg monolithic aluminum structure that had been involved in a high-speed accident. As a result of the collision, this object contains numerous cracks. A diagonal line of cracks, each more than 5cm long, can be seen in the top of the IR image in Figure 6b. The bright pattern at the lower right of the IR image is where the ultrasonic transducer was placed, and other cracks below the transducer were also detected at the bottom of the image.

The question of possible deleterious effects of foreign materials in defects is sometimes raised by critics of Sonic IR imaging. The object described above had been previously inspected with dye penetrant, and the penetrant was seeping out of the cracks. Although the penetrant was invisible to the naked eye, the flash from the digital camera made the dye fluoresce, and appear to be green as shown in Figure 7a, which is a close-up optical image of part of the object shown in Figure 6. We imaged this part with the Sonic IR technique, and, as shown in Figure 7b, the crack still lit up, even though it was flooded with penetrant. We also have imaged other crack samples in the laboratory after having soaked them in lubricating oil, etc., and we did not observe decreasing IR signal.

Figure 6a. An 84kg monolithic aluminum structure that had been involved in a high-speed accident. This object contains many cracks as a result of the collision.

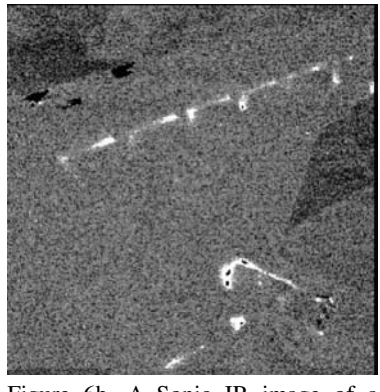


Figure 6b. A Sonic IR image of a section of the massive object. A diagonal line of cracks, each more than 5cm long, can be seen in the top of this image.

Figure 7a. A close-up optical image of part of the object shown in Fig. 6 showing penetrant seeping from a crack (Green).

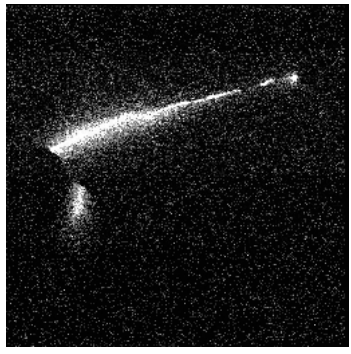


Figure 7b. Crack still lights up although flooded with penetrant.

Delaminations/disbonds can be easily detected in composite materials by Sonic IR imaging. Such an example is presented in Figure 8, in which we show a graphite/epoxy composite sample that had been subjected to impact damage. The impact caused disbonding between laminated layers at different depths, but no surface damage was visible. In Figure 8a, a Sonic IR image taken at an early time during the ultrasonic excitation pulse shows the delaminations closer to the front surface, and in Figure 8b, another Sonic IR image taken at a later time shows delaminations at deeper depths as well. Furthermore, Sonic IR imaging has the advantage of detecting "kissing disbonds", which are often invisible to the conventional thermal wave imaging technique. A thermal wave image is made by flash heating of the sample surface, and an IR camera follows the time evolution of the surface temperature after the flash. If the heat flow into the sample is blocked by a subsurface air gap, the thermal wave launched by the flash is reflected, and causes the area over the gap to be higher in temperature than the background.[11] Air gaps caused by delaminations along four sequential plies in this composite sample are seen to make a pattern much like a chicken's foot shown in the thermal wave image in Figure 8c. However, the Sonic IR images show that the sample contains much more extensive damage in the form of so-called, "kissing disbands". These disbonds do not show in the thermal wave image because they do not block the heat flow from the surface. However they are very efficient generators of frictional heat when a sound wave passes through.

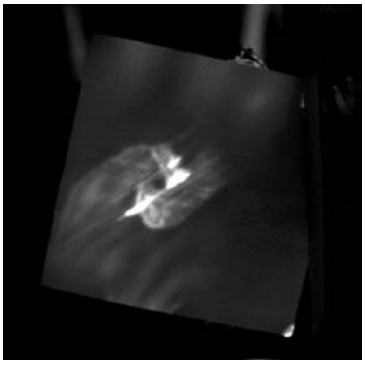


Figure 8a. A Sonic IR image taken at an early time during the ultrasonic excitation pulse shows the delaminations closer to the front surface.

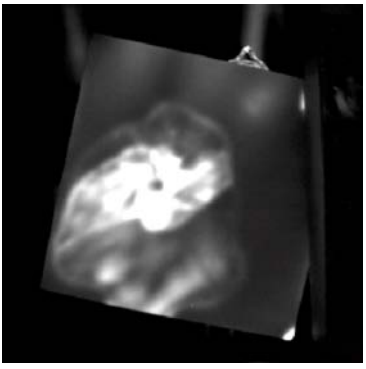


Figure 8b. Another Sonic IR image taken at a later time shows the delaminations at deeper depth as well.

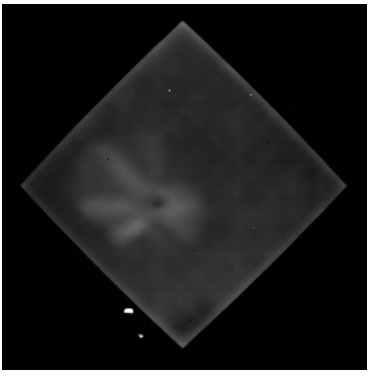


Figure 8c. A thermal wave image made with flash heating shows delaminations that result in air gaps in the sample. This is to be contrasted with the Sonic IR images above that show the existence of "kissing disbands" as well.

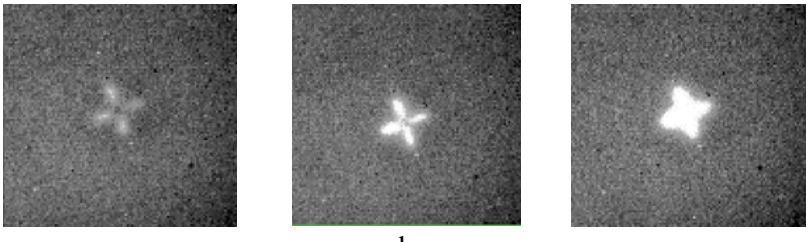


Figure 9. Sonic IR images of cracks resulting from a Vickers™ indentation in an Al_2O_3 ceramic sample taken at different stages of the ultrasonic excitation pulse. a) A Sonic IR image taken at an early time during the ultrasonic excitation pulse. b) A Sonic IR image taken in the middle of the ultrasonic excitation pulse. c) A Sonic IR image taken at a later time during the ultrasonic excitation pulse.

Although extra care is needed to avoid damage to the sample, we have demonstrated that Sonic IR imaging can also be employed to detect cracks in brittle materials. In Figure 9, Sonic IR images taken at different stages of the ultrasonic excitation pulse show four cracks radiating out in the form of a cross from a Vickers™ indentation in an Al_2O_3 ceramic sample.

In addition to crack detection, it is possible to detect corrosion using Sonic IR imaging. In Figure 10, we show cracks and corrosion on the belly of B737 aircraft at the FAA's NDI Validation center in Albuquerque. Figure 10a is an IR image taken shortly after the sonic excitation was turned on. The two cracks extend from both sides of a rivet located on the left of the image center, one small crack above the rivet and one longer crack below the rivet. Figure 10b is an IR image taken some time later during the excitation. Corroded regions on the inside surface of the aircraft skin are now beginning to show in the image. Meanwhile, two tape markers on the surface of the belly (one on the right edge, and one at the lower left) have also lit up.

As a far-field imaging technology, Sonic IR is limited by the wavelength of the radiation detected by the IR sensors for its resolution of images. The smallest crack we have detected with Sonic IR imaging so far is a 20 micron long crack in a titanium fatigue sample. This is not too far above the limit imposed by the 3 - 5 micron wavelength detectability band of our IR camera.

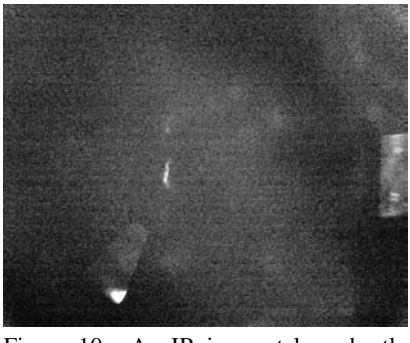


Figure 10a. An IR image taken shortly after the sonic excitation was turned on. There are two cracks extending from a rivet located on the left of the image center, a longer one underneath, and a shorter one above.

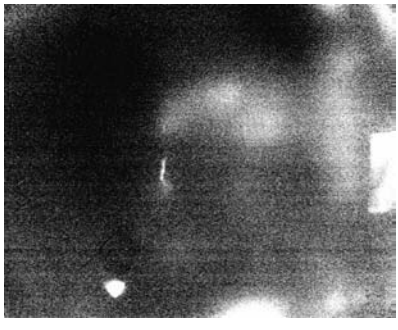


Figure 10b. An IR image taken some time later during the excitation. Heat from corroded regions on the inside surface of the aluminum skin has begun to appear in the image.

In some cases, we observe acoustic mode patterns that are built up during the course of the acoustic energy input. These patterns can be especially evident on samples with simple geometries and/or surface coatings that are acoustically

dissipative. An example of such behavior is given in Figure 11, which shows four (non-consecutive) images from a Sonic IR image sequence of a blade. The first one was taken before the ultrasonic pulse was turned on, which shows the background of part of the blade. The second image is taken at the beginning of the ultrasonic pulse, which shows the crack on the leading edge of the blade. The third one was taken at early stage of the ultrasonic pulse, when a mode pattern had built up on the edge, although the crack is also clearly shown. The fourth image was taken at later stage of the excitation pulse, when the vibration in the sample changed to a different mode that we call "acoustic chaos" (described in the next section). The change of vibration mode spoiled the mode pattern in the blade edge and therefore only the crack lights up. This illustrates the fact that the heating in the vicinity of cracks ordinarily occurs before mode patterns have time to build up, thus providing a means of distinguishing between a crack and a mode pattern. In addition, acoustic chaos can be introduced into Sonic IR imaging to eliminate or reduce the mode pattern.

There are several factors that affect the IR signals in Sonic IR imaging. These factors include ultrasonic frequency,[12] coupling between the transducer and the sample, boundary conditions, material characteristics of the sample, etc.

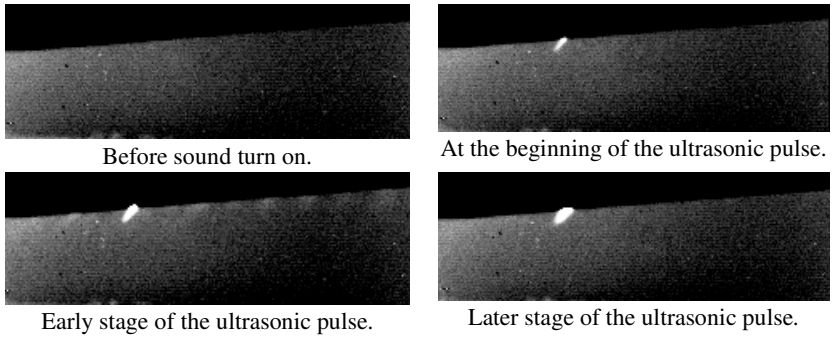


Figure 11. Four (non-consecutive) images from a Sonic IR image sequence of a blade. These images show both a crack and mode pattern in the leading edge of the blade. They illustrate the timings of the heating in a crack and a mode pattern.

4. Acoustic Chaos in Sonic IR Imaging

Although the sonic excitation source provides a nominal single driving frequency, the vibration in a sample coupled to it may contain many frequency components in addition to the excitation driving frequency. In many situations, some of these different frequency components can be much stronger than the driving frequency. The vibration waveform measured on the tip of the transducer by means of the

Doppler laser vibrometer during its excitation, when it is not loaded by acoustic coupling to the test object, has a single frequency in its spectrum. When the Sonic IR experiment is underway, the tip of the transducer is coupled through a layer of softer material to the object under inspection with enough pressure to maintain contact without causing the sonic source to shut down. The excitation of the part itself, when strongly coupled to, and excited by, the transducer, is not constant in amplitude, nor is it a single frequency. Fourier analysis shows that the waveform typically has several stages, with the frequency components of each stage being slightly different. However, in each stage, the spectrum consists of an integral fraction, say for instance one thirteenth or one seventh, etc., of the transducer frequency, plus very many harmonics of that fractional frequency. For example, Figure 13a shows the waveform from a titanium specimen when 40kHz sonic excitation was used. Figure 13b shows the spectrum of the time interval marked by the lines in Figure 13a. The spectrum contains a series of frequencies that are integral multiples of one fourteenth of the transducer frequency. Even though the driving frequency from the sonic source is pure 40kHz, numerous different frequency components are present in the vibration of the sample, both lower frequencies and much higher frequencies. These additional frequencies are the signature of acoustic chaos. The chaotic behavior, which is caused by the nonlinear coupling between the transducer and the object, has been studied analytically and by finite element analysis, in addition to the experimental results.[13, 15-16]

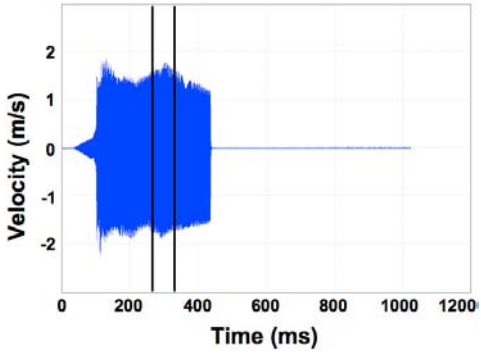


Figure 13a. Waveform from a titanium specimen when 40kHz sonic excitation was used. The vibration in the sample switched to different stages in which its frequency components are very different. The spectrum of the region marked is shown in Figure 13b.

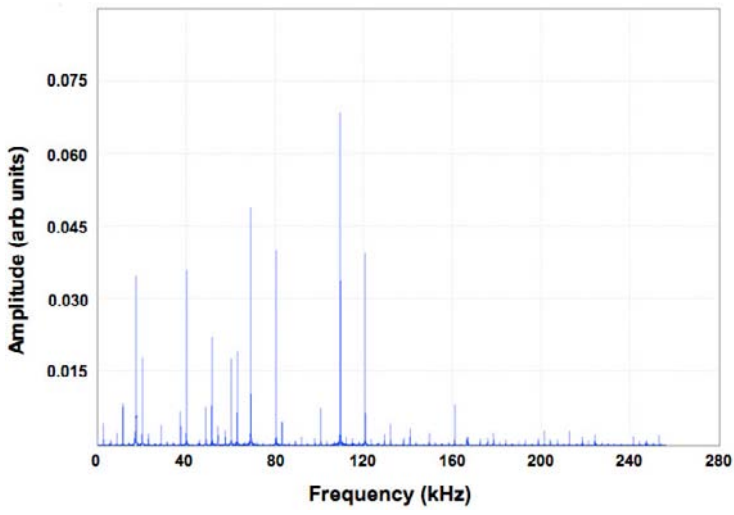


Figure 13b. The spectrum of the stage marked by the lines in Figure 13a. The spectrum contains a series of frequencies that are integral multiples of one fourteenth of the fundamental frequency.

The discovery of acoustic chaos is important, because it appears that chaotic vibrations are much more efficient at heating cracks, than vibrations with the pure transducer frequency, or with that frequency and its harmonics.[14] To illustrate the effects of chaotic sound, we show the results from two carefully controlled successive experiments of chaotic and non-chaotic Sonic IR imaging on an Al plate containing a fatigue crack on one of its edges. In Figure 14a we show the time dependence of the acoustic power, as measured by the mean squared vibrational velocity, averaged over three millisecond intervals. One of the pulses was chaotic, and the other was non-chaotic. The average power of the non-chaotic pulse is much higher than that of the chaotic pulse. The corresponding Sonic IR temperature versus time curves in a region near the crack are shown for comparison in Figure 14b. The non-chaotic and low-power chaotic curves have essentially the same maximum temperature. Sonic IR images corresponding to the peaks of the two temperature-time curves are shown in Figure 14c. It is apparent that the chaotic waveform produces the same or even higher temperature with less acoustic power. Thus, Sonic IR imaging can be carried out with lower acoustic power with the same image contrast if chaotic sound is used, or alternatively, one can significantly improve the detectability of a crack using chaotic sound of the same power. Sometimes, acoustic chaos is absolutely essential to detect defects in certain materials and structures. Since chaotic vibrations are much more efficient at heating cracks than vibrations with a pure frequency, or with a frequency and its harmonics, acoustic chaos promises to make Sonic IR a much better crack detector.

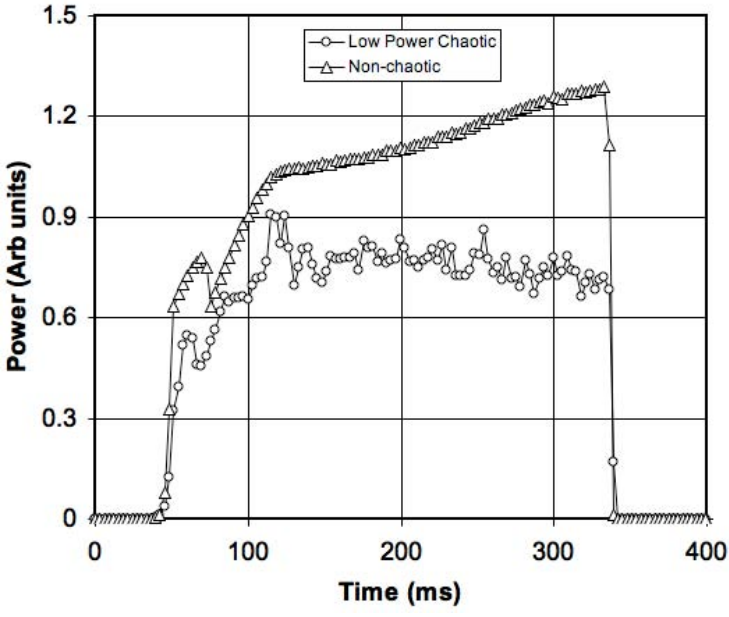


Figure 14a. Time dependence of the acoustic power for a chaotic and a non-chaotic pulse.

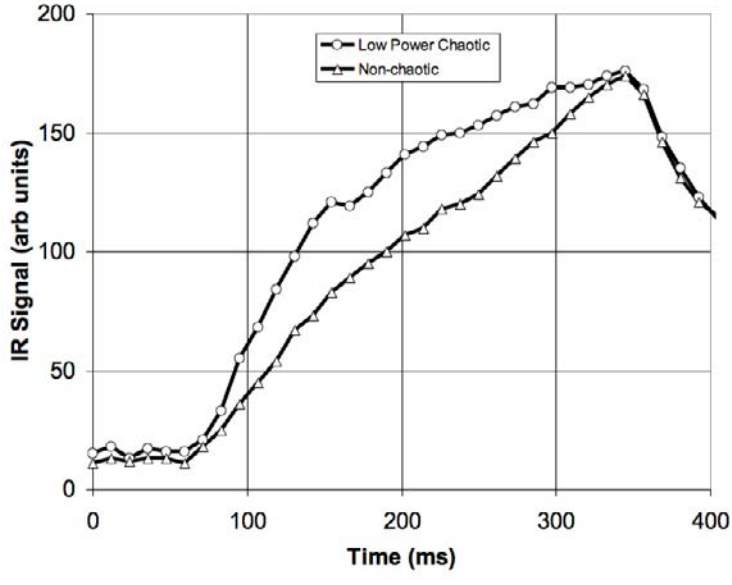


Figure 14b. Sonic IR temperature vs. time curves near the crack for the corresponding chaotic and non-chaotic cases above.

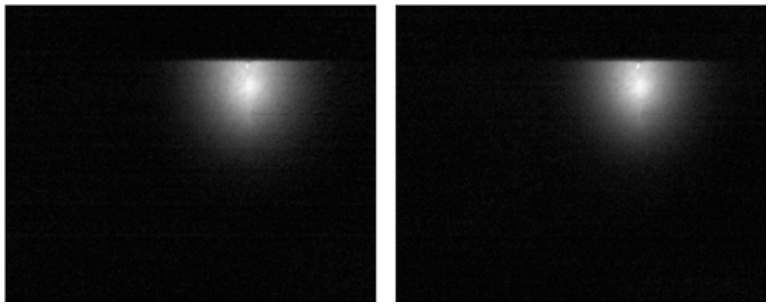


Figure 14c. Sonic IR images corresponding to the peaks of the two temperature-time curves in Figure 14b. The one on the left is from the low power chaotic case, and the one on the right is from the higher power non-chaotic case. The images are indistinguishable.

From the investigation on chaotic sound and its effect on heating through both experimental work and finite element analysis, it appears that acoustic chaos produces much bigger relative motion between crack faces than that in the non-chaotic case for the same input acoustic power.[17] Thus, more frictional heating is produced in the crack for the chaotic case than for the non-chaotic case.

5. Conclusions

Sonic IR imaging is a fast, and wide area NDE imaging technique. It can be used to detect either surface or subsurface defects, including cracks, delaminations and disbonds, in materials such as metals/metal alloys, composites, ceramics, coatings, and adhesives. Acoustic chaos can be used to enhance IR signals since it increases frictional heating in defects.

References

1. Edmund G. Henneke, II, Kenneth L. Reifsnider, and Wayne W. Stinchcomb, J. Metals 31, 11 (1979).
2. K.L. Reifsnider, E.G. Henneke, II, and W.W. Stinchcomb, in The Mechanics of Nondestructive Testing, W.W. Stinchcomb, ed., Plenum Press, New York, pp. 249-276 (1980).
3. R.B. Mignogna, R.E. Green, Jr., J.C. Duke, Jr., E.G. Henneke, II, and K.L. Reifsnider, Ultrasonics 19, 159 (1981).
4. D. Wu, J. Rantala, W. Karpen, G. Zenzinger, B. Schoenbach, W. Rippel, R. Steegmueller, L.Diener and G. Busse, Review of Progress In Quantitative Nondestructive Evaluation, eds., Donald O. Thompson and Dale E. Chimenti, Vol. 15, 511 (1996).
5. L.D. Favro, Xiaoyan Han, Zhong Ouyang, Gang Sun, Hua Sui, and R.L. Thomas, Review of Scientific Instruments 71, 2418 (2000).
6. Xiaoyan Han, L.D. Favro, Zhong Ouyang, and R.L. Thomas, Journal of Adhesion 76, 151 (2001).

7. L.D. Favro, Xiaoyan Han, Li Li, Zhong Ouyang, Gang Sun, Austin Richards, and R.L. Thomas, Review of Progress in Quantitative NDE Vol. 20, ed. by D.O. Thompson and D. Chimenti, CP557, Am. Inst. Phys., 478-482 (2001).
8. L.D. Favro, R.L. Thomas, X. Han, Z. Ouyang, G. Newaz, and D. Gentile, Int. Journal of Fatigue 23, S471 (2001).
9. Xiaoyan Han, L.D. Favro, Zhong Ouyang, and R.L. Thomas, Review of Progress in Quantitative Nondestructive Evaluation, Am. Inst. Phys. CP615, Vol. 21, pp. 552-557, 2002.
10. Xiaoyan Han, L.D. Favro, and R.L. Thomas, Review of Progress in Quantitative Nondestructive Evaluation 22, ed. By D.O. Thompson and D.E. Chimenti, Am. Inst. Phys. CP657, pp. 500-504 (2003).
11. Xiaoyan Han, L.D. Favro, T. Ahmed, Xun Wang, and R.L. Thomas, Review of Progress in Quantitative Nondestructive Evaluation, 18, ed. By D.O. Thompson and D.E. Chimenti, Am. Inst. Phys. pp593-596, Vol. 18, (1999).
12. Xiaoyan Han, Review of Scientific Instruments 74 (2003).
13. Xiaoyan Han, Wei Li, Zhi Zeng, L.D. Favro, and R.L. Thomas, Applied Physics Letters 81, 3188 (2002).
14. Xiaoyan Han, Zhi Zeng, Wei Li, Md. Islam, Jianping Lu, Vera Loggins, E. Yitamben, L.D. Favro, G. Newaz, and R.L. Thomas, J. Appl. Phys., 3792 (2004).
15. Xiaoyan Han, V. Loggins, Zhi Zeng, L.D. Favro, and R.L. Thomas, Applied Physics Letters 85, 1332 (2004).
16. Xiaoyan Han, Md. Islam, G. Newaz, L.D. Favro, and R.L. Thomas, J. Appl. Phys., 98 (2005).
17. Xiaoyan Han, Md. Islam, G. Newaz, L.D. Favro, and R.L. Thomas, J. Appl. Phys., 99 (2006).

This page intentionally left blank

CHAPTER 17

A DATA FUSION FRAMEWORK FOR MULTIPLE NONDESTRUCTIVE INSPECTION IMAGES

Zheng Liu*, David S. Forsyth†, Pradeep Ramuhalli‡, and Abbas Fhar§

I i e f Ae ace Re ea ch, Na i a Re ea ch C ci Ca ada, M ea R ad 1200, B i di g M-14, O a a, O a i K1A 0R6 Ca ada
he g. i @ c-c c.gc.ca

Multiple nondestructive inspection (NDI) techniques were applied to detect and quantify the hidden corrosion in aircraft lap joints. The inspection data is presented in raster-scanned images. In this chapter, the use of Dempster-Shafer (DS) theory to fuse multi-frequency and pulsed eddy current inspection data is presented. The NDI images are first discriminated by iteratively trained classifiers. The basic probability assignment (BPA) is defined based on the conditional probability of information classes and data classes, which are obtained from a teardown inspection followed by a X-ray thickness mapping process and the NDI measurements respectively. The DS rule of combination is applied to fuse multiple NDI inputs. The remaining thickness is estimated by applying locally weighted regression of the DS-fused results.

1. Introduction

The development of image sensors makes it possible to combine information across the electromagnetic spectrum by the fusion of multi-modal images. This technique is known as image fusion, which is deemed to provide a more accurate and reliable result than any single image. The applications of image fusion have been found in various fields.[1] Two dimensional image is a common format to present nondestructive inspection (NDI) results. When multiple NDI techniques are involved, it is possible to fuse multiple NDI images to achieve an improved evaluation result.

Data fusion is a more general concept than image fusion. The research on applying data fusion techniques to nondestructive inspection started almost two

*Dr. Zheng Liu is with the Institute for Research in Construction, National Research Council Canada since June 2006.
†Nondestructive Testing Information Analysis Center, 415 Crystal Creek Drive, Austin, Texas, USA
‡NDE Laboratory, Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA
§Institute for Aerospace Research, National Research Council Canada, Montreal Road 1200, Building M-14, Ottawa, Ontario K1A 0R6 Canada

decades before.[2,3] The challenges from the inspection make it necessary to employ multiple NDI techniques. In order to take advantage of all the techniques, data fusion plays an important role for interpreting the acquired signals. Data fusion itself is not isolated science or technology; it encompasses all the techniques and tools, which can integrate data originating from different sources. The evaluation of fusion results remains an application-dependent issue.

The general procedure for image fusion is depicted in Figure 1, where the procedure starts with a preprocessing stage in which any essential operations like noise removal operation could be applied. Second step is the registration of multiple images, i.e. corresponding pixels should be associated with the same physical points on the object so that the input images can be compared pixel by pixel. Next comes the image fusion process. Depending on the fusion algorithm, the fusion result could be a composite image, a thematic map, or a three-dimensional graphic. Usually, when the fusion is implemented at pixel level, a post-processing step may be applied to the fused image, such as classification, segmentation, enhancement, etc. Quantitative results may be achieved through the post-processing stage. The efficiency of the fusion algorithm can be assessed through the predefined metrics, which depend on the requirements of the application. The quantitative results from the evaluation may be used as feedback to guide the fusion process where applicable. An adaptive fusion process can be achieved, but this still remains a challenge for most applications.

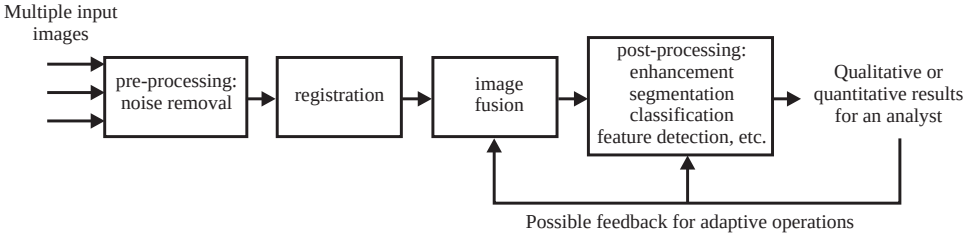


Fig. 1. The general procedure for image fusion.

There are a number of algorithms available to fuse NDI images. Reports on using optimal filtering, neural networks, multiresolution analysis (MRA), Bayesian inference, and Dempster-Shafer (DS) theory for fusing NDI images can be found in the references.[4–10] Details about these methods will not be repeated here. Among all the approaches, the Dempster-Shafer theory provides a framework to handle multi-source information fusion.[11] The flexibility of this method is attributed to the basic probability assignment (BPA) or the definition of a probability mass function. The fusion process is implemented through an updating mechanism, where the information from each sensor is modeled with a BPA value, $m(A)$, between zero and one in a discernment frame based on θ mutually exclusive and exhaustive propositions. Thus, there are 2^θ singletons in the frame, i.e. $m : 2^\theta \to [0, 1]$. The

subset $A \subset 2^{\theta}$ can be either a singleton or a combination of multiple propositions. Therefore, the belief function is defined as the summation of all the BPAs assigned to the subsets of proposition A, i.e. $Bel(A) = \sum_{B \subset A} m(B)$. $Bel(A)$ represents a degree of support that all the available evidence provides for A.

Lap joint is a common structural element in transport aircraft fuselarge design and it is one of the common airframe structural elements across a variety of aircraft models. A typical lap joint structure consists of multiple sheets of aluminum overlapped and riveted. If the sealants or protection systems break down, corrosion may happen on the faying surfaces of these sheets. This type of corrosion is also known as hidden corrosion. According to Fahr et al.,[12] the material loss by layer can serve as one of the corrosion metrics for structural integrity and life prediction analysis of aircraft lap joint structures. This metric is obtained by estimating the remaining thickness of each layer. Therefore, nondestructive inspection of lap joints was carried out to map the hidden corrosion.[13,14] The purpose of fusing multiple NDI images was to generate thickness maps of the lap joints for better estimation of the remaining thickness. The DS-based fusion is employed to integrate data/information obtained from the inspection with multiple NDI techniques.

In this chapter, a data fusion framework is proposed for fusing multiple NDI images of lap joints. The data fusion operation starts with training classifiers for multiple NDI inputs (images). A probability model between the data class and information class is set up based on the information retrieved from the available training data sets. The combination rule of the DS theory is applied to fuse the classification results of the new (testing) set of NDI inputs and the fusion result is obtained. To convert this fusion result into a numeric value of thickness, a locally weighted regression is employed.[15]

2. The Fusion Strategy

As described in reference,[12] the material loss by layer serves as a corrosion metric for lap joint structural integrity analysis or fatigue life prediction. To fully characterize the hidden corrosion, it is necessary to have information on the location of corrosion and the remaining thickness by layer. Locating corrosion by layers accurately still remains a challenge, especially for corrosion between the faying surfaces. The pulsed eddy current (P-ET) technique is a promising NDI method that may provide such information.[16] Unlike the conventional ET that uses single frequency continuous-wave excitation, the P-ET probe is excited by a broadband pulse. Therefore, the P-ET technique in principle is able to provide information from different depths of the specimen.

A data fusion-based analysis system for corrosion characterization in a two-layer structure is shown in Figure 2. The P-ET technique is to identify the location of corrosion, e.g. top of bottom layer or bottom of top layer. The results from P-ET inspection can further be fused with other NDI measurements such as multi-

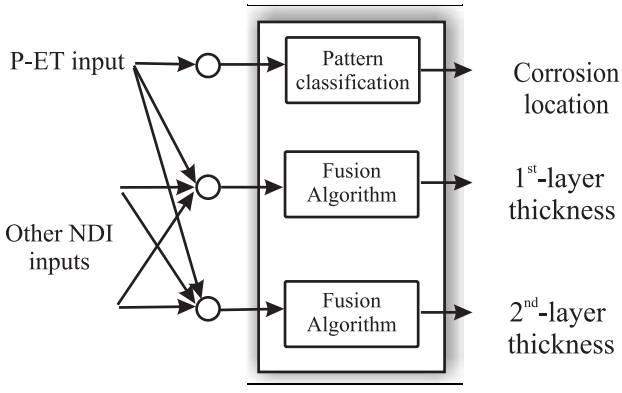


Fig. 2. A data fusion approach to characterize the aircraft multi-layer structures.

frequency eddy current to quantify the remaining thickness for the top and bottom layer respectively.

Earlier work on characterizing corrosion in multi-layered structures is reported in references,[17,18] where the P-ET signal is transformed to time-frequency domain with Wigner-Ville distribution as a 2D pattern. This pattern is mapped into a feature space by the principle component analysis (PCA) method. The P-ET signal is then classified in the feature space. Although the results are promising, the lack of representative data to train a good classifier still remains a problem.

Once the corrosion location in a multi-layer structure is identified, the remaining thickness by layer can be estimated from the multiple NDI measurements, i.e. multi-frequency eddy current and pulsed eddy current. Thus, two issues remain: what is the best combination of the NDI inputs or features for certain layer and how to find the optimal parameters for the data fusion module. In other words, the task is to optimally map the measurement inputs to the target characterization, e.g. remaining thickness. The proposed data fusion framework provides a solution to these problems.

As depicted in Figure 2, the goal of data fusion is to find the mapping function $f : \Re^m \rightarrow \Re$, when the input vector is $\vec{x} \in \Re^m$. Obviously, the individual measurements can be related to the target outputs with a simple linear or nonlinear regression approach. A smaller error exhibits a better fit of the input and the corresponding output. Previous work[19] indicates that nonlinear regression can improve the individual estimation result while a Generalized Additive Model (GAM) achieves a better performance by fusing multiple inputs. The regression analysis reveals the relation between the inputs and outputs and therefore helps to use these data.

Another way to address the characterization problem of a two-layer lap joint is by subtracting the first-layer thickness from the total of two-layer's.[13,20] However, this approach still needs a procedure to match the images before any pixel-based

operation is carried and this method is limited to two-layer structures only. This chapter will focus on the study of building a mapping function for multiple inputs, i.e. the data fusion algorithms to optimally integrate those inputs.

2.1. *The Procedure for DS-Based Fusion*

As shown in Figure 3, the proposed data fusion architecture consists of three parts. The first part is the classifier training and data modeling, where a labeled data set is needed. The initial label can be found by a clustering method, which is based on the characteristic of the data itself. These labels define the "data class". Suppose there is M data classes denoted by $\omega_i, i = 1, ..., M$ into which the pixels from an NDI image are classified. An iterative training process is implemented to train the pre-selected classifiers.

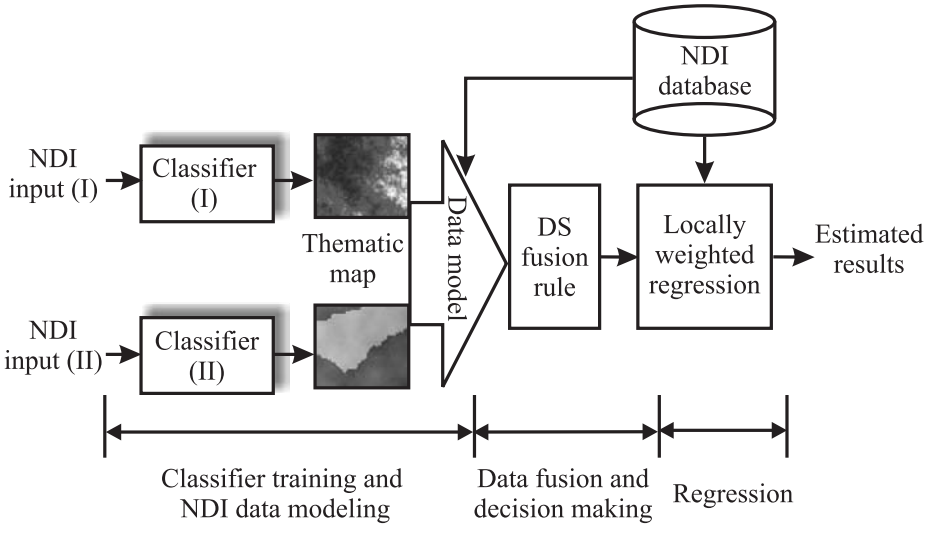


Fig. 3. The data fusion framework.

The ground truth data, which is obtained by the digital X-ray mapping technique, is used to generate the "information class", which represents the remaining thickness range.[13] A list of such classes is shown in Table 1. In this case, there are totally seven corrosion classes and each class number stands for a thickness range. Because the pixel value in a X-ray thickness map reflects the remaining thickness directly, simply applying the thickness range in Table 1 can get the "information classes". The relation between the "data class" and the "information class" is described by a statistical model. The definition of the basic probability assignment is based on this model. Using the DS combination rule to update the classification results of the inputs, we can obtain a thematic map representing information classes. To obtain a numeric result, a locally weighted regression can be applied.[21]

Table 1. The definition of information classes in term of material loss.

Corrosion class	Thickness range (inch)
C_1	$0.024 - 0.027$
C_2	$0.027 - 0.030$
C_3	$0.030 - 0.033$
C_4	$0.033 - 0.036$
C_5	$0.036 - 0.039$
C_6	$0.039 - 0.042$
C_7	$0.042 - 0.045$

2.2. *The Training of Classifiers*

An iterative classifier training strategy is proposed as shown in Figure 4. The NDI image is first segmented with the fuzzy k-means clustering method.[22] A labeling image Λ_j is obtained. The result is used as the initial label to train a classifier. Applying the classifier to the NDI image generates a new labeling image Λ_{j+1}. The differences between the two images Λ_j and Λ_{j+1} are counted. If the difference exists, the new labeling image is used to train the classifier in the next turn, i.e. $\Lambda_j = \Lambda_{j+1}$. The process will stop until no difference between two labeling images is observed. Figure 5 demonstrates the convergence of the classifier training process.

The problem with the fuzzy k-means algorithm is the choice of the cluster number. An optimal value can be selected by carrying out a validity analysis.[23,24] In our case, a cluster number 8 and 7 are selected for the P-ET and ET data respectively. Moreover, the nearest mean classifier (NMC) outperforms the other classifiers, such as normal density based linear and quadratic classifiers in terms of the classification error.[22] Therefore, the NMC is used in the classifier training process.

2.3. *Defining BPA*

Assuming that C_i $(i = 1, 2, ..., N)$[a] represents the information classes, i.e. thickness range, and vector $\vec{x}$ contains the multi-modal NDI measurement values, the BPA can be defined as the probability of being a certain information class based on the statistical information from available training data sets, i.e. $m_s(C_i) = p\left(C_i \left| \vec{x} \right.\right)$.[25] Herein, s indicates the different data sources $(s = 1, 2, ..., S)$ or the number of the inputs. The vector $\vec{x}$ is mapped to data class d_j $(j = 1, 2, ..., M)$ by a classification operation. Thus, the BPA can be defined as:

$$m_s(C_i) = p\left(C_i \left| \vec{x} \right.\right) p_s(d_j | C_i) \tag{1}$$

where $p_s(d_j | C_i)$ means the probability for a pixel from input s to be classified as a "data class" d_j, given the thickness range ("information class") C_i. This

[a]As shown in Table 1, N should be 7 in this case.

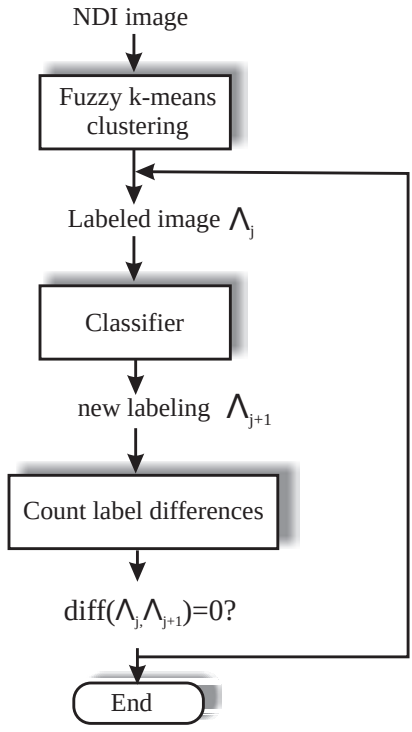


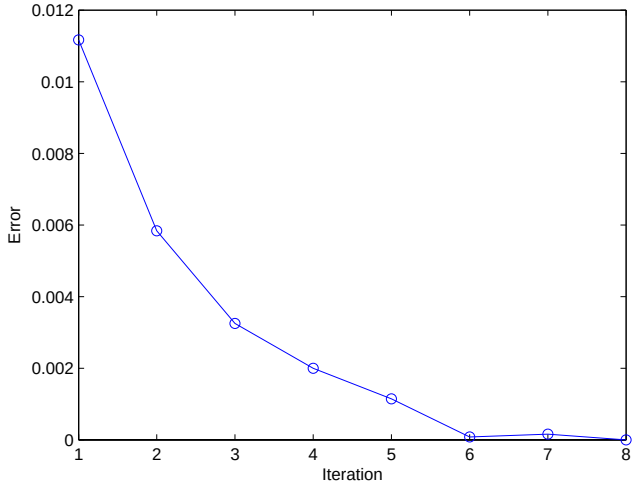
Fig. 4. The iterative classifier training.

$p_s(d_j|C_i)$ value is regarded as a measure of the capability of each data source in discriminating the "information classes". According to DS theory, the BPA values must be normalized to meet the requirement $\sum_i m_s(C_i) = 1$ before the updating operation is applied.

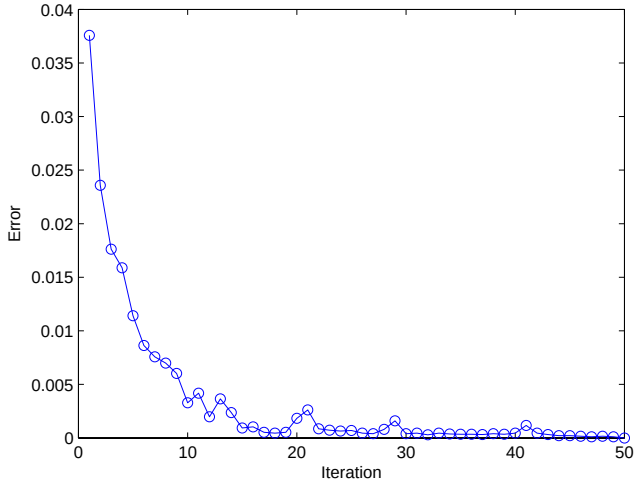
2.4. *Fusion with DS Updating Rule*

The mass value can be updated according to equation (2), where the BPA values from the P-ET inspection m_{P-ET} and eddy current (ET) m_{ET} are combined. The final decision is made based on the maximum belief value, i.e. the maximum mass output.[26] Therefore, the fusion result is represented by a thematic map, where the class number indicates a different thickness range as listed in Table 1. Meanwhile, the value $m_{ET,P-ET}(C_k)$ can be used as an indication to the reliability of the fusion result.

$$[ht]\, m_{ET,P-ET}(C_k) = \frac{\sum_{C_i \cap C_j = C_k} m_{ET}(C_i)\, m_{P-ET}(C_j)}{1 - \sum_{C_i \cap C_j = \phi} m_{ET}(C_i)\, m_{P-ET}(C_j)} \tag{2}$$



(a) The converge of classifier for P-ET data.



(b) The converge of classifier for ET data.

Fig. 5. The convergence of iterative training.

The equation (2) can also be expressed with Table 2. In our case, we only assign the mass function value to a singleton in the discernment frame, e.g. C_i or C_j, which are mutually exclusive. Therefore, the intersection of the two sets of C_i and C_j is $\{\phi\}$. The decision will be made based on the value $m_{ET,P-ET}\{C_i\}$ and $m_{ET,P-ET}\{C_j\}$, i.e. the pixel will be eventually classified as either C_i or C_j. The class number with the larger mass value will be chosen.

Table 2. The updating mechanism of DS theory.

		m_{ET}	
		$\{C_i\}$	$\{\theta\}$
m_{P-ET}	$\{C_j\}$	$\{C_k\} = \{\phi\}$	$\{C_j\}$
	$\{\theta\}$	$\{C_i\}$	$\{\theta\}$

2.5. *Locally Weighted Regression*

So far, what we obtain from the fusion process is a thematic map, which indicates the extent of corrosion in different areas. However, a numeric value of thickness loss due to corrosion is needed as the input to the structural integrity model. Thus, locally weighted regression (LWR) is employed to estimate this value using information pertaining only to a neighborhood of the input query, which is determined by the DS fusion result.

The LWR method is a form of lazy learning or memory-based learning, which defers processing of training data until a query needs to be answered.[21] The training data is stored in memory or a database. Relevant data is found from the database when a query is presented. The LWR method uses locally weighted training to average, interpolate, extrapolate, or combine training data.[27]

Suppose the variable represents the material thickness, we have the NDI measurement vector $\vec{x} \in \Re^m$ and need to find the mapping function $f : \Re^m \to \Re$ with the training data set $\left\{\left(\vec{x}_p, y_p\right)\right\}_{p=1}^{n}$. There is:

$$y_p = f\left(\vec{x}_p\right) + \varepsilon_p \tag{3}$$

where ε_p is a random variable such that $E[\varepsilon_p] = 0$ and $E[\varepsilon_p \varepsilon_q] = 0$, $\forall p \neq q$. Given a query point $\vec{x}_p$, to obtain the regression applicable to this query point, the following cost function is minimized:

$$J = \sum_{p=1}^{n} \left(f\left(\vec{x}_p, \beta\right) - y_p\right)^2 K\left(\frac{\left|\vec{x}_p - \vec{x}_q\right|}{h}\right) \tag{4}$$

where, $K(\cdot)$ is a weight function and h is the bandwidth. X is matrix of data samples with $\left[\vec{x}_p, 1\right]^T$ in the p^{th} row and y $= [y_1, y_2, \cdots, y_n]^T$. The solution to the above cost function is:[21]

$$\hat{\beta} = \left(\mathbf{Z}'\mathbf{Z}\right)^{-1} \mathbf{Z}'\mathbf{v} \tag{5}$$

where Z = WX and v = Wy . W is a diagonal matrix with the i^{th} diagonal element $w_{ii} = \sqrt{K\left(\left|\vec{x}_p, \vec{x}_q\right| / h\right)}$. The estimation is then given by:

$$\hat{y}\left(\vec{x}_p\right) = \hat{\beta}\vec{x}_p \tag{6}$$

Therefore, the fusion result obtained in section 2.4 can be further quantified with equation 6.

3. Experimental Results

3.1. *Inspection of Aircraft Lap Joints*

The specimen used in the experiments comes from a 30-year-old service-retired Boeing 727 aircraft. It is a two-layer lap joint cut out from below the cargo floor, near the belly of the fuselage as shown in Figure 6. This is one of the three common structural elements susceptible to hidden corrosion damage: multi-layer thin structures, thick section fastener holes, and internal structural components.[12] The material is aluminum 2024-T3, except for the stringer, which is aluminum 7075-T6. The thickness of each layer is 1.14mm (0.045 inch). Following inspections, the specimen was dissembled and cleaned of all corrosion products. Then, accurate thickness loss measurement was carried out by using a digital X-ray mapping technique on each layer.[13] The lap joint specimen was finally cut into seven sections, resulting in fourteen pieces in total including both the top and bottom plates.

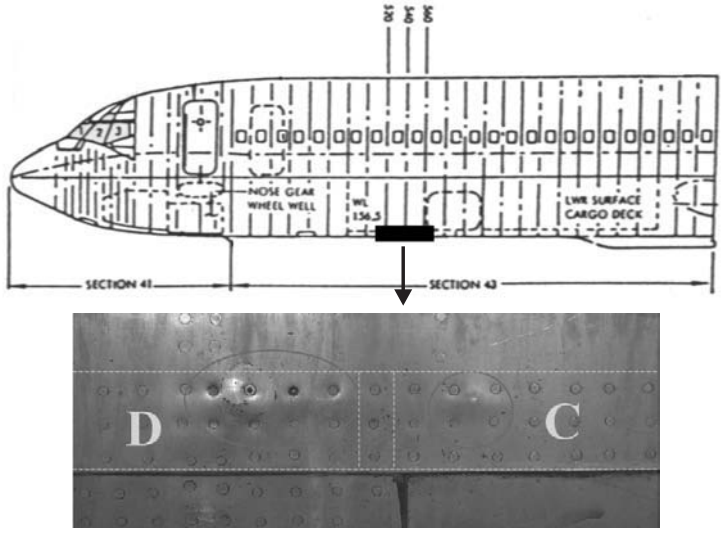


Fig. 6. A lap joint from a service-retired Boeing 727 aircraft.

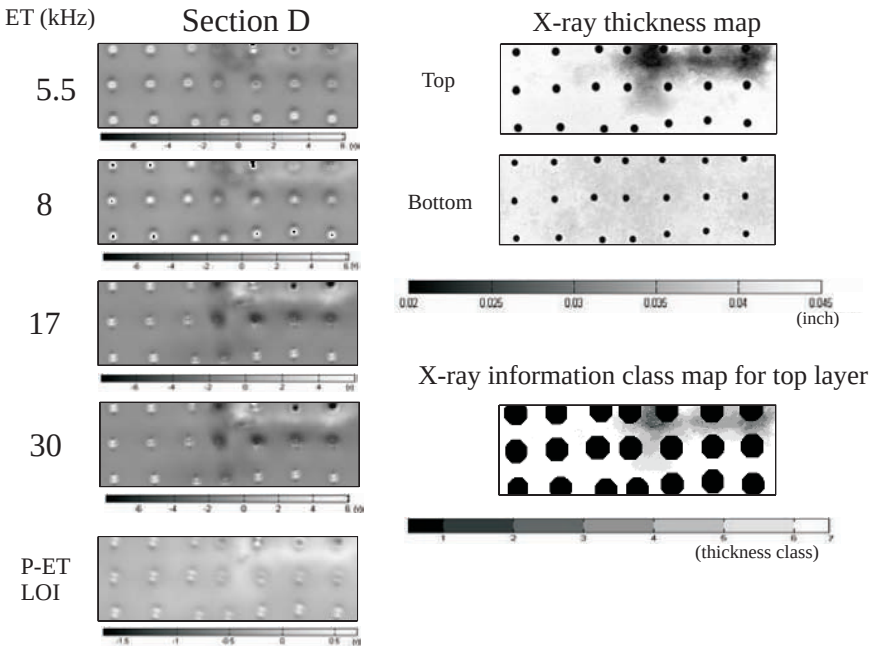


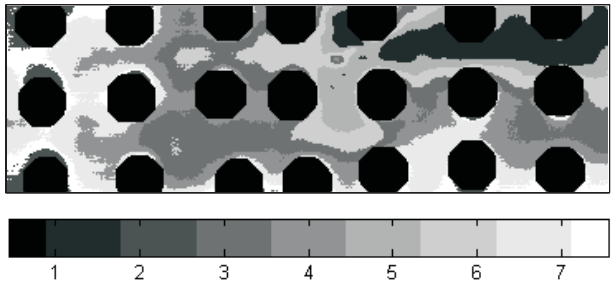
Fig. 7. The NDI inspection data.

As described in section 2, multi-frequency ET and P-ET techniques were employed in the inspection. Based on the penetrating thickness needed, the ET probe was operated at four excitation frequencies, i.e. $5.5kHz$, $8kHz$, $17kHz$, and $30kHz$. Only the real component of the ET measurement was used in the quantification process. The lift-off noise was assigned to the real part through instrumental setup to eliminate the effect of probe lift off. For the P-ET measurement, a time-domain feature called "lift-off intersection" (LOI) point was employed to quantify the hidden corrosion.[28,29] Previous studies[16,28] report that the LOI feature is not sensitive to lift-off effect. Figure 7 gives the inspection results of the two techniques. The X-ray thickness maps for the top- and bottom-layer are also presented. The information class map of the top layer was obtained from the top-layer X-ray thickness map by applying the definition in Table 1 and digitally removing the images of the fastener holes to avoid interference in analysis.

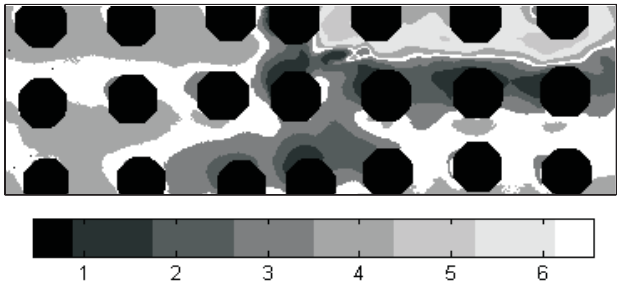
3.2. *Fusion Results*

In the experiment, we fuse the P-ET LOI image with two groups of ET images. The first group includes ET $17kHz$ and $30kHz$ images and the second group includes ET $8kHz$ and $30kHz$. The classification results for P-ET and ET data are shown in Figure 8 and the DS-based fusion results are presented in Figure 9. The X-ray thickness reference is also given at the bottom for comparison in this figure. In this

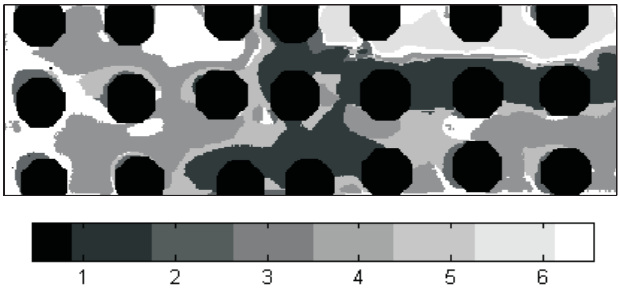
work, only the fused classification results for the top layer are presented. Locally weighted regression is then used to obtain a numerical estimate of the thickness.[15] The results are shown in Figure 10. The estimation can be validated using the available X-ray thickness map as references. Similarly, the same procedure can be applied to the second or third layer, where applicable.



(a) The P-ET classification result.



(b) The ET (17kHz and 30kHz) classification result.



(c) The ET (8kHz and 30kHz) classification result.

Fig. 8. The classified results from NDI inputs.

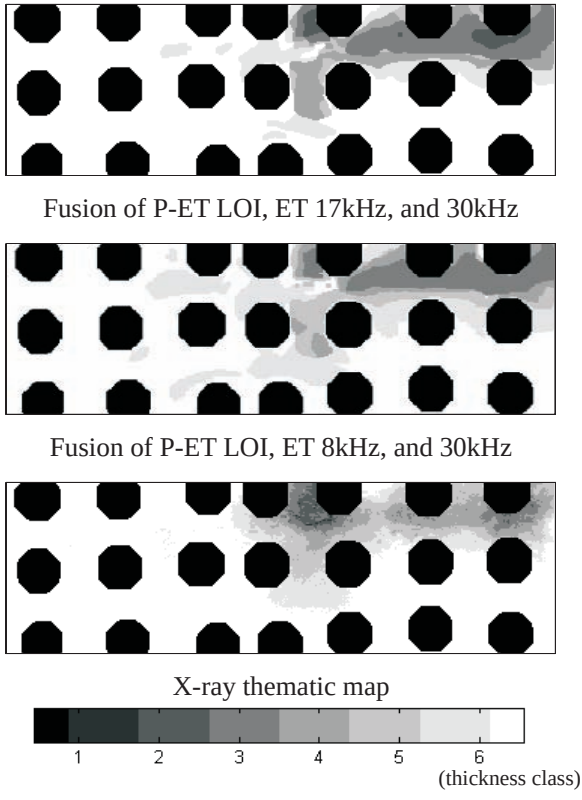


Fig. 9. The DS fusion results.

To compare the estimated results with the X-ray thickness maps, the following metrics are employed, i.e. root mean square error (RMSE), cross-correlation (CORR), peak signal-to-noise ratio (PSNR), difference entropy (DE), and mutual information (MI) and the results are provided in Table 3. The equation (7) to (11) are the mathematical expressions for these metrics.

Table 3. The comparison of the fusion results.

	RMSE	CORR	PSNR	DE	MI
P-ET LOI + ET 17kHz + ET 30kHz	2.8572	0.9965	15.4177	1.9816	0.9763
P-ET LOI + ET 8kHz + ET 30kHz	3.1790	0.9956	9.7635	1.8961	0.9685

Root mean square error:

$$RMSE = \sqrt{\frac{\sum_{m=1}^{M}\sum_{n=1}^{N}\left[I\left(m,n\right)-F\left(m,n\right)\right]^{2}}{MN}} \tag{7}$$

final result of P-ET LOI, ET 17kHz, and 30kHz

final result of P-ET LOI, ET 8kHz, and 30kHz

X-ray thickness map

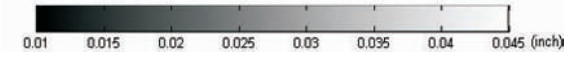


Fig. 10. The final results.

Peak signal-to-noise ratio:

$$PSNR = 10\log_{10}\left(\frac{L^2}{\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[I\left(m,n\right)-F\left(m,n\right)\right]^2}\right) \tag{8}$$

Cross-correlation:

$$CORR = \frac{2\sum_{m=1}^{M}\sum_{n=1}^{N}I\left(m,n\right)F\left(m,n\right)}{\sum_{m=1}^{M}\sum_{n=1}^{N}I\left(m,n\right)^2+\sum_{m=1}^{M}\sum_{n=1}^{N}F\left(m,n\right)^2} \tag{9}$$

Difference entropy:

$$DE = \left|\sum_{g=0}^{L-1}P_I\left(g\right)\log_2 P_I\left(g\right)-\sum_{g=0}^{L-1}P_F\left(g\right)\log_2 P_F\left(g\right)\right| \tag{10}$$

Mutual information:

$$MI = \sum_{i=1}^{L} \sum_{j=1}^{L} h_{IF}(i,j) \log_2 \frac{h_{IF}(i,j)}{h_I(i)\, h_F(j)} \tag{11}$$

The notation for these metrics is given in Table 4. Here, the reference image is the X-ray thickness map. The fused image is the final numerical result obtained from the locally weighted regression. These metrics compare two images from a different perspective. A larger value of CORR, PSNR, and MI indicates a better match between the estimation and the reference X-ray map while a smaller value of RMSE and DE indicates a better match. In our case, four of the five metrics show a better result from the fusion of ET $17kHz$ and $30kHz$ images with P-ET LOI image as shown in Table 3.

Table 4. The notation for fused image evaluation metrics.

$I(m,n)$	reference image
$F(m,n)$	fused image
L	maximum pixel value
$P_I(g)$	probability of value g of the reference image
$P_F(g)$	probability of value g of the fused image
$h_{IF}(i,j)$	normalized joint histogram of the reference image and he fused image
$h_I(i)$	normalized marginal histogram of the reference image
$h_F(j)$	normalized marginal histogram of the fused image

4. Discussion

As mentioned in earlier in Table 1, the classification range chosen for thickness is around three thou. A smaller value does not necessarily indicate a more precise result at the classification stage. The choice of an appropriate range is more important as it can accommodate relatively more points for the locally weighted regression. If a large number of "information classes" are defined, the classification errors will be introduced to the regression analysis.

The experiments described in previous section used the whole image to train a classifier. By doing so, more pixels from corrosion-free regions than from corroded regions were involved in the training process. Even in the corrosion-free regions, the pixel values varied due to a number of reasons. Such variance did not improve the classifier training. If only pixels above a threshold are used in the classifier training, the performance could be further improved.

The quality of the training data sets will eventually affect the final estimation results. The choice of representative data for training still remains a challenge. When more NDT data are available in a later time, the performance of the classifier could be improved through introducing new data sets in the training. A reinforced learning algorithm should be implemented to achieve this goal and this remains a topic for our future research.

A fused image can be evaluated by comparing it with a reference image. A number of image comparison metrics are available for this purpose.[15] Usually the comparison is done for the whole image. In this application, the analysis can also focus on the corroded region only. The data from the corrosion-free region will not be used for the training and estimation. Also, NDI reliability assessment can be carried out by calculating probability of detection (POD) using estimated thickness value ($\hat{a}$) and actual X-ray thickness value (a) as the inputs.

It should also be noted that the capability of the P-ET inspection has not been fully explored in this study as only the LOI feature is employed in the experiment. As described in section 2, a mechanism to identify the location of the corrosion on the faying surfaces should be introduced to fully characterize the hidden corrosion of aircraft lap joint structures.

5. Conclusion

In this chapter, an image fusion framework for interpreting NDI images based on Dempster-Shafer theory is presented. The procedure consists of three major steps, i.e. classification, fusion, and regression. In the proposed method, the basic probability assignment is implemented by considering the relation between the "data classes" and the "information classes" as well as the capability of each NDI technique. The use of DS theory to fuse the inspection results of aircraft lap joints with conventional eddy current and pulsed eddy current is demonstrated. A numerical estimation is obtained with the locally weighted regression of the DS-fusion outputs.

Acknowledgements

The work was supported by the Defence Research and Development Canada, Air Vehicles Research Section and National Research Council Canada. Dr. S. M. Safizadeh is acknowledged for helpful discussion.

References

1. R. S. Blum and Z. Liu, Eds., *Multi-Sensor Image Fusion and Its Applications*. Signal Processing and Communications, (CRC Press, Taylor & Francis, 2005).
2. X. E. Gros, *NDT Data Fusion*. (Arnold, 1997).
3. X. E. Gros, *Applications of NDT Data Fusion*. (Kluwer Academic Publishers, 2000).
4. M. Mina, J. Yim, S. S. Udpa, and L. Udpa. Two dimensional multi-frequency eddy current data fusion. In *Review of Progress in QNDE*, vol. 15, pp. 2125–2132, (1996).
5. J. Yim, S. S. Udpa, L. Udpa, L. Udpa, and W. Lord. Neural network approaches to data fusion. In eds. D. O. Thompson and D. E. Plenum, *Review of Progress in QNDE*, vol. 14, pp. 819–826, (1995).
6. G. Simone and F. C. Morabito, NDT image fusion using eddy current and ultrasonic data, *The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*. **20**(3), 857–868, (2001).

7. P. Ramuhalli and Z. Liu. Wavelet neural network based data fusion for improved thickness characterization. In eds. D. O. Thompson and D. E. Chimenti, *Review of Progress in Quantitative NDE*, vol. 23, (2003).
8. Z. Liu, K. Tsukada, K. Hanasaki, and M. Kurisu, Two-dimensional eddy current signal enhancement via multifrequency data fusion, *Research in Nondestructive Evaluation.* **11**, 165–177, (1999).
9. X. E. Gros, Z. Liu, K. Tsukada, and K. Hanasaki, Experimenting with pixel level ndt data fusion techniques, *IEEE Transactions on Instrumentation and Measurement.* **49** (5), 1083–1090 (October, 2000).
10. O. Dupuis, V. Kaftandjian, D. Babot, and Y. Zhu, Automatic detection and characterization of weld defects: Determination of confide3nce levels for data fusion of radioscopic and ultrasonic images, *Insight.* **41**(3), 170–172 (March, 1999).
11. D. L. Hall, *Mathematical Techniques in Multisensor Data Fusion.* (Artech House, Norwood, MA, USA, 1992).
12. A. Fahr, D. S. Forsyth, and C. E. Chapman. Suvery of nondestructive evaluation (NDE) techniques for corrosion inaging aircraft. Technical Report, LTR-ST-2238, National Research Council Canada (October, 1999).
13. D. S. Forsyth and J. P. Komorowski, *Applications of NDT Data Fusion*, chapter NDT Data Fusion for Improved Corrosion Detection, pp. 205–225. Kluwer Academic Publishers, (2001).
14. Z. Liu, D. S. Forsyth, M. S. Safizadeh, M. Genest, A. Fahr, and A. Marincak. Fusion of visual and eddy current inspection results for the evaluation of corrosion damage in aircraft lap joints. In *SPIE Proceedings 5768*, San Diego, CA, USA (March 6-11, 2005).
15. Z. Liu, D. S. Forsyth, M. S. Safizadeh, M. Genest, and C. M. A. Fahr. The use of locally weighted regression for the data fusion with dempster-shafertheory. In ed. X. P. V. Maldague, *V International Workshop on Advances on Signal Processing for NondestructiveEvaluation of Materials*, vol. 7 (August, 2005).
16. B. A. Lepine, J. S. R. Giguere, D. S. Forsyth, and J. M. S. D. A. Chahbaz. Applying pulsed eddy current NDI to the aircraft hidden corrosion problem. In *Proceedings of the 5th NASA/FAA/DoD Conference on Aging Aircraft*, Orlando, Florida (September, 2001).
17. M. S. Safizadeh, Z. Liu, D. S. Forsyth, and A. Fahr. Automatic classification and characterization of hidden corrosion usingpulsed eddy current data. In *Proceedings of 16th World Conference on NDT*, Montreal, Canada (August, 2004).
18. M. S. Safizadeh, Z. Liu, C. Mandache, D. S. Forsyth, and A. Fahr. Intelligent pulsed eddy current method for detection and classificationof hidden corrosion. In *6th International Workshop on Adavances in Signal Processing for NDE andMaterials*, Quebec, Canada (August, 2005).
19. Z. Liu, D. S. Forsyth, B. A. Lepine, S. Safizadeh, and A. Fahr. Quantifying aircraft hidden corrosion by using multi-modal NDI. In eds. D. Thompson and D. Chimenti, *Review of Progress in Quantitative NDE*, pp. 1355–1362, Green Bay, Wisconsin, (2003). American Institute of Physics.
20. S. J. Hurlbut, J. M. S. Dubois, and V. Lefebvre. Quantification of superposed defects in multi-layered metal structures. Montreal (April, 2003). Canadian Aeroautics and Space Institute 50th AGM Conference.
21. C. G. Atkeson, A. W. Moore, and S. Schaal, Locally weighted learning, *Artificial Intelligence Review.* **11**, 11–73, (1997).
22. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification.* (John Wiley and Sons Inc., 2002), second edition.

23. X. L. Xie and G. A. Beni, Validity measure for fuzzy clustering, *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **13**(8), 841–847, (1998).
24. A. M. Bensaid, L. O. Hall, J. C. Bezdek, L. P. Clarke, M. L. Silbiger, J. A. Arrington, and R. F. Murtagh, Validity-guided (re)clustering with apllications to image segmentation, *IEEE Transactions on Fuzzy Systems*. **4**, 112–123, (1996).
25. D. S. Forsyth, Z. Liu, J. P. Komorowski, and D. Peeler. An application of NDI data fusion to aging aircraft structures. In *6th Joint FAA/DoD/NASA Conference on Aging Aircraft*, San Francisco, CA, USA. (September, 2002).
26. I. Bloch, Information combination operatiors for data fusion: A comparative review with classification, *IEEE Transactions on Systems, Man, and Cybernetics*. **26**(1), 52–67, (1996).
27. M. Birattari, G. Bontempi, and H. Bersini. Lazy learning meets the recursive least squares algorithm. In *Proceedings of the 1998 Conference on Advances in Neural Information ProcessingSystems II*, pp. 375–381. MIT Press, (1999).
28. B. A. Lepine, J. S. R. Giguere, D. S. Forsyth, J. M. S. Dubois, and A. Chahbaz. Interpretation of pulsed eddy current signals for locating and quantifying metal loss in skin lap splice. In eds. D. Thompson and D. Chimenti, *Review of Progress in Quantitative NDE*, vol. 21. American Institute of Physics, (2001).
29. J. H. V. Lefebvre and J. M. S. Dubois. Lift-off point of intercept (LOI) behavior. In eds. D. Thompson and D. Chimenti, *Review of Progress in Quantitative NDE*, vol. 24. American Institute of Physics, (2004).

CHAPTER 18

SMART STRUCTURES APPROACH: ACHIEVING GLOBAL STRUCTURAL HEALTH MONITORING THROUGH THE APPLICATION OF DISTRIBUTED, EMBEDDED MICRO-SENSOR SYSTEMS

Dennis Roach

Sandia National Laboratories
Albuquerque, NM 87185

Multi-site fatigue damage, hidden cracks in hard-to-reach locations, disbonded joints, erosion, impact, and corrosion are among the major flaws encountered in today's extensive array of aerospace vehicles and civil structures. These damage scenarios, coupled with new and unexpected phenomena, have placed greater demands on the application of advanced nondestructive inspection (NDI) and health monitoring techniques. Reliable, structural health monitoring systems can automatically process real-time data, assess structural condition, and signal the need for human intervention. Prevention of unexpected flaw growth and structural failure could be improved if on-board health monitoring systems are used to continuously assess structural integrity. Such systems would be able to detect incipient damage before catastrophic failures occurs. Condition-based maintenance practices could be substituted for the current time-based maintenance approach. Other advantages of on-board distributed sensor systems are that they can eliminate costly, and potentially damaging, disassembly, improve sensitivity by producing optimum placement of sensors with minimized human factors concerns in deployment, and decrease maintenance costs by eliminating more time-consuming manual inspections. This chapter focuses on developments in mountable sensors and how they can be integrated into such a Structural Health Monitoring (SHM) system to guide condition-based maintenance activities.

Introduction

Health monitoring of structures is a major concern of the engineering community. This need is even more intense in the case of aging aerospace and civil structures many of which are operating well beyond their initial design

lives. The current damage tolerance design philosophy requires that a structure be capable of sustaining small damage without failure, and that an inspection program be instituted to detect such flaws before they grow to a critical size. This damage tolerance approach recognizes the impossibility of establishing complete structural redundancy – the fail safe design premise – and places greater emphasis on inspection to ensure safety and reliability. Multi-site fatigue damage, hidden cracks in hard-to-reach locations, disbonded joints, erosion, and corrosion are among the major flaws encountered in today's extensive fleet of aging aircraft, bridges, buildings, and civil and space transport vehicles. Examples of some common flaws found in aircraft structure are pictured in Figure 1.

The costs associated with the increasing maintenance and surveillance needs of our aging infrastructure are rising at an unexpected rate. Aircraft maintenance and repairs represent about a quarter of a commercial fleet's operating costs. The application of distributed sensor systems may reduce these costs by allowing condition-based maintenance practices to be substituted for the current time-based maintenance approach. In the near future, it may be possible to quickly, routinely, and remotely monitor the integrity of a structure in service. A series of expected maintenance functions will already be defined, however, they will only be carried out as their need is established by the health monitoring system [1-3].

The replacement of our present-day manual inspections with automatic health monitoring would substantially reduce the associated life-cycle costs. Motivated by these pressing needs, considerable research efforts are currently being directed towards: (1) development of new and better NDI techniques; and (2) development of health monitoring sensors and systems. Whether the sensor network is hardwired to an accessible location within the aircraft or monitored in a remote, wireless fashion, the sensors can be interrogated in a real-time mode. The important item to note is that the ease of monitoring an entire network of distributed sensors means that structural health assessments can occur more often, allowing operators to be even more vigilant with respect to flaw onset. In addition, innovative deployment methods can be employed to overcome a myriad of inspection impediments stemming from accessibility limitations, complex geometries, and the location and depth of hidden damage. Figure 2 depicts a sensor network deployed on an aircraft to monitor critical sites over the entire structure.

Figure 1. Sample Disbond, Crack and Corrosion Damage in Aircraft Structure.

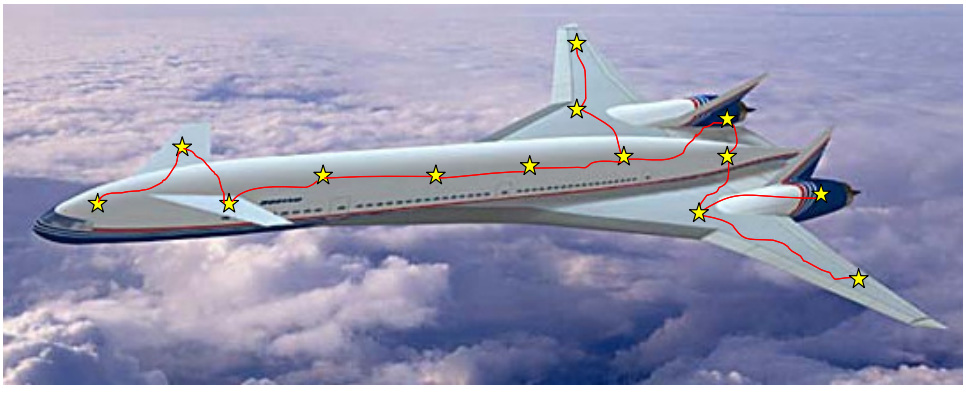


Figure 2. Depiction of Distributed Network of Sensors to Monitor Structural Health.

Comparative Vacuum Monitoring

Comparative Vacuum Monitoring (CVM) has been developed on the principle that a small volume maintained at a low vacuum is extremely sensitive to any ingress of air [4]. Figure 3 shows top-view and side-view schematics of the self-adhesive, elastomeric sensors with fine channels on the adhesive face along with a sensor being tested in a lap joint panel. When the sensors are adhered to the structure under test, the fine channels and the structure itself form a manifold of galleries alternately at low vacuum and atmospheric pressure. When a crack develops, it forms a leakage path between the atmospheric and vacuum galleries, producing a measurable change in the vacuum level. This change is detected by the CVM monitoring system shown in Figure 4. This material is able to withstand the high loading stresses that result during the riveting process.

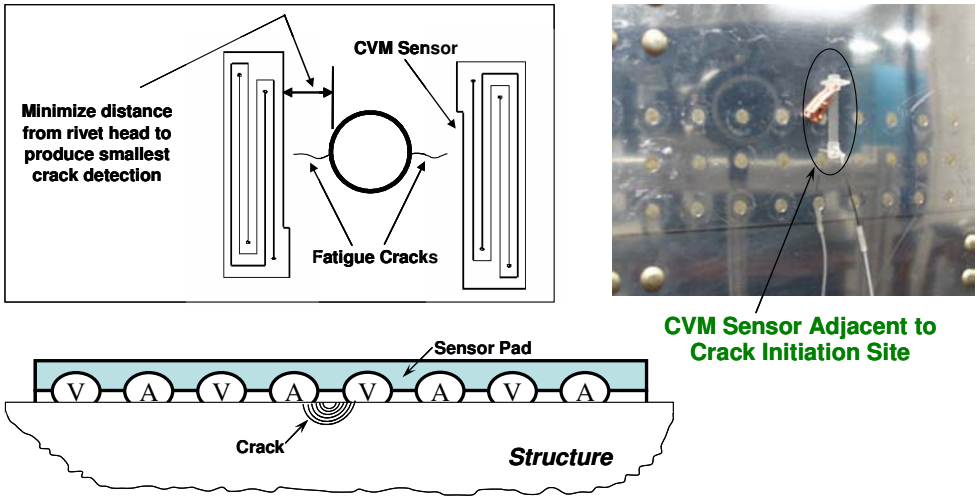


Figure 3. Schematics Depicting Operation of CVM Sensor and Polymer Sensor Mounted on Outer Surface of Riveted Lap Joint.

These sensors can be attached to a structure in areas where crack growth is known to occur. On an established engineering interval, a reading will be taken from an easily accessible area. Each time a reading is taken, the system performs a self-test. This inherent fail-safe property ensures the sensor is attached to the structure and working properly. Figure 4 also shows sample CVM sensors mounted on aircraft structure as part of a field performance validation effort. Fatigue tests were conducted on simulated aircraft panels to grow cracks in riveted specimens (see Fig. 3) while the vacuum pressure within the various sensor galleries were simultaneously recorded. The fatigue cracks were propagated until they engaged, and fractured, one of the vacuum galleries such that crack detection was achieved (sensor indicates the presence of a crack by its inability to maintain a vacuum). The test program produced a statistically-relevant set of crack detection levels for a range of materials at different thicknesses. The calculated probability of detection (POD) values for 90% POD and 95% confidence were achieved in 2024-T3 material at crack lengths of: a) 0.021” for 0.040” thick plate, 0.042” for 0.071” thick plate, and 0.090” for 0.100” thick plate. This indicates that the CVM sensors will reliably detect these cracks when they reach the lengths indicated.

Figure 5 shows a photo of a fatigue crack as it engages the first vacuum gallery of a CVM sensor. The pressure rise, corresponding to a rupture in the gallery and a leakage path to atmospheric pressure, is shown on the right side of Figure 5. The large increase in the pressure corresponds to crack detection. One

signal (blue curve) corresponds to vacuum levels produced when there is no crack indication and the other signal (red curve) occurs when a vacuum is not achievable. This latter signal is produced when the CVM detects a crack.

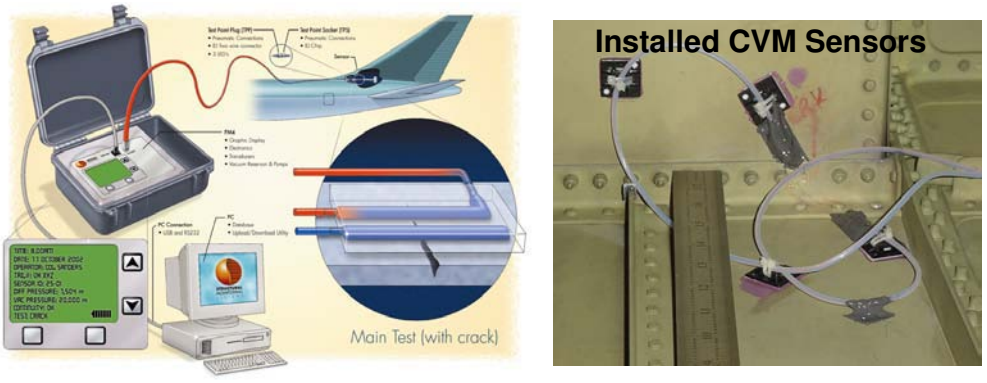


Figure 4. Crack Detection Via CVM System and Aircraft Test Installations of Sensors.

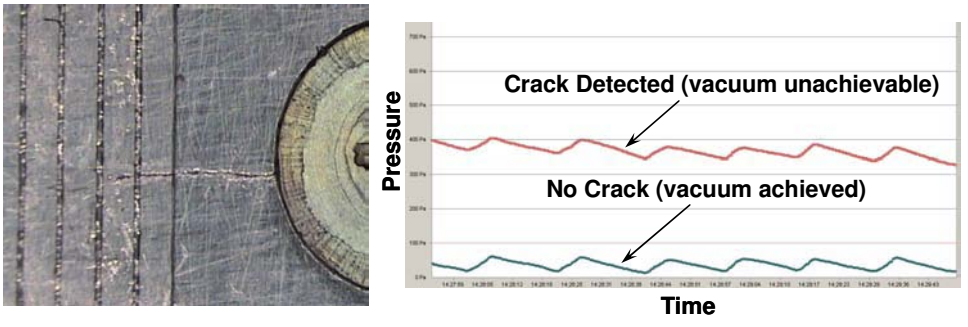


Figure 5. Fatigue Crack Crossing Into CVM Galleries; Typical Pressure vs. Time Plots from CVM Sensors Indicating 'No Crack' and 'Crack Detection'.

Piezoelectric Transducers (PZT)

Prime candidates for sensors based on active-material principles utilize thin piezoelectric wafers of 0.125"-0.25" diameter with thicknesses of 0.010" - 0.030". They can be easily attached to existing aging structures without changing the local and global structural dynamics [5-7]. PZT sensors can also be embedded inside composite structures to closely monitor for internal flaws. These sensors can act as both transmitters and receptors. As transmitters, piezoelectric sensors generate elastic waves in the surrounding material. As receptors, they receive elastic waves and transform them into electric signals. It

is conceivable to imagine arrays of active-sensors, in which each element would take, in turn, the role of transmitter and receptor, and thus scan large structural areas using ultrasonic waves. The structural interrogation strategies using active piezoelectric sensors are two fold:

(a) For local area detection, the electro-mechanical (E/M) impedance method is applied to detect changes in the point wise structural impedance resulting from the presence and propagation of structural damage.
(b) For large area detection, wave propagation techniques using Lamb and Love waves methods are used to identify zones in the monitored area that have undergone changes in their structural integrity.

In the high-frequency E/M impedance approach, pattern recognition methods are used to compare impedance signatures taken at various time intervals and to identify damage presence and progression from the change in these signatures. In the Lamb/Love waves approach, the acousto-ultrasonic methods identifying changes in transmission velocity, phase, and additional reflections generated from the damage site are used. Both approaches can benefit from the addition of artificial intelligence neural network algorithms that can extract damage features based on a learning process.

Validation Testing of Piezoelectric Sensors in Electro-Mechanical Impedance Mode - In order to evaluate this health monitoring approach, test specimens representative of aircraft lap joints were produced. The specimens contained realistic, engineered crack and corrosion flaws. Three specimens were constructed: (1) pristine; (2) with cracks only; (3) with cracks and corrosion. The specimens were instrumented with several piezoelectric wafer active sensors, 0.25” square and 0.0075” thick. Figure 6 shows a set of 4 such active sensors equidistantly placed in a row at right angles to a 0.75” simulated crack (notch). An impedance analyzer for the E/M impedance testing, and wave generator, digital oscilloscope, pulser-receiver and MHz range A/D boards for wave propagation testing were used in the experiments. Initial E/M impedance measurements were made to detect cracks at rivet locations. Figure 6 shows the superposed results of E/M impedance testing performed on pristine (PZT-Ref) and cracked (PZT) specimens. The sensors further away from the crack registered less change in the E/M impedance spectrum than sensors closer to the crack. The sensor closest to the crack showed the largest change in the E/M impedance. The experimental results in Figure 6 clearly show the increase in impedance produced by the structural flaws. Additional testing is underway to study strategic sensor placement, to quantify piezoelectric probe sensitivity, and to optimize data analysis algorithms.

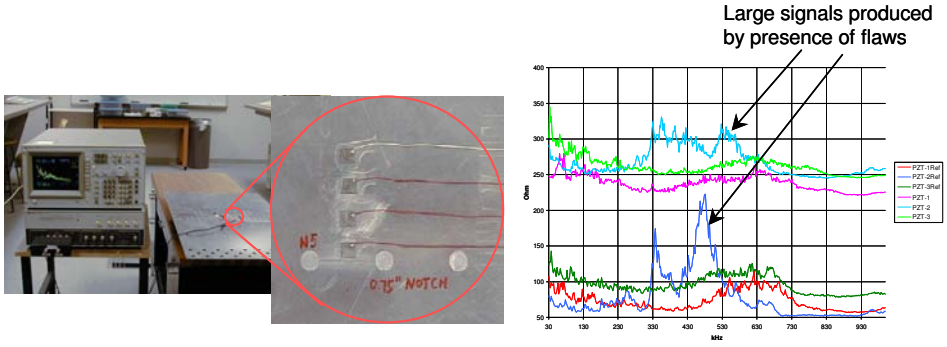


Figure 6. Piezoelectric Sensors on an Aircraft Panel Containing Crack and Corrosion Damage (left). Results Obtained with the E/M Impedance Technique Using Active Sensors (right).

Damage Identification through Electro-Mechanical Impedance Method - Consider an array of 4 active sensors as presented in Figure 7. Since piezoelectric active sensors can act as both sensors and actuators, the strategy assumes that one active sensor acts as actuator (1) while the others act as sensors (2, 3, and 4). Active sensor 1 generates elastic waves that propagate through the material and are sensed at active sensors 2, 3, and 4. The properties of these waves are affected by the presence of damage, and can be interpreted to yield damage location and amplitude. In the coupling between the sensor and the structure, the electrical admittance of the piezoelectric materials can be directly related to the structure's mechanical impedance. If the mechanical impedance of a structure varies, as a result of damage for example, the electrical admittance will be affected as well. The electro-mechanical impedance method is applied by scanning a predetermined frequency range and recording the complex impedance spectrum. Signal processing methods include pattern recognition methods that compare impedance signatures taken at various time intervals to identify the presence of damage or to monitor its growth. To maximize the amount of data and mitigate experimental error, a round-robin process is applied, whereby active sensors 2, 3, and 4 take turns acting as the wave generators with the rest of the active sensors being wave receptors. The signals are subsequently transmitted to the diagnostic hardware for comparison with previously recorded baseline test signals. The difference between current and baseline signals is used to identify the location and extent of damage or other structural anomalies. This method can be applied to detect the damage such as cracks, disbonds, and corrosion. In addition, the network of sensors coupled with their interactive transmission of signal allows for wider area coverage and more global flaw detection.

Mountable PZT Networks and Lamb Wave Interrogation Methods - Another structural health monitoring approach was evaluated using a built-in network of piezoelectric transducers embedded in a think dielectric carrier film. The SHM system included the PZT network connected to portable, diagnostic hardware and software developed by Acellent Technologies, Inc. The system performs in-situ monitoring, data collection, signal processing, and real-time data interpretation to produce a two-dimensional image of the structure being interrogated. The Acellent software instructs the actuators to generate pre-selected diagnostic signals and transmit them to neighboring sensors. Multiple diagnostic wave types can be generated including 3-peak, 5-peak, and 10-peak narrow band frequency waveforms, chirp, random, and user defined excitations. The software links each sensor with its neighbors to form a web, or network, covering the structure. The system then collects the total set of responses from each of the sensor sets as each PZT takes its turn as the actuator. Changes in the Lamb waves generated within the structure are used in concert with triangulation methods to detect the presence of structural anomalies and to determine the size and location of the flaws.

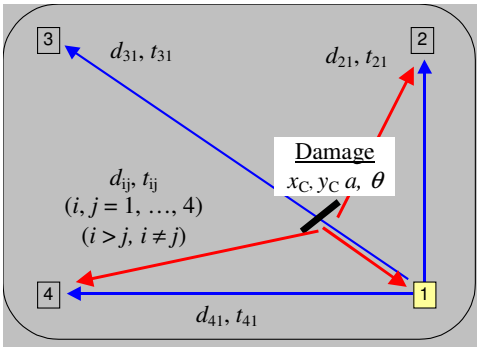


Figure 7. Damage Detection Strategy Using a Network of Piezoelectric Sensors and Wave Propagation Techniques for All Paths Between Each Actuator and Receiver Set.

Damage Identification through Elastic Wave Propagation - The wave propagation approach uses the pitch-catch method for detecting damage in a structure. Acousto-ultrasonic methods are used to identify changes in wave transmission. Figure 8 shows some of the wave motion from sensors (1) and (9) when they are used as the source of excitation for the structure. The mechanical vibration is introduced into the structure by the PZT element and travels by wave motion through the test piece at the velocity of sound, which depends on the material. If the pulses encounter a reflecting surface, some or all of the energy is

reflected and monitored by adjacent PZT sensors in the network. The reflected beam, or echo, can be created by any normal (e.g. in multi-layered structures) or abnormal (flaw) interface. Figure 8 highlights the interaction of the UT waves with a flaw within the structure. The degree of reflection depends largely on the physical state of the materials forming the interface. Cracks, delaminations, shrinkage cavities, pores, disbonds, and other discontinuities that produce reflective interfaces can be detected. Complete reflection, partial reflection, scattering, or other detectable effects on the ultrasonic waves can be used as the basis for flaw detection.

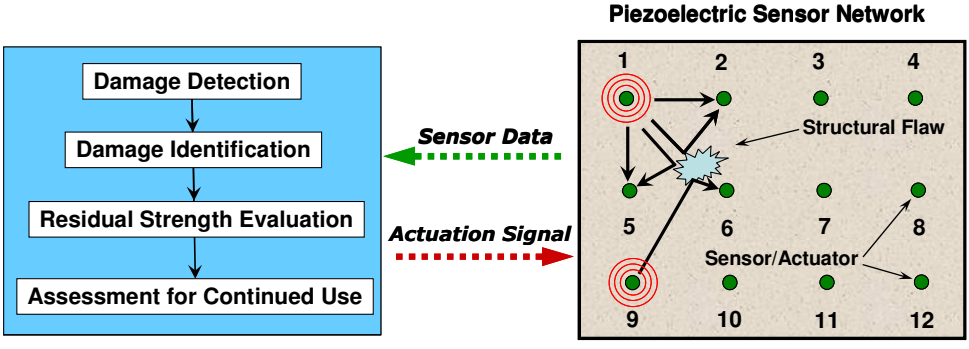


Figure 8. Flaw Detection Using the Wave Propagation Method.

Validation Testing - A network of PZT sensors was deployed to assess bonded joints and crack growth in a composite doubler repair installation. Figure 9 shows a schematic and photos of the Boron-Epoxy laminate repair on a metal parent structure along with the set of PZTs distributed over the structure to be monitored. Note that the network of sensors/actuators is embedded in a custom polyamide film to allow for accurate placement of the network and eliminating the need for each sensor to be installed individually. The test specimen, containing engineered disbonds and a central crack, was subjected to constant-amplitude fatigue loads with maximum stresses in excess of 80% of yield.

Similar to conventional ultrasonic testing, the PZT data analysis can include one or more of the following measurements: time of wave transit (or delay), path length, frequency, phase angle, amplitude, and angle of wave deflection (reflection and refraction). In this test series, the pitch-catch method studied the transmission of sound waves as they traveled from each actuator to all other receiving sensors. The sum total of received beams were then analyzed to define the presence and location of flaws. In order to optimize flaw detection, a series of excitation frequencies were used: 50 KHz, 200 KHz, 350 KHz, and 500 KHz. Overall test results revealed that disbond flaws were most strongly detected with

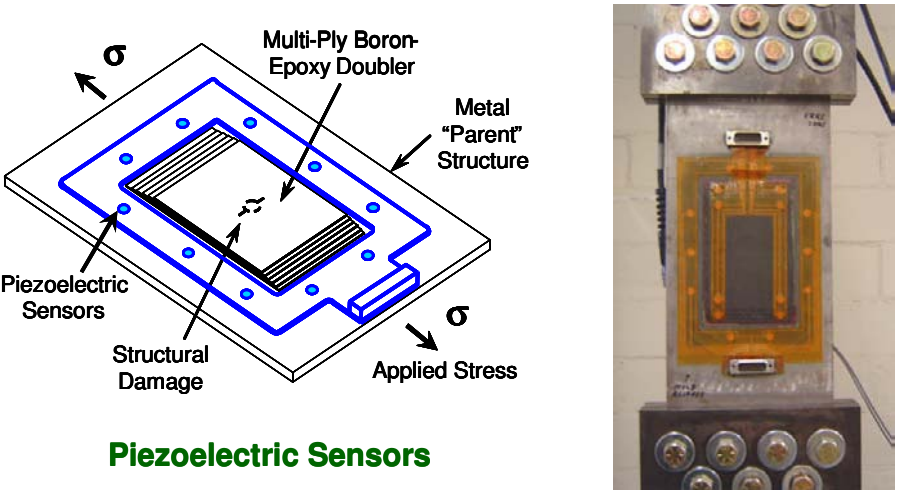


Figure 9. Set of Piezoelectric Sensors Used to Monitor Crack Growth and Disbonds in a Composite Doubler Bonded to a Metal Plate.

the lower, 50 KHz excitation while the crack growth was monitored best with the highest, 500 KHz excitation. Figure 10 shows raw PZT response data produced during the Lamb wave interrogation method. Signal attenuations, corresponding to disbonds between the laminate and parent skin, are apparent. When all of the signals are analyzed with the Acellent imaging software and flaw locations are determined by using the time base and triangulation methods, a two dimensional image of the disbond flaws was produced. Figure 11 shows the engineered disbonds in the test specimen along with the image produced by the PZT sensor network. Note that both disbond flaws were clearly imaged even though one is a weak bond produced by a mold release agent and one is a complete disbond produced by a Teflon insert.

Crack detection and growth was monitored using the same approach. PZT data was acquired at discrete intervals during the crack growth process. In addition, eddy current and microscopic inspections were conducted to measure the crack lengths at each cycle count. Figure 10 also shows PZT response signals before and after crack growth occurred into the sensor path. A set of images produced by the PZT network are shown in Figure 12. The crack growth (two fatigue cracks emanating from a central hole) can be clearly seen. The PZT crack growth data was analyzed further to produce crack length predictions. The Acellent software contains an algorithm that allows for system learning. After inputting several crack lengths to match with the PZT data at discrete fatigue

intervals, it was possible for the system to predict all subsequent crack lengths using the PZT data alone. Table 1 compares the crack lengths predicted by the PZT sensor network with the crack lengths determined from eddy current and microscopic measurements. The PZT predictions were all within 5% of the actual crack lengths for data taken at max load (34 kips) and, for the most part, within 10% of actual values for PZT data taken in the unloaded condition.

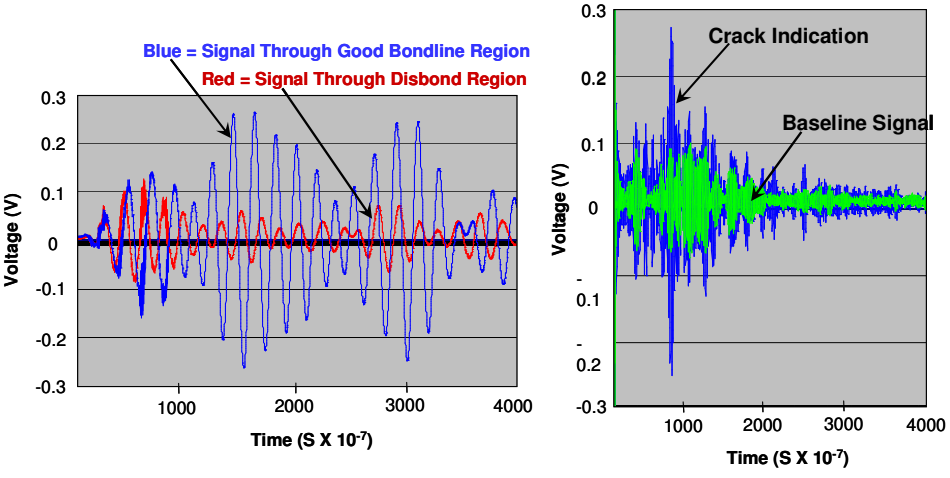


Figure 10. Sample Signals Observed by PZTs During 50 KHz Lamb Wave Interrogation Showing the Attenuation Corresponding to Disbonds in the Structure (left) and PZT Signals Showing the Indication of a Fatigue Crack with a 500KHz Excitation (right).

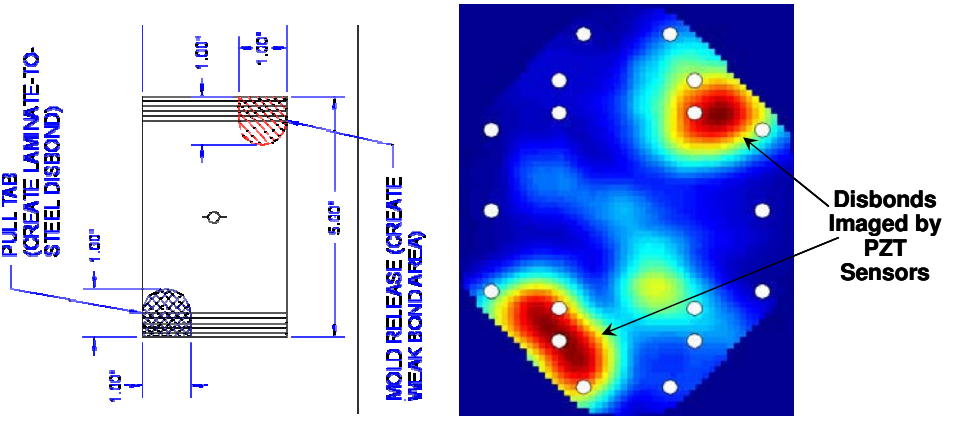


Figure 11. Color-Coded Image of Disbond Flaws Produced by the PZT Sensor Network.

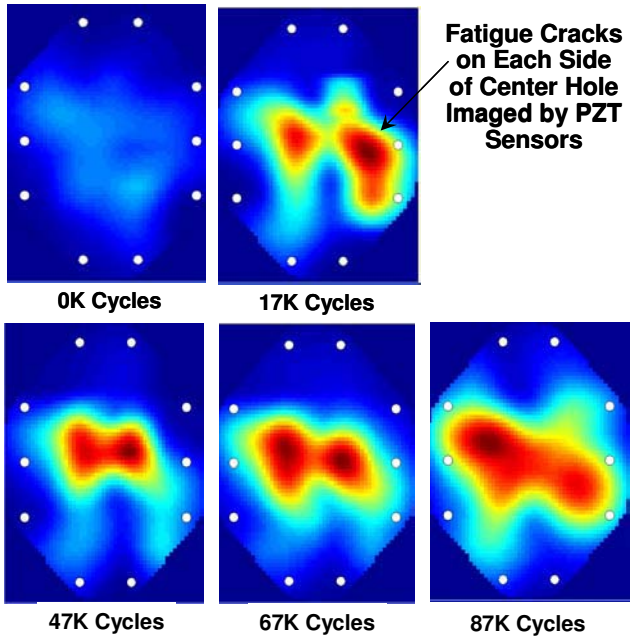


Figure 12. Color-Coded PZT Images Showing Crack Growth.

Fiber Optics

Rapid growth in the optoelectronics and telecommunications industries has resulted in the evolution of highly-sensitive fiber optic sensors. Fiber optic (FO) sensors have been developed for a wide variety of applications including the measurement of rotation, acceleration, vibration, strain, temperature, pressure, electric and magnetic fields, moisture, and humidity. Fiber optic sensors are light weight, low profile (typically 145 microns in diameter with a polyimide coating), are corrosion resistant, and multiple sensors are easily multiplexed into a single fiber. These factors make them ideal for embedding in or surface mounting on composite and metallic structures without affecting structural performance. Other advantages of fiber optic sensors include their high sensitivity, wide bandwidth, EMI resistance, low power requirements, and environmental ruggedness. Fiber optic sensors can be configured to monitor crack growth and corrosion in civil and aerospace structures [8-10]. Omni-directional fiber optic sensors can be used to measure large strains (up to 150% strain), small displacements (10 μm range), and crack growth in any material. The introduction of low-cost laser diodes, the alternative use of LEDs as light sources, and the development of inexpensive, single mode optical fibers have greatly reduced the costs associated with deploying FO sensors.

Table 1. Comparison of Crack Lengths Predicted by PZT Sensors with Actual Crack Lengths Measured Using Eddy Current and Microscopic Methods.

Composite Doubler with PZT Health Monitoring			
Fatigue Cycles	**Measured Total Crack Length**	**Estimated Crack Length from PZT Sensor Data (0 lbs. load)**	**Estimated Crack Length from PZT Sensor Data (34 kips load)**
Specimen 1 - Unflawed Composite Doubler			
0	0.00		
26,218	0.32	PZT Learning Data	PZT Learning Data
47,000	0.70	PZT Learning Data	PZT Learning Data
67,000	1.50	1.274	1.385
87,000	2.44	1.956	2.367
Specimen 2 - Composite Doubler with Disbond Flaws			
0	0.00		
19252	0.16	PZT Learning Data	PZT Learning Data
29274	0.32	PZT Learning Data	PZT Learning Data
38064	0.48	PZT Learning Data	PZT Learning Data
51576	0.80	PZT Learning Data	PZT Learning Data
60438	1.08	0.981	1.099
66439	1.34	1.35	1.349
76444	1.76	1.567	1.762
82446	2.02	1.909	2.08

Information about the environment to which a fiber optic is exposed can be inferred by analyzing the guided light transmitted through the optical filaments. In this approach the entire length of the fiber acts as a continuous sensor. The fiber can be mounted in a serpentine path or fiber optic tentacles can be created to provide full coverage over a large area of concern. The presence of a crack can be determined by monitoring changes in the magnitude and phase of the returned light in the fiber. Optical time domain reflectometry can then be applied to determine the exact location of the crack. Fiber optic sensors can also aid in the detection of structural corrosion. Sacrificial corrosion sensors can be mounted on the end face of the fiber optic. These sensors are designed to corrode at the same rate as the parent material to which they are mounted. Degradation in these sensors, and thus the parent structure, can be calibrated to changes in the intensity of the reflected light within the fiber optics.

Optical Fibers as Uniaxial Sensors - Fiber optics work on the principle that light can be guided by an interface between materials of different indices of refraction. The components of a fiber optic line consist of a core with a higher index of refraction surrounded by a cladding with a lower index of refraction. The fiber is made mostly of silicone dioxide (SiO_2), with the core containing germanium dopants. The dopants in the core create a differential in the index of

refraction between the core and the cladding. These differences in the index of refraction in the core and the cladding cause light to be guided through the core. To better protect the fiber, a Polyimide coating around the cladding is used. Fiber optic sensors (core and cladding) are 80 μm – 120 μm in diameter which is about the size of a human hair. With the protective coating added, the fiber is approximately 125 - 150 μm in diameter.

If a short wavelength light is used to illuminate the core region of the fiber, the core material is rearranged to produce an increase in the index of refraction in that region. A quartz phase mask can be used, as in Figure 13, to image an interference pattern through the side of the optical fiber. This results in a localized periodic index of refraction modulation along the length of the optical fiber (sensor grating). A typical sensor grating consists of thousands of grating lines with approximately 0.5 micron spacing. The mask shown in Figure 13 controls the laser exposure to induce a customized interference pattern on the optical core. This fiber grating can then be used to measure strain since axial stretching or compression of the fiber will change the grating period and, thus, the peak wavelength reflected back from the sensor. When the fiber grating is illuminated by a broadband light source or a tunable laser source the fiber grating will reflect a narrow band of light corresponding to its period. The reflected light will be centered in a narrow band at the Bragg wavelength according to the relationship:

$$\lambda_B = 2n_e\Lambda \tag{1}$$

where n_e is the effective index of refraction at the core of the fiber and Λ is the grating line spacing. Measuring the shift in wavelength of light reflected back from the grating can thus be used to determine strain or temperature changes. If strain is applied to the Fiber Bragg Grating, the resonant reflected wavelength λ_B will shift by an amount $\Delta\lambda_B$ given by:

$$\Delta\lambda_B / \lambda_B = (1-P_\varepsilon)\varepsilon \tag{2}$$

where P_ε is the photoelestic constant for the fiber core. By tracking $\Delta\lambda_B$ it is possible to accurately measure strain levels in the structure.

Multi-Parameter Fiber Grating Strain Sensors – A method of generating multiple fiber gratings capable of measuring multi-directional strain is to write a fiber grating into a special type of fiber called polarization maintaining or birefringent fiber [10-11]. The birefringence (difference in index in refraction depending on the polarization state of light) in the fiber is created with a built in

residual stress introduced during the fiber draw (manufacture). This stress results in a slight change in the index of refraction along two mutually orthogonal directions (termed the polarization axes). This creates two spectral peaks for each optical grating written into the optical fiber, one associated with each polarization axis. By measuring the wavelength of these peaks it is possible to generate two equations in two unknowns effectively allowing two environmental parameters (in this case, two strain tensors) to be monitored.

For this type of fiber grating strain sensor, a single fiber grating results in two distinct spectral peaks. These peaks correspond to each of the polarization axes of the polarization preserving fiber, which differ slightly in index of refraction. The two spectral reflection peaks, corresponding to the effective fiber gratings along each birefringent (polarization) axis, will move apart or together uniformly providing a means to measure transverse strain. Figure 14 shows how changes in the spectral output from birefringent fiber gratings can be used to measure both axial and transverse strain. When the fiber is loaded transversely, the relative index of refraction of the polarization axes of the fiber changes and the net result is that the difference in wavelength between the spectral peaks changes as well (peak-to-peak separation changes). When the fiber is strained axially, the fiber elongates or compresses, changing the fiber grating spectral period. As a result the output spectrum shifts to longer or shorter wavelengths, respectively. Private companies, such as Blue Road Research have developed techniques to accurately align the polarization axes in these fibers to produce dual axis fiber grating sensors.

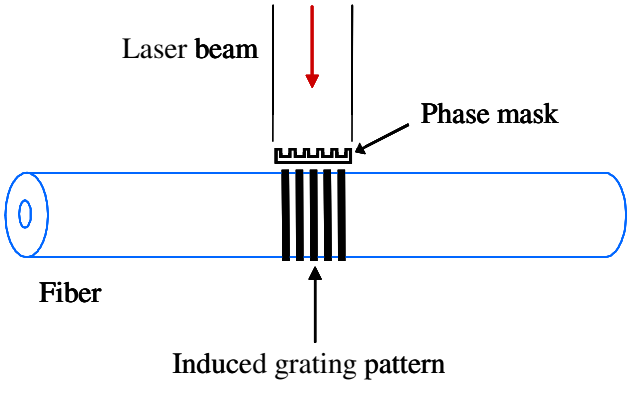


Figure 13. Process to Install a Fiber Grating onto an Optical Fiber Using a Short Wavelength Laser in Combination with a Phase Mask.

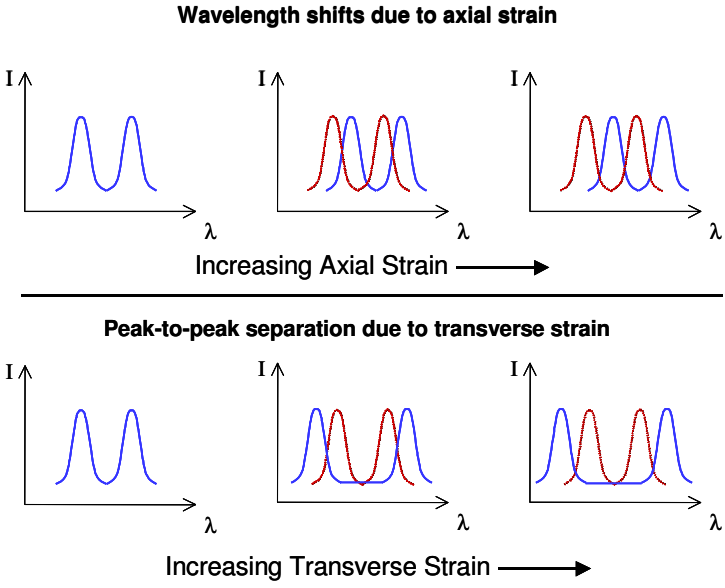


Figure 14. Change in Spectral Reflection of a Multi-Axis Fiber Grating Strain Sensor When Subjected to Axial and Transverse Loads.

Measuring the Complex Grating Reflection Profile with a Scanning Laser Interferometer – Traditionally, fiber grating strain measurements are made with high reflectivity fiber gratings, which allow detailed information about the spectral peaks to be accessed directly via illuminating the fiber gratings with broadband light sources and then recording the reflected grating spectrum with an optical spectrum analyzer. By peak-fitting the spectrum and applying wavelength-to-strain conversion coefficients it is possible to obtain strain values. Multiplexing large numbers of high reflectivity fiber gratings can be limited by shading effects wherein gratings with nearby wavelengths are permitted to operate in the same wavelength space such that their spectra may become confused. To avoid shading effects, the wavelengths of the set of gratings being monitored should be spaced to avoid overlaps of wavelength space given the expected strain levels and the corresponding shift in FBG wavelengths. This may require the use of a larger spectral band or tunable laser diodes to illuminate the fiber gratings. Another approach to high density multiplexing while avoiding shading effects is to use Optical Frequency Domain Reflectometry (OFDR) to characterize low reflectivity fiber grating strain sensors [9, 11]. This method employs a tunable light that is used to scan a low-reflectivity fiber grating. A

Fourier transform of the detector signals will give the locations of each fiber-grating sensor. With the knowledge of the sensor locations, the detector signal can be passed through narrow band frequency filters in order to separate out spectral information from individual grating sensors, even if the sensors occupy the same wavelength space. The OFDR approach was used in the tests described below to interrogate the FBG sensors and calculate the axial and transverse strains as a function of grating position along the fiber.

Use of Fiber Optic Strain Measurements to Monitor Cracks and Disbonds - The sensors for these proof-of-concept tests consisted of long chirped Fiber Bragg gratings (FBG) in polarization maintaining fiber. For chirped gratings, the line width function varies linearly with grating length. Thus, for a chirped grating, the wavelength of light which is reflected from the grating is a function of grating position. This allows the FBG to have an extremely long gage length such that strain gradients over long distances of the fiber can be measured. A scanning laser interferometer was used to readout the grating period as a function of grating position. Both strain along the fiber axis and the differential transverse strain could be deduced from the grating period.

Blue Road Research participated in a study with Sandia Labs to evaluate the use of embedded chirped FBG sensors for detecting crack growth in metal structures and disbonds in composite-to-metal bonded joints. The test specimens, similar to the ones shown in Figure 9, were cycled in tension to grow cracks and disbonds while the sensors were monitored. The objectives of the tests were to: 1) show that these sensors could be successfully installed into the specimen bond line and survive the cure process, 2) demonstrate that the distributed strain profile along the gratings could be measured, and 3) prove that the distributed axial and transverse differential strain profiles could be used to track the presence and location of the engineered flaws in the bond line and crack growth. The test article was subjected to extreme fatigue loads (peak loads = 80% of yield) to rapidly grow fatigue cracks from notches placed in each side of a center hole in the metal. To implant disbonds, a Teflon insert was placed in the bond line at one location and a mold release agent was applied in a controlled region to produce a weak bond in another location.

Figure 15 shows the overall design of the instrumented test specimen and placement of the FBG sensors within the bond line between the composite laminate repair and the parent metal structure. A series of 110 mm long FBG were embedded into the bond lines between the metal panel and the boron fiber composite doublers. Gratings A and C were routed between the plate and adhesive across the doubler end (and across the Teflon insert), while gratings B and D were routed in a parabolic path roughly parallel to the long axis of the

doubler. The optical fiber was taped down to the steel surface prior to the application of the adhesive and the doubler and survived the high temperature (225°F) and vacuum bag process used to bond the composite doubler to the metal plate. The gratings were purposely oriented so that the polarization axes of the fiber were at 45 degrees to the plane of the metal panel. This orientation gave the sensor sensitivity to changes in shear in the bond line between the doubler and the panel. A series of strain gages were also bonded to the surface of the metal and composite laminate to monitor load transfer into the repair, as well as stress risers along the doubler edge and in the vicinity of the fatigue cracks.

Figure 16 shows a close-up of the FBG sensors, and associated end connectors, as they enter and exit the bond line of the test specimen. Typical data acquisition equipment used to monitor the FBG sensors is also pictured. During the course of the structural tests, the cyclic loading was periodically paused to allow the grating sensors to be interrogated. At these intervals, crack lengths were measured using eddy current and optical magnification methods. The presence of disbonds was monitored using hand-held ultrasonic pulse-echo inspections.

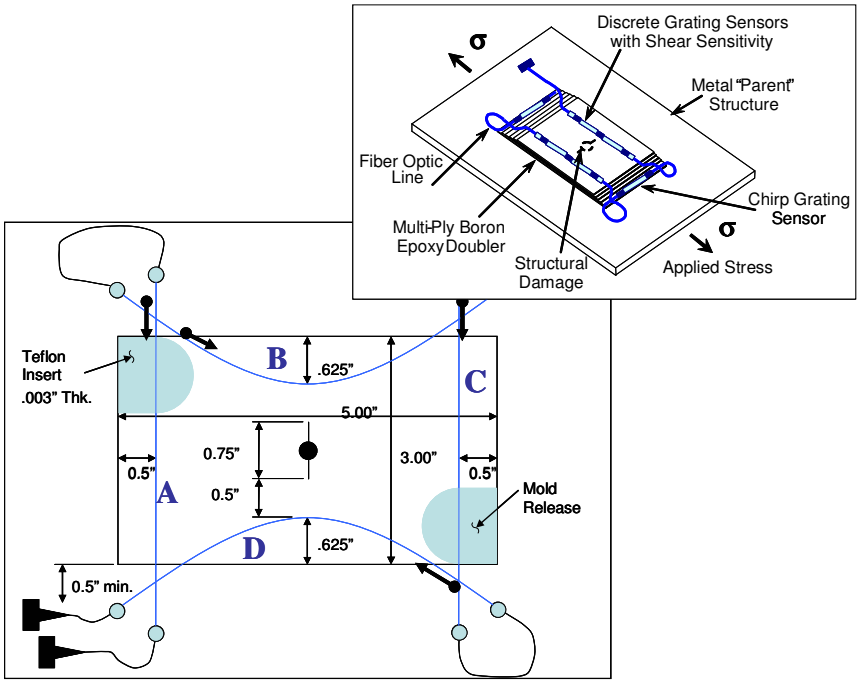


Figure 15. Fiber Bragg Grating Sensor Layout and Engineered Flaws in Bonded Composite Repair Test Specimens.

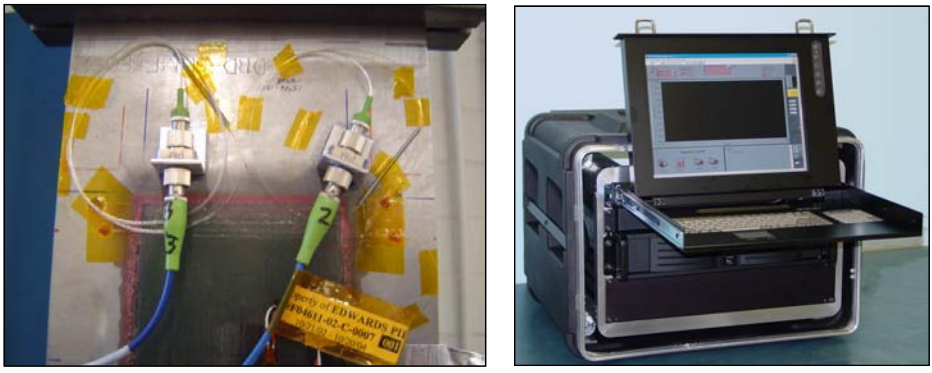


Figure 16. Fiber Optic Sensors in Adhesive Bond line and FO Monitoring Equipment.

Figure 17 contains sample results from the FBG sensors (sensors B & D) as they monitored fatigue crack growth. Both plots show the strain distribution along the sensors as a crack tip approaches. The plot on the left shows the strain levels produced by the crack when there was no load on test specimen 1. Specimen 1 contained a central crack but no disbond flaws. Note that the strain levels are close to 0 με in regions away from the cracks in the center of the plate. The plot on the right shows the strain levels along the axis of sensor B when specimen 2 was loaded to its maximum tension load of 34,000 lbs. A similar strain profile is observed, however, the strain levels produced away from the center cracks is now 1000 με. Normal stress-strain calculations determined that the maximum strain in the plate at 34,000 lbs. load was 1000 με. In addition, strain gage readings in the uniform strain region away from the cracks were between 950 με and 1020 με. Data such as this demonstrated the accuracy of the strain levels obtained from the FBG sensors. As the cycle number increased and the crack propagated further out from the center of the panel, the strain near the center position of the grating started to rise dramatically. At 63,000 and 67,000 cycles the strain at the 45 mm point in Sensor B (right side plot) finally falls to 0 as the crack passes through the fiber and relieves the strain in that portion of the sensor. Although the FBG sensors were 0.5" or further away from the crack tip, significant and localized strain changes were observed near the center of the sensors with less than 0.100" crack growth.

Sensors A and C were used to interrogate the bond line along the edge of the test specimens. As shown in Figure 15, these edges contain the tapered portion of the doubler and are critical, high shear strain areas because the majority of the load transfer into the composite repair occurs here. As a result, it is important to

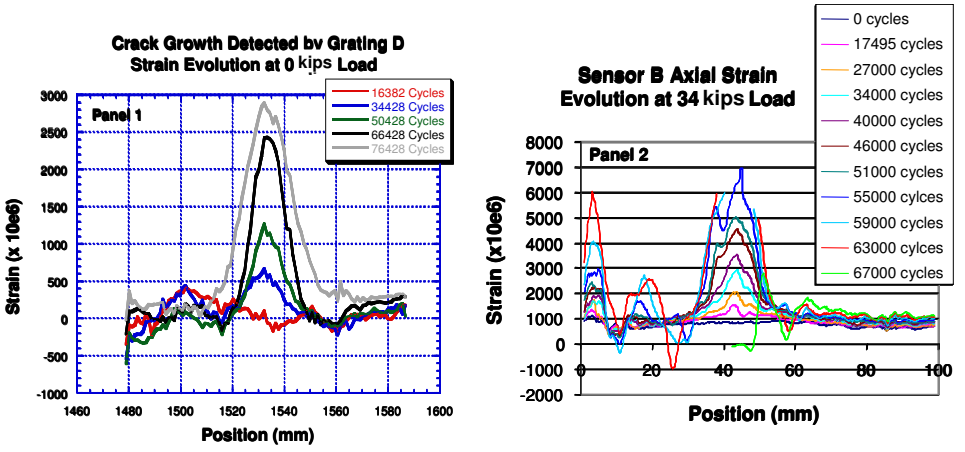


Figure 17. Strain Distribution Along Sensors *B* and *D* Determine Strain Levels in the Specimen and Indicate Crack Growth During Fatigue Testing.

monitor the taper region for disbonds. Since the fibers containing sensors A and C run across the width of the bond line (perpendicular to the direction of load), the shear strain component provided the best measure for identifying disbonds. Figure 18 shows the shear strain distribution along sensor A while the specimen was under load. Note the 1000 με level across most of the taper region (from 1185 mm to 1235 mm where the bond line is intact). As observed in the sensor B data above, the maximum strain of 1000 με in the direction of the load is confirmed by sensor A. However, in the region of the disbond produced by the Teflon insert, the shear strains are almost 0 indicating the presence of an extreme strain relief (i.e. disbond eliminates load transfer into doubler in this region). This plot indicates that the disbond extends from the 1160 mm position to the 1185 mm position correctly predicting the size of the one inch (25 mm) Teflon insert.

Remote Field Eddy Current

There are two fundamental approaches that have been developed to improve the performance of eddy current NDI. The first approach attempts to cancel, or compensate for, the large surface signal that is directly coupled to the pick-up coil without first entering the part being inspected. This portion of the overall eddy current signal has no useful information and serves to increase the noise floor. It can be removed by altering the probe design or by applying special data analysis to the captured signal. The Self-Nulling Probe (SNP) and the Remote Field Eddy Current (RFEC) method are examples of the probe alteration

approach [12-13]. The sensor and the drive coil of the SNP are separated by a ferromagnetic cylinder called a "flux focusing lens" to minimize the sensor signal that arrives directly from the drive coil. The RFEC approach blocks the unwanted, directly-coupled signal through the use of shields and magnetic circuitry. This produces a signal that is dominated by the indirect coupling path and contains the information of interest.

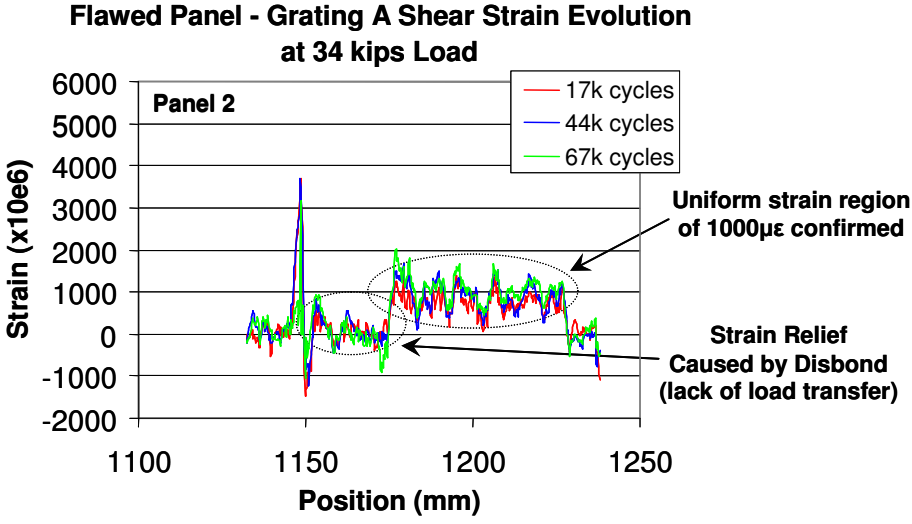


Figure 18. Shear Strain Levels Monitored by Sensor *A* Indicating the Presence of a Disbond.

The inspection of multi-layered structures requires detection of fine cracks and shallow corrosion pits that are embedded deep below the inspection surface. If there are air gaps between the layers, ultrasonic and thermography methods are not good candidates since it is difficult for them to penetrate beyond the first layer. Conventional eddy current (EC) techniques are often limited by depth of penetration effects wherein deeper penetration requires larger EC coils that produce a corresponding loss in sensitivity. A relatively new breakthrough in the eddy current arena involves the use of "remote fields" to interrogate conductive structures. The Remote Field Eddy Current (RFEC) method is especially suited for detecting deep flaws.

Figure 19 depicts the physics behind the RFEC approach. When a coil is excited by an alternating current and is placed on a structure the energy diffuses along two different paths. The interaction between the two fields results in what is often referred to as the remote field eddy current effect. Studies have shown that the energy diffusing via the direct path attenuates very rapidly. The signal

received by the pick-up coil that is located a certain distance away from the excitation coil is primarily due to the energy diffusing via the indirect path. This portion of the energy passes through the structure twice before arriving at the pick-up coil. Thus, the RFEC probe is designed to focus on the indirect coupling path, so that the signal measured by the pickup coil carries the information of the whole wall-thickness. It produces deeper penetration for detection of deep, subsurface flaws. Traditional eddy current approaches depend on changes in the electromagnetic flux to signal the presence of a flaw. Typical values of such changes are very small and may be less than 0.01%. Therefore, it is very difficult to separate this change from the quiescent signal generated by an unflawed structure. This limits the maximum gain that one can employ when using these instruments. RFEC is based on the measurement of the voltage induced in a pick-up coil by the flux that has passed through the test object twice (see Figure 19). This set-up creates a significant change in voltage so that a stronger flaw signal is produced. The end result is a high ratio of flaw-signal-to-quiescent-signal strength. This allows higher gain settings to be used so that deeper flaws can be detected.

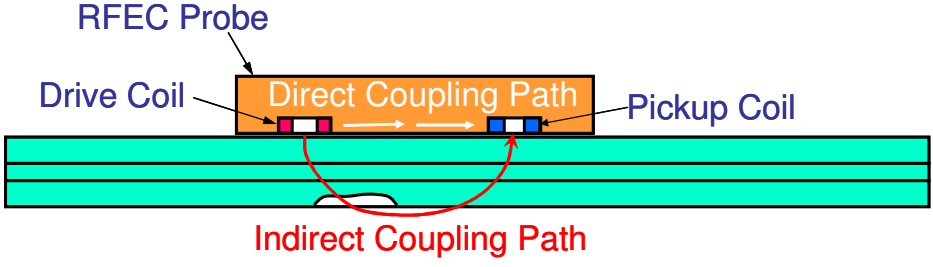


Figure 19. Schematic of Remote Field Eddy Current Approach.

A super-sensitive eddy-current system has been developed to accommodate the RFEC probes. The system is capable of providing an extremely high gain, up to 135 dB, for amplifying the remote-field probe signal. This allows the inspection device to provide higher signal-to-noise ratios thus increasing its flaw detection sensitivity. Sample crack detection signals from a 2 KHz RFEC probe inspection are shown in Figure 20. These plots demonstrate that the voltage amplitude-based data presentation requires minimal signal interpretation. Experimental measurements have shown the ability of the RFEC method to detect a 0.03” long fatigue crack that is located 0.446” below the inspection surface.

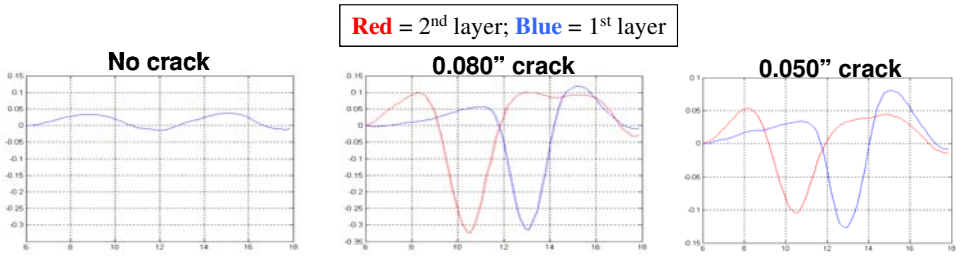


Figure 20. First and Second Layer Crack Detection in 0.040" Thick Plates with Cracks Hidden Under Raised Head Fasteners.

When scanning a multiple layer aluminum sample with a total thickness of 0.29", the RFEC system was able to detect wall thinning as small as 1% of the total thickness. Figure 21 shows a hand-held RFEC probe scanning a riveted aluminum joint containing various levels and sizes of corrosion. The dotted lines show the locations of the corrosion sites that were engineered into the specimens. Sample signals from the hidden, second layer corrosion regions are shown in Figure 22. The RFEC system produced signal-to-noise ratios that were well in excess of 5 to clearly indicate the presence of corrosion. Note that the size of the second-layer corrosion detected with the signals shown in Figure 22 was 0.125" in diameter and ranged from 0.0008" to 0.004" deep (2% to 10% corrosion thinning in 0.040" thick plate).

RFEC Deployed with In-Situ Sensors – A sensor called the RFEC Magnetic Carpet Probe (MCP) was developed to deploy this inspection for in-situ health monitoring. This novel leave-in-place MCP can be installed to detect and monitor fatigue cracks or corrosion in inaccessible regions on structures. Networks of surface mounted and embedded eddy current sensor arrays permit rapid, large area damage monitoring with the high resolution of localized probe testing. The RFEC-MCP is made using a piece of thin, flexible printed circuit board (PCB) with densely populated printed coil sensors on the PCB as shown in Figure 23. A sensor for detecting cracks emanating from a hole and a sensor for monitoring larger areas for corrosion are shown. Figure 23 also shows a prototype sensor, with co-located electronics, measuring corrosion in an NDI Reference Standard.

The sensors work on the same principal as the RFEC method described above to provide deep penetration for locating subsurface flaws. Multiple-phase AC currents are applied to each sensor on the PCB. This generates a traveling magnetic wave that is stronger than those generated by conventional eddy current excitation. Each sensor on the PCB alternately acts a transmitter and receiver so

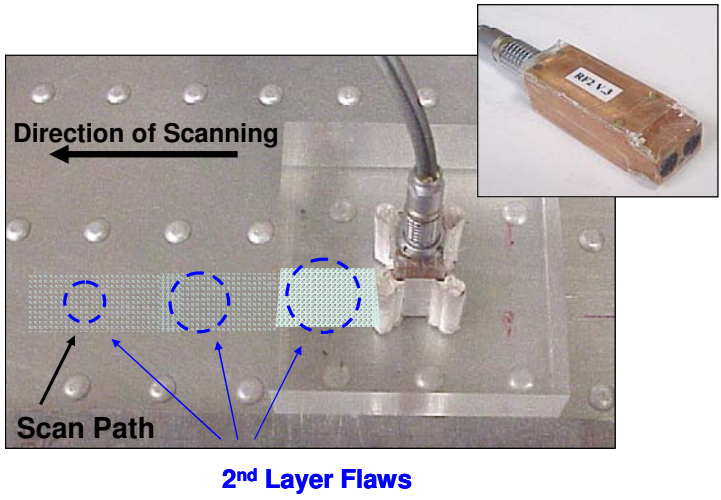


Figure 21. RFEC Probe Inspecting for Corrosion in a Riveted Aircraft Joint.

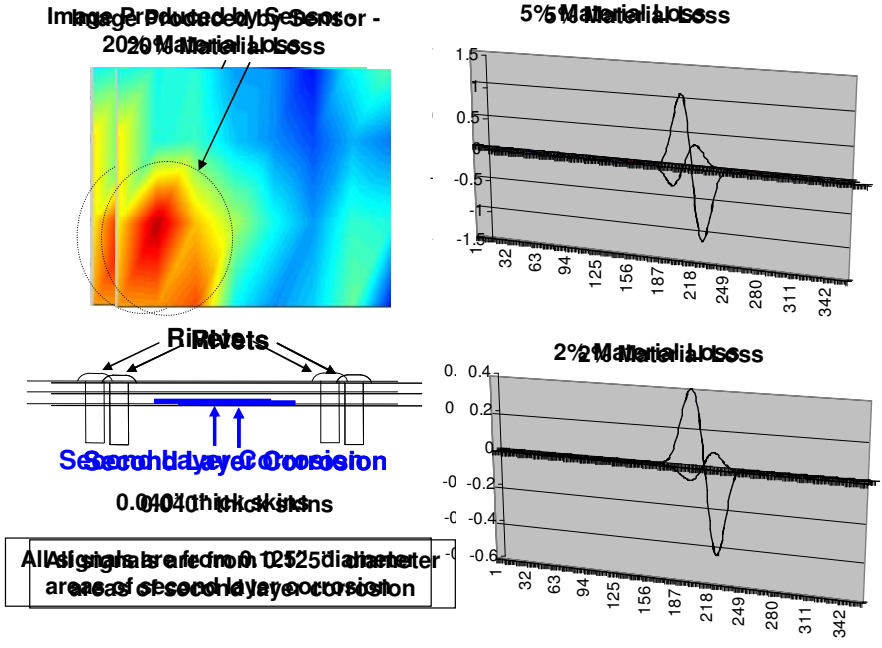


Figure 22. Large Signals Produced by RFEC Probe Indicating the Presence of Corrosion.

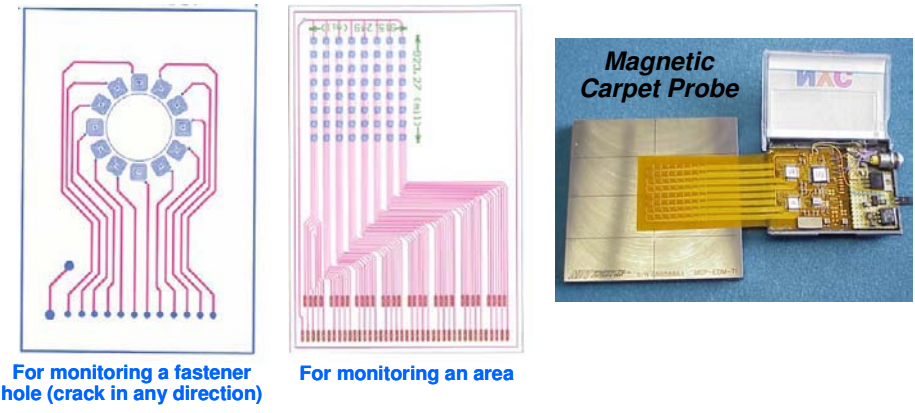


Figure 23. Sample RFEC Printed Circuit Board Designs and Prototype Corrosion Sensor with Co-Located Electronics.

that a uniform magnetic field is set-up over the entire MCP area and the probe can detect cracks in any orientation. In addition, the RFEC-MCP method is insensitive to the geometry variations of a component while competing ultrasonic-based sensors must carefully account for the effects of complex geometries on the transmission of UT waves.

Meandering Winding Magnetometer Probes

JENTEK Sensors has developed mountable, electromagnetic sensors for fatigue and corrosion monitoring of aircraft structures. The Meandering Winding Magnetometer (MWM) and the segmented-field, spatial wavelength Interdigitated Dielectrometer (IDED) have been demonstrated for a wide variety of health monitoring applications. The MWM is an inductive, eddy current, sensor and the IDED is a capacitive sensor [14-16]. MWMs are suitable for metals, graphite fiber composites, reinforced carbon composites, and low observable coatings (nonconducting, magnetizable coatings using magnetic particle suspensions), while the IDED is suitable for characterization of glass fiber composites, corrosion protection coatings, sealants, glass, paint, and wood, as well as for detection/monitoring of corrosion products, moisture ingress and monitoring of cure states of epoxies and adhesives. The corrosion monitoring technology is relatively new while the MWM-Array fatigue monitoring technology is already a commercial product to support component fatigue testing. It is being transitioned to use on aircraft for specific applications. Through the combined use of IDED and MWM sensors it is possible to not only detect metal

loss, as with eddy current-based NDI methods for corrosion, but also to detect corrosion products and moisture ingress within the joint itself.

In both inductive and capacitive sensors, a single drive element is used with multiple sense elements to facilitate parallel data acquisition of sensor responses and to eliminate cross-talk or interference between multiple sense elements. For these sensors, the responses can be modeled accurately. This allows calibration information to be reused each time electrical contact to the sensor is established prior to performing a measurement. These sensing modalities are advantageous for leave-in-place sensors, where it is important to be able to perform self-diagnostics on the sensor to confirm correct sensor operation. Example geometries for MWM sensors are shown in Figure 24. Electrical current is passed through one or more drive winding loops to impose a near-surface magnetic field in the test material. Secondary windings located between the drive winding segments are used to sense the perturbations in the magnetic field due to the presence and condition of the test material.

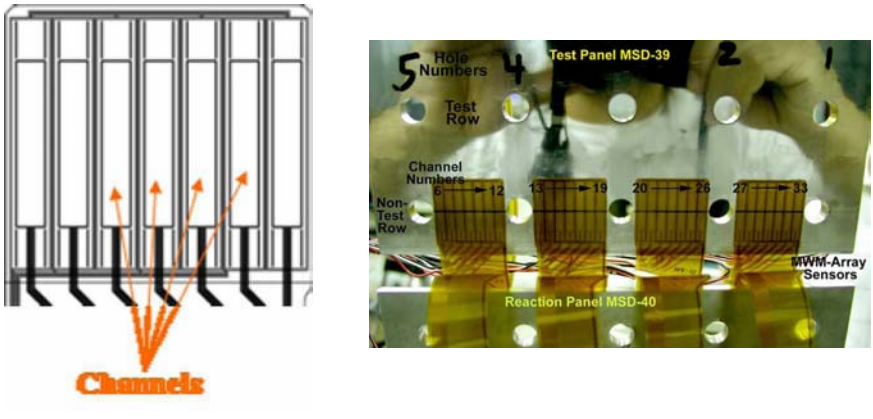


Figure 24. Seven Channel MWM Array and Surface Mounted MWM Arrays Monitoring for Cracks Around Fastener Holes.

Hidden Corrosion Imaging Using MWM Array Sensors **-** Measurements were performed on a test panel containing areas of simulated hidden corrosion. The panel was fabricated from two sheets of 7075-T6 aluminum alloy, each 0.040 in. thick. The sheets were joined by six rows of removable fasteners. There were five rows of flat bottom holes located between the fastener rows. The flat bottom holes ranged from 0.125” to 1.0” in diameter and from 0.0008” to 0.008” in depth. During the tests, the corrosion sites were located on the bottom sheet at the interface between the sheets. Figure 25 provides a drawing of a section of the

panel. Scans were produced using an existing FA24 (37 channel) sensor, JENTEK impedance instrument, and the GridStation inversion system. An MWM Array image obtained by scanning the row of 0.008" deep (10% corrosion levels) flat bottom holes is shown in Figure 26. The color scale represents remaining total thickness of material which agrees well with the known values for the higher volume areas. Since the fasteners had raised heads, only the response of the sense elements positioned between the fasteners is shown. Similar results were obtained for the other rows.

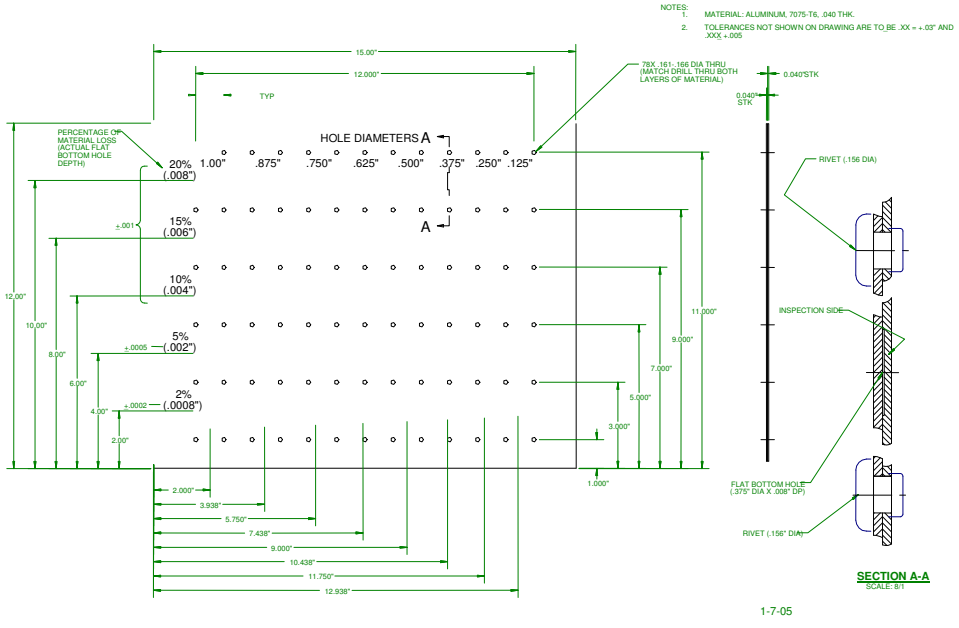


Figure 25. Riveted Panel with Corrosion Sites Ranging from 1/8" to 1" in Diameter and 5% to 20% in Depth.

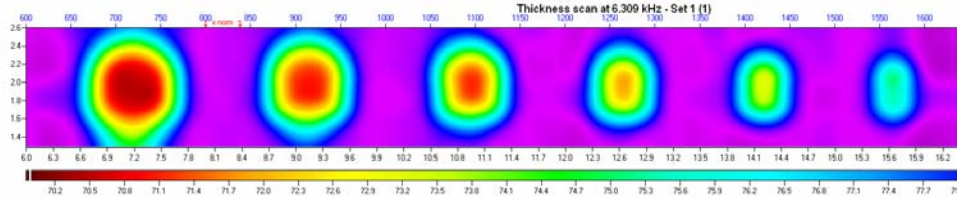


Figure 26. MWM-Array Results for the Row of 0.008" Deep (10% thickness loss) Corrosion Sites at the Hidden Interface Between the Two Plates.

Other SHM Sensors

Breakthroughs in microelectronics are being used to miniaturize conventional hand-held probes into mountable, in-situ devices. Miniature flaw detection sensors, combined with easily fieldable diagnostic electronics, are the backbone of any global SHM system. An extensive amount of research is ongoing to develop and mature a wide variety of new sensors that are applicable to in-situ structural health monitoring. The sensors may directly detect flaws, such as those discussed here, or other physical parameters, such as acceleration, force, pressure, flow, displacement (capacitive proximity sensors), temperature and the presence of chemicals, from which flaws or faulty operation can be identified. Other leave-in-place sensors that may produce direct indications of structural flaws include: magnetoresistors, capacitive micromachined ultrasonic transducers (cMUT) [17], flexible eddy current array probes [18], nickel-foil magnetostrictive sensors (MsS) [19], optical fiber long-period grating (LPG) chemical sensors [20], adhesive bond degradation sensors (ABDS) [21], microwave antenna, acoustic emission transducers, spectroscopic instrumentation; optoelectric sensors such as injection lasers, photovoltaic diodes, photoconductors, and polycrystalline silicon for measuring an array of physical properties including transmission of ultrasonic waves and variations in electric fields.

Following are brief descriptions of other sensors being pursued for SHM.

1. *Micro Eddy Current Sensor* (Sandia Labs) – To support civil engineering applications, Sandia has developed a customized eddy current sensor for crack detection in thick steel structure. The probe can detect deep, subsurface cracks that may be as much as 0.5” below the surface. Impedance bridge and other differential circuits were explored to maximize the magnetic flux density and corresponding eddy current strength. Successful crack detection was achieved with a dual coil configuration that combines a pancake excitation inductor with a co-located pickup coil to produce a transducer that requires very little drive current (75 mA) and operates in the desired 10 KHz range. Figure 27 shows the prototype sensor along with the large signal it produced when it was scanned over a crack. Excellent crack detection was achieved even when inspecting through surface materials approaching 0.5” thick (EC lift-off impediment). Ongoing activities are developing a more

sophisticated electronics package - including digital signal processing to filter and detect phase shifts - to further improve probe sensitivity. The goal is to produce a dual coil configuration that packages conditioning electronics, shielding, data logging and data transmission to remote sites into a rugged surface mountable sensor as shown in Figure 28.

2. *Capacitive Micromachined Ultrasonic Transducers* (Stanford, General Electric) – This is an electrostatic transducer that utilizes semiconductor fabrication techniques to allow for mass production and low cost (see Figure 29 example). An applied voltage produces a displacement of the membrane (generates ultrasonic wave) due to electrostatic attraction. When used to receive input signals (monitor ultrasonic waves), the displacement of the membrane causes a change in capacitance of the device. Advantages of this device include completely integrated electronics and a large operation bandwidth.
3. *Flexible Eddy Current Array Probes* (General Electric) – This probe allows the advantages of eddy current arrays to be implemented in a conformable, printed circuit board (see Figure 30). The sensor incorporates magneto-resistors and spiral coils in two-dimensional arrays. This allows for simultaneous, multi-frequency inspections and real-time imaging of cracks and corrosion. Successful demonstrations of GE flexible EC sensors have been completed on testbed aircraft at Sandia Labs.

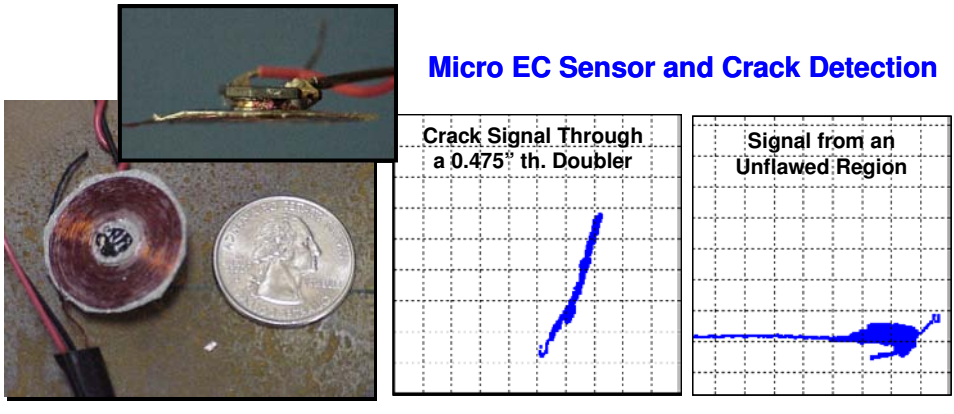


Figure 27. Sandia Pancake EC Sensor and Large Signals Showing Crack Detection Through Thick Composite Doublers.

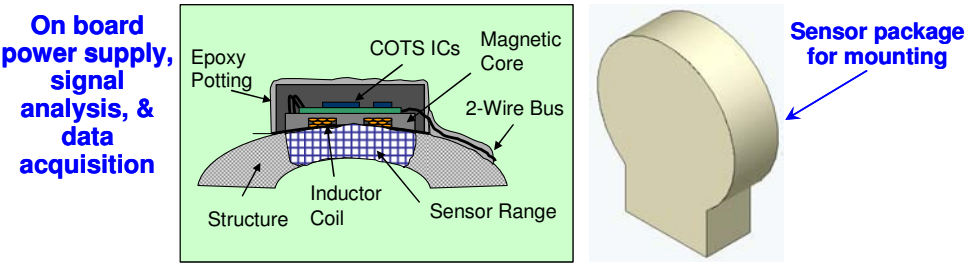


Figure 28. EC Sensor Packaged with Electronics for Automated Signal Conditioning and Data Logging.

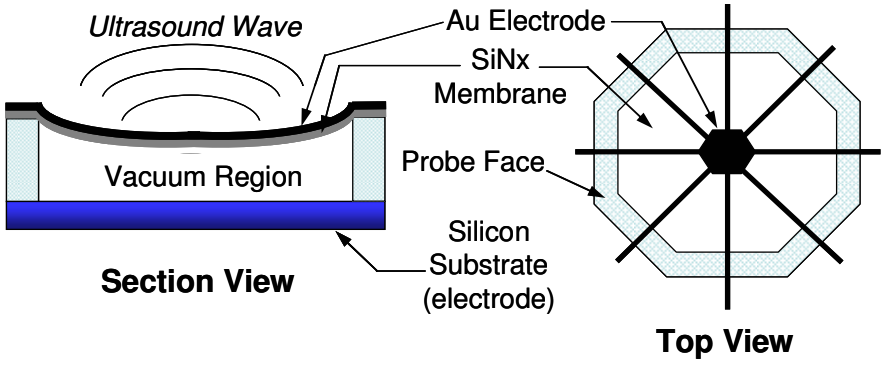


Figure 29. Schematic of Capacitive Micromachined Ultrasonic Transducer (cMUT).

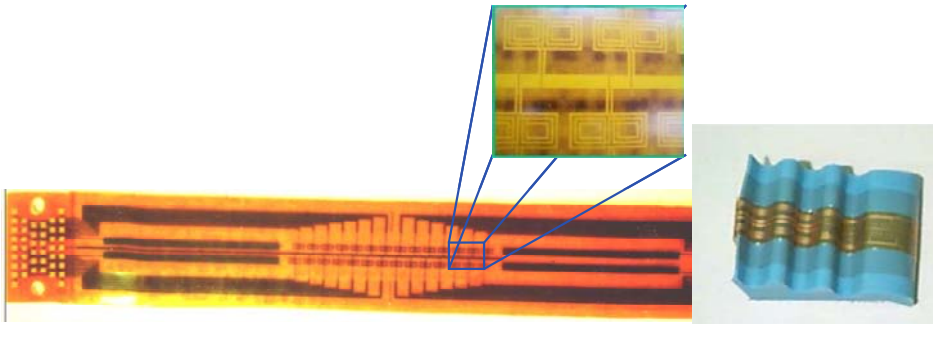


Figure 30. General Electric's Flexible Eddy Current Array.

4. *Nickel-Strip Magnetostrictive Sensors* (Southwest Research Institute) - This guided wave probe consists of a thin nickel substrate (foil) and a thick MsS coil in a printed circuit board that is placed on the nickel strip. The strip is mounted on the structure being monitored. It generates pulses

of guided waves and detects the signals being returned from within the structure. Data comparisons with baseline signals (undamaged configuration) are used to identify defects. A single probe can work over long distances allowing for economical SHM of large structures.

5. *Adhesive Bond Degradation Sensor* (Lockheed-Martin, Tetra Tech) – This sensor is used in conjunction with fiber optic lines. The ABDS is being developed to detect bond line decay by measuring water ion concentration in the adhesive. These ions are by-products of decay at the adherend-to-adhesive interface. Features of light reflection from the ABDS are altered in response to the concentration of water ions. Research is underway to calibrate the Adhesive Bond Degradation Sensor in order to quantitatively relate the level of water ions to the decay in bond strength.
6. *Long Period Grating (LPG)-Based Chemical Sensors* (Luna Innovations, Boeing, Navy) – This is an optical fiber sensor that can detect precursors to corrosion such as moisture, pH levels, and metal-ion corrosion by-products. By applying special coatings that change refractive index with absorption of target molecules on the LPG surface, the optical fiber sensor becomes the transducer for chemical measurement. For detection of moisture, for example, the LPG sensor is coated with a PEO hydrogel that swells in the presence of moisture.

Alternatively, sacrificial corrosion sensors – for installation on the ends of fiber optic lines – have also been developed (Luna Innovations Inc., BAE Systems). These sensors are designed to deteriorate in the same manner and at the same rate as the parent structure they are monitoring. Characteristics of light reflected back through the fiber optic line can be calibrated to indicate different levels of corrosion.

Deployment of Health Monitoring Sensor Networks – Distributed sensor networks can be deployed in any of the three major approaches listed below. These options are listed in the order of increasing complexity, however, less labor is required to monitor the systems as they become more complex.

1. *In-Situ Sensors Only* – The sensors are the only items permanently installed on the structure. At the desired inspection intervals, power, signal conditioning, and data acquisition electronics are manually transported to the structure to be monitored. The sensors are linked to the monitoring electronics via an electrical connector and flaw detection is completed by an inspector at the site.

2. *Sensor Network with In-Situ Data Acquisition* – In this system, miniature, packaged electronics is also placed in-situ with the sensor network. The electronics contains the necessary power, memory and programmable circuitry for automated data logging. The data is periodically downloaded to a laptop through manual hook-ups at the site.
3. *Sensor Network with Real-Time Data Transmission to a Remote Site* – This approach is similar to item #2 with the addition of a telemetry system that allows for continuous, wireless transmission of data to a remote site. A web site can be programmed to interrogate critical aspects of the data and use pre-set thresholds to provide continuous green light/red light information regarding the health of the structure. The web site can even be programmed to automatically send an e-mail to maintenance personnel if the condition monitoring process indicates the need for repairs or other maintenance. In this mode of operation it may be desirable to incorporate interface electronics to condition the signals and analyze data in-situ. This eliminates the need to transmit larger raw data files.

The latter approach allows for true condition-based maintenance in lieu of maintenance checks based on time of operation. A series of expected maintenance functions will already be defined, however, they will only be carried out as their need is established by the health monitoring system. The use of condition-based maintenance coupled with continuous on-line structural integrity monitoring could significantly reduce the cost of inspection, maintenance, and repair. In order to endure long periods of operation without any direct interface with maintenance personnel, it may be desirable to use energy harvesting methods that can power sensors systems by converting structural stresses (strain energy harvesting) into electrical power via piezoelectric transducers [22].

Conclusions

Detection of unexpected flaw growth and structural failure could be improved through the use of on-board health monitoring systems that could continuously assess structural integrity. Such systems would be able to detect incipient damage before catastrophic failures occur. Local sensors, such as the ones described in this chapter, can be used to directly detect the onset of crack, corrosion, or disbond flaws. Global SHM, achieved through the use of sensors such as accelerometers coupled with structural dynamics assessments, can be used to assess overall performance (or deviations from optimum performance) of large

structures such as bridges and buildings. Whether the health monitoring approach is local or global, the key element in a SHM system is a calibration of sensor responses so that damage signatures can be clearly delineated from sensor data produced by unflawed structures.

This chapter focused on local flaw detection using embedded sensors. While some of these leave-in-place sensors are able to produce wide area inspections, their use is predicated on the identification of primary flaw regions to be monitored. The replacement of our present-day manual inspections with automatic health monitoring would substantially reduce the associated life-cycle costs. The ease of monitoring an entire network of distributed sensors means that structural health assessments can occur more often, allowing operators to be even more vigilant with respect to flaw onset. When accessibility issues are considered, distributed sensors systems may also represent significant time savings by eliminating the need for component tear-down. In addition, corrective repairs initiated by early detection of structural damage are more cost effective since they reduce the need for subsequent major repairs. Aerospace structures have one of the highest payoffs for SHM applications since damage can quickly lead to expensive repairs and aircraft routinely undergo regular, costly inspections.

In general, SHM sensors should be low profile, lightweight, easily mountable, durable, and reliable. To reduce human factors concerns with respect to flaw identification, the sensors should be easy to monitor with minimal need for users to complete additional data analysis. For optimum performance of the in-situ sensor based approaches, the signal processing and damage interpretation algorithms must be tuned to the specific structural interrogation method. For example, in the high-frequency E/M impedance approach, pattern recognition methods can be used to compare impedance signatures taken at various time intervals and to identify damage presence and progression from the change in these signatures. These approaches can benefit from the use of artificial intelligence and neural networks algorithms that can extract damage features based on a learning process.

Recent failures of aircraft and civil structures have compelled the engineering community to take a fresh look at the fail-safe, safe-life, and damage tolerance design philosophies. The effect of structural aging and the dangerous combination of fatigue, corrosion, and other environmentally-induced deterioration is now being reassessed. The end result of these assessments has been a greater emphasis on the application of sophisticated, health monitoring systems. Recent advances in on-board structural health monitoring sensors have proven that distributed and autonomous health monitoring systems can be

applied to reliably detect incipient damage. Such systems have wide range of uses in aerospace, automotive, civil infrastructure and other industrial applications.

References

1. Bartkowicz, T. J., Kim, H. M., Zimmerman, D. C., Weaver-Smith, S. (1996) "Autonomous Structural Health Monitoring System: A Demonstration", *Proceedings of the AIAA/ASME Structures, Structural Dynamics, and Materials Conference*, April 1996.
2. Beral, B., Speckman, H., "Structural Health Monitoring for Aircraft Structures: A Challenge for System Developers and Aircraft Manufacturers," 4th International Workshop on Structural Health Monitoring, September 2003.
3. Roach, D., "*Use of Distributed Sensor Systems to Monitor Structural Integrity in Real-Time*," Quality, Reliability, and Maintenance in Engineering, Professional Engineering Publishing Ltd., Oxford, UK, 2004.
4. Wheatley, G., Kollgaard, J., Register, J., Zaidi, M, "Comparative Vacuum Monitoring as an Alternate Means of Compliance," FAA/NASA/DOD Aging Aircraft Conference, September 2003.
5. Giurgiutiu, V., Redmond, J., Roach, D., and Rackow, K, "Active Sensors for Health Monitoring of Aging Aerospace Structures," *International Journal of Condition Monitoring and Diagnostics Engineering*, February 2001.
6. Lin, M., Beard, S., Kumar, A., Qing, X., "Development of Structural Health Monitoring Systems Using SMART Layer Technology," American Society for Composites 17th Technical Conference, October 2002.
7. Kumar, A., Beard, S., Hannum, R. (Acellent Technology), and Roach, D. (Sandia Labs), **"**In-Situ Monitoring of the Integrity of Bonded Repair Patches on Civil Infrastructures," SPIE Conference on NDE for Health Monitoring and Diagnostics, February 2006.
8. Roach, D., Wanser, K., and Griffiths, R., "Application of Fiber Optics to Health Monitoring of Aircraft," Advanced Aerospace Materials and Processes Conference, June 1994.
9. E. Udd, W.L. Schulz, J.M. Seim, E. Haugse, A. Trego, P.E. Johnson, T.E. Bennett, D.V. Nelson, A. Makino, "Multidimensional Strain Field Measurements using Fiber Optic Grating Sensors", SPIE Proceedings, Vol. 3986, p. 254, 2000
10. M. Froggatt and J. Moore, "Distributed Measurement of Static Strain in an Optical Fiber with Multiple Bragg gratings at Nominally Equal Wavelengths", Applied Optics, Vol. 37, p. 1741, 1998.
11. Udd, E., Kreger, S., Calvert, S., Kunzler, M., and Davol, K., "Usage of Multi-Axis Fiber Grating Strain Sensors to Support Nondestructive Evaluation of Composite Parts and Adhesive Bond Lines," 4th International Workshop on Structural Health Monitoring, September 2003.
12. Sun, Y., and Roach, D., "New Advances in Detecting Cracks in Raised-Head Fastener Holes Using Rotational Remote Field Eddy Current Technique," American Society of Nondestructive Testing Research Symposium, October 2005.

13. Sun, Y. and Ouyang, T., "Application Of Remote-Field Eddy-Current Technique to Aircraft Corrosion Detection", presented at Tri-Service Corrosion Conference, January 2002, *Corrosion* Journal 2002.
14. Goldfine, N., et al, "Method of detecting Widespread Fatigue and Cracks in a Metal Structure," U.S. Patent Number 6,420,876, July 16, 2002.
15. Goldfine, N. 1993, "Magnetometers for Improved Materials Characterization in Aerospace Applications," *Materials Evaluation*, March 1993.
16. Goldfine N., Washabaugh, A., Schlicker , D., Sheiretov, Y., Huguenin, C., Lovett, T. (JENTEK Sensors), and Roach, D. (Sandia Labs), "Corrosion and Fatigue Monitoring Sensor Networks," 5th International Workshop on Structural Health Monitoring, September 2005.
17. Barshinger, J., "GE Nondestructive Evaluation and Introduction to Capacitive Micromachined Ultrasonic Transducers," General Electric presentation to Sandia Labs, April 2005.
18. Plotnikov, Y., and Nath, S., "Real Time Imaging of Subsurface Flaws Using the Pulse Eddy Current Array Probe," American Society of Nondestructive Testing Conference, November 2004.
19. Kwun, H., Kim, S., Light, G., "Long-Range guided Wave Inspection of Structures Using Magnetostrictive Sensor," *Journal of Korean Society of Nondestructive Testing*, Vol. 21, 2001.
20. Elster, J., Trego, A., Jones, M., Tulou, P., FitzPatrick, B., Perez, I., "Corrosion Monitoring in Aging Aircraft Using Optical Fiber-Based Chemical Sensors," FAA/NASA/DOD Aging Aircraft Conference, September 2000.
21. Jackson, A., Schaafsma, D., "Aircraft Fail-Safe Self Monitoring System," FAA/NASA/DOD Aging Aircraft Conference, September 2000.
22. Pfeifer, K., Leming, S., and Rumpf, A., "Embedded Self-Powered MicroSensors for Monitoring the Surety of Critical Buildings and Infrastructures," Dept. of Energy SAND Report SAND2001-3619, November 2001.

This page intentionally left blank

CHAPTER 19

NONDESTRUCTIVE TESTING FOR BEARING CONDITION MONITORING AND HEALTH DIAGNOSIS

Robert X. Gao and Shuangwen Sheng

Department of Mechanical & Industrial Engineering，University of Massachusetts
Amherst, MA, 01003, USA
{gao, ssheng}@ecs.umass.edu

Nondestructive Testing (NDT) techniques assess the conditions of rolling bearings without affecting their structural functionality, thus have played an important role in bearing health diagnosis. Among the various NDT techniques investigated for bearing applications, vibration and acoustic-based techniques are the most popular, since vibration and acoustic signals measured from bearings contain physical information that directly reflect the defect severity. In this chapter, a review of such NDT techniques is presented, and related sensors and sensing principles are discussed.

1. Introduction

Rolling bearings are fundamental mechanical elements providing rotational support to a wide range of machines [1]. Despite significant advancement in bearing design and maintenance over the past decades, unexpected failure of bearings still occur before they reach the designed life, due to a variety of reasons: material fatigue, adverse operating conditions, faulty installation, inappropriate lubrication, etc. To maximize the service life of bearings and subsequently, minimize downtime of the machines which the bearings are supporting, it is of critical importance to replace the traditional schedule-based maintenance with a new, condition-based maintenance strategy. For both bearing end users and designers, accurate evaluation of the present working status of the bearing being monitored and reliable prediction of its remaining service life are of key interest. To achieve this goal, various types of sensing and diagnosis techniques have been developed over the past several decades to identify the existence of bearing structural defects, their severity level, their possible causes,

and based on the diagnosis, to predict the time to failure of the bearings of concern [2], [3].

A bearing condition monitoring system typically consists of sensing elements (converting physical symptom to electrical voltage or current signals for ease of post processing), data acquisition devices (for data transmission from sensors to processing unit), software for feature extraction (separating characteristic information from the rest of the raw data), pattern classification (diagnosis of defect types and severity level) and remaining life prediction (prognosis of time left to total failure) [4]. Depending on whether the function of the monitored bearing is affected, the monitoring techniques are classified into Nondestructive Testing (NDT) and Destructive Testing (DT). Generally, NDT techniques are more suited for online monitoring whereas DT techniques provide more clues to bearing design and material property specification [5]. In this chapter, the discussion is limited to NDT techniques for bearing condition monitoring. Based on the type of the input signals or the *measurands*, the NDT techniques are grouped into a total of five types: Mechanical, Acoustic, Thermal, Chemical, and Electromagnetic [6], as shown in Table 1.

Table. 1. NDT techniques for bearing condition monitoring

NDT Techniques	**Sensors**	**Measurands**
Mechanical	Strain Gauge	Force
	Displacement Probe	Displacement
	Velocity Transducer	Velocity
	Accelerometer	Acceleration
Acoustic	Microphone	Sound
	Acoustic Emission Sensor	Acoustic Emission
Thermal	Infrared Thermometer	Temperature
Chemical	Chip Plugs	Wear Debris Ingredients and Size
	Ultrasonic Transducer	Lubricant Film Thickness
Electromagnetic	Current Transducer	Electric Current

Among the five types of sensors and their corresponding measurands, mechanical and acoustic techniques are the most popular due to the rich, defect-related physical information contained in the signals, their relatively low implementation cost, and the ease of implementation. They are reviewed in this chapter, with a focus on sensing. The reason for focusing on sensing is that, without accurate measurements, reliable and quantitative assessment of the

bearing conditions will not be feasible [7], even when sophisticated signal processing algorithms are applied.

2. Symptoms Related to Bearing Operation

Rolling element bearings vibrate no matter how perfectly they are designed and manufactured [8]. The reason is that bearing load is supported by a discrete number of rolling elements, which results in the periodical variation of bearing assembly stiffness when the bearing rotates. The vibrations, compounded by non-ideal operating conditions, such as inappropriate mounting or lubrication, lead to initiation of structural defects in a bearing during operation and may finally lead to bearing failure. During the entire bearing service life, various symptoms related to bearing operation, as shown in Fig. 1, can be observed and used for bearing condition monitoring.

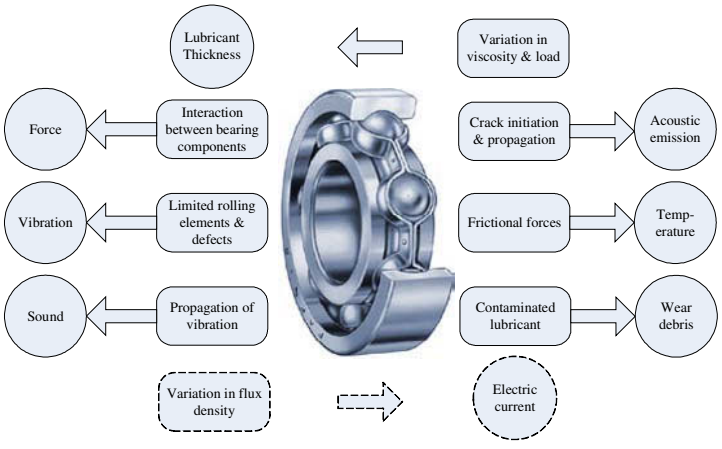


Figure 1. Symptoms observed during a bearing service life.

A structurally sound, "healthy" bearing generates vibrations due to the varying compliance (VC) [8]. The amplitude of VC vibrations depends mostly on the number of rolling elements supporting the applied load. In general, VC vibration increases with clearance (the gap between rolling element and raceway). The clearance leads to impulsive forces [9] between the rolling element and the inner/outer raceway, every time when they interact. Such forces cause material fatigue and deformation within a bearing structure, which subsequently leads to the initiation of microscopic cracks. These cracks will in turn induce stronger impulsive forces, leading to a self-deteriorating process that finally results in bearing failure, if no remedial measure is taken.

Vibration serves as a measure of bearing condition, and can be measured in three forms: displacement, velocity, and acceleration. Generally, the amplitude and frequency component of a vibration signal will increase with the deterioration of bearing condition. When vibrations propagate through the various media within a bearing assembly, such as solids and lubricants, audible sound waves are generated. As cracks are initiated and spread through the bearing structure, acoustic emissions are generated. The amplitude of acoustic emissions will increase with the growth of crack size.

In addition to the generation of vibrations, audible sound, and acoustic emissions, contaminated lubricant and increased temperature can also be observed during a bearing's service life. With the initiation of micro-cracks, small pieces of bearing materials fall into the lubricant and contaminate the lubrication oil. While wear debris is generally the object of chemical NDT, lubricant contamination increases frictional forces between the rolling elements and the inner/outer raceways, resulting in increased temperature in bearing structure. The increased temperature subsequently changes the viscosity of the lubricant and leads to variations in the lubricant film thickness [10]. If the monitored bearing is located inside of an electric motor, electric current can be used as an additional indicator [11]. This is because bearing defect will produce relative motion in the radial direction between the rotor and stator, and result in anomalies in the air gap flux density [3]. The spectrum of electric current is a function of motor power supply frequency and bearing characteristic vibration frequency. The amplitude of electric current and its number of harmonics will increase when the bearing condition deteriorates.

3. Mechanical Testing

Mechanical NDT techniques based on force and vibration measurements have been frequently applied for bearing condition monitoring and fault diagnosis. In this section, commonly used sensors such as strain gauges, displacement probes, velocity transducers, and accelerometers are reviewed. For each type of the sensors, pertinent sensing principle and the application to bearing condition monitoring are presented.

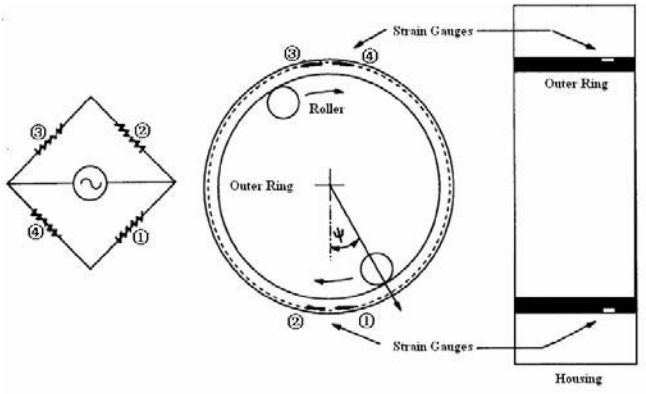
3.1. *Force Sensing*

Impact force generated during bearing operation is a direct measure for bearing condition. Abnormal increase in the impact force amplitude usually indicates structural deterioration. Impact force measurement is commonly performed by integrating a force sensor into the chain of force application (e.g. between a

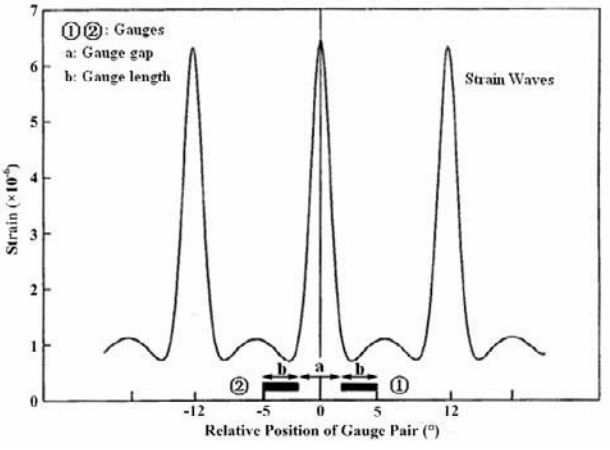
bearing and the housing) so that the sensors measure the same force experienced by the bearing. Because of the relatively high energy content contained in the impacts generated during the interactions between the rolling element and structural defect, impact force signals generally have a high signal-to-noise ratio, making them a good candidate for measurement [12]. Several efforts have been made on using forces for bearing condition monitoring and the sensors investigated included strain gauges [13]-[16], fiber optic displacement probes [17], and current carrying plates [18].

A strain gauge is a resistive elastic sensor whose resistance is a function of the applied strain (unit deformation) [19]. Such a relationship describes the resistance change over unit strain (dimensionless), by means of a gauge factor. For metallic wires, the gauge factor ranges from 2 to 6. For semiconductor, it can be as high as 40 to 200. To obtain accurate strain data, small resistance changes in the range of micro-Ohms must be measured. A Wheatstone bridge circuit is used to convert the microstrain in gauges into voltage changes [20]. The bridge circuit can be employed in full, half or quarter configurations, among which the full bridge circuit has the highest sensitivity (e.g., twice as that for a half-bridge) and signal-to-noise ratio (e.g., four times of a quarter bridge) and lowest error-prone components (e.g., all four branches can be temperature-compensated to reduce thermal errors) [20]. Factors to be considered when selecting a strain gauge include the gauge length (generally 0.2 mm to 100 mm), gauge pattern (uniaxial or multiaxial), nominal resistance (120 Ohm or 350 Ohm), and materials (e.g., silicon or platinum alloy) [21].

Strain gauges are normally sandwiched between the bearing housing and its outer ring to measure elastic deformations in the bearing outer ring. These deformations are directly related to the applied force that the bearing is subjected to. As illustrated in Fig. 2 (a), four strain gauges (numbered by ① to ④, each with a length b, forming two pairs) were mounted on the bearing outer ring, with a gap of a between the two gauges within each pair [13]. Each gauge serves as one branch of the bridge circuit. The strain in the bearing outer ring causes the gauge to change resistance. This leads to changes in the voltage across the bridge. The change in strain measured by a pair of gauges when a roller passes over them is shown in Fig. 2 (b). Further investigation of strain gauges for spindle bearings' initial preloads, thermally induced preloads and grinding loads monitoring can be found in [14]. The study shows that by evenly distributing strain gauges around the bearing outer ring, the outputs of the gauges are less sensitive to variations of the external load applied on the spindle housing.



(a) Two pairs of strain gauges mounted on a bearing outer ring



(b) Strain outputs from one pair of gauges

Figure 2. Strain gauges for force sensing (adapted from [13]).

Besides strain gauges, some other nondestructive sensing approaches for bearing force measurement include optical fibers [17] and current carrying plates [18]. In [17], a fiber-optic displacement transducer was mounted in the bearing housing to measure the relative motion between the bearing outer ring and its housing. As illustrated in Fig. 3, such a transducer is operated by using a pair of optical fibers that are adjacent to each other: a transmitting fiber and a receiving fiber. The transmitting fiber is used to carry light from a remote source to an object or target whose displacement or motion is to be measured, e.g. the bearing surface. The receiving fiber receives the light reflected from the object and carries it back to a remote photo sensitive detector [22]. The displacements on the target surface modulate the intensity of the reflected light, which is detected by

the photo detector. By examining the change in the intensity of the transmitting and receiving light, the displacements of the target surface are determined. Using the stiffness (28.8 lb/microinch) of the monitored bearing obtained from its load/deflection chart, the forces causing the motion of bearing outer ring can be determined.

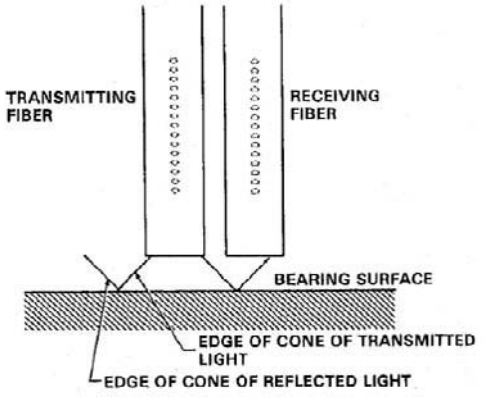
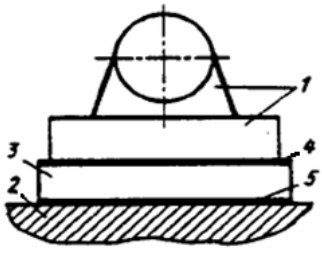


Figure 3. Fiber optic displacement probes for force sensing [17].

In [18], a pair of current-carrying plates was used for bearing load measurement. As illustrated in Fig. 4, one of the two current-carrying plates (object 4) was inserted between the body of the bearing and an elastic element (a rubber layer), and the other (object 5) was inserted between the elastic element and the housing. The plates are electronically insulated from the metal structures and are connected to an electronic monitoring system. By measuring the change in the resistance and reactance of the packet (formed by objects 3, 4 and 5) due to deformation of the elastic element by the external load, the change in the thickness of the packet is evaluated. Subsequently, time-varying forces applied to the bearing are determined. Such a design was experimentally verified on one bearing in a rotary furnace. Through continual monitoring of the loads on the bearing rollers and redesigns to reduce the loads applied, the service life of the rollers were reported to have increased by a factor of 1.5 to 2.



1. Bearing body
2. Bearing housing
3. Elastic element
4. Current-carrying plate
5. Current-carrying plate

Figure 4. Current carrying plates for force sensing [18].

Among the nondestructive sensors investigated for bearing force measurement, strain gauges are the most popular, because they have a wider measurement range (load cells built based on strain gauges can measure up to 500k lbs force) [23] and are easily miniaturized in size to fit in a bearing structure.

3.2. *Vibration Sensing*

Bearings vibrate because of their time-varying compliance rooted in the design. For undamaged, "healthy" bearings, such vibrations are relatively low in amplitudes. As the bearing condition deteriorates, vibrations increase in amplitude. Illustrated in Fig. 5 is the history of the peak acceleration values measured on a ball bearing [24].

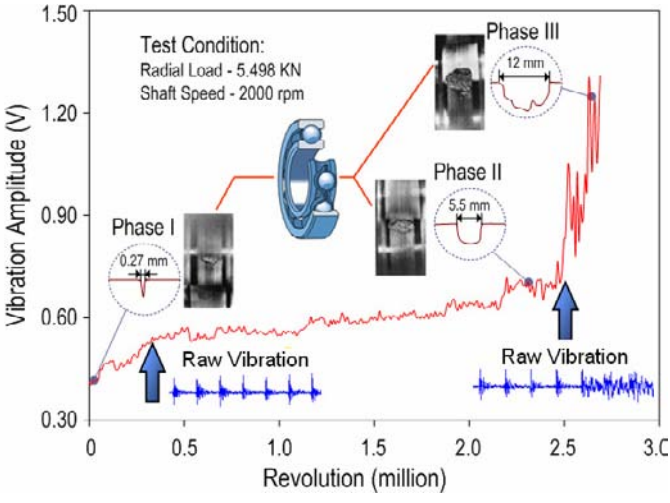


Figure 5. Vibration amplitude increase during a bearing life time (adapted from [24]).

The initial defect is a 0.27 mm groove cut across the bearing outer raceway. During the course of the test, the size of the defect has grown to approximately 5.5 mm, after the bearing has run for 2.3 million revolutions. Upon reaching 2.7 million revolutions, the defect has propagated throughout the entire raceway and rendered the bearing practically non-functional. Three defect propagation phases (I - III) are identified for the bearing life, each associated with a specific defect size.

Besides the increase in amplitude, the number of frequency components contained in the vibration signals also increases with the deterioration of bearing condition. For a healthy bearing, vibration is mainly in the low frequency range (up to 3 times of bearing rotating frequency). Such vibrations are generally

related to bearing misalignment, which is 2 times of the bearing rotating frequency. Comparatively, when a bearing is defective, components in the middle and high frequency range (from 500 Hz to 25 kHz) will start appearing, e.g., due to the occurrence of ball passing frequency on the inner raceway and its harmonics [25], [26].

Vibration can be measured in terms of displacement, velocity and acceleration. The commonly used sensors include displacement probes [27], [28], velocity transducers [29], [30] and accelerometers [31], [32]. These are introduced in the following sections.

A. *Displacement Probe*

Displacement probes are non-contact measuring devices. The physical principles that they are built upon differ widely, from eddy current [33] over fiber optics [34] to capacitance [35]. Of the various principles, eddy current is most commonly used, as such probes have high resolution (on the order of nanometer) and are highly resistant to nonmetallic environmental contaminants such as dust, cutting fluids, and oil [36]. They operate by detecting changes in the sensor coil impedance, which varies as a result of the changing distance between the sensor and the surface of the monitored conductive material [33], as illustrated in Fig. 6. A radiofrequency (RF) oscillator is frequently used as part of the interface circuit to read the effective impedance (Z_{eff}) change in the sensor coil. The diameter of practical eddy current probes varies from a few millimeters to a meter [37]. They have a sensing range from 0.5 mm to 100 mm and a resolution range from 0.2 nm to 3 μm. A bandwidth of 100 kHz is readily achievable [38].

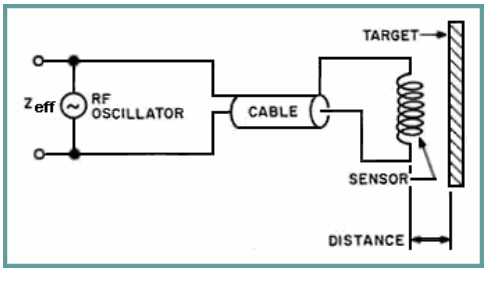
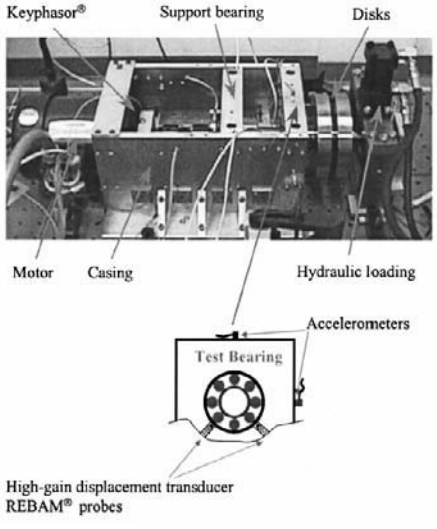


Figure 6. Eddy current probe sensing principle [33].

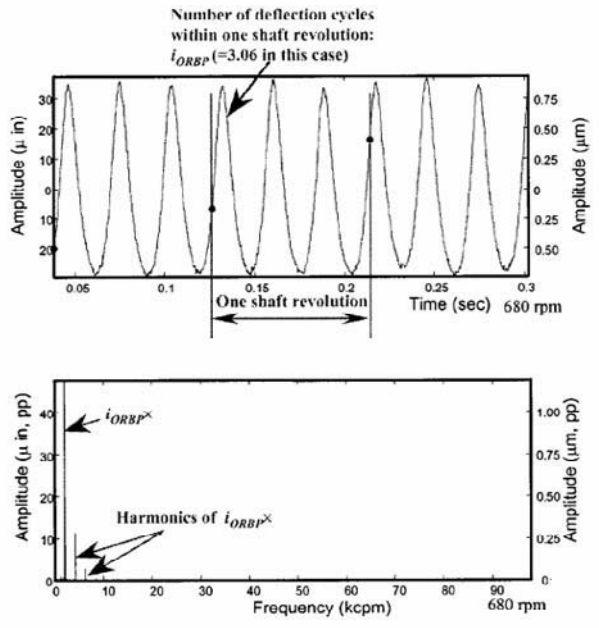
Two eddy current probes (REBAM® probes) with a gain of 78.7 mV/μm and 7.87 mV/μm, respectively, have been investigated in [28] for bearing defect detection. The experimental test rig is shown in Fig. 7 (a), where two eddy current probes were installed to the left and right side of the bearing bottom section, 45° apart from the middle line.

To compare the effectiveness of these probes with accelerometers for bearing condition monitoring, two accelerometers were installed on the bearing housing, measuring vibrations in the vertical (top) and horizontal (right) directions, respectively. The measurement results from the left probe and the accelerometer that was installed on top of the housing to measure the vertical acceleration are illustrated in Fig. 7 (b) and (c), where the subscript *ORBP/IRBP* represents the ratio of the outer/inner race ball passing frequency to shaft speed. It is seen that the probe provided a clearer signal than the accelerometers did. The reason can be attributed to the fact that the accelerometers were installed on the housing and are therefore subject to structural-borne vibrations from the structures, e.g., rotor and foundation. Besides eddy current, optical fibers have also been applied for displacement measurement in nondestructive bearing condition monitoring [39], [40]. As shown in Fig. 8, two fiber optic probes: "bearing probe" and "shaft probe", were used for bearing defect monitoring. The "bearing probe" detects displacements in the bearing, and the "shaft probe" measures the rotation speed and provides a reference signal for the phase relationship. They operate under the same principle as the optical sensor illustrated in Fig. 3. Based on the "bearing probe", a defect factor was defined using the ratio between the peak and root mean square values of the displacement signal. For a healthy bearing, the factor has a value of unity. Values above 1.5 indicate defects that can shorten the bearing life. The experimental results showed that this monitoring technique can detect an inner raceway defect nine weeks earlier than using other techniques based on vibration levels.

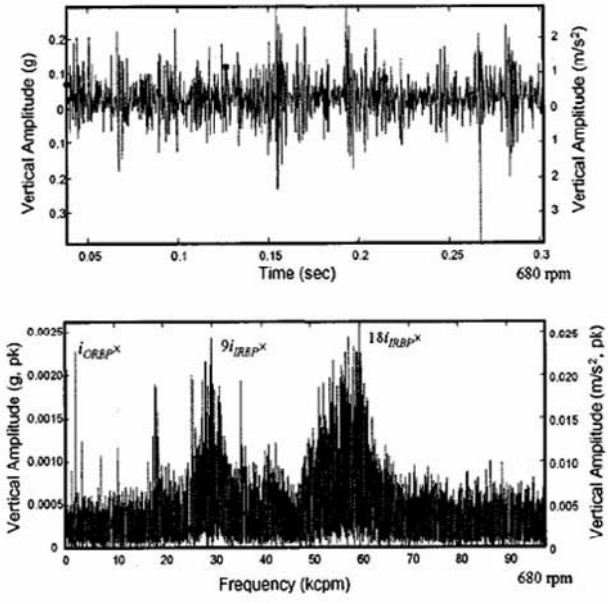


(a) High-gain eddy current probes test rig

Figure 7. Eddy current probes for displacement monitoring (adapted from [28]).



(b) Outputs from left probe: time (up) and frequency domain (down)



(c) Outputs from the top accelerometer: time (top) and frequency domain (down)

Figure 7. (*Continued*)

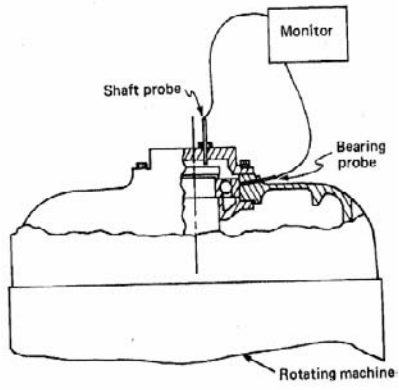
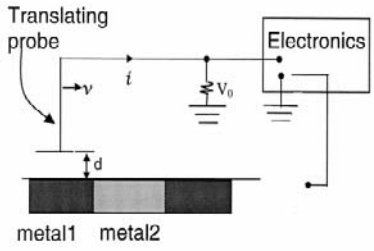


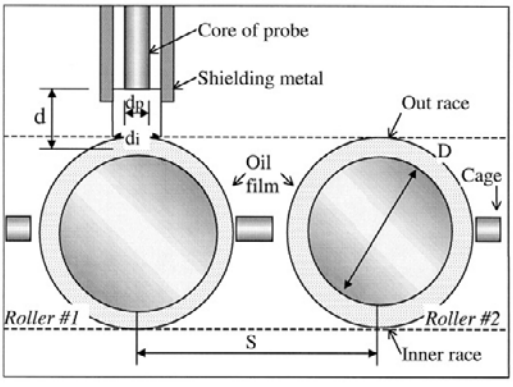
Figure 8. Fiber optic probes for displacement measurement [39].

Another technique for displacement measurement is based on capacitance measurement [41], [42]. A Kevin probe operating in a non-vibrating mode was investigated, which is a capacitive sensor that measures the contact potential difference between two dissimilar metals. Its non-vibrating operation mode uses the relative linear motion between the probe and the surface-of-interest, and generates a current that varies according to the physical and chemical conditions of the surface. The operating principle of such a probe is illustrated in Fig. 9 (a), where one plate of the capacitor is positioned above a moving surface (e.g., bearing rolling element), and the other plate is electronically connected to the moving surface. A current is produced as the probe moves over the surface that has physical or chemical irregularities. The detailed position of a probe relative to a bearing roller is shown in Fig. 9 (b). The spacing between the probe head and the roller surface is controlled with a resolution of 0.01 mm. Such probes were successfully applied to the diagnosis of a seeded defect (approximately 0.25 mm by 0.25 mm) on a roller and a missing roller [41], as well as for measuring rolling element skews (from 0.1 degree to 0.6 degree) [42].

Of the three sensing mechanisms for structural displacement measurements, the eddy current probes have the highest resolution (up to 0.2 nm [38]), but they only work on conductive materials. Fiber optic probes allow for measurement of the largest gaps (15 mm to 1,000 mm) with relatively low resolution (2 μm). Capacitance probes have good resolution (2.5 nm [43]), but they are vulnerable to contamination between the sensor plate and the target. For a given bearing monitoring task requiring a displacement probe. Thus choosing appropriate type of probe is critical in terms of target material, measuring range, and environmental influences [36].



(a) Operating principle of a capacitive probe



(b) Location of the capacitive probes

Figure 9. Capacitive probes for displacement measurement [41].

B. *Velocity Transducer*

Before accelerometers became popular, velocity transducers had been widely applied for vibration measurement, since they could produce relatively high voltage output with good sensitivity (up to 500 mV/in/sec) [44], [45]. The two types of velocity transducers most often used for bearing condition monitoring are magnetic and laser-based.

Magnetic velocity transducer makes use of a varying magnetic field to induce electric current in a wire [19]. This is based on the Faraday's law, which states that the induced voltage is equal to the rate at which the magnetic flux, through the circuit, changes. One way to achieve such a changing magnetic flux is to adjust the effective area exposed to the magnetic field, which is the approach adopted in a magnetic velocity transducer, implemented by a relative movement between the coils and the magnets. As illustrated in Fig. 10 [44], a spring-mass system is suspended in the magnetic field, formed by a pair of magnets arranged above and below the mass. When the magnet and the housing vibrate with the bearing structure at a frequency above the natural frequency of the spring-mass

system, relative movements between the system and the housing are generated [46]. The induced electrical output is proportional to the velocity of the coil moving through the magnetic field.The frequency response of such velocity transducers is limited [44], typically between 10 Hz and 2,000 Hz [46].

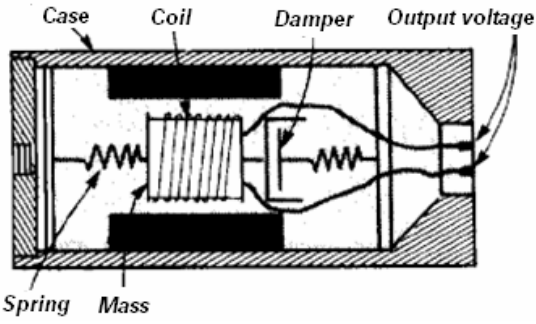


Figure 10. Magnetic velocity transducer [44].

Another type of velocity transducers are the laser vibrometers or laser velocimeters [46]. Their operation is based on the Doppler effect, which describes the change in frequency of a light that is perceived by an observer moving relative to the source light [47]. Such a change in frequency, normally called Doppler frequency shift, is a function of the velocity of the moving target. By comparing the reflected beam with the original beam, the Doppler frequency shift, which is a measure for the velocity of the vibrating surface, is determined. One such vibrometer was used in [47] for ball surface profile measurement and surface crack detection, and the experimental setup is shown in Fig. 11. Cracks down to 22 μm in width were detected using a 150 x 500 μm wide laser spot size. By decreasing the laser spot size, smaller cracks were detected. The laser vibrometer was also investigated for vibration measurement in other rotating machineries [48], [49].

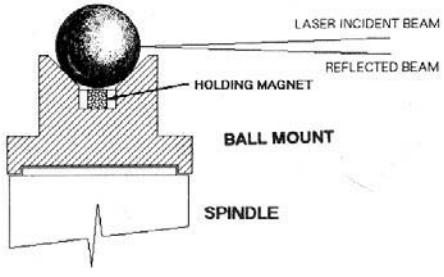


Figure 11. Laser vibrometer in ball surface profiling [47].

As compared with the magnetic velocity transducers, the laser vibrometers are capable of providing high resolution (0.02 μm/s) and frequency range (up to 22 kHz) [50], but their applications are limited because of special requirements on mounting location, sensing direction, and sensing distances.

C. *Accelerometer*

Accelerometers are sensors that measure vibration in terms of acceleration. For nondestructive bearing condition monitoring, they have been more widely applied than displacement probes and velocity transducers. This is because they are generally easy to mount, low in cost, and available in a wide variety of sizes, shapes (cylinder or rectangular), and performance characteristics (measurement range up to 2000 g, bandwidth up to 60 kHz, and sensitivity up to 1,000 mV/g) [51]. Another desirable property of accelerometers is that displacement and velocity can be readily obtained by integrating the measured accelerations.

An accelerometer is generally composed of a proof mass suspended by a spring structure in a case. When the case is accelerated, the proof mass is accelerated by the force transmitted through the spring structure. This force is converted into an electrical signal proportional to the acceleration [46]. The three popular sensing mechanisms used in accelerometers are: piezoelectric [46], piezoresistive [52], and capacitive effects [53]. They can be found in both Micro-Electro-Mechanical Systems (MEMS) accelerometers and non-MEMS accelerometers (e.g., strain gauge-based accelerometers) [54]. In MEMS accelerometers, capacitive effect has been more frequently explored for sensor design than other effects such as piezoelectric [55], piezoresistive [56], [57], tunneling current, and heat transfer [58] due to the relative simplicity in sensor configuration, low power consumption, and good stability over a wide temperature range [53].

Piezoelectric effect refers to the generation of electric charges by a crystalline material upon being subjected to a stress [19]. It is described by the piezoelectric coefficients, d_{mn}, which expresses the charge output along the m direction by the applied unit force in the n direction [54]. Depending on the different combination of force input and charge output directions, the piezoelectric material can be configured to measure longitudinal, transversal, and shear accelerations. The values of the pieozoelectric coefficients vary with materials. For example, a commonly used piezoelectric coefficient in the longitudinal mode is d_{33}, where the subscript 3 indicates the direction that is parallel to the polarization direction of the piezoelectric material. It ranges from 2 pC/N for quartz crystal [59], to 215 ~ 620 pC/N for lead zirconate titanate (PZT) [60], and to 1,400 ~ 3,500 pC/N

for lead magnesium niobate-lead titanate solid (PMN-PT) [61]. These coefficients are also temperature dependent. For some materials, e.g., quartz crystal, the coefficient may drop with a slope of -0.016%/ºC [19] as temperature increases. Popular interface circuits used with piezoelectric accelerometers are charge amplifiers, which convert electrical charge output from the accelerometer to voltage readings. Compared to voltage amplifiers, the charge amplifiers have a sensitivity and time constant that are independent of the stray capacitance of the connecting cables. As a result, long cables will not result in a reduced sensitivity or a variation in their frequency response [54]. The amplifier can be integrated in the sensor package, as shown in Fig. 12 [46]. The final output from this sensor is voltage, which is obtained by letting the charge output from the quartz plates go through the integrated electronics (impedance converter).

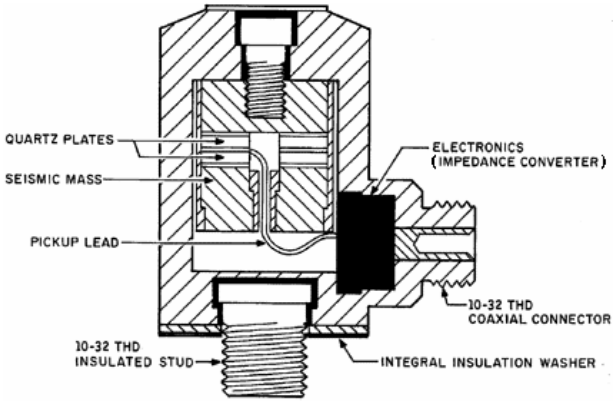


Figure 12. A piezoelectric accelerometer with voltage output [46].

Piezoresistive effect has been explored in the design of strain gauges, as introduced previously, under force sensing. As shown in Fig. 13, the acceleration output of the sensor is obtained from the force detected by the pair of strain gauges. The overload stop prevents possible damage to the gauges.

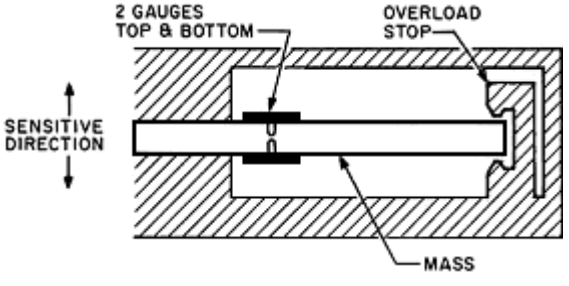


Figure 13. A strain gauge-based accelerometer [46].

Capacitive accelerometers sense the change in capacitance between a proof mass and a fixed conductive electrode, which is separated from the proof mass by a narrow gap. For non-MEMS capacitive accelerometers, their structure is normally composed of two fixed plates sandwiching a diaphragm, as shown in Fig. 14 [62]. Such a structure creates two capacitors, each with an individual fixed plate and sharing the diaphragm as a movable plate. These accelerometers are inherently insensitive to strain variations caused by the mounted structure and transverse acceleration effects. They are able to conduct low-amplitude (as low as 30 μg) and low-frequency (down to DC) acceleration measurement.

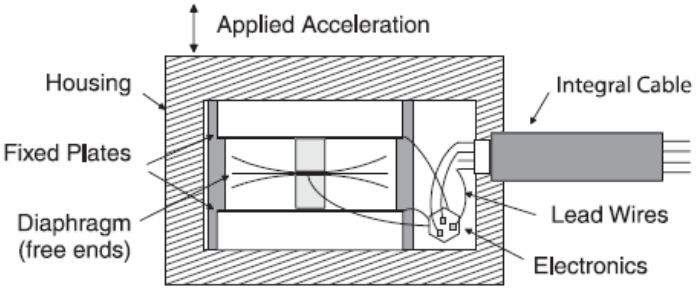
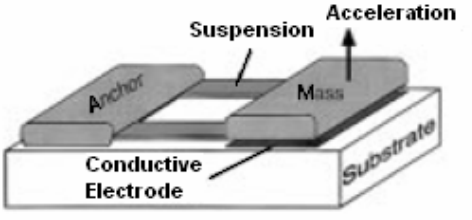
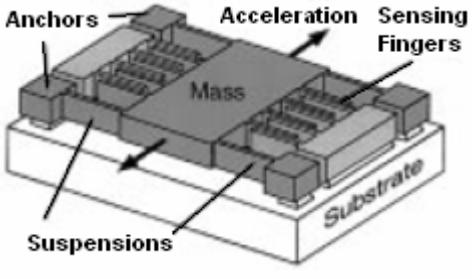


Figure 14. A non-MEMS capacitive accelerometer [62].

MEMS capacitive accelerometers have vertical or lateral structures [58], as shown in Fig. 15.



(a) Vertical structure



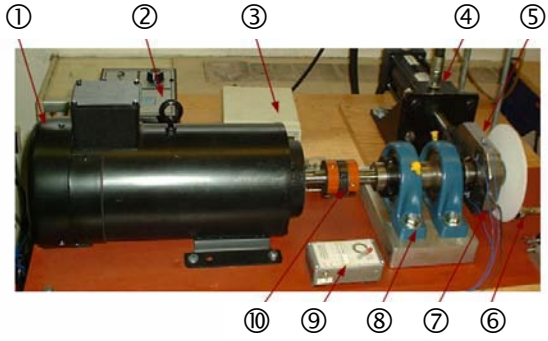
(b) Lateral structure

Figure 15. MEMS capacitive accelerometer (adapted from [58]).

In the vertical design in Fig. 15 (a), the proof mass is separated by a narrow air gap from a fixed plate, forming a parallel plate sensing capacitor. The proof mass moves in the direction perpendicular to its plane and changes the air gap. In the lateral structure, Fig. 15 (b), a number of moving sensing fingers are attached to the proof mass, and the sensing capacitance is formed between these and the fixed fingers parallel to them. MEMS capacitive accelerometers are advantageous in terms of high sensitivity (2 pF/g), low temperature dependency (less than 30 μg/ºC), low power dissipation (0.5 mW), and simple structure [58].

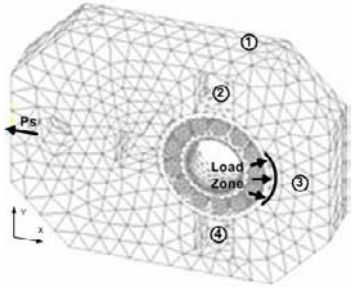
Accelerometers have been applied for bearing condition monitoring [31], [32], [63]-[65]. As illustrated in Fig. [16], four accelerometers were placed in a test rig for detecting bearing outer raceway defect. The test bearing was seeded with a 0.1 mm defect in the outer raceway. Preload (P_s) is applied to the test bearing by a hydraulic system, which resulted in a load zone on the bearing outer raceway as labeled in Fig. 16 (b). The extracted ball passing frequency outer raceways (BPFOs) at the four sensing locations are shown in Fig. 16 (c). It is seen that the BPFO measured by the accelerometers at locations 2 and 3 is about 1.5 to 2.5 times stronger than other frequency components. The BPFO at sensor location 1 is about 1.0 to 1.8 times weaker than that at locations 2 and 3, but comparatively still 25% stronger than that at location 4, where no distinct peak exists at BPFO. This indicates that the bearing outer raceway defect could not be detected. The defect is thus more detectable by placing sensors at locations 2 and 3 than locations 1 and 4. These results demonstrate the necessity of systematical selection of sensing locations when conducting NDT.

Besides vibration sensing techniques discussed above, other methods have also been investigated for vibration measurement in bearing condition monitoring. One such example is shock wave sensors [66] for acceleration measurement. It uses the piezoelectric effect but does not contain a proof mass. Impulsive impact from the rolling element-raceway defect interactions generates shock waves that compress the sensors and induce a microscopic deformation within the sensor, in the form of mechanical strain. Because of the strain development, an electrical charge is produced that is directly proportional to the magnitude of the compressive wave. Such a charge generation mechanism is different from that of an accelerometer, where a proof mass physically compresses the piezoceramic plate placed beneath it, when subject to an accelerated motion. With a customized charge amplifier, a sensitivity of 11 mV/g and bandwidth up to 5 kHz were obtained for the shock wave sensor. The effectiveness of the shock wave sensing technique was experimentally verified for the detection of unbalance, misalignment, and common defects on the bearing inner and outer raceways.

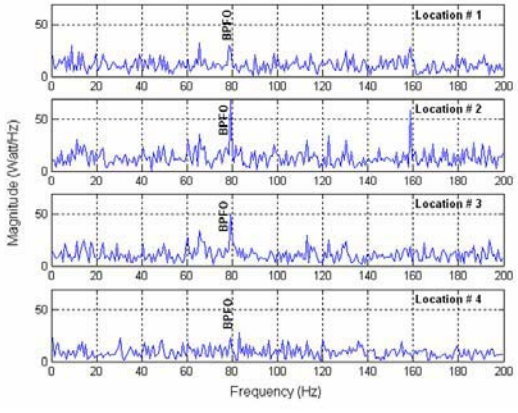


1. DC Motor
2. Motor Speed Controller
3. RPM Recorder
4. Hydraulic Cylinder
5. Accelerometer
6. Optic Encoder
7. Bearing with Housing Plate
8. Pillow Blocks
9. Wireless Receiver
10. Shaft and Coupler

(a) Experimental test rig



(b) Experimental sensor locations



(c) Defect characteristic frequency at four sensor locations

Figure 16. Accelerometers used in bearing condition monitoring [65].

The various types of force and vibration sensors discussed above provide the tools to obtain raw force and vibration data (i.e. displacement, velocity and acceleration) from the bearing being monitored. These data are subsequently processed by various signal processing techniques, such as Statistical Analysis [67], [68], Spectral Analysis [69], [70], Enveloping Analysis [71], [72], Wavelet Transform [73]-[76], Hilbert-Huang Transform [77], and Combined Wavelet and Spectral Analysis [78]. The outputs from these signal processing techniques are often referred to as features, such as amplitude of vibration at bearing defect characteristic frequencies, complexity values [79], and approximate entropy [80], and can be used for assessing bearing condition. The features containing time domain information are useful for assessing the overall condition of the bearings being monitored, whereas features containing frequency information are useful for both overall condition evaluation and specific defect diagnosis, since they indicate which part of a bearing structure went wrong. Four representative bearing defective frequencies identified in [81] include: 1) Ball passing frequency at the inner raceway (BPFI); 2) Ball passing frequency at the outer raceway (BPFO); 3) Ball spinning frequency (BSF); and 4) Fundamental train frequency (FTF). For a defective bearing, they correspond to the defects located in the inner raceway, outer raceway, rolling element, and cage, respectively. For common bearing defective conditions, such as unbalance, misalignment, and wear in bearing raceway, their corresponding frequency characteristics are summarized in [82].

4. Acoustic Testing

Acoustic NDT techniques are presented in this section through discussion on sensors for measuring both audible sound waves and acoustic emissions (AEs). Audible sound waves are generated when vibrations propagate in a bearing structure, and relate to the alternate physical compression and expansion of medium (solids, fluids, and gases) [19]. They are detectable by human ear, in the frequency range of 20 Hz to 15 kHz. Acoustic emissions are defined as transient elastic waves generated from a rapid release of strain energy caused by a deformation or damage within or on the surface of a material [83]. These emissions cover a wide frequency range, from 20 kHz up to 2 MHz, and have two types of time domain waveforms: burst and continuous. A continuous type AE refers to a waveform where transient bursts are not discernable. Both types are seen in rotating machine components, such as bearing, where background

noise is of the continuous type and typical breakage of material is of the burst type. The sensors commonly used for monitoring sound waves and AEs are microphones [84], [85] and acoustic emission sensors [86]-[88], respectively. In the following sections, representative sensing principles and applications are introduced.

4.1. *Microphone*

Microphones are essentially pressure sensors, which convert pressure exerted on a sensitive element into an electric signal [19]. The difference between a microphone and a pressure sensor is that the stimuli for a microphone are broad-band inputs. Besides, a microphone does not measure constant or very slow changing pressure. Various mechanisms can be used to design a microphone, such as piezoresistive effect, capacitance, fiber-optics, and piezoelectric effect. The most popular type is the condenser microphone [89]. These microphones generally have high sensitivity (50 mV/Pa), wide linearity range (4 Hz to 20 kHz), and are insensitive to environmental temperature (0.01 dB/ºC), pressure (0.001 dB/mbar), and humidity. The configuration of a condenser type microphone is shown in Fig. 17 [90]. The capacitor is formed between the diaphragm and the fixed electrode. Sound waves pass through the cloth and reach the diaphragm, whose shape and subsequently, the capacitance, is altered. Such capacitance changes are detected by an interface circuit that may contain a field effect transistor (FET) as a pre-amplifier. With the advances in MEMS technology, micromachined microphones have been developed in recent years [91], [92].

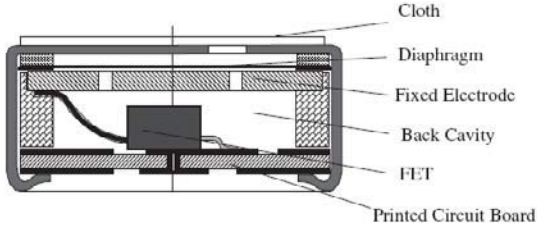


Figure 17. Structure of a condenser microphone [90].

For bearing condition monitoring, microphones have been mainly used for sound pressure [85], [93] and sound intensity [84] measurement. Sound pressure is referred to the difference between the actual pressure at any point in the field

of a sound wave and the average pressure at that point. While sound pressure is a scalar quantity, sound intensity is a vector which describes the amount and direction of time-averaged flow of acoustic power per unit area at a given position [54]. In [85], a condenser microphone was used for sound pressure measurement, as shown in Fig. 18. Bearing defects with sizes ranging from 1.40 to 2.40 mm in width and from 0.44 to 1.50mm in depth were tested. The experiment was conducted under 1.5 kN radial load with speed varying in 500-3,500 rpm. Statistical parameters, such as kurtosis and crest factor, were extracted for bearing condition evaluation. It showed that these statistical parameters from a defective bearing are higher than those obtained from a healthy bearing.

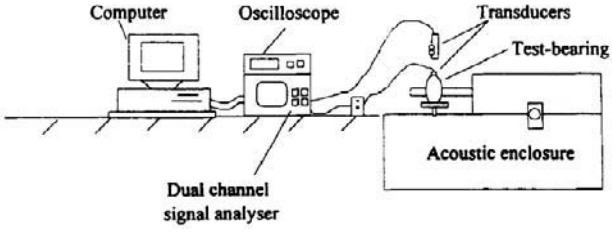
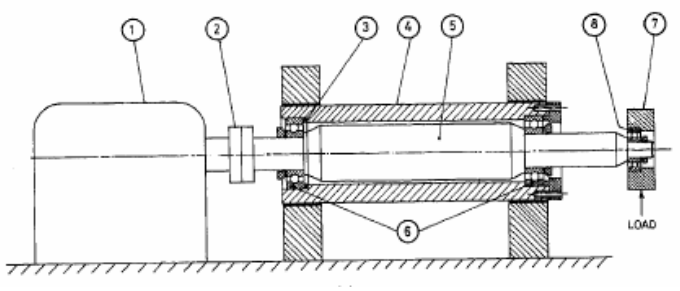


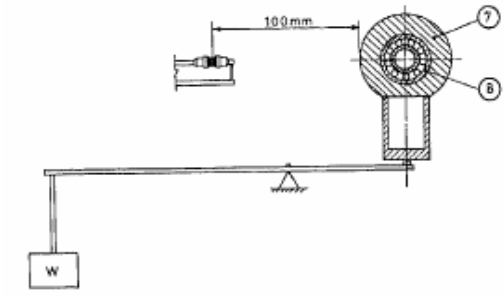
Figure 18. Sound pressure measurement for bearing monitoring [85].

In [84], sound intensity was calculated as the time-averaged product of instantaneous acoustic pressure and particle velocity, which was obtained by measuring the pressure gradient with two closely spaced microphones (the arrangement is often called "two-microphone sound intensity probe") [89]. A probe composed of 12.7 mm phase-matched condenser microphones separated by a 12 mm spacer was used in [84]. The experiments were conducted based on the test rig as shown in Fig. 19 (a), where the "sound intensity probe" was placed on one side of the test bearing housing with the distance between the middle point of the probe and the edge of the test bearing housing of 100 mm, (Fig. 19 (b). The probe was arranged at about the same level as the middle line of the shaft. The frequency range from 250 Hz to 6.4 kHz was used for obtaining the overall levels of sound-intensity. The sound intensity levels measured for varying defect diameter and depth are shown in Fig. 19 (c) and (d), respectively. Figure 19 (c) indicates poor or no detection for a defect diameter less than 200 μm for defects in the outer raceway and ball. The sound intensity levels increase substantially when the defect diameter and depth are greater than 100 and 250 μm, as shown in Fig. 19(c) and 19(d).

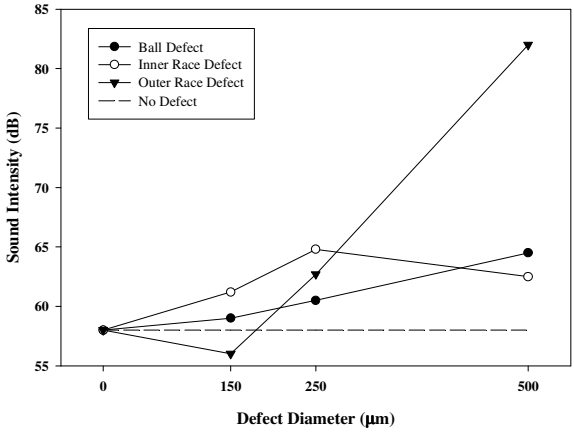


1. DC Motor
2. Flexible coupling
3. Rubber 'O' ring
4. Outer casing
5. Shaft
6. Support bearing
7. Test bearing housing
8. Test bearing

(a) Experimental test rig

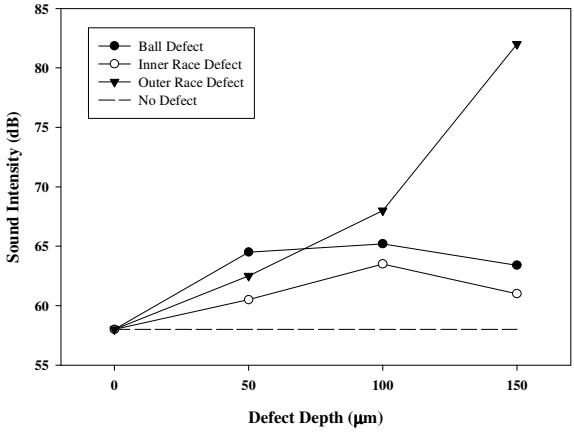


(b) Sound intensity probe installation



(c) Overall sound intensity vs. defect diameter (defect depth fixed at150 μm)

Figure 19. Microphones used in bearing condition monitoring (adapted from [84]).



(d) Overall sound intensity vs. defect depth (defect diameter fixed at 500 μm)

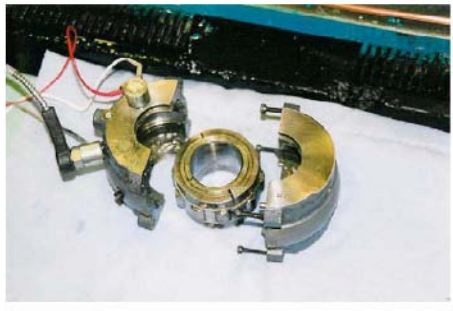
Figure 19. (*Continued*)

4.2. *Acoustic Emission Sensor*

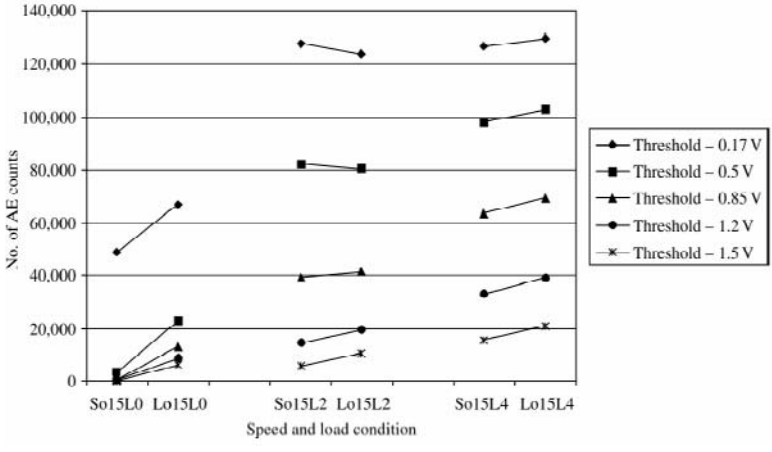
Acoustic emission (AE) sensors are sensing devices for the detection of acoustic emissions generated in stressed materials. Typical operating principle for an AE sensor is the piezoelectric effect. The piezoelectric material is often operated at its resonant mode [94]. When the AE waves strike the sensor surface, the piezoelectric element will oscillate at its resonant frequencies. To achieve accurate measurement, the AE sensor must have a good acoustic contact with the monitored structure. Typically, they are coupled with the test piece using a fluid couplant and are secured with tape, an adhesive bond, or a magnetic hold-down device. It is different from a piezoelectric accelerometer, which operates below the resonant frequencies of the piezoelectric material [54]. With the advancement of MEMS technology, some MEMS AE sensors have been developed [95]. Various parameters extracted from an AE signal can be used in bearing condition monitoring, such as AE counts, peak, root mean square (RMS), kurtosis, crest factor, duration, rising time, and energy [94]-[97]. Spectrum analysis on AE signals [88] has found only limited success, which is primarily due to the broad frequencies associated with the sources of generation of AE signals in a bearing structure [83].

In [86], a wideband AE sensor with frequency range from 100 kHz to 1 MHz was investigated for a radially loaded bearing. The AE counts, peak, RMS, and energy extracted from the AE signals are used for evaluating the bearing condition. The energy is computed by trapezoidal numerical integration. The

experiment is conducted based on a specially designed test rig, whose test bearing with installed AE sensor is shown in Fig. 20 (a). The test bearings are seeded with line defects on the inner and outer raceways. The nominal width, depth and length of the defect is measured at 1 mm, 75 μm and 5mm for a "small defect", while the "large defect" has a length of approximately 15mm. For each rotational speed tested (600, 1,500, and 3,000 rpm) three load cases were considered: 0, 2.4 and 4.8 kN. As shown in Fig. 20 (b), apart from two cases (speed 1500 rpm, load 2.4 kN, trigger level 0.17 V and 0.5 V), for most outer race defects there is an increase in AE counts with increasing defect size (So-small defect, Lo-large defect). It was concluded in [86] that AE is more effective for outer raceway defect detection.



(a) AE sensor and test bearing [98]



(b) AE counts at various thresholds vs. speed and load conditions [86]

Figure 20. Acoustic emission sensors used in bearing condition monitoring .

An alternative sensing mechanism using piezoelectric effect in non-resonant mode was introduced for AE sensing in [88]. The sensor design consists of a conical element bonded to a backing material and allows the sensor to act as a pointlike receiver over the frequency range of20 kHz to 2 MHz. Defects as small as 0.076 mm were successfully detected. Comparatively, accelerometer-based monitoring could not detect defects until they were larger than 1.651 mm in width and 0.127 mm in depth.

To compare the performance of sound, AE and vibration measurement techniques, a study was conducted in [99], where vibrations were measured by an accelerometer, and sound intensity and pressure were measured using the probe shown in Fig. 19 (b). Bearings with no defect and a large defect (500 μm in diameter and 150 μm in depth) were tested under varying loads (200 – 1000 N) and speeds (100 – 1500 rpm). The study showed that sound is a good measure for the overall bearing condition, whereas AE reflects better on crack initiation and propagation. Such findings were also supported by reports from other investigations [88], [98]. Generally speaking, AE signals have a lower signal-to-noise ratio, and establishing reliable calibration procedures to quantitatively correlate defect severity levels with the root causes of the defect has remained a challenging topic that continues to inspire researchers to search for new techniques to advance the science base of bearing condition monitoring and health diagnosis [94].

5. Conclusions

Nondestructive Testing techniques have played a critical role in diagnosing the health conditions of rolling bearings. In this chapter, commonly used NDT techniques based on mechanical and acoustic sensing are reviewed, and the related sensing mechanisms and sensor installation scenarios are described. Vibration has shown to be an informative and convenient variable for revealing bearing conditions. In acoustic measurements, acoustic emission signals have shown to be more effective as compared to audible sound signals in revealing bearing defects at the initial stage. Synergistically combining various sensing techniques may improve sensing accuracy and robustness. Recent advancement of wireless sensor networks [100] has provided the platform for acquiring various types of sensing data in real time, and relaying the information wirelessly to a remote central data processing station, for distributed sensing and improved decision making. In addition to rolling bearings, the NDT techniques presented in this chapter can also be extended to the condition monitoring and health diagnosis of other machines commonly used in manufacturing.

Acknowledgement

This work has been supported by the National Science Foundation under grants DMI-9624353 and DMI-0218161. Experimental work was supported by the SKF and Timken companies.

References

[1] T. A. Harris, *Rolling Bearing Analysis (3rd Edition)*. New York: John Wiley & Sons, 1991.

[2] Y. Choi and C. R. Liu, "Spall progression life model for rolling contact verified by finish hard machined surfaces", *Wear*, vol. 262, no. 1-2, pp. 24-35, 2007.

[3] N. Tandon, G. S. Yadava, and K. M. Ramakrishna, "A comparison of some condition monitoring techniques for the detection of defect in induction motor ball bearings", *Mechanical Systems and Signal Processing*, vol. 21, no. 1, pp. 244-256, 2007.

[4] R. Gao, R. Yan, S. Sheng, and L. Zhang, "Sensor placement and signal processing for bearing condition monitoring", in *L. Wang and R. Gao (Eds)*, *Condition Monitoring and Control for Intelligent Manufacturing*, pp. 167-191. Berlin: Springer, 2006.

[5] C. Hellier, *Handbook of Nondestructive Evaluation*. New York: McGraw-Hill, 2001.

[6] The American Society for Nondestructive Testing, *Introduction to Nondestructive Testing*. [Online]. Available: http://www.asnt.org/ndt/primer5.htm.

[7] A. Davies, P. W. Prickett and R. I. Grosvenor, "Future developments in condition monitoring techniques and systems", in *A. Davies (Ed), Handbook of Condition Monitoring Techniques and Methodology*, pp. 541-560. London, UK: Chapman & Hall, 1998.

[8] C. S. Sunnersjo, "Varying compliance vibrations of rolling bearings", *Journal of Sound and Vibration*, vol. 58, no. 3, pp. 363-373, 1978.

[9] A. V. Barkov and N. A. Barkova, "The artificial intelligence systems for machine condition monitoring and diagnosis by vibration". [Online]. Available: http://www.vibrotek.com/articles/intelect-eng/index.htm

[10] J. Zhang, B. W. Drinkwater, and R. S. Dwyer-Joyce, "Monitoring of lubricant film failure in a ball bearing using ultrasound", *Journal of Tribology*, vol. 128, no. 3, pp. 612-618, 2006.

[11] M. J. Devaney and L. Eren, "Detecting motor bearing faults: Monitoring an induction motor's current and detecting bearing failure", *IEEE Instrumentation and Measurement Magazine*, vol. 7, no. 4, pp. 30-35, 2004.

[12] B. T. Holm-Hansen and R. Gao, "Vibration Analysis of a Sensor-integrated Ball Bearing", *ASME Journal of Vibration and Acoustics*, vol. 122, no. 4, pp. 384-392, 2000.

[13] J. F. Tu, "Strain field analysis and sensor design for monitoring machine tool spindle bearing force", *International Journals of Machine Tools and Manufacture,* vol. 36, no. 2, pp. 203-216, 1996.

[14] J. F. Tu and J. G. Katter, "Bearing force monitoring in a three-shift production environment", *Tribology Transactions*, vol. 39, no. 1, pp. 201-207, 1996.

[15] H. Gloeckner, "Force-monitoring bearings for machine tool spindles", *Ball Bearing Journal*, pp. 19-25, 1985.

[16] J.-S. Chen and K.-W. Chen, "Bearing load analysis and control of a motorized high speed spindle", *International Journal of Machine Tools and Manufacture*, vol. 45, no. 12-13, pp. 1487-1493, 2005.

[17] G. J. Phillips, "Using the ball bearing as a force transducer", *Experimental Techniques*, vol. 12, no. 12, pp. 32-34, 1988.

[18] V. I. Zinchenko, "Method of monitoring the loads transmitted to the bearing of a rotary furnace", *Metallurgist (English Translation of Metallurg)*, vol. 32, no. 7-8, pp. 263, 1989.

[19] J. Fraden, Handbook of Modern Sensors, Physics, Designs, and Applications (2nd Edition). New York: Springer-Verlag Inc., 1996.

[20] J. R. Gyorki, "Strain sensor basics and signal conditioning tips", *Sensors*, vol. 22, no. 6, pp. 18-23, 2005.

[21] Efunda Engineering Fundamentals, *Strain Gauge Selection.* [Online]. Available: http://www.efunda.com/DesignStandards/sensors/strain_gages/strain_gage_selection.cfm

[22] R. O. Cook and C. W. Hamm, "Fiber optic lever displacement transducer", *Applied Optics*, vol. 18, no. 19, pp. 3230-3241, 1979.

[23] Omega Engineering, Inc., *Introduction to Load Cells*. [Online]. Available: http://www.omega.com/prodinfo/loadcells.html

[24] R. Gao and R. Yan, "Non-stationary signal processing for bearing health monitoring", *International Journal of Manufacturing Research*, vol. 1, no. 1, pp. 18-40, 2006.

[25] J. Sandy, "Monitoring and diagnostics for rolling element bearings", *Sound and Vibration*, vol. 22, no. 6, pp.16-20, 1988.

[26] D. Carter, Rolling element bearing condition testing method and apparatus, U.S. Patent #5,477,730.

[27] P. Y. Kim, "A review of rolling element health monitoring (III): preliminary test results on eddy current proximity transducer technique", *Institution of Mechanical Engineers Conference Publications*, no. C311/84, pp. 119-125, 1984.

[28] J. J. Yu, D. E. Bently, P. Goldman, K. P. Dayton, and B. G. Van Slyke, "Rolling element bearing defect detection and diagnostics using displacement transducers", *Journal of Engineering for Gas Turbines and Power*, vol. 124, no. 3, pp. 517-527, 2002.

[29] H. Ohta and S. Satake, "Vibrations of the all-ceramic ball bearing", *Journal of Tribology*, vol. 124, no. 3, pp. 448-460, 2002.

[30] R. Valliapan and D. K. Lieu, "Defect characterization of roller bearing surfaces with laser Doppler vibrometry", *Precision Engineering*, vol. 14, no. 1, pp. 35-42, 1992.

[31] T. I. Liu, J. H. Singonahalli, and N. R. Iyer, "Detection of roller bearing defects using expert system and fuzzy logic", *Mechanical Systems and Signal Processing*, vol. 10, no. 5, pp. 595-614, 1996.

[32] F. Honarvar and H. R. Martin, "New statistical moments for diagnostics of rolling element bearings", *Journal of Manufacturing Science and Engineering, Transactions of the ASME,* vol. 119, no. 3, pp. 425-432, 1997.

[33] S. D. Welsby and T. Hitz, "True position measurement using eddy current technology", *Sensors*, vol. 14, no. 11, 1997.

[34] S. J. Prosser, "The evolution of proximity, displacement, and position sensing", *Sensors*, vol. 15, no. 4, 1998.

[35] S. J. Prosser, "Capacitive and conductive noncontact measurement", *Sensors*, vol. 20, no. 3, 2003.

[36] B. Duval, "Advances in analog distance sensing", *Sensors*, vol. 21, no. 9, pp. 36-40, 2004.

[37] S. D. Roach, "Designing and building an eddy current position sensor", *Sensors*, vol. 15, no. 9, 1998.

[38] MICRO-EPSILON International, *General Catalog*. [Online]. Available: http://www.micro-epsilon.com/staticcontent/PDF/Prod_EN/Catalog_MICRO-EPSILON_-_Products_-_en.pdf

[39] G. J. Philips, "The fiber optic bearing monitor", in *Proceedings of Instrument Society of America*, vol. 28, pp. 379-395, Las Vegas, NV, USA, 1982.

[40] G. J. Philips, "Monitoring and analysis of rolling element bearings using the figer optic bearing monitor", in *Proceedings of the 33rd International Instrumentation Symposium*, Las Vegas, NV, USA, 1987.

[41] Y. Yang, T. Kurfess, S. Liang, and S. Danyluk, "Application of a specialized capacitance probe in bearing diagnosis", *Wear*, vol. 225-229, no. II, pp. 1215-1221, 1999.

[42] Y. Yang, S. Danyluk, and M. Hoeprich, "Study on rolling element skew measurement in a tapered roller bearing with a specialized capacitance probe", *Journal of Tribology, Transactions of the ASME*, vol. 122, no. 3, pp. 534-538, 2000.

[43] MTI Instruments Inc., *Accumeasure™ System 9000 Brochur*e. [Online]. Available: http://www.mtiinstruments.com/products/accumeasure9000.aspx

[44] J. Reason, "Continuous vibration monitoring moves into diagnosis", *Power*, vol. 131, no. 1, pp. 47-51, 1987.

[45] Sales Technology Inc., *Velocity Transducer Installation.* [Online]. Available: http://www.stiweb.com/appnotes/velocity.htm

[46] J. Wilson, "A practical approach to vibration detection and measurement part 1: physical principles and detection techniques", *Sensors*, vol. 16, no. 2, 1999.

[47] R. Valliapan and D. K. Lieu, "Defect characterization of roller bearing surfaces with laser Doppler vibrometry", *Precision Engineering*, vol. 14, no. 1, pp. 35-42, 1992.

[48] S. J. Rothberg and N. A. Halliwell, "Vibration measurements on rotating machinery using laser Doppler velocimetry", *Journal of Vibration and Acoustics, Transactions of the ASME*, vol. 116, no. 3, pp. 326-331, 1994.

[49] R. A. Cookson and P. Bandyopadhyay, "Fiber-optic laser Doppler probe for vibration analysis of rotating machines", *Journal of Engineering for Power, American Society of Mechanical Engineers*, vol. 102, pp. 607-612, 1980.

[50] Polytech, Inc., *Digital Measuring*. [Online]. Available: http://www.polytec.com/usa/158_827.asp?highlightSubMenu=%20Applications&highlightPopupMenu=Digital%20Measuring

[51] PCB Piezotronics, Inc., *Accelerometer Product Specifications.* [Online]. Available: http://www.pcb.com/products/browse_productlist.asp?RequestType=Filter&CategoryType=Product%20Type&CategoryId=316&CategoryDescription=Accelerometer

[52] J. T. Suminto, "A simple, high performance piezoresistive accelerometer", in *Proceedings of Transducers '91*, pp. 104-107, 1991.

[53] J. Bernstein, "An overview of MEMS inertial sensing technology", *Sensors*, vol. 20, no. 2, 2003.

[54] E. O. Doebelin, Measurement Systems: Application and Design (5th Edition). New York: McGraw Hill, 2004.

[55] L.-P. Wang, R. A. Wolf Jr, Y. Wang, K. K. Deng, L. Zou, R. J. Davis, and S. Trolier-McKinstry, "Design, fabrication, and measurement of high-sensitivity piezoelectric microelectromechanical systems accelerometers", *Journal of Microelectromechanical Systems*, vol. 12, no. 4, pp. 433-439, 2003.

[56] J. P. Lynch, A. Partridge, K. H. Law, T. W. Kenny, A. S. Kiremidjian, and E. Carryer, "Design of piezoresistive MEMS-based accelerometer for integration with wireless sensing unit for structural monitoring", *Journal of Aerospace Engineering*, vol. 16, no. 3, pp. 108-114, 2003.

[57] A. Partridge, J. K. Reynolds, B. W. Chui, E. M. Chow, A. M. Fitzgerald, L. Zhang, N. I. Maluf, and T. W. Kenny, "High-performance planar piezoresistive accelerometer", *Journal of Microelectromechanical Systems*, vol. 9, no. 1, pp. 58-66, 2000.

[58] N. Yazdi, F. Ayazi, and K. Najafi, "Micromachined inertial sensors," *Proceedings of the IEEE*, vol. 86, no. 8, pp. 1640-1658, 1998.

[59] J. W. Waanders, Piezoelectric Ceramics Properties and Applications. Eindhoven, NL: Philips, 1991.

[60] APC International Ltd., *Physical and Piezoelectric Properties of APCI Materials*. [Online]. Available: http://www.americanpiezo.com/materials/apc_properties.html

[61] H. C. Materials Corporation, *Properties of Piezoelectric PMN-PT Crystal*. [Online]. Available: http://www.hcmat.com/Pmn_Properties.html

[62] PCB Piezotronics Inc., *Capacitive accelerometers brochure*. [Online]. Available: http://www.pcb.com/Linked_Documents/Vibration/Cap_Accels_0704.pdf

[63] B. A. Paya, I. I. Esat, and M. N. M. Badi, "Artificial neural network based fault diagnostics of rotating machinery using wavelet transforms as a preprocessor", *Mechanical Systems and Signal Processing*, vol. 11, no. 5, pp. 751-765, 1997.

[64] B. Samanta and K. R. Al-Balushi, "Artificial neural network based fault diagnostics of rolling element bearings using time-domain features", *Mechanical Systems and Signal Processing*, vol. 17, no. 2, pp. 317-328, 2003.

[65] S. Sheng, L. Zhang, and R. X. Gao, "A systematic sensor-placement strategy for enhanced defect detection in rolling bearings", *IEEE Sensors Journal*, vol. 6, no. 5, pp. 1346-1354, 2006.

[66] R. X. Gao, C. Wang, and S. Sheng, "Sensor placement strategy for high quality sensing in machine health monitoring", *Smart Structures and Systems, An International Journal of Mechatronics, Sensors, Monitoring, Control, Diagnostics, & Life Cycle Eng.*, vol. 1, no. 2, pp. 121-140, 2005.

[67] R. B. W. Heng and M. J. M. Nor, "Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition", *Applied Acoustics*, vol. 53, no. 1-3, pp. 211-226, 1998.

[68] X. Niu, L. Zhu, and H. Ding, "New statistical moments for the detection of defects in rolling element bearings", *International Journal of Advanced Manufacturing Technology*, vol. 26, no. 11-12, pp. 1268-1274, 2005.

[69] N. Tandon, "Comparison of some vibration parameters for the condition monitoring of rolling element bearings", *Measurement: Journal of the International Measurement Confederation*, vol. 12, no. 3, pp. 285-289, 1994.

[70] J. P. Dron, L. Rasolofondraibe, C. Couet, and A. Pavan, "Fault detection and monitoring of a ball bearing benchtest and a production machine via autoregressive spectrum analysis", *Journal of Sound and Vibration*, vol. 218, no. 3, pp. 501-525, 1998.

[71] I. M. Howard, "Higher-order spectral techniques for machine vibration condition monitoring", *Proceedings of the Institution of Mechanical Engineers Part G-Journal of Aerospace Engineering*, vol. 211, sec. G4, pp. 211-219, 1997.

[72] D. M. Yang, A. F. Stronach, P. MaCconnell and J. Penman, "Third-order spectral techniques for the diagnosis of motor bearing condition using artificial neural networks", *Mechanical Systems and Signal Processing*, vol. 16, no. 2-3, pp. 391-411, 2002.

[73] N. G. N. Ikolaou and I. A. A. Ntoniadis, "Demodulation of vibration signals generated by defects in rolling element bearings using complex shifted Morlet wavelets", *Mechanical Systems and Signal Processing*, vol. 16, no. 4, pp. 677-694, 2002.

[74] G. Y. Luo, D. Osypiw and M. Irle, "On-line vibration analysis with fast continuous wavelet algorithm for condition monitoring of bearing", *Journal of Vibration and Control*, vol. 9, no. 8, pp. 931-947, 2003.

[75] B. T. Holm-Hansen, R. Gao, and L. Zhang, "Customized wavelet for bearing defect detection", *ASME Journal of Dynamic Systems, Measurement, and Control*, vol. 126, no. 4, pp. 740-745, 2004.

[76] R. Yan and R. Gao, "Harmonic wavelet packet transform: a new approach to machine condition monitoring and health evaluation", *Robotics and Computer Integrated Manufacturing*, vol. 21, pp. 291-301, August, 2005.

[77] R. Yan and R. Gao, "Hilbert-Huang Transform-based vibration signal analysis for machine health monitoring", *IEEE Transactions on Instrumentation and Measurement*, vol. 55, no. 6, pp. 2320-2329, 2006.

[78] C. Wang and R. Gao, "Wavelet transform with spectral post-processing for enhanced feature extraction", *IEEE Transactions on Instrumentation and Measurement,* vol. 52, no. 4, pp. 1296-1301, 2003.

[79] R. Yan and R. Gao, "Complexity as a measure for machine health evaluation", *IEEE Transactions on Instrumentation and Measurement,* vol. 53, no. 4, pp. 1327-1334, 2004.

[80] R. Yan and R. Gao, "Machine health monitoring based on approximate entropy", *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 824-839, 2007.

[81] A. Malhi and R. Gao, "PCA-based feature selection scheme for machine defect classification", *IEEE Transactions on Instrumentation and Measurement,* vol. 53, no. 6, pp. 1517-1525, 2004.

[82] Y. A. Azovtsev, A. V. Barkov, and I. A. Yudin, "Automatic diagnostics and condition prediction of rolling element bearings using enveloping methods". [Online]. Available: http://vibrotek.com/article.php?article=articles/new94vi/index.htm

[83] D. Mba and R. B. Rao, "Development of acoustic emission technology for condition monitoring and diagnosis of rotating machines: bearings, pumps, gearboxes, engines, and rotating structures", *Shock and Vibration Digest*, vol. 38, no. 1, pp. 3-16, 2006.

[84] N. Tandon and B. C. Nakra, "The application of the sound-intensity technique to defect detection in rolling-element bearings", *Applied Acoustics*, vol. 29, no. 3, pp. 207-217, 1990.

[85] R. B. W. Heng and M. J. M. Nor, "Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition", *Applied Acoustics*, vol. 53, no. 1-3, pp. 211-226, 1998.

[86] A. Morhain and D. Mba, "Bearing defect diagnosis and acoustic emission", *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, vol. 217, no. J4, pp. 257-272, 2003.

[87] N. Tandon and B. Nakra, "Defect detection in rolling element bearings by acoustic emission method", *Journal of Acoustic Emission*, vol. 9, no. 1, pp. 25-28, 1990.

[88] M. W. Hawman and W. S. Galinaitis, "Acoustic emission monitoring of rolling element bearings", presented at *IEEE 1988 Ultrasonics Symposium*, Chicago, IL, USA, 1988.

[89] Hewlett Packard Company, Sound Power Measurements Application Note 1230, 1992.

[90] Y. Yasuno and J. Ohga, "Temperature characteristics of electret condenser microphones", *Acoustical Science and Technology*, vol. 27, vol. 4, pp. 216-224, 2006.

[91] W. T. Chee and J. Miao, "Analytical modeling for bulk-micromachined condenser microphones", *Journal of the Acoustical Society of America*, vol. 120, no. 2, pp. 750-761, 2006.

[92] R. Dieme, G. Bosman, T. Nishida, and M. Sheplak, "Sources of excess noise in silicon piezoresistive microphones", *Journal of the Acoustical Society of America*, vol. 119, no. 5, pp. 2710-2720, 2006.

[93] M. S. Carney, J. A. III. Mann, and J. Gagliardi, "Adaptive filtering of sound pressure signals for monitoring machinery in noisy environments", *Applied Acoustics*, vol. 43, no. 4, pp. 333-351, 1994.

[94] A. A. Pollock, "Acoustic emission inspection", in *Metals Handbook (9th Edition),* vol. 17, pp 278-294, 1989.

[95] I. J. Oppenheim, D. W. Greve, D. Ozevin, D. R. Hay, T. R. Hay, S. P. Pessiki, and N. L. Tyson, "Structural tests using a MEMS acoustic emission sensor", in *Proceedings of SPIE - The International Society for Optical Engineering*, v 6174 I, *Smart Structures and Materials 2006 - Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems,* San Diego, CA, United States, 2006.

[96] D. Mba, "Acoustic emissions and monitoring bearing health", *Tribology Transactions,* vol. 46, no. 3, pp. 447-451, 2003.

[97] J. Shiroishi, Y. Li, S. Liang, T. Kurfess, and S. Danyluk, "Bearing condition diagnostics via vibration and acoustic emission measurements", *Journal of Engineering and Applied Science*, vol. 11, no. 5, pp. 693-705, 1997.

[98] A. M. Al-Ghamd and D. Mba, "A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size", *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1537-1571, 2006.

[99] N. Tandon and B. C. Nakra, "Comparison of vibration and acoustic measurement techniques for the condition monitoring of rolling element bearings", *Tribology International*, vol. 25, no. 3, pp. 205-212, 1992.

[100] J. M. Rybak, "Remote condition monitoring using open-system wireless technologies", *Sound and Vibration*, vol. 40, no. 2, pp. 16-20, 2006.

CHAPTER 20

ADVANCED SENSORS FOR HEALTH MONITORING SYSTEMS

Nezih Mrad

Air Vehicles Research Section, Defence R&D Canada
Department of National Defence, National Defence Headquarters
Ottawa, Ontario, K1A 0K2, Canada
Nezih.Mrad@drdc-rddc.gc.ca

Recent advances in sensors, microelectronics, adaptive signal processing and predictive technologies have significantly shaped the fundamental approach to dealing with traditional problems within the aerospace community. Aircraft diagnostics, prognostics and health management (DPHM) is increasingly becoming a main stream approach to aircraft maintenance within an advanced operational autonomic logistics structure. Driven by the requirement for increased safety, reliability, enhanced performance and aircraft availability at reduced cost, key sensor technologies such as electrorheological fluids, shape memory alloys, piezoelectric materials, magnetostrictive and electrostrictive materials, triboluminescent materials, optical fibres, carbon nanotubes, comparative vacuum monitoring and MEMS/NEMS are expected to play a major role in the development of such DPHM systems. Such advanced sensors often referred to as smart sensors are further expected to provide functionality that is not matched by current technologies and nondestructive evaluation techniques within in an on-line *in-situ* environment. Regardless of the extensive number of sensors and sensory systems that could potentially be employed or integrated within a DPHM system, only a selected few illustrating near commercial exploitation are introduced in this chapter. Furthermore, the DPHM and on-line structural health monitoring concepts and its impact on aircraft operations and maintenance is briefly introduced.

1. Introduction

The increasing operational complexity and stringent performance requirements of legacy and emerging air vehicles have made on-line health monitoring and management a critical component of aircraft operational logistics and maintenance programs. From an operator's perspective, the goal of such on-line

health management system is to improve operational safety, to lower life cycle costs, to increase combat readiness and availability of military systems and to improve asset management. It is estimated[1] that about 60% of the overall ownership cost of an aircraft is attributed to Operation and Support (O&S). Most of which is linked to personnel and materials used to support scheduled and unscheduled maintenance actions. Because it is difficult to diagnose faults in some components, such as engine compressors, more than 40% of compressors replaced on aircraft are not defective resulting in unnecessary false repairs. Realizing the significant savings and the increased performance that could result from the development of on-line integrated health management systems, the Joint Strike Fighter (JSF) Program Management Office (PMO) is set to deploy an advanced capability into its new fighter jet program to reduce O&S costs to below 50%. Such advanced capability, known as Diagnostics, Prognostics and Health Management (DPHM), was initially conceptualized during the introduction of the JSF program for engine diagnostics and prognostics.

Traditionally, real-time operating parameters and damage information are captured and compared against both test and field data to assess the potential damage severity and required maintenance actions of vehicle's components. Such global health monitoring approach is well known as Condition-Based Maintenance (CBM). This approach is intended to help manage the frequency and extent of maintenance procedures for various components within a vehicle. The incorporation of CBM with predictive tools is expected to increase the level of effectiveness of on-line health monitoring and management systems' capability. Such capability is known as the PHM capability and is expected to provide the ability to predict and pinpoint potential damage and problem areas in vehicles; thus, reducing or perhaps even eliminating the need for periodic inspections. This capability will not only reduce air vehicles' ownership cost, that is considered to be a critical concern for civilian aircraft operators, but also increase air safety, that is mandated by the Federal Aviation Administration (FAA) (e.g. 80% accident reduction by year 2007.) This adjustment and shift in maintenance philosophy employing DPHM methodology can be reflected in the PHM evolution chart shown in Figure 1.

Faas et al.[2] reported that the ultimate results of a DPHM capability are to be able to reliably predict failures with high probability and high confidence. However, they claim that realistically one hopes that such systems provide tools to integrate with existing systems to break the ambiguity groups and then provide some insight into the predictive side. Furthermore, by deploying such capability, they identified that the specific war fighter benefits include reduced maintenance man-hours, restored aircraft to operational status sooner, reduced consumption of

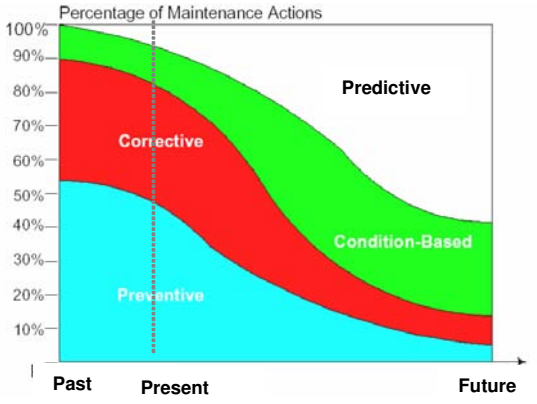


Figure 1. Evolution of maintenance actions – Preventive to Predictive.

spare parts, reduced mission aborts due to system failures, reduced deployment footprint and increased warning time to order spares and invoke proactive repairs. Results reflecting the return on investment (ROI) for the deployment of such on-board system are illustrated in Table 1.[2]

Table 1: ROI for the deployment of on-board on-line monitoring system

	100% On-Board	75% On-Board	50% On-Board	25% On-Board
Return On Investment (ROI)	9.2	11.5	16.5	24.5
Time to Break Even (Years)	10	10	9	7

2. Definition of Diagnostics, Prognostics and Health Management

A Diagnostics, Prognostics and Health Management (DPHM) system is considered as the central repository system of desired information for on-line collaborative environment in aircraft autonomic logistics. DPHM has the capability to determine the state of health of a component to perform its function(s) (Diagnostics). It is also a predictive diagnostics tool that includes determining the remaining life or time span of proper operation of a component (Prognostics). It has the capability to make appropriate decisions about component maintenance based on diagnostics/prognostics information, available resources and operational demand (Health Management).[3] DPHM evolved from a vision of advanced diagnostics that uses advanced sensors to monitor and manage aircraft health, and quite often is referred to as on-line structural health monitoring or advanced on-line nondestructive evaluation.

3. Methodology of Diagnostics, Prognostics and Health Management

Depending on the application and desired outcome, a DPHM system varies in complexity from simple (e.g. detect and alert) to more complex (e.g. detect and advise). Figure 2[4] illustrates a conceptual implementation approach of a PHM system having moderate complexity; whereas, Figure 3[5] illustrates the expected complexity for integrating an aircraft PHM system within an operational autonomic logistics structure.

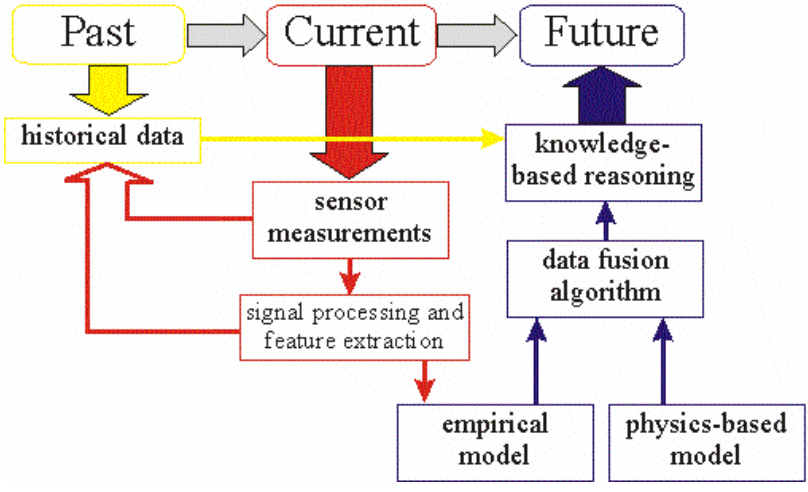


Figure 2. Conceptual structure of a PHM system.

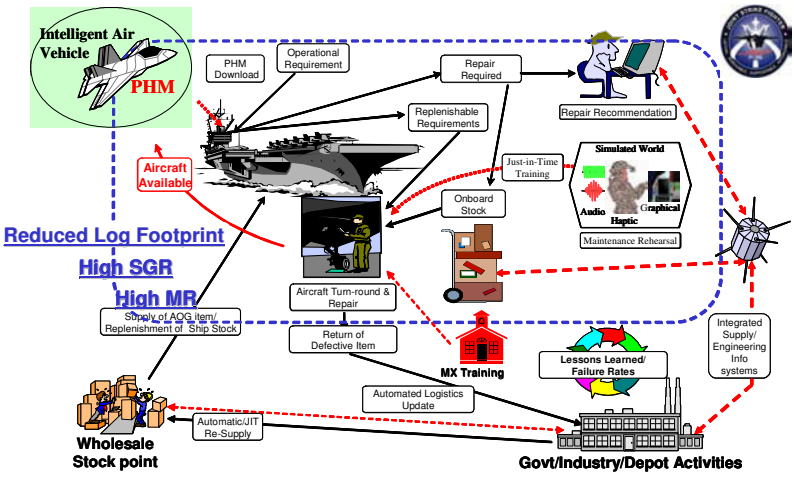


Figure 3. A concept of aircraft autonomic logistics structure.

Regardless of its complexity, four major building blocks constitute the core of the DPHM system's architecture and structure. These blocks are sensor networks, usage and damage monitoring reasoner (diagnostics), life management reasoner (predictive and prognostics), and decisions making and assets management. One possible approach to describe the functioning of such a system is that usage and damage parameters, acquired via wired and wireless sensor network, are transmitted to an on-board data acquisition and signal processing system. The acquired data is developed into information related to damage, environmental and operational histories as well as aircraft usage employing information processing algorithms embedded into the usage and damage monitoring reasoner. This information, when provided to the life management reasoner and through the use of predictive diagnostic and prognostics models, is converted into knowledge about the state of operation and health of the aircraft. This knowledge is then disseminated and transmitted to flight crew, operations and maintenance services, regulatory agencies, and or Original Equipment Manufacturers (OEM) for decision making and asset management.

Generally, it is expected that the critical and perhaps the most significant and limiting factor in the implementation of DPHM systems is the array of sensors and sensor networks. Sensors must be accurate, reliable, robust, small size, lightweight, immune to radio frequency and electromagnetic interferences, easily networked to on-board preprocessing capabilities, capable of withstanding operational and environmental conditions, require no or low power for both passive and active technologies and possess self-monitoring and self-calibrating capabilities. Several sensing technologies and NonDestructive Evaluation (NDE) techniques are currently in use or under development for DPHM applications. The NDE techniques refer to an array of nondestructive evaluation techniques and processes to monitor, probe, measure, and assess material response to internal or external stimuli. The measured response, generally conducted off-line, is related to a desired material property and characteristics reflecting the state of its health. The main NDE methods are:[6] visual inspection, liquid penetrant inspection, magnetic particle inspection, radiographic inspection (X-ray and gamma ray), Eddy current inspection, ultrasonic inspection, and thermographic inspection. Although, each of these methods is dependent on different basic principles in both application and output, repeatability and reproducibility depends significantly on specific understanding and control of several factors including human factors[6]. These techniques do not lend themselves as the best candidates for real-time on-line environment; however, due to their maturity, they are significant tools for off-line components damage detection and assessment.

Advanced sensors often referred to as smart sensors or sensor nodes perform several functions delivered by NDE techniques in a real-time on-line environment with added integrated capabilities such as signal acquisition, processing, analysis and transmission. These highly networked sensors (passive or active) are suitable for large platforms and wide area monitoring and exploit recent development in micro and nano technologies. Regardless of the extensive number of sensors and sensory systems that could potentially be employed or integrated within a DPHM system, only a selected few illustrating near commercial exploitation are introduced in this chapter. These are identified as advanced non-traditional sensors employing advanced concepts and are generally developed for potential integration into existing or emerging advanced systems. These sensors include MEMS sensors,[7] fibre optic sensors,[8] piezoelectric sensors,[9] piezoelectric wafer active sensor,[10] triboluminescent sensors,[11] SMART layer sensor networks,[12] nitinol fibre sensors,[13] carbon nanotube sensors[14] and comparative vacuum sensors[15]. Reflecting on the significance of such advanced sensing technology; in the US, the market for industrial sensor technologies in 2004 was $6.1 billion and expected to grow to $7.7 billon by 2009. While an average annual growth (AAGR) of 4.6% is achieved for these sensors, MEMS are expected to observe an AAGR of 7.6% leading to a market value of $2.5 billion by 2009. It is also expected that the semiconductor based world market will reach $21.8 billion in $2008 as opposed to the $12.6 billion of 1998. Fibre optic sensors of several microns in size are no exception. The 1998 world market was $175 million and expected to reach $600 million by 2011.[16]

4. Smart and MEMS Based Sensor Networks

The word smart is defined as “intelligent or able to think and understand quickly in difficult situations”.[17] For example, smart devices are defined as ones that operate using computers (e.g. smart bombs and smart cards.) The same source defines the word intelligence as “the ability to understand and learn well and to form judgments and opinions based on reason.” Furthermore, the definition of advanced is said to be “highly developed or difficult.” According to the IEEE 1451 standard,[18] a smart or intelligent sensor is defined as “one chip, without external components, including the sensing, interfacing, signal processing and intelligence (self-testing, self-identification or self-adaptation) functions”. Figures 4,[18] illustrates the smart sensor concept as defined by IEEE 1451.

Sensors based on this concept generally exploit development in MEMS and nano technologies along with advanced wireless devices with radio frequency communications technologies. Figure 5[19] depicts such a smart sensor, known as sensor node, for multi-parameters sensing. In this case, the sensor node contains

four major components: 3M's MicroflexTM tape carrier thinned MEMS strain sensors, Linear Polarization Resistor (LPR) sensors to detect wetness and corrosion and electronics module. The electronics module consists of a Micro Controller Unit (MCU), a signal conditioning unit, a wireless Integrated Circuit (IC) unit, a battery and an antenna. Employing this node design, Niblock et al.[20] developed an Arrayed Multiple Sensor Networks (AMSN) for materials and structural prognostics. Some of the observed benefits employing smart sensors system include the wealth of information that can be gathered from the process leading to reduced downtime and improved quality; increased distributed intelligence providing complete knowledge of a system, subsystem, or component's state of awareness and health for 'optimal' decision making.[21] Additionally, due to their significant small size and integrated structure, these sensors can easily be embedded into composites structures or sandwiched between metallic components for remote wireless and internet based monitoring. Intelligent signal processing and decision making protocols can also be implemented within the node structure to provide ready to use decisions for reduced downtime and increased maintenance efficiency. Figure 6 illustrates a Graphic User Interface (GUI) display of three sensors output in the most desired format (pass/fail). Even though the promise of this sensor technology is significant; wireless communication, power requirement, and system integration meeting practical size and weight requirements within an economical manufacturing environment require further consideration.

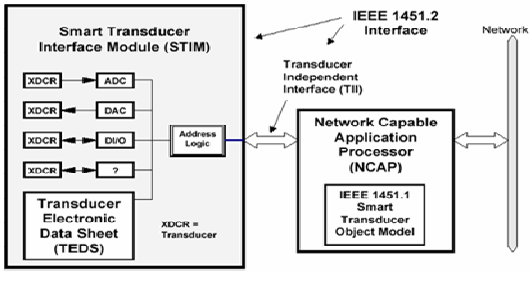


Figure 4. Smart sensor concept defined by IEEE 1451.

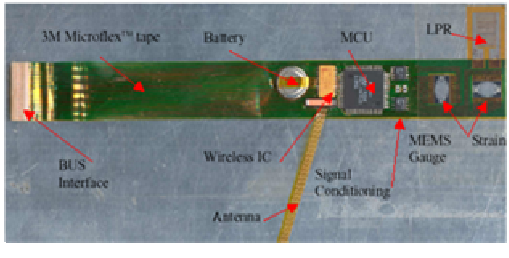


Figure 5. Smart MEMS based smart sensor node.

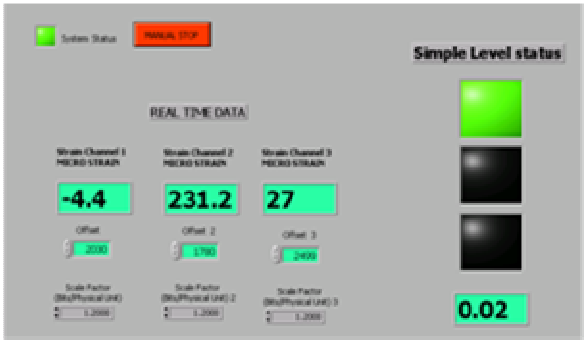


Figure 6. Smart sensor host display for three sensors with a graphical decision making capability.

5. Piezoelectric Based Sensor Networks

Piezoelectricity dates back to its discovery in 1880 by the Curie brothers. Materials possessing these characteristics are generally referred to as dual function or smart materials and have been used extensively in the development of innovative small size and effective actuators and sensors technologies. Piezoelectric actuators are generally made of ceramics and employ the indirect effect; whereas piezoelectric sensors are made of thin polymers and employ the direct effect. Both materials are exploited in the development of MEMS based sensors and actuators as well as advanced structural health monitoring and prognostic health management systems.[22-24]

In recent years, significant research was devoted to the development of structural health monitoring capabilities based on piezoelectric materials. Masson et al.[25] demonstrated a number of modeling tools and damage detection strategies in low and medium frequency ranges. A number of sensing and actuation technologies including arrays of piezoelectric sensors and actuators, shape memory alloys, and micro-accelerometers were also demonstrated for application to structural health monitoring at higher frequencies. Employing piezoelectric transducers' arrays, shown in Figure 7, and a time reversal approach, damage detection and localization were demonstrated. Furthermore, Mrad et al.[26] exploited these smart materials characteristics to detect exfoliation damage in metallic structures through the use of arrayed sensors, as shown in Figure 8. These novel transducers, analogous to nanocoatings, are known as integrated thin film paint-on Ultrasonic Transducers Array (UTA). The fabrication novelty and flexibility in producing UTA with different sizes, thicknesses, for different aerospace applications is documented and demonstrated by Kobayashi et al.[27]

Piezoelectric material can be used both for active and passive defect detection employing network of sensors. As illustrated in Figure 9,[28] in the active mode, an electric pulse is sent to a piezoelectric actuator that produces Lamb waves within the structure under evaluation. The array of piezoelectric sensors will pick up the resultant Lamb waves for processing and analysis. If defects such as cracks, delamination, disbond or corrosion exist within the array of sensors, a change in the signal results, which is distinct from the reference healthy or non-defect component's signature. These systems rely on a reference signal in the structure before they are placed in service. The location and the size of the defect can generally be determined from the degree of signal change. In the passive mode, the sensors are used continuously as "listening" devices for any possible damage initiation or propagation. Sensors within the network can detect energies emitted from impact events and defects generation, including crack formation, delamination, disbond, and possibly non-visible impact damage. In this mode, arrays of sensors are used to monitor or detect such events that generally have a particular energy signal associated with them. Cracks of length 0.005 inches from a distance of six inches on flat plate samples[29] were detected using these sensors. The latter mode is generally desired as it requires no power for its operation. Systems based on this dual concept of passive and active monitoring have been developed.[30,31] Stanford Multi-Actuator-Receiver Transduction (SMART) Layer based system has been developed and demonstrated for several aerospace structural health applications. The system is designed and built around a set of piezoelectric sensors networked and embedded into desired configurations in a single polyimide layer, known as SMART Layer®. Using dedicated diagnostics software, analysis tools and graphics user interface packaged into a SMART Suitcase™ portable diagnostic hardware, the damage severity and location can be easily identified. Figure 10[30,31] depicts this approach with two applications: composite bonded patch repair health monitoring and integrity assessment and damage detection under fasteners.

This sensor technology provides significant potential in the development and implementation of DPHM capability due to the high sensor multiplexing capability, the suitability of the sensor array in harsh environment and its sensitivity to pressure, temperature, vibration, and strains. However, this technology, used in either mode of operation (passive or active), comes with drawbacks that require further research. In the passive mode, background noise (acoustic emission from non-defect events) needs to be accommodated for requiring significant experience and additional expertise to accurately diagnose the damage. In the active mode, sensors/actuators must be spaced properly and

excited with certain frequencies at selected energies to be able to detect damage with certain sizes and regions as demonstrated by Pinsonnault.[32] Tremendous progress was reported in this area; however, significant research is needed to bring this technology to practical use and to facilitate its qualification and certification on air vehicles.

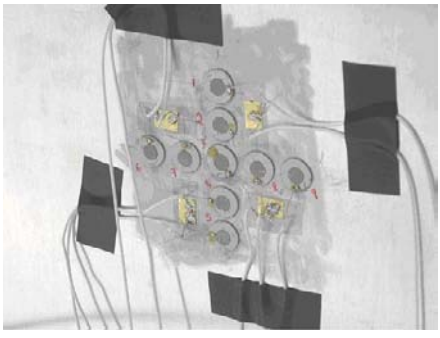


Figure 7. Arrayed piezoceramic sensors used for damage (crack) detection and localization.

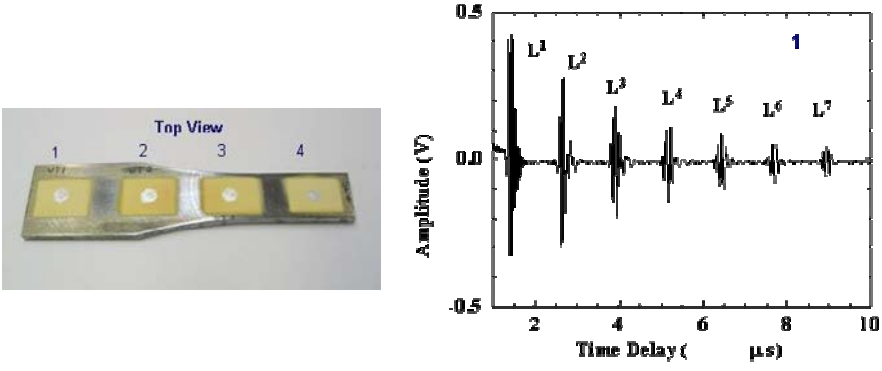


Figure 8. Exfoliated test article with four integrated UTAs and a typical UTA response.

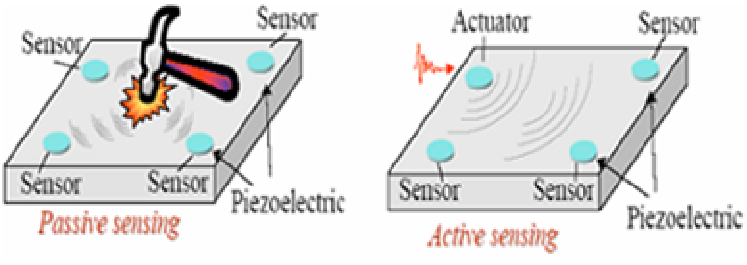


Figure 9. Passive and active sensing mode using piezoelectric materials.

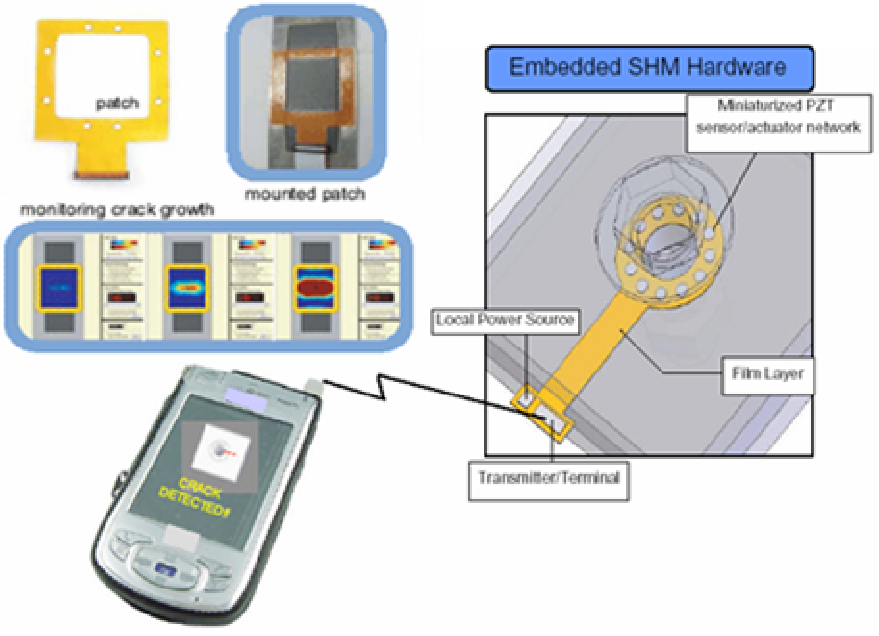


Figure 10. A SMART technology based structural health monitoring system.

6. Fibre Optic Based Sensor Networks

Fibre optic sensor development that capitalized on the 1950's successful discovery of communication optical fibres has been underway since the early 1970's. Only in recent years, accelerated progress was experienced due to the significant development of new, low-cost materials and devices, the emergence of micro and nano technologies for the telecommunications industry and the increased interest in the development and implementation of DPHM systems. The shape and form of optical fibres are similar to those reinforcing fibres used in fibre-reinforced composite materials. However, the diameter of optical fibres is much larger, usually in the order of 40 to 250 microns, compared to glass and carbon fibres used in composites that are typically 10 microns or smaller. Optical fibres consist of a light waveguide inner silica-based core surrounded by an annular doped silica cladding that is protected by a polymer coating. This optical fibre can also be made using other materials, such as plastic.[33] The fibre core refractive index is relatively large compared to that of the cladding index. The change in refractive indices, between the core and the cladding, provides the required mechanics for light propagation within the fibre core. Depending on the wavelength of the light input, waveguide geometry and distribution of its

refractive indices, several modes can propagate through the fibre, resulting in the so-called, single and multi-mode optical fibres. Both fibre types are used in the construction of fibre optic sensors. However, single mode fibres are more sensitive to strain variation and are thus the preferred choice for DPHM applications. Krohn[34] provides a summary of the typical properties of various optical fibre types.

