New Algorithms for Ultrasonic Non-Destructive Evaluation

Timothy Lardner

Department of Electronic and Electrical Engineering

University of Strathclyde

A thesis submitted for the degree of

Doctor of Philosophy

2016

Copyright

Date:

Signed:

This thesis is dedicated to my loving parents,

Bernadette and David

Acknowledgements

Foremost, I would like to thank my supervisor, Anthony Gachagan, for his advice, encouragement and continued patience.

My thanks extend to all I have worked with throughout this study, in particular Jerzy Dziewierz. The members of the Centre for Ultrasonic Engineering have made my time as a PhD student both enjoyable and fulfilling.

I would like to thank Weidlinger Associates, Royal Dutch Shell, Rolls-Royce, National Nuclear Laboratory, Amec Foster Wheeler, Siemens and E.ON for their contributions to this body of work. This work was funded by the aforementioned companies via the Research Council for Non-Destructive Evaluation and the Engineering and Physical Sciences Research Council.

My gratitude extends to all of my family, and in particular my parents, Roseanne and Patrick, Kathleen and of course Mary O'Neill.

Thanks to Adam Crawford, Michael Ferguson and Gary McLelland for their support and advice, but mostly for the pints.

Finally, I am eternally grateful to my girlfriend Adele who has been at my side throughout this process and without whom the completion of this work certainly would not have been possible.

Abstract

The push for more efficient operation of power generation stations has led to the development of advanced alloys designed to cope with the stresses of running at elevated temperatures. The micro-structure of these new alloys makes the inspection process difficult due to large grains that scatter ultrasonic energy. Aerospace components such as aircraft engine turbine blades are made from similar materials and pose the same difficulties for inspection. In addition, the complex geometries of many of these parts hinder the use of existing advanced imaging methodologies.

The current inspection process involves using both individual transducers and phased arrays to collect pulse-echo data from structures. This process is not sufficient for such difficult materials and a new process must be devised, tested and deployed.

This thesis presents an investigation of new practical techniques to process ultrasonic array data collected via a Full Matrix Capture. Two novel signal processing techniques are presented, evaluated and compared to the Total Focusing Method, which is currently considered as the gold standard in ultrasonic array processing. A study into efficient imaging has also been completed, which involved development of an algorithm to focus upon any point through an arbitrary refracting interface. This algorithm was implemented on a commercially available graphics card and is able to account for a curved interface in real time with no prior knowledge of the surface profile.

Spatially Averaged Sub-Aperture Correlation Imaging splits the full matrix of data into a set of sub-apertures which are imaged independently from each other. These images are then combined into two

sets and are input to a two-dimensional cross-correlation algorithm that outputs a weighting matrix that can be applied to the sum of all images. Signals that are from legitimate reflectors are highly correlated while less-correlated indications are the result of noise from scattering and multi-path propagation. SASACI has been shown to perform well experimentally through inspection of defects within multiple highly scattering welds at a frequency of 5 MHz.

Correlation for Adaptively Focused Imaging aims to correct for anisotropy within difficult materials. The longitudinal velocity within a difficult material can vary with position and using an average velocity does not guarantee a well-focused image. For each pixel in an image, CAFI calculates which samples will be used to calculate the amplitude of the pixel before cross-correlating the signals from adjacent array elements and shifting the delay to the point of maximum focus. This methodology is effective when a small area with a known reflector is being imaged, and for this reason the algorithm is suited to characterisation of reflectors. This technique was experimentally validated on a block of Inconel 625 with a number of side-drilled holes.

Contents

List of Figures

List of Tables

List of Symbols

α	Unit of attenuation $(Np \cdot m^{-1})$
β	Bulk modulus of elasticity (Pa)
Δx	Pixel pitch (m)
Δ	Wave propagation delay (s)
$\dot{u}$	Particle velocity $(m \cdot s^{-1})$
ϵ	Strain
κ	Lamé's first parameter (Pa)
λ	Wavelength (m)
ν	Poisson's ratio
ω	Angular velocity $(rad \cdot s^{-1})$
ψ	Sample index
ρ	Material density $(Kg \cdot m^{-3})$
σ	Stress (Pa)
θ	Angle (rad)
A	Amplitude

a	Acceleration $(m \cdot s^2)$
$A(w)$	Attenuation $(dB \cdot m^{-1})$
A_c	A complex number representing amplitude and phase
A_{cs}	Cross sectional area (m^2)
B	Amplitude
B_D	Beam divergence
B_j	Phase delay
C	Centre of the signal to be cross-correlated
c	Wave frequency (Hz)
D	Delay to be applied for CAFI
d	The number of lags required to maximise p
d_c	The value of d for the centre element(s) of the array
$D_F(\omega)$	Directivity of transducer
E	Young's modulus (Pa)
F	Axial force (N)
G	Modulus of rigidity (Pa)
H	Additional weighting for SASACI if $p(x,y) < T$
h	Full Matrix Capture dataset

k	Wavenumber
m	Mass (Kg)
n_{el}	Number of elements in a given array
P	Pressure (Pa)
p	Cross-correlation coefficient
R	Reflection coefficient
$r_j(x,z)$	Distance to array element j from point (x,z) (m)
RX	The input to a cross-correlation algorithm
S	Number of pixels used for cross correlation
T	Transmission coefficient
t	Time (s)
$TxPx$	A given transmitting element and a specific pixel to be imaged
u	Particle displacement (m)
V	An ideal point reflector
v_L	Longitudinal wave propagation velocity $(m \cdot s^{-1})$
v_S	Shear wave propagation velocity $(m \cdot s^{-1})$
W	Weighting applied to a dataset
w	Transducer width (m)

X	SASACI threshold
x	Distance (m)
x'_j	Distance from array element j to the array centre (m)
Z	Acoustic impedance ($Pa \cdot s \cdot m^{-3}$)

Chapter 1

Introduction

1.1 Motivation for the Work

Non-destructive testing is an essential component of modern engineering. The structural integrity of safety critical components such as aircraft wings, turbine blades and high-pressure pipes must be verified at both the time of manufacture and throughout the life of the component[1]. There is a strong drive in industry to save costs wherever possible while maintaining safety[2].

For ‘next generation’ fossil-fuelled power generation stations, greater efficiencies can be gained through burning fuel at an increased temperature[3, 4]. A greater efficiency of operation leads to reduced fuel costs and less carbon emission[5, 6]. Generating steam at this increased temperature means that the infrastructure will need to be resistant to the potential corrosion that will occur at increased pressures[7, 8]. New ‘superalloys’ have been developed that are suited to environments that will subject the metal to extremes of pressure and heat due to their mechanical strength and resistance to corrosion[9, 10, 11]. These alloys typically have grains that are large enough to interact with the ultrasonic waves used for non-destructive testing[12]. These grains scatter ultrasonic energy and hinder inspection[13]. Aircraft engine turbines also need to be resistant to a large amount of stress and are manufactured from similar alloys that are difficult to inspect ultrasonically. It must be stated that while these advanced alloys are more resistant to the stresses of high pressure operation, flaws do appear in these materials and thus they must be regularly inspected[14, 15, 16].

The materials that 'current generation' power generation stations use in their steam piping pose different inspection challenges compared to new power stations. Many existing power stations are reaching their end-of-life that was designated at the time of manufacture[17]. A large proportion of these stations are undergoing life-extension programmes to ensure their structural integrity for continued operation[18, 19, 20]. These power stations require regular inspections to find any flaws[21]. Potential defects in these structures include voids and creep corrosion, which is small cavities in the micro-structure of a material and is caused by prolonged exposure to high temperatures or pressure[22]. Creep damage spreads slowly throughout a structure and does not generally pose a risk to the structural integrity of a component until the voids reach a critical size[23]. For this reason, it is important to accurately monitor the growth of creep damage and to fastidiously record all defects within a component so that structural engineers can make a decision about whether or not to replace it[24]. In order to detect and size potentially small defects, a sufficiently high interrogation frequency (and therefore small wavelength, given the two properties are inversely proportional) must be used to ensure that these defects will reflect a significant portion of the incident energy[25, 26]. Furthermore, there are a number of welded components with complex geometries that hinder contact inspection.

For the aforementioned cases there are difficulties with conventional ultrasonic inspection. The problem of high rates of attenuation can be overcome by using probes with a larger surface area to receive more energy and higher sensitivity of receivers coupled with low-noise amplifiers[27]. The problem of scattering can be overcome by inspecting components at a lower frequency, but at the expense of resolution. Imaging components through complex geometries can be done via pre-calculation of focal laws. This process is time consuming and does not allow for modifying these focal laws once set. This thesis presents a number of signal-processing-based approaches for overcoming these difficulties without significantly reducing resolution, sensitivity or the time taken to generate results.

1.2 Contributions to Knowledge

- A novel process has been developed that allows rapid generation of images from recorded ultrasonic signals, while accounting for refraction. This process was developed in collaboration with Dziewierz and McGilp.

  My contribution to this work is the idea to split the depth in the imaging volume to a number of discrete points and interpolating for each pixel. I also worked on the optimisation algorithm that calculates the time of flight for a given set of parameters. These times of flight are input to a function, that converts them to a set of polynomial coefficients that are passed to the imaging algorithm. This imaging algorithm was initially developed by Dziewierz for inspecting materials through a flat interface, and was modified to use the aforementioned coefficients. McGilp developed the surface recognition methodology that allows for the arbitrary surface imaging process.

- A new imaging algorithm, Spatially Averaged Sub-Aperture Correlation Imaging (SASACI), was developed to produce images with reduced structural noise compared to standard TFM images. This process was inspired by medical imaging literature where cross-correlation has been successfully used to improve signal-to-noise ratio.

  This technique is novel, as cross-correlation of images from different array sub-apertures to improve TFM-based imaging has not been reported in the literature.

- Correlation for Adaptively Focused Imaging (CAFI) is a second novel methodology that uses cross-correlation to improve ultrasonic imaging of NDE datasets.

  This technique uses cross-correlation to correct focusing in cases where the speed of sound in a material is not well defined. This can occur in anisotropic materials where a ray of sound may not take a straight path through a component, or simply in structures where the speed of sound varies throughout. Using an average velocity will result in a poor focus and an inability to accurately size defects. CAFI is shown to overcome this limitation.

1.3 List of Publications

[1] R. H. Brown, S. G. Pierce, I. Collison, B. Dutton, J. Dziewierz, J. Jackson, **T. Lardner**, C. MacLeod, and M. Morozov, "Automated full matrix capture for industrial processes," in *AIP Conference Proceedings*, vol. 1650, pp. 1967–1976, AIP Publishing, Mar. 2015.

[2] A. McGilp, **T. Lardner**, J. Dziewierz, A. Gachagan, J. Mackersie, and C. Bird, "Inspection of complex components using 2d arrays and TFM," in *53rd Annual Conference of the British Institute of Non-Destructive Testing (NDT 2014)*, (Manchester, UK), Curran Associates, Inc., Jan. 2015.

[3] R. Gongzhang, M. Li, B. Xiao, **T. Lardner**, and A. Gachagan, "Robust frequency diversity based algorithm for clutter noise reduction of ultrasonic signals using multiple sub-spectrum phase coherence," in *Proceedings of the 40th Annual Review of Progress in Quantitative Nondestructive Evaluation: Incorporating the 10th International Conference on Barkhausen Noise and Micromagnetic Testing*, pp. 1948–1955, 2014.

[4] **T. Lardner**, M. Li, and A. Gachagan, "Using phase information to enhance speckle noise reduction in the ultrasonic NDE of coarse grain materials," in *AIP Conference Proceedings*, vol. 1581, pp. 1061–1068, AIP Publishing, Feb. 2014.

[5] A. McGilp, J. Dziewierz, **T. Lardner**, J. Mackersie, and A. Gachagan, "Inspection design using 2d phased array, TFM and cueMAP software," in *AIP Conference Proceedings*, vol. 1581, pp. 65–71, AIP Publishing, Feb. 2014.

[6] J. Dziewierz, **T. Lardner**, and A. Gachagan, "A design methodology for 2d sparse NDE arrays using an efficient implementation of refracted-ray TFM," in *IEEE International Ultrasonics Symposium (IUS)*, pp. 136–138, IEEE, July 2013.

[7] **T. Lardner**, M. Li, R. Gongzhang, and A. Gachagan, "A new speckle noise suppression technique using cross-correlation of array sub-apertures in ultrasonic NDE of coarse grain materials," in *Review of Progress in Quantitative Non-Destructive Evaluation: Volume 32*, vol. 1511, pp. 865–871, Jan. 2013.

[8] R. Gongzhang, M. Li, **T. Lardner**, and A. Gachagan, "Robust defect detection in ultrasonic nondestructive evaluation (NDE) of difficult materials," in *IEEE International Ultrasonics Symposium (IUS)*, pp. 467–470, Oct. 2012.

1.4 Structure of Thesis

This body of work is presented over six chapters. This chapter serves as an introduction to the work and introduces the concept of 'difficult materials' and the problems that arise during their inspection. It puts the thesis into context and outlines the novelties of the work completed.

Chapter two is a review of traditional imaging techniques. It starts with the wave equation being derived from first principles. Well-known laws that govern waves, such as Snell's law and Huygens' principle, are also explained. Basic single-element ultrasonic probes are introduced and the concept of arrays explained. Array imaging is also considered and conventional imaging methodologies are explained. From here, more advanced techniques are discussed. Finally, the concepts of both mathematical and finite element modelling are introduced and the benefits of each discussed.

Chapter three tackles the issue of performing the Total Focusing Method through an interface, accounting for refraction and delivering real-time imaging performance. An interpolation and curve-fitting method is used to calculate propagation times with improved accuracy compared to traditional methods. This methodology is evaluated for both its imaging speed and its ability to create an accurate image through a curved or arbitrarily shaped interface.

Chapter four introduces a novel signal processing technique that can be used to reduce speckle noise in ultrasonic images. The technique operates on the premise that grain noise differs when inspecting from different locations whereas the response from legitimate flaws will remain the same. An existing imaging process is used as a benchmark and both the methodology and results are compared to that of the Total Focusing Method. The signal-to-noise ratio is used as a metric of performance and the proposed technique shows improvement over the benchmark.

Chapter five expands on the premise of using cross-correlation to improve ultrasonic imaging within difficult materials. A technique is proposed that can compensate for a varying wave propagation velocity within a material. The problem, which occurs in anisotropic materials, is explained and the proposed solution thoroughly explored. The solution is verified experimentally with an industrially relevant sample and results show the proposed technique is able to improve on currently used imaging methods.

Chapter six discusses the results of the thesis as a whole. The findings from each of the three novel methodologies introduced in this thesis are summarised and potential improvements to each are evaluated. Finally, potential areas of future work are discussed.

Chapter 2

Review of Conventional Array Imaging Approaches

2.1 Mechanical Wave Propagation

To understand how a wave propagates through an elastic medium, first consider a system that consists of a mass connected to a spring, as shown in Figure 2.1. If the mass is moved from its resting state and then released, restoring forces will act upon the system until it returns to equilibrium, and these will follow the laws of harmonic oscillation. During this oscillation, potential energy in the spring is transferred to kinetic energy of the mass and vice versa. Some of this energy will be lost as it is converted to heat through friction. The mechanical properties of the mass and the spring will determine the frequency of this damped oscillation[28].

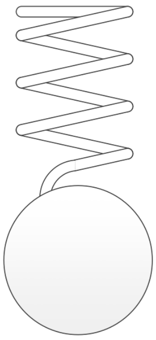


Figure 2.1: *A mass connected to a spring*

Now consider the same system, but instead of the system being moved out of its resting state and then released, the system will be continually excited by a sinusoidally varying wave. In this case, the oscillation of the spring and the mass will match the frequency of the driving force and will maintain its amplitude while the excitation force exists.

An elastic medium can be considered as a network of ‘masses’ (molecules) connected to each other by ‘springs’ (elastic binding forces). If this system is excited by a sinusoidal wave, all of the ‘masses’ in the system will oscillate with the same frequency. The ‘springs’ will transfer the motion to each ‘mass’ and will also introduce a delay as the kinetic energy is being transferred. This delay is known as propagation delay and is one of the fundamental and constant laws that bound wave mechanics[29].

In an infinite solid or elastic medium, or one so large that its boundaries can be ignored, there are two kinds of stress that the medium can undergo. Therefore there are two methods of wave propagation that are possible.

The first are longitudinal waves, named as such because the wave propagates in the same direction as the particle motion. They have the greatest wave propagation velocity in any material. The velocity of a longitudinal wave, v_L, can be calculated using Equations 2.1 and 2.2 where κ is Lamé's constant, G is the modulus of rigidity, ρ is the material density and β is the bulk modulus of elasticity for a fluid[30].

$$v_L = \sqrt{\frac{\kappa + 2G}{\rho}} \tag{2.1}$$

$$v_L = \sqrt{\frac{\beta}{\rho}} \tag{2.2}$$

For a homogeneous isotropic solid with two independent material constants, the modulus of rigidity, G, also known as the shear modulus, can be calculated using Equation 2.3[31] where E is the Young's modulus (the measure of the stiffness of an elastic material) and ν is Poisson's ratio (the ratio of transverse strain to longitudinal strain)[32].

$$G = \frac{E}{2(1 + \nu)} \tag{2.3}$$

The second method of wave propagation, shear waves, propagate in a direction perpendicular to particle motion and can only exist in materials with shear elasticity (i.e. solids). The shear wave velocity, v_S, can be calculated using Equation 2.4.

$$v_S = \sqrt{\frac{G}{\rho}} \tag{2.4}$$

2.1.1 The Wave Equation

Wave propagation has been introduced and explained. Propagation can be represented in a mathematical way so that it can be analysed and problems involving mechanical waves can be solved.

Consider a one-dimensional model of longitudinal waves in an elastic bar. Let the distance along the bar equal to x, the time equal to t, the cross sectional area equal to A_{cs}, the particle displacement equal to $u(x,t)$ and the axial force equal to $F(x,t)$[33].

At any instance in time, a small element in the bar can be represented by a forces diagram, as shown in Figure 2.2.

From Figure 2.2, the net force to the right can be observed. It is expressed in Equation 2.5.

$$(F + \delta F) - F = \delta F \tag{2.5}$$

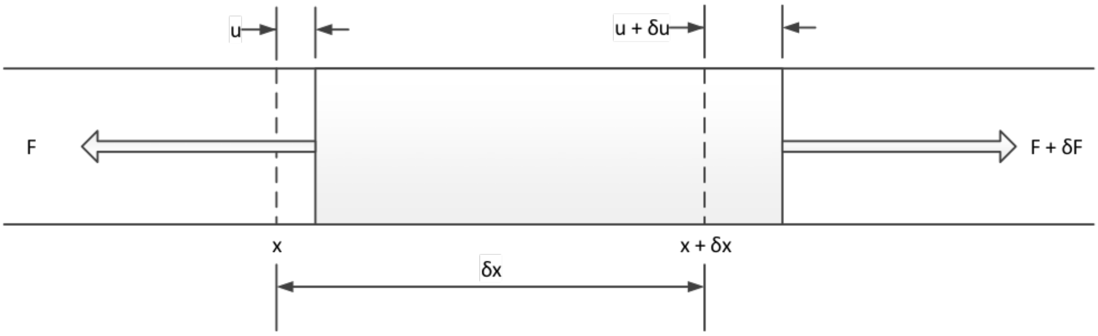


Figure 2.2: *A 'forces diagram' of a one-dimensional bar*

From Newtons second law, $F = ma$, Equation 2.6 can be derived.

$$\delta F = \rho \frac{\delta^2 u}{\delta t^2}(A_{cs}\delta x) \tag{2.6}$$

Hooke's law states that the elasticity in a bar can be calculated using Equation 2.7 where σ is stress, E is Young's modulus and ϵ is strain[34].

$$\sigma = E\epsilon \tag{2.7}$$

The strain displacement relationship is given in Equation 2.8.

$$\epsilon = \frac{\delta u}{\delta x} \tag{2.8}$$

Hence, Equation 2.9 can be derived.

$$F = A_{cs}\sigma = A_{cs}E\epsilon = A_{cs}E\frac{\delta u}{\delta x} \tag{2.9}$$

Knowing that δF must equal $\frac{\delta F}{\delta x}\delta x$, Equation 2.10 can be derived from Equations 2.9 and 2.6.

$$\frac{\delta}{\delta x}(A_{cs}E\frac{\delta u}{\delta x})\delta x = \rho\frac{\delta^2 u}{\delta t^2}(A_{cs}\delta x) \tag{2.10}$$

Equation 2.10 can be simplified into the expression shown in Equation 2.11.

$$E\frac{\delta^2 u}{\delta x^2} = \rho\frac{\delta^2 u}{\delta t^2} \tag{2.11}$$

If $\frac{E}{\rho} = {v_L}^2$ and given Equation 2.11, the formula for a one-dimensional wave can be derived. It is represented in Equation 2.12.

$$\frac{\delta^2 u}{\delta x^2} = \frac{1}{{v_L}^2}\frac{\delta^2 u}{\delta t^2} \tag{2.12}$$

The one-dimensional wave equation shows the classic feature of any wave equation: the second derivative with respect to time on one side, and the second derivative with respect to space on the other[35, 36].

To solve the wave equation, a displacement field must be found that satisfies the wave equation itself, as well as the appropriate boundary conditions. Analytically, it is generally impossible to solve this for both conditions, but if an infinite material is considered (i.e. a material with no boundaries), a basic solution can be found.

Equation 2.13 shows a possible solution for the wave equation, where A and B are unknown constants representing the amplitude of forward and backwards propagating waves respectively, f is wave frequency, and ω is the angular velocity calculated by $\omega = 2\pi \cdot f$.

$$u(x,t) = A\cos\omega(\frac{x}{c} - t) + B\cos\omega(-\frac{x}{c} - t) \tag{2.13}$$

Equation 2.13 can be analysed further, given that the wavenumber can be calculated using $k = \frac{\omega}{c}$, and the solution as shown in Equation 2.14 can be derived.

$$u(x,t) = A\cos(kx - \omega t) + B\cos(-kx - \omega t) \tag{2.14}$$

The complex exponential equation, derived from Euler's work is shown in Equation 2.15.

$$e^{i\theta} = \cos\theta + i\sin\theta \tag{2.15}$$

Only the real part of this will exist physically, and therefore Equation 2.14 can be simplified and written as shown in Equation 2.16.

$$u(x,t) = Re[Ae^{i(kx-\omega t)} + Be^{i(-kx-\omega t)}] \tag{2.16}$$

Since the subject of Equation 2.16 is u, the term for displacement, the previous derivations can be used to rearrange this equation for strain.

$$\epsilon = \frac{\delta u}{\delta x} = \frac{\delta}{\delta x} A e^{i(kx-\omega t)} = iku \tag{2.17}$$

Using Hooke's law, shown in Equation 2.7, the stress-strain relationship can be written as shown in Equation 2.18.

$$\sigma = E\epsilon = ikEu \tag{2.18}$$

The particle velocity is the derivative of the particle displacement with respect to time.

$$\dot{u} = \frac{\delta u}{\delta t} = \frac{\delta}{\delta t} A e^{i(kx-\omega t)} = -i\omega u \tag{2.19}$$

The ratio between the particle velocity, $\dot{u}$ and σ is the acoustic impedance, Z. Acoustic impedance is the key factor in determining energy transfer from one medium to another.

$$Z = \frac{\sigma}{\dot{u}} \tag{2.20}$$

The wave equation is linear and therefore superposition can be applied. The principle of superposition will be explored in greater depth in Section 2.2.2. Solutions to physical problems are found by combining a number of simple solutions to satisfy the boundary conditions of the problem.

If a boundary condition is set such that the conceptual bar ends at $x = 0$, the one-dimensional wave equation can be solved for this. From Equations 2.16 and 2.18 an expression can be formulated to calculate the stress at any point in the bar (Equation 2.21).

$$\sigma(x,t) = ikEAe^{i(kx-\omega t)} + ikEBe^{i(-kx-\omega t)} \tag{2.21}$$

The stress at $x = 0$ is shown in Equation 2.22:

$$\sigma(0,t) = ik(A - B)e^{-i\omega t} \tag{2.22}$$

Hence $A = B$. The final solution in terms of displacement is written in Equation 2.23[37].

$$u(x,t) = A(e^{i(kx-\omega t)} + e^{i(-kx-\omega t)}) \tag{2.23}$$

If Equation 2.23 is considered, it can be seen that there are two waves of equal amplitude travelling in opposite directions. One of the waves is the reflection from the end of the bar. Equation 2.23 can be simplified to show that a standing wave will exist in this bar.

$$u(x,t) = 2A\cos(kx)e^{-i\omega t} \tag{2.24}$$

If the waves in the bar are considered, a standing wave is expected if a wave is being reflected from a perfect reflector (i.e. no absorption) as the reflected wave will travel backwards exactly 180° out of phase.

2.1.2 Waves at Boundaries

In ultrasonic testing, pulse echo inspection is a common method for attempting to identify flaws in materials. Reflections, and therefore echoes of incident waves, occur when a wave reaches a boundary. An ultrasonic transducer will transmit a pulse and then detect any incident waves, i.e. reflections from the transmitted signal. The signals received by the transducer can then be analysed to determine the locations of boundaries.

Consider a scenario where a longitudinal wave is travelling in a first medium towards a boundary beyond which there is a second medium. This boundary can be considered as planar and of infinite length. Equations exist to calculate the amplitude of a reflected signal given that the amplitude of the incident signal, as well as the acoustic impedance in both materials, is known.

Using the formula derived in Equation 2.23, an incident wave can be represented by the equation written in 2.25.

$$u_i = A_i e^{i(k_1 x - \omega t)} \tag{2.25}$$

The reflected wave can be written as shown in Equation 2.26.

$$u_r = A_r e^{i(k_1 x - \omega t)} \tag{2.26}$$

The wave transmitted into the second medium can be represented by Equation 2.27.

$$u_t = A_t e^{i(k_2 x - \omega t)} \tag{2.27}$$

It must be noted the wavenumber of the transmitted wave changes, as the propagation velocity is different in the new material. For Equations 2.25 to 2.27, k_1 and k_2 represent the wavenumber of the longitudinal wave in the first medium and the second medium respectively. For the boundary between the two materials, the expression shown in Equation 2.28 is used to relate the amplitude between each of the waves.

$$A_i + A_r = A_t \tag{2.28}$$

The pressures acting upon the boundary are also continuous and therefore Equation 2.29 is also true.

$$P_i - P_r = P_t \tag{2.29}$$

Taking Equations 2.30 and 2.31 into account, Equation 2.32 can be derived.

$$v_L = \sqrt{\frac{E}{P}} \tag{2.30}$$

$$E = P{v_L}^2 \tag{2.31}$$

$$Ek = P{v_L}^2 \frac{\omega}{v_L} = P v_L \omega = Z\omega \tag{2.32}$$

The stresses in the system can be represented by the expressions in Equations 2.33 to 2.35.

$$P_i = iE_1 k_1 A_i e^{i(-\omega t)} \tag{2.33}$$

$$P_r = iE_1 k_1 A_r e^{i(-\omega t)} \tag{2.34}$$

$$P_t = iE_2 k_2 A_t e^{i(-\omega t)} \tag{2.35}$$

Equation 2.36 can be derived from this.

$$Z_1 A_i - Z_1 A_r = Z_2 A_t \tag{2.36}$$

Combining this with the previous derivations to cancel out A_t, Equation 2.37 can be derived, which is the definition of a reflection coefficient.

$$R = \frac{A_r}{A_i} = \frac{Z_1 - Z_2}{Z_1 + Z_2} \tag{2.37}$$

Equation 2.37 is the proportion of the wave reflection from a boundary, which is commonly written as the reflection coefficient, R.

The transmission coefficient can also be calculated from this and is given in Equation 2.38.

$$T = \frac{A_t}{A_i} = \frac{2Z_1}{Z_1 + Z_2} \tag{2.38}$$

Equation 2.39 must be true due to conservation of energy.

$$T = R + 1 \tag{2.39}$$

If a second medium is completely rigid, it will not support any particle motion. All of the energy must be reflected. $T = 0$ and therefore R must be equal to -1. If the second medium has exactly the same acoustic impedance as the first, 100% of the wave will be transmitted. In this case $T = 1$ and $R = 0$, meaning that there will be no reflection from this interface.

If the incident wave reaches the boundary, propagating in a direction perpendicular to the boundary, then the above equations are the only ones needed to calculate what will occur. The transmitted wave will continue in a direction perpendicular to the boundary. If, however, the wave reaches the boundary at an angle, refraction will occur.

Refraction is calculated using Snell's law[38]. Reflections are explained first as they the most simple interaction between a wave and an interface.

The angle of incidence is equal to the angle of reflection.

$$\theta_i = \theta_r \tag{2.40}$$

Equation 2.40 is true for both longitudinal and shear waves. For transmission, the angle can be calculated using the expression written in Equation 2.41. Note that L stands for longitudinal, and S for shear waves.

$$\frac{v_{L1}}{\sin\theta_{L1}} = \frac{v_{S1}}{\sin\theta_{S1}} = \frac{v_{L2}}{\sin\theta_{L2}} = \frac{v_{S2}}{\sin\theta_{S2}} \tag{2.41}$$

Figure 2.3 provides a visual representation of the refraction that occurs at a material boundary. The subscripts $_I$, $_R$ and $_T$ relate to incident, reflected and transmitted waves respectively.

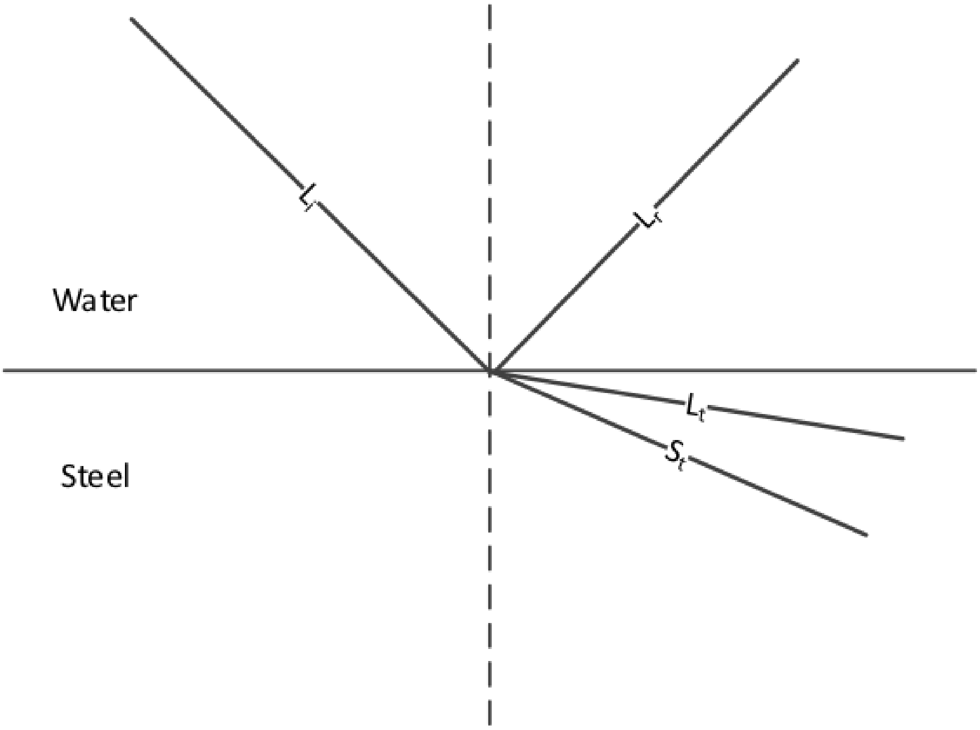


Figure 2.3: *An illustration of Snell's law*

When the θ_2 reaches 90°, transmission does not occur. The incident angle at which it will happen is known as the critical angle. This is important for non-destructive testing applications as if there is no longitudinal wave transmitted (and assuming that the material is a solid) only a shear wave will be transmitted. Since there is only propagation in one mode, there are many less unwanted reflections. Figure 2.4 shows what happens at the first critical angle with respect to longitudinal and shear waves; only the shear wave is transmitted.

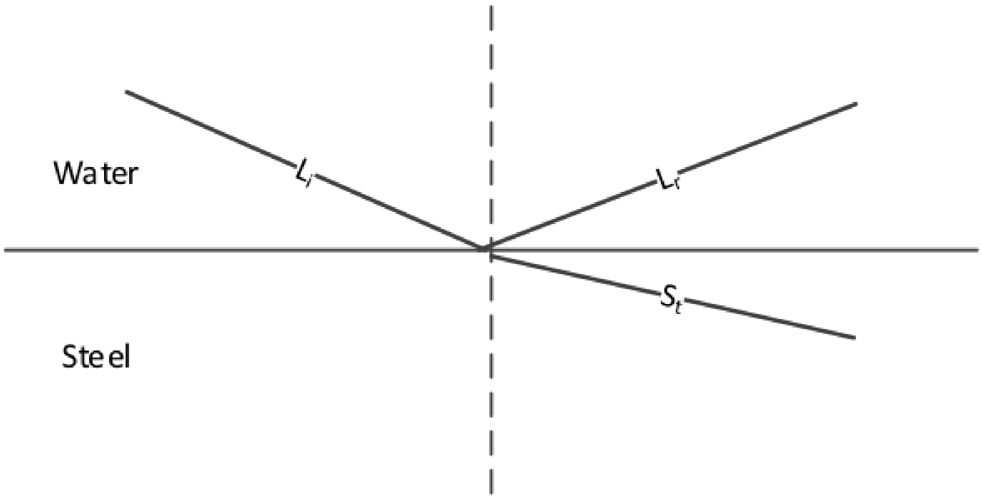


Figure 2.4: *The first critical angle*

Equation 2.41 can be rewritten to calculate the critical angle for longitudinal waves.

$$\frac{v_{L1}}{\sin\theta_{L1}} = \frac{v_{L2}}{\sin 90} \tag{2.42}$$

The equation can then be rearranged for θ_{L1}, written in Equation 2.43.

$$\theta_{L1} = \sin^{-1}\frac{v_{L1}}{v_{L2}} \tag{2.43}$$

To find the second critical angle, when no shear or longitudinal waves are transmitted, Equation 2.41 can be rearranged in the following way:

$$\frac{v_{S2}}{\sin 90} = \frac{v_{L1}}{\sin\theta_{L1}} \tag{2.44}$$

$$\theta_{L1} = \sin^{-1} \frac{v_{L1}}{v_{S2}} \tag{2.45}$$

Figure 2.5 illustrates what happens at the second critical angle where no energy is transferred from one medium to the other.

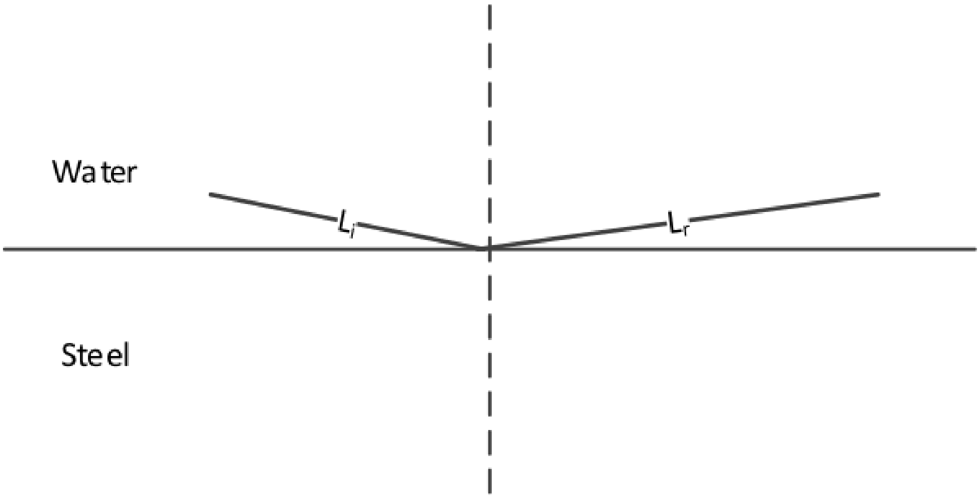


Figure 2.5: *The second critical angle*

2.2 Ultrasonic Devices

2.2.1 Ultrasound Transducers

The definition of the word ‘transducer’ is a device that converts one physical quantity into another[39].

An electromechanical ultrasonic transducer is an electromechanical device that converts an electrical voltage into a mechanical pressure wave or vice versa. To be classed as ultrasound the frequency of the pressure wave must be greater than the range of human hearing, which is around 20 kHz[40].

There are a number of different types of transducer, each using a different method of converting the electrical signal into mechanical motion. The most common type of transducer incorporates a piezoelectric material, which will be discussed in more detail, but there are also electrostatic, magnetostrictive and moving coil transducers. These transducers are not limited to ultrasound, but for mechanical waves in general. The moving coil transducer is similar to a commercial loudspeaker and has been used in the study of ultrasonic absorption in gases[41]. The electrostatic transducer has been used to induce ultrasonic waves in air[42], but is more commonly known for driving high-end loudspeakers[43]. Magnetostrictive transducers rely on the property of some materials where they deform when magnetised[44].

Piezoelectric transducers are manufactured from materials that exhibit the piezoelectric effect. When a piezoelectric plate is deformed, a voltage develops between the two faces of the plate. Conversely, if a voltage is applied between two faces of a piezoelectric plate, a deformation will occur. This is known as the inverse piezoelectric effect[45, 46].

Commercial piezoelectric ultrasonic transducers are manufactured with a piezoceramic material at their core. Figure 2.6 shows an annotated drawing of an ultrasound transducer.

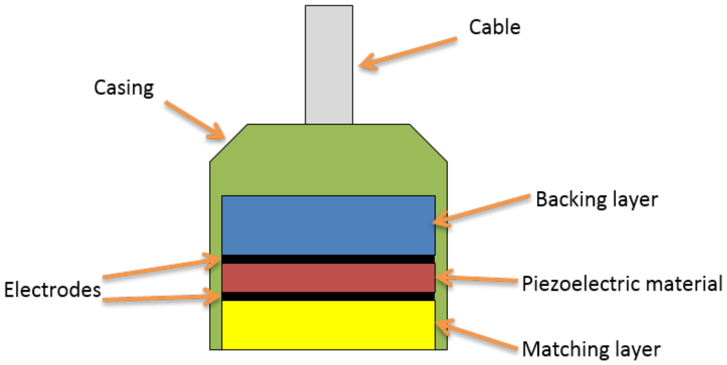


Figure 2.6: *An annotated drawing of an ultrasound transducer*

In Figure 2.6, the piezoelectric material can be seen in the middle of the transducer, with electrodes on each face. Electrical wires connect the electrodes to the cable and to external instrumentation. It is the piezoelectric material that deforms when a voltage is applied between the electrodes[47]. The backing layer serves to absorb any energy that is not output from the front of the device. This is necessary to avoid excess reverberation within the piezoelectric material, which would potentially inhibit inspection due to the reverberations inducing additional voltages in the piezoelectric material. These signals would be considered coherent noise. The matching layer serves to better match the acoustic impedance of the inspection medium to the piezoelectric plate itself. This will be discussed in more depth in Section 2.2.5. The cable is required to connect the device to instrumentation, and the case protects the device as it is often mechanically sensitive. These two components also serve to shield the transducer from electrical interference. A wear layer, not depicted in the figure, is also present and protects the transducer from damages during general use. It also ensures that the device is waterproof, if appropriate for the desired application.

2.2.2 Huygens' Principle

A problem still to be considered is that of superposition. It has been mentioned briefly and can be used to simplify the solving of the wave equation. Superposition will now be considered in greater depth.

The wave equation is linear and this means that if A and B are solutions, $A+B$ is also a solution[48]. This is the basis for many mathematical methods that consider wave fields and is can be used to explain how a group of transducers, known hereafter as an 'array', can be used to manipulate a beam.

Huygens, a prominent Dutch mathematician, discovered wave superposition when he was attempting to determine, given that a wavefront is known, where subsequent wavefronts would occur. He proved his theory graphically and it was later explained mathematically by considering the known wavefront to be an infinite number of point sources[49].

The principle of superposition can be used to calculate the wave field from a set of transducers. If a two dimensional image is considered, with length x and width z, the pressure from a single source can be represented using Equation 2.46 where r_j is the distance from the source, $P_j(x, z)$ is the pressure exerted by source, j, at a given point and A_c is a complex number representing the amplitude and phase of the source[27]. The $\frac{1}{\sqrt{r}}$ term is necessary for conservation of energy for a cylindrical wave coming from a line source. A line source has a similar profile to an element of an ultrasonic array. It should be noted that this equation is an approximation and is not valid in the near field (calculated by $\frac{D^2}{4\lambda}$ where D is the source width) due to the rapid fluctuations of amplitude that occur in this region.

$$P_j(x,z) = A_c \frac{1}{\sqrt{r_j}} e^{i(kr_j - \omega t)} \tag{2.46}$$

This can be summed for all point sources to get the total wave field with all sources considered. In Equation 2.47, n is the number of sources in the system.

$$P(x,z) = \sum_{j=1}^{n} A_c \frac{1}{\sqrt{r_j}} e^{i(kr_j - \omega t)} \tag{2.47}$$

For an array of point sources, Equation 2.47 can be written as an integral, where r' is the distance from the point (x, z) to the point on the array, that is distance x' from the centre. a is the width of the transducer.

$$P(x,z) = \int_{-\frac{a}{2}}^{\frac{a}{2}} A_c \frac{1}{\sqrt{r'}} e^{i(kr' - \omega t)} dx' \tag{2.48}$$

Equation 2.48 can often be solved by adopting assumptions, making the solution only valid in the far field, but greatly simplifying it to the point where it can be solved analytically. Equations 2.49 to 2.54 show a solution for the field of a transducer, valid only for the far field.

The integral in Equation 2.48 can be expressed in terms of polar co-ordinates where $r' = \sqrt{R^2 + x'^2 - 2Rx'\cos(90 - \phi)}$. R is the distance from the array centre and ϕ is the angle of the array centre with respect to the normal.

$$P(R, \phi) = \int_{-\frac{a}{2}}^{\frac{a}{2}} \frac{1}{\sqrt{r'}} e^{i(kr' - \omega t)} dx' \tag{2.49}$$

It is also assumed that P is in the far field and it is assumed that R is much greater than x' and so, due to these assumptions, the following is true.

$$r' \approx R - x' \sin\phi \tag{2.50}$$

$$\frac{1}{\sqrt{r'}} \approx \frac{1}{\sqrt{R}} \tag{2.51}$$

The integral can then be rewritten taking these assumptions into account.

$$P(R, \phi) \approx \int_{-\frac{a}{2}}^{\frac{a}{2}} \frac{1}{\sqrt{R}} e^{i(k(R - x' \sin\phi) - \omega t)} dx' \tag{2.52}$$

$$P(R, \phi) \approx \frac{1}{\sqrt{R}} e^{i(kR - \omega t)} \int_{-\frac{a}{2}}^{\frac{a}{2}} e^{ikx' \sin\phi} dx' \tag{2.53}$$

If Equation 2.53 is integrated, the final equation for the field from a signal rectangular element in the far field is derived.

$$P(R,\phi) \approx \frac{a}{\sqrt{R}} e^{i(kR-\omega t)} \frac{sin(\frac{1}{2}ka\sin\phi)}{\frac{1}{2}ka\sin\phi} \tag{2.54}$$

Huygens' principle has shown that a line source can be integrated across and a solution derived for the pressure field of this transducer. An array of ultrasonic transducers can be considered as a group of sources and thus the pressure field from an array can be calculated using the above methodology. The field from each element in the array can be determined individually and then summed to find the field for the full array.

The directivity function defines how the pressure in a field varies with the angle from the transducer. The directivity of an element is a function of the width of the element and the wavelength of the emitted wave.

$$D_F(\phi) = \frac{\sin(\frac{1}{2}ka\sin\phi)}{\frac{1}{2}ka\sin\phi} = \mathrm{sinc}(\frac{1}{2}ka\sin\phi) = \mathrm{sinc}(\frac{\pi a\sin\phi}{\lambda}) \tag{2.55}$$

The delay required from each element in a transducer array required to achieve focus at any point in the far field of the image can be calculated. The centre element is chosen to have a delay of 0 and the relative time delays for each other element can be calculated.

$$t_j = \frac{d_j - d_0}{v} \tag{2.56}$$

Equation 2.56 shows a simple equation to calculate each element's time delay, where v is the velocity in the medium, d_j is the distance to element j from the focal point, and d_0 is the distance to the reference element (the centre element, in this case). This is then converted to a phase delay, B_j, so that the term can be used in the frequency domain and applied as a transfer function.

$$B_j = e^{i\omega t_j} \tag{2.57}$$

With this knowledge, an equation can be written to calculate the wave field in the far field from an array element, complete with beam focusing and directivity.

$$P(x,z) = \sum_{j=1}^{n} B_j D_f(\phi_j) \frac{1}{\sqrt{r_j}} e^{i(kr_j - \omega t)} \tag{2.58}$$

Figure 2.7 shows the wave field of a 32 element array employing beam steering. This pressure field was computed using a software implementation of Equations 2.55 to 2.58. Note that the pressure field from the beam is off-centre due to the beam steering, and also the complexity of the field within the 'near-field' which shows graphically why simple assumptions cannot be made for modelling the near-field.

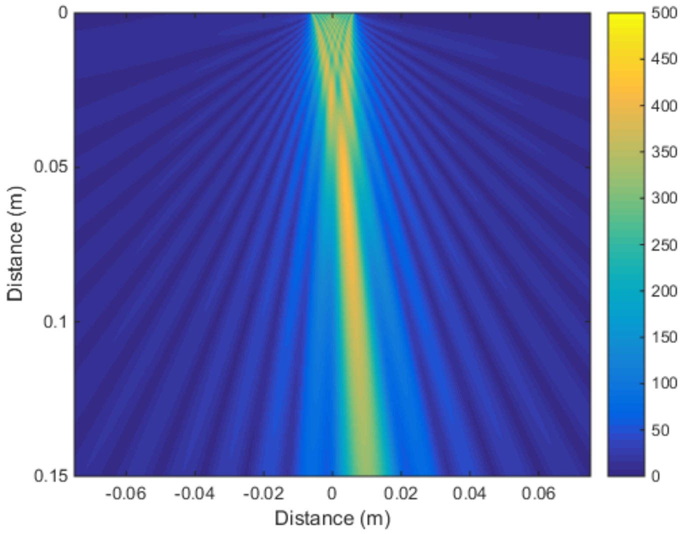


Figure 2.7: *A wave field, calculated programmatically using an implementation of the theory in this section. Shown with a linear scale.*

2.2.3 Ultrasound Arrays

An ultrasonic array can be simply defined as an arrangement of two or more ultrasonic transducers[50]. Each transducer is referred to as an element of the array. Each element in the array can be driven independently of, or along with, any other element. This allows for phased array ultrasonic testing. Arrays are available in either 1D, 2D or annular configurations. Figures 2.8 to 2.10 show examples of element layouts for each configuration. Additionally, a photograph of a 2D commercial array (Imasonic, France) is shown in Figure 2.11.

In this thesis, a commercial linear array (Vermon, France) was used for all experimentation. The parameters of this probe are shown in Table 2.1.

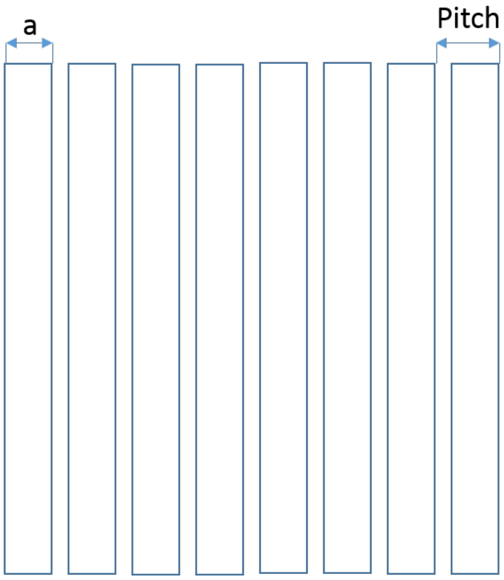


Figure 2.8: *An example element layout for a 1D array*

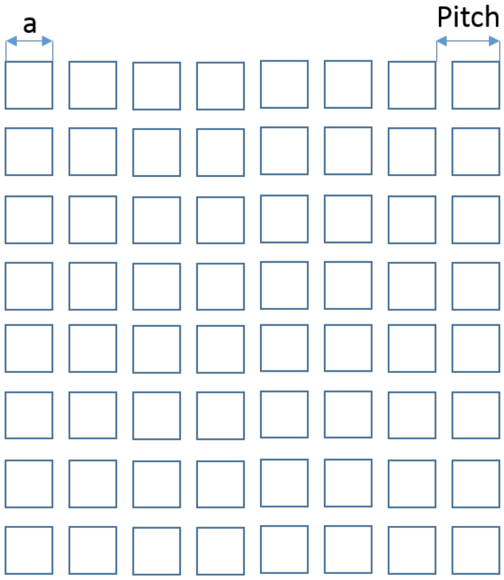


Figure 2.9: *An example element layout for a 2D array*

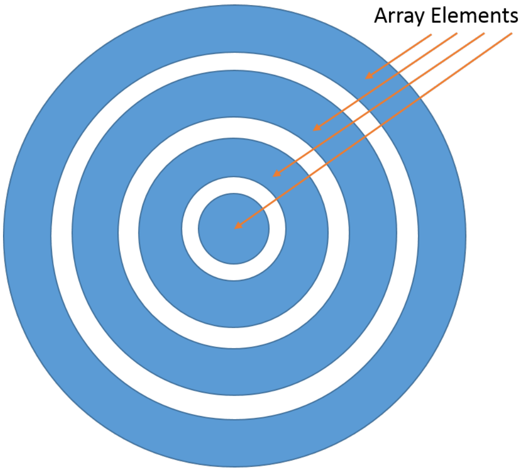


Figure 2.10: *An example element layout for an annular array*

Figure 2.11: *A commercial two-dimensonal array*

Parameter	**Value**
Centre Frequency	5 MHz
Element Count	128
Pitch	0.7 mm
Layout	Linear

Table 2.1: *Vermon array parameters for step wedge inspection*

2.2.4 Ultrasound Arrays Versus Single Element Probes

Single element probes exert a pressure load onto the inspection medium, creating a mechanical wave that will disperse in the medium. The probe will pick up any received waves and convert them to electrical energy which can then be manipulated for observation. Due to the wave propagating spherically, reflectors do not necessarily have to be directly under the probe[51]. This makes it difficult to locate a defect within a medium if there are a number of off-axis reflectors[52]. This can be partly solved by using a directional probe. In this case the surface of the probe is slightly concave and will have a depth at which all the emitted energy is focused at one point[50]. The energy from off-axis signals is reduced and one can be more certain that any reflections are within the expected wave path of the probe[53]. With each of these methods the standard way to interpret this data is to view or process the one dimensional amplitude signal in the time domain, also known as an A-scan[54].

The benefit of using an array for these problems is that the focus is not fixed and can be changed on the fly, during an inspection. The wavefront can be dynamically steered, focused, or both. Furthermore, the use of Full Matrix Capture allows a user to generate datasets and apply signal processing techniques in post-processing. The use of phased array also allows for beam steering, meaning a larger area can be inspected in the same amount of time when compared to single element inspection.

There are few downsides to using an array over a single element probe. The first is cost. To exploit ultrasonic arrays to their full potential, a Phased Array Controller (PAC) is required. These range in cost from thousands of pounds to hundreds of thousands of pounds. The connectors for these arrays are also expensive. While almost all single element probes use standard co-axial connectors, commercial arrays will use one of a number of proprietary connectors.

Most PACs are bundled with the manufacturer's software which is used to drive the array. It is often difficult to export the data from this software into a file that can be opened by other programs. This is in contrast to a single element probe that can be operated with a simple signal generator and an oscilloscope.

The final drawback of arrays is the typical element size. Arrays need to be of a size and weight such that operators can use them comfortably. It means that a large number of elements are required to fit in a comparatively small space, leading to a small element size. The narrow spacing is also necessary to avoid grating lobes which occur when the element spacing is greater than half the wavelength in a periodic array. This has a direct effect on how much energy each element can impart to the inspection medium, as well as the sensitivity of the element.

2.2.5 Key Array Parameters for Design and Inspection

Similar to single element inspection, an appropriate ultrasonic array must be selected for a given inspection. Some parameters are more important than others depending on the type of inspection.

The majority of commercial arrays for NDE are either linear (one-dimensional) or matrix (two-dimensional), and are periodic, meaning that the element spacing is the same for each element. 1D arrays have larger element sizes leading to an increased sensitivity and it is generally simpler to create focal laws for these devices. 2D arrays often offer a complex arrangement of elements and can used for volumetric imaging. Annular arrays are used when a variable focus is required, and can be used in place of single element transducers that are focused using hardware[27].

‘Sparse’ layouts are also available, where the elements are more spread out than a standard probe and may not be spaced linearly. Care must be taken when choosing a sparse array as if an regular, periodic array has an element spacing greater than $\frac{\lambda}{2}$ (where λ is the wavelength) grating lobes will be present[55]. Energy distribution patterns, or lobes, will be discussed in greater detail in Section 2.3.1. 2D arrays allow for a complex layout of elements such as the one shown in Figure 2.12, which is the element layout of the array shown in the photograph in Figure 2.11. These complex array layouts can allow for a greater sensitivity over a larger area. The amplitude of sidelobes can often be reduced in post-processing[56].

The array must also be properly matched to the inspection medium. As with high frequency electromagnetic signals transmitted along a wire, a significant impedance mismatch between two media can result in the majority of the energy being reflected at the boundary and little transfer of energy into the sample itself. This can be approached in two ways. Arrays generally have a matching layer, which is an intermediate step between the acoustic impedance of the device and of the target material. In order to ensure an efficient transfer of energy, matching

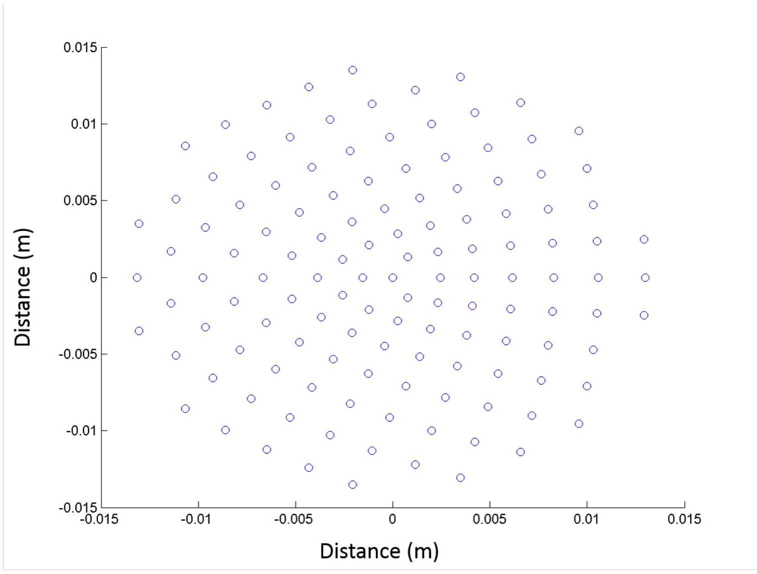


Figure 2.12: *An example of element locations on a 128 element sparse array*

layers are often complex and include multiple layers of materials. Wedges, which can be used to steer a beam, can also be used as a tool for matching. It is common for commercial arrays to be matched to Rexolite[57] or other such wedge materials.

The centre frequency of a probe is a key parameter in determining the performance of an ultrasonic testing system. High frequency waves are attenuated to a greater extent and interact with much smaller particles within a medium of inspection when compared to waves of a lower frequency. Difficult materials often have grains that are of a size that interferes with the propagation of high frequency ultrasonic signals. The wavelength of the wave determines the size of particles that will interact with the wave. Generally, grains or particles with a size greater than half-wavelength ($\frac{\lambda}{2}$) will interact with incident waves. This interaction is desirable when attempting to locate small defects. The drawback of using a lower frequency probe is twofold. Lower frequencies lead to a lower spatial resolution as small features will not be resolved[58]. Lower frequency probes take a longer time to reach an equilibrium after an excitation. The time taken to reach this equilibrium after the initial excitation is known as the ringdown time. While the transducer is returning to a steady state after the excitation, any energy received will not appear as a significant contribution to the energy within the probe. Hence, a long ringdown time means that areas close to the transducer cannot be inspected using conventional methods. Figure 2.13 shows an A-scan of a low frequency probe emitting a signal into a Rexolite sample. The long ringdown time of this probe can be clearly observed in the first $20\mu s$ of this image, where saturation can also be observed. Problems with long ringdown times can be overcome using a stand-off, but unless the stand-off has the same acoustic impedance and ultrasound propagation velocity as the sample, reflections will occur and refractions will add complexity to the focal law generation.

All ultrasonic transducers have a finite frequency response giving rise to an effective bandwidth. The bandwidth is measured relative to its maximum amplitude response in the frequency domain and a transducer is considered responsive until the magnitude of the response falls below half of the maximum value. A typical frequency response curve for a commercial array is shown in Figure 2.14. Bandwidth can be measured in Hertz or as a fraction of the centre frequency of the array. Commercial probes have a fractional bandwidth ranging between 30% to over 100%[59]. Bandwidth becomes important for advanced post-processing where the frequency domain signal is manipulated in order to remove noise and

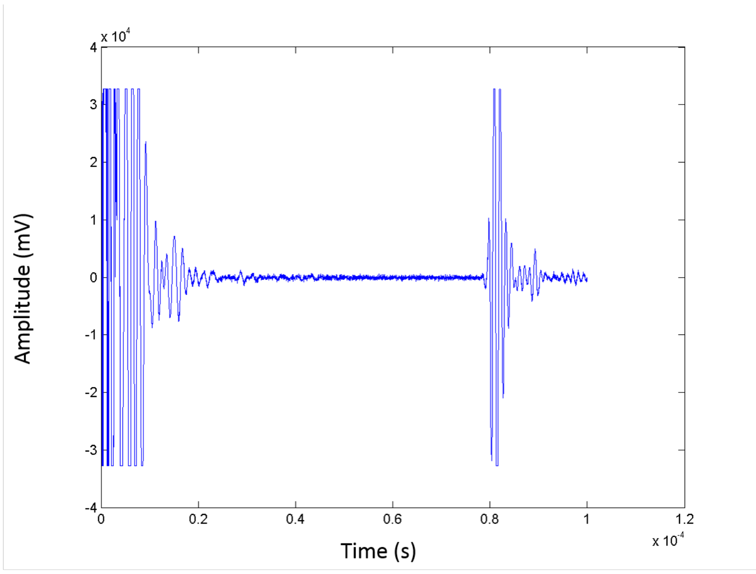


Figure 2.13: *The response of an element from interrogating a Rexolite block with 1 MHz linear array*

other unwanted signals. A higher bandwidth probe means that there is more frequency content to work with. It also means that the signal can be more truly reconstructed.

It is often difficult to find an off-the-shelf commercial probe for specialised applications, however a number of array manufacturers will manufacture custom probes to a customer's specification. In these cases, the customer is expected to provide the desired element locations, centre frequency and required bandwidth. Arrays that have a high fractional bandwidth (>100%) are well suited to acquisition for the purpose of frequency-domain specific post-processing.

The surface geometry of elements is often not square. The Vermon 5MHz linear 128 element 1D array used in this thesis has elements with a pitch of 0.7mm, length of 0.5mm in the direction of the primary axis and width of 10mm in

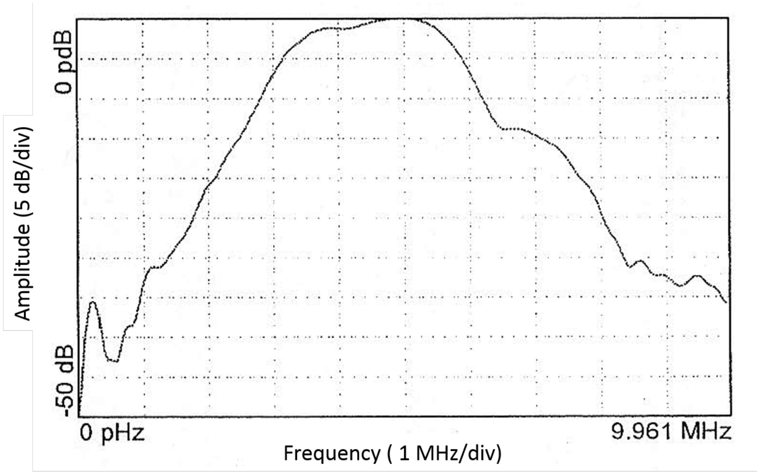


Figure 2.14: *The frequency response for a commercial 5 MHz linear array, scanned from a datasheet (Vermon, France).*

the direction of the secondary axis. This rectangular shape makes the elements directional. This property is beneficial for a one dimensional array as one does not want to receive off-axis signals. The physical properties of each element also reduce the sensitivity to signals arriving from extreme angles. The angular sensitivity for an array element primarily depends on its width and the operating frequency.

2.3 Array Imaging

There are many ways to create an image from data acquired from an ultrasonic array. This section will deal with the concept of beamforming and introduce array imaging to support understanding of advanced array processing discussed in subsequent chapters.

Fermat stated that the path a ray of light takes between two points is the path that can be traversed in the least amount of time[60] (it is this that gives rise to Snell's law, introduced in Section 2.1.2). This principle also holds true for ultrasound[61].

Instead of considering ultrasound energy as a moving wavefront, it can instead be thought of as a ray. This simplifies many imaging algorithms.

It must be noted that while images can be created more easily when taking this principle into account, the wavefront still exists and will contribute to noise if there are any off-axis reflectors. The algorithms in this section will treat the ultrasonic wave as a wavefront if more than one element is being excited at any time, and a ray if only one element is being excited at a time.

The most simple way to display ultrasonic data is via an amplitude scan (A-Scan). It is a plot of amplitude against time for a single ultrasonic receiver. A typical A-Scan is shown and annotated in Figure 2.15.

The A-Scan in Figure 2.15 shows an initial excitation, two reflections from features within the medium and a larger reflection from the back wall of the medium. The second reflection from the back wall is also visible.

A-Scans are useful for detecting the presence of a reflector in a material which is ultrasonically clean (that is, one which does not have large grains). The depth of a reflector can be calculated, given the knowledge of the wave propagation velocity and the time where the reflection occurred.

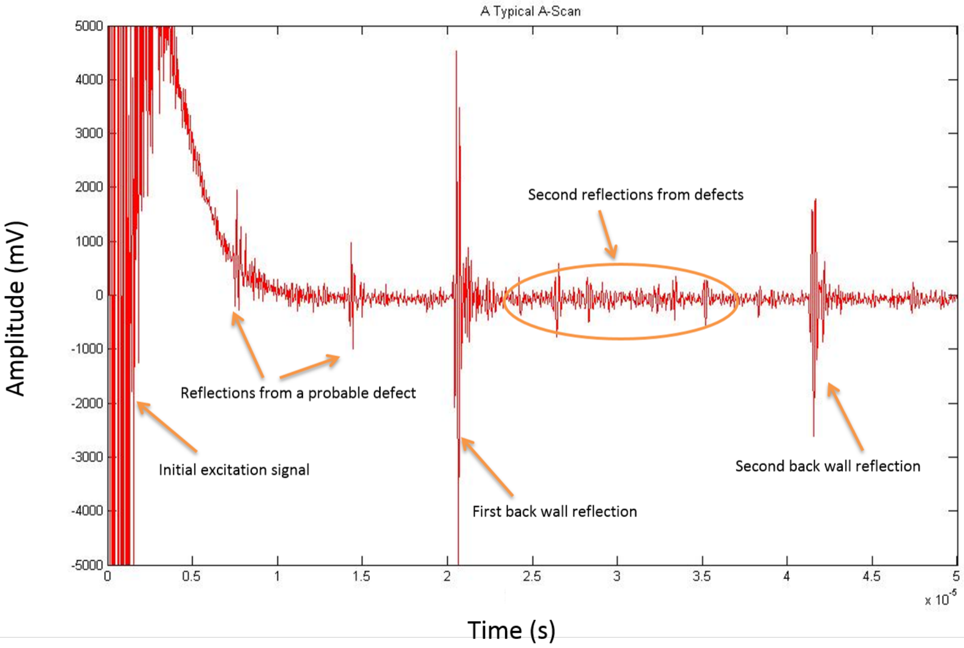


Figure 2.15: *An annotated A-Scan*

More information about a defect can be gained through the use of arrays. Arrays allow a larger volume to be inspected from a single position and also allow for defect sizing and characterisation. They also allow for representing data using B and C scans.

The quality of images is often quantitatively measured using a metric known as Signal to Noise Ratio (SNR). SNR is a ratio of the amplitude of the desired signal to the amplitude of the background noise and is commonly represented in decibels (dB).

2.3.1 Phased Array Imaging

Phased array imaging at its most basic is the use a phased array controller to excite a number of elements in an array with a set of delay laws intended to focus or steer the beam. The data received by each element is then processed and combined in order to create an A-scan. When the transmitted signal from more than one element is summed to generate a focused wavefront, a set of lobes are formed[62]. Figure 2.16 shows a typical lobe pattern from an array. This image was generated by Kummer et al[63]. The largest lobe, travelling in the primary steering direction, is known as the main lobe. Side lobes are also generated as an additional effect of generating the interference pattern. These are attached to the main lobe but radiate energy in undesired directions. These often contribute to noise by causing off-axis reflections. Grating lobes are outputs of focused energy created when the element spacing is over $\frac{\lambda}{2}$. These can also contribute to noise. The choice of element layout can help to reduce the amplitude of these undesirable lobes[55].

2.3.1.1 The Focused B-Scan

By applying delays on transmission, ultrasonic energy can be focused at a single point. This will increase the amplitude of reflections from this point, increasing the signal-to-noise ratio[64] (SNR). By applying these same delays at the time when the data is received, the transmission delays can be accounted for, and the contribution from each element summed to create a A-Scan where the energy is focused.

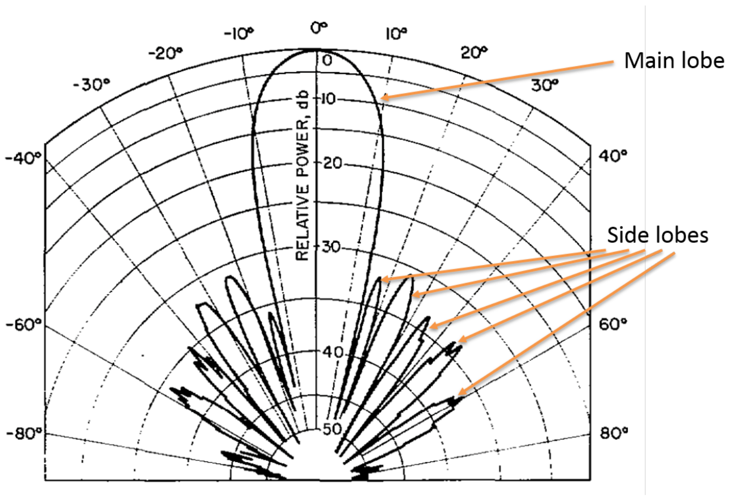


Figure 2.16: *A typical lobe pattern generated from an array*

Figure 2.17 illustrates how a focused B-scan is applied to inspect a range of locations using phased array. A sub-aperture of the array is used to generate a focused wavefront which can be moved by shifting the delay laws to different elements along the entire aperture of the array.

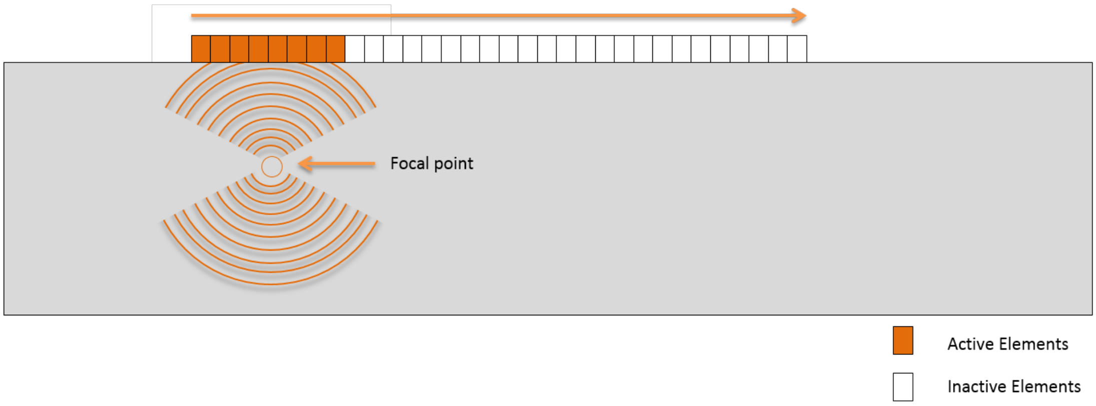


Figure 2.17: *An illustration of the focused B-scan*

To focus at a single point a frame of reference must first be established, from which all delays are calculated with respect to the reference point. This is usually chosen to be the centre of the sub-aperture. The time taken for a wave to propagate from the centre of the array to the point at which the beam will focus is calculated using Equation 2.59, where x is the the distance of the focal point parallel to the array, z is the distance to the focal point perpendicular to the array, and v_L is the longitudinal propagation velocity. The point can be off-axis, relative to the array, but is limited by the directionality of the array elements.

$$T = \frac{\sqrt{x^2 + z^2}}{v_L} \tag{2.59}$$

For every element on the array other than the centre one, the time taken for a wave to propagate to the focal point is calculated and subtracted from the time calculated in Equation 2.59. This gives the delay to be applied to each element. A negative delay indicates that the pulse should lead the reference time (i.e. should be excited before the reference $t = 0$).

On reception, the opposite delays can be applied to that the received A-scans may be summed together and plotted. This style of processing can be referred to as delay-and-sum. It is the basis of all of the other focused B-Scan techniques discussed in this section.

2.3.1.2 Dynamic Depth Focusing

The classic delay-and-sum methodology will result in a processed A-scan which is focused at the point of interest. All other points will be defocused.

It is possible to amend the delays applied on receiving in order to focus at multiple depths in a medium[65]. This is called dynamic depth focusing (DDF) and is an example of the kind of post-processing available on ultrasonic arrays. With a single transmission with fixed focus, variable focal laws on receive allow the modification of focus through a range of points.

DDF quantises the path the beam travels into a number of focal zones. These zones can be infinitely divided, allowing for a different focal law for each sample in the processed A-scan, or can cover a large area to limit the number of focal laws. The latter can be useful when implementing DDF into hardware where memory and processing time is a concern.

For a focal zone, the focal point is defined as the centre of the zone. From there, the receive portion standard delay-and-sum algorithm is applied. This is repeated for every focal zone defined along the beam's path. In this case, the only true focus is where the beam is focused at the time of transmission but the pseudo-focusing performed in DDF makes the a larger portion of the A-scan in focus, as opposed to only a small range of samples.

2.3.1.3 Sector Scanning

After DDF has been applied, the result will be a well focused A-scan valid for the beam path from the point of reference on the array, to beyond the initial point of focus.

This can be repeated across an arc to create a sweep of focal points throughout a slice of a volume. Combined with dynamic depth focusing, a focused radial image can be created of a 2D slice of a volume with relatively few transmission events. It is usually displayed as a polar plot, illustrating a function of angle versus depth.

Figure 2.18 illustrates how delays can be applied to an array in order to steer a wavefront.

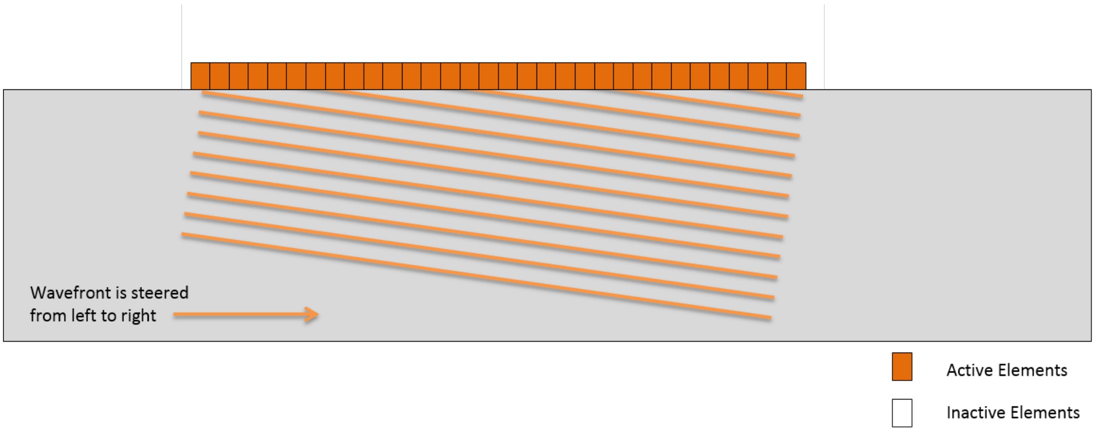


Figure 2.18: *An illustration of beam steering, used for the sector B-scan*

2.3.1.4 Advanced Phased Array Imaging

The techniques described in the previous section are well established and have been used in ultrasonics for many years in both the NDE and medical fields. This subsection will introduce a number of advanced phased array signal processing methods that are designed to improve ultrasonic imaging via increasing both SNR and resolution.

The Phase Coherence Factor (PCF) is a technique that originated in medical imaging[66] that has since been applied to NDE[67]. The instantaneous phase of the signal can be calculated using the Hilbert transform. Using a standard delay-and-sum process, the standard deviation of phases at each point of focus is found. The standard deviation of the phases is used as an indicator of focal quality. For a well-focused beam, the standard deviation of the phases is very low. For an unfocused beam, the opposite is true. Using this knowledge, a weighting factor can be derived for the image so that the contributions from elements that are not well focused are minimised.

Minimum Variance beamforming, also known as Capon beamforming[68], applies an adaptive spatial filter to reshape the lobes of the signal[69, 70]. This technique has been well established in telecommunications and radar but has only comparatively recently been applied to ultrasonic imaging[71]. The adaptive beamforming algorithm minimises the energy received from the medium while maintaining unity gain in a set direction. The minimum varience beamforming method was expanded upon through application of the Wiener beamformer[72]. A range of adaptive weighting functions including Adaptive Sidelobe Reduction and Multiple Signal Classification (MUSIC) were investigated by DeGraaf[73]. Computational complexity for these methods are a concern in the medical field[74] where real-time imaging is of utmost importance. Wang presented a minimum variance beamformer suited to high frame-rate imaging[75], though he does not address computing power in his paper.

In delay-and-sum imaging, resolution can be improved through the minimisation of the main lobe width. This width affects the point spread function of an imaging system[76]. Jeong presented a method to scale received signals based on the ratio of main lobe to side lobe width[77]. Sakhaei investigated a similar frequency-domain technique in order to reduce sidelobe levels[78]. A limitation of standard phased array processing techniques is that they are generally applied at the data acquisition stage and commercial phased array instrumentation often only outputs a processed signal. It is desirable to have the raw data so that data can be re-processed with modified parameters or entirely new algorithms.

2.3.2 Full Matrix Capture

In the previous section, delay-and-sum beamforming was introduced and reviewed and the methodologies for constructing an image evaluated. To provide a true focus for each pixel in an image, a transmission event would need to take place for each pixel. For a 500 by 500 pixel image, 250×10^3 transmissions are required. For a 1 kHz pulse repetition frequency (PRF), it would take 250 seconds to capture the relevant data to generate the image.

There is another methodology for collecting ultrasonic data from a medium. The Full Matrix Capture (FMC) process involves pulsing on a single element and recording the received signal at every element in the array. This is repeated, pulsing on each element of the array in turn. At the end of the process the time trace from every combination of transmit-receive pairs has been recorded[79]. An illustration of the FMC acquisition procedure is shown in Figure 2.19. The data is stored in as a matrix of A-scans, h, and individual A-scans are referenced using the transmit-receive pair (tx, rx) they are associated with. An illustration of the matrix is shown in Figure 2.20. Individual samples of A-scans are referred to by their sample index, (ψ).

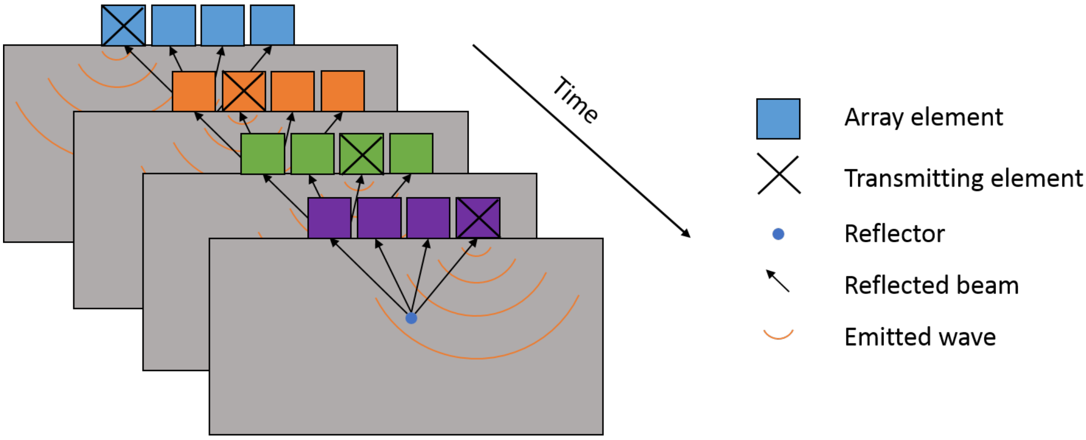
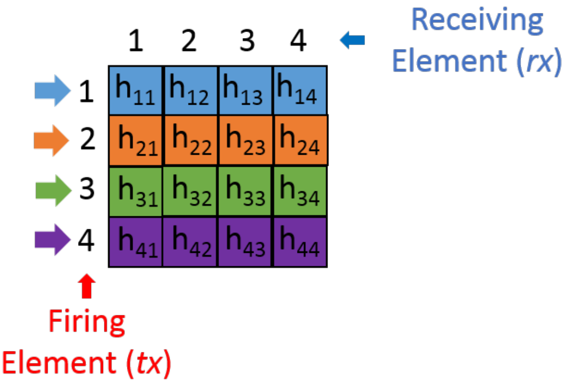


Figure 2.19: *An graphic representation of Full Matrix Capture*



	1	2	3	4
1	h_{11}	h_{12}	h_{13}	h_{14}
2	h_{21}	h_{22}	h_{23}	h_{24}
3	h_{31}	h_{32}	h_{33}	h_{34}
4	h_{41}	h_{42}	h_{43}	h_{44}

Figure 2.20: *An illustration of an FMC dataset*

Assuming that a medium has a linear response, and the principle of superposition holds true, the maximum amount of information able to be collected via an array has been recorded in the full matrix capture. Theoretically, any beamforming technique can be applied in post-processing with the Full Matrix Capture.

FMC is also more efficient when creating a high resolution image. For an array with 128 elements, assuming the same PRF as before, it would take 0.13 seconds (128 elements ÷ 1kHz) to record an FMC dataset. Any other limitations are a result of processing power, memory access speed and data transfer between instrument and PC.

2.3.2.1 The Total Focusing Method

One of the largest benefits of using FMC is that previously impractical imaging methods are now possible. The Total Focusing Method (TFM) is an example of this. It is an imaging algorithm that processes an FMC dataset in such a way that every pixel in an image is individually focused upon.

For every pair of transmitters and receivers, the round-trip propagation time is calculated to a specific pixel in a given image. This propagation time can be converted to an index for a sample (ψ) in an FMC dataset, given knowledge of the FMC's sampling rate and acquisition delay. The samples from each transmit-receive pair are summed and this process is repeated for every pixel in the image. The result is a TFM image. TFM images generally are superior in both SNR and resolution to focused or sector B-scan images.

An example the TFM imaging process is shown in Figure 2.21. Every element is used to focus upon a single point and the response recorded for each element, as if there was a physically focused wavefront at that point. This is repeated for every pixel in the image to be generated.

The equation for calculating a 2D TFM image is shown in Equation 2.60[79].

$$TFM(y,z) = |\sum h_{tx,rx}(\frac{\sqrt{(y_{tx}-y)^2+z^2}+\sqrt{(y_{rx}-y)^2+z^2}}{v_L})| \tag{2.60}$$

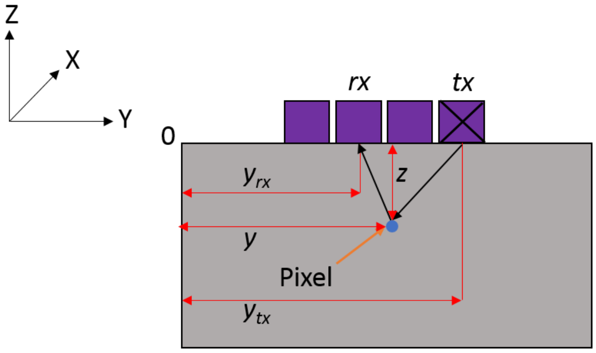


Figure 2.21: *An graphical representation of the variables used in the TFM imaging process*

In Equation 2.60, y_{tx} and y_{rx} are the x co-ordinates of the transmitting element and the receiving element of the probe, respectively. The z co-ordinate of the probe are assumed to be zero i.e. that the image is centred around the interface of the probe and the load medium. $h_{tx,rx}$ represents the FMC data set and v_L is the longitudinal velocity of propagation in the load medium. Finally, y and z represent the discrete location of the pixel to be processed. This is represented graphically in Figure 2.21.

Attention must be drawn to the orientation of the Cartesian plane which is depicted in Figure 2.21 alongside the TFM scenario. This orientation is constant throughout this thesis and all reconstruction images are to be assumed to be in the 2D Y-Z plane unless otherwise stated.

TFM is a straight-forward technique which provides a synthetic focus at every discrete point in a given area (or volume) of the component under investigation. It does this by delaying the response from each array element to simulate a virtual focus at the point of interest. The contributions from each element are then summed and the TFM result is the signal magnitude at that point. The application of additional signal processing is common both before and after the reconstruction of the data. Filtering is often performed to remove electronic noise and any frequencies present in the signal that are outside the operational range of the array. After TFM imaging, a Hilbert transform can be applied to the resultant image in order to produce a smoother image[80].

2.3.2.2 Advanced Applications of the Full Matrix Capture

The Vector Total Focusing Method (VTFM) allows additional information about the orientation of a crack to be gained[81].

The process is similar to TFM but for the contributions from each transmit element the amplitude and direction is recorded. These are summed as a vector and the resulting vector shows the direction of the source from where the most energy was reflected. This allows a greater understanding of reflector shape, orientation and type.

Camacho's Phase Coherence Imaging, introduced in Section 2.3.1.4, can be applied to FMC datasets. The principle remains the same, and in fact the process produces better results due to the fact that there are more samples contributing to a pixel, meaning there are more samples from which to calculate the standard deviation of phases.

Li and Li's Generalised Coherence Factor Imaging is similar to Phase Coherence Imaging[82]. The Generalised Coherence Factor (GCF) calculates the ratio of energy in a defined area of the frequency spectrum to the energy in the full spectrum and uses this ratio to weight the TFM (or focused) image. For this method it assumes that the frequency profile of the coherent signals is known, and in fact it assumes it to be the low-frequency area of the spectrum. For materials that are difficult to inspect ultrasonically, this may not always be the case. This is a limitation of the GCF technique.

Gongzhang et al investigated methods of exploiting the larger range of frequencies recorded with wideband arrays. Frequency diversity techniques were used to remove speckle noise. This technique had the drawback of inserting a number of artefacts into both A-scans and images, affecting the probability of detection of genuine defects[83, 84].

2.4 Modelling Platforms

Modelling is an important factor in designing arrays or post-processing methods. It allows multiple scenarios to be simulated both quickly and with less effort, compared to setting up experiments. The manufacture of numerous of arrays with differing configurations to test design parameters is not economically viable due to manufacturing time and costs.

2.4.1 Finite Element

The Finite Element method (FE) is a technique for simulating physical effects on models. Finite Element simulation is applicable in many areas of engineering due to its ability to solve complex problems and the fact that without it, many simulations would not be possible and experiments would have to be run every time data was required[85]. The benefits of being able to gather information from a simulation as opposed to experimentally are numerous. The first is repeatability. Each time the experiment is run, the same data will be returned if the easily controllable parameters are kept the same. The second is ease of use; no complex phased array controllers need to be worked with, and a suitable sample does not have to be found. Parameters such as wave frequency can easily be changed and defects can be added to existing models.

The drawbacks to Finite Element modelling involve validity and time. There are many external factors when performing an ultrasonic inspection, including ambient temperature, thickness of couplant and array degradation. It is not practical to fully account for all potential factors and assumptions must be made. Because of this, results from experiment and modelling are rarely identical. The time taken to run large models can be an issue. Even fairly modest models can take hours to run with larger NDE models taking days, even on high performance computers. This pales in comparison to the field of geology where simulations may be run for many weeks to gather results.

Weidlinger Associates' PZFlex[86] was validated by Dobson et al to prove that the approximations that PZFlex makes will result in a valid model by comparing the results of a model to experimentally gathered data[87].

2.4.2 Mathematical Modelling

In Section 2.1.1 the wave equation was derived, with the following section dealing with the interaction of waves and boundaries. Section 2.2.2 introduced Huygens' principle of superposition and equations introduced which can be used to model the pressure field of an individual transducer of an array. The following section introduces mathematical modelling in the frequency domain, and a methodology for determining signal received by a transducer given an input signal and a knowledge of the system.

2.4.2.1 The Frequency Domain

The French mathematician, Fourier, demonstrated that frequency components in a linear system could be separated and analysed individually[88]. Fourier's theorem states that any signal or image can be de-constructed and represented by a summation of a series of sinusoids[89].

The Fourier Transform is used to convert between the time and frequency domain analytically. The Fast Fourier Transform (FFT) is used in digital systems to perform an approximation of the Fourier Transform.

The Fourier Transform and its inverse can be expressed analytically as shown in Equations 2.61 and 2.62.

$$g(\omega) = F[f(t)] = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}dt \tag{2.61}$$

$$f(t) = F^{-1}[g(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(\omega)e^{i\omega t}d\omega \tag{2.62}$$

Any linear system can be represented as a transfer function and therefore the effects of propagation through a material can be considered as a transfer function applied to the input signal.

This transfer function can be represented as the product of individual effects of propagation. Figure 2.22 shows a graphical representations of the individual transfer functions, many of which are a function of frequency, that can be used to model wave propagation.

$$H(\omega) = T_x(\omega)A(\omega)B_D X(\omega)\Delta(\omega)R_x(\omega) \tag{2.63}$$

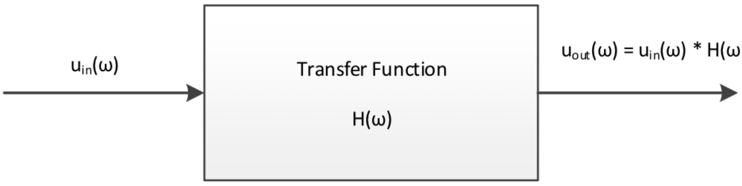


Figure 2.22: *A graphical representation of a transfer function system*

In Equation 2.63[90], $T_x(\omega)$ represents the output of the transmitter; $R_x(\omega)$ is the characteristics of the receiver; $A(\omega)$ is the amplitude reduction due to attenuation; B_D is the component representing the beam divergence; $X(\omega)$ represents the effects of boundaries encountered during propagation and $\Delta(\omega)$ is the time delay due to propagation. For a clean material with no features or reflectors, $X(\omega)$ may be set to 1.

Both $T_x(\omega)$ and $R_x(\omega)$ are defined by the characteristics of the transmitter and receiver respectively. They are each the product of $I(\omega)$, which is the frequency response of the transducer, and $D_F(\omega)$, which is the angular sensitivity.

The characteristics of the transmitting and receiving transducers can be considered as the product of the respective frequency response and angular sensitivity (directivity) of each transducer.

The transfer function used to calculate propagation delay can be calculated as follows where d is the propagation distance. It does not modify the amplitude of the signal, only the phase. v_L can also be a function of ω as different frequencies may travel at different velocities. If this is the case, the signal will distort over distance. This is known as dispersion.

$$\Delta(\omega) = e^{\frac{-i\omega x}{v_L}} \tag{2.64}$$

Attenuation rises exponentially with propagation distance and can be represented by Equation 2.65, where is α the unit of attenuation, measured in Nepers per metre.

$$A(\omega) = e^{-\alpha x} \tag{2.65}$$

The loss from beam spread (divergence) can be calculated simply, using Equation 2.66, where x is the propagation distance.

$$B_D = \frac{1}{\sqrt{x}} \tag{2.66}$$

It is now possible to calculate the expected output at a receiving element using Equation 2.63, given that the transducer characteristics are known, the material properties of the medium that the signal will propagate through, and the input wave packet itself.

2.4.2.2 Application of Mathematical Modelling

A number of commercial software packages use mathematical modelling to calculate focal laws for inspection, such as UltraVision[91] (Zetec, USA) or CIVA[92] (CEA, France). More complex mathematical models have been achieved, simulating anisotropic materials in order to test image reconstruction methods[93].

Dzieweirz and McGilp designed a process, using the mathematical modelling technique, to model an array and a point reflector in multiple locations to evaluate the performance of the array[94]. The mathematical modelling in this paper uses the same principles as the ones introduced in Section 2.1.1. The Full Matrix Capture is simulated, the TFM image generated, and key metrics such as sensitivity and resolution are extracted. The data is compiled to a graph that shows the effective area of an array for a set of prerequisite metrics.

This process is computationally intensive, requiring thousands of separate simulation and imaging procedures. The process becomes more complex when coupling changes. When the array is not in contact with the medium to inspected, and water coupling is used, refraction will occur. A need arose to develop a rapid imaging platform which would take refraction into account and is the subject of the work developed in Chapter 3.

Chapter 3

Focusing Ultrasound Through An Arbitrary Interface

3.1 Theory

3.1.1 Overview

This chapter introduces a novel approach to dealing with refraction for real time imaging through a refractive interface. This chapter introduces a novel approach to dealing with refraction for real time imaging through a refractive interface. The methodologies and results in this chapter were developed by a research team within the Centre for Ultrasonic Engineering and the University of Strathclyde. This work builds upon the original research performed by Dziewierz[95] and was subsequently extended by McGilp[96] to produce the results presented in Section 3.3.3.2.

In industry, inspection of welds is often hampered by their complex geometries. A simple solution to this is to use an array on a conformable wedge[97]. Calculating wave propagation times through an irregular interface is an iterative process and can take a significant amount of time to complete for a real-time interactive implementation of TFM. A new method for interpolation of time delays is presented where the entire imaging algorithm is implemented in Compute Unified Device Architecture (CUDA) C++ and iterative processes are minimised through a curve fitting procedure. In this chapter, an implementation of the TFM algorithm that allows rapid, low-latency imaging through an arbitrary 3D curved refracting interface is presented.

3.1.2 Introduction

The use of arrays in ultrasonic non-destructive evaluation has opened a door to the application of many different diagnostic methods. Phased arrays can be used to change the effective aperture size of an array, create an angled wavefront or focus all the energy emitted from the array at a chosen point.

Focusing energy on an individual point improves the SNR when the point is imaged but generally decreases the SNR for areas outside the focus. Focusing at multiple points requires multiple transmissions and altering of focal laws for each point. This is time consuming and becomes infeasible when an image requires either a large area to be covered or a high resolution.

Delay-and-sum beamforming, introduced in Section 2.3.1, can be implemented in post-processing[79]. For this, the data has to be acquired in format known as Full Matrix Capture (FMC). An FMC is the complete set of time domain data from every combination of transmit and receive elements. Theoretically, any focusing that could be done in transmission can instead be implemented in post-processing.

A popular method of focusing data acquired via FMC is known as the Total Focusing Method (TFM). This method uses delay-and-sum calculations in post-processing to focus energy at discrete points on a high resolution grid in a digital image. This method allows a higher resolution and SNR than the traditional focused and sector scans. Both FMC and TFM were introduced previously, in Section 2.3.2.

The imaging process begins with acquiring a data set. Hereafter, the time domain signal of an echo will be referred to as an 'A-scan'. For example, with a 128-element probe there are $128 * 128$ transmit-receive pairs, therefore 16,384 A-scans are recorded, forming a Full Matrix Capture (FMC) dataset. If only some combination of transmit-receive pairs are used, the resulting dataset is known as a Sparse Matrix Capture[98]. For example, if transmission occurs on 32 elements and on reception 128 elements are used, this results in 4,096 A-scans; however, in this chapter this second style of data will also be referred to as an FMC dataset.

In the imaging stage, for each pixel in the image, the following operations have to be performed for each transmit-receive pair:

a) Calculate time of flight of sound along the path between the transmitting element, through the refracting interface, to the physical location of the pixel of interest and back to the receiving element;

b) Identify the sample at which the data from the round distance trip resides and accumulate the recorded amplitude along with the contributions to this pixel from every other transmit-receive pair.

In this work, an imaging case is considered where the probe is positioned in, or in contact with, medium 1 of wave velocity $v1$, the waves are sent and received through a curved interface into medium 2 with wave velocity $v2$, and the image of reflectors in medium 2 is sought, as depicted in Figure 3.1.

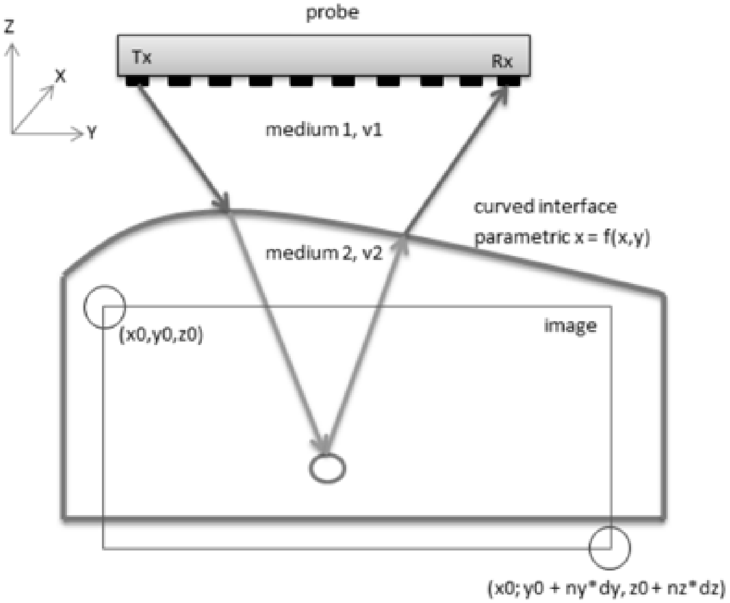


Figure 3.1: *Example geometry of the refraction problem*

The most basic implementation of TFM has two major drawbacks. The first of these is the fact that the classic TFM algorithm accounts for only one wave propagation speed[79]. This means that if the wave travels through an interface and reaches a reflector while travelling at a different velocity, the reflector will not appear correctly in the image. This limitation restricts the use of common tools in NDE, such as wedges, in conjunction with TFM imaging.

When imaging, a single sample will be selected from an A-scan for each transmit-receive pair. For a homogeneous material, one can calculate the distance between the element and the point of focus and divide it by the longitudinal velocity of propagation in order to calculate the first arrival of the wavefront at the focal point. For a case where a wave will travel between media the calculation is not so simple as one must account for the change in velocity between the two materials. As the wavefront is somewhat spherical for a typical array configuration, when it refracts through a surface it will do so across a large area as different parts of the wavefront will reach different parts of the surface first. The point at which the wave crosses the surface for the shortest propagation time to the point of interest must be found in order for the wavefront to be treated as a beam to simplify post-processing.

This problem has been well explored and is known as Fermat's principle. The wave that follows this path can be referred to as a ray, as is common in optics[60]. Cassereau et al investigated two ways to approach this problem. One technique, known as the Fermat's Surface (FS) method, calculates delays which are applied to the transmitting elements of an array[99]. This method requires recalculation of focal laws for every focal point, and is therefore inefficient. Artefacts are often common in images generated using FS due to abnormally large sidelobes[100]. Time reversal was also investigated. It is an adaptive self-focusing method that requires no a priori knowledge of the propagating medium[101]. The time reversal method shows promise but requires transmission of a pulse which is then analysed and used to calculate accurate focal laws for the next transmission. As the aim is to produce a TFM-like image, which will reduce the impact of beam divergence[102], multiple focal points are required. Re-transmission for each focal point is not feasible due to the number that would be required.

An iterative solution to Fermat's principle was presented by Weston et al[103]. Weston used a curve fitting method to learn the shape of the interface before using an optimisation algorithm to find the point where the ray crosses the interface. The limitation of this approach is the time taken for an iterative solution to be found.

A virtual array simulating each array element was proposed by Fritsch et al[104]. This technique allows rapid computation of focal laws through curved interfaces. Once the positions of the virtual array elements have been determined, the area of inspection can be treated as a single homogeneous medium. The estimated propagation times from the method generally had an error of around 2ns, which would produce sampling errors for an FMC dataset sampled at 100MHz. The maximum sampling frequency of the Phased Array Controllers used to capture experimental data for this thesis is 100MHz. Ideally, the error would be consistently <1ns so that sampling errors would not become statistically relevant.

Tweedy et al applied the techniques described by Brekhovskikh[105] to perform volumetric imaging using the vector total focusing method[106][107]. Brekhovskikh focuses on flat layers and only discusses surface waves in curved surfaces. Matuda et al combined a synthetic transmit aperture (STA) with the Sign Coherence Factor method (SCF) to minimise grating lobe noise in images[66]. Fermat's standard approach was taken to find the point of refraction[108].

Another popular solution to Fermat's principle when waves are propagating through a non-homogeneous or anisotropic media is ray tracing. Ray tracing involves quantising the 3D space through which the wave propagates. For each discrete volume, the wave equation is solved so that for each point in time, the position of the wavefront is known[109]. Jurado et al developed a mathematical model that uses a ghost interface to approximate bending of rays through complex structures[110]. The method is based on ray tracing, and the time taken to calculate the path of a single ray is around 1 second. For an ultrasound array configuration scenario, there are multiple sources and multiple receivers each of which would require solving individually. For a 32 element array and a 500 by 500 pixel image, the paths of 8 million rays would need to be calculated. Jurado's method is not feasible at this scale. Ray tracing also provides the positional information of the wavefront at any point in time. For this application, the only knowledge required is the time when the wavefront reaches a point of interest. Another solution that deals with arbitrary interfaces was presented by Sutcliffe et al[111]. They used an iterative minimisation algorithm to pre-calculate wave paths before capturing data to give pseudo real-time imaging. The huge drawback of this approach is the fact that the pre-calculation takes over an hour to complete, rendering it impractical for true real-time imaging applications.

The second limitation of TFM is its computational cost. Consider the start of the imaging process when a full collection of data from every transmit-receive pair of elements (an FMC dataset) is available. In the imaging stage, for each pixel of the image, the following operations have to be performed. For each transmitter/receiver pair:

a) Calculate time of flight of sound along the path between the transmitting element to the pixel and back to the receiving element

b) Accumulate (i.e. sum) the echo value from respective A-scan memory. In the simplest implementation of TFM, one would take each pixel, calculate the times of flight (ToFs) for all combinations of transmit/receive signals, and coherently accumulate respective A-scan samples from the FMC data set. This process results in a large number of loops (for each transmit element, for each receive element, for all x-pixels, for all y-pixels). Both Lambert et al[112] and Yiu et al[113] have presented papers which use GP-GPU software to generate TFM images. While neither of these papers deal with the problem of refraction, their results will be used as a benchmark for the presented algorithm.

This chapter presents an efficient approach to TFM imaging through an arbitrary interface. Both the imaging methodology and its implementation are novel. Iterative processes are computationally inefficient but they are unavoidable when calculating ray paths through non-flat interfaces. Furthermore, once the intermediary point of the ray path is found, two separate calculations need to take place to find the propagation time from the originating point to the interface, and from the interface to the point of interest within the second material. Considering that this process must be done for each transmit-receive pair for each pixel, there are a large number of calculations to be performed.

It must also be noted that Zhang et al have investigated a similar technique[114] that involves using a smaller dataset then extrapolating this data in order to generate a high-quality image. Zhang proposes generating a sparse grid of known time-of-flights then interpolating to generate a final image. An in-depth study of the errors incurred via this method was performed in order to establish the maximum point separation of the grid before the image quality was affected.

General Purpose Graphical Processing Units (GP-GPUs) are gaining popularity in the field of ultrasonic imaging due to the number of parallel processes that they can support. While a modern high-end Intel i7 CPU has 4 cores with 8 compute threads, modern NVidia graphics cards contain GPUs with 2880 cores, though only one thread can run on each core. While the CUDA cores are less complex than a CPU core, it is perfectly suited to do simple mathematical operations.

A GPU's main drawback is memory latency; that is the time taken to access a location in memory and deliver its contents to the program. While numerical operations are very fast, fetching stored numbers from memory takes a comparatively long time. Using all of the features available on a GPU, the memory latency can be minimised, thus accelerating the procedure. Another drawback of GPUs is that they do not deal with diverging threads (i.e. processes performing differing operations) particularly efficiently. Cores are clustered in groups of 32, called a warp. Each warp can only perform the same operation per clock cycle, be it reading memory or performing a mathematical operation. If threads on a warp diverge, cores will spend valuable computing time doing nothing waiting for other threads to catch up. It is for this reason that the use conditional statements need to be minimised in parallel GPU computing.

Due to the need to minimise thread divergence, and the fact that the code takes a long time to run regardless of divergence, the iterative solver for Fermat's principle needs to be run as few times as possible.

While propagation time is not proportional to the depth in the material, it is possible to describe the relationship with a curve. Although it is not known if the coefficients of the curve can be found directly given the parameters of the experiment, it is possible to find a small number of propagation times and from this, calculate a set of coefficients that will describe a depth-time curve for any transmit-receive pair.

The presented method does this, and as such is fast and efficient while retaining the accuracy of non-refractive TFM imaging. This enables an implementation of refractive TFM which is able to produce rapid and accuate results.

3.2 Methodology

TFM imaging is a process which can be referred to as 'embarrassingly parallel'. This means that the methodology can be easily split into a number of identical independent processes. With TFM, each pixel and each ToF can be calculated independently, and in parallel at the same time as others, from a single FMC data set. However, as it will be shown later in this work, the order of calculations does matter in terms of computational efficiency. Given a relatively simple set of operations needed for each pixel, GP-GPU computing cards are a good candidate for realisation of the TFM process. With their compute-dense architecture, the memory interface bandwidth becomes the limiting factor in utilising large look up tables. In fact, due to their architecture, some results are faster to be re-calculated on the chip as needed, rather than calculated once and stored[115]. When considering implementing the TFM process on the GPU, one should consider taking maximum advantage of various subsystems of the GPU architecture. In particular, there are a number of memory subsystems, varying in functionality, bandwidth and latency, and special function units, like texturing units, that can work in parallel with the main streaming processors responsible for the bulk of computation.

The classic way of calculating the time of propagation of a sound ray through a refracting interface is by use of Fermat's principle of shortest time of propagation. To solve this problem, various small-scale optimisation algorithms can be used. Dziewierz described a method for obtaining a computationally efficient, closed-form solution for the equation describing the time of flight of acoustic ray through a 3D planar interface[116]. However, the problem now has to be generalized for arbitrary interfaces. Solving this problem for a planar interface comes down to finding roots for a 4th order polynomial equation. It was therefore anticipated that curved interfaces would involve equations such as a 5th order polynomial, or higher. Although there exist methods for finding roots of such equations[117], a different approach is used here due to the desire to find a general solution.

The propagation time between neighbouring pixels is a non-linear, but smooth function of space. The approach adopted here approximates the exact solution of the equation that describes propagation delay with an interpolating polynomial. A polynomial was selected as the interpolating function because of the high efficiency with which it can be evaluated on the GPU, and controllable accuracy of the solution, as discussed later in this chapter.

3.2.1 Stage 1: Prototype time of flight probing points calculation

a) The imaging space is divided into straight z-lines, as depicted in Figure 3.2[118], with z-coordinates going in a general direction away from the probe. In this implementation, z is the ‘depth’ coordinate related to some arbitrary coordinate system in which the probe elements, refracting surface and TFM image volume is described.

b) In order to iteratively find the point of refraction on a surface, the surface must first be defined. Early versions of this process required the surface to be input as a mathematical equation, for example $z = 0$ for a flat surface. A methodology has since been developed that automatically finds the shape of a surface and incorporates this into the solver[119]. This methodology involves creating a basic TFM image of the surface and using thresholding to isolate the area where the first reflections occur. It is assumed that coupling takes place through water and that the first large reflection will be that of the material’s surface. The estimated surface profile is extracted from the image and input

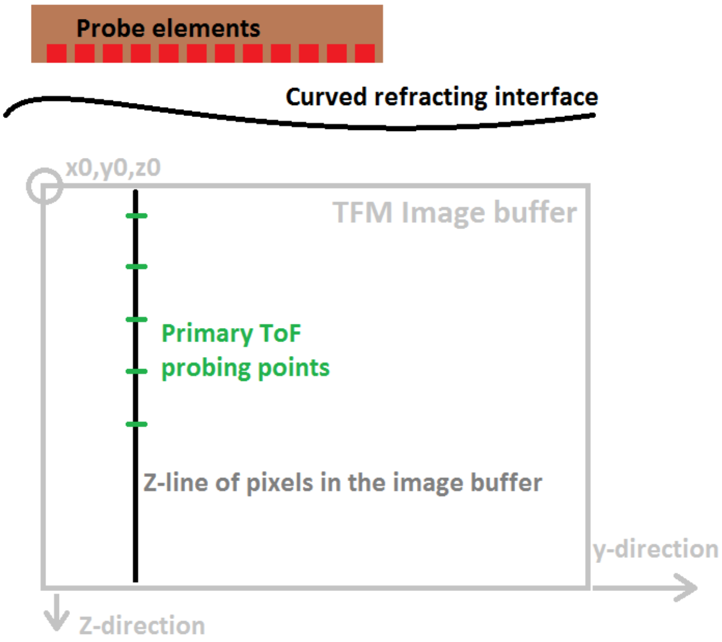


Figure 3.2: *Location of probe, refracting material interface, image buffer, and z-line of pixels inside the buffer. The linear memory locations progress down-first, then right. The third dimension is x.*

to a curve-fitting algorithm that smooths the result. Finally, the surface is converted to a lookup table and uploaded to the graphics card. Due to the fact that curve-fitting is used on the surface profile, it is expected that this algorithm will not perform well on specimens with a rapidly changing surface profile.

c) For each z-line and probe element combination, propagation times for a number of points along the z-line are calculated using one of the iterative methods based on Fermat's principle. The number of points was chosen to be 16, and this value remains constant throughout this chapter. Zhang et al explored a similar technique[114], and investigated the maximum spacing permitted between points before errors were introduced. The Nelder-Mead simplex optimisation method was chosen to find the minima due to its simplicity. Such an algorithm can be optimised for speed. The Newton-Raphson method was also investigated but discounted due to the fact it is prone to converging to local minima. MATLAB (MathWorks, USA) is a technical computing language and IDE. Its inbuilt optimisation function[1] is a generalised implementation of the Nelder-Mead method. It was used as basis for a low-level implementation of the methodology that is far more efficient than a general implementation. A number of performance enhancements techniques were used, such as loop unrolling and optimising the code for the desired number of coefficients. This allowed removal of many conditional statements that will impact the speed of the process. This is especially important when porting the code to GP-GPU as the architecture of GPUs does not allow for efficient processing of branches. The spatial distribution of the points can be regular, or generated using affine transformation of Chebyshev nodes as in Equation 3.1 where n is number of points required and x_k is the distance of each point from the origin. These are known as the primary time of flight $= f(z)$ points, where $f(z)$ is some unknown function that describes the time of flight along the z-line.

$$x_k = \cos\left(\frac{2k-1}{2n}\pi\right), k = 1, ..., n \tag{3.1}$$

---

[1] `www.mathworks.com/help/matlab/ref/fminsearch.html`

d) The physical time of flight (in seconds) value is scaled to a normalized value that maps into the appropriate fraction of the A-scan, assumed to be coincident with spatial location inside the TFM image buffer. This single multiplication/offset operation at this point saves thousands of processor cycles that would be needed in Stage 2 to execute the same operation.

3.2.2 Stage 2: Pre-calculation of interpolant coefficients

The primary time of flight data points are divided into two groups. One group is used to fit an interpolating polynomial, and the second group to estimate the error of the approximation to ensure accuracy. This transformation of $f(z)$ samples into interpolating polynomial coefficients is executed using a specialized, non-branching solver as detailed later in this chapter. At this point, the minimal order of polynomial is selected in such way as to fulfil the requirement for accuracy of the approximation, as detailed later in this chapter. The interpolant coefficient database is segmented and optimised for GPUs' on-chip cache, enabling rapid retrieval of the coefficients to the streaming processors.

For a given probe, surface and scene parameter combination, a set of interpolant coefficients can be calculated once and stored for reuse in delay-and-sum imaging. If the probe-refracting surface configuration is changing (for example, when the probe probe is moving over the refracting surface causing the surface profile to change), re-calculation of this database is necessary. This operation is comparatively inexpensive in terms of computing complexity compared to the imaging process.

3.2.3 Stage 3: Delay-and-Sum Imaging

For most imaging problems, this part of the process is the most computationally expensive, and therefore has been designed to benefit from details of the GPU architecture.

In the GPU, during parallel processing, each thread is made responsible for a single pixel of the image. The thread block size is configured in such way that all the active threads in a block process consecutive pixels from the same Z-line of the image. This allows the threads of the thread block to take a couple of advantages from the GPU's memory architecture, as described below.

For this implementation, a list is employed that describes which A-scans from the FMC buffer will be used in each iteration. This allows use of a single code base to effectively generate images from Sparse Matrix Capture datasets. This table is loaded through the constant cache - also benefiting from the broadcast mechanism, in a similar way to the interpolant coefficient table.

Inside the thread, the outermost and only loop is over the list of A-scans that contribute to a given pixel. For a given A-scan, the thread utilizes the polynomial coefficients obtained in Stage 1 (Section 3.2.1) to calculate the respective transmitter-to-pixel and pixel-to-receiver propagation times.

At this point, one should be aware a particular feature of GPU architecture: shared memory broadcast. The interpolant coefficients corresponding to a single Z-line of the final are loaded from device memory into shared memory only once at the beginning of the thread block, and from there, the values are repeatedly broadcast to all threads as needed, using a pointer in memory. This is possible because consecutive A-scans are processed simultaneously across all active threads of the multiprocessor, in lockstep. One could treat this operation as user-managed caching.

The threads evaluate the time of flight using an interpolating polynomial with their respective pixel's coordinate fed into the Horner method[120]. This operation maps to a short series of very efficient FMA (Fused Multiply Accumulate) instructions with no loop. In effect, the required time of flight value for each pixel is recreated as appropriate. As detailed in Section 3.2.1, the resultant value is already scaled such that this value is a pointer that the texturing unit can use to load the appropriate A-scan sample value from the FMC buffer.

Here another feature of GPU becomes useful to accelerate the retrieval of data. It would normally appear that random memory access is required for the FMC sample retrieval step. However, since all threads evaluate times of flight for a given transmit/receive pair (A-scan) simultaneously, and beginning from the top pixel down, they will request A-scan data that reside at neighbouring and progressing memory locations, albeit with non-integer stride. The texturing units can load requested A-scan data through the L2 cache, which is a high-speed memory unit on the GPU itself. This is very efficient because, typically, an entire A-scan will fit in the L2 cache. In this study an NVidia GTX 580 was used which has approximately 1.5 Megabytes of L2 cache memory. From there, the data block is broadcasted to the texturing units that request it. This way the data is efficiently distributed to the texturing units, which in turn also operate their own, smaller texture cache (depending on the GPU model). The data from L2 cache is reused as many times as there are pixels per Z-line, or less, depending on the effect of texturing unit cache.

The texturing units are hardware optimised for this kind of operation as this is essentially the same operation as 3D graphics texturing itself, which a major factor upon which the competitive relative performance between various graphics card models is evaluated. The texturing units select appropriate samples from the A-scan buffer, and provide additional service of interpolation between A-scan samples (super sampling), if necessary, maximising the parallel spread of the computations across the GPU chip. This operation can be referred to as ‘free’ in terms of computational time because as it progresses, the main CUDA cores are executing other threads, preparing a new batch of requests for the texturing units.

Finally, the threads integrate the A-scan samples they received from the texturing units. At this point, the integration operation can be summation (as in classic TFM) but it can also be a different operation; for example multiplication[121] or Phase Coherence Factor calculations[66].

Upon completion of integrating all the A-scans, threads store the computed pixel value in the image buffer, which resides in the global memory. This again is an operation that benefits from the fact that neighbouring threads process neighbouring pixels. The memory write operation is well coalesced and parallel, and fully exploits the main memory bus. Importantly, no synchronisation nor atomic operations are ever needed because by design, no memory write race ever occurs.

Since the image is processed line-by-line, only the interpolant coefficients related to that z-line need to be stored in the quickly-accessible on-chip shared memory. For example, for a 128-element probe and interpolating polynomial of maximum order of 7, and single-precision floating point format of the coefficients, $128*7*4 = 4480$ bytes of information are needed to be loaded from global memory per z-line, but only once for all threads. Each thread working in this z-line re-uses this information using the broadcast mechanism, reading $7*4 = 28$ bytes of coefficients from this buffer, up to $2*128*128$ times per pixel. This saves a significant amount of time that would otherwise be spent loading data from global memory if the broadcast mechanism was not used[122].

However, this also means that running more than few blocks of threads per GPU is somewhat counter-productive, as there is a very limited amount of on-chip shared memory available. Therefore this process works best if there are enough pixels in the z-line, and not many z-lines processed in parallel. For images with a small number of pixels in the z-direction the GPU can potentially be under-utilised. This is due to the fact that there are a number CUDA cores on a GPU and multi-threading will not take place by default. If there are fewer z-pixels than the number of CUDA cores, these cores will remain idle during the computation process.

Incidentally, it is of practical benefit to have the highest image resolution along the depth axis (z-line), because in such cases the constructive/destructive interference between A-scans produces the best TFM process gain (contrast improvement) and phase accuracy. If fewer than 7 pixels per wavelength resolution are used, the TFM process may not achieve its peak process gain due to sampling phase error[123]. The same applies to other TFM-like algorithms like Phase Correlation Factor algorithm.

Overall, the process described is extremely efficient and takes approximately 48 hardware cycles per integrated A-scan (this measured value includes all exposed latencies) per pixel, and is the main source of performance of our implementation.

3.2.4 The non-branching polynomial interpolant coefficient solver

In order to find the coefficients for interpolants needed for this imaging algorithm, a well-known least-square method of fitting a polynomial into a set of points by solving a matrix-quadratic equation can be used.

Typical CPU implementations of this method utilize loops and conditional jumps to allow single implementations of a code to solve for an arbitrary order of polynomial and arbitrary number of data points, sometimes even reordering the data first to minimize the numerical errors. However, in this case, the speed of execution of the solver being of utmost importance, it was decided to compile a range of specialized solvers, each taking exclusively a fixed number of data points and returning a fixed order of polynomial. This approach allows a linear, non-branching code for each case to be obtained. A subset of statistics on the addition and multiplication operation count versus number of inputs and polynomial order has been gathered in Table 3.1.

Polynomial order	**Number of input points**	**FMA ops count**	**MUL ops count**	**ADD + SUB ops count**	**Total op count**
4	6	120	182	44+35	381
4	10	192	260	76+36	564
5	8	330	445	76+99	950
5	12	453	571	116+98	1238
6	10	743	976	119+296	2134
6	14	935	1152	167+296	2550

Table 3.1: *Computational operations required to compute polynomials*

Table 3.1 shows statistics of computation cost for calculating the interpolating polynomial coefficient as a function of polynomial order and number of contributing data points. Each version of the algorithm also requires a single reciprocal operation.

The code has been obtained using Wolfram Mathematica Computer Algebra System, using a method similar to the one Dall'Osso described in detail in Computer algebra systems as mathematical optimizing compilers[124].

The code obtained only requires multiply-accumulate operations and a single reciprocal, and uses no jumps or conditional statements whatsoever. The benefit of such an approach is that such code will execute efficiently on a GPU, solving multiple z-lines in parallel. It is appreciated that such approach is not well suited for poorly conditioned inputs, and will allow the numerical errors to surface in the results, even when using double precision arithmetic. However, as argued in the next section, in this application, the inputs are always well conditioned and the observed numerical errors are acceptable.

3.3 Results

3.3.1 Selection of polynomial order and error analysis

In order to establish confidence in the proposed method, it is necessary to perform a detailed error analysis of the algorithm. The calculated time of flight errors come primarily from 3 sources:

a) Inaccurate primary time of flight solver;

b) Inherent method inaccuracy of polynomial approximation;

c) Numerical inaccuracy of polynomial coefficient solver.

In this work, it is assumed that a) is exact; errors incurred by the primary time of flight solver have been reduced to the point where they negligible. Any additional error added may be caused by the way computers handle floating point numbers. Accounting for this eventuality is considered out of the scope of this thesis. Here only b) and c) are considered as source of errors.

To estimate the end-to-end time of flight calculation error of this process, the following method has been applied:

1) Prepare a set of pre-calculated times of flight for a typical imaging scenario;

2) Convert the subset of pre-calculated times of flight to interpolant coefficients, and then back into full grid of times of flight;

3) Calculate the peak difference between original and processed time of flight data.

The summary of results for the imaging scenario depicted in Figure 3.2 is gathered in Figure 3.3 and are shown as a function of the polynomial order used.

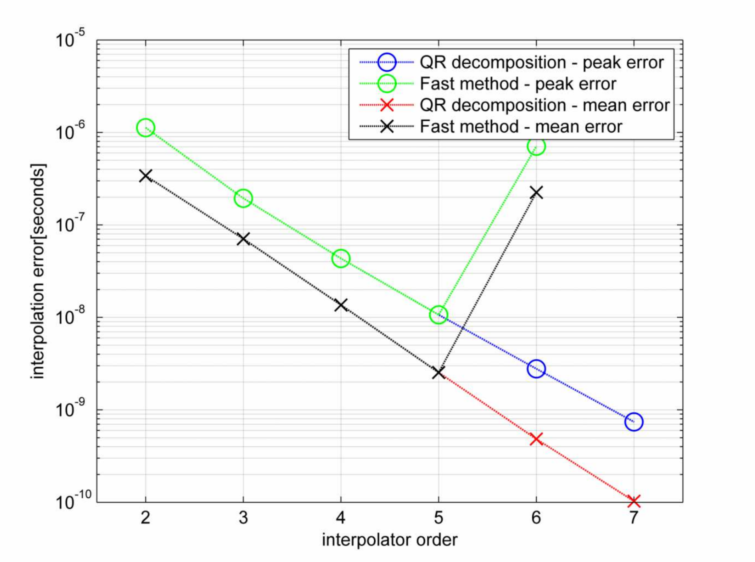


Figure 3.3: *Mean and peak time of flight calculation errors resulting from using a polynomial of a given order as interpolant for a given imaging scenario.*

The proposed direct no-branch solver returns values that are close enough to the results of MATLAB's QR solver to be considered negligible. This holds true up to an interpolant order of 5. If a higher order interpolant is needed, the QR solver is recommended.

The significance of these results are as follows. For a source signal sampling of 50MHz, the useful bandwidth of the sampled signal is considered no better than approximately 17MHz; full sine wave cycle at that frequency is 58.8×10^{-9} seconds long; therefore to obtain phase accuracy better than 1/4 cycle (worst case scenario), the error must be less than 1.47×10^{-8} seconds. From Figure 3.3, it can be observed that a polynomial order of 5 or higher must be selected for interpolation. In this case, the phase accuracy for a 5MHz signal will be better than 1/15 cycle.

Note that the cited timing errors are peak errors and the average timing error will be much lower. In any case, it is possible to obtain an arbitrarily low peak error estimate by raising the interpolant order (up to the limit of numerical representation accuracy). This will be at the cost of a minor decrease in Stage 2 performance.

It is appreciated that the evaluation method presented here does not give strict upper bound for error; however, it can be easily repeated for any practical imaging scenario and the minimum required order calculated for a specific imaging scenario.

3.3.2 Implementation benchmark

The proposed process consists of several stages that can be executed independently, and thus, benchmarked independently. In other publications, the overall performance in practical scenario is typically expressed in frames per second for a given specific scenario[125, 126, 127]; here a more synthetic approach is taken. Each stage is benchmarked independently and then combined to estimate overall performance for a given set of input parameters. These results are shown in Table 3.2.

	Stage Name	**Options**	**Platform**	**Performance Unit**	**Result**
1	Stage 1: Calculating prototype Time of Flight points	Planar interface z=0	GPU	Points/second $\times 10^6$	77.9
2	Stage 2: Transform from time points set into interpolant coefficients	8 points into 5th order, fast solver	GPU	Lines/second $\times 10^3$	23 538
3		8 points into 5th order, fast solver	CPU		2 960
4		8 points into 5th order, QR solver	CPU		70
5		Double precision, 14 points into 7th order, QR solver	CPU		68.5
6	Stage 3: TFM integration	Nearest sample interpolation	GPU	Paths/second $\times 10^9$	26.7

Table 3.2: *Results of benchmarking for each stage*

For each stage of the process, an appropriate measure of performance is introduced. For finding the initial times of flight, the metric is time of flight points calculated per second; for calculating the polynomials, it is lines per second (as each atomic transform deals with entire image line); for creating the final image, paths per second are used - as each atomic operation deals with estimating time of flight over a specific path.

With such metrics, one can trade number of pixels, number of probe elements, and the Tx/Rx firing scheme against the performance of the particular GPU system used and frame rate achieved. For any TFM image resolution and number of elements of the probe, the total calculation cost can be obtained. For example, for an image with 1024^2 pixels, and 64-element probe operating in full FMC firing scheme, the computational cost of the TFM method is $2 * 1024^2 * 64^2 = 8.6 \times 10^9$ paths.

In the example above, assuming that a polynomial order of 5 is selected, with 8 sampling points, there are $8 * 1024 * 64 = 524 \times 10^3$ primary times of flight to be calculated initially, and $1024 * 64 = 65.5 \times 10^3$ interpolant coefficient sets to be obtained in the second stage. To create an image from this data, for each pixel, Tx-to-pixel time of flight is calculated once and then for each of these, pixel-to-Rx is calculated, resulting in $1024^2 * 64^2 = 4.3 \times 10^9$ paths. The symmetry of the time of flight in the FMC is exploited, halving the actual times of flight combination count. Assuming that the FMC data is uploaded to the GPU asynchronously, each frame will take 0.2 sec to compute; image generation takes 82% of the total time. Therefore any future improvement has to be concentrated in this stage of the process. One obvious improvement would be to introduce partial on-chip caching of the calculated times of flight; this option offers potential for future research. The performances cited in this chapter scale almost linearly over multiple GPUs, for example, processing times are approximately halved when comparing an NVidia GTX 580 to an NVidia GTX590, which has two of the GPUs used in the former. This comparison is not intended to highlight the differences between hardware platforms, but to show the effects of algorithm choice (see rows 3 and 4 of Table 3.2).

3.3.3 Experimental validation

3.3.3.1 Cylindrical Surface

The experimental setup is shown in Figure 3.4. The probe is a 128-element, 5MHz, linear phased array probe (Vermon, France), and the Phased Array Controller is a Dynaray (Zetec, USA). The probe is placed over a half-cylinder of solid PVC material, in which a flat bottom hole has been drilled out. In this process, the location and shape of the surface have already been measured and input to the imaging algorithm.

The algorithm returns the expected image; the reflection coming from the hole is blurred out when the refraction is not taken into account (Figure 3.5). When the refracting surface is taken into account (Figure 3.6), the reflection is properly located and focused and the amplitude of the feature is increased by 4.57dB. The large black patches on the sides of the cylinder reflection are side lobes, as expected for this probe type and experimental configuration. Figure 3.7 shows the refracted TFM result in Figure 3.6 with a higher dynamic range.

