UNIVERSITY OF CAMBRIDGE
DEPARTMENT OF ENGINEERING

3D Freehand Ultrasound: Reconstruction and Spatial Compounding

Robert Nicholas Rohling

Churchill College

September 1998

A dissertation submitted to the University of Cambridge
for the degree of Doctor of Philosophy.

Summary

In 3D freehand ultrasound, a position sensor is attached to a conventional ultrasound probe so that a set of 2D B-scans are acquired along with their relative locations. This allows the set of B-scans to be reconstructed into a regular 3D voxel array for visualisation and analysis. This thesis is concerned with two topics: reconstruction and spatial compounding of 3D freehand ultrasound data.

A key requirement for reconstruction is calibration of the freehand system. Four different calibration techniques are described using a common mathematical framework, then evaluated according to their repeatability, precision and accuracy. The new Cambridge phantom and the conventional cross-wire phantom are shown to offer the best performance.

Following calibration, interpolation of the set of irregularly spaced B-scans is required to reconstruct a regular voxel array. A radial basis function interpolation method is adapted for reconstruction of ultrasound data and compared with three standard reconstruction methods. The new functional method exploits the underlying shape of the ultrasound data to fill gaps and average overlapping data, without introducing the kind of artifacts common with the standard methods.

The second topic, spatial compounding, aims to improve the visualisation and volumetric analysis of the reconstructed data, typically inhibited by artifacts such as speckle. The basic principle of spatial compounding is to average multiple images of the same object taken from different viewpoints. With averaging, the signal to noise ratio increases, but accurate image registration is required to avoid blurring. Although a number of 2D studies can be found in the literature, this is the first to fully investigate 3D spatial compounding.

An initial series of tests are performed *in vitro*. The statistical theory to predict the increase in signal to noise ratio is developed and verified empirically. Using a novel registration method based on image landmarks, spatial compounding is shown to improve visualisation and volume estimation. Having proven the benefits of spatial compounding *in vitro*, tests with higher levels of compounding are performed *in vivo*. A more robust correlation-based registration technique is used on an examination of a human gall bladder and, again, visualisation and volume estimation are clearly improved.

Key words: 3D ultrasound, freehand, reconstruction, calibration, interpolation, radial basis function, spatial compounding, registration.

Acknowledgements

Thanks are owed to a number of people who have helped me during my three years of study in Cambridge. I must thank Andrew Gee, my supervisor, for giving me the freedom to investigate the topics that interested me, challenging me to prove my arguments, and encouraging me to keep on schedule. My wife, Josée, should be thanked too, for sharing in the highs and lows that are inevitably part of research. My friends should also be thanked for ensuring I took much-needed tea breaks, and for making my time in Cambridge enjoyable. Of course, none of this would have been possible without the continuous support from my parents through all the years of my studies.

Several people have also contributed to the software and apparatus I used in my research. In particular, the freehand acquisition system was developed by Richard Prager and Patrick Gosling. Ultrasound scanning was done with the help of Laurence Berman at the University of Cambridge Department of Radiology. Jonathon Carr designed the three-wire phantom, took some of the photographs in this thesis, and has been a great source of information. Calculations of the 3D gradients were performed using a separable, recursive 3D edge detector, kindly provided by the researchers at INRIA, Sophia Antipolis, France.

The original program code for the regularized splines with tension was kindly donated by Lubos Mitáš and Helena Mitášová. Lubos Mitáš is at the National Center for Supercomputing Applications at the University of Illinois. Helena Mitášová is at the Geographic Modeling Systems Laboratory at the University of Illinