GENERATION OF HIGH RESOLUTION TOMOGRAPHIC IMAGES FOR NDT APPLICATIONS.

Ian D. Hall

Submitted for the degree of
Doctor of Philosophy

2001

Ultrasonics Research Group
Department of Electrical and Electronic Engineering
University of Strathclyde

Acknowledgements.

First, I would like to express my thanks to Gordon Hayward and Alistair McNab for giving me the opportunity to conduct this research. I would also like to express my gratitude to Alistair McNab for his guidance and support as my supervisor over the duration of my study.

I would like to thank all the other members of the Ultrasonic Research Group both past and present who have given me a helping hand along the way.

And last but not least I would like to thank my parents and my long suffering girlfriend Joanna for the support they have given me throughout my long university career.

Contents.

Abstract.

Techniques for the generation of quantitative ultrasonic images in non-destructive testing have generally involved a substantial cost in terms of data storage and computational time, and have thus found limited application. Preference has therefore been given to the more straightforward imaging methods, such as main beam projection, which detect the presence of defects and provide a limited flaw sizing capability.

The relatively small number of flaws requiring detailed examination, coupled with substantial increases in available data storage and computational power, have made it possible to use a number of straightforward tomographic reconstruction methods to produce high resolution images of flaws contained within the material under examination. A set of these images are then fused together using a novel fuzzy logic image fusion technique into a single image from which more accurate measurements of flaw size, shape and orientation can be made. However, if the quality of the raw A-scan data is not sufficiently high then the data will be filtered using Maximum Likelihood Deconvolution (MLD). The aim of this blind deconvolution method is to improve the time resolution and Signal to Noise Ratio of the A-scan data with only knowledge obtained from the data, this is in contrast to the majority of techniques currently used for this purpose.

The three tomographic methods which have been implemented in this work are Reflection tomography, Time-of-flight Diffraction tomography and Transmission tomography. In addition a Single Bounce Image Enhancement method has been developed to improve the images. Selection of images used in the fusion process depends on the nature of the flaw, as each of these methods identifies different characteristics of the flaw shape. The components of the imaging system have been validated experimentally by the generation of high resolution images from a variety of flaws contained within cylindrical aluminium test specimens.

Chapter 1.

INTRODUCTION.

1.1. Background.

Ultrasonic inspection is one of the most powerful and widely used methods for the interrogation of a component interior in Non-Destructive Testing (NDT) today. There are many advantages of using ultrasound to perform inspections and some of them will be briefly outlined here. Ultrasound and radiography are the only two methods apart from thermography that are available for the examination of a component's interior. Ultrasound is the preferred method of inspection as it is good at detecting planar defects such as cracks even if they are very thin. Radiography on the other hand is good at detecting volumetric defects that represent a lower risk to the component's integrity than planar flaws such as cracks or lack of sidewall fusion defects. An added advantage of ultrasound is that it presents no danger to the material under test, or the operator of the test equipment (i.e. no source of ionising radiation, toxic chemicals). Generally NDT can be utilised to predict the failure of structures and therefore, helps to increase productivity and prevent accidental damage or injury.

Ultrasonic data acquisition can be performed in a number of ways. If the area to be examined is relatively small then a manual operator can perform the acquisition. However, if the area under examination is large or in a hazardous environment then the acquisition will be performed using automatic scanning systems. The subsequent interpretation of the obtained data can be a very complex task. For even modest sized test structures, the amount of ultrasonic data generated can be large. If manual interpretation of this data is used, then the time taken can be prohibitively long. With long processing times, manual operators will often suffer from boredom and fatigue, which can lead to possible errors in defect identification [1][2]. Therefore, there is great potential for computers to aid in the interpretation of ultrasonic data [3]. If a number of operator tasks can be performed by a computer system, then the defect detection and characterisation process is made more reliable and less time consuming.

Computer systems can aid the interpretation in two ways: by processing the data and by improving its presentation. This thesis will concentrate on processing the

ultrasonic data to improve the inspection task. The objective of data processing is to increase the signal-to-noise ratio of the data; that is to enhance defect signals and suppress noise signals.

The relatively small number of defects requiring detailed examination, coupled to substantial increases in available data storage and computational power have made it possible to use a number of tomographic reconstruction techniques to generate images of defects contained within the material under examination. These images can then be fused together into a single image from which more accurate measurements of the defect size, shape and orientation can be made. However, if the quality of the raw A-scan data is not sufficiently high then the data will require filtering. This can be achieved by utilising a Maximum Likelihood Deconvolution (MLD) / Minimum Variance Deconvolution (MVD) filter, the aim being to improve the time resolution and Signal to Noise Ratio of the A-scan data.

The methodology of reconstructing a number of simple tomographic images and fusing them was chosen over the development of a single more complex reconstruction algorithm, using all of the acquired data. This maintains theoretical simplicity, leading to lower computational requirements and greater flexibility for the user.

1.2. Existing techniques for interpretation of ultrasonic data in NDT.

Traditionally techniques for the generation of quantitative ultrasonic images have involved substantial cost in terms of data storage and computational time, and thus have found limited application. Preference has therefore been given to more straightforward imaging methods, such as main beam projection, which detect the presence of defects but provide a limited flaw sizing capability.

A brief overview of some of the currently used NDT imaging methods is presented in Section 1.2.1. This will be followed by a discussion of applicable image fusion methods in Section 1.2.2. In many NDT applications the quality of the data is not sufficient to obtain the desired quality of image. In order to improve the quality of

the data it is necessary to filter the raw data prior to image reconstruction. This will be performed using deconvolution. A brief discussion of deconvolution for NDT applications will be given in Section 1.2.3.

1.2.1. Commonly used Imaging methods in NDT.

Main Beam Projection.

Main beam projection is one of the most basic imaging methods currently in use. Signals from each A-scan are simply projected along a line starting from the transmitter position and following the central probe axis. Since the beam is not perfectly focused, but has a finite beam spread, targets in the far-field become distorted. Another drawback is that a target detected by the edge of the spreading beam will be incorrectly positioned along the central beam at that position, for example a point reflector will become an arc and specular reflectors appear oversize [4]. However, the reconstruction process is very fast and flaws can be readily identified as in the characteristic arcs from crack tips.

TOFD (time-of-flight diffraction).

TOFD utilises the edge waves that are generated by defects. This method requires the use of at least two probes, one as a transmitter and one or more receivers. The most commonly used set-up is a two probe, pitch-catch set-up. Usually angled probes are used producing either longitudinal or transverse waves. This set-up is only suitable for scanning objects with simple geometry i.e. thick plates [5].

The most common application of this imaging method is in the inspection of welds [6]. This method is so successful for weld testing because of the relatively short distance from the receiver to the defect allowing the use of focused beams from probes of a handy size. This is an advantage for weld testing but for general NDT applications can cause problems [5].

Synthetic Aperture Focusing Technique (SAFT).

Synthetic Aperture Focusing Technique (SAFT) is the most commonly used technique for the generation of high quality NDT images for point and edge defects.

SAFT takes the A-scan data acquired from transducers at different locations and combines the data to simulate a single transducer with a large aperture, which can be accurately focused at all depths.

For each transmitter location every pixel in the image is provisionally considered a possible reflector position. All echoes with transit times corresponding to the distance from any particular transmitter position to each separate pixel are added up, having equal phase only if they are genuine echoes from a reflector in this pixel [4]. This image reconstruction takes place in either the time [7][8][9][10] or frequency domain [11]. The drawbacks of using this imaging method are that the reconstruction time is often long, especially when the method is extended to any data acquisition set-up more complex than pulse-echo.

Work has been done by Lorenz [9] on extending SAFT imaging to include information after the incident ultrasonic beam has been reflected by the backwall of the test specimen. This allows the maximum amount of defect information to be obtained from the available ultrasonic data.

Tomographic Image Reconstruction Techniques for NDT.

Transmission tomography methods for non-destructive testing applications are fundamentally different from the methods used in medical [12][13][14] and seismic applications [15] mainly due to the physical properties of the materials under test. In medical and seismic applications the human body and earth contain regions of differing acoustic velocity, so a velocity profile of the test object can be reconstructed from time-of-flight measurements. This is not the case for the majority of engineering components, as they are by nature made of a single material and therefore the velocity changes across the object cross section are negligible. Moreover, there is usually a large acoustic mismatch between the defect and surrounding media so the majority of ultrasound is either reflected or diffracted by the defect. For NDT methods that utilise the energy reflected or diffracted by the defect have to be considered [8][16][17][18][19][20][21][22][23].

1.2.2. Image Fusion for NDT.

Data fusion is a process dealing with data and information from multiple sources to achieve refined/improved information for decision-making [24]. Image fusion is a method of data fusion and can be defined as *'Image fusion is the combination of two or more different images to form a new image by using a certain algorithm'* [25]. Image fusion can be performed at one of three levels depending at what level the fusion is performed. The three processing levels are; pixel, feature and decision. Pixel level fusion is the lowest level of fusion and refers to the merging of physical parameters. Feature level fusion requires the extraction of objects recognised in the various data sources e.g. segmentation. Decision level fusion represents a method that uses value-added data where the input images are processed individually for information extraction [25].

The image fusion in this thesis is a pixel fusion technique based on fuzzy logic set theory. Some work in this area has been undertaken by Nejatalli and Ciric for application to medical images [26]. A similar image fusion approach will be presented for application to NDT images.

1.2.3. Deconvolution of Ultrasonic Data.

In the majority of ultrasonic systems the resolution of the data is dependent on the bandwidth of the transducers. A number of design methods have been utilised to improve the range resolution of the transducer, though the transducer still remains the weak link. Other approaches for increasing the range resolution of the ultrasonic data include the use of pulse compression [27] or signal processing [28] techniques.

The use of deconvolution techniques for resolution enhancement in ultrasonic applications is widespread. A wide variety of deconvolution techniques have been presented in the literature [28][29][30][31][32][33]. Many of these techniques have the disadvantage that they require detailed information about the measurement system and the test specimen.

For the development of a highly flexible imaging system this is not ideal. However, a number of blind deconvolution techniques do exist. By blind what is implied is that the parameters required for deconvolution of the ultrasonic A-scans are obtained solely from the ultrasonic data itself and the deconvolution model. These methods have the disadvantage that they often require longer processing times [34] and may not produce results that are comparable to conventional deconvolution techniques. They generally do produce an increase in the resolution of the raw data and have the advantage that they are highly flexible.

1.3. Overview of the thesis.

This thesis describes a new technique for the generation of high-resolution ultrasonic images of defects embedded within engineering structures. The imaging method has been applied to a number of representative defects contained within aluminium cylindrical specimens. The imaging technique employed involves the use of ultrasonic A-scan data to reconstruct tomographic images of the object cross-section and then fusion of a set of the images to obtain a high-resolution image of the defect under examination. The three ultrasonic tomographic imaging methods employed along with the low-level fuzzy logic image fusion method are outlined in Chapter 2. Additionally an image enhancement method that utilises the ultrasonic energy reflected by the specimen's back-wall is presented.

Chapter 3 describes the technique of Maximum Likelihood Deconvolution (MLD) and Minimum Variance Deconvolution (MVD) for processing of the A-scan data prior to reconstruction, the purpose being to improve the time resolution and Signal-to-Noise Ratio (SNR) of the ultrasonic A-scan data. In addition, the performance of Maximum Likelihood Deconvolution on experimentally obtained A-scan data is evaluated.

Chapter 4 experimentally evaluates the tomographic image reconstruction algorithms, the single bounce image enhancement method and the fuzzy logic image fusion technique all described in Chapter 2. This chapter also examines the effect on

tomographic image reconstruction of MLD / MVD filtering of the A-scan data prior to image reconstruction.

1.4. Contributions to Ultrasonic imaging for NDT applications.

The main achievements of this thesis are:

- The benefits of using a number of different but complementary tomographic image reconstruction methods for defect characterisation have been investigated.
- It has been demonstrated that high-resolution tomographic images of defects can be obtained using relatively small amounts of ultrasonic A-scan data.
- The effectiveness of using the shadow cast by a defect has been demonstrated in the reconstruction of transmission tomography defect images.
- The effectiveness of utilising fuzzy set theory in fusion of ultrasonic tomographic images for NDT applications has been demonstrated.
- The improvement in defect characterisation utilising a new fuzzy logic fusion method has been demonstrated for a number of representative defect types.
- It has been shown that defect characterisation is improved by utilising the single bounce image enhancement technique presented, this is especially of note when the amount of data for image reconstruction is severely limited.
- The resultant improvement in the image quality due to pre-processing of the ultrasonic A-scan data utilising Maximum Likelihood Deconvolution (MLD) / Minimum Variance Deconvolution (MVD) filtering has been presented.

Publications Arising.
The conference and journal publication arising from this thesis are as follows:

Generation of ultrasonic images by tomographic image fusion.

I.D. Hall, A. McNab.

NDT '98 Conf. Proc., The British Institute of Non-Destructive Testing annual Conference, 15-17 September 1998. pp 198-201.

Ultrasonic arrays for monitoring cracks in an industrial plant at high temperatures.

K.J. Kirk, A. McNab, A. Cochran, I. Hall and G. Hayward.

IEEE trans. UFFC, Vol. 46 no. 2 March 99 pp 311-320.

Improved ultrasonic image generation through tomographic image fusion.

I.D. Hall, A. McNab, G. Hayward.

Accepted for publication by the journal Ultrasonics.

Chapter 2:

TOMOGRAPHIC IMAGE RECONSTRUCTION AND FUSION.

2.1. Introduction.

Reconstruction is the process of forming 2-D or 3-D images from raw A-scan data, transforming amplitudes at times-of-flight to amplitudes at spatial positions. Images have the advantage over the raw A-scan data in that they show the spatial relationships between signals. An image can provide an understanding of the entire A-scan acquisition, whereas in the study of the A-scans it is possible to loose sight of the global picture [4].

Many methods have been used to obtain images of defects embedded within engineering type components. Some of the most common currently used methods were outlined in Chapter 1.

The relatively small number of flaws requiring detailed examination, coupled to substantial increases in available data storage and computational power, have made it possible to use a number of relatively simple tomographic reconstruction methods to produce images of defects contained within the material under examination. A fuzzy-logic fusion-technique can then be used to obtain a single defect image.

The three methods chosen were Reflection tomography, Time-of-flight Diffraction tomography and Transmission tomography. Reflection tomography is discussed in Section 2.3, Time-of-flight Diffraction tomography in Section 2.4 and Transmission tomography in Section 2.5. The low-level fuzzy logic fusion method used will be outlined in Section 2.6. The methodology of reconstructing a number of simple tomographic images was chosen over the development of a single more complex reconstruction algorithm, using all of the acquired data, in order to maintain theoretical simplicity. This leads to lower computational requirements and greater flexibility for the user, and will be further discussed in Section 2.7. The final Section of this chapter considers the problem of reconstruction from the incomplete data set and the expected effect on the image quality. Before the image reconstruction methods are described the data acquisition model for the reconstruction methods will be examined.

2.2. Fan-beam Insonification.

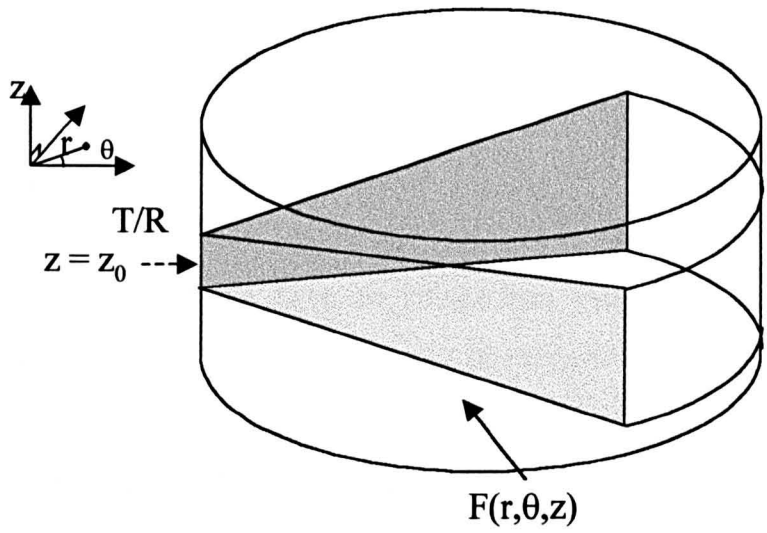


Figure 2.1. *Fan beam insonification of a cylindrical test specimen from a point source / receiver.*

The geometry for fan-beam isonification is shown in figure 2.1, where the scalar reflectivity field function is, $F(r,\theta,z)$, and a two-dimensional cross section at $z = z_0$ is to be imaged. The sound beam is fan-shaped to diverge cylindrically and cover the two dimensional slice of the object of interest, but is collimated in the z direction perpendicular to the slice. Ideally, the thickness of the beam would be infinitesimally small in the plane of interest. Obviously in practice this is not possible, so a beam of finite thickness has to be used. This means that the received signals arise from a finite-slice of the reflectivity field [20]. Let the beam profile in the z direction be b(z), then the two-dimensional reflectivity function, $f(r,\theta)$, can be defined as a slice along z_0, through the three-dimensional reflectivity field $F(r,\theta,z)$ in which the z dependence has been averaged over the vertical dimensions of the slice.

$$f(r,\theta) = \int_{-\infty}^{\infty} F(r,\theta,z)b(z-z_0)dz. \qquad (2.2\text{-}1)$$

Or

$$f(r,\theta) = F(r,\theta,z) * b(z), \quad \text{evaluated at } z = z_0 \quad (2.2\text{-}2)$$

where * denotes one-dimensional convolution.

As mentioned the beam cannot have an infinitesimal small thickness because of diffraction. Thus the averaged slice, $f(r,\theta)$, must be reconstructed, and not $F(r,\theta,z)$. With this fact in mind, if three-dimensional images are required then images of slices of the object are reconstructed and stacked to give a three-dimensional overview of the test specimen. The alternative to stacking 2-D images is to insonify the whole object and reconstruct a three-dimensional image from the data [35]. This method was not deemed worthy of consideration here due to the computational requirements of the image reconstruction.

Reflection tomography is the simplest of the tomographic reconstruction algorithms and will be the first to be discussed in Section 2.3.

2.3. Reflection Tomography.

This is the most commonly used method for reconstructing a tomographic image in NDT and is suitable for obtaining images from a wide variety of defect types [16][17][19][20][21][36]. Reflection tomography employs ultrasonic pulses to excite echoes from the boundaries of flaws contained within the cross section of the object field, illuminated in sequence by a number of ultrasonic beams. The beams may come from a variety of sources, a single transducer (Pulse-Echo), or a pair of transducers (Pitch-Catch), depending on the user requirements. A-scan data is then collected and used to reconstruct a pseudo image of the cross sectional reflectivity function.

The most widely used reflection imaging method is the B-scan. This uses parallel beams to insonify the test specimen. However, this data acquisition set-up results in images with poorer transverse resolution than range resolution. This can be overcome

by averaging a number of B-scan images and then performing some digital processing on the resultant image. The disadvantage of this method is the prohibitive amounts of data to be processed and long scanning times are required to obtain the B-scan data [20]. The type of imaging considered here uses a fan-beam insonification of the object, the received signals being processed and an ultrasonic reflectivity image reconstructed.

The Pulse-Echo Reflection method utilised here will be outlined in Section 2.3.1 and the Pitch-Catch Reflection method will be discussed in Section 2.3.2.

2.3.1. Pulse-Echo Reflection Tomography.

The Pulse-Echo configuration employs a single transducer as the transmitter and receiver. The image reconstruction geometry for this case is shown in Figure 2.2. The probe transmits a broadband ultrasonic pulse into the test specimen, with the probe then receiving the backscattered echoes as a function of time. This Pulse-Echo is repeated for each transmitter location around the specimen, resulting in N A-scans, where N is the number of transmitter locations. These N A-scans are then used to reconstruct the tomographic image. Since only a single transmitter is active at a time, a given point on the A-scan corresponds to the sum of all scatterers lying on an arc centred at the transmitter. This can be interpreted as generating intersecting line integrals of the reflectivity defined over families of concentric circular arcs centred at the transmitter. When this pulse-echo process is performed for all N transmit positions, attempts can be made to reconstruct the unknown reflectivity function. Because the integration paths are circular, the image reconstruction will be straightforward to implement.

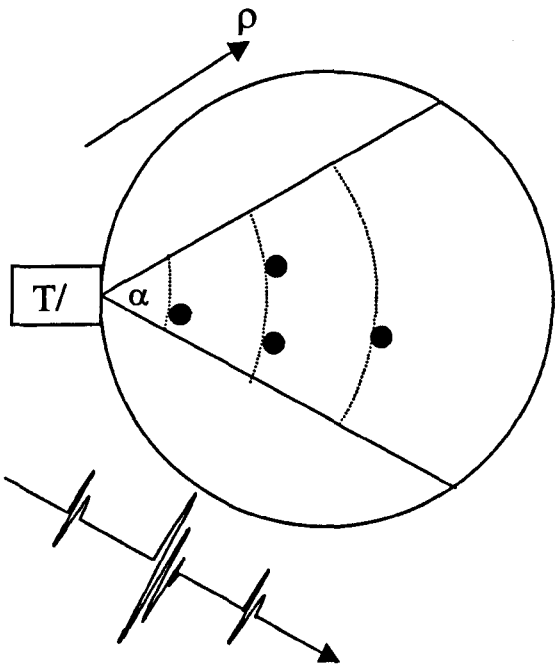


Figure 2.2. *Data acquisition geometry for Pulse-Echo Reflection tomography.*

In order to illustrate this approach to reflection imaging; consider a reflector situated a distance ρ from the transmitter. To obtain a measure of the reflectivity at this point, the A-scan is examined at time 2ρ/c, where c is the speed of sound in the specimen. This A-scan value represents the sum of echoes arising from scattering points along an arc of radius ρ centred at the transmitter. If this process is repeated for N transmitter locations (where ρ is always the distance from the transmitter position and the point under reconstruction), then the coherent summation of the A-scan values at the image point will yield the reflectivity value at the image point. Thus, each A-scan will contribute one value to each image point. To reconstruct all of the image points in this way, the amplitudes of the A-scan evaluated at time 2ρ/c are uniformly distributed or smeared back over an arc of radius ρ in the image space [20][37]. This process is repeated for each point of the N A-scans, and at each point in the image space, the results are coherently summed. The smearing back process is referred to as backprojection. The form of backprojection used here differs from the classical meaning as defined in x-ray tomography. In x-ray tomography the recorded data is back projected along straight lines connecting the source to receiver. However, here

the data is backprojected along circular paths at right angles to the propagation direction.

Here the only concern is with the mathematical problem of reconstructing reflectivity. To allow this analysis to be performed, a highly idealised image reconstruction is considered, being based on the following assumptions [18]:

- The object is weakly reflecting such that the energy contained in the propagating sound wave is much larger that the total energy backscattered over the propagation path. Also second order reflections are considered to be insignificant.
- The absorptivity of the medium is uniform, thus attenuation due to absorption can be compensated for by varying the gain of the receiver exponentially.
- The velocity of sound in the specimen under test is uniform. Specifically there are no variations in the sound velocity large enough to produce errors in the round trip delay time between the transmitter and any elementary reflecting point that are a significant fraction of the reciprocal of the signal bandwidth. This assumption ensures that the integration paths are essentially circular or elliptical, depending on the mode of operation.
- The object can be modelled as a collection of isotropic scatterers, each acting like a Huygen's-like source. Under this assumption, the reflectivity of the specimen can be modelled as a scalar function of space. Its value at a point in space is assumed independent of the direction from which the ultrasound impinges on the point.

In practice all of these assumptions will break down to some degree, therefore image degradation will occur. The mathematical analysis of reflection tomography over circular paths will now be presented.

2.3.1.1. Image Reconstruction of Reflectivity Images over Circular Backprojection Paths.

Let the reflectivity of the object in the plane of the circular test specimen be, $f(r,\theta)$, as outlined in Section 2.2. The centre of the specimen represents the origin of the polar co-ordinate system and the specimen has a radius, R. If a transmitter at angle ϕ emits a short pulse of ultrasound into the test specimen, which diverges into the specimen with an angle of divergence given by α, as shown in Figure 2.2. The reconstruction model based on these assumptions is shown in Figure 2.3.

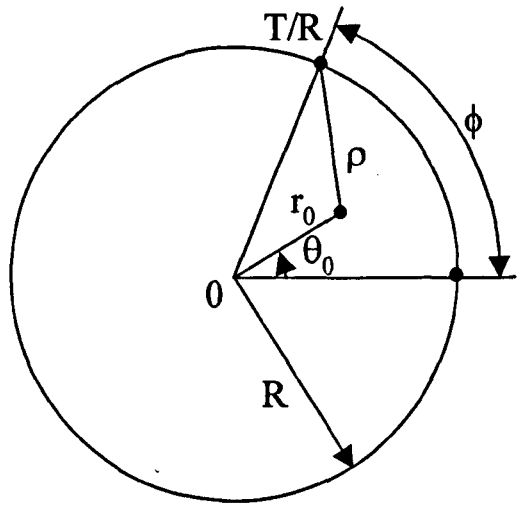


Figure 2.3. *Data acquisition model for Pulse-Echo Reflection Tomography.*

Let the omni-directional echo data (A-scan) recorded at the receiver as a function of time be denoted by $g_\phi(t)$. For clarity, $\bar{t} = tc$, where c is the speed of sound in the specimen (time delay is measured in terms of distance rather than time). In this analysis it is assumed that A-scan data is collected for N transmitter locations equally spaced around the test specimen i.e. ϕ ranges from 0 to 2π radians. The reconstruction process described will now be examined as a method of reconstructing the reflectivity function $f(r,\phi)$.

Let the reconstruction of the reflectivity function bedenoted by $\hat{f}(r,\phi)$. Then the backprojection of the A-scan signals, $g_\phi(t)$, can be expressed mathematically [18] as

$$\hat{f}(r,\phi) = \int_0^{2\pi} g_\phi(2\rho(\phi;r,\theta))d\phi \qquad (2.2\text{-}3)$$

where

$$\rho(\phi;r,\theta) = \left[R^2 + r^2 - 2Rr\cos(\theta-\phi)\right]^{1/2} \qquad (2.2\text{-}4)$$

is the distance between the transmitter/receiver at (R,ϕ) and the image point at (r,θ). To reconstruct the reflectivity at a point (r,θ), the wavefront recorded at position ϕ around the specimen circumference is evaluated at time $2\rho(\phi;r,\theta)/c$ (the round trip delay time). This process being repeated for all transmitter positions and for a given (r,θ), with the results summed to obtain $\hat{f}(r,\theta)$.

2.3.2.2. Point Spread Function.

The reconstruction of a point target in the image space is referred to as the point spread function (PSF) of the imaging system. The PSF can be regarded as the two-dimensional analogue of the one-dimensional impulse response of a linear system. The point spread function is usually characterised by a shape lobe of non-zero width at the position corresponding to the point target position, and sidelobes that trail off to some distance from the main lobe. The PSF of an imaging system is a useful measure of image reconstruction performance because it provides information about the inherent resolving capabilities of the system. It also provides an insight into the problem of artefact formation and Signal-to-Noise Ratio (SNR).

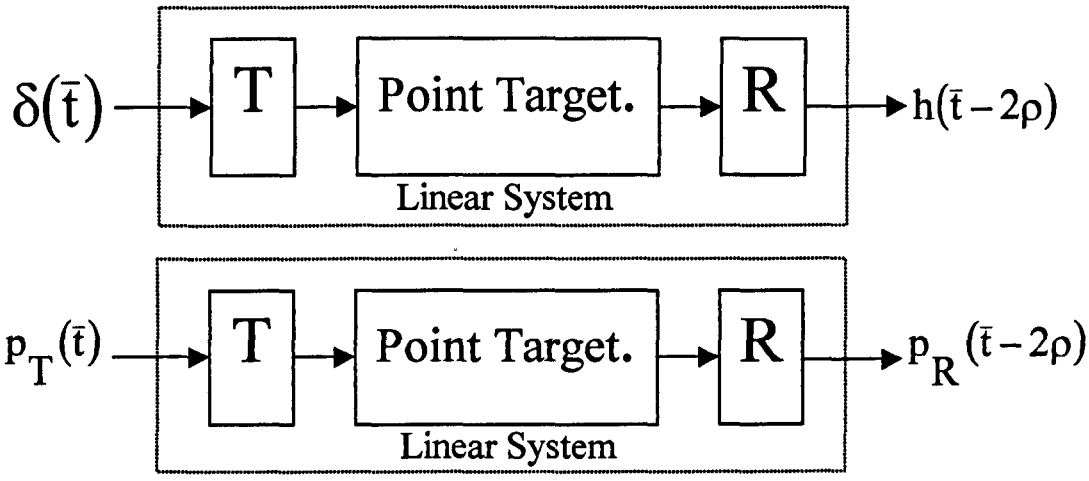


Figure 2.4. *Schematic of the convolutional relationship for the linear system; encompassing the transmitter, point target and the receiver.*

The point spread function associated with the backprojection process will now be derived [18]. Initially, the shape of the received pulse reflected from a point target will be defined in terms of the temporal impulse response of a linear system, which consists of the transmitter/receiver, the point target and the intervening medium. Let the impulse response of this linear system be denoted by $h(\bar{t})$, which is illustrated in Figure 2.4. If a voltage impulse $\delta(\bar{t})$ is applied to the transmit transducer, then in response to the resultant echo reflected back from the target at a distance ρ from the transmitter, the received A-scan will be $h(\bar{t}-2\rho)$. Usually the shape of the impulse response, $h(\bar{t})$, is governed by the characteristics of the transducer and the propagating medium.

However, if the highly idealised system based on the stated assumptions is considered, it is in theory possible to design a system where the functional form of $h(\bar{t})$ remains essentially independent of target position. In particular, it is independent of target azimuth due to the omni-directionality of the source, and independent of range through the employment of time-gain compensation at the receiver to account for echo attenuation due to beam-spreading and absorption. Thus, it will be assumed that

$h(\bar{t})$ can be made invariant with respect to the target position, with the amplitude of $h(\bar{t})$ being proportional to the target reflectivity, which is the parameter of interest.

Under these conditions, the system is both linear and time invariant. Therefore, for an arbitrary transmitted pulse $p_T(\bar{t})$, the received pulse is given by $p_R(\bar{t}-2\rho)$

$$p_R(\bar{t}-2\rho) = p_T(\bar{t}) * h(\bar{t}-2\rho) \qquad (2.2\text{-}5)$$

where * denotes convolution. The next stage in the derivation of the PSF is to reconstruct the image of a point reflector of unit reflectivity at (r_0,θ_0). From this equation an expression for the point-spread-function for backprojection along circular arcs will be derived.

As described earlier, the system generates A-scan data for N transmit locations evenly distributed around the test specimen. If the element in the image space at (R,ϕ) is excited by a transmit pulse $p_T(\bar{t})$, then using equation (2.2-5) the signal received at this element is

$$g_\phi(\bar{t}) = p_R[\bar{t}-2\rho(\phi;r_0,\theta_0)] \qquad (2.2\text{-}6)$$

where $2\rho(\phi;r_0,\theta_0)$ is the round trip distance between the target at (r_0,θ_0) and the receiver at (R,ϕ), and $\bar{t}=0$ is the time the pulse is transmitted. Putting equation (2.2-6) into equation (2.2-3) gives

$$\hat{f}(r,\theta) = \int_0^{2\pi} p_R[2\rho(\phi;r,\theta)-2\rho(\phi;r_0,\theta_0)]d\phi \qquad (2.2\text{-}7)$$

The next step is to evaluate a point spread function, $\hat{f}(r,\theta)$, for a point target at an arbitrary location, i.e. (r,θ). Using equation (2.2-4), the argument of $p_R(\bullet)$ in equation (2.2-7) can be written as

$$2\rho(\phi;r,\theta)-2\rho(\phi;r_0,\theta_0)=2\left[R^2+r^2-2R\cos(\theta-\phi)\right]^{1/2} - 2\left[R^2+r_0^2-2R\cos(\theta_0-\phi)\right]^{1/2} \qquad (2.2\text{-}8)$$

To carry out further analysis a couple of assumptions need to be made, namely that $(r/R^2)<<1$ and $(r_0/R^2)<<1$. The resultant PSF will yield an approximation to the true PSF which will bear a close resemblance to the PSF when the target point is situated near the centre of the specimen, but will generally be poorer as the target point moves towards the periphery of the specimen. It has been shown [18], that even when r and r_0 are not much less than R, the resultant Point spread function does not deviate considerably, especially around the mainlobe region, from the result when the target is near to the centre of the test specimen. As such, the analysis presented should still give useful information on system lateral resolution (mainlobe width) and other important PSF properties (sidelobe width).

Expanding equation (2.2-8) and dropping higher than second order terms in r/R and r_0/R, gives

$$2\rho(\phi;r,\theta)-2\rho(\phi;r_0,\theta_0)\cong -2r\cos(\theta-\phi)+2r\cos(\theta_0-\phi)+\frac{r^2-r_0^2}{2R} - \frac{r^2}{2R}\cos 2(\theta-\phi)+\frac{r_0^2}{2R}\cos 2(\theta_0-\phi) \qquad (2.2\text{-}9)$$

The approximations made here are similar to those made in the paraxial approximation used in optics. This can be written in the equivalent form

$$2\rho(\phi;r,\theta)-2\rho(\phi;r_0,\theta_0)=2X\cos(\phi-Y)+\gamma_1+\gamma_2\cos 2(\phi-\alpha) \qquad (2.2\text{-}10)$$

where

$$X = \left(r_0^2 + r^2 - 2r_0 r\cos(\theta - \theta_0)\right)^{1/2} \qquad (2.2\text{-}11)$$

$$\tan Y = \frac{r_0 \sin\theta_0 - r\sin\theta}{r_0\cos\theta_0 - r\cos\theta} \qquad (2.2\text{-}12)$$

$$\gamma_1 = \frac{1}{2R}\left(r^2 - r_0^2\right) \qquad (2.2\text{-}13)$$

$$\gamma_2 = \frac{1}{2R}\left(r_0^4 + r^4 - 2r^2 r_0^2\cos 2(\theta - \theta_0)\right)^{1/2} \qquad (2.2\text{-}14)$$

$$\tan\alpha = \frac{r_0^2 \sin 2\theta_0 - r^2 \sin 2\theta}{r_0^2\cos 2\theta_0 - r^2\cos 2\theta} \qquad (2.2\text{-}15)$$

The trigonometrical manipulation required to obtain equation (2.2-10) form (2.2-9) is presented in Appendix A. Note that X represents the distance between the reflector, at (r_0, θ_0), and the point under reconstruction, at (r, θ). Substituting equation (2.2-10) into equation (2.2-7), gives

$$\hat{f}(r,\theta) = \int_0^{2\pi} p_R\left[2X\cos(\phi - Y) + \gamma_1 + \gamma_2\cos 2(\phi - \alpha)\right]d\phi \qquad (2.2\text{-}16)$$

Now letting $p_R(\bar{\nu})$ denote the Fourier transform of $p_R(\bar{t})$, where $\bar{\nu} = \nu/c$ (i.e. the frequency is measured in units of reciprocal length). Then,

$$p_R(\bar{t}) = \int_{-\infty}^{\infty} p_R(\bar{\nu}) e^{j2\pi\bar{\nu}\bar{t}} d\bar{\nu} \qquad (2.2\text{-}17)$$

Substituting equation (2.2-17) into equation (2.2-16) gives

$$\hat{f}(r,\theta) = \int_0^{2\pi} d\phi \int_{-\infty}^{\infty} d\bar{\nu}\, p_R(\nu) e^{2\pi j\bar{\nu}\left[\gamma_1 + \gamma_2\cos 2(\phi - \alpha)\right]} e^{4\pi j\bar{\nu}X\cos(\phi - Y)} \qquad (2.2\text{-}18)$$

To simplify this expression the first exponential term is expanded in a power series. For convenience the result is divided into two parts.

$$\hat{f}(r,\theta) = \hat{f}_0(r,\theta) + \varepsilon(r,\theta) \qquad (2.2\text{-}19)$$

where

$$\hat{f}_0(r,\theta) = \int_0^{2\pi} d\phi \int_{-\infty}^{\infty} d\bar{\nu}\, p_R(\bar{\nu}) e^{4\pi j \bar{\nu} X \cos(\phi - Y)} \qquad (2.2\text{-}20)$$

$$\varepsilon(r,\theta) = \int_0^{2\pi} d\phi \int_{-\infty}^{\infty} d\bar{\nu}\, p_R(\bar{\nu}) \left[2\pi j \bar{\nu}\xi + \frac{1}{2}(2\pi j \bar{\nu}\xi)^2 + \cdots \right] e^{4\pi j \bar{\nu} X \cos(\phi - Y)} \qquad (2.2\text{-}21)$$

and

$$\xi = \gamma_1 + \gamma_2 \cos 2(\phi - \alpha) \qquad (2.2\text{-}22)$$

The function denoted by $\hat{f}_0(r,\theta)$ on the right hand side of equation (2.2-19) results from the zero-order term in the expansion of the first exponential in equation (2.2-18). This equation can be simplified by interchanging the order of integration with respect to ϕ and $\bar{\nu}$, and using the Bessel identity as follows [18],

$$J_0(p) = (1/2\pi) \int_0^{2\pi} d\phi\, e^{jp\cos(\phi - q)} \qquad (2.2\text{-}23)$$

The result of this operation is

$$\hat{f}(r,\theta) = 2\pi \int_{-\infty}^{\infty} d\bar{\nu}\, p_R(\bar{\nu}) J_0(4\pi\bar{\nu}X) \qquad (2.2\text{-}24)$$

The next step in deriving the point spread function is to access the contribution both of the terms in equation (2.2-19), namely $\hat{f}_0(r,\theta)$ and $\varepsilon(r,\theta)$, make to the PSF. It will be shown that the first term, $\hat{f}_0(r,\theta)$, represents a much larger contribution to the point spread function, $\hat{f}(r,\theta)$. For convenience the PSF, $\hat{f}(r,\theta)$, will be approximated by $\hat{f}_0(r,\theta)$, with $\varepsilon(r,\theta)$ regarded as an error term. For this assumption to be justified, the relationship $|\varepsilon(r,\theta)| << |\hat{f}_0(r,\theta)|$ must be examined, and the conditions in which it is valid determined, with particular attention being paid to what domain of (r,θ) it is valid over. To begin this analysis, utilise the fact from equation (2.2-21) that

$$\varepsilon(r_0,\theta_0) = 0. \qquad (2.2\text{-}25)$$

Since, $\xi = 0$ when $r = r_0$ and $\theta = \theta_0$, i.e. $\varepsilon(r,\theta)$ is identically zero at the point spread function mainlobes peak. Moreover, equations (2.2-11), (2.2-13) and (2.2-14) show that the terms X, γ_1 and γ_2 are all small compared to R when the point target at (r,θ) is close to (r_0,θ_0), the point at which the reflection occurs. Therefore, the expression for $\hat{f}_0(r,\theta)$ should provide a good approximation to the true PSF, $\hat{f}(r,\theta)$, at least in the vicinity of the reflectors actual position, (r_0,θ_0).

The relative size of ε and $\hat{f}_0$ as X becomes large when compared to the mainlobe width (i.e. when (r,θ) is far from (r_0,θ_0)) is more difficult to obtain by simple inspection of equation (2.2-21). Therefore, further examination is required to show how well, and under what conditions $\hat{f}_0(r,\theta)$ approximates the asymptotic behaviour of the sidelobes of the true PSF, $\hat{f}(r,\theta)$, relatively far away from the mainlobe. Norton *et al* has shown that for particular circumstances the magnitudes of ε and $\hat{f}_0$ do not adhere to the relationship given in equation (2.2-25). Under certain conditions the magnitude of ε can become comparable to the magnitude of $\hat{f}_0$, this usually only

happens for X very large when compared to the mainlobe width, in which case the sidelobes contribute a negligible amount to the total PSF.

All of the ultrasonic pulses, $p_R(\bar{t})$, considered here are wideband pulses. If $p_R(\bar{t})$ is considered to be very wideband then, approximately, $p_R(\bar{v}) = P = \text{constant}$. It then follows from equation (2.2-24) that

$$\hat{f}(r,\theta) = 2\pi P \int_{-\infty}^{\infty} J_0(4\pi\bar{v}X)d\bar{v} = \frac{P}{|X|} \qquad (2.2\text{-}26)$$

This shows clearly that the PSF falls off inversely with distance from the point target location.

2.3.2. Pitch-Catch Reflection Tomography.

In this case, the transmitter and receiver are separate transducers on the circumference of the test specimen, separated by an angular distance of 2β as shown in Figure 2.5. For a given round-trip delay-time, the receiver integrates the echoes arising over an elliptical path whose foci correspond to the transmitter and receiver locations. It can easily be shown that for a round trip delay time, τ, the length of the semi-major axis of the ellipse is denoted by $\tau c/2$. When a short pulse is emitted at point T and data is continuously recorded as a function of time at R, line integrals are generated over an entire family of ellipses whose foci are at T and R [18].

Assume that the separation between the transmitter and receiver remains constant at 2β and let the line that bisects this angle have an angle ϕ measured from the positive x-axis.

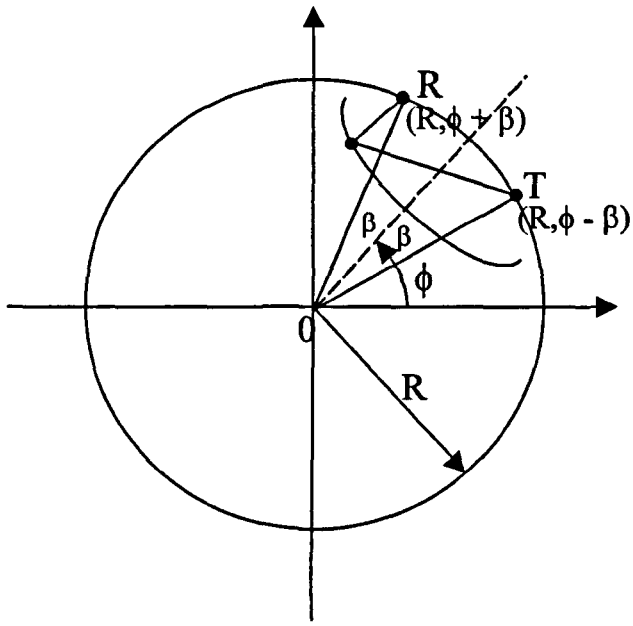


Figure 2.5. *Data acquisition model for Pitch-Catch Reflection Tomography.*

The angle ϕ represents the angle of the minor axis of the family of ellipses for the chosen transmitter and receiver locations. If ϕ is varied from 0 to 2π radians as in the pulse echo case, and β remains constant for all chosen transmit locations, then it is possible to attempt to reconstruct the unknown reflectivity function $f(r,\theta)$.

Let $g_{\phi,\beta}(\tilde{t})$ denote the received echo data (A-scan) recorded as a function of time for a transmit and receive pair separated by an angle of 2β and orientated at an angle ϕ, as shown in Figure 2.5. Obviously if $\beta = 0$, then the problem is reduced to the pulse-echo case examined in section 2.2.1. When this function is recorded and backprojected over elliptical paths for all ϕ, the resultant reflectivity function [18] is given by

$$\hat{f}(\beta;r,\theta) = \int_0^{2\pi} g_{\phi,\beta}[\rho(\phi+\beta;r,\theta)+\rho(\phi-\beta;r,\theta)]d\phi \qquad (2.2\text{-}27)$$

where

$$\rho(u;r,\theta)=\left[R^2+r^2-2Rr\cos(\theta-u)\right]^{1/2} \qquad (2.2\text{-}28)$$

Equation (2.2-27) is analogous to equation (2.2-3), the only difference being that $\rho(\phi+\beta;r,\theta)+\rho(\phi-\beta;r,\theta)$ represents the delay time from the transmitter to the separate receiver via the point target. Thus, the locus of curves in the integrand of equation (2.2-27) represents an elliptical path as opposed to a circular path in the pulse-echo case.

2.3.2.1. Alternative Reconstruction method for Pitch-Catch Reflection Tomography.

An alternative method [36] to the one described in the previous section will now be outlined. This problem can be reduced to finding the intersection points between two ellipses, the first ellipse being centred at the origin with axis $2a_0$ and $2b_0$, the second ellipse is rotated and displaced and has axis 2a and 2b. The equation for the ellipse centred at the origin is given by

$$\frac{u^2}{a_0^2}+\frac{v^2}{b_0^2}=1 \qquad (2.2\text{-}29)$$

The second (rotated and translated) ellipse is given by

$$A_0u^2+A_1v^2+A_2uv+A_3u+A_4v+A_5=0 \qquad (2.2\text{-}30)$$

where A_0, A_1, etc. are coefficients depend on θ, and the axis of the two ellipses. This expression can then be rearranged to give a quartic in u, which can be solved to give the intersection points of in u and v, which will be complex. The real values only are used, and then transformed back to the (x,y) co-ordinate for plotting. This reconstruction technique is illustrated in Figure 2.6.

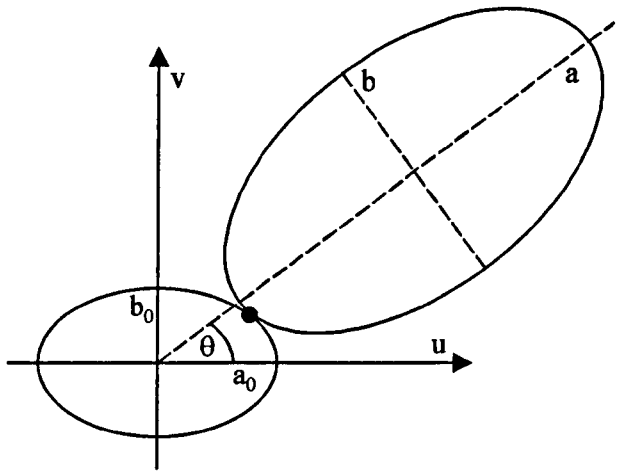


Figure 2.6. *Reduction of the reconstruction technique to finding the intersection of two ellipses.*

2.3.2.2. Point Spread Function.

Now that an equation for the reflectivity function has been obtained, the next step is to determine the PSF of the system [18]. This is achieved in much the same way as for the pulse-echo case. The point spread function for a point target situated at location (r_0, θ_0) will now be derived.

As before, the first step is to assume that the received waveform originating from a point target is $p_R(\bar{t})$. Then for a point target at (r_0, θ_0), it can shown that

$$g_{\phi,\beta}(\bar{t}) = p_R[\bar{t} - \rho(\phi + \beta; r, \theta) - \rho(\phi - \beta; r, \theta)] \quad (2.2\text{-}31)$$

substituting equation (2.2-31) into equation (2.2-27) gives

$$\hat{f}(r,\theta) = \int_0^{2\pi} p_R[\Delta_{\phi,\beta}(r,\theta; r_o, \theta_0)]d\phi \quad (2.2\text{-}32)$$

where

$$\Delta_{\phi,\beta}(r,\theta;r_0,\theta_0) \equiv \rho(\phi+\beta;r,\theta)+\rho(\phi-\beta;r,\theta)- \rho(\phi+\beta;r_0,\theta_0)-\rho(\phi-\beta;r_0,\theta_0) \qquad (2.2\text{-}33)$$

As before, assuming that $(r/R)^2 << 1$ and $(r_0/R)^2 << 1$, then equation (2.2-28) can be expanded to the second order in r/R and r_0/R, giving

$$\rho(u;r,\theta) \cong R - r\cos(\theta-u)+\frac{r^2}{4R}-\frac{r^2}{4R}\cos 2(\theta-u). \qquad (2.2\text{-}34)$$

Therefore,

$$\begin{aligned}\Delta_{\phi,\beta}(r,\theta;r_0,\theta_0) \cong{}& \frac{1}{2R}\left[r^2-r_0^2\right]-r\cos(\theta-\phi-\beta)-r\cos(\theta-\phi+\beta)+\\ & r_0\cos(\theta_0-\phi-\beta)+r_0\cos(\theta_0-\phi+\beta)+\\ & \frac{1}{4R}[-r^2\cos 2(\theta-\phi-\beta)--r^2\cos 2(\theta-\phi+\beta)+\\ & r_0^2\cos 2(\theta_0-\phi-\beta)+r_0^2\cos 2(\theta_0-\phi+\beta)]\end{aligned} \qquad (2.2\text{-}35)$$

After a large amount of trigonometric manipulation the above equation can be reduced to

$$\Delta_{\phi,\beta}(r,\theta;r_0,\theta_0) = 2X\cos(\phi-Y)+\gamma_1+\gamma_2\cos 2\beta\cos 2(\phi-\alpha) \qquad (2.2\text{-}36)$$

where X, Y, γ_1, γ_2 and α are as defined in equations (2.2-11), (2.2-12), (2.2-13), (2.2-14) and (2.2-15), respectively. The trigonometrical manipulation performed to obtain equation (2.2-36) form (2.2-35) is presented in Appendix A

This last expression is similar to equation (2.2-10) in section 2.2.1, the only difference being the presence of the two cosine terms, cosβ and cos2β, multiplying the first and third terms of the right hand side of the equation. The next stage in the derivation is to

substitute the expression for $\Delta_{\phi,\beta}$ into equation (2.2-32) and use the Fourier transform relation given in equation (2.2-17) to arrive at an expression similar to that given in equation (2.2-18) [18]. This result can be conveniently be decomposed into two parts:

$$\hat{f}(\beta;r,\theta) = \hat{f}(\beta;r,\theta) + \varepsilon(\beta;r,\theta) \qquad (2.2\text{-}37)$$

where

$$\hat{f}_0(\beta;r,\theta) = \int_{-\infty}^{\infty} d\bar{\nu}\, p_R(\bar{\nu}) J_0(4\pi\bar{\nu}X\cos\beta) \qquad (2.2\text{-}38)$$

As in the pulse-echo case $\varepsilon(\beta;r,\theta)$ is considered to be an error term analogous to equation (2.2-19). This assumption can be justified in the same way as in the proceeding section, leading to the same result. That is the error term, $\varepsilon(\beta;r,\theta)$, is small when compared to $\hat{f}_0(\beta;r,\theta)$, so $\varepsilon(\beta;r,\theta)$ is neglected and $\hat{f}(\beta;r,\theta)$ is replaced by $\hat{f}_0(\beta;r,\theta)$.

Upon comparison of $\hat{f}_0(\beta;r,\theta)$ and $\hat{f}_0(r,\theta)$ for the pulse-echo situation (equation (2.2-24)), it is evident from inspection of the two equations that the only difference is the presence of the $\cos\beta$ term in the Bessel function argument.

For a wideband excitation pulse the same assumptions as for the pulse-echo case are made, that is $p_R(\bar{t})$ is very wideband in nature. Then, $p_R(\bar{t})$ = P = constant, and equation (2.2-36) yields

$$\hat{f}(r,\theta) = 2\pi P \int_{-\infty}^{\infty} J_0(4\pi\bar{\nu}X\cos\beta)d\bar{\nu} = \frac{P}{|X\cos\beta|} \qquad (2.2\text{-}39)$$

Note that $\hat{f}_0(\beta;r,\theta)$ is identical in form to $\hat{f}_0(r,\theta)$, defined in equation (2.2-24) except for the presence of the cosβ term in the argument of the Bessel function. Upon examination of equation (2.1-39) it is evident that the minimum mainlobe width will occur at $\beta = 0$, also the mainlobe width is proportional to $(\cos\beta)^{-1}$.

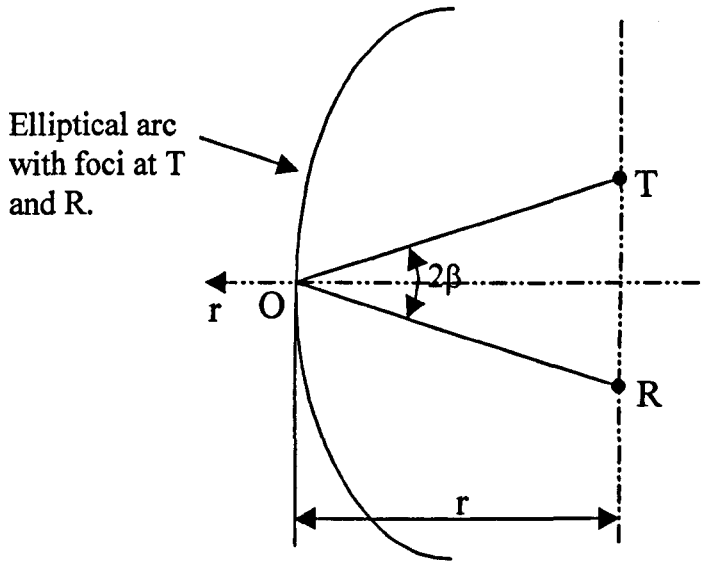


Figure 2.7. *The rate of change of the distance r with respect to the time delay over path TOR increases as* $(\cos\beta)^{-1}$.

This can be more easily understood by examination of Figure 2.7. For simplicity, consider an elliptical path intersecting a point at the array centre (r_0,θ_0). The point is to consider is the resolving power of the transmitter and receiver in the r direction, which is as defined in the figure. From the geometry of the figure, it is not difficult to derive that the rate of change of the round trip delay (over the path TOR) with respect to r decreases as cosβ. This implies that for a fixed time delay resolution, if a reflecting target moves an increment of Δr along the r axis, the minimum possible Δr that can be resolved increases as $(\cos\beta)^{-1}$. Thus, the maximum resolving capability of the imaging system is obtained when $\beta = 0$. When the recorded signals from the N transmit locations are back projected over the paths distributed symmetrically around 2π radians, the same $(\cos\beta)^{-1}$ dependence reveals itself in the equation for the PSF (equation (2.2-36)). These considerations suggest an upper bound on the value of β.

If this upper limit is chosen to be $\pi/4$ radians (i.e. the receiver would never be greater that one quarter of the circumference away from the transmitter), then mainlobe broadening would not exceed a factor of $\sqrt{2}$.

2.4. Time-of-flight Diffraction Tomography.

Time-of-flight Diffraction tomographic [22] techniques provide information on the position, shape and size of defects by the interpretation of ultrasonic signals received with either a single or twin probe system. This reconstruction method is more specialised than the simpler reflection tomography methods and may therefore give additional information about the defect structure. Diffraction tomography uses the diffracted signals from defect boundaries to reconstruct flaw images. The data is collected using transducers that are sufficiently small as to approximate point sources. The objective is to transmit ultrasound into the specimen from a single transmit location and to receive the signals diffracted from the defect at a number of points along the receiver aperture.

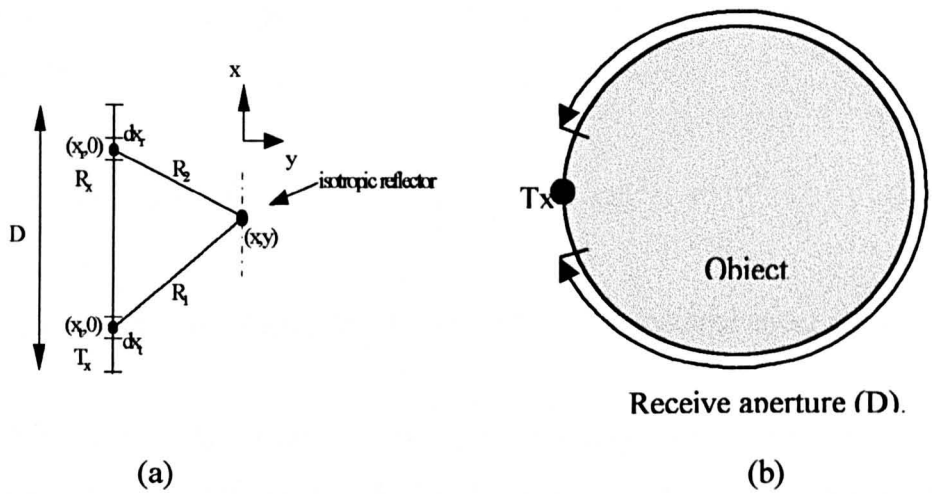


Figure 2.8. *a) Reconstruction model for time-of-flight diffraction tomography for a linear aperture D. b) Representation of the reconstruction model for a cylindrical geometry.*

The image is reconstructed in the spatial domain by a coherent summation of elliptic functions whose parameters depend on the transmitter and receiver locations and the time-of-flight of the diffracted signals. Figure 2.8(a) shows a typical data acquisition

model for a linear aperture. This algorithm is easily adaptable to any sampling geometry; the data acquisition system used on a cylindrical object for a single transducer position is shown in Figure 2.8(b). For each transmitter position on the aperture D, the scattered signals are detected with the receiver scanning the aperture.

For example, let $(x_t,0)$ and $(x_r,0)$ be the positions of the transmitter and receiver respectively, both of which are assumed to approximate point sources. Since the system is based on time-of-flight diffraction, the broadband signal transmitted by an elementary source of width dx_t is in the form $\mathrm{Re}\left[s(t)e^{j\omega t}\right]$, where s(t) is the pulse envelope and ω is the angular frequency. The signal, du, scattered from an isotropic point reflector placed at (x,y) and received by an elementary receiver width dx_r is [22]

$$du(t,x_t,x_r) = \mathrm{Re}[s(t - d(x_r,x_t,x,y)/c) \times e^{j\omega(t-d(x_r,x_t,x,y)/c)}]dx_r dx_t \qquad (2.3\text{-}1)$$

where $d(x_r,x_t,x,y)$ is the path length from transmitter to receiver via the scatterer and is given by

$$d(x_r,x_t,x,y) = ((x_r - x)^2 + y^2)^{1/2} + ((x_t - x)^2 + y^2)^{1/2} \qquad (2.3\text{-}2)$$

and c is the ultrasonic propagation velocity within the material under examination.

Note that the time factor $d(x_r,x_t,x,y)/c$ defines a time dependent function, that in the space domain is an ellipse with foci at $(x_r,0)$ and $(x_t,0)$, passing through the point reflector (x,y). The assumptions are that the material is non-attenuating and that all scatterers are in the far field. Hence the amplitude decreases as 1/(R1R2) [22], where R_1 and R_2 are as shown in Figure 2.7(a) and correspond to the separation between the defect and the transmitter and receiver, respectively. Image formation consists of processing all received signals for all chosen transmitter positions in the aperture D.

2.4.1. Point Spread Function.

When the data is recorded in this multiple receiver configuration, the generated line integrals of reflectivity are along elliptical paths exhibiting a variety of shapes and sizes. It is useful for the derivation of an expression for the PSF to interpret the reconstructed image as the result of coherently summing many Pitch-Catch Reflection tomography images obtained for differing values of β. Therefore, the PSF of the imaging method can be computed by effectively averaging $\hat{f}(\beta;r,\theta)$, given by equation (2.2-36), over the chosen range of β. The resulting PSF can then be expressed as,

$$\hat{f}(r,\theta) = \int_{-\pi}^{\pi} w(\beta)\hat{f}(\beta;r,\theta)d\beta \qquad (2.3\text{-}3)$$

2.5. Single Bounce Image Enhancement.

To increase the reliability of imaging methods it is desirable to account for a single bounce [9] before and/or after the flaw in the reconstruction algorithms. Thus giving more information about the flaw for the same number of A-scans used presently. The indirectly reflected beams can be used to improve the defect image quality when all parts of the defect cannot be directly isonified.

This involves using reflections from the backwall of the test specimen to obtain more detail about any defects embedded in the specimen. The fact that the backwall geometry must be known is a drawback of this image enhancement method.

For simplicity and to obtain low reconstruction times the method chosen does not take account of mode conversions when the ultrasound is reflected, i.e. all of the reflected waves in the specimen are assumed to be longitudinal.

The approach that will be used to perform the backprojection after the single bounce will now be outlined for the three data acquisitions considered. Initially the method

used for Pulse-Echo Reflection tomography will be outlined, then adapted for Pitch-Catch Reflection tomography and finally for Time-of-flight Diffraction tomography.

2.5.1. Pulse-Echo Reflection Tomography.

As stated, the backwall geometry of the specimen needs to be known; if the backwall is a flat surface then backprojection of the echo delay times, corresponding to distances greater than the backwall distance, is along circular arcs. However, if the backwall has a more complex geometry (i.e. for the cylindrical test specimens considered here) then backprojection will be along more complex curves determined by the backwall geometry, where each A-scan value represents the sum of all echoes from scatterers positioned along the curve. Due to the complex shape of these curves a ray tracing approach has been chosen to achieve the backprojection.

If a pulse is transmitted into the specimen from a transmitter at T and diverges within the specimen at an angle α, as shown in Figure 2.2, then rays are cast from the transmitter for all angles within the angle of divergence, α. There is no backprojection performed along these rays until the ray is incident on the backwall, as this has already been performed. Using the backwall geometry and Snell's law (assuming no mode conversions) then it is relatively straightforward to calculate the ray path after reflection by the backwall. The first step of the backprojection is to determine the reflection paths for all rays in the fan beam. Once the reflected ray paths have been determined, then each A-scan value corresponding to a echo delay time greater than the backwall echo delay time can be backprojected. For each of the relevant A-scan values the distance along the reflected rays are determined using the round trip delay time, this is shown in figure 2.9 for a single ray at angle, θ, where the direction of propagation of the ultrasound is immaterial.

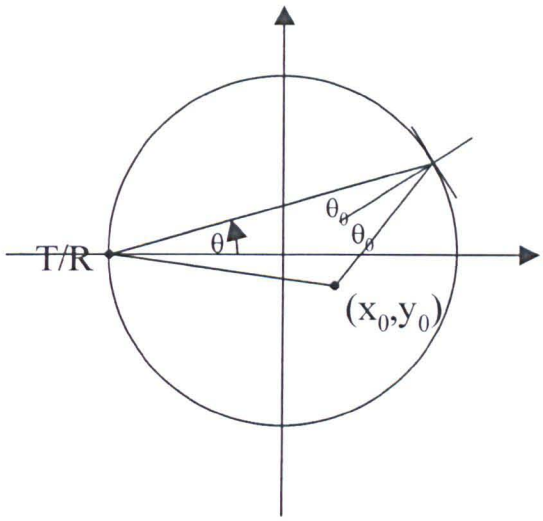


Figure 2.9. *Single bounce backprojection for a single ray at angle θ.*

In practice a finite number of rays are chosen and it is necessary to interpolate between the scatterer positions determined for each A-scan echo delay time along each of the rays to obtain a continuous backprojection curve. This will not give the exact backprojection curve, but if the number of rays chosen is sufficiently large then it will closely resemble the actual curve. This family of curves can be seen in Figure 2.10.

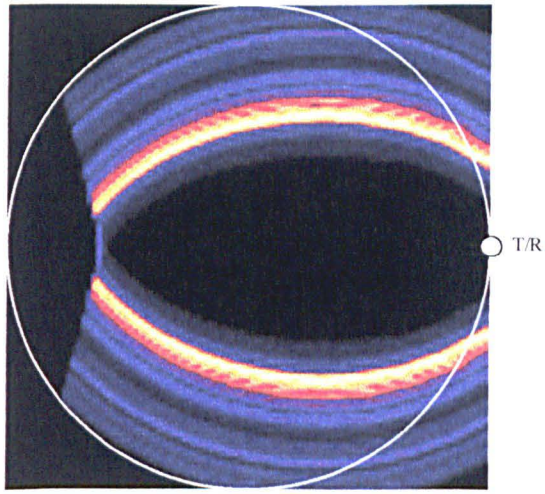


Figure 2.10. *Single Bounce Image Enhancement Backprojection curves for Pulse-Echo Reflection tomography.*

2.5.2. Pitch-Catch Reflection Tomography.

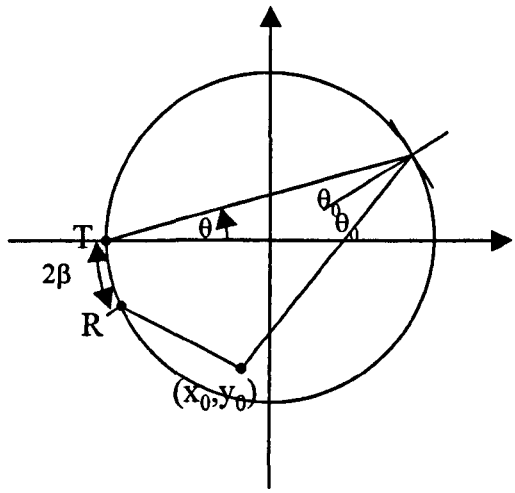


Figure 2.11. *Single bounce backprojection for separate transmitter and receiver; for a single ray at angle θ.*

The approach taken for the pitch-catch situation is analogous to the method used for the pulse-echo configuration, the only difference being that backprojection will be along elliptical paths in the case where the backwall is a flat surface. Therefore for more complex backwall geometries a different family of curves than in the pulse-echo configuration will result. The backprojection path for a single ray is analogous to the pulse-echo case, the only difference being the use of a separate transmitter and receiver with an angular separation of 2β, as shown in Figure 2.11. The family of backprojection curves is shown in Figure 2.12.

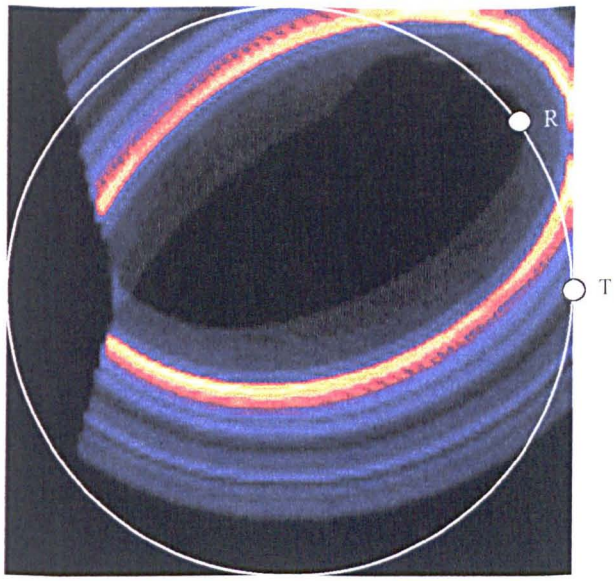


Figure 2.12. *Single bounce image enhancement backprojection curves for Pitch-Catch Reflection tomography and also Time-of-flight Diffraction tomography for a single transmit and receive pair.*

2.5.3. Time-of-flight Diffraction Tomography.

The addition of single bounce information to the time-of-flight diffraction images is analogous to the pitch catch case, the difference being that the transmitter to receiver separation, 2β, can be varied between 0 and 2π radians. This means that for each transmit location there is a family of single bounce backprojection curves for each receiver location.

2.6. Transmission Tomography.

Transmission tomography methods for non-destructive testing applications are fundamentally different from the methods used in medical applications [12][13], mainly due to the physical properties of the materials under test. In medical applications the human body contains regions of differing acoustic velocity, so a velocity profile of the body can be reconstructed from time-of-flight measurements. This is not the case for the majority of engineering components, as they are by nature made of a single material and therefore the velocity changes across the object cross section are negligible. Moreover, there is usually a large acoustic mismatch between

the defect and surrounding media so the majority of ultrasound is either reflected or diffracted by the defect. For NDT applications other methods have to be considered.

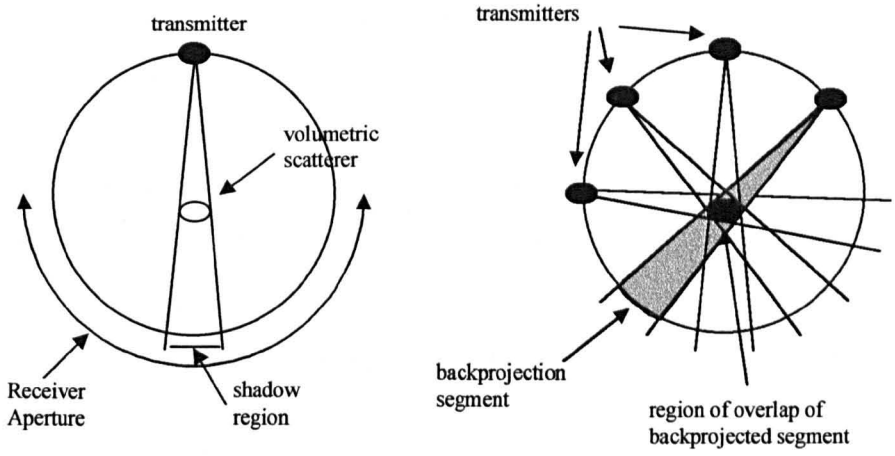


Figure 2.13. *Transmission Tomography based on the shadow cast by the defect. a) shadow cast by the defect assuming it totally reflects all incident ultrasound. b) backprojection of shadow region to reconstruct defect images.*

Transmission tomography methods for NDT applications are usually based on the shadow projected by the flaw [5]. Transmission tomography has been used in the present work to improve images reconstructed using other tomographic methods. An obscuration method [38] has been implemented based on the principle that the defect will totally reflect an incident ultrasonic beam, which does not propagate to the receiver position. Therefore a flaw will cast a shadow which depends on its size and position with respect to the transmitter, as illustrated in Figure 2.13. The interaction of the defect and the incident ultrasound for the transmission scenario will now be examined.

2.6.1. Shadow Cast by Defect.

The propagation of a wave is not only disturbed by an obstacle in generating a reflected wave, but the shadow cast by the obstacle is also present. If the obstacle is much larger that the wavelength, then diffraction effects in the shadow region will not be significant. However, if the obstacle is not much bigger than the wavelength then

diffraction phenomena will be present in the shadow region. This situation will now be examined. To compute the distributed sound field behind the reflector, the following propositions need to be considered. The shadow field will be built up of the undisturbed original field and from interference caused by a disturbing wave travelling from the rear of the reflector.

In the case of a flat, circular, thin disc reflector, perpendicular to the wave fronts, the characteristics of the interfering wave are easily identified [5] because the overall excitation of the rear wall must be zero. In this area the interference and the undisturbed wave must cancel each other out entirely.

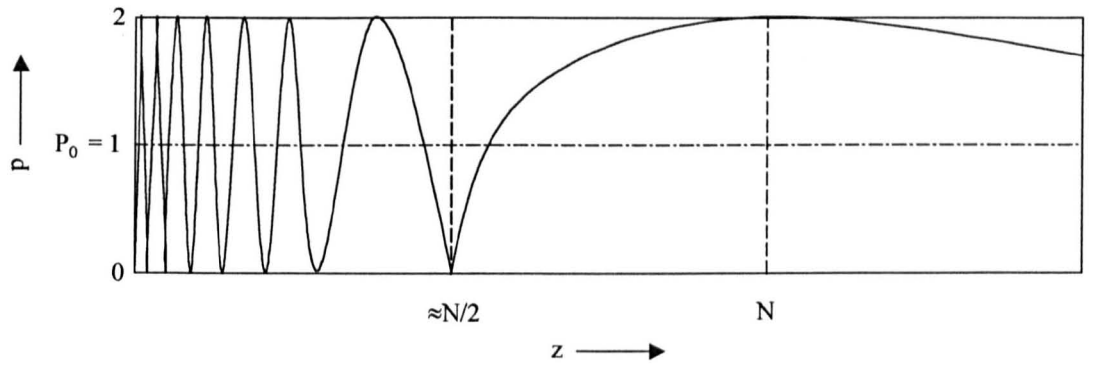


Figure 2.14. *An example of the sound pressure along the axis of a ideal piston oscillator, where N is the near field distance, z the distance along the axis of the disc, if the disc is much larger than the wavelength then* $N \approx D^2/4\lambda$, *p is the sound pressure with* p_0 *being the initial sound pressure of the transmitter.*

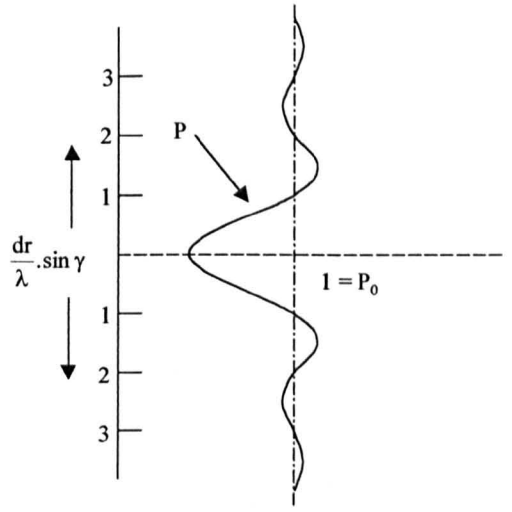


Figure 2.15. *Sound pressure directivity behind a circular reflector placed at a distance of six near field lengths from the transmitter, where γ is the angle from the disc axis.*

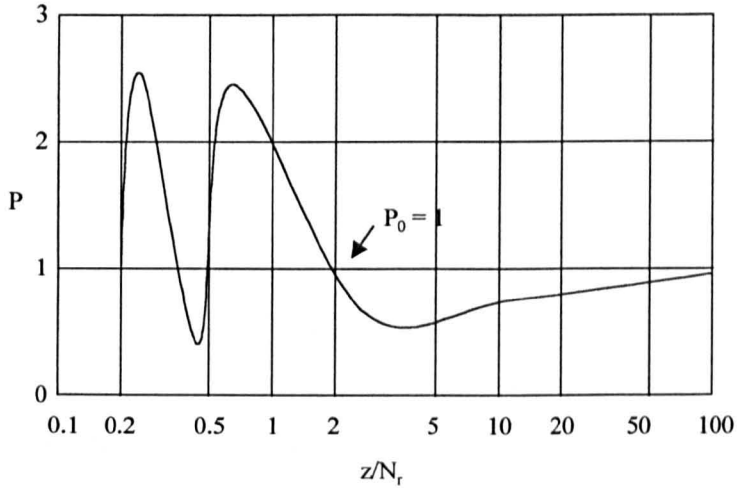


Figure 2.16. *The sound pressure amplitude in the shadow field along the axis of a circular disc. Where N_r is the near field distance of the reflector (reflector acting as a secondary source).*

Thus the interfering wave must have the amplitude of the primary wave over the whole reflector area, but of opposite phase, so that it is in fact the well known piston oscillator wave, shown in figure 2.14 travelling in the same direction and coaxial with

the primary wave. As examples consider Figure 2.15 and Figure 2.16. These show the sound pressure directivity and the pressure, along the axis, respectively. They are computed on the assumption that the primary field is a plane wave or in the distant far field of the transmitter [5]. An example of the sound pressure in the far field with no reflector in between the transmitter and receiver can be seen in Figure 2.17. These three figures show the pressure as if measured by a small source. If larger receivers are used then these pressure values will be averaged out.

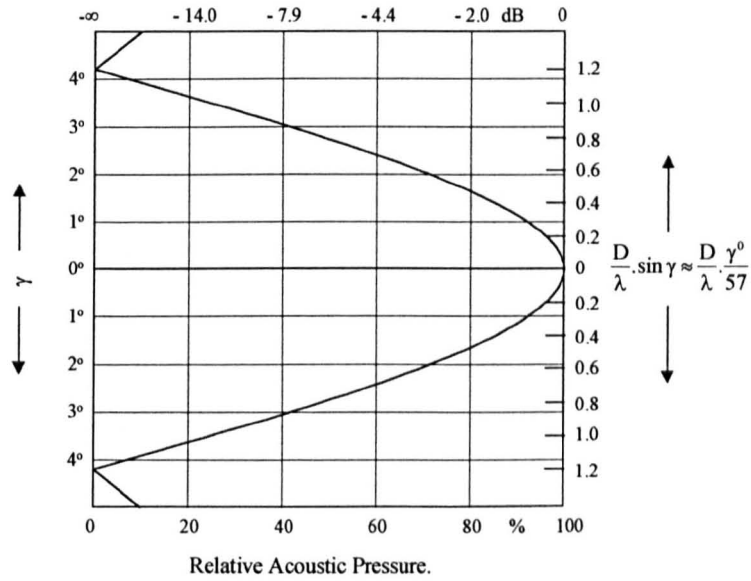


Figure 2.17. *Relative far-field sound pressure plotted against angle γ for $D/\lambda = 16$. Where D is the oscillator diameter and λ is the wavelength of sound in the material under test and γ is the angle from the disc's axis.*

This represents the case where the reflector is perpendicular to the axis of the ultrasonic beam. In practice this will not always be the case so oblique angles of incidence need to be considered. At moderately oblique defect orientations, the shadow wave will take roughly the same form as that of a disc reflector at right angles, if edge effects are neglected [5]. In this situation, upon backprojection the disc reflector will appear to the receiver as a disc at right angles to the incident beam, but will appear to have a smaller diameter. If a number of shadow regions are backprojected from differing angles then the true orientation and size of the reflector

should become evident due to the nature of the summation of the backprojection segments.

This information about the interaction of ultrasound and ideal defects can be utilised to design filter functions for the backprojection enabling the resolving capability of the image reconstruction to be increased. The image reconstruction methods will now be examined.

2.6.2. Transmission Tomography Image Reconstruction.

In the simplest form a binary system is used typically, so that if a signal is received then a value of 1 is assigned to the receiver [38]. If no signal is detected then a zero is assigned to the receiver. In practice, a signal is always detected at the receiver, so it is necessary to set a threshold to identify any shadow regions. For the defects examined in this thesis, a threshold value of between 10 % and 50 % of the expected receiver amplitude is utilised. Thresholds within this region were chosen in order to minimise the chance of incorrectly identifying shadow regions due to variations in the transducer coupling and any effects on the received signal caused by diffraction effects around any of the defects.

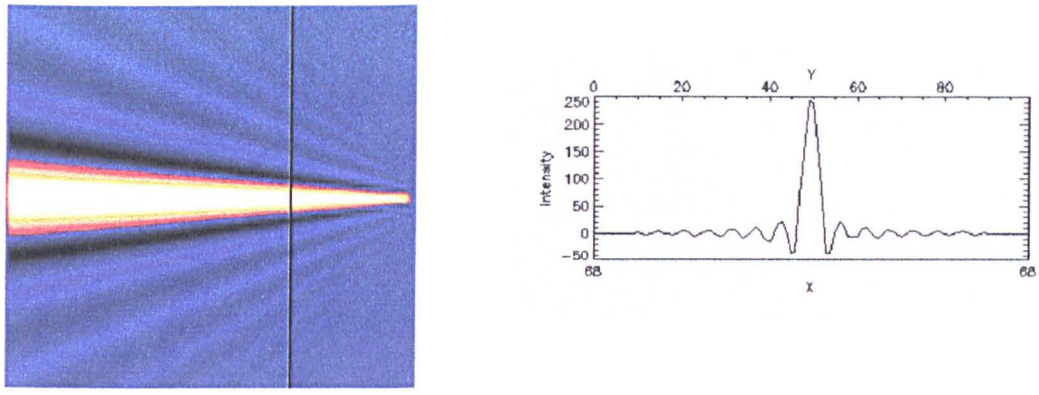
Once any regions of zeros have been identified, they can then be backprojected, and an image of the flaw reconstructed. The back projection consists of using segments, as shown in Figure 2.13(a), with the receiver permitted to be at any location around the specimen under test. If a number of transmitter positions are used, then the backprojected sectors will overlap in the region of the flaw and superimpose to give a darker region, which will correspond to the flaw.

A straightforward example is shown in Figure 2.13(b). The minimum size of the flaw that can be detected depends largely on the beam width of the transducer used; it is not possible to detect small defects when utilising transducers with a relatively wide beamwidth. However, it is possible to detect larger defects using transducers with a small beamwidth.

In order to improve the quality of the reconstructed images, the information about the interaction of the sound and ideal disc reflector, as outlined in Section 2.5.1, can be combined with the binary backprojection segments to generate a filtered backprojection reconstruction algorithm. This technique for improvement of the transmission images will now be outlined.

2.6.3. Filtered Backprojection.

The use of filter functions is common in transmission tomography [12][13] in order to improve the resolution of the reconstructed images. The filter function used here will be based on the sound field distribution behind an ideal disc reflector at right angles to the axis of the ultrasonic beam as shown in Figure 2.15. This simple model for the filter design has been favoured in order to reduce the theoretical complexity and therefore maintain low image reconstruction times. The backprojection method used is similar to the non-filtered backprojection in a number of respects; the shadow regions are identified in the same way. However, the nature of the backprojection segments is fundamentally different. The backprojection segments involve combination of the linear backprojection segment and a window function based on the actual shadow cast by an ideal disc type defect as described in Section 2.6.1. A single filtered backprojection segment can be seen in Figure 2.18.



(a) (b)

Figure 2.18. *Filtered backprojection segment based on the sound pressure directivity behind a circular reflector as in Figure 2.15. Figure (a) shows the backprojection segment in the image plane with (b) showing the amplitude of the backprojection segment along the line marked on the image in (a).*

This method of backprojection should yield better quality images due to the nature of the segment outside the region that is definitely identified as the shadow region. The phase of the filter function will be negative in certain areas and will therefore lead to a phase cancellation within the image in close proximity to the defect, thus improving the image of the object boundary. The improvement in image quality using this backprojection method will be experimentally validated in Chapter 4.

2.7. Tomographic Image Fusion.

The fusion of NDT data is a fast growing set of signal processing techniques which represent the possibility of reducing signal uncertainty and improving the overall performance of NDT systems. The majority of work in NDT data fusion has concentrated on fusion of signals from different NDT inspection techniques i.e. multi-sensor data fusion [24][25][39]. The approach chosen here is to look at the fusion of images rather than the signals used to reconstruct the image. Another difference is that all of the images considered originate from ultrasonic inspections of the object under test.

A composite image can be generated from the set of reconstructed flaw images. This has been achieved using a low level fuzzy logic pixel fusion technique [26][39]. Fuzzy logic represents a powerful framework for data processing, as it allows processing commands to be expressed as a set of rules which bear a resemblance to the human decision making process [40]. Fuzzy sets are generally defined by characteristic membership functions. These functions give the level of membership of a particular object to the fuzzy set. This set being used to determine the level of membership of the pixel to the final flaw image.

If the pixel is definitely a member of the flaw image, then the weighting associated with the pixel will be 1. If the pixel is definitely not contained in the flaw image then the weighting will be 0 and for all other pixels, which could be contained in the final image, a weighting of between 0 and 1 is applied depending on the membership function. If the pixel amplitude is greater or less than the chosen threshold, then it may or may not be part of the final image. If the pixel amplitude is close to the threshold, then the probability of the pixel being contained in the flaw image is higher than if the amplitude is small when compared to the threshold. This is reflected in the pixel weighting, as seen in the membership function.

A number of membership functions were evaluated with the final choices being shown in Figure 2.19, 2.20 and 2.21. All functions examined took a similar form to the ones shown in Figure 2.19, but with the number of pixels receiving a pixel weighting of one being increased (i.e. a number of pixel amplitudes have a weighting of one rather than a single pixel amplitude in figure 2.19). This has the effect of increasing the number of pixels that are definitely contained within the final defect image, which can be desirable in many of the fusion problems considered here. This method of image fusion has the advantages that it is quick and simple to implement.

The threshold value, T_c, for determining which pixels represent the defect and those that do not was determined using the one of three of the thresholding methods, that will be outlined in Section 2.7.1.

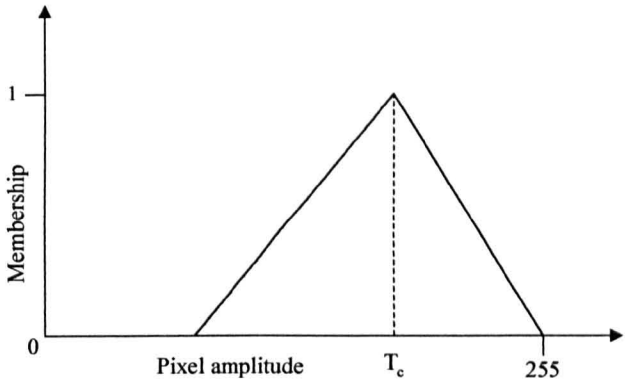


Figure 2.19. *Fuzzy logic fusion membership function.*

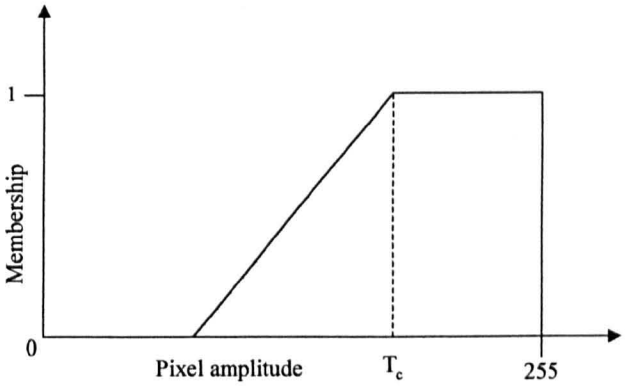


Figure 2.20. *Second Fuzzy logic fusion membership function.*

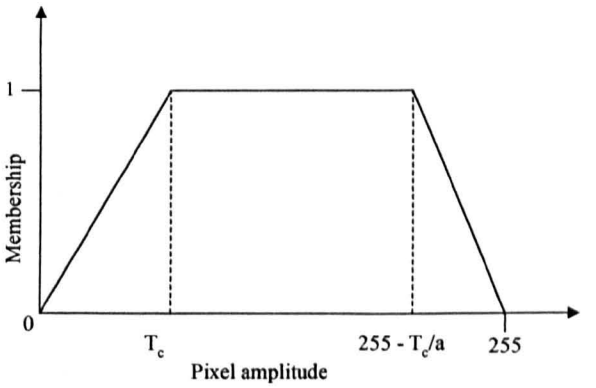


Figure 2.21. *Third fuzzy logic fusion function, where a is a user defined ramp variable between 2 and 5.*

A fuzzy logic function is determined for each image and then the weighted pixel amplitudes are summed to give a composite image of the flaw. For each composite image generated only a single type of fuzzy logic membership function is utilised. Another factor which has a large influence on the final flaw image is the pixel amplitude at which the pixel weighting is zero. For the reflection and diffraction images, the lower cut-off point is typically chosen between $0.25T_c$ and $0.5T_c$. For the transmission image, the lower cut-off is typically chosen between $0.5T_c$ and $0.75T_c$. Generally this increase in the cut-off value is necessary to remove the large quantity of information contained within the image that is not related to the flaw under examination.

The choice of image fusion function is dependent on the nature of the defect under examination. The first fuzzy logic membership is generally used to generate composite images of point reflectors and uses the Ostu thresholding method. The second of the fuzzy logic fusion functions is usually used for defect images of inclusion type defects and for best results utilises either the Ostu or Kapur, Sahoo and Wong method. The final image-fusion membership-function is used most frequently on planar type defects and gives the best results with the Kapur, Sahoo and Wong or the Pun method. These three thresholding methods used to determine T_c will now be outlined.

2.7.1. Thresholding Methods.

Three thresholding methods have been used to obtain the optimal threshold value, T_c, for the fuzzy logic image fusion functions. These are the Ostu method, Pun method and Kapur, Sahoo and Wong method [41]. All of these thresholding operations are global point dependent techniques, that is a single optimal threshold value is used on the whole image with the threshold value being determined from only the pixel amplitudes. The underlying theory of these methods will now be outlined.

For the threshold determination the images are assumed to be grey scale images. Let G be a set of positive integers representing grey levels, where 0 is assumed to be the

darkest and (max-1) the lightest pixel amplitude, t denotes the threshold value of the image and T_c to represent the optimal image threshold.

The underlying theory of the three thresholding methods will now be outlined. Let N be the set of natural numbers, (x,y) define the spatial co-ordinates of the image i.e. the pixels of the image, and Let G = {0,1,2,....,max-1} be a set of positive integers representing the pixel amplitudes. Following from this an image function can be defined as the mapping $f : N \times N \rightarrow G$. The amplitude of a pixel at position (x,y) is denoted as f(x,y). Let $t \in G$ be a threshold value and the optimal threshold determined using some predefined criteria be $T_c \in G$.

A number of elements are common to all three methods, these will now be outlined. Let the number of pixels with greylevel i be n_i. Then the total number of pixels in the chosen image is given by

$$n = \sum_{i=0}^{max-1} n_i \qquad (2.7\text{-}1)$$

From this the probability of occurrence of grey level i is defined as

$$p_i = \frac{n_i}{n} \qquad (2.7\text{-}2)$$

2.7.1.1. Ostu Thresholding method.

This global thresholding technique is based on discriminant analysis [41]. The thresholding operation is regarded as the partitioning of the pixels into two classes C_o and C_1 (e.g. object and background) at grey level t. That is, C_o = {0,1,....,t) and C_1 = {t + 1,t + 2,....,max - 1}. Let σ_W^2, σ_B^2 and σ_T^2 be the within-class, between class and total variance, respectively. An optimal threshold can be determined by minimising one of the three possible criterion functions with respect to t, the probabilities used to compute these functions is given in Equation (2.7-2).

$$\psi = \frac{\sigma_B^2}{\sigma_W^2},\ \eta = \frac{\sigma_B^2}{\sigma_T^2},\ \kappa = \frac{\sigma_T^2}{\sigma_W^2} \qquad (2.7\text{-}3)$$

Of the above three criterion functions, η is the simplest. Thus, the optimal threshold T_c is

$$T_c = \operatorname{Arg}\operatorname*{Min}_{t \in G} \eta, \qquad (2.7\text{-}4)$$

where

$$\sigma_T^2 = \sum_{i=0}^{l-1}(i-\mu_T)^2 p_i, \quad \mu_T = \sum_{i=0}^{l-1} i p_i,$$

$$\sigma_B^2 = \omega_0\omega_1(\mu_0\mu_1)^2, \quad \omega_0 = \sum_{i=0}^{t} p_i, \quad \omega_1 = 1-\omega_0 \qquad (2.7\text{-}5)$$

$$\mu_1 = \frac{\mu_T - \mu_t}{1-\omega_0}, \quad \mu_0 = \frac{\mu_t}{\omega_0}, \quad \mu_t = \sum_{i=0}^{t} i p_i$$

2.7.1.2. Pun Thresholding Method.

This thresholding method is known as an entropic method, that is it is based on the application of information theory to the image histogram to obtain the optimal image threshold [41], where the histogram is considered an max-symbol source.

Let t represent the optimal threshold value and define two posteriori entropies, given by

$$H_b' = -\sum_{i=0}^{t} p_i \log_e p_i \quad (2.7\text{-}6)$$

$$H_b' = -\sum_{i=t+1}^{max-1} p_i \log_e p_i \quad (2.7\text{-}7)$$

where $H_b^{'}$ and $H_w^{'}$ are measures of the a posteriori information associated with the black and white pixels after thresholding, respectively. The probabilities used in Equation (2.7-6) and (2.7-7) as are defined in Equation (2.7-2). If the a-priori entropy of the grey level histogram is known, Pun proposed an algorithm to identify the optimal threshold by maximising the upper bound of the aposteriori entropy,

$$H^{'} = H_b^{'} + H_w^{'} \qquad (2.7\text{-}8)$$

Pun has shown that maximising H' is equivalent to maximising the evaluation function, given by

$$f(t) = \frac{H_t}{H_T}\frac{\ln P_t}{\ln \max\{p_0,.....,p_t\}} + \left[1 - \frac{H_t}{H_T}\right]\frac{\ln(1-P_t)}{\ln \max\{p_{t+1},.....,p_{max-1}\}} \qquad (2.7\text{-}9)$$

with respect to t, where

$$H_t = -\sum_{i=0}^{t} p_i \ln p_i , \qquad (2.7\text{-}10)$$

$$H_T = -\sum_{i=0}^{max-1} p_i \ln p_i , \qquad (2.7\text{-}11)$$

$$P_t = \sum_{i=0}^{t} p_i . \qquad (2.7\text{-}12)$$

2.7.1.3. Kapur, Sahoo, and Wong Thresholding Method.

This method is also entropic in nature, and is based on the derivation of two probability distributions (i.e. object distribution and background distributions) from the original grey level distribution of the image [41] as follows:

$$\frac{p_0}{P_t}, \frac{p_1}{P_t}, ..., \frac{p_t}{P_t} \qquad (2.7\text{-}13)$$

and

$$\frac{p_{t+1}}{1-P_t}, \frac{p_{t+2}}{1-P_t}, \ldots\ldots, \frac{p_{max-1}}{1-P_t} \qquad (2.7\text{-}14)$$

where t is the threshold value, p_0, p_1, ..., p_{max-1} are probabilities computed using Equation (2.7-2) and

$$P_t = \sum_{i=0}^{t} p_i \qquad (2.7\text{-}15)$$

Define

$$H_b(t) = -\sum_{i=0}^{t} \frac{p_i}{P_t} \ln\left(\frac{p_i}{P_t}\right) \qquad (2.7\text{-}16)$$

and

$$H_w(t) = -\sum_{i=0}^{t} \frac{p_i}{1-P_t} \ln\left(\frac{p_i}{1-P_t}\right) \qquad (2.7\text{-}17)$$

Then the optimal threshold T_c is defined as the grey level that maximises $H_b(t) + H_w(t)$, i.e.

$$T_c = \operatorname{Arg}\operatorname{Max}_{t \in G} \{H_b(t) + H_w(t)\} \qquad (2.7\text{-}18)$$

2.8. Computational Requirements of the Tomographic Imaging System.

The computational requirements of the imaging system components will now be considered. In the introduction to this chapter it was stated that the idea of choosing a number of tomographic reconstruction methods and then fusing the resultant images was favoured over the development of a single reconstruction algorithm to maintain a

low theoretical complexity. For this reason, the computation requirements of the reconstruction methods and the image fusion tool are important system parameters and will now be discussed in more detail.

These methods differ in their computational complexity and the quality of the images that can be reconstructed. This is evident in the differing point-spread function for the Reflection and Time-of-flight Diffraction tomography methods. The Pulse-Echo Reflection and the Transmission reconstruction algorithms are not very computationally intensive, as they involve backprojection of ultrasonic data along arcs and sections of the image respectively. The Time-of-flight Diffraction and Pitch-Catch Reflection reconstruction methods require the backprojection of data over elliptical paths and are therefore more computationally intensive. In addition, the Time-of-flight Diffraction reconstruction algorithm utilises a more complex system of transmit and receive locations than the other reconstruction algorithms, thus adding to the computational complexity. For this reason the Time-of-flight Diffraction method is restricted to only a small number of transmitter locations in order to obtain an acceptable balance between image quality and reconstruction time. For the other methods the reconstruction time and image quality are deemed acceptable for the amount of data used. The image reconstruction times will be discussed in more detail for all of the image reconstruction algorithms in Chapter 4.