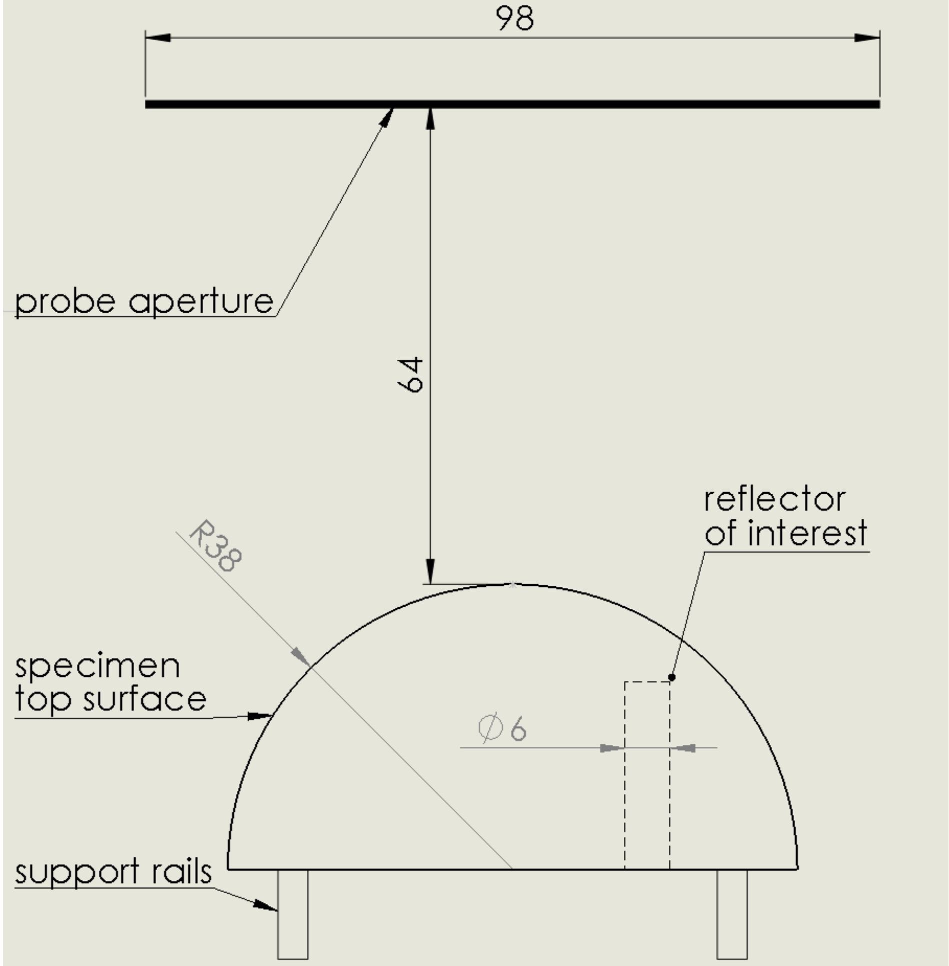


Figure 3.4: *Schematic drawing of the PVC specimen submerged in water, and aperture of the 128-element 5MHz linear array. All dimensions are in mm.*

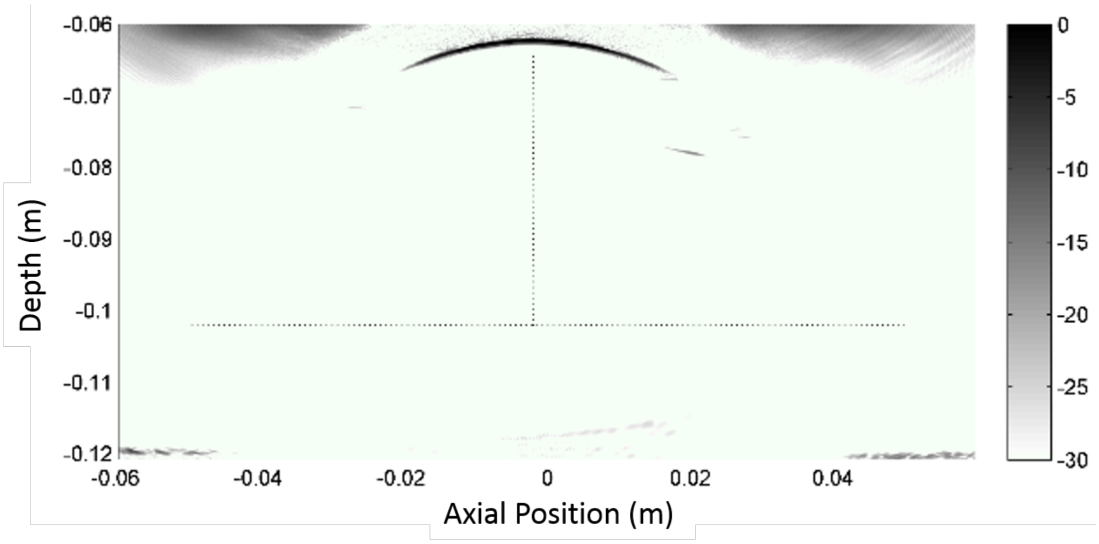


Figure 3.5: *The image of the flat bottom hole inside the specimen - no refraction applied. Guides are shown in the image to illustrate the centre and bottom of the semi-cylinder*

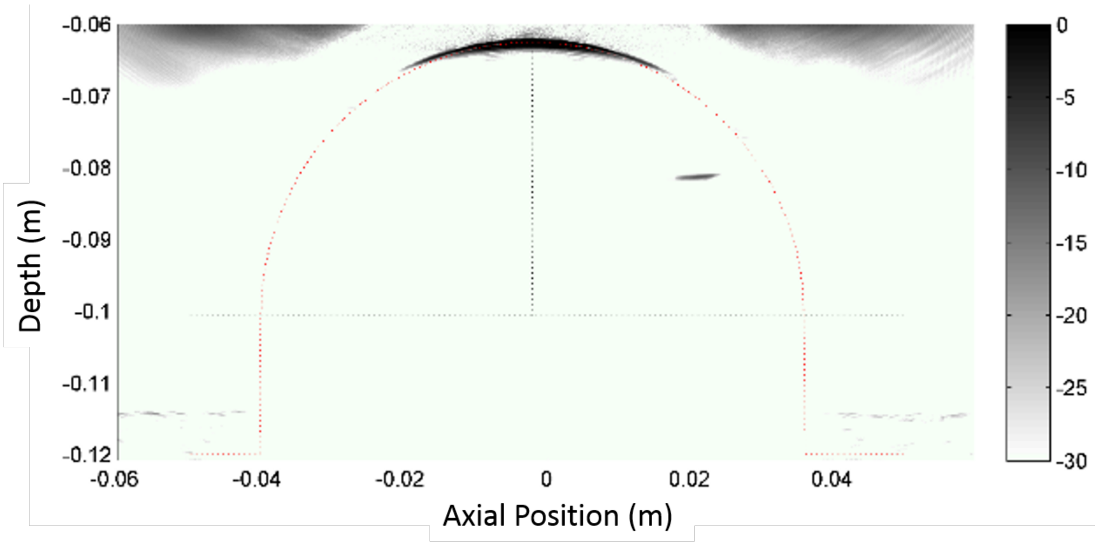


Figure 3.6: *The image of the reflector as in Figure 3.5, but with correct refracting surface taken into account. The amplitude of the reflector is 4.57dB higher and the reflector is correctly positioned and flat. The overlay depicting the centre and bottom of the semi-cylinder sample are also shown in this image.*

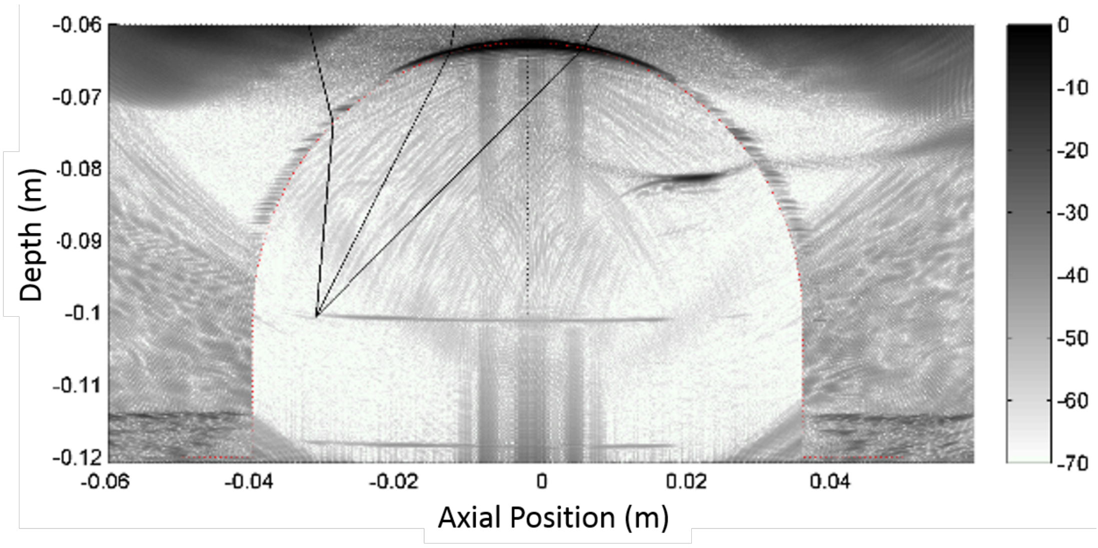


Figure 3.7: *The corrected image with a high dynamic range. The correctly flat back wall is visible at -42dB from the top surface. The 3 black straight lines exemplify calculated ray paths between the 3 of the probe elements and a pixel in the image.*

3.3.3.2 Machined Block

The surface of a stainless steel block was machined in order to create a surface through which inspection would be challenging. A photograph of the block is shown in Figure 3.8. The area that is to be imaged is highlighted in the figure via a blue rectangle. The experiments using this block did not use a pre-input surface profile and instead relied on the software being capable of recognising and accounting for the surface automatically[119]. All experiments using this block used a Dynaray (Zetec, USA) phased array controller and a 128 element linear 5 MHz array (Vermon, France).

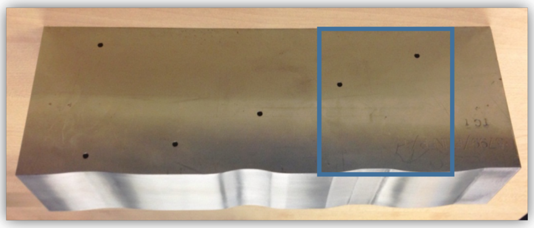


Figure 3.8: *A machined stainless steel 316L block, showing the area to be imaged. The array was placed on a wedge coupled to the top of the block, relative to the image.*

The block was initially tested from the underside. This was to present a simple scenario to test the surface recognition algorithm before inspecting the specimen through the complex surface. A probe was placed on a 14.5° wedge and Full Matrix Capture data recorded.

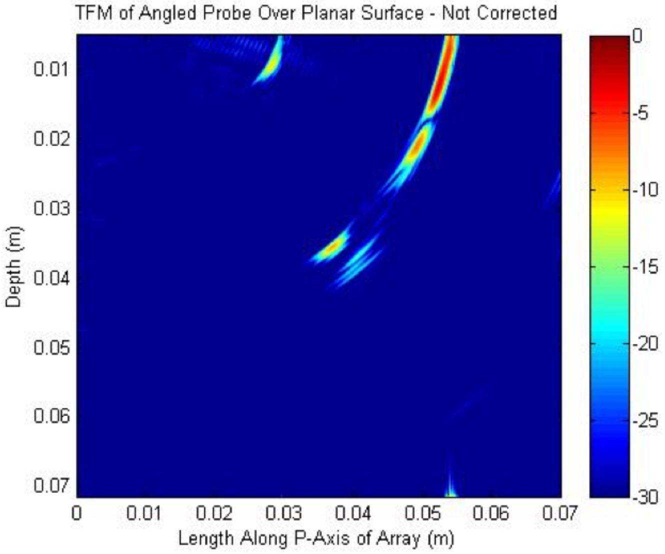


Figure 3.9: *TFM after incorrect planar refraction*

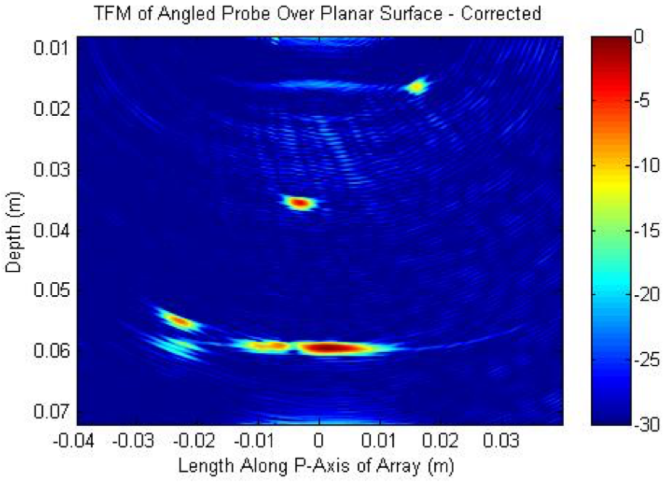


Figure 3.10: *TFM after corrected planar refraction*

Figures 3.9 and 3.10 show the uncorrected and corrected TFM images respectively. The automatic surface correction was found to work well with planar surfaces, as the back wall became visible in the correct position at a depth of 60mm and the side drilled holes can also been seen in the image.

An array was then fixed to a stand, and the block inspected from the machined side via water coupling. The area of the block inspected is shown in Figure 3.11.

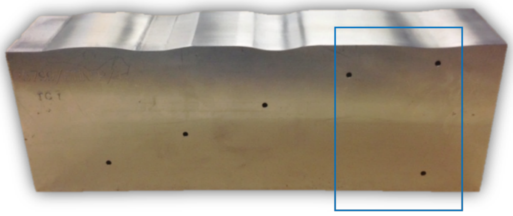


Figure 3.11: *The area of the machined block to be inspected*

The output of the surface recognition algorithm was checked to ensure that the surface was being correctly identified. The surface profile is shown overlaid on the TFM image of the surface in Figure 3.12. It can be observed from the image that the programmatically identified surface conforms well to the surface seen in the experimental image. McGilp has performed additional research into the accuracy of the surface recognition algorithm, but this is considered out of scope for this body of work[96].

Finally, the Total Focusing Method is applied to the FMC dataset. The resulting image is shown in Figure 3.13 and with a dynamic range of 20dB. The dynamic range has been shifted to maximise the reflections from the side-drilled holes.

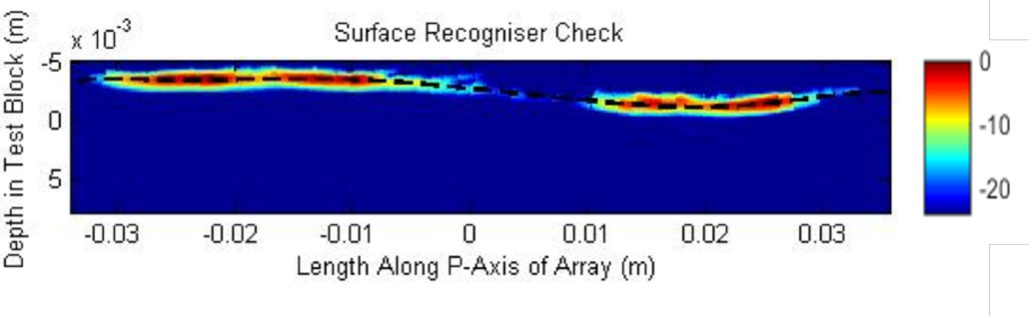


Figure 3.12: *The surface profile of the machined block, shown with dB compression*

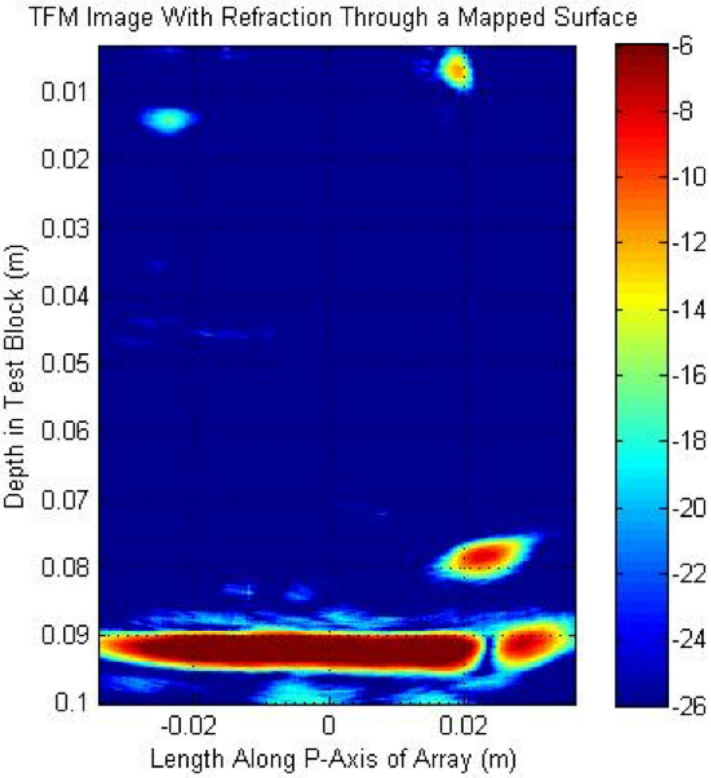


Figure 3.13: *The corrected TFM image of the area shown in Figure 3.11*

The three side-drilled holes can be seen clearly in Figure 3.13 and in the expected locations. From this image, the conclusion can be drawn that the automatic surface recognition algorithm works well, alongside the TFM algorithm that has been designed to cope with an arbitrary surface.

To further ensure that the results are accurate datasets were captured from across the entire sample, processed, and the results stitched together to form an image of the complete block. The results of this are shown in Figure 3.14.

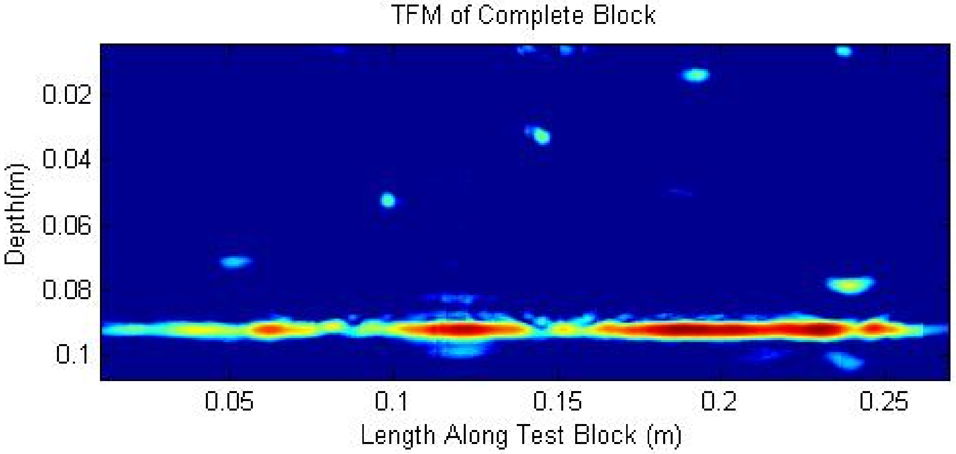


Figure 3.14: *The corrected TFM image of the entire block*

3.4 Discussion

The solution described in this chapter can be viewed as a way of lossy data compression, where a time of flight look-up table is replaced with a function taking from a limited set of data, to reconstruct parts of the content of the original look-up table when and where needed. The efficiency of this solution comes from transforming the time of flight calculation problem in a way that maps well to the hardware resources that are available in GPU processors. Due to the speed of modern processors compared to the memory transfer speed, it is computationally more efficient to regenerate the data from a ‘compressed’ version than to transfer the excess data to GPU memory and spend time accessing it.

This chapter demonstrates that advanced imaging algorithms that were previously considered to be of research interest only, or limited to military or nuclear applications only due to their cost, can be currently implemented in affordable consumer grade hardware.

This work also allows the possibility of bringing advanced ultrasonic imaging to automated applications where classic phased array imaging is currently prevalent. The aerospace industry routinely conducts non-destructive testing using wheel probes to couple an array to a component as a robot co-ordinates with a phased array controller for acquisition. The wheel is deformable and will therefore change shape as the coupling pressure alters throughout the course of a scan. This is a problem easily solved with conventional phased array but presents a problem when varying focal laws are required to perform TFM or similar imaging algorithms. The use of the presented methodology allows real time TFM imaging in inspection scenarios like this.

Chapter 4

Sub-Aperture Correlation Imaging

4.1 Background

The objective of developing this novel advanced signal processing method is to combine time and spatial domain based algorithms to increase the SNR of ultrasound images while maintaining an acceptable resolution for defect detection and characterisation during the inspection of materials with high levels of structural noise.

TFM works very well on homogeneous, isotropic materials and is often referred to as the 'gold standard' in ultrasonic NDE[128]; however the results produced from TFM in highly scattering materials can suffer from degradation[129]. The large grain boundaries within these materials cause incident ultrasonic energy to scatter which gives rise to speckle noise and lowers the SNR of the resultant image.

Phase Coherence Imaging proposes using the Phase Coherence Factor (PCF) to determine whether or not a signal at a given point is predominantly noise or from a legitimate reflector[66]. PCF calculates the instantaneous phase of each

A-scan over the whole FMC dataset before applying delays. Unlike delay-and-sum, PCF calculates the standard deviations of the phases for a focal point. The lower the standard deviation, the less likely it is that the signal is speckle noise. The PCF can be used to weight delay-and-sum images, or even TFM. Experiments in the medical field have shown the PCF is also able to reduce main lobe width as well as sidelobe intensity. Reduction of the main lobe width is desirable as it can lead to a potential increase in resolution. In NDE, PCF has been shown to lower sidelobe levels and improve resolution, but the improvement is not significant for some materials of interest, high nickel alloys for example[67]. PCF has limitations when applied to data from materials in which multipath propagation is an issue. Multipath propagation causes phase aberration with resultant problems in generating an accurate phase coherence factor.

The Generalised Coherence Factor (GCF), proposed by Li and Li[82] uses a similar principle to generate a weighting function based on the statistical likelihood that a signal is predominantly noise. For each point to be analysed, GCF performs the requisite delays before summing for a range of time samples around the desired sample. An FFT is performed and the frequency spectrum of the signal is analysed. The GCF is defined by the ratio of low frequency energy in the signal to total energy. Furthermore, GCF as a weighting function has been shown to effectively reduce speckle noise.

Dual Apodisation with Cross-Correlation (DAX) has shown promise in reducing speckle noise in medical ultrasound. Introduced by Seo and Yen [56], DAX uses two apodisation functions alongside basic Delay and Sum beamforming to create opposing signals with out of phase grating lobes, which are then cross-correlated.

Seo proposed four different apodisation methods in his paper: Alternating, Common Midpoint, Random and Hamming Windowed. The alternating method created two opposing apertures consisting of the odd and even numbered apertures respectively. The random method selects a random set of $\frac{n}{2}$ elements from the n element array. The common midpoint method uses the central array elements in both apertures and assigns the elements from each end of the array their respective signals. The windowing method applies a window across all of the array elements for one set of signals and uses a rectangular window across the other set. A high correlation coefficient at a certain point means that both of the signals are similar and that the signal at the point is likely to be from a reflector. Conversely, a low correlation coefficient will mean that the signals are unrelated and are likely to be either contributions from the sidelobe, or speckle noise. DAX has shown promise in medical ultrasound imaging and is the inspiration for the technique presented in this chapter.

The motivation behind developing a new method is straightforward. The aforementioned techniques are designed to be applied using an approach based on the focused B-scan imaging methodology. The Full Matrix Capture records a larger volume of data that can be used in post-processing and TFM-based imaging techniques can be applied that will offer a higher resolution and SNR while reducing the point spread when comparing to focused B-scan inspired techniques[79].

4.2 Theory

4.2.1 Overview

This chapter presents a novel imaging algorithm designed to reduce the contributions from sidelobes when inspecting ultrasonically noisy materials. It has been called Sub-Aperture Spatially Averaged Correlation Imaging (SASACI) and combines the focusing qualities of TFM with the sidelobe reducing properties of DAX[56].

Any ultrasound signal received from a pulse-echo measurement can be thought of as the sum two discrete signals, reflections from the main lobe and reflections from sidelobes. The energy reflected from sidelobes is considered to be noise which will reduce the image contrast and possibly add speckle to an image. SASACI minimises the contributions from the reflection of sidelobes. To do this, contributions from the main lobe must be differentiated from contributions from reflections in the clutter region. As seen from the flowchart in Figure 4.1, SASACI operates on an FMC data set and is strongly linked to the Total Focusing Method.

4.2.2 SASACI

SASACI is based on the same principle as TFM, introduced in Section 2.3.2.1, but aims to extract additional information from the same FMC data. SASACI creates an image for every transmitting element for which data has been acquired. The set of images is generated using Equation 4.1, which uses the same general

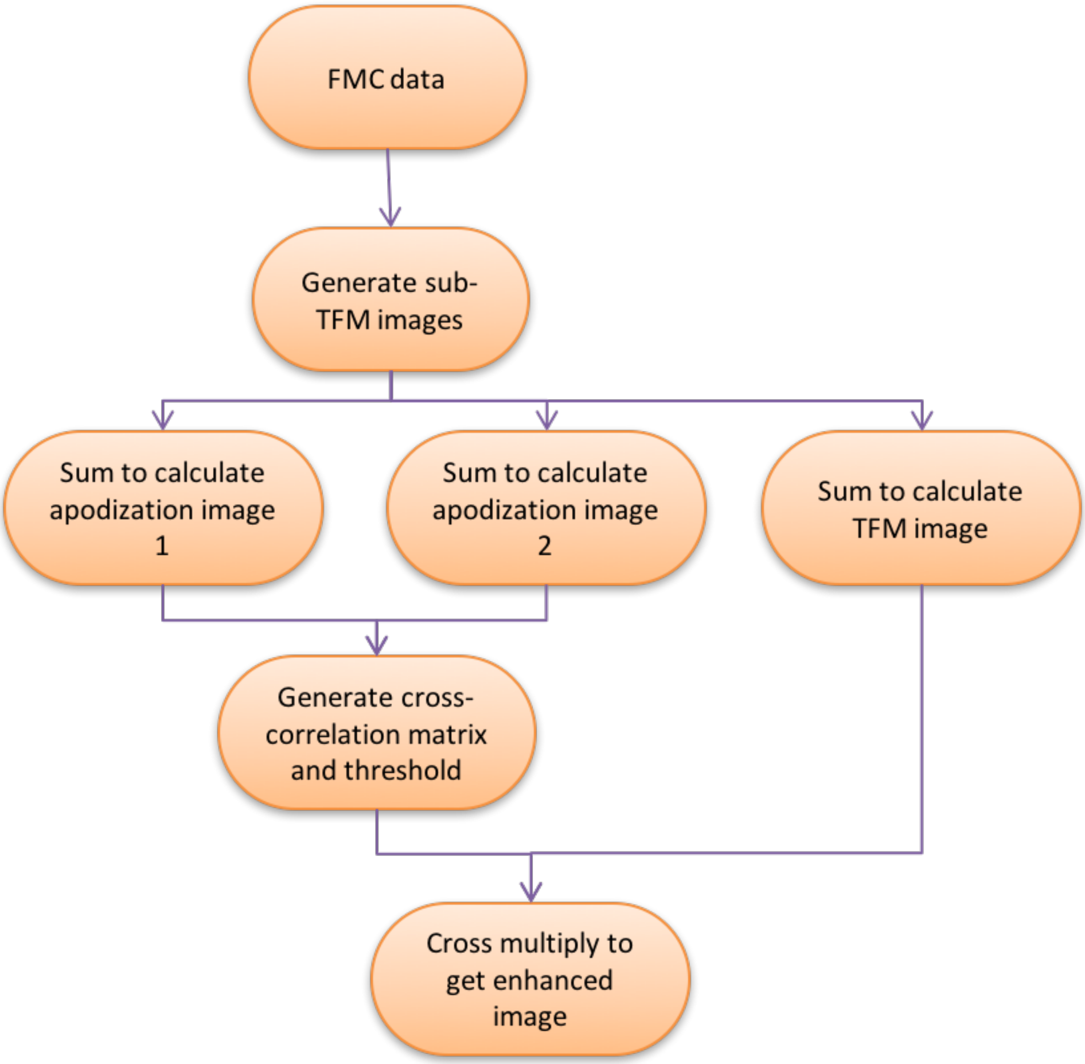


Figure 4.1: *SASACI Methodology*

concept as the Total Focusing Method. The only difference between Equations 2.60 and 4.1 is that the data from the transmitting elements will contribute to individual images, as demonstrated by the variable $SubTFM_{tx}$ which is a matrix of images. These images will be referred to as Sub-TFM images and the sum of all these images will be equal to a standard TFM image. Figures 4.2 and 4.3 show a graphical representation of this concept. In Figure 4.2, a TFM image is generated by summing the contributions for each transmit-receive pair of elements. In Figure 4.3, a separate image is generated for each element, using every contribution from individual transmitting elements for each separate image.

$$SubTFM_{tx}(y,z) = |\sum h_{tx,rx}(\frac{\sqrt{(y_{tx}-y)^2+z^2}+\sqrt{(y_{rx}-y)^2+z^2}}{v_L})| \tag{4.1}$$

In the SASACI process (see Figure 4.1), these Sub-TFM images are summed together to create two new images. One of the ways to do this is to sum all the odd-numbered Sub-TFM images to create the first new image, and all of the even-numbered Sub-TFM images to create the second new image. Section 4.4 explores methodologies for summing the Sub-TFM images in more detail and proposes a number of differing ways to apply a weighting to the images prior to summation.

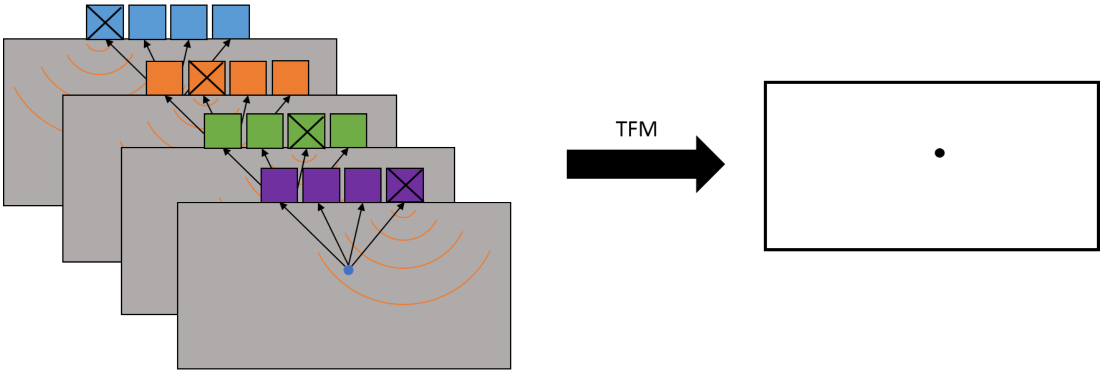


Figure 4.2: *A simple illustration of the TFM imaging algorithm*

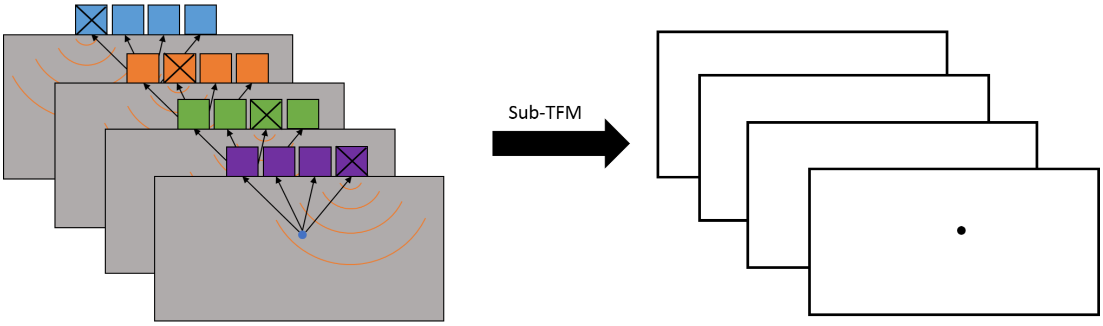


Figure 4.3: *A simple illustration of how Sub-TFM images are generated*

4.2.3 Cross-Correlation

Once two differing images are generated, they are input to a cross-correlation function, as shown in Equation 4.2, which generates a matrix of scalars that range between -1 and 1, depending on the level of correlation.

When the images are weighted optimally, any reflections from defects and structural features will be in the same position in each image. Noise from grain structure reflections and multipath propagation will differ between the two images and can therefore be reduced through appropriate signal processing.

In Equation 4.2, p is the resulting cross-correlation matrix. $RX1$ and $RX2$ are the matrices to be cross-correlated and S defines the number of neighbouring pixels to be analysed for one matrix element in the cross-correlation.

$$p(x,y) = \frac{\sum\limits_{i=x-S}^{x+S} \sum\limits_{j=y-A}^{y+S} (RX1(i,j) \cdot RX2(i,j))}{\sqrt{\sum\limits_{i=x-S}^{x+S} \sum\limits_{j=y-S}^{y+S} RX1(i,j)^2} \cdot \sqrt{\sum\limits_{i=x-S}^{x+S} \sum\limits_{j=y-S}^{y+S} RX2(i,j)^2}} \tag{4.2}$$

S can be calculated by the method shown in Equation 4.3 and is a measure of the number of pixels used for cross-correlation. v_L is the longitudinal propagation velocity, c is the centre frequency of the array, l distance from the centre of each pixel to analyse in wavelengths, and Δx is the pixel pitch of the quantised image measured in meters per pixel. Figure 4.4 shows a graphical representation of the pixels used for the cross-correlation function. l is set to one wavelength in this work.

$$S = \frac{v_L \cdot l}{c \cdot \Delta x} \tag{4.3}$$

Thresholding is also applied at this stage, meaning points that are highly uncorrelated will be further minimised, the thresholding equation being shown in Equation 4.4.

$$W(x,y) = \begin{cases} p(x,y) \cdot H, & \text{if } p(x,y) < X; \\ p(x,y), & \text{otherwise.} \end{cases} \tag{4.4}$$

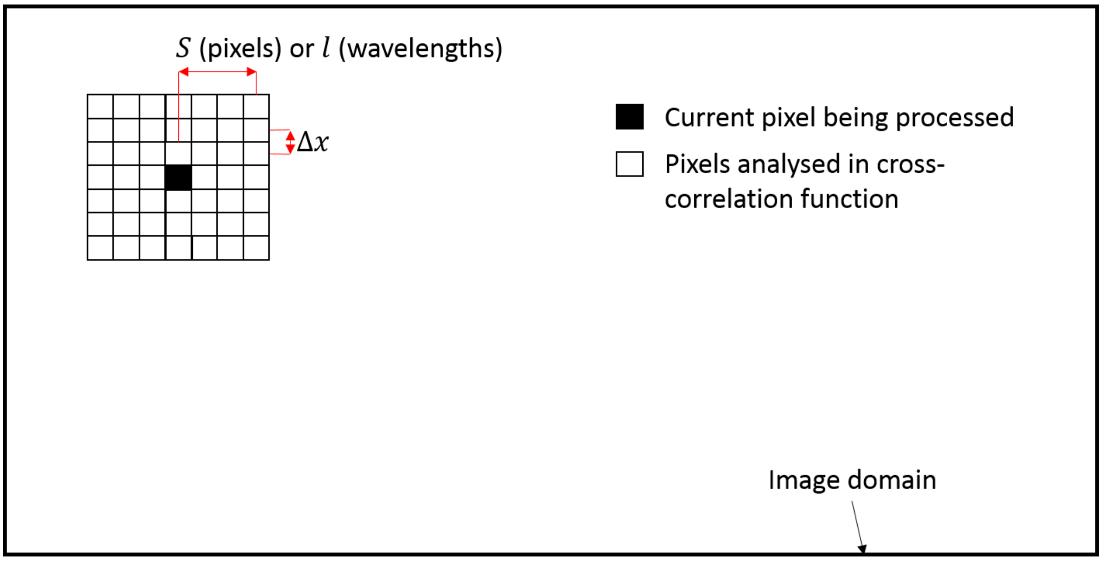


Figure 4.4: *A graphical representation of how S is represented for SASACI*

In Equation 4.4, W is the weighting matrix derived from p, the cross-correlation matrix. X is the threshold and H is the weight applied to any pixel whose value is under the threshold. H was set to 0.01 for all of the experimental results presented in this thesis.

Finally, the weighting matrix W is multiplied by the basic TFM image, as shown in Equation 4.5 where $SASACI$ is the resultant image.

$$SASACI = TFM \cdot W \tag{4.5}$$

The final SASACI image has the same properties as the TFM image with regards to sensitivity and resolution but should also effectively minimise reflections from sidelobes.

4.3 Experimental Evaluation

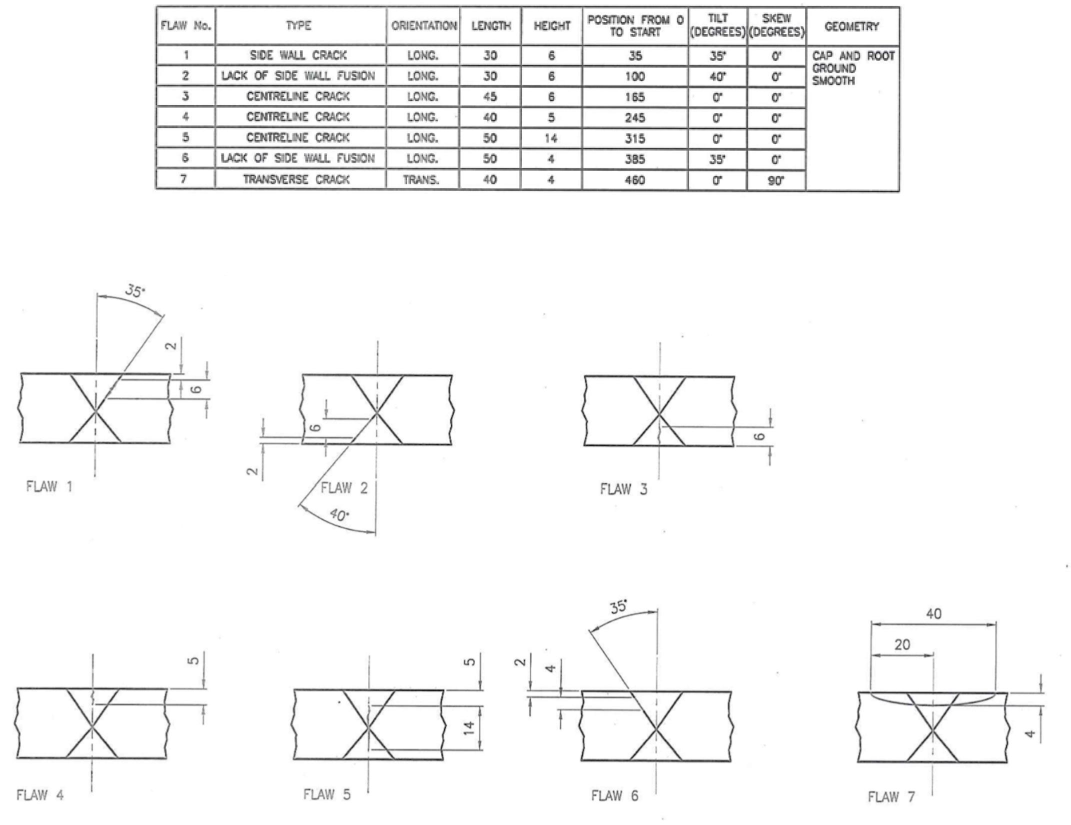
4.3.1 Steel Weld

A welded 316L stainless steel sample (Amec Foster Wheeler, UK) was inspected. This sample has a number of defects within the weld and each of these defects was inspected in turn. A photograph of the sample is shown in Figure 4.5 and a scan of the datasheet detailing the flaws is shown in Figure 4.6.

Figure 4.5: *A photograph of the welded plate sample*

4.3.1.1 Experimental Setup

Data was collected from this sample using a Dynaray phased array controller (Zetec, USA) and a sub-aperture of a 128 element commercial linear array (Vermon, France). Table 4.1 shows the parameters used for this experiment.



FLAW No.	TYPE	ORIENTATION	LENGTH	HEIGHT	POSITION FROM 0 TO START	TILT (DEGREES)	SKEW (DEGREES)	GEOMETRY
1	SIDE WALL CRACK	LONG.	30	6	35	35°	0°	CAP AND ROOT GROUND SMOOTH
2	LACK OF SIDE WALL FUSION	LONG.	30	6	100	40°	0°	
3	CENTRELINE CRACK	LONG.	45	6	165	0°	0°	
4	CENTRELINE CRACK	LONG.	40	5	245	0°	0°	
5	CENTRELINE CRACK	LONG.	50	14	315	0°	0°	
6	LACK OF SIDE WALL FUSION	LONG.	50	4	385	35°	0°	
7	TRANSVERSE CRACK	TRANS.	40	4	460	0°	90°	

Figure 4.6: *Excerpts from the datasheet relating to AMEC's 316L welded sample, showing each defect's type, shape and location within the weld.*

Parameter	**Value**
Array Centre Frequency	5 MHz
Array Aperture	32
Array Pitch	0.7 mm
Longitudinal Wave Velocity	5560 m/s
Filter Passband	0.5 - 10 MHz
SASACI Threshold	0.95

Table 4.1: *Parameters for stainless steel weld inspection*

Due to the sample not being completely flat, it was not possible to directly couple all 128 elements of the array to the sample above each of the defect locations. Instead, a 32 element sub-aperture of the array was used to collect a Full Matrix Capture dataset. For each defect, the array sub-aperture was centred over the defect of interest.

Two defect-free datasets were also recorded; one above the weld and another in a clean, non-welded, region of the stainless steel sample. These were used as baseline to which the other results can be compared.

4.3.1.2 Results

Results from each defect will be presented individually as a series of three images. The leftmost image is the result of the Total Focusing Method on the otherwise unprocessed data. Regions with defects will show the defect location overlaid on this image. The centre image is the TFM result after being applied to the filtered Full Matrix Capture data using the bandpass filter specified in Table 4.1. The image on the right is the result of the SASACI process. The input to the SASACI algorithm is a filtered FMC dataset, and the filtered TFM image has been included as a fair comparison between the two methods.

Figure 4.7 shows the three images from the non-welded area of the stainless steel sample away from the weld. Figure 4.8 shows the images from the weld, but from a region free of defects.

In these sets of images, the backwall can be seen clearly at a depth of approximately 22mm with no other features in the images. For areas with no defects, it can be observed that SASACI does not offer an improvement compared to TFM on a filtered dataset. This is an expected result as SASACI only tends to improve the images from datasets collected in ultrasonically noisy specimens. Figures 4.9 to 4.15 show the results for defects numbered 1 to 7. These results are presented and discussed in this section.

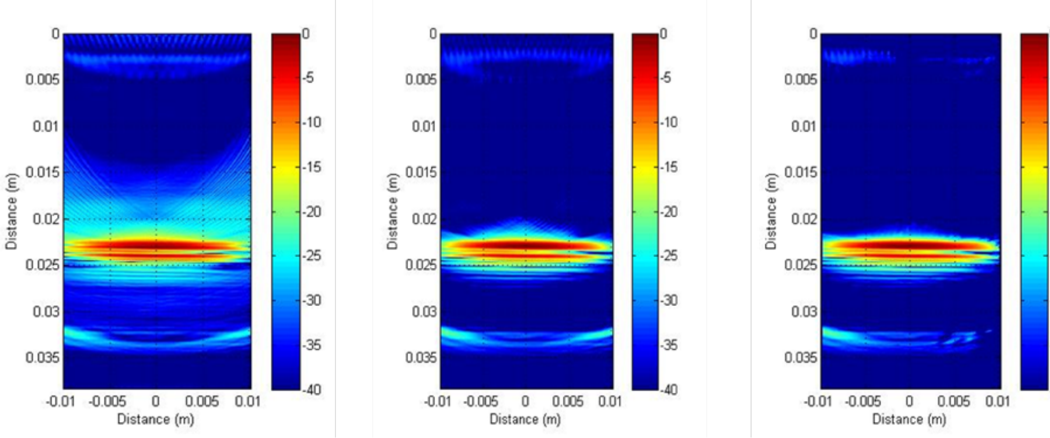


Figure 4.7: *TFM, Filtered TFM and SASACI images of a defect-free, non-welded area of the stainless steel sample.*

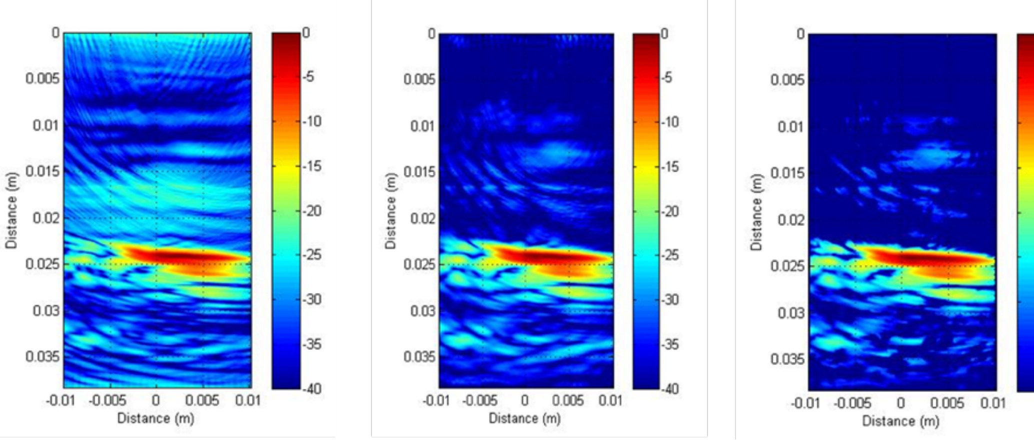


Figure 4.8: *TFM, Filtered TFM and SASACI images of a defect-free area of the weld in the stainless steel sample.*

4.3.1.2.1 Flaw 1

Flaw 1 is a side-wall crack angled at 35° with respect to the normal. It has a length of 30mm, is 6mm tall and the processed images of this flaw can be seen in Figure 4.9. The reflection from the flaw is strong and the defect can be seen even in the unfiltered TFM image. As expected, the filtered TFM image reduces a significant portion of the noise and allows the defect to be seen more clearly. The SASACI process reduces this noise (measured by taking the mean amplitude in the highlighted area - this area is consistent throughout the following results) by a further 5dB, though the defect is already visible in the previous image and accurate sizing is possible.

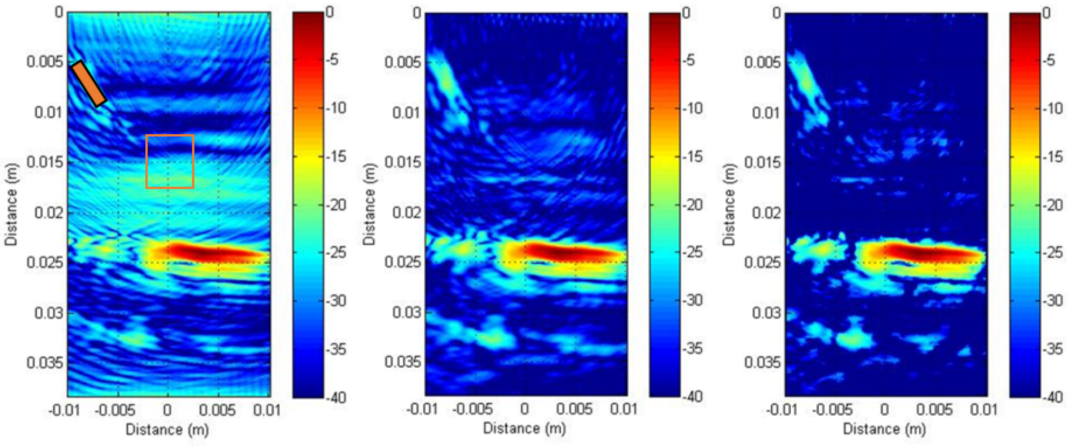


Figure 4.9: *TFM, Filtered TFM and SASACI images of Flaw 1 within the stainless steel weld. The side wall crack can be seen clearly at the top left of all three images.*

4.3.1.2.2 Flaw 2

Flaw 2 is a lack of side wall fusion. It is angled at 40° with respect to the normal and has a length and height of 30mm and 6mm respectively. The image results of this flaw are shown in Figure 4.10. It can be immediately observed from the first of the three images that this section of the weld has much less grain noise compared to the section in which Flaw 1 is located. The background noise level is significantly lower and the defect can be clearly seen. The filtered TFM reduces some of the noise surrounding the backwall and SASACI is able to further reduce some of the noise surrounding the defect.

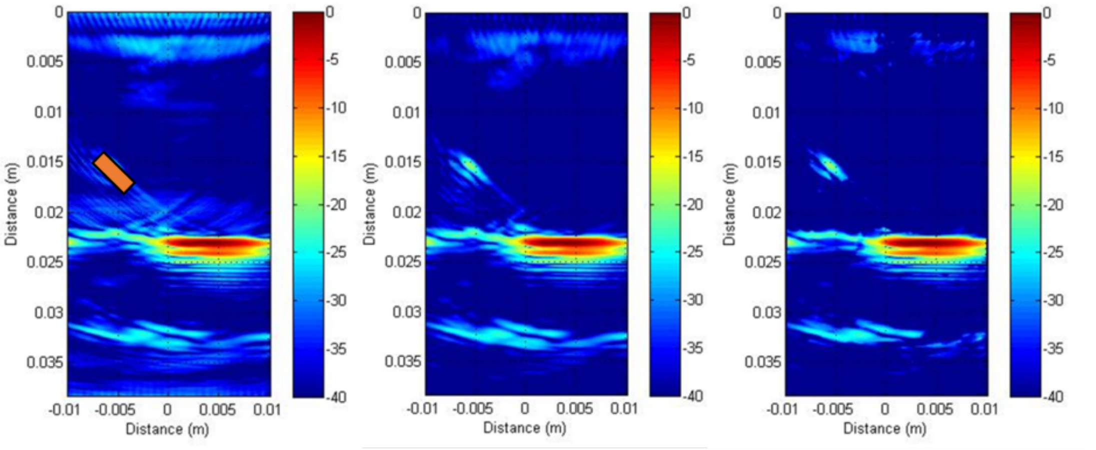


Figure 4.10: *TFM, Filtered TFM and SASACI images of Flaw 2 within the stainless steel weld. This defect is a lack of side wall fusion and can be seen approximately 6mm above the back wall on the left hand side of each of the three images.*

4.3.1.2.3 Flaw 3

Flaw 3 is a vertical centreline crack within the weld. It is 45mm long and 6mm tall. It starts at the bottom of the sample and protrudes upwards. Figure 4.11 shows images of the region in which the flaw is located. No useful information can be observed in the unfiltered TFM image. There is a significant amount of noise surrounding the back wall which hinders analysis of the image. The filtered TFM image reduces noise but it is still not clear where the defect is located. The SASACI image shows further reduced noise, to a point where the top of the defect can be observed. It should be stated that the experimental setup is not optimal for detecting Flaws 3, 4 and 5. To detect vertical flaws, inspection should take place with an angled beam which allows for more energy to be reflected from the defect. A wedge would allow for this.

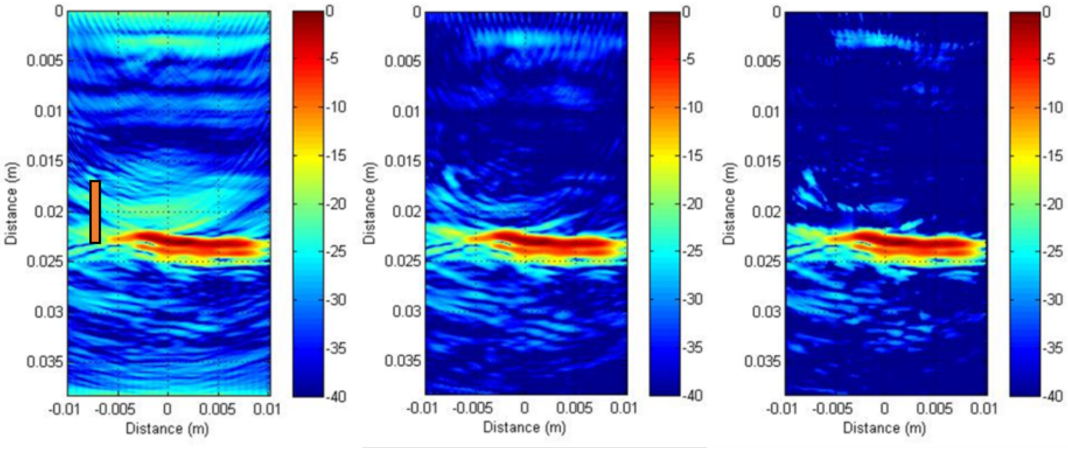


Figure 4.11: *TFM, Filtered TFM and SASACI images of Flaw 3 within the stainless steel weld. The centreline crack is expected to be visible in the left hand side of the images and appear as an artefact just above the back wall reflection. It is more prominent in the SASACI image.*

4.3.1.2.4 Flaw 4

Flaw 4 is also a vertical centreline crack. It is smaller than Flaw 3, being 40mm long and only 5mm tall. This flaw is starts at the top of the sample and protrudes downwards towards the centre of the weld. This flaw cannot be seen in any of the three images shown in Figure 4.12.

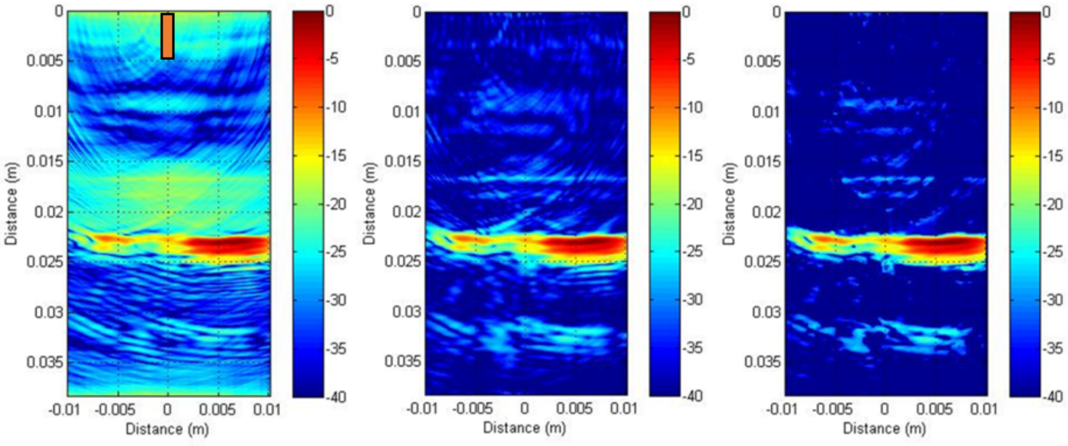


Figure 4.12: *TFM, Filtered TFM and SASACI images of Flaw 4 within the stainless steel weld. Flaw 4 is a centreline crack propagating from the top of the weld. Although this defect is not visible in any of the three images, the effects of the crack can be seen as second and third reflections are visible at depths of 8mm and 16mm.*

4.3.1.2.5 Flaw 5

Flaw 5 is a vertical centreline crack in the centre of the weld. It is larger than the previous flaws and is 50mm in length and 14mm in height. The top of the crack can be seen in the unfiltered TFM image in Figure 4.13, however the bottom cannot be seen due to the noise around the back wall. The filtered image reduces noise, allowing the top of the crack to be seen clearly. In this case, the bottom of the crack still cannot clearly be seen due to noise. The SASACI image reduces this noise further to a point where the location of the bottom of the crack can be estimated to allow sizing.

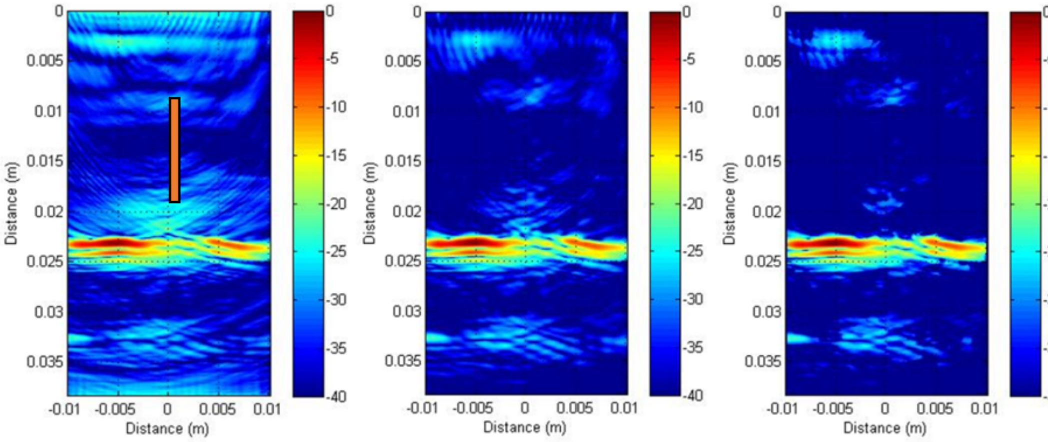


Figure 4.13: *TFM, Filtered TFM and SASACI images of Flaw 5 within the stainless steel weld. Flaw 5 is a centreline crack that starts just below the surface and ends just above the back wall of the specimen. The top of the crack can be seen at a depth of 8mm and the bottom is visible at a depth of 20mm. These reflections are most clear in the SASACI image.*

4.3.1.2.6 Flaw 6

Flaw 6 is a lack of side wall fusion and is at an angle of 35° with respect to the normal. The crack is sized 50mm in length and 4mm tall. Figure 4.14 shows the three images related to this defect. The unfiltered TFM image is very noisy due to the grain structure within the weld. The bandpass filtering reduces noise to the point where the defect can be seen above the noise and SASACI further reduces this noise by around 10dB.

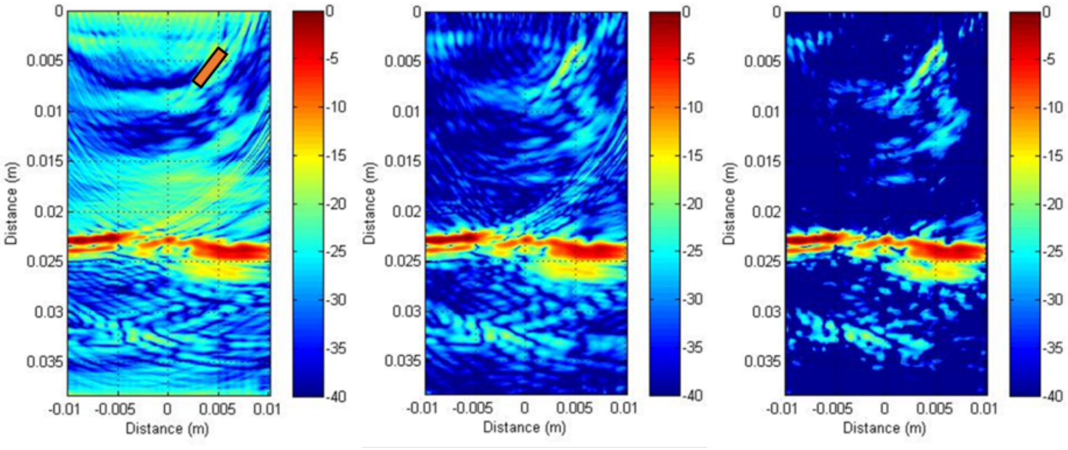


Figure 4.14: *TFM, Filtered TFM and SASACI images of Flaw 6 within the stainless steel weld. This defect is a lack of side wall fusion and can be seen at the top-right of both the filtered and SASACI images. The defect is clearest in the SASACI image due to the reduced noise in the image.*

4.3.1.2.7 Flaw 7

Flaw 7 is a transverse crack at the top of the sample and is 40mm long and 4mm in height. This defect cannot be seen in any of the three images represented in Figure 4.15. The filtered TFM and SASACI images show clear horizontal artefacts that are also present in Flaw 4. While the defects near the top of the sample cannot be seen in any of the images, it is possible that the flaws are contributing to these artefacts. The orientation of this defect is such that it is expected to be found. There is currently no explanation for the fact that the defect cannot be seen in any of the images.

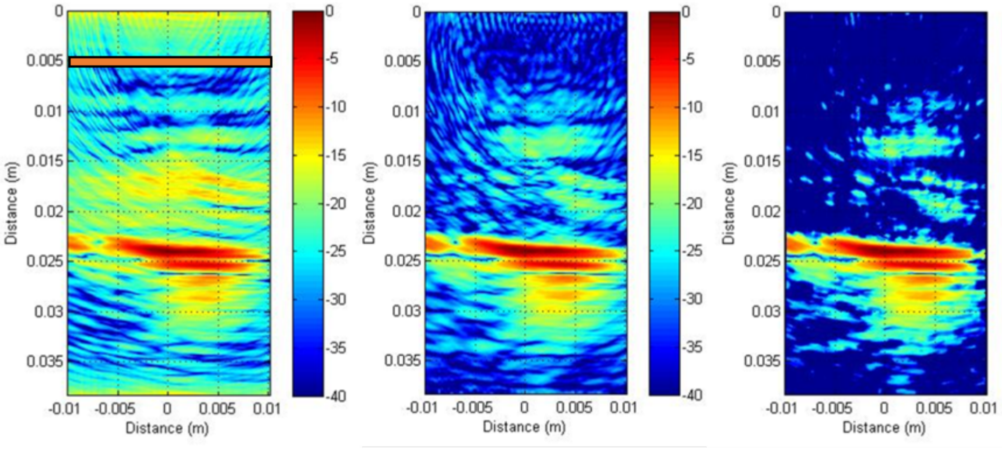


Figure 4.15: *TFM, Filtered TFM and SASACI images of Flaw 7 within the stainless steel weld. Flaw 7 is a transverse crack propagating 4mm into the material from the top surface. It is not visible in any of the three images, though the effects can be seen as a large artefact is visible at a depth of around 13mm in both the filtered and SASACI images.*

4.3.2 Dissimilar Weld

A dissimilar weld specimen (TWI, UK) with implanted flaws was inspected. The specimen is shown in Figure 4.16. This specimen has a number of defects, the locations of which can be seen marked on the surface of the sample. In this section only one is investigated for the evaluation of this algorithm.

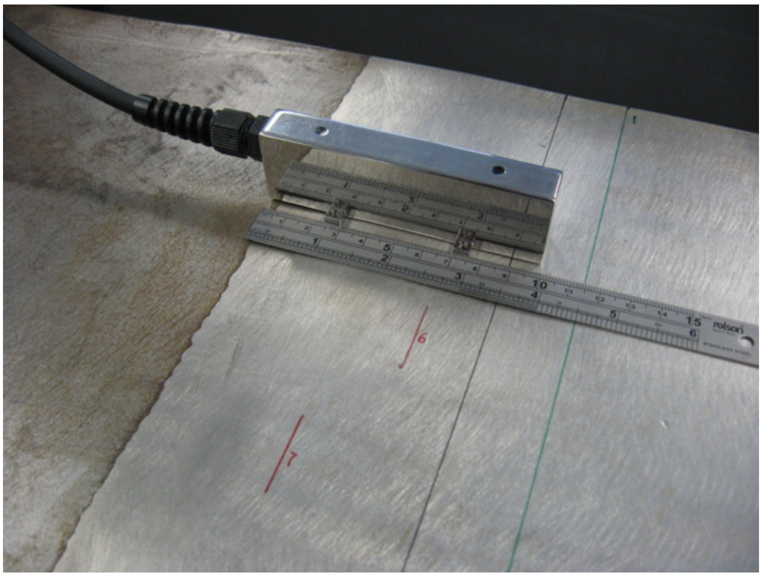


Figure 4.16: *A dissimilar weld specimen*

4.3.2.1 Experimental Setup

The defect focused upon for this experiment was a 20mm long vertical crack that began approximately 35mm from the top of the sample, protruding downwards.

A commercial 128 element linear 5MHz array (Vermon, France) was used to record an FMC dataset using a 45 element sub-aperture of the array. A sub-aperture was used because, similar to the previous experiment, the specimen was not flat and a contact inspection using the full aperture of the array was not possible. The array was driven by a Dynaray phased array controller (Zetec, USA). Table 4.2 shows additional experimental parameters.

Parameter	**Value**
Array Centre Frequency	5 MHz
Array Aperture	45
Array Pitch	0.7 mm
Longitudinal Wave Velocity	5700 m/s
Filter Passband	N/A
SASACI Threshold	0.95

Table 4.2: *Parameters for stainless dissimilar weld inspection*

4.3.2.2 Results

The Full Matrix Capture dataset was input to both the TFM and SASACI algorithms and the results compared. The TFM and SASACI results are shown in Figures 4.17 and 4.18 respectively. The images are shown with a 30dB dynamic range. The SNR was measured in each of these images. The signal was defined as the area where the back wall and extremities of the crack were expected to be, and the noise was defined as all other areas. Table 4.3 shows the calculated signal to noise ratio for each of the algorithms.

Algorithm	**SNR**
TFM	7 dB
SASACI	20 dB

Table 4.3: *Signal to Noise Ratio for the dissimilar weld sample*

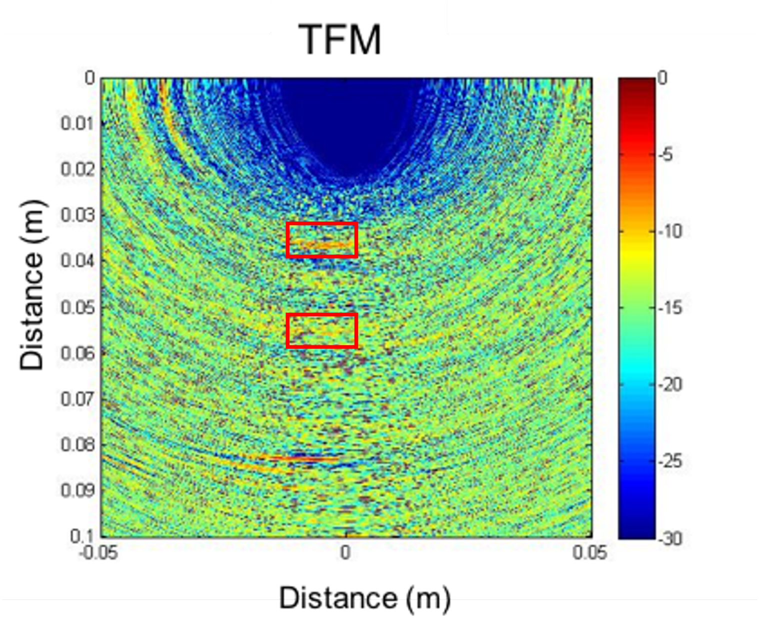


Figure 4.17: *TFM result from the dissimilar weld specimen. The region where the reflections from the top and bottom of the crack are expected to appear are highlighted in red.*

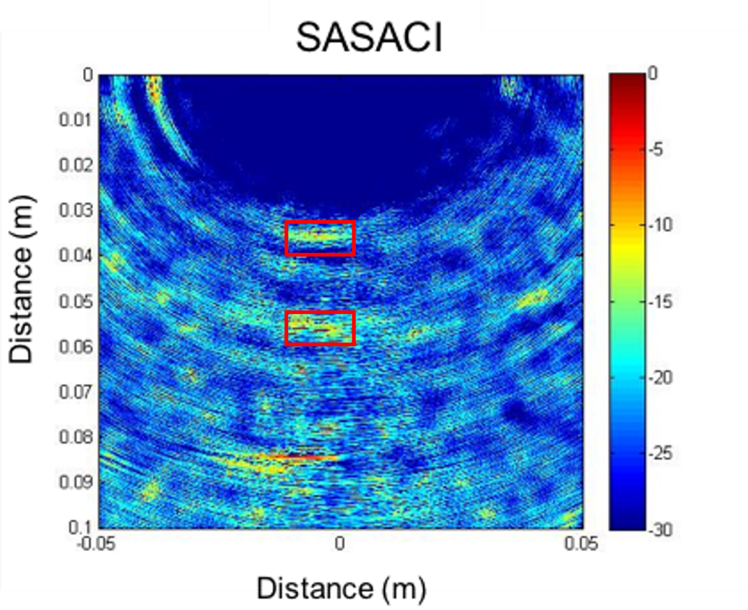


Figure 4.18: *SASACI result from the dissimilar weld specimen. The region where the reflections from the top and bottom of the crack are expected to appear are highlighted in red.*

In each of the images, the back wall can be clearly seen at a depth of 85mm. The top of the vertical crack can also be seen in both the TFM and SASACI images, at a depth of 35mm and centred. The SASACI algorithm reduces the background noise of the image to the point where the bottom of the crack can also be observed. This is visible at a depth of 55mm in the SASACI image and is not visible in the TFM image. These regions have been highlighted via a red rectangle in each of the images.

Although the noise is greatly reduced in the SASACI image, there are a number of remaining areas with an amplitude response close to that of the defect. Tuning of the SASACI parameters may reduce the amplitude of these indications.

4.4 Apodisation Analysis

In this section, different apodisation methods will be evaluated using results from experimental data to fine-tune the parameters of the algorithm. Five different apodisation techniques were trialled and used to weight an FMC dataset in order to create different sets of sub-TFM images.

4.4.1 Experimental Setup

The SASACI algorithm was implemented in CUDA C++ running on NVidia graphics cards using the efficient imaging methods described in Chapter 3. An Inconel 625 step wedge was provided by Siemens (via E.ON) for these tests. An image of the specimen is shown in Figure 4.19.

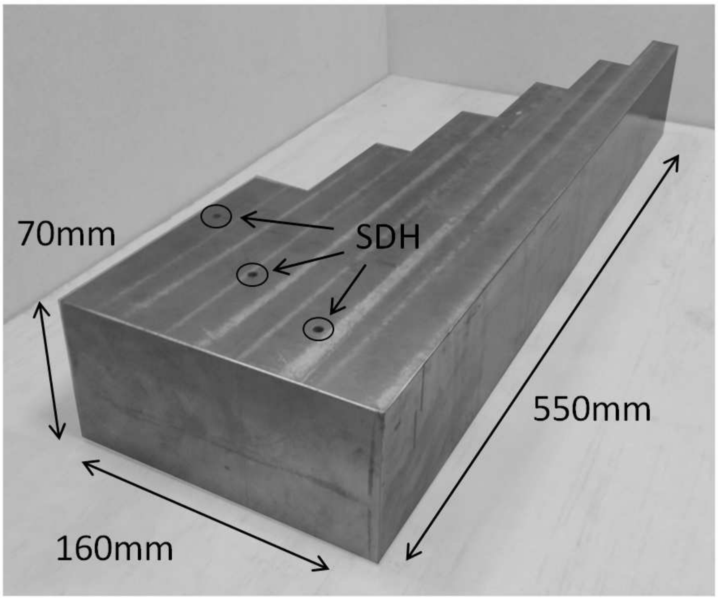


Figure 4.19: *Inconel 625 Specimen*

A commercial array (Vermon, France) was used to record ultrasonic data. The parameters are shown in Table 2.1. The array was driven with a Dynaray phased array controller (Zetec, USA). The Dynaray's set-up parameters are shown in Table 4.4.

Parameter	Value
Pulse Voltage	50 V
Hardware Gain	40dB
Pulse Width	140 ns
Sampling Frequency	100 MHz

Table 4.4: *Dynaray setup parameters for step wedge inspection*

Due to the size of the sample, it is not possible to generate an image from a single scan that shows all of the sample's features. Instead, a single feature will be selected. The feature that the experimental results will concentrate upon is the Side Drilled Hole (SDH) that is at a depth of 60mm from the top of the specimen. This is the middle hole in Figure 4.19.

The FMC dataset also had a bandpass filter applied to remove noise in the pre-processing state. The applied filter was a second order Butterworth bandpass filter with -3dB cut-offs at 2MHz and 6MHz. A Butterworth filter was chosen due to its flat frequency response in the passband.

4.4.2 Experiment and Results

4.4.2.1 Benchmark TFM

The resultant image from the experiment defined above is shown in Figure 4.20 with a dynamic range of 40dB. In Figure 4.20, the SDH is central and is at a depth of 60mm. Due to multipath propagation and the fact that the speed of sound is not constant in the material, the hole appears as a line in the image. This is a limitation of the imaging process that SASACI cannot overcome. It is evident that the longitudinal propagation velocity is set correctly, as another SDH can be seen in Figure 4.20 at a depth of 10mm. This feature is resolved correctly in the image due to the fact that is is close to the surface and the sound energy has less distance to travel and therefore less opportunity for error to be accumulated in estimating the distance travelled.

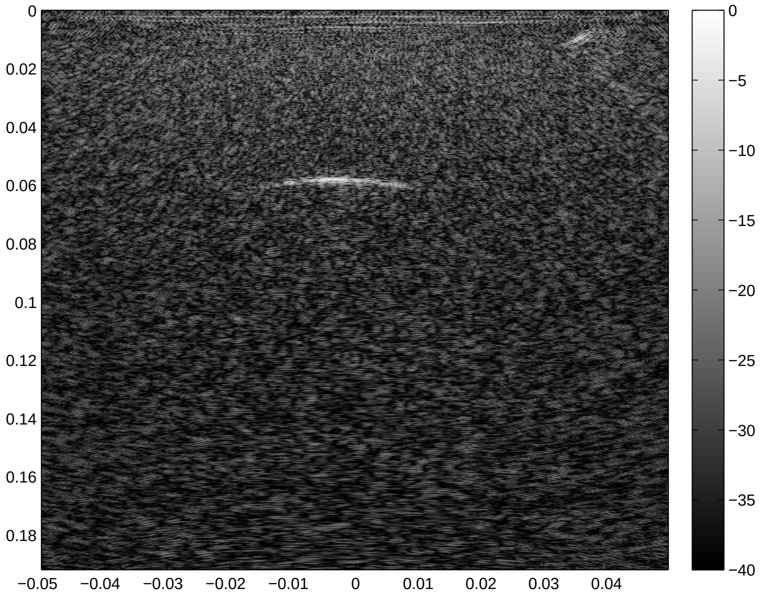


Figure 4.20: *A Conventional TFM Image*

There are a number of different ways to weight the sub-TFM images before combining them. Seo and Yen[56] investigated four methods of doing this for a one-dimensional cross-correlation approach. Each of the approaches, as well as a novel hamming windowing approach, will be evaluated for the current methodology. The Hamming window is designed to reduce the amplitude of the sidelobe closest to the main lobe, but at the expense of the rest of the side lobes being much higher when compared to a Hann window.

For the SASACI processing a number of variables need to be set, as described in Section 4.2.2. These values are defined in Table 4.5 and remain constant in all the results presented in this Chapter. The value of A was chosen to compare 2 wavelengths of data for the cross-correlation algorithm. Multiple values of H

and T were investigated. The values in Table 4.5 were used in this experiment as they produce an optimal image. Future work will involve generating an adaptive processing methodology to set these values algorithmically.

Parameter	**Value**
X	0.9
H	0.1
S	6
l	1

Table 4.5: *SASACI Parmeters*

4.4.2.2 Alternating Elements Apodisation

Figure 4.21 shows the weighting applied to the transmitting elements to create the modified aperture. Although 32 elements are shown, any number of elements can be used with this method and the pattern expanded or reduced as necessary. In this experiment, a 128 element array was used and the weighting shown is simply repeated as necessary.

The SASACI image after using the weighting shown in Figure 4.21 is shown in Figure 4.22. Compared to the TFM image there is reduced speckle noise and the second, shallow SDH appears clearer.

4.4.2.3 Hamming Window Apodisation

Hamming windowing is an apodisation technique untested by Seo and Yen, but is thought to have potentially improved results due to it being designed to reduce the nearest sidelobe to the main lobe. The hamming window weightings are shown in Figure 4.23 and the resultant SASACI image in Figure 4.24.

4.4.2.4 Hann Window Apodisation

Compared to the Hamming window, the Hann windowing function has a much steeper frequency roll-off and will therefore minimise sidelobes. The Hann window weightings are shown in Figure 4.25 and the results from this apodisation method shown in Figure 4.26.

4.4.2.5 Common Midpoint Apodisation

Figure 4.27 shows the apertures used for the common midpoint apodisation. Essentially, the two apertures have a number of elements disabled at opposing sides. This is done to create a pair of images where the position and amplitude of noise differs between them. The result image is depicted in Figure 4.28.

4.4.2.6 Random Element Apodisation

The random element apodisation weighting is shown in Figure 4.29. This method has the limitation of yielding differing results each time the random element permutations are chosen. Is is possible to generate permutations over and over until an preferred set of elements are found and this research could be completed in future. A random (non-optimised) result is shown in Figure 4.30

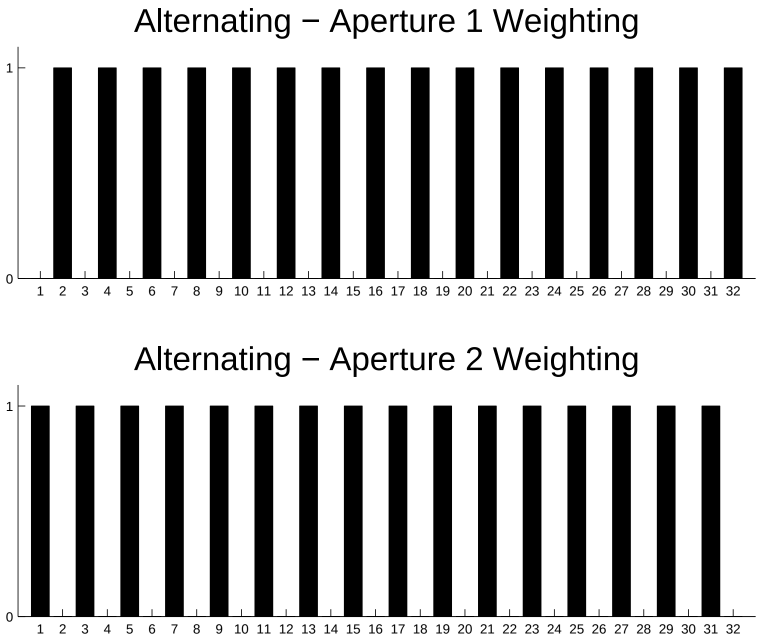


Figure 4.21: *Alternating Apodisation*

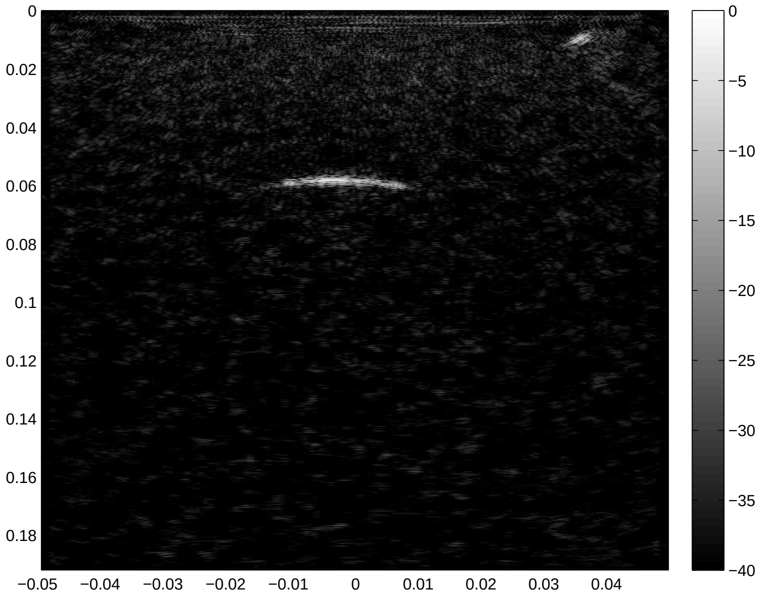


Figure 4.22: *Alternating Apodisation SASACI Result*

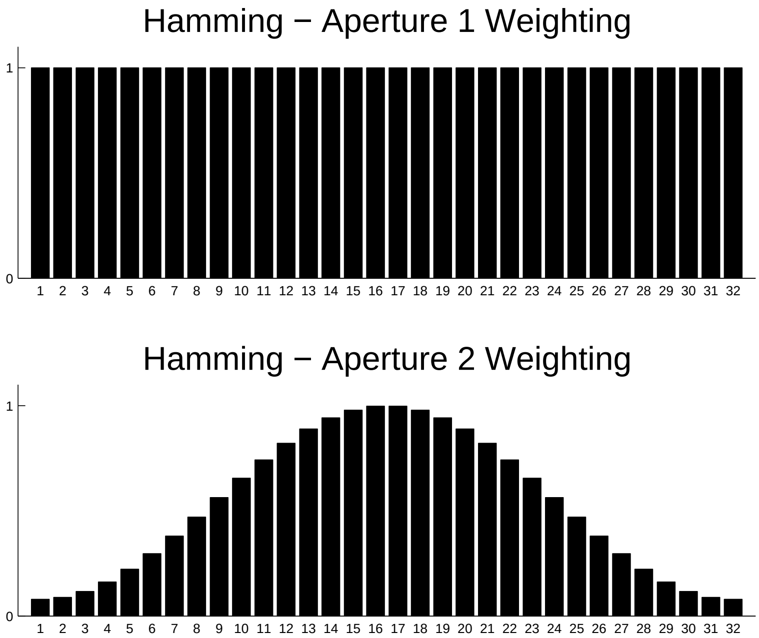


Figure 4.23: *Hamming Apodisation*

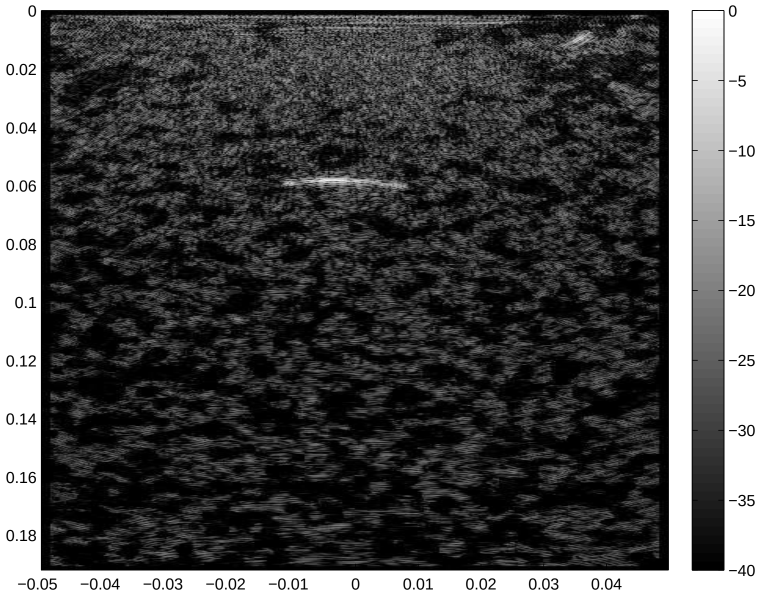


Figure 4.24: *Hamming Apodisation Result*

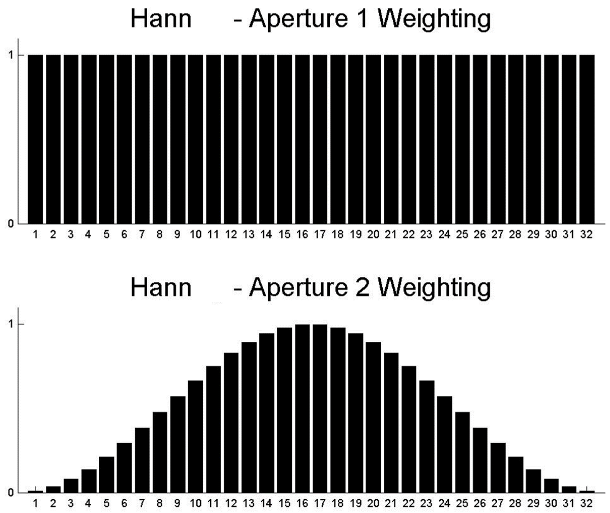


Figure 4.25: *Hann apodisations*

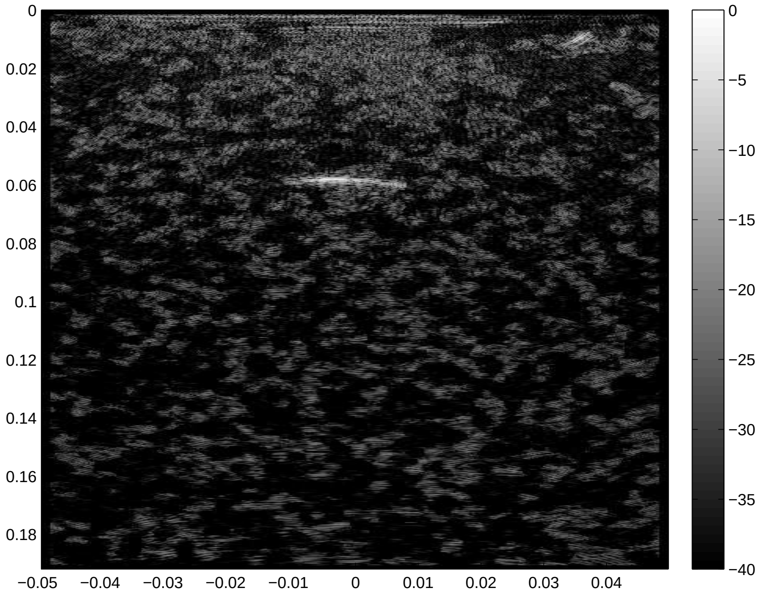


Figure 4.26: *Hann apodisations*

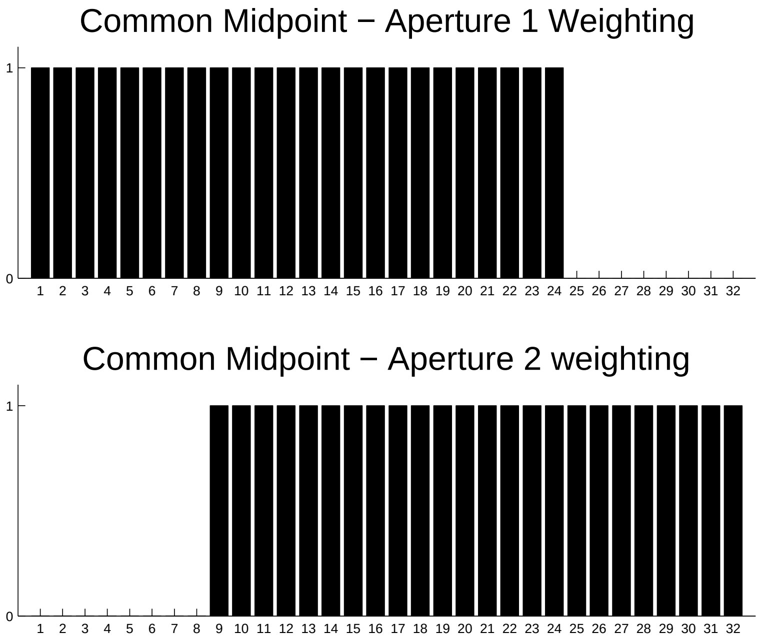


Figure 4.27: *Common Apodisation*

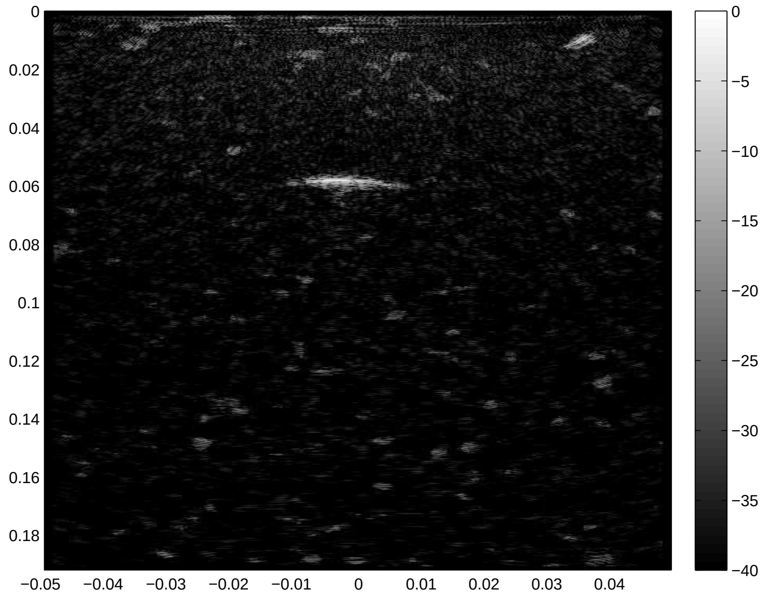


Figure 4.28: *Common Midpoint Apodisation Result*

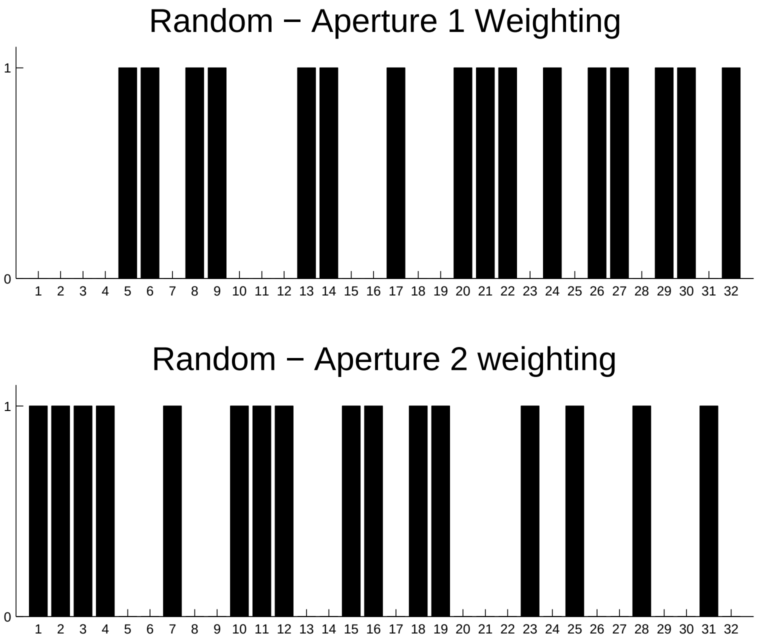


Figure 4.29: *Random Apodisation*

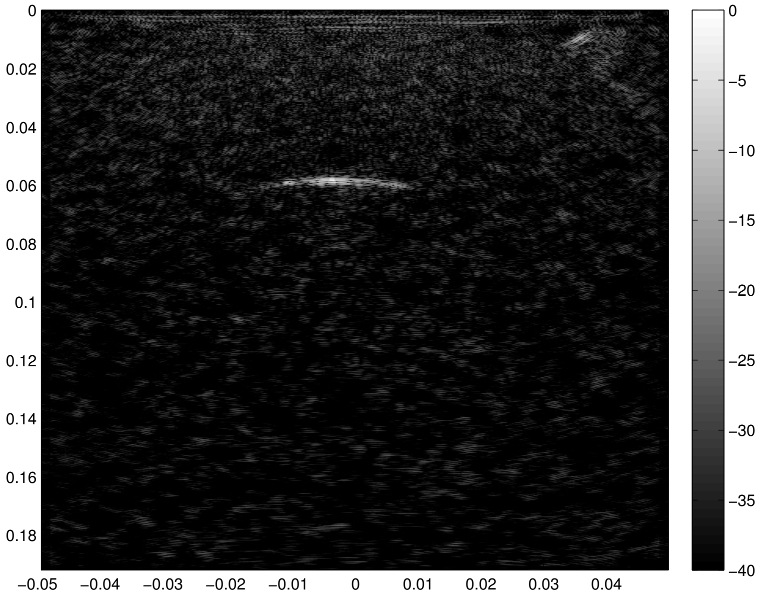


Figure 4.30: *Random Apodisation Result*

4.4.3 Evaluation of Apodisation Methods

The results from the apodisation methods discussed in the Section 4.4.2 were evaluated to quantise their effectiveness. The signal-to-noise ratio (SNR) for each image was measured and tabulated (shown Table 4.6) so that the best of the image weightings could be identified. Using the known position for both side-drilled holes in the test block, the area where the reflection from the holes is above -6dB in the TFM image is considered signal. Everything else is considered noise. This method of differentiating signal and noise was chosen so that the poor focus on the 60mm side drilled hole would not be detrimental to the SNR in either the TFM or SASACI images. Instead, the SNR will be representative of the mean amplitude of the speckle noise in the images. A look-up matrix was generated for the current dataset so that signal and noise can quickly be calculated. This matrix is shown in Figure 4.31, where the black areas are considered signal and the white areas considered noise.

In Table 4.6 it can be seen that the alternating apodisation technique is the best performing in this test scenario. It outperforms TFM by approximately 12dB. The random apodisation method is also effective, but due to the nature of the randomly generated apodisation weights the results differ each time that the algorithm runs. A number of averages could be taken to minimise the random component of these results, but this has not been investigated. Section 6.3.2 discusses this in more detail.

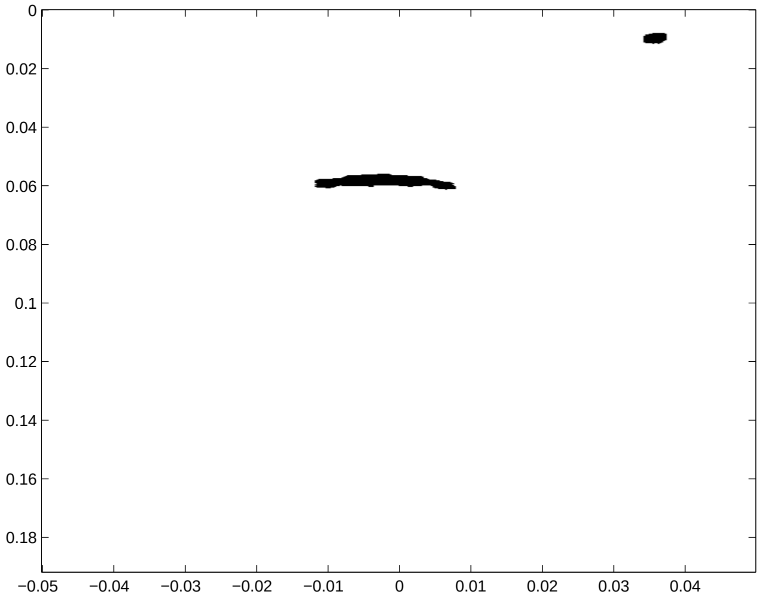


Figure 4.31: *Locations of Signal and Noise*

Method	**SNR**
TFM	27.58dB
Alternating	39.55dB
Common	38.78dB
Hamming	31.80dB
Hann	34.03dB
Random	38.68dB

Table 4.6: *Comparison of Apodisation Methods*

The alternating apodisation method was further investigated by altering the aperture widths. Alternating apertures with aperture widths of 1, 2, 4, 8, and 16 were created and images generated. The aperture shapes and resulting images can be seen in Figures 4.32 to 4.39. The resulting SNRs from these methods are shown in Table 4.7. From this table, it can be observed that there appears to be an optimal aperture width, with SNR becoming poorer either side of the optimal value. In this case, the optimal aperture width was found to be 4 elements wide.

Aperture Width (elements)	**SNR**
1	39.55dB
2	40.29dB
4	40.59dB
8	32.30dB
16	31.30dB

Table 4.7: *Comparison of Alternating Apodisation Aperture Widths*

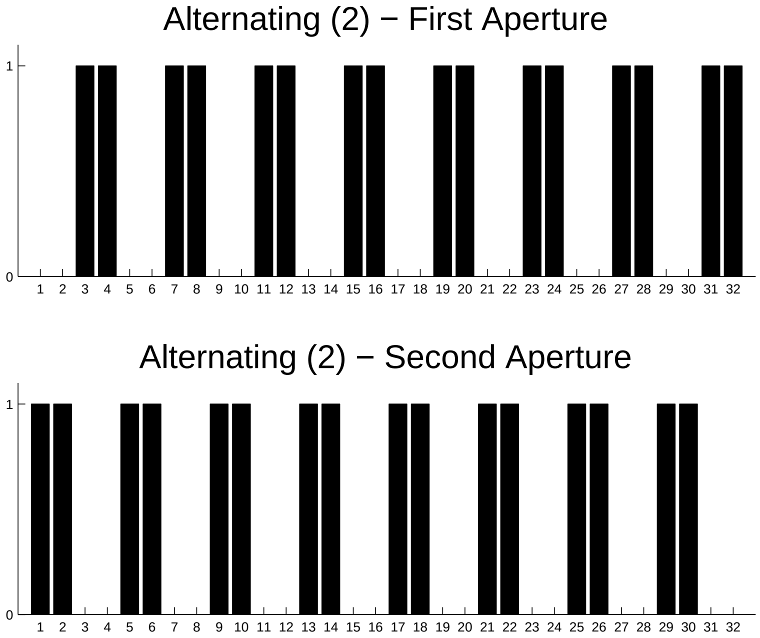


Figure 4.32: *Alternating Apodisation with 2x Aperture Width*

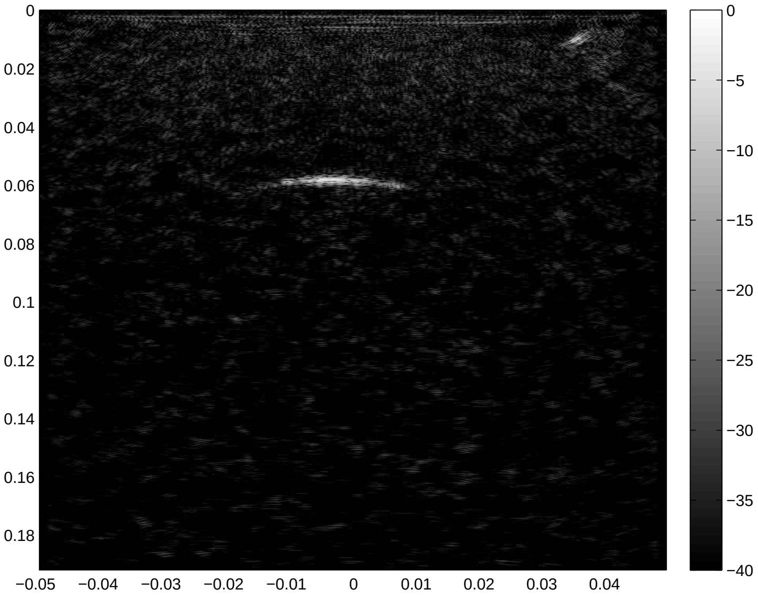


Figure 4.33: *Alternating Apodisation with 2x Aperture Width Result*

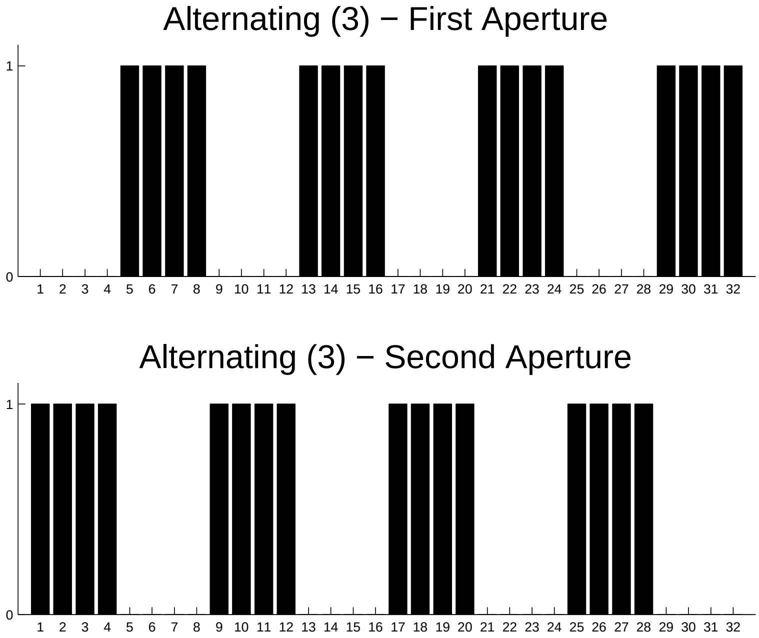


Figure 4.34: *Alternating Apodisation with 4x Aperture Width*

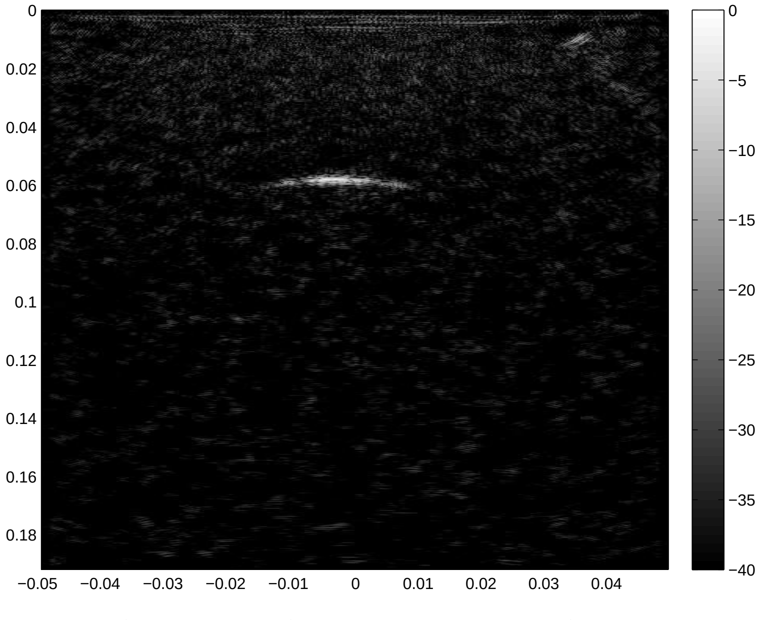


Figure 4.35: *Alternating Apodisation with 4x Aperture Width Result*

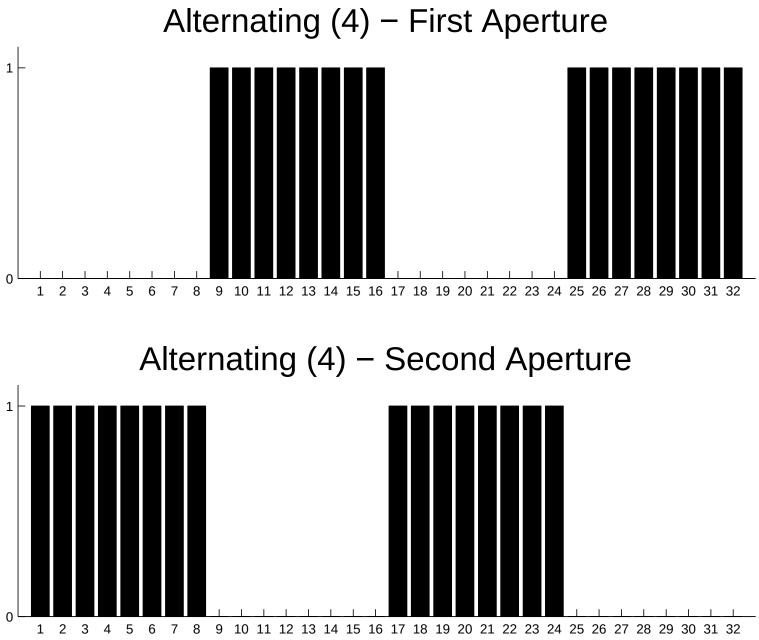


Figure 4.36: *Alternating Apodisation with 8x Aperture Width*

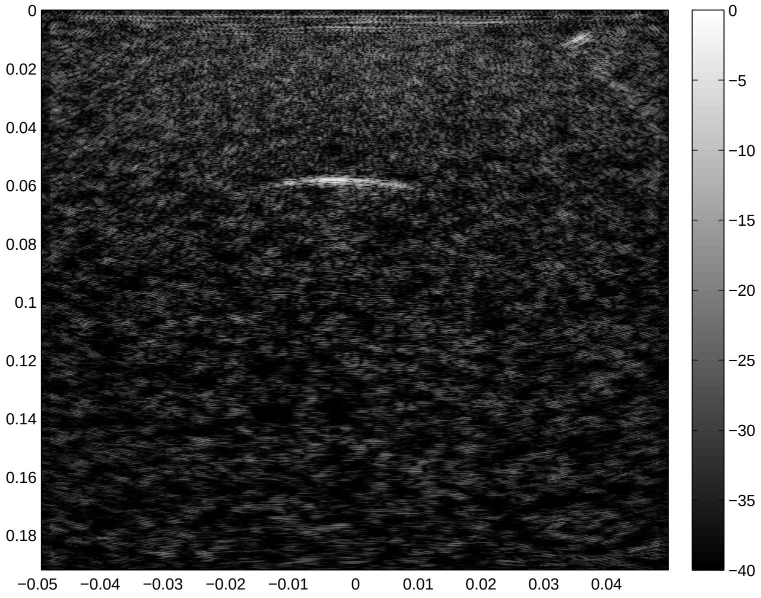


Figure 4.37: *Alternating Apodisation with 8x Aperture Width Result*

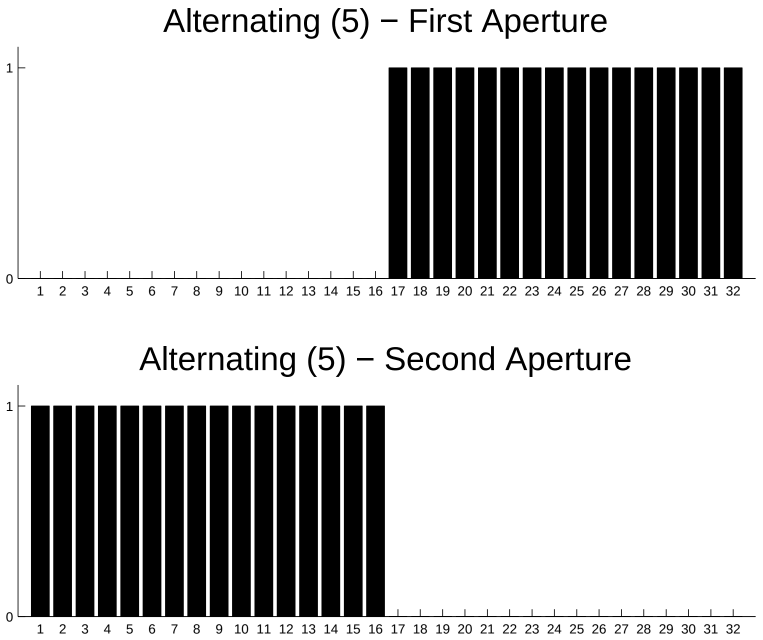


Figure 4.38: *Alternating Apodisation with 16x Aperture Width*

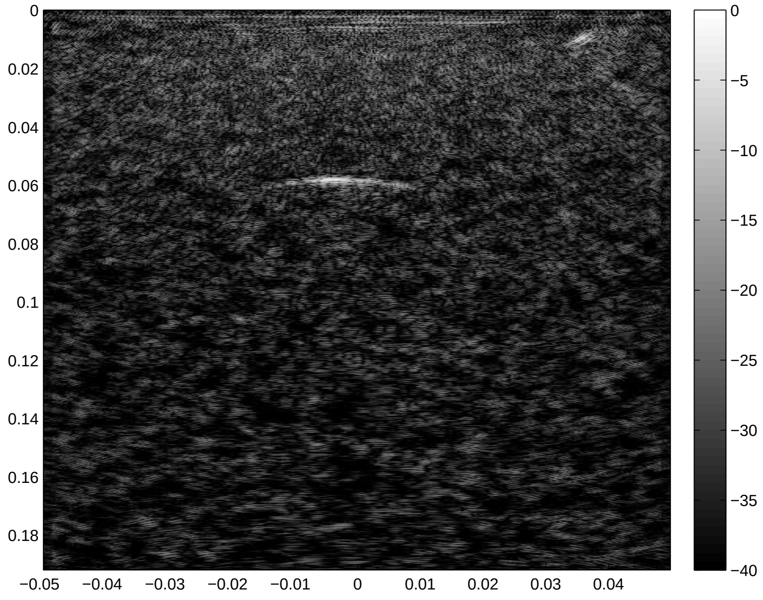


Figure 4.39: *Alternating Apodisation with 16x Aperture Width Result*

4.5 Discussion

A new signal processing technique for ultrasonic NDE has been proposed and experimental results demonstrated on a range of industrially relevant samples.

The SASACI approach involves creating two images of the same region. In this chapter, an investigation into improving imaging using sub-apertures from a single FMC dataset. It is anticipated that there are other ways to improve imaging using this methodology, using a set of arrays with differing centre frequencies, for example.

It was found the optimal way to weight the sub-TFM images is with an alternating scheme with a 4 element wide aperture. An maximum SNR improvement of 13dB has been achieved compared to a standard TFM image. Visually, defects appear clearer in SASACI images in a number of test cases, when compared to a bandpass-filtered TFM image. It can be concluded that this signal processing method is both novel and effective within the field of non-destructive evaluation of difficult materials.

Chapter 5

Correlation for Adaptively Focused Imaging

5.1 Introduction

Spatially Averaged Sub-Aperture Correlation Imaging (SASACI), described in Chapter 4, has shown potential in reducing noise caused by large grains[130] but this technique has a number of limitations. The first is due to the need for a range of input variables. Hence, the resulting output image is sensitive to the parameter settings and these parameters require a degree of trial and error in order to select the values that will give optimal results for a given inspection scenario. The second major limitation is resolution. While speckle noise is reduced, resolution is decreased, preventing characterization of any genuine flaws.

A new technique, Correlation for Adaptively Focused Imaging (CAFI), is now proposed which builds upon the SASACI technique[131]. This technique aims to solve the problem of generating images of materials with variable propagation velocities, while not being sensitive to parameter settings. CAFI is inspired by the Nearest Neighbour Cross-Correlation (NNCC) technique, introduced by Flax & O'Donnell in 1988[132].

5.2 Methodology

Correlation for Adaptively Focused Imaging (CAFI) processes FMC datasets to create images in the same manner as TFM and SASACI, but utilizes cross-correlation to improve focusing in materials where there are local velocity variations. It does this by amending the delays applied to each element so that the cross-correlation coefficient of neighbouring signals are maximised.

A flowchart outlining the cross-correlation methodology for CAFI is shown in Figure 5.1. The technique processes each transmitting element and pixel individually and each iteration of the loop is referred to as a TxPx. First, for each receive element, the estimated delay (for delay-and-sum imaging) is calculated. A range of samples is then extracted from the A-scan centred on the calculated delay. Once this has been completed for every receiving element for a given TxPx combination, the contributions from neighbouring receiving elements are cross-correlated using the equation shown in Equation 5.1, where C is the centre of the signal to be cross-correlated, S is the number of pixels on each side of C to be analysed and $RX1$ and $RX2$ are the signals to be cross-correlated[56]. The cross-correlation is performed at a range of delays, known as lags, for every pair of receiving elements. The aim of this stage is to find the number of lags, d, that results in the highest value of cross-correlation coefficient, p. This value is recorded for each pair of adjacent elements. The original delay adjusted by the calculated lags will result in the optimal focus for each TxPx. The corrected delays for a single TxPx are normalized to the centre element so that imaging can take place using the corrected delays.

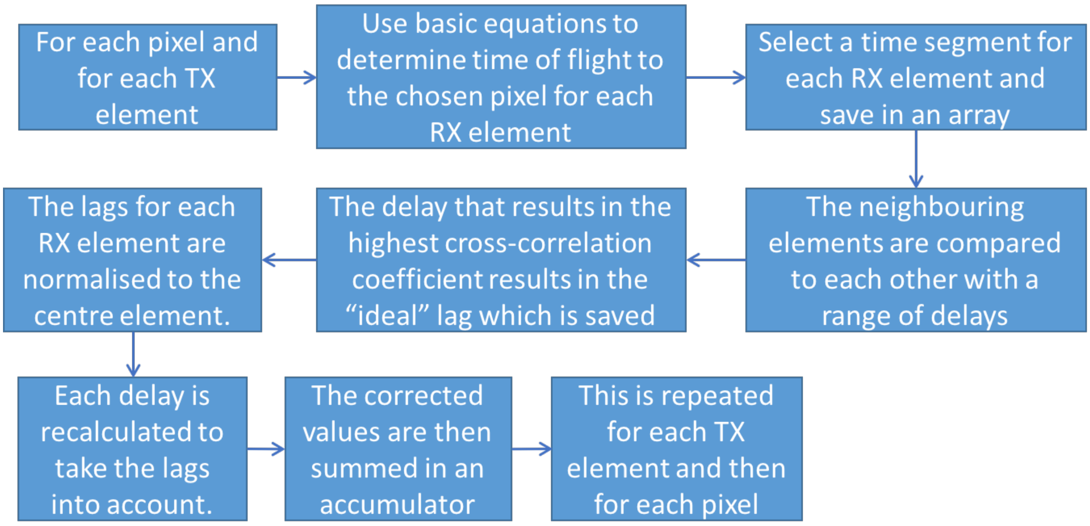


Figure 5.1: *The flowchart depicting the cross-correlation process of Correlation for Adaptively Focused Imaging*

$$p(x) = \frac{\sum_{i=x-S}^{x+S} RX1(i) \cdot RX2(i-d)}{\sqrt{\sum_{i=x-S}^{x+S} RX1(i)^2} \cdot \sqrt{\sum_{i=x-S}^{x+S} RX2(i-d)^2}} \tag{5.1}$$

Equation 5.1 is repeated for different values of d. d is typically set to a range of integers, centred around 0, that relate to the number of samples per wavelength in the signal. Once d has been calculated for each pair of elements within a TxPx combination, the lags are normalised to the centre element using Equation 5.2

where n_{el} is the number of array elements. This equation is valid only for the second half of the array, and must be repeated iterating from the centre element towards the first in order to calculate D for the first half of the array.

$$D_{rx} = \sum_{x=\frac{n_{el}}{2}}^{rx-1} d_x - d_{x+1} \tag{5.2}$$

$$TFM_{TxPx} = |\sum h_{tx,rx}(\frac{\sqrt{(y_{tx}-y)^2+z^2}+\sqrt{(y_{rx}-y)^2+z^2}}{v_L} - D_{rx})| \tag{5.3}$$

The modified TFM algorithm is shown in Equation 5.3 where $h_{tx,rx}$ is the FMC matrix, y_{tx} is the location of the transmitting element, y_{rx} is the location of the receiving element, y and z are the co-ordinates of the pixel of interest, v_L is the propagation velocity in the inspection medium, and D_{rx} is the delay adjustment calculated through cross-correlation. The delay adjustment is unique for each transmit-receive pair. tx is constant for Equation 5.3 and the process is repeated for every TxPx.

5.2.1 An Introduction to Nearest Neighbour Cross-Correlation

To examine how Nearest Neighbour Cross-Correlation (NNCC) works, a simple problem with a transmitter and two receivers is considered. A point reflector in a perfect medium is modelled and it is assumed that there is no energy loss due to absorption or dispersion. The diagram for this model is shown in Figure 5.2. A hypothetical scenario will be explored using this model and examine methodologies of combining the responses from the receivers.

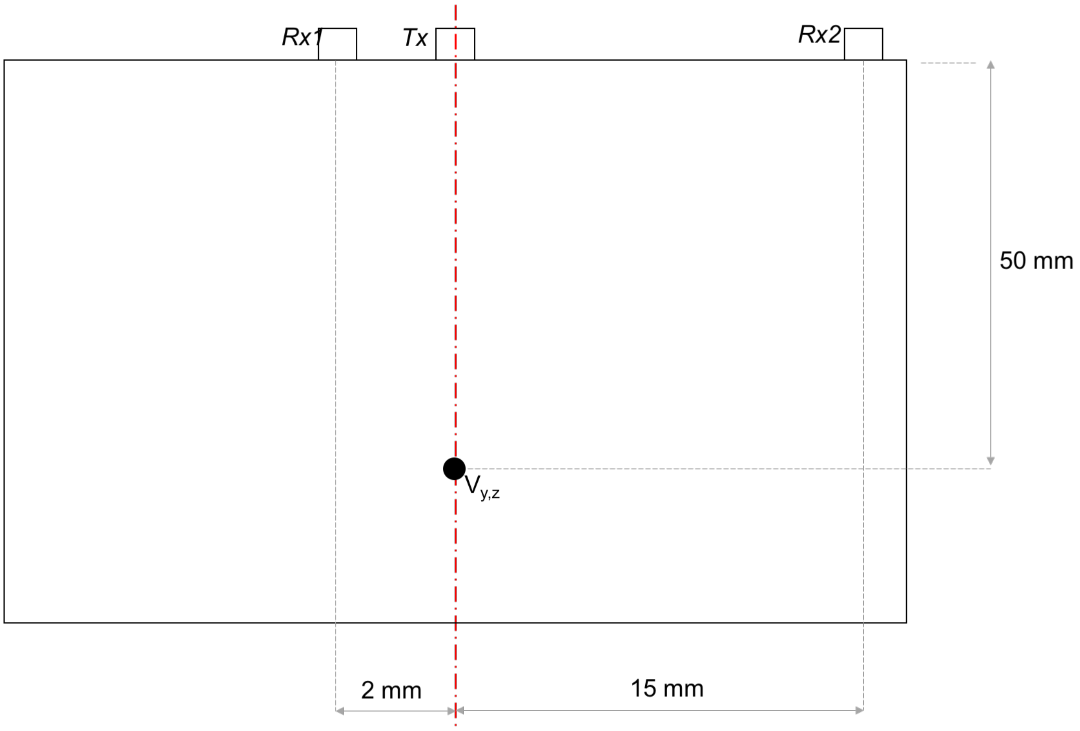


Figure 5.2: *An hypothetical setup with a transmitter, two receivers and a point reflector. Not to scale.*

For reference in this chapter, the leftmost receiving element will be referred to as $Rx1$ and the rightmost, $Rx2$. For equations 5.4 to 5.6 a general solution will be sought, referring to any receiving element only as $r_{x,y}$. When a signal is referred to as being combined, it is the summation of the responses from $Rx1$ and $Rx2$.

The input to the system is a Gaussian windowed tone-burst. There are 5 cycles of a 1MHz sine wave, sampled at a frequency of 20MHz and weighted using a standard Gaussian window generated via MATLAB. This is shown in Figure 5.3.

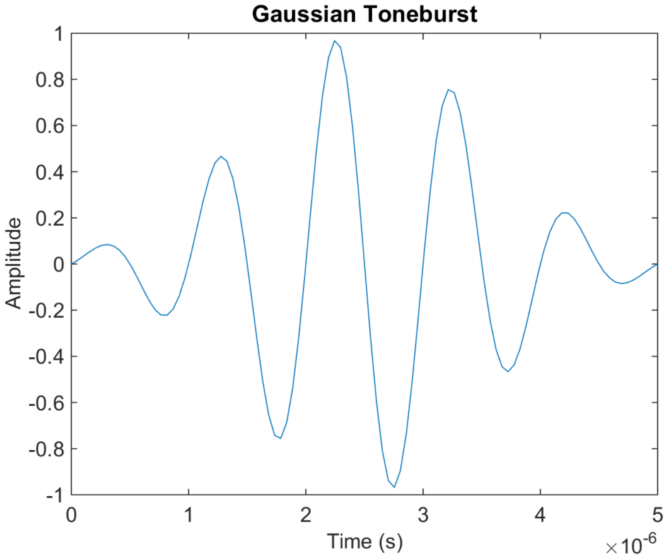


Figure 5.3: *A Gaussian windowed toneburst of frequency 1MHz*

The moment when the tone-burst is applied to the system is labelled as time $t = 0$. The combined A-scan trace without any correction for propagation delays is shown in Figure 5.4 below. The red trace represents $Rx1$ and the green trace represents $Rx2$.

The combined signal is shown in Figure 5.5. Instead of overlaying the signals on top of each other, as in Figure 5.4, they are now being summed together. It is apparent that the signals are interacting but constructive interference is not taking place due to the fact that the signals are not in phase. The maximum amplitude in the system is still less than 1 which is the maximum amplitude of our initial tone-burst. Using the knowledge of the location of the receiving sensors with respect to the transmitter, the round-trip delay can be calculated for each receiver. A lag (calculated by taking the difference between round-trip delays) can be applied to one of the signals, thus delaying it with respect to the other.

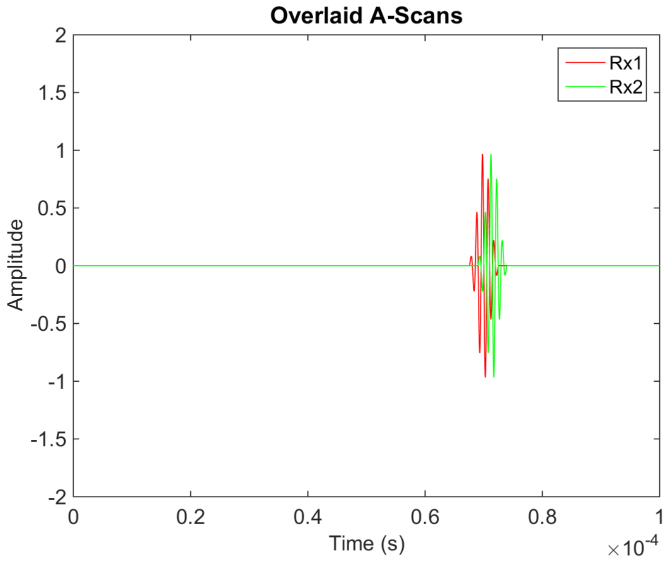


Figure 5.4: *Two A-Scans overlaid without delay correction*

Equations 5.4 and 5.5 are used to calculate each of the delays where $V_{y,z}$ is the point reflector defined by its Cartesian coordinates. $t_{y,z}$ is the location the point of transmission and $r_{y,z}$ is the location of the receiver. These points can be related to the hypothetical experimental setup shown in Figure 5.2. Equation set 5.4 calculates the distance between the point reflector and the other points of interest. t_{path} and r_{path} are the distances to and from the transmitting element to a pixel, p, and from the pixel to the receiving element, respectively.

$$t_{path} = \sqrt{(t_y - V_y)^2 + (t_z - V_z)^2} \tag{5.4a}$$

$$r_{path} = \sqrt{(r_y - V_y)^2 + (r_z - V_z)^2} \tag{5.4b}$$

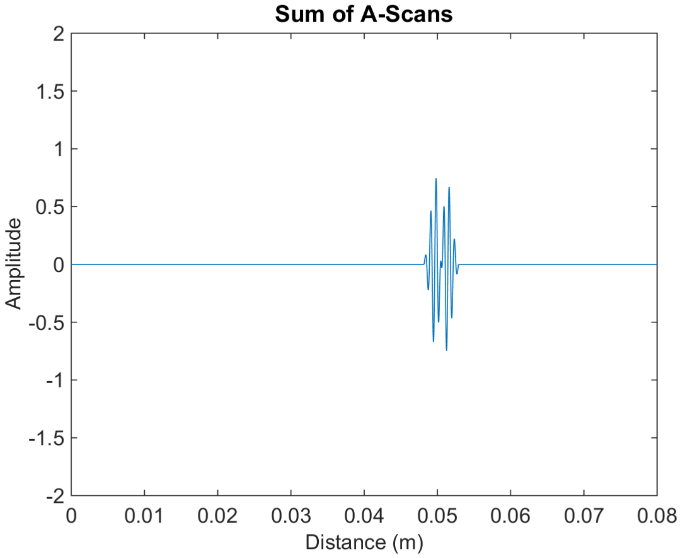


Figure 5.5: *Two A-Scans combined without delay correction*

Equation set 5.5 uses the times calculated in 5.4 to calculate the delays. Note that the delay is equal to the difference in time between the receive path and the transmit path, as everything is calculated relative to the point of transmission. This becomes essential when a number of receiving elements are employed as a static point of reference is required.

$$delay = \frac{(r_{path} + t_{path}) - (t_{path} + t_{path})}{v_L} \tag{5.5a}$$

$$delay = \frac{r_{path} - t_{path}}{v_L} \tag{5.5b}$$

$$delay = \frac{\sqrt{(r_y - V_y)^2 + (r_z - V_z)^2} - \sqrt{(t_y - V_y)^2 + (t_z - V_z)^2}}{v_L} \tag{5.5c}$$

For the example in Figure 5.2, Equation 5.5c can be simplified further. It is known that the transmitter and receiver have the same y co-ordinate and that the point reflector and the transmitter have the same x co-ordinate. Taking this into account, the simplification shown in Equation 5.6 can be derived. Note that this equation only holds true when the points of interest are orthogonal.

$$delay = \frac{\sqrt{V_z^2 + r_y^2} - V_z}{v_L} \tag{5.6}$$

The equation is essentially the propagation time from the transmitter to the point of interest minus the propagation time from the receiver to the point of interest. When applied to both receiving elements, the delayed signal can be seen as depicted in Figure 5.6. Again, the red line is $Rx1$ and the green line is $Rx2$. It should be noted that due to the corrected delays, the lines are on top of each other, rendering the red line completely invisible. This is the expected and ideal case.

The next image, in Figure 5.7, shows the sum of both the individual A-scans. It can be observed the maximum amplitude is now 2, indicating that the signals are now summing constructively. The distance is also at the correct location, at 0.05 meters which is the location of the simulated point reflector.

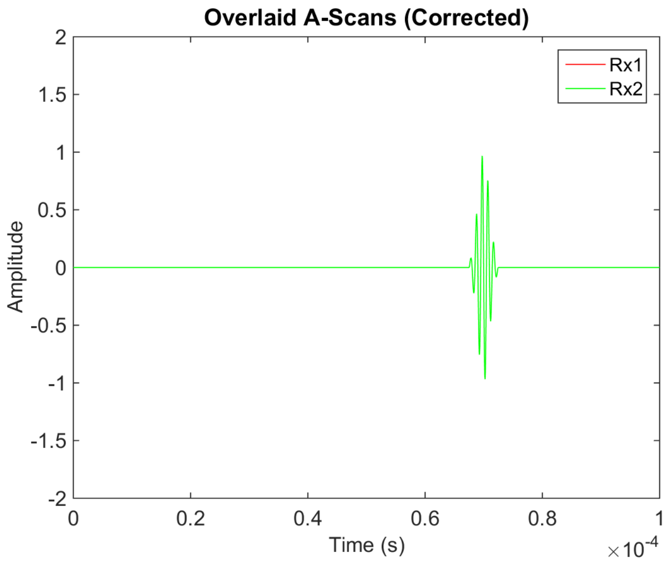


Figure 5.6: *Two A-Scans overlaid with delay correction*

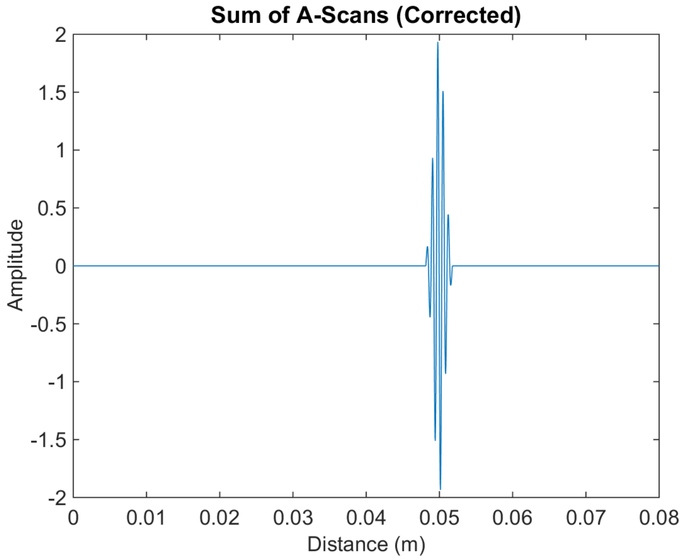


Figure 5.7: *Two A-Scans combined delay correction*

5.2.1.1 Nearest Neighbour Cross-Correlation in Anisotropic Materials

Delay-and-sum processing has been considered for a simple system. Let us now add a layer of complexity to our system. Difficult materials, besides being highly scattering and attenuating, are generally anisotropic. That is, having properties that are directionally dependent. Often, a wave will propagate in one direction at a differing speed from other directions. This hinders the ability to effectively delay and sum contributions from different receivers. An example of this is now presented.

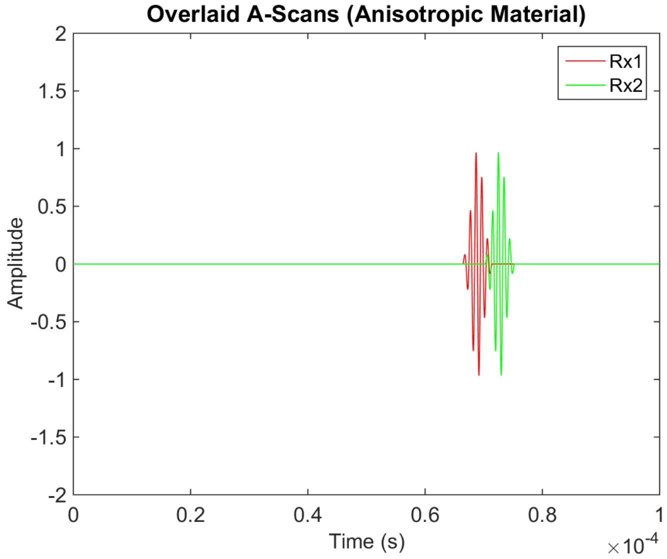


Figure 5.8: *Two A-Scans overlaid without delay correction (anisotropic model)*

The same hypothetical arrangement is used as shown in Figure 5.2, only now the wave velocity from p to $Rx1$ is 1530m/s and the wave velocity from p to $Rx2$ is 1430m/s. Figures 5.8 and 5.9 show that without accounting for propagation time everything looks much the same as in the more simple scenario. Figure

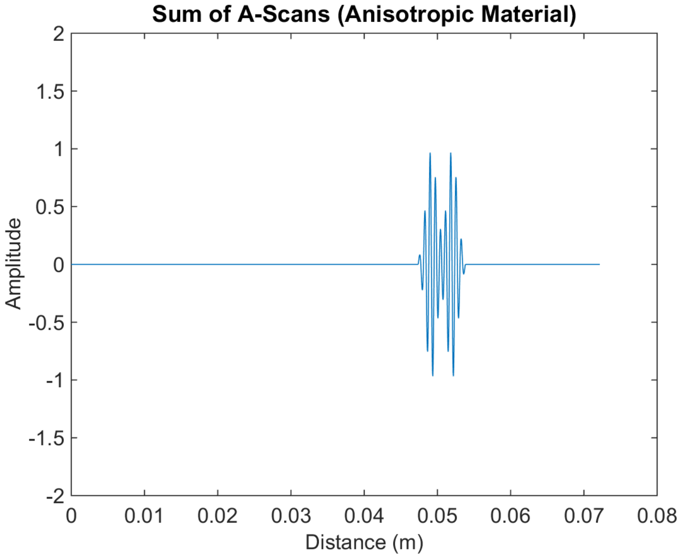


Figure 5.9: *Two A-Scans combined without delay correction (anisotropic model)*

5.10 is different compared to Figure 5.6, as both A-scans are visible now. The waves are no longer co-incident with each other and when Figure 5.11 is observed, the waves are no longer summing constructively and the maximum amplitude is reduced to 1. It can be observed that a deviation of just 3% from the assumed propagation velocity (1480m/s) is enough to stop delay and sum imaging from working correctly.

Figure 5.12 shows a snapshot of the two signals. The estimated delays of the signal are known and with this knowledge, the area where the reflections occur can be extracted. Within this plot there is no knowledge of where the reflections are, but the anisotropy is assumed to be mild enough that the desired reflection will be in the area selected for cross-correlation.

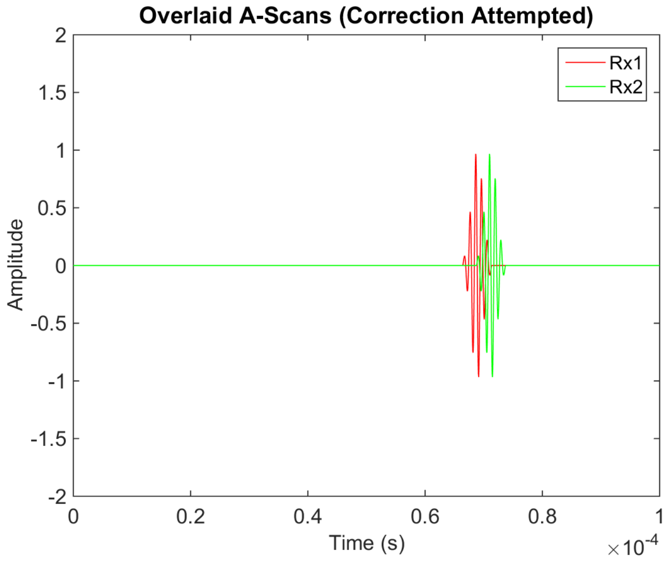


Figure 5.10: *Two A-Scans overlaid with attempted delay correction (anisotropic model)*

The signals can be cross-correlated to find out the delay correction that would result in the highest cross-correlation coefficient between the two signals. This is found using Equation 5.1 over a range of delays. Figure 5.13 shows a plot of lags against cross-correlation coefficient. The maximum value of this plot is taken to be the optimum refined delay for the pair of A-Scans, which for this case was found to be 18.

Figure 5.14 shows the overlaid A-scans when the delay of 18 samples was applied to the second A-scan. The signals recombine and overlap as expected. Applied to the full A-scan, the final expected result can be seen in Figure 5.15 and the combined result in Figure 5.16. Note that the location of the reflector has appeared to have moved slightly, from 0.5m to 0.49m. This is due to the fact that

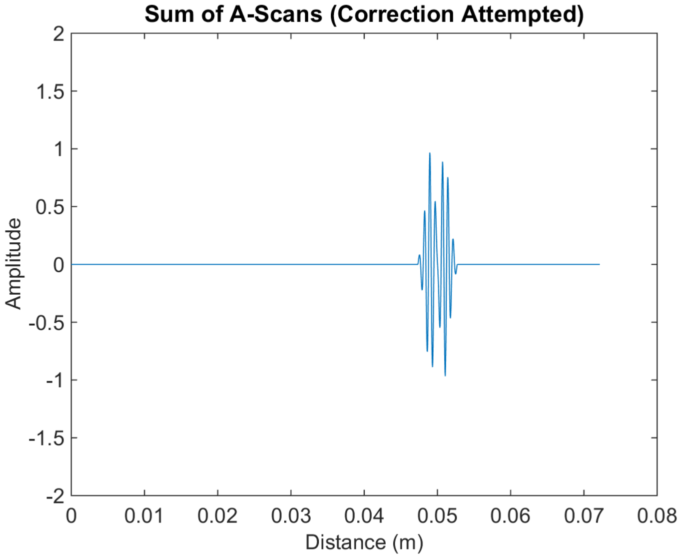


Figure 5.11: *Two A-Scans combined with attempted delay correction (anisotropic model)*

perfect delays cannot be applied due to the unknown precise velocities in the model. This error is low (2%) and is a consequence of using NNCC to inform the CAFI algorithm.

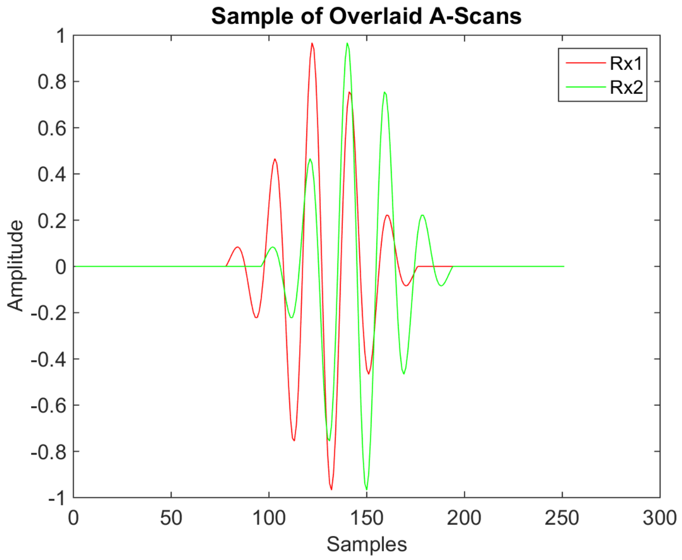


Figure 5.12: *A section of the overlaid A-Scans before cross-correlation*

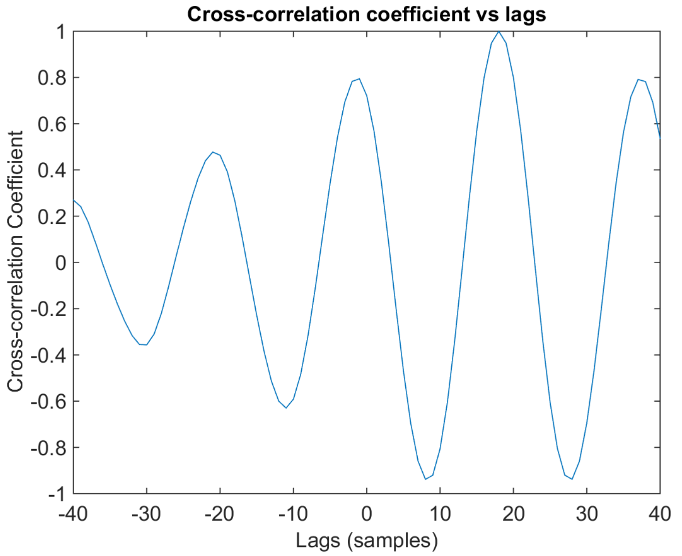


Figure 5.13: *A plot of tested delays vs cross-correlation coefficient for the A-scans in Figure 5.12*

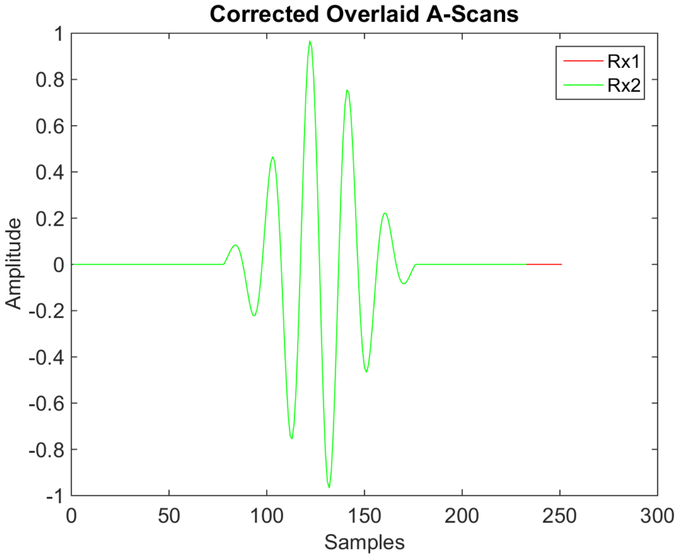


Figure 5.14: *A section of the overlaid A-Scans after cross-correlation and delay correction*

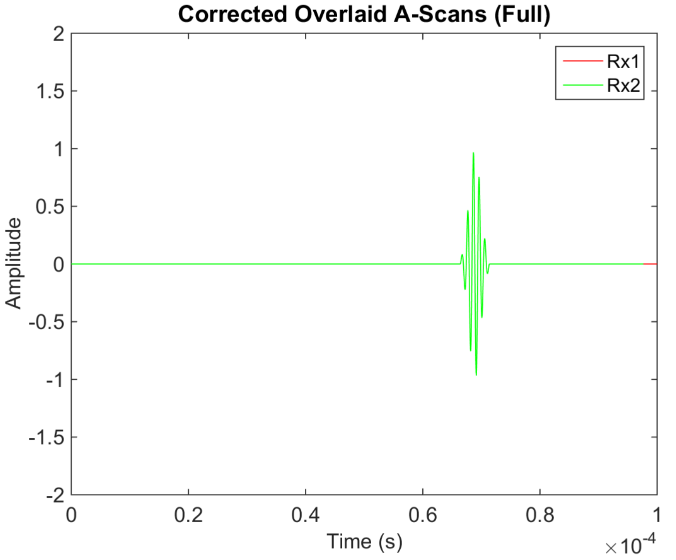


Figure 5.15: *Two A-Scans overlaid with CAFI correction*

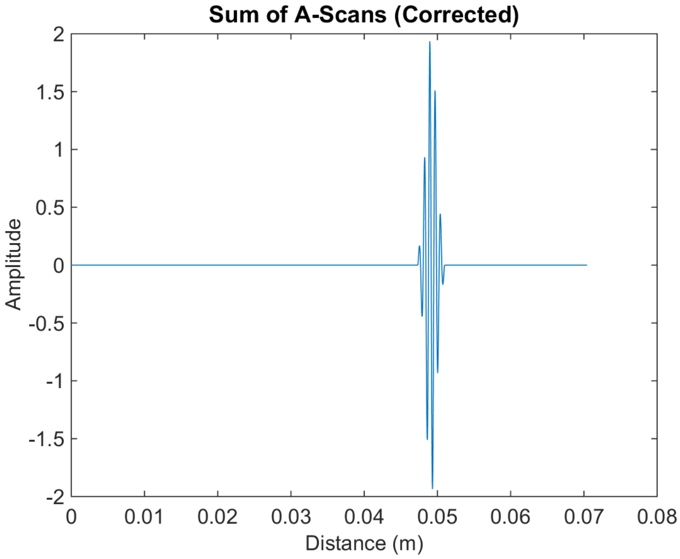


Figure 5.16: *Two A-Scans combined with CAFI correction*

5.2.2 Nearest Neighbour Cross-Correlation in the Presence of Noise

For NNCC-based imaging algorithms to be applicable in this work, they need to be robust in the presence of noise. In the next simulation, random noise is added to each A-scan. The peak amplitude of the toneburst is 1 and the peak amplitude of the noise is 1.5. This yields a signal to noise ratio of -3dB.

The combined delay-and-sum signal before NNCC correction is attempted is shown in Figure 5.17.

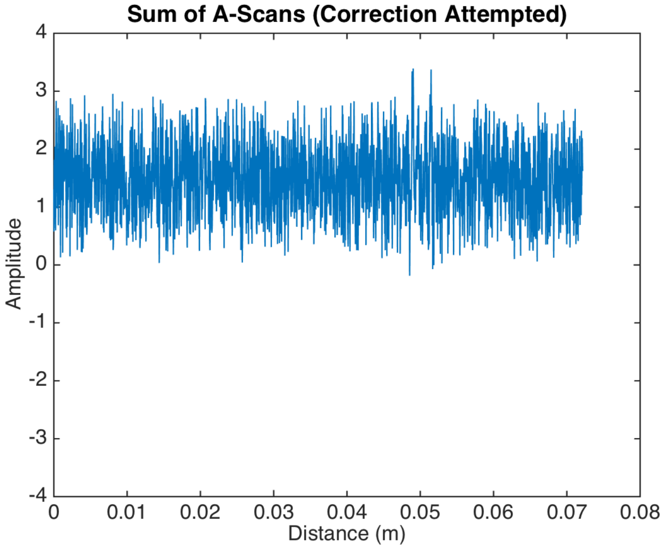


Figure 5.17: *Two A-Scans combined with added noise*

In Figure 5.17 there is a peak at 0.05m, but it is not clear if that is the result of noise or an actual feature. In this case, it is possible for the noise in each signal to undergo constructive interference and appear as a legitimate reflector in the combined A-scan. In the next set of images, Figures 5.18 to 5.22, the CAFI technique is applied to the noisy signal to attempt to refocus the delay-and-sum A-Scan on the reflector.

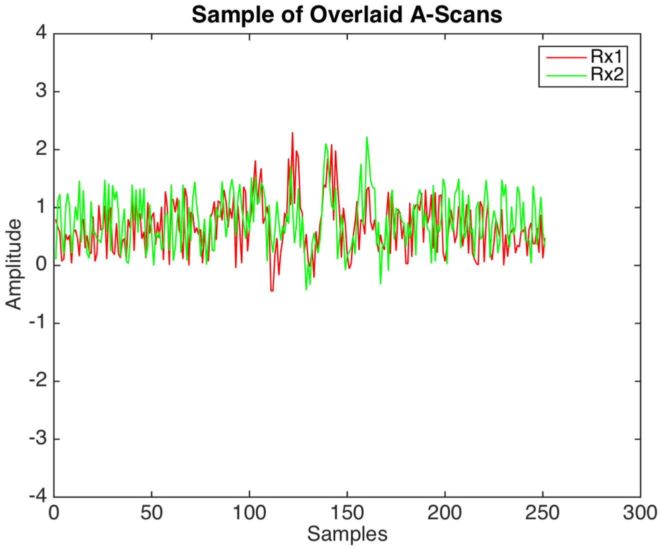


Figure 5.18: *A subset of two A-Scans overlaid with added noise*

Figure 5.19 shows the cross-correlation coefficient against lags for the noisy signal. When compared with a clean signal such as the one in the previous example, shown in Figure 5.13, the peak indicating the correct number of lags can be seen to be less prominent. The ideal delay can be observed from the graph to be 19 samples. This differs from the previous example by only one sample, or 50ns.

Comparing the before and after, Figures 5.17 and 5.22 respectively, there is a marked improvement in effective SNR. The SNR in the after image (peak signal to peak noise) is 3.0, while in the before image the SNR is 1.4. This may seem

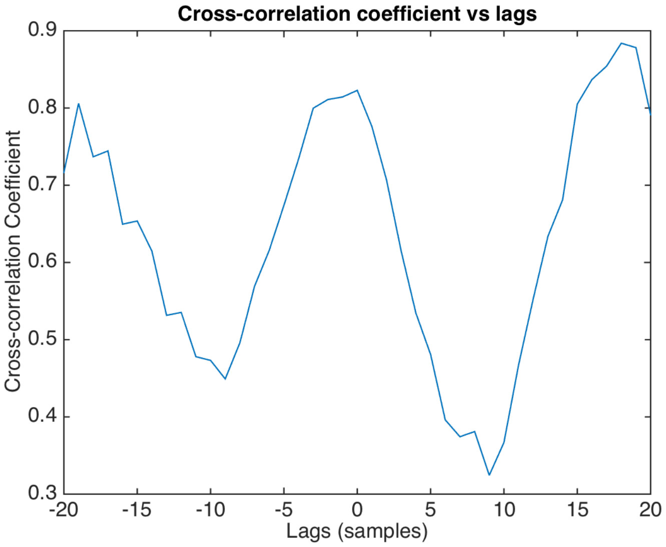


Figure 5.19: *A plot of tested delays vs cross-correlation coefficient for the A-scans in Figure 5.18*

like a small difference but it shows an increase of SNR of 110%, over double the original SNR. Qualitatively, there is clearly a reflector shown at 0.05m. This is an improvement over the 'before' image.

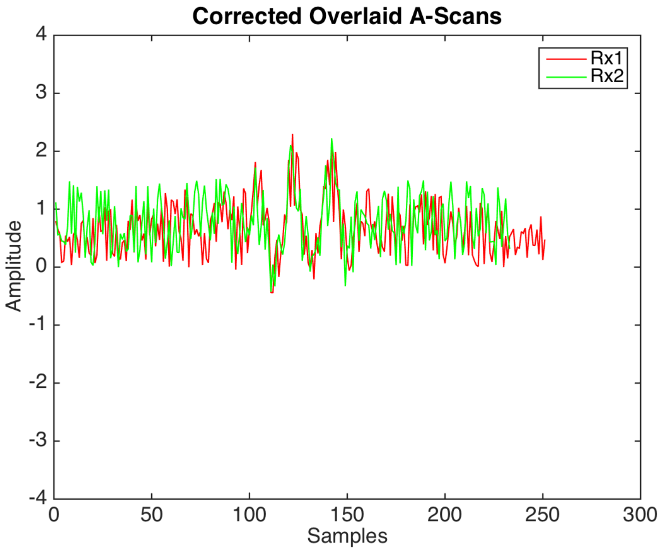


Figure 5.20: *A subset of two A-Scans overlaid with added noise and corrected using CAFI*

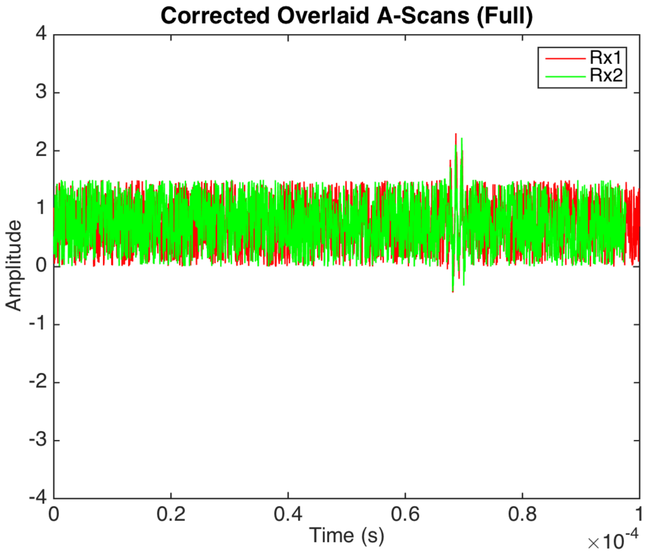


Figure 5.21: *Two A-Scans overlaid with added noise and corrected using CAFI*

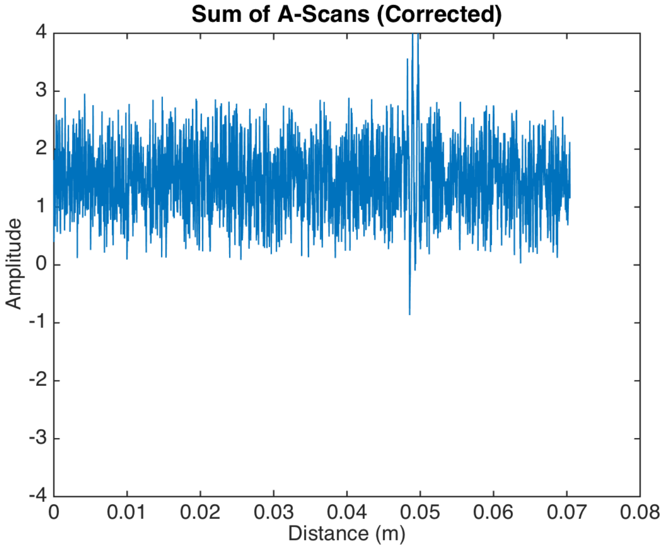


Figure 5.22: *Two A-Scans combined with added noise and corrected using CAFI*

5.2.3 Phase Enhanced Correlation for Adaptively Focused Imaging

The focus correction can be altered by using a different way to measure the signal in the cross-correlation. Using the instantaneous phase of a signal for comparison means that the cross-correlation algorithm no longer depends on signal amplitude. This makes CAFI more resilient to high amplitude noise that is typically found when inspecting coarse grain materials[67]. Applying the Hilbert transform to an RF signal yields the analytical signal with both real and imaginary components, as seen in 5.23. The instantaneous phase is calculated for each sample by calculating the phase angle,ϕ_0, between the real and imaginary components of said sample.

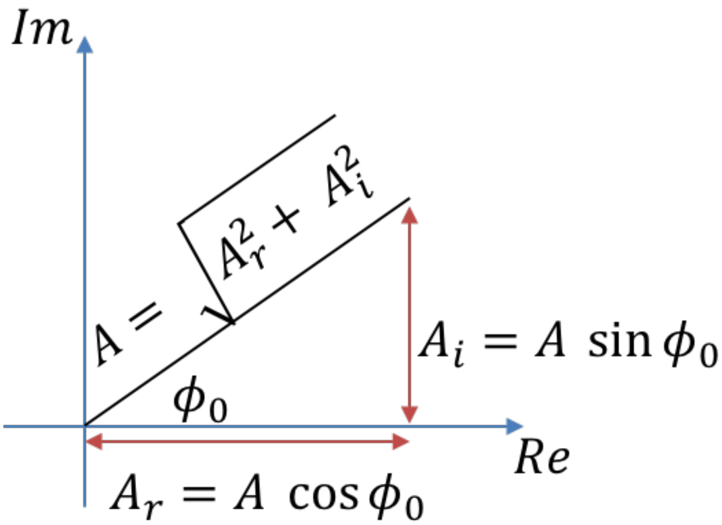


Figure 5.23: *Finding the phase of an analytical signal using the real and imaginary components*

The CAFI process can then be repeated using the instantaneous phases as an input to the cross-correlation algorithm, in place of a signal's amplitude. Once the ideal delays have been calculated, the delay and sum imaging can take place as before using the original A-Scans. Again, this concept will be demonstrated on the noisy anisotropic model.

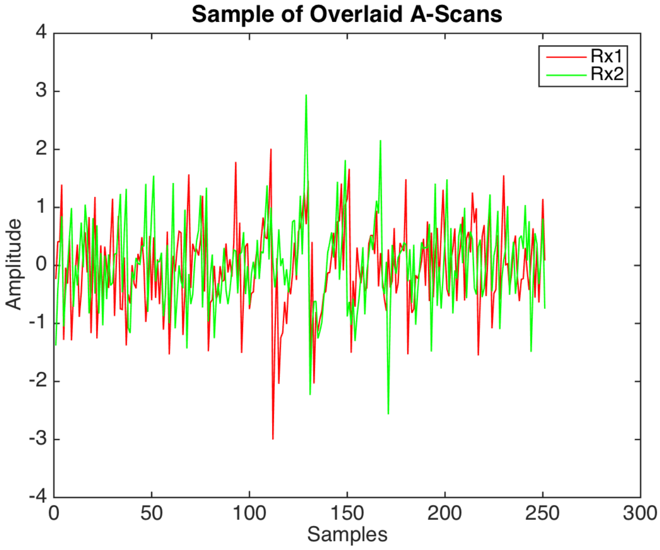


Figure 5.24: *Overlaid subset of instantaneous phases*

Figure 5.24 shows a sample of the A-scans, overlaid on each other, after the instantaneous phase angle has been calculated. These signals are then input to the cross-correlation function, the results of which are shown in Figure 5.25. The graph shows that the ideal delay is 18 samples, which is the expected value, given knowledge of the parameters of the hypothetical imaging scenario. This is an improvement over the standard 'amplitude-only' CAFI using result which yielded an 'ideal' delay of 19 samples. Furthermore, the graph is much steeper than the one generated from the amplitude-only CAFI. This is a property which can help to reduce false positives. Delays which are not optimal now have a much lower cross-correlation coefficient in the phase-based CAFI implementation.

This correct delay is then applied and can be seen in Figures 5.26 to 5.28.

The maximum signal value in Figure 5.28 is 3.8 while the peak noise value is 2.8, meaning that the signal to peak noise ratio is 2.7dB. This is less than the SNR of the amplitude-based CAFI even though the anisotropy has been properly corrected for. The reason for the lower SNR is due to the noise destructively interfering with the reflected signal.

If more A-Scans were modelled and used to build a new delay-and-sum signal, the phase-based CAFI would eventually outperform the amplitude-based due to the fact that the noise would start to be averaged out and the signals would combine more optimally.

In some cases both the phase-based and amplitude-based versions of CAFI fail to calculate the correct delays. In these cases, the calculated ideal delays are generally different and will produce different images when each of the delays are used. This can be used to our advantage as the two different generated images can be used as an input to SASACI's cross-correlation algorithm[130] to attempt to pick out the coherent component of the images.

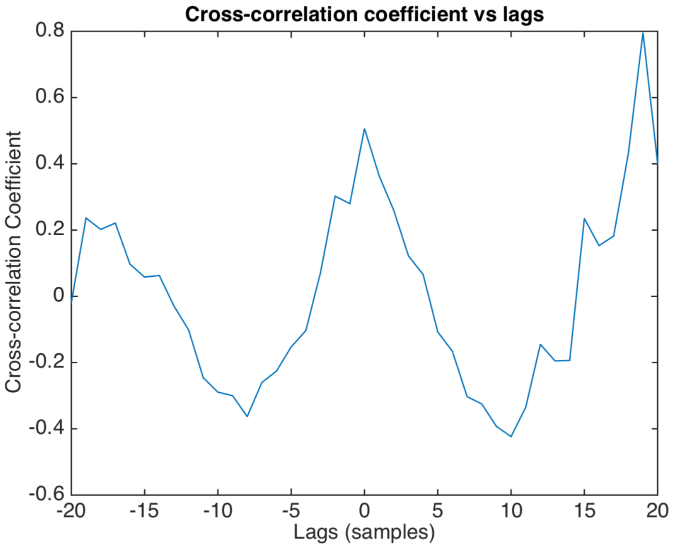


Figure 5.25: *A plot of tested delays vs cross-correlation coefficient for the A-scans in Figure 5.24*

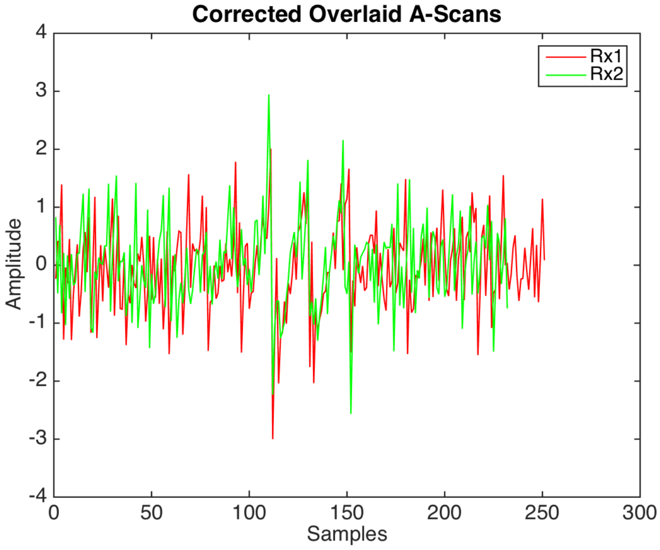


Figure 5.26: *A subset of two A-scans' instantaneous phases overlaid with added noise and corrected using CAFI*

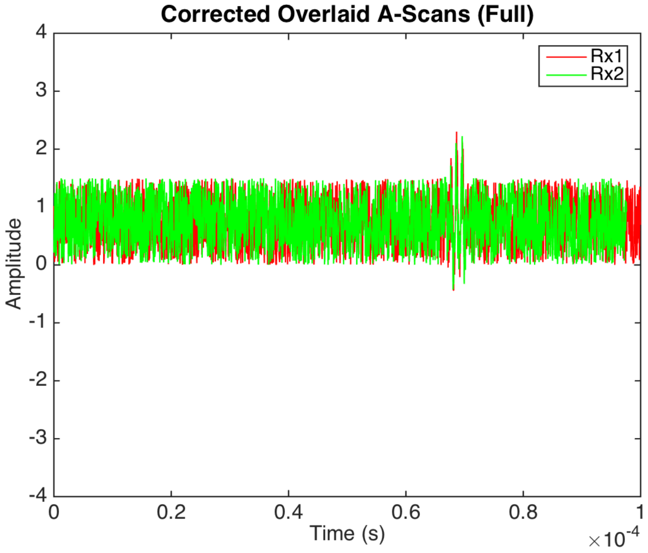


Figure 5.27: *Two A-Scans overlaid with added noise and corrected using phase-based CAFI*

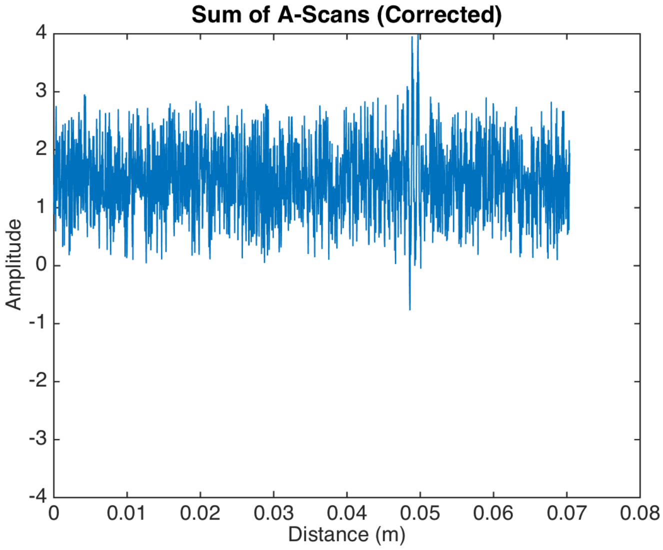


Figure 5.28: *Two A-Scans combined with added noise and corrected using phase-based CAFI*

5.2.4 CAFI Imaging

In previous examples, a model was used with one transmitting and two receiving elements. A Full Matrix Capture dataset can have 64 transmitting elements and 64 receiving elements, though both greater and fewer elements are also commonly used. A slightly different approach is required in order to refine the delays for each A-scan.

For imaging with CAFI, each transmitter and each pixel are treated independently. CAFI will attempt to focus on one pixel for one transmitting element at a time. For an FMC dataset collected from a 64 element array, there will be 64 A-scans to apply the technique to. In the previous examples, there were only 2 A-scans used. It does not matter in these examples if $Rx1$ is correlated with respect to $Rx2$, or the other way around. The ideal delay, if reversed, will simply be the inverse. For multiple A-scans, it does matter which A-scan is correlated with respect to another.

First, the expected delays are calculated for each receiving element using the speed of sound (generally measured with a pulse-echo test on a block of known thickness) within the material. A subset (normally around 300 samples) of each A-scan, centred around where the reflection would be expected (calculated using the estimated velocity), is copied into an array for each receiving element. Each of the A-scans is then correlated with its neighbouring A-scan, always with respect to the element closest to the centre of the array. Once each of the A-scans has been correlated to each other $n - 1$ delays will have been calculated, where n is the number of elements in the array.

The delays are then normalised to the centre element (i.e. the centre element has a delay of 0) and the delays are then calculated relative to the centre element. This is done using Equation 5.7 where *delay* is the delay to apply to each A-scan, x is the A-scan of interest, n_{el} is the number of array elements and $d_{ci} = d_i - d_c$ is the normalised delay for A-scan, i, with respect to its neighbouring element closest to the centre element. If two centre elements exist, d_c is calculated by taking the average of the two.

$$delay = \begin{cases} \sum\limits_{i=x}^{\frac{n_{el}}{2}} d_{ci}, & \text{if } x < \frac{n}{2}; \\ \\ \sum\limits_{i=\frac{n_{el}}{2}}^{x} d_{ci}, & \text{otherwise.} \end{cases} \tag{5.7}$$

After these delays have been calculated, standard TFM imaging takes place and a unique delay is calculated for each receive element while looping through each transmitting element and each pixel.

This process can then be repeated with phase-based CAFI in order to generate two different sets of delays and therefore two different images. If this is the case, then two images have been created, each using TFM with a form of focus correction to account for phase aberration. Any legitimate reflectors should be subject to enhanced focusing in the corrected images. Any noise may differ between the two images since most of the noise in a TFM image is from grain reflections that will suffer more from multipath propagation. SASACI cross-correlates images created from the two array sub-apertures to remove grain noise. If two images have been generated then the final stage of the CAFI process is to cross-correlate the focus-corrected images using the two-dimensional cross-correlation algorithm in Equation 4.2. Once p is found for each pixel, using this equation, any pixel where $p < 0$ is set to 0. p is then cross-multiplied with the original TFM image to give the final CAFI result.

5.3 Results

CAFI was validated experimentally using an ultrasonically noisy material. Two metrics were used to measure the performance of the technique. The number of pixels with amplitude >-6 dB are counted. This is an approximate measure of point spread when imaging a small defect. The area within these pixels is determined to be signal and anything outside of this is noise. The mean values of these areas are taken to give the Signal to Noise Ratio (SNR) in dB.

5.3.1 Experimental Set-up

Figure 5.29 shows a block of Inconel 625 which was inspected from the top face with a Vermon 128 element 5MHz linear array driven by a Zetec Dynaray Phased Array Controller. An FMC dataset was captured using a 32 element sub-aperture of the array. The target for inspection is the side drilled hole indicated in Figure 5.29, which is 60mm from the top surface.

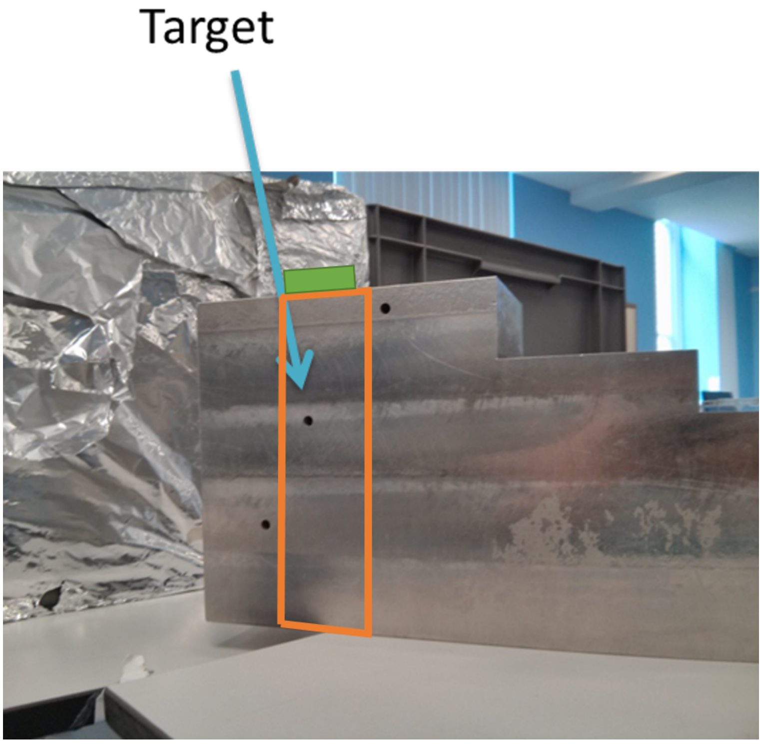


Figure 5.29: *An Inconel 625 step wedge of height 180mm. The area to be imaged is highlighted by an orange box, and the green rectangle represents the array position.*

The Inconel block is the same specimen used in Section 4.4.1, however array used to record the data used in this section was positioned at a different point. Due to the internal structure of the material, the velocities and scatterers within the sample change through its cross-section. This gives rise to significantly different results depending on where the array is placed on the sample.

The TFM image of the full depth of the block is shown in Figure 5.30. The image is shown at a dynamic range of 20dB and has multiple visible defects, one at a depth of 60mm and another at 105mm. Due to the computational complexity of the algorithm (in the order of minutes per pixel), only a small subset of this area will have CAFI applied to it, as depicted by the green rectangle in the image.

5.3.2 CAFI Images and Quantitative Results

Figure 5.31 shows the TFM image at a dynamic range of 50dB. It has significant contributions of high amplitude noise and the resolution is considered poor. The SNR of this image is 8.16dB and the number of pixels >-6dB are 179. The main signal is at the correct depth, which is 60mm although it is not clear that there is a side drilled hole in the image. The side drilled hole is clearer in Figure 5.30 due to the 20dB dynamic range in the image. The apparent SNR would also be different due to the size of the area sampled when the noise is measured. This is due to the maximum amplitude within the area shown being normalised to 0dB. Sampling a larger area, where noise would be low due to the lack of reflectors, would result in a lower average noise being measured.

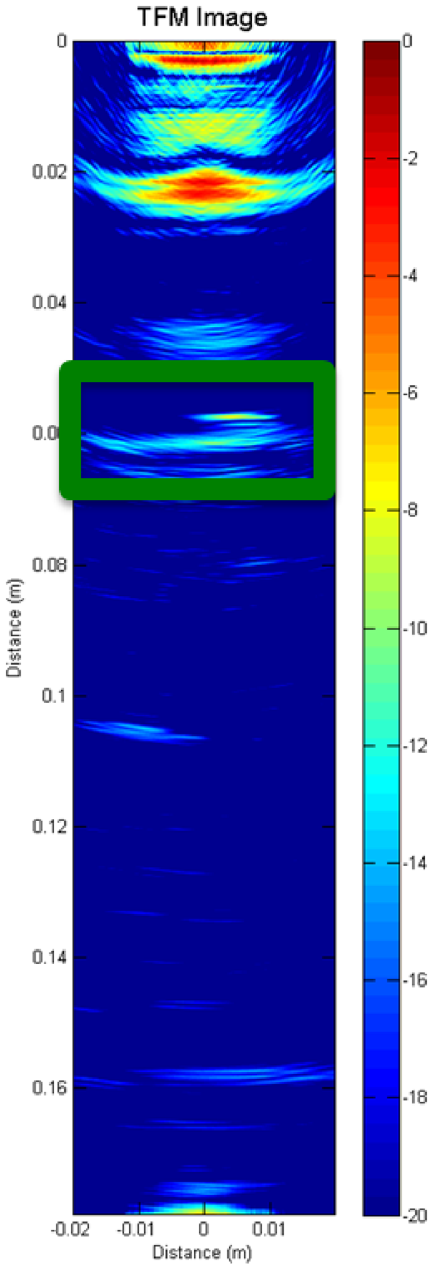


Figure 5.30: *A TFM image of the step wedge sample. Reflections from side drilled holes are visible at 60mm and 110mm. The back wall is at 0.16m.*

As long as the area where the noise is sampled is the same in each image, a fair comparison can still be drawn between the imaging techniques used. For the following results, all areas in the image >-6dB are considered signal while the rest are considered noise.

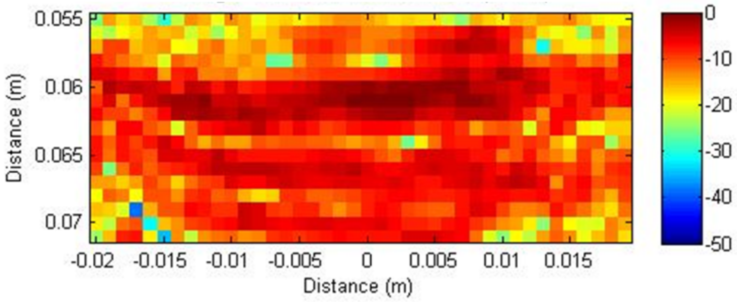


Figure 5.31: *TFM of the side drilled hole*

Figure 5.32 shows the CAFI image with amplitude-based focus correction. The SNR of this image is 8.59dB which is slightly better than the TFM image. The pixels >-6dB count is 226 which is poorer than the original TFM image.

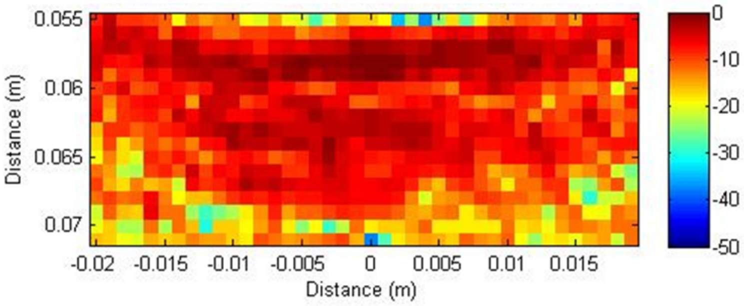


Figure 5.32: *The side drilled hole with amplitude-based focus correction*

Figure 5.33 shows CAFI with phase-based focus correction. The SNR of this image is 7.93dB which is again poorer than the original TFM image but the number of pixels >-6dB is 164 which represents a slight improvement over TFM. The results from the previous two images show that cross-correlation can be potentially used to improve images due to the differences in both SNR and resolution that can be achieved via standard focus correction algorithms.

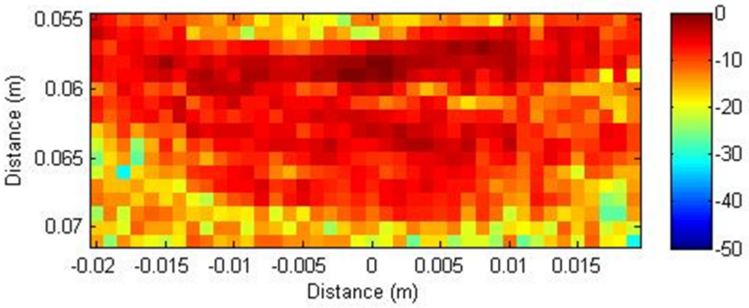


Figure 5.33: *The side drilled hole with phase-based focus correction*

Figure 5.34 shows the final CAFI images after the cross-correlation of both the phase-based and amplitude-based corrections. Visually, the hole is present in the image at the expected depth of 60mm, although there are some other artefacts. Quantitatively, the SNR of the image is 36.3dB which is significantly higher than the original TFM image, as well as the intermediary CAFI images. The number of pixels >-6dB is 6 which shows a large improvement in resolution over TFM.

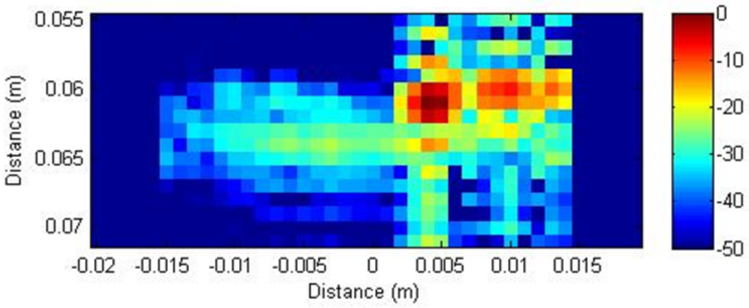


Figure 5.34: *The side drilled hole with the full CAFI method applied*

5.3.3 Discussion

The Correlation for Adaptively Focused Imaging algorithm was inspired by both Spatially Averaged Sub-Aperture Correlation Imaging and Nearest Neighbour Cross-Correlation. CAFI has a number of advantages over SASACI, including a potential for higher resolution images and robustness to parameter setting which was one of the most significant issues with the practical implementation of SASACI.

The results shown in this chapter show a clear improvement over TFM, although the visible artefacts in the final image demonstrate that the technique still requires improvement.

Method	**SNR**	**Pixels >6dB**
TFM	8.16dB	179
Amplitude-based CAFI	8.59dB	226
Phase-based CAFI	7.93dB	164
CAFI with SASACI	36.3dB	6

Table 5.1: *Comparison of TFM and different CAFI methodologies*

Looking at the quantitative results in Table 5.1, it can be seen that the phase-based CAFI and amplitude-based CAFI have different values of SNR and point spread. This is expected due to the fact that the material has large grains which contribute to coherent noise. This coherent noise will affect the outcome of the cross-correlation function and hinder the identification of a defect. As discussed in the methodology section, combining SASACI and CAFI has the potential to create images superior to standard CAFI and this is demonstrated in both the table and in Figure 5.34.

A limitation of CAFI is that when there is no defect to focus upon, the cross-correlation function can end up choosing delays where reflections from the grain will sum constructively and appear similar to a genuine defect. For this reason, when CAFI is applied to a large area, the SNR can be comparatively low compared with TFM. It is therefore recommended that CAFI is used to characterise, instead of to determine the location of defects. It is here that CAFI excels and this can be seen when comparing TFM in Figure 5.31 with the full CAFI method in Figure 5.34.

CAFI also has a technical drawback, which is its computational complexity, resulting in significant time required to process images. This can be alleviated by utilizing parallel computing. CUDA based imaging has been discussed in Chapter 3 and this technique of accelerating the imaging process can also be applied to CAFI as each ‘loop’ of the algorithm is independent from the last.

Chapter 6

Outlook and Conclusion

6.1 Analysis of Results

A number of novel computational methodologies have been introduced and evaluated in this thesis. Each methodology has its advantages and disadvantages as well as potential for future research and improvements on the technique. These methodologies will be discussed in the context of the thesis as a whole and related to the industrial problem described in Section 1.1.

The rapid TFM implementation introduced in Chapter 3 is a solution to two problems encountered in NDE within industry. Firstly, it allows for advanced signal processing methods to run in real-time for directly coupled inspections. This means that operators using these advanced imaging techniques receive immediate feedback. This is a feature currently offered by commercial equipment manufacturers, but only for traditional imaging approaches. The second problem it solves is the inspection of curved components using either immersion or a conformable wedge. In these scenarios, the array is not in direct contact with the surface and refraction will occur. Unlike traditional delay-and-sum imaging, pre-calculating focal laws for TFM-based imaging approaches was not previously practical due to the time required for computation.

A major drawback of this method is the specific hardware required. The presented technique is written in the NVidia's proprietary CUDA programming language and is therefore confined to running on NVidia graphics cards. Algorithms that make use of this approach to imaging are therefore limited to hardware supporting NVidia GPUs, i.e. desktop PCs or high-end laptops. Section 6.3.1 discusses ways that this technique can be further developed and improved, and will address this limitation. Another drawback of using the CUDA programming language is the expertise required to modify and debug the code. The work presented in Chapter 3 demonstrates the TFM algorithm running at a high speed and with refraction, but SASACI uses a modified version of this in order to exploit the acceleration offered by using commercial graphics cards. A PCI[66] implementation for CUDA also exists using this approach. Many more existing imaging approaches can benefit from this, but modification of the CUDA kernel is not trivial.

SASACI, introduced in Chapter 4, is shown to perform better than the current ‘gold standard’ of ultrasonic imaging on an industrially relevant sample. It has been shown to improve on SNR using a range of different parameters for SASACI. The presented imaging algorithm reduces the difficulties associated with inspecting materials with coarse grains. It has reduced background noise in a number of different imaging scenarios leading to increased clarity of defects. SASACI requires parameter settings which need to be fine-tuned in order to achieve an optimal result. At the moment, a trial-and-error approach is used to optimise these parameters.

The CAFI algorithm, introduced in Chapter 5, has advantages and disadvantages compared to SASACI and other imaging algorithms. CAFI uses no thresholding and therefore is not sensitive to the setting of this parameter. It also offers the potential of a sharper image resulting from an increased quality of focusing. CAFI is, however, sensitive to the speckle noise exhibited by grains in difficult materials. This can cause the algorithm to attempt to adjust focus onto the speckle. The best results have been observed when the CAFI algorithm has been used for image enchancement of an area already known to contain a defect.

6.2 Acoustic Research Toolbox

A software package has been created as both an inspection tool and a platform to demonstrate the imaging algorithms discussed in this thesis. This package has been named cueART (Centre for Ultrasonic Engineering Acoustic Research Toolbox). cueART is a LabVIEW-based program designed to enable non-expert users to use advanced imaging algorithms and view results in real-time. It supports a multitude of phased array controllers and is designed to be easily modified. Multiple advanced post-processing algorithms have already been implemented, including SASACI and PCF, and this software is a work-in-progress with functionality being continually added.

A key weakness of this thesis is the fact that the algorithms presented are not compared to each other. If SASACI and CAFI were implemented within an imaging platform, such as cueART, an in-depth comparison can be made between the two techniques and they can be evaluated against each another.

In its current state, SASACI and TFM are both functional and use the rapid imaging process described in Chapter 3, thus they can be used while accounting for refraction. The software allows the user to modify any imaging parameters on the fly to enable fine-tuning of algorithms. Users can also switch seamlessly between the imaging techniques for quick and easy comparison between output images. This software also interfaces with a robotic system (Kuka, Germany) and can be used for automated inspection of large components[133].

Future work for this platform involves integrating live surface recognition so that arbitrary surfaces can be accounted for in real-time on a moving array. This will allow for rapid sector scanning of industrial components.

6.3 Future Work

6.3.1 Rapid Ultrasonic Imaging

Detailed error analysis shows that the peak inherent error of the interpolant always occurs near the top of the Z-line, that is, where the non-linearity of the $f(z)$ function is the highest. One could exploit this fact and split the TFM volume into two zones, with separate interpolant coefficient databases. Since Phases 1 and

2 take a small fraction of time taken by Phase 3, this would allow the interpolant for each zone to better adhere to the original $f(z)$ function, possibly reducing the required interpolant order, at a very small extra computational cost.

In the current implementation, the full time of flight is calculated for each A-scan as a sum of Tx-to-pixel and pixel-to-Rx time of flight. At this point, if the database contains a series of A-scans gathered using the same Tx, the Tx-to-pixel time of flight value could be cached and only recalculated when advancing to the next Tx element. In general, depending on the exact Tx/Rx firing scheme, a range of interpolant caching scenarios or symmetry scenarios could be employed. It is anticipated that in a practical NDE imaging scenario, only a subset of the FMC will be actually needed to obtain sufficient image contrast; in such a case, the field is open for research on optimal firing schemes and related time of flight caching schemes.

This approach to imaging currently uses a lot of proprietary CUDA techniques that limit the platforms on which this code can run. In the future, the methodology could be ported to OpenCL, a parallel programming language that supports multiple platforms and architectures.

6.3.2 Spatially Averaged Sub-Aperture Correlation Imaging

The SASACI imaging technique has been shown to increase SNR in noisy images, but has a user defined threshold. Any areas with a cross-correlation coefficient below the threshold are further minimised. The threshold is set manually and requires some trial-and-error to find an ideal value for each inspection. There is potential for the threshold parameter to be automatically tuned so that the algorithm can intelligently select a value for this parameter by sweeping through a range of values and selecting the one which results in the most optimal image.

The Receiver Operator Characteristic (ROC) can be used to assess the performance of inspection methods. The performance of SASACI has not yet been assessed in this way but for a detailed comparison between different imaging methods, ROC should be used[134].

One of the aspects of SASACI not fully explored is the random apodisation technique. A number of random array configurations have been created and used for imaging. As expected, the results vary depending on the array elements chosen for each permutation. Data could be collected from a calibration block with known reflectors and imaged using a range of permutations until the SNR has been optimised. This is possible by using either a genetic algorithm or a brute force approach.

6.3.3 Correlation for Adaptively Focused Imaging

Correlation for Adaptively Focused Imaging has shown potential to improve focus in scenarios where the speed of sound is not constant throughout a material. Currently the algorithm is implemented in MATLAB and takes a significant amount of time (in the order of hours for a small image) is required to generate results. This limits the size and resolution of images being generated by the algorithm. SASACI has been implemented successfully in CUDA and can be run using GP-

GPU processing. CAFI can be implemented in the same way, removing the limitation of resolution and allowing for real-time processing of data. Real-time processing will make the technique industrially viable.

Finite element simulation can be used to model a range of scenarios quickly and efficiently. It can be used to test how CAFI copes with different situations. CAFI can be refined and improved by analysing the result from these modelling scenarios.

Using the full CAFI methods, different corrected images are produced using phase-based correction and amplitude-based correction. At the moment, the technique uses cross-correlation to extract information from this pair of images. There are other comparison methods which have not been investigated in this thesis, such as the Best Linear Unbiased Estimator (BLUE) technique[135].

6.4 Overall Conclusion

The aim of this thesis is to present an investigation into signal-processing based solutions for the problem of inspecting difficult materials using ultrasound arrays.

An approach to imaging through non-flat surfaces was presented which allows compensation for refraction in real-time. This is a significant improvement on similar techniques that require either prior knowledge of the sample or more data than a single FMC.

Two signal processing methods for image generation have been presented which are novel and tackle the problem of inspecting materials with large grains or anisotropy. Experimental results have been obtained from industrially relevant samples for each of these techniques, and the results discussed with varying degrees of success.

In summary, signal processing can be used with traditional data-gathering techniques to improve the quality of ultrasonic imaging and increase the probability of locating defects. Each of the methodologies presented in this thesis are effective and have potential to be explored further via future research.

References

[1] X. Zhao, H. Gao, G. Zhang, B. Ayhan, F. Yan, C. Kwan, and J. L. Rose, "Active health monitoring of an aircraft wing with embedded piezoelectric sensor/actuator network: I. Defect detection, localization and growth monitoring," *Smart Materials and Structures*, vol. 16, p. 1208, Aug. 2007. 1

[2] P. Roscoe, C. Waites, and M. J. Whittle, "Cost-Effective Inspection Qualification," in *Proceedings of the 13. international conference on NDE in the nuclear and pressure vessel industries*, (Kyoto), ASM International, Aug. 1995. 1

[3] R. Viswanathan, J. F. Henry, J. Tanzosh, G. Stanko, J. Shingledecker, B. Vitalis, and R. Purgert, "U.S. program on materials technology for ultra-supercritical coal power plants," *Journal of Materials Engineering and Performance*, vol. 14, pp. 281–292, June 2005. 1

[4] I. I. Novikov, "The efficiency of atomic power stations (a review)," *Journal of Nuclear Energy*, vol. 7, pp. 125–128, Aug. 1958. 1

[5] R. Viswanathan, K. Coleman, and U. Rao, "Materials for ultra-supercritical coal-fired power plant boilers," *International Journal of Pressure Vessels and Piping*, vol. 83, pp. 778–783, Nov. 2006. 1

[6] C. Descamps, C. Bouallou, and M. Kanniche, "Efficiency of an Integrated Gasification Combined Cycle (IGCC) power plant including CO2 removal," *Energy*, vol. 33, pp. 874–881, June 2008. 1

[7] P. J. Ennis and A. Czyrska-Filemonowicz, "Recent advances in creep-resistant steels for power plant applications," *Sadhana*, vol. 28, pp. 709–730, June 2003. 1

[8] R. Viswanathan and W. Bakker, "Materials for ultrasupercritical coal power plantsBoiler materials: Part 1," *Journal of Materials Engineering and Performance*, vol. 10, pp. 81–95, Feb. 2001. 1

[9] G. Hayner and W. Windes, "Next Generation Nuclear Plant Materials Research and Development Program Plan," Sept. 2004. 1

[10] F. Masuyama, "History of Power Plants and Progress in Heat Resistant Steels," *ISIJ International*, vol. 41, no. 6, pp. 612–625, 2001. 1

[11] R. L. Klueh and A. T. Nelson, "Ferritic/martensitic steels for next-generation reactors," *Journal of Nuclear Materials*, vol. 371, pp. 37–52, Sept. 2007. 1

[12] H. G. Jiang, M. L. Lau, and E. J. Lavernia, "Grain growth behavior of nanocrystalline Inconel 718 and Ni powders and coatings," *Nanostructured Materials*, vol. 10, pp. 169–178, Feb. 1998. 1

[13] G. S. Was, H. H. Tischner, and R. M. Latanision, "The influence of thermal treatment on the chemistry and structure of grain boundaries in inconel 600," *Metallurgical Transactions A*, vol. 12, pp. 1397–1408, Aug. 1981. 2

[14] A. Diboine and A. Pineau, "Creep Crack Initiation and Growth in Inconel 718 Alloy at 650c," *Fatigue & Fracture of Engineering Materials & Structures*, vol. 10, pp. 141–151, Feb. 1987. 2

[15] F. P. Ford and P. L. Andresen, "Development and use of a predictive model of crack propagation in 304/316l, A533b/A508 and Inconel 600/182 alloys in 288^0C water," in *Proceedings of the third international symposium on environmental degradation of materials in nuclear power systems*, pp. 789–800, 1988. 2

[16] P. J. Ennis, A. Zielinska-Lipiec, O. Wachter, and A. Czyrska-Filemonowicz, "Microstructural stability and creep rupture strength of the martensitic steel P92 for advanced power plant," *Acta Materialia*, vol. 45, pp. 4901–4907, Dec. 1997. 2

[17] M. de Witte, "Power plant life estimation and extension: The Belgian experience from the users' point of view," *International Journal of Pressure Vessels and Piping*, vol. 39, no. 12, pp. 41–55, 1989. 2

[18] S. Thomas, "Power-plant life extension," *Energy*, vol. 13, pp. 767–786, Oct. 1988. 2

[19] N. Dominelli, A. Rao, and P. Kundur, "Life extension and condition assessment: techniques for an aging utility infrastructure," *IEEE Power and Energy Magazine*, vol. 4, pp. 24–35, May 2006. 2

[20] D. Ellerman, "Note on the seemingly indefinite extension of power plant lives, a panel contribution," *The Energy Journal*, vol. 19, pp. 129–132, Jan. 1998. 2

[21] A. K. Ray, Y. N. Tiwari, R. K. Sinha, S. Chaudhuri, and R. Singh, “Residual life prediction of service exposed main steam pipe of boilers in a thermal power plant,” *Engineering Failure Analysis*, vol. 7, pp. 359–376, Oct. 2000. 2

[22] J. Tong, S. Dalby, J. Byrne, M. B. Henderson, and M. C. Hardy, “Creep, fatigue and oxidation in crack growth in advanced nickel base superalloys,” *International Journal of Fatigue*, vol. 23, pp. 897–902, Nov. 2001. 2

[23] J. R. Rice, *Mechanics of Crack Growth.* ASTM International, Jan. 1976. 2

[24] J. Saniie, B. Panda, T. Wang, and D. Nagle, “Life assessment of creep degraded super alloy materials using ultrasound,” in *Ultrasonics Symposium, 1990. Proceedings., IEEE 1990*, pp. 987–990 vol.2, Dec. 1990. 2

[25] C. M. Sehgal and J. F. Greenleaf, “Diffraction of Ultrasound by Soft Tissues:The Inhomogeneous Continuous Model,” in *Acoustical Imaging* (M. Kaveh, R. K. Mueller, and J. F. Greenleaf, eds.), pp. 217–231, Springer US, 1984. 2

[26] G. Sposito, C. Ward, P. Cawley, P. B. Nagy, and C. Scruby, “A review of non-destructive techniques for the detection of creep damage in power plant steels,” *NDT & E International*, vol. 43, pp. 555–567, Oct. 2010. 2

[27] B. W. Drinkwater and P. D. Wilcox, “Ultrasonic arrays for non-destructive evaluation: A review,” *NDT & E International*, vol. 39, pp. 525–541, Oct. 2006. 3, 28, 38

[28] J. L. Rose and B. B. Goldberg, *Basic Physics in Diagnostic Ultrasound.* New York: John Wiley & Sons Inc, Nov. 1979. 9

[29] J. Achenbach, *Wave Propagation in Elastic Solids.* Elsevier, Dec. 2012. 10

[30] R. Halmshaw, *Non-destructive testing.* London: E. Arnold, 1991. 11

[31] A. J. Bush and G. Y. Baladi, *Nondestructive Testing of Pavements and Backcalculation of Moduli.* ASTM International, 1989. 11

[32] A. P. Boresi, R. J. Schmidt, and O. M. Sidebottom, *Advanced mechanics of materials*, vol. 5. Wiley New York, 1993. 11

[33] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics, Part I.* Boston, Mass: McGraw-Hill Science/Engineering/Math, June 1953. 12

[34] J. Rychlewski, "On Hooke's law," *Journal of Applied Mathematics and Mechanics*, vol. 48, no. 3, pp. 303–314, 1984. 13

[35] K. F. Graff, *Wave Motion in Elastic Solids.* Courier Corporation, Apr. 2012. 14

[36] D. Royer and E. Dieulesaint, *Elastic Waves in Solids II: Generation, Acousto-optic Interaction, Applications.* Springer Science & Business Media, Jan. 2000. 14

[37] J. Feldman, "Derivation of the Wave Equation," 2000. 17

[38] E. Kuhnicke, "The limitations of Snell's law in the design of ultrasound transducers," in *1999 IEEE Ultrasonics Symposium, 1999. Proceedings*, vol. 2, pp. 1067–1070 vol.2, 1999. 22

[39] Transducer, *Concise Oxford English Dictionary: Eleventh edition.* Oxford: OUP Oxford, 11th revised edition edition ed., June 2006. 25

[40] S. Rosen and P. Howell, *Signals and Systems for Speech and Hearing 2nd edition.* United Kingdom: Emerald Group Publishing Limited, 2 edition ed., Dec. 2010. 26

[41] H. W. St Clair, "An Electromagnetic Sound Generator for Producing Intense High Frequency Sound," *Review of Scientific Instruments*, vol. 12, pp. 250–256, May 1941. 26

[42] W. M. D. Wright, D. W. Schindel, and D. A. Hutchins, "Studies of lasergenerated ultrasound using a micromachined silicon electrostatic transducer in air," *The Journal of the Acoustical Society of America*, vol. 95, pp. 2567–2575, May 1994. 26

[43] R. Heydt, R. Pelrine, J. Joseph, J. Eckerle, and R. Kornbluh, "Acoustical performance of an electrostrictive polymer film loudspeaker," *The Journal of the Acoustical Society of America*, vol. 107, pp. 833–839, Feb. 2000. 26

[44] J. Ryu, A. V. Carazo, K. Uchino, and H.-E. Kim, "Magnetoelectric Properties in Piezoelectric and Magnetostrictive Laminate Composites," *Japanese Journal of Applied Physics*, vol. 40, p. 4948, Aug. 2001. 26

[45] G. L. Gooberman, *Ultrasonics: Theory and Application.* Hart Publishing Company, 1969. 26

[46] A. A. Vives, *Piezoelectric Transducers and Applications.* Springer Science & Business Media, Mar. 2013. 26

[47] A. Fukumoto, M. Kawabuchi, and J. Sato, "Design of Ultrasound Transducers Using New Piezoelectric Ceramic Materials," *Ultrasound in Medicine and Biology*, vol. 7, no. 3, pp. 275–281, 1981. 27

[48] T. Freegarde, *Introduction to the Physics of Waves.* Paris, France : Amsterdam: Cambridge University Press, pap/psc edition ed., Nov. 2012. 28

[49] B. B. Baker and E. T. Copson, *The Mathematical Theory of Huygens' Principle.* American Mathematical Soc., 2003. 28

[50] F. L. Thurstone, N. I. Kjosnes, and W. M. McKinney, "Ultrasonic Scanning of Biologic Tissue by a New Technique," *Science (New York, N.Y.)*, vol. 149, pp. 302–303, July 1965. 33, 36

[51] D. N. Alleyne and P. Cawley, "Optimization of lamb wave inspection techniques," *NDT & E International*, vol. 25, no. 1, pp. 11–22, 1992. 36

[52] J. C. Lazaro, J. L. San Emeterio, A. Ramos, and J. L. Fernandez-Marron, "Influence of thresholding procedures in ultrasonic grain noise reduction using wavelets," *Ultrasonics*, vol. 40, pp. 263–267, May 2002. 36

[53] E. Buddemeyer, "Physics of Diagnostic Ultrasound," *Radiologic Clinics of North America*, vol. 13, no. 3, pp. 391–402, 1975. 36

[54] J. Santodomingo-Rubido, E. a. H. Mallen, B. Gilmartin, and J. S. Wolffsohn, "A new non-contact optical device for ocular biometry," *British Journal of Ophthalmology*, vol. 86, pp. 458–462, Apr. 2002. 36

[55] R. Harrington, “Sidelobe reduction by nonuniform element spacing,” *IRE Transactions on Antennas and Propagation*, vol. 9, pp. 187–192, Mar. 1961. 38, 45

[56] C. H. Seo and J. Yen, “Sidelobe suppression in ultrasound imaging using dual apodization with cross-correlation,” *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 55, pp. 2198–2210, Oct. 2008. 38, 101, 103, 127, 145

[57] San Diego Plastics Inc., “Rexolite,” 1996. [http://www.sdplastics.com/rexolite.html; accessed 27-June-2015]. 39

[58] F. Simonetti, “Multiple scattering: The key to unravel the subwavelength world from the far-field pattern of a scattered wave,” *Physical Review E*, vol. 73, p. 036619, Mar. 2006. 40

[59] A. Nowicki, I. Trots, P. A. Lewin, W. Secomski, and R. Tymkiewicz, “Influence of the ultrasound transducer bandwidth on selection of the complementary Golay bit code length,” *Ultrasonics*, vol. 47, pp. 64–73, Dec. 2007. 40

[60] A. Schuster, *An introduction to the theory of optics.* London, E. Arnold, 1904. 43, 67

[61] G. D. Connolly, M. J. S. Lowe, J. A. G. Temple, and S. I. Rokhlin, “The application of Fermat’s principle for imaging anisotropic and inhomogeneous media with application to austenitic steel weld inspection,” *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 465, pp. 3401–3423, Nov. 2009. 43

[62] S. W. Smith, H. G. Pavy Jr, and O. T. von Ramm, "High-speed ultrasound volumetric imaging system. I. Transducer design and beam steering," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 38, no. 2, pp. 100–108, 1991. 45

[63] W. Kummer, A. Villeneuve, T. Fong, and F. Terrio, "Ultra-low sidelobes from time-modulated arrays," *IEEE Transactions on Antennas and Propagation*, vol. 11, pp. 633–639, Nov. 1963. 45

[64] R. Long, J. Russell, and P. Cawley, "Ultrasonic phased array inspection using full matrix capture," *Insight - Non-Destructive Testing and Condition Monitoring*, vol. 54, pp. 380–385, July 2012. 45

[65] S. Chatillon, S. Mahaut, and P. Dubois, "Simulation of Advanced Ut Phased Array Techniques with Matrix Probes and Dynamic Settings for Complex Component Inspections," in *Review of Progress in Quantitative Nondestructive Evaluation* (D. O. Thompson and D. E. Chimenti, eds.), vol. 1096, (Melville), pp. 864–871, Amer Inst Physics, 2009. 48

[66] J. Camacho, M. Parrilla, and C. Fritsch, "Phase coherence imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, pp. 958–974, May 2009. 50, 69, 80, 100, 182

[67] J. Camacho and C. Fritsch, "Phase coherence imaging of grained materials," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 58, pp. 1006–1015, May 2011. 50, 101, 166

[68] F. Vignon and M. R. Burcher, "Capon beamforming in medical ultrasound imaging with focused beams," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 55, pp. 619–628, Mar. 2008. 50

[69] M. Sasso and C. Cohen-Bacrie, "Medical ultrasound imaging using the fully adaptive beamformer," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05)*, vol. 2, pp. 489–492, Mar. 2005. 50

[70] B. M. Asl and A. Mahloojifar, "Minimum variance beamforming combined with adaptive coherence weighting applied to medical ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, pp. 1923–1931, Sept. 2009. 50

[71] J.-F. Synnevg, A. Austeng, and S. Holm, "Minimum variance adaptive beamforming applied to medical ultrasound imaging," in *2005 IEEE Ultrasonics Symposium*, vol. 2, pp. 1199–1202, Sept. 2005. 50

[72] C.-I. Nilsen and S. Holm, "Wiener beamforming and the coherence factor in ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 57, pp. 1329–1346, June 2010. 50

[73] S. R. Degraaf, "Sidelobe reduction via adaptive FIR filtering in SAR imagery," *IEEE Transactions on Image Processing*, vol. 3, no. 3, pp. 292–301, 1994. 50

[74] C.-I. C. Nilsen and I. Hafizovic, "Beamspace adaptive beamforming for ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 56, pp. 2187–2197, Oct. 2009. 50

[75] S.-L. Wang and P.-C. Li, “MVDR-based coherence weighting for high-frame-rate adaptive imaging,” *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 56, pp. 2097–2110, Oct. 2009. 50

[76] M. Sakhaei, A. Mahloojifar, and A. Malek, “Optimization of point spread function in ultrasound arrays,” *Ultrasonics*, vol. 44, pp. 159–165, Feb. 2006. 51

[77] M. K. Jeong, “A Fourier transform-based sidelobe reduction method in ultrasound imaging,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 47, no. 3, pp. 759–763, 2000. 51

[78] S. M. Sakhaei, “Optimum beamforming for sidelobe reduction in ultrasound imaging,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 59, pp. 799–805, Apr. 2012. 51

[79] C. Holmes, B. Drinkwater, and P. Wilcox, “Post-processing of the full matrix of ultrasonic transmitreceive array data for non-destructive evaluation,” *NDT & E International*, vol. 38, pp. 701–711, Dec. 2005. 51, 54, 64, 66, 102

[80] M. Ali, D. Magee, and U. Dasgupta, “Signal processing overview of ultrasound systems for medical imaging,” *SPRAB12, Texas Instruments, Texas*, 2008. 55

[81] P. D. Wilcox, C. Holmes, and B. W. Drinkwater, “Advanced reflector characterization with ultrasonic phased arrays in NDE applications,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 54, pp. 1541–1550, Aug. 2007. 55

[82] P. C. Li and M. L. Li, “Adaptive imaging using the generalized coherence factor,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 50, pp. 128–141, Feb. 2003. 56, 101

[83] R. Gongzhang, M. Li, B. Xiao, T. Lardner, and A. Gachagan, “Robust frequency diversity based algorithm for clutter noise reduction of ultrasonic signals using multiple sub-spectrum phase coherence,” in *Proceedings of the 40th Annual Review of Progress in Quantitative Nondestructive Evaluation: Incorporating the 10th International Conference on Barkhausen Noise and Micromagnetic Testing*, pp. 1948–1955, 2014. 56

[84] R. Gongzhang, M. Li, T. Lardner, and A. Gachagan, “Robust defect detection in ultrasonic nondestructive evaluation (NDE) of difficult materials,” in *Ultrasonics Symposium (IUS), 2012 IEEE International*, pp. 467–470, Oct. 2012. 56

[85] S. Banerjee and T. Kundu, “Elastic wave propagation in sinusoidally corrugated waveguides,” *Journal of the Acoustical Society of America*, vol. 119, pp. 2006–2017, Apr. 2006. 57

[86] Weidlinger Associates Inc., “PZFlex,” 2014. [http://www.pzflex.com/; accessed 27-June-2015]. 58

[87] J. Dobson, A. Tweedie, G. Harvey, R. OLeary, A. Mulholland, K. Tant, and A. Gachagan, “Finite element analysis simulations for ultrasonic array NDE inspections,” in *AIP Conference Proceedings*, vol. 1706, p. 040005, AIP Publishing, Feb. 2016. 58

[88] R. N. Bracewell, *The Fourier Transform and Its Applications.* McGraw Hill, 2000. 59

[89] J. Szilard, *Ultrasonic Testing: Non-conventional Testing Techniques.* Chichester Eng. ; New York: Wiley-Blackwell, Apr. 1982. 59

[90] University of Bristol, "The Frequency Domain," Nov. 2011. 60

[91] Zetec Inc., "UltraVision," 2015. [http://ultravision.zetec.com/; accessed 27-June-2015]. 62

[92] Commissariat à l'énergie atomique et aux énergies alternatives, "CIVA: solutions for NDT," 2015. [http://www-civa.cea.fr/en/; accessed 27-June-2015]. 62

[93] Y. Humeida, V. Pinfield, R. Challis, P. Wilcox, and C. Li, "Simulation of ultrasonic array imaging of composite materials with defects," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 60, pp. 1935–1948, Sept. 2013. 62

[94] A. McGilp, J. Dziewierz, T. Lardner, J. Mackersie, and A. Gachagan, "Inspection design using 2D phased array, TFM and cueMAP software," in *AIP Conference Proceedings*, vol. 1581, pp. 65–71, AIP Publishing, Feb. 2014. 62

[95] J. Dziewierz, *2D ultrasonic phased arrays for quantitative characterisation of complex defects.* PhD thesis, University of Strathclyde, 2015. 63

[96] A. McGilp, *Adaptation of TFM for Inspection of Complex Components.* EngD thesis, University of Strathclyde, Jan. 2016. 63, 95

[97] J. Russell, R. Long, and P. Cawley, “Development of a Membrane Coupled Conformable Phased Array Inspection Capability,” in *Review of Progress in Quantitative Nondestructive Evaluation: Volume 29*, vol. 1211, pp. 831–838, Feb. 2010. 63

[98] F. Reverdy and G. Ithurralde, “Advanced Ultrasonic 2d Phased-Array Probes,” in *18th World Conference on Non-Destructive Testing*, (Durban, South Africa), 2012. 65

[99] D. Cassereau, N. Chakroun, M. Fink, and F. Wu, “Theoretical and experimental analysis of focusing techniques through liquid-solid interfaces,” in *, 1994 IEEE Ultrasonics Symposium, 1994. Proceedings*, vol. 2, pp. 1075 –1080 vol.2, Nov. 1994. 67

[100] J. P. Jones, *Acoustical Imaging.* Springer Science & Business Media, Dec. 2012. 67

[101] L. Cunningham, A. Mulholland, G. Harvey, and C. Bird, “Ultrasonic wave propagation in heterogenous media,” in *Proceedings of Acoustics*, 2012. 67

[102] J. Kitze, J. Prager, R. Boehm, U. Vlz, and H.-J. Montag, “Saft-reconstruction in ultrasonic immersion technique using phased array transducers,” in *Review of Progress in Quantitative Nondestructive Evaluation*, vol. 1430, pp. 825–832, May 2012. 68

[103] M. Weston, P. Mudge, C. Davis, and A. Peyton, “Time efficient auto-focussing algorithms for ultrasonic inspection of dual-layered media using Full Matrix Capture,” *NDT & E International*, vol. 47, pp. 43–50, Apr. 2012. 68

[104] C. Fritsch, J. F. Cruza, J. Brizuela, J. Camacho, and J. M. Moreno, "A new technique for fast dynamic focusing law computing," in *AIP Conference Proceedings*, vol. 1433, pp. 165–168, AIP Publishing, May 2012. 68

[105] L. M. Brekhovskikh, *Waves in layered media.* Academic Press, 1980. 68

[106] J. Zhang, B. W. Drinkwater, P. D. Wilcox, and A. J. Hunter, "Defect detection using ultrasonic arrays: The multi-mode total focusing method," *NDT & E International*, vol. 43, pp. 123–133, Mar. 2010. 68

[107] A. Tweedie, R. O'Leary, G. Harvey, A. Gachagan, C. Holmes, P. Wilcox, and B. Drinkwater, "Total Focussing Method for Volumetric Imaging in Immersion Non Destructive Evaluation," in *IEEE Ultrasonics Symposium, 2007*, pp. 1017–1020, Oct. 2007. 68

[108] M. Y. Matuda, F. Buiochi, and J. C. Adamowski, "Imaging through a convex interface with unknown position and shape using an ultrasonic linear array," in *AIP Conference Proceedings*, vol. 1433, pp. 177–180, AIP Publishing, May 2012. 69

[109] J. Zhang, F. Li, and G. Yang, "Ray-based Simulations of Received Signals from Ground Penetrating Radar," in *International Conference on Radar, 2006. CIE '06*, pp. 1 –4, Oct. 2006. 69

[110] F. Jurado, P. Lailly, and A. Ehinger, "Fast 3D two-point ray tracing for travel-time tomography," in *Mathematical Methods in Geophysical Imaging V*, vol. 3453, pp. 70–81, 1998. 69

[111] M. Sutcliffe, M. Weston, B. Dutton, I. Cooper, and K. Donne, "Real-time full matrix capture with auto-focusing of known geometry through dual layered media," in *51st Annual Conference of The British Institute of Non-Destructive Testing*, 2012. 69

[112] J. Lambert, A. Pedron, G. Gens, F. Bimbard, L. Lacassagne, and E. Iakovleva, "Performance evaluation of total focusing method on GPP and GPU," in *Design and Architectures for Signal and Image Processing (DASIP), 2012 Conference on*, pp. 1–8, 2012. 70

[113] B. Y. S. Yiu, I. K. H. Tsang, and A. C. H. Yu, "GPU-based beamformer: Fast realization of plane wave compounding and synthetic aperture imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 58, pp. 1698–1705, Aug. 2011. 70

[114] J. Zhang, B. Drinkwater, and P. Wilcox, "Efficient immersion imaging of components with nonplanar surfaces," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 61, pp. 1284–1295, Aug. 2014. 71, 76

[115] NVIDIA Corporation, "CUDA C Programming Guide," 2015. [http://docs.nvidia.com/cuda/cuda-c-programming-guide; accessed 20-May-2015]. 73

[116] J. Dziewierz and A. Gachagan, "Computationally efficient solution of snell's law of refraction," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 60, pp. 1256–1259, June 2013. 73

[117] E. Weisstein, “Quintic equation,” 2004. [http://docs.nvidia.com/cuda/cuda-c-programming-guide; accessed 20-May-2015]. 73

[118] J. Dziewierz, T. Lardner, and A. Gachagan, “A design methodology for 2D sparse NDE arrays using an efficient implementation of refracted-ray TFM,” in *Ultrasonics Symposium (IUS), 2013 IEEE International*, pp. 136–138, IEEE, July 2013. 74

[119] A. McGilp, T. Lardner, J. Dziewierz, A. Gachagan, J. Mackersie, and C. Bird, “Inspection of complex components using 2D arrays and TFM,” in *53rd Annual Conference of the British Institute of Non-Destructive Testing (NDT 2014)*, (Manchester, UK), Curran Associates, Inc., Jan. 2015. 74, 93

[120] W. Dorn, “Generalizations of Horner’s Rule for Polynomial Evaluation,” *IBM Journal of Research and Development*, vol. 6, pp. 239–245, Apr. 1962. 79

[121] D. I. A. Lines, “Rapid distributed data collection with arrays the next step beyond full waveform capture,” *Insight - Non-Destructive Testing and Condition Monitoring*, vol. 48, pp. 84–88, Feb. 2006. 80

[122] D. Lee, I. Dinov, B. Dong, B. Gutman, I. Yanovsky, and A. W. Toga, “CUDA optimization strategies for compute- and memory-bound neuroimaging algorithms,” *Computer Methods and Programs in Biomedicine*, vol. 106, pp. 175–187, June 2012. 81

[123] M. F. Wagdy and M. S. Lucas, “Errors in Sampled Data Phase Measurement,” *IEEE Transactions on Instrumentation and Measurement*, vol. 34, pp. 507–509, Dec. 1985. 82

[124] A. DallOsso, “Computer algebra systems as mathematical optimizing compilers,” *Science of Computer Programming*, vol. 59, pp. 250–273, Feb. 2006. 83

[125] J. W. Choe, A. Nikoozadeh, m. Oralkan, and B. T. Khuri-Yakub, “GPU-Based Real-Time Volumetric Ultrasound Image Reconstruction for a Ring Array,” *IEEE Transactions on Medical Imaging*, vol. 32, pp. 1258–1264, July 2013. 86

[126] T. Y. Phuong and J.-G. Lee, “Design space exploration of SW beamformer on GPU,” *Concurrency and Computation: Practice and Experience*, vol. 27, pp. 1718–1733, May 2015. 86

[127] S. U. Gjerald, R. Brekken, T. Hergum, and J. D’hooge, “Real-time ultrasound simulation using the GPU,” in *Ultrasonics Symposium (IUS), 2011 IEEE International*, pp. 258–261, Oct. 2011. 86

[128] C. Fan, M. Caleap, M. Pan, and B. W. Drinkwater, “A comparison between ultrasonic array beamforming and super resolution imaging algorithms for non-destructive evaluation,” *Ultrasonics*, vol. 54, pp. 1842–1850, Sept. 2014. 100

[129] C. Lane, A. K. Dunhill, B. Drinkwater, and P. Wilcox, “The inspection of anisotropic single-crystal components using a 2-D ultrasonic array,” *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 57, pp. 2742–2752, Dec. 2010. 100

[130] T. Lardner, M. Li, R. Gongzhang, and A. Gachagan, “A new speckle noise suppression technique using cross-correlation of array sub-apertures in ultrasonic NDE of coarse grain materials,” in *Review of Progress in Quantitative Non-Destructive Evaluation*, vol. 1511, pp. 865–871, Jan. 2013. 144, 168

[131] T. Lardner, M. Li, and A. Gachagan, “Using phase information to enhance speckle noise reduction in the ultrasonic NDE of coarse grain materials,” in *AIP Conference Proceedings*, vol. 1581, pp. 1061–1068, AIP Publishing, Feb. 2014. 144

[132] S. Flax and M. O’Donnell, “Phase-aberration correction using signals from point reflectors and diffuse scatterers: basic principles,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 35, pp. 758–767, Nov. 1988. 145

[133] R. H. Brown, S. G. Pierce, I. Collison, B. Dutton, J. Dziewierz, J. Jackson, T. Lardner, C. MacLeod, and M. Morozov, “Automated full matrix capture for industrial processes,” in *AIP Conference Proceedings*, vol. 1650, pp. 1967–1976, AIP Publishing, Mar. 2015. 184

[134] A. Van Pamel, C. Brett, and M. Lowe, “A methodology for evaluating detection performance of ultrasonic array imaging algorithms for coarse-grained materials,” *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 61, pp. 2042–2053, Dec. 2014. 186

[135] R. Fjortoft, A. Lopes, P. Marthon, and E. Cubero-Castan, “An optimal multiedge detector for SAR image segmentation,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, pp. 793–802, May 1998. 187