Compared to more traditional measurement techniques, fibre optic sensors offer unique capabilities such as monitoring the manufacturing process of composite and metallic parts, performing non-destructive testing once fabrication is complete, enabling structural and component health monitoring for prognostics health management, and structural control for component life extension. Because of their very low weight, small size, high bandwidth and immunity to electromagnetic and radio frequency interferences, fibre optic sensors have significant performance advantages over traditional sensors. In contrast to classical sensors that are largely based on measurement of electrical parameters such as resistance or capacitance, fibre optic sensors make use of a variety of novel phenomena inherent in the structure of the fibre itself. Some of these phenomena are extensively discussed in the literature.[35,36] In general these sensors can be classified into two classes, the discrete and distributed class. The distributed class of sensors includes Michelson and Mach-Zhender interferometer as well as sensors based on Brillouin scattering. These are generally seen in infrastructure applications where spatial resolution, system's weight and size are not as critical and long range sensing is desired.[37] The discrete class of sensors includes cavity-based and grating-based designs. Cavity-based designs utilize an interferometric cavity in the fiber to create the sensor and define its gauge length. Extrinsic and Intrinsic Fabry-Perot interferometers (EFPI, IFPI), along with In-Line Fiber Etalon (ILFE) are the most known ones. Grating-based designs utilize a photo-induced periodicity in the fiber core refractive index to create a sensor whose reflected or transmitted wavelength is a function of the periodicity that is indicative of the parameter being measured. Any shift in the reflected wavelength indicates a change in the monitored parameter. The principle of operation of Bragg gratings based sensors is shown in Figure 11.[38]

Fiber Bragg gratings (FBG) are the most commonly used sensors in this class and are the most attractive for integration into advanced diagnostics and prognostics capabilities due to their sensitivity, small size and high multiplexing capabilities forming a highly effective sensor networks. As shown in Figures 12,[39] these sensors can be used to monitor bondline integrity in bonded joints, acoustic emission resulting from structural damage and corrosion

monitoring. When integrated with a centralized monitoring system, as shown in Figure 13, on-line real-time corrosion sensing and health prediction can be performed using very high number (up to 1000) of sensors networked on a single strand fiber.

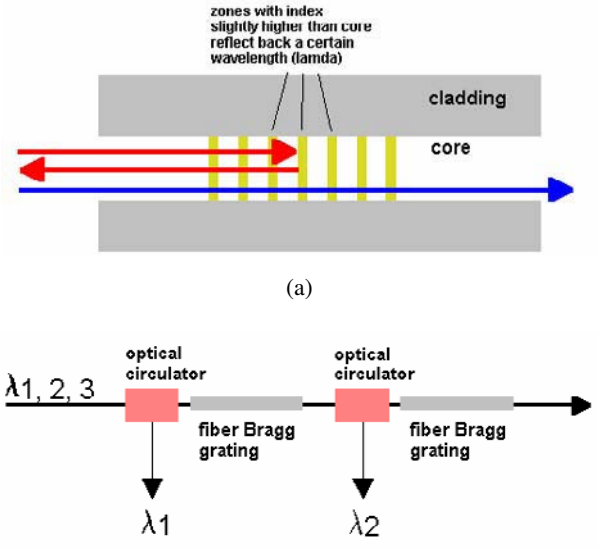


Figure 11. Fibre Bragg gratings principle of operation.

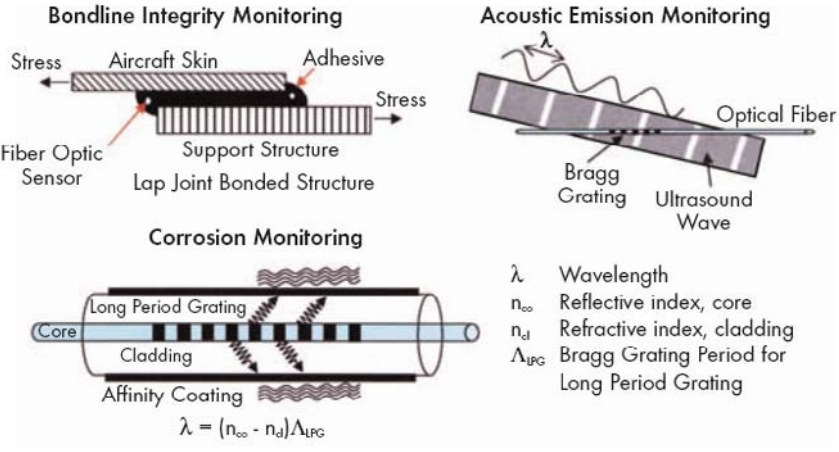


Figure 12. Bragg grating based sensing.

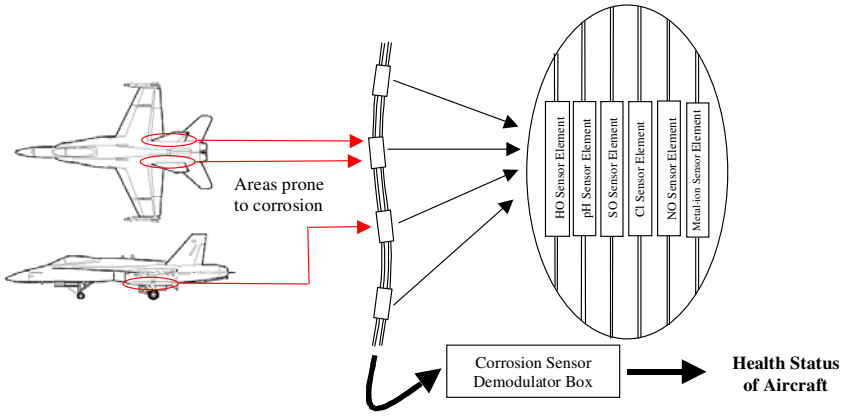


Figure 13. Illustration of a fighter jet health monitoring system.

Several sensor output interrogation techniques and their associated modulation systems have been developed and implemented.[40] These systems are found to be costly and impractical for in-flight applications due to their significant size, weight and power requirement. Recent efforts have focused on developing an increased level of understanding of the issues impeding the implementation of fiber Bragg gratings (single or multiplexed) into aerospace and military air platforms[41,42] and developing a micro fiber optic interrogation system for increased sensor networks and reduced size and weight. Figure 14[43] illustrates a developmental effort in exploiting the enabling area of MEMS technology to develop a micro arrayed highly multiplexed FBG interrogation system providing significant advances in air platforms diagnostics and prognostics. The FBG sensor is embedded into the composite lay-up and initially used for component process monitoring (potential process control) during manufacturing.[42] It is subsequently used as diagnostics and prognostics tool in the integrated DPHM. Regardless of the extensive and successful outcome of several investigations supporting aerospace platform DPHM requirements, research effort continue to address the critical issues for practical implementation that include adhesive selection, bonding procedures, quality control for surface mounted fiber optic sensors, optimum selection of sensor configuration, sensor material and host structure for embedded configurations; characterization of embedded fiber optic sensors at elevated and cryogenic temperatures, resolution optimization for desired parameters from multi-gratings as well as sensitivity to transverse and temperature effects and development of an integrity assurance procedure for embedded sensors, particularly sensor protection at egress/ingress points.

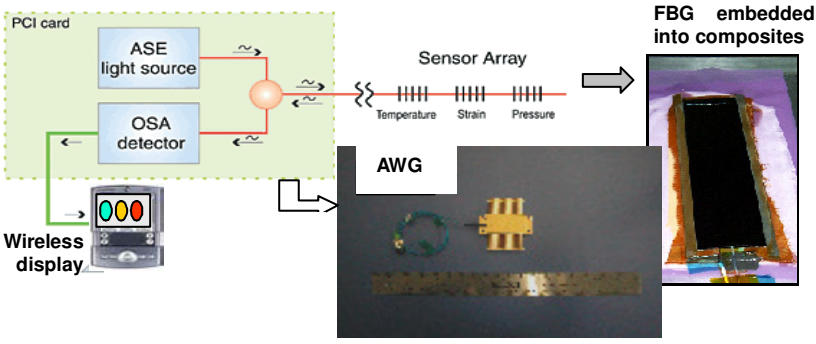


Figure 14. Conceptual illustration of a micro arrayed highly multiplexed FBG interrogation system.

7. Comparative Vacuum based Sensor Networks

Among the wide range of sensor technologies that have emerged in recent years as damage detection sensors, Comparative Vacuum Monitoring (CVM) technology is assessed to be a mature technology, for both Non Destructive Testing / Inspection (NDT/NDI) and Condition Monitoring (CM) of structural integrity, that is ready for deployment onto operational platforms. The state of maturity of this technology coupled with the desire of the aerospace industry to deploy an automated inspection method that does not have reduced accuracy and that could take place with personnel remote to the inspection area has triggered a strong desire to evaluate this technology within an aerospace environment. Recently, Boeing, FAA, Airbus, Northwest Airlines, United States (US) Navy and Royal Australian Air Force (RAAF) initiated independent verification trials with this technology. These trials involve laboratory, environmental and on-aircraft tests. Validation trial within the US Navy has successfully demonstrated the detection of a crack.[44,45] Based on these trials, it is expected that the CVM technology be approved as an alternate means of compliance (AMOC).[44] This CVM technology has been developed for crack initiation and propagation detection. A CVM system consists of three primary components: a sensor, a pressure differential flow meter and a stable host reference vacuum source. Figures 15 and 16[46] illustrate the conceptual model of a typical CVM system and an example of crack detection process on a component, respectively. The sensor or sensor networks is linked to a reference vacuum source through a flow meter, which contains impedance to the flow of air molecules through the system. The flow meter measures the pressure difference across the impedance. The vacuum source provides the flow meter with a continuous stable reference vacuum and power to maintain vacuum at a stable level.

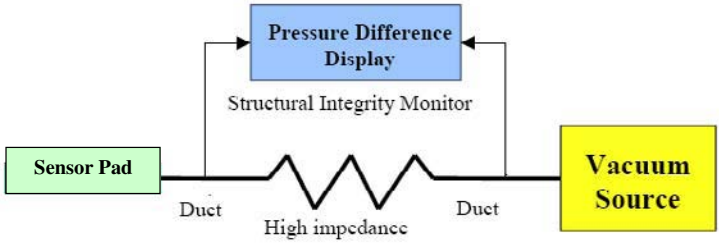


Figure 15. Conceptual model and principle of operation of a typical CVM system.

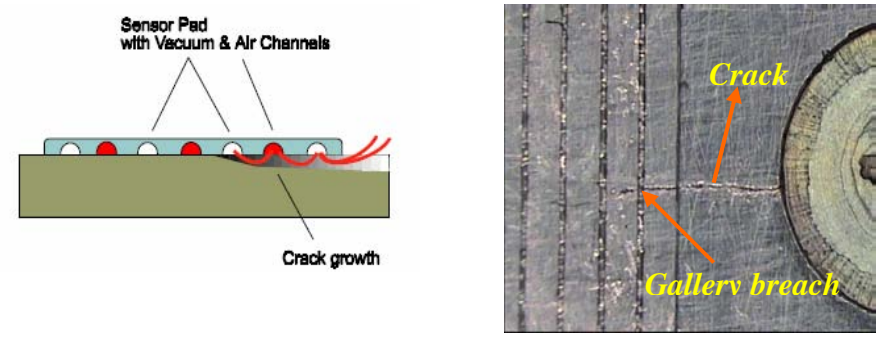


Figure 16. Illustration of crack detection using CVM technology.

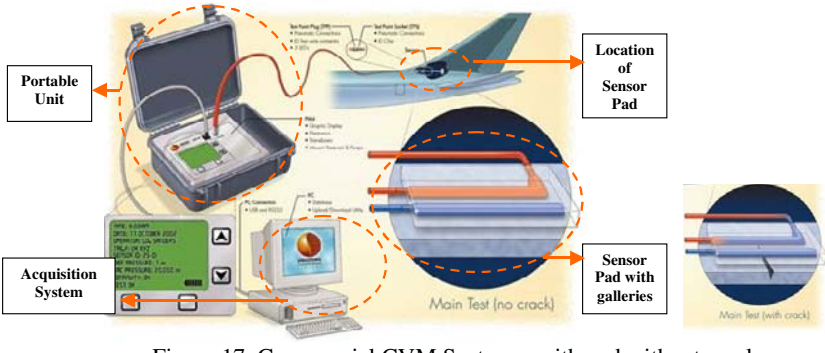


Figure 17. Commercial CVM System – with and without crack.

The fundamental principle of the CVM operation is the detection of pressure difference caused by a crack in a vacuum channel, known as ‘gallery’ within the sensor pad. If there is no damage and since the sensor is sealed to the test article there should be no leaks or change in pressure difference and hence a balance is maintained between the sensor and the vacuum source. If there is damage, a leak occurs at the location of the sensor pad and a pressure increase is detected by the flow meter. The rate of growth of the damage (crack) is determined by the rate of

pressure change. The increase in pressure is measured as a differential pressure in relation to the reference vacuum level. This is significantly more sensitive to micro fluidic air flow than conventional air flow meters.

An example of the operation of the commercially available system is also shown in Figure 17.[47] As illustrated by the extensive work in the open literature, this technology is experiencing tremendous growth for both military and civilian aircraft.[48,49] This technology, however, is very well suited for localized damage detection but not for global systems' and components' health monitoring. Sensors employed for this technology are often compared to resistive based crack detection gauges. Even though sensor multiplexing is demonstrated, implementation remains limited to hot spots and applications where false positives and system weight can be tolerated. Other limitations experienced by this technology are common to sensors discussed in this chapter and require further consideration. The technology is seen to be commercially more mature than the earlier presented technologies.

8. Conclusions

Diagnostics, Prognostics and Health Management (DPHM) concept that has been around for a long time has emerged in recent years as the ultimate approach for aircraft enhanced operational envelope, reduced Life Cycle Costs (LCC), increased aircraft availability, enhanced operational logistics and reduced complexity for new acquisition programs and legacy platforms alike. The implication of employing a DPHM system that includes continuous operational monitoring and damage detection, assessment and prediction technologies is monumental. However, its impact is underestimated as it enables lesser access to aircraft, reduces reliance on statistical based scheduled maintenance, increases maintenance efficiency and aircraft availability and moves from condition-based to predictive maintenance approach which can substantially reduce ownership costs. One major challenge in the implementation of the DPHM system is the distribution, networking and communication of sensors providing minimal impact onto vehicle's operation, maintenance and repair. The presented sensors and sensory systems based on fibre optic, piezoelectric and comparative vacuum technologies are assessed to be near commercial exploitation and are considered to be advanced in addressing some of the challenges. However, numerous sensing devices (in addition to the ones covered here) are emerging with the introduction of new technologies that are expected to help overcome the technical and operational challenges and to ensure the implementation of high-performance DPHM systems.

References

1. J. Toon, *Research Horizons*, (2001).
2. P. Faas, J.B. Schroeder and G. Smith, *IEEE AESS Systems Magazine*, 10 (2002).
3. A. Hess, retrieved from *http://www.dtic.mil/ndia/2001systems/hess.pdf*
4. Z. Liu, M. Genest, N. Mrad and A. Fahr, *National Research Council Canada, LTR-SMPL-2006-0085*, (2006).
5. A. Hess, *DARPA/DSO Prognosis Bidder's Conference*, (2002)
6. http://www.corrosion-doctors.org/
7. S. Beeby, G. Ensell, M. Kraft and N. White, MEMS Mechanical Sensors: Micromechanical System series, *Artech House Publishers* (2004).
8. E. Udd, *Review of Scientific Instruments*, **66(8)**, 4015 (1995).
9. J. F. Tressler, S. Alkoy and R. E. Newnham, *Journal of Electroceramics*, **2(4)**, 257 (1998).
10. V. Giurgiutiu, A. N. Zagrai, *Transactions of the ASME: Journal of Vibration and Acoustics*, **124(1)**, 116 (2002).
11. I. Sage, L. Humberstone, I. Oswald, P. Lloyd and G. Bourhill1, *Smart Mater. Struct*. **10**, 332 (2001).
12. S. Beard, X. Q. Peter, M. Hamilton and D. C. Zhang, *Acellent Technologies, Inc*. (2005).
13. K. Otsuka and X. Ren, *Intermetallics*, **7**, 511 (1999).
14. I. Kang, M. J. Schulz, J. H. Kim, V. Shanov, D. Shi, *Smart Mater. Struct*. **15**, 737 (2006).
15. M. Wishaw and D. P. Barton, *10th Asia-Pacific Conference on Non-Destructive Testing* (2001).
16. BCC Market reports research, retrieved from *http://www.bccresearch.com/*
17. http://dictionary.cambridge.org/results.asp?searchword=smart
18. National Institute of Standards and Technology, *IEEE 1451 Website*, retrieved from http://ieee1451.nist.gov/
19. T. Niblock, B. Laskowski, H. Surangalikar and J.Moreno, *Analatom Inc.*, retrieved from http://analatom.com
20. T. Niblock, H.S. Surangalikar, J. Morse, B.C. Laskowski, J. Moreno and B. Caldwell, *Materials Damage Prognosis: Proceedings of a Symposium*, 259 (2005).
21. P. Cleaveland, *Control Engineering Asia* (2004).
22. Y. J. Yan and L. H. Yam, *Composite Structures*, **58(1)**, 29 (2002).
23. V. Giurgiutiu, A. N. Zagrai and J. J. Bao, *Structural Health Monitoring*, 1(1), pp. 41-61, 2002.
24. G. Park, H. Sohn, C.R. Farrar, D.J. Inman, *The Shock and Vibration Digest*, **35(6)**, 451(2003).
25. P. Masson, P. Micheau, Y. Pasco, M. Thomas, V. Brailovski, M. Meunier, Y.-A. Peter, D. Mateescu, A. Misra, N. Mrad, J. Pinsonnault, and A. Cambron, *Proceedings of the CANSMART 2006*, 191 (2006).
26. N. Mrad, Z. Liu, M. Kobayashi, M. Liao, C. K. Jen, *Insight – Non-Destructive Testing & Condition Monitoring*, **48(12)** (2006).
27. M. Kobayashi, C.-K. Jen, J.-F. Moisan, N. Mrad and S. B. Nguyen, *Proceedings of the CANSMART 2005*, 181 (2005).
28. S. Beard, A. Kumar, P. X.Qing, I. H. L. Chan, D. C. Zhang, T. K. Ooi, *12th SPIE International Symposium*, **5762** (2005).

29. C. Marantidis, J.D. Gentry, and J.N. Kudva*, Proc. SPIE - Smart Structures and Materials,* **1917,** 972 (1993).
30. A. Kumar, S. J. Beard, P. X. Qing, H.-L.Chan, T. Ooi, S. A. Marotta, F. K. Chang, *SEM X International Congress*, (2004).
31. S. Beard, P. X. Qing, M. Hamilton, D. Zhang, *2nd European Workshop on Structural Health Monitoring*, (2004).
32. J. Pinsonnault, P. Masson, P. Micheau, N. Mrad, *Proceedings of the 5th International Workshop on Structural Health Monitoring,* 961 (2005).
33. G. Zhou and L.M. Sim, *Smart Mater. Struct*., **11**, 925 (2002).
34. D.A. Krohn, Fiber optic sensors: Fundamentals and Applications, *Instrument Society of America,* (1988).
35. N. Mrad, *Encyclopedia of Smart Materials*, **2**, 715 (2002).
36. E. Udd, Fiber *Optic Sensors - An Introduction for Engineers and Scientists, John Wiley & Sons*, (2006).
37. L. Zou, X. Bao, F. Ravet and L. Chen, *Applied Optics*, **45(14),** 3372 (2006).
38. PC magazine, retrieved from *computer desktop encyclopedia* (2006).
39. I. Perez, *Fiber Sensors for Aircraft Monitoring*, Naval Air Warfare Center Aircraft Division, 1999. DTIC Doc: AD-A375814.
40. N. Mrad, *National Research Council Canada (NRC), Publication No. LTR-SMPL-2001-0091* (2001).
41. N. Mrad, S. Sparling and J. Laliberté, *Proceedings of SPIE 6th Annual meeting* **3670**, 82 (1999).
42. N. Mrad, T. Marincak, B. Moyes and K. McRae, *Fourth Canadian International CANCOM 2003*, paper No. 084, (2003).
43. N. Mrad and G. Z. Xiao, *Proceedings of ICMENS 2005,* 359 (2005).
44. G Wheatley, J Kollgaard, J Register, M Zaidi, *Insight - Non-Destructive Testing and Condition Monitoring,* **47(3)**, 153 (2005).
45. Structural monitoring systems, *Investor Newsletter* (2006)
46. D. Roach and K. Rackow, *Sandia National Laboratories*.
47. C. M. Doherty, M. Lindroos, D. P. Barton, *Proceedings of the 4th Australian Pacific Veritflite Conference on Helicopter Technology* (2003).
48. D. Roach, J. Kollgaard, S. Emery, J.Register, K. Colavito, D. Galella, Retrieved from *http://www.airlines.org/NR/rdonlyres/F7515E99-C984-46FA-970B-7DDD5BFAD8E6/0/Thurs0900a_Roach.pdf*
49. H. Speckmann and R. Henrich, *16th WCNDT 2004 - World Conference on NDT*, (2004).

This page intentionally left blank

Part 3

Signal and Image Processing in NDE

This page intentionally left blank

CHAPTER 21

COMPUTATIONAL INTELLIGENCE METHODOLOGIES FOR NON DESTRUCTIVE TESTING/EVALUATION APPLICATIONS

M. Cacciola, F. C. Morabito, and M. Versaci*

Via Graziella Feo di Vito,
89100 Reggio Calabria (RC), Italy,
{*matteo.cacciola, morabito, mario.versaci*}*@unirc.it*†

This chapter aims to show how some relevant industrial applications of Non Destructive Testing and Evaluation (NDT/NDE) can benefit of some recently emerged Computational Intelligence (CI) techniques. The problems under study are generally formulated as inverse problems whose typical ill-posedness and ill-conditioning (in their numerical counterpart) can be dealt with such tools like innovative signal processing techniques (i.e., Independent Component Analysis, ICA, wavelet representation, Support Vector Machines, Fuzzy Systems, Neural Networks). Some practical examples for the solution of relevant inspection problems will be discussed. The advantages of using the CI approach will be clarified and assessed on applications of NDT/NDE in both civil engineering and industrial problems.

1. Introduction

The deep interest of the society to advance knowledge and to improve technologies often faces the need of ensuring safety in both relevant scientific challenge (like nuclear fusion reactors and space applications) and industrial equipments in the field of the energy production and in biomedical applications as well as in public transport. These important goals can be best achieved through the involvement of different competencies deriving from sefveral scientific disciplinary fields. From one hand, researchers experiment continuously novel materials and deepen knowledge in nanotechnologies, on the other hand, the advancement of methodologies and techniques for the NDT/NDE of materials to ensure quality of relevant processes is largely justified. For many of the above mentioned objectives, there is also a quest for the in line and non invasive character of the procedures of inspection, whereas the economic-financial aspect of reducing maintenance cost of research

*Researches described in this chapter have been mainly developed within the Project of National Relevant Interest 2004, co-funded by the Italian Ministry of University and Research, and the APEX Project in collaboration with Alenia Aeronautica SpA.

†Authors are with University "Mediterranea" of Reggio Calabria, DIMET, Italy.

and/or technologic plants is to be considered as a significant additional constraint. The development of affordable, secure, reproducible and quick methodologies of inspections can then be considered as a matter of primary importance, in order to draw conclusions on the state of the system. The elements from which making decisions are invariably some sets of reading measurements whose acquisition does not compromise the integrity of the specimen and/or apparatus under test and the safety of the operation as well. The intention of this chapter is to introduce the reader to some new innovative approaches in NDT/NDE of materials and structures of both industrial and civil engineering interest. In some cases, some advancements of well known approaches are just shown without going in analytical details of the methodologies involved for which the reader is redirected towards the huge literature on the different subjects here discussed. The faced problems are formulated and generally solved within the framework of the inverse problem: starting from some measurements, properly taken, for instance, along the border of the system, it would like to be able to figure out the configuration of the system from which the pattern of measurements originate. Some details and clues could support the decision on the piece or at least suggest some closer examinations. CI, among which the "learning" aspects of the tasks deserve a special care, could help at different levels in this tasks. In particular, techniques to solve the inverse problems by dealing with the very common ill-posedness (and the numerical ill-conditioning) of the problem.[1] The use of advanced signal processing methods (see Sec. 3.1) can be useful for making a suitable pre-processing of the available raw data also in order of speeding up the learning procedures.[2]

This chapter is structured as follows: Sec. 2 is devoted to the description of the application problems of interest together with the assessment of methodological aspects and formulation of the direct problem. The CI framework of investigations is presented in Sec. 3 with special emphasis on the NDT/NDE relevant details. A particular reference to the Eddy Current Technique (ECT) is made. Section 4 reports the solution of significant application problems (lift-off effect reduction, characterisation of shape and in-depth defects, identification and classification of defects on aeronautic-relevant materials determination of mechanical stresses, and imaging of corrosion in stainless steel specimens). Most interesting problems cannot be reported in this chapter and some reference to them will be made. A selected list of references closes the work.

2. Flaw Characterization in Industrial Applications, and in Civil and Structural Engineering

Integrity of materials has clearly a great relevance in many industrial applications where a trade-off is typically necessary between the need of material re-using for reducing costs and the obvious safety requirements. The structural elements used in these technological applications are of very different nature (i.e., aluminium or

structural steel plates, aeronautic composites, rivets, and so on); each one of them shows its peculiar characteristics of mechanical resistance and can be subjected to limited solicitations or decay. From this viewpoint, qualitative information, such as rough shape, kind and deepness of defects, can be of interest to technicians. The NDE/NDT problem can be easily formulated as an inverse problem: how to detect and to characterize a defect starting from non invasive (external) measurements. It will be examined thoroughly in this section by splitting it up in three different sub-problems: 1) characterization of defects in metallic plates in terms of their shapes as well as their in-depth locations; 2) identification of imagery of corrosion in metallic plates; 3) classification of flaws in aeronautic composites. In the first two cases, ECT is mainly used as means of inspection and to achieve a suitable non destructive electromagnetic representation of the problem, whereas UTs are preferably used in the third one. ECT is significantly affected by noise. The different noise sources can be classified as usual as systematic noise (the so called lift-off effect is one relevant example) and uncorrelated or content independent noise, like the acquisition measurement noise. The presence of noise may reduce the reliability of numerical algorithms basically used to solve the inverse problem. In order to improve the success of inspection tests the pattern of the reading data should emphasize the informative content properly filtering the noise. However, because of the typical level of the signals acquired during scans it is also critical to maximize the information without reducing it through the filtering process. This can be achieved, for example, by generalizing the band-pass filtering concept, as it is carried out either by the wavelet transform approach or the extraction of independent components of the signal. In ECT, one of the noise source is the so called lift-off effect. It non-linearly influences the measured data by introducing a typical low frequency and variable spatial distributed noise. It will be shown how to cope with it.[3–5]

2.1. *Shape and In-Depth Location of Defect in Metallic Plates*

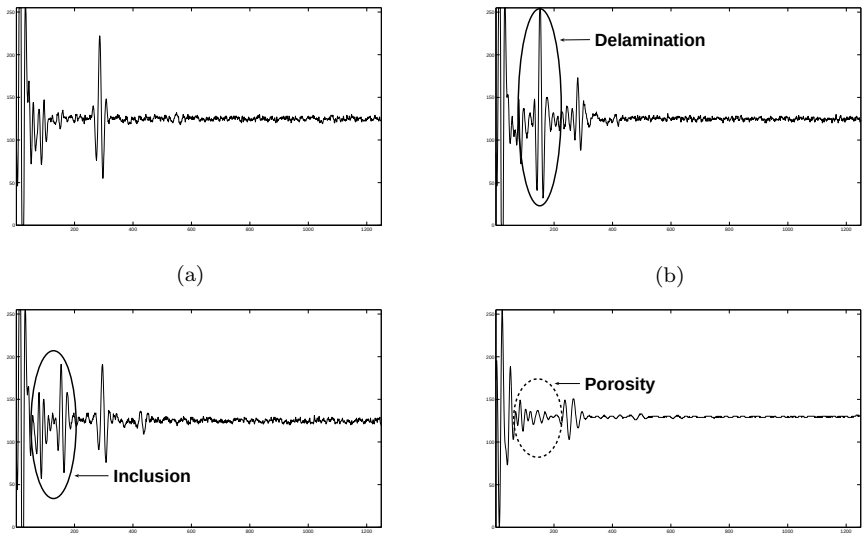
Metallic plates are commonly used in structural engineering as well as in some industrial applications. In these problems, it is of interest the detection of in-depth location of a defect and the characterization of the defect's shape. As far the determination of defect's deepness is concerned, a good calibration of the probe, and a wise setting of the frequency f_{exc} for the exciting ECT signal is crucial. This is related to the well known *skin effect*, i.e., the current density near the surface of the conductor is greater than that at its core. This phenomenon allows to detect sub-superficial flaws. On the other hand, using a low f_{exc}, deep flaws can be in principle detected. As a consequence, the detection and characterization of "middle-depth" defects is remarkable,[6] which requires a precise calibration of the measurement instruments. The solution of the related inverse problem is further complicated by the presence of noise. The main contribution of the present work is the use of CI in order to implement real-time procedures that can help to assess the deepness of a possible defect. Another interesting subject is the determination

of the shape of a defect, in order to reach its complete characterization through ECT. This impacts on the definition of a suitable threshold for the life cycle of the inspected materials. The reading is here often similar for possibly very different sources. In the literature, a lot of methods have been proposed to approach this ill-posed inverse problem.[7] These works always focused on the emergence of the links between measured signals and shape of the defect.[8]

2.2. *Characterization of Flaws in Composites for Aeronautic Applications*

Within the framework of aeronautic industry, it is needed to ensure the quality check of the manufacture process, with special regards to the composite materials used for the airplanes. According to the current regulations, monitoring of material integrity is even carried out during their operative life, in terms of both reliability and safety. The basic materials used for constructing the main part of modern airplanes are the Carbon Fiber Reinforced Polymers (CFRPs). Today, air companies exploit airplanes of their fleets as much as possible, in order to amortize purchase and maintenance costs. These financial requirements have, however, to be balanced with high safety. Moreover, it is easy to understand how mechanic solicitations and atmospheric agents are responsible of a relatively rapid degradation of airplanes' structures. Therefore, it is evidently of interest to not introduce other a priori dangerous risk factors. Notwithstanding, a number of characteristic flaws can be induced into a CFRP during the manufacture process or along the life cycle of airplane. Thus, it is needed to accurately detect even the smallest defects in real-time by using a suitable NDT/NDE techniques, and, where required, to classify the different kind of flaws starting from the measurements. Among the others, the defects of main industrial significance are: delaminations (i.e. separation between two close plies along a horizontal plane), inclusions (i.e. accidental leaving of foreign material into the CFRP during the lay-up) and porosities (i.e. spherical/elliptical micro-voids distributed within a volume or in a plane parallel to a lay-up).

Delaminations are characterized by a peak due to a defect back-scattering, whereas the back-wall echo, due to the reflection on the back medium surface, could not be present as the delamination almost reflects the signal. The inclusions usually show a shape similar to delamination, even if the ultrasonic signal is generally not completely reflected by the defect and the back-wall echo appears. As far as the porosity is concerned, volumes with few small pores (tens or also hundreds of microns) could not be detected by the pulse-echo techniques. Examples of ultrasonic A-scan are depicted in Fig. 1. CI techniques can be again of help starting from the availability of sufficient in-lab measurements useful to learn the solution of the problem. Selected features must be able to assure a good class separation without getting trapped in the "curse of dimensionality" problem, i.e. an excessive number of training inputs which increases the system complexity without any remarkable advantages in terms of classification performances.



(a) (b) (c) (d)

Fig. 1. Some A-scan exemplifying the problem of defect characterization in CFRP: ultrasonic pulses (a) without defect; (b) with delamination; (c) with inclusion; (d) with porosity.

2.3. *State of the Art for Imaging of Corrosion in Metallic Plates*

Corrosion is a material degradation that also involves geometric modifications. Imaging corrosion techniques allow to retrieve an imagery representation of analysed plate showing the loss material profile when sub-superficial corrosion takes place. A map of the impedance variation in metallic plate is generated by an ECT scan. State of the art of image of corrosion systems allows the correlation of corrosion profile with the impedance variation[9] and the assumption of thin plate[10] by the minimization of a suitable functional. Let $(\mathbf{E}^\varepsilon, \mathbf{H}^\varepsilon)$ represent the electric and magnetic fields generated by the probe when the plate is damaged, whereas $(\mathbf{E^0}, \mathbf{H^0})$ are fields in the undamaged plate. On the boundary of the plate, there is the continuity of the tangential components of $(\mathbf{E}^\varepsilon, \mathbf{H}^\varepsilon)$. From Faraday's law we obtain, for the measured impedance, when the probe is at (x_0, y_0):

$$Z^\varepsilon(x_0, y_0) = \frac{i\omega\mu}{I} \int_{(\Delta x)^2+(\Delta y)^2 \leq r_0^2} \mathbf{H}^\varepsilon{}_z(x, y, h+t)dxdy \tag{1}$$

where $\Delta x = (x - x_0)$, $\Delta y = (y - y_0)$ and H_z is the z-component of the magnetic field. The inverse problem can be stated as follows: *find the loss profile* $\varepsilon\theta(x, y)$ *from the impedance measurement* $Z^\varepsilon(x_0, y_0)$ *for* $-\infty < (x_0, y_0) < \infty$. Using Maxwell's equation and Stoke's theorem, and modelling the current source (i.e. the eddy current probe) by a circular loop, it is possible to derive the following relationship

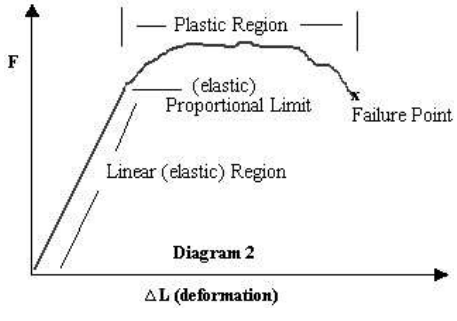
between impedance changes on the plate surface and the loss profile:

$$\delta Z^{\varepsilon}(x_0, y_0) = -\varepsilon \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(x_0 - x, y_0 - y)\theta(x, y)dxdy \tag{2}$$

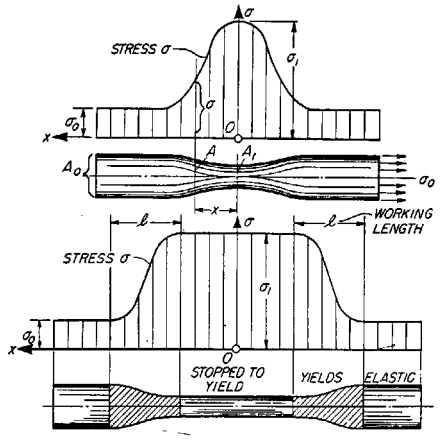
where K is a so called *kernel* given by $\delta K(x, y) = \frac{\sigma_0}{I^2} E^0(x, y, 0) \cdot E^0(x, y, 0)$. Since both $\widehat{E_x^0}$ and $\widehat{E_y^0}$ decay exponentially, $K(x, y)$ is a very smooth kernel. It is expected that a small high wave number error in the data δZ^{ε} would lead to large errors in $\theta(x, y)$ when inverting the relation Eq. (2). The ill-positioning can be reduced by a careful examination of $K(x, y)$ shape and a right determination of a regularization strategy. A competitive regularization method based on evolutionary computation is therefore useful for our aims.[11]

2.4. *The Problem of Material Integrity in Civil Engineering*

From an experimental point of view, materials subjected to solid-solid phase changes show discontinuous deformations. In particular, metallic materials have a stress-deformation step function, similarly close to Savart-Masson effect. Among the well-known mechanisms of deformation, the "microstructure" exhibited by the solid under stress is to be considered. It is usually known as "shear",[12] from which the "shear-band" definition descends, referring to surfaces in which the deformation is mostly (highest values) concentrated along bands. Within this framework, the stress-strain law affirms that the deformation ΔL caused by an applied force $\mathbf{F}$ depends on the same load (Fig. 2). In the elastic region, Hooke's law is applicable in order to calculate the applied force starting from the deformation of inspected



(a)



(b)

Fig. 2. (a) Representative diagram of stress-strain law. (b) Working lengths obtained by a specimen under traction.[12] Both necking and trends of strain state on each section of the specimen are evidenced.

material, which has an isotropic linear-elastic behaviour. It is the so called necking phenomenon, where the solid becomes homogeneously deformed until a given value of the hard or soft device (Fig. 2(b)). After enough force has been applied the material enters the plastic region, where the force and the deformation are not proportional, but rather a small amount of increase in force produces a large amount of deformation. In this region, the material is about to fail. The above described phenomena can involve also metallic plates, used for instance in reinforcement of beams and pillars in old civil buildings, or magnetic steels, used to fabricate rotating electrical machines. At the same time, it is well known how mechanical stresses can influence the microscopic structure of a metallic object, with a consequent variation of its electromagnetic properties, i.e. local electric conductivity.[13] Thus, if a deformed plate is inspected by ECT, the path of the eddy currents is disturbed, and a variation of the magnetic field is produced. In this way, the stress field, which is not otherwise directly measurable, can be calculated just measuring the eddy current related voltage variation over a specific area and applying the relevant constitutive law. Experimental evaluations, carried out on a structural steel model subject to a one- or bi-axial load having an intensity beyond the yielding point up to the plastic region, have proven the qualitative equivalence between stress and eddy current maps, thus suggesting the use of ECT to early detect structural degradation on the inspected plates (see Sec. 4.4 for details).

3. Computational Intelligence in NDT: Methodological Aspects

In this section, advanced signal processing techniques useful for solving the above proposed NDE/NDT inspection problem will be described making reference to the widely available literature on different subjects. The aim of this section is to resume mathematical foundations of the involved techniques.

3.1. *Advanced Signal Processing Approaches Able to Emphasize Features in NDT/NDE Signals*

Signal processing techniques are nowadays used to emphasize peculiar features of one- and two-dimensional signals. The referred techniques involve, for instance, extraction of statistical features, but they are not sufficient to describe the signal trend in many cases, above all when the signal is produced by a non-linear phenomenon. Thus, advanced signal processing must be introduced to extract a maximum information content from the reading, or else to reduce complexity of the extracted data. In the following paragraphs, some techniques useful to achieve these aims will be described.

3.1.1. *Multi-Resolution Analysis by Using Wavelet Transform*

In mathematics, Wavelets, Wavelet Analysis, and the Wavelet Transform refers to the representation of a signal in terms of a finite length or fast decaying oscillating waveform (known as the mother wavelet). This waveform is scaled and translated to match the input signal. In this way, it is possible to split local and global dynamics for a signal, i.e. low- and high-frequency contribute in a frequency domain, by a Multi-Resolution Analysis (MRA) in the wavelet domain, less sensitive to noise than Fourier Transform. MRA can be applied to the analysis of both one- and two-dimensional signals, in order to retrieve features useful to characterize relative trend into the time-, spatial- or frequency-domain. The global dynamics of a NDT-related signal $f(x)$ can be condensed in the Wavelet Approximation Coefficient (WACs) at the higher multi-resolution level J; if inspected signal is measured in time- or spatial-domain then it is possible to affirm that WACs are related to the low frequencies. On the other hand, local oscillations of $f(x)$ are depicted in a set of so called Wavelet Details Coefficients (WDCs) at the different scales $j = 1,\ 2,\ ...,\ J$; they are related to the high frequencies of a time- or spatial-domain signal. Moreover, let us remark how such noising effects as lift-off effect are restricted to low-frequencies; therefore, a suitable MRA could be also useful in ECT. In formal terms, this representation is a wavelet series, which is the coordinate representation of a square integrable function with respect to a complete, orthonormal set of basis functions for the Hilbert space of square integrable functions.[14] Wavelet Transform guarantees the possibility of not specifying in advance the key signal features and the optimal basis functions needed to project the signal in order to highlight the features. MRA takes origin from the Wavelet analysis.[15]

Wavelet functions are piecewise continuous with compact support. They belong to the space of quadratic integral Lebesgue measurable functions $L^2(\Re)$, and they satisfy the relation $\int_{-\infty}^{+\infty} \psi(x)\ dx = 0$. MRA is defined as a family of subspaces $\{V_j\}_{j\in\mathcal{Z}}$ of $L^2(\Re)$, satisfying the following properties: $\bigcup_{j\in\mathcal{Z}} V_j$ is dense in $L^2(\Re)$; $\bigcap_{j\in\mathcal{Z}} V_j = \{0\}$; $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1},\ \forall f \in L^2(\Re),\ \forall j \in \mathcal{Z}$. So, a so called scaling function $\phi(x)$ exists; it constitutes a Riesz basis for $V_j,\ \forall j \in \mathcal{Z}\ :\ \left\{\phi_{jk}(x) = 2^{-j/2}\phi(2^{-j}x - k)\right\}_{k\in\mathcal{Z}}$. Orthogonal projection $P_j(f)$ of a function f into the space V_j is given by $f_j = P_j f = \sum_k a_{jk}\phi_{jk}$ and is known as the approximation of f at a resolution $2 - j$. If the difference between two successive approximations $m_j = f_{j-1} - f_j$ is considered, it is possible to affirm that lost details are preserved when one passes from a (2-j+1)-th to a (2-j)-th approximation. The so-defined space $W_j = \{m_j = f_{j-1} - P_j f :\ P_j f \in V_j\}$ is a subspace of V_{j-1} and is complementary to V_j. Thus, a function $\psi(x)$ (i.e. a mother wavelet) exists such that the family $\left\{\psi_{jk}(x) = 2^{-j/2}\psi(2^{-j}x - k)\right\}_{k\in\mathcal{Z}}$ is a basis of W_j. It is obtained by a k-translation and a j-scaling of function ψ. Spaces $\{W_j\}$ are called wavelet subspaces of $L^2(\Re)$ related to scaling function $\phi(x)$; they satisfy the following properties: $V_{j-1} = V_j \oplus W_j$; $W_j \perp V_j$; $W_j \perp W_k,\ j \neq k$; $L^2(\Re) = \oplus_{j\in\mathcal{Z}} W_j$. Functions $\phi(x)$ and $\psi(x)$ satisfy the follow-

ing relationships: $\psi = \sqrt{2}\sum_{k\in\mathcal{Z}} g_k(2x-k)$; $\phi = \sqrt{2}\sum_{k\in\mathcal{Z}} h_k\phi(2x-k)$. Series $\{h_k\}$ and $\{g_k\}$ are considered as filters: in the orthogonal case the relationship $g_k = (-1)^k h_{1-k}$ is valid. This orthogonality permits the WDCs $d_j(k) = \langle f, \psi_{jk}\rangle$ and the WACs $a_j(k) = \langle f, \phi_{jk}\rangle$ of any function $f(x)$ to be obtained by inner product with the corresponding basis functions. In practice, the decomposition is carried out just over a finite number of scales J. Wavelet Transform with a depth J is then given by:[16]

$$f(x) = \sum_{j=1}^{J}\sum_{k\in\mathcal{Z}} d_j(k)\psi_{jk} + \sum_{k\in\mathcal{Z}} a_J(k)\phi_{Jk}. \tag{3}$$

Considering the case of frequency domain analysis, sequence $\{h_k\}$ is known as low-pass filter, whereas $\{g_k\}$ is called high-pass filter. Their lengths depend on the support of basis functions ψ and ϕ.

3.1.2. *Independent Component Analysis*

In NDT/NDE it is possible to find such problems involving mixed signals and blind source separation, with special reference to noise reduction in both ECT and UT. For the related non-linear phenomena, and for non-gaussian signals (i.e., the lift-off effect in ECT), it is very difficult to separate the informative signal from the noise. In this case, blind source separation techniques, and Independent Component Analysis (ICA) could be an interesting aid to scientists. The general framework for ICA was introduced by Jeanny Herault and Christian Jutten in 1986 and was most clearly stated by Pierre Comon in 1994. In 1995, Tony Bell and Terry Sejnowski introduced a fast and efficient ICA algorithm based on infomax, a principle introduced by Ralph Linsker in 1992. In 1997, Shun-ichi Amari realized that the infomax ICA algorithm could be improved by using the natural gradient, which was independently discovered by Jean-Francois Cardoso. However, the original infomax ICA algorithm with sigmoidal nonlinearities was only suitable for super-Gaussian sources. Te-Won Lee developed an efficient extended version of the infomax ICA algorithm that is suitable for general non-Gaussian signals.[17] Let us consider a signal $\mathbf{s}(\cdot)$ in time- or spatial-domain, which is registered by j receivers and is the result of mixing of j sources, such that $\mathbf{s}(\cdot) = \{\mathbf{s}_1(\cdot), \mathbf{s}_2(\cdot), ..., \mathbf{s}_j(\cdot)\}$, where $\mathbf{s}_k(\cdot) = \sum_{h=1}^{j} a_{kh}\mathbf{x}_k(\cdot)$, being $\mathbf{x}_k(\cdot)$ the h-th source. Under some assumption, it is possible to recovery the set of j sources by calculating a suitable mixing matrix $\mathbf{A}$, i.e. the matrix with elements a_{kh}. It is the aim of ICA, where each mixture $\mathbf{s}_k$ as well as each Independent Component (IC) $\mathbf{x}_h$ is a random variable. Mixtures must have zero mean and unitary variation (whitening process). Once $\mathbf{A}$ is calculated, it is possible to obtain its inverse $\mathbf{A}^{-1}$ and retrieve ICs having non-gaussian distributions. There are many algorithms available in the literature which do ICA. A largely used one, including in industrial applications, is the FastICA algorithm, based on the fixed-point algorithm.[18] Since it is possible to demonstrate that ICs must be non-gaussian,

FastICA exploits kurtosis as cost function measuring the non-gaussianity. But, since kurtosis is not the best parameter to evaluate the non-gaussianity of the ICs, FastICA algorithm uses also the concept of differential entropy H, also known in this framework as negentropy $\hat{H}$. It is so defined: $\hat{H}(\mathbf{y}) = H(\mathbf{y_{gauss}}) - H(\mathbf{y})$, where $\mathbf{y_{gauss}}$ is a gaussian random variable of the same covariance matrix as $\mathbf{y}$. Calculation of $\hat{H}$ is very difficult, and Hyvärinen[17] introduced an approximation based on the maximum-entropy principle: $\hat{H}(\mathbf{y}) \approx \sum_{i=1}^{p} \kappa_i \left[E\{G_i(\mathbf{y})\} - E\{G_i(\nu)\}\right]^2$, where κ_i are some positive constants, and ν is a Gaussian variable of zero mean and unit variance (i.e., whitened or standardized). The variable $\mathbf{y}$ is assumed to be whitened, and the functions G_i are some nonquadratic functions. The following choices of G have proved very useful: $G_1(u) = \frac{1}{\alpha_1} log\left(cosh(\alpha_1 u)\right)$; $G_2(u) = -exp\left(-\frac{u^2}{2}\right)$, where $1 \leq \alpha_1 \leq 2$ is a suitable constant. ICA can be also considered as an extension of the well-known Principal Components Analysis (PCA).

3.2. *Soft Computing Techniques Useful to Solve Inverse Problems in NDT*

Soft computing techniques like Neural Networks (NNs), Fuzzy Logic (FL), Evolutionary Computation, are nowadays widespread in engineering. Starting from large database, they can give the key to solve ill-posed inverse problems, by optimizing the mapping between physical quantities, or by minimizing suitable functional as in the case of Genetic Algorithms (GAs) or Particle Swarm Optimization (PSO). The following paragraphs give a brief theoretical description of some soft computing techniques useful to our NDT/NDE purposes.

3.2.1. *Fuzzy Inference System and Adaptive Neuro-Fuzzy Inference System*

The Fuzzy Inference Systems (FISs) and the Adaptive Neuro-Fuzzy Inference Systems (ANFISs) are popular computing frameworks based on the concepts of fuzzy sets theory, fuzzy if-then rules and fuzzy reasoning.[19] FISs could be useful in order to implement algorithms able to solve inverse problems based on experts' suggestions. The basic structure of a fuzzy system consists of three conceptual components: a rule base, which contains a selection of fuzzy rules; a database or dictionary, which defines the membership functions used in the fuzzy rules and also called Fuzzy Membership Functions (FMFs); a reasoning mechanism, which performs the inference procedure to derive a reasonable output or conclusion. A FIS implements a nonlinear mapping from its input space to output space. The inputs of the procedure are considered as fuzzy variables, and characterized by FMFs. Each FMF expresses a membership measure for each of the linguistic properties. FMFs are usually scaled between zero and unity, and they overlap.

The method usually used to built a FIS, i.e. the fuzzy modelling, defines the optimal number of clusters and their initial values for initializing an iterative optimization algorithm that minimizes a suitable cost function in order to reach a high

quality representation of the original data. Once the optimal cluster configuration has been selected, the set of input/output data can be used for model identification, by considering each cluster centre as fuzzy rule. The ANFIS is a way of applying to FISs learning techniques, as it can be found in NNs literature.[20]

FL can also be interpreted as a pattern classification problem-solver (a pattern defect to belong to different classes at the same time, according to the concept of partial membership $0 \leq u_{jk} \leq 1$). According to Kosko,[21] the concept of Fuzzy Entropy (FE), H, can be introduced is defined as follows: $H = \sum_{k=1}^{N} u_{jk} \ln(u_{jk})$ where $\ln(u_{jk}) = 0$ when $u_{jk} = 0$. Kosko also introduced a typical fuzzy operator, i.e. the so called Subsethood Operator (SO).[21] It measures how a particular pattern belongs to a given fuzzy set; therefore, $0 \leq SO \leq 1$. This operator can be helpful for the discrimination of middle-depth defect in metallic plates.

3.2.2. *Short Summary on Support Vector Machines*

Support Vector Machines (SVMs) are another useful soft computing technique applicable to the solution of ill-posed classification or regression problems. SVMs was introduced by Vapnik within the framework of the Statistical Learning Theory,[22] which describes the principle of Structural Risk Minimization (SRM). It has been demonstrated to be more effective than the traditional Empirical Risk Minimization principle employed by Artificial Neural Networks.[23] A SVM-based classifier attempts to locate a hyperplane that maximises the distance from the members of each class to the same hyperplane. Due to SRM, the error probability is upper bounded by a quantity depending by both the error rate achieved on the training set and a measure of the "richness" of the set of decision functions it can implement (named "capacity", or Vapnik Chervonenkis dimension). SVMs can be also applied to regression problems (see[24] and references therein).

3.2.3. *Particle Swarm Optimization: A Brief Overview*

Particle Swarm Optimization (PSO) is a population based stochastic optimization technique developed by Eberhart and Kennedy,[25] inspired by social behaviour of bird flocking and used with nonlinear continuous functions. PSO-based systems learn from the scenario how to solve the optimization problems. In PSO, each single solution is defined as a particle in the multidimensional search space. The PSO algorithm evolves a population of particles called swarm. All of the particles have fitness values which are evaluated by the fitness function to be optimized, and have velocities which direct the convergence of the particles. During iterations, each particle adjusts its position making use of the best position encountered by itself (the so called *pbest*) and its neighbours (*gbest*). Each particle updates its current velocity and position in the search space using historical information regarding its own previous best position as well as the best position discovered by all other particles or neighbouring particles.[26]

4. Study-Cases of CI in NDT/NDE

This section will present some selected applications of the above introduced CI methodologies at the cases of study previously formulated. Of course, it is not the intent of the authors to give an exhaustive report of the applications that can be found in the literature. The idea is to convey the message that an interdisciplinary study could be the key to go beyond the limits of each technique. The use of suitable sensors and probes could really motivate on its own the use of a specific algorithm and, in turn, suggest some guidelines to its implementations.

For example, the Fluxset sensor®,[27] provided by the Research Institute for Technical Physics and Materials Science (Fig. 3), has been used in the ECT laboratory measurements. In this sensor, a phase demodulation of the pick-up voltage V_p gives us a measurement proportional to the magnetic field parallel to the plane of the specimen. The amount of exciting coil turns is 93. The lift-off of the sensor core is about 0.75 mm.

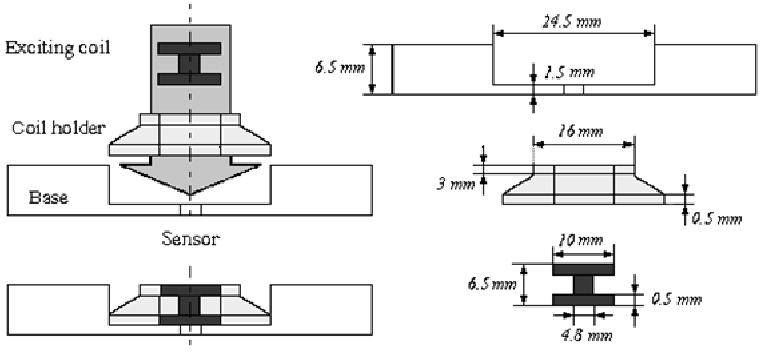


Fig. 3. Fluxset® probe with its 3 coils: exciting (to induce eddy currents in the specimen), driving (to saturate the core), pick-up (sensor).

4.1. *Lift-Off Effect Reduction in ECT*

The lift-off effect can be reduced by using the Wavelet Transform and the ICA. The layout of the experiments included an inspected specimen which is a square plate 8x8x0.125 cm made of INCONEL600 ($\sigma = 10^5$, $\mu = \mu_0$). It presents a rectangular thin crack (9 mm in length and 0.2 mm in width) in the central region; depth of crack is 20% of the plate thickness. The specimen was inspected by both inner defect (ID) and outer defect (OD) analysis, using a 110.7 mA (rms) exciting current at frequency 20 KHz. The measured quantity is $\|V_p\|$. Scanned area was a 40x40 mm region, referred to the plate center, with 0.5 mm spacing along x and y direction ($-20mm \leq X \leq 20mm$, $-20mm \leq Y \leq 20mm$).[28] Figure 4 show the influence of the lift-off effect: the defect signal is hardly extractable.

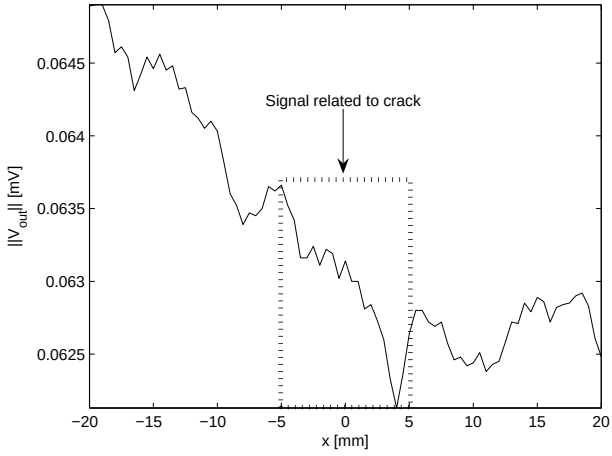


Fig. 4. $\|V_{out}\|$ measured at $Y = 0$, $-20 \leq X \leq 20$ in the ID analysis: it is possible to denote the lift-off noise into the signal trend.

Thus, MRA can be used to separate the noise from the useful components by decomposition of measured signal and thresholding the WACs in the subsequent reconstruction. For instance, a Daubechies10 Wavelet[29] decomposition into five frequency levels can assures to restrict the lift-off noise into a_5 (cfr. Eq. 3 with $J = 5$). Thus, inverse transforming by excluding a_5 can reduce the lift-off effect on eddy current signal.[5] ICA can be also exploited to solve the same problem. In the study-case, signal $\mathbf{s}$ is considered as the mixing of $\|V_p\|$ for both ID and OD analysis. Then the component related to lift-off noise can be discriminated to the one representing the useful denoised $\|V_p\|$ signal. Experimental tests proved that the ICA method shows less micro-oscillations but higher lift-off at the top and bottom side of inspected area.[3] Further advantages can be obtained by a joint use of MRA and ICA, especially by firstly applying ICA on the mixture of ID and OD ECTs and subsequently processing the useful component by means of Daubechies10 at five multi-resolution scales (ICAWAV approach). The specular method firstly applies MRA on both ID and OD analysis and subsequently extracts their ICs (WAVICA approach). Results, shown in Fig. 5, emphasize the influence of defect in $\|V_p\|$ and strongly reduce the lift-off noise even at the borders and corners of examined area. Table 1 resumes all obtained results in terms of SNR values.

Table 1. SNR (in dB) obtained by the exploited pre-processing techniques.

SNR_{ID}	SNR_{MRA}	SNR_{ICA}	SNR_{WAVICA}	SNR_{ICAWAV}
14.623	19.129	16.954	19.127	19.192

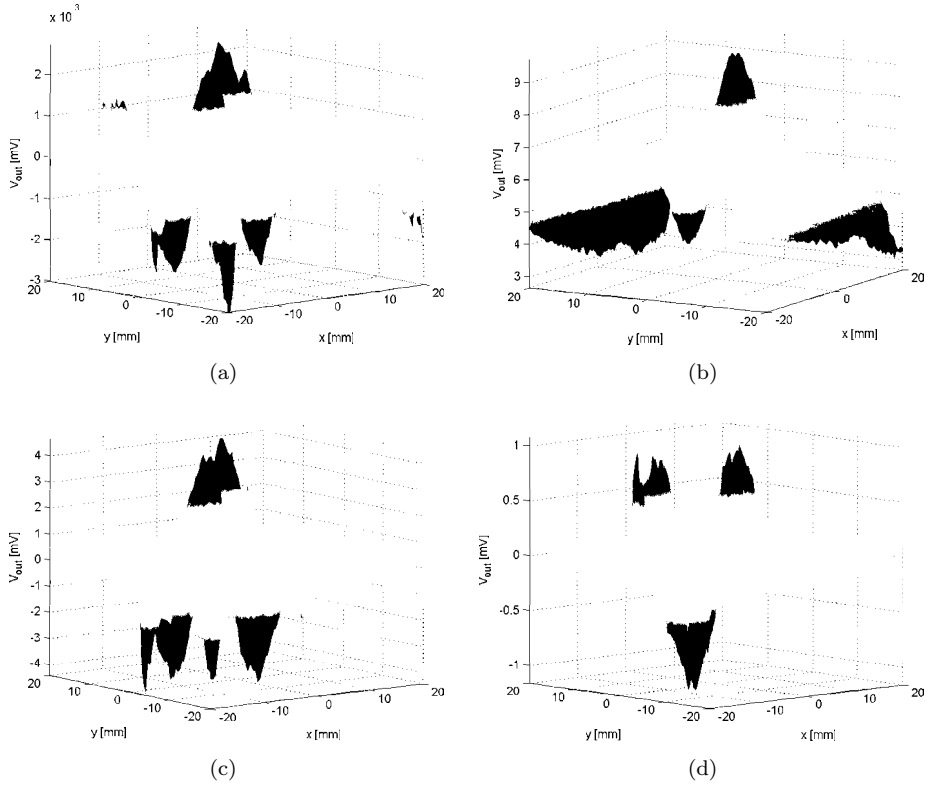


Fig. 5. Results of lift-off denoising: (a) MRA on ID signal; (b) IC representing the useful signal obtained by ICA approach; (c) useful IC obtained by proposed hybrid WAVICA method; (d) ICAWAV result.

4.2. *Characterization of In-Depth Location of Defects*

FL can be exploited by implementing a suitable FIS and subsequently using FE and SO, in order to solve the problem of detection of middle-depth cracks for such materials as INCONEL600[a] used in various industrial frameworks.

In our experimentations, a number of INCONEL600 plates have been considered (each one has been tagged with a suitable identification code). Here, there are either sub-superficial and deep cracks, i.e. cracks having a depth z into the range $20\% \leq z \leq 60\%$ or $40\% \leq z \leq 100\%$ of plates' thickness, respectively. ECTs have been carried out by using the Fluxset® sensor. An external sinusoidal exciting field has to be generated in the vicinity of each specimen; a current of 107 mA rms at

[a]INCONEL600 is an alloy composed by 70% nichel, 15% chrome and 8% iron, having an excellent resistance to corrosion and high temperatures. It is used, for instance, to build steam generators, but is easily subjected to cracking, pitting, denting and other type of degradation.

a frequency of 1 kHz has been considered adequate in our experimentations. A driving signal - triangular shape, 125 kHz frequency, 2Vpp amplitude - was applied to saturate the core material inside the probe. 48 full scans were run.

An MRA, in particular using a Daubechies2 at level 4, turns out useful to select suitable features (i.e. the FIS' inputs) of the defect deepness starting from the experimental measurements. In the present case of study, the WDCs have been considered; into the same eddy current signal, their values vary at least of one order of magnitude between presence or absence of defect. Moreover, WDCs appears able to discriminate also the deepness of the crack.

Therefore, the 8 FIS inputs are: the peak-value of $\|V_p\|$ ($|V_{peak}|$, [mV]); the current applied by the signal generator $|i_{gen}|$, [mA]; the frequency of the exciting sinusoidal signal f_{signal}, [kHz]; the five Wavelet Detail Coefficients (WDC) concerning the region where a crack is found. Let us remark how crack's width does not influence either the value of $|V_{peak}|$, only depending by the crack depth, or WDCs, since it involves macro-oscillations, (i.e. WACs) of each ECT-related signal. The FIS' output is the following class codification: 1 for the cracks having a depth between 20% and 60%, 2 for the cracks with $40\% \leq z \leq 100\%$. In our experimentation, the initial database has been split up into 40 training patterns ad 8 test patterns. The FE has been considered as the classifying quantity, in order to evaluate the fuzziness degree of a fuzzy classification. Table 2 shows the classifica-

Table 2. FE-based simulation of classification.

Specimen	Observed classification	Simulated classification	FE value
jsaem#41	1	1	$FE_{1,41}$=0.318; $FE_{2,41}$=0.351
jsaem#42	1	2	$FE_{1,42}$=0.322; $FE_{2,42}$=0.102
jsaem#43	2	2	$FE_{1,43}$=0.052; $FE_{2,43}$=0.032
jsaem#44	1	2	$FE_{1,44}$=0.094; $FE_{2,44}$=0.008
jsaem#45	2	1	$FE_{1,45}$=0.004; $FE_{2,45}$=0.181
jsaem#46	2	1	$FE_{1,46}$=0.003; $FE_{2,46}$=0.116
jsaem#47	1	1	$FE_{1,47}$=0.009; $FE_{2,47}$=0.200
jsaem#48	2	1	$FE_{1,48}$=0.007; $FE_{2,48}$=0.195

Table 3. Identification of middle-depth cracks by subsethood values.

Specimen	$S_{1,i}$	$S_{2,i}$	Middle-depth crack
jsaem#41	0.636	0.408	No
jsaem#42	0.568	0.521	Yes
jsaem#43	0.392	0.816	No
jsaem#44	0.554	0.481	Yes
jsaem#45	0.539	0.524	Yes
jsaem#46	0.590	0.550	Yes
jsaem#47	0.783	0.348	No
jsaem#48	0.569	0.536	Yes

tion of 8 test patterns; 5 patterns have been wrongly classified. Maybe, it can be due to location of cracks inside the transition region $40\% \leq z \leq 60\%$. In order to prove this assumption and then improve classification quality, SO of each i-th test pattern for each k-th class can be used. Table 3 shows SO values for misclassified signals; it proves how errors were due to the fact that misclassified cracks lay in middle-depth of inspected specimens.

4.3. *Identification of the Kind of Defect in CFRP*

Characterization of defects on CFRP starting from UTs is a typical pattern classification in which CI can turns out very useful. Please be careful that heuristic machines do not have to suffer of the "curse of dimensionality" (see the end of Sec. 2.2), in order to work in real time. In this sense, it is possible to jointly use MRA and PCA as feature extractors, and SVM as classifier. Proposed experimental approach exploits a database provided by the Alenia Aeronautica SpA. It is composed by 300 ultrasonic pulses for each delamination-like and inclusion flaws (equally partitioned according to location at top, middle or bottom of the inspected specimen's thickness), 206 signals of porosity and 100 signals showing absence of defects. Values of samples for each pulse are between 0 and 255, i.e. they was acquired by using a 8 bit AD converter. A 5 MHz UT probe has been used, with 100 MHz sampling frequency; speed of ultrasonic wave on CFRP is 3000 m/s. Firstly, a suitable MRA has been carried out on each ultrasonic pulse by means of a Daubechies6-level3 DWT.

The considered 58 WACs have been subsequently reduced to 10 elements (inputs of SVM) by means of PCA, cutting out all the Principal Components (PCs) whose contributions to total variation of the whole set of PCs are less than 2%. The threshold imposed on PCs causes an average informative loss equal to $0.68 * 10^{-2}\%$ (see Fig. 6).

560 pulses have been used to obtain the training patterns (210 delaminations, 210 inclusions, 70 porosities and 70 having no defect). Remaining signals (i.e. 90 delamination, 90 inclusion, 136 porosity and 30 no defect) compose the test set. The output of the SVM is a number $n \in \mathcal{N}$ which codifies each kind of defect as follows: 50 for delamination, 100 for inclusion, 200 for porosity and 0 if no defect is present. SVM has been trained by using a 2-degree polynomial kernel $K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma * \langle \mathbf{x}_i, \mathbf{x}_j \rangle + \chi)^2$, where $\gamma = 1$ and $\chi = 1$ for the case of study. At the end of the training phase, SVM returns a particular matrix, so called Confusion Matrix, which evaluates the goodness of a trained classifier. In fact, generally speaking, the element CM_{ij} of a Confusion Matrix is the probability that a single pattern belonging to the i-th class could be classified as belonging to the j-th class (sum of elements of each rows is therefore equal to 1). Thus, the more the Confusion Matrix is similar to the identity matrix, the better classification performances are. In the case of study, the norm of Confusion Matrix is equal to 1.007, very close to the norm of a four-rank identity matrix (i.e. equal to 1).

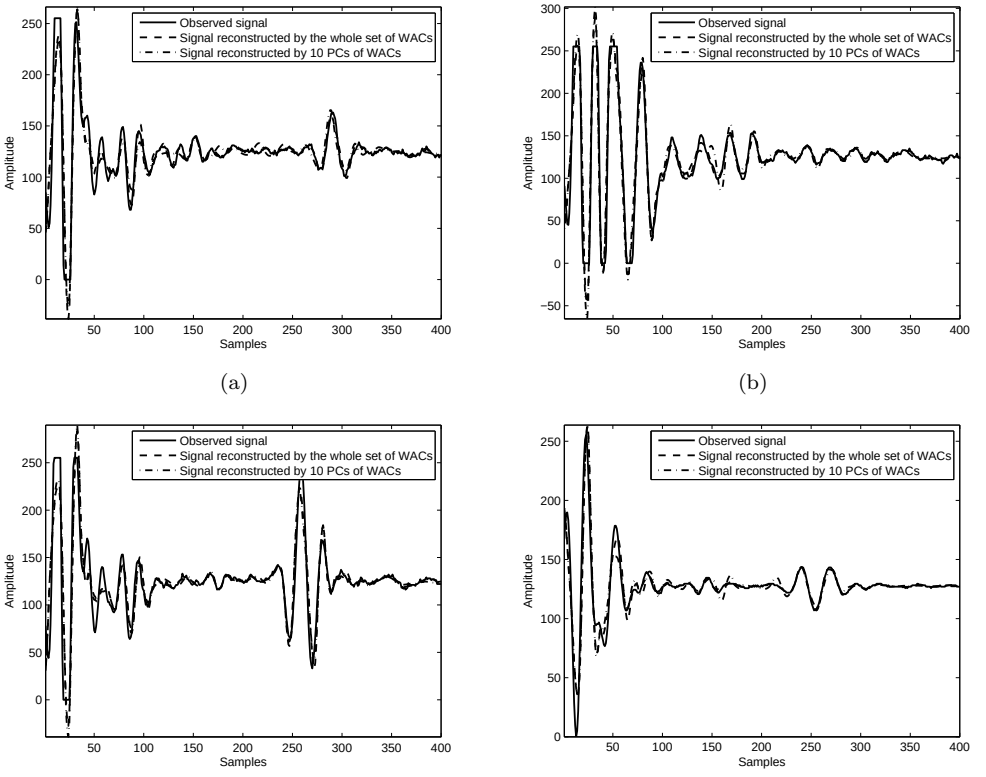


Fig. 6. Comparation of reconstruction of ultrasonic signals by using the whole WACs set or the 10 PCs; example of signals detecting: (a) delamination-like defect, (b) inclusion, (c) porosity, (d) no defect.

Obtained classification reliability is equal to 99.711%. The unique error is a so called false warning, that is a delamination-like flaw detected by the SVM classifier where instead no defect is. In general, false warnings are not a so big problem in an industrial framework, if compared to non-detected flaws and if they stay in a limited number of cases. Figure 7(b) shows a quantitative description of classification results.

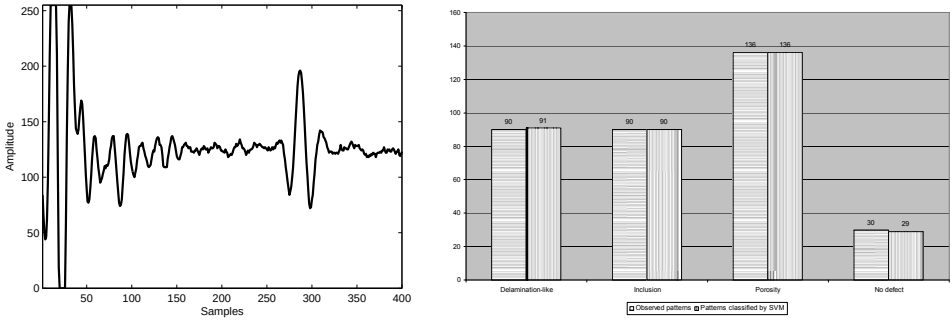


Fig. 7. (a) The ultrasonic pulse misclassified by the implemented SVM; (b) quantitative diagram of classification results.

4.4. *Determination of Mechanical Stresses*

Two sets of specimens, named SETMONOLOAD and SETBILOAD respectively, made by structural steel have been used to prove what described in Sec. 2.4. SETMONOLOAD and SETBILOAD were deformed under a 172 kN mono-axial and bi-axial loads respectively (Fig. 8).

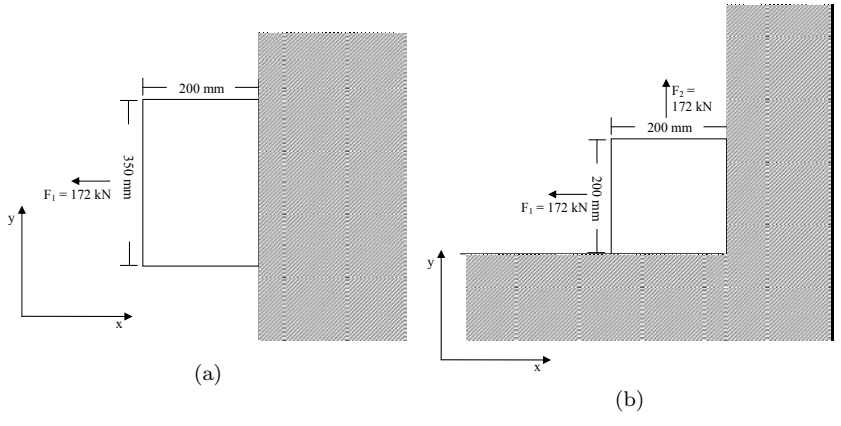


Fig. 8. Examples of applied tractions: (a) mono-axial and (b) bi-axial loads.

Table 4. Experimental specifications for inspected plates and exciting signals.

Range of dimensions [mm]	Range of applied exciting frequencies [kHz]	Range of applied sinusoidal exciting current [mA rms]
[120x120x5 ÷ 350x350x10]	[1 ÷ 1.2]	[100 ÷ 170]

Thus, ECTs have been carried out on each plate of both sets, applying a sinusoidal excitation (see Table 4 for experimental specifications). Let us denote how the plane-case hypothesis is applicable since the thickness of specimens is irrelevant compared to length and width. The Fluxset® coil was used to map a central portion of each specimen. The driving signal was set to have triangular shape, 100 kHz frequency, 2 Vpp amplitude. Since structural steel does not suffer of piezoelectric or magnetostrictive phenomena, variation of V_p is due to the local variation of electric conductivity caused by the necking phenomenon. Thus, complete eddy current maps have been experimentally obtained, whereas a Finite Element Method based software (i.e. Comsol Multiphysics® 3.2a) has been exploited in order to reconstruct the respective stress maps. Generally, the assumption of equivalence has been experimentally tested. Figure 9 depicts the comparisons between eddy current and stress maps for the specimen SETMONOLOAD#7 (in terms of $\|V_p\|$), whereas "shear bands" on the specimen SETBILOAD#1 are shown by the eddy current maps of Fig. 10, with most likely incipient damages at top and bottom-left zone of inspected area.

A suitable Soft Computing approach can be useful to estimate the eddy current map starting from easily calculable mechanical quantities when ECTs are unfeasible. The deformation ΔL_x for a mono-axial load, or the single x- and y-component ΔL_x and ΔL_y for a bi-axial load, are the considered mechanical quantities. They can be calculated as $\Delta L_x = \frac{P * L_{0_x}}{E}$ and $\Delta L_y = \frac{P * L_{0_y}}{E}$, where P is the intensity

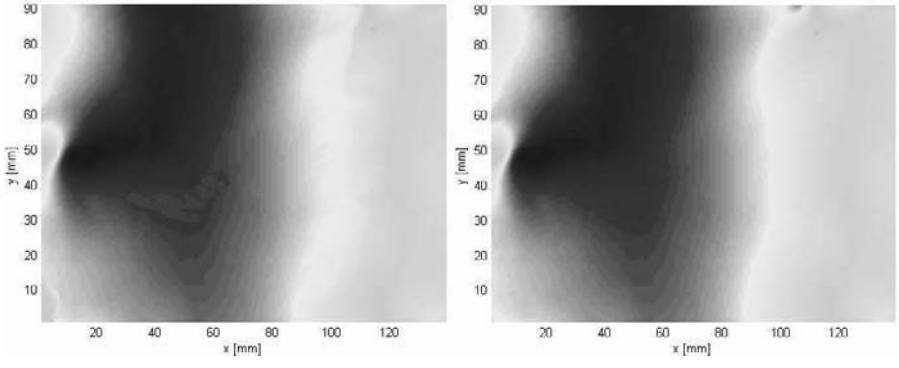


Fig. 9. Comparison between the stress map (at left) and the ECT map (at right) for SETMONOLOAD#7. Color hues represent intensity of mechanical stress (stress map) or $\|V_p\|$ (ECT map): dark hues are used for high values, light hues for low values.

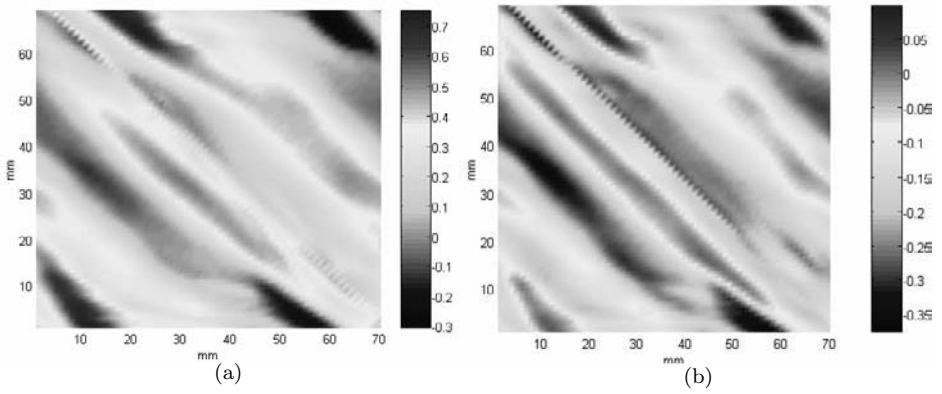


Fig. 10. Maps for the inspected SETBILOAD#1: (a) 2D map of $Re(V_p)$; (b) 2D map of $Im(V_p)$.

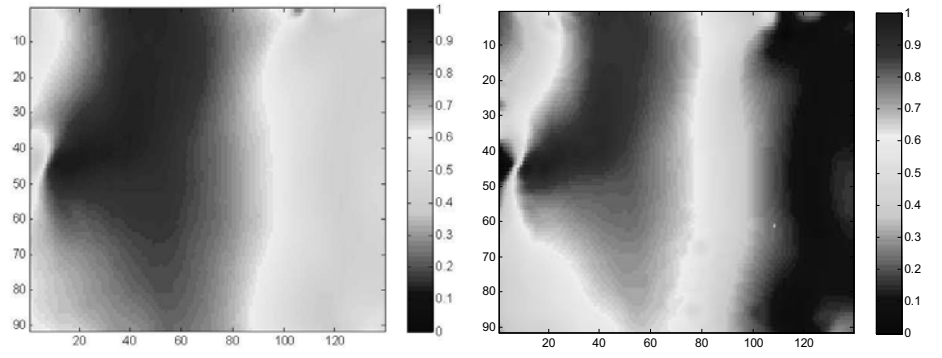


Fig. 11. Maps of: $\|V_p\|$ ([V]) of the measured signal (at left); magnitude of the signal reconstructed by SVRM-based estimation (at right). They have been normalized between 0 and 1 in order to improve the comparison.

of mechanical pressure in Pa, E is the Young's elastic modulus (i.e. 210 GPa for the structural steel), and L_{0_x} and L_{0_y} are initial lengths along x- and y-axis respectively. By dividing each plate in 5x5 mm squares and selecting them randomly, the tuples $\{x_i, y_i, \Delta L_x\}$ and $\{x_i, y_i, \Delta L_x, \Delta L_y\}$ have been considered as the inputs of the heuristic system for mono- and bi-axial loads respectively. On the other hand, $\|V_p\|$ and the couple $(Re(V_p), Im(V_p))$ represent the correspondent output, respectively. For the case of mono-axial traction, a ε-SVM based regressor has been exploited. Generally, obtained results are very encouraging, with an average estimative performances of about 98%. The worst estimation is depicted on Fig. 11 (SETMONOLOAD#7): in this case performances slightly decrease to 97.6%.

Here, the high-stressed area has been reconstructed in an acceptable way, showing the differently hued shear bands and the presence of incipient damages at top-left of inspected area. Similar results have been obtained by exploiting the SETBILOAD database and using an ANFIS-based approach as heuristic method.[30]

4.5. *Imaging of Corrosion by Means of PSO*

Theoretical approaches described in Sec. 2.3 and Sec. 3.2.3 can be exploited on experimental measurements. A 104.5x95x6 mm stainless steel plate has been used in our experimentation. Here, a galvanic corrosion has been simulated by a circular milling with a 2 mm diameter and an in-depth extension equal to 2 mm (see Fig. 12 for details). Inspection has been carried out on the surface at the opposite side than corroded one. A 40x40 mm portion (centered on defect) of the specimen has been investigated this way by means of the the Fluxset® probe, with a 123 mA rms external sinusoidal exciting current at a frequency of 877 Hz. A driving signal - triangular shape, 125 kHz frequency, 2 Vpp amplitude, 50 Ω output impedance - has been applied to saturate the core material inside the probe. The magnitude map of impedance variation has been reconstructed (Fig. 13).

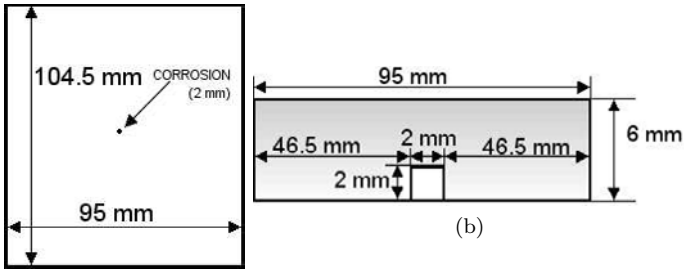


Fig. 12. A particular of stainless steel specimen used for our experimentations: (a) bottom-view, (b) front-view.

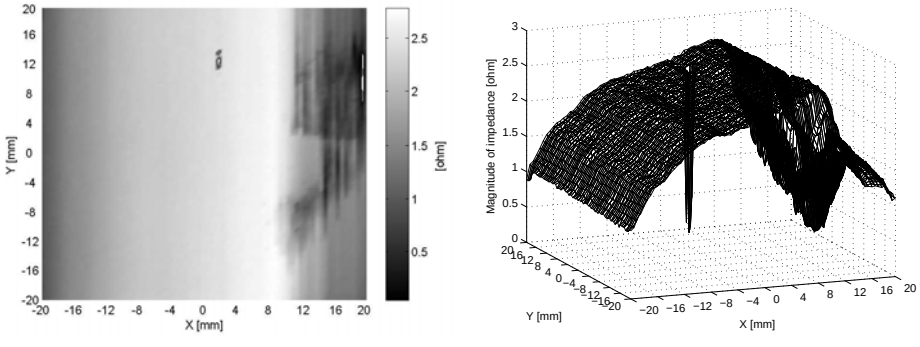
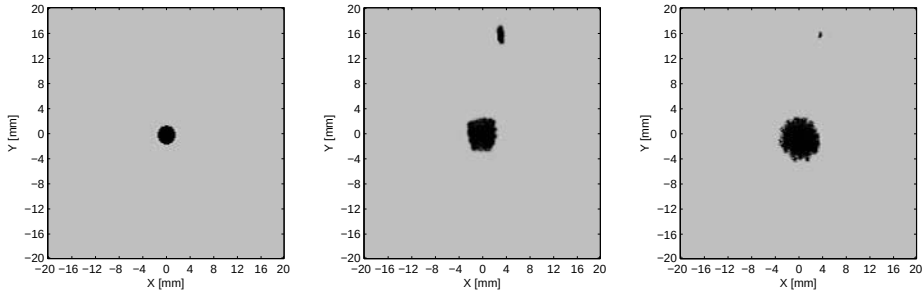


Fig. 13. (a) 2D and (b) 3D reconstructions of the $\|Z(x_0, y_0)\|$ trend over the whole of inspected surface; pitch clearly visible in 3D reconstruction is maybe due to a local roughness of material generating a marked lift-off effect.

The problem of the imaging of corrosion can be solved by an opportune regularization of Eq. (2), i.e., by a PSO based minimization of the misfit $J(\theta) = \sum_j \left| \int\int K(x_{0j} - x, y_{0j} - y)\theta(x, y)dxdy - g_j \right|^2$,[9] with $g_j = Z(x_{0j}, y_{0j})$. Iterations have been early stopped if a likely solution is reached up, i.e. *gbest* value does not change by at least the fixed lowest error gradient tolerance for the fixed max number of epochs. Obtained results have been compared with the Least Square Minimization (LSM) method, and are depicted in Fig. 14. LSM as well as PSO detect a corrosion where the pitch due to a locally misfit measurement (maybe due to a local roughness of material causing a marked lift-off effect) is located. Therefore, it is not a misclassification of regularization methods: both appear very robust and efficient, but PSO has a lower computational time (105.17s against the 830.87s of LSM method), and is able to minimize better than LSM the measurement error.



(a) (b) (c)

Fig. 14. Scanned area of analysed plate and graphical results of corrosion detection: (a) 2D schematization of scanned area; reconstruction of corrosion by (b) LSM, (c) PSO.

5. Acknowledgments

Authors thank Prof. Gasparics, for providing the dataset used in Sec. 4.1. We also acknowledge Alenia Aeronautica SpA, for data and technical support needed to Sec. 4.3.

References

1. F. C. Morabito, Independent component analysis and feature extraction techniques for ndt data, *Mater. eval.* **58**(1), 85–92, (2000). ISSN 0025-5327.
2. K. Hwang, S. Mandayam, S. S. Udpa, L. Udpa, W. Ford, and M. Afzal, Characterization of gas pipeline inspection signals using wavelet basis function neural networks, *NDT&E International.* **33**, 531–545, (2000). ISSN 0963-8695.
3. M. Cacciola, A. Gasparics, F. C. Morabito, and M. Versaci, Advances in signal processing to reduce lift-off noise in eddy current tests, *PIERS On line.* ISSN 1931-7360. accepted.
4. G. Simone and F. C. Morabito. ICA-NN based data fusion approach in ECT signal

restoration. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN2000)*, vol. 5, pp. 59–64, Como, Italy, (2000).
5. G. Chen, A. Yamaguchi, and K. Miya, A novel signal processing technique for eddy-current testing of steam generator tubes, *IEEE Trans. on Magnetics*. **34**(3), 642–648, (1998). ISSN 0018-9464.
6. M. Buonsanti, M. Cacciola, S. Calcagno, F. C. Morabito, and M. Versaci, Fuzzy computation for classifying defects in metallic plates, *International Journal of Applied Electromagnetics and Mechanics*. ISSN 1383-5416. in press.
7. S. Calcagno, F. C. Morabito, and M. Versaci, A novel approach for detecting and classifying defects in metallic plates, *IEEE Transactions on Magnetics*. **39**(3), 1531–1534, (2003). ISSN 0018-9464.
8. M. Cacciola, D. Costantino, F. C. Morabito, and M. Versaci. Dynamical analysis for flaw shape identification in non linear eddy current tests. In *XIX Symposium Electromagnetic Phenomena in Nonlinear Circuits (EPNC'06)*, pp. 65–66, Maribor, Slovenia, (2006).
9. B. Luong and F. Santosa, Quantitative imaging of corrosion in plates by eddy current methods, *SIAM Journal of Applied Mathematics*. **58**(5), 1509–1531, (1998). ISSN 0036-1399.
10. P. G. Kaup, F. Santosa, and M. Vogelius, A method for imaging corrosion damage in thin plates from electrostatic data, *Inverse Problems*. (12), 279–293, (1996). ISSN 0266-5611.
11. M. Cacciola, S. Calcagno, F. C. Morabito, and M. Versaci, Swarm optimization for imaging of corrosion by impedance measurements in eddy current test, *IEEE Transactions on Magnetics*. ISSN 0018-9464. in press.
12. A. Nadai, *Theory of flow and fracture of solids, Vol. I.* (McGraw-Hill, New York, 1950).
13. J. Blitz, *Electric and Magnetic Methods of Non-destructive Testing*. (Chapman & Hall, London, 1995), 2nd edition.
14. C. K. Chui, *An Introduction to Wavelets*. (Academic Press, San Diego, 1992).
15. B. B. Hubbard, *The World According to Wavelets: The Story of a Mathematical Technique in the Making*. (AK Peters, Wellesley, USA, 1998).
16. A. Graps, An introduction to wavelets, *IEEE Comp. Science and Engineering*. **2**(2), 50–61, (1995). ISSN 1070-9924.
17. A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. (John Wiley & Sons, New York, 2001).
18. A. Hyvärinen, Fast and robust fixed-point algorithms for independent component analysis, *IEEE Trans. on Neural Networks*. **10**(3), 626–634, (1999). ISSN 1045-9227.
19. L. A. Zadeh, Fuzzy logic, *Computer*. **1**(4), 83–93, (1988). ISSN 0018-9162.
20. R. Jang and C. Sun, Neuro-fuzzy modeling and control, *IEEE Transaction on Systems, Man and Cybernetics*. **23**(3), (1993). ISSN 0018-9472.
21. B. Kosko, *Fuzzy Engineering*. (Prentice Hall, Upper Saddle River, New Jersey, 1997).
22. V. N. Vapnik, *The Nature of Statistical Learning Theory*. (Springer Verlag, New York, 1995).
23. C. Cortes and V. N. Vapnik, Support vector network, *Machine Learning*. **20**, 273–297, (1997). ISSN 0885-6125.
24. A. J. Smola. Regression estimation with support vector learning machines. Master's thesis, Technische Universitat, Munchen, Germany, (1996).
25. R. C. Eberhart and J. Kennedy. A new optimizer using particle swarm theory. In *Proceedings of the Sixth International Symposium on Micromachine and Human Science*, pp. 39–43, Nagoya, Japan, (1995).

26. B. Al-Kazemi, Velocity analysis for particle swarm optimization, *WSEAS Trans. on Information Science & Applications*. **3**(4), 751–757, (2006). ISSN 1790-0832.
27. A. Gasparics, C. S. Daroczi, G. Vertesy, and J. Pavo, *Improvement of ECT probes based on Fluxset-type magnetic field sensor*, In eds. T. T. R. Albanese, G. Rubinacci and S. Udpa, *Electromagnetic Nondestructive Evaluation (II)*, vol. 14, *Studies in Applied Electromagnetics and Mechanics*, pp. 146–151. IOS Press, Amsterdam, (1998).
28. A. Gasparics. Probe_98/6 - surface scan results within manodet project, (1997). `http://alag3.mfa.kfki.hu/ect/manodet/results/FS06B/index.htm`.
29. I. Daubechies, *Ten Lectures on Wavelets*. (SIAM: Society for Industrial and Applied Mathematics, London, 1992).
30. M. Buonsanti, M. Cacciola, S. Calcagno, F. Morabito, and M. Versaci. Fuzzy inference models for ec-ndt experimental tests: the domain of bi-phase materials. In *Proceedings of First Euro Mediterranean in Advances on Geomaterials and Structures (AGS'06)*, pp. 725–730, Hammamet, Tunisia, (2006).

CHAPTER 22

SPLIT SPECTRUM PROCESSING AND AUTOMATIC FLAW DETECTION

Kevin D. Donohue

Department of Electrical and Computer Engineering, University of Kentucky
Lexington, KY, 40506, USA
donohue@engr.uky.edu

The integrity of materials used in applications such as nuclear power plants and aircraft engines is tested periodically for defects that could potentially lead to part failures during critical operations. Ultrasonic flaw detection plays an important part in the nondestructive evaluation (NDE) of these materials. This chapter presents signal processing approaches for enhancing defect detection in materials consisting of grain-like microstructures. Grain scatterers create noise processes with properties similar to those created by flaw scatterers, which makes for a challenging detection problem. This chapter highlights a particular flaw detection approach, referred to as split-spectrum processing (SSP), and explains its performance relative to classical approaches, such as Wiener and matched filtering. Analytical and simulation results are presented to demonstrate the ways in which SSP uses statistical properties of scatterer phase spectra to enhance flaw detection. In addition, examples are included to demonstrate SSP implementations for automatic flaw detection.

1. Introduction

Flaw detection in large-grained materials, such as austenitic stainless-steel, has attracted significant attention in the NDE community [1-12]. Materials composed of large-grained structures pose significant challenges for ultrasonic pulse-echo techniques because scattering from the grain microstructure often obscures material flaw echoes. Figure 1a shows an example of a simulated pulse-echo scan (A-scan) through material with subresolution grain scatterers and a flaw scatterer consisting of a strong coherent scatterer embedded in grain structures. In this example the flaw echo corresponds to the largest A-scan peak; however, other

grain noise peaks can also be considered potential flaws. The discrimination between grain and flaw echoes often requires additional processing or scanning.

Several approaches have been proposed to exploit key differences between the flaw and grain scatterers. Since the flaw and grain scatterers often have different sizes, their scattering responses have different spectral characteristics [13]. These differences have motivated filtering approaches such as Weiner filtering [14-16]. In addition, the flaw echo typically has a coherent structure, where scattered energy results from a single spatial point, while grain echoes are spatially distributed throughout the insonified resolution volume. These differences have motivated spatial and frequency diversity techniques [17-22]. If the insonifying pulse remains consistent as it propagates through the material, and the flaw scatterer echo does not undergo significant distortion as a part of the scatterer response, then a matched filtering approach can be used [8].

This chapter focuses on a frequency diversity approach referred to as split-spectrum processing (SSP), which involves subband filtering A-scans into ensembles of frequency diverse signals and applying novel processing schemes over the spectral bands. Early speech researchers attempted to analyze signals separated into subbands in the hopes of unraveling the signal complexities for automatic speech recognition [23]. While the split-spectrum ensembles looked intriguing, a successful speech recognizer never resulted from this analysis. However, subband analysis of information signals is used with remarkable success in nature. Both the auditory and visual sensory organs in mammals divide the incoming signal into spectral bands creating neural channels, which are sent to various parts of the brain for processing [24,25]. The bat processes acoustic information in this way. It uses broadband echoes to navigate and perform complex tasks, such as detecting food and predators [26,27]. Therefore, based on the wealth of information extracted by biological systems, there is strong motivation to apply SSP in computer-based programs for detection and classification problems.

Flaw detection in the presence of grain noise lends itself well to a split-spectrum approach. The properties of signals from distributed grain and coherent flaw echoes decomposed over multiple spectral bands are well-understood. Early applications of a split spectrum to the flaw detection [28] observed a peak amplitude correlation over the spectral bands at the flaw location, whereas grain echoes exhibited decorrelated amplitudes over the spectral bands. Figure 1 shows a simulated flaw echo embedded in grain noise along with a corresponding subband ensemble. In this case the grain echoes were formed from many unresolvable scatterers per resolution cell, and the flaw echo was generated by a single scatterer. The radio frequency (RF) signal for the A-scan is shown in

Fig. 1a, and the envelopes of 3 subband outputs are shown in Fig. 1b. The subband outputs were normalized for energy to reduce the frequency variation due to the transducer's frequency response. Note the relative amplitude stability at the flaw peak position over all bands. (The RF subbands also exhibit consistent peak amplitudes; however, it is easier to observe this in the signal envelopes as plotted here). It is interesting to note that the dominant grain echo located toward the end of the scan has large amplitudes for the high frequency band, and smaller values for the other bands. This pattern results from key flaw and grain signal properties that has led to several novel processing schemes in SSP.

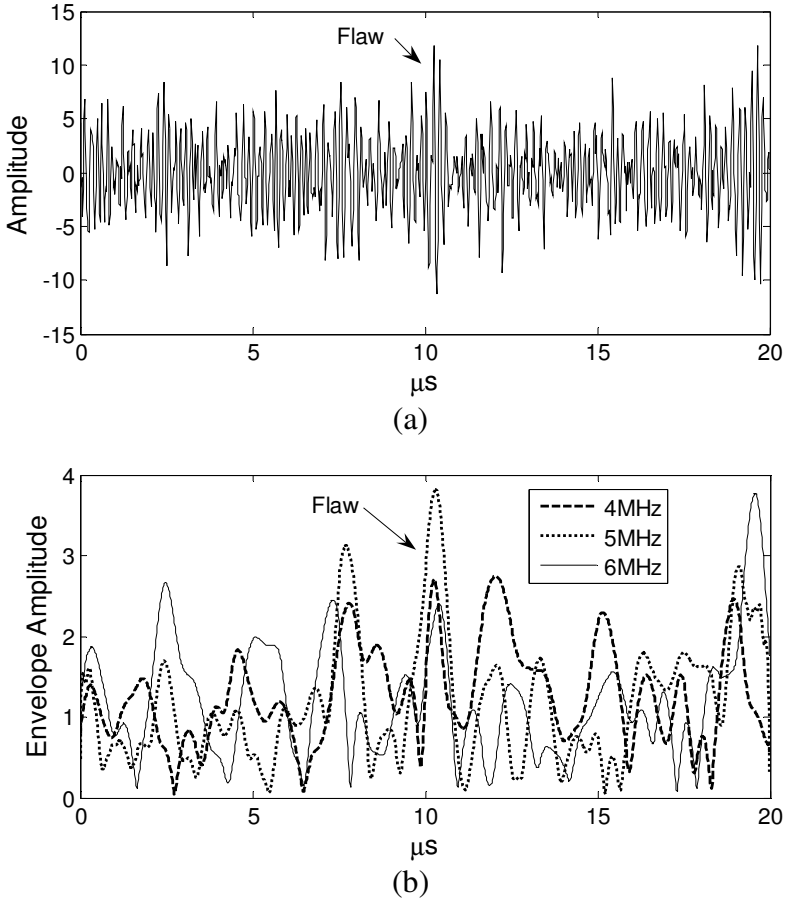


Figure 1. (a) RF A-scan with flaw and grain echoes from a 5 MHz transducer (b) envelopes of subband ensemble (rms normalized) from Gaussian pass-bands at 0.6 MHz bandwidths and center frequencies listed in the legend.

The signal amplitude information that SSP algorithms rely on differs from the signal properties used by classical methods, such as the matched and Wiener filter. This chapter presents these fundamental signal characteristics in order to better understand SSP performance and enhance its application over a broad range of problems. Monte Carlo simulations are used to demonstrate performance differences between approaches that use different types of signal information. The next section briefly reviews previous work related to SSP critical to its development. Section 3 presents a signal model for the flaw and grain echoes along with descriptions of the matched and Wiener filters. Section 4 presents implementation details of the various SSP algorithms and Section 5 discusses setting parameters for creating the subband ensemble. Section 6 presents simulations demonstrating relative performance differences between SSP algorithms and classical approaches. Section 7 shows examples of implementing SSP in an automatic flaw detection system. The chapter concludes with a summary of SSP algorithm performance and application.

2. SSP Development

The early literature presenting SSP techniques presented compelling results for real data based on scans of flat-bottom holes in large grained material. However, because SSP involved nonlinear processes, the signal properties driving SSP performance were not clearly identified. In addition, practical issues such as, determining the best processing parameters, and knowing the limitations of SSP were not clearly explained. Therefore, much of the literature that followed focused on resolving these issues.

Papers introducing SSP cited that the main reason for performance enhancement was result of frequency diversity, which decorrelated the grain echo amplitudes, while leaving the target echo amplitudes highly correlated over the frequency bands [28]. This observation led to the application of various compounding algorithms that involved nonlinear operations (such as envelope detection) and averaging [17-22]. Similar ideas had already been applied in related areas such as radar [29]. A unique step in the development was the introduction of the minimization algorithm [2]. The minimization algorithm selected the minimum value over the absolute values of the SSP subbands. This was later called absolute minimization (AM) [30] to denote that the absolute value was taken before selecting the minimum value. The examples from real data showed impressive grain noise suppression results relative to simple compounding algorithms. The SSP-AM algorithm effectively performed a nonlinear statistical processing operation on the amplitude distributions, and

marked a significant departure from the weighted averages that had been applied previously.

The minimization operation was later generalized as belonging to a class of statistical operations referred to as order statistics (OS). An OS operation ranks input values according to magnitude and passes the value at a specified rank to the output [7,31]. This includes the minimum, median, and maximum values as special cases. The OS operation changes the distribution function of the input, such that it emphasizes a particular part of the amplitude distribution in a given quantile range [32,33]. So if the grain and flaw amplitude distributions had greater differences in the tails of their distributions (rather than their central location) then either the minimum or maximum values should be used to enhance their discrimination. The median value emphasizes the central location of the distribution (similar to the mean value) deemphasizing differences in the tails. Thus, cases where the SSP-AM algorithm performed better than the averaging algorithms implied that the amplitude distribution tails (skewness) had greater differences than their central locations. This was an important difference to exploit, especially for cases where the flaw and grain scatterer energy were at similar power levels. The SSP-OS operations effectively used amplitude distribution differences of the SSP ensemble, which distinguished it from linear filtering or averaging operations that simply affected the central part of the distribution.

Closely related to SSP-AM was the polarity thresholding (PT) algorithm [34,35]. In this case the A-scan value at a particular time sample was passed to the output when the RF SSP ensemble values were of the same sign. If this was not the case, the output value was set to 0. This differed from the SSP-AM algorithm in that the absolute value was not applied, so the polarity of the RF signal could be used. As in the case of SSP-AM, dramatic results were shown also for this algorithm. The good results, however, were highly dependent on selecting high signal-to-noise (SNR) regions of the spectral range [9,34]. Therefore, improvements in performance focused on finding the range of high SNR bands (usually the lower end of the transducer bandwidth) [34,36].

An SSP algorithm that used signal information differently from the AM and PT algorithm was the geometric mean (GM) algorithm [30]. The GM algorithm effectively multiplied values over the SSP ensemble at each time sample. Results from experimental samples presented in [30] showed similar results to those of PT and AM, and also included A-scans containing multiple targets. The SSP-GM represented a fundamentally new operation relative to earlier applications. The multiplication of RF ensembles corresponded to convolutions in the frequency domain. So from this viewpoint, the GM was responding to matches of frequency

domain patterns (convolutions or correlations exhibit maximum values when patterns match or have the same shape). These notions lead to the development of spectral redundancy techniques, which utilized this principle more efficiently. Spectral redundancy methods used the generalized spectrum (i.e. spectral autocorrelation) to classifying more complex scattering structures, such as complex flaw scatterers in NDE and tissue structures in medical imaging [37-39].

Many related works examined SNR enhancement as a performance criterion and searched for the right combinations of SSP parameters, such as bandwidth, center frequency spacing, and subband weight factors [3,7,9,15,36,40-42]. The most critical parameter identified in these works was frequency range over which the SSP ensemble spanned. Analyses indicated that the lower ranges of the transducer bandwidth exhibited less grain noise power due to the relative increase in scattering power at the higher frequencies for smaller scatterers [13]. If the Wiener spectrum (or power spectral density) of the target and grain were known, they could be used to either select the high SNR range or derive relative weights for each band. Ideally, this weighting should deconvolve the flaw spectrum [14] for enhanced SSP performance. For cases where such information could not be obtained, techniques were developed to automatically find the high SNR range [15,41,42]. Other works looked at variations creating the SSP ensemble through more advanced time-frequency approaches [10-12,16]. A wavelet decomposition application with multi-resolution bands reported some success, but a loss in resolution affecting cases with closely spaced targets [43]. In another study, an adaptive frequency range was proposed based on the material depth of the return echo using grain scattering models [44]. In this case the models helped select the widest band possible at each depth to maximize potential resolution.

Later works considered probability of flaw detection as a more critical metric than SNR enhancement for evaluating performance [9,34,46]. While impressive results were presented using SNR enhancement, ambiguity still existed for discriminating between flaw and grain signal peaks even after the enhancement, especially in scans with no reference peaks. The identification of flaw and grain peaks was further complicated by material attenuation and various normalization schemes used to obtain SSP ensembles. Therefore, methods were considered for designing thresholds to discriminate between signals due to grain and flaws, especially for SSP outputs [46-49]. These works modeled the SSP outputs statistically such that a threshold could be adaptively set to maintain a fixed false-alarm probability. The algorithm's ability to enhance the probability of detection was a more meaningful performance metric than SNR, especially when nonlinear SSP algorithms were involved.

Several works identified the linearity of the phase spectrum as the key component for enabling the SSP to perform well. This led to advanced studies to better explain and enhance SSP performance [8,9,15,36], and works focused on applying SSP to multiple targets [43,45,46]. The dependence of the algorithm on linear phase limits the nature of the target scatterer; however, the coherent scatterer model applies to many structures of interest. Therefore SSP has been considered for other applications, such as detecting tissue boundaries in medical imaging [50,52,53], tissue parameter estimation [51], and defect detection and layer thickness estimation in concrete [54]. The works in [50,52] extended SSP to 2 dimensions where the signal is split over both frequency and orientation (angle in the 2-D B-scan images). Superior boundary detection results were reported in phantom studies comparing SSP to other popular image processing approaches [50].

The literature presents SSP as a method well-suited for detecting flaws embedded in large-grained material. The rest of this chapter presents results to both confirm and complement previous work. In particular, signal model analyses are presented to identify critical signal features and characterizations upon which SSP and classical methods rely for improving flaw detection probability. While the relationships of SSP algorithms to other optimal detectors and filters have been discussed in various papers, this chapter directly compares their operation and performance for simple single scatterer flaws and complex multi-scatterer flaws through Monte Carlo simulation and probability of detection estimates. The results demonstrate the unique way in which SSP uses phase information and adds to the understanding of how SSP performs for different flaw types and at different SNR levels.

3. SSP and Classical Signal Processing Approaches

This section describes the grain and flaw signal model and identifies key elements that suggest the usefulness of an SSP approach. Features related to classical approaches are identified also. The matched [8,9] and Wiener filters [14-16] represent distinct classical approaches, where both use spectral magnitude characterizations of the grain and flaw echoes, the matched filter also uses deterministic phase information for the flaw. The SSP approach falls between these in that it indirectly uses a statistical characterization of the phase.

Denote the pulse-echo signal received from a sample containing grain and flaw structures with the convolution integral given by:

$$r(t) = \int_{-\infty}^{\infty} \sum_{m=1}^{M} d_m(\lambda - \tau_m) h(t - \lambda; \tau_m) d\lambda + \int_{-\infty}^{\infty} \sum_{n=1}^{N} g_n(\lambda - \tau_n) h(t - \lambda; \tau_n) d\lambda, \quad (1)$$

where $d_m(t)$ is the response of the m^{th} flaw scatterer, τ_m is the time delay associated with flaw positions, and $h(t;\tau_m)$ represents the insonifying energy at the flaw positions, as well as the propagation effects for the returning signal. Similarly, $g_n(t)$, τ_n, and $h(t;\tau_n)$ are the associated parameters for grain scatterers. For the analysis in this chapter, the received A-scan will be denoted as a function of time. If the material speed of sound is known, the time axis can be directly converted to depth in spatial coordinates.

The transducer response (center frequency and bandwidth) largely determines the convolution kernel $h(t;\tau_m)$. Propagation effects (losses due to diffraction, absorption, and scattering) result in a time/depth varying kernel [44]. To practically relate the classical signal processing approach to the signal model, a stationarity assumption is made over the A-scan interval being processed. Therefore, signals must be processed over sufficiently small segments such that the kernel remains approximately the same.

The keys to identifying flaw echoes in the presence of grain noise lies in scatterer strength differences in addition to differences in frequency magnitude and phase response. In many cases the flaw echo is stronger than individual grain echoes because of the material discontinuities it introduces. However, the large number of grain echoes resulting from the insonified resolution cell can generate sufficient energy to mask a flaw echo. In these cases it is useful to extract additional information from the phase and magnitude spectra characteristics to enhance the detection of the flaw echoes.

To better examine the scattered energy differences, Eq. (1) can be expressed in the frequency domain. Consider the case of a single flaw scatterer in Eq. (1), which can be expressed in the frequency domain as:

$$\hat{R}(f) = \hat{H}(f) \left(\hat{D}_1(f) \exp(-j2\pi f \tau_1) + \sum_{n=1}^{N} \hat{G}_n(f) \exp(-j2\pi f \tau_n) \right), \quad (2)$$

where $\hat{H}(f)$ is the Fourier transform of the interrogating pulse, $\hat{D}_1(f)$ is the frequency response of the flaw scatterer at a location corresponding to delay τ_1, and $\hat{G}_n(f)$ is the grain scatterer frequency response at locations with delays τ_n.

The flaw typically has a flat magnitude spectrum (if the scatterer is larger than the wavelengths of insonifying pulse) and the phase spectrum will be dominated by the linear term associated with delay τ_1. The linear phase causes energy over the whole band to coherently sum up at location τ_1 and create a peak [8,15,36,39]. This is in contrast to the grain scatterers, which due to their small size and distributed locations, have a frequency magnitude shift in emphasis toward high frequencies [13,44,47] and random (uniformly distributed) phase [39].

The most direct flaw detection approach involves computing the envelope of a signal and comparing the peak values to a threshold. This relies entirely on the relative flaw scatterer and grain scatterer strengths and does not take into account spectral and spatial distribution differences. This is the simplest model for discriminating between flaw and grain scatterers. It typically exhibits a robust, but suboptimal performance. The simulations presented later in this chapter use the envelope detector for comparing performance. The envelope detector is implemented with the Hilbert transform for the simulation results and is given by:

$$y(t) = \left| r(t) + j\hat{r}(t) \right|, \tag{3}$$

where $\hat{r}(t)$ denotes the Hilbert transform of $r(t)$.

The Weiner and matched filter approaches for detecting the flaw echoes embedded in grain noise exploit spectral magnitude and phase differences. The Wiener filter approach exploits only the frequency magnitude differences between the flaw and grain echoes. The filter can be expressed in the frequency domain as [14-16]:

$$\hat{S}_W(f) = \frac{\mathrm{E}\left[\left| \hat{H}(f)\hat{D}(f) \right|^2 \right]}{\mathrm{E}\left[\left| \hat{H}(f)\hat{G}(f) \right|^2 \right] + \mathrm{E}\left[\left| \hat{H}(f)\hat{D}(f) \right|^2 \right] + \varepsilon}, \tag{4}$$

where $\mathrm{E}[\cdot]$ is the expected value operator, and ε is a regularization factor to ensure stability when the spectral values of the estimated grain and flaw spectra are both close to zero (i.e. falls below the noise floor). In general, ε should be set just greater than the noise floor. A larger ε value results in greater stability (robustness) at the expense of losing filter sensitivity for separating gain and flaw signals.

The Weiner filter approach assumes the grain and target signals can be modeled as wide-sense stationary processes. In this case phase is not used. The Weiner filter is considered an optimal linear filter for a random signal in random noise. Note that for frequencies where the grain noise spectrum goes to 0, the filter goes to 1, and when grain noise spectral values become very large with respect to the flaw spectrum the filter approaches 0. This filter effectively emphasizes the high SNR regions and deemphasizes low SNR regions.

The matched filter approach assumes the flaw echo is deterministic, where both the magnitude and phase spectrum is known. If the grain noise is considered stationary, the match filter can be written in the form [8]:

$$\hat{S}_M(f) = \frac{\hat{H}^*(f)\hat{D}_1^*(f)}{\mathrm{E}\left[\left|\hat{G}(f)\hat{H}(f)\right|^2\right] + \varepsilon}, \tag{5}$$

where ε serves the same purpose as in the Wiener filter, and the linear phase factor from Eq. (2) can be dropped since it only designates the flaw echo position. The filtering can be implemented so that every A-scan sample is treated as a point of interest (τ_1 becomes 0 at some point through a convolution of the filter kernel). The remaining phase component is due to the insonifying pulse and the scatterer response, and is approximately linear for most common pulses, with some nonlinearities near the bandlimits of the pulse that result from nonsymmetries in the pulse envelope. The denominator of Eq. (5) effectively whitens (deconvolves) the colored grain noise and modifies the flaw echo template to match the whitened flaw echo. The output from the filter based on Eq. (5) is locally peak detected and each peak used as a detection statistic.

SSP has an advantage over the Wiener filter in that it incorporates phase information and does not need explicit knowledge of the grain and flaw spectra. It differs from the matched filter in that it uses statistical characteristics of the phase rather than an explicit phase spectrum. In practical applications, however, the SSP algorithms are applied over a portion of the transducer's frequency range, so this bandwidth has to be known. In summary, the key advantages of the SSP algorithms relative to the classical approaches include:

- Explicit knowledge of scatterer spectra is not required.
- Phase information can be used based on general characteristics.

The first bullet implies that the SSP algorithm is more practical to design, since less information needs to be known about the material scatterers. It also

implies that performance should be relatively robust to spectral shape changes. The second bullet implies that while explicit phase spectra models are not used, useful statistics about the phase behavior can still be used to enhance the flaw detection performance.

4. SSP Algorithms

SSP techniques operate on an ensemble obtained by filtering received broadband echoes into signals with different center frequencies. Consider the discrete received signal spectrum from Eq. (2) and let *k* denote the discrete frequency sample index and *n* denote the time sample index. The SSP ensemble signals are given by:

$$x_i(n) = \sum_{k=1}^{N} a_i W(k; b_i, f_i) \hat{R}(k) \exp\left(j2\pi \frac{kn}{N} \right) \quad \text{for } 0 \le i < M \ , \tag{6}$$

where $W(k; b_i, f_i)$ is a normalized band-pass filter window with 3 dB bandwidth b_i, center frequency f_i, and band scaling factor a_i. M is the total number of filters in the SSP filter bank and N is the number of DFT points. An example SSP ensemble relative to a transducer spectrum is shown in Fig. 2. In general, the band-pass filter bank covers the transducer spectrum with some adjustments to compensate for frequency dependent attenuation and scattering [44, 47].

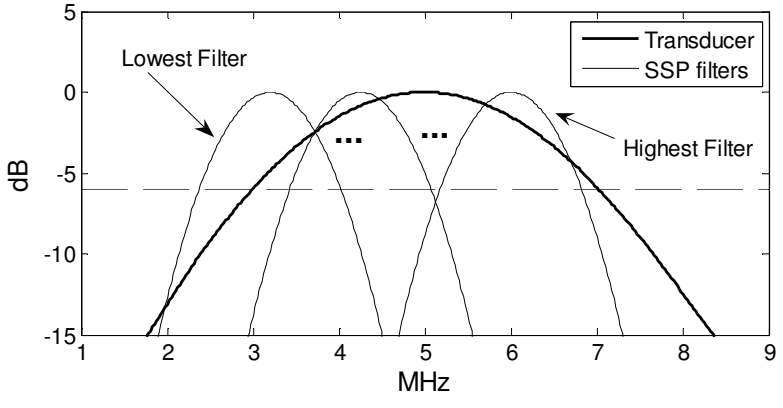


Figure 2. SSP ensemble subband filter magnitudes plotted with transducer magnitude response. Dashed line at 6 dB down point.

Statistical processing for SSP is performed in the time domain at each time instant over the ensemble. To explicitly denote this set of points define the following vector:

$$\mathbf{x}_n = \left[x_0(n),\ x_1(n),\ \cdots\ x_{M-1}(n)\right]^T. \tag{7}$$

For various SSP algorithms, the final output can be expressed as a decision statistic, which is function of $\mathbf{x}_n$.

Frequency Compounding

The simplest SSP operation involves taking the absolute value of each subband output and performing a weighed average, referred to as absolute averaging, and is denoted by:

$$y(n) = \sum_{i=0}^{M-1} \left| x_i(n) \right|, \tag{8}$$

where the a_i weights of Eq. (6) are used to give spectral emphasis for enhancing high SNR spectral regions (if such information is known).

Order Statistics and the Minimization Algorithm

The absolute-minimization (AM) algorithm simply passes the minimum absolute value of $\mathbf{x}_n$ to the output. If some bands could potentially exhibit a low SNR, the performance and robustness could be improved through the OS operation denoted as:

$$y(n) = \left| x \right|_{r:M} = \mathrm{OS}_{r:M}\left(\left| \mathbf{x}_n \right| \right), \tag{9}$$

where $|x|_{r:M}$ is the absolute value of the elements of $\mathbf{x}_n$ such that it has rank r among the M absolute values of its other elements. So the minimization algorithm is included in this notation with $r = 1$ (1 denotes smallest and M denotes largest rank). For many of the simple flaws tested, SSP-AM typically worked the best; however, for some complex flaws other values of r showed good performance [7].

Polarity Thresholding

The PT algorithm deemphasizes amplitude variations in the subbands and just considers the consistency of polarity. The PT algorithm acts as a gate to zero out

samples in the original A-scan that did not exhibit a consistent polarity in its subband decomposition. This operation is expressed as:

$$y(n) = \begin{cases} r(n) & \text{for } x_i(n) > 0 \text{ or } x_i(n) < 0, \text{ for all } i \\ 0 & \text{Otherwise} \end{cases} . \tag{10}$$

Both the PT and AM algorithms suppress grain echoes based on their frequent zero crossings over short spectral ranges in the SSP ensemble.

Geometric Mean Filtering

In order to address the robustness problems of the AM and PT algorithms, the Geometric Mean (GM) algorithm was introduced. This algorithm simply multiplies all the subbands together in the time domain. This operation is expressed as:

$$y(n) = \left| \prod_{i=0}^{M-1} x_i(n) \right|^{\frac{1}{M}} . \tag{11}$$

Note that for this operation, the subband weighting through the a_i parameters has no effect on the relative band contributions to the final statistic, since the individual weights factor out of the product operation and result in a single scale factor that cannot discriminate between bands. As mentioned earlier this operation results in a convolution of the subbands in the frequency domain.

5. SSP Parameters and Performance

The performance of the SSP algorithms depend on the subband ensemble parameters that include bandwidths, center frequencies, subband weights, and the frequency windowing function $W(k; b_i, f_i)$. From a statistical point of view, the more independent channels containing information on flaw scatterer the better for *averaging* out the random grain noise fluctuations. However, the bandwidth of usable information is limited by the pulse echo system and the attenuation properties of the material.

Selecting the set of bandwidths and center frequencies to cover the transducer bandwidth is a good start in most applications, since the spectrum outside this range is dominated by noise and electronic artifacts. If information is known about the material properties such as the frequency dependent attenuations due to scattering, absorption, and diffraction, adjustments can be made to the spectral

range to compensate for the effective SNR differences over the transducer spectrum [15,40,42,44,47]. Typically the frequencies associated with the smaller wavelengths result in more scattering from the grain structures, so the low SNR regions typically occur at the upper end of the transducer bandwidth. In addition, there is often a downward shift in the spectrum due to greater attenuation for the higher frequencies (relative to the lower frequencies). Therefore, the frequency range is typically set a little lower than the transducer spectral range to compensate for the frequency dependent attenuation of the insonifying pulse. If such adjustment is not done, losses in efficiency can be expected since high SNR regions may be lost due to the downward shifting of the spectral emphasis and low SNR regions included at the upper end of the bandwidth.

The relative weights for each subband can be adjusted to give more weight to the high SNR regions. This is particularly helpful for the frequency compounding algorithms that rely on a linear combination of spectral energies for the final detection statistics. For algorithms, such as the SSP-GM, these weights have no impact on the detection performance. In the case of SSP-AM and PT, the weighting is particularly helpful if it effectively deconvolves (equalizes) the flaw spectrum thereby minimizing the distribution of the peak at the target location.

To a limited degree the filter shape controls independence between the subband channels, while the majority of this is controlled by the spectral overlap due to the bandwidth and center frequency spacing. The filter shape also controls the relationship between the spectral bandwidth and the axial resolution in the time domain. Ideally it is best to have both good time and frequency resolution; however, these are inversely related to each other. So by making the SSP filter narrower, a smearing of the detail (lowering resolution) occurs in the time domain. The window used in many SSP works has been the Gaussian (or Gabor) window. One reason for this is that according to the uncertainty principle [55], the Gaussian window has the minimum time-frequency bandwidth product, and results in a good time-frequency resolution compromise. Therefore, it has been used in most SSP applications.

If the 3dB bandwidth is considered the effective bandwidth of the Gaussian window spectrum, then the relationship to the effective time resolution (3dB envelop in time) is given by [8]:

$$\Delta_t \approx \frac{2\sqrt{\ln(\sqrt{2})\ln(4)}}{\pi b_{3\text{dB}}}. \tag{12}$$

As an example, consider the case of Fig. 1 where the original pulse width was approximately 0.2 μs. The bandwidth of the subbands was .6 MHz, which

resulted in a spreading of the pulse widths associated with each subband to 1.7 μs. If the flaw signal is isolated, this effective lowering of the axial resolution includes more unresolvable backscattered energy from grain scatterers in neighboring spatial cells. If this is done to excess, the weaker flaw envelope peaks can be lost. For many broadband ultrasound applications where the primary interest is in isolated flaws, extending the spatial resolution by a factor of 3 to 6 typically preserves the spatial resolution of critical features. This implies a subband filter bandwidth selection of about one third to one sixth the transducer bandwidth. Center frequencies should be chosen so the SSP ensemble efficiently covers the bandwidth of interest.

6. Simulation Results

This section presents Monte Carlo simulations to demonstrate SSP algorithm performance relative to the matched and Wiener filter approaches. Since flaw and grain spectra are known explicitly from the simulation, the Wiener and matched filters are designed very close to optimal form. While this knowledge is difficult to achieve in practice, the filters are tuned to their optimal performance for the sake of comparisons with the SSP approach. The primary purpose of the simulation is to show that the use of statistical phase properties (as done in SSP) is better than not using phase at all, as in the case of the Wiener filter, and at the same time is not as good as using complete deterministic information about the phase, as in the case of the matched filter (at least for simple flaw scatterers).

For the following simulations, ensembles of grain-only and flaw-only A-scans were generated consisting of 1000 A-scans each. Each grain-only scan consisted of about 40 resolution cells. The signals were created through a discrete implementation of Eq. (1), where flaw and grain terms were computed separately, scaled, and added together to achieve a specific SNR. The impulse response for the flaw scatterer, $d(t)$, was taken to be an impulse, so the flaw echo had the same characteristics as the insonifying pulse $h(t;\tau_m)$. The insonifying pulse was simulated with a Rayleigh envelope modulated by a 5 MHz sinusoid (about 0.2 μs axial resolution). Simulations that use symmetric envelope pulses, like the Gaussian, do not include phase nonlinearities typically found in a real (causal) pulses. An example of the insonifying pulse used is shown in Fig. 3a.

The grain response is assumed to have high-pass characteristics to model Rayleigh region scattering [13]. Figure 3b shows the resulting spectra for simulating the flaw and grain scatterers magnitude responses ($\left|\hat{D}(f)\right|$ and $\left|\hat{G}(f)\right|$) filtered by the transducer response. Without a magnitude response

difference, the Wiener filter would show no significant performance improvements over simple envelope detection.

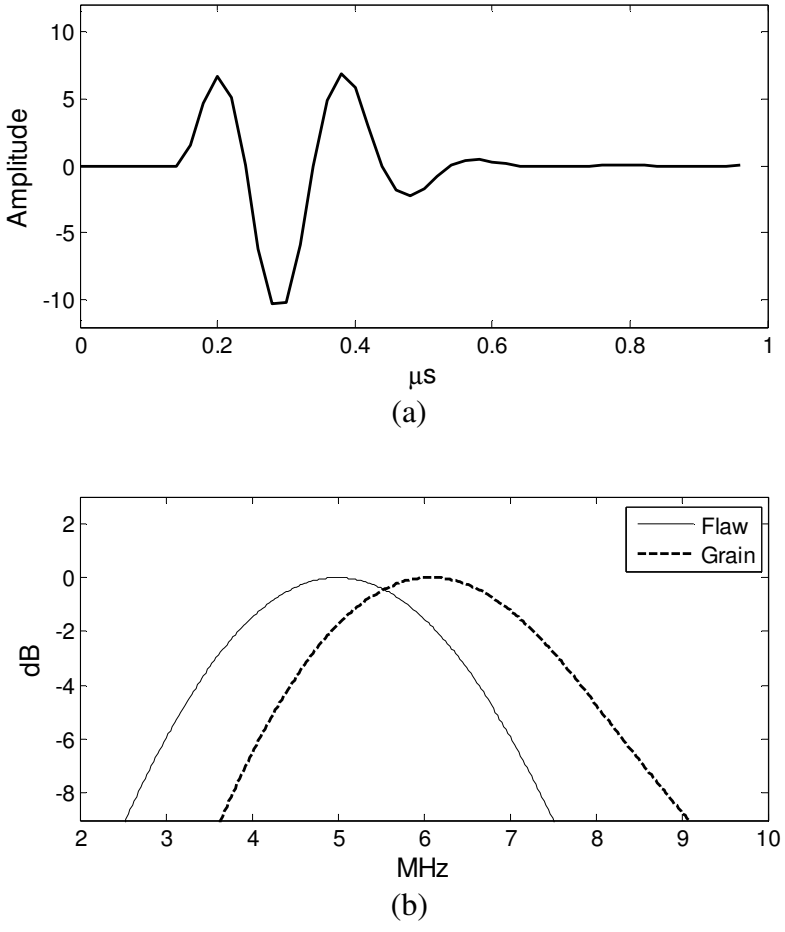


Figure 3. (a) Pulse-echo system impulse response. (b) Magnitude response of flaw and grain scatterers.

For each simulated grain-only A-scan the positions of the grain scatterers were uniformly distributed (spatially) over the duration of the A-scan with an average of 16 scatterers per resolution cell. The relative scatterer strengths were also uniformly distributed between 0.1 and 1.0 (i.e. all on the same order of magnitude). The scatterer responses were then superimposed at randomly generated positions (convolved) to obtain the final A-scan.

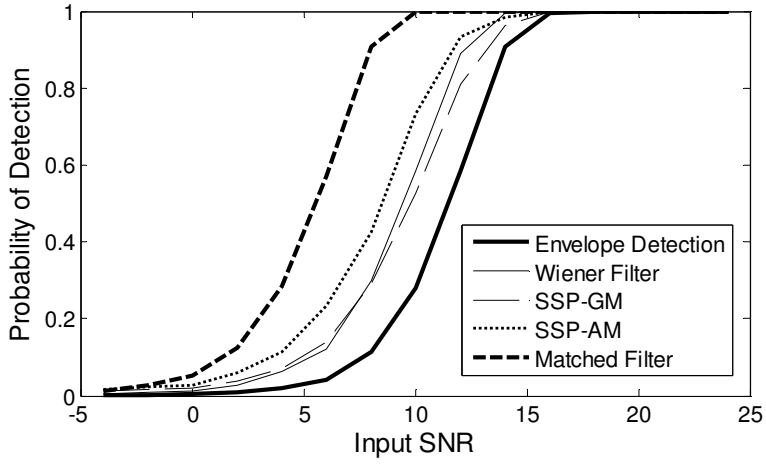
The grain-plus-flaw A-scans were generated with 2 flaw types. The first was a simple single scatterer with response $d(t)$, whose position was uniformly distributed within one resolution cell about a position near the center of the A-scan. The second case was a complex flaw scatterer consisting of 3 scatterers uniformly distributed within one resolution cell of each other. The complex flaw scatterer is used to demonstrate the effects of phase corruption (deviation from linear phase), a property exploited in the matched filter and SSP. Finally, the flaw-only signal r_f, was corrupted by adding it to the grain-only A-scan, r_g, according to:

$$r(n) = \beta_{\text{SNR}}\, r_f(n) + r_g(n), \tag{13}$$

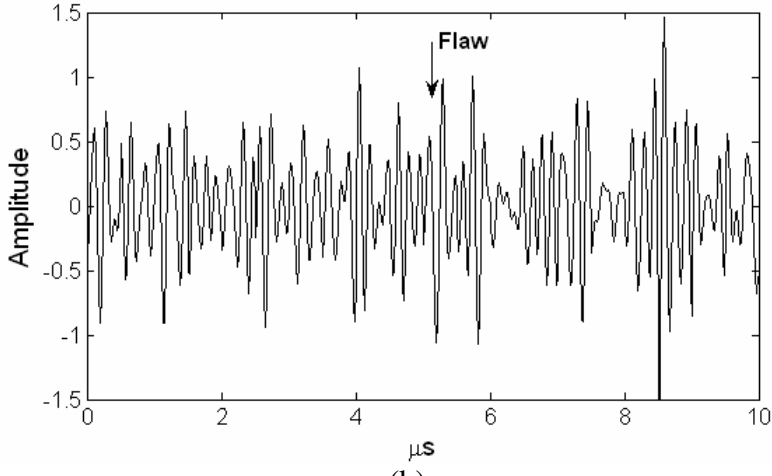
where β_{SNR} is the scale factor to achieve the desired SNR, which is defined for this study as the ratio between the peak flaw amplitude and the root mean square (rms) value of the grain noise.

The probability of detection at a fixed false positive probability was estimated from each set of 1000 Monte Carlo runs (without flaw). From the 1000 A-scans (about 40,000 resolution cell) the false-alarm threshold was estimated for $p_f = 10^{-3}$ with a 95% confidence range of $\pm 0.3\times10^{-3}$. For each set of combined flaw-plus-grain A-scans at a fixed SNR (1000 independent flaw signals) the probability of detection was estimated using a false-alarm probability threshold. The probability of detection was estimated over a range of SNR value and plotted as a function of SNR in Figs. 4 and 5. All probability of detection estimates had an upper bound of $\pm 3\times10^{-2}$ on the 95% confidence ranges.

Figure 4 shows relative performance for 5 different detection strategies. The SSP-AM and SSP-GM were implemented directly using Eqs. (9) and (11) with an SSP ensemble created from Eq. (6) using 5 Gaussian-shaped filters with constant bandwidths of b =0.7MHz, and center frequencies ranging from 3.1 MHz to 5.9 MHz (compare to flaw spectrum in Fig. 3b), in increments of 0.7MHz. All band scaling factors a_i were set to 1. The Wiener filter was implemented directly in the frequency domain on each A-scan with a small ε value approximately 100 times smaller than the SNR values. The only out-of-band noise in the simulation was from quantization error and windowing. The small ε values maintained high filter sensitivity so the comparisons could be made to these filters achieving their best possible performance. In practice, where the characteristic spectra are estimated, ε would need to be larger. The envelope of the Wiener filter output was computed and used as the detection statistic. The matched filter was implemented in the frequency domain according to Eq. (5).

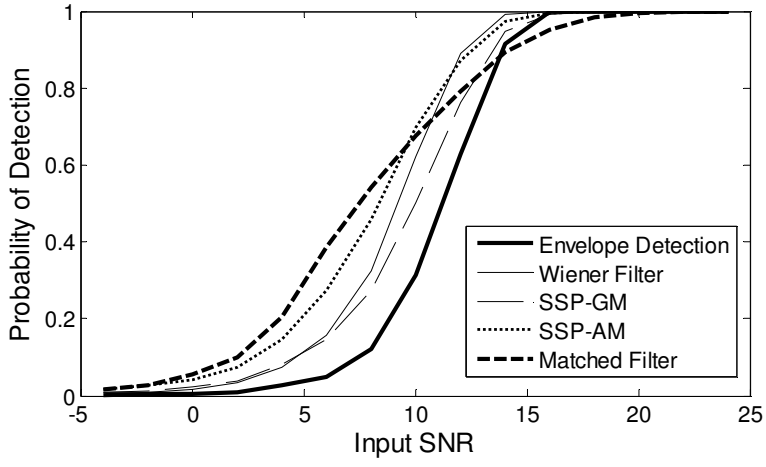


(a)

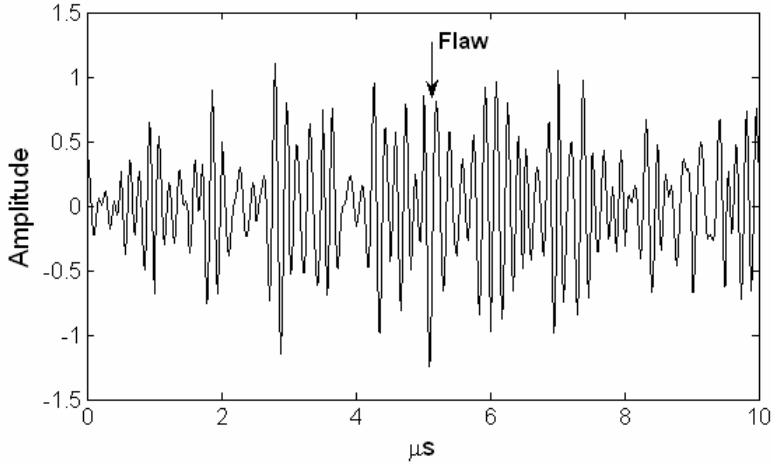


(b)

Figure 4. (a) Probability of detection estimates for a single-scatterer (simple) flaw using a false-alarm probability threshold of 10^{-3} for 5 different processing strategies. (b) Example A-scan containing a flaw at an 8 dB SNR.



(a)



(b)

Figure 5. (a) Probability of detection estimates for a 3-scatterer (complex) flaw at a false-alarm probability threshold of 10^{-3} for 5 different processing strategies. (b) Example A-scan containing a complex flaw at an 8 dB SNR.

For the sake of comparing the graphs in Figs. 4a and 5a and discussing the relative performance, a single dB point on the *x*-axis is used. For the following discussions the probability of detection at an SNR of 8 dB is used since it is in a critical region where all approaches have detection probabilities with significant differences between 0 and 1. An example of simulated A-scans with flaws at 8dB SNR is shown in Figs. 4b and 5b.

Figure 4a shows the envelope detector outperformed by all other processors because it relies only on echo strength and exploits no other differences. The Wiener filter, which exploits the spectral differences, performs better than the envelope detector with an 8dB detection probability almost 3 times greater than that of the envelope detector. The best performance is the Matched filter with an 8 dB detection probability of 0.91. The added performance is the result of accurately characterizing the flaw signal (including its phase spectrum). The SSP approaches exhibit suboptimal performance relative to the matched filter. The SSP-AM performs between the matched and Weiner filters, with an 8 dB detection probability of 0.43. Its relative performance is explained by the fact that SSP-AM uses phase information effectively, and therefore outperforms the Wiener filter. However, it does not use the complete deterministic model of the phase, and therefore does not perform as well as the matched filter. The SSP-GM performs on par with the Wiener filter with an 8 dB detection probability of 0.29. As the SNR increases the Wiener filter begins to outperform the SSP-GM detector. While these results suggest that the SSP-GM does not utilize phase information as efficiently as the SSP-AM or matched filter, the fact that it uses phase information to achieve a performance near the Wiener filter is significant. Note that the SSP techniques do not require specific knowledge of the spectral differences between the grain and flaw scatterers.

One problem with SSP-GM is that it effectively convolves the subband spectra through multiplication in the time domain. The convolution operation reverses the spectral axes in adjacent subbands and then correlates the 2 bands in a sliding window fashion. The correlation works well when the flaw spectra (both phase and magnitude) are relatively symmetric (so the axis reversal does not create a significant mismatching). A better approach is to directly correlate the subbands to match patterns in different bands due to the transient nature of the scatterer echo (linear phase pattern). This approach was applied to complex scatterer structures in both NDE [37,39] and medial imaging problems [38] with good success.

Figure 5 shows simulation results for the case where the flaw scatterer is similar to the grain, in that it consists of multiple unresolvable scattering centers. This case corresponds to the A-scans of Eq. (1) for M=3, where the flaw structure

starts to exhibit phase distortion over the SSP ensemble. And as a result the detectors of Fig. 5a that depended on phase spectra show degradations relative to the simple scatterer case for Fig. 4a. The greatest change is observed for the matched filter, with the 8dB detection probability dropping to 0.54, and even falling slightly below the envelope detector for SNR above 14 dB. This indicates that the deterministic phase model, which was accurate for the simple scatterer case, no longer accurately describes the flaw and grain differences. The phaseless characterizations, such as the envelope detector and Wiener filter maintain a consistent performance as expected (8dB detection probability for the Wiener filter is 0.32). A drop in performance is also observed for the SSP algorithms; however, the drop is not significant as in the case of the matched filter. The SSP-GM has an 8dB detection probability of 0.27, and the SSP-AM has an 8-dB detection probability of 0.46. The reason for the small SSP performance losses relative to the matched filter is that the statistical phase models for the SSP algorithms were not as restrictive as that for the matched filter. So the explicit scatterer model for the matched filter that resulted in superior performance for the simple scatterer case was also the reason for the significant degradation when the complex scatterer caused a deviation from the linear phase model.

7. Automatic Flaw Detection

The chapter closes with examples of implementing an SSP algorithm as an automatic flaw detector [47-49]. A critical issue for automatic flaw detection schemes is determining a threshold to yield expected probability of detection and false alarm. Since grain echoes are present throughout every scan, thresholds are usually set to maintain a fixed false detection probability, p_f, based on a statistical model. In the case of the simulations in the last section, ensembles of grain-only echoes were generated to empirically determine a threshold with a specific p_f. In actual material scans, it cannot be assumed that grain noise is stationary over the scanning range. Therefore, the threshold must vary over the entire scan according to parameters estimated from local data.

Threshold determination for a desired false-alarm probability requires a statistical model for the grain-only processed signals. The detection process can then be modeled as a null hypothesis test [48,49], where the null hypothesis is that an A-scan section under test resulted from grain-only scattering. The test section (usually the resolution cell) is processed and compared to a threshold associated with a fixed probability of false alarm (probability of rejecting the null hypothesis when it is true). If it exceeds the threshold, then the null hypothesis is rejected and a flaw is assumed present (detected). Consider a data point at sample

n_0. The probability of the null hypothesis being true based on a decision statistic is denoted as:

$$P_0(n_0) = 1 - P_{\mathbf{y}}\left(y(n_0) \;\middle|\; n_0 \in \text{ Grain-only cell }\right), \tag{14}$$

where $P_{\mathbf{y}}(\cdot)$ is the cumulative distribution function for the detection statistic, given it belongs to a grain-only resolution cell. A practical implementation of this test requires that $P_{\mathbf{y}}(\cdot)$ be estimated from the local data.

The occurrence of a flaw is typically a low probability event and occurs in isolation (surrounded by normal material grain). Therefore, the data surrounding the resolution cell under test most likely results from grain-only material, and it can be used to estimate probability distribution parameters to get an estimate of P_0. In cases where this cannot be assumed, data censoring can be used over the surrounding areas [47].

If possible, it is best to assume a distribution form that only requires the estimation of a few parameters to fit the distribution to the data (i.e. the mean and standard deviation in the case of a Gaussian). If the statistical form is accurate, then this is a more efficient use of the data rather than non-parametrically estimating the distribution (i.e. using a histogram).

A neighborhood of data points surrounding the cell of interest can be used to estimate parameters for the null hypothesis distribution. Denote these data in terms of the following vector:

$$\begin{aligned}\mathbf{y}_0 = [&y(n_0 - K\Delta), y(n_0 - (K-1)\Delta), \cdots \\ &y(n_0 - \Delta), y(n_0 + \Delta), y(n_0 + 2\Delta), \cdots y(n_0 + K\Delta))]\end{aligned}, \tag{15}$$

where Δ is the number of samples corresponding to the resolution cell length (Eq. (12)), and K is the number of resolution cells before and after the cell associated with n_0. Note that not every sample in this neighborhood is used, but only samples at integer intervals of the resolution cell, which results in independence between the $2K$ samples of $\mathbf{y_0}$. Sampling more densely may not improve the estimates because of the high correlation of samples within a resolution cell.

The data range in Eq. (15) includes samples from neighboring resolution cells on both sides of the test cell. In cases requiring strict real-time, then only the K cells previous to the test cell can be used. In A-scans with strong attenuation, this may cause a bias in the estimate and would need to be accounted for. The value of K depends on how much data is needed for an accurate estimation (the larger the K the more precise the estimate), and the interval over which data are effectively stationary (the larger the K, the more likely it includes nonstationary data).

A statistical model must be selected for use in Eq. (14). For the examples presented here, an exponential model will be used for the SSP-AM grain noise given by [47]:

$$P_y(y) = 1 - \exp(-y/\sigma), \tag{16}$$

where σ is both the mean and standard deviation of the distribution. Therefore for each A-scan point under test the data of $\mathbf{y_0}$ can be used to estimate σ. In the examples of this section, the mean of elements in $\mathbf{y}_0$ was used for the σ estimate. The exponential model is a convenient approximation for the SSP-AM grain noise samples. More accurate models can be derived based on processing the original grain distribution with order statistics [7,31,32], or determined empirically from many data samples [33]. For the sake of illustrating the automatic detection process in this chapter, however, the simple exponential mode is sufficient.

Consider the A-scan in Fig. 6 consisting of 3 simple flaw scatterers at the indicated positions. The simulated data is similar to the previous experiment except that attenuation as a function of depth was applied to model the progressive losses from scattering and diffraction. This further complicates the automatic detection process and creates more ambiguities for visual inspection, since energy in the insonifying pulse is not constant over the A-scan. However, the estimation of the mean of $\mathbf{y_0}$ effectively creates a scaling to remove the impact of the nonstationarity over the A-scan. Each flaw signal was added to the attenuated grain noise with different strengths relative to the RMS noise immediately surrounding their locations. For the examples presented here, the $\mathbf{y_0}$ neighborhood was determined by the resolution cell, which was approximately .2 μs, or Δ=10 samples at the 50MHz sampling rate, and a K value of 9.

Figure 6b shows the P_0 values of Eq. (14) for each sample point. Instead of applying the threshold to each cell, the reciprocal of P_0 was plotted on a log base 10 scale. In this way, the small probabilities corresponding to a greater likelihood of a flaw are shown as large values for easier identification in the plot. In addition, the y-axis of the graph directly corresponds to the p_f threshold, and therefore the effect of moving the threshold can be directly observed. A threshold corresponding to a false positive probability of 10^{-3} is marked on the plot with a broken line for reference. The A-scan contains about 100 resolution cells, so out of 10 A-scans, one crossing due to a grain only resolution cell is expected for this threshold. In this example none of the cells containing grain only exceeded this threshold. However, if this threshold is moved to a value of 2 (corresponding to $p_f = 10^{-2}$), then 2 grain samples exceed the threshold. The threshold associated with $p_f = 10^{-2}$ results 1 false alarm per 100 resolution cells on average, but this

also depends on the accuracy of the model in Eq. (16) and the quality of the σ estimate. The probabilities corresponding to flaw show significant deviation from the grain. The 15 dB target would still be detected even at a threshold corresponding to $p_f = 10^{-8}$, while the other 2 targets would be missed for thresholds exceeding $p_f = 10^{-4}$.

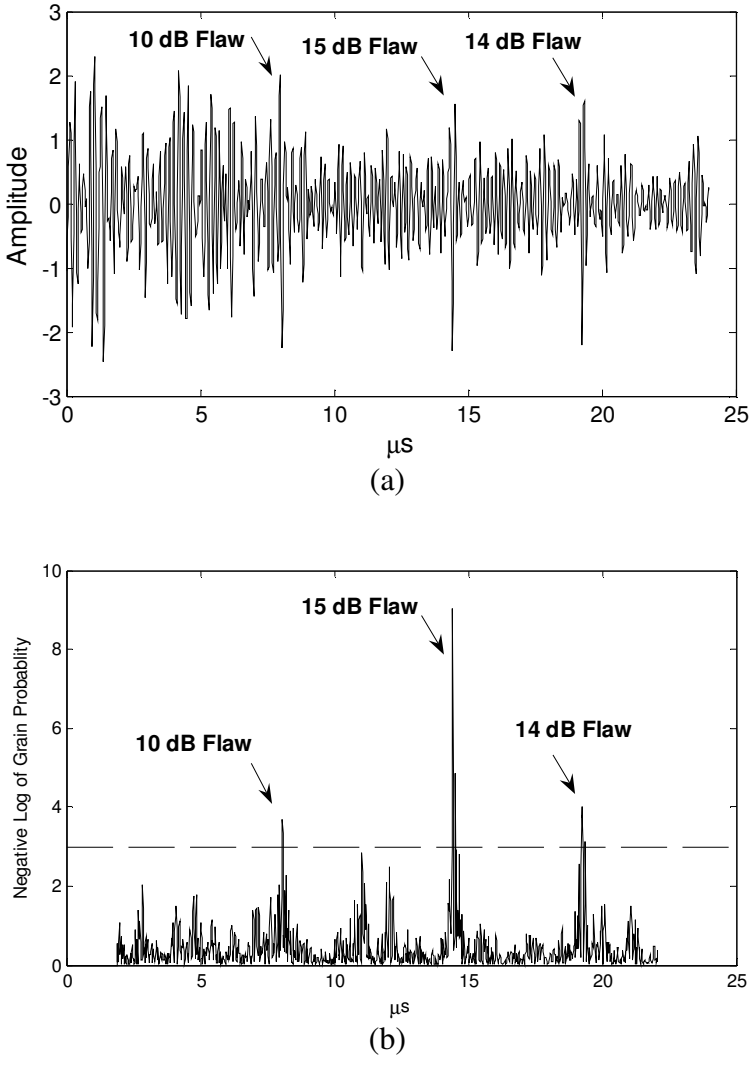


Figure 6. (a) Simulated A-scan with 3 simple flaw scatterers embedded in grain scatterers. (b) Negative log plot of grain-likelihood probability for SSP-AM processed samples using an adaptive exponential distribution model for processed grain samples.

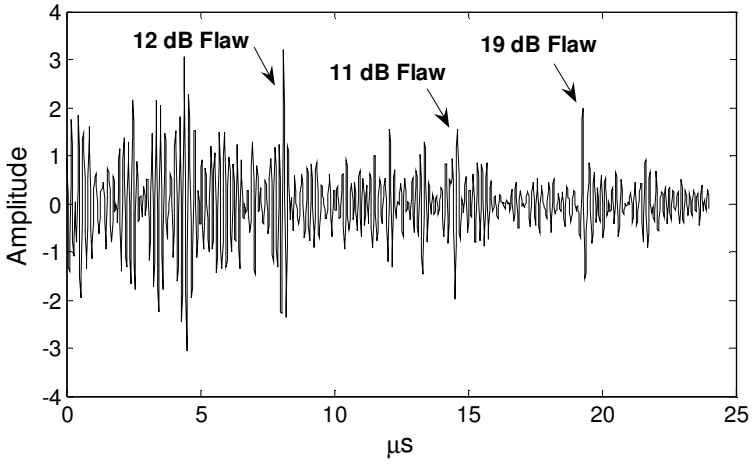
Figure 4a indicates that for flaw strengths ranging between 10 and 15 dB, a detection probability of .7 to 1 is expected. If this same flaw detection scheme was applied to the envelope detector output (assuming a reasonable statistical model was used), the detection probabilities would range from 0.25 to 0.95. Therefore, SSP increases the sensitivity to detecting weaker flaw echoes as a result of exploiting additional flaw and grain scatterer characteristics.

Similar results are shown in Fig. 7 for complex flaw scatterers (3 scattering centers on the same order of magnitude randomly distributed within the resolution cell). For flaw scattering strengths ranging between 11 and 19 dB, Fig. 5a suggests a detection rates between .75 and 1, which is a significant increase over envelope detection that has a detection range of .35 and 1. In summary, Figs. 6 and 7 demonstrate the potential advantages of SSP-AM with an adaptive distribution model for processed grain noise for reducing the ambiguities in discriminating between flaw and grain noise echoes, and automating the process for computer analysis of the A-scans.

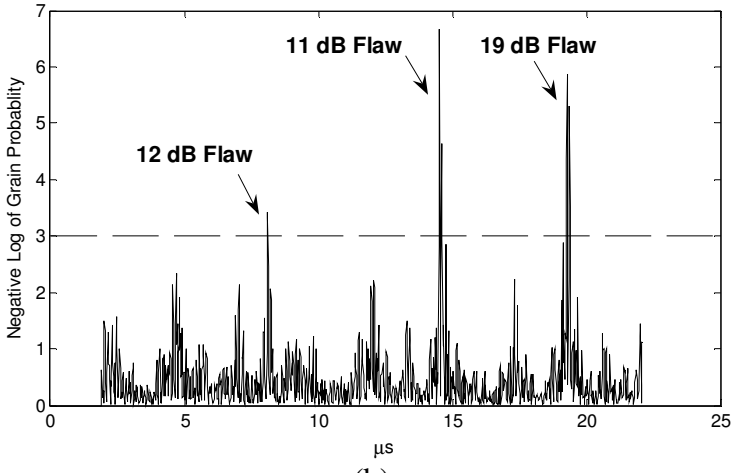
8. Summary

This chapter reviewed the development and application of SSP for the ultrasonic flaw detection problem. It was shown that the SSP using nonlinear processing schemes represents a unique approach relative to classical methods in that it employs a statistical characterization of the phase. Monte Carlo simulations demonstrated superior detection performance over approaches not using phase information, such as the envelope detector and the Wiener filter. While the Wiener filter requires explicit knowledge of the grain and flaw echo spectra, the SSP approach only requires knowledge of a useful frequency range, which is typically close to the spectrum of the transducer. Some adjustments to the frequency range are important to account for propagation effects. If explicit knowledge of the target and flaw spectrum is known, it can be incorporated in the SSP channel weighting to further enhance performance (a uniform weighing was used for the results in this chapter).

The matched filter showed superior performance to the SSP algorithms for the simple flaw case, because it incorporated explicit information about the target scatterer (phase and magnitude) along with the grain magnitude spectrum. In practice, accurate information of this nature is typically not available. When complex scatterers were used for flaws in the simulation, a deviation from the near-linear phase patterns resulted in performance losses for algorithms using phase information. In this case, however, the matched filter performance dropped below that of the SSP-AM algorithm. This demonstrated the relative robustness



(a)



(b)

Figure 7. (a) Simulated A-scan with 3 complex flaw scatterers embedded in grain scatterers. (b) Negative log plot of grain-likelihood probability for SSP-AM processed samples using an adaptive exponential distribution model for processed grain samples.

of the SSP algorithms and the advantages using statistical characterizations of the phase information.

Finally, examples were presented demonstrating an application of SSP in an automatic flaw detection scheme for multiple flaws embedded in nonstationary grain noise. The nonlinearities of the SSP algorithm effectively change the flaw and grain echo distribution to enhance the separation of their amplitudes beyond that of simple envelope detection. The application of a parametric statistical model for the grain echoes allowed thresholds to be adaptively set at a fixed false alarm probability. Thus, taking advantage of the impact that the nonlinear SSP had on the output samples, and reducing ambiguity in the interpretation of the peaks.

References

[1] S. Kraus and K. Goebbles, "Improvement of signal-to-noise ratio for the ultrasonic testing of coarse grained materials by signal averaging techniques", *Ultrasonic Materials Characterization*, edited by Berger and Linzer, Special Publication 596, pp. 551-559, National Bureau of Standards Gaithersburg MD, 1980.

[2] V. L Newhouse, N. M. Bilgutay, J. Saniie, and E. S. Furgason, "Flaw-tograin echo enhancement by split-spectrum processing," *Ultrason.,* pp. 59-68, Mar. 1982.

[3] N. M. Bilgutay and J. Saniie, "The effect of grain size on flaw visibility enhancement using split-spectrum processing", *Materials Evaluation*, Vol. 42, pp. 808-814, May 1984.

[4] R. B. Thomson and D. O. Thompson, "Ultrasonics in nondestructive evaluation", *Proc. of the IEEE*, Vol. 73, No. 12, Dec. 1985.

[5] B. Balingand, M. Grozellier and D. Romy, "Improvement in ultrasonic examination of austenitic steels", *Materials Evaluation*, Vol. 44, pp. 577-581, Apr. 1986.

[6] G. E. Kechter, J. D. Achenbach, Combined linear and homomorphic deconvolutions for processing bandpass measurements, *IEEE Trans. Signal Processing*, Vol. 39, No. 6, pp. 1300-1304, June 1991.

[7] J. Saniie, D. T. Nagle, and K. D. Donohue, "Analysis of OS Filters Applied to Ultrasonic Flaw Detection Using Split-Spectrum Processing," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.,* Vol. 38, No. 2, pp. 133-140, Mar. 1991.

[8] K. D. Donohue, "Maximum-Likelihood Estimation for A-Scan Imaging of Coherent Targets in Media of Unresolvable Scatterers," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.,* Vol. 39, No. 3, pp. 422-431, May 1992.

[9] M. G. Gustafsson, "Nonlinear clutter suppression using split spectrum processing and optimal detection," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol. 43 No. 1, pp. 109-124, Jan. 1996.

[10] A. Abbate, J. Koay, J. Frankel, S. Schroeder, P. Das, Signal detection and noise suppression using a wavelet transform signal processor: application to ultrasonic flaw detection, *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol. 44, No. 1, pp. 14-26, Jan. 1997.

[11] M. Rashmi, N. M. Bilgutay K. Kaya, O. Kagan, "Detection of ultrasonic anomaly signals using wavelet decomposition," *Materials evaluation*, pp. 1274-1279, Nov. 1997.

[12] M. Robini, I. Magnin, H. Benoit-Cattin, A. Baskurt, "Two-dimensional ultrasonic flaw detection based on the wavelet packet transform," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol. 44, No. 6, pp. 1382-1394, June 1997.

[13] E. P. Papadakis, "Ultrasonic attenuation caused by scattering in poly-crystalline metals," *J. Acoust. Soc. Amer.,* Vol. 37, pp. 703-710, 1965.

[14] S. P. Neal, P. L. Speckman, M. A. Enright, "Flaw signature estimation in ultrasonic nondestructive evaluation using the Wiener filter with limited prior information," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol. 40, No. 4, pp. 347–353, July 1993.

[15] X. Li and N. M. Bilgutay, "Wiener filter realization for target detection using group delay statistics," *IEEE Trans. Signal Processing,* Vol. 41, No. 6, pp. 2067-2074, 1993.

[16] M. A. G. Izquierdo, M. G. Hernández, O. Graullera and L. G. Ullate, "Time–frequency Wiener filtering for structural noise reduction," *Ultrasonics,* Vol. 40, No. 1-8, pp. 259-261, May 2002.

[17] C. B. Burckhardt, "Speckle in ultrasound B scans," *IEEE Trans. Sonics Ultrason.,* Vol. SU-25, pp. 1-6, Jan. 1978.

[18] J. G. Ahbott and F. L. Thurstone, "Acoustic speckle: Theory and experimental analysis," *Ultrason. Imaging*, Vol. 1, pp. 303-324, 1979.

[19] P.N . T. Wells and M. Halliwell, "Speckle in ultrasonic imaging," *Ultrason.,* Vol. 19, pp. 225-229, 1981.

[20] P. M. Shankar, "Speckle Reduction in Ultrasound B-Scans Using Weighted Averaging in Spatial Compounding," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol. 33, No. 6, pp. 754-758, Nov 1986.

[21] U. Bencharit, J. L. Kaufman, N. M. Bilgutay and J. Saniie, "Frequency and spatial compounding techniques for improved ultrasonic imaging", *Proc. 1984 IEEE Ultrasonic Symposium*, pp. 1021-1026, 1984.

[22] I. Amir, N. M. Bilgutay, V. L. Newhouse, "Analysis and comparison of some frequency compounding algorithms for reduction of ultrasonic clutter", *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol. 33, No. 44, pp. 402-411, Jul. 1986.

[23] B. Gold, N. Morgan, *Speech and Audio Signal Processing: Processing and Perception of Speech and Music*, John Wiley & Sons, New York, NY, pp. 28-38, 2000.

[24] B. C. J. Moore, *An Introduction to the Psychology of Hearing*, Academic Press Inc. San Diego, CA, 2003.

[25] B. Wandell, *Foundations of Vision*. Sinauer Associates, Inc. Sunderland, MA, 1995.

[26] D. R. Griffin, F. A. Webster, C. R. Michael, "The echolocation of flying insects by bats." *Animal Behav.,* Vol. 8, pp. 151-154, 1960.

[27] J. A. Simmons, "Perception of echo phase in bat sonar," *Science,* Vol. 204, pp. 1336-1338, 1979.

[28] N. M. Bilgutay, J. Saniie, E. S. Furgason, V. L. Newhouse, "Flaw-to-grain echo enhancement", *Proc. Ultrasonics International*, pp. 152-157, May 1979.

[29] E. W. Beasley, H. R. Ward, "Quantitative analysis of sea clutter decorrelation with frequency agility", *IEEE Trans. of Aerospace and Electronic Systems*, Vol. AES-4, No. 3, pp. 468-473, May 1968.

[30] J. Xin, K. D. Donohue, N. M. Bilgutay, X. Li*.,* "Frequency-diverse geometric- and arithmetic-mean filtering for ultrasonic flaw detection," *Materials Evaluation,* pp. 987-992, Aug. 1991.

[31] H.A. David, Order Statistics, John Wiley and Sons, New York, NY, 1981.

[32] J. Saniie, K. D. Donohue, N. M. Bilgutay, "Order Statistic Filters as Postdetection Processors," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. 39, No. 10, pp. 1722-1732, Oct. 1990.

[33] K. D. Donohue, N. M. Bilgutay, "OS Characterization for Local CFAR Detection," *IEEE Trans. on Systems Man and Cybernetics*, Vol. 21, No. 91, pp. 1212-1216, Sept. 1991.

[34] P. M. Shankar, P. Karpur, V. L. Newhouse, J. L. Rose, "Split-spectrum processing: analysis of polarity threshold algorithm for improvement of signal-to-noise ratio and detectability in ultrasonic signals," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol. 36, No. 1, pp. 101–108, Jan. 1989.

[35] N. M. Bilgutay, U. Bencharit, J. Saniie, "Enhanced ultrasonic imaging with split-spectrum processing and polarity thresholding," *IEEE Trans. Acoustics, Speech, and Signal Processing*, Vol. 37, No. 10, pp. 1590-1592, Oct. 1989.

[36] M. G. Gustafsson, T. Stepinski, "Split spectrum algorithms rely on instantaneous phase information-a geometrical approach," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol 40, No. 6, pp 659-665, Nov. 1993.

[37] K. D. Donohue, J. M. Bressler, T. Varghese, N. M. Bilgutay, "Spectral Correlation in Ultrasonic Pulse-Echo Signal Processing," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol. 40, No. 3, pp. 330-337, July 1993.

[38] K. D. Donohue, and L. Huang, T. Burks, F. Forsberg, and C. W. Piccoli, "Tissue Classification with Generalized Spectrum Parameters," *Ultrasound in Medicine and Biology*, Vol. 27, No. 11, pp. 1505-1514, 2001.

[39] K. D. Donohue, and L. Huang "Ultrasonic scatterer structure classification with the generalized spectrum," *Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, Vol. 6, pp. 3401 – 3404, 2001.

[40] P. Karpur, P. Shankar, J. L. Rose, V. L. Newhouse, "Split spectrum processing; Optimizing the processing parameters using minimization", *Ultrasonics*, Vol. 25, pp. 204-208, Jul. 1987.

[41] X. Li, R. Murthy, K. D. Donohue, N. M. Bilgutay, "Adaptive and robust filtering techniques for ultrasonic flaw detection," *Proc. IEEE 1989 Ultrason. Symp.,* pp. 1145-1149, 1989.

[42] X. Li, N. M. Bilgutay, R. Murthy, "Spectral histogram using the minimization algorithm-Theory and applications to flaw detection," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.,* Vol. 39, No. 2, pp. 279-284, Mar. 1992.

[43] K. Kaya, N. M. Bilgutay, R. Murthy, "Flaw detection in stainless steel samples using wavelet decomposition," *Proc. 1994 Ultrasonics Symposium,* Vol. 2, No. 1-4, pp. 1271-1274, Nov. 1994.

[44] C. Cudel, M. Greyillot, J.-J. Meyer, L. Simonim, "Detecting echoes of different spectral characteristics in absorbing media by variable moving bandwidth SSP minimization," *Proc.2002 Ultrasonics Symposium,* Vol. 1, pp. 785-788, Oct. 2002.

[45] T. Qi, L. Xing; N. M. Bilgutay, "Multiple target detection using split spectrum processing and group delay moving entropy," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol. 42, No. 6, pp. 1076-1086, Nov. 1995.

[46] T. Qi, N. M. Bilgutay, "Statistical analysis of split spectrum processing for multiple target detection," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol. 45, No 1, pp. 251-256 Jan. 1998

[47] J. Saniie, D. T. Nagle, "Analysis of order-statistic CFAR threshold estimators for improved ultrasonic flaw detection," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, Vol. 39, No. 5, pp. 618-630, Sept. 1992.

[48] K. D. Donohue, N. M. Bilgutay, "Split-spectrum processing with nonparametric methods for ultrasound imaging," in *Proc. 1990 Bilcon Int. Conf. New Trends Commun., Contr., Signal Processing,* Ankara, Turkey, pp. 1567-1574, 1990.

[49] N. M. Bilgutay, K. D. Donohue, and X. Li, "Nonparametric Flaw Detection in Large-Grained Materials," *Proc. 1990 Ultrasonics Symposium*, pp. 1137-1141, 1990.

[50] P. C. Pedersen, V. Mitra, J. Dey, "Boundary detection in 3D ultrasound reconstruction using nearest neighbor map," *Progress in Biomedical Optics and Imaging, Proc. SPIE,v 6147, Medical Imaging 2006: Ultrasonic Imaging and Signal Processing*, p. 614707, 2006

[51] P. J. Chen, T. Chen, "Measurements of acoustic dispersion on calcaneus using spilt spectrum processing technique," *Medical Engineering and Physics*, Vol. 28, No. 2, pp. 187-193, March, 2006.

[52] R. G. Dantas, S. Leeman, E. T. Costa, J. P. Jones, E. J. V. Oliveira, "Phase diversity for speckle reduction," *Proc. of the SPIE - The International Society for Optical Engineering*, Vol. 5035, pp. 414-422, 2003.

[53] K. H. Uslu, N. M. Bilgutay, R. Murthy, K. Kaya, "Medical image enhancement using Split Spectrum Processing," *Proc. of the IEEE Ultrasonics Symposium*, Vol. 2, pp. 993-997, 1993.

[54] S. Popovics, N. M. Bilgutay, M. Karaoguz, T. Akgul "High-frequency ultrasound technique for testing concrete," *ACI Materials Journal*, Vol. 97, No. 1, Jan/Feb, pp 58-65, 2000.

[55] S. Qian, and D. Chen, *Joint Time Frequency Analysis: Methods and Applications*, Prentice Hall, Upper Saddle River, NJ, 1996.

CHAPTER 23

AN APPLICATION OF THE EM ALGORITHM TO THICKNESS ESTIMATION OF THIN LAYERED MATERIALS

C. H. Chen and Xiaojun Wu

University of Massachusetts Dartmouth
Electrical and Computer Engineering Dept.
N. Dartmouth, MA 0274

For thin layered materials, such as fiber reinforced titanium, or some composite structure, the determination of thickness of individual layers can be difficult with ultrasonic testing method because the reflected signals from thin layers are highly overlapped. The popular methods using cross-correlation and homomorphic deconvolution are reviewed. It is suggested that the model based method can be more suitable for this problem. The EM (expectation-maximization) maximization) algorithm is proposed for estimating the model parameters A detailed implementation of the algorithm is presented and comparison with the homomorphic method clearly shows that the EM algorithm is more suitable for an accurate estimate of thickness of multi-layered materials.

1. Introduction

An important application of ultrasonic nondestructive testing (NDT) techniques is to acquire the quantitative material properties, such as the velocity of ultrasound in unknown material, the thickness of the material, and the attenuation of the material, etc. Most of the formulas describing the material properties in ultrasonic NDT technique include time of arrival (TOA) and time-difference of arrival (TDOA) parameters. TOA parameter can give us the information of the location of the target (flaw, layer … etc.). The thickness of the inspected specimen, the velocity of ultrasound in the unknown material, and some other information about the material require the estimation of TDOA between corresponding layer reflective echoes.

The cross-correlation method is widely used to obtain the high resolution TOA and TDOA [1]. This method uses a linear time invariant filter, which can be viewed as a "matched filter," to produce maximum value at the TOA of the received echo. However, to estimate the thickness of thin-layered materials is a challenging task in ultrasonic nondestructive testing. This is because when ultrasonic testing is applied to a thin-layered material, the echoes that come from the front surface and back wall of the layer will be overlapped in the time domain. We cannot use the traditional cross-correlation method in this case to find out the high resolution TOA and TDOA.

C. Martin, et al. [2] advanced the homomorphic deconvolution method to obtain the thickness of thin-layered material. This method assumes there to be only two reflections from the specimen. In other words, it only gives the time delay of echoes from two reflections. This method can only be used to calculate the TDOA of thin-layered material whose attenuation is large, such as thin-layered fiber. For the thin-layered material, which has multiple reflections, such as thin metal material, we need to use other methods to calculate the TDOA between different layers.

Recently, R. Demirli, et al. [3] presented the model-based method to obtain the high resolution TOA of echoes from thin-layered material that has multiple reflections. In this chapter, we introduce the Gaussian model based space alternating generalized EM (SAGE) algorithm to estimate the parameters in the Gaussian model and system model to obtain the high-resolution estimate of the system response in thin-layered material. By combining the system model based SAGE algorithm with the Gaussian model base SAGE algorithm, the TDOA of thin-layered material is estimated.

The organization of the chapter is as follows. Ultrasonic pulse-echo immersion testing scheme and the formula to compute the material thickness in NDT are introduced in section 2. In sections 3 and 4, we present the backscatter echo model and pulse-echo system model. The EM algorithms are introduced to estimate model parameters. In sections 6 and 7, the EM algorithms are used to estimate the thickness of thin-layered fiber/titanium and draw a conclusion.

2. Ultrasonic Pulse-echo Immersion Testing

To obtain a precise result of the ultrasonic material characterization, we need to study the waveforms in the A-scan signal and their corresponding echoes. Figure 1 is the scheme of ultrasonic pulse-echo immersion testing and the corresponding waveforms are shown in Figure 2 [4].

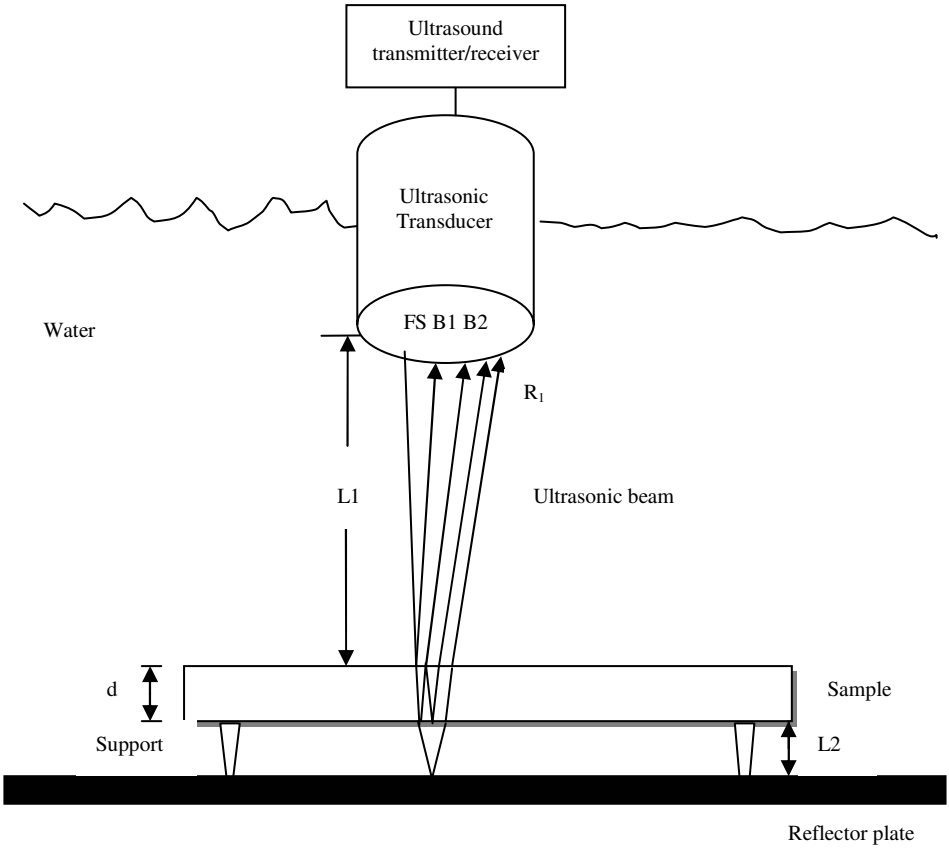


Figure 1. Ultrasonic pulse-echo immersion testing system.

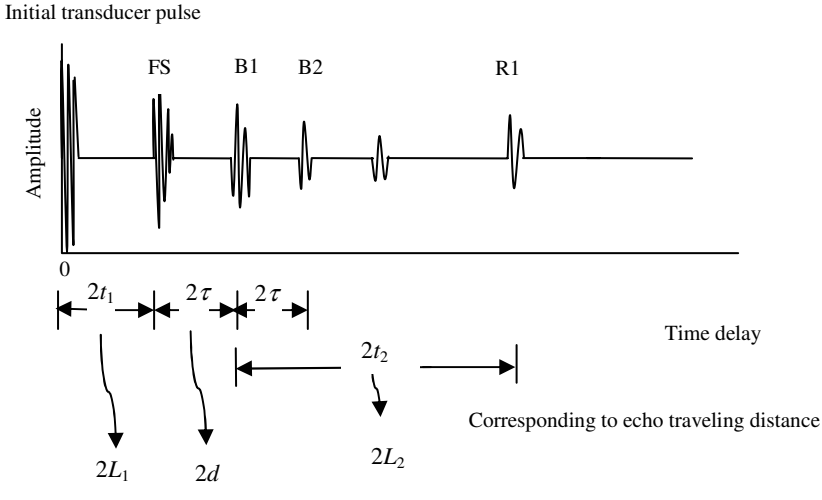


Figure 2. Corresponding waveforms of Figure 1.

Here, L_1 is the distance between ultrasonic transducer and the sample; L_2 is the height of the support; d is the thickness of the sample; FS represents the ultrasonic front surface echo of the sample; B_1 represents ultrasonic first back wall echo of the sample; B_2 is the ultrasonic second back wall echo of the sample; R_1 is ultrasonic front surface echo of the reflector plate with the sample present; $2t_1$ represents the pulse-echo time delay between the front surface of the sample and the transducer; 2τ represents the pulse-echo time delay between a front surface and back wall echoes or two successive back wall echoes of the sample; $2t_2$ represents the pulse-echo time delay between the back wall of the sample and the front surface of the reflector plate.

The relationships among the above parameters are:

$$2t_1 = \frac{2L_1}{V_{water}} \tag{1}$$

$$2\tau = \frac{2d}{V_{sample}} \tag{2}$$

$$2t_2 = \frac{2L_2}{V_{water}} \tag{3}$$

We will notice that, 2τ is pulse-echo TDOA between a front surface and back wall echoes or two successive back wall echoes of the specimen, that is

$$2\tau = TOA(B_1) - TOA(FS) \tag{4}$$

or

$$2\tau = TOA(B_2) - TOA(B_1) \tag{5}$$

The thickness of the material can be determined by using equation (6),

$$d = \frac{2\tau \times V}{2} \tag{6}$$

Here, d is the thickness of the specimen. 2τ is the pulse-echo TDOA between a front surface and back wall echoes or two successive back wall echoes of the specimen, which is given in equation (4) and (5). V is the velocity of ultrasound in the material.

3. Ultrasonic Backscattered Echo Model

The ultrasonic backscattered echo from a flat surface reflector can be written as in equation (7) [3],

$$y(t) = s(\theta; t) + v(t) \tag{7}$$

Here, $s(\theta;t)$ is a Gaussian echo model, and $s(\theta;t) = \beta e^{-\alpha(t-\tau)^2}\cos(2\pi f_c(t-\tau)+\phi)$, $\theta = [\alpha \quad \tau \quad f_c \quad \phi \quad \beta]$. α is bandwidth factor that determines the bandwidth of the echo or the pulse duration in the time domain; τ is arrival time, indicates the location of the reflector; f_c is center frequency, governed by the transducer center frequency; ϕ is phase; β is amplitude; both ϕ and β are related to the impedance, size and orientation of the reflector. $v(t)$ is the additive white Gaussian noise process.

Equation (8) is an altered version to describe the multiple echoes from the reflector,

$$y(t) = \sum_{m=1}^{M} s(\theta_m;t) + v(t) \tag{8}$$

Here, M denotes the number of superimposed Gaussian echoes; echo vector θ_m indicates the shape and position of each echo. The summation term represents a signal model of multiple reflections while the number of reflections is unknown.

4. Ultrasonic Pulse-echo System model

The transducer pulse-echo wavelet can be represented by altering equation (7) with M superimposed Gaussian echo wavelets [5],

$$h(t) = \sum_{m=1}^{M} c_m e^{-\alpha_m(t-\lambda_m)^2}\cos(2\pi f_m(t-\lambda_m)+\phi_m) \tag{9}$$

$$\theta_m = [\alpha_m \quad \lambda_m \quad f_m \quad \phi_m \quad c_m]$$

Here, α_m is bandwidth factor; λ_m is arrival time; f_m is center frequency; ϕ_m is phase; c_m is amplitude.

An ultrasonic echo reflected from a flat surface can be represented as the model:

$$s(t) = \beta h(t-\tau) \tag{10}$$

where, $h(t)$ is the transducer pulse-echo wavelet defined in equation (9); τ corresponds to the arrival time of the ultrasonic echo traveling through the different reflective layers, and it can indicate the position of the reflector; β is the amplitude of the reflective echo, and is affected by the impedance, size, inner structures, and orientation of the reflector.

The system model (10) can be generalized to the M-echoes with noise effect:

$$y(t) = h(t) * \{\sum_{n=1}^{N} \beta_n \delta(t-\tau_n)\} + v(t) \tag{11}$$

$$\psi_n = [\tau_n \quad \beta_n]$$

where, $\{\sum_{n=1}^{N} \beta_n \delta(t-\tau_n)\}$ is the impulse train, which denotes the unknown system response.

5. The EM Algorithm

The Expectation Maximization (EM) algorithm is an iterative optimization method to estimate a set of unknown parameters Θ of given measurement data [6, 7]. It is an algorithm for finding maximum likelihood estimates of parameters in probabilistic models, where the model depends on a set of unknown parameters. It includes two steps; one is expectation step (E-step), and the other is maximization step (M-step). In the E-step, we compute the expected value from the model using current parameters. In the M-step, we compute the maximum likelihood estimation of the parameters from the given data and the model, and set the resulting parameters as the current parameters for the E-step.

Now, we introduce the EM algorithms to estimate the parameters in ultrasonic backscatter echo model and pulse-echo system model as presented in sections 3 and 4.

For computational purpose, equation (8) can be rewritten as in equation (12):

$$y = \sum_{m=1}^{M} s(\theta_m) + v \tag{12}$$

We need to estimate parameters θ_1, θ_2,…, θ_M to construct the estimated M-superimposed echoes. We first define x_m as in equation (13):

$$x_m = s(\theta_m) + v_m \tag{13}$$

Here, $s(\theta_m)$ is the m th Gaussian echo; v_m is the WGN sequence.

Then equation (12) can be rewritten as

$$y = \sum_{m=1}^{M} x_m \tag{14}$$

where x_m and y are Gaussian random sequences.

In the E-step, we can use equation (15) to obtain the expected value $\hat{x}_m$ [3]:

$$\hat{x}_m^{(k)} = s(\theta_m^{(k)}) + \beta_m (y - \sum_{l=1}^{M} s(\theta_l^{(k)})) \text{, where } \sum_{m=1}^{M} \beta_m = 1 \tag{15}$$

In the M-step, we use equation (16) to obtain the next parameter:

$$\theta_m^{(k+1)} = \theta_m^{(k)} + (H^t(\theta_m^{(k)})H(\theta_m^{(k)}))^{-1} H^t(\theta_m^{(k)})(\hat{x}_m^{(k)} - s(\theta_m^{(k)})) \qquad (16)$$

where, parameter $\theta = [\alpha \ \ \tau \ \ f_c \ \ \phi \ \ \beta]$; $H(\theta)$ is the gradient of the model with respect to parameters in θ.

Then we can use the EM algorithm to estimate the M-superimposed Gaussian echoes with WGN. However, E-step and M-step are computed in parallel in EM algorithm, that is, we get all the required current parameters $\Theta^{(k)}$ and the estimated data $\hat{X}^{(k)}$ in the E-step; then we use them to calculate the $\Theta^{(k+1)}$ in the M-step. One alternative to this parallel method is to update the $\Theta^{(k+1)}$ in the M-step right after estimating $\hat{x}_m^{(k)}$ in the E-step, without waiting for the other parameter vector to be estimated. This method is known as the generalized EM. It can help the computation to be faster. In the case of WGN, the generalized EM is also known as space alternating generalized EM (SAGE) algorithm [8]. The use of SAGE algorithm to estimate the M-superimposed Gaussian echoes with WGN is shown below.

Algorithm 1 (Space Alternating Generalized EM to estimate Θ)

Step 1. Initialize $\Theta^{(0)} = [\theta_1^{(0)}; \theta_2^{(0)}; ...; \theta_M^{(0)}]$, Norm

Step 2. Set $k = -1$ and tolerance.

Step 3. Check the norm, while the norm greater than the tolerance, go to step 4, otherwise stop.

Step 4. Set $k = k+1$, $m = 1$.

Step 5 Check m, while $m \le M$, go to Step 6, otherwise go to Step 9.

Step 6. E-step: compute the expected echoes

$$\hat{x}_m^{(k)} = s(\theta_m^{(k)}) + \frac{1}{M}(y - \sum_{l=1}^{M} s(\theta_l^{(k)}))$$

Step 7. M-step: using the corresponding parameters to calculate equation (16):

$$\theta_m^{(k+1)} = \theta_m^{(k)} + (H^t(\theta_m^{(k)})H(\theta_m^{(k)}))^{-1} H^t(\theta_m^{(k)})(\hat{x}_m^{(k)} - s(\theta_m^{(k)}))$$

Step 8. Set $m = m+1$, and then go to Step 5.

Step 9. Compute $\| \Theta^{(k+1)} - \Theta^{(k)} \|$ to obtain norm, set $\Theta^{(k)} = \Theta^{(k+1)}$, then go to Step 3.

The flow chart for SAGE algorithm is shown in Figure 3.

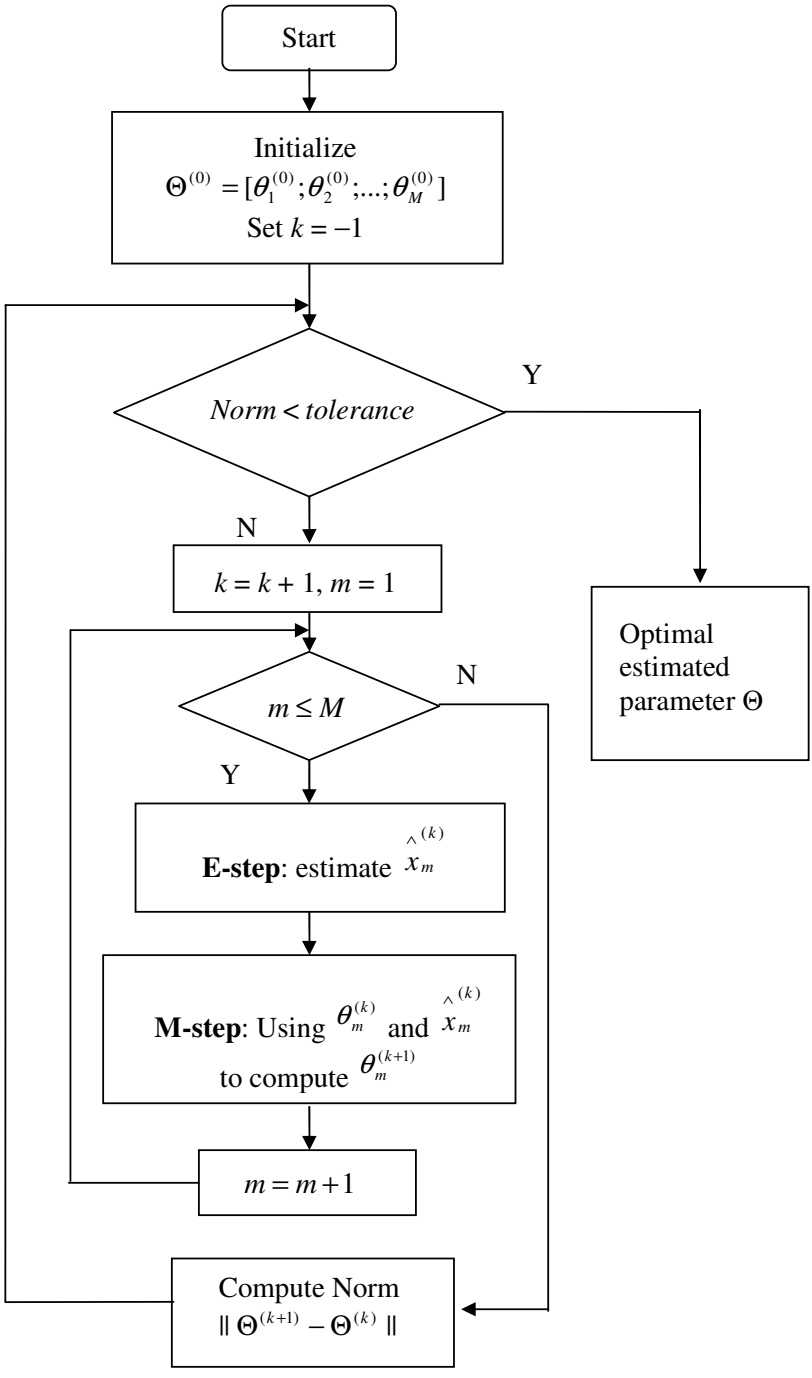


Figure 3. Flow chart for SAGE algorithm.

We can also use the SAGE algorithm to estimate the parameter $\psi_n = [\tau_n \quad \beta_n]$ from give data y in ultrasonic pulse-echo system model shown in equation (11).

Algorithm 2 (Space Alternating Generalized EM to estimate Ψ)

Step 1. Initialize $\Psi^{(0)} = [\psi_1^{(0)}; \psi_2^{(0)}; ...; \psi_N^{(0)}]$, Norm
Step 2. Set $k = -1$ and tolerance.
Step 3. Check the norm, while the norm greater than the tolerance, go to step 4, otherwise stop.
Step 4. Set $k = k+1$, $n = 1$.
Step 5. Check n, while $n \le N$, go to Step 6, otherwise go to Step 9.
Step 6. E-step: compute the expected echoes

$$\hat{x}_n^{(k)} = s(\psi_n^{(k)}) + \frac{1}{N}(y - \sum_{l=1}^{N} s(\psi_l^{(k)})),$$

where $\psi_n = [\tau_n \quad \beta_n]$, $s(\psi_n) = \beta_n h(t - \tau_n)$, and $h(\cdot)$ is given by equation (9)
Step 7. M-step: using the corresponding parameters to calculate equation (17):

$$\psi_n^{(k+1)} = \psi_n^{(k)} + (H^t(\psi_n^{(k)}) H(\psi_n^{(k)}))^{-1} H^t(\psi_n^{(k)})(\hat{x}_m^{(k)} - s(\psi_n^{(k)})) \tag{17}$$

$H(\psi)$ is the gradient of the model with respect to the parameters in $\psi = [\tau \quad \beta]$. The process to get $H(\psi)$, and $H^t(\psi)H(\psi)$ is given in [9].
Step 8. Set $n = n+1$, then go to Step 5.
Step 9. Compute $\| \Psi^{(k+1)} - \Psi^{(k)} \|$ to get norm, set $\Psi^{(k)} = \Psi^{(k+1)}$, then go to Step 3.

6. The Application of EM Algorithm to Estimate the Thickness of Thin Layered Materials

We can use the ultrasonic pulse-echo system model to obtain an accurate TOA of different layer reflections of the thin layer. After computing the TOA, we can calculate the TDOA between the front surface and the back wall of the thin layer. For a given pulse-echo signal generated from the thin layer, first we estimate M sets of the five parameters $\theta_m = [\alpha_m \quad \lambda_m \quad f_m \quad \phi_m \quad c_m]$ using Algorithm 1, and substitute them into equation (9) to obtain $h(t)$; then assume $h(t)$ is known and unchanged to estimate N sets of the system parameters $\psi_n = [\tau_n \quad \beta_n]$ using Algorithm 2.

The above two algorithms are used to obtain the thickness of thin-layered titanium. Figure 4 (a) and (b) depict the A-scan signal of the thin titanium (using a 5-MHz transducer) and its magnitude spectrum:

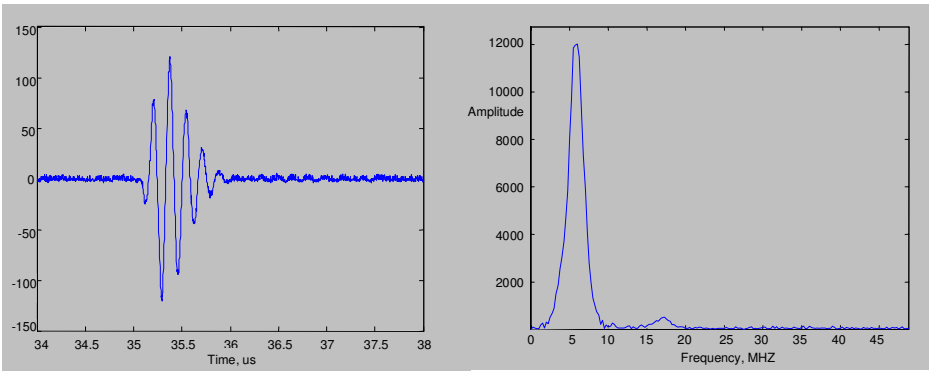


Figure 4. (a) A-scan signal of the thin-layered titanium and (b) its magnitude spectrum

We use the model order M=2 in Algorithm 1 to estimate the parameters. Figure 5 (a) is the estimated echo for Figure 4 (a). And Figure 5 (b) is the magnitude spectrum of Figure 5 (a).

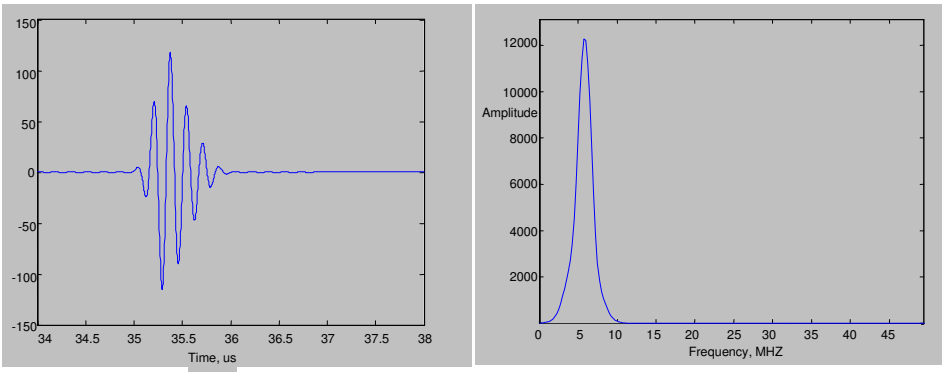


Figure 5. (a) Estimated echo of the thin titanium and (b) its magnitude spectrum.

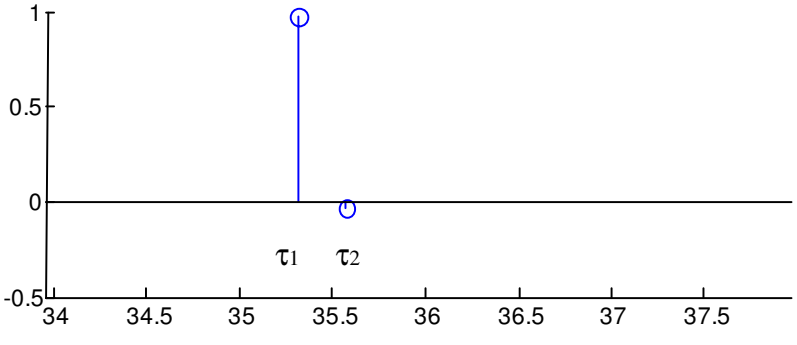


Figure 6 . The system response of the thin-layered titanium.

Then we use Algorithm 2 to estimate the time delay between the front surface and back wall of the thin titanium. Figure 6 shows the system response of the thin layer titanium using Algorithm 2.

The difference between the two estimated parameters arrival time is

$$\text{TDOA} = \text{TOA (Back wall)} - \text{TOA (Surface)} = \tau_2 - \tau_1 = 0.265\mu s$$

and the thickness of the thin-layered titanium calculated by the estimated TDOA fits the actual thickness well. If the homomorphic deconvolution method is used the time difference estimated [9] is 0.23μs, which is slightly off the correct value.

It is useful at this point to make some comments on the EM algorithms.

1) To estimate the thickness of each layer in multi-layered materials, it is necessary to have reflections from all layers, though they may be highly overlapped. In the fiber reinforced titanium, with fiber layer on each side of titanium, the reflection from the last layer may be very weak making detection and estimation difficult. The recent experiments conducted by Manandhar [10] and Chiang [11] showed that the EM algorithm may not converge, or may converge very slowly, in such situation.
2) There can be various alternative assignments of the reflections in the derived TOA data, which may lead to different thickness estimation. Some good heuristic may be helpful to resolve the ambiguity.
3) The EM algorithm can be computationally fast if the process converges. In fact it can be faster than the homoorphic deconvolution for a single layer case.

   The algorithm has also been shown to be much less sensitive to additive noise.

7. Conclusion

In this chapter, we introduced a model-based method to estimate thickness of a thin layered material, which may exhibit many reflections in ultrasonic testing. This model-based method includes two parts, the Gaussian echo model and the system model. By using the Estimation-Maximization (EM) algorithm, we estimate the parameters of the Gaussian echo model, and then use the estimated results to estimate the parameters in the system model. Then we obtain the TOA and TDOA of a fiber reinforced titanium, based on which the layer thickness can be estimated with good accuracy.

Acknowledgement

This work was supported by Information Research Foundation, and it made use of the facility at the Ultrasonic NDE Lab. of Advanced Technology and Manufacturing Center at the University of Massachusetts Dartmouth.

References

[1] D. Pagodinas, K. Barsauskas, "Ultrasonic signal processing methods for detection of defects in composite materials", *Ultragarsas Journal*, Vol. 45, No. 4, 2002.

[2] C. Martin, J.J. Meister, M. Arditi, and P.A. Farine, "A novel homomorphic processing of ultrasonic echoes for layer thickness measurement", *IEEE Trans. Signal Processing*, Vol. 40, No. 7, pp. 1819-1825, July 1992.

[3] R. Demirli and J. Saniie, "Model-based estimation of ultrasonic echoes Part I: Analysis and Algorithms", *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.,* Vol. 48, No. 3, pp. 787-802, May 2001.

[4] D. J. Roth and D. V. Carney, Thickness-independent ultrasonic imaging applied to ceramic materials, NDTnet, Vol. 2, No. 10, October 1997.

[5] R. Demirli and J. Saniie, "Model-based estimation of ultrasonic cchoes Part II: Nondestructive Evaluation Application", *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.,* Vol. 48, No. 3, pp. 803-811, May 2001.

[6] F. Dellaert, "The Expectation Maximization algorithm", College of Computing, Georgia Institute of Technology, Technical Report number GIT-GVU-02-20, Feb., 2002.

[7] T. K. Moon, "The Expectation-maximization algorithm", *IEEE Signal Processing Mag.,* pp. 47-60, Nov. 1996.

[8] J. A. Fesslor and A. O. Hero, "Space alternation generalized expectation maximization algorithm", *IEEE Trans. Signal Processing*, Vol. 42, No. 10, pp. 2664-2677, Oct. 1994.

[9] Xiaojun Wu, "Signal analysis and modeling for ultrasonic characterization of materials", *Master Thesis of ECE Dept.*, University of Massachusetts Dartmouth, Sep 2005.

[10] P. Manadhar, "Investigation of the possibility of using the SAGE algorithm to deconvolve ultrasound A-scan data of fiber reinforced titanium for determining the thickness of various layers in multi-layered materials" internal report prepared for UMD Ultrasonic NDE Lab., Sept. 2006.

[11] M.C. Chiang, "Ultrasonic signal and image processing for materials characterization and MEMS defect detection", MS Project Report, ECE Dept., University of Massachusetts Dartmouth, Dec. 2006.

CHAPTER 24

HOMOMORPHIC DECONVOLUTION OF ULTRASONIC IMAGES

Radovan Jiřík

De . f Bi edica E gi ee i g
B U i e i f Tech g
K ej 2906/4
61200 B , C ech Re b ic
ji i @feec. b .c

Torfinn Taxt

De . f Bi edici e
U i e i f Be ge
J a Lie ei 91
5009 Be ge , N a
T .Ta @f . ib.

The chapter deals with restoration of medical ultrasonic images by means of deconvolution. The so far published relevant approaches are reviewed and shortly described. The main focus is given to homomorphic deconvolution. Basic concepts and practical realization of homomorphic filtering are given for both one-dimensional and two-dimensional signal processing. Finally, application of homomorphic filtering to medical ultrasound signals is explained and examples of deconvolved phantom and clinical images are given.

1. Introduction

This chapter is focused on deconvolution of medical ultrasonic images. In medical imaging, the imaged objects, i.e. body tissues, are very complex compared to materials imaged in nondestructive material testing. The tissues consist of structures with a wide range of size. On one hand, the cellular structures and cells are smaller than or comparable to the wavelength of the used ultrasound pulse (in medical ultrasound the pulse frequency is 1 – 20 MHz, depending on application, corresponding to the wavelength 75 – 1500 μm). On the other hand, the size of organs is much larger than the pulse wave length. Thus, the received radiofrequency signal is mostly formed by scattered and reflected wave. Scattering corresponds to structures comparable to the wave length and reflection is due to large well defined interfaces between the tissue layers and organs. The dense scatterers / reflectors in

the tissue are the main difference from the objects imaged in nondestructive material testing which consist mainly of sparse scatterers / reflectors. That is why the methods described in this chapter may perform in a different way when applied to nondestructive material testing. Sparse deconvolution techniques[1–3] may be more suitable for this application field.

Medical diagnostics is, to a large extent, dependent on medical imaging modalities which enable early disease diagnostics and thus increase the probability of successful treatment. Ultrasonic imaging is a widely used imaging modality. In contrast to other medical imaging modalities (e.g. magnetic resonance imaging and X-ray computed tomography), ultrasound imaging scanners are easily movable, have a fairly low cost and standard real-time imaging capabilities. On the other hand, ultrasound imaging scanners have a low image quality caused by fairly low spatial resolution and the presence of a special type of noise, called speckle. The low spatial resolution is due to the fairly large extent of the interrogating ultrasound pulse in both the radial and lateral directions. The speckle is caused by the coherent character of the imaging process, and is unavoidable.

Although the ultrasound imaging quality has been improved substantially within the last decades, there is a large potential for further improvements by new hardware features and by postprocessing of the radiofrequency signals. The need for such improvements comes from the clear limitations ultrasound imaging shows in clinical practice, mainly caused by too low spatial resolution. Deconvolution of ultrasonic image data is one way to improve the spatial resolution.

The chapter is organized as follows. First, a linear convolutional model of the processed signal is described (Sec. 2). Then, the one-dimensional (1-D) and two-dimensional (2-D) approaches to the deconvolution of medical ultrasonic images are reviewed (Sec. 3). The readers' attention is then drawn to homomorphic deconvolution, as the authors believe that it is one of the most reliable approaches to the deconvolution of medical ultrasonic images. The basic concepts of the homomorphic transform, its properties and computational variations are given in Sec. 4.1 for 1-D processing and in Sec. 4.2 for 2-D processing. In Sec. 4.3, the theory of homomorphic transform is related to the properties of ultrasonic image data. Arguments are given for suitability of homomorphic filtering for deconvolution of this type of signals. Finally, to illustrate the potential of homomorphic deconvolution, several deconvolution results are shown in Sec. 5 for phantom and clinical images.

2. Model of Ultrasonic Image Data

To understand better the ultrasonic imaging process, the description by means of mathematical modeling is necessary. A detailed description of the physical phenomena taking place in ultrasonic imaging can be found in.[4–9]

The published models are derived for radiofrequency signal, i.e. the signal measured directly at ultrasound probe. As will be shown later in this section, a con-

volutional model can be derived for radiofrequency signals. Due to high technical requirements on the digital data acquisition units, several early papers on deconvolution of ultrasound images[10] reported on processing of video data, i.e. images derived from the radiofrequency signals by envelope detection, decimation and compression of the signal dynamics (e.g. by logarithmic transform). As these include nonlinear operations, the convolutional model no more holds for video data. Thus, it is necessary to deconvolve the ultrasonic images as raw radiofrequency signals. This has mostly been the approach of the recent and current research.

A commonly used model of the received radiofrequency signal is[6–8,11–14]

$$y(\vec{r}_2, t) = \nu_{pea}(t, |\vec{r}_1 - \vec{r}_2|) *_t x(\vec{r}_1) *_r h_{pe}(\vec{r}_1, \vec{r}_2, t) + w(\vec{r}_2, t), \tag{1}$$

where

- t stands for time, which can also be viewed as a spatial coordinate in the axial direction (supposing constant ultrasound propagation velocity),
- $\vec{r}_1$ is the position of the reflectors/scatterers and $\vec{r}_2$ the position of the transducer,
- $\star_t$ denotes temporal and $\star_r$ spatial convolution,
- ν_{pea} is the pulse-echo wave that accounts for the transducer excitation, the electromechanical impulse response during emission and reception of the pulse, and for the attenuation phenomena.
- x represents the original underlying image. It accounts for inhomogeneities in the tissue due to density and propagation perturbations, which give rise to the reflected and scattered signal. This signal will be further referred to as the tissue reflectivity function,
- h_{pe} is the modified pulse-echo spatial impulse response that relates the transducer geometry to the spatial extent of the scattered field,
- $w(\vec{r}_2, t)$ represents the measurement noise and the phenomena of the image formation process not included in the model.

The model Eq. (1) can be simplified by combining the terms $\nu_{pea}(t, |\vec{r}_1 - \vec{r}_2|)$ and $h_{pe}(\vec{r}_1, \vec{r}_2, t)$ in one spatially and temporally varying ultrasonic system impulse response (point spread function - PSF) $h(\vec{r}_1, \vec{r}_2, t)$. Supposing constant ultrasound propagation velocity, the time coordinate can be transformed to the spatial coordinates. For a fixed transducer location at the origin of the coordinate system, the radiofrequency signal is modeled as

$$y(\vec{r}_1) = x(\vec{r}_1) *_r h(\vec{r}_1) + w(\vec{r}_1). \tag{2}$$

For computational purpose, discrete spatial coordinates are more convenient. Considering the model in 2-D, the distortion model of B-scan radiofrequency data is

$$y(m, n) = x(m, n) * h(m, n) + w(m, n). \tag{3}$$

Assuming separability of the distortion process, 1-D models in the axial and lateral direction can be used

$$\begin{aligned} y(m) &= x(m) * h(m) + w(m) \\ y(n) &= x(n) * h(n) + w(n). \end{aligned} \tag{4}$$

3. Deconvolution

In general, deconvolution aims at restoration of the original signal from its linearly distorted noisy realization. In classical deconvolution methods the PSF is supposed to be known. Its accurate knowledge is essential for a successful deconvolution. In ultrasonic imaging, the PSF changes with depth and position. The tissue layers that are interposed between the transducer and the target behave as a transmission channel introducing unknown distortions in the ultrasonic signal. This is a result of several physical phenomena taking place in ultrasonic imaging process: the tissue dependent attenuation, dispersion, diffraction and wave front aberration.[4,5] Thus, the PSF can not be well measured in advance, before the image acquisition.

Nevertheless, there have been published several approaches where the PSF was estimated before the image acquisition. A rough PSF estimate can be obtained by measuring the signal reflected from a line target oriented perpendicularly to the scanning plane.[10,15] This way, the spatial variance of the PSF and the influence of the tissue on the PSF are simply ignored. Another approach is to mathematically model the PSF, either using a space-invariant model,[16–18] or using a more complex model which includes the spatial variance of the PSF.[16,19] Again, the modeled PSFs do not include distortions of the pulse caused by the tissues.

More advanced techniques are based on estimation of the PSF directly from the measured radiofrequency signals,[11–14,20–30] based on some a priori assumptions about the PSF and tissue reflectivity function. Only these methods will be considered further.

Another aspect of the published deconvolution methods is the deconvolution algorithm itself. The most frequently used approach has been Wiener deconvolution. The general form of the 1-D Wiener deconvolution filter is

$$M(\omega) = \frac{1}{H(\omega)} \frac{|H(\omega)|^2}{|H(\omega)|^2 + \frac{S_{ww}(\omega)}{S_{xx}(\omega)}}, \tag{5}$$

where $H(\omega)$ is the Fourier transform of the estimated PSF, $S_{ww}(\omega)$ and $S_{xx}(\omega)$ are the power spectra of noise $w(n)$ and the original tissue signal $x(n)$, respectively. An example of the power-spectra estimation can be found in.[20] As it is practically a difficult task, an approximation is mostly used instead. The ratio of the power spectra is supposed constant through the whole frequency band. This leads to a simplified form of the Wiener filter:[15,23–25,31]

$$M(\omega) = \frac{1}{H(\omega)} \frac{|H(\omega)|^2}{|H(\omega)|^2 + \Gamma}, \tag{6}$$

where Γ is the squared noise-to-signal ratio. This simplified Wiener filter deconvolution gives a maximum posterior estimate of the original image if the degraded signal and noise are independent and Gaussian distributed.

A straightforward extension of Eq. (6) to 2-D has been applied in 2-D deconvolution of medical ultrasonic images.[22,26,30]

A modified version which improves the results of the Wiener filter by calling it in an iterative algorithm was used in.[29]

In,[12] Mendel's deconvolution algorithm[32,33] was employed. It enables to consider the spatial variance of the PSF. The Mendel's deconvolution procedure consists of a Kalman filter followed by an estimation step performed time-recursively.

3.1. *1-D deconvolution*

This section describes 1-D deconvolution approaches in the axial and lateral directions, used for medical ultrasonics. In axial deconvolution, the radiofrequency signal along each A-line is processed. In lateral deconvolution, the processed sequences are composed from samples of A-scans corresponding to the same depth.

In,[34–36] the axial PSF was modeled with an AutoRegressive (AR) model. The parameters of the model were estimated from the recorded radiofrequency signal. The model was then used to estimate the center frequency of the pulse from which the attenuation coefficient was derived.

In[11] a similar approach to modeling of the pulse was used as a preceding step of deconvolution. An AutoRegressive-Moving-Average (ARMA) model was used for the axial PSF. It was estimated using a prediction error method. A short description of the algorithm follows:

As the pulse is changing with depth, short segments of A-lines are used for estimation of the PSFs. The PSF is considered stationary within each segment. Using the ARMA model, the observed radiofrequency signal $y(n)$ can be written as

$$\begin{aligned} y(n) = c_1 e(n-1) + c_2 e(n-2) + ... + c_N e(n-N) - \\ -a_1 y(n-1) - a_2 y(n-2) - ... - a_M y(n-M) + e(n), \end{aligned} \tag{7}$$

where $e(n)$ is the tissue reflectivity function, a_i and c_i are the coefficients of the ARMA model to be determined and N, M are the orders of the nominator and denominator of the ARMA model.

The "quality" of the ARMA parameters is expressed as the prediction capability of the model assuming that all samples up to the time index $n-1$ are known. In Eq. (7), the output at time n is written as a function of the previous input and output plus the current input $e(n)$. Assuming zero mean value of $e(n)$, the best

prediction of the output $\hat{y}(n)$ is

$$\begin{gathered}\hat{y}(n) = c_1\hat{e}(n-1) + c_2\hat{e}(n-2) + ... + c_N\hat{e}(n-N) - \\ -a_1y(n-1) - a_2y(n-2) - ... - a_My(n-M).\end{gathered} \tag{8}$$

The prediction error is then

$$y(n) - \hat{y}(n) = \hat{e}(n). \tag{9}$$

The ARMA parameters are found by minimizing the prediction error over the whole segment of the A-line, formulated as

$$err(a_1, ..., a_M, c_1, ..., c_N) = \sum_{k=1}^{K} \hat{e}^2(k), \tag{10}$$

where K is the number of samples in the A-line segment. The optimization problem is solved in an iterative algorithm with initial values set using a simplified model.

The prediction error algorithm was further extended in[13] by including data from a number of adjacent A-lines, which has decreased the variance of the estimated parameters and reduced the risk of terminating the algorithm at a local minimum instead of the global one.

This PSF estimation method is computationally very demanding and is not realizable in real-time applications.

In[37] a nonparametric PSF estimation technique based on homomorphic filtering[38] was proposed. A detailed description of homomorphic filtering will be given in Sec. 4. The method is based on filtering of the radiofrequency signal in the cepstrum domain. In general, the radiofrequency signal $y(t)$ is mapped to the cepstrum domain using the following sequence of operations:

$$\begin{gathered}Y(\omega) = \mathcal{F}\{y(t)\} \\ \hat{Y}(\omega) = logY(\omega) \\ \hat{y}(t) = \mathcal{F}^{-1}\{\hat{Y}(\omega)\},\end{gathered} \tag{11}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the forward and inverse Fourier transform and log is the logarithm of a complex function. In the cepstrum domain, the PSF and the tissue reflectivity function reside in different bands (discussed later in Sec. 4.3). The ultrasound PSF is then extracted by simple lowpass filtering.

The logarithm of a complex function poses a difficult problem, due to the discontinuities in the phase (see section 4). There have been suggested several methods to cope with the problem.

In[37] only the amplitude of the signal is considered under the assumption of the minimum-phase property of the PSF (see section 4.1.3). In,[28] the same assumption was used and, in addition, orthogonal compactly supported wavelet bases were used instead of the Fourier transform bases.

In[14,20] higher order statistics[39] is used to estimate the PSF in the axial and lateral direction. The homomorphic transform is applied to the third order cumulant. The third order cumulant of a 1-D time-domain signal $y(n)$ is defined as the following 2-D function

$$c(\tau, \rho) = E\{y(n) \cdot y(n+\tau) \cdot y(n+\rho)\}, \tag{12}$$

where E denotes the mean. Given the bispectrum of $c(\tau, \rho)$, defined as

$$C(\omega_1, \omega_2) = DFT\{c(\tau, \rho)\}. \tag{13}$$

The homomorphic transform of the estimated third order cumulant is called bicepstrum:

$$b_y(m_1, m_2) = DFT\{logC(\omega_1, \omega_2)\}. \tag{14}$$

The model Eq. (4) can be written as

$$b_y(m_1, m_2) = b_x(m_1, m_2) + b_h(m_1, m_2), \tag{15}$$

where $b_y(m_1, m_2)$, $b_x(m_1, m_2)$ and $b_h(m_1, m_2)$ are the bicepstra of $y(n)$, $x(n)$ and $h(n)$, respectively.

Considering the original tissue reflectivity function $x(n)$ as a white stationary stochastic signal with non-Gaussian distribution, the axial PSF can be separated in the bicepstrum domain as[14,20,23]

$$h(n) = DFT^{-1}\{e^{DFT\{\tilde{b}_h(m)\}}\}, \tag{16}$$

where

$$\tilde{b}_h(m) = \begin{cases} b_y(m, 0) & m > 0 \\ 0 & m = 0 \\ b_y(-m, 0) & m < 0. \end{cases} \tag{17}$$

The same procedure is done to estimate the lateral PSF. The estimated PSFs are then used in 1-D Wiener deconvolution in both directions.

A comparison of homomorphic filtering methods applied in 1-D deconvolution of medical ultrasonic images in the axial direction can be found in.[23] The methods

being compared include the minimum-phase and bicepstrum methods mentioned above, and other cepstrum based methods using phase unwrapping, logarithmic derivative, polynomial rooting, spectral root cepstrum and generalized cepstrum (see Sec. 4.1.3 for description of the methods). The results of the radial deconvolution of in-vivo ultrasonic images show the superiority of the generalized cepstrum and phase unwrapping methods.

In[24,25] the generalized cepstrum method and the phase unwrapping method were further developed. An essential part of both methods is the frequency domain phase unwrapping, which is a very noise sensitive task. A robust Bayesian phase unwrapping method using Markov random field model was used here. The approach has improved the deconvolution in both the axial and lateral directions. The phase unwrapping method and the generalized cepstrum method were reported to have very similar performance, with the former being less demanding on the computational time and thus the method of preference.

3.2. *2-D deconvolution*

2-D deconvolution is a more general and more complex procedure. It does not impose the constraint of separability to the PSF. There have been considerably fewer publications devoted to the 2-D deconvolution of medical ultrasonic images.

Bayesian 2-D deconvolution was presented in.[17] The approach is based on a statistical model for diffuse scattering in 2-D ultrasonic images, incorporating both spatial smoothness constraints and physical model for diffuse scattering. The emphasis is put on Bayesian modeling of the ultrasound radiofrequency signal, while the PSF is simply modeled as a sine modulated circular Gaussian function, constant through the whole image. In[18] the effect of this simplification is studied.

In[22,26] the 2-D homomorphic filtering was applied to estimate the 2-D PSF for deconvolution of radiofrequency data. The idea is the same as in 1-D homomorphic filtering, i.e. making use of the fact, that the 2-D PSF and the original 2-D tissue reflectivity function reside in different bands of the 2-D cepstrum representation of the recorded radiofrequency image data (the assumption will be discussed in Sec. 4.3). The 2-D PSF is obtained by simple lowpass filtering in the cepstrum domain and used in the subsequent Wiener filtering.

As in 1-D homomorphic filtering, the problem of false discontinuities in the 2-D phase image, caused by the wraps around $\pm\pi$, has to be solved (see Sect. 4.2 for more details). The published approaches are based on 2-D phase unwrapping,[22,26,29] logarithmic derivative method[22] and assumption of minimum-phase PSF[22] or symmetric PSF[30] (see Sec. 4.2 for description of the methods).

In[27] the method of homomorphic filtering using phase unwrapping has been further extended to 3-D. The 3-D radiofrequency image was created by free-hand, slow, constant speed manual tilting of a linear array transducer.

4. Homomorphic filtering

4.1. *1-D homomorphic filtering*

There are many signal processing problems where signals are combined by convolution. A common approach to separation of the signal components, deconvolution, is based on accurate knowledge of one of the components (usually the PSF). Homomorphic filtering is an alternative approach where only some a priori knowledge about the components is necessary to separate the two components.

An excellent and widely cited description of homomorphic filtering can be found in.[38] The approach has been used with success in many fields, for example in deconvolution of speech,[38] seismic recordings[40] and DNA sequencing films.[41]

The method is based on a homomorphic system,[38,42] which transforms convolution of two signals

$$y(n) = x(n) * h(n) \tag{18}$$

into the sum of the signals derived from the original ones:

$$\hat{y}(n) = \hat{x}(n) + \hat{h}(n). \tag{19}$$

The principle of the homomorphic deconvolution is shown in Fig. 1.

The method is based on the assumption that the two components of the convolution have different properties, so that they can be separated from each other by simple filtering in the cepstrum domain, as will be shown in Sec. 4.3.

The homomorphic deconvolution schema includes logarithm of a complex-valued function (Fig. 1). This is given as

$$\log Y(z) = ln|Y(z)| + jArgY(z) = \hat{Y}(z). \tag{20}$$

For computational purposes, the Z-transform on the unit circle is considered, i.e. the discrete Fourier transform (DFT) is used. The spectrum phase obtained by the DFT is given only as principal values, i.e. values between $+\pi$ and $-\pi$, which causes false discontinuities in the imaginary part of $\hat{Y}(z)$ at the phase wrap positions. This poses a problem, because $\hat{Y}(z)$ is the Z-transform of the cepstrum domain signal $\hat{y}(n)$ and the Z-transform is defined as a continuous function.[38,42] The same problem concerns the Z-transform on the unit circle i.e. the Fourier transform and its sampled version obtained by DFT.

There are several methods to overcome this inconsistency. A straightforward approach is to unwrap the phase. This is done by adding integral multiple of 2π to its values. An example illustrating the operation is given in Fig. 2.

The Fourier transform of a discrete signal is a continuous periodic function of frequency, meaning that it is also continuous at the borders of the periods. The unwrapped phase often conveys a linear component which leads to violation of this

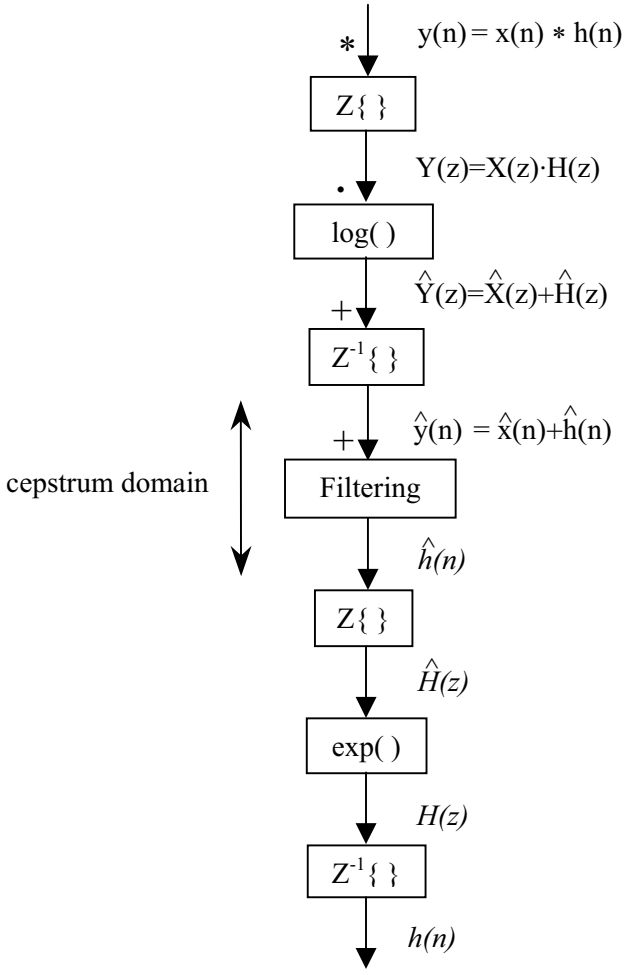


Fig. 1. The principal of homomorphic deconvolution

property, as illustrated in Fig. 2(b). That is why, the linear component has to be subtracted from the unwrapped phase (Fig. 2(c)). This operation corresponds to the shift of the input signal in the time domain.

4.1.1. *Te i g f h hic dec i*

There has been established a specific terminology in homomorphic deconvolution.[38] The term "cepstrum" is a paraphrase of the word spectrum. It indicates the fact that the roles of the time and frequency domains are in a sense interchanged, as will be shown later. That is why the filtering of the cepstrum domain signal is sometimes denoted as "liftering".

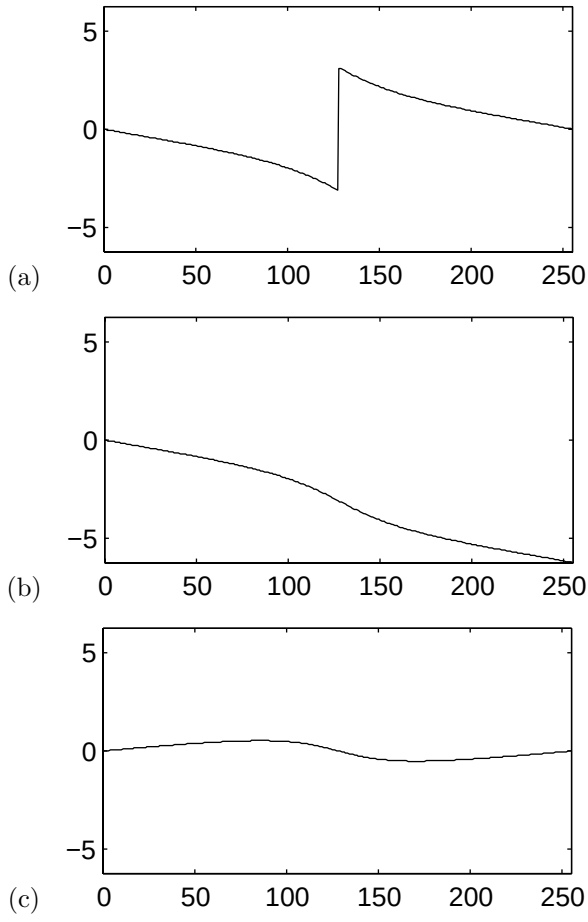


Fig. 2. Illustration of the phase corrections (phase given in radians): (a) principal phase values given by the DFT, (b) unwrapped phase, (c) unwrapped phase after subtraction of the linear component

Some approaches to homomorphic filtering are based on applying the logarithmic function to the power spectrum rather than to its spectrum. In such case only the logarithm of a real-valued function is used. Another way to avoid the complex logarithm is to consider only the spectrum amplitude. In all such cases, the resulting cepstrum is then called the "real cepstrum". If the homomorphic transform includes the logarithm of a complex-valued function, the term "complex cepstrum" is used. It should be noted that both the real and complex cepstrum can be real- or complex-valued. The attribute stands for the way of computing, not for its values. The term "cepstrum" should be used for the general cases which apply to both, the real and complex cepstrum.

4.1.2. *P e ie f c e ce*

The properties of complex cepstrum signals will be explained for discrete real-valued finite-length signals. The Z-transform of such signal is

$$Y(z) = Az^r \prod_{k=1}^{m_i}(1 - a_k z^{-1}) \cdot \prod_{k=1}^{m_o}(1 - b_k z), \tag{21}$$

where $|a_k|$ and $|b_k|$ are less than unity. The terms with z^{-1} correspond to zeroes inside the unit circle, while terms with z correspond to zeroes outside the unit circle.

After applying the logarithm we obtain

$$\hat{Y}(z) = log(A) + log(z^r) + \sum_{k=1}^{m_i} log(1 - a_k z^{-1}) + \sum_{k=1}^{m_o} log(1 - b_k z). \tag{22}$$

The term $log(A)$ contributes only to $\hat{y}(0)$. The term z^r only corresponds to the delay or the advance of the sequence $y(n)$ (i.e. a time shift). It contributes to the complex cepstrum. However, for $r = 0$ the term vanishes. So, the first two terms of Eq. (22) can be neglected, because if two transforms of the form Eq. (21) are multiplied together, it is not possible to determine how each component contributes to A and r.

Supposing the complex cepstrum signal is stable, the convergence regions of $\hat{Y}(z)$ must include the unit circle.[38] Then, the power-series expansions of the terms in Eq. (22) are

$$\begin{array}{ll} log(1 - a_k z^{-1}) = -\sum_{n=1}^{\infty} \frac{a_k^n}{n} z^{-n}, & |z| > |a_k| \\ log(1 - b_k z) = -\sum_{n=1}^{\infty} \frac{b_k^n}{n} z^n, & |z| < |b_k^{-1}|. \end{array} \tag{23}$$

Grouping together terms with the same power of z, the complex cepstrum signal is

$$\hat{y}(n) = \begin{cases} log|A| & n = 0 \\ -\sum_{k=1}^{m_i} \frac{a_k^n}{n}, & n > 0 \\ \sum_{k=1}^{m_o} \frac{b_k^{-n}}{n}, & n < 0. \end{cases} \tag{24}$$

Several properties of the complex cepstrum can be derived from Eq. (24):

(1) The complex cepstrum decays at least as fast as $\frac{1}{n}$, specifically

$$|\hat{x}(n)| < C|\tfrac{\alpha^n}{n}|, -\infty < n < \infty, \tag{25}$$

where C is a constant and α is the maximum of $|a_k|$ and $|b_k|$.

(2) If $y(n)$ is minimum phase (i.e. no zeroes outside the unit circle), then

$$\hat{x}(n) = 0, n < 0. \tag{26}$$

(3) If $y(n)$ is maximum phase (i.e. all zeroes outside the unit circle), then

$$\hat{x}(n) = 0, n > 0. \tag{27}$$

(4) Even for finite-duration signals $y(n)$, the complex cepstrum $\hat{y}(n)$ has always infinite duration.

So far, the zeroes were assumed not to lie on the unit circle. Assuming that $\hat{Y}(z)$ (Eq. (22)) is the Z-transform of a stable sequence, the region of convergence (i.e. the region of z where the z-transform converges) must contain the unit circle. This condition is violated for zeroes on the unit circle as the logarithmic function introduces discontinuities at these points.

To avoid this problem, the sequence can be multiplied by an exponential sequence (exponential weighting)

$$w(n) = \alpha^n \cdot y(n), \tag{28}$$

where α is real and positive. Then the Z-transform of the resulting sequence is

$$W(z) = X(\alpha^{-1}z). \tag{29}$$

4.1.3. *C a i a ea i a i f h hic a f*

Realization using phase unwrapping The method follows the schema in Fig. 1. Instead of the Z-transform, the DFT is used. The Transform is realized using the following sequence of transforms:

$$\begin{aligned} Y(k) &= DFT\{y(n)\} \\ \hat{Y}(k) &= logY(k) \\ \hat{y}_p(n) &= DFT^{-1}\{\hat{Y}(k)\}. \end{aligned} \tag{30}$$

Before the logarithm operation is applied, the phase of $Y(k)$ is unwrapped, as illustrated in Fig. 2. An example of such algorithm is given in.[38] In general, the phase unwrapping algorithms require the spectrum to be sampled at a high-enough rate to interpret correctly fast changes in the spectrum.

The infinite duration of complex cepstrum signals causes aliasing. It corresponds to aliasing caused by sampling of a broadband signal, only the roles of the Fourier and time domain are interchanged. As a result, $\hat{y}_p(n)$ is a time-aliased version of the exact complex cepstrum signal $\hat{y}(n)$.[38] However, according to Eq. (25) the complex cepstrum decays faster than an exponential sequence. To avoid the time aliasing effect, zeros are appended to the time-domain signal before the homomorphic transform.

Realization using polynomial rooting This approach is the direct realization of Eq. (24). The zeroes a_k and $1/b_k$ are found by a numerical method for finding

roots of a polynomial,[43] where the coefficients of the polynomial are the samples of the input signal. The advantage of the method is that it avoids time aliasing in the complex cepstrum. On the other hand, for longer signal sequences the computational load is high.

Realization using logarithmic derivative The approach based on logarithmic derivative is used to calculate the complex cepstrum without the need for phase unwrapping. The Fourier transform of the complex cepstrum is

$$logY(e^{j\omega}) = \sum_{n=0}^{N-1} \hat{y}(n)e^{-j\omega n}. \tag{31}$$

Deriving both sides of Eq. (31) with respect to ω, we get

$$\frac{\frac{\partial Y(e^{j\omega})}{\partial \omega}}{Y(e^{j\omega})} = \sum_{n=0}^{N-1} \hat{y}(n)e^{-j\omega n}(-jn), \tag{32}$$

which can be written using the Fourier transform operator $\mathcal{FT}$ as

$$\frac{\frac{\partial Y(e^{j\omega})}{\partial \omega}}{Y(e^{j\omega})} = -j\mathcal{FT}\{\hat{y}(n)n\}. \tag{33}$$

After the inverse Fourier transform and rearranging, we get

$$\frac{j}{n}\mathcal{FT}^{-1}\{\frac{\frac{\partial Y(e^{j\omega})}{\partial \omega}}{Y(e^{j\omega})}\} = \hat{y}(n). \tag{34}$$

Using the DFT instead of $\mathcal{FT}$ the complex cepstrum is computed as

$$\hat{y}(n) = \frac{j}{n}DFT\{\frac{Y'(k)}{Y(k)}\}, \tag{35}$$

where

$$Y(k) = DFT\{y(n)\} \tag{36}$$

and

$$Y'(k) = \frac{\partial Y(e^{j\omega})}{\partial \omega}\bigg|_{\omega=\frac{2\pi k}{N}} = \sum_{n=0}^{N-1} y(n)e^{-j\omega n}(-jn)\bigg|_{\omega=\frac{2\pi k}{N}} = -jDFT\{y(n)n\}. \tag{37}$$

For $n = 0$ it follows directly from Eq. (30) that

$$\hat{y}(0) = \frac{1}{N} \sum_{k=0}^{N-1} log|Y(k)|. \tag{38}$$

The advantage of the homomorphic transform using Eq. (35)–Eq. (38) is that no phase unwrapping is needed, however at the cost of more severe aliasing than in the case of correctly performed phase unwrapping.[38]

Minimum phase realization The approach is based on the assumption that the sequence to be transformed is minimum-phase,[38,42] i.e. the zeroes of its Z-transform are inside the unit circle. For such signals the phase of their spectrum is related to the amplitude by the Hilbert transform. Thus, it is sufficient to consider only the spectrum amplitude, which substantially simplifies the homomorphic transform:

$$\hat{y}_e(n) = DFT^{-1}\{log|Y(k)|\}. \tag{39}$$

As $log|Y(k)|$ is the real part of the log-spectrum signal representation $\hat{Y}(k)$, its inverse DFT, $\hat{y}_e(n)$, is the even part of the real cepstrum $\hat{y}(n)$. The complete sequence can be obtained from its even part using the following Hilbert transform relation:[38]

$$\hat{y}(n) = \begin{cases} \hat{y}_e(n), & n = 0 \\ 2\hat{y}_e(n), & 1 \le n < N/2, \\ 0, & otherwise. \end{cases} \tag{40}$$

Spectral root cepstrum The spectral root cepstrum approach[44] is a method of homomorphic mapping where the logarithm of a complex function is replaced with a fractional power operation $(\ )^{\gamma}$. The exponential function in the inverse homomorphic transform is correspondingly replaced by $(\ )^{1/\gamma}$. It is a generalization of the common logarithmic homomorphic mapping, in a sense, since for $|\gamma| \to 0$ the spectral root cepstrum approaches the complex cepstrum. The spectral root cepstrum can be used in the same manner as the ordinary complex cepstrum to estimate the PSF in ultrasound data.[23]

A time-domain convolution $y(n) = h(n) * x(n)$ is mapped to the convolution in the spectral root cepstrum domain $\tilde{y}(n) = \tilde{h}(n) * \tilde{x}(n)$, which is different from the complex cepstrum, where a convolution is mapped to a sum.

Since the fractional power of a complex number is multiple valued, the phase of the argument has to be unwrapped first in the same way as in the ordinary logarithmic homomorphic mapping.

Generalized cepstrum Another generalization of the common logarithmic homomorphic mapping is the generalized cepstrum approach.[45] Here, the "generalized

logarithmic" operation is applied to the spectrum instead of a logarithmic operation. It is defined as

$$\check{Y}(e^{j\omega}) = \begin{cases} \frac{1}{\gamma}[Y^{\gamma}(e^{j\omega}) - 1], \gamma \neq 0 \\ log\ Y(e^{j\omega}) \qquad \gamma = 0. \end{cases} \tag{41}$$

In the inverse homomorphic mapping, an inverse function is used instead of the exponential operation:

$$Y(e^{j\omega}) = \begin{cases} [1 + \check{Y}(e^{j\omega})]^{1/\gamma}, \gamma \neq 0 \\ exp\ \check{Y}(e^{j\omega}) \qquad \gamma = 0. \end{cases} \tag{42}$$

A time-domain convolution $y(n) = h(n) * x(n)$ is mapped to the following expression in the generalized cepstrum domain:

$$\check{y}(n) = \check{h}(n) + \check{x}(n) + \gamma \check{h}(n) * \check{x}(n). \tag{43}$$

Similarly to the spectral root cepstrum, the generalized cepstrum approaches the ordinary complex cepstrum for $|\gamma| \to 0$. Again, to obtain a unique solution of the fractional power operation, phase unwrapping has to be used.

4.2. *2-D homomorphic filtering*

The concept of 1-D homomorphic filtering can be extended to 2-D under certain conditions. The existence of a 2-D complex cepstrum representation of images is discussed in.[46] As in 1-D, the 2-D log-spectrum must be a continuous and periodic function, so that the 2-D complex cepstrum exists. This condition is violated, if the 2-D spectrum is zero at any spatial frequency, as the logarithmic operation causes discontinuity at this point. These points are also denoted as phase residua[47] in the theory of 2-D phase unwrapping. The phase residua cause discontinuities in the phase image of the spectrum.

In the following, the concepts of computational realization of 2-D homomorphic mapping are described.

Realization using phase unwrapping Similarly to the 1-D phase unwrapping approach, the homomorphic mapping includes the following sequence of transforms:

$$\begin{gathered} Y(k,l) = DFT\{y(m,n)\} \\ \hat{Y}(k,l) = logY(k,l) \\ \hat{y}(m,n) = DFT^{-1}\{\hat{Y}(k,l)\}. \end{gathered} \tag{44}$$

To solve the problem of false discontinuities in the phase image, 2-D phase unwrapping is done. It is a very noise sensitive procedure. 2-D phase unwrapping has

been an active research field. A detailed description of the 2-D phase unwrapping theory is given in.[48] The description is beyond this chapter's scope. Thus, only the 2-D phase unwrapping methods used in homomorphic filtering of ultrasonic image data are shortly reviewed.

A heuristic method was applied in.[22] The reference point is selected to be in the middle of the phase image (i.e. at zero spatial frequency). The phase is unwrapped only for the two right quadrants. Since real-valued original images are considered, the left quadrants are the antisymmetric counterparts. The phase is unwrapped line by line (in the horizontal direction) so that the phase value being computed $\phi(m,n)$ is given by its original wrapped phase value $\psi(m,n)$ and the already unwrapped phase value of the left adjacent pixel $\phi(m-1,n)$ as follows

$$\phi(m,n) = \psi(m,n) + 2\pi k, \tag{45}$$

where

$$k = round\Big[\frac{\psi(m,n) - \phi(m-1,n)}{2\pi}\Big]. \tag{46}$$

The first pixels in the rows are unwrapped using the unwrapped pixel adjacent in the vertical direction (except for the reference pixel).

After having applied the described phase unwrapping algorithm, the 2-D linear phase component is estimated and subtracted from the resulting phase image and the phase unwrapping is repeated. However, this time it is done by columns rather than by rows. Subsequently, the remaining 2-D linear component is estimated and subtracted.

Another 2-D phase unwrapping algorithm, the Block least squares method, was presented in[49] and applied in 2-D homomorphic deconvolution in.[26] The input phase image is first tessellated into small square blocks, which are then unwrapped separately, using the following operator on every pixel (m,n) of the block

$$\phi(m,n) = W[\psi(m,n) + \rho] - \rho. \tag{47}$$

$\psi(m,n)$ is the observed wrapped phase value at point (m,n). W is a wrap function defined as $W[\psi] = \psi + 2\pi r$, where r is an integer such that $-\pi < W[\psi] \leq \pi$. The parameter ρ in Eq. (47) is a constant specific for each block. ρ must be in the interval $[0, 2\pi)$. It is estimated for each block by minimizing the sum of the absolute values of the first order differences in the block. The unwrapped blocks are then merged by minimizing the sum of squared differences in the adjacent border points of the neighboring blocks.[49]

It was shown[49] that this method makes it possible to directly unwrap the phase in blocks under the assumption that the true phase values of the blocks are in an interval smaller than 2π, i.e. each block contains a maximum of one phase wrap contour.

The phase unwrapping was further improved[26] by a preprocessing step, aiming at removing the phase wraps as much as possible before phase unwrapping. It is done by shifting the recorded spatial-domain image in the radial direction so that the center of the mass is in origio, followed by the same operation in the lateral direction.

In,[29] least-squares phase unwrapping[48] was used. It is based on minimization of the term

$$\Omega = \sum_{i=0}^{M-2}\sum_{j=0}^{N-1}(\phi_{i+1,j} - \phi_{i,j} - \Delta_{i,j}^{x})^2 + \sum_{i=0}^{M-1}\sum_{j=0}^{N-2}(\phi_{i,j+1} - \phi_{i,j} - \Delta_{i,j}^{y})^2, \qquad (48)$$

where $\phi_{i,j}$ denotes the unknown unwrapped phase values. $\Delta_{i,j}^{x}$ and $\Delta_{i,j}^{y}$ are the wrapped phase differences in x and y direction calculated from the known phase values. M and N stand for the size of the processed phase image. In other words, the unwrapped phase values $\phi_{i,j}$ are determined so that their differences agree with the corresponding wrapped differences of the known wrapped phase (i.e. $\Delta_{i,j}^{x}$ or $\Delta_{i,j}^{y}$) in the least-squares sense. To include information about the reliability of phase values in different regions of the phase image, the phase values can have different weights in Eq. (48) (weighted least-squares phase unwrapping).

Realization using logarithmic derivative This alternative to the 2-D phase unwrapping is a direct generalization of the 1-D logarithmic derivative algorithm (section 4.1.3).

The Fourier transform of a complex cepstrum is

$$logY(e^{j\omega_1}, e^{j\omega_2}) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\hat{y}(m,n)e^{-j(\omega_1 m + \omega_2 n)}. \qquad (49)$$

Differentiating both sides of Eq. (49) with respect to ω_1, we get

$$\frac{\frac{\partial Y(e^{j\omega_1}, e^{j\omega_2})}{\partial \omega_1}}{Y(e^{j\omega_1}, e^{j\omega_2})} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\hat{y}(m,n)e^{-j(\omega_1 m + \omega_2 n)}(-jm), \qquad (50)$$

which can be written using the 2-D Fourier transform operator $\mathcal{FT}$ as

$$\frac{\frac{\partial Y(e^{j\omega_1}, e^{j\omega_2})}{\partial \omega_1}}{Y(e^{j\omega_1}, e^{j\omega_2})} = -j\mathcal{FT}\{\hat{y}(m,n)m\}. \qquad (51)$$

After the inverse Fourier transform and rearranging

$$\frac{j}{m}\mathcal{FT}^{-1}\{\frac{\frac{\partial Y(e^{j\omega_1},e^{j\omega_2})}{\partial\omega_1}}{Y(e^{j\omega_1},e^{j\omega_2})}\}=\hat{y}(m,n),\qquad m\neq 0. \tag{52}$$

Using the 2-D DFT instead of the 2-D $\mathcal{FT}$, the complex cepstrum is computed as

$$\hat{y}(m,n)=\frac{j}{m}DFT^{-1}\{\frac{Y_1'(k,l)}{Y(k,l)}\}\qquad m\neq 0, \tag{53}$$

where

$$Y(k,l)=DFT\{y(m,n)\} \tag{54}$$

and

$$\begin{aligned} Y_1'(k,l)&=\left.\frac{\partial Y(e^{j\omega_1},e^{j\omega_2})}{\partial\omega_1}\right|_{\omega_1=\frac{2\pi k}{M},\omega_2=\frac{2\pi l}{N}}= \\ &=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}y(m,n)e^{-j(\omega_1 m+\omega_2 n)}(-jm)\Big|_{\omega_1=\frac{2\pi k}{M},\omega_2=\frac{2\pi l}{N}}= \\ &=-jDFT\{y(m,n)m\}. \end{aligned} \tag{55}$$

To compute $\hat{y}(m,n)$ for $m=0$, differentiation with respect to ω_2 has to be used. Using the same procedure, it follows

$$\hat{y}(m,n)=\frac{j}{n}DFT^{-1}\{\frac{Y_2'(k,l)}{Y(k,l)}\}\qquad n\neq 0, \tag{56}$$

where

$$Y_2'(k,l)=-jDFT\{y(m,n)n\}. \tag{57}$$

For $m=0,n=0$ it follows directly from Eq. (44) that

$$\hat{y}(0,0)=\frac{1}{MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1}log|Y(k,l)|. \tag{58}$$

It should be noted that Equations (53) and (56) both give the complex cepstrum values for $m\neq 0$, $n\neq 0$. However, the values are equal only in case of no phase residua.

Real cepstrum realization Similarly to the 1-D homomorphic transform, the logarithm can be applied only to the 2-D amplitude spectrum by assuming that the estimated convolution component (e.g. PSF) is minimum-phase.[22,29] Another approach to avoid processing of phase is the assumption that the estimated convolution component is 2-D symmetric.[30] Then the spectrum phase is zero.

4.3. *Homomorphic filtering of ultrasonic images*

This section deals with the application of homomorphic filtering in estimation of the PSF in ultrasonic radiofrequency images, as introduced in.[23,37] For sake of simplicity, only the estimation of the 1-D PSF in the axial direction is considered first. The results will be related to the 1-D PSF in the lateral direction and the 2-D PSF at the end of this section.

To illustrate homomorphic deconvolution of ultrasound radiofrequency signals, synthetic radiofrequency signal was generated, simulating a signal recorded for one beam. The 1-D PSF (Fig. 3 left) was modeled in a simple way[17,18] as a cosine modulated Gaussian function:

$$cos(\omega_0 n) \cdot e^{-\frac{(n-n_0)^2}{2\sigma^2}}, \tag{59}$$

where ω_0, n_0 and σ denote the parameters of the pulse.

The original tissue reflectivity function (Fig. 3 right) was modeled as randomly distributed reflectors of random reflection coefficients. Thus each reflector was represented by Dirac pulse of random weight at a random time position.

The synthetic radiofrequency signal (Fig. 4 left) was obtained as a convolution of the PSF and the tissue reflectivity function.

The log-spectrum domain plots of the PSF and the tissue reflectivity function (Fig. 3a,b) show that the log-spectrum of the PSF is fairly smooth while rapid changes are dominant in the log-spectrum representation of the tissue reflectivity function. The difference can also be observed in the complex cepstrum domain, where the PSF cepstrum mostly resides close to the zero index position, while the tissue reflectivity function cepstrum is spread further to outer parts.

When observed in vitro (e.g. the pulse-echo signal obtained as the echo from a planar reflector orthogonal to the beam in a water tank), the PSF is a smooth, continuous function of time, having a fairly smooth spectrum. The absorption of the pulse energy by tissues is also a process with smooth frequency response. Hence, it is reasonable to suppose that the ultrasound PSF in tissues is also a rather smooth function in the spectrum and log-spectrum domain.

A typical in vivo radiofrequency line has a very "spiky" amplitude spectrum. As the radiofrequency signal is a convolution of the PSF and the tissue reflectivity function (neglecting noise), i.e. the product of their spectra in the Fourier domain, the rapid changes in the spectrum are ascribed to the tissue reflectivity function.

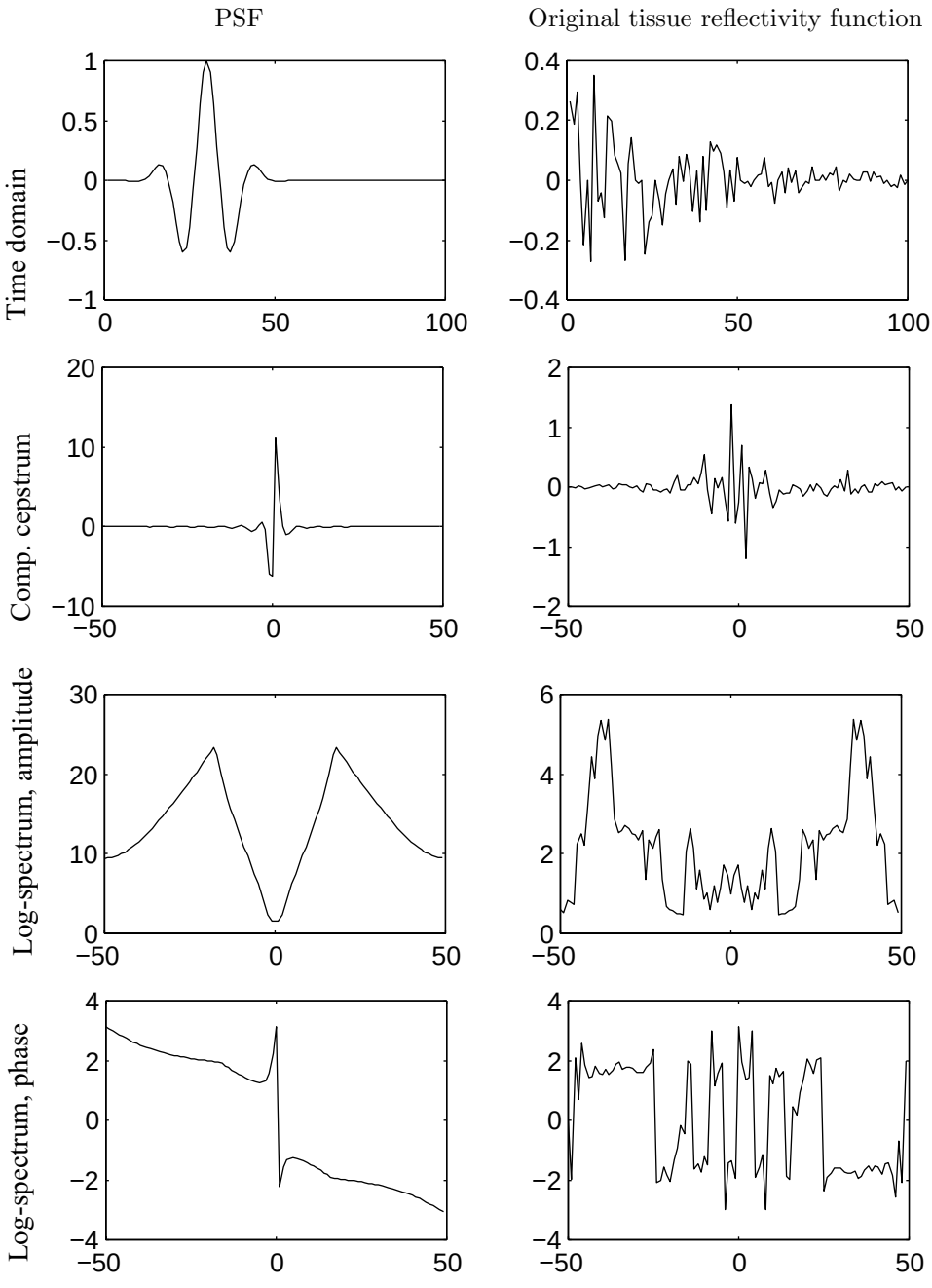


Fig. 3. Illustration of homomorphic filtering on synthetic radiofrequency signal. Left column: PSF. Right column: original tissue reflectivity function.

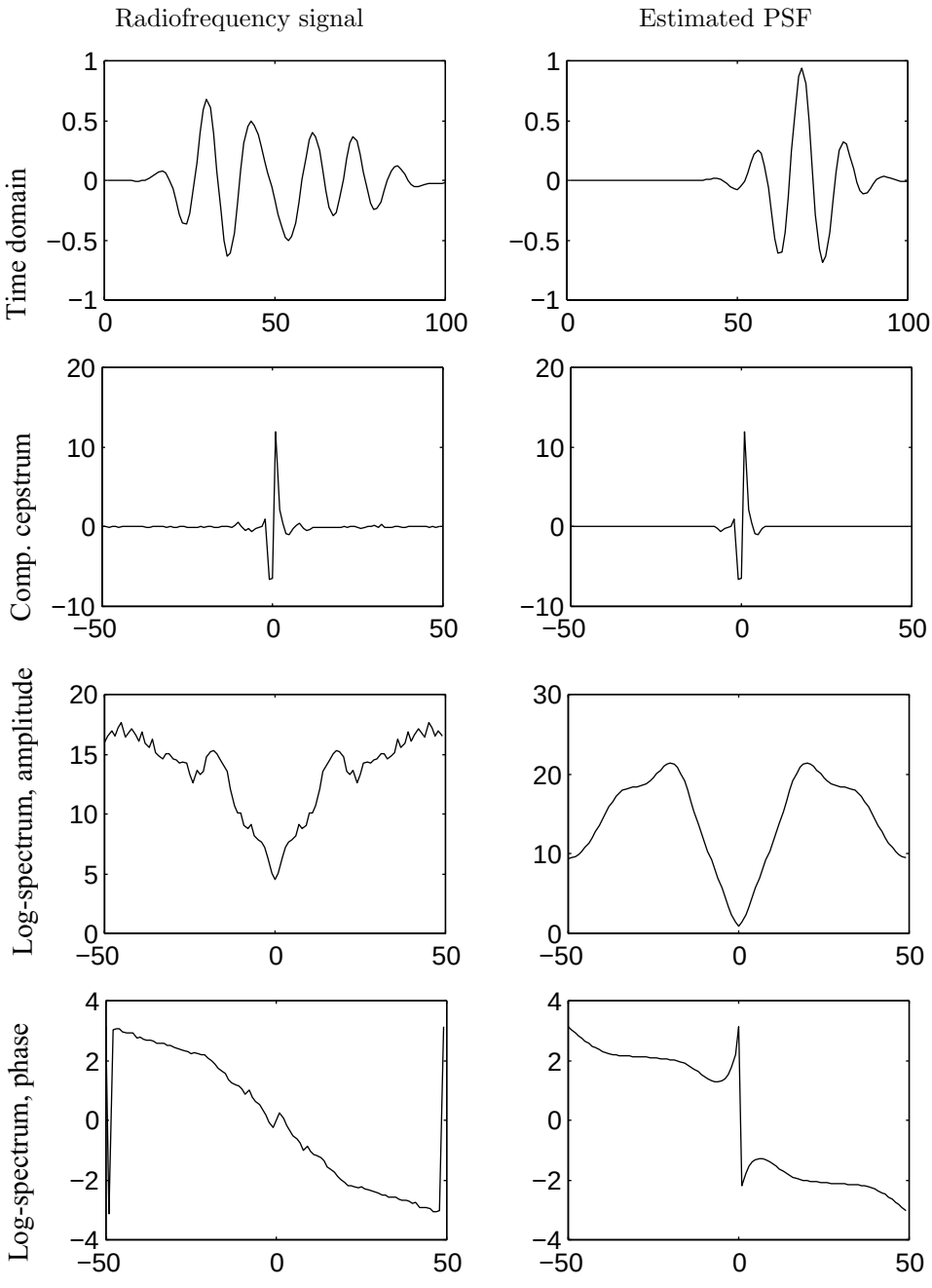


Fig. 4. Illustration of homomorphic filtering on synthetic radiofrequency signal. Left column: radiofrequency signal. Right column: estimated PSF.

From the above described a priori assumptions about the PSF and the tissue reflectivity function it follows, that the operation of estimating the PSF component can be interpreted as smoothing the log-spectrum of the radiofrequency signal. In the complex cepstrum domain, the operation corresponds to masking out the values lying in the outer parts of the time axis. In the example (Fig. 4), the complex cepstrum of the radiofrequency signal was weighted by a rectangular window.[22,23,38]

Thinking of the complex cepstrum and the log-spectrum as the time domain and the Fourier domain representations of signals, respectively, the operation can be viewed as low-pass filtering with the time and Fourier domain interchanged. Despite of the swapped roles of the time and Fourier domains, the term linear filtering is widely used in the literature.

To avoid the Gibbs ringing phenomenon, a Butterworth low-pass filter, with the time and Fourier domains interchanged, is used in[24–26] instead of weighting by a rectangular window. The Butterworth filter has the form

$$h_{LP}(n) = \frac{1}{1 + (n/D_0)^{2r}}, \tag{60}$$

where r is the order and D_0 is the locus. The low-pass filtering is done by multiplication of the complex cepstrum by $h_{LP}(n)$.

A procedure of determining the cut-off value of the cepstrum domain low-pass filter (the width of the rectangular window or the locus value of the Butterworth filter) was given in.[37] The cut-off value is determined from the log-spectrum of the recorded radiofrequency signal as the fastest change that can be ascribed to the PSF. It is usually the first uprising flank of the log-spectrum. First, the "rise-time" Δf is found as the frequency span between 10% and 90% of the maximum log-spectrum amplitude. Then it is normalized by the sampling frequency f_s. Taking its inverse value yields directly the cepstrum sample value where the cut-off should be made. The estimate can be improved by combining more segments of radiofrequency data in order to smooth the log-spectrum.

The cut-off value estimated using the above described procedure is usually considered as a rough estimate, which is further adjusted by trial and error.[22,26]

To further support the appropriateness of this type of homomorphic filtering in estimation of the PSF in ultrasonic images, a mathematical reasoning,[38] concerning a special class of signals, will be given. It shows the suitability of homomorphic filtering in so called reverberant environments. In such cases a signal can be represented as a sum of a number of overlapping delayed replicas, or echoes, of the same basic waveform. Ultrasonic imaging is an example of such cases.

A signal $y(n)$ containing scaled, delayed echoes (e.g. radiofrequency ultrasound signal) can be modeled as a convolution of the echo wave $h(n)$ (ultrasound PSF in ultrasonic imaging) and a signal $x(n)$, which is a train of M impulses (modeling

tissue reflectors or scatterers in ultrasonic imaging):

$$\begin{array}{c} y(n) = h(n) * x(n), \\ x(n) = \delta(n) + \sum_{k=1}^{M} a_k \delta(n - n_k) \end{array} \tag{61}$$

Let us first consider the case of a single echo, i.e.

$$x(n) = \delta(n) + a_1 \delta(n - n_1). \tag{62}$$

The Fourier transform of $x(n)$ is

$$X(e^{j\omega}) = 1 + a_1 e^{-j\omega n_1} \tag{63}$$

and the log-spectrum

$$\hat{X}(e^{j\omega}) = log(1 + a_1 e^{-j\omega n_1}). \tag{64}$$

If $|a_1| < 1$, the corresponding complex cepstrum is obtained using the power-series expansion and the inverse DFT as

$$\hat{x}(n) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{a_1^k}{k} \delta(n - kn_1). \tag{65}$$

If the log-spectrum of $h(n)$ is slowly varying, its complex cepstrum values $\hat{h}(n)$ lie close to the zero-time index, while $\hat{x}(n)$ has zero values for $n < n_1$. Thus the signals can be easily separated in the complex cepstrum domain.

In the general case of more echoes,

$$x(n) = \delta(n) + \sum_{k=1}^{M} a_k \delta(n - n_k) \tag{66}$$

and the log-spectrum

$$\hat{X}(e^{j\omega}) = log(1 + \sum_{k=1}^{M} a_k e^{-j\omega n_k}). \tag{67}$$

In this case, the complex cepstrum $\hat{x}(n)$ consists of impulses at times that are complicated functions of the original delays. However, in the special case when $x(n)$ is minimum-phase, it can be shown (using the power series expansion of Eq. (67)) that $\hat{x}(n) = 0$ for $n < n_m$, where n_m is the shortest original delay.

The given derivation only applies to a special class of signals, the minimum-phase signals. The discussion in Sec. 4.1.2 suggests the use of exponential weighting to transform the input signal into a minimum-phase signal. Indeed, the exponential

weighting is used to increase the degree of separation between the two components of the deconvolution,[22,24,25,38] to a certain extent. However, a low value of α (Eq. (28)) causes smoothing in the frequency domain.[38] Thus, the exponential weighting factor α is usually found by trial and error. Values between 0.95 and 1.00 are used.

So far, only 1-D homomorphic filtering in the axial direction was considered in this section. The same reasoning can, however, be applied to the estimation of the PSF in the lateral direction. The lateral PSF is also supposed to be a fairly smooth, continuous function in the time and Fourier domains and thus, in the log-spectrum domain as well. As in the axial direction the tissue properties in the lateral direction exhibit rapid changes in the spectrum and log-spectrum. Consequently, the 1-D homomorphic deconvolution is done in the same way as in the axial direction.

The 2-D homomorphic deconvolution is a straightforward extension of the 1-D homomorphic deconvolution, where the assumption of separability of the PSF is not made. The same properties of the 2-D PSF and the 2-D tissue reflectivity function hold as in the 1-D case.

5. Examples of Homomorphic Deconvolution

To show the potential of homomorphic filtering in deconvolution of ultrasonic images, some illustrating results are given.[30] The radiofrequency signals were recorded from a tissue-mimicking phantom Gammex 403GS as well as from clinical examinations of pancreas, aorta and kidney. All image sequences were recorded using GE Vingmed Ultrasound System 5 scanner. For clinical images, a curved linear array probe with 192 elements (CLA 192 3.5) was used; nominal frequency 3.5 MHz. For phantom images, a phased array probe with 64 elements (FPA 64 2.5c) was used; nominal frequency 2.5 MHz. The image data were recorded using harmonic imaging mode, i.e. the radiofrequency signal contained both the first and second harmonics.

The imaged phantom consisted of cystic objects and nylon threads perpendicular to the scanning plane (i.e. point targets in the image). The deconvolved phantom images of the cystic objects (Fig. 5) have sharper border outline, compared to the unprocessed input images. Also the point targets (Fig. 6) of the phantom are shown as smaller spots for the deconvolved images.

The clinical images are shown in Figs. 7, 8, 9. In general, the deconvolved images are sharper and show the imaged structures more clearly than the original ones. The delineations of the tissue borders are improved. Some structures visible in the deconvolved images are only hardly visible in the original images or not visible at all. The speckle pattern of the deconvolved images is more distinct with shorter spatial correlation. In the images of pancreas area (Fig. 7), the deconvolution performance can be assessed on several structures. The delineation of aorta is considerably improved in the deconvolved image. The inner and outer border of the

aorta wall becomes clearly visible. The delineation of other vessels in the image as well as the delineation of the vertebral bone becomes sharper. Similar improvements can be observed in the deconvolved images of kidney (Fig. 8) and aorta (Fig. 9). Especially, the border of the kidney cortex and kidney pelvis becomes sharper and more precise, as well as the border of aorta.

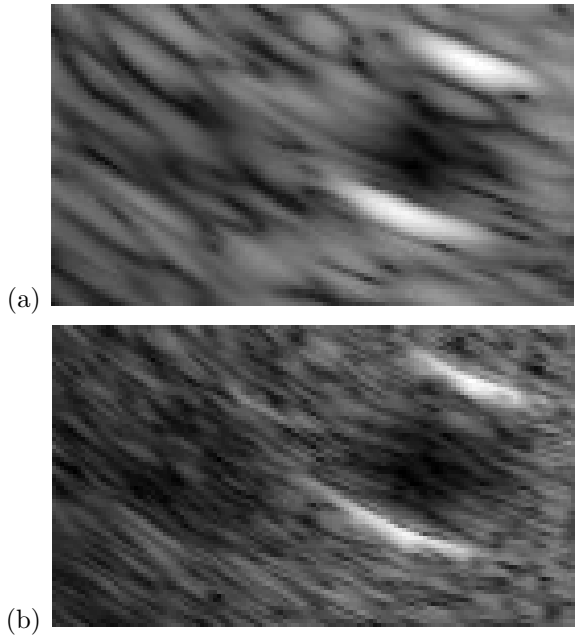


Fig. 5. Phantom images in the region of two cystic objects of different acoustic properties. (a) Input image. (b) deconvolved image. After Ref.[30] (©2006 IEEE).

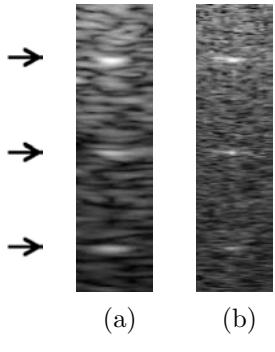


Fig. 6. Phantom images in the region of three point targets (nylon pins), at positions shown by arrows. (a) Input image. (b) deconvolved image. After Ref.[30] (©2006 IEEE).

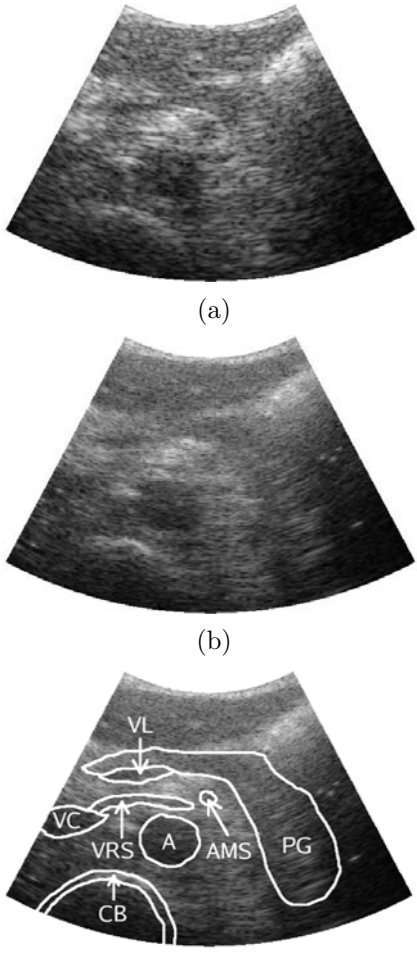


Fig. 7. Images of pancreas. (a) Original image. (b) Deconvolved image. (c) Structures in the image: VL - vena lienalis, VC - vena cava inferior, VRS - vena renalis sinistra, A - aorta, AMS - arteria mesenterica superior, PG - pancreatic gland, CB - cortical bone of lumbal vertebral body. After Ref.[30] (©2006 IEEE).

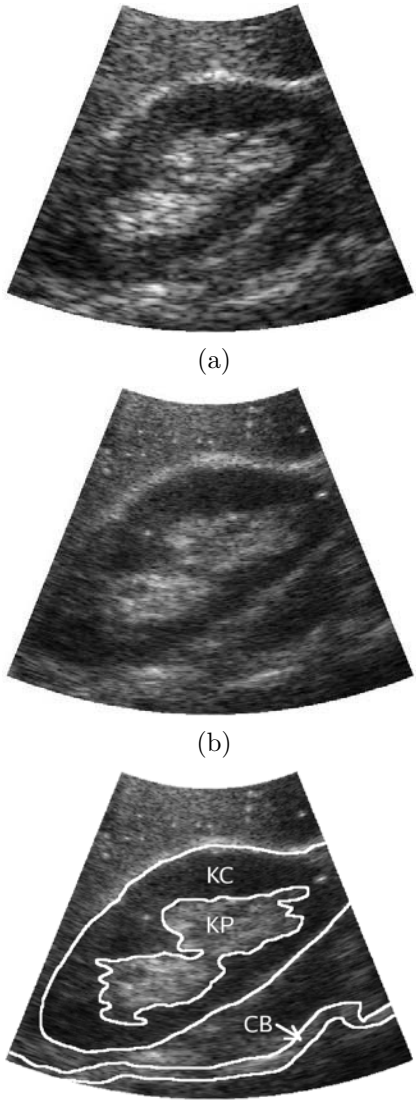


Fig. 8. Images of kidney. (a) Original image. (b) Deconvolved image. (c) Structures in the image: KC - kidney cortex, KP - kidney pelvis, CB - cortical bone of lumbal vertebral bodies. After Ref.[30] (©2006 IEEE).

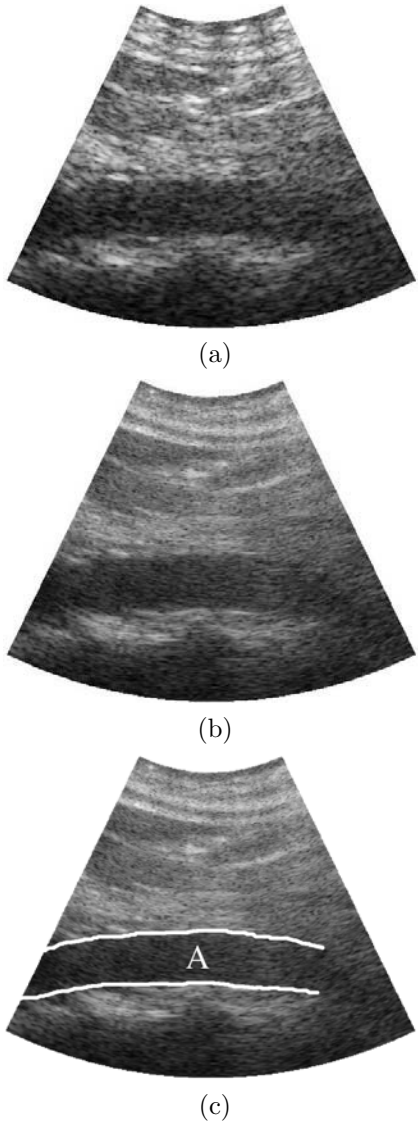


Fig. 9. Images of aorta. (a) Original image. (b) Deconvolved image. (c) Structures in the image of aorta: A - aorta. After Ref.[30] (©2006 IEEE).

References

1. Y. Goussard, G. Demoment, and J. Idier. A new algorithm for iterative deconvolution of sparse spike trains. In *Proc. IEEE Ultrasonics Symposium*, pp. 1547–1550, (1990).
2. K. F. Kåresen and E. Bølviken, Blind deconvolution of ultrasonic traces accounting for pulse variance, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **46**(3), 564–573, (1999).
3. K. F. Kåresen, Evaluation and applications of the iterated window maximization method for sparse deconvolution, *IEEE Trans. Signal Processing.* **46**(3), 609–624, (1998).
4. C. R. Hill, J. C. Bamber, and G. R. ter Haar, *Physical Principles of Medical Ultrasonics.* (John Wiley & Sons, 2002).
5. B. Angelsen, *Ultrasound Imaging: Waves, Signals, and Signal Processing.* (Emantec, Trondheim, Norway, 2000). ISBN 82-995811-0-9.
6. J. A. Jensen. Ultrasoundn fields in an attenuating medium. In *Proc. IEEE Ultrasonics Symposium*, vol. 2, pp. 943–946, (1993).
7. J. A. Jensen. A new approach to calculating spatial impulse responses. In *Proc. IEEE International Ultrasonics Symposium*, pp. 1755–1759, (1997).
8. M. Fink, W. A. Kuperman, J. P. Montagner, and A. T. (Eds.), *Imaging of Complex Media with Acoustic and Seismic Waves.* Topics in Applied Physics, (Springer Verlag, 2002).
9. J. Ng, R. Prager, N. Kingsbury, G. Treece, and A. Gee, Modeling ultrasound imaging as a linear, shift-variant system, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **53**(3), 549–563, (2006).
10. J. Jan and P. Kilián. Modified Wiener approach to restoration of ultrasonic scans via frequency domain. In *Proc. 9th Scandinavian Conference on Image Analysis*, pp. 1173–1181, (1995).
11. J. A. Jensen, Estimation of pulses in ultrasound B-scan images, *IEEE Trans. Med. Imag.* **10**(2), 164–172, (1991).
12. J. A. Jensen, Deconvolution of ultrasound images, *Ultrasonic Imaging.* **14**(1), 1–15, (1992).
13. J. A. Jensen, Estimation of in vivo pulses in medical ultrasound, *Ultrasonic Imaging.* **16**(3), 190–203, (1994).
14. U. R. Abeyratne, A. P. Petropulu, and J. M. Reid, Higher order spectra based deconvolution of ultrasound images, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **42**(4), 1064–1075, (1995).
15. C. N. Liu, M. Fatemi, and R. C. Waag, Digital processing for improvement of ultrasonic abdominal images, *IEEE Trans. Med. Imaging.* **MI-2**(2), 66–75, (1983).
16. G. Cincotti, R. Carotenuto, G. Cardone, P. Gori, and M. Pappalardo, Real-time deconvolution in ultrasonic imaging systems, *Journal of Computational Acoustics.* **9** (3), 745–755, (2001).
17. O. Husby, T. Lie, T. Langø, J. Hokland, and H. Rue, Bayessian 2-d deconvolution: A model for diffuse ultrasound scattering, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **48**(1), 121–130, (2001).
18. T. Langø, T. Lie, O. Husby, and J. Hokland, Bayessian 2-d deconvolution: Effect of using spatially invariant ultrasound point spread functions, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **48**(1), 131–141, (2001).
19. M. Fatemi and A. C. Kak, Ultrasonic B-scan imaging: Theory of image beam formation and a technique for restoration, *Ultrasonic Imaging.* **2**(2), 1–47, (1980).
20. U. R. Abeyratne, A. P. Petropulu, J. M. Reid, T. Golas, E. Conant, and F. Forsberg, Higher order versus second order statistics in ultrasoundn image deconvolution, *IEEE*

Trans. Ultrason. Ferroelec. Freq. Cont. **44**(6), (1997).
21. A. Yamani, M. Bettayeb, and L. Ghouti. High-order spectra-based deconvolution of ultrasonic NDT signals for defect identification. In *Proc. 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 1457–1460, (1999).
22. T. Taxt, Restoration of medical ultrasound images using two-dimensional homomorphic deconvolution, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **42**, 543–554, (1995).
23. T. Taxt, Comparison of cepstrum based methods for radial blind deconvolution of ultrasound images, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **44**(3), 666–674, (1997).
24. T. Taxt and G. V. Frolova, Noise robust one-dimensional blind deconvolution of medical ultrasoundn images, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **46**(2), 291–299, (1999).
25. G. V. Frolova and T. Taxt. Homomorphic deconvolution of medical ultrasoundn images using a Bayesian model for phase unwrapping. In *Proc. 1996 IEEE Ultrasonics Symposium*, pp. 1371–1376, (1996).
26. T. Taxt and J. Strand, Two-dimensional noise-robust blind deconvolution of ultrasound images, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **48**(4), 861–866, (2001).
27. T. Taxt, Three-dimensional blind deconvolution of ultrasoundn images, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **48**, 867–871, (2001).
28. D. Adam and O. Michailovich, Blind deconvolution of ultrasound sequences using nonparametric local polynomial estimates of the pulse, *IEEE Trans. Biomed. Eng.* **49** (2), 118–131, (2002).
29. T. Taxt and R. Jiřík, Superresolution of ultrasound images using the 1st and 2nd harmonic signal, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **51**(2), 163–175, (2004).
30. R. Jiřík and T. Taxt, High-resolution medical ultrasonic images using fast two-dimensional homomorphic filtering, *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.* **53** (8), 1440–1448, (2006).
31. A. Yamani, M. Bettayeb, and L. Ghouti, High-order spectra-based deconvolution of ultrasonic NDT signals for defect identification, *Ultrasonics.* **35**(7), 525–531, (1997).
32. J. M. Mendel and J. Kormylo, New fast optimal white-noise estimators for deconvolution, *IEEE Trans. Geosci. Electron.* **GE-15**(1), 32–41, (1977).
33. J. M.Mendel, Minimum-variance deconvolution, *IEEE Trans. Geosci. Remote Sens.* **GE-19**(3), 161–171, (1981).
34. R. Kuc and H. Li, Reduced order autoregressive modeling for center-frequency estimation, *Ultrasonic Imaging.* **7**(3), 244–251, (1985).
35. L. Y. Shih, C. W. Barnes, and L. A. Ferrari, Estimation of attenuation coefficient for ultrasonic tissue characterization using time-varying state-space model, *Ultrasonic Imaging.* **10**(2), 99–109, (1988).
36. G. Salomonsson and B. Löfström, Analysis of a system for ultrasonic imaging of attenuation and texture in soft tissue, *Ultrasonic Imaging.* **7**(3), 225–243, (1985).
37. J. A. Jensen and S. Leeman, Nonparametric estimation of ultrasound pulses, *IEEE Trans. Biomed. Eng.* **41**(10), 929–936, (1994).
38. A. V. Oppenheim and R. W. Schafer, *Discrete Time Signal Processing.* (Prentice Hall, 1989).
39. J. G. Proakis, C. M. Rader, F. Ling, and C. L. Nikias, *Advanced Digital Signal Processing.* (Macmillan Publishing Company, New York, USA, 1992).
40. T. J. Ulrych, Application of homomorphic deconvolution to seismology, *Geophysics.* **36**(4), 650–660, (1971).
41. J. T. Ives, R. Gesteland, and J. T. G. Stockham, An automated film reader for DNA sequencing based on homomorphic deconvolution, *IEEE Trans. Biomed. Eng.* **41**(6), 509–519, (1994).

42. J. Jan, *Digital Signal Filtering, Analysis and Restoration*. (IEE, London, United Kingdom, 2000).
43. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*. (Cambridge University Press, 1993). ISBN 0521431085. available on-line: http://www.library.cornell.edu/nr/bookcpdf.html; accessed August 19, 2003.
44. J. S. Lim, Spectral root homomorphic deconvolution system, *IEEE Trans. Acoust., Speech, Signal Processing*. **ASSP-27**(3), 223–233, (1979).
45. T. Kobayashi and S. Imai, Spectral analysis using generalized cepstrum, *IEEE Trans. Acoust., Speech, Signal Processing*. **ASSP-32**(5), 1087–1089, (1984).
46. D. E. Dudgeon, The existence of cepstra for two-dimensional rational polynomials, *IEEE Trans. Acoust., Speech, Signal Processing*. **ASSP-23**(2), 242–243, (1975).
47. D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms and Software*. (John Wiley & Sons, Inc., 1998). ISBN 0-471-24935-1.
48. D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms and Software*. (John Wiley & Sons, Inc., 1998). ISBN 0-471-24935-1.
49. J. Strand, T. Taxt, and A. K. Jain, Two-dimensional phase unwrapping using a block least-squares method, *IEEE Trans. Image Processing*. **8**(3), 375–386, (1999).

CHAPTER 25

FREQUENCY-WAVENUMBER DOMAIN FILTERING FOR IMPROVED DAMAGE VISUALIZATION

Massimo Ruzzene

Sch f Ae ace E gi ee i g
Ge gia I i e f Tech g
A a a GA, 30332

A technique for the analysis of full wavefield data in the wavenumber/frequency domain is presented as an effective tool for damage detection, visualization and characterization. Full wavefield data contain a wealth of information regarding the space and time variation of propagating waves in damaged structural components. Such information can be used to evaluate the response spectrum in the frequency/wavenumber domain, which effectively separates incident waves from reflections caused by discontinuities encountered along the wave paths. This allows removing the injected wave from the overall response through simple filtering strategies, thus highlighting the presence of reflections associated to damage. The concept is first illustrated on analytical and numerically simulated data, and then tested on experimental results. In the experiments, full wavefield measurements are conveniently obtained using a Scanning Laser Doppler Vibrometer, which allows the detection of displacements and/or velocities over a user-defined grid, and it is able to provide the required spatial and time information in a timely manner. Tests performed on an aluminum plate with artificially seeded slits simulating longitudinal cracks show the effectiveness of the technique and its potential for application to the inspection of a variety of structural components.

1. Introduction

The need to reduce the life-management costs of structural components drives the development of health monitoring and damage prognosis technologies that enable more accurate component failure predictions. Structural Health Monitoring (SHM) is the first step to reducing uncertainty in component life prediction. The objectives of SHM include the determination of existence, type , location and size of damage. In particular, damage assessment is a pre-cursor to the complete evaluation of the state of the health of the structure to permit its continued use or proceed with maintenance and repairs. A subtle, but important distinction differentiates SHM from Non-Destructive Evaluation (NDE). NDE supports the notion of examination per scheduled maintenance, with focus limited to predetermined critical locations, or with knowledge that damage has already occurred. Established practices of

NDE such as die-penetrant, ultrasonics, and X-ray generally are used to confirm the presence of damage, or to estimate its extent. The NDE concept does not support the overall inspection of the health of a structure, where a basic question must be answered on the existence of the damage. The technique presented in this chapter aims at transferring some of the concepts developed for NDE to a SHM context.

Among the various technologies explored for SHM, guided waves testing has shown great potentials.[1,2] The attractive features of guided waves include sensitivity to a variety of damage types, as well as the capability of traveling relatively long distances within the structure under investigation. In most applications, guided waves are generated and received by transmit/receive transducer pairs distributed over the test specimen according to convenient patterns. The fundamentals of this type of operation consist in evaluating the characteristics of propagation along the wave path between each transducer/receiver pair. This technology has demonstrated its effectiveness in detecting small damage and discontinuities in the structure. The precise localization of damage however requires the application of additional algorithms which process the information obtained by a distributed sensors array.[3] Alternatively, piezoelectric phased arrays have been proposed to focus guided waves and directly obtain information regarding damage direction[4] or to tune a specific wave mode in a multi-mode signal.[5] Phased arrays have been also proposed as directional filters to remove boundary reflections and directly identify the location of damage.[6]

A similar objective is here pursued, without however the application of phased arrays. A Scanning Laser Doppler Vibrometer (SLDV) is used to measure the velocity of the inspected surface in points belonging to a predefined grid. Scanning the grid and post-processing the data allows the detection and the visualization of the full wavefield as it propagates in the structure. The resulting images describe the main features of the propagating wave and show its interactions with discontinuities that may be encountered along the wave path.[7,8] Damage can be thus immediately detected and localized through limited processing. Moreover, wavefield time-domain data can be fed to signal processing algorithms to obtain detailed images of the defect. For example, the application of multi-dimensional Fourier Transforms (FT)s in space and time provide the representation of the component response in the frequency/wavenumber domain. In this domain, all wave components propagating in directions opposite to the direction of propagation of the main injected pulse are clearly highlighted, so that the presence of reflections and mode conversions caused by the presence of damage are immediately noticeable.[9] The application of two-dimensional FT (2D FT) has previously been proposed in[10] for the analysis of multi-mode wave signals for the identification of dispersion relations and the characterization of the various modes. In this work, the frequency/wavenumber representation is used to devise simple filtering strategies which are able to eliminate the applied excitation and corresponding propagating wave from the recorded

response, while highlighting the presence of any reflections along the wave path. A similar concept has been applied in the past in the biomedical field, where filtering tools allowing identification of patterns of the diagnostic images which provide vital health information about the screened individual are extensively in use and in constant development.

This chapter is organized in 5 sections including this introduction. Section 2 illustrates the concept of frequency/wavenumber filtering through its application to simple analytical configuration, while section 3 presents the results of tests performed on numerically simulated examples. Experimental results are presented in section 4, while section 5 summarizes the main results of the work and provides recommendation for future research.

2. Concept of Frequency/Wavenumber Filtering

A promising technique, which is readily applicable to SHM is a mode filtering technique developed in support of Magnetic Resonance Elastography (MRE).[11,12] MRE measures the displacements caused by a shear wave propagating within the tissue. From the displacement data, MRE evaluates the propagation characteristics of the shear wave which are then used to quantify elastic properties of the tissue. When processing the MRE cine images it is difficult to separate the different modes of tissue motion, which generally includes vibrations, guided waves, shear waves, reflections, and ambient background noise. The shear wave motion, from which tissue properties are estimated, can be effectively isolated from the overall displacement, using a wave mode filtering procedure implemented with a three dimensional (3D) FT. The MRE technique consists of three main components: methods for generating acoustic shear waves, magnetic resonance imaging sequences capable of imaging cyclic displacements, and image processing techniques. The SHM technology presented in this work operates in a very similar fashion, as it includes an exciter used to generate elastic waves, the scanning laser system used as an imaging tool and the filtering technique to be presented below acting as a processing tool. It is a well known fact that damage is a source of secondary waves which emanate from the damage location. The frequency/wavenumber representation can be used to effectively isolate the secondary waves associated with damage from the overall response which is a combination of primary wave generated by the exciter and background noise.

2.1. *Application to a One Dimensional Propagating Wave*

The basic concept behind the filtering technique under consideration is illustrated for a stress wave in a one dimensional (1D) non-dispersive waveguide. The considered configuration is depicted in Figure 1, and assumes the presence of a discontinuity in material at location $x = x_0$. The 1D stress in the region $x < x_0$ can be expressed as a sum of incident and reflected stress waves, according to the following

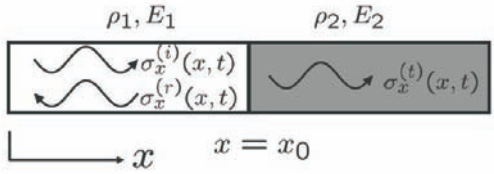


Fig. 1. Wave propagation in 1D elastic medium with material discontinuity.

expression:

$$\sigma_x(x,t) = \sigma_x^{(i)}(x,t) + \sigma_x^{(r)}(x,t). \tag{1}$$

where

$$\sigma_x^{(i)}(x,t) = f(t - \frac{x}{c_{L_1}}). \tag{2}$$

and

$$\sigma_x^{(r)}(x,t) = g(t + \frac{x}{c_{L_1}} - 2\frac{x_0}{c_{L_1}}). \tag{3}$$

respectively define the incident and the reflected waves. Also in eq.s (2,3), f, g are the functions defining the waves propagating in the considered media, while c_{L_1} is the phase velocity in region 1. Imposing velocity and stress interface conditions:

$$\dot{u}^{(i)}(x_0^{(-)},t) + \dot{u}^{(r)}(x_0^{(-)},t) = \dot{u}^{(t)}(x_0^{(+)},t), \tag{4}$$

$$\sigma_x^{(i)}(x_0^{(-)},t) + \sigma_x^{(r)}(x_0^{(-)},t) = \sigma_x^{(t)}(x_0^{(+)},t). \tag{5}$$

yields the well known relationships between the functions f, g:

$$g(\bullet) = \mathcal{R} f(\bullet) \tag{6}$$

The reflection coefficient $\mathcal{R}$ is defined in terms of density ρ and wave speed c_L of the material of the two domains composing the waveguide:

$$\mathcal{R} = \frac{\rho_2 c_{L_2}/\rho_1 c_{L_1} - 1}{\rho_2 c_{L_2}/\rho_1 c_{L_1} + 1} \tag{7}$$

In eq. (5), the transmitted stress wave $\sigma_x^{(t)}$ is given by:

$$\sigma_x^{(t)}(x,t) = \mathcal{T} f(t - \frac{x}{c_{L_2}}). \tag{8}$$

where the transmission coefficient $\mathcal{T}$ is defined as:

$$\mathcal{T} = \frac{2\rho_2 c_{L_2}/\rho_1 c_{L_1}}{\rho_2 c_{L_2}/\rho_1 c_{L_1} + 1} \tag{9}$$

We consider the propagation of a harmonic stress wave of frequency ω_0 and amplitude A_0 so that:

$$f(\tau) = A_0 e^{i\omega_0 \tau}. \tag{10}$$

Accordingly, the stress distribution in the $0 \leq x \leq x_0$ region is given by:

$$\begin{aligned} \sigma_x(x,t) &= A_0[e^{i\omega_0(t-\frac{x}{c_{L_1}})} + \mathcal{R}e^{i\omega_0(t+\frac{x}{c_{L_1}}-2\frac{x_0}{c_{L_1}})}] \\ &= A_0[e^{i(\omega_0 t - k_0 x)} + \mathcal{R}e^{i(\omega_0 t + k_0 x - 2k_0 x_0)}]. \end{aligned} \tag{11}$$

where $k_0 = \omega_0/c_{L_1}$ is the wavenumber. It is assumed that both spatial and time information are available for the stress distribution in the domain of interest. Under this assumption, one can evaluate the two-dimensional Fourier Transform (2D FT) of $\sigma_x(x,t)$:

$$\begin{aligned} \Sigma_x(k,\omega) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sigma_x(x,t)e^{-i(\omega t + kx)}dxdt, \\ &= \mathcal{F}_{2D}[\sigma_x(x,t)]. \end{aligned} \tag{12}$$

which for the considered propagating wave is given by:

$$\Sigma_x(k,\omega) = A_0\delta(\omega-\omega_0)[\delta(k-k_0) + \mathcal{R}e^{2ik_0x_0}\delta(k+k_0)]. \tag{13}$$

where δ denotes the delta function. The magnitude of the 2D FT expressed in eq. (13) features two peaks in the frequency/wavenumber domain at $\omega = \omega_0, k = \pm k_0$, corresponding to incident and reflected waves (Figure 2.a). The 2D FT effectively separates the two wave components and allows the application of simple window function to filter out one of the components. Upon filtering, the residual signal can be transformed back through an inverse FT to the space/time domain for visualization and further processing. This procedure is particularly useful for damage detection purposes, where reflected components carry information regarding the presence and the nature of damage. In many occasions, the reflections are small in amplitude and are often overshadowed by noise or by the incident wave. This makes the identification of damage, and its potential characterization difficult. From this perspective, 2D FT-based filtering represents an attractive approach to separate and highlight the presence of reflections in a given stress wave. Mathematically, the windowing process can be simply expressed as a function product between the wave's 2D FT and a 2D window function:

$$\Sigma_x^{(r)}(k,\omega) \approx [1 - H(k - k_0, \omega - \omega_0)]\Sigma_x(k,\omega) \tag{14}$$

where $H(k - k_0, \omega - \omega_0)$ denotes the windowing function centered at k_0, ω_0. The process is depicted graphically in Figure 2.b,c, which respectively show a Hanning window overlapped to the signal's 2D spectrum and the residual signal upon filtering. The space/time domain approximation of the reflected wave can be expressed as:

$$\sigma_x^{(r)}(t,x) = \mathcal{F}_{2D}^{-1}[\Sigma_x^{(r)}(k,\omega)] \tag{15}$$

where $\mathcal{F}_{2D}^{-1}$ denotes the inverse 2D FT.

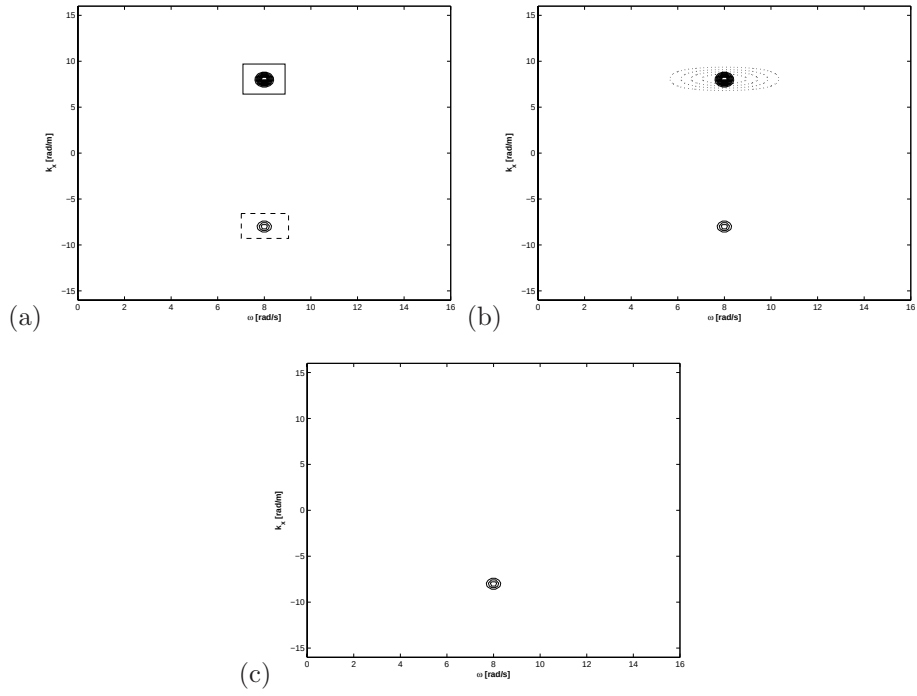


Fig. 2. Magnitude of 2D FT of the stress distribution in the $0 < x < x_0$ region: contour plot (Solid box highlights incident wave component, dashed box highlights reflected wave component) (a), contour plot with overlapped Hanning window applied for filtering (b), and residual signal spectrum upon filtering (c).

2.2. *Analysis of 2D Wave Propagation*

The previous concept can be extended to the case of waves propagating in a 2D domain. We first consider a signal composed of a plain harmonic wave propagating

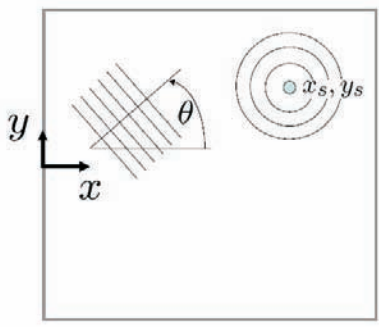


Fig. 3. Wave propagation in a 2D elastic medium: interaction between plane and spherical waves.

in the $x > 0$ direction, and of a secondary spherical wave generated at location x_s, y_s as depicted in Figure 3. The wave may be expressed as:

$$\begin{aligned} w(x,y,t) &= w_p(x,y,t) + w_s(x,y,t) \\ &= W_0[e^{-i(k_{x0}x+k_{y0}y-\omega_0 t)} + \mathcal{S}e^{i(k_0 r+\omega_0 t-2k_0 r_s)}] \end{aligned} \tag{16}$$

where W_0 is the amplitude, $\mathcal{S}$ is a generic scattering coefficient, and where $k_0 = \omega_0/c_0, k_{x0} = k_0 \cos\theta_0, k_{y0} = k_0 \sin\theta_0, r_s = \sqrt{x_s^2 + y_s^2}$.

The considered waveform can be represented in the frequency/wavenumber domain through the application of the 3D FT, defined as:

$$W(k_x,k_y,\omega) = \mathcal{F}_{3D}[w(x,y,t)]$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} w(x,y,t)e^{-i(\omega t+k_x x+k_y y)}dxdydt, \tag{17}$$

$$\tag{18}$$

The 3D FT of the wave signal defined in eq. (23) is given by:

$$W(k_x,k_y,\omega) = W_p(k_x,k_y,\omega) + W_s(k_x,k_y,\omega). \tag{19}$$

where

$$W_p(k_x,k_y,\omega) = W_0\delta(\omega-\omega_0)\delta(k_x+k_{x0})\delta(k_y+k_{y0}). \tag{20}$$

and

$$W_s(k_x,k_y,\omega) = W_0\mathcal{S}e^{-2ik_0 r_s}\delta(k-k_0)\delta(\omega-\omega_0),\ k_x \geq 0 \tag{21}$$

Figure 4.a shows the displacement distribution $w(x,y,t)$ over the considered domain at a specific instant of time, while Figure 4.b shows a cross section of the magnitude of the 3D FT evaluated at frequency ω_0. The plane wave component

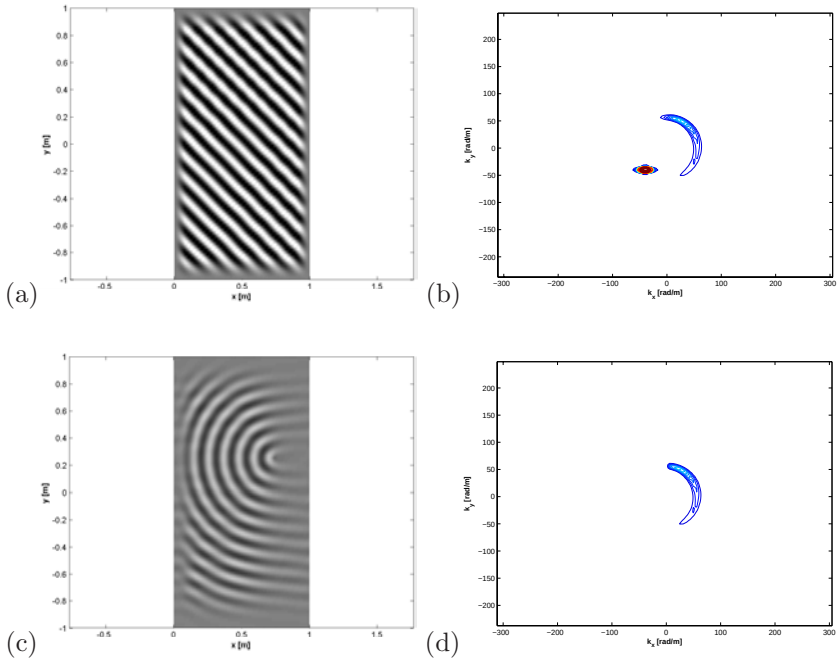


Fig. 4. Snapshot of interaction between plane and spherical wave $w(x, y, t_0)$ (a), and corresponding 3D FT at $\omega = \omega_0$ $|W(k_x, k_y, \omega_0)|$ (b). Residual spherical wave obtained upon filtering $w_s(x, y, t_0)$ (c) and corresponding 3D FT ($|W_s(k_x, k_y, \omega_0)|$) (d).

appears as a peak located at coordinates $k_x = -k_{x0},\ k_y = -k_{y0}$, while the spherical wave appears as a semicircular contour limited to the $k_x > 0$ half-plane. While the presence of the spherical wave as part of the considered displacement distribution is very evident in the wavenumber domain, it appears completely overshadowed by the plane wave component in the spatial displacement distribution of Figure 4.a. As discussed in the case of 1D wave propagation, the plane wave component can be conveniently removed through the simple application of a window function centered at the peak corresponding to the wavenumbers $k_x = -k_{x0}, k_y = -k_{y0}$, so that the 3D FT of the spherical wave can be approximated as:

$$W_s(k_x, k_y, \omega) \approx [1 - H(k_x - k_{x0}, k_y - k_{y0}, \omega)]W(k_x, k_y, \omega) \tag{22}$$

where $H(k_x - k_{x0}, k_y - k_{y0}, \omega)$ is the considered window function. The result of the filtering process is to isolate the spherical wave, which can then be visualized separately from the overall measured response. The spatial distribution of the filtered signal and the corresponding 3D FT at frequency ω_0 are shown in Figures 4.c,d. The analysis of the residual signal displayed in Figure 4.c allows the identification of location and potentially the characterization of the characteristics of the

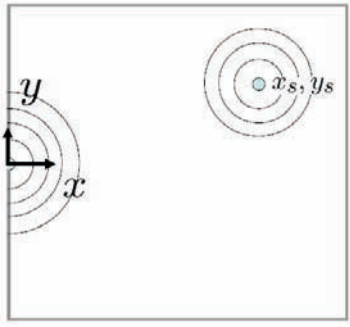


Fig. 5. Wave propagation in a 2D elastic medium: interaction between two spherical waves.

secondary wave source. Similar results can be illustrated by considering the interaction between two spherical waves according to the configuration depicted in Figure 5. The wave can be expressed as:

$$w(x,y,t) = w_{s1}(x,y,t) + w_{s2}(x,y,t), \tag{23}$$

$$= W_0[e^{-i(k_0 r - \omega_0 t)} + \mathcal{S}e^{i(k_0 r + \omega_0 t - 2k_0 r_s)}] \tag{24}$$

The corresponding 3D FT is given by:

$$W(k_x,k_y,\omega) = W_{s1}(k_x,k_y,\omega) + W_{s2}(k_x,k_y,\omega). \tag{25}$$

where

$$W_{s1}(k_x,k_y,\omega) = W_0\delta(|k| + k_0)\delta(\omega - \omega_0),\ k_x > 0 \tag{26}$$

$$W_{s2}(k_x,k_y,\omega) = W_0\mathcal{S}e^{-2ik_0 r_s}\delta(|k| - k_0)\delta(\omega - \omega_0). \tag{27}$$

Spatial distribution at a given instant of time, and 3D FT at $\omega = \omega_0$ are shown in Figure 6, for a case where $\mathcal{S} = 0.1$. Again, the presence of the secondary spherical wave is evident from the spectral representation of the signal, but hard to identify from the spatial distribution plot. The application of the window function to eliminate the primary wave can be here replaced by the elimination of the part of the 3D spectrum corresponding to $k_x > 0$. This portion of the spectrum in fact contains all information regarding waves propagating in the $x > 0$ direction, and therefore, given the configuration depicted in Figure 5, includes the primary wave. Removal of the negative part of the spectrum can be done through the evaluation of the Hilbert Transform (HT)[13] of the signal in terms of the x coordinate. This operation can be expressed mathematically as follows:

$$w^{\mathcal{H}_x}(x,y,t) = \mathcal{H}_x[w(x,y,t)]. \tag{28}$$

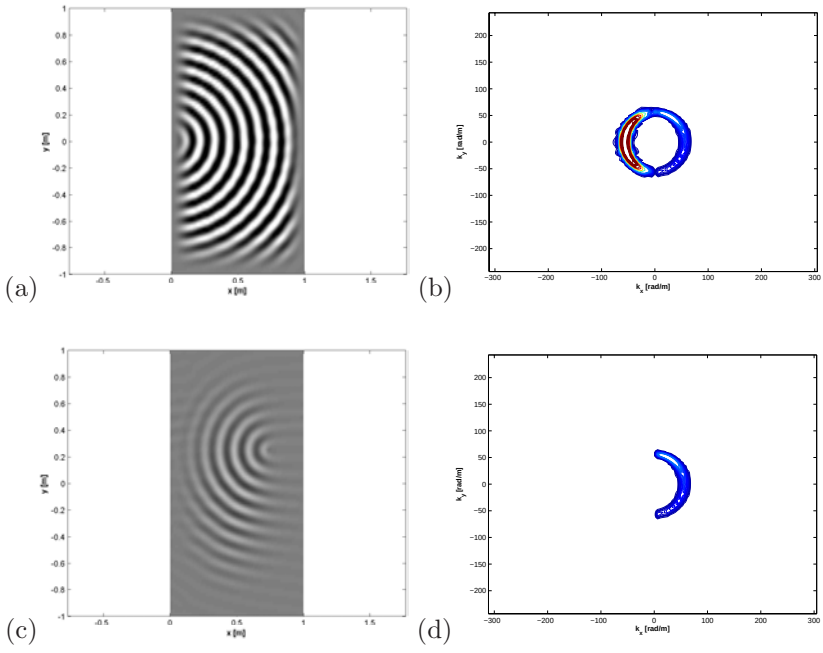


Fig. 6. Snapshot of interaction between two spherical waves $w(x, y, t_0)$ (a), and corresponding 3D FT at $\omega = \omega_0$ $|W(k_x, k_y, \omega_0)|$ (b). Residual spherical wave obtained upon filtering $w_s(x, y, t_0)$ (c) and corresponding 3D FT ($|W_s(k_x, k_y, \omega_0)|$) (d).

where $\mathcal{H}_x$ denotes the HT performed in terms of the x coordinate. The result of this filtering process is presented in Figure 6.c,d, which respectively show the spatial distribution at a given instant of time and the 3D FT spectrum evaluated at ω_0. The residual signal as in the previous case, clearly shows the presence of the secondary wave and the location of its origin. In a practical setting, the excitation signal generate to interrogate the state of health of the structure can be considered as a primary wave, while the secondary wave can be associated to any damage or in general to any discontinuity in the structure. As the primary wave is selected for optimal excitation of the desired wave modes, one can consider it as fully known. As in the cases considered above, the nature of the filtering procedure can be selected on the basis of the type of excitation signal used for inspection.

3. Numerical Examples

The concept described in the previous section is first tested on numerically simulated signals. The simple case of wave propagation in a damage rod is presented first, followed by a 2D elastic domain with simulated defects in the form of longitudinal cracks.

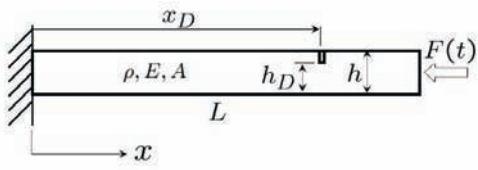


Fig. 7. Configuration for simulation of wave propagation in a damaged rod.

3.1. *Wave Propagation in a Damaged Rod*

The configuration considered in the first set of simulations is depicted in Figure 7. The rod, which has a length $L = 1$ m, thickness $h = 5$ cm, and it is made of aluminum ($E = 7.1 \times 10^{10}$ N/m^2, $\rho = 2700$ kg/m^3), is excited at its free end by a 5-cycle sinusoidal burst at 50 kHz. The rod is discretized using 160 bar elements, and its response is computed through numerical integration of the equation of motion using Newmark method. The simulation is performed over a time interval which corresponds to the time required for the injected pulse to reach the clamped end of the rod. The applied perturbation propagates and interacts with a damage located at $x_D = L/2$, modeled as a thickness reduction corresponding to $h_D/h = 0.9$ occurring over one element of the considered mesh. The response at a specific point ($x = 3/4L$) is plotted in Figure 8.a, which shows the response characterized by a main pulse, corresponding to the injected wave, and by a smaller secondary pulse corresponding to the reflection caused by damage. The 2D FT of the rod response is shown in Figure 8.b, which highlights the presence of the main pulse propagating along the $x > 0$ direction, and of the reflected pulse. This pulse has lower amplitude, it is characterized by the same frequency, and appears in the $k < 0$ region of the wavenumber/frequency domain. Again, the 2D representation effectively separates incident and reflected wave components, and allows effective filtering of the main wave. A 2D Hanning window is applied to isolate the reflected wave, and to obtain the frequency/wavenumber spectrum shown in Figure 8.d. The residual space-time domain signal can be then reconstructed through inverse FT transformation, as discussed in the previous section. The residual signal at $x = 3/4L$ is shown in Figure 8.c, to demonstrate how the windowing procedure is effective in removing the main propagating pulse, while leaving the reflected response corresponding to damage. This filtering procedure can be also applied to reduce the effects of noise in the data. Figure 9.a shows for example the rod response at the considered location when a random noise is added to the simulated data. The noise, which has an amplitude corresponding to 10% of the signal Root Mean Square (RMS) value, hides almost completely the reflected signal thus making any damage identification very difficult. A narrow band filter can be applied on the filtered signal to eliminate the effect of noise and highlight the reflected signal. A 2D window (Hanning) can be easily applied to the frequency/wavenumber spectrum of the residual signal

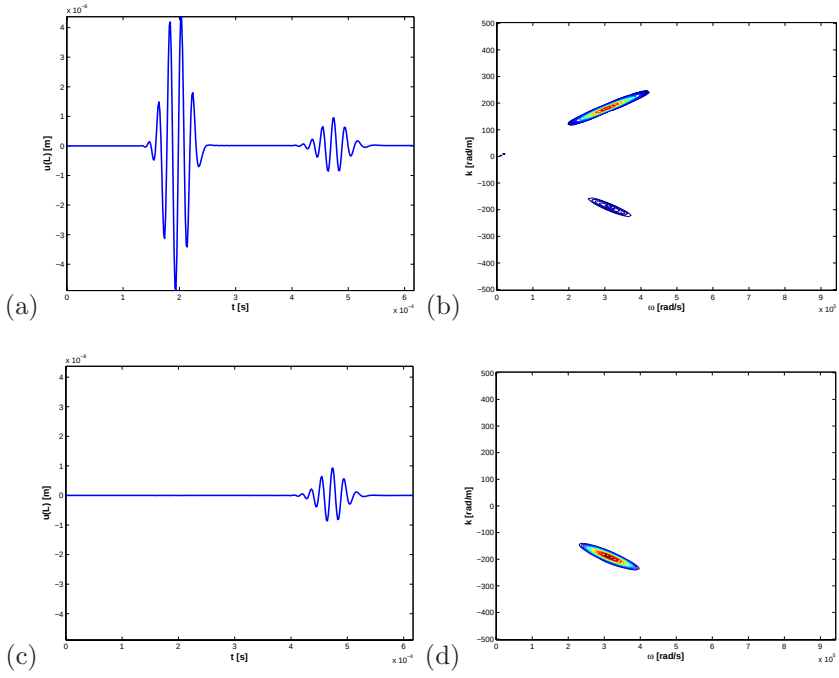


Fig. 8. Time response of damaged rod at $x = 3/4L$ (a), and corresponding 2D FT representation(b). Filtered time response (c), and 2D FT (d).

(Figure 8.d) in order to act as a narrow band filter and to mitigate the effects of background noise. The comparison of the estimated residual signal reconstructed with and without narrow band filtering in the frequency/wavenumber domain is presented in Figures 9.b,c, to demonstrate how how narrow band filtering mitigates the disturbances due to noise and allows visualizing the damage signature.

3.2. *Simulation of Wave Propagation in a Homogeneous Medium*

The propagation of elastic waves in a homogeneous medium is governed by the well known elastodynamic wave equation, which can be expressed as:[16]

$$\mu\nabla^2 u + (\lambda + \mu)\nabla\nabla u = \rho\ddot{u} \tag{29}$$

where μ, λ are the Lamé constants, and $u = [u \quad v \quad w]$ is the displacement vector. The mass spring lattice model (MSLM) presented in[15] is used to simulate the propagation of elastic waves in the considered homogeneous medium. MSLM utilizes the formalism of finite difference methods to discretize the medium into a lattice, for which elastodynamic equations are formulated. Discontinuities and

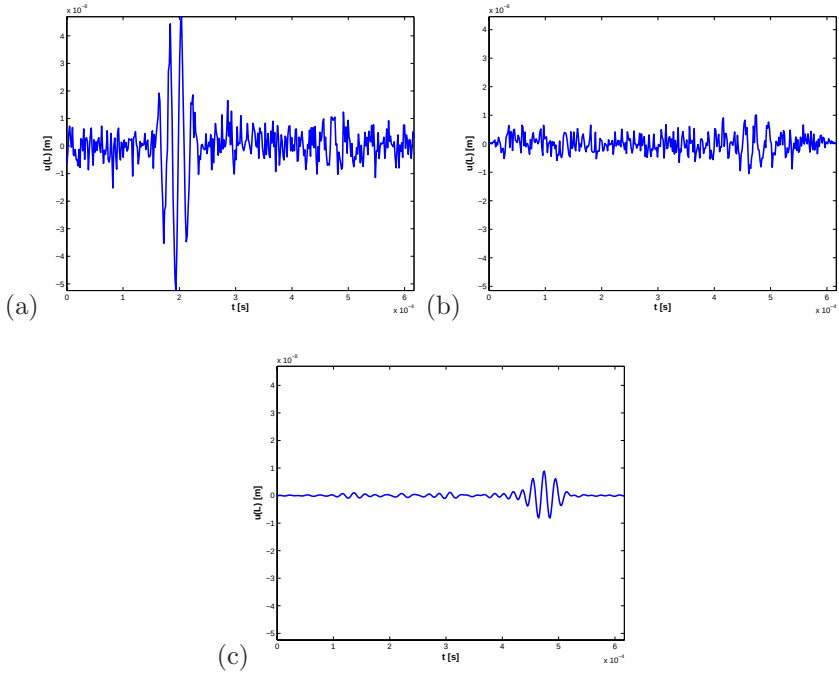


Fig. 9. Time response at $x = 3/4L$ with added noise (a), residual signal (b), and residual signal after narrow band filtering around reflected peak in frequency/wavenumber domain (c).

in general changes to the material properties of the medium are treated by simply modifying the properties of the lattice at the corresponding locations. The approach is very simple, easy to implement, and therefore very convenient for simulating wave propagation in a damaged specimen. Here, the approach is used for simulating ultrasonic waves in a thin two-dimensional elastic domain, with the assumption that all displacement components are constant through the thickness, i.e. $u = u(x, y), v = v(x, y), w = w(x, y)$. This allows the out-of-plane displacement component to be decoupled and evaluated separately from the in-plane behavior of the domain. The goal of the simulations is to test the frequency/wavenumber domain filtering process as applied to simulated 2D wave propagation data. The domain in Figure 10 has dimensions $L_x = 0.5$ m and $L_y = 0.25$ m and it is discretized using a 200×100 lattice. Damage is modeled as a 30% reduction in stiffness over the length of the simulated damage L_D. In the first configuration depicted in Figure 10.a, damage is located at $y_D = 0.125$ m, $x_D = 0.235$ m, and has a length $L_D = 2.5$ cm, while the configuration of Figure 10.b features two longitudinal damages respectively located at $y_{D_1} = 0.125$ m, $x_{D_1} = 0.160$ m, and at $y_{D_2} = 0.125$ m, $x_{D_2} = 0.375$ m of length $L_{D_1} = 1.5$ cm and $L_{D_2} = 2.5$ cm. The domain is excited by an imposed displacement applied at $x_E = 0.25$ m, $y_E = 0$, varying as

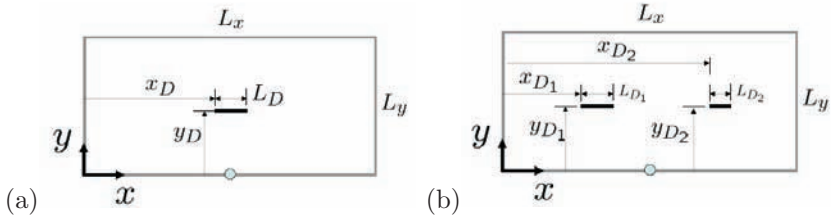


Fig. 10. Schematic of 2D elastic domain and simulated damage configurations.

a 5-cycle sinusoidal burst of frequency $f_0 = 100$ kHz. Results for the first damage configuration are shown in Figure 11. Specifically, Figure 11.a displays a snapshot of the propagating wave at a time instant corresponding to approximately $t_0 = 0.7T_s$, where T_s denotes the total simulation time, here selected as the time required for the induced wave to reach the top edge of the domain. The time snapshot of Figure 11.a shows the main pulse with small traces of scattering corresponding to damage. The cross section of the domain response evaluated at $f = f_0$ shown in Figure 11.b effectively decomposes the main wave propagating along $y > 0$ from the reflected wave traveling in the opposite direction. In this situation, the negative part of the spectrum can be easily removed to isolate the scattered wave corresponding to damage. This can be effectively done through the application of the Hilbert Transform along one of the spatial directions. Specifically, the reflected wave $w^{(r)}(x, y, t)$ can be estimated as:

$$w^{(r)}(x, y, t) \approx w^{\mathcal{H}_y}(x, y, t) = \mathcal{H}_y[w(x, y, t)]. \tag{30}$$

where $\mathcal{H}_y$ denotes the Hilbert Transform evaluated in terms of the y coordinate. The results of this process are shown in Figure 11.c,d which respectively present a snapshot at $t = t_0$ of the reconstructed time response upon filtering, and the corresponding cross section of the 3D FT at f_0. Given the approximation of the reflected wave $w^{(r)}(x, y, t)$, an effective visualization of presence and location of damage can be obtained through simple means such as the representation of the Root Mean Square (RMS) distribution of the reflected signal defined as:

$$w^{(r)}_{RMS}(x, y) = \Big[\int_0^{T_s} (w^{(r)}(x, y, t))^2 dt \Big]^{\frac{1}{2}} \tag{31}$$

The RMS distribution corresponding to the filtered response of the domain with a single damage is shown in Figure 11.e, while results for the domain with the two damage sites are shown in Figure 12, which indicates how the technique can effectively provide indications of the presence of multiple damages.

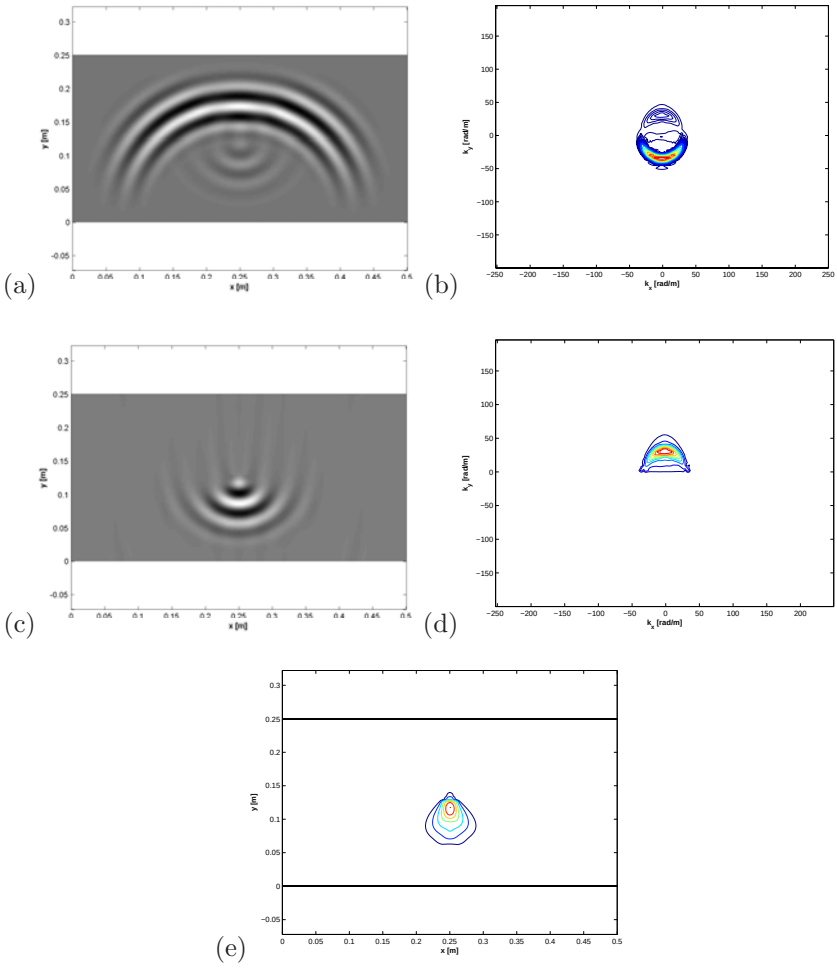


Fig. 11. Snapshot of propagating wave in 2D domain with single damage (a), corresponding 3D FT at the excitation frequency $f_0 = 100$ kHz, snapshot of filtered wave (c), corresponding 3D FT at $f_0 = 100$ kHz, and RMS distribution of reflected signal (d).

4. Experimental Results

4.1. *Set-Up*

The approach presented in the previous sections is validated experimentally on a simple plate structure. Piezoceramic discs are employed as actuators, and they are placed at locations which are selected so that elastic waves suitable for damage detection are generated.[1,4] The SLDV utilized (Polytec PI, Model PSV400M2) allows high frequency sampling (≈ 1 MHz), which enables ultrasonic waves detection

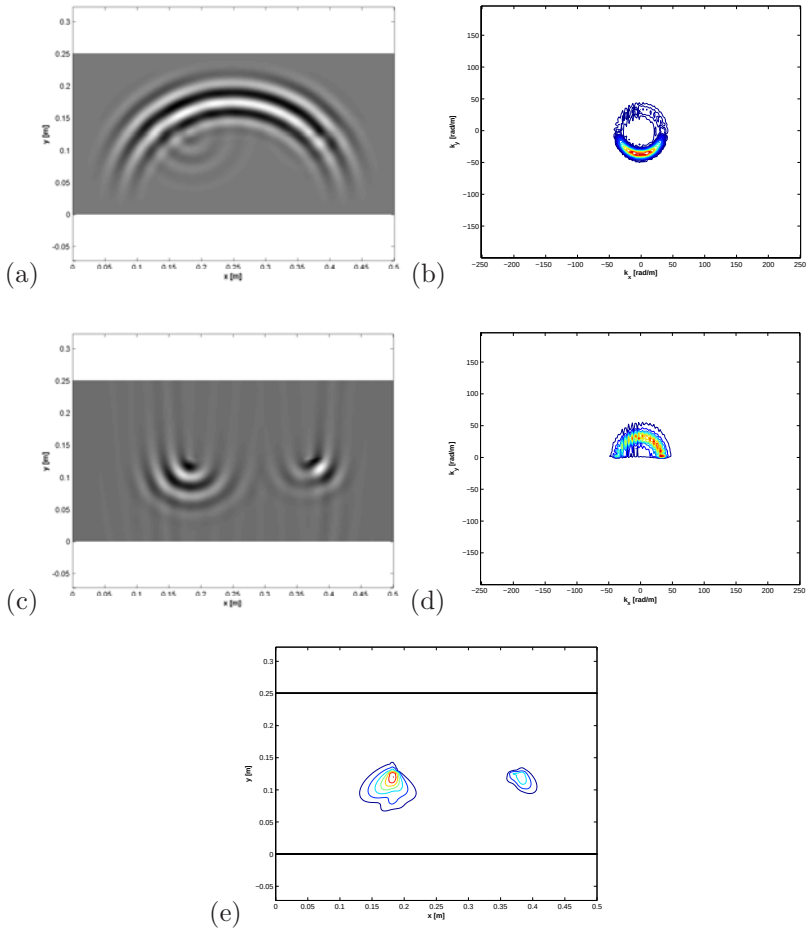


Fig. 12. Snapshot of propagating wave in 2D domain with two damages (a), corresponding 3D FT at the excitation frequency $f_0 = 100$ kHz, snapshot of filtered wave (c), corresponding 3D FT at $f_0 = 100$ kHz, and RMS distribution of reflected signal (d).

and visualization. In the wave propagation tests, the piezo discs are driven by a sinusoidal burst generated by a signal generator (Stanford Research Mod. SR360). The resulting elastic waves are recorded at the measurement grid points. The operation of the SLDV requires the generation of a pulse at each grid point in order to record the corresponding response. Phase information is retained by triggering the excitation signal through a low frequency signal (10 Hz), which also defines the scanning rate. Upon completion of measurements at all grid points, the recorded responses are then post-processed to obtain full images of the propagating wavefield within the region of inspection. A schematic of the set-up considered for wavefield measurement is shown in Figure 13.

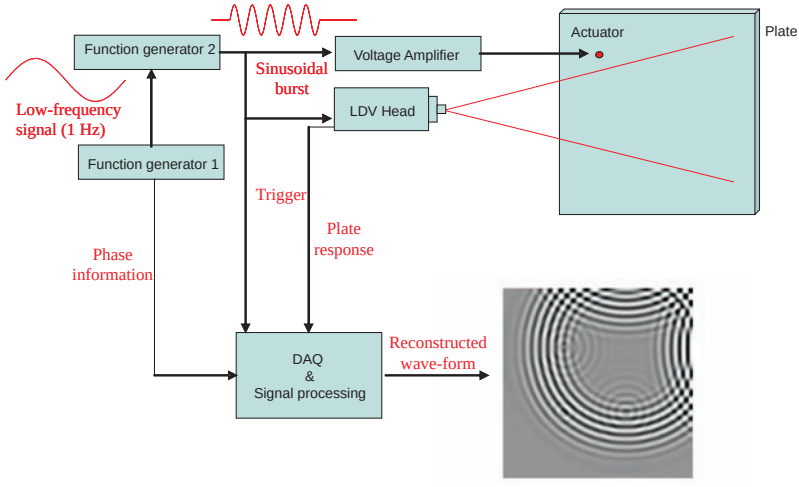


Fig. 13. Schematic of experimental set-up for wavefield measurement.

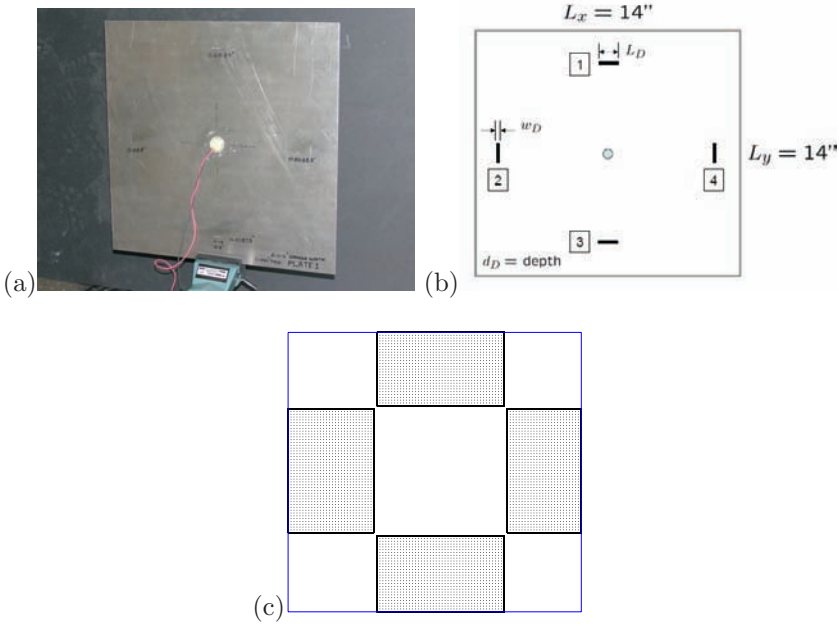


Fig. 14. Picture of plate with 4 damages and actuator (a), schematic of damage configuration (b), and detail of scanned region and measurement grid (c).

4.2. *Plate Experiments*

A first set of experiments is performed on a simple aluminum plate artificially damaged at 4 locations (Figure 14.a). The plate measures 14" × 14" and it is 0.05" thick. The considered damages are longitudinal slits of the same length and width, but varying thickness, so that the sensitivity of the proposed technique with respect to increasing levels of damage can be evaluated. A schematic of the damaged configuration is given in Figure 14.b, while Table 1 provides details of the dimensions of the 4 longitudinal slits. The plate is excited by a piezoceramic disc of 0.6" of diameter and 0.030" of thickness (Steminc mod. SMD28T07F300). The actuator is placed approximately in the middle of the plate, equidistant from the 4 damages. The actuator is simply bonded to the plate using a Loctite Quick Set epoxy. The disc is also fully encased in a layer of epoxy to provide it with an adequate backing to impart sufficient force to its base. The four regions of the plate surrounding the damages are scanned with the laser vibrometer, at the grid points shown in Figure 14.c. The excitation signal is a 5-cycle sinusoidal burst at $f_0 = 40$ kHz. This frequency was selected after a manual sweep was performed in the $20-70$ kHz range to evaluate the frequency of maximum response. The estimation of the optimal frequency of excitation based on a model of the piezoceramic disc bonded to the plate is currently under investigation and will support future developments of this study.

Table 1. Damage dimensions in plate experiment.

	L_D	w_D	d_D
1	0.5"	0.05"	0.03125"
2	0.5"	0.05"	0.025"
3	0.5"	0.05"	0.01875"
4	0.5"	0.05"	0.00625"

Figure 15 shows a cross section at $f = f_0$ of the 3D FT of the plate response in region 1. The plot shows a strong component propagating along the $x < 0$ direction and a reflected signal of same wavelengths propagating along $x > 0$. This reflection is associated to the presence of damage, as no boundary reflections are observed during the time interval considered for the response acquisition. A closer look at the spectrum of Figure 15 also shows the presence of a secondary wave component characterized by lower wavenumbers. As the identification of the wave modes excited by the piezo exciter considered in this configuration is underway, we here speculate that the main excited component corresponds to the first anti-symmetric Lamb wave mode A_0, while the faster secondary wave of lower amplitude is the symmetric mode S_0.[16] Figure 16 shows snapshots of the plate response evaluated at $t = 0.67T_s$, where T_s denotes the total acquisition time. The plot clearly shows the propagating wavefront and some evident reflections at the location

of the two most extensive damages (1 and 2). These reflections anticipate the main wavefront which suggests that they correspond to the interaction of the faster mode S_0 with damage. The presence of damages 3 and 4 cannot be easily observed from the displacement map of Figure 16.a. The time response is transformed in the

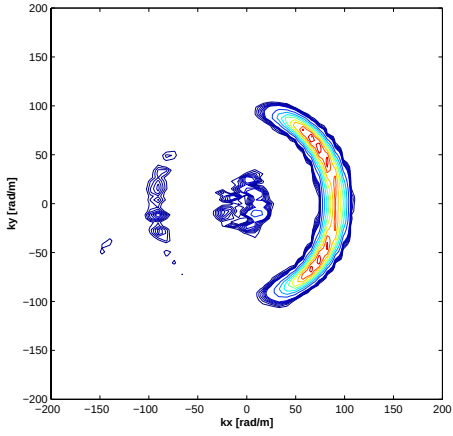


Fig. 15. Example of 3D FT for damaged plate

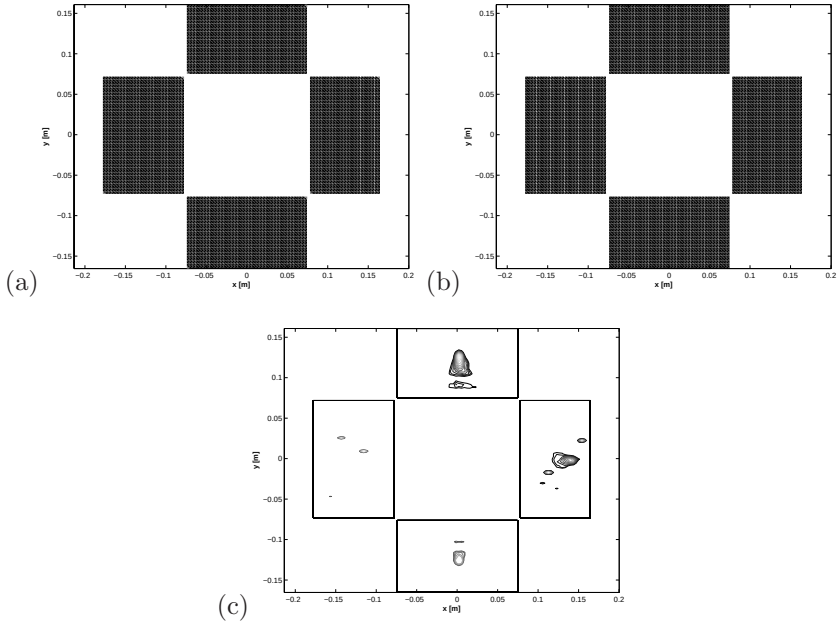


Fig. 16. Snapshots of plate response before (a) and after filtering (b). RMS plot of filtered response (c).

wavenumber frequency domain and is then filtered to isolate reflections according to the procedure outlined in the previous sections. A snapshot of the reconstructed time response at $t = 0.67T_s$ (Figure 16.b) clearly highlights the reflections caused by three of the four longitudinal cracks. The RMS map corresponding to the filtered time response (eq. 31) is shown in Figure 16.c, which clearly highlights the location of 3 of the 4 damage sites, and provides indications of their extent through the intensity of the corresponding contour lines. As in the case of Figure 16.b no indication regarding the smallest damage is obtained through this representation. This limitation in sensitivity can be the result of noise in the data, or more likely can be associated to the relatively low excitation frequency used for the tests. These issues will be addressed in future refinements of the technique.

5. Conclusions

This chapter presents a filtering technique aimed at improving the damage visualization capabilities of full wavefield measurements. The technique operates in the frequency/wavenumber domain, where the presence of reflections associated to damage can be easily observed and separated from the main signal generated for inspection. Upon filtering, the residual signal highlights the location of the source of reflections and may be used for the characterization of the source itself. The technique is applied as part of a Structural Health Monitoring technique based on guided wave inspection. The propagating waves are measured over a refined grid through a Scanning Laser Doppler Vibrometer which is able to provide the spatial and time resolution required to perform multidimensional Fourier Transforms. The application of the technique is first illustrated on a set of analytical and numerical examples and then it is successfully applied to analyze experimental data. The presented results demonstrate the potential of the technique and highlight future research needs. In particular, the technique needs to be supported by the optimal selection of frequency and mode of excitation, through methodologies that are currently under development. Moreover, the filtering process can be used to estimate the response of the undamaged component which can be used as a baseline for the formulation of damage indexes defining shape and size of the damage.

Acknowledgments

The author wishes the thank the work of Mr. V.K. Sharma of Millennium Dynamics Corporation for supplying the experimental data used in this work.

References

1. Staszewski, W.J., Boller, C., Tomlinson, G., *Health Monitoring Of Aerospace Structures. Smart Sensors And Signal Processing*, Wiley & Sons Ltd., 2004.

2. Rose, J.L. "A Baseline and Vision of Ultrasonic Guided Wave Inspection Potential" *Journal of Pressure Vessel Technology*, Vol. 124, pp. 273-282.
3. Michaels, T.E., Michaels, J.E., "Ultrasonic Signal Processing For Structural Health Monitoring" *Review of Progress in QNDE*, 23, eds. D.O. Thompson, D.E. Chimenti, American Institute of Physics, 2004.
4. Giurgiutiu, V., "Tuned Lamb Wave Excitation and Detection With Piezoelectric Wafer ACtive Sensor of Structural Health Monitoring" *Journal of Intelligent Material Systems and Structures*, Vol. 16, No. 4, April 2005, pp. 291-305.
5. Wooh S.C., Shi Y., "Synthetic Phase Tuning of Guided Waves", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 48, No. 1, January 2000, pp. 209-223.
6. Purekar A.S., Pines D.J., Sundararaman, S., Adams, D.E., "Directional Piezoelectric Phased Array Filters for Detecting Damage in Isotropic Plates", *Smart Materials and Structures*, 13(2004), pp. 838-850.
7. Ruzzene, M., Jeong, S.M., Michaels, T.E., Michaels, J.E., Mi, B., "Simulation and Measurement of Ultrasonic Waves in Elastic Plates Using Laser Vibrometry." *Review of Progress in QNDE*, 23, Eds. D.O. Thompson, D.E. Chimenti, Am. Inst. of Phys., 2004.
8. Sharma, V.K., Hanagud, S., Ruzzene, M., "Damage Index Estimation in Beams and Plates Using Laser Vibrometry", Proceeding of the *2005 International Workshop on Structural Health Monitoring (IWSHM)*, Stanford University, Stanford CA, September 2005.
9. Basri, R., Chiu, W.K., "Numerical analysis on the interaction of guided Lamb waves with a local elastic stiffness reduction in quasi-isotropic composite plate structures", *Composite Structures*, 66 (2004), pp. 8799.
10. Alleyne, D., Cawley, P., "A Two-dimensional Fourier Transform Method ofr the Measurement of Propagating Multimode Signals", *Journal of the Acoustical Society of America*, Vol. 89, March 1991, pp. 1159-68.
11. Muthupillai, R., Lomas, D.J., Rossman, P.J., Greenleaf, J.F., Manduca, A., Ehman, R.L., "Magnetic Resonance Elastography by Direct Visualization of Acoustic Strain Waves," *Science*, Vol. 269, 1854-1857 (1995).
12. Muthupillai, R., Rossman, P.J., Lomas, D.J., Greenleaf, J.F., Reiderer, S.J., Ehman, R.L., "Magnetic Resonance Imaging by of Transverse Acoustic Strain Waves," *MRM*, Vol. 36, 266-274 (1996).
13. Feldman, M., "Nonlinear system vibration analysis using Hilbert transform: I Free vibration analysis method FREEVIB", *Mechanical Systems and Signal Processing*, **8**, pp. 119-127.
14. Graff, K.F., *Wave Motion In Elastic Solids*, Dover, NY, 1991.
15. Yim, H., Sohn, Y. "Numerical Simulation and Visualization of Elastic Waves Using Mass-Spring Lattice Model", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, Vol. 47, No. 3, May 2000, pp. 549-558.
16. Graff, K.F., *Wave Motion In Elastic Solids*, Dover, NY, 1991.

This page intentionally left blank

CHAPTER 26

RECENT ADVANCED SIGNAL PROCESSING METHODS FOR WAVEFORM-BASED NDE METHODS

Don J. Roth[1], John C. Aldrin[2], Jeff P. Seebo[3], Long B. Trinh[3], James L. Walker[4], and William P. Winfree[5]

[1]*NASA Glenn Research Center, Cleveland, OH 44135*
[2]*Computational Tools, Inc., Gurnee, IL 60031*
[3]*Lockheed Martin Space Systems Company, 13800 Old Gentilly Road New Orleans, LA 70129–2218*
[4]*NASA Marshall Space Flight Center, Huntsville, AL 35812*
[5]*NASA Langley Research Center, Hampton, VA 23681*
Donald.j.roth@nasa.gov

The critical NASA problem of foam release from the foam-based thermal protection system of the Space Shuttle external tank during ascent has spurred the advanced development and tailoring of NDE methods to inspect the foam for flaws that might facilitate release. NDE methods under consideration include terahertz, microwave, shearography and xray backscatter imaging. In parallel with the development of waveform-based methods such as terahertz, a sophisticated broadband signal and image processing software package has been developed. One of the unique aspects of the software is the inclusion of a merit assessment algorithm to "grade" various signal processing methods by providing a quantitative measure well correlated to the ability to subjectively distinguish flaw data from noise or background data. This chapter describes the software and provides a case history for its focus and use to process and analyze terahertz imaging results on a foam sample standard containing seeded voids.

1. Introduction

Figure 1 shows a portion of the Space Shuttle external fuel tank (ET) and the coverage of it by the thermal protection system (TPS) sprayed-on foam insulation (SOFI). The foam is necessary to provide insulation for the cryogenic propellants in the ET while the shuttle sits awaiting liftoff with cryogenic propellants in the ET. Four specially engineered low-density, closed cell foams are used on the

tank. Flaws present in the foam can result in initiation sites for foam loss. Manually-sprayed foam areas are especially susceptible to flaws. A major effort is underway at NASA to find and improve inspection methods to locate cracks, voids, delaminations (between foam and underlying tank, or between foam and knitlines), and crushed foam areas of the SOFI to minimize foam debris release during ascent. The latest external tank for the STS–121 July 2006 Shuttle launch had the protuberance air load ramp windshield removed to minimize large foam debris release upon launch. As a result, the highest priority areas for inspection are now the ice frost ramps (IFRs) (Fig. 2) which prevent ice from forming on underlying aluminum brackets used to fasten fuel-pressurization lines and a tray of electrical cables to the tank's exterior. It is impossible to completely prevent foam from coming off of the IFRs during launch but inspection methods can help locate flaws in the ice frost ramps prior to launch and also aid in the understanding of conditions and flaws causing foam release. Recent studies have indicated that thermal cycling due to fueling and de-fueling especially appears to be a significant flaw creator. The largest chunk of foam debris ever detected coming off an ice-frost ramp was 0.121 lbs.

As seen in Fig. 2, the front (top) and sides of the IFR are sloped, while the back (bottom) is almost blunt. Pressurization lines pass through holes in the middle of the IFR. The cutout off to the right side of the tall part of the IFR is the location of the cable tray. With the cable tray in place, foam subsections cannot be seen.

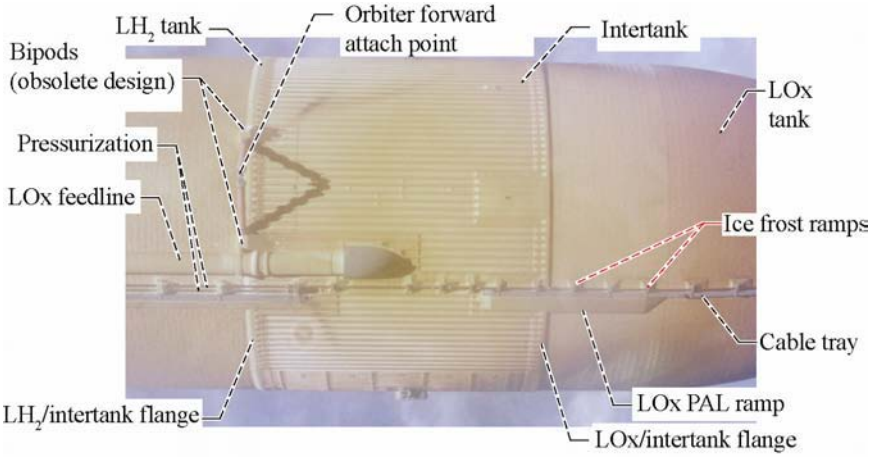


Figure 1. A Portion of the Space Shuttle External Tank. Foam covers the entire tank. PAL ramp has been removed as of shuttle flight STS–121. Ice Frost Ramps are considered to be the riskiest areas in terms of foam loss once the PAL ramp was removed.

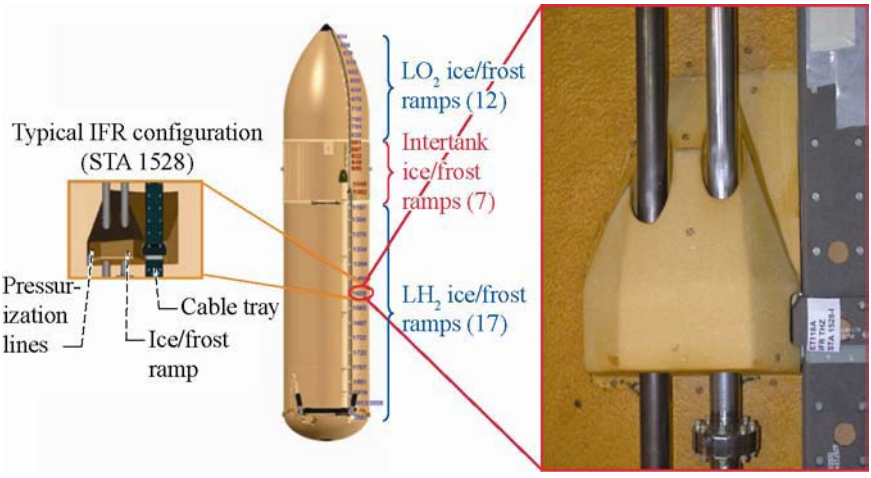


Figure 2. Ice Frost Ramp configuration and location on Shuttle external tank.

2. Terahertz Nondestructive Evaluation

Foam inspection methods being applied and researched for improvements at NASA include terahertz, microwave, shearography, and x-ray backscatter imaging. Terahertz inspection has shown significant promise for detection of voids and crushed foam. Background information on pulsed terahertz measurements can be obtained from several references.[1–3] Briefly, terahertz waves are electromagnetic waves with wavelengths on the order of 200–1000 μm. Metallic materials reflect terahertz waves while non-polar liquids, dielectric solids, and gases are transparent to terahertz energy. Reflections also occur at interfaces between materials with dissimilar dielectric properties (difference in indices of refraction). Thus, small reflections occur at voids in foam and the signal transmitted through the void has reduced signal amplitude. Continuous wave (narrowband) and pulsed (broadband) terahertz systems exist. This study is based on results from a reflection-mode broadband terahertz method with a collinear source-detector configuration shown schematically in figure 3. Samples investigated for study at NASA generally involve the use of a SOFI foam-covered metallic article. For this configuration, the typical terahertz waveform is characterized by a large echo reflecting off of the metal substrate followed by ringing.

Significant time- and frequency-domain information is available in these signals. Currently applied signal processing techniques for terahertz data are generally based on simple parameter extractions such as peak detection of

the time-domain signal or the Fourier-transformed signal. The focus of this research effort is to develop advanced signal analysis techniques that capitalize on all the information that is available in the terahertz signal to improve the resolution of flaws and differentiation of flaws from normal SOFI variations such as knit lines. Particularly, this chapter focuses on the use of the centroid of the power spectral density of the substrate reflection and how it compares to simple parameter extraction methods such as peak-to-peak amplitude (time-domain) and peak magnitude (frequency-domain). Both subjective and quantitative merit assessment protocol are used in the comparisons.

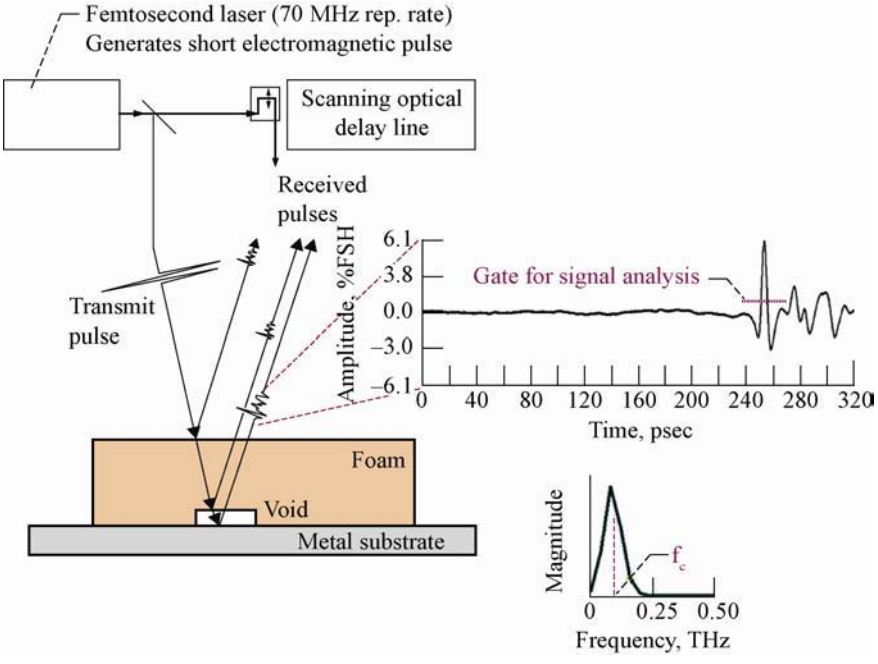


Figure 3. Basic schematic diagram of reflection-mode terahertz methodology.[4] Reflections will be received off of the various interfaces. Reflection from metal will be the strongest. The horizontal dotted line over the large, initial portion of the wave shows the time gate of the echo typically used during signal processing. A sample power spectral density is shown with centroid (f_c) denoted.

3. Software Development

Complementary commercial-grade terahertz wave processor software packages (Fig. 4) were developed at NASA Glenn (GRC) and Langley (LaRC) Research centers so that software algorithms could be developed and tested at the various

NASA facilities with researchers and operators using terahertz NDE. The software can also be used to analyze pulsed broadband ultrasonic, microwave and radar signal data sets if users can store the data in the specified custom format. (The authors can be contacted for further information by readers interested in use of this software.) Previously-applied signal processing techniques for terahertz data were generally based on simple parameter extractions such as the peak of the time-domain signal received off of a metal substrate located beneath foam. Using band-pass filtering methods where an image is formed from the highest practical frequency is probably the most significant and accepted improvement since the work began. However, the software team has gone beyond this method to also look at many other image types such as total energy, moment analysis, cross-correlation with respect to a wave reference, first harmonic nonlinearity parameter, power loss, centroid of the power spectral density, and biomedically-based routines. The software packages allow for extremely flexible gating routines where only a portion of the waveform is analyzed to create images. Gate options include static gate, dynamic gate (which can be set to track a feature of the waveform such as positive peak amplitude at each scan location to account for inexact experimental setups such as nonuniform distance between sample and terahertz head), and multiple gates to develop image time series over the course of the wavetrain. Additional time was spent developing standard and custom image-processing functions for the software packages that allow b-scan display, three-dimensional displays, custom color tabling, smoothing, edge detection, image and waveform-based deconvolution, feature size measurement, line profiles, region-of-interest selection and statistics, image math, principal component analysis, and image averages as applied to time series and frequency series of images. The user also has the ability save all setup information to a configuration file for recording exact signal and image processing setup parameters, and also has the ability to save images to various formats including ASCII text for further post-processing.

One particular unique capability of the software was focused on developing a quantitative assessment protocol to evaluate the various potential signal processing methods. Automated feature extraction methods were developed to replicate the important visual classification characteristics used by inspectors and a new merit assessment approach was developed to identify the most promising combination of signal processing algorithms. This work lays the framework for automating the characterization and quality of images obtained using the various signal processing parameters. This capability is described in a later section of this chapter.

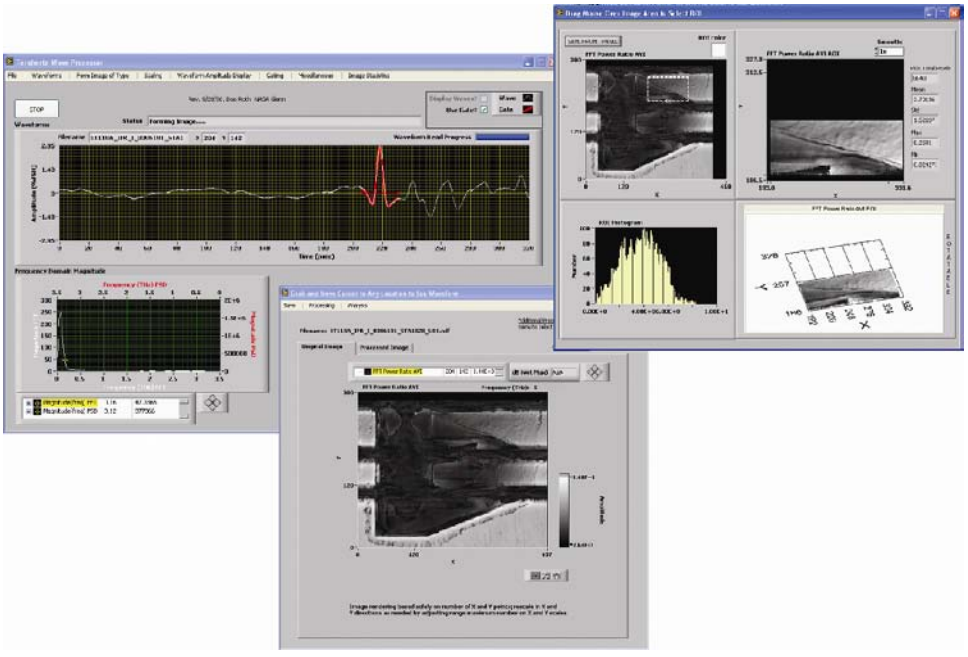


Figure 4. Commercial grade GUI screenshots for Terahertz Wave Processor.

4. Signal Processing

This chapter primarily focuses on comparison of one alternative signal processing method with those conventionally extracted from the waveform. Specifically, images formed from the centroid frequency (f_c) of the power spectral density PSD(f) are compared against those formed from time-domain peak-to-peak amplitude, peak magnitude of the PSD(f), and magnitude at highest practical frequency of the PSD(f). The PSD(f) describes how the power of a signal or time series is distributed with frequency. f_c is defined as the frequency (f) demarcating the point of balance in the power spectral density and is determined from:

$$f_c = \frac{\int_{f_{low}}^{f_{high}} PSD(f)\,fdf}{\int_{f_{low}}^{f_{high}} PSD(f)\,df} \tag{1}$$

A dynamic gate of short duration (25 sec) was used to track the peak of the large reflection off of the metallic substrate. This gate is characterized by a fixed time width with variable start time that tracks the peak of the signal up to a certain start time tolerance to allow for variations in geometry. The gate centers itself over the peak of the echo and the start time is determined by subtracting one-half the gate length from the peak center time. For such a gate, a typical PSD(f) and subsequently-determined f_c might look like that shown at the lower right corner of Fig. 3. f_c is reasonably simple computationally and has been shown to take about 60% longer to compute than time-domain peak-to-peak amplitude. For example, for a scan having nearly 30000 data points, post-processing time to compute the entire image might take 5 sec for f_c versus 3 sec for peak-to-peak amplitude.

Prior to image comparisons, images were contrast enhanced to the greatest extent possible by effectively removing outlier values that caused a reduction in image contrast. Image comparisons were performed by subjective evaluations with the aid of line profile and region-of-interest zoom examination for the hardest to detect flaws, and by quantitative evaluations using the merit assessment protocol.

5. Materials

Table 1 describes the SOFI specimens examined. These include 1) a 3-layer square SOFI foam sample containing 36 seeded spherical voids (produced using milling procedure) of various sizes and depths located in the bottom layer and 2) mock-up and actual ice frost ramps. The latter are shown to provide the reader with images of actual components or simulated components. Other foam samples with defects were also inspected and characterized over the course of NASA investigations.[5]

Figure 5 presents a diagram of the manufactured void foam panel used for the study. The cylindrical voids varied from 3.2–19.1 mm in diameter and 3.1–19.1 mm in depth. The end condition of the holes was milled with a spherical profile to avoid an exaggerated NDE response seen for holes with a pure cylindrical shape. Note, for some holes with depth greater than the diameter, a portion of the hole length will still have a sharp cylindrical transition from the pristine foam.

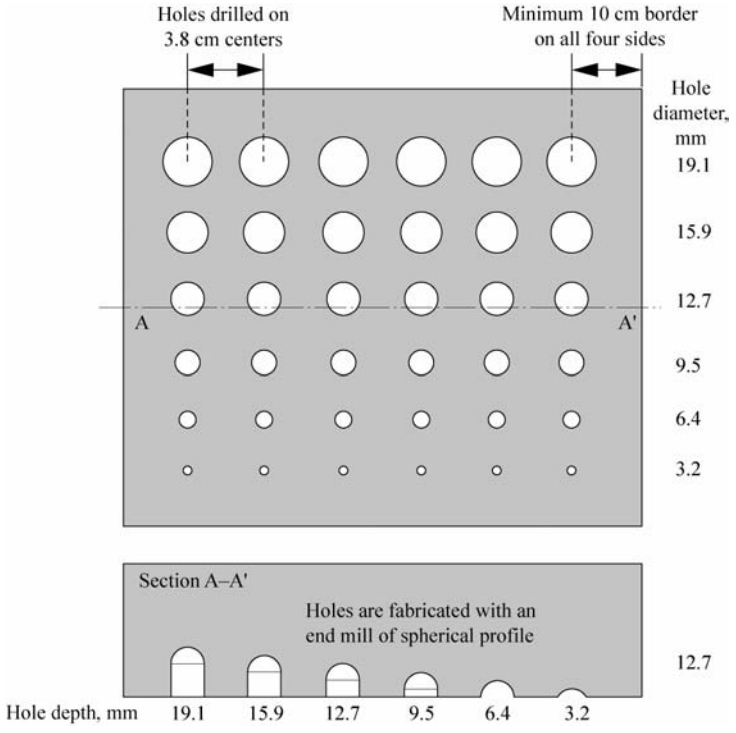


Figure 5. Diagram of 36 hole foam test sample with varying hole diameter and depth.

Table 1. Sample and scan parameters.

Void Sample Material	BX–265
Dimensions (length × width × thickness) (cm)	35 × 30 × 5 (per layer)
Spherical void sizes and depths (cm)	0.31–1.9
Scan Dimensions (X by Y) (cm)	43.4 × 42.9
Scan, Step Increment (X, Y) (cm)	0.254, 0.254
Further Description	3 stacked layers, spherical voids in bottom layer. THz focused on substrate.
Ice Frost Ramp (Mock-up and Actual) Material	PDL foam over 1.27 cm thick NCFI foam sub-layer
Dimensions (length × width × thickness) (cm)	52 × 37 × up to 24
Natural flaws of various size and depth	Observed during mock-up dissection
Scan Dimensions (X by Y) (cm)	61.9 × 43.2
Scan, Step Increment (X, Y) (cm)	0.152, 0.152
Further Description	IFR area consisting of two separate foam pours without (mock-up) and with (actual) pressurization lines running through openings in the interface

6. Terahertz Experimental Parameters

Table 2 describes the broadband 1 THz system and associated experimental parameters.

Table 2. Experimental parameters for terahertz system and signal processing.

Perpendicular Distance and Angle between Send Head and Sample (cm)	29.85
Typical Received Frequencies (THz)	≤1.0
Data Acquisition Rate (1/sec)	6.4×10^{12}
Signal Acquisition: Width of dynamic (peak-centered) gate off echo of metal substrate (psec)	25
Waveform Length Acquired (psec/points)	320 psec/2048 points
Waveform Acquisition Rate (Hz)	100
Collinear source-detector	Yes
Spatial Resolution (at FWHM Point Spread Function) (cm)	0.5

7. Results

7.1. *Sample 1 – Seeded Voids*

Figure 6 shows a direct comparison of images formed from time-domain peak-to-peak amplitude (TDPTP), power spectral density peak magnitude (PSD(f)PM) (occurring at approximately 0.08 THz), power spectral density magnitude at the highest practical frequency (0.16 THz) (PSD(0.16 THz)) and centroid frequency of the power spectral density (f_c) image. The voids are relatively deeper (farther from the front surface) as one moves to the right in the image. The first four rows (referenced from the top) of voids in all images are detected clearly. TDPDT and PSD(f)PM do not reveal unambiguously any of the voids of the bottom two rows. To aid in determining detectibility for the bottom two (fifth and sixth) rows of voids (with the sixth row location indicated by the white arrow on the images), image line profiles are used. The line profiles show normalized contrast for that particular image line and will show a peak-valley characteristic where the voids are located if clearly detected. It is seen from the line profiles in Fig. 6 that only the f_c image clearly resolves the smallest and deepest (rightmost) void as indicated by the dotted circle. PSD(0.16 THz) is able to resolve all but the rightmost void in the bottom row. It is also useful to view two-dimensional (2D) and three-dimensional (3D) zoomed-in regions-of-interest (ROI) to characterize the detectibility differences between the PSD(0.16 MHz) and f_c images (Fig. 7). In this case, not much difference is seen in terms of sensitivity except that the f_c image appears to show less background noise. The use of f_c appears to show

promise versus image formation from more conventional parameters of the wave signal for image formation. Similar results have been seen with other foam samples having seeded flaws.[6]

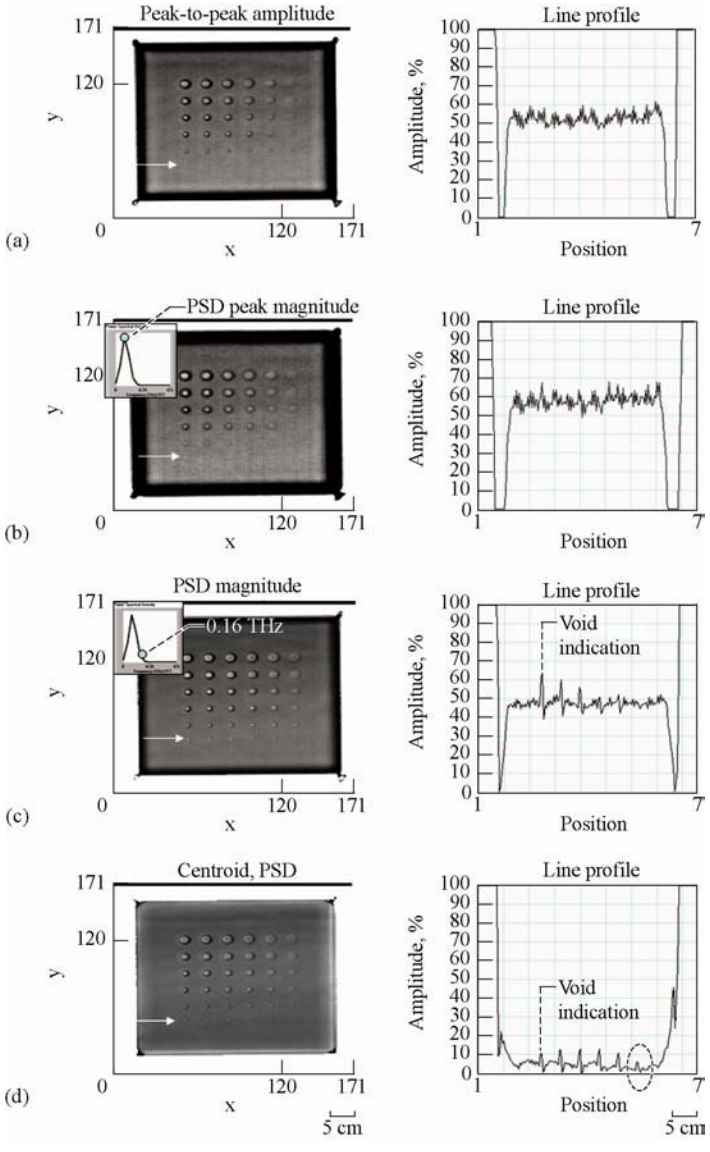


Figure 6. A comparison of different terahertz images including line profiles at the location of the bottom (smallest) row of voids. The voids are relatively deeper (farther from the surface) as one moves to the right in the image. (a) Time-domain peak-to-peak amplitude. (b) Power spectral density peak magnitude. (c) Power spectral density at 0.16 MHz. (d) Centroid (f_c) of the power spectral density.

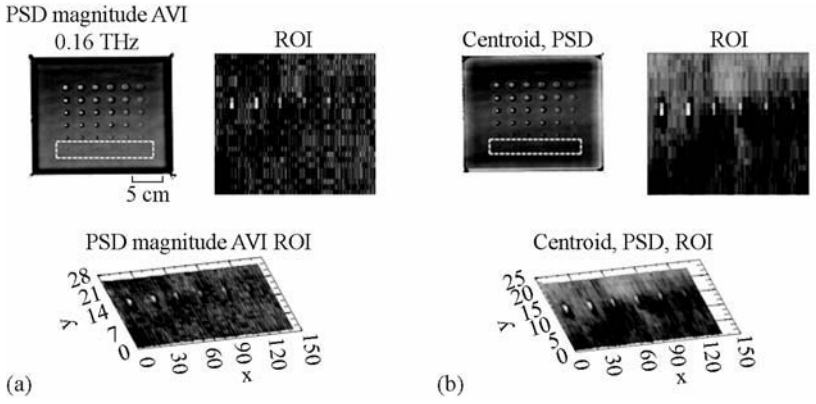


Figure 7. Two-dimensional and three-dimensional zoomed-in image views of the bottom row of voids as indicated by the dashed box for PSD(0.16 THz) (a) and f_c (b) images. ROI = region of interest.

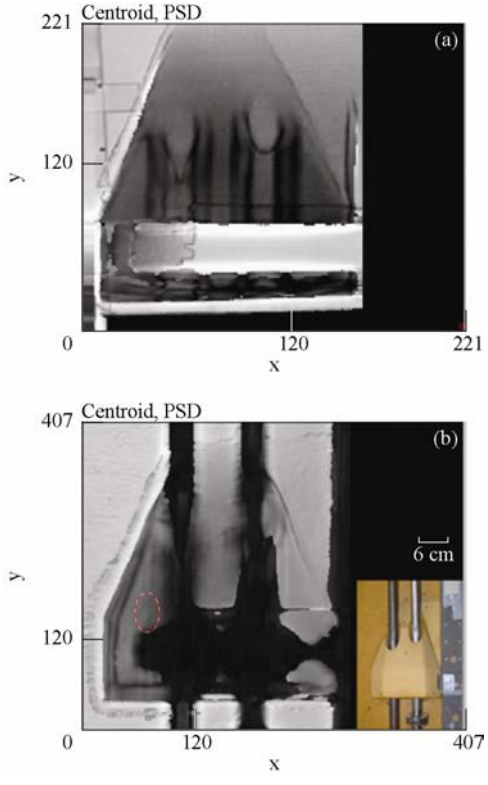


Figure 8. Terahertz f_c images of ice frost ramps. (a) Mock-up ramp with no pressurization lines. (b) Actual ramp. Inset on (b) shows photograph of ice frost ramp. Indications noted within dashed circle on (b).

7.2. *Ice Frost Ramp*

Because of the findings on the seeded void samples, the f_c signal analysis method was applied to obtain images of mock-up and actual ice frost ramps (Fig. 8). The part of the actual ramp that does not return much of a signal in the middle of the thickest foam section is the location of the metal bracket that holds the pressurization lines. Terahertz reflections received here are outside the time-delay window and account for the very dark areas in the image. Indications are noted on the images but in general, these would be quite difficult to draw conclusions about until dissection of the mock-up occurs.

8. Protocol for Quantitative Evaluation of THz NDE Data

A general protocol to perform a quantitative assessment for evaluating the merit of various signal processing algorithms is presented and applied for the void sample. The steps are listed here.

(1) Define the sample data and flaw conditions to test. Definition of the test samples must consider the flaw type and range of expected characteristics. For SOFI foam applications, voids, delaminations, cracks, knit-lines and compressed foam are possible defects of interest. A range of expected sample conditions found in the field must also be replicated in the study providing representative structures, density variations and porosity.
(2) The image data is acquired for the representative sample set using the selected method according to the standard acquisition procedure, in this case terahertz. The acquired data must be properly labeled and stored for post-processing and analysis.
(3) Define the signal processing algorithms or combination of algorithms to evaluate. Both select algorithms and key parameters of the signal can be considered in the designed experiment.
(4) Define the important visual classification characteristics for each defect/noise feature type (such as size, shape, contrast with background, edge characteristics) of interest. Also, local image analysis feature extraction methods must be developed to replicate visual classification of indications by an inspector.
(5) Define the quantitative statistical methodology for the assessment of merit for each of the proposed signal processing algorithms or combination of algorithms for a select set of conditions of interest (for example embedded voids of different sizes). The objective of merit evaluation is to provide a

quantitative measure well correlated to the ability to subjectively distinguish the void data from the noise or background data.

(6) Validation of quantitative analysis procedure with visual subjective interpretation results in a designed experiment is needed prior to implementation.

The next paragraphs will present details on the quantitative analysis protocol including feature extraction and signal processing algorithm merit assessment. For the terahertz image data shown in Fig. 6, several visual characteristics were identified with association to the seeded voids. First, there are regions of adjacent pixels with similar magnitude present that differ from the mean background level. Both plateau and ring features are observed associated with the void center and edge regions respectively. Quantifying these plateau and ring features through statistical measures such as maximum, minimum, mean and standard deviation with respect to background levels is proposed. Also, connected edge features are present that may also be quantified through slope and advanced edge image measures. Ideally, a good measure will correlate well with all of the key visual indications used by inspectors. Note, although ring features are present in these data, they may not be particularly reliable for detecting natural voids. Thus, separate analyses for each feature type were investigated.

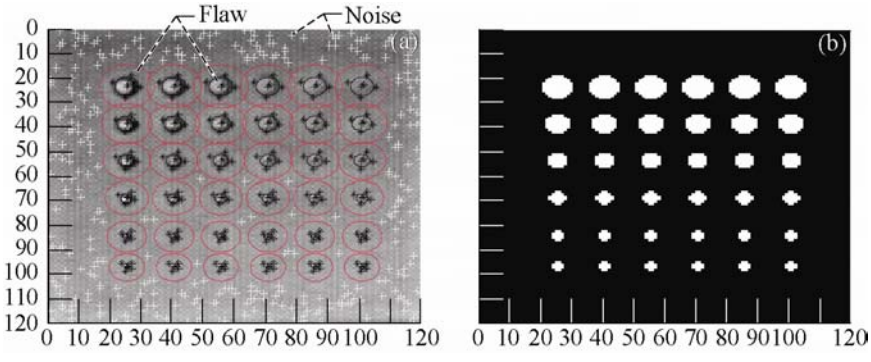


Figure 9. (a) A point map associated with flaw and noise data for the Centroid (f_c) of the power spectral density and (b) a region map associated with flaw data after filtering.

To acquire a statistical distribution of data on void and background regions for quantitative evaluation, two local feature extraction methods are presented. The first method is a line-based method that uses selected data points in known areas of interest and performs a search for the most significant response (e.g., largest grey value, greatest change/slope in pixel grey scale, etc.) in four directions (vertical, horizontal, two diagonals). Figure 9(a) displays an example

image including select points for void regions and randomly selected points for unflawed regions. Data vectors are constructed from measurements at a series of points in each direction through these center points. Five distinct measures were then performed to replicate the discriminant capability of an inspector: (A1) maximum signal along a line, (A2) minimum signal along a line, (A3) maximum peak to peak signal response along a line, (A4) maximum slope change along line (using a 5 point fit), and (A5) variation of signals along line.

The second method for obtaining a statistical distribution of data instead uses region-based evaluation in 2D of both void and background (noise) in regions. Figure 9(b) displays an example image filter defined for each of the known void regions. Background noise regions, not shown in Fig. 9(b), were defined as a series of circular regions randomly placed across the non-flaw region. Three purely statistical measures sensitive only to the extreme points in the regions: (B1) max in region, (B2) min in region, (B3) peak to peak in region were used. In addition, a fourth measure was developed to quantify the significance of the edge features in the region. The proposed edge measure (B4) incorporates a Canny edge detection algorithm, evaluates the magnitude at each edge pixel and sums the total edge pixel grey level response as a quantitative region measure. An example of the quantitative edge measure applied to the centroid of the PSD data is shown in Fig. 10. Alternative statistical measures comparing the distribution of points within a flaw region with respect to the average distribution for the entire background noise level were also investigated. Although both parametric (F-test, t-test) and non-parametric (Kolmogorov–Smirnov test, Wilcoxon test) methods were investigated, none of the hypothesis test measures performed well in initial studies. The line-based analysis through select points provides more data points

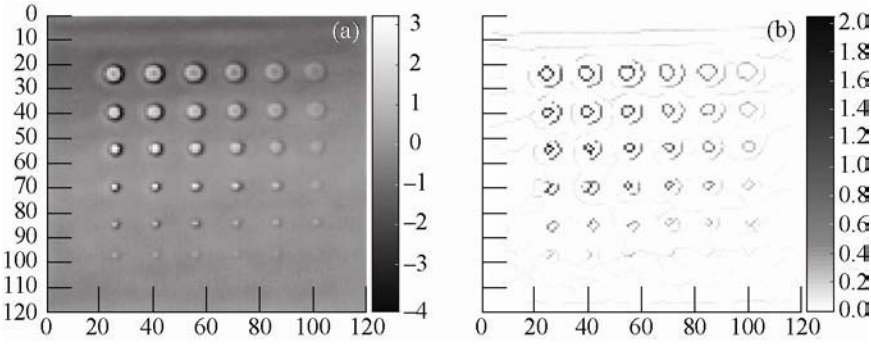


Figure 10. (a) Example THz image using the Centroid (f_c) of the power spectral density for 36 hole panel, and (b) the corresponding THz image with the application of the quantitative edge analysis.

for statistical evaluation, while region based analysis provides area measures that likely correlate well with the visual queues used by inspectors. Prior studies with cylindrical flat-bottom holes have shown the best results with the maximum slope measure (A4) for line-based analyses and quantitative edge quality measure (B4) for regions.[6]

The next step is to evaluate the merit of local measures for a feature of interest to best discriminate between the flaw and no flaw conditions. For this case study example, the merit evaluation step uses the feature extraction measure results for a selected number of void regions (or point-line searches) and a selected number of no-flaw (background noise) regions. The set of measure values for the no-flaw regions will be defined as class 1 and the set of measure values for the void regions will be defined as class 2. Several approaches can be used to evaluate the classification merit depending on the characteristics of the data. Often, a linear (Fisher) discriminator can be used to quantitatively assess the ability to separate a measure into two distinct classifications through evaluating the means (m_1,m_2) and standard deviations (s_1,s_2) of the two conditions (no flaw = 1, flaw = 2) with the following relation:[7]

$$J_1 = \frac{\left|m_1 - m_2\right|^2}{s_1^2 + s_2^2} \tag{2}$$

For example, the mean and standard deviation measurements might be taken from the greatest slope/change in pixel grey scale of the lines in the different directions using the line-based approach. For our study, the first column of Table 3 identifies the specific measures used in this study to calculate the means and standard deviations. This approach is valid when the distributions are normal or the sample number is high. Thus, with the application of new signal processing algorithms that improve the visual characteristics of the image to better resolve the flaw (void) and no flaw data, the value of J_1 increases.

When the distribution of the no-flaw and flaw (void) data set measure results are non-parametric (non-normal), other approaches can be used to quantify the merit of an algorithm to differentiate two classes. The receiver operating characteristic (ROC) curve presents a relationship between the detection rate (sensitivity) and the 1 minus the false call rate (specificity) for a varying detection threshold level.[8] It is constructed by varying the detection threshold from zero to the maximum data value and plotting the corresponding (1—false call rate) and (1—missed flaw) results. Figure 11(a) displays an example ROC curve plot. The Euclidean distance (d) can be defined between a perfect classifier ($MF = 0$, $FC = 0$) and the ROC curve such that

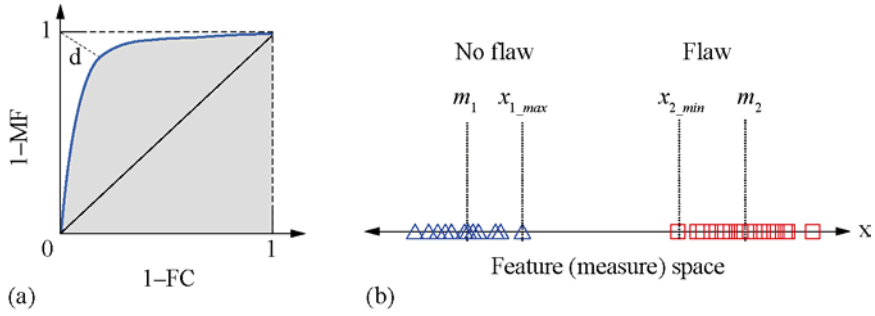


Figure 11. (a) Example receiver operating characteristic (ROC) curve in terms of specificity (1—missed flaw (MF) rate) and sensitivity (1—false call (FC) rate) and (b) a plot of a feature measure data presenting fully separable classes.

$$d = \sqrt{MF^2 + FC^2} \tag{3}$$

where *MF* is the missed call rate and *FC* is the false call rate for a given operator-defined detection threshold level. This threshold level involves setting a demarcation between the flaw and no flaw (noise) for the specific feature extraction measures described above. It is not directly applied to the grey scale image data but is instead applied to the feature extraction measure calculated from the grey scale image data that has been processed by a particular signal processing algorithm. Through optimization of the detection threshold to equally minimize the missed call and false call rate, the minimum of the Euclidean distance (d_{min}) can be found and used as a quantitative measure of the ability to classify/discriminate the void data set from the background data set for non-normal data. Although the ROC curve measure does address non-parametric distributions, it fails to differentiate the quality of classes of data where 100% classification is achieved for a specific threshold setting. To supplement the Euclidean distance measure, an estimate of the relative degree of separation between the no-flaw and void data sets is proposed. Shown in Fig. 11(b), the degree of separation can be evaluated using the relative distance between the two class extreme points with respect to the mean difference:

$$r = \frac{x_{2\_\min} - x_{1\_\max}}{m_2 - m_1} \tag{4}$$

where $x_{2\_\min}$ and $x_{1\_\max}$ represent the minimum value from the flaw data set and maximum value from the unflawed data set, respectively. For this study, $x_{2\_\min}$ would typically correspond with the smallest void in terms of volume in the flaw data set. For the maximum noise feature, $x_{1\_\max}$, this value corresponds with the

strongest background noise feature from the randomly selected set of background regions. A new merit measure is proposed for non-parametric data based on the ROC curve measure and the degree of separation between the no flaw and void conditions:

$$J_2 = -\sqrt{2 \cdot d_{\min}} \qquad d_{min} > 0, \tag{5}$$

$$J_2 = r \qquad d_{min} = 0. \tag{6}$$

The value for J_2 will vary from –1 to 1 with 0 being the point at which the data sets become separable with no false calls and no missed flaws. Thus, the best algorithms will produce merit values for J_2 greater than zero (and ideally approaching one) which correspond with a significant degree of separation between the entire flaw and no-flaw data sets. Since this measure is not independent of the sample size, it is proposed only as a relative measure for data sets of equal size. As a compromise between the two measures, a non-linear combination of the two measures was also used in this study

$$J_3 = \alpha J_1^n + \beta J_2^m \tag{7}$$

Parameter values of $\alpha = 1$, $\beta = 16$, $n = 0.5$, and $m = 1$ were chosen to both linearize and equally weight J_1 and J_2. Similar to J_1 and J_2, the ability of the no-flaw and flaw (void) data sets to be separated increases as the value of J_3 increases. In this study, all three merit measures were used to investigate the ability of the proposed signal processing algorithms to correctly separate the selected void and unflawed data sets.

9. Evaluation Results

A quantitative evaluation of the merit of several signal processing algorithms was performed. Results from six algorithms are presented in this study: (a) time-domain peak to peak amplitude, (b) centroid of time-domain signal (in time steps), (c) power spectrum density peak magnitude, (d) PSD(0.12 THz), (e) PSD(0.16 THz), and (f) centroid (f_c) of power spectral density. The corresponding images for the six signal processing algorithms are presented in Fig. 12. Table 3 presents the quantitative results for the six signal processing algorithms for separating all 6.4 (mm) diameter holes (the second row of voids from the bottom) from the background noise. The second smallest diameter was chosen for the quantitative evaluation since it is a challenging detection problem while also providing visual results for at least some of the voids for each of the

proposed algorithms in the study. The results for two of most promising quantitative measures from prior studies are presented: the peak slope for a line search (A4) and the edge measure across a region (B4).[6] In addition, the maximum response (largest grey level value) at the center of the void regions was also evaluated to determine the best signal processing method for irregular void detection, where the sharp edge features due to the cylindrical edge of the artificial voids are not present.

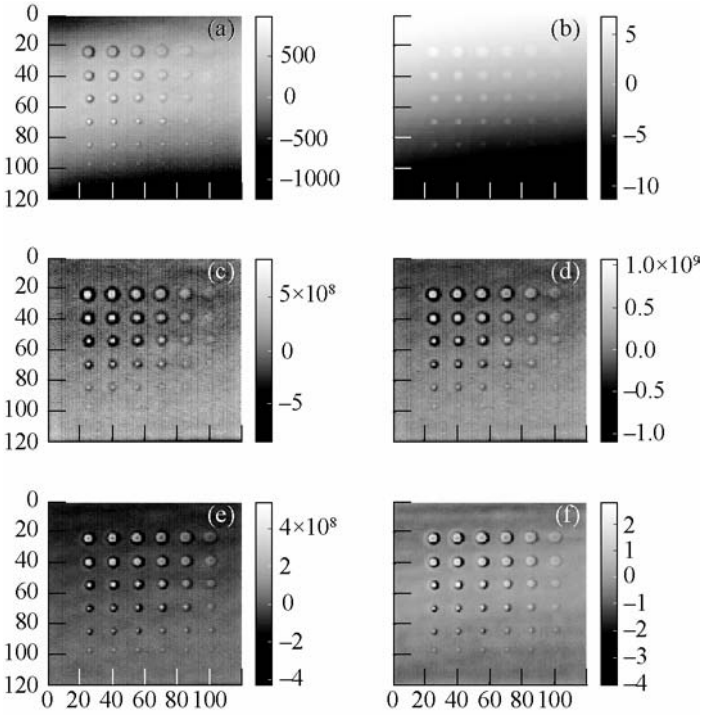


Figure 12. THz images for various signal processing methods: (a) Time-domain peak to peak amplitude, (b) centroid of time-domain signal, (c) power spectrum density peak magnitude, (d) PSD(0.12 THz), (e) PSD(0.16 THz), (f) centroid (f_c) of PSD(f).

All three statistical measures, J_1, J_2 and J_3 are presented. Overall, the centroid of the power spectrum density (f_c) for the peak signal was found to perform best. The high frequency component PSD(0.16 THz) was also found to significantly improve the detectability of small voids with respect to conventional peak to peak measurements. The centroid of the time domain peak signal[9] also scored well but only for the peak slope measure. The time domain centroid is calculated similarly to the centroid of the power spectrum density (f_c) but using "t" (time) rather than "f" (frequency) in equation (1). This was due to the sharp edge

Table 3. Quantitative results from merit assessment for six signal processing algorithms for separating all 6.4 (mm) diameter voids (second row from bottom) from noise. Bold indicates top two scores.

Quantitative measure	**Statistical measure**	**(a) Time peak-peak**	**(b) Time centroid**	**(c) PSD(f) peak**	**(d) PSD(f) (f=0.12 THz)**	**(e) PSD(f) (f=0.16 THz)**	**(f) PSD(f) centroid**
Line search, peak slope[a]	J_1	2.29	3.71	0.53	1.61	**5.05**	**5.66**
Line search, peak slope[a]	J_2	–0.29	**0.14**	–0.43	–0.31	–0.05	**0.09**
Line search, peak slope[a]	J_3	–3.08	**4.22**	–6.19	–3.63	1.37	**3.54**
Region search, edge measure [b]	J_1	2.12	3.39	2.71	3.28	**5.27**	**6.02**
Region search, edge measure [b]	J_2	–0.26	**0.15**	–0.04	–0.04	–0.01	**0.16**
Region search, edge measure [b]	J_3	–2.75	**4.24**	1.02	1.18	2.17	**5.07**
Center, max measure[c]	J_1	1.89	0.01	1.34	2.70	**4.39**	**5.43**
Center, max measure[c]	J_2	–0.40	–0.76	–0.43	–0.24	**0.00**	**–0.04**
Center, max measure[c]	J_3	–5. 03	–12.05	–5.76	–2.13	**2.03**	**1.76**

[a]Maximum grey level slope change along line (using a 5 point fit).

[b]The proposed edge measure (B4) incorporates a Canny edge detection algorithm, evaluates the magnitude at each edge pixel and sums the total edge pixel grey level response as a quantitative region measure.

[c]The maximum response (maximum grey level value) at the center of the void regions was also evaluated to determine the best signal processing method for irregular void detection, where the sharp edge features due to the cylindrical edge of the artificial voids are not present. The center of the void was selected near the middle of the void features and a 'local' line search (only searching two pixels away from the center point in each direction) was used to evaluate the center maximum.

observed between the hole and background and the relatively slow change in the background response. When the edge feature is eliminated from the evaluation to address irregular voids by only considering the maximum response at the center of the void, the centroid of the power spectrum density (f_c) for the peak signal again was again found to perform best. These results are in agreement with those observed subjectively, thereby validating the automated quantitative protocol for this case. To gain statistical confidence, more data and cases need to be explored.

To better illustrate the methodology, example calculations of the quantitative merit assessment for the PSD centroid algorithm using the region-based edge measure are presented in Table 4. Values are presented for all parameters used in the evaluation of the three statistical measures (J_1, J_2, J_3).

Table 4. Example of calculations for quantitative merit assessment for the PSD centroid algorithm for separating all 6.4 (mm) diameter voids from background noise using a region search with an edge measure.

Parameter	**Variable**	**Value**
number of background (noise) samples	n_1	360
number of flaw (void) samples	n_2	6
mean of background (noise) samples	m_1	0.25
mean of flaw (void) samples	m_2	6.26
standard deviation of background (noise) samples	s_1	0.26
standard deviation of flaw (void) samples	s_2	2.44
measure J_1: linear (Fisher) discriminator	J_1	**6.02**
minimum missed flaws for best threshold by ROC curve	$MF_{\_min}$	0.0
minimum false calls for best threshold by ROC curve	$FC_{\_min}$	0.0
minimum of the Euclidean distance from ROC curve	$d_{\_min}$	0.0
maximum value from background (noise) samples	x_{1_max}	1.28
minimum value from flaw (void) samples	x_{2_max}	2.26
the relative distance between extremes of flaw / no flaw data	r	0.16
measure J_2 for non-parametric data	J_2	**0.16**
measure J_3 composite measure	J_3	**5.07**

A plot is also presented in Fig. 13 showing the distribution of the sample data points for both the 6.4 (mm) diameter voids and the background noise resulting from a region-based search with an edge measure. The J2 merit assessment was used to evaluate the PSD centroid algorithm. Each of the background (noise) regions is defined by circular regions (with a radius of 5 pixels) placed at the center of randomly selected points in the background. The random distribution of the center of the no-flaw data regions is shown in Fig. 9(a) as indicated by the light-colored cross-hair symbols. Due to the limited void samples ($n_2 = 6$) used in this calculation and the character of the distributions resulting from this edge based feature extraction method, nonparameteric (J2) statistical measures were found to be appropriate. Line-based search methods that use selected data points in known areas of interest can be used to provide more statistical data samples. For this study, four points for each void site were used (as shown in Fig. 9(a) per the dark-colored cross-hair symbols around the edges of the flaws) resulting in a total number (n_2) of 24 flaw (void) data samples.

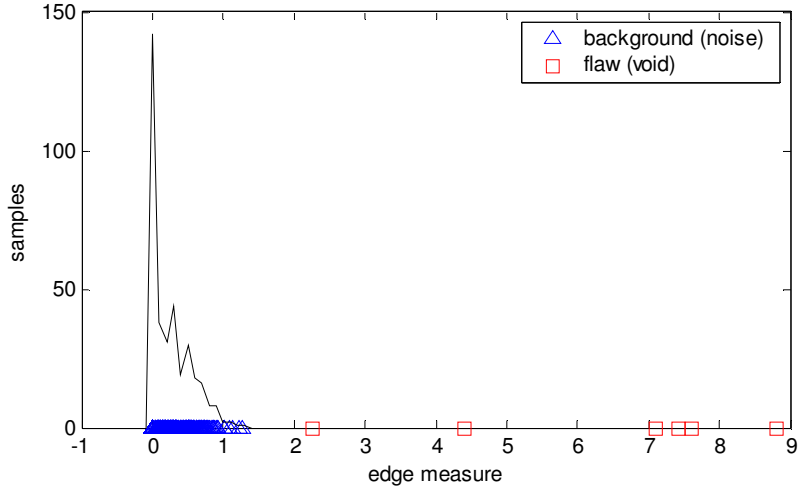


Figure 13. Plot of region-based edge measure results for separating all 6.4 (mm) diameter voids from background noise using J_2 non-parametric merit measure for the PSD centroid algorithm.

10. Conclusions and Recommendations

In this study, results from various signal processing approaches to improve resolution of defects in terahertz images of external tank foam were described. Centroid of the power spectral density for the main (substrate) reflection off of a metallic substrate under the foam shows promise versus conventional parameters of the signal used to form images. The centroid is reasonably simple and rapid computationally. Line profile and region-of-interest examination are suggested as ways to compare flaw resolvability/sensitivity when evaluating various signal processing methods. Automated feature extraction methods were implemented to replicate the important visual classification characteristics of the defect and noise used by inspectors. Both global and local image analysis methods were introduced for quantitative evaluation. A methodology was also developed for merit assessment of particular classification features in combination with the most promising signal processing algorithms. The protocol also addresses the evaluation of the best combination of features and signal processing algorithms for optimal discrimination. This analysis protocol was applied to a case study data set and the performance of both the feature extraction measures and the signal processing algorithms was studied where several signal processing methods were found to improve the discrimination of defects. The results from the analysis protocol agreed with those observed subjectively in terms of

identifying the signal processing methods that produced the best discrimination of voids.

Several questions remain concerning the evaluation of the best signal processing algorithms for practical application to real defects. In particular, additional work is needed to evaluate the impact of irregular void profiles on signal processing algorithm performance. Prominent edge (ring) features around voids will likely not be present for irregular voids. Second, additional signal processing development may be needed to address new inspection requirements for detection and characterization of delaminations, cracks and crushed foam. Quantitative evaluation and fusion of methods including terahertz, back-scatter X-ray, microwave, and shearography for optimum sensitivity is of interest. As a new protocol is developed, it is expected that these methods will be made available in the NASA software, which is applicable to any pulsed broadband signal data sets used to form images.

References

1. Hu, B. -B. and Nuss, M. C., "Imaging with terahertz waves," Opt. Lett., Vol. 20, p. 1716 (1995).
2. Mittleman, D. M., Jacobsen, R. H., and Nuss, M. C., "T-ray imaging," IEEE J. Sel. Top. Quant. Elec., Vol. 2, p. 679 (1996).
3. Mittleman, D. M., Gupta, M., Neelamani, R. G., Baraniuk, J. V., Rudd and Koch, M., "Recent advances in terahertz imaging," Appl. Phys. B Vol. 68. pp. 1085–1094 (1999).
4. Winfree, W. P. and Madaras, E. I., "Detection and Characterization of Flaws in Sprayed on Foam Insulation with Pulsed Terahertz Frequency Electromagnetic Waves," Proceedings of the 41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Tuscon, Arizona, July 10–13, 2005.
5. Roth, D. J., Seebo, J. P. Trinh, L. B., Walker, J. L., Aldrin, J. C., "Signal processing approaches for terahertz data obtained from inspection of the shuttle external tank thermal protection system foam," Proceedings of the 33rd Annual Review of Progress in Quantitative Nondestructive Evaluation. Hilton Portland & Executive Tower Portland, Oregon July 30–August 4.
6. Aldrin, J. C., Roth, D. J., Seebo, J. P., Winfree, W. P., "Protocol and Assessment of Signal Processing and Feature Extraction Methods for Terahertz NDE for Spray-on Foam Insulation," Proceedings of the 33rd Annual Review of Progress in Quantitative Nondestructive Evaluation, Portland, OR, July 30–Aug. 4, 2006.
7. K. Fukunaga, An Introduction into Statistical Pattern Recognition, Academic Press, (1990).
8. Hanley, J. A. and McNeil, B. J. "The meaning and use of the area under a receiver operating characteristic curve," Radiology, 143, (1982), pp. 29–36.
9. Roth, D. J., Verrilli, M. J., Cosgriff, L. M., Martin, R. E., and Bhatt, R. T., "Microstructural and Discontinuity Characterization in Ceramic Composites Using an Ultrasonic Guided Wave Scan System," Materials Evaluation, Vol. 62, No. 9, pp. 48–953., Sept. 2004.

CHAPTER 27

A ROBUST ALGORITHM FOR NONDESTRUCTIVE TESTING OF WELD SEAMS

Miguel A. Carrasco and Domingo Mery

Departamento de Ciencia de la Computación
Pontificia Universidad Católica de Chile.
Av. Vicuña Mackenna 4860 (143). Santiago de Chile
mlcarras@puc.cl; dmery@ing.puc.cl

The detection of welding flaws by means of nondestructive inspection methods remains open to the development of new algorithms and methods of inspection. One of the most widely used techniques is radiographic analysis, which requires interpretation by trained inspectors. Unfortunately, manual inspection is subject to various factors that can alter performance in the detection of the faults. An automated welding fault segmentation algorithm is presented using a set of digitized radiographic images. The result of the study has allowed the development of the following scheme: first, use the median filter to reduce noise; second, apply the bottom-hat filter to separate the hypothetical faults from the background; third, determine the segmented regions by binary thresholding; fourth, use the filters provided by morphological mathematics to eliminate over segmentation; and fifth, use the watershed transform to separate the internal regions. The results of the study have generated a general ROC curve on a set of 10 images with an area A_z=93.6%.

1. Introduction

Quality control of manufactured products has become one of the main objectives of production processes. For its evaluation several inspection and analysis techniques are available that can be applied during the manufacturing process, but all of them depend on the fulfillment of safety standards imposed by the manufacturer or by some regulatory standard. Although some manufacturers tolerate faulty products, for others products safety plays a critical role. Traditionally, quality control has been carried out manually by an inspector, because human inspection is flexible and adaptable to new situations that have not been considered, but this process itself has serious drawbacks such as i) Time

consuming: it takes a large amount of time and it is dependent on weariness and monotony at work; ii) Inconsistency: it depends on the inspector's capacity and experience for its analysis. These factors have led industry to gradually replace human inspection by automatic methods. One of the conventional forms of applying automatic inspection has been through machine vision systems. The objective is to determine automatically if the product fulfills a set of previous specifications using only visual information. This methodology is known as Automatic Visual Inspection (AVI), a name that encompasses a large set of analyses and algorithms that are divided into a series of processing stages that include image formation, preprocessing, segmentation, extraction of characteristics, and classification.

AVI has largely solved visual quality control, setting precise and objective control policies [1]. The main objective is to determine whether a product falls within or outside the range of acceptance in the manufacturing process. To that end it has to fulfill two basic conditions to improve product quality: efficiency and speed [2]. i) Efficiency means detecting the large number of defective products, and at the same time rejecting the smallest number of products in good condition. ii) Speed means that the production process is not affected by the time taken for the inspection, i.e. that production speed is maintained or increased. Current methods aim at decreasing the number of false positives and false negatives. Ideally these variables should be 0%, however this is difficult to achieve because it depends on the kind of product that is being analyzed. The decision to decrease some of these variables affects the performance of AVI quality control. Some of the effects of the adjustment can be: to detect 100% of products free from defects and reject a margin of products in good condition; or to accept a margin of defective products together with all the products in good condition.

Table 1. Examples of recent automatic visual inspection techniques

Textiles	Mamic *et al* (2000); Zhang, *et al* (2005)	[3, 4]
Welding	Liao (2003); Carrasco & Mery (2004)	[5, 6]
Bearings	Liang-Yu L. *et al* (2005)	[7]
Automobile doors and glass	Yun Koo Chung (1998)	[8]
Vehicle wheels	Mery *et al* (2002); Carrasco & Mery (2006)	[9, 10]
Aircratf turbines	Nguyen *et al* (1998)	[11]
Printed circuits	Park *et al* (2006)	[12]
Food	Pedreschi *et al* (2004)	[13]
Glass bottles	Mery & Medina (2004); Shafait, *et al* (2004)	[14, 15]
Wood	Lihari & Obac (2005)	[16]

In general, the applications of AVI are found in different areas of industry, such as detection of faults in printed circuits, welds, glass, textiles, food, aluminum wheels, bearings, aircraft turbines, etc. Therefore, every inspection is directly related to the kind of material that it is desired to analyze (see Table 1). A method designed for recognizing faults in one material may not be applicable to another [2].

One of the most important quality control areas for determining faults in welded structures has traditionally been done with radiographs, and in recent years with time-of-flight diffraction (TOFD) ultrasound techniques. Radiography allows the inspection of the quality of structures made from metals or other materials by making X-rays or γ-rays go through them. The faults are shown by the different light intensities generated by capturing the energy transmitted by the radiation on the radiographic images [17]. Ultrasound techniques use high frequency sound waves to detect the faults. TOFD is based on measuring the time and amplitude taken by the wave to travel from the emitter to a receptor along the weld. The faults are revealed by an alteration of the waves if they are found between the surface and the background of the material [18]. The above techniques have the advantage that the analyzed material is not destroyed during the inspection process, but both share the same problems of manual quality control, because the quality of the inspection is largely dependent on the quality and training of the inspector.

The field that covers nondestructive tests is known as NDT, and it involves various inspection techniques whose objective is to ensure that the design and operation requirements are safe and reliable. For that purpose different inspection techniques have been developed [19], the most common of which are the methods of radiographic analysis and of ultrasound described above. According to Moura *et al* [20], ultrasound methods have comparative advantages over radiographic methods as they avoid the effects of ionizing radiation on the operators and the use of radiographic films or plates which in some cases degrade the quality of the images [21]. However, classification of defects through ultrasound methods is often questioned because it is highly subjective, since the analysis and identification depend exclusively on the experience and knowledge of the operator. On the other hand, radiographic methods are the most widely used as inspection methods of welds in NDT. Their development and research continue being applied in industry and in science, leading to technical and/or economic advantages [22].

According to Liao [5], most radiographic inspection procedures have three parts: segmentation of the weld joints on the background,[a] segmentation of the flaws on the weld joints, and classification of the flaws. Our research begins with the second step, i.e. the position of the weld joint is known. Therefore, our approach consists in segmenting the largest number of flaws in the welds. Segmentation is commonly considered one of the most complex tasks in image processing [24]. Research in this field is extensive, but it is specific to the material that is being analyzed. In this research, different strategies and methods have been evaluated, oriented at the detection of faults in images of welds supplied by BAM[b].

This chapter is organized as follows: Section 2 presents the state of the art in flaw detection in welds; Section 3 presents the segmentation methods and the solution scheme developed in this study; Section 4 presents the results obtained; and Section 5 presents our conclusions. An extended version of this article may be found in [25].

2. State of the Art

There are structures that use welds for critical functions, such as high pressure equipment, chemical compounds, etc, where any kind of flaw can trigger catastrophic consequences. The conventional forms for detecting welding flaws are by means of visual inspection of radiographic images. The images are generated by using X-rays and γ-rays which penetrate the material generating a radiological image on a photographic plate. Flaws are detected due to variations in the density of the material (see Fig. 1). Nonetheless, manual interpretation of flaws can generate subjective and imprecise results which require a great deal of time and are inconsistent in that they depend on an inspector for their analysis [5].

Due to the problems associated with manual detection, there is currently a great deal of work and research on non-destructive testing (NDT) methods for detecting welding defects. The objective is to develop an automated method for the detection of defects that is precise and objective. Some of the most important achievements in this area are presented below.

[a] An application to determine the position and orientation of welded joints is available in [23]

[b] The images for this research are part of a set of radiographic images of welds generated by the Federal Institute for Materials Research and Testing, Berlin (BAM).

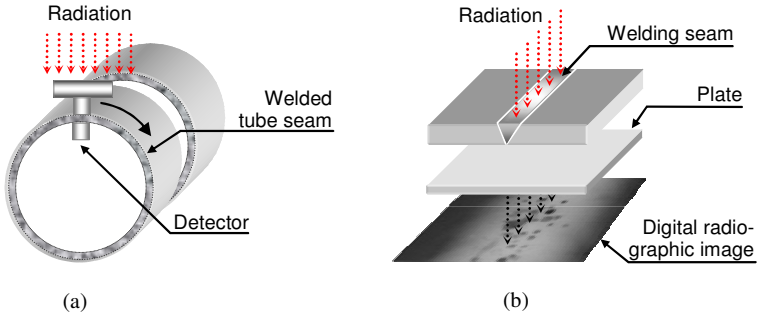


Figure 1. (a) Flaw detection scheme through radiation of weld. (b) Radiation over material flaw captured on a radiographic plate.

Gayer *et al* [26] proposed a method that can be summarized as having two steps: i) A quick search for potential defects in the X-ray image: Assuming that the defects will be smaller than the regular structure of the test piece, potential defects are classified as those regions of the image where higher frequencies are significant. The spectrum of the X-ray image is determined with the help of a fast Fourier transformation, which is calculated either row by row or column by column in small 32×32 windows. When the sum of the higher frequencies of a window is greater than a given threshold value, the entire window is marked as potentially defective. Another possibility is suggested by the authors as part of this task: A window is selected as potentially defective when the sum of the first derivative of the rows and columns of a window is large enough. ii) Identification and location of the true defect: Because of the time-consuming nature of this step, only those regions which were previously classified as being potentially defective were studied here. Two algorithms were developed here as well. The first leads to a matching between the potential defect and typical defects, which are stored in a library as templates. Whenever a large resemblance between the potential defect and a template is found, the potential defect is classified as a true defect. The second algorithm estimates a defect-free X-ray image of the test piece by modeling every line of an interpolated spline function without special consideration for the potentially defective region. Following this, the original and the defect-free images are compared. True defects are identified when large differences occur compared to the original input image.

Lawson and Parker [27] proposed that artificial neural networks (ANN) be used for the automated detection of defects in X-ray images. The method generates a binary image from the test image where each pixel is either 0 when a regular structure feature of the piece exists or 1 when a defect is detected. This entails the supervised learning of a multi-layer perceptron network (MLP) where

an attempt is made to obtain detection from training data. A back propagation algorithm is used for the assignment of weights within the MLP. The authors use one or two hidden layers in the network topography of the ANN, where the input signal corresponds to a window of $m \times m$ grey values in the X-ray image. The output signal is the pixel at the image centre in the binary image. Since the threshold value function for the neurons is sigmoidal in this method, a threshold is used to obtain a binary output signal. The desired detection in the training data was obtained with a segmenting procedure based on an adaptive threshold. During the experiments with five X-ray images, the authors showed that detection using ANN is superior to the segmenting method using adapted thresholds.

Sofia and Reduane [28] proposed a method for automated recognition of welding defects. The detection follows a pattern recognition methodology: i) Segmentation: regions of pixels are found and isolated from the rest of the X-ray image using a watershed algorithm and morphological operations (erosion and dilation). ii) Feature extraction: the regions are measured and shape characteristics (diameter variation and main direction of inertia based on invariant moments) are quantified. iii) Classification: the extracted features of each region are analyzed and classified using a k-nearest neighbor classifier. According to the authors, the method is robust and achieves a good detection rate.

Silva *et al* [29] proposed another welding defect classification method. In a first step, called image pre-processing, the quality of the X-ray image is improved using a median filter and a contrast enhancement technique. The defect detection follows the pattern recognition scheme mentioned above: i) Potential defects are segmented in the X-ray image. ii) Geometric and grey value features (contrast (C), position (P), aspect ratio (a), width-area ratio (e/A), length-area ratio (L/A) and roundness(R)) are extracted. The correlation between features and each defect class considered (slag inclusion, porosity, lack of penetration and undercutting) was evaluated by analyzing the linear correlation coefficient. iii) The most relevant features were used as input data on a hierarchic linear classifier [29]. In order to achieve a higher degree of reliability of the results, radiographic standards from the International Institute of Welding were used, with 86 films containing the main defect classes. The experimental results show that features P and e/A are able to classify the undercutting and lack of penetration classes. Nevertheless, the six mentioned features are required to obtain high performance by classifying the porosity and inclusion defects.

Liao and Li [30] proposed a detection approach based on curve fitting. The key idea of this work is to simulate a 2D background for a normal welding bead

characterized by low spatial frequencies in comparison with the high spatial frequencies of defect images. Thus, a 2D background is estimated by fitting each vertical line of the weld to a polynomial function, and the obtained image is subtracted from the original image. The defects are detected where the difference is considerable. Wang & Liao [31] and Liao [5] proposed a fuzzy k-nearest neighbor, multi-layer perceptron neural network and a fuzzy expert system for the classification of welding defect types. The features used for the classification are distance from centre, circularity, compactness, major axis, width and length, elongation, Heywood diameter, and average intensity and standard deviation of intensity. Finally the (K-NN) and the (MLP) methods are used for classification. The results indicate that the (MLP) method is superior to the (K-NN) method, classifying 92.39% and 91.57% respectively.

Mery and Berti [32] presented a new methodology based on texture analysis. Texture is one of the most important characteristics in pattern recognition, but it has seen limited use in the analysis of digital images in NDT. The referenced study examines the analysis of two types of texture features: those based on the occurrence matrix, and those based on the Gabor function. The proposed approximation uses the following methodology: i) Segmentation: the LoG edge detector is used. ii) Extraction of characteristics: the features of potential defects are extracted. iii) Classification: the most relevant features are used as input data for a statistical classifier. The best results have been achieved with a polynomial classifier, with 91% defect detection, and 8% false alarm rate.

Li *et al* [33] proposed the development of an adaptive segmentation algorithm through genetic algorithms (GA) [34]. One of the major problems of segmentation is the determination of an appropriate threshold value that allows the separation of the relevant objects from the background. One of the most widely used methods is that of Otsu [35], because it determines an optimum threshold value for two classes. Unfortunately, the weld images contain more than two classes due to the surface of the object, light interference, etc. According to Li *et al*, the solution consists in generalizing Otsu's method as a multiple class problem that can be tackled through a genetic algorithm. The results indicate that the designed method is adaptive and efficient at generating a segmentation in welded joints.

Wang & Wong [36] presented a welding flaw segmentation method by means of the Fuzzy C-Means algorithm. The method consists of three steps: First, the top-hat and bottom-hat filters are applied. These filters extract the light from the objects generating low changes over the background of the image. The defects can thus be stressed and the background regions can be eliminated. Second, apply an adaptive wavelet thresholding filter (proposed by Donoso [37]). The purpose

of this filter is to eliminate the noise present in the signal while preserving the characteristics of the signal. According to Chang *et al* [38], this filter retains the sharpness of the edges of the defects better than the median filter. Third, use a fuzzy c-means (FCM) clustering algorithm. The clustering allows the assignment of a class depending on the degree of similarity that there is in the patterns that make it up [39]. One of the advantages of using this technique is that it allows the use of any number of characteristics and assign them to any number of classes, in addition to being applicable to instances in an unsupervised way. According to Wang & Wong, this technique has a more efficient performance compared to the method of Otsu, so the fuzzy c-means algorithm can detect a greater number of flaws in the welds.

Movafeghi *et al* [40] proposed improving the quality of the digitized radiographic images to intensify the location and recognition of the flaws. The research proposes the independent use of three techniques, two in the space domain, filters through morphologic mathematics and pseudo-color, and one in frequency, the wavelet filter. The morphologic mathematics technique consists in first applying a median filter to reduce the noise level and retain the edges, and then using a bottom-hat and top-hat filter operation. The pseudo-color technique converts the image from grey levels to a new color image using color mapping. The objective is to improve the visualization of the flaws. The wavelets technique makes it possible to decompose the image into subcomponents in the frequency and time domains. The idea is to select only those coefficients that have more information and eliminate the high frequency coefficients, which contain most of the noise. Then the reverse transformation is carried out and the image is reconstructed. This methodology has been used in radiological applications in mamography [41]. The results indicate that the morphologic mathematics technique has the best performance, with S_n=90% and 1-S_p=0%. According to the authors, the application of wavelets has greater complexity and must have supervised training.

Kleber *et al* [23] developed a methodology for extracting and limiting the weld joints in radiographic images. In various researches, analysis of the welds starts with a segmentation between the weld regions, discarding the background, but this requires a previous knowledge because it is necessary to determine the location, width, length and angle of the weld. Furthermore, the problem is increased due to markers and indicators that are inserted in the radiographic image. The methodology proposed by Kleber *et al.* uses a genetic algorithm (GA) [34] to locate the position of the welds using a model stored previously in the system. The objective is for the GA to determine the best position that matches

the stored model. The results have a 94.4% yield for detecting weld joints, either in different location, width, length and angle.

The literature reviewed includes a large number and variety of methodologies for the detection of welding defects, such as interpolating the image's background curves, neuronal networks, geometric characteristics, application of mathematical morphology and the watershed algorithm, texture analysis, etc. Research in this area continues largely because there are, as yet, no satisfactory results that allow the detection of all flaws without false alarms. Moreover, it is not possible to determine which of the research directions will improve the overall results, because each of them has room for improvement. The research presented in this paper is justified as it develops a new segmentation method which improves the process of automated detection of welding defects. Moreover, it also involves the development of a methodology that brings together some of the best characteristics of the reviewed studies, such as application of a median filter [29], comparison between a real image and a defect-free image [30], the use of mathematical morphology, and the watershed algorithm [28].

3. Steps of the Proposed Segmentation Algorithm

Segmentation consists in partitioning the image into disjoint regions, where each region is homogenous with respect to a given property, such as texture, grey level, type of color, etc., with the purpose of separating regions of interest for their later recognition [24]. Thus, the segmentation problem can be approached with different methods, which generally can be categorized into three types of techniques: those oriented to the detection of edges, pixel detection, and region detection [42].

In our study the edge detection was done first using gradient approaches. The Roberts, Prewitt, Sobel, and Laplace of Gaussian [42] filters belong to this category. These calculate the gradient, and according to some threshold level determine if a possible edge exists. Subsequently the Canny filter [43] was reviewed. This filter uses a combination of techniques, such as the Gaussian filter, for the elimination of noise, as well as making use of directional gradients, thus allowing the selection of only those edges that are found within the specified threshold. In pixel detection we have thresholding [42], which is a widespread technique as it allows the conversions of an image in grey scale to a binary image in such a way as to separate background objects according to a specified threshold. The watershed transform [44] is used in region detection. This technique makes use of morphological mathematics and allows the generation of regions based on cavity filling, simulating a valley filling with water; as the water

level rises, adjacent regions start forming unions. We also studied the distance transform [45] and the manner in which it can improve segmentation quality by using the watershed transform. In this case segmentation is used with the purpose of separating known regions from "hypothetical defects". These regions are made of true defects as well as false alarms, and it is the task of the classifier to correctly separate these two groups according to a set of geometric properties such as area, perimeter, invariant moments, etc., which are analyzed in the characteristic extraction process [9].

The proposed segmentation process uses a combination of digital processing techniques which have been selected on the basis of experimental tests and analyses, the development of which is presented in Fig. 2. Each stage has been analyzed independently and globally, allowing the study of different variations and strategies.

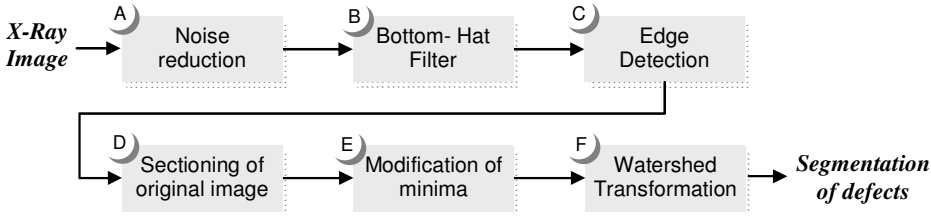
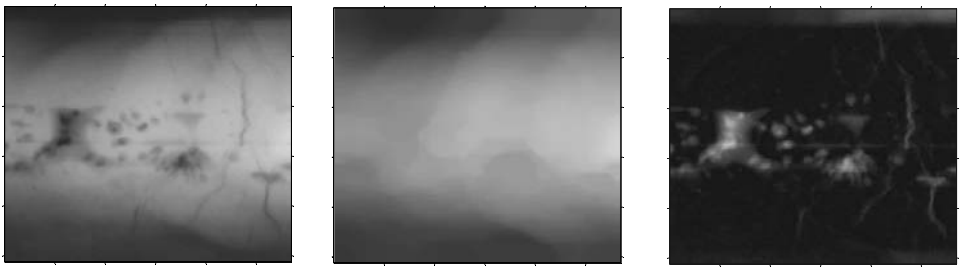


Figure 2. Proposed segmentation process for the detection of flaws in radiographic images.

A) Noise Reduction: The purpose of this phase is to attenuate the largest amount of noise in order to improve the segmentation process. There are different types of noise in an image, such as Gaussian, impulsive, frequential and multiplicative noise [42], and for that reason reduction is difficult, especially in radiographic images. Even though there is a large number of noise reduction filters, in this work we analyzed three techniques: the average, the Gaussian, and the median filters. Each filter has the purpose of attenuating in the best way some kind of noise, and they are therefore evaluated specifically in the BAM images. According to the experimental tests carried out, the median filter has at least two advantages with respect to the average and Gaussian filters: i) it is a more robust indicator, because an unrepresentative pixel does not introduce its value in the result, and ii) it does not generate new pixels when working on the edges. The median filter is therefore much better for preserving the edges clearly. Other research has shown the advantage of using the median filter. Such is the case of flaw detection in glass bottles [14].

B) Bottom-Hat Filter: The bottom-hat filter is used to highlight structures with hypothetical flaws and separate them from the background [46]. This process consists of two stages: i) First, use the morphologic lock operator on the original image, for which it is required that the size of the nucleus applied to the lock must allow the elimination of most hypothetical flaws, and as a consequence a background similar to that of the original image is generated, but without flaws. ii) Second, use the subtraction operator between the original image and the modified image from the first stage. As a result, the smaller structures, which in general are flaws, are revealed because of their separation from the background (Fig. 3).



(a) (b) (c)

Figure 3. (a) Original image; (b) Application of the lock operator to the original image; (c) Difference between images (a) and (b).

C) Edge Detection: The purpose of this stage is to determine a separation of the edges of each potential flaw. At this point we define two strategies: edge detection with the Canny filter, and binary thresholding. i) The first strategy uses the method of Canny and the dilation, filling and erosion operations. It is seen that lines have formed around the detected structures (Fig. 4a.), but it is necessary to close them, since the objective consists in generating a set of closed regions from those lines and in that way segment the hypothetical flaws from the image's background. The process continues with the application of the dilation shown in Fig. 4b. The use of this operator allows closing most of the open regions. Then the closed regions from the previous operation are filled; this process is illustrated in Fig. 4c. Finally, the erosion operator is applied to reverse the effect of the dilation and also determine more precisely the zones with flaws from already closed regions.

The method of Canny detects more precisely the edges of the structures because it is less sensitive to noise. Since it uses a Gaussian filter to decrease it; however, it generates a large amount of uncertain edges, and for that reason it is

necessary to carry out the process of dilation, filling and erosion. The disadvantage lies precisely in the generation of open edges, as seen in Fig. 4d; the dotted arrow B indicates a line that has not closed and was detected by the Canny method, but that line does not form a structure that can be closed, even though during the process of dilation and filling it has increased its thickness; the erosion operation returns it to its original state. In another zone of Fig. 4, arrow A indicates the formation of a curve, which suggests the presence of a structure with flaws, however once again the dilation process does not generate a closed region.

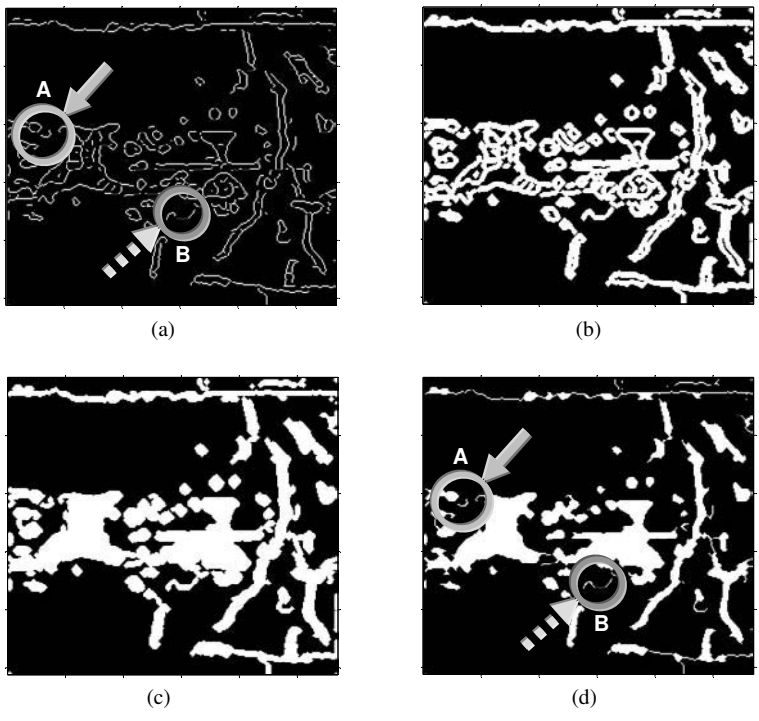
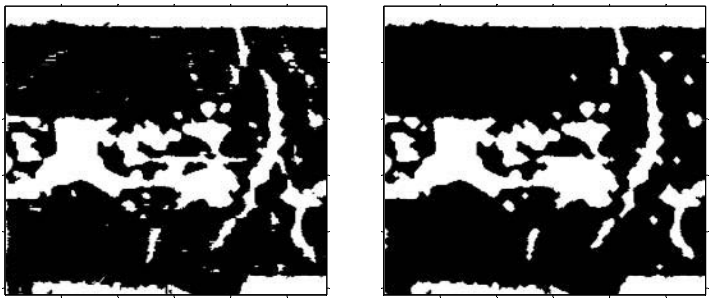


Figure 4. (a) Generation of edges with the method of Canny, (b) Dilation of the edges, (c) Filling of the closed regions, (d) Erosion of the structures.

The binary thresholding strategy consists in applying this operator to the result generated by the bottom-hat filter (Fig. 3c). In this way closed regions are obtained that can be defined by a given threshold. If it is desired to increase the collection of regions, the value of the threshold is decreased so that it allows a greater number of grey levels to pass; conversely, if the threshold is increased, a smaller number of regions are collected. However, the use of this operator generates noise.

The solution consists in applying the opening operator, which results in the generation of an image with low noise level and with uniform and closed structures. The opening operator can be adjusted by changing the size of the structures with which erosion and dilation are applied (Fig. 5b).

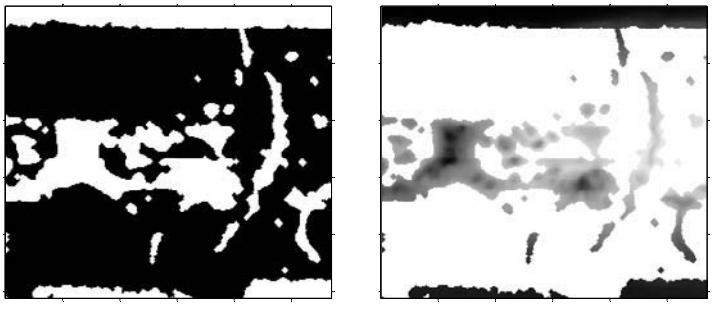


(a) (b)

Figure 5. (a) Binary image produced by thresholding. (b) Application of aperture to image (a).

The main disadvantage compared to the previous strategy is due to the lower detection of edges; the results show that the Canny method generates a large amount of edges in the form of lines without the possibility of closing them. On the other hand, the advantage of binary thresholding consists in generating most of the regions with already closed flaws.

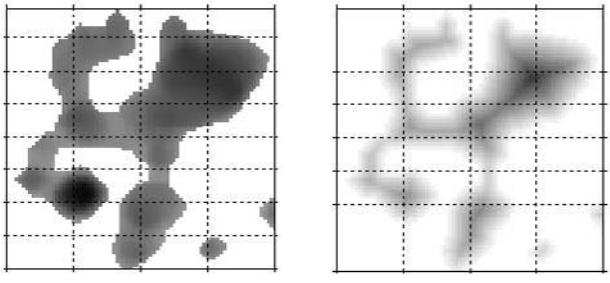
D) Sectioning of original image: Sectioning requires the application of binary thresholding. The idea consists in highlighting and segmenting only those regions with hypothetical flaws. In this point we define two strategies: i) The first consists in rescuing the pixels of the original image from the result of the binary thresholding. The previous result is used as a template that is superimposed on



(a) (b)

Figure 6. (a) Binary image from binarization with a given threshold. (b) Copy of pixels of the original image in levels of grey.

the original image (Fig. 6), and in this way it is defined that all the zones or pixels that are black are turned into white, and the white pixels of the binary image are used to copy all the pixels of the original image in the same position. ii) The second strategy uses the distance transform over the white regions, to generate a grey scale surface depending on the distance of its center with respect to its edges (Fig. 7). The purpose is to have a map of the structures with flaws, generated either through the distance transform or as a copy of the original image, because the next process uses the grey levels for segmentation through the watershed transform.



(a) (b)

Figure 7. (a) Original pixels from binary mask, (b) Distance transform applied to a binary mask.

When the distance transform is applied to a binary image that has an anthropomorphic shape, undesired results are obtained for later watershed analysis, because the resultant grey scale may not necessarily represent the surface that it is desired to divide. If there are closed circular figures, the differences can be smaller because the resultant formation of the distance operator has the same characteristic as the original image. However, the images that are being studied usually have quite varied shapes and grey scales, and do not represent exact shapes like those mentioned above. From this standpoint the application of the distance transform may not be recommended for the kinds of flaws that are being analyzed, because it does not represent the shape of the grey levels of the original image.

E) Modification of minima: The modification of minima is part of the strategy called “Homotopy modification” [47], which has allowed the avoidance of over segmentation and the generation of more precise segmentation. However, it is important to determine the number of pixels that make up the minimum because a variation in this figure can generate a greater or smaller number of hypothetical flaws (Fig. 8).

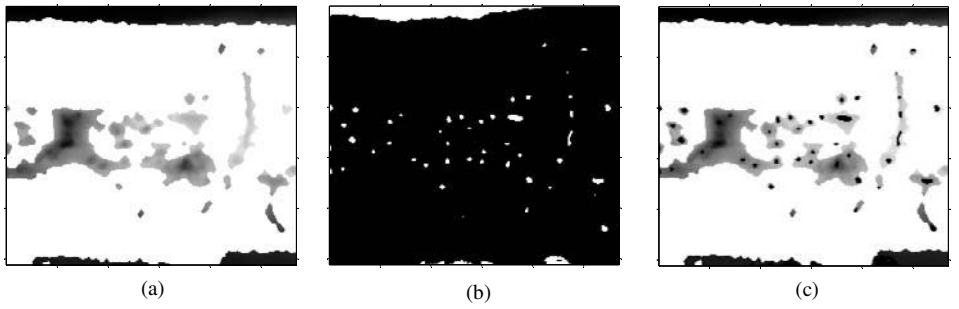


Figure 8. (a) Binary image. (b) Sections of the binary image on the original image, (c) Calculation of the minimum points of the cut image, (d) Superposition of the minima on image (b).

F) Watershed transformation: The final tool used is the watershed transform. This technique, in combination with the modification of minima, allows the segmentation of structures in the interior of the flaws, because the external segmentation has been carried out previously in the process of binary thresholding. The general segmentation process, especially the binary thresholding stage, has facilitated the generation of the majority of the structures and their edges. It has been shown that the watershed transform by itself does not

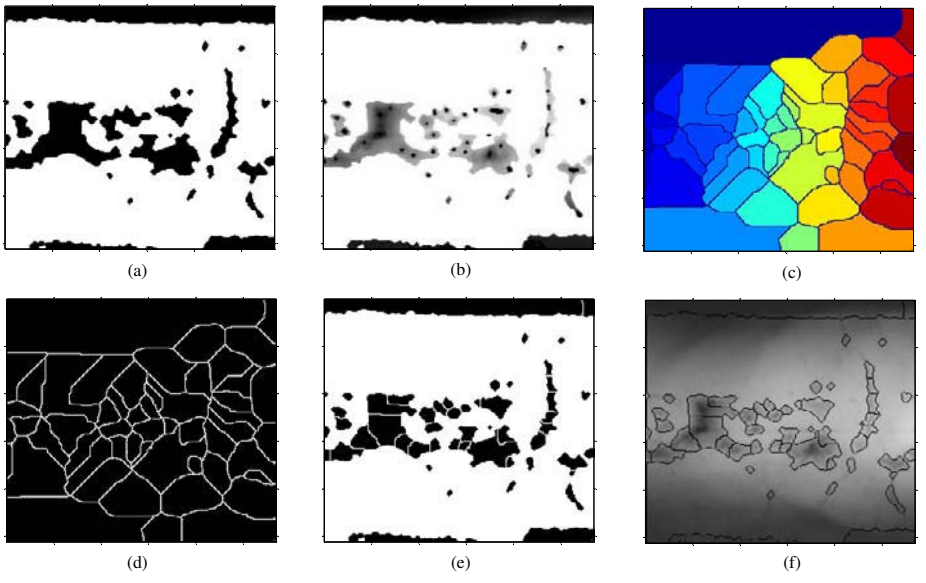
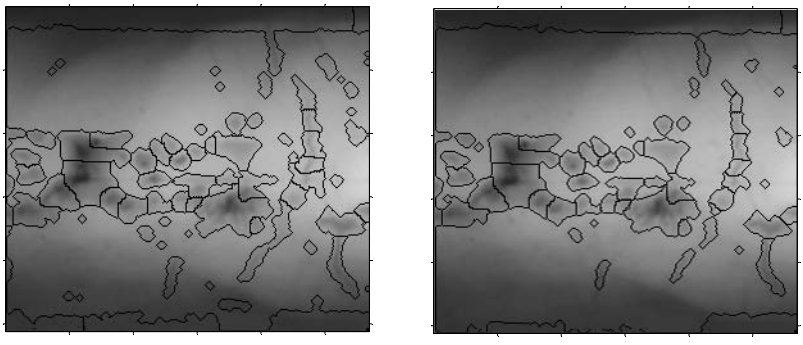


Figure 9. (a) Denied binary image. (b) Image with minima superposition. (c) Separation of regions by watershed. (d) Watershed lines. (e) Segmented binary image. (f) Superposition of the cuts on the original image.

generate the segmentation of the defects as it segments the entire image, thus highlighting the importance of the step prior to its application. For that reason, we apply the watershed transform on the structures that can be flawed, because only those regions are segmented and the zones adjacent to the flaws, which have been discarded in the bottom-hat process, are not considered. The whole process is described in Fig. 9.

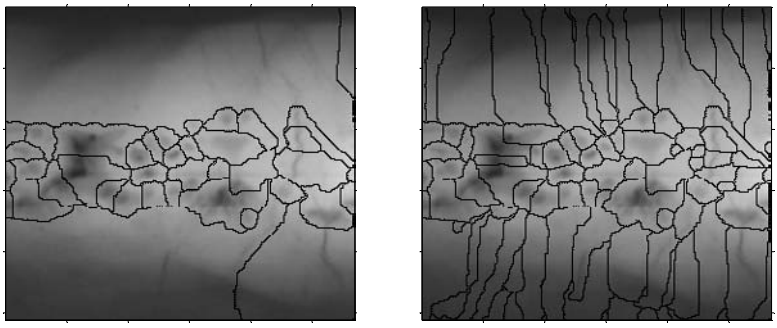
As the threshold is increased, the segmented regions decrease, and moreover they tend to have a smaller internal area (Fig. 10). Other process variables, such as bottom-hat filter, cut-off point selection, and dilation and erosion operations, influence the number and shape of the regions. However, modifying the variables generates different results, and in some cases this change can lead to imprecise segmentation, and thus it is important to analyze the effect of the modification of each variable, and how it affects the final result (Fig.10).



(a) (b)

Figure 10. (a) Segmentation of the image with threshold = 3; (b) Segmentation of the image with threshold = 5.

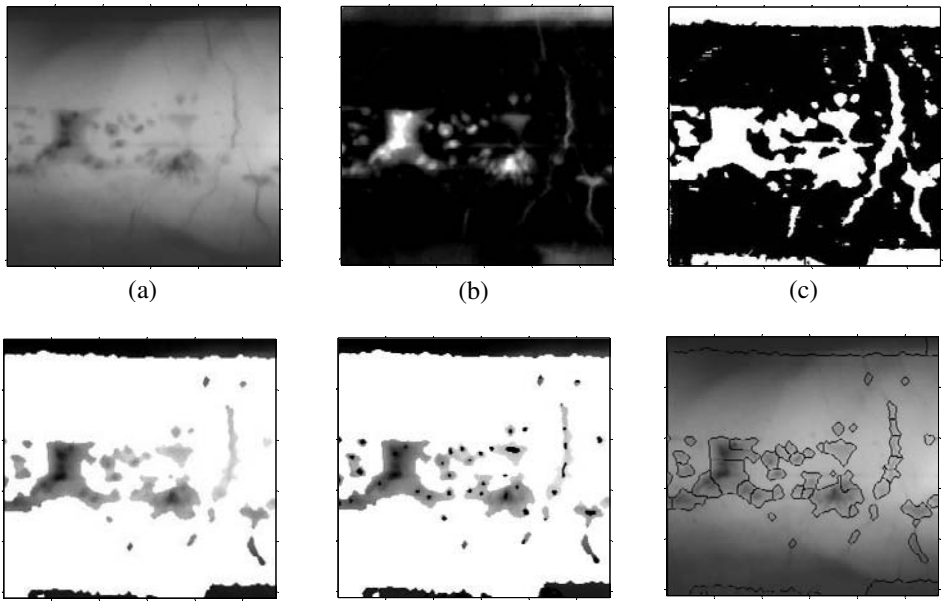
The resultant segmentation generates an erroneous division not shown in the previous process. For example, if the modification of the minima process is carried out directly on the original image, the result generates a segmentation with respect to the whole image and not on the flawed regions (Fig. 11a). Also, in the case of applying the watershed transform to the original image, without the cuts processing and without the modification of the minima, an over segmentation is generated that does not allow either the determination of the flawed regions or the determination of their location (Fig. 11b).



(a) (b)

Figure 11. (a) Watershed segmentation without separation of the regions. (b) Watershed segmentation without the "Homotopy modification" processing.

Summary of the proposed methodology. Every stage of the method proposed in Fig. 2 is respresented in each of the six images of Fig. 12. The image generated at the end of the process with the watershed transform represents the internal and external segmentation of each region.



(a) (b) (c)

(d) (e) (f)

Figure 12. Summary of the proposed segmentation process: (a) Image after application of the median filter. (b) Application of the bottom-hat filter. (c) Application of binary thresholding. (d) Application sectioning process. (e) Modification of minima. (f) Application of the watershed transform.

4. Results

The receiver operation characteristic (ROC) analysis is commonly used to measure the performance of a two-class classification. In our case, each feature is analyzed independently using a threshold classifier. In this way a hypothetical flaw is classified as a 'no-defect' (or 'defect') if the value of the feature is below (or above) a threshold value. The ROC curve represents a 'sensitivity' (S_n) versus '1-specificity' (1-S_p), defined as

$$S_n = \frac{TP}{TP+FN} \qquad 1-S_p = \frac{FP}{TN+FP} \tag{1}$$

where TP is the number of true positives (correctly detected defects), TN is the number of true negatives (correctly detected no-defects), FP is the number of false positives (false alarms, or no-defects detected as defects), and FN are false negatives (flaws detected as no-defects). Ideally, $S_n = 1$ and $1-S_p = 0$, which means that all defects were found without any false alarms. The ROC curve makes it possible to evaluate the performance of the detection process at different points of operation (as defined for example by means of classification thresholds). The area under the curve (A_z) is normally used as a measure of this performance as it indicates how flaw detection can be carried out: a value of $A_z = 1$ indicates an ideal detection, while a value of $A_z = 0.5$ corresponds to random classification [48].

The analysis of the ROC curve is carried out according to the method proposed in [49]. For each image, an ideal detection was achieved using visual interpretation. The methodology was to create an ideal binary image ('1' is defect and '0' is non-defect) according to the visual information with the Microsoft Paint software, using the largest scale (zoom = 800%). The results obtained with our algorithm were then compared with the ideal binary image. Thus, the values for TP, TN, FP and FN were tabulated as shown in Fig. 13.

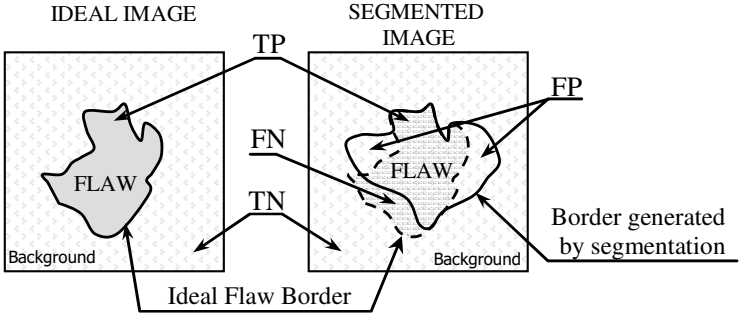


Figure 13. Representation of the differences between the ideal and segmented image.

Table 2. Summary of the segmented images and their best sensitivity and 1-specificity values

Image Name	Number of ideal regions	Regions segmented by the process	Best operational points		A_Z
			Sensitivity	1- Specificity	
BAM5.tif	273	495	90.04%	7.62%	93.08%
12R_M.tif	36	933	96.93%	3.87%	98.03%
13R_M.tif	62	520	90.38%	7.47%	95.22%
22R_M.tif	23	678	94.21%	4.29%	97.44%
28R_M.tif	47	1062	91.08%	5.10%	94.34%
31R_M.tif	11	179	99.51%	0.58%	99.52%
39R_M.tif	90	1110	87.68%	7.86%	94.36%
40R_M.tif	30	731	82.28%	3.74%	91.31%
106R_M.tif	97	836	81.24%	6.17%	89.90%
107R_M.tf	59	2461	80.65%	10.96%	89.41%

Table 2 presents a summary of the images analyzed in this study. The column labelled "Number of ideal regions" presents the number of real regions that must be segmented. The sensitivity and 1-specificity analysis is carried out pixel by pixel using the method proposed in [49]. For this reason, the results differ from those presented in [32], where the regions are analyzed as sets, and not pixel by pixel. The images used correspond to ten of the images catalogued by BAM for non-destructive testing in weld seams. These have been captured using a Lumysis LS85 SDR scanner in high density mode. An LUT linear fit has been used in order to reduce the grey levels of the original images from 12 to 8 bits, and they have been resized to leave only those areas that must be segmented.

In order for the study of the images not to depend on the set analyzed, a generalized ROC curve has been developed. This curve incorporates tests that use the same values with the purpose of obtaining a set of optimum parameters that will allow the use of this filter on an image that is not part of the study. In order to calculate the curve presented in Fig. 14, the values for TP, FN, FP and TN have been added to each test in the study, taking into account that the values being added must be generated on the basis of the same test for each image. Subsequently the 'sensitivity' and '1-specificity' values are calculated.

The ROC curve in Fig. 14 has an area of 93.58%; its best point has an 87.83% sensitivity and a 9.40% 1-specificity. The next-best point has 86.72% sensitivity and a 7.98% 1-specificity. In the latter case the decrease in false alarms is due to the increase in the area of detection, which means that regions with less than 27 pixels are not considered as 'hypothetical flaws.'

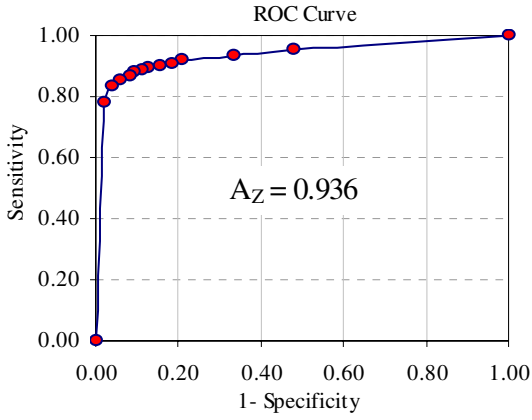


Figure 14. ROC Curve for *TP*, *FN*, *FP*, and *TN* values determined from a set of tests with the same parameters.

The proposed filter has been compared with the method developed in [32]. To carry out the comparison a binary image was been generated manually which contains all the real flaws of the original image. Once the segmented regions have been determined with the use of the proposed filter and the texture analysis filter, the results can be compared against the ideal binary image. As can be seen in Table 3, the results show that the sensitivity for the method developed in this research is 90.94%, compared with 64.13% for the texture analysis method. Moreover, the 1-specificity is 7.62% compared to 4.86% for the texture analysis method. This shows that the proposed method detects a greater number of false alarms, but at the same time detects a greater number of real flaws.

Table 3. Comparison between the developed filter and the texture analysis filter for the BAM-5 image

Image	Process	Segmented regions	Classified regions	Sensitivity	1-Specificity
Bam5.tif	Bottom-hat segmentation (proposed filter)	495	*(na)*	90.04%	7.62%
Bam5.tif	Texture Analysis [32]	1419	187	64.13%	4.86%

The purpose behind having a smaller number of segmented regions is to reduce processing time for the later AVI stages. The type of region being analyzed is important for the process, and consequently the objective must be to decrease the number of segmented regions while simultaneously having a low number of false alarms. To this end the segmentation process must be as precise

as possible in detecting real defects. These results indicate that the proposed filter has two advantages: the first is a reduction in the number of regions, which implies that during subsequent characteristic-extraction and classification processes, fewer regions need to be analyzed, thereby reducing processing time. The second advantage consists in having a high percentage of regions with real flaws and a low percentage of false alarms. This implies that when the segmentation process is applied, the majority of real flaws are included.

5. Conclusions

One of the main filters used in the detection of hypothetical flaws is the bottom-hat filter. Should the flaw have a minimum of contrast with the background, this filter will not detect it. Noise plays an important role in the result of segmentation because if it is too high, a larger mask must be applied, and thus low intensity flaws disappear. On the other hand, if a smaller mask is used, a large number of regions will be generated which will, for the most part, correspond to noise regions.

The analysis of the results indicates that the proposed filter is sensitive to noise: as noise increases, a larger number of regions is detected, and when noise decreases fewer regions are segmented. This effect can be minimized by using a median filter because of its noise attenuating properties and preservation of edge structure.

This work has been compared with the method developed in [32]. The latter method carries out segmentation by means of the LoG filter, extraction is carried out through the co-occurrence matrix and the Gabor function, and finally the classification process is implemented through the use of the polynomial, Mahalanobis and near-neighbor methods. Clearly, comparison with the proposed filter is not entirely fair, as only the segmentation process has been carried out. However, the results may improve on application of the next stages of AVI such as extraction of characteristics and classification. The results indicate that the sensitivity of the proposed method is 90.94%, compared to 64.13% for the texture analysis method. Also, 1-specificity is 7.62% compared to 4.86%, respectively.

Acknowledgments

The authors wish to thank the Federal Institute for Materials Research and Testing, Berlin (BAM) for the radiographic material used in this research. This work was supported in part by Fondecyt Project Number 1040210.

References

[1] E. R. Davis, *Machine Vision*, Third ed. Amsterdam: Morgan Kaufmann Publishers, 2005.

[2] T. S. Newman and A. K. Jain, "A survey of Automated Visual Inspection," *Computer Vision and Image Understanding,* vol. 61, pp. 231-262, 1995.

[3] G. Mamic and M. Bennamoun, "Automatic flaw detection in textiles using a Neyman-Pearson detector," in *Proc. of the international Conference on Pattern Recognition (ICPR)*, Washington, DC, 2000.

[4] L. W. Zhang, A. Dehghani, Z. W. Su, T. King, B. Greenwood, and M. Levesley, "Real-time Automated visual inspection system for contaminant removal from wool," *Real-Time Imaging,* vol. 11, pp. 257-269, 2005.

[5] T. W. Liao, "Classification of welding flaw types with fuzzy expert systems," *Expert Systems with Applications,* vol. 25, pp. 101-111, 2003.

[6] M. Carrasco and D. Mery, "Segmentation of welding discontinuities using a robust algorithm," *Materials Evaluation,* vol. 62, pp. 1142-1147, 2004.

[7] L.-Y. Lei, X.-J. Zhou, and M.-W. Pan, "Automated Vision Inspection System for the Size Measurement of Workpieces," in *Proceedings of the IEEE Instrumentation and Measurement Technology Conference, IMTC 2005*, Ottawa, Canada, 2005, pp. 872-877.

[8] Y.-K. Chung and K.-H. Kim, "Automated visual inspection system of automobile doors and windows using the adaptive feature extraction," in *Second International Conference on Knowledge-Based Intelligent Electronic Systems, KES 98*, Adelaide, Australia, 1998, pp. 286-293.

[9] D. Mery and D. Filbert, "Automated flaw detection in aluminum castings based on the tracking of potential defects in a radioscopic image sequence," *IEEE Trans. Robotics and Automation,* vol. 18, pp. 890-901, Aug. 21 2002.

[10] M. Carrasco and D. Mery, "Automated Visual Inspection using trifocal analysis in an uncalibrated sequence of images," *Materials Evaluation,* vol. 64, pp. 900-906, September 2006.

[11] V.-D. Nguyen, A. Noble, J. Mundy, J. Janning, and J. Ross, "Exhaustive detection of manufacturing flaws as abnormalities," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1998, pp. 945-952.

[12] T. H. Park, H. J. Kim, and N. Kim, "Path Planning of automated optical inspection machines for PCB assembly systems," *International Journal of Control Automation and Systems,* vol. 4, pp. 96-104, 2006.

[13] F. Pedreschi, D. Mery, F. Mendoza, and J. M. Aguilera, "Classification of potato chips using pattern recognition," *Journal of Food Science,* vol. 69, pp. E264-E270, 2004.

[14] D. Mery and O. Medina, "Automated Visual Inspection of Glass Bottles using Adapted Median Filtering," *Lecture Notes in Computer Science (LNCS),* vol. 3212, pp. 818-825, 2004.

[15] F. Shafait, S. M. Imran, and S. Klette-Matzat, "Fault detection and localization in empty water bottles through machine vision," in *IEEE, Emerging Technology Conference (E-Tech)*, San Diego, CA, 2004, pp. 30-34.

[16] E. Linhari and V. Obac, "Methodology for Classification of Images based on the characteristics of a new interpretation of Shannon's Entropy," *Wood Science Technology,* vol. 39, pp. 113-128, April 2005.

[17] C. Hayes, "The ABC's of nondestructive weld examination," *Weld. J. (Miami),* vol. 76, pp. 46–51, 1997.

[18] W. Al-Nuaimy and O. Zahran, "Time-of-flight diffraction from semi-automatic inspection to semi-automatic interpretation," *Insight,* vol. 47, pp. 639-644, October 2005.

[19] G. Edward, *Inspection of welded joints. ASM handbook, welding, brazing and soldering* vol. 6. OH: ASM International, 1993.

[20] E. P. de Moura, M. H. S. Siqueira, R. R. da Silva, J. M. A. Rebello, and L. P. Calôba, "Welding defect pattern recognition in TOFD signals. Part 1. Linear classifiers," *Insight,* vol. 47, pp. 777-782, December 2005.

[21] B. Blakeley, "Digital Radiography," *Insight - Non-Destructive Testing and Condition Monitoring,* vol. 46, pp. 403-406, July 2004.

[22] E. Deprins, "Computed radiography in NDT applications," *Insight - Non-Destructive Testing and Condition Monitoring,* vol. 46, pp. 590-593, October 2004.

[23] M. Kleber F., H. Silverio L., T. Mezzadri C., and L. Valeria R. de A., "An object detection and recognition system for weld bead extraction from digital radiographs," *Computer Vision and Image Understanding,* vol. 102, pp. 238 - 249, June 2006.

[24] R. González and R. E. Woods, *Digital Image Processing*, 2nd ed.: Prentice Hall Inc., 2002.

[25] M. Carrasco, "Segmentation of welding defects using digital image processing techniques (In Spanish)," in *Departamento de Ingeniería Informática*. vol. Master of Informatics Engineering Santiago, Chile: Universidad de Santiago de Chile, 2004.

[26] A. Gayer, A. Saya, and A. Shiloh, "Automatic recognition of welding defects in Real-Time radiography," *NDT International,* vol. 23, pp. 131–136, June 1990.

[27] S. W. Lawson and G. A. Parker, "Intelligent Segmentation Of Industrial Radiographic Images Using Neural Networks," in *Proc. of SPIE Machine Vision Applications And Systems Integration III*, 1994, pp. 245–255.

[28] M. Sofia and D. Redouane, "Shapes recognition system applied to the non destructive testing," i*n Proceedings Of The 8th European Conference On Non-Destructive Testing (Ecndt 2002)*, Barcelona, Spain, 2002.

[29] R. R. Silva, M. H. S. Siqueira, L. P. Calôba, I. C. Da Silva, A. A. De Carvalho, and J. M. A. Rebello, "Contribution To The Development Of A Radiographic Inspection Automated System," in *Proceedings of The 8th European Conference On Non-Destructive Testing (ECNDT 2002)*, Barcelona, Spain, 2002.

[30] T. W. Liao and Y. Li, "An Automated Radiographic NDT System For Weld Inspection," *NDT&E International,* vol. 31, pp. 183-192, 1998.

[31] G. Wang and T. W. Liao, "Automatic identification of different types of welding defects in radiographic images," *NDT&E International, Elsevier Science,* vol. 35, pp. 519-528, 2002.

[32] D. Mery and M. A. Berti, "Automatic detection of welding defects using texture features," *Insight, X-Ray Image Processing,* vol. 45, pp. 676-681, October 2003.

[33] G. Li, G. Wang, and Z. Jiguang, "A Genetic Algorithm on Welding Seam Image Segmentation," in *IEEE Proceedings of the 5th World Congress on Intelligent Control and Automation*, Hangzhou. P.R. China, 2004.

[34] K. S. Simunic and S. Loncaric, "A genetic search-based partial image matching," in *Proceedings ICIPS'98 IEEE international conference on Intelligent Processing Systems*, 1998, pp. 119-122.

[35] N. Otsu, "A Threshold Selection Method from Gray level Histogram," *IEEE Transactions on System, Man and Cybernetics,* vol. 9, pp. 62-67, 1979.

[36] X. Wang and B. S. Wong, "Segmentation of radiographic images using fuzzy c-means algorithm," *Insight,* vol. 47, pp. 631-633, October 2005.

[37] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory,* vol. 41, pp. 613-627, 1995.

[38] S. G. Chang, Y. Bin, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Transactions on Image Processing,* vol. 9, pp. 1522-1531, 2000.

[39] W. Pedrycz and J. Waletzky, "Fuzzy clustering with partial supervision," *Systems, Man and Cybernetics, Part B, IEEE Transactions on,* vol. 27, pp. 787-795, 1997.

[40] A. Movafeghi, M. H. Kargarnovin, H. Soltanian-Zadeh, M. Taheri, F. Ghasemi, B. Rokrok, K. Edalati, and N. Rastkhah, "Flaw detection improvement of digitized radiographs by morphological transformations," *Insight,* vol. 47, pp. 625-630, 2005.

[41] T. C. Wang and N. B. Karayiannis, "Detection of microcalcifications in digital mammograms using wavelets, " *IEEE Transactions on Medical Imaging,* vol. 17, pp. 498-509, 1998.

[42] K. R. Castleman, *Digital Image Processing*. New Jersey: Prentice-Hall, 1996.

[43] J. Canny, "A Computational Approach To Edge Detection," *IEEE Transactions Pattern Analysis And Machine Intelligence,* vol. 8, pp. 679-698, 1986.

[44] S. Beucher and C. Lantuejoul, "Use of Watersheds in Contour Detection," in *International Workshop on image processing, Real-time edge and Motion detection/estimation*, Rennes, France, 1979, pp. 17-21.

[45] H. Breu, J. Gil, D. Kirkpatrick, and M. Werman, "Linear time Euclidean distance transform algorithms," *IEEE Transactions On Pattern Analysis And Machine Intelligence,* vol. 17, pp. 529-533, May 1995.

[46] L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *Image Processing, IEEE Transactions on,* vol. 2, pp. 176-201, 1993.

[47] S. Beucher, "The Watershed Transformation applied to Image Segmentation," in *10^{th} Pfefferkorn Conf. on Signal and Image Processing in Microscopy and Microanalysis*, 1991, pp. 299-314.

[48] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Second ed. New York: John Wiley & Sons, Inc., 2001.

[49] D. Mery and F. Pesdreschi, "Segmentation of color food images using a robust algorithm," *Journal of food engineering,* vol. 66, pp. 353-360, 2005.

SUBJECT INDEX