2.9. Partial Access to the Test Specimen.

In practice, access to the complete test specimen circumference is not always available. Therefore, it will be advantageous to consider the effect on the image point spread function of reconstructing an image from an incomplete set of data. This will obviously have a detrimental effect on the point spread function and all of the expressions derived previously will be invalid to a varying degree depending on the proportion of the full data set available. The performance of the image reconstruction algorithms and single bounce image enhancement tool will be evaluated experimentally when access to the test specimen is restricted in Chapter 4.

2.10. Concluding Summary.

This chapter presents the underlying theory for the majority of the main imaging system components. The tomographic reconstruction methods namely Reflection tomography, Time-of-flight Diffraction tomography and Transmission tomography have been explained in detail and the point spread functions for Pulse-Echo Reflection tomography, Pitch-Catch Reflection tomography and Time-of-flight Diffraction tomography derived. The data pre-processing component of the imaging system will be outlined in Chapter 3.

Two transmission tomography algorithms based on the shadow cast by a defect have been presented. They both involve the backprojection of the shadow regions, with one of the algorithms incorporating information about the defect type in the backprojection.

An image enhancement technique based on the ultrasonic energy reflected by the test specimen backwall has been presented for the two Reflection tomography and Time-of-flight Diffraction tomography reconstruction algorithms. The intention being to use the technique to improve characterisation of any defects present, this should be especially useful when access to the test specimen is restricted or the amount of data available for image reconstruction is small.

In order to facilitate the generation of high resolution images an image fusion technique based on fuzzy logic set theory has been given. The Fusion technique takes a set of the tomographic image reconstructed and applies the fusion function to the images to generate a high resolution image of any defects present. The imaging systems components presented here will be evaluated experimentally in Chapter 4.

Chapter 3:

MAXIMUM-LIKELIHOOD DECONVOLUTION.

3.1. Introduction.

In Non-Destructive Testing (NDT) image resolution is the key factor that dominates the quality of the image and also the amount of information it contains. Means to improve the resolution of ultrasonic images [28][29][30][31][32][33] fall into one of two categories these are hardware and software based methods. In hardware based methods a wideband transducer with shorter pulse width is utilised to improve axial resolution or pulse compression techniques are used [27]. To improve lateral resolution using hardware a focused transducer with a smaller focal spot is utilised. In practice, the bandwidth and the depth of focus are inherently limited [12]. These facts along with the nature of the imaging system proposed in Chapter 2 point to the use of software techniques based on digital signal processing for the improvement of the image quality. The approach that has been chosen is to use a deconvolution method to improve the SNR and time resolution of raw A-scan data utilised for image reconstruction, thus improving the quality of the reconstructed image.

In NDT the ultrasonic data received from scanning the test component is degraded by a number of system components, among these are:

- Spatial blurring due to the finite beamwidth of the scanning transducer.
- Spatial distortion due to the anisotropy scattering function of defects within the components.
- Noise introduced by scattering of energy from metal grains.

Since it has been decided to concentrate on software deconvolution methods, these will be examined in more detail with discussion focusing on the chosen method. In the majority of deconvolution methods employed accurate, a priori knowledge of the transducer impulse response, w(t), is required. This information can be obtained using either a measurement or model based approach.

The measurement approach is achieved by using a process in which the transducer response is characterised under conditions that closely resemble the intended operational environment. This method has the advantages that little knowledge of the transducer parameters and associated electrical and mechanical loading is required,

the transducer being considered a black box. However, this methodology has a number of drawbacks; the measured response is a direct function of the reflector type and position within the sound field, a separate measurement being required for different reflectors and positions within the sound field. Moreover, the experimental procedure can be time consuming and prone to error due to incorrect positioning or contamination noise. Also for some wideband applications high frequency absorption within the propagation channel reduces the effective bandwidth for the deconvolution process. It is apparent from this that effective deconvolution of ultrasonic data is severely constrained by experimental characterisation of the transducer impulse response and an alternative procedure is required. A number model based approaches have been examined in the literature, some of these will be outlined.

One method involves the use of transducer modelling techniques to closely approximate the transducer impulse response [34]. Apart from being significantly faster than the measurement approach, this approach is independent of measurement error and provides an enhanced degree of flexibility. However, detailed knowledge of the transducer construction, material parameters and electrical loading is required. Due to these constraints a third model based approach will be utilised.

The method considered here [34] for the deconvolution of ultrasonic data utilises a model of the experimental conditions under which the ultrasonic data is captured, this model is presented in Section 3.3. The chosen method uses the ultrasonic data to obtain an initial estimate for the parameters of the model. Optimisation techniques are then used to refine the model parameters until such time as a minimum is obtained. Once an acceptable estimate of the model parameters has been found, then these parameters are used in a deconvolution filter to improve the resolution of the ultrasonic data.

Before the techniques used here for the deconvolution of ultrasonic data are presented the basics of convolution will be described in Section 3.2. In subsequent sections the deconvolution model, the model parameter estimation methods and the

Maximum-Likelihood Deconvolution (MLD) algorithm used for optimisation of the model parameters are described in detail. The implementation of these algorithms is also examined as in many cases they are computationally intensive and require the used of equivalent recursive algorithms to obtain acceptable processing times. In Section 3.11 the Minimum Variance Deconvolution (MVD) filter which utilises the obtained model parameters to filter the A-scan data is presented. In Section 3.13 a technique for the estimation of the ultrasonic data's noise variance is described as it cannot be estimated using the MLD technique. Finally, in Section 3.15 the deconvolution techniques presented here are applied to experimentally obtained data to show the resultant improvement in the time resolution and Signal-to-Noise Ratio of the A-scan data.

3.2. Convolution.

Convolution is the most important method for the description of Linear-Time-Invariant (LTI) systems, since the convolution operator allows the output of the system to be calculated from the input and the system Impulse Response (IR), as

$$\text{output} = \text{input} * \text{IR} \qquad (3.2\text{-}1)$$

Where * denotes convolution.

Convolution is associated with the forward problem, that is identification of the system output given the input and the impulse response of the LTI system. Deconvolution is the inverse process, that is generating the system input from the output and the LTI system IR. If IR^{-1} represents the mathematical inverse of the system's impulse response. $IR^{-1}*IR$ is a delta function, and a delta function convolved with another function (output in this context) equals that other function, i.e.

$$\text{input} = \text{output} * IR^{-1} \qquad (3.2\text{-}2)$$

Deconvolution is used in many areas of science and engineering, from communication systems, seismic systems and non-destructive testing (NDT) [34][42]. The area of interest here is that of ultrasonic NDT signals.

If convolution is the most important and widely used LTI system operation, then deconvolution is the second most important. However, deconvolution is a far more difficult operation to perform. This can be justified by explaining a number of LTI system attributes; in practice it is not possible to obtain the inverse IR by simply inverting the IR. For example, if the system is non-minimum phase, so that some of the zeros of the z-transform of the system impulse response lie outside the unit circle in the complex z-domain, then the IR^{-1} will be unstable because some of its poles will be unstable. Therefore, direct inversion would not be advantageous. In addition to this the output of a LTI system is often corrupted by noise, so that

$$\text{measured output} = \text{input} * \text{IR} + \text{noise}, \qquad (3.2\text{-}3)$$

and,

$$\text{input} = \text{measured output} * \text{IR}^{-1} - \text{noise} * \text{IR}^{-1} \qquad (3.2\text{-}4)$$

The noise values are unknown so the input cannot be calculated from equation (3.2-2). If the noise is neglected, serious errors in reconstructing the input may occur.

The final reason is that, because the exact inverse cannot be computed, it may be very difficult to obtain a high resolution version of the input opposed to a blurred version, and so

$$\begin{aligned} \text{input} &= \text{measured output} * \text{IR}^{-1} - \text{noise} * \text{IR}^{-1} \\ &= (\text{input} * \text{IR} + \text{noise}) * \text{IR}^{-1} - \text{noise} * \text{IR}^{-1} \\ &= \text{input} * (\text{IR} * \text{IR}^{-1}) = \text{input} * \text{resolution function} \end{aligned} \qquad (3.2\text{-}5)$$

where

$$\text{resolution function} = IR * IR^{-1} \qquad (3.2\text{-}6)$$

If IR^{-1} is a perfect inverse of IR, then the resolution is a delta function. However if the inversion is not perfect then the resolution function is a smeared out delta function, where the degree of smearing depends on the bandwidth (BW) of the IR operation and the Signal-to-Noise Ratio (SNR) [34]. The design of a deconvolution operator requires a careful balancing of the SNR and BW effects. A deterministic approach is not suitable as it neglects the SNR effects. However, a stochastic design process can be used as it can be designed to take account of BW and SNR effects.

Stochastic design procedures can lead to two types of deconvolution operators, linear and non-linear. Non-linear methods generally give rise to higher resolution deconvolution operators, so an approach that generally leads to non-linear deconvolution operators has been chosen.

The deconvolution method to be examined here is based on the maximum-likelihood method developed by R. A. Fischer. This method, in the context of ultrasonic NDT inspection is; *given the A-scan data and an assumed model for that experiment, determine values for the parameters of the model which most probably led to the observed data.*

Probability is the statistical process associated with the forward experiment, and likelihood is associated with an inverse experiment (probability is proportional to likelihood). The maximum-likelihood method is based on the simple idea that different probability models generate different samples and that given any chosen sample, it is more likely to have come from some probability models than from others. In order to apply the maximum-likelihood method to the design of a deconvolution operator (i.e. filter), the following steps must be undertaken;

1. Specify a probability model for the measured output;
2. Determine a formula for the likelihood function; and,
3. Maximise the likelihood function.

These steps will now be examined for the NDT system described in chapter 2.

3.3. Convolution model.

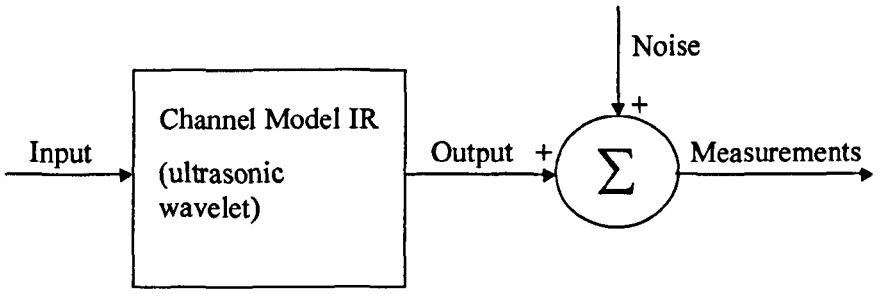


Figure 3.1. *Convolutional Model.*

The basic convolutional model is

$$\text{measured output} = \text{output} + \text{noise} = \text{input} * \text{IR} + \text{noise} \qquad (3.3\text{-}1)$$

This is shown in Figure 3.1. The three components of the model, i.e. the input, the system impulse response and the noise will now be outlined.

3.3.1. Input.

The fundamental assumption made about the input is that it is random and white in nature. White means that values of the input signal are independent from one time value to the next. This implies that the present input value gives no information about the past or future values of the input, and vice versa. However, if the input sequence is not white but still random then this signal is referred to as coloured noise. Coloured noise can be generated by applying white noise to a colouring filter. If the impulse response of such a filter is IR_{CF}, then the

$$\text{coloured input sequence} = \text{white input sequence} * IR_{CF} \qquad (3.3\text{-}2)$$

so that

$$\begin{aligned} \text{measured output} &= \text{coloured input sequence} * IR + \text{noise} \\ &= \text{white input sequence} * IR_{CF} * IR + \text{noise} \\ &= \text{white input sequence} * IR' + \text{noise} \\ &= \text{input sequence} * IR' + \text{noise} \end{aligned} \tag{3.3-3}$$

where

$$IR' = IR_{CF} * IR \tag{3.3-3}$$

This shows that the convolutional model can be expressed in terms of a white input sequence that is applied to a more complex channel model. Thus illustrating that assuming a white input sequence is not a restrictive assumption.

There are many models that can be used to approximate white sequences. The methods considered here will be: Gaussian, Bernoulli, Bernoulli Gaussian and Bernoulli Gaussian plus backscatter.

3.3.1.1. Gaussian White Sequence.

An example of a Gaussian sequence is shown in figure 3.2, where a Gaussian random number generator has been used to generate the sequence elements.

Figure 3.2. *Gaussian input sequence.*

The values of this sequence will be denoted by u(k), the numbers

$$u(1),u(2),.....,u(k),.....,u(N)$$

denote a Gaussian white sequence with N elements, with the sequence elements occurring at time points 1,2,.....,k,.....,N. This sequence is completely characterised by its mean and variance. For simplicity the mean is assumed to be zero, thus the sequence is described by a single parameter, its variance v_r. Entropy is defined as the uncertainty about a random signal and also the information gained when a signal is observed. Gaussian sequences are known to be maximum entropy signals.

3.3.1.2. Bernoulli White Sequence.

A random sequence of ones and zeros is a Bernoulli sequence. If the sequence has an equal number of ones and zeros, then the probability of a one or a zero occurring is 1/2. If we require a sparse sequence then obviously the probability of a zero occurring must be made significantly higher than the probability of a one occurring. Using this philosophy it is possible to generate any random sequence containing any number of ones and zeros simply by altering the two probabilities. An example of such a sequence is shown in figure 3.3.

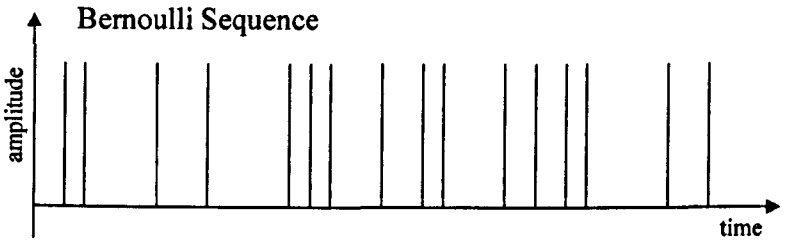


Figure 3.3. *Bernoulli Sequence.*

The Bernoulli sequence elements will be denoted by q(k), the numbers

$$q(1),q(2),.....,q(k),.....,q(N)$$

denote a Bernoulli sequence with N elements, where the sequence elements occur at time points 1,2,.....,k,.....,N. There are a possible 2^N Bernoulli sequences, each of these being known as a realisation. Any of the realisations can be chosen simply be choosing at which time points a value of one is desired. Each element of the Bernoulli sequence is described by a probability mass function that is characterised by a single parameter, λ, this value equals the mean value of the sequence and also the variance of the sequence. In general, λ takes on a value between zero and one, and $Pr[q(k)] = \lambda$ if $q(k) = 1$ or $Pr[q(k)] = 1 - \lambda$ if $q(k) = 0$. The entropy of the Bernoulli sequence q(k) is $\lambda\ln(\lambda) - (1 - \lambda)\ln(1 - \lambda)$ i.e. it has a minimum when $\lambda = 0$ or 1 and a maximum when $\lambda = 1/2$. Consequently, the most entropic values of the sequence occur when λ is close to 1/2, whereas the least entropic values occur when λ is close to zero or one.

3.3.1.3. Bernoulli-Gaussian Sequence.

Non-Gaussian input signals have a very important role to play in Maximum-Likelihood Deconvolution, as they lead to non-linear deconvolution operators. A Bernoulli-Gaussian sequence is such a non-Gaussian sequence. It can be obtained by multiplying the elements of a Gaussian sequence with the values of a Bernoulli sequence. An example of such a sequence can be seen in Figure 3.4 and has been determined by the multiplication of the Gaussian sequence in Figure 3.2 and the Bernoulli sequence shown in Figure 3.3. Another way to generate a Bernoulli-Gaussian sequence is to consider the nature of the Bernoulli sequence. If the fact that the sequence only contains two values, zero and one, is used then a Bernoulli-Gaussian sequence can be generated by only switching on the Gaussian random number generator when the Bernoulli sequence has a value of unity.

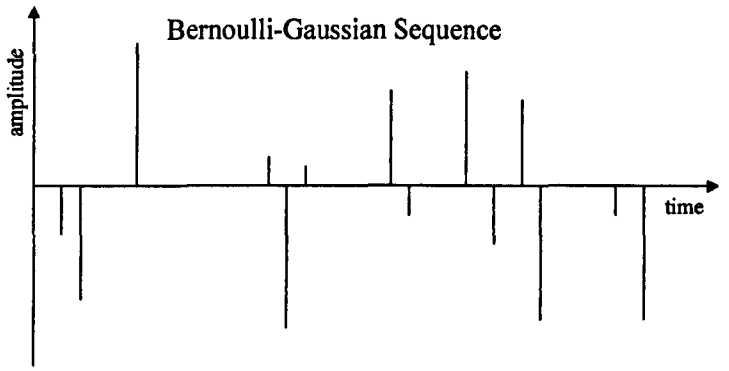


Figure 3.4. *Bernoulli-Gaussian sequence.*

This product model for the generation of a Bernoulli-Gaussian sequence can be described as follows. If q(k) denotes the elements of a Bernoulli sequence and r(k) denotes the elements of the Gaussian sequence, then the Bernoulli-Gaussian sequence equals r(k)q(k). So letting u(k) represent the input sequence,

$$u(k) = r(k)q(k) \qquad (3.3\text{-}4)$$

The numbers

$$r(1)q(1),r(2)q(2),.....,r(k)q(k),.....r(N)q(N)$$

denote the Bernoulli-Gaussian sequence with N elements, at time points 1,2,.....,k,.....,N. This sequence requires two parameters to completely characterise it; λ, which is the probability parameter value associated with the Bernoulli sequence, and v_r, which is the variance of the Gaussian sequence.

3.3.1.4. Bernoulli-Gaussian plus Backscatter Sequence.

For the ultrasonic NDT applications of Maximum-Likelihood Deconvolution considered here the input signal is the reflectivity function that characterises the specimen under test. When we model the input as a Bernoulli-Gaussian sequence we are accounting for any defects contained within the test specimen, however some

components, by the nature of their material properties will cause grain noise to be present in the input sequence. This effect is modelled by using another zero mean Gaussian sequence, $u_B(k)$, the backscatter sequence. The Bernoulli-Gaussian sequence plus backscatter input model is

$$u(k) = r(k)q(k) + u_B(k) \qquad (3.3\text{-}5)$$

where k = 1,2,.....,N. An example of such a sequence is shown in Figure 3.5. It was obtained by adding a Gaussian white sequence to a Bernoulli-Gaussian sequence.

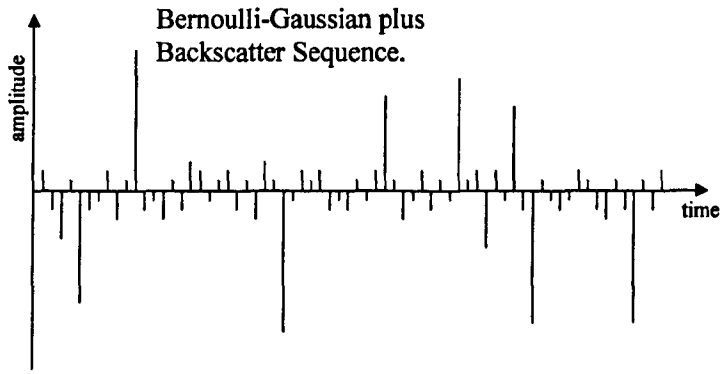


Figure 3.5. *Bernoulli-Gaussian plus backscatter sequence.*

As outlined previously in equation (3.1-3),

$$\text{measured output} = \text{input}*\text{IR} + \text{noise}$$

However, the input is now viewed as a sum of two terms

$$\text{input} = \text{specimen message} + \text{grain noise} \qquad (3.3\text{-}6)$$

The backscatter sequence modelling the component of the input referred to as grain noise, is a convolutional parameter. The Bernoulli-Gaussian plus backscatter input sequence requires three terms to completely characterise it; λ, and v_r the parameters associated with the Bernoulli-Gaussian sequence, and v_B, the variance of the

Gaussian backscatter sequence. The variance v_B has an interesting effect on the input sequence. Decreasing or increasing it allows us to generate input sequences ranging from less to more entropic.

3.4. Channel Model Impulse Response (ultrasonic wavelet).

The next part of the model to be examined is the channel's impulse response, i.e. the wavelet generated by the ultrasonic source. The sampled values of the wavelet are denoted by

$$w(0),w(1),w(2),.....$$

where w(0) is usually zero. One model for a wavelet is to use its sampled values directly. For example, if a finite length wavelet is considered, i.e. for w(k) = 0 for all values of k greater than M. Then

$$W(z) = w(0) + w(1)z^{-1} + w(2)z^{-2} + ...+ w(M)z^{-M} \qquad (3.4\text{-}1)$$

where W(z) is the z-transform of w(k). This model is commonly referred to as a moving average model (MA), the parameters of the model being w(0),w(1),.....,w(M). If the wavelet is known, then the number M will also be known. However, in many cases the wavelet is unknown, so M is consequently unknown. This model has the draw back that it can only be used to model wavelets with power spectral densities containing peaks. If the wavelet spectral density contains both peaks and troughs then a more complex model is required [43], such a model is an Auto-Regressive Moving Average (ARMA) [44] and takes the form of a ratio of two polynomials, i.e.

$$W(z) = \frac{b_1 z^{n-1} + b_2 z^{n-2} + ...+ b_{n-1} z + b_n}{z^n + a_1 z^{n-1} + ...+ a_{n-1} z + a_n} \qquad (3.4\text{-}2)$$

And is described by the 2n parameters

$$a_1,a_2,.....,a_n \text{ and } b_1,b_2,.....,b_n.$$

Where n typically ranges from 2 to 12.

Sometimes the wavelet samples are known. In this case the ARMA model parameters can be obtained in a number of ways. This is an approximation problem that involves determining the model parameters, such that some measure between the approximate wavelet and the given M samples is minimised.

When the wavelet samples are unknown then some other method must be used to determine the ARMA wavelet parameters, many such methods exist the majority of them using higher order statistics [44] of the discrete system output [43][45][46][47]. A method based on these higher order statistics and Singular Value Decomposition has been used to determine the model order and parameters. This will be outlined in Section 3.10.

3.5. Measurement Noise.

The third component of the convolution model is measurement noise, which is assumed to be additive. The assumptions made about this noise are that it is zero mean, white and Gaussian. This noise will be denoted by n(k) where k = 1,2,.....,N and is characterised by the variance v_n.

3.6. Mathematical Model.

Now that the underlying theories of the convolution model have been presented, they will now be collected to form a single system model [34]. All of components of the convolutional model; the input, the channel mode IR, output, noise and measurements are now represented by mathematical symbols these will now be summarised for clarity [34]:

u(k) = random input signal sequence which will be modelled as a Bernoulli-Gaussian or a Bernoulli-Gaussian plus backscatter sequence.

w(k) = channel model IR which will be modelled as an ARMA wavelet.

y(k) = output of convolutional model.

n(k) = additive measurement noise, which will be modelled as white and Gaussian.

z(k) = measurements which are available at the output of the receive transducer at time points 1,2,.....,N.

The basic convolutional model can now be expressed mathematically, as

$$z(k) = u(k) * w(k) + n(k), \quad k = 1,2,.....,N \qquad (3.6\text{-}1)$$

or alternatively as

$$z(k) = [r(k)q(k) + u_B(k)] * w(k) + n(k), \quad k = 1,2,.....,N \qquad (3.6\text{-}2)$$

for a Bernoulli-Gaussian plus Backscatter sequence. For convenience these N measurements are represented by a vector **z**, where

$$\mathbf{z} = \text{col}(z(1),z(2),.....,z(N)),$$

where col() denotes a column vector. Using the fact that

$$u(k) * w(k) = \sum_{j=1}^{k} u(j)w(k-j) \qquad (3.6\text{-}3)$$

it is observed that

$z(1) = w(0)u(1) + n(1)$

$z(2) = w(1)u(1) + w(0)u(2) + n(2)$

.....

$z(N) = w(N\text{-}1)u(1) + w(N\text{-}2)u(1) + ... + w(0)u(N) + n(N).$

For convenience, these N equations can be written as the following vector measurement equation.

$$\mathbf{z} = \mathbf{W}\mathbf{u} + \mathbf{n} \qquad (3.6\text{-}4)$$

where

$$\mathbf{u} = \text{col}(u(1),u(2),.....,u(N))$$
$$\mathbf{n} = \text{col}(n(1),n(1),.....n(N))$$

and

$$\mathbf{W} = \begin{bmatrix} w(0) & 0 & ..... & 0 \\ w(1) & w(0) & ..... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ w(n-1) & w(n-2) & ..... & w(0) \end{bmatrix}$$

If $w(k) \cong 0$ for $k > M$, then many of the terms in **W** will be zero. However there is no loss in generality of showing these terms not equal to zero. Using the fact that $u(k) = r(k)q(k) + u_B(k)$, we can extend the vector measurement equation to

$$\mathbf{z} = \mathbf{W}\mathbf{Q}\mathbf{r} + \mathbf{W}\mathbf{u}_B + \mathbf{n} \qquad (3.6\text{-}5)$$

where

$$\mathbf{Q} = \text{diag}[q(1),q(2),.....,q(N)]$$
$$\mathbf{r} = \text{col}[r(1),r(2),.....,r(N)]$$
$$\mathbf{u_B} = \text{col}[u_B(1), u_B(2),....., u_B(N)]$$

in the equation for **Q**, diag[] denote a diagonal matrix. The mathematical convolutional model is now complete and is shown in Figure 3.6.

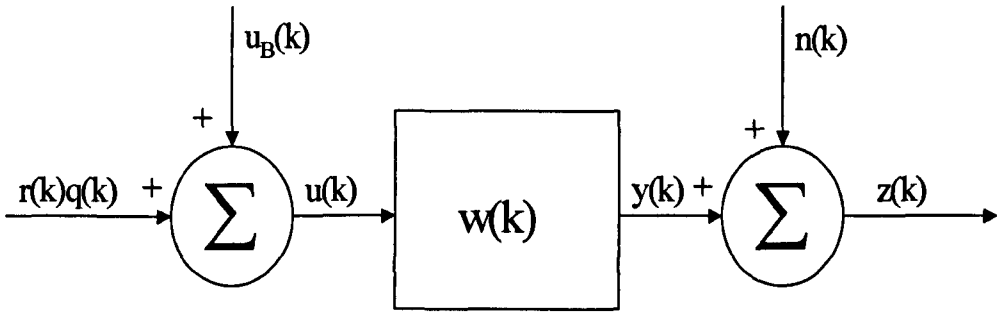


Figure 3.6. *Mathematical convolution model.*

The convolutional model is a parametric one. It is completely described by parameters that are associated with the three model components, these can be either random or deterministic in nature. There are 2n + 4 deterministic and 3N random parameters [34]. These being:

Deterministic parameters:

Wavelet parameters: $\mathbf{a} = \text{col}(a_1, a_2, \ldots\ldots, a_n)$

$\mathbf{b} = \text{col}(b_1, b_2, \ldots\ldots, b_n)$

statistical parameters: $\mathbf{s} = \text{col}(v_r, v_B, v_n, \lambda)$

Random parameters:

Event parameters: $\mathbf{q} = \text{col}[q(1), q(2), \ldots\ldots, q(N)]$

Amplitude Parameters: $\mathbf{r} = \text{col}[r(1), r(2), \ldots\ldots, r(N)]$

Backscatter Parameters: $\mathbf{u_B} = \text{col}[u_B(1), u_B(2), \ldots\ldots, u_B(N)]$

This model requires the determination of 3N + 2n + 4 parameters from only N measurements, this can be achieved because in-addition to the N measurements there are the convolution model and the probabilistic models for the 3N random parameters.

3.7. Likelihood.

The 3N random parameters and the 2n + 4 deterministic parameters to be estimated have been established, the next step is to derive an expression for the likelihood function. R.A. Fischer [34] developed the method of maximum-likelihood for problems characterised by deterministic parameters. The method, associated with Thomas Bayes, called a maximum aposteriori method (MAP) was developed for problems characterised just by random parameters. The deconvolution problem considered here has to account for both types of parameters in the correct way; i.e. our likelihood function will treat **a**, **b**, **s** as deterministic parameters and **r**, **q**, $\mathbf{u}_B$, as random. The resultant likelihood function is called an unconditional likelihood function, because the random parameters have been accounted for properly. For simplicity this will be referred to as the likelihood function [34].

The likelihood function contains exponential functions, so taking the natural logarithm of the likelihood function removes inherent complexity. Due to this the log likelihood function is used here, this can be done because the logarithm of a function is a monotonic transformation of the function, i.e. the loglikelihood function mirrors the likelihood function [34]. Consequently the maximum of the log likelihood function corresponds to the maximum of the likelihood function.

3.7.1. Likelihood Function.

Whether the likelihood or loglikelihood functions are considered the underlying deconvolution problem remains the same, that is given the measurements, **z**, the unknowns **a**, **b**, **s**, **r**, **q** and $\mathbf{u}_B$ must be estimated. The derivation of the likelihood function will now be presented. Obviously it must be a function of **a**, **b**, **s**, **r**, **q** and $\mathbf{u}_B$. The function takes the form

$$L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$$

Using the fact that likelihood is proportional to probability, and assuming a value of one for the constant of probability, it can be stated that

$$L\{\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B\} = p(\mathbf{z},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{a},\mathbf{b},\mathbf{s}) \qquad (3.7\text{-}1)$$

Function $p(\mathbf{z}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{a}, \mathbf{b}, \mathbf{s})$ is the joint probability density function of random vectors **z**, **r**, **q** and $\mathbf{u}_B$ given the deterministic vectors **a**, **b** and **s**. This likelihood function is conditional with respect to **a**, **b** and **s** and is unconditional with respect to **r**, **q** and $\mathbf{u}_B$. Applying the elementary rules of conditional probabilities to the likelihood function, the following is obtained

$$\begin{aligned} L\{\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{z}\} &= p(\mathbf{z},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{a},\mathbf{b},\mathbf{s}) \\ &= p(\mathbf{z} \mid \mathbf{q},\mathbf{r},\mathbf{u}_B,\mathbf{a},\mathbf{b},\mathbf{s})p(\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{a},\mathbf{b},\mathbf{s}) \\ &= p(\mathbf{z} \mid \mathbf{q},\mathbf{r},\mathbf{u}_B,\mathbf{a},\mathbf{b},\mathbf{s})p(\mathbf{q} \mid \mathbf{a},\mathbf{b},\mathbf{s})p(\mathbf{r} \mid \mathbf{a},\mathbf{b},\mathbf{s})p(\mathbf{u}_B \mid \mathbf{a},\mathbf{b},\mathbf{s}) \end{aligned} \qquad (3.7\text{-}2)$$

For the last expression, the fact that **r**, **q** and $\mathbf{u}_B$ are statistically independent is utilised. Now, using the fact that **q** is a vector of discrete random variables, we express $p(\mathbf{q} \mid \mathbf{a}, \mathbf{b}, \mathbf{s})$ as $\Pr(\mathbf{q} \mid \mathbf{a}, \mathbf{b}, \mathbf{s})$; hence,

$$\begin{aligned} L\{\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{z}\} = p(\mathbf{z} \mid \mathbf{q},\mathbf{r},\mathbf{u}_B,\mathbf{a},\mathbf{b},\mathbf{s})\Pr(\mathbf{q} \mid \mathbf{a},\mathbf{b},\mathbf{s})p(\mathbf{r} \mid \mathbf{a},\mathbf{b},\mathbf{s}) \\ p(\mathbf{u}_B \mid \mathbf{a},\mathbf{b},\mathbf{s}) \end{aligned} \qquad (3.7\text{-}3)$$

Next, because **r**, **q** and $\mathbf{u}_B$ do not depend on **a** and **b**, $L\{\ \}$ can be expressed as

$$L\{\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{z}\} = p(\mathbf{z} \mid \mathbf{q},\mathbf{r},\mathbf{u}_B,\mathbf{a},\mathbf{b},\mathbf{s})\Pr(\mathbf{q} \mid \mathbf{s})p(\mathbf{r} \mid \mathbf{s})p(\mathbf{u}_B \mid \mathbf{s}) \qquad (3.7\text{-}4)$$

The probability functions on the right hand side will now evaluated, the key to this is the vector measurement equation,

$$\mathbf{z} = \mathbf{WQr} + \mathbf{Wu}_B + \mathbf{n} \qquad (3.7\text{-}5)$$

Due to the Bernoulli nature of the elements of **q**, and the Gaussian nature of **r**, $\mathbf{u}_B$ and **n** it can be established that [34]:

$$p(\mathbf{z}\,|\,\mathbf{q},\mathbf{r},\mathbf{u}_B,\mathbf{a},\mathbf{b},\mathbf{s}) = (2\pi v_n)^{-N/2}\exp\{-(\mathbf{z}-\mathbf{WQr}-\mathbf{Wu}_B)' (\mathbf{z}-\mathbf{WQr}-\mathbf{Wu}_B)/2v_n\} \tag{3.7-6}$$

$$\Pr(\mathbf{q}\,|\,\mathbf{s}) = \prod_{k=1}^{N}\Pr[q(k)|\,\lambda] = \lambda^{m(q)}(1-\lambda)^{[N-m(q)]}, \tag{3.7-7}$$

In which

$$m(q) = \sum_{q=1}^{N} q(k); \tag{3.7-8}$$

$$p(\mathbf{r}\,|\,\mathbf{s}) = (2\pi v_r)^{-N/2}\exp(-\mathbf{r}'\mathbf{r}/2v_r); \tag{3.7-9}$$

and,

$$p(\mathbf{u}_B\,|\,\mathbf{s}) = (2\pi v_B)^{-N/2}\exp(-\mathbf{u}_B'\mathbf{u}_B/2v_B) \tag{3.7-10}$$

Combining these equations the expression obtained for L{ } is

$$\begin{aligned} L\{a,b,s,q,r,u_B\,|\,z\} = (2\pi)^{-3N/2}(v_r v_n v_B)^{-N/2}\exp\{&-\mathbf{r}'\mathbf{r}/2v_r \\ &-(\mathbf{z}-\mathbf{WQr}-\mathbf{Wu}_B)'(\mathbf{z}-\mathbf{WQr}-\mathbf{Wu}_B)/2v_n \\ &-\mathbf{u}_B'\mathbf{u}_B/2v_B\}\lambda^{m(q)}(1-\lambda)^{[N-m(q)]} \end{aligned} \tag{3.7-11}$$

As mentioned earlier is often easier to work with the loglikelihood function, *L{}*, this is shown in equation (3.7-12).

$$\begin{aligned} L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B\,|\,\mathbf{z}) = &-\frac{3N}{2}\ln 2\pi - \frac{N}{2}\ln(v_r v_n v_B) - \mathbf{r}'\mathbf{r}/2v_r \\ &-(\mathbf{z}-\mathbf{WQr}-\mathbf{Wu}_B)'(\mathbf{z}-\mathbf{WQr}-\mathbf{Wu}_B)/2v_n \\ &-\mathbf{u}_B'\mathbf{u}_B/2v_B + m(q)\ln(\lambda) \\ &+[N-m(q)]\ln(1-\lambda) \end{aligned} \tag{3.7-12}$$

Usually, the first term is dropped, as it does not depend on any of the model parameters. In the case where no backscatter is included, then this equation becomes:

$$\begin{aligned} L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r}\mid\mathbf{z}) = & -\frac{3N}{2}\ln 2\pi - \mathbf{r}'\mathbf{r}/2v_r \\ & -(\mathbf{z}-\mathbf{WQr})'(\mathbf{z}-\mathbf{WQr})/2v_n \\ & +m(q)\ln(\lambda) \\ & +[N-m(q)]\ln(1-\lambda) \end{aligned} \qquad (3.7\text{-}13)$$

3.8. Maximising Likelihood.

The deconvolution problem has been shown to be an optimisation problem. There are many methods that can be used to maximise $L\{\}$. A Block Component (BCM) search algorithm will be used to maximise $L\{\}$ [34].

3.8.1. Block Component Search Algorithms.

Two factors complicate the maximisation of the likelihood function. The first is the large number of parameters (3N + 2n + 4), and the second is the discrete nature of the N elements of **q** [34]. Application of popular gradient search algorithms to the likelihood needs some modification, due to the nature of **q**. The derivative of $L\{\}$ with respect to these zero/one parameters does not exist.

One method that can be used is a recursive block optimisation method [34]. For example, if a function of two variables is considered say, f(x,y). The first step is to fix x and maximise f(x,y) with respect to y. The second step is to keep y constant and maximise f(x,y) with respect to x. These two steps are repeated recursively. Let x_i and y_i denote the values of x and y at iteration i, and x_{i+1}, y_{i+1} denote the values at the i+1 st iteration. For this method to be successful;

$$f(x_i,y_i) \le f(x_{i+1},y_i) \le f(x_{i+1},y_{i+1}) \qquad (3.8\text{-}1)$$

This approach to optimisation is equally valid for functions of more than two variables. For the optimisation of $L\{\}$ one possible approach is to fix the generic variables, x and y, as follows:

$$x = \text{elements of } \mathbf{q}, \quad \text{and} \quad y = \text{elements of } \mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{r}, \mathbf{u}_B.$$

Using this approach means a total of 2^N possible $\mathbf{q}$ vectors have to be generated, and 2^N value of $L\{\}$ have to be computed. Even for small values of N this leads to prohibitive computation requirements, so another approach must be used. Keeping $\mathbf{q}$ constant and maximising $L\{\}$ with respect to y can be accomplished in many ways. $\mathbf{a}$, $\mathbf{b}$, $\mathbf{s}$, $\mathbf{r}$ and $\mathbf{u}_B$ are continuous in nature, so it is possible to use a gradient search algorithm to update the entire collection of these y-parameters at the same time. This would still be extremely costly in term of computational resources, as there are 2N + 2n + 4 elements in y. However, the block component search algorithm does not have to be reduced to 2 steps; y can be decomposed further, for example into

$$y_1 = \text{elements of } \mathbf{r}, \mathbf{u}_B$$
$$y_2 = \text{elements of } \mathbf{a}, \mathbf{b}$$
$$y_3 = \text{elements of } \mathbf{s}$$

This approach is summarised in Figure 3.7. It guarantees that

$$\begin{aligned} L(\mathbf{a}_i, \mathbf{b}_i, \mathbf{s}_i, \mathbf{q}_i, \mathbf{r}_i, \mathbf{u}_{B,i} \mid \mathbf{z}) &\le L(\mathbf{a}_i, \mathbf{b}_i, \mathbf{s}_i, \mathbf{q}_{i+1}, \mathbf{r}_i, \mathbf{u}_{B,i} \mid \mathbf{z}) \\ &\le L(\mathbf{a}_i, \mathbf{b}_i, \mathbf{s}_i, \mathbf{q}_{i+1}, \mathbf{r}_{i+1}, \mathbf{u}_{B,i+1} \mid \mathbf{z}) \\ &\le L(\mathbf{a}_{i+1}, \mathbf{b}_{i+1}, \mathbf{s}_i, \mathbf{q}_{i+1}, \mathbf{r}_{i+1}, \mathbf{u}_{B,i+1} \mid \mathbf{z}) \\ &\le L(\mathbf{a}_{i+1}, \mathbf{b}_{i+1}, \mathbf{s}_{i+1}, \mathbf{q}_{i+1}, \mathbf{r}_{i+1}, \mathbf{u}_{B,i+1} \mid \mathbf{z}) \end{aligned} \qquad (3.8\text{-}2)$$

The above illustrates that the likelihood is increased from one iteration to the next. The generic variable can be partitioned in many ways; this particular partition has lead to a block component search algorithm in which the parameters are updated in the following order; random parameters, wavelet parameters and finally statistical

parameters. There is no reason to assume that this particular search algorithm is any better than any other.

A block component search algorithm provides an optimisation strategy. Within each block of the algorithm a free choice of optimisation algorithms exists, and sub iteration may be performed within each block. For example algorithms such as steepest descent, conjugate gradient or Marquardt-Levenberg algorithms may be used to update the y_2 and y_3 parameters from their iteration i values to their iteration i+1 values. An MVD algorithm may be used to update the y_1 parameters from their ith iteration values to their ith+1 iteration values, a variety of recursive detectors can be used to update the x parameters. The i+1 iteration values are only chosen to replace the i iteration values of the generic variable if there is an increase in the likelihood value when the i+1 iteration values are used to calculate it.

The block component Search algorithm can be stopped in a number of ways: the most common being to run for a set number of iterations. Alternatively the algorithm is run until the likelihood does not increase by more than a small pre-set value from iteration i to iteration i+1.

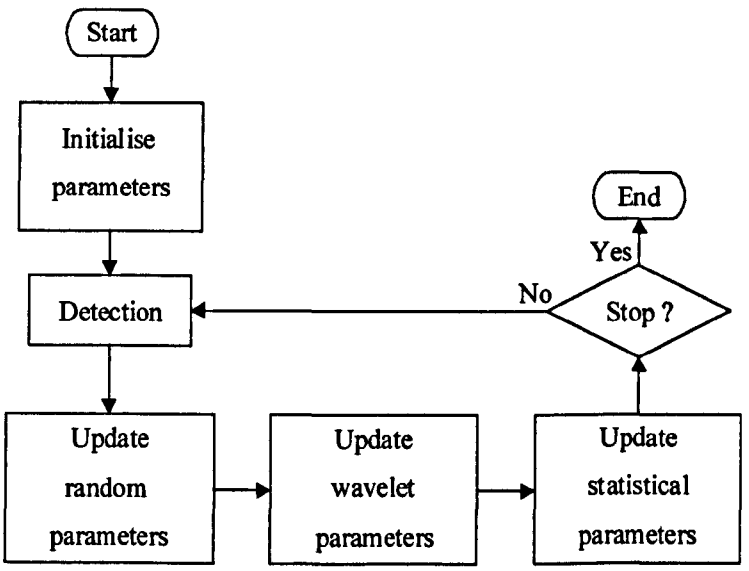


Figure 3.7. *Block Component Search algorithm.*

3.8.2. Update Random parameters.

When a block component search algorithm is used to maximise $L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$, then $\mathbf{r}$ and $\mathbf{u}_B$ are computed using,

$$\mathbf{r}^{ML} = A\left(\mathbf{a}^{ML}, \mathbf{b}^{ML}, \mathbf{s}^{ML}, \mathbf{q}^{ML}\right)\mathbf{z} \qquad (3.8\text{-}3)$$

and

$$\mathbf{u}_B^{ML} = B\left(\mathbf{a}^{ML}, \mathbf{b}^{ML}, \mathbf{s}^{ML}, \mathbf{q}^{ML}\right)\mathbf{z} \qquad (3.8\text{-}4)$$

in which the Maximum-Likelihood quantities are replaced by those obtained at the i th iteration, i.e.

$$\mathbf{r}_i = A\left(\mathbf{a}_i, \mathbf{b}_i, \mathbf{s}_i, \mathbf{q}_i\right)\mathbf{z} \qquad (3.8\text{-}5)$$

and

$$\mathbf{u}_{B,i} = B\left(\mathbf{a}_i, \mathbf{b}_i, \mathbf{s}_i, \mathbf{q}_i\right)\mathbf{z} \qquad (3.8\text{-}6)$$

These algorithms for updating $\mathbf{r}$ and $\mathbf{u}_B$ are known as Minimum Variance Deconvolution (MVD) algorithms [34], specific algorithms for updating the random parameters will be described later.

3.8.3. Binary Detection.

Each element of the event sequence $\mathbf{q}$ is a binary variable. When q(k) =1 an event has taken place at time point k, whereas no event has taken place if q(k) = 0. At each time point a decision must be made as to whether an event has taken place or not. This decision must be made in such a way as to lead to an increase in the likelihood function $L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$.

Many strategies exist for solving this binary decision process. Some process all of the data just once and lead to a decision about whether q(k) = 1 or 0 for all values of k = 1,2,.....,k,.....,N. These detectors are known as one shot detectors. Other

detectors process all of the data in an iterative fashion. At each iteration there is a decision rule, about whether q(k) = 1 or 0. These are recursive detectors and are usually not self starting (they require an initial set of values for q(k), k = 1,2,.....,N). The subscript 0 is used to denote the initial **q** values, e.g. $\mathbf{q}_0$. For the application to non-real-time data considered here these two classes of detectors are adequate [34].

Regardless of the detector chosen, the structure of the detection procedure is the same and contains three fundamental steps. The first step is to process the measurements **z**, the next step is to create the decision function and the final step is to specify the decision strategy that allows the value of q(k) to be determined. Let D(z; k) denote the decision function and let $S_k[D(z; k)]$ denote the decision strategy. The subscript k on S indicates that a decision must be made for all the values of k in the range. The two detectors used in the maximum-likelihood framework will now be described.

3.8.3.1. Threshold Detector.

The threshold detector is a one shot detector, and is used to generate the initial values of the event vector $\mathbf{q}_0$. In order to run this detector the initial values of $\mathbf{a}_0$, $\mathbf{b}_0$ and $\mathbf{s}_0$ must be provided. The first step is to create a non-linear function of the data. This is achieved by first processing all N measurements linearly using a Minimum Variance Deconvolution filter, to obtain $u^{MV}(k|N)$, k = 1,2,.....,N, and then squaring $u^{MV}(k|N)$. The next step is to compute a time varying threshold function t(k) that depends on the error variance between $u^{MV}(k|N)$ and u(k) as follows:

$$t(k) = \{A_0(k)A_1(k)/[A_1(k)-A_0(k)]\}\{\ln[A_1(k)/A_0(k)]-2\ln[\lambda/(1-\lambda)]\} \qquad (3.8\text{-}7)$$

where, for q = 0 and 1,

$$\begin{aligned} A_q(k) = &\{1-\mathrm{Var}[u_{ERR}(k|N)]/v_u\}^2[qv_r+v_B] \\ &+\mathrm{Var}[u_{ERR}(k|N)]\{1-\mathrm{Var}[u_{ERR}(k|N)]/v_u\} \end{aligned} \qquad (3.8\text{-}8)$$

$Var[u_{ERR}(k|N)]$ is the covariance matrix that describes the error between $\mathbf{u}$ and $\mathbf{u}^{MV}(N)$. The decision function is the difference between $[u^{MV}(k|N)]^2$ and t(k). i.e.

$$\begin{aligned} &\text{If } [u(k|N)]^2 - t(k) > 0 \text{ decide } q(k) = 1; \text{ or} \\ &\text{If } [u(k|N)]^2 - t(k) < 0 \text{ decide } q(k) = 0. \end{aligned} \qquad (3.8\text{-}9)$$

The threshold detector algorithm is as shown in Figure 3.8, below.

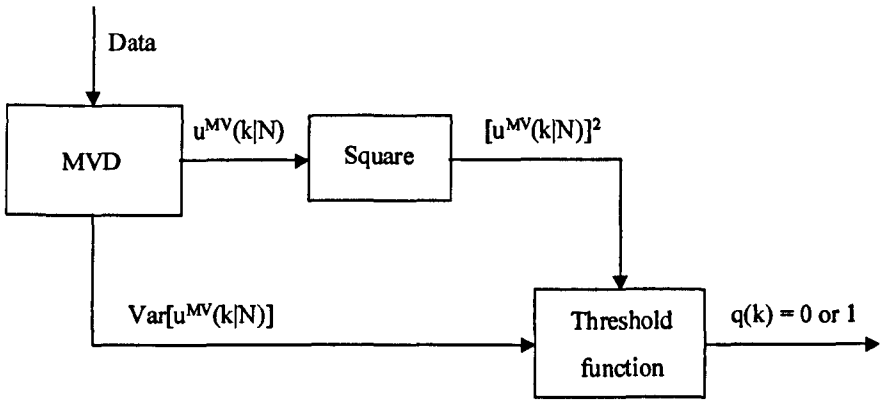


Figure 3.8. *Threshold Detection algorithm.*

3.8.3.2. Single Most-Likely Replacement Detector.

A Single Most-Likely Replacement detector (SMLR detector) is a recursive detector [34], therefore it cannot be used to generate the initial value of $\mathbf{q}$. This detector is used for detection once these initial values are available. The SMLR detector utilises the vectors $\mathbf{a}_0$, $\mathbf{b}_0$ and $\mathbf{s}_0$ and a reference $\mathbf{q}$, that is denoted by $\mathbf{q}_r$. The first choice of $\mathbf{q}_r$ is the value of $\mathbf{q}$ obtained from the threshold detector, $\mathbf{q}_{TD}$. N test vectors which differ from $\mathbf{q}_r$ in a single location are generated, these are denoted

$$q_{t,1}, q_{t,2}, \ldots\ldots, q_{t,N}$$

Note that the kth test vector, $\mathbf{q}_{t,k}$, only differs from q_r at the kth time point. By evaluating a likelihood ratio between $\mathbf{q}_{t,k}$ and $\mathbf{q}_r$, a decision function D(z; k) can be computed, this is given in equation (3.8-10).

$$\ln D(z;k) = \left(u^{MV}(k \mid N)\right)^2 / \left\{v_r^{-1}\left[q_{t,k}(k) - q_r(k)\right]^{-1} + \left[v_u - \mathrm{var}\left[u_{ERR}(k \mid N)\right]/v_u^2\right]\right\} + 2\left[q_{t,k}(k) - q_r(k)\right]\ln\left[\lambda/(1-\lambda)\right] \quad (3.8\text{-}10)$$

The decision strategy is to examine the decision function, D(z; k) and find its maximum value. Call this point k', this represents the single time point at which the reference sequence has been changed. This winning sequence $\mathbf{q}_{t,k}$ replaces the reference sequence for the next iteration of the detector. The detector is run until D(z; k) is less than zero for all values of k = 1,2,.....,N or for a set number of iterations. The algorithm is shown in Figure 3.9.

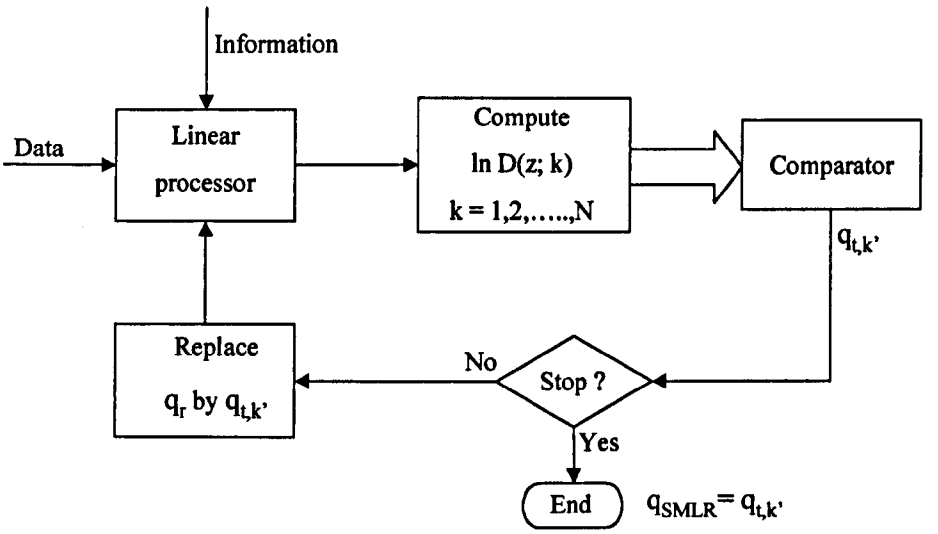


Figure 3.9. *SMLR (Single Most Likely Replacement) Detector algorithm.*

3.8.4. Update Wavelet Parameters.

The function $L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$ is a complicated function that depends on the wavelet parameters **a** and **b** in a highly non-linear fashion. The only methods available to maximise this function with respect to **a** and **b** are mathematical programming ones, the most commonly used in Maximum-likelihood Deconvolution are methods that use both first and second derivative information about *L*, such as the Newton-Raphson and Marqdaurt-Levenberg algorithms. The algorithm that will be used to maximise *L* will be the Marqdaurt-Levenberg algorithm [34].

Let $\mathbf{a}_i$ and $\mathbf{b}_i$ be the values of **a** and **b** for the ith iteration of the Marqdaurt-Levenberg algorithm, and $\mathbf{a}_{i+1}$ and $\mathbf{b}_{i+1}$ be values of **a** and **b** obtained for the i+1 th iteration of the algorithm.

The Marqdaurt-Levenberg algorithms for **a** and **b** are:

$$\mathbf{a}_{i+1} = \mathbf{a}_i - \left(\mathbf{H}_{a,i} + \mathbf{D}_{a,i}\right)^{-1} \mathbf{g}_{a,i} \qquad (3.8\text{-}11)$$

$$\mathbf{b}_{i+1} = \mathbf{b}_i - \left(\mathbf{H}_{b,i} + \mathbf{D}_{b,i}\right)^{-1} \mathbf{g}_{b,i} \qquad (3.8\text{-}12)$$

In these equations $\mathbf{g}_{a,i}$ is the gradient of $L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$ with respect to **a** evaluated at $\mathbf{a} = \mathbf{a}_i$ and $\mathbf{b} = \mathbf{b}_i$, and $\mathbf{g}_{b,i}$ is the gradient of $L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$ with respect to **b** evaluated at $\mathbf{a} = \mathbf{a}_i$ and $\mathbf{b} = \mathbf{b}_i$, i.e.

$$\mathbf{g}_{a,i} = \mathrm{col}\left[\partial L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\} / \partial \mathbf{a}_j \mid \mathbf{a} = \mathbf{a}_i \text{ and } \mathbf{b} = \mathbf{b}_i ; j = 1, \ldots\ldots, n\right] \qquad (3.8\text{-}13)$$

$$\mathbf{g}_{b,i} = \mathrm{col}\left[\partial L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\} / \partial \mathbf{b}_j \mid \mathbf{a} = \mathbf{a}_i \text{ and } \mathbf{b} = \mathbf{b}_i ; j = 1, \ldots\ldots, n\right] \qquad (3.8\text{-}14)$$

$\mathbf{H}_{a,i}$ is the Hessian matrix of $L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$ with respect to **a** and $\mathbf{H}_{b,i}$ is the Hessian matrix of $L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{u}_B \mid \mathbf{z}\}$ with respect to **b**, i.e.

$$\mathrm{H}_{a,i} = \left\{\partial^2 L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\} / \partial \mathrm{a}_j \partial \mathrm{a}_m \mid \mathbf{a} = \mathbf{a}_i \text{ and } \mathbf{b} = \mathbf{b}_i ; j, m = 1, \ldots\ldots, n\right\} \qquad (3.8\text{-}15)$$

$$\mathrm{H}_{b,i} = \left\{\partial^2 L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\} / \partial \mathrm{b}_j \partial \mathrm{b}_m \mid \mathbf{a} = \mathbf{a}_i \text{ and } \mathbf{b} = \mathbf{b}_i ; j, m = 1, \ldots\ldots, n\right\} \qquad (3.8\text{-}16)$$

Finally, $\mathbf{D}_{a,i}$ and $\mathbf{D}_{b,i}$ are diagonal stabilisation matrices to ensure the invertibility of the matrices $(\mathbf{H}_{a,i} + \mathbf{D}_{a,i})$ and $(\mathbf{H}_{b,i} + \mathbf{D}_{b,i})$, respectively. These matrices must be positive define for their inverses to exist. The flow chart for this algorithm is shown in Figure 3.10.

If the matrices $(\mathbf{H}_{a,i} + \mathbf{D}_{a,i})$ and $(\mathbf{H}_{b,i} + \mathbf{D}_{b,i})$ matrices are invertible, then $\mathbf{a}_i$ and $\mathbf{b}_i$ are updated using the just described Marqdaurt-Levenberg algorithm. However, if this is not the case then a modified Marqdaurt-Levenberg algorithm is used:

$$\mathbf{a}_{i+1} = \mathbf{a}_i + \mathbf{D}_{a,i}^{-1}\mathbf{g}_{a,i} \qquad (3.8\text{-}17)$$

$$\mathbf{b}_{i+1} = \mathbf{b}_i + \mathbf{D}_{b,i}^{-1}\mathbf{g}_{b,i} \qquad (3.8\text{-}18)$$

Due to the diagonal structure of the matrices $\mathbf{D}_{a,i}$ and $\mathbf{D}_{b,i}$ computing their inverses is straightforward. Once $\mathbf{a}_{i+1}$ and $\mathbf{b}_{i+1}$ have been computed, the next step is to evaluate $L\{\mathbf{a}_{i+1}, \mathbf{b}_{i+1}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$ and compare it with $L\{\mathbf{a}_i, \mathbf{b}_i, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$. The objective is to maximise L so that if $L\{\mathbf{a}_{i+1}, \mathbf{b}_{i+1}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\} > L\{\mathbf{a}_i, \mathbf{b}_i, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$ then the i+1 iteration values of **a** and **b** replace the ith iteration values of **a** and **b**. However, if there is no increase in L from iteration i to i+1 then we reject $\mathbf{a}_{i+1}$ and $\mathbf{b}_{i+1}$ and generate new values for these quantities. This is achieved by increasing the elements of the diagonal stabilisation matrices, $\mathbf{D}_{a,i}$ and $\mathbf{D}_{b,i}$. This causes the direction of $(\mathbf{H}_{a,i} + \mathbf{D}_{a,i})$ and $(\mathbf{H}_{b,i} + \mathbf{D}_{b,i})$ to change, in addition the length of the resulting vectors change so the newly computed values of $\mathbf{a}_{i+1}$ and $\mathbf{b}_{i+1}$ will be closer to the previous $\mathbf{a}_i$ and $\mathbf{b}_i$. This testing and modification process continues until either L has increased or a set number of iterations has elapsed. If the latter occurs then $\mathbf{a}_i$ and $\mathbf{b}_i$ are accepted as the maximum-likelihood values, $\mathbf{a}^{ML}$ and $\mathbf{b}^{ML}$, of **a** and **b**.

When $\mathbf{a}_{i+1}$ and $\mathbf{b}_{i+1}$ are accepted there are two possible options; stop or continue. To choose between the two the difference between $L\{\mathbf{a}_{i+1}, \mathbf{b}_{i+1}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$ and $L\{\mathbf{a}_i, \mathbf{b}_i, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$ is examined. If $|\, L\{\mathbf{a}_{i+1}, \mathbf{b}_{i+1}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\} - L\{\mathbf{a}_i, \mathbf{b}_i, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}|$ is less than some small pre-set value then the algorithm is stopped and $\mathbf{a}_i$ and $\mathbf{b}_i$ are accepted as the maximum-likelihood values, $\mathbf{a}^{ML}$ and $\mathbf{b}^{ML}$, of **a** and **b**. If $|\, L\{\mathbf{a}_{i+1}, \mathbf{b}_{i+1}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\} - L\{\mathbf{a}_i, \mathbf{b}_i, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}|$ is larger that the chosen pre-set value, new values of $\mathbf{a}_{i+1}$ and $\mathbf{b}_{i+1}$ are computed using the Marqdaurt-Levenberg algorithm. For this the gradient vectors, Hessian matrices and Diagonal matrices require updating before $\mathbf{a}_{i+1}$ and $\mathbf{b}_{i+1}$ are calculated.

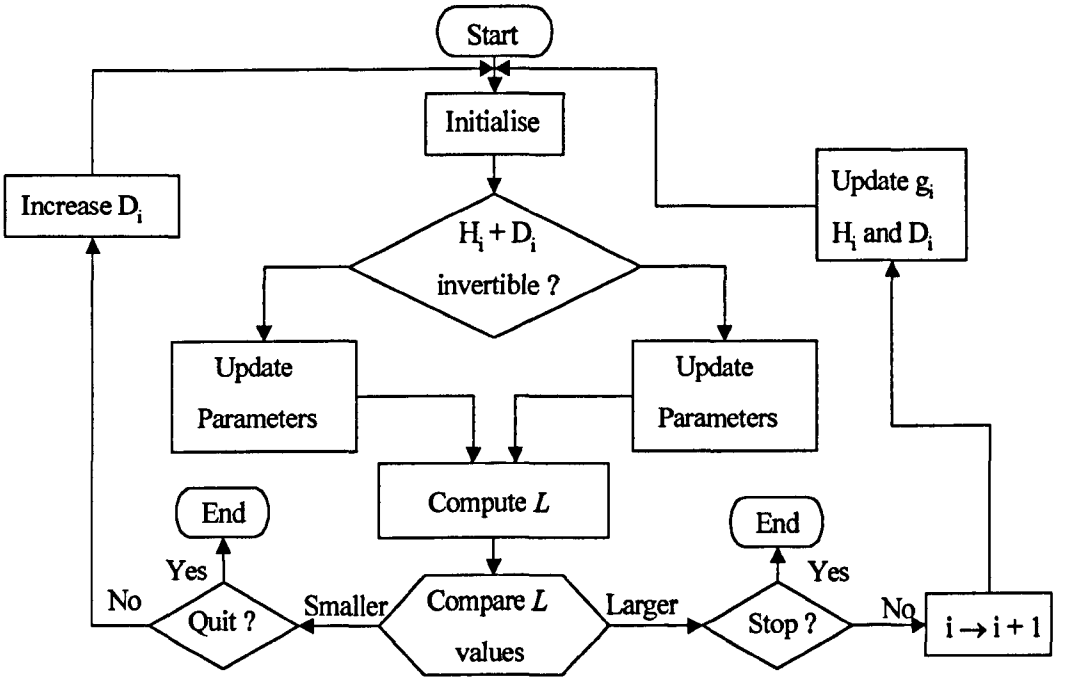


Figure 3.10. Marquardt-Levenberg algorithm.

3.8.5. Update Statistical Parameters.

There are four statistical parameters that require updating: v_r the variance of the amplitude sequence, v_B the amplitude of the backscatter sequence, v_n the variance of the measurement noise sequence and λ the average number of spikes in the event sequence. The three variances are always treated differently than the parameter λ.

The variances must always be positive, so optimising them is a constrained optimisation problem, this makes realisation of the optimisation task more difficult than a non-constrained optimisation problem. Therefore, the direct approach is not chosen, instead the standard deviations are optimised [34]. Standard deviations can be positive or negative, and represent the square root of the variances. This is permissible due to the fact that maximum-likelihood quantities have the following invariance property: Functions of maximum-likelihood quantities are themselves maximum-likelihood quantities.

Let the variances v_r, v_B and v_n be denoted by p_1, p_2 and p_3, respectively. The aforementioned invariance property allow the following conclusion about the variance parameters to be made,

$$v_r = \left(p_1^{ML}\right)^2, v_B = \left(p_2^{ML}\right)^2 \text{ and } v_n = \left(p_3^{ML}\right)^2 \qquad (3.8\text{-}18)$$

When L is to be maximised; p_1, p_2 and p_3 are all updated using a Marquardt-Levenberg algorithm. Let $\mathbf{p} = \text{col}(p_1, p_2 \text{ and } p_3)$. Then

$$\mathbf{p}_{i+1} = \mathbf{p}_i + \left(\mathbf{H}_{p,i} + \mathbf{D}_{p,i}\right)^{-1}\mathbf{g}_{p,i} \qquad (3.8\text{-}19)$$

where, $\mathbf{g}_{p,i}$ is the gradient of $L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$ with respect to the three elements of $\mathbf{s}$ contained in $\mathbf{p}$ evaluated at $\mathbf{p} = \mathbf{p}_i$. $\mathbf{H}_{a,i}$ is the Hessian matrix of $L\{\mathbf{a}, \mathbf{b}, \mathbf{s}, \mathbf{q}, \mathbf{r}, \mathbf{u}_B \mid \mathbf{z}\}$ with respect to $\mathbf{p}$ evaluated at $\mathbf{p} = \mathbf{p}_i$ and $\mathbf{D}_{p,I}$ is a diagonal stabilisation matrix, to insure the invertibility of $(\mathbf{H}_{p,i} + \mathbf{D}_{p,i})$. The flow chart in Figure 3.8 is also applicable here.

When both the wavelet and amplitude variance v_r are unknown, then v_r cannot be determined within a maximum-likelihood framework [34].

The statistical parameter, λ, is updated using a relaxation algorithm, that is

λ = (Number of elements in last detected event sequence that equal one, when λ_i)/N.

It is not necessary to update all four statistical parameters in unison, λ is usually updated after each detection stage, with the other statistical parameters, v_n and v_B, being updated together using the Marquardt-Levenberg optimisation algorithm.

3.9. Recursive Programming.

The ARMA model for the wavelet given in equation (3.4-2) is clearly recursive in nature. The recursive signal processing algorithms presented in this section exploits the recursive nature of this wavelet [34]. The first stage in the development of recursive processing algorithms is to develop a recursive wavelet model. The ARMA wavelet transfer function implies the following ARMA finite difference equation:

$$\begin{aligned} y(k+n)+a_1y(k+n-1)+.....+a_{n-1}y(k+1)+a_ny(k)=b_1u(k+n-1) \\ +b_2u(k+n-2)+.....+b_{n-1}u(k+1)+b_nu(k) \end{aligned} \qquad (3.9\text{-}1)$$

This equation is an nth order finite difference equation. It can also be represented as a collection of n first-order, difference equations. These n first-order equations can then be collated into a more compact form known as a state equation. A state variable representation of the ARMA model will now be presented.

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \vdots \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_n \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_1(k) \\ \vdots \\ x_1(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} u(k) \qquad (3.9\text{-}2)$$

and

$$y(k)=(b_n,b_{n-1},.....,b_1)\mathbf{x}(k) \qquad (3.9\text{-}3)$$

Equation (3.9-2) is the state equation and equation (3.9-3) is the output equation, because y(k) is the output of the wavelet model. These equations can be expressed in a more compact form as,

$$\mathbf{x}(k+1)=\phi\mathbf{x}(k)+\gamma u(k) \qquad (3.9\text{-}4)$$

and

$$y(k) = \mathbf{h}'\mathbf{x}(k) \qquad (3.9\text{-}5)$$

In these equations; $\mathbf{x}(k)$ is the $n\times 1$ state vector; ϕ is an $n\times n$ state transition matrix; γ is a $n\times 1$ input distribution vector (it distributes the input into those equations in which the input should appear); and $\mathbf{h}$ is an $n\times 1$ observation vector. This state variable model provides exactly the same relationship between u(k) and y(k) as the convolutional model. These two equations are now used to form the basis of the recursive algorithms.

3.9.1. Recursive MVD algorithm.

MVD is at the heart of the entire Maximum-Likelihood Deconvolution procedure. There are a number of methods used to realise the MVD filter. The recursive algorithm is derived directly from mean squared estimation theory applied to the state variable model [34]. The state variable model used can be seen in Figure 3.6, in Figure 3.11. The wavelet model block of the state variable model is interpreted as the interconnection of a state equation and an output equation.

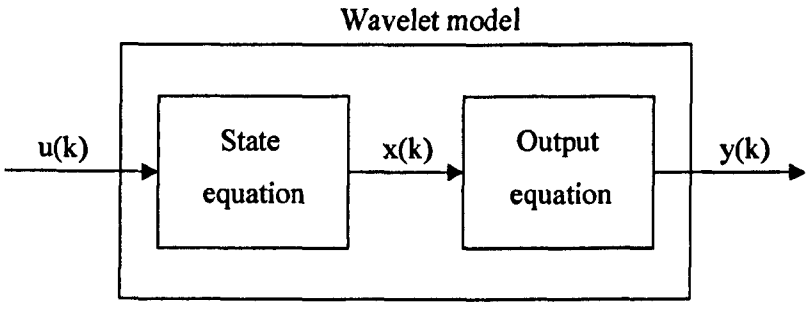


Figure 3.11. *Interpretation of the wavelet model as an interconnection of a state equation and output equation.*

The recursive MVD algorithm presented here has four parts and is detailed in Figure 3.12.

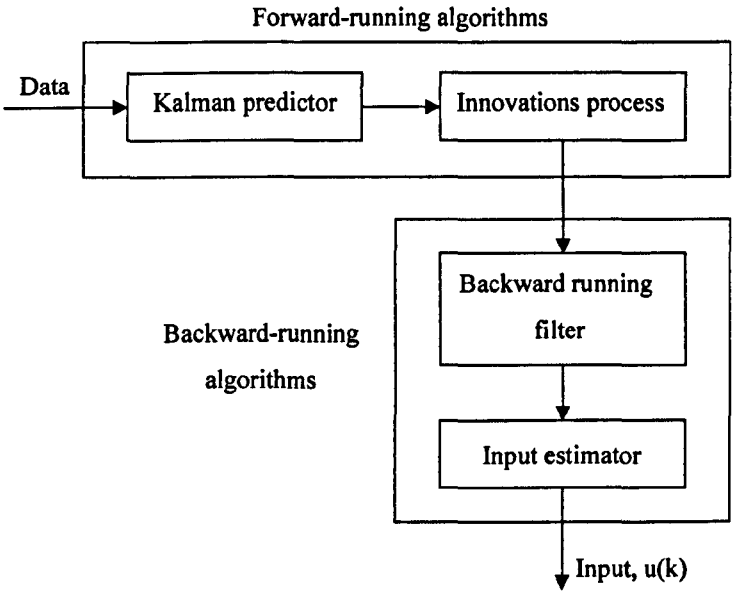


Figure 3.12. *Recursive MVD algorithm.*

3.9.1.1. Input Estimator.

Let $\mathbf{r}(j|N)$ denote an n×1 backward running (i.e. j = N,N-1,.....,0) state vector, and $\mathbf{S}(j|N)$ denote the state vectors associated n×n covariance matrix. Then

$$u^{MV}(k|N) = v_u \gamma' \mathbf{r}(k+1|N) \qquad (3.9\text{-}6)$$

$$\mathrm{Var}[u_{ERR}(k|N)] = v_u - v_u \gamma' \mathbf{S}(k+1|N) \gamma v_u \qquad (3.9\text{-}7)$$

where k = N,N-1,.....,0.

3.9.1.2. Backward-Running Filter.

The backward running equations for $\mathbf{r}(j|N)$ and $\mathbf{S}(j|N)$ are:

$$\mathbf{r}(j|N) = [\mathbf{I} - \mathbf{K}(j)\mathbf{h}']'\phi' \mathbf{r}(j+1|N) + \mathbf{h}\tilde{z}(j|j-1)/\mathrm{Var}[\tilde{z}(j|j-1)] \qquad (3.9\text{-}8)$$

and

$$\mathbf{S}(j|N)=[\mathbf{I}-\mathbf{K}(j)\mathbf{h}']'\phi'\mathbf{S}(j+1|N)\phi[\mathbf{I}-\mathbf{K}(j)\mathbf{h}']+\mathbf{h}\mathbf{h}'/\mathrm{Var}[\tilde{z}(j|j-1)] \qquad (3.9\text{-}9)$$

where j = N,N-1,.....,1; $\mathbf{r}(N+1|N) = \mathbf{0}$ and $\mathbf{S}(N+1|N) = \mathbf{0}$. In these equations $\tilde{z}(j|j-1)$ is a scalar process known as the innovations and $\mathbf{K}(k)$ is an n×1 Kalman gain matrix and $\mathbf{I}$ is the n×n identity matrix.

3.9.1.3. Innovations Process.

The innovations and its variance are the outputs of a forward-running Kalman predictor. They are given by the following equations:

$$\tilde{z}(k+1|k)=z(k+1)-\mathbf{h}'\hat{\mathbf{x}}(k+1|k) \qquad (3.9\text{-}10)$$

and

$$\mathrm{Var}[\tilde{z}(k+1|k)]=\mathbf{h}'\mathbf{P}(k+1|k)\mathbf{h}+v_n \qquad (3.9\text{-}11)$$

where k = 0,1,.....N-1; $\hat{\mathbf{x}}(k+1|k)$ denotes a n×1 mean squared prediction of state vector $\mathbf{x}(k+1)$, based on the measurements z(1),z(2),...,z(k); and $\mathbf{P}(k+1|k)$ is the n×n covariance matrix that describes the estimation error between $\hat{\mathbf{x}}(k+1|k)$ and $\mathbf{x}(k+1)$. Quantities $\hat{\mathbf{x}}(k+1|k)$ and $\mathbf{P}(k+1|k)$ are generated by the Kalman Predictor.

3.9.1.4. Kalman Predictor.

Let $\hat{\mathbf{x}}(k|k)$ and $\mathbf{P}(k|k)$ denote a n×1 mean squared filtered estimate of state vector $\mathbf{x}(k)$, that is based on the measurements z(1),z(2),...,z(k), and its associated n×n covariance matrix. $\mathbf{P}(k|k)$ describes the estimation error between $\hat{\mathbf{x}}(k|k)$ and $\mathbf{x}(k)$.

Then

$$\hat{\mathbf{x}}(k+1|k) = \phi\hat{\mathbf{x}}(k|k) \qquad (3.9\text{-}12)$$

and

$$\mathbf{P}(k+1|k) = \phi\mathbf{P}(k|k)\phi' + v_u\gamma\gamma' \qquad (3.9\text{-}13)$$

where

$$\hat{\mathbf{x}}(k+1|k+1) = [\mathbf{I} - \mathbf{K}(k+1)\mathbf{h}']\hat{\mathbf{x}}(k+1|k) + \mathbf{K}(k+1)z(k+1) \qquad (3.9\text{-}14)$$

$$\mathbf{K}(k+1) = \mathbf{P}(k+1|k)\mathbf{h}[\mathbf{h}'\mathbf{P}(k+1|k)\mathbf{h} + v_n]^{-1} \qquad (3.9\text{-}15)$$

$$\mathbf{P}(k+1|k+1) = [\mathbf{I} - \mathbf{K}(k+1)\mathbf{h}']\mathbf{P}(k+1|k) \qquad (3.9\text{-}16)$$

and k = 0,1,.....,N - 1. The Kalman predictor can be initialised by $\hat{\mathbf{x}}(0|0) = \mathbf{0}$ and $\mathbf{P}(0|0) = \mathbf{0}$.

3.10. ARMA Parameter Estimation.

The problem here is to estimate the parameters of the ARMA wavelet model from noisy observations of its output in response to excitation by an unknown independent identically distributed (i.i.d.) sequence [43][45][46][47].

3.10.1. Model and Assumptions.

The observed time series is modelled as the output of an ARMA system that is excited by an unobservable input [47]. The output is assumed to be corrupted by additive white Gaussian noise.

The time series is described by

$$\sum_{j=1}^{n} a(k)y(k-j) = \sum_{j=1}^{n} b(k)u(k-j) \qquad (3.10\text{-}1)$$

$$z(k) = y(k) + n(k) \qquad (3.10\text{-}2)$$

where u(k) is the system input, y(k) is the noiseless system output, n(k) is the additive white Gaussian noise and z(k) is the observed output noisy time series. The assumptions made are that the model order n is known; u(k) is an observable i.i.d. non-Gaussian process with at least one finite, non-zero cumulant $\gamma_{m,u}, m > 2$; the additive noise n(k) is independent of the input u(k), and hence, the output y(k); the additive noise n(k) is a coloured Gaussian process, of unknown power spectral density; the system is casual and exponentially stable; and a(0) and b(0) = 1, this fixes the inherent scale ambiguity.

Since y(k) and n(k) are independent, the cumulant of z(k) is the sum of the y(k) and the n(k) cumulant, due to the fact that n(k) is Gaussian its cumulants equal zero for orders greater than 2. This indicates that the cumulants of the noisy output z(k) are equal to cumulants of the output y(k).

Impulse response h(k) of the model satisfies the recursion

$$\sum_{j=0}^{n} a(k)h(k-j) = \sum_{j=0}^{n} b(k)\delta(j-k) = b(k) \qquad (3.10\text{-}3)$$

The mth order cumulants, that are (m - 1) dimensional sequences, are related to the input response by

$$C_{m,y}(t_1, t_2, t_3, \ldots\ldots, t_{m-1}) = \gamma_{m,u} \sum_{i=0}^{\infty} h(i)h(i+t_1)\ldots\ldots h(i+t_{m-1}) \qquad (3.10\text{-}4)$$

$$\sum_{j=0}^{n} a(k)C_{m,y}(\tau_1 - j, \tau_1 - \tau_2, \ldots\ldots, \tau_1 - \tau_{m-1}) = \gamma_{m,u} \sum_{j=0}^{n} b(k) \prod_{i=1}^{m-1} h(k-\tau_i) \qquad (3.10\text{-}5)$$

With $\tau = \tau_1, t = \tau_1 - \tau_2, \tau_3 = ... = \tau_{m-1} = \tau_1$ in (3.10-5), the following is obtained

$$f_m(t;\tau) := \sum_{j=0}^{n} a(j) C_{m,y}(\tau - j, t, 0, ..., 0)$$
$$= \gamma_{m,u} \sum_{j=0}^{n} b(j) h(j - \tau + t) h^{m-2}(j - \tau) \qquad \text{(3.10-6a and b)}$$

Equations (3.10-6a) and (3.10-6b) are the basis of the MA parameter estimation algorithm presented in Section 3.10.3. These algorithms assume that values of the AR parameters are available. An algorithm for obtaining estimations of the AR parameters will now be presented.

3.10.2. AR Parameter Estimation.

3.10.2.1. Model Order Determination.

The problem of estimating the order of the ARMA model will now be discussed. Whether exact values of the autocorrelation lag or cumulants are available or sampled estimates of these quantities are available, determination of the model order, n, remains an open question [43]. Here the method that will be considered makes uses of Singular Value Decomposition (SVD) on an extended autocorrelation matrix. A useful by-product of this process will be an AR parameter estimation procedure, which has been shown to give good spectral estimation performance [43].

When the ARMA model order is not known a priori then the usual practice is to select an order, n_e, which is much larger than the anticipated order of the model. Although n is usually not known a priori, it is generally possible to make an educated guess about the value of n such that $n_e > n$.

It has been shown that for a ARMA (n_e, n_e) model the $t \times (n_e+1)$ extended order autocorrelation matrix, where t is an integer value that is greater than $n_e + 1$, although larger values of t tend to lead to better model representations.

The extended order autocorrelation matrix may be expressed as

$$R_e = \begin{bmatrix} r_x(n_e+1) & r_x(n_e) & \cdots & r_x(1) \\ r_x(n_e+2) & r_x(n_e+1) & \cdots & r_x(2) \\ \vdots & \vdots & \cdots & \vdots \\ r_x(n_e+t) & r_x(n_e+t-1) & \cdots & r_x(t) \end{bmatrix} \quad (3.10\text{-}6)$$

If the autocorrelation lag entries, $r_x(k)$, correspond to the ARMA (n,n) model for which $n_e \geq n$, then it follows that the rank of the extended order matrix, R_e, will be n. In practice these autocorrelation lag entries are replaced by autocorrelation lag estimates, $\hat{r}_x(k)$, characterised by the time series. These are given by

$$\hat{r}_x(k) = \frac{1}{N}\sum_{j=1}^{N} x(j+k)\overline{x}(j) \quad (3.10\text{-}7)$$

where $\overline{\ }$ denotes the operation of complex conjugate. These sampled estimates will inevitably be in error compared to the actual autocorrelation lag values, it follows that the matrix $\hat{R}_e$ formed using the sampled estimates will be of full rank (i.e. $\min(t,n_e+1)$) even when the time series corresponds to an ARMA (n,n) process. Nonetheless, even though $\hat{R}_e$ will be of full rank it will still tend to have an effective rank of n. Effective rank will be quantified by introducing the concept of singular value decomposition.

3.10.2.2. Singular Value Decomposition (SVD).

In many applications the primary use of SVD is to solve sets of linear equations [43][48]. The matrix associated with the system of equation also conveys dynamical property information. With this in mind for determining the ARMA model order, it is advantageous to examine the salient properties of this characterising matrix. The following theorem for Singular Value Deconvolution fits this purpose particularly well.

Let A be a m×p matrix of generally complex valued elements. Then there exists m×p and p×p matrices U and V, respectively such that

$$A = U\Sigma V' \qquad (3.10\text{-}8)$$

where Σ is a p×p matrix whose elements are zeros except along the leading diagonal. These non-negative elements are ordered such that

$$\sigma_{11} \geq \sigma_{22} \geq \ldots\ldots \sigma_{hh} \geq 0 \qquad (3.10\text{-}9)$$

where h = min(m,p)

The diagonal elements of σ_{kk} are commonly referred to as the singular values of the matrix A. These singular values convey valuable information about the rank characterisation of A. This becomes apparent upon examining the problem of determining the m×p matrix of rank k, which is the best approximation of A in the Frobenious norm sense. The Frobenious norm of the m×p matrix difference A – B is defined to be

$$\|A - B\| = \left[\sum_{i=1}^{m}\sum_{j=1}^{p}|a_{ij} - b_{ij}|^2\right]^{1/2} \qquad (3.10\text{-}10)$$

Now the m×p matrix, B, of rank k that renders the Frobenious norm a minimum is sought. The solution to this approximation problem is outlined in the following theorem. The unique m×p matrix of rank k, where rank k ≤ rank A which best approximates A in the Frobenious norm sense is given by

$$A^{(k)} = U\Sigma_k V' \qquad (3.10\text{-}11)$$

where U and V are as specified previously and Σ_k is obtained from Σ by setting all the singular values to zero that are less than the kth singular value. The quality of this optimum approximation is given by

$$\left\|A - A^{(k)}\right\| = \left[\sum_{j=k+1}^{h} \sigma_{ij}^{1/2}\right] \quad ,1 \le k \le h. \qquad (3.10\text{-}12)$$

The degree to which $A^{(k)}$ approximates A is dependent on the sum of the (h - k) smallest singular values squared. In order to provide a convenient measure for this behaviour that is independent of the size of A, consider the normalised ratio

$$\begin{aligned} v(k) &= \frac{\left\|A^{(k)}\right\|}{\|A\|} \\ &= \left[\frac{\sigma_{11}^2 + \sigma_{22}^2 + \ldots\ldots + \sigma_{kk}^2}{\sigma_{11}^2 + \sigma_{22}^2 + \ldots\ldots + \sigma_{hh}^2}\right]^{1/2}, \quad 1 \le k \le h. \end{aligned} \qquad (3.10\text{-}13)$$

Clearly this ratio approaches 1 as k approaches h. For matrices of low effective rank v(k) is close to one for values of k significantly lower than h. On the other hand matrices of high effective rank v(k) approaches one as k approaches h. To determine the order of an ARMA model, the $t \times n_e+1$ autocorrelation matrix is used. That is,

$$\hat{R}_e = U \Sigma V' \qquad (3.10\text{-}14)$$

where U and V are t×t and $(n_e+1)\times(n_e+1)$ unitary matrices respectively, and Σ is a $t\times(n_e+1)$ matrix as outlined in equation (3.10-9). The required AR order n is obtained be examining the normalised ration v(k). The value of n is chosen to be the smallest value of k for which v(k) is deemed adequately close to one. This criteria is subjective and will depend on the particular application. The result of this will be a rank n optimum approximation of the $t\times(n_e+1)$ matrix, R_e that is,

$$\hat{R}_e^{(n)} = U \Sigma_n V' \qquad (3.10\text{-}15)$$

where Σ_n is Σ with all but its largest n singular value set to zero. Using matrix manipulation it can be shown that this rank n approximation may be represented as:

$$\hat{R}_e^{(n)} = \sum_{j=1}^{n} \sigma_{jj} u_j v_j' \qquad (3.10\text{-}16)$$

where u_k and v_k are the kth column vectors of the t×t and $(n_e+1)\times(n_e+1)$ unitary matrices U and V respectively. A procedure for estimating the AR parameters from this rank n autocorrelation matrix will now be presented [43]. For this approach, the rank n approximation is interpreted as providing an improved estimate of the underlying autocorrelation matrix. It is convenient to decompose the rank n approximation matrix, as follows;

$$\hat{R}_e^{(n)} = \left[\hat{r}_1^{(n)} : \hat{R}_a^{(n)}\right] \qquad (3.10\text{-}17)$$

where $\hat{r}_1^{(n)}$ is the leftmost t×1 column vector of $\hat{R}_e^{(n)}$ and $\hat{R}_a^{(n)}$ is a $t\times(n_e+1)$matrix composed of the n_e rightmost t×1 column vectors of $\hat{R}_e^{(n)}$. The task now is to find a $(n_e+1)\times1$ AR parameter vector **a**, with first component one, that satisfies the relationship

$$R_e^{(n)} a = \theta \qquad (3.10\text{-}18)$$

where $\mathbf{a} = [1, a_1, a_2, \ldots, a_n]$ and θ denotes a t×1 zero vector.

Since the rank of $\hat{R}_e^{(n)}$ is less than full, there will be an infinite number of possible solutions. The minimum norm solution will be chosen, as specified by

$$\begin{bmatrix} a_1^0 \\ a_2^0 \\ \vdots \\ a_{n_e}^0 \end{bmatrix} = -\left[\hat{R}_a^{(p)}\right]^{\#} \hat{r}_1^{(p)} \qquad (3.10\text{-}19)$$

In which the superscript # denotes the operation of pseudo matrix inversion. It is readily shown that the minimum norm solution can be simplified to

$$a^0 = \frac{e_1 - \sum_{k=1}^{n} \bar{v}_k(0) v_k}{1 - \sum_{k=1}^{n_e} |v_k(0)|^2} = \frac{\sum_{k=n+1}^{n_e} \bar{v}_k(0) v_k}{\sum_{k=n+1}^{n_e} |v_k(0)|^2} \qquad (3.10\text{-}20)$$

where v_k is the kth column vector of the unitary matrix V.

An alternative method is to use the best rank n approximation matrix (3.10-15) to estimate the AR parameters [43]. This matrix contains within its column structure the characteristics required to generate an estimate the AR parameters of the ARMA model. In particular the submatrices of the matrix $\hat{R}_e^{(n)}$ consisting of the matrix columns from column k through k+n inclusive will yield n rank approximations to the t×(n+1) autocorelation matrices R_k for $1 \le k \le n_e - n + 1$ as specified in (3.10-21)

R_k = [submatrix of R_e composed of its kth through n + kth column vectors inclusively] for $1 \le k \le n_e - n + 1$. (3.10-21)

The rank n approximations will be denoted by, $\hat{R}_e^{(n)}$. Due to the SVD operation and errors inherent in the generation of $\hat{R}_e$, there will generally not be a unique AR parameter vector (with the first term being unity) that will satisfy all of the n_e+n+1 estimates of (3.10-18). Nonetheless, it is still desirable to obtain an estimate for the AR parameters for which these relationships are almost satisfied. A function that measures the degree to which this is satisfied is given by,

$$f(\mathbf{a}) = \mathbf{a}^* S^{(n)} \mathbf{a} \qquad (3.10\text{-}22)$$

where

$$S^{(n)} = \sum_{k=1}^{n_e-n+1} \hat{R}_k^{(n)*} \hat{R}_k^{(n)} \quad (3.10\text{-}22a)$$

The (n+1)×(n+1) matrix $S^{(p)}$ is non-negative define Hermitian and may be conveniently computed using

$$S^{(n)} = \sum_{j=1}^{n} \sum_{k=1}^{n_e-n+1} \sigma_{jj}^2 v_j^k v_j^{k*} \quad (3.10\text{-}23)$$

in which v_j^k denotes the (n+1)×1 vector as specified by

$$v_j^k = \left[v_j(k), v_j(k+1), \ldots\ldots, v_j(k+n)\right] \quad \begin{matrix} 1 \le k \le n_e - n + 1 \\ 1 \le j \le p \end{matrix} \quad (3.10\text{-}24)$$

This vector represents a windowed segment of the nth column vector (v_j) of the unitary matrix V. Due to the nature of v_j^k and v_j^{k+1} it is possible to devise an iterative procedure for updating the (n+1)×(n+1) matrices $v_j^k v_j^{k*}$ as k evolves. This involves (n+1) computation for each chosen value of k.

Once the generation of $S^{(n)}$ is complete the next step is to select the AR parameter vector, **a**, with its first term equal to one so as to minimise (3.10-22). This constrained minimisation will result in the best least squares approximation of the theoretical relationship (3.10-18). Using standard procedures, the required optimum AR parameter vector is determined by solving the following system of linear algebraic equations:

$$S^{(n)}\mathbf{a} = \alpha \mathbf{e}_1 \quad (3.10\text{-}25)$$

where α is a normalising factor to assure the first term of the AR parameter vector, **a**, is one.

3.10.3. MA Parameter Estimation.

The MA parameter estimation algorithm that has been used to obtain the MA parameters once the AR parameters have been determined, is a recursive parameter estimation algorithm. The algorithm is outlined in Section 3.10.3.1.

3.10.3.1. Recursive MA parameter estimation.

In order to complete the model identification the MA parameters need to be estimated. This is done using a technique to generate MA model parameters and using Residual time series on the system output to take account of the fact that an ARMA model is being estimated [45]. These Residual time series are computed as follows:

$$\tilde{y}(k) = y(k) + \sum_{j=1}^{n} \hat{a}(j) y(k-j) \quad (3.10\text{-}32)$$

A two step algorithm can be then be used to generate the MA parameters of the ARMA model.

Step 1: From the output or residual time series $\tilde{y}(k)$, k = 1,2,....,N, form the estimates

$$\hat{r}(m) = \frac{1}{N} \sum_{k=0}^{N-m} \tilde{y}(k)\tilde{y}(n+m); \quad m = 0,1,.....,n \quad (3.10\text{-}33)$$

and

$$\hat{c}(m) = \frac{1}{N} \sum_{k=\max(0,-m)}^{\min(N,N-m)} \tilde{y}(k)\tilde{y}^2(n+m); \quad m = -n,...,0,...,n. \quad (3.10\text{-}34)$$

Step 2: Using the estimates $\hat{r}(m)$ and $\hat{c}(m)$, iteratively solve the system of equations given below for m = -n,.....,0,.....,n to obtain $\{b(k)\}_{k=1}^{n}$, σ^2 and γ. This will now be outlined.

The approach used to solve this system of equations is shown in the flowchart shown in figure 3.13 [45]. All of the equations needed for the algorithm are stated below in equations (3.10-35) to (3.10-40).

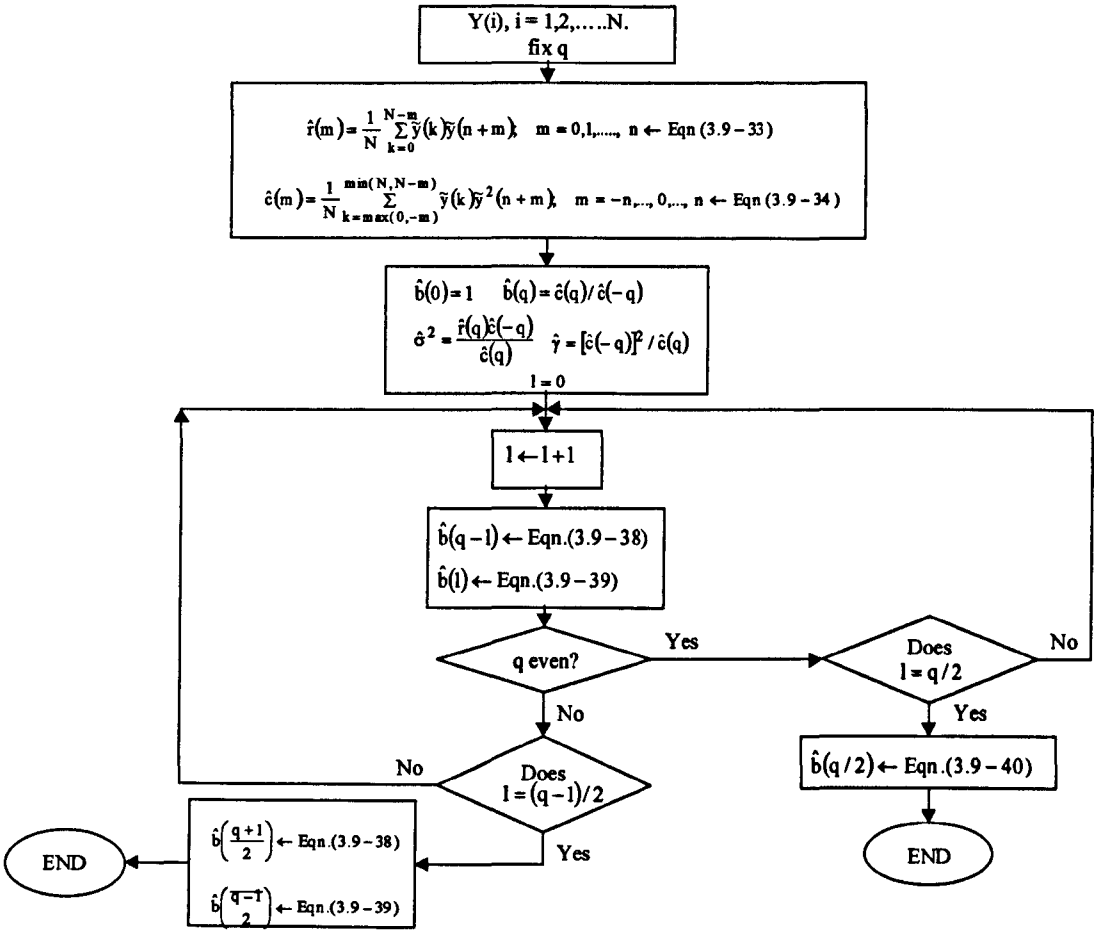


Figure 3.13. *Flow chart for recursive MA parameter estimation algorithm.*

Let

$$s(l-1)=\sum_{k=1}^{l-1}b(k)b(n-l+k) \quad (3.10\text{-}35)$$

$$s_{12}(l-1)=\sum_{k=1}^{l-1}b(k)b^2(n-l+k) \quad (3.10\text{-}36)$$

and

$$s_{21}(l-1)=\sum_{k=1}^{l-1}b^2(k)b(n-l+k) \quad (3.10\text{-}37)$$

Then

$$b(n-l)=N(n-l)/D(n-l) \quad (3.10\text{-}38a)$$

where

$$\begin{aligned}N(n-l)=\;& r(n)\left[r(n)c(-n+l)-\sigma^2c(n-l)\right]\\ &-\gamma r(n-l)\left[r(n-l)-r(n)\right]b(n)\\ &-\gamma r(n)\left[\begin{array}{l}r(n)s_{21}(l-1)-\sigma^2s_{12}(l-1)+\\ \sigma^2s^2(l-1)-r(n)s(l-1)-2r(n-l)s(l-1)\end{array}\right]\end{aligned} \quad (3.10\text{-}38b)$$

and

$$D(n-l)=2\gamma r(n)\left[\sigma^2s(l-1)-r(n-l)\right] \quad (3.10\text{-}38c)$$

The coefficient b(l) is computed using

$$b(l)=\frac{r(n-l)-\sigma^2\left[b(n-l)+s(l-1)\right]}{r(n)} \quad (3.10\text{-}39)$$

When n is even the middle coefficient is computed using

$$b(n/2)=\left[r(n/2)-\sigma^2\sum_{k=1}^{n/2}b(k)b(n/2+k)\right]/\sigma^2 \qquad (3.10\text{-}40)$$

3.11. Minimum Variance Deconvolution Filter.

The purpose of the proceeding system parameter estimation algorithms is to provide the information to implement the MVD filter on the noisy system output.

The MVD filter is, in general a time varying digital filter [34][42]. In order to develop some insight into the operation of the filter it will be assumed to be a constant coefficient filter. This assumption leads to a different derivation of the filter from the case where the coefficients are not constant. Consider the situation shown in figure 3.14 below, where the data, z(k), is processed by a two sided digital filter. During the design of this filter its output, y(k), is compared to the desired output, d(k), and the filter's coefficients chosen to minimise the mean squared error between d(k) and y(k). In the design of a deconvolution filter, d(k) is chosen to be the input u(k).

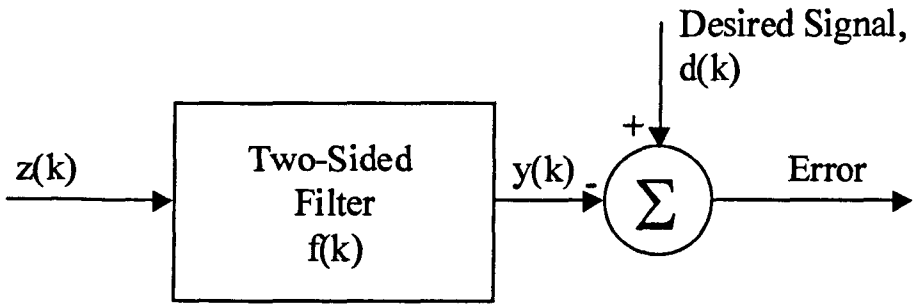


Figure 3.14. *Two Sided MVD digital filter.*

The filter output can be expressed as

$$y(k)=f(k)*z(k)=\sum_{i=-\infty}^{\infty}f(i)z(k-i)=...+f(-2)z(k+2)+f(-1)z(k+1)+f(0)z(k)+f(1)z(k-1)+f(2)z(k-2)+... \qquad (3.11\text{-}1)$$

The filter considered here is a doubly infinite two sided filter, this means the filter is described by the coefficients f(0), f(±1), f(±2), ...

An equation for these filter coefficients will now be derived. The filters coefficients are chosen to minimise the mean squared error, I(**f**), between the filters output, y(k), and the desired signal u(k), i.e. **f** = col[f(0), f(±1), f(±2),.....] is chosen to minimise I(**f**) given in equation (3.11-2).

$$I(\mathbf{f})=E\left\{\sum_{k=0}^{\infty}\left[u(k)-\sum_{m=-\infty}^{\infty}f(m)z(k-m)\right]^2\right\} \qquad (3.11\text{-}2)$$

Taking the partial derivative of I(**f**) with respect to the jth filter coefficient, F(j), interchanging the order of expectation and summation, and setting the result equal to zero, it can be shown that

$$\sum_{m=-\infty}^{\infty}f(m)\phi_{zz}(j-m)=\phi_{zu}(m), \quad j=0,\pm1,\pm2,..... \qquad (3.11\text{-}3)$$

This equation is now solved to obtain an expression for the filters coefficients. In order to achieve this the discrete-time Fourier transform must be taken, i.e.

$$F(\omega)\phi_{zz}(\omega)=\phi_{zu}(\omega) \qquad (3.11\text{-}4)$$

The next step is to determine $\phi_{zz}(\omega)$ and $\phi_{zu}(\omega)$. Initially, given by $\phi_{zu}(m)$ and $\phi_{zz}(j)$. In order to evaluate $\phi_{zu}(m)$ the convolutional model given in equation (3.6-1) needs to be utilised, i.e.

$$\begin{aligned}\phi_{zu}(m) &= E\{z(k-m)u(k)\} = \sum_{i=1}^{k} E\{u(i)u(k)\}w(k-m-i) + E\{n(k-m)u(k)\} \\ &= v_u w(-m)\end{aligned} \qquad (3.11\text{-}5)$$

because

$$E\{u(i)u(k)\} = v_u \delta(i-k) \qquad (3.11\text{-}6)$$

and

$$E\{n(k-m)u(k)\} = E\{n(k-m)u(k)\} = E\{n(k-m)\}E\{u(k)\} \qquad (3.11\text{-}7)$$

To evaluate $\phi_{zz}(j)$ the starting point is again the convolutional model, but in order to simplify the limits of summation the convolutional model is embedded.

$$z(k) = \sum_{i=1}^{k} u(i)w(k-i) + n(k), \quad k = 1,2,.....,N \qquad (3.11\text{-}8)$$

is transformed into the two sided model

$$z(k) = \sum_{i=-\infty}^{\infty} u(i)w(k-i) + n(k), \quad k = 1,2,.....,N \qquad (3.11\text{-}9)$$

For casual systems and inputs, the second reduces to the first form. Using the fact that u(k) and n(k) are stationary, white, uncorrelated, and of zero mean. Hence,

$$\phi_{zz} = \mathbf{E}\{z(k)z(k-j)\} = \mathbf{E}\left\{\sum_{i,l=-\infty}^{\infty}\sum u(i)u(l)w(k-i)w(k-1+j)\right\}$$
$$+\mathbf{E}\{n(k)n(k-j)\}$$
$$= \sum_{i,l=-\infty}^{\infty}\sum v_u\delta(i-l)w(k-i)w(k-1+j)+v_n\delta(j) \qquad (3.11\text{-}10)$$
$$= v_u\sum_{i=-\infty}^{\infty}w(k-i)w(k-i+j)+v_n\delta(j)$$
$$= v_u\sum_{n=-\infty}^{\infty}w(n)w(n+j)+v_n\delta(j)$$

which is the desired result.

From (3.11-5), it can be shown that

$$\phi_{zu} = F\{\phi_{zu}(m)\} = F\{v_u w(-m)\} = v_u\overline{W}(\omega) \qquad (3.11\text{-}11)$$

and from (3.11-10) it can be shown that

$$\phi_{zz}(\omega) = F\{\phi_{zz}(m)\} = F\left\{\left[v_u\sum_{n=-\infty}^{\infty}w(n)w(n+j)+v_n\delta(j)\right]\right\} \qquad (3.11\text{-}12)$$
$$= v_u W(\omega)W(-\omega)+v_n$$

but, W(-ω) is the complex conjugate of W(ω), so

$$\phi_{zz}(\omega) = v_u\left|W(\omega)^2\right|+v_n \qquad (3.11\text{-}13)$$

substitute (3.11-12) and (3.11-13) into (3.11-4) , to obtain the result below,

$$F(\omega) = v_u\overline{W}(\omega)/\left[v_u|W(\omega)|^2+v_n\right] \qquad (3.11\text{-}14)$$

3.11.1. Resolution Function.

The resolution function, $\rho(k)$, for this deconvolution filter is obtained by convolving the filter's input, $z(k)$, with $f(k)$, i.e.

$$\begin{aligned} y(k) &= f(k) * z(k) = [f(k) * w(k)] * u(k) + f(k) * n(k) \\ &= \rho(k) * u(k) + f(k) * n(k) \end{aligned} \qquad (3.11\text{-}15)$$

where

$$\rho(k) = f(k) * w(k) \qquad (3.11\text{-}16)$$

and using the fact that $R(\omega) = F(\omega)W(\omega)$ and equation (3.11-14).

$$R(\omega) = \left| W(\omega) \right|^2 v_u / v_n \Big/ \left| W(\omega) \right|^2 v_u / v_n \Big| \qquad (3.11\text{-}17)$$

This resolution function is a zero phase function. This means that the noise free portion of the deconvolution filter's output, $\rho(k) * u(k)$, is a zero waveshaped version of u(k). Therefore, if the input is a stationary white sequence and the deconvolution filter is two sided doubly infinite, then MVD zero waveshapes as well as deconvolves.

SNR is directly proportional to the ration v_u/v_n. From equation (3.11-17) it is evident that as v_u/v_n approaches ∞, $R(\omega)$ approaches one, so that the resolution function, $\rho(k)$, approaches the unit spike function. Thus, for signals with a high SNR the noise free portion of the deconvolution filter's output approaches u(k). This represents perfect resolution. Also when $|W(k)|^2 v_u / v_n >> 1$, $R(\omega)$ again approaches unity, so that once again perfect resolution is obtained. Broadband wavelets often satisfy this condition. The inequality $|W(k)|^2 v_u / v_n >> 1$ clearly demonstrates the interrelationship between SNR and bandwidth.

3.12. Determination of Hessian matrices and gradients for the Mardquart-Levenberg algorithm.

3.12.1. Gradients of L with respect to a and b.

A nine step algorithm is used to compute the gradients of L with respect to **a** and **b** [34]. Each of the steps in the algorithms will occur for each iteration, i. The algorithm is as follows:

1. Initialise $\mathbf{a}_i$ and $\mathbf{b}_i$. When i = 0 the initial values of **a** and **b** must be provided. For all other values of i, $\mathbf{a}_i$ and $\mathbf{b}_i$ will be available from the most recent iteration of the Mardquardt-Levenberg algorithm.

2. Compute the matrix $W=W(a_i,b_i)$. This is achieved by solving the finite difference equation

$$\begin{aligned} &w(k+n)+a_1w(k+n-1)+.....+a_{n-1}w(k+1)+a_nw(k) \\ &\quad =b_1\delta(k+n-1)+b_2\delta(k+n-2)+.....+b_{n-1}\delta(k+1) \\ &\quad +b_n\delta(k) \end{aligned} \qquad (3.12\text{-}1)$$

for w(0),w(1), ...,w(N-1), where w(-n) = w(-n+1) = ... = w(-1) = 0. This equation can also be written as

$$A(z)w(k)=B(z)\delta(k) \qquad (3.12\text{-}2)$$

where

$$A(z)=\left(z^n+a_1z^{n-1}+a_2z^{n-2}+.....+a_{n-1}z+a_n\right) \qquad (3.12\text{-}3)$$

and

$$B(z)=\left(b_1z^{n-1}+b_2z^{n-2}+.....+b_{n-1}z+b_n\right) \qquad (3.12\text{-}4)$$

it follows from Section 3.4 that

$$W(z)=B(z)/A(z)=\left(b_1z^{n-1}+b_2z^{n-2}+.....+b_{n-1}z+b_n\right)/\left(z^n+a_1z^{n-1}+a_2z^{n-2}+.....+a_{n-1}z+a_n\right) \qquad (3.12\text{-}5)$$

and from the fact that a transfer function, such as W(z), is the z-transform of a systems impulse response.

3. Compute $\partial w(k)/\partial a_1$ for k = 1,2,.....,N. This is done by solving the linear finite difference equation

$$A(z)[\partial w(k)/\partial a_1]=-w(k+n-1) \qquad (3.12\text{-}6)$$

Equation (3.12-6) was obtained by taking the partial derivative of (3.12-1) with respect to $\mathbf{a}_1$. Observing that the right hand side of equation (3.12-6) is generated from the results obtained in step 2. Letting

$$s_{a1}(k)=\partial w(k)/\partial a_1 \qquad (3.12\text{-}7)$$

Equation (3.12-2) can be expressed as

$$s_{a1}(k+n)+a_1s_{a1}(k+n-1)+.....+a_ns_{a1}(k)=-w(k+n-1) \qquad (3.12\text{-}8)$$

This equation clearly shows the linear nature of $\partial w(k)/\partial a_1$. With the term $s_{a1}(k)$ denoting the partial derivative, $\partial w(k)/\partial a_1$, this partial derivative is often referred to as a sensitivity function. Since w(-n) = w(-n+1) = ... = w(-1) = 0, $s_{a1}(-n)$ = $s_{a1}(-n+1)$ = ... = $s_{a1}(-1)$ = 0.

4. Compute $\partial w(k)/\partial a_j$ for j = 2,3,.....,n and k = 1,2,.....,N. This is achieved using equation (3.12-9)

$$\partial w(k)/\partial a_j = \partial w(k-j+1)/\partial a_1 \qquad \text{for all j and k} \qquad (3.12\text{-}9)$$

This equation is obtained by taking the partial derivatives of equation (3.12-1) with respect to a_j, giving

$$A(z)[\partial w(k)/\partial a_j] = -w(k+n-j) \qquad (3.12\text{-}10)$$

then set $k = k-j+1$ in equation (3.12-6), i.e.

$$A(z)[\partial w(k-j+1)/\partial a_1] = -w(k+n-j) \qquad (3.12\text{-}11)$$

Comparing equations (3.12-10) and (3.12-11) the result shown in equation (3.12-9) is self evident.

5. Compute $\partial w(k)/\partial b_1$ for k = 1,2,.....,N. This is done by solving the linear difference equation

$$A(z)[\partial w(k)/\partial b_1] = -\delta(k+n-1) \qquad (3.12\text{-}12)$$

which was obtained by taking the partial derivative of equation (3.12-1) with respect to $\mathbf{b}_1$.

6. Compute $\partial w(k)/\partial b_j$ for j = 2,3,.....,n and k = 1,2,.....,N. This is done using

$$\partial w(k)/\partial b_j = \partial w(k-j+1)/\partial b_1 \qquad (3.12\text{-}13)$$

7. Construct the 2n N×N matrices $\partial \mathbf{W}/\partial a_j$ and $\partial \mathbf{W}/\partial b_j$ for j =1,2,.....,N. Recall that the matrix **W**, is as follows

$$\mathbf{W}=\begin{bmatrix} w(0) & 0 & \cdots & 0 \\ w(1) & w(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ w(N-1) & w(N-2) & \cdots & w(0) \end{bmatrix}$$

then taking the partial derivative with respect to a_j, i.e.

$$\partial\mathbf{W}/\partial a_j=\begin{bmatrix} \partial w(0)/\partial a_j & 0 & \cdots & 0 \\ \partial w(1)/\partial a_j & \partial w(0)/\partial a_j & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \partial w(N-1)/\partial a_j & \partial w(N-2)/\partial a_j & \cdots & \partial w(0)/\partial a_j \end{bmatrix} \quad (3.12\text{-}14)$$

A similar equation exists for $\partial\mathbf{W}/\partial b_j$.

8. Compute $\partial\mathbf{W}'\mathbf{W}/\partial a_j$ and $\partial\mathbf{W}'\mathbf{W}/\partial b_j$ for j = 1,2,.....,N. This is accomplished using the formulas

$$\partial\mathbf{W}'\mathbf{W}/\partial a_j=[\partial\mathbf{W}'/\partial a_j]\mathbf{W}+\mathbf{W}'[\partial\mathbf{W}'/\partial a_j] \quad (3.12\text{-}15)$$

and

$$\partial\mathbf{W}'\mathbf{W}/\partial b_j=[\partial\mathbf{W}'/\partial b_j]\mathbf{W}+\mathbf{W}'[\partial\mathbf{W}'/\partial b_j] \quad (3.12\text{-}16)$$

9. Compute $\partial L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B\,|\,\mathbf{z})/\partial a_j$ and $\partial L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B\,|\,\mathbf{z})/\partial b_j$. This is done using the equations

$$\begin{aligned}\partial L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B\,|\,\mathbf{z})/\partial a_j=&-\mathbf{z}'[\partial\mathbf{W}/\partial a_j]\mathbf{u}-\mathbf{u}'[\partial\mathbf{W}/\partial a_j]\mathbf{z}\\&+\mathbf{u}'[\partial\mathbf{W}'\mathbf{W}/\partial a_j]\end{aligned} \quad (3.12\text{-}17)$$

and

$$\partial L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{z})/\partial b_j = -\mathbf{z}'\left[\partial \mathbf{W}/\partial b_j\right]\mathbf{u} - \mathbf{u}'\left[\partial \mathbf{W}/\partial b_j\right]\mathbf{z} + \mathbf{u}'\left[\partial \mathbf{W}'\mathbf{W}/\partial b_j\right] \quad (3.12\text{-}18)$$

for j = 1,2,.....,n. These equations were obtained by taking the partial derivatives of equation (3.7-11) with respect to $\mathbf{a}_j$ and $\mathbf{b}_j$, respectively and letting

$$\mathbf{u} = \mathbf{Q}\mathbf{r} + \mathbf{u}_B \quad (3.12\text{-}19)$$

In equations (3.12-17) and (3.12-18) **u** is required, the only **u** available is the most recently computed MVD estimates. $\mathbf{u}^{MV}(N \mid q)$ is updated when the random parameters are updated in the block component search algorithm, and also when *L* is calculated.

3.12.2. Derivatives of L with respect to the variances.

When both the wavelet and the statistical parameters are unknown, there is an inherent scale ambiguity that cannot be resolved [34]. Hence it is not possible to estimate the variance of the amplitude sequence, v_r. For this reason no derivation of equations for v_r is undertaken. The derivatives with respect to the variance v_B and v_n are computed using a four-step algorithm.

1. Initialise $v_{B,i}$ and $v_{n,i}$. When i = 0 the initial values $v_{B,0}$ and $v_{n,0}$ must be provided, usually small values are chosen. For all other values of i $v_{B,i}$ and $v_{n,i}$ will be available from the most recent iteration of the Mardquardt-Levenberg algorithm.

2. Compute the matrix $\mathbf{W}=\mathbf{W}(\mathbf{a}_i,\mathbf{b}_i)$.

3. Compute $\mathbf{u}^{MV}(N \mid q)$. This is done using the recursive MVD algorithm outlined in Section 3.9.1.

4. Compute $\partial L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{z})/\partial v_B$ and $\partial L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{z})/\partial v_n$. This is accomplished using the following formulas

$$\partial L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{z})/\partial v_B = -(N/2v_B)+\mathbf{u}_B'\mathbf{u}_B/2v_B^2 \qquad (3.12\text{-}20)$$

and

$$\partial L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{z})/\partial v_n = -(N/2v_n)+(\mathbf{z}-\mathbf{W}\mathbf{u})'(\mathbf{z}-\mathbf{W}\mathbf{u})/2v_n^2 \qquad (3.12\text{-}21)$$

These equations being obtained by taking the partial derivative of equation (3.7-11) with respect to v_B and v_n, respectively, and letting $\mathbf{u} = \mathbf{Q}\mathbf{r} + \mathbf{u}_B$ as before. In order to calculate $\partial L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{z})/\partial v_n$ it is necessary to replace $\mathbf{u}$ by $\mathbf{u}^{MV}(N \mid q)$. To compute $\partial L(\mathbf{a},\mathbf{b},\mathbf{s},\mathbf{q},\mathbf{r},\mathbf{u}_B \mid \mathbf{z})/\partial v_B$ $\mathbf{u}_B$ must be replaced by $\mathbf{u}_B^{MV}(N \mid q)$. This is given by

$$\mathbf{u}_B^{MV}(N \mid q) = v_B(v_r\mathbf{Q} + v_B\mathbf{I})^{-1}\mathbf{u}^{MV}(N \mid q) \qquad (3.12\text{-}22)$$

Therefore, it is straight forward to compute $\mathbf{u}_B^{MV}(N \mid q)$ from $\mathbf{u}^{MV}(N \mid q)$.

3.12.3. Pseudo-Hessian Matrices.

In this section the procedures for computing the second derivatives of L with respect to the wavelet parameters $\mathbf{a}$ and $\mathbf{b}$ and the two variance v_r and v_B are described [34]. The pseudo-Hessian matrix represents the approximate second derivative of the variable. The pseudo-Hessian matrix is used in the Mardquardt-Levenberg algorithm in place of the Hessian matrices to save on computation time.

3.12.3.1. Pseudo-Hessian of L with respect to a and b.

Starting with equations (3.12-17) and (3.12-18) it is straightforward to show that

$$\begin{aligned}\partial^2 L/\partial a_i \partial a_j = &-\mathbf{z}'\left[\partial^2 \mathbf{W}/\partial a_i \partial a_j\right]\mathbf{u} - \mathbf{u}'\left[\partial^2 \mathbf{W}/\partial a_i \partial a_j\right]\mathbf{z} \\ &+ \mathbf{u}'\left[\partial^2 \mathbf{W}'\mathbf{W}/\partial a_i \partial a_j\right]\mathbf{u}\end{aligned} \tag{3.12-23}$$

and

$$\begin{aligned}\partial^2 L/\partial b_i \partial b_j = &-\mathbf{z}'\left[\partial^2 \mathbf{W}/\partial b_i \partial b_j\right]\mathbf{u} - \mathbf{u}'\left[\partial^2 \mathbf{W}/\partial b_i \partial b_j\right]\mathbf{z} \\ &+ \mathbf{u}'\left[\partial^2 \mathbf{W}'\mathbf{W}/\partial b_i \partial b_j\right]\mathbf{u}\end{aligned} \tag{3.12-24}$$

where from equations (3.11-15) and (3.11-16),

$$\begin{aligned}\partial \mathbf{W}'\mathbf{W}/\partial a_i a_j = &\left[\partial^2 \mathbf{W}'/\partial a_i \partial a_j\right]\mathbf{W} + \left[\partial \mathbf{W}'/\partial a_i\right]\left[\partial \mathbf{W}/\partial a_j\right] \\ &+ \mathbf{W}'\left[\partial^2 \mathbf{W}'/\partial a_i \partial a_j\right] + \left[\partial \mathbf{W}'/\partial a_j\right]\left[\partial \mathbf{W}/\partial a_i\right]\end{aligned} \tag{3.12-25}$$

$$\begin{aligned}\partial \mathbf{W}'\mathbf{W}/\partial b_i b_j = &\left[\partial^2 \mathbf{W}'/\partial b_i \partial b_j\right]\mathbf{W} + \left[\partial \mathbf{W}'/\partial b_i\right]\left[\partial \mathbf{W}/\partial b_j\right] \\ &+ \mathbf{W}'\left[\partial^2 \mathbf{W}'/\partial b_i \partial b_j\right] + \left[\partial \mathbf{W}'/\partial b_j\right]\left[\partial \mathbf{W}/\partial b_i\right]\end{aligned} \tag{3.12-26}$$

As before if the exact Hessian matrices are to be computed then $\partial^2 \mathbf{W}/\partial a_i a_j$ and $\partial^2 \mathbf{W}/\partial b_i b_j$ need to be computed. This represents a large computational burden, so these second derivatives are removed in the proceeding equations. The results are (i,j = 1,2,.....,n):

$$\partial \mathbf{W}'\mathbf{W}/\partial a_i a_j \cong \left[\partial \mathbf{W}'/\partial a_i\right]\left[\partial \mathbf{W}/\partial a_j\right] + \left[\partial \mathbf{W}'/\partial a_j\right]\left[\partial \mathbf{W}/\partial a_i\right] \tag{3.12-27}$$

$$\partial \mathbf{W}'\mathbf{W}/\partial b_i b_j \cong \left[\partial \mathbf{W}'/\partial b_i\right]\left[\partial \mathbf{W}/\partial b_j\right] + \left[\partial \mathbf{W}'/\partial b_j\right]\left[\partial \mathbf{W}/\partial b_i\right] \tag{3.12-28}$$

$$\partial L^2/\partial a_i \partial a_j \cong \mathbf{u}'\left[\partial^2 \mathbf{W}'\mathbf{W}/\partial a_i \partial a_j\right]\mathbf{u} \tag{3.12-29}$$

$$\partial L^2 / \partial b_i \partial b_j \cong \mathbf{u}' \left[\partial^2 \mathbf{W}'\mathbf{W} / \partial b_i \partial b_j \right] \mathbf{u} \qquad (3.12\text{-}30)$$

where **u** must be replaced by $\mathbf{u}^{MV}(N \mid q)$.

3.12.3.2. Second Derivatives of *L* with respect to variances.

The second derivatives of *L* with respect to the variances v_B and v_n are obtained by differentiating equations (3.12-20) and (3.12-21) with respect to v_B and v_n. It follows that

$$\partial^2 L / \partial v_B = N / \left(2 v_B^{\,2} \right) - \mathbf{u}'_B \, \mathbf{u}_B / v_B^{\,3} \qquad (3.12\text{-}31)$$

and

$$\partial^2 L / \partial v_n = N / \left(2 v_n^2 \right) - (\mathbf{z} - \mathbf{W}\mathbf{u})'(\mathbf{z} - \mathbf{W}\mathbf{u}) / v_n^3 \qquad (3.12\text{-}32)$$

where MVD estimates of **u** and $\mathbf{u}_B$ must be used.

3.13. Estimation of the Noise Variance.

For the majority of time series to be examined here the noise variance of the series is not known. In order to try and improve the performance of the Maximum-Likelihood Deconvolution algorithm it would be advantageous to have an initial estimate of the noise variance that bears a close proximity to the actual noise variance. A method based on higher order statistics has been used to obtain an estimate for the noise variance [50]. This approach will hopefully lead to value of the noise variance that is a close approximation to the unknown noise variance, thus leading to improved performance of the maximum-likelihood algorithm and MVD filter for a set number of iterations. For the estimation of the noise variance the noise is assumed to be white.

The noise variance can be obtained by modelling the linear time invariant discrete time process with an Auto-Regressive (AR) model. The steps taken to identify the model parameters from the discrete time series will now be outlined.

In order to uniquely identify the AR model of order n, two sets of parameters are required: the AR coefficients $a_1, a_2, \ldots, a_n$ and the variance v_r of the white noise v(k) used as excitation.

The first step in the determination of the noise variance, v_n, is to determine the AR model parameters. These AR parameters can be determined from the Yule-Walker equations of the autocorrelation function as specified in equation (3.13-1) below

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(n-1) \\ r^*(1) & r(0) & \cdots & r(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ r^*(n-1) & r^*(n-2) & \cdots & r(0) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} r^*(1) \\ r^*(2) \\ \vdots \\ r^*(n) \end{bmatrix} \qquad (3.13\text{-}1)$$

where $w_k = -a_k$. In the practice the exact autocorrelation lags are unavailable, so statistical estimates determined from the Higher Order Statistics of the time-series have to be utilised. These estimates will introduce a certain degree of error into the determination of the AR parameters.

Once the AR parameters have been determined, the noise variance can be computed. Starting with the AR model difference equation below:

$$u(k) + a_1^* u(k-1) + \cdots + a_n^* u(k-n) = n(k) \qquad (3.13\text{-}2)$$

where u(k) is the time series time series to be modelled, n(k) is a white noise process and * denotes the complex conjugate.

Assuming that the condition for asymptotic stationarity is satisfied, a recursive relationship for the autocorrelation of the AR process can be derived:

$$E\left[\sum_{j=0}^{n} a_j^* u(k-j)u^*(k-l)\right] = E[v(k)u^*(k-l)] \qquad l = 0,1,\ldots,n. \quad (3.13\text{-}3)$$

By making a number of assumptions this can relationship can be simplified to

$$\sum_{j=0}^{n} a_j^* r(l-j) = 0 \quad l > 0 \qquad (3.13\text{-}4)$$

For l = 0 the expectation on the right hand side of equation (3.13-3) assumes the special form

$$E[v(k)u^*(k)] = E[v(k)v^*(k)] = v_n \qquad (3.13\text{-}5)$$

Accordingly, setting l = 0 in equation (3.13-3) and taking the complex conjugate of both sides, the following relationship for the variance of the white noise can be derived.

$$v_n = \sum_{j=0}^{n} a_j r(j) \qquad (3.13\text{-}6)$$

where $a_0 = 1$.

Hence the white noise variance can be determined.

3.14. Implementation and Computational Requirements of the MVD filter.

Now that the entire theory for Maximum-Likelihood Deconvolution and MVD filtering has been presented the computational requirements for all of the steps in the Block Component Search algorithm will be examined in more detail [34]. It will be

demonstrated that the recursive algorithms presented are necessary to obtain acceptably low processing times for the algorithm.

The computational requirements of the individual components of the Block Component Search algorithm shown in Figure 3.7 are presented in Table 1.

COMPONENT.	Flops
Initialise Parameters	
MVD (needed for threshold detection).	$3N^3$
Threshold detector.	--
Adaptive Detection.	
SMLR Detection.	$3N^3$
λ-Update.	--
Update Wavelet Parameters.	
Gradients.	N^3
Pseudo-Hessians.	--
Marquardt-Levenberg.	--
Update Standard Deviations.	
Gradients.	N^3
Pseudo-Hessians.	--
Marquardt-Levenberg.	--
Update Random Parameters.	N^3
Stop?	N^3

Table 1. *Computational requirements of the BCM algorithm shown in Figure 3.7 using batch programming techniques. Only $O(N^3)$ calculations are listed. A blank indicates $O(N^2)$ or $O(N)$ calculations.*

Assuming α entries to the Adaptive detection block, β iterations in the update statistical parameters block, γ iterations in the update Wavelet parameters block, and

one calculation of L per iteration in each block, the total number of Flops for batch processing is as follows:

$$\begin{aligned}\text{Flops}_b(L) &= \left(3N^3 + N^3\right) + \alpha\left(3N^3 + N^3\right) + \beta\left(N^3 + N^3\right) + \gamma\left(N^3 + N^3\right) + \left(N^3 + N^3\right) \\ &= 4N^3\alpha + 2N^3\beta + 4N^3\gamma + 6N^3\end{aligned} \quad (3.14\text{-}1)$$

As an illustrative example consider N=1000, n = 6 (i.e. a sixth order ARMA wavelet) and $\alpha = \beta = \gamma = 1$, then $\text{Flops}_b(L) = 14\times10^9$. If one Flop requires a micro second then to evaluate $\text{Flops}_b(L)$ requires almost 4 hours of computation time. In practice a number of iterations of each block in the algorithm is required, therefore the already unacceptably long computational time will be become massive. Recursive processing algorithms were presented in Section 3.9 as an alternative to batch processing, the computational requirements of these methods are shown in Table 2.

COMPONENT.	Flops
Initialise Parameters	
MVD (needed for threshold detection).	$4n^3N$
Threshold detector.	--
Adaptive Detection.	
SMLR Detection.	$4n^3N$
λ-Update.	--
Update Wavelet Parameters.	
Gradients.	$2n^3N$
Pseudo-Hessians.	--
Marquardt-Levenberg.	--
Update Standard Deviations.	
Gradients.	$2n^3N$
Pseudo-Hessians.	--
Marquardt-Levenberg.	--
Update Random Parameters.	$2n^3N$
Stop?	--

Table 2. *Computational requirements of the BCM algorithm shown in Figure 3.7 using recursive programming techniques. Only $O(n^3N)$ calculations are listed. A blank indicates $O(n^2N)$ or $O(nN)$ calculations.*

Assuming the same number of iterations per block as in equation (3.14-1), the number of Flops required for recursive processing is,

$$\begin{aligned} \text{Flops}_r(L) &= 4n^3N + \alpha\left(4n^3N\right) + \beta\left(2n^3N\right) + \gamma\left(2n^3N\right) + 2n^3N \\ &= 6n^3N + 4n^3N\alpha + 2n^3N\beta + 2n^3N\gamma \end{aligned} \qquad (3.14\text{-}2)$$

Utilising the same example as before (N=1000, n=6 and $\alpha = \beta = \gamma = 1$) then the processing time taken for $\text{Flops}_r(L)$ is approximately 3 seconds. Therefore, the use of recursive programming makes it practical to implement Maximum-Likelihood Deconvolution.

For this application the algorithms outlined have been implemented in C within a UNIX environment with all processing will be done on a SUN Ultra 10 workstation. All of the numerical algorithms that are required e.g. SVD, FFT, etc. have been taken from *Numerical Recipes in C: The art of scientific computing* [48].

Now that the theory of MLD / MVD filtering has been presented and the computation requirements of the methods investigated the performance of the techniques will be investigated by application to experimentally obtained A-scan data.

3.15. Performance of MLD / MVD filter for experimentally obtained NDT A-scans.

The application of the MVD filter to improve the quality of the A-scan data will now be examined. The performance of the filter will be illustrated examining the effect that application of the MVD filter has on an individual A-scan. The effect of the deconvolution techniques described here on tomographic image reconstruction will be examined in Chapter 4.

Before the performance of this filter can be considered the initial conditions for the estimation of the ARMA wavelet and statistical parameters have to be examined. The statistical parameter v_r cannot be estimated by maximising the likelihood function, L, where both the wavelet and amplitude variance, v_r, are unknown. The reasons for this are related to the nature of L, and depends not only on the structure of the convolutional model but also the apriori statistics. Maximising L turns out to be an ill posed problem with respect to the parameter v_r. For different values of v_r different solutions for the other parameters will be obtained [34]. Recalling the convolution model from Equation (3.6-2), i.e.

$$z(k) = [r(k)q(k) + u_B(k)] + n(k), \quad k = 1,2,.....,N \qquad (3.15\text{-}1)$$

This equation can be rewritten as

$$\begin{aligned} z(k) &= \left\{ \left[r(k)/v_r^{1/2} \right] + u_B(k)/v_r^{1/2} \right\} * v_r^{1/2} w(k) + n(k) \\ &= \left[r'(k) q(k) + u_B'(k) \right] * w'(k) + n(k), \quad k = 1,2,.....,N \end{aligned} \qquad (3.15\text{-}2)$$

where

$$r'(k) = r(k)/v_r^{1/2} \qquad (3.15\text{-}3)$$

$$u_B'(k) = u_B(k)/v_r^{1/2} \qquad (3.15\text{-}4)$$

and

$$w'(k) = v_r^{1/2} w(k) \qquad (3.15\text{-}5)$$

From these equations it can be seen that, when w(k), r(k) and $u_B(k)$ are all unknown, then the model we are working with is a scaled version of the model in equation (3.5-2). The scaled wavelet $w'(k)$, the scaled amplitude sequence $r'(k)$, and the scaled backscatter sequence $u_B'(k)$ are the model parameters that can be determined. Observing the fact that the variance of the scaled amplitude sequence equals one, i.e. $v_r'=1$. Hence when all of these quantities are unknown then the variance of the amplitude sequence may as well be fixed, as it cannot be estimated. This is the case here, so the amplitude sequence variance, v_r, will be fixed by the user.

Also the initial values of the variances v_n and v_B will not be known. Usually these variances are assumed initially to have small values, or in the case of v_n Higher Order Statistics (HOS) can be used to obtain an estimate as described in Section 3.13.

3.15.1 MVD Filter Performance.

Initially the filter's performance will be evaluated by applying the MVD filter to a single A-scan extracted from three complete image reconstruction sets. In order to

demonstrate the effectiveness of MVD filtering for tomographic images reconstructed using the algorithms described in Chapter 2, A-scans obtained under different experimental configurations need to be examined. To demonstrate the effectiveness of the MVD filter the following transmitter and receiver configurations were chosen,

1. Pulse-Echo.
2. Pitch-Catch configuration with a forty degree separation between transmitter and receiver.
3. Pitch-Catch configuration with a sixty degree separation between transmitter and receiver.

The A-scans used were obtained from a 1.2mm inclusion through a 75mm aluminium specimen as shown in Figure 4.3 in Chapter 4. The same experimental set-up as described in Chapter 4 has been used to obtain the experimental A-scan data. The filtering operation will be described in detail for the first A-scan and for the remaining A-scans the filter parameters will be presented, followed by the result of the filtering operation and a brief discussion of the improvements in the A-scan.

The first A-scan that will be considered was obtained using a Pulse-Echo experimental configuration. The A-scan prior to processing is shown in Figure 3.15.

MISSING PAGES ARE UNAVAILABLE

As described previously the ideal resolution function for the MVD filter is a delta function. The resolution function for the deconvolution performed on the A-scan in Figure 3.15 can be seen in Figure 3.17.

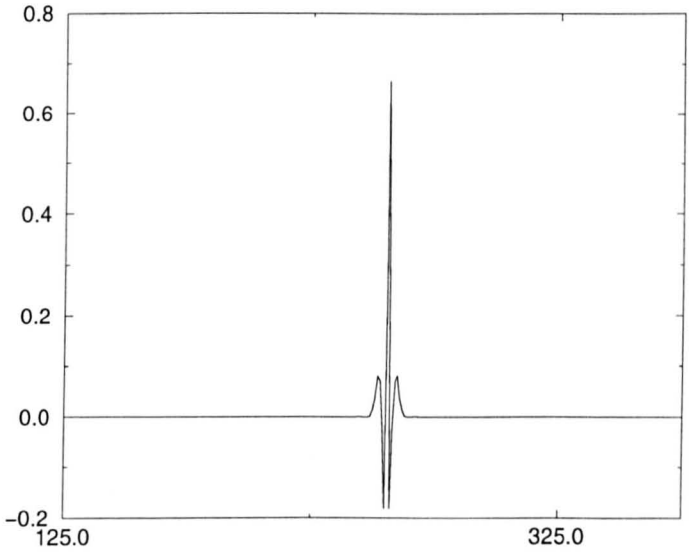


Figure 3.17. *The resolution function for the deconvolution performed on the A-scan shown in Figure 3.15.*

It is clear upon examination of the resolution function that the resolution of the MVD filter operation approaches the ideal.

The second A-scan to be examined was obtained with a forty degree separation between the transmitter and receiver, the unfiltered A-scan is shown in Figure 3.18.

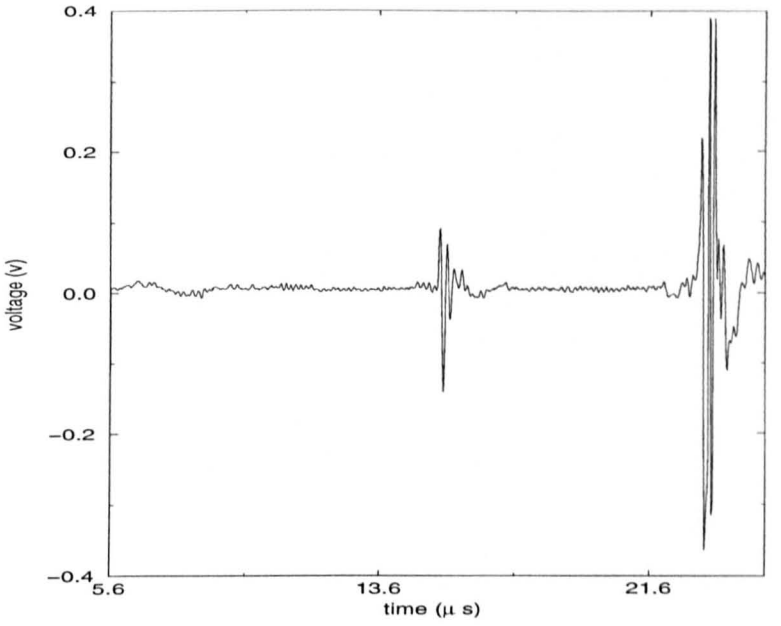


Figure 3.18. *A-scan prior to filtering, the reflection from the defect and the backwall can clearly be seen.*

The same ARMA estimation technique and parameters as the first A-scan were used and yielded an ARMA model order of 4. The noise variance, v_n, was determined using higher order statistics as before. The amplitude variance, v_r, was chosen to be 0.01. The initial value of λ was chosen to be 0.05, and there was assumed to be no backscatter in this case (i.e.$v_B = 0$). The BCM algorithms were run for the same number of iterations as the first A-scan (i.e. 5). The filtered A-scan can be seen in Figure 3.19.

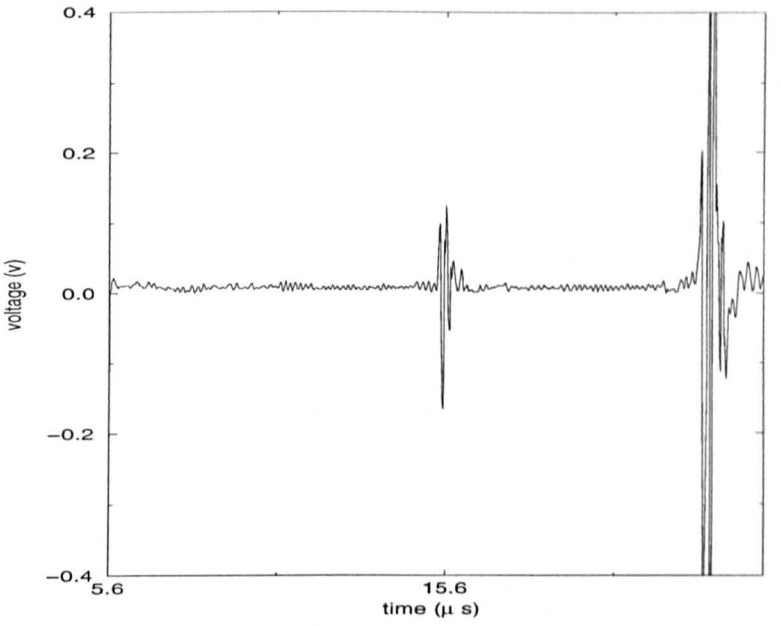


Figure 3.19 *A-scan after application of the MVD filter, The parameters for the filter being obtained using MLD as described.*

Upon examination of the filtered A-scan it is evident that the level of noise present has been reduced in a number of regions of the A-scan, notably at the beginning of the A-scan and immediately after the pulse reflected from the 1.2 mm hole. However, of more significance for this A-scan is that the pulse reflected from the defect has been decreased by approximately 15% in length resulting in a higher time resolution, in addition the signal amplitude of the reflected pulse has been increased with no increase in the noise amplitude. This figure was obtained by measuring the pulse width of the pulse reflected by the simulated defect at the 3dB point before and after the filtering operation. The resolution function for the filtering operation is presented in Figure 3.20.

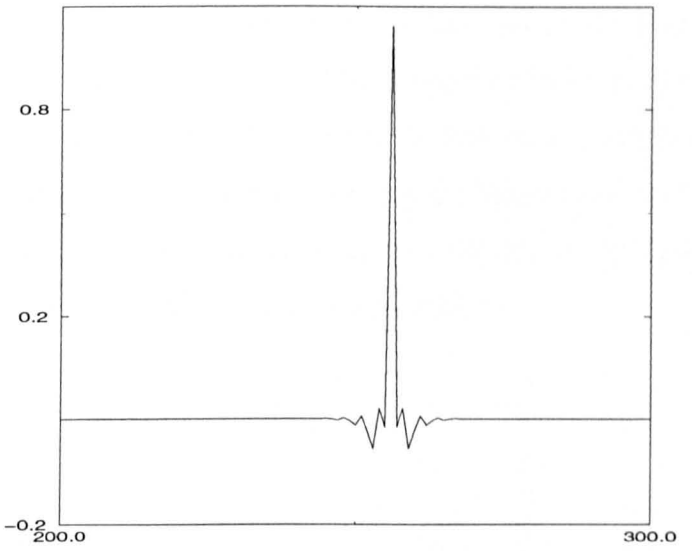


Figure 3.20 *The resolution function for the deconvolution performed on the A-scan shown in Figure 3.18.*

It is clear upon examination of the resolution function in Figure 3.20 that the resolution of the MVD filter operation approaches the ideal.

The third and final A-scan to be examined was obtained with a sixty degree separation between the transmitter and receiver, this A-scan is shown in Figure 3.21.

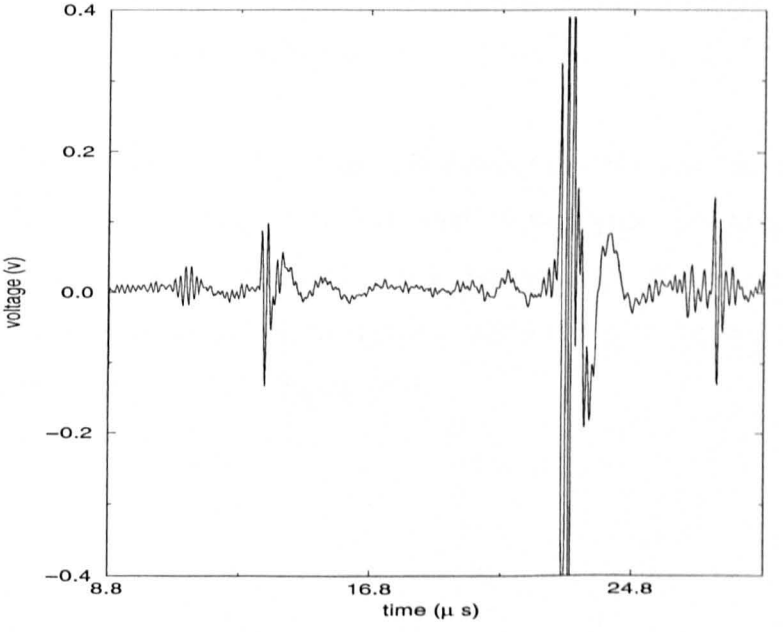


Figure 3.21 *A-scan prior to filtering, the reflection from the defect and the backwall can clearly be seen.*

The same ARMA estimation technique and parameters as the first and second A-scans was used here and produced an ARMA model order of 2. The noise variance, v_n, was determined using higher order statistics as before. The amplitude variance, v_r, was chosen to be 0.05. The initial value of λ was chosen to be 0.05, and there was assumed to be no backscatter in this case (i.e.$v_B = 0$). The BCM algorithms were run for the same number of iterations as the first A-scan.

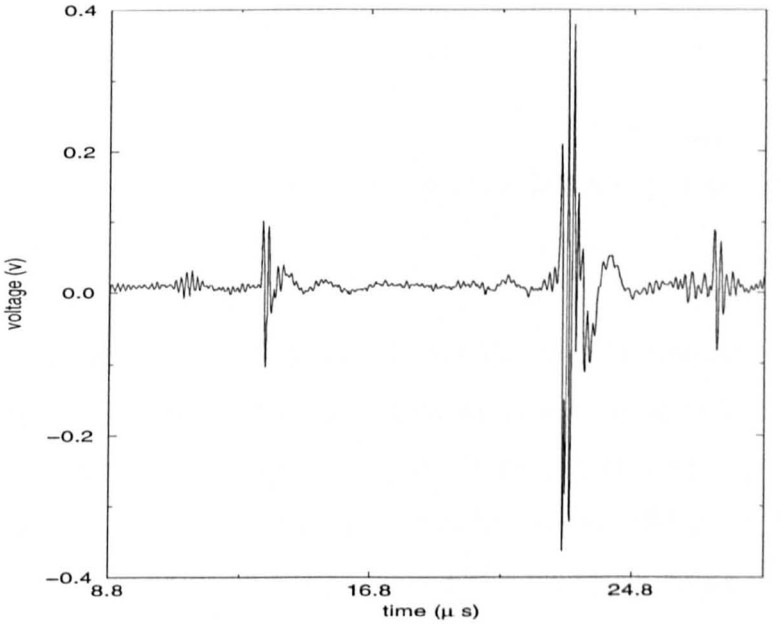


Figure 3.22 *A-scan after application of the MVD filter, The parameters for the filter being obtained using MLD as outlined above.*

Upon examination of the filtered A-scan it is evident that the level of noise present has been significantly reduced. There has been no noticeable improvement in the time resolution of any of the reflected pulses, however, the filtering operation has improved the A-scan significantly enough for the filtering to be worthwhile. The resolution function can be seen in Figure 3.23.

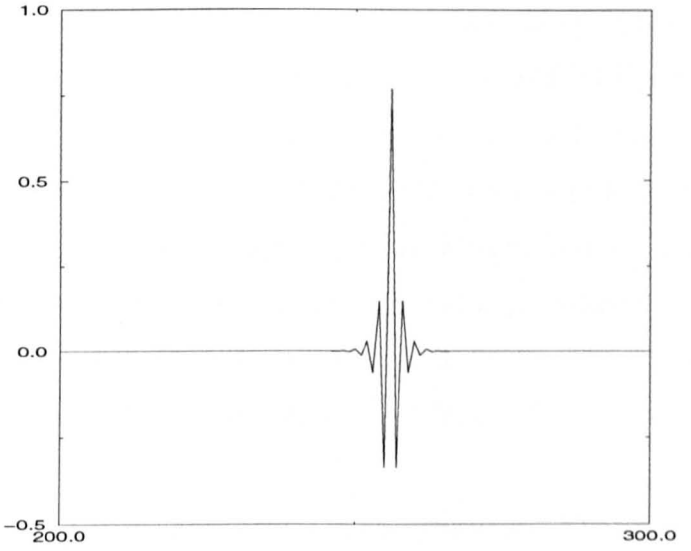


Figure 3.23 *The resolution function for the deconvolution performed on the A-scan shown in Figure 3.21.*

It is clear upon examination of the resolution function that the resolution of the MVD filter operation is not as ideal as the other resolution functions shown in Figures 3.20 and 3.18. This can possibly be explained by the larger amount of noise present in the unfiltered A-scan, thus resulting in poorer estimation of the ARMA model and other MLD parameters. The increased noise levels will particularly affect the parameters estimated using HOS.

The performance of Maximum-Likelihood Deconvolution and Minimum Variance Deconvolution with respect to the reconstruction of tomographic images will be investigated in Chapter 4.

3.16. Concluding Summary.

In NDT imaging the resolution and SNR of the data to be reconstructed is of critical importance to the quality and resolution of the image reconstructed. To this end a deconvolution filter has been implemented to try and improve the time resolution and SNR of the raw A-scan data used for the reconstruction algorithms.

In order to maintain maximum flexibility within the imaging system, it is desirable to obtain the wavelet and statistical parameters for the implementation of the Minimum

Variance Deconvolution (MVD) filter without knowledge of the measurement system or test specimen. This can be achieved for the MVD filter using Maximum-Likelihood Deconvolution (MLD). The underlying theory for MLD has been presented along with a discussion of requirements for actual implementation of the algorithms presented. It has been necessary to develop recursive algorithms to obtain acceptable processing times, these algorithms are also presented here. MLD requires initial estimate of the wavelet and statistical parameters to be given. The methods utilised here based on Higher Order Statistics (HOS) have been outlined in detail.

The underlying theory of the MVD filter is also discussed. The operation of utilising MLD to obtain estimates for the wavelet and statistical parameters required for MVD filtering has been discussed and then evaluated for experimentally obtained A-scan data. The performance of the MVD filter has been evaluated using the parameters obtained. It has been demonstrated that the filter results in an improvement in the time resolution and signal to noise ratio of the raw A-scan data.

It has been shown that the algorithms presented here can be highly computationally intensive. To this end the processing times required to obtain the final wavelet and statistical parameter estimates have been investigated and an improvement in both the time resolution and signal to noise ratio of the raw A-scan data has been obtained for relatively low computational times.

Chapter 4:
EXPERIMENTAL EVALUATION OF THE IMAGING SYSTEM.

4.1. Introduction.

The objective of this chapter is to evaluate the performance of the four main components of the imaging system outlined in Chapters 2 and 3. The components of the system being the four tomographic image reconstruction algorithms, the low level fuzzy logic image fusion technique, the single bounce image enhancement method, and maximum likelihood deconvolution.

These five system components will be evaluated experimentally. The approach that has been chosen will be to evaluate their performance for a number of simulated, representative defects embedded within cylindrical engineering type components. The representative defects that have been chosen are inclusions of varying dimensions, planar defects and to assess the resolving capacity of the imaging system on a line of small inclusions of differing separation.

The reconstruction algorithms have been implemented to be independent of the test specimen geometry. However, to evaluate the performance of the system it is necessary to have access to the defect from all angles, therefore all of the defects will be contained within objects with a cylindrical geometry. All of the specimens used here were made of aluminium and had a diameter of 75 mm. In Section 4.1 the data acquisition set-up used to capture the raw ultrasonic data for image reconstruction will be outlined. Section 4.2 describes the Linear Cross-Correlation method that has been used to assess the quality of the images generated using the tomographic reconstruction algorithms and the image fusion utility. The first components of the imaging system that will be examined are the tomographic reconstruction algorithms and the image fusion utility, this is dealt with in Section 4.4 for a number of simulated defects. The next system component to be evaluated will be the single bounce image enhancement method, this is presented in Section 4.5. In Section 4.6 the performance of the image reconstruction and the single bounce image enhancement method are evaluated where complete access to the test specimen is not available. The final component of the imaging system to be examined is the pre-processing of the A-scan data using MVD filtering, the effect of this filtering operation on image reconstruction is examined in Section 4.7.

4.2. Data Acquisition.

A data acquisition system was constructed to acquire tomographic time domain data (A-scan data) from cylindrical objects up to a diameter of 100 mm. The system consists of a pair of transducers mounted on two rings of 160 mm surrounding the object, to facilitate the collection of the A-scan data. This positioning system allows the transducers to be placed anywhere on the circumference of the object under test. The transducers are held in spring loaded holders to keep them lightly in contact with the object. Good coupling between the transducers and the object was obtained using a thin layer of coupling gel. This component of the imaging system is shown in Figure 4.1.

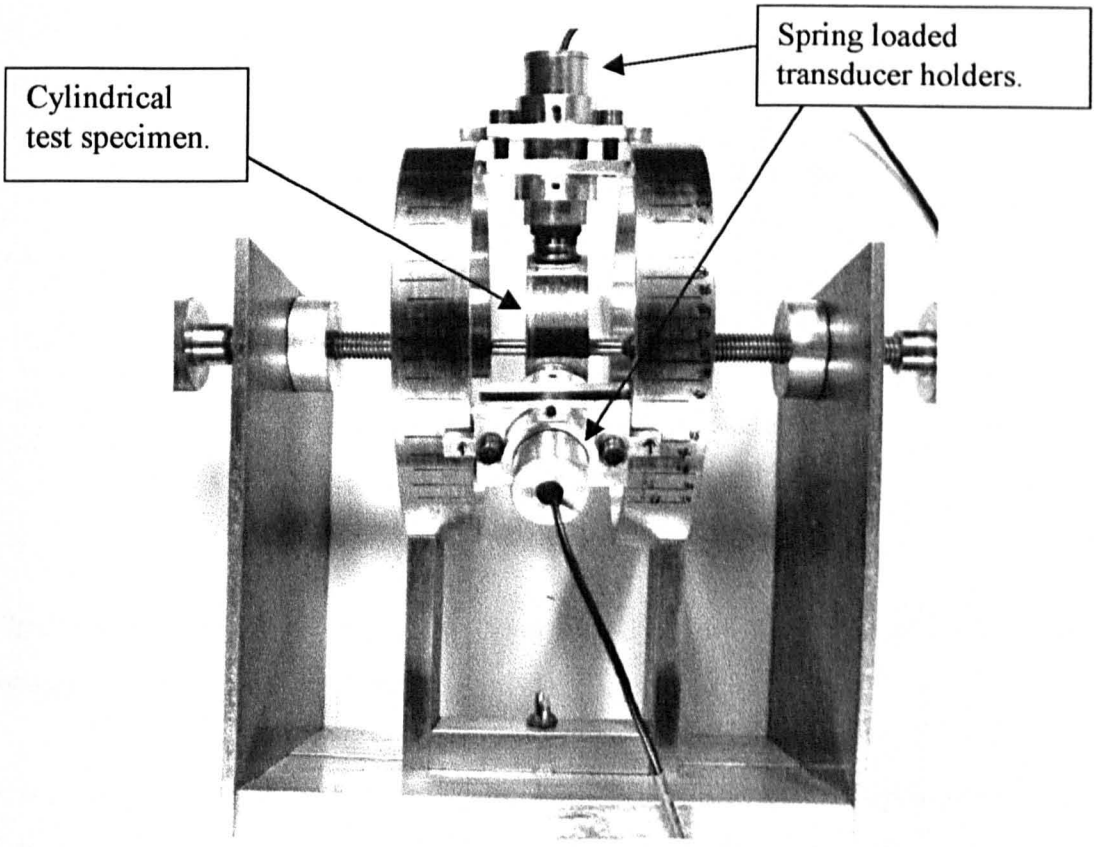


Figure 4.1. *The A-scan data acquisition component of the imaging system. The test specimen can be clearly seen in the centre of the scanning rig, the two transducers held within spring loaded holders and positioned on two rings surrounding the test object can also be seen in the figure.*

The transducers used for data acquisition were a pair of Panametrics broadband Videoscan transducers, with a centre frequency of 5 MHz. These were driven using a Panametrics 5052PR pulser/receiver. The received signals were digitised using a HP 54502A 400 MHz digitising oscilloscope, which was controlled and waveform data obtained using a GPIB interface linked to a personal computer. The transducers were positioned manually in the desired location around the object with the resulting data being stored before being used by the components of the image reconstruction process.

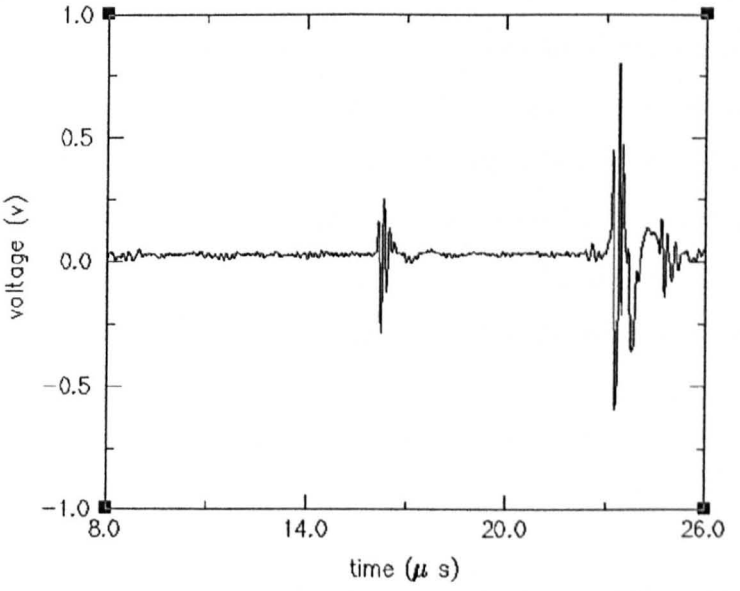


Figure 4.2. *Sample A-scan resulting from Pulse-Echo isonification of a 1.2mm hole through 75 mm diameter aluminium cylinder.*

A sample A-scan can be seen in Figure 4.2, this waveform was captured using the data acquisition set-up described. The system is in Pulse-echo configuration where the simulated defect being insonified is a 1.2mm side drilled hole through 75 mm diameter aluminium cylinder.

4.2.1. Point Source/Receiver.

For the image reconstruction methods described in Chapter 2, namely Reflection tomography, Time-of flight Diffraction tomography and Transmission tomography, the transmitter and receiver are assumed to be point sources for 2D imaging. This assumption will now be justified for the data acquisition set-up described in Section 4.2.

The transducers employed were cylindrical in shape but can be considered to be point sources due to the transducer/object contact conditions and the nature of the image reconstruction. That is, as the object is cylindrical and the transducer face flat, then the region of contact between the two can be considered a line. In addition, the reconstructed image is a 2-D slice of the object cross section, so the line approximates to a point source in the reconstruction plane. This assumption is consistent with the model for fan-beam isonification presented in Chapter 2. Moreover, the frequency of the ultrasound used was 5 MHz, which has a wavelength of approximately 1.28 mm in aluminium. This is small when compared to the object size of 75 mm.

4.2.2. Beam Divergence.

It can be shown that the angle of divergence of the beam within the object under test is an important parameter of the imaging system. Obviously if the beam has a narrow beamwidth, then for each of the N chosen transmitter locations only a small area of the test specimen cross section will be illuminated. Therefore, the chance of defects in the specimen being illuminated is considerably smaller than when a transducer with a wide beamwidth is used. This means the choice of a narrow beamwidth transducer will lead to sub optimal image reconstruction of defects that are not close to the object centre. This degradation of the reconstructed image is additional to the image degradation described in Section 2.2, where the point spread function was examined with respect to the position of the point target. Moreover, the further the defect is from the centre of the test specimen the more both of these effects will effect the image. Therefore, it is highly advantageous to use transducers with a wide

beamwidth in order to minimise at least one of the causes of image degradation. For the data acquisition set-up described here, the divergence of the beam within the object was measured to be ±35 degrees for an aluminium cylinder of 75mm diameter, providing adequate coverage of the samples.

4.3. Image Quality.

In order to assess the quality of the images reconstructed and to identify any improvement in the image quality resulting from the image fusion and single bounce image enhancement, Pearson's linear cross-correlation technique has been applied [48]. The intention is to use cross-correlation to gain a measure of the similarity of the reconstructed images to that of an ideal image of the defect. Here the term ideal defect image refers to an image where all pixels contained within the defect boundary have maximum amplitude and all other pixels have zero amplitude. The chosen cross-correlation method will now be briefly outlined.

4.3.1. Pearson's Linear Cross-Correlation.

For this method the images are considered to consist of two 1-D arrays, represented by x_i and y_i respectively, where i = 1,2,....,M (M is the number of pixels in the image). The Pearson's or linear cross-correlation coefficient, r, is determined using the following formula [48].

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2}\sqrt{(y_i - \bar{y})^2}} \qquad (4.3\text{-}1)$$

where $\bar{x}$ is the mean of the x_i 's, and $\bar{y}$ is the mean of the y_i 's. For this application the value of r varies from zero to one. If r equals one then there is complete correlation between the two images.

4.4. Image Reconstruction and Fusion.

In this section the experimental performance of the tomographic reconstruction algorithms and the image fusion process will be investigated. The methods will be evaluated for the set of simulated representative defects. The set of defects chosen are a 1.2 mm inclusion simulated by a 1.2mm side drilled hole; a 7mm inclusion simulated by a 7mm side drilled hole; a planar flaw simulated by a slot in the test specimen; and a line of holes at various separations to assess the resolving capacity of the imaging system. All of these defects are embedded within aluminium cylinders of 75 mm diameter.

In order to validate the improvement in image quality obtained through fusing a set of tomographic images, a two step approach has been used. The first step is to reconstruct the individual defect images and assess there correlation with the ideal defect image. The second step is to fuse a set of the tomographic images to generate a single composite image of the defect and assess the correlation of this image with the ideal image. However, before any image reconstruction is performed the computational requirements of the algorithms will be discussed.

4.4.1. Computational Requirements of Tomographic Image Reconstruction.

For the tomographic image reconstruction and fusion methods used here to be useful for the non-destructive testing of components, the reconstruction times must be acceptably low. All of the images reconstructed here have been reconstructed on a SUN ultra 10 workstation. For Pulse-Echo Reflection, which is the simplest of the reconstruction algorithms computationally, the reconstruction times for the images presented here is below 30 seconds.

For the Pitch-Catch Reflection algorithm, which involves backprojection of the ultrasonic data over elliptical paths rather than circular paths, the reconstruction times are in the order of a minute, this being due to the increased theoretical complexity of the algorithm. Reconstruction times of this order for Time-of-Flight Diffraction tomography have been achieved by limiting the number of transmitter

locations, however, it will be shown that acceptable defect images can be reconstructed for most of the defects considered.

Obviously, with the addition of single bounce image enhancement these reconstruction times will increase. The addition of this reconstruction feature adds up to an extra two minutes to the reconstruction times when equal amounts of data are backprojected both before and after the backwall. This is due to the complex shape of the backwall considered. For test specimens which are less geometrically complex the time taken to backproject the single bounce data would be considerable less.

The Transmission tomography reconstruction times depend on the method of backprojection used and the amount of transmission data utilised for the reconstruction. For linear backprojection and the amount of data used for all of the transmission images generated in this chapter, reconstruction times are approximately 60 seconds. For the filtered backprojection case the reconstruction times are slightly longer, being in the order of 90 seconds.

The final component of the imaging system to be considered in this section is the fuzzy logic fusion utility. For all applications of the fusion process considered here the processing time taken was negligible.

4.4.2. 1.2 mm Inclusion.

The first defect that will be examined will be a small inclusion 1.2mm in diameter and positioned through the test specimen as shown in Figure 4.3.

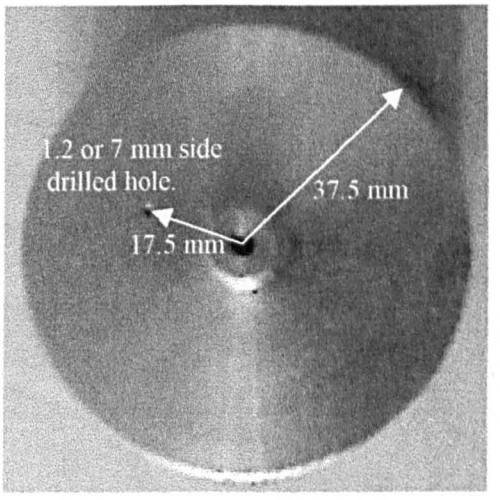


Figure 4.3. *Position of the simulated Inclusion type defects within an aluminium cylinder of 75 mm diameter.*

For this particular defect three tomographic images were reconstructed, these being; Pulse-Echo Reflection, Pitch-Catch Reflection and Time-of-flight Diffraction tomography. No Transmission tomography image was reconstructed for this defect due to its small size compared to the beamwidth of the transducers used.

4.4.2.1. Pulse-Echo Reflection Tomography.

The Pulse-Echo Reflection tomography image was reconstructed from 72 A-scans obtained from transmitters equally distributed around the circumference of the aluminium cylinder, giving a 5° angular separation between transmit locations. The reconstructed image is shown in Figure 4.4. The colour table for all images presented in this thesis is given below the reconstructed image. The linear cross-correlation with the ideal image was determined to be 0.466.

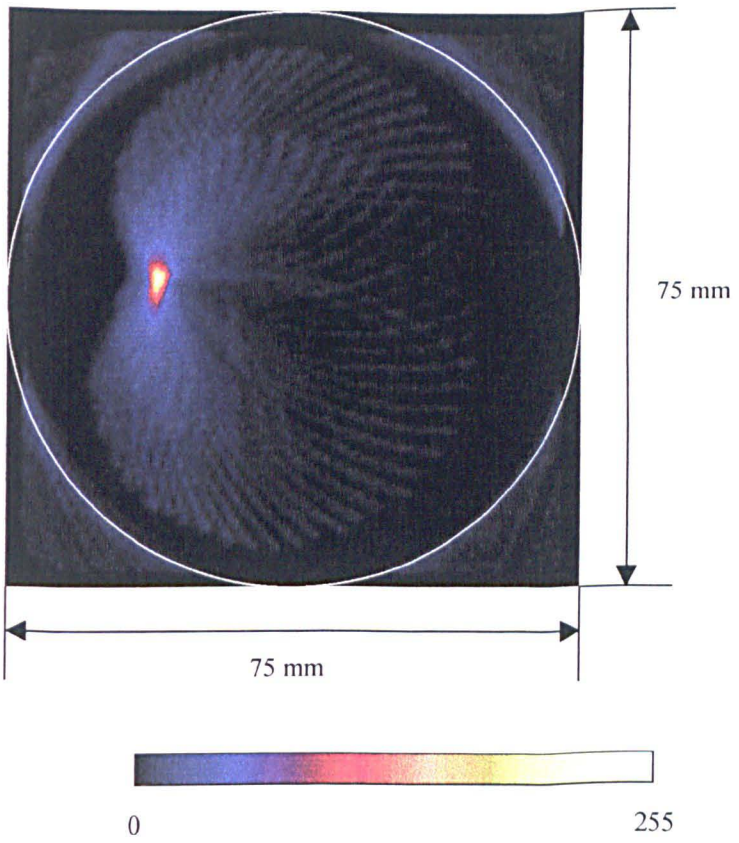


Figure 4.4. *Image of the 1.2mm inclusion where the centre of the image corresponds to the centre of the test specimen. The colour table for all of the tomographic images presented here is shown below the image (image scale 1:1). The backwall of the cylindrical test specimen is marked on the image in white.*

4.4.2.2. Pseudo Point Spread Function.

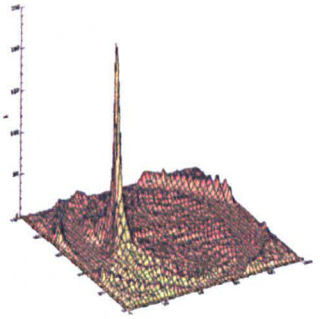


Figure 4.5. *Pseudo point spread function of Pulse-Echo Reflection tomography.*

In order to evaluate the performance of the reconstruction algorithms the experimental point spread function will be examined. The theoretical expressions for the Point Spread Functions (PSF) were derived in Chapter 2. A direct comparison cannot be made between the theoretical PSF and the experimentally obtained pseudo PSF. The main reasons are that the target being imaged is not a point reflector but an inclusion of finite size. Also in the theoretical expression, the aperture is assumed to be sampled over its entirety (here the aperture is sampled at a relatively small number of points), and finally the defect is not situated in the centre of the test specimen as was assumed in the derivation of the PSF. However, the pseudo PSF can still give a useful indication to the performance of the image reconstruction algorithms relative to each other. The Pulse-Echo Reflection pseudo point spread function is shown in Figure 4.5.

4.4.2.3. Pitch-Catch Reflection Tomography.

The Pitch-Catch Reflection tomography image reconstructed is shown in Figure 4.6. A-scans were captured at 72 equally spaced positions around the circumference of the cylindrical sample, giving a 5^{o} angular separation between transmitter locations. The separation between the transmitter and receiver for this image was chosen to be 50^{o} ($\beta = 25$). The linear cross-correlation of this image with the ideal image is 0.365.

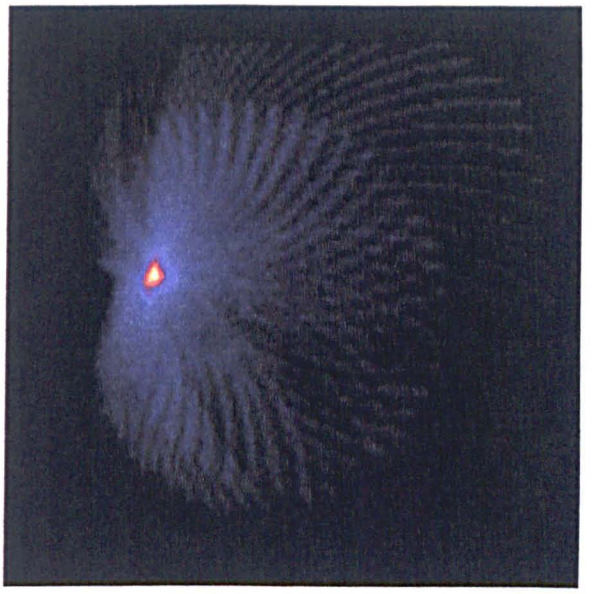


Figure 4.6. *Pitch-Catch Reflection tomography image, the transmitter angular separation is 5^{o} and the transmitter to receiver separation is 50^{o} (image scale 1:1).*

4.4.2.4. Pseudo Point Spread Function.

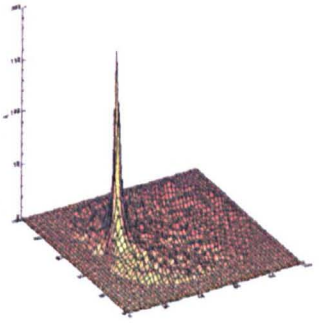


Figure 4.7. *Pseudo point spread function for Pitch-Catch Reflection tomography.*

The pseudo point spread function for Pitch-Catch Reflection tomography is shown in Figure 4.7. The discussion regarding the PSF for Pulse-Echo Reflection tomography in Section 4.4.1.2 is valid here. Upon comparing the point spread functions for the Pulse-Echo and Pitch-Catch data acquisition set-ups, it is evident that for small transmitter to receiver separations (β less than 45 degrees) the pseudo point spread functions have a similar rate of main lobe decay. This can be seen clearly upon examination of the point spread functions shown in Figure 4.5 and 4.7.

4.4.2.5. Time-of-flight Diffraction Tomography.

The Time-of-flight Diffraction tomography image was reconstructed from time domain data taken from 4 transmitter locations equally spaced around the test specimen. The receiver was scanned over the complete circular aperture for each of the four transmitter positions, with the angular separation between the receiver locations being chosen to be 10°. This amount of data was used in order to obtain reconstruction times comparable with the other two reconstruction methods, the reconstructed image being shown in Figure 4.8. The image has a cross-correlation with the ideal image of 0.396.

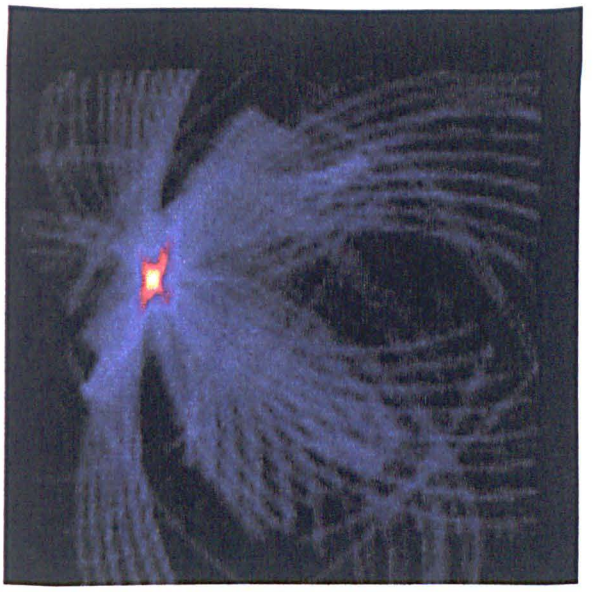


Figure 4.8. *Time-of-flight Diffraction tomography image of the 1.2mm inclusion. The transmitter separation was chosen to be 90^0 with the receiver scanning the aperture with an angular separation of 10^0 (image scale 1:1).*

4.4.2.6. Pseudo Point Spread Function.

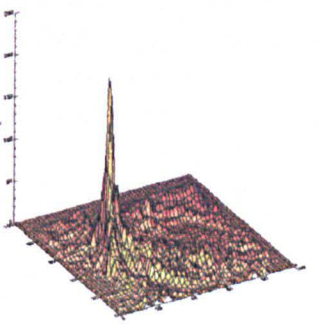


Figure 4.9. *Pseudo Point Spread Function for Time-of-flight Diffraction Tomography.*

The pseudo point spread function for the Time-of-flight Diffraction tomography is shown in Figure 4.9. This clearly does not possess the same rate of mainlobe decay as the other two reconstruction methods. This can be partially explained by the need to use significantly fewer transmitter positions in the reconstruction in order to obtain

an acceptable reconstruction time when compared to the other tomographic reconstruction methods. The varying transmitter to receiver separation (β will often be greater than 45^{o}) will also have a detrimental effect on the point spread function, as stated in Chapter 2. That is, as the separation between the receiver and transmitter becomes larger, the more degraded the PSF will become. Finally the choice of transmitter locations with reference to the defect position will have considerable effect on the point spread function due to the small number of locations used for reconstruction.

4.4.2.7. Fused Images.

In order to assess the improvement in the image quality due to image fusion, a linear cross-correlation with the ideal image is employed. To put the improvement in the image quality in context, the average of the constituent images has been generated and assessed using linear cross-correlation. The three images used in the fusion are the Pulse-Echo Reflection, Pitch-Catch Reflection and Time-of-flight Diffraction tomography, the average of these three images can be seen in Figure 4.10.

The fused image is shown in Figure 4.11. For this fusion the Ostu image thresholding method was applied with the third ramp function. A zero cut-off of $0.45T_c$ was used, where T_c is the computed threshold value and the ramp parameter, a, was chosen to be 3.

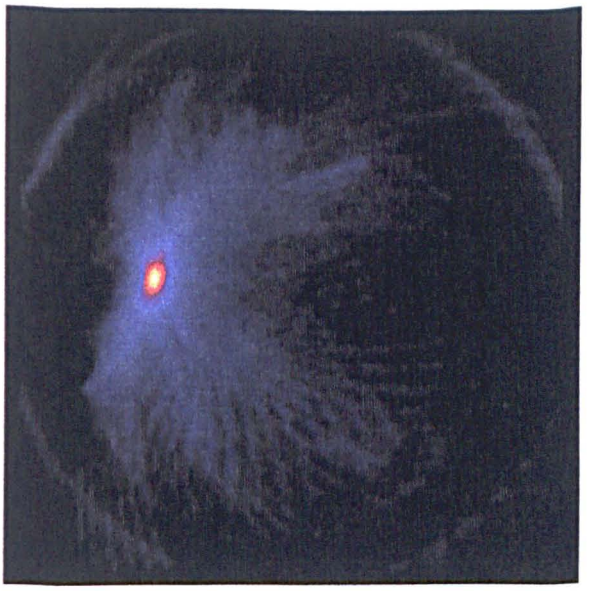


Figure 4.10. *Average of the three images used for fusion, these images can be seen in Figures 4.4,4.6 and 4.8 (image scale 1:1).*

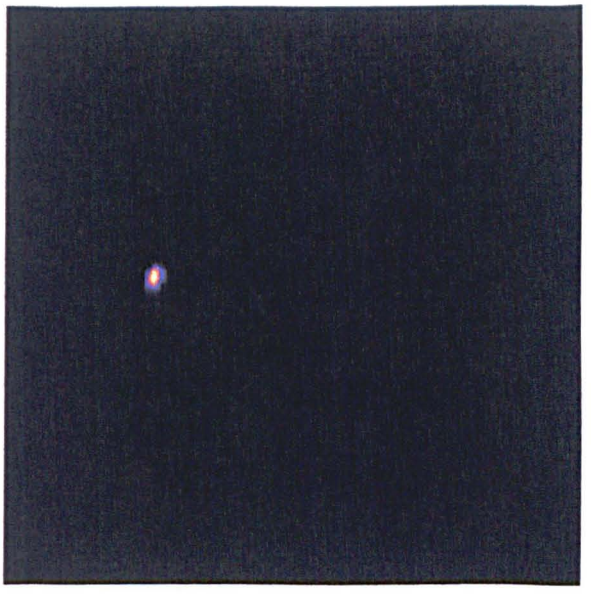


Figure 4.11. *Optimal fusion of the Three tomographic images. The 1st fusion function was utilised and the Ostu thresholding technique employed, with a zero cut-off of 045 T_c and a = 3 (image scale 1:1).*

The correlation values for these images are given in Table 1. The characterisation of the defect is improved considerably with an increase in correlation over the average image of 0.313.

Tomographic Image	Linear Cross-correlation with ideal image, r.
Pulse-Echo Reflection Tomography.	0.466
Pitch-Catch Reflection Tomography.	0.365
Time-of-flight Diffraction Tomography.	0.396
Summation Image.	0.436
Fused image	0.749

Table 1. *Correlation values for all of the tomographic images reconstructed. These correlation values were obtained using Pearson's linear correlation equation (4.4-1).*

The pseudo point spread function can be seen in Figure 4.11a. The PSF has a mainlobe rate of decay that is higher than the images reconstructed using the individual tomographic reconstruction techniques. This goes some way to explaining why the cross-correlation value for the fused image is significantly higher than the average image. In addition, the majority of noise present in the constituent has been removed from the fused image.

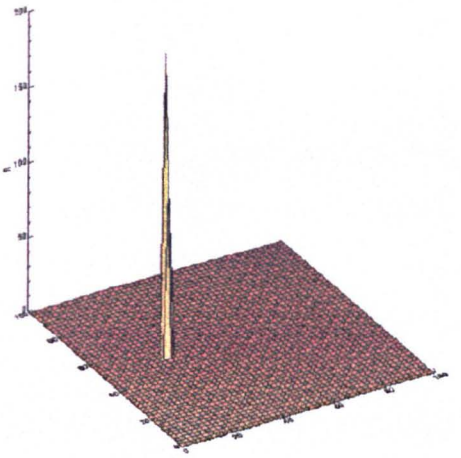


Figure 4.11a. *Pseudo Point Spread Function of the fused image presented in Figure 4.11.*

4.4.3. 7mm Inclusion.

The position of the 7mm inclusion within the aluminium test specimen is given in Figure 4.3. The images reconstructed using the four image reconstruction methods described in Chapter 2 will now be presented.

4.4.3.1. Pulse-Echo Reflection Tomography.

The Pulse-Echo Reflection tomography image was reconstructed using A-scan data obtained at seventy-two transmitter locations equally spaced around the circumference of the test specimen. The resultant image can be seen in Figure 4.12, below. The physical boundary of the defect is shown in white, it can be seen that the boundary has not been uniformly reconstructed in its entirety. This is due to the location of the defect within the test specimen and the data acquisition set-up. Because the transducer used to acquire the A-scans is held in contact with the specimen, the near field is within the test specimen (in immersion testing the set-up is designed so the specimen is in the far field to avoid these problems). This means that the defect is not always situated in the far field of the transducer. Therefore, the

energy incident on the defect varies considerably depending on the transmitter position leading, to poor reconstruction of certain parts of the defect boundary.

The image quality was assessed using linear cross-correlation with the ideal image. For this particular defect three of the reconstruction methods use the ultrasound reflected or diffracted from the defect boundary, notably Pulse-Echo Reflection, Pitch-Catch Reflection and Time-of-flight Diffraction tomography. For this reason, the ideal image was chosen to have zero pixel amplitude everywhere except along the defect boundary. This ideal image was chosen because it represents the best image of the defect that can be reconstructed using those algorithms. For the other defects examined, their dimensions are such that there is no difference between the two ideal images.

This image has a linear cross-correlation value of 0.229, this low value of correlation was caused by the presence of noise within the image and does not truly reflect the quality of the defect image reconstructed. However, the correlation value is still meaningful for analysis of the image fusion tool.

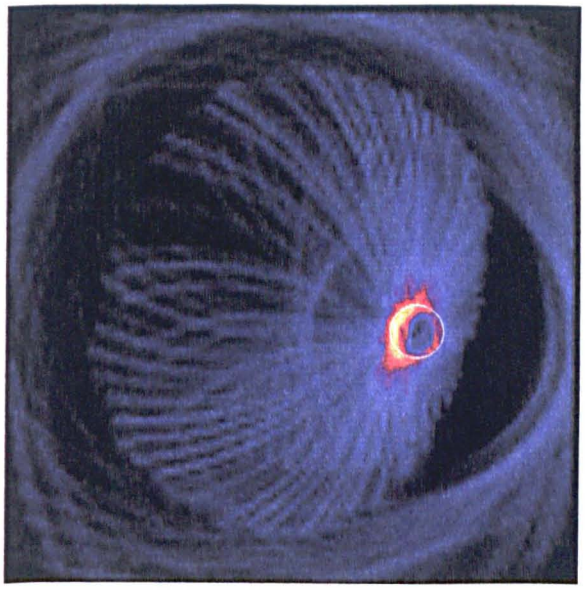


Figure 4.12. *Pulse-Echo Reflection tomography image with the actual defect boundary marked. The image was reconstructed with A-scan data obtained with a 5° angular transmitter separation (image scale 1:1).*

4.4.3.2. Pitch-Catch Reflection Tomography.

The pitch-Catch Reflection tomography image reconstructed of the 7mm inclusion is shown in Figure 4.13. This image was reconstructed with an angular transmitter separation of 5^{o} and a constant transmitter to receiver separation of 40^{o} (β=20).

The boundary of the defect is shown in black and it can be seen that the boundary has not been uniformly reconstructed in its entirety. This is due to the reasons discussed in the previous section for the Pulse-Echo reflection image. This image has a linear cross-correlation with the ideal image of 0.313.

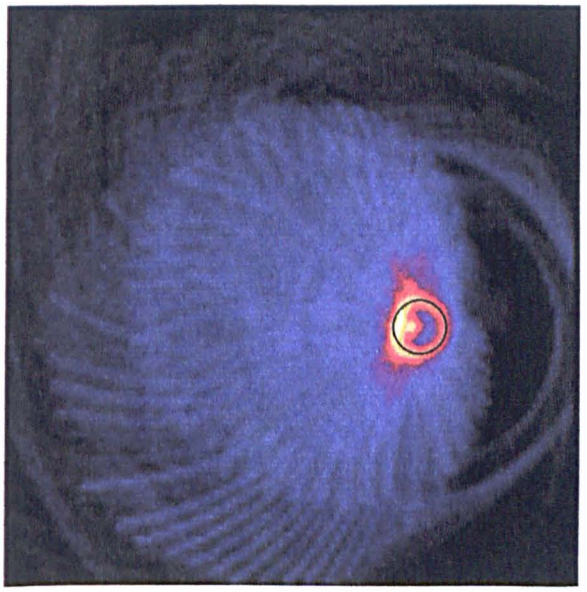


Figure 4.13. *Pitch-Catch Reflection tomography Image of the 7mm circular inclusion with a 5^{o} transmitter separation and a 40^{o} transmitter to receiver separation (image scale 1:1).*

4.4.3.3. Time-of-flight Diffraction Tomography.

The Time-of-flight Diffraction tomography image is shown in Figure 4.14. For this image, four transmit locations equally spaced around the object circumference were chosen. The ultrasonic energy scattered by the defect was then collected, over the circumference of the sample, at 10^{o} intervals. The reconstructed image does not give a good indication of the flaw shape or orientation, the actual defect boundary being

shown in white on the image. This can be explained by the fact that strong a diffracted signal can only be obtained from the defect when the angle of incidence to the defect is favourable. In this case, a number of the diffracted pulses received were not of sufficient quality, resulting in a deterioration of the image. If more transmitter and receiver locations are utilised then the image obtained would more closely resemble the ideal scenario. However, a balance has to be struck between the amount of data used and the computational time required to reconstruct the image. In the majority of cases this will depend on the user requirements. The linear cross-correlation with the ideal image is 0.335.

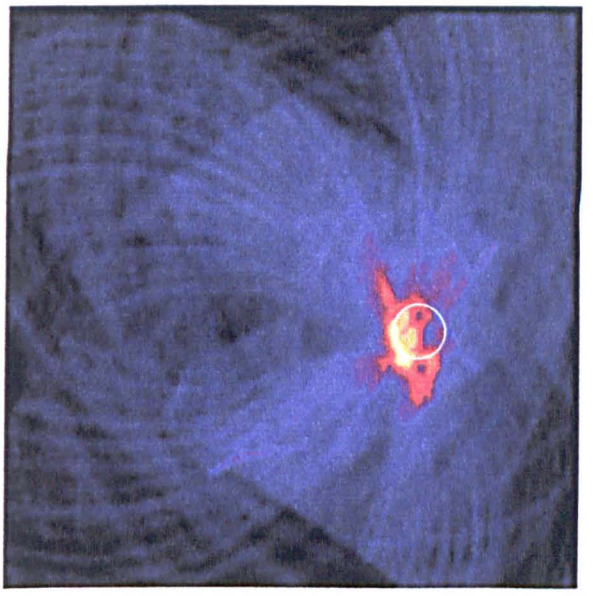


Figure 4.14. *Time-of-flight Diffraction Tomography image with a transmitter separation of 90° and a receiver separation of 10° (image scale 1:1).*

4.4.3.4. Transmission Tomography.

Transmission images of the 7mm inclusion were reconstructed using both the linear and filtered backprojection algorithms presented in Chapter 2. The same set of A-scan data was used for both the image reconstruction algorithms. The A-scan data being collected with a transmitter separation of 10° and a 5° receiver separation. For the reconstruction of this image it was decided that a shadow region is present if the

received signal amplitude was less than 20% of the amplitude expected at the backwall when no defect is present between the transmitter and receiver.

4.4.3.4.1. Linear Backprojection.

The first method to be evaluated is the straightforward linear backprojection technique, the resultant image is shown in Figure 4.15 on the next page with the defect boundary marked. This image gives a reasonably good indication of the defect size, shape and position. The linear cross-correlation with the ideal image having a value of 0.580.

4.4.3.4.2. Filtered Backprojection.

The image reconstructed utilising the filtered backprojection algorithm can be seen in Figure 4.16 on the next page. This image gives a slightly better indication of the defect position, shape and size compared to linear backprojection image, this is reflected in the correlation scores for the two images. The linear backprojection image has a correlation value of 0.580 whereas the filtered backprojection image has a value of 0.610.

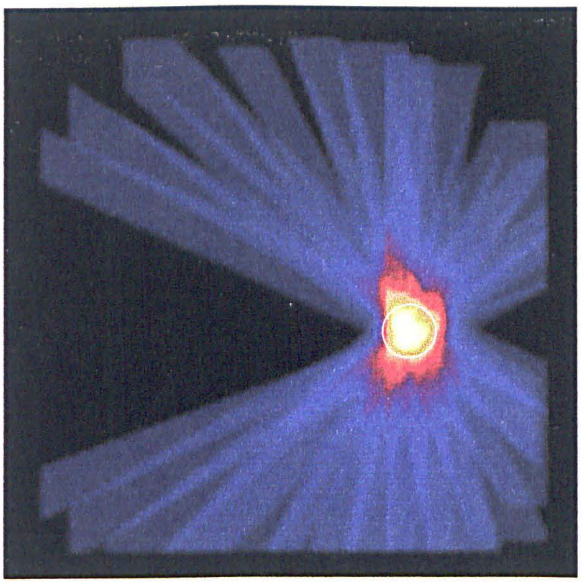


Figure 4.15. *Transmission tomography image reconstructed using the linear backprojection algorithm. The transmitter separation has been chosen to be 10^0 and a receiver separation of 5^o (image scale 1:1).*

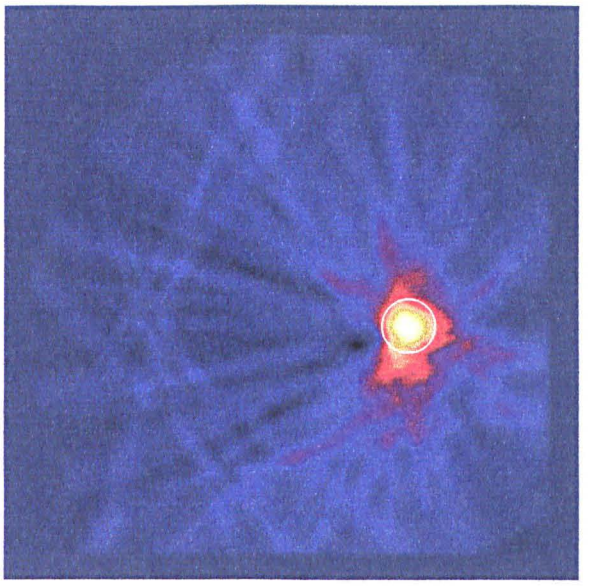


Figure 4.16. *Transmission tomography image reconstructed using the filtered backprojection algorithm. The transmitter separation has been chosen to be 10^0 and a receiver separation of 5^o (image scale 1:1).*

4.4.3.5. Fused Images.

In order to assess the improvement in the image quality due to image fusion linear cross-correlation with the ideal image was employed. In order to put the improvement in the image quality in context, the average of the constituent images has been generated and assessed using linear cross-correlation. The four images used in the fusion are the Pulse-Echo Reflection, Pitch-Catch Reflection, Time-of-flight Diffraction and Filtered Backprojection Transmission images, the average of these four images is shown in Figure 4.17.

For the generation of the fused image the second fusion function was utilised with the Kapur, Wong and Sahoo thresholding method. The zero cut-off was chosen to be $0.75T_c$, where T_c is the computed threshold value. The fused image can be seen in Figure 4.18. The cross-correlation values for all of the images used in the fusion are given in Table 2. The improvement in the linear cross-correlation values between the average and the fused image is 0.112. The improvement is relatively small compared to the 1.2 mm inclusion but still represents a worthwhile improvement in the image quality.

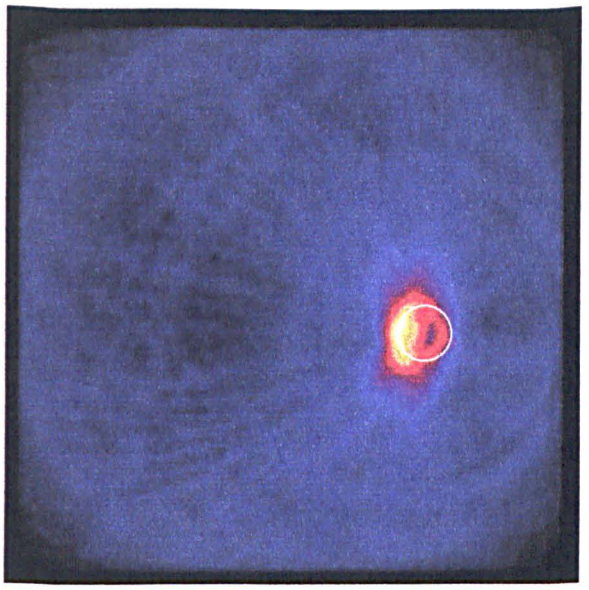


Figure 4.17. *Average image generated from the four tomographic images to be used in the image fusion: Pulse-Echo Reflection Tomography, Pitch-Catch Reflection Tomography, Time-of-flight Diffraction Tomography and Filtered Backprojection Transmission Tomography (image scale 1:1).*

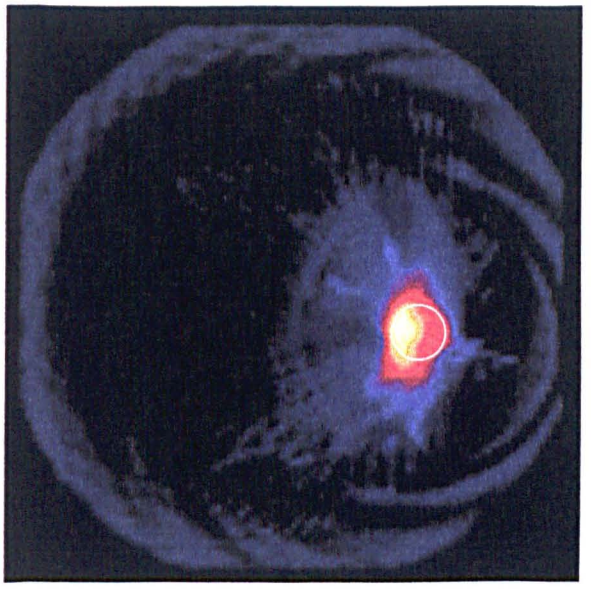


Figure 4.18. *Optimal fused tomographic image generated from the four tomographic images shown in the Figures 4.13,4.14,4.15 and 4.16, respectively. The 2nd fusion function used with the Kapur, Sahoo and Wong thresholding technique. The zero cut-off was chosen to be $0.75T_c$ for all four images (image scale 1:1).*

Tomographic Image.	Linear Cross-correlation with Ideal Image, r.
Pulse-Echo Reflection.	0.229
Pitch-Catch Reflection.	0.313
Time-of-flight Diffraction.	0.335
Linear Backprojection Transmission.	0.580
Filtered Backprojection Transmission.	0.610
Summation.	0.604
Fused Image.	0.716

Table 2. *Correlation values for all of the tomographic images reconstructed. These correlation values were computed using Pearson's linear correlation equation (4.4-1).*

4.4.4. Simulated Planar Defect.

The third type of representative defect that was investigated is a simulated planar type defect. The defect is a 25×2mm slot machined through the 75mm diameter aluminium cylinder, its orientation and position being shown in Figure 4.19.

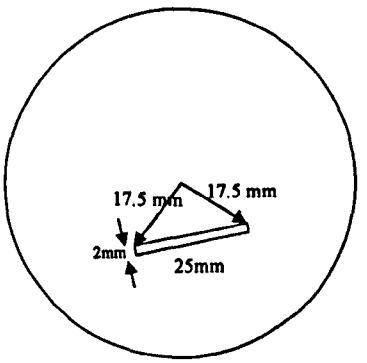


Figure 4.19. *The position of the simulated 25×2 mm planar type defect within the cylindrical test specimen.*

4.4.4.1. Pulse-Echo Reflection Tomography.

The Pulse-Echo Reflection tomography image reconstruction is shown in Figure 4.20 with the planar defect's actual boundary being displayed in white. As before, the image was reconstructed with A-scans collected with a transmitter angular separation of 5^{o}.

The image gives a good indication of the position of the defect end points but has not reconstructed the body of the defect with the same level of confidence, however it is still evident that the defect is planar in nature. In addition there are a couple of artifacts in the image, the biggest artifact is on the backwall of the test specimen so will not interfere with defect detection. However, the other artifact which is probably caused by diffraction from the ends of the slot and is within the test specimen so could interfere with defect reconstruction. The linear cross-correlation value with the ideal image is 0.427.

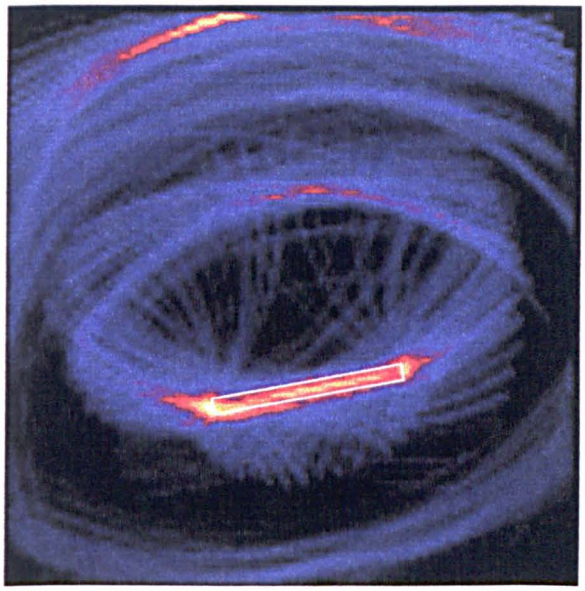


Figure 4.20. *Pulse-Echo Reflection tomography image of the simulated planar defect with a transmitter separation of 5^{o}. The actual defect boundary is marked on the image in white (image scale 1:1).*

4.4.4.2. Pitch-Catch Reflection Tomography.

The Pitch-Catch Reflection image reconstructed of the simulated planar defect is given in Figure 4.21. The image was reconstructed with A-scan data acquired with an angular transmitter separation of 5^{o} and a constant transmitter to receiver separation of 40^{o}.

As before, the defect boundary has been marked on the image in white for comparison purposes. The defect endpoints can clearly be seen, the body of the defect is more pronounced than in the corresponding Pulse-Echo Reflection tomography image although it is still not reconstructed in its entirety, an artifact is also present in the image in roughly the same location as the Pulse-Echo image. The improved reconstruction of the defect body is reflected in the higher linear cross-correlation score of 0.461.

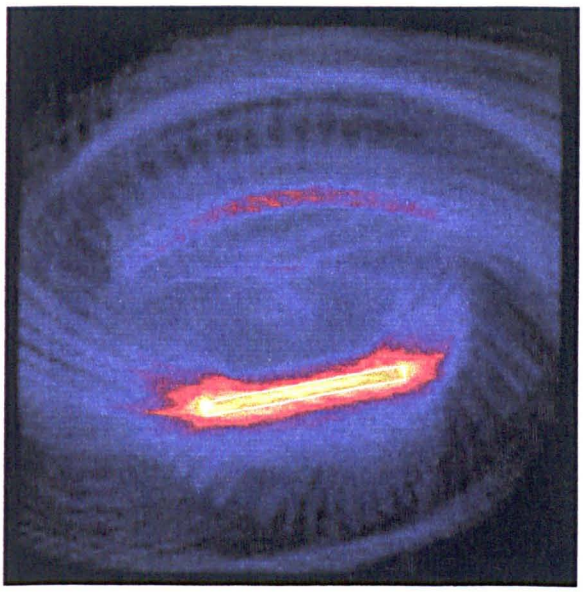


Figure 4.21. *Pitch-Catch Reflection image of the planar type defect. The transmitter angular separation was chosen to be 5^{o} and the separation between the transmitter and receiver was 40^{o} (image scale 1:1).*

4.4.4.3. Time-of-flight Diffraction Tomography.

The Time-of-flight Diffraction image is shown in Figure 4.22. For this image, four transmit locations equally spaced around the object circumference were chosen. The ultrasonic energy scattered by the defect was then collected, over the circumference of the sample, at 10^{o} intervals.

The Time-of-flight Diffraction image identifies the end points of the defect but does not give a definite indication as to whether there are two inclusion type defects present or just a single planar defect. This poor reconstruction of the defect body can be attributed to the small number of transmitter location chosen in order to give comparable reconstruction times with the other reconstruction algorithms.

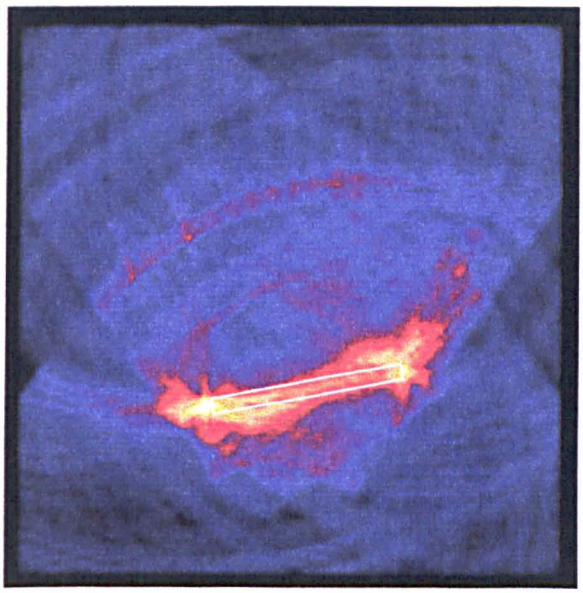


Figure 4.22. *Time-of-flight Diffraction tomography image. The transmitter separation is 90^{o} and the receiver separation is 10^{o}. The defect boundary is marked in white (image scale 1:1).*

4.4.4.4. Transmission Tomography.

Transmission images of the planar defect were reconstructed using both the linear and filtered backprojection algorithms presented in Chapter 2. The same set of A-scan data was used for both the image reconstruction algorithms, the A-scan data being collected with a transmitter separation of 10^{o} and a 5^{o} receiver separation. For

the reconstruction of this image it was decided that a shadow region is present if the received signal amplitude was less than 30% of the amplitude expected at the backwall when no defect is present between the transmitter and receiver.

4.4.4.4.1. Linear Backprojection.

The first method to be evaluated is the straightforward linear backprojection technique and the reconstructed image is shown in Figure 4.23. This method does not give a good indication of the defect size or shape. The fact that the defect appears to be shorter than it actually is can be explained by the ultrasound diffracting around the ends of the defect. If the transmitter and receiver separation were decreased then it is expected that the defect image would give a more accurate indication of the defect shape, but at the expense of reconstruction time. It does however give a reasonably good indication of the defect orientation. The linear cross-correlation with the ideal image having a value of 0.320.

4.4.4.4.2. Filtered Backprojection.

The image reconstructed utilising the filtered backprojection algorithm can be seen in Figure 4.24. This image clearly gives a poorer indication of the defect position, shape and size than the linear backprojection image. This can be explained by the choice of the filter function. The design of the filter function is based on the shadow cast by an ideal disc reflector. This planar defect is too far removed from this assumption resulting in the poor defect reconstruction. This is reflected in the correlation scores for the two images; the linear backprojection image has a correlation value of 0.320, whereas the filtered backprojection image has a value of 0.280.

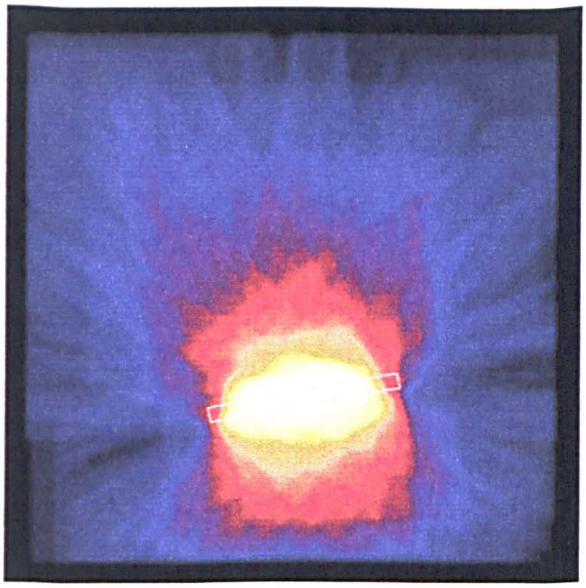


Figure 4.23. *Linear backprojection Transmission Tomography image with the defect boundary indicated by the rectangle. The transmitter separation was chosen to be 10^{o} with a 5^{o} receiver separation (image scale 1:1).*

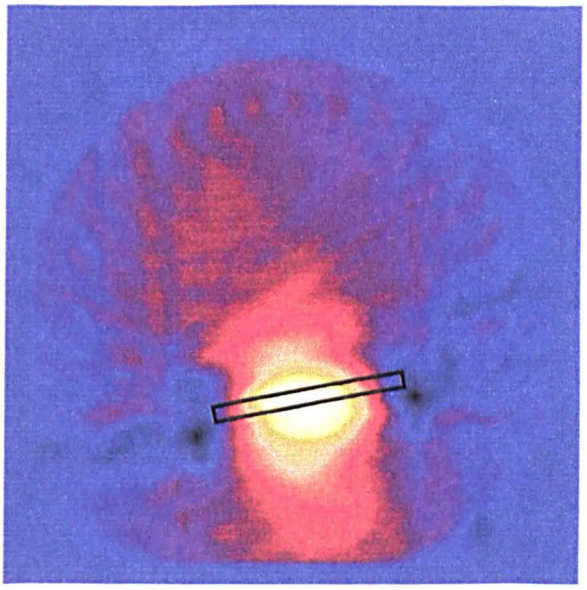


Figure 4.24. *Filtered Backprojection Transmission image of the planar type defect utilising the same A-scan data as the Linear Backprojection reconstruction (image scale 1:1).*

4.4.4.5. Fused Images.

In order to assess the improvement in the image quality due to image fusion, linear cross-correlation with the ideal image was employed. In order to put the improvement in the image quality in context the average of the constituent images has been generated and assessed using linear cross-correlation. The four images used in the fusion are the Pulse-Echo Reflection, Pitch-Catch Reflection, Time-of-flight Diffraction and Linear Backprojection Transmission tomography images, the average of these four images is shown in Figure 4.25.

The fused image generated can be seen in Figure 4.26. For the generation of the fused image the third fusion function was utilised with the Pun thresholding method. The zero cut-off was chosen to be $0.75T_c$, where T_c is the computed threshold value. The ramp fusion variable, a, was chosen to have a value of 4.

The cross-correlation values for all of the images used in the fusion are given in Table 3. The improvement in the linear cross-correlation values between the average and the fused image is 0.093, this appear to be very small but upon visual inspection of the image it is evident that the defect image is close to the actual defect (marked in black).

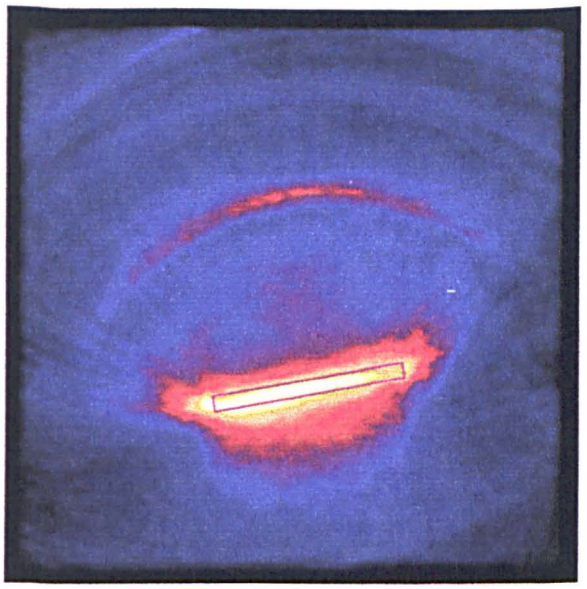


Figure 4.25. *Average of four of the reconstructed images: Pulse-Echo Reflection Tomography, Pitch-Catch Reflection tomography, Time-of-flight Diffraction Tomography and Linear Backprojection Transmission tomography (image scale 1:1).*

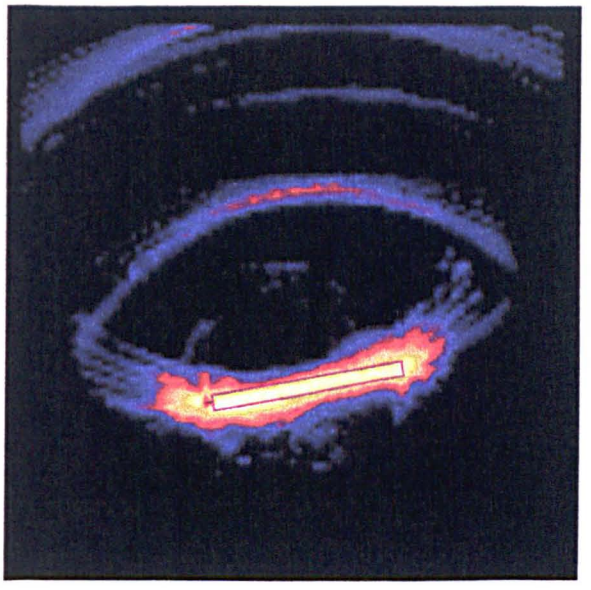


Figure 4.26. *Optimal fused tomographic image generated from the tomographic images shown in the Figures 4.20,4.21 and 4.22. The 3rd fusion function used was and the Pun thresholding method was used with a $0.75T_c$ zero cut-off with a = 4 (image scale 1:1).*

Tomographic Image.	Linear Cross-correlation with Ideal Image, r.
Pulse-Echo Reflection.	0.427
Pitch-Catch Reflection.	0.461
Time-of-flight Diffraction.	0.341
Linear Backprojection Transmission.	0.320
Filtered Backprojection Transmission.	0.280
Average.	0.443
Fused Image.	0.536

Table 3. *Correlation values for all of the tomographic images reconstructed. These correlation values were computed using Pearson's linear correlation equation (4.4-1).*

4.4.5. Investigation of the Imaging System Resolution.

The first aim of this section is to experimentally investigate the resolution of the first three tomographic reconstruction techniques detailed in Chapter 2, notably Pulse-Echo Reflection tomography, Pitch-Catch Reflection tomography and Time-of-flight Diffraction tomography. No transmission tomography image will be reconstructed due to the small size of the defects when compared to the transducer beamwidth. The second aim of this section is to assess the improvement, if any, in image resolution from the application of the fuzzy logic fusion technique.

In order to investigate the image resolution a 75 mm diameter aluminium test specimen containing a line of inclusions with differing separations, as shown in Figure 4.27, has been used.

The distances between the circular inclusions are as follows;

- distance between the centres of defect number 1 and 2 is 2.4 mm
- distance between the centres of defect number 2 and 3 is 3.6 mm
- distance between the centres of defect 3 and 4 is 4.8 mm
- distance between the centres of defect 4 and 5 is 6 mm.

The speed of sound in aluminium is 6420 ms^{-1}, which corresponds to a wavelength of approximately 1.28mm. Therefore, the two defects that are of primary interest are the two that have a wavelength spacing between their boundaries, i.e. simulated defects numbered 1 and 2 on the figure. It was determined in Chapter 2 that the tomographic reconstruction algorithms should give near wavelength resolution. However, in practice this is often not corroborated due to a break down in the assumptions made in the derivation, and more importantly the relatively small number of transmitter locations chosen (the Point Spread Function calculations were based on a completely sampled aperture).

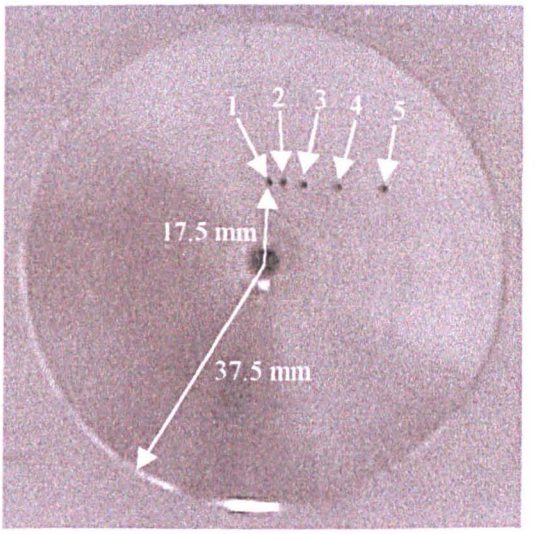


Figure 4.27. *Position of simulated defects within the test specimen. The longitudinal velocity in aluminium is 6420 ms^{-1}, which corresponds to a wavelength of 1.28 mm. All five of the holes through the test specimen have a diameter of 1.2 mm, The distances between the hole centres are as follows: 1 to 2: 2.4 mm, 2 to 3: 3.6 mm, 3 to 4: 4.8 mm and finally 4 to 5: 6 mm.*

4.4.5.1. Pulse-Echo Reflection Tomography.

The Pulse-Echo Reflection tomography image reconstruction is shown in Figure 4.28 with defect boundaries marked in black and white. The image was reconstructed with transmitter positions separated by a 5^{o} angular separation.

All of the defects present have been reconstructed, some more strongly than others. This is due to a number of factors. The two defects on the far left of the image (4 and 5) have not been imaged clearly. This can be explained by considering the orientation of the other defects in relation to defects 4 and 5, (i.e. the other defects (1,2 and 3) prevent isonification of the defects (4 and 5) from a significant number of transmitter locations). Of more concern is the fact that the two most closely spaced defects (1 and 2) have been reconstructed as a single defect, thus implying that the reconstruction does not give wavelength resolution for this reconstruction algorithm. The linear cross-correlation value with the ideal image is 0.273.

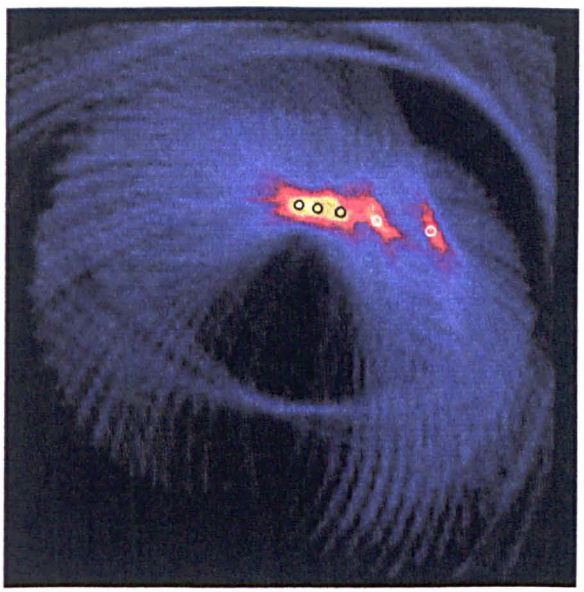


Figure 4.28. *Pulse-Echo Reflection Tomography image. Reconstructed with A-scan data obtained with a transmitter separation of 5^{o}. The defect boundaries are marked in black and white (image scale 1:1).*

4.4.5.2. Pitch-Catch Reflection Tomography.

The Pitch-Catch Reflection image reconstruction of the line of holes is shown in Figure 4.29. This image was reconstructed with an angular transmitter separation of 5^{o} and a constant transmitter to receiver separation of 40^{o}.

This image exhibits the same characteristics as the Pulse-Echo Reflection image, only the failure to reconstruct defects 4 and 5 being more pronounced. However, there is a slight improvement in the reconstruction of the two closely spaced defects, although it is still not clearly evident that if there is a single or multiple defects present. The linear cross-correlation value was calculated to be 0.292.

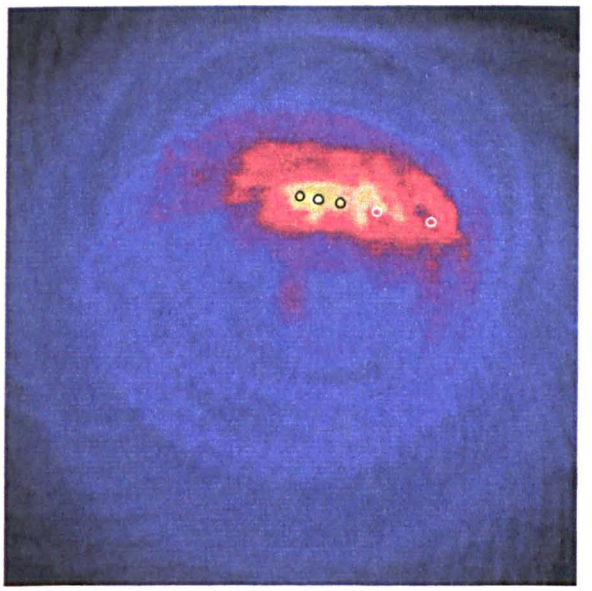


Figure 4.29. *The Pitch-Catch Reflection tomography image. The A-scan data was obtained with a 5^{o} transmitter separation and a 40^{o} transmitter to receiver separation. The defect boundaries are marked in black and white (image scale 1:1).*

4.4.5.3. Time-of-flight Diffraction Tomography.

The Time-of-flight Diffraction image is shown in Figure 4.30. For this image, four transmit locations equally spaced around the object circumference were chosen. The ultrasonic energy scattered by the defect was then collected, over the circumference of the sample, at 10^{o} intervals.

The Time-of-flight Diffraction tomography image gives a poor reconstruction of all the defects apart from one (defect 2). The primary reason for this is the choice of transmitter locations. As only four transmitter locations were chosen the position of the defects with respect to the transmitter becomes important. However the image does indicate that further examination of the region where the defects are present is worthwhile.

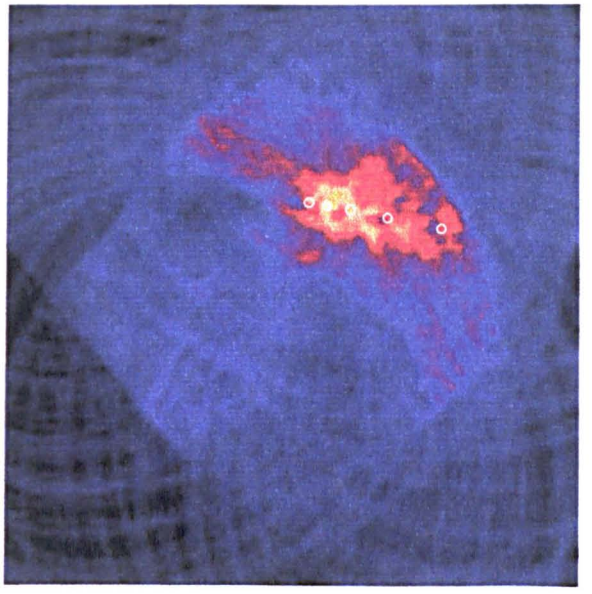


Figure 4.30. *Time-of-flight Diffraction tomography image. The transmitter separation was 90° and the receiver was 10°. The defect boundaries are marked in white (image scale 1:1).*

4.4.5.4. Fused Images.

In order to assess the improvement in the image quality due to image fusion linear cross-correlation with the ideal image will be employed. In order to put the improvement in the image quality in context the average of the constituent images has been generated and assessed using linear cross-correlation. The two images used in the fusion are the Pulse-Echo Reflection and the Pitch-Catch Reflection tomography images, the average of these two images is shown in Figure 4.31.

The fused image generated can be seen in Figure 4.32. For the generation of the fused image the second fusion function was utilised with the Ostu thresholding method. The zero cut-off was chosen to be 0.

The cross-correlation values for all of the images used in the fusion are given in Table 4. The improvement in the linear cross-correlation values between the average and the fused image is 0.188, which represents a worthwhile improvement in the image quality.

For this set of simulated defects the resolution of the fused image is also extremely important since the three reconstructed tomographic images gives less than wavelength resolution. Upon closer examination of the defect region it is clear that two defect are present in the image, thus the resolution of the image has been improved.

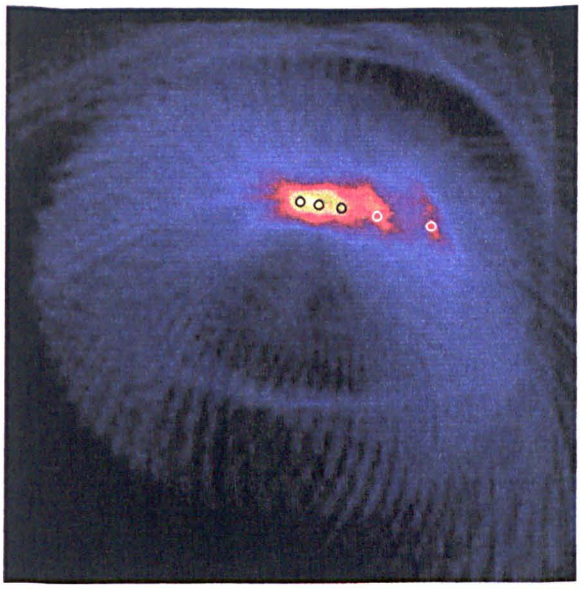


Figure 4.31. *Average of the Tomographic images shown in figure 4.28 and 4.29 (image scale 1:1).*

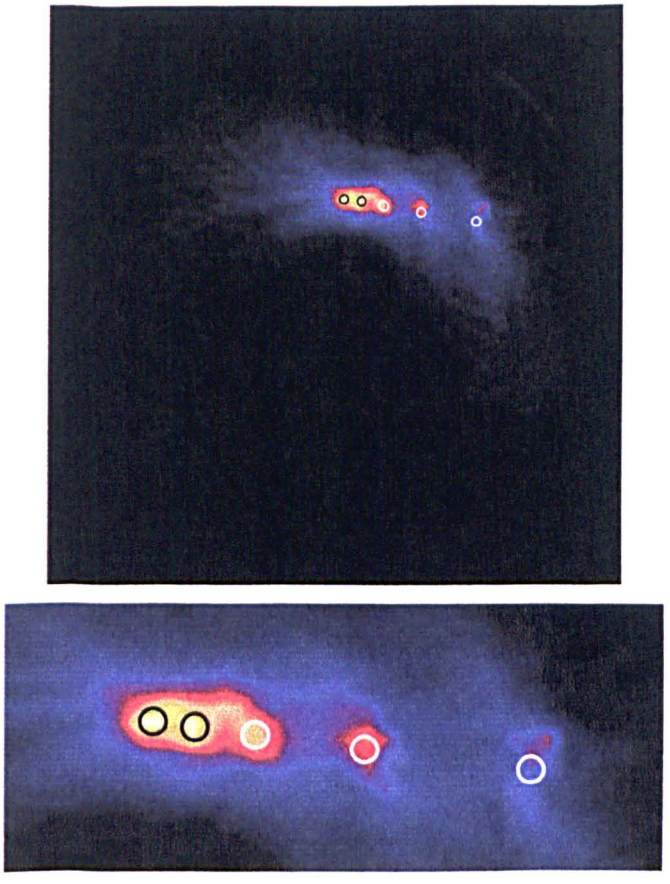


Figure 4.32. *Optimal fused image of the two Reflection tomography images in Figure 4.28 and Figure 4.29, respectively. The fusion function utilised was the second function with a zero cut-off value of* $0T_c$ *and the Ostu threshold method (image scale 1:1). The second image being a magnified section of the first image showing that there are indeed five distinct defects present in the image.*

Tomographic Image.	Linear Cross-correlation with Ideal Image, r.
Pulse-Echo Reflection.	0.273
Pitch-Catch Reflection.	0.292
Time-of-flight Diffraction.	0.160
Average.	0.323
Fused Image.	0.511

Table 4. *Correlation values for all of the tomographic images reconstructed. These correlation values were computed using Pearson's linear correlation equation (4.4-1).*

4.5. Single Bounce Image Enhancement.

The proceeding section detailed the optimum scenario for image reconstruction, that is the generation of defect images where A-scan data can be acquired from any position on the object circumference. However, frequently in practice this will not be the case. Either the component being examined will not have a regular geometry making access to all aspects of the component difficult or its situation may restrict access. In order to improve the practical applicability of the image reconstruction, the application of single bounce image enhancement will now be examined for the two Reflection tomography and Time-of-flight Diffraction tomography reconstruction algorithms.

Initially the single bounce image enhancement method will be validated by examining the scenario where full access to the object circumference is available. Once the effectiveness of the technique has been demonstrated the partial access scenario will be considered. This will be undertaken in Section 4.6.

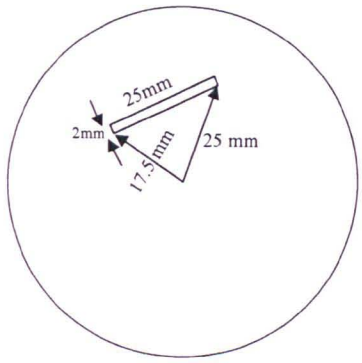
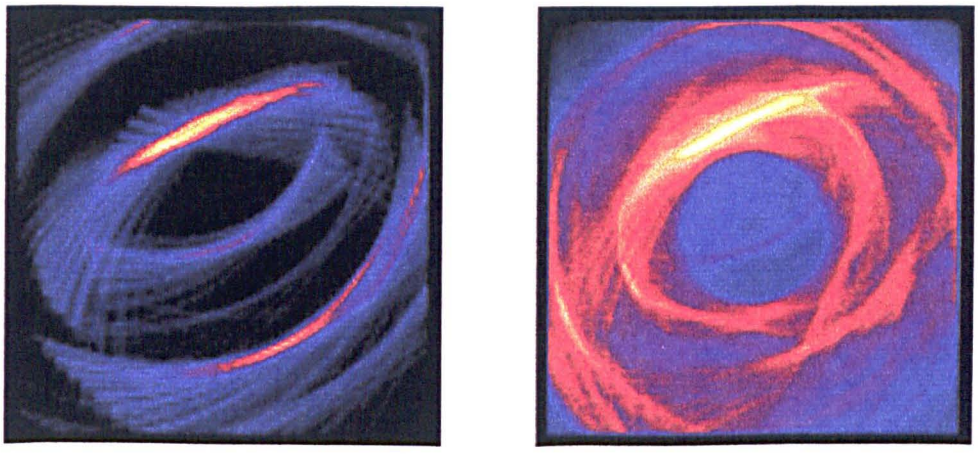


Figure 4.33. *Position of the 25mm × 2mm slot used to assess the single bounce image enhancement method. This simulated defect will also be used to assess the performance of the single bounce image enhancement method where only partial access to the test specimen is available.*

4.5.1. Pulse-Echo Reflection Tomography.

The first of the tomographic image reconstruction methods to be considered is Pulse-Echo Reflection tomography. Both the Pulse-Echo Reflection image and the single bounce image reconstructed from A-scan data acquired at 72 transmitter locations equally spaced around the circumference of the test specimen are shown in figure 4.34 (a) and (b), respectively.



(a) (b)

Figure 4.34. *Pulse-Echo Reflection tomography Images for the evaluation of the single bounce image enhancement feature. (a) Pulse-Echo Reflection Image reconstructed with data taken at 5° angular separation (b) Single Bounce image reconstructed using the same A-scan data as image (a). (image scale 1:2)*

Both of the images in Figure 4.34 give a good indication of the defect orientation, size and position. However, the single bounce image does contain a significant amount of noise. In addition, both images contain a number of artifacts. In Figure 4.34(a) the artifacts being caused by multiple reflections form the defect to the receiver. In Figure 4.34(b) the artifacts are caused by the focusing effect of the test specimen backwall. The combination of these two images is shown in Figure 4.35. This image gives an improved indication of defect shape, size and position, this is reflected in an increase in the linear cross-correlation value for the combined image over the Pulse-Echo Reflection tomography image, as given in Table 5. However, there are still a couple of small artifacts in the final image, although the they have been reduced in magnitude from the two images presented in Figure 4.34.

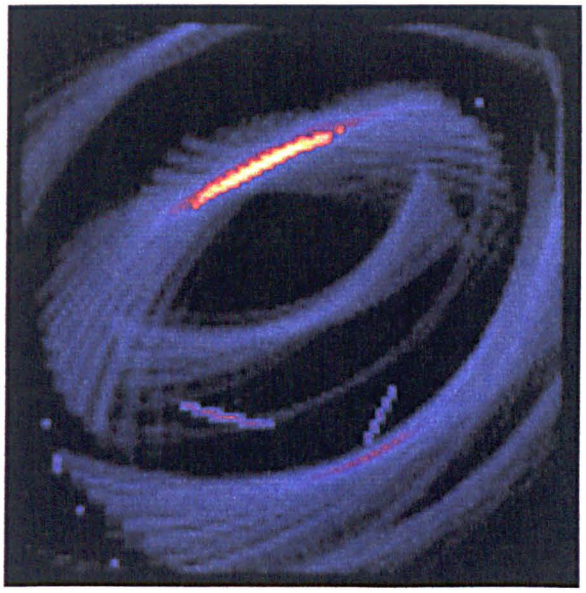


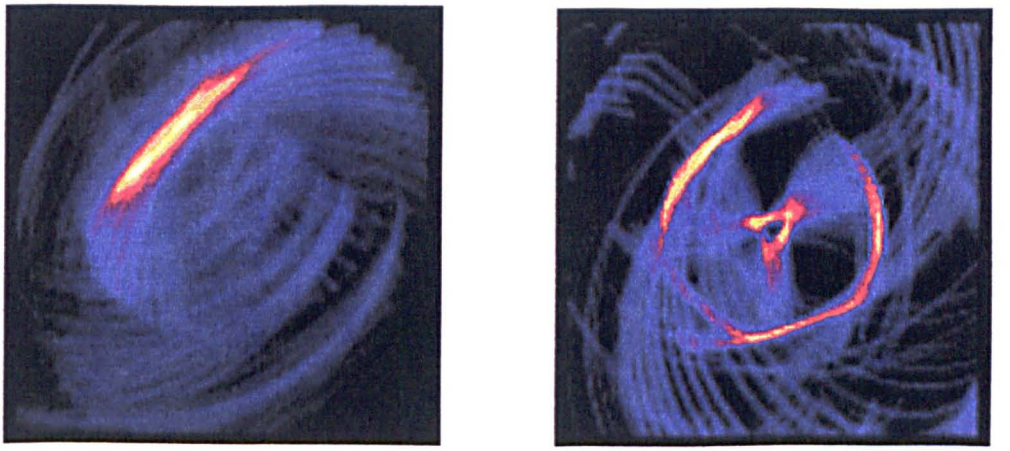
Figure 4.35. *The combination of the Pulse-Echo Reflection tomography and the Single Bounce Image. With the actual defect boundary marked (image scale 1:1).*

Tomographic Image.	Linear Cross-correlation with ideal image, r.
Pulse-Echo Reflection.	0.394
Single Bounce.	0.272
Combination.	0.534

Table 5. *Correlation values for all of the tomographic images reconstructed. These correlation values were computed using Pearson's linear correlation equation (4.4-1).*

4.5.2. Pitch-Catch Reflection Tomography.

The next tomographic reconstruction technique to be examined is the Pitch-Catch Reflection technique. The Pitch-Catch Reflection image and the single bounce image reconstructed from A-scan data acquired with a transmitter separation of 5° and a transmitter to receiver angular separation of 60° can be seen in Figure 4.36 (a) and (b), respectively.



(a) (b)

Figure 4.36. *Pitch-Catch Reflection tomography Images for the evaluation of the single bounce image enhancement feature. (a) Pitch-Catch Reflection Image reconstructed with data acquired with a 5° transmitter angular separation and a 60° transmitter to receiver separation (b) Single Bounce image reconstructed using the same A-scan data as image (a). (image scale 3:4).*

The images in Figure 4.36(a) gives a good indication of the defect orientation, size and position, however, the single bounce image in Figure 4.36(b) contains a number of artifacts. The artifacts are a consequence of the focusing effect of the backwall and the transmitter / receiver separation by choosing a different separation these artifacts could be removed. The combination of these two images is shown in Figure 4.37. This image gives an improved indication of defect shape, size and position, however, some of the artifacts contained in the single bounce image are present in the final image. As before the single bounce image (Figure 4.36(b)) does contain significantly more noise than the signal. The improvement in defect characterisation is reflected in the increase in the linear cross-correlation values for the combined image over the Pitch-Catch Reflection image, as given in Table 6.

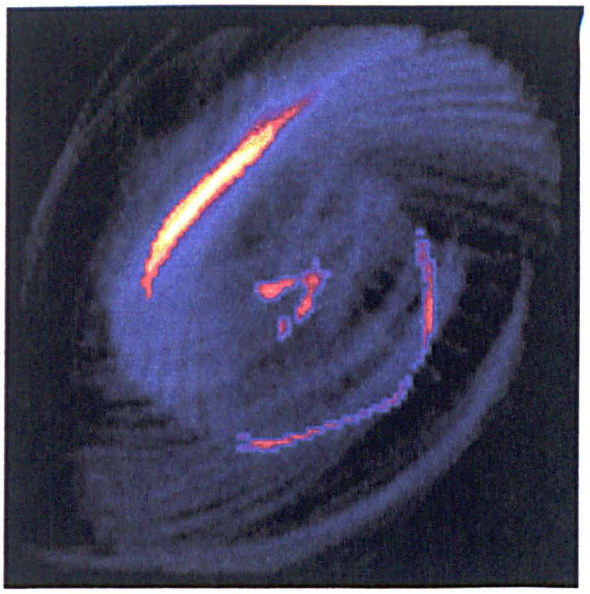


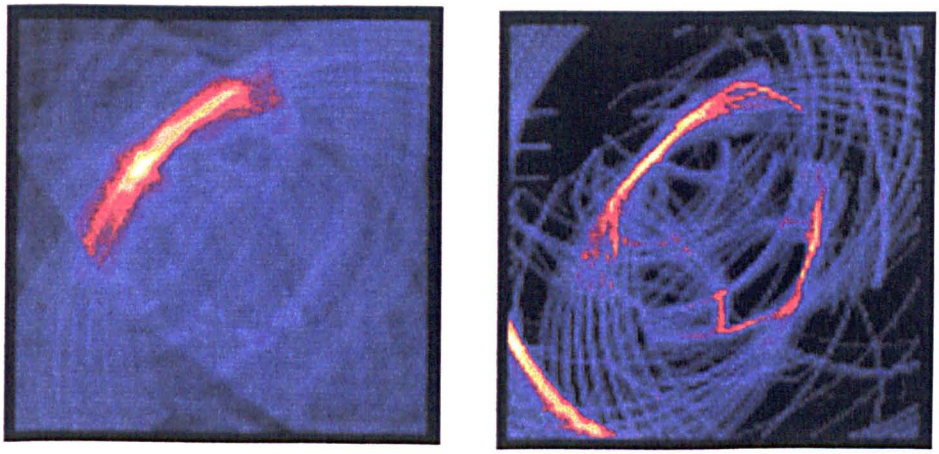
Figure 4.37. *The combination of the Pitch-Catch Reflection Tomography and the Single Bounce Images. The actual defect boundary is marked in black (image scale 1:1).*

Tomographic Image.	Linear Cross-correlation with ideal image, r.
Pitch-Catch Reflection.	0.402
Single Bounce.	0.303
Combination.	0.501

Table 6. *Correlation values for all of the tomographic images reconstructed. These correlation values were computed using Pearson's linear correlation equation (4.4-1).*

4.5.3. Time-of-flight Diffraction Tomography.

The Time-of-flight Diffraction tomography image is shown in Figure 4.38(a) and the single bounce image is shown in Figure 4.38(b). For these images, four transmit locations equally spaced around the object circumference were chosen. The ultrasonic energy scattered by the defect was then collected, over the circumference of the sample, at 10^{o} intervals.



(a) (b)

Figure 4.38. *Time-of-flight Diffraction tomography Images for the evaluation of the single bounce image enhancement feature. (a) Time-of-flight Diffraction Image reconstructed with data acquired with a 90^{o} transmitter angular separation and a 5^{o} receiver separation (b) Single Bounce image reconstructed using the same A-scan data as image (a). (image scale 1:2).*

Both of the images in Figure 4.38 give a poor indication of the defect orientation, size and position and contain a number of artifacts. The combination of these two images is shown in Figure 4.39. This image gives a slightly improved indication of defect shape, size and position, with the artifacts contained within the test specimen being decreased in severity. As before, the single bounce image (Figure 4.38(b)) does contain significantly more noise than the Time-of-flight Diffraction image. The improvement in defect characterisation is reflected in the increase in the linear cross-correlation values for the combined image over the Time-of-flight Diffraction image, as given in Table 7. However, the single bounce images contain a number of artifacts, these are due to the fact that the sound reflected by the defect could have originated from a number of locations within the test specimen. Generally these tend to disappear when the tomographic and single bounce images are combined as is the case here, the situation may not be so clear in specimens containing multiple defects so care will need to be taken when using this image enhancement technique.

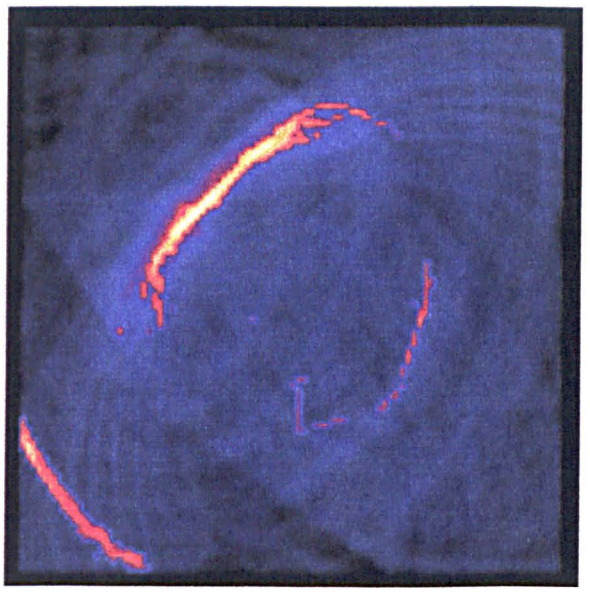


Figure 4.39. *The combination of the Time-of-flight Diffraction Tomography and the Single Bounce Images (image scale 1:1).*

Tomographic Image.	Linear Cross-correlation with ideal image, r.
Time-of-flight Diffraction.	0.402
Single Bounce.	0.309
Combination.	0.470

Table 7. *Correlation values for all of the tomographic images reconstructed. These correlation values were computed using Pearson's linear correlation equation (4.4-1).*

4.6. The Effect of Partial Access to the Test Specimen on Image Reconstruction.

Now that the operation of the single bounce image enhancement technique has been validated for complete access to the test specimen's circumference the next stage is to investigate its performance when only partial access to the object circumference is available. The region of access to the test specimen is as shown in figure 4.40.

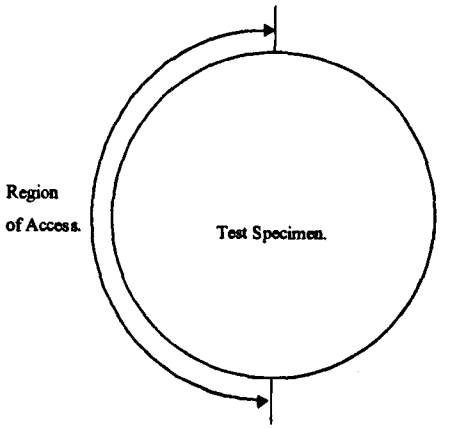


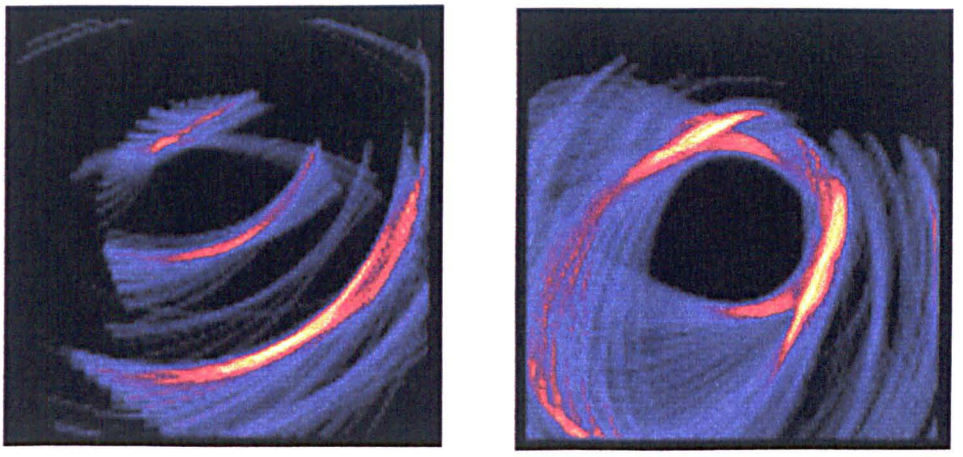
Figure 4.40. *Region of access to test specimen for the evaluation of the single bounce image enhancement technique.*

4.6.1. Pulse-Echo Reflection Tomography.

The first of the tomographic image reconstruction methods to be considered is Pulse-Echo Reflection tomography. The Pulse-Echo Reflection image and the single

bounce image reconstructed from A-scan data acquired at 36 transmitter locations equally along the access region of the test specimen are shown in figure 4.41(a) and (b), respectively.

The Pulse-Echo Reflection tomography image, shows multiple reflections from the defect. This characteristic is not present in the single bounce image. However the single bounce image does indicate two possible locations for a large planar defect. The next step is to combine the two images to determine if there is any improvement in defect characterisation.



(a) (b)

Figure 4.41. *Pulse-Echo Reflection tomography Images for the evaluation of the single bounce image enhancement feature. (a) Pulse-Echo Reflection Image reconstructed with data taken at 5° angular separation (b) Single Bounce image reconstructed using the same A-scan data as image (a). (image scale 1:2)*

The combination of these two images is shown in Figure 4.42. This image gives an improved indication of defect shape, size and orientation. This is born out by the linear cross-correlation values given in Table 8. However, the image does contain two artifacts. One is on the specimen backwall so can be ignored, the other artifact can be show to be an artifact by examining the Pulse-Echo Reflection image. There is no defect at the artifacts position within the pulse-echo image so the artifact can be identified as such.

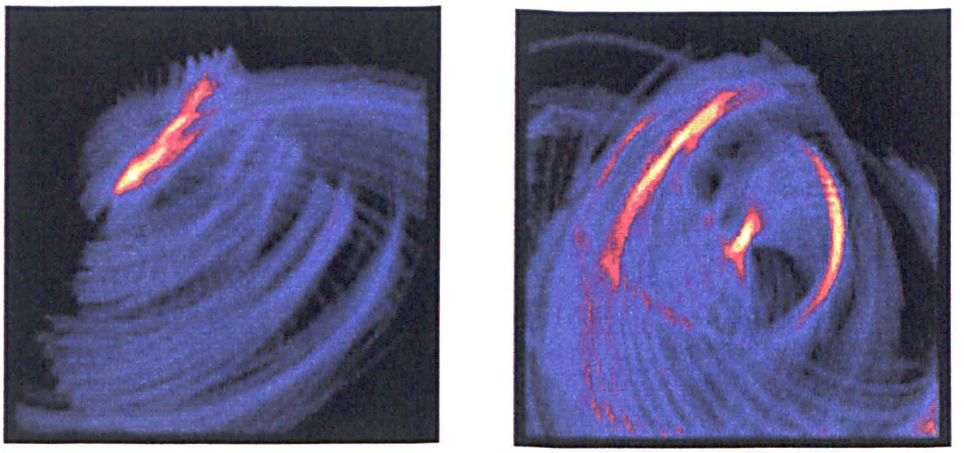
Figure 4.42. *The combination of the Pulse-Echo Reflection Tomography and the Single Bounce Images (image scale 1:1).*

Tomographic Image.	Linear Cross-correlation with ideal image, r.
Pulse-Echo Reflection.	0.104
Single Bounce.	0.292
Combination.	0.534

Table 8. *Correlation values for all of the tomographic images reconstructed. These correlation values were computed using Pearson's linear correlation equation (4.4-1).*

4.6.2. Pitch-Catch Reflection Tomography.

The next tomographic reconstruction technique to be examined is the Pitch-Catch Reflection technique. The Pitch-Catch Reflection image and the single bounce image reconstructed from A-scan data acquired with a transmitter separation of 5° and a transmitter to receiver angular separation of 60° over the access region, these can be seen in Figure 4.43 (a) and (b), respectively.



(a) (b)

Figure 4.43. *Pitch-Catch Reflection tomography Images for the evaluation of the single bounce image enhancement feature. (a) Pitch-Catch Reflection Image reconstructed with data acquired with a 5^o transmitter angular separation and a 60^o transmitter to receiver separation (b) Single Bounce image reconstructed using the same A-scan data as image (a). (image scale 1:2).*

Both of the images in Figure 4.43 give an indication of the defect orientation, size and position. However, the image in Figure 4.43(a) is blurred and shows two closely spaced defects rather than a single planar defect. The image in Figure 4.43(b) contains a number of artifacts caused by the focusing effect of the test specimen backwall. The combination of these two images is shown in Figure 4.44. This image gives an improved indication of defect shape, size and position. The single bounce image (Figure 4.43(b)) does contain a number of artifacts, which can be identified as such, by examining the Pitch-Catch Reflection image and observing that the image does not contain any defects at the artifact positions. The improvement in defect characterisation is reflected in the increase in the linear cross-correlation value for the combined image over the Pitch-Catch Reflection image, as given in Table 9.

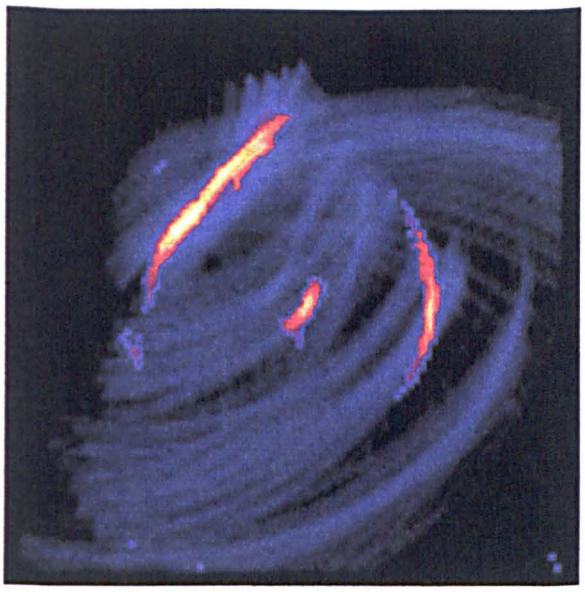


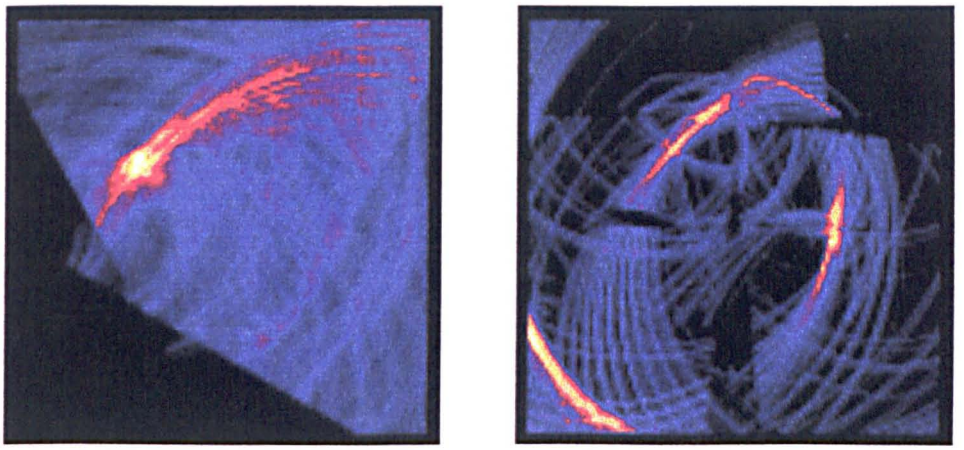
Figure 4.44. *The combination of the Pitch-Catch Reflection tomography and the Single Bounce Images (image scale 1:1).*

Tomographic Image.	Linear Cross-correlation with ideal image, r.
Pitch-Catch Reflection.	0.373
Single Bounce.	0.266
Combination.	0.458

Table 9. *Correlation values for all of the tomographic images reconstructed. These correlation values were computed using Pearson's linear correlation equation (4.4-1).*

4.6.3. Time-of-flight Diffraction Tomography.

The Time-of-flight Diffraction image is shown in Figure 4.45(a) and the single bounce image is shown in Figure 4.45(b). For these images, three transmit locations equally spaced along the region of access were chosen. The ultrasonic energy scattered by the defect was then collected, along the region of access, at 10^{o} intervals.



(a) (b)

Figure 4.45. *Time-of-flight Diffraction tomography Images for the evaluation of the single bounce image enhancement feature. (a) Time-of-flight Diffraction Image reconstructed with data acquired with a 90^{o} transmitter angular separation and a 10^{o} receiver separation over the region of access (b) Single Bounce image reconstructed using the same A-scan data as image (a). (image scale 1:2)*

Both of the images in Figure 4.46 give a poor indication of the defect orientation and position this is a consequence of the extremely small amount of data used to reconstruct the images. However, they do not give an indication that the defect is planar. The combination of these two images is shown in Figure 4.46. This image gives an improved indication of defect shape, size and position over the images in Figure 4.46. As with the previous reconstruction methods the single bounce image (Figure 4.46(b)) does contain two artifacts, one of which lies outwith the test specimen boundary, the other can be identified as an artifact as it is not present in the Time-of-flight Diffraction image. The improvement in defect characterisation is

reflected in the increase in the linear cross-correlation values for the combined image over the Time-of-flight Diffraction image, as given in Table 10.

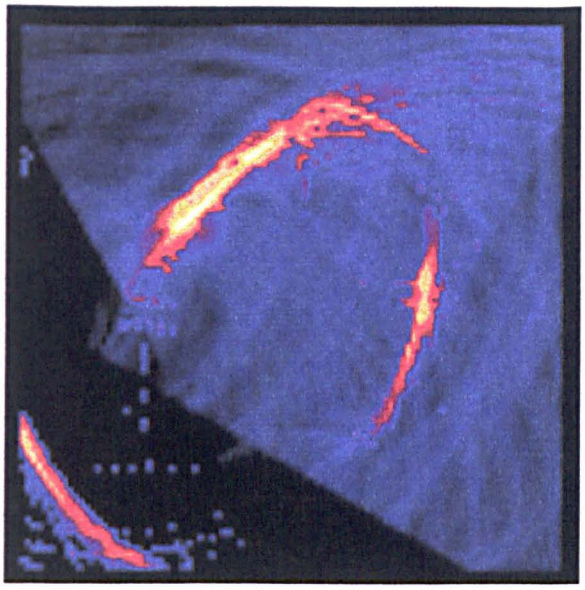


Figure 4.46. *The combination of the Time-of-flight Diffraction Tomography and the Single Bounce Images (image scale 1:1).*

Tomographic Image.	Linear Cross-correlation with ideal image, r.
Time-of-flight Diffraction.	0.279
Single Bounce.	0.242
Combination.	0.357

Table 10. *Correlation values for all of the tomographic images reconstructed. These correlation values were computed using Pearson's linear correlation equation (4.4-1).*

4.7. Effect of MLD/ MVD filtering on Tomographic Image Reconstruction.

The purpose of this Section is to evaluate the effect of the Maximum Likelihood Deconvolution (MLD)/ Minimum Variance Deconvolution (MVD) filtering presented in Chapter 3 on tomographic image reconstruction.

The effect on the image quality will be examined for pulse-echo A-scan data captured for a 1.2mm hole through an aluminium test specimen. The A-scan data was captured at 72 equally spaced transmit/receive locations around the test specimen with the same data acquisition set-up that was used in Section 4.2. The data set chosen was one that was particularly noisy, in order to demonstrate the effectiveness of the filter on A-scan data that it would be used on in practice. The image reconstructed using the Pulse-Echo Reflection tomography algorithm is shown in Figure 4.47.

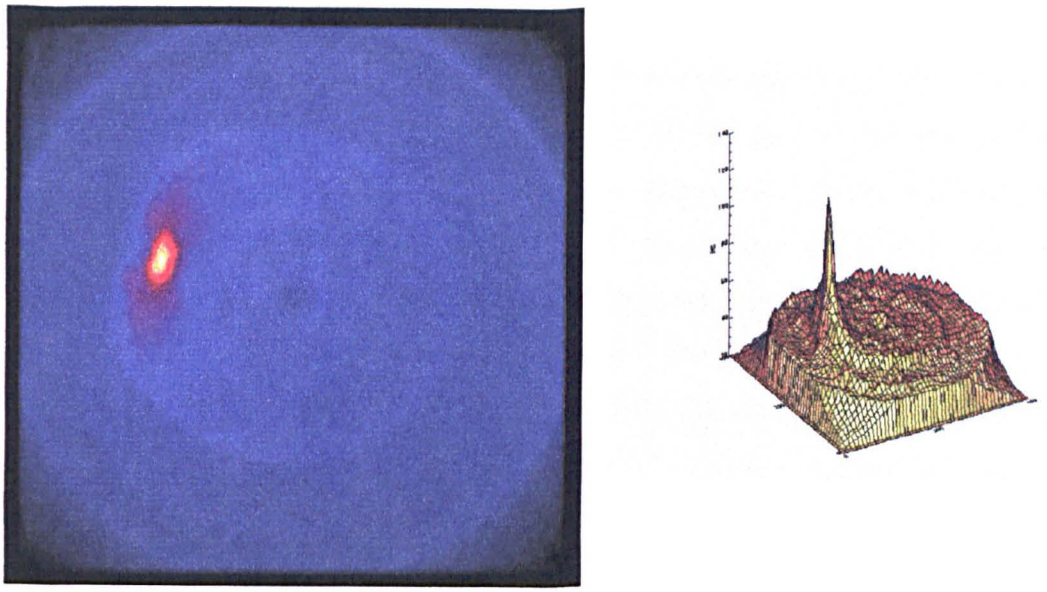


Figure 4.47. *The reconstructed Pulse-Echo Reflection Tomography image prior to MVD filtering of the raw data. The pseudo point spread function can be seen adjacent the reconstructed image (image scale 1:1).*

The initial estimates for the system parameters were obtained using the same procedure outlined in section 3.15. Prior to image reconstruction the A-scans filtered rectified and smoothed.

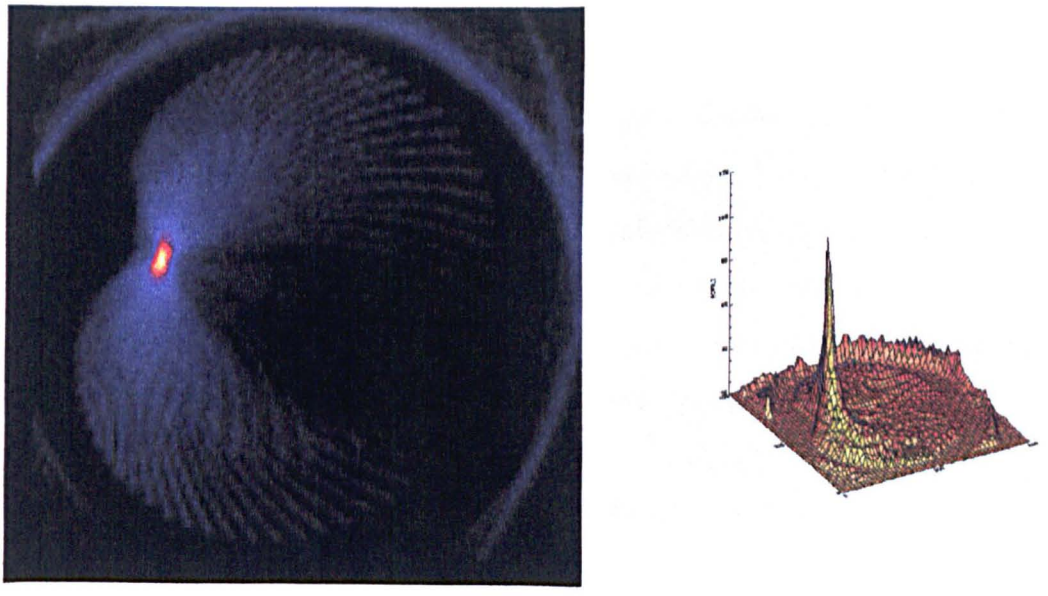


Figure 4.48. *The reconstructed Pulse-Echo Reflection Tomography image after MVD filtering of the raw data. The pseudo point spread function can be seen alongside the reconstructed image (image scale 1:1).*

The image reconstructed using the filtered data is shown in Figure 4.48 along with the pseudo point spread function. It is evident that the level of noise present in the image in Figure 4.48 is considerably lower than in Figure 4.47, the rate of decay of the main lobe is also higher, this is corroborated by the linear cross-correlation values for the images. The linear cross-correlation values for the images reconstructed with the filtered and raw data can be seen in Table 11.

Tomographic Image.	Linear Cross-correlation with ideal image, r.
Raw A-scan data.	0.194
MVD filtered data.	0.444

Table 11. *Linear Cross-correlation values for the images reconstructed with raw and filtered A-scan data.*

4.8. Concluding Summary.

The four components of the imaging system presented in Chapters 2 and 3 have been investigated experimentally. The four tomographic image reconstruction methods have been shown to give acceptable reconstruction of a number of representative defect types contained within aluminium cylindrical components. With the addition of the fuzzy logic fusion process it has been shown that characterisation of the defects considered is improved. The image resolution has been investigated for each of the reconstruction algorithms and it has been shown that they give poorer than wavelength resolution for the amount of raw data utilised. However, with fusion of the images it has been shown that wavelength resolution can be obtained.

The second imaging system component that has been evaluated is the Single Bounce Image Enhancement tool. This has been done for Pulse-Echo Reflection tomography, Pitch-Catch Reflection tomography and Time-of-flight Diffraction Tomography for complete and incomplete access to the test specimen circumference. It has been demonstrated that in the vast majority of cases the defect characterisation is improved.

The final section of this chapter investigates the effect that MLD / MVD filtering of the A-scans prior to image reconstruction has on the quality of the reconstructed images. It has been demonstrated that the image reconstructed with the filtered data gives improved characterisation of the defect.

Chapter 5.

CONCLUSIONS AND FUTURE WORK.

5.1. Conclusions.

The aim of this work has been to develop an ultrasonic imaging system to generate high quality images of defects embedded within engineering type components. This section summarises the imaging system components and draws conclusions on their overall performance.

Image Reconstruction.

The reconstruction of high resolution ultrasonic images using a number of different tomographic images of the defect and then fusing then together, was found to be justified from the results obtained. Four straightforward tomographic reconstruction algorithms were chosen. These being; Pulse-Echo Reflection tomography, Pitch-Catch Reflection tomography, Time-of-flight Diffraction tomography and Transmission tomography.

This approach was chosen over the development of a single reconstruction algorithm in order to maintain low theoretical complexity, therefore allowing low reconstruction times. Another advantage is that it presents the user with a highly flexible system for generation of high resolution images.

The reconstruction of tomographic images featuring a range of representative simulated defects has been presented for the tomographic reconstruction methods implemented. It has been shown that good characterisation of defects can be obtained using a number of different tomographic techniques even for relatively small amounts of ultrasonic data.

Image Enhancement.

In order to try to obtain the maximum amount of defect information from the A-scan data available, a novel image enhancement method that utilises the ultrasound reflected by the backwall of the test specimen has been successfully developed and was found to improve image quality in all situations examined. The image enhancement method is expected to be of greatest merit for scenarios where access to

the test specimen is restricted or the number of A-scans available is insufficient to obtain a satisfactory image of any defects present.

The single bounce image enhancement method has been implemented for the two Reflection tomography and the Time-of-flight Diffraction tomography algorithms. It has been demonstrated experimentally that the single bounce image enhancement generally results in an increase in the defect characterisation for both the cases where the access to the component is restricted and also the case when there is full access to the test specimen. In the case where full access to the specimen is available, the application of the single-bounce, image-enhancement method resulted in an average increase in the linear cross-correlation for the ideal image of approximately 0.10. However, for the case where access was restricted the average increase in the cross correlation was 0.20. This indicates that application of the single-bounce, image-enhancement method may not always be advised when the full data set is available in terms of the improvement to the image versus the increased image reconstruction times. However, for the case where access is restricted the results indicate that the image improvement resulting from the additional reconstruction usually produces an improvement in the defect characterisation.

Image Fusion.

In order to obtain high resolution, high quality tomographic defect images a novel fusion technique, based on fuzzy logic set theory, has been implemented. The fusion method utilises a set of the tomographic images generated using the reconstruction algorithms. The fusion method consists of utilising one of three fuzzy logic membership functions depending on the nature of the defect being examined to generate a single, high-resolution image of the defect. The exact nature of the membership functions is determined by using image thresholding theory on the images involved in the fusion process.

The operation of the fusion tool has been experimentally evaluated for a range of simulated defects embedded within cylindrical test specimens. It has been demonstrated that the resolution and quality of the images generated by utilising the

fuzzy logic fusion technique on a set of the tomographic images is significantly greater than the average of the constituent images. The fusion method results in an average increase of 0.20 in image cross correlation for the ideal image when compared to the cross correlation of the summation and ideal image.

MVD filtering of A-scan data.

The final component of the imaging system to be examined is processing of the raw ultrasonic data prior to reconstruction in order to try to improve the signal-to-noise ratio and time resolution of the data. The technique chosen was Minimum Variance Deconvolution (MVD) filtering. This method was selected because it is possible to determine the filter parameters without prior knowledge of the measurement system or the specimen under test. The filter requires knowledge about the ultrasonic wavelet used as well as statistical information about the material under test. These can be estimated from the raw data by the use of higher order statistics and Maximum Likelihood Deconvolution (MLD). This method of filtering has been used for a number of seismic and NDT applications. However, in these applications the user obtains the initial system-parameters from the test equipment and specimen rather than by estimation from the raw data as is the case here. This has only recently been made possible by improvements in computer performance, as the algorithms are computationally intense.

The effectiveness of this form of MVD filtering has been experimentally validated in two stages. The first stage was to examine the performance of MVD filtering on individual A-scans obtained using a number of different data acquisition set-ups. The second stage was to evaluate the performance of the filter on complete sets of A-scans.

The acquisition of the wavelet and statistical parameters has been performed on three experimentally obtained A-scans obtained using Pulse-Echo and Pitch-Catch acquisition set-ups. It has been shown that MVD filtering leads to an increased signal-to-noise ratio for all the A-scan and also an increase in the time resolution of any pulses reflected from defects.

The filtering technique was applied to complete sets of A-scans that were then used to reconstruct tomographic images. It has been experimentally demonstrated that there is a significant increase in the quality of the image reconstructed using the filtered data. This is reflected in the cross correlation values for the images reconstructed using both the filtered and unfiltered data. The filtering operation results in an increase of approximately 0.25 in the cross correlation with the ideal image, which is a significant improvement in overall image quality.

5.2. Future Work.

The program of work outlined above has led to the identification of several specific areas where further research would be beneficial.

Specimen Geometry.

All of the reconstructed images presented here are of defects embedded within cylindrical objects. However, the reconstruction algorithms have been implemented to be applicable to any test specimen geometry with the intention of incorporating the imaging methods with the NDT workbench developed within the Ultrasonics Research Group[4]. It would be advantageous to investigate the performance of the reconstruction algorithms and the fuzzy logic fusion technique presented here for a larger range of test specimens with differing geometries.

CAD model.

In order to aid interpretation and visualisation of the reconstructed tomographic images it would be advantageous to incorporate a CAD model of the test specimen with the image. This is of particular interest when the sample geometry is of greater complexity than has been used here. The presence of a CAD model of the component would allow better assessment of any defects present.

A system to allow visualisation of images together with a CAD model of the component has been developed by members of the Ultrasonics Research Group [4] so incorporation of the imaging system presented here with the visualisation and interpretation tools should be straightforward.

Investigation of System Performance for Real Defects.

The reconstructed images in this thesis are all of simulated defects contained within the test specimens. It would be extremely valuable to evaluate the system performance on actual defects.

Parallel Processing.

At present the image reconstruction times are longer than ideal for practical application. This is more of a problem with the more complex algorithms i.e. Pitch-Catch Reflection and Time-of-flight Diffraction tomography due to the nature of the backprojection although the other reconstruction methods would also benefit from parallel processing implementation. All of the tomographic algorithms presented here can easily be adapted for parallel processing [37].

3-D Image Reconstruction.

At present the reconstruction algorithms only reconstruct images of the object's cross section. This has the disadvantage that defects that are not at a favourable position and orientation within the object with respect to the object slice being imaged, will not get reconstructed satisfactorily. This can be overcome by extending the image reconstruction to 3-D.

There are two obvious ways this can be achieved. The first method would be to use the existing reconstruction algorithms to reconstruct images of a number of slice images and then stack them to obtain a 3-D image. The second method would be to develop tomographic algorithms that take the data at all positions about the test specimen and reconstruct a true 3-D image i.e. the backprojection of the ultrasonic data would be over spherical surfaces rather than along circular paths. These approaches present a trade off between the processing power required to reconstruct the image and the quality of the generated image. The first method is likely to be the least computationally intensive but would give images whose quality depends on the number of slices utilised. The second method would be extremely computationally intensive, but would give better quality images.

5.3. Closing Remarks.

This program of work has presented a number of new techniques to aid the characterisation of defects. The main contributions of this work are:

- The development and verification of an image fusion technique based on fuzzy logic set theory. It has been demonstrated that significant improvement can be made in defect characterisation by the fusion of a number of ultrasonic images. This methodology applied to the generation of high resolution images also provides a high degree of flexibility for the user.

- In order to maximise the amount of defect information that can be obtained from the set of A-scans available for reconstruction of tomographic images, a single-bounce, image-enhancement method has been developed. It has been demonstrated that the method gives improved defect characterisation.

- Since a system that is required to be flexible is extremely desirable, a blind deconvolution method has been implemented in order to obtain the information required to implement a deconvolution filter to improve A-scan resolution. Blind deconvolution was chosen to overcome the problem associated with the majority of deconvolution methods currently used in NDT; that is in many cases specific information is needed about the scanning geometry, probe utilised for data acquisition and also the component under test. This type of blind deconvolution has not been possible in NDT until now due to the computational complexity of the algorithms involved. It has been demonstrated that the blind deconvolution of A-scan data prior to reconstruction gives improved defect characterisation.

When these three innovative powerful ultrasonic signal-processing tools are integrated into a single imaging system along with tomographic image reconstruction tools the result is a highly flexible and powerful tool for the characterisation of defects contained within engineering components.

References.

1. R.A. Murgatroid, G.M. Worrall, S. Cruten, *Lessons learned from the PISCIII study on the influence of human factors on inspection reliability,* Proceedings for the 6th European Conference on Non-Destructive Testing, Nice, pp 989-993 (1994).
2. G.S. Passi, Y.Shoef, M.V. Kritsky, *Reducing the influence of human factors on the reliability of manual ultrasonic weld inspection.* Insight, Vol. 37 no 10 pp788-791 (1995).
3. M. G. Silk. *Benefits of signal processing in ultrasonic inspection.* Insight, Vol. 36 pp 776-781 (1994).
4. I.H. Cornwell. *Software systems for interpretation of ultrasonic NDT data.* Ph.D. Thesis University of Strathclyde (1997).
5. J. Krautkramer, H. Krautkramer. *Ultrasonic testing of materials (4th edition)* Springer Verlag, London (1990).
6. M. G. Silk. *An Evaluation of the performance of the TOFD technique as a means of sizing flaws, with particular reference to flaws with curved profiles.* Insight Vol. 38 pp 280-287 (1996).
7. W. Muller, V.Smitz, G. Schafer. *Reconstruction by the synthetic aperture focusing technique (SAFT).* Nuclear engineering and design, Vol. 94 pp 393-404 (1986).
8. S.F. Burch, J.T. Burton. *Ultrasonic synthetic aperture focusing using planar pulse echo transducers.* Ultrasonics, November pp 275-281 (1984).
9. M. Lorenz, T.S. Wielinga. *Ultrasonic characterisation of defects in steel using MULTI-SAFT imaging and neural networks.* ND & E International, Vol. 26 no. 3, pp 127-133 (1993).
10. R.Y. Chaio, L.J. Thomas. *Analytic Evaluation of Sampled aperture Ultrasonic Imaging techniques for NDE.* IEEE trans. Ultrasonics, Ferroelectrics and Frequency Control, Vol. 41 pp 484-493 (1994).
11. L.J. Busse. *3-dimensional imaging using a frequency-domain synthetic aperture focusing technique.* IEEE Trans. UFFC, Vol. 39 pp174-179 (1992).

12. R.K. Mueller, Mostafa Kaveh, G. Wade. *Reconstructive tomography and applications to ultrasonics.* Proc. IEEE Vol. 67 pp567-587 (1979).
13. H.W. Jones. *Recent Activity in Ultrasonic Tomography.* Ultrasonics, Vol. 31 pp 353-360 (1993).
14. P.B. Abraham, C.F. Gaumond. *Reflection Tomography.* J. Acoust. Soc. Am., Vol. 82 (4) pp1303-1315 (1987).
15. R.P. Bording, A. Gersztenkorn.,L.R. Lines,J.A. Scales,S. Treitel. *Applications of v seismic travel time tomography.* Geophysics Journal, Vol. 90 pp 285-303 (1987).
16. G.S. Kino, D. Corl, S. Bennett, K. Peterson. *Real Time Synthetic Aperture Imaging System.* 1980 Ultrasonics Symposium, pp 722-731 (1980).
17. K. Liang, B.T. Khuri-Yakub, G.S. Kino. *Reflection Tomography at 50 and 300 MHz.* 1981 Ultrasonics Symposium, pp 925-929 (1981).
18. S.J.Norton, M. Linzer. *Ultrasonic reflectivity tomography: Reconstruction with circular transducer arrays.* Ultrasonic imaging 1, pp154-184 (1979).
19. P.C. Pederson. *Inverse Methods in Ultrasonics.* Ultrasonics international 87 Conf. Proc. pp372-380 (1987).
20. M. Moshfenghi. *Ultrasound Reflection-Mode tomography using fan-beam-shaped insonification.* IEEE trans. Ultrasonics, Ferroelectrics and frequency control, Vol. UFFC-33 pp 299-314 (1986).
21. P.D. Hanstead. *Imaging with reduced data* Ultrasonics International 87 Conf. Proc. pp 12-19 (1987).
22. L. Capineri, H.G Tattersal, J.A.G. Temple, M.G. Silk. *Time-of-flight diffraction tomography for NDT applications.* Ultrasonics, Vol. 30, pp 275-288 (1992).
23. A.Cochran. *Ultrasonic arrays used to monitor the condition of solid test subjects.* Ph.D.Thesis University of Strathclyde (1990).
24. X.E. Gross *NDT Data Fusion* Arnold (1997).
25. C. Pohl, J. L. Van Genderen. *Multisensor image fusion in remote sensing: concepts, methods and applications.* Intl. J. Remote Sensing, Vol. 19 pp 823-854 (1998).
26. Abdolhossien Nejatali, Ioan R. Ciric. *Novel image fusion methodology using fuzzy set theory.* Opt. Eng., Vol. 37(2) pp 485-491 (1998).

27. I.M. Skolnik. Introduction to Radar Systems (2nd edition). *McGraw Hill (1981).*

28. R.N. Carpenter, P.R. Stepanishen. *An improvement in the range resolution of ultrasonic pulse echo systems by deconvolution.* J. Acoust. Soc. Am., Vol. 75 pp 1084-1091 (1984).

29. G. Hayward, J.E. Lewis. *A theoretical approach for inverse filter design in ultrasonic applications.* IEEE Trans. UFFC, Vol. 36 pp 356-364 (1989).

30. G. Hayward, J.E. Lewis. *Comparison of some non-adaptive deconvolution techniques for resolution enhancement of ultrasonic data.* Ultrasonics, Vol. 27 pp 155-164 (1989).

31. Sheng-Wen Cheng, Min-Kang Chao *Resolution improvement of ultrasonic C-scan images by deconvolution using the monostatic point-reflector spreading function (MPSF) of the transducer.* NDT&E international, Vol. 29 pp 293-300 (1996).

32. Burch S.F. *Comparison of SAFT and two dimensional deconvolution methods for the improvement of resolution in ultrasonic B-scan images* Ultrasonics Vol. 25 pp259-266 (1987).

33. J. Zhao, P. A. Gaydecki, F. M. Burdekin. *Investigation of block filtering and deconvolution for the improvement of lateral resolution and flaw sizing accuracy in ultrasonic testing.* Ultrasonics, Vol. 33 pp187-194 (1995).

34. J.M. Mendel. *Maximum likelihood deconvolution: A journey into model based signal processing.* Springer-Verlag (1990).

35. S.J. Norton, M. Linzer. *Ultrasonic Reflectivity imaging in three dimensions: Reconstruction with spherical transducer arrays.* Ultrasonic Imaging, Vol. 1 pp 210-231 (1979).

36. D.A. Hutchins, D.K. Mak, J.K. Hu. *EMAT ultrasonic reflection tomography.* Ultrasonic International 87 Conf. Proc. pp 860-864 (1987).

37. B.S. Hoyle, F. Wiegand. *Real time parallel tomographic ultrasound imaging using transputers.* Electronic letters, Vol. 24, pp 605-606 (1988).

38. Mohamed Ali Seiraffi. *Two-phase measurement in pipelines using ultrasonic tomography.* Acoustical imaging, Vol. 19 pp 933-937 (1992).

39. F. Russo, G. Ramponi. *Fuzzy Methods for Multisensor Data Fusion.* IEEE Trans. on Instrumentation and Measurement, Vol. 43 pp 288-294 (1994).

40. H. Bandemar, S. Gottwald. *Fuzzy sets, Fuzzy logic, Fuzzy methods: with applications.* Wiley (1993).

41. P.K. Sanoo. *A survey of thresholding techniques.* Comput. Vis. Grap. Image process, Vol. 41 pp 233-260 (1988).

42. J.M. Mendel. *Minimum variance deconvolution.* IEEE Trans. Geoscience Remote Sensing, Vol. 19, pp 161-171 (1991).

43. J.A. Cadzow. *Spectral estimation: An overdetermined rational model equation approach.* Proc. IEEE, Vol. 70 pp 907-939 (1989).

44. J.M. Mendel. *Tutorial on higher order statistics (spectra) in signal processing and system theory: Theoretical results and some applications.* Proc IEEE Vol. 79 pp278-305 (1991).

45. G.B. Giannakis, J.M. Mendel. *Identification of Non-minimum phase systems using higher order statistics.* IEEE trans. Acoust., Speech and signal processing Vol 37, pp 360-377 (1989).

46. G.B. Giannakis, J.M. Mendel. *Cumulant-Based order determination of non-Gaussian ARMA models.* IEEE trans. Acoust., Speech and signal processing Vol. 38, pp1411-1422 (1990).

47. Ananthram Swami, J.M. Mendel. *ARMA Parameter Estimation using only output cumulants.* IEEE Trans. Acoust., Speech and Signal Processing, Vol. 38, pp1411-1422 (1990).

48. *Numerical Recipes in C: The art of scientific computing.* Cambridge University Press (1992).

49. S. Haykin. Adaptive filter theory: 3rd edition, *Prentice Hall* (1996).

50. U.R. Abeyratne, A.P. Petrropulu, J.M. Reid, T. Golas, E. Conant. *Higher order versus second order statistics in ultrasound image deconvolution.* IEEE Trans. UFFC, Vol. 44 pp 1409-1416 (1997).

Appendix A.

TRIGONOMETRICAL MANIPULATIONS.

In Chapter 2 some trigonometrical manipulation of equations (2.2-9) and (2.2-35) was required. This manipulation will be presented in this appendix. Initially the more straightforward Pulse-echo Reflection case will be examined. Once this trigonometric manipulation has been outlined the more complex Pitch-Catch Reflection situation will be examined.

Pulse-Echo Reflection Tomography.

The aim of this section is to showing the trigonometrical manipulation performed on equation (1) to obtain equation (2). For clarity the manipulation will be performed in two parts. The first part will be the manipulation of the first two Cosine terms, Then the remaining Cosine terms will be examined, with the remaining term not requiring and manipulation to be in the form in equation (2).

$$2\rho(\phi;r,\theta)-2\rho(\phi;r_0,\theta_0)\cong -2r\cos(\theta-\phi)+2r\cos(\theta_0-\phi)+\frac{r^2-r_0^2}{2R} -\frac{r^2}{2R}\cos 2(\theta-\phi)+-\frac{r_0^2}{2R}\cos 2(\theta_0-\phi) \quad (1)$$

$$2\rho(\phi;r,\theta)-2\rho(\phi;r_0,\theta_0)=2X\cos(\phi-Y)+\gamma_1+\gamma_2\cos 2(\phi-\alpha) \quad (2)$$

where

$$X=\left(r_0^2+r^2-2r_0r\cos(\theta-\theta_0)\right)^{1/2} \quad (3a)$$

$$\tan Y=\frac{r_0\sin\theta_0-r\sin\theta}{r_0\cos\theta_0-r\cos\theta} \quad (3b)$$

$$\gamma_1=\frac{1}{2R}\left(r^2-r_0^2\right) \quad (3c)$$

$$\gamma_2=\frac{1}{2R}\left(r_0^4+r^4-2r^2r_0^2\cos 2(0-0_0)\right)^{1/2} \quad (4a)$$

$$\tan\alpha=\frac{r_0^2\sin 2\theta_0-r^2\sin 2\theta}{r_0^2\cos 2\theta_0-r^2\cos 2\theta} \quad (4b)$$

The first two Cosine terms of equation (1) can be expanded as follows:

$$\begin{aligned} & 2[-r\cos(\theta-\phi)+r_0\cos(\theta_0-\phi)] \\ & \Rightarrow 2[-r(\cos\theta\cos\phi+\sin\theta\sin\phi)+r_0(\cos\theta_0\cos\phi+\sin\theta_0\sin\phi)] \\ & \Rightarrow 2[(r_0\cos\theta_0-r\cos\theta)\cos\phi+(r_0\sin\theta_0-r\sin\theta)\sin\phi] \end{aligned} \qquad (5)$$

Substituting

$$A = r_0\cos\theta_0 - r\cos\theta \qquad (6)$$

and

$$B = r_0\sin\theta_0 - r\sin\theta \qquad (7)$$

into equation (5), then the following relationship is obtained,

$$2[A\cos\phi + B\sin\phi] \qquad (8)$$

Using standard trigonometrical relationships, this can be written in the form $2X\cos(\phi - Y)$, i.e.

$$2[A\cos\phi + B\sin\phi] = 2X\cos(\phi - Y) = 2[X\cos\phi\cos Y + X\sin\phi\sin Y] \qquad (10)$$

Then equating Cosine and Sine terms in equation (10) an expression for Y can be obtained

$$\tan Y = \frac{B}{A} = \frac{r_0\sin\theta_0 - r\sin\theta}{r_0\cos\theta_0 - r\cos\theta} \qquad (11)$$

All that remains is to obtain an expression for X. Squaring and adding the expression obtained from equating the cosine and sine terms, then

$$A^2 + B^2 = X^2 \cos^2 Y + X^2 \sin^2 Y = X^2 \qquad (12)$$

$$X = \sqrt{(A^2 + B^2)}$$

where

$$A^2 = r_0^2 \cos^2 \theta_0 - 2rr_0 \cos\theta \cos\theta_0 + r^2 \cos^2 \theta \quad (13a)$$

$$B^2 = r_0^2 \sin^2 \theta_0 - 2rr_0 \sin\theta \sin\theta_0 + r^2 \sin^2 \theta \quad (13b)$$

It follows from this that

$$X = \left(r^2 + r_0^2 - 2rr_0 \cos(\theta - \theta_0)\right)^{1/2} \qquad (14)$$

Now all that remains is to perform the manipulation on the third and fourth trigonometric terms of equation (1). The two trigonometrical terms can be expanded as follows.

$$\begin{aligned}
&-\frac{r^2}{2R}\cos 2(\theta - \phi) + \frac{r_0^2}{2R}\sin 2(\theta_0 - \phi) \\
&\Rightarrow \frac{1}{2R}\left[-r^2 \cos 2(\theta - \phi) + r_0^2 \sin 2(\theta_0 - \phi)\right] \\
&\Rightarrow \frac{1}{2R}\left[-r^2 \cos 2\theta \cos 2\phi - r^2 \sin 2\theta \sin 2\phi + r_0^2 \cos 2\theta_0 \cos 2\phi + r_0^2 \sin 2\theta_0 \sin 2\phi\right] \\
&\Rightarrow \frac{1}{2R}\left[\left(r_0^2 \cos 2\theta_0 - r^2 \cos 2\theta\right)\cos 2\phi + \left(r_0^2 \sin 2\theta_0 - r^2 \sin 20\right)\sin 2\phi\right]
\end{aligned} \qquad (15)$$

This is now in a similar form to equation (5), with

$$A = r_0^2 \cos 2\theta_0 - r^2 \cos 2\theta \qquad (15a)$$

and

$$B = r_0^2 \sin 2\theta_0 - r^2 \sin 2\theta \qquad (15b)$$

Using the same trogonometrical relationship as before the following equation is obtained

$$\frac{1}{2R}[A\cos 2\phi + B\sin 2\phi] = \frac{1}{2R}\gamma_2 \cos 2(\phi - \alpha)$$
$$= \frac{1}{2R}[\gamma_2 \cos 2\phi \cos 2\alpha + \gamma_2 \sin 2\phi \sin 2\alpha] \qquad (16)$$

The same method of equating the Cosine and Sine terms is now adopted to determine γ_2 and α. It follows that

$$\gamma_2 = \frac{1}{2R}\left(r_0^4 + r^4 - 2r^2 r_0^2 \cos 2(\theta - \theta_0)\right)^{1/2} \qquad (17)$$

and

$$\tan\alpha = \frac{r_0^2 \sin 2\theta_0 - r^2 \sin 2\theta}{r_0^2 \cos 2\theta_0 - r^2 \cos 2\theta} \qquad (18)$$

Pitch-Catch Reflection Tomography.

The simpler Pulse-echo case has been examined, now all that remains is to examine the Pitch-catch Reflection tomography case. The task is to perform the trigonometrical manipulation on equation (19) to obtain equation (20).

$$\Delta_{\phi,\beta}(r,\theta;r_0,\theta_0) \cong \frac{1}{2R}[r^2 - r_0^2] - r\cos(\theta - \phi - \beta) - r\cos(\theta - \phi + \beta) +$$
$$r_0 \cos(\theta_0 - \phi - \beta) + r_0 \cos(\theta_0 - \phi + \beta) +$$
$$\frac{1}{4R}[-r^2 \cos 2(\theta - \phi - \beta) - -r^2 \cos 2(\theta - \phi + \beta) +$$
$$r_0^2 \cos 2(\theta_0 - \phi - \beta) + r_0^2 \cos 2(\theta_0 - \phi + \beta)] \qquad (19)$$

and

$$\Delta_{\phi,\beta}(r,\theta;r_0,\theta_0) = 2X\cos(\phi - Y) + \gamma_1 + \gamma_2 \cos 2\beta \cos 2(\phi - \alpha) \qquad (20)$$

where X, Y, γ_1, γ_2 and α are as defined in equations (3) and (4).

The same approach taken for the Pulse-echo manipulation will be undertaken here. The first step is to split equation (19) into two parts. The first four trigonometrical terms can be expanded as follows:

$$\begin{aligned}&-r\cos(\theta-\phi-\beta)-r\cos(\theta-\phi+\beta)+r_0\cos(\theta_0-\phi-\beta)+r_0\cos(\theta_0-\phi+\beta)\\&=-r\cos(D-\beta)-r\cos(E-\beta)+r_0\cos(D-\beta)+r_0\cos(E+\beta)\\&=-2r\cos D\cos\beta+2r_0\cos E\cos\beta\end{aligned} \qquad (21)$$

where

$$D=\theta-\phi \qquad (22a)$$

and

$$E=\theta_0-\phi \qquad (22b)$$

Now substituting D and E back into equation (21) it becomes

$$\begin{aligned}&-2r[\cos\theta\cos\phi+\sin\theta\sin\phi]\cos\beta+2r_0[\cos\theta_0\cos\phi+\sin\theta_0\sin\phi]\cos\beta\\&=2[(r_0\cos\theta_0-r\cos\theta)\cos\phi+(r_0\sin\theta_0-r\sin\theta_0)\sin\phi]\cos\beta\end{aligned} \qquad (23)$$

By multiplying equation (10) by $\cos\beta$

$$\begin{aligned}2[A\cos\phi+B\sin\phi]\cos\beta&=2X\cos\beta\cos(\phi-Y)\\&=2[X\cos\phi\cos Y+X\sin\phi\sin Y]\cos\beta\end{aligned} \qquad (24)$$

where

$$A=r_0\cos\theta_0-r\cos\theta \qquad (25a)$$

and

$$B = r_0 \sin\theta_0 - r\sin\theta \qquad (25b)$$

Equating the cosine and sine as before the same expressions for X and Y are obtained as expected, gives the same expressions for X and Y as obtained previously, Notably

$$\tan Y = \frac{r_0 \sin\theta_0 - r\sin\theta}{r_0 \cos\theta_0 - r\cos\theta} \qquad (26)$$

and

$$X = \left(r^2 + r_0^2 - 2rr_0\cos(\theta - \theta_0)\right)^{1/2} \qquad (27)$$

The second step of the manipulation is to examine the other four trigonometrical terms of equation (19). Notably

$$\frac{1}{4R}[-r^2\cos 2(\theta-\phi-\beta)--r^2\cos 2(\theta-\phi+\beta)+ r_0^2\cos 2(\theta_0-\phi-\beta)+r_0^2\cos 2(\theta_0-\phi+\beta)] \qquad (28)$$

A similar approach is taken to the manipulation of the first four cosine terms of equation (19). These terms can be expanded as follows:

$$\begin{aligned}&\frac{1}{4R}[-r^2\cos 2(\theta-\phi-\beta)-r^2\cos 2(\theta-\phi+\beta)\\ &+r_0^2\cos 2(\theta_0-\phi-\beta)+r_0^2\cos 2(\theta_0-\phi+\beta)]\\ &=\frac{1}{4R}\left[-r^2\cos 2(A-\beta)-r^2\cos 2(A+\beta)+r_0^2\cos 2(B-\beta)+r_0^2\cos 2(B+\beta)\right]\\ &=\frac{1}{4R}\left[-2r^2\cos 2A\cos 2\beta+2r_0^2\cos 2B\cos 2\beta\right]\end{aligned} \qquad (29)$$

where

$$A = \theta - \phi \qquad (30a)$$

and

$$B = \theta_0 - \phi \qquad (30b)$$

Now substituting A and B back into equation (29)

$$\begin{aligned} &\frac{1}{4R}\left[-2r^2\cos 2A\cos 2\beta + 2r_0^2\cos 2B\cos 2\beta\right] \\ &= \frac{1}{2R}\begin{bmatrix} -r^2\left[\cos 2\theta\cos 2\phi + \sin 2\theta\sin 2\phi\right]\cos 2\beta \\ +r_0^2\left[\cos 2\theta_0\cos 2\phi + \sin 2\theta_0\sin 2\phi\right]\cos 2\beta \end{bmatrix} \\ &= \frac{1}{2R}\left[\left[\left(r_0^2\cos 2\theta_0 - r^2\cos 2\theta\right)\cos 2\phi + \left(r_0^2\sin 2\theta_0 - r^2\sin 2\theta_0\right)\sin 2\phi\right]\cos 2\beta\right] \end{aligned} \qquad (31)$$

As before use the standard trigonometrical relationship stated in equation (32) to reduce equation (31) to a single term.

$$\begin{aligned} \frac{1}{2R}\left[A\cos 2\phi + B\sin 2\phi\right]\cos 2\beta &= \frac{1}{2R}\gamma_2\cos 2\beta\cos 2(\phi - \alpha) \\ &= \frac{1}{2R}\left[\gamma_2\cos 2\phi\cos 2\alpha + \gamma_2\sin 2\phi\sin 2\alpha\right]\cos 2\beta \end{aligned} \qquad (32)$$

where

$$A = r_0^2\cos 2\theta_0 - r^2\cos 2\theta$$

and

$$B = r_0^2\sin 2\theta_0 - r^2\sin 2\theta$$

Equating the Cosine and Sine terms the expressions for the γ_2 and α can be obtained, and are as specified in equations (3) and (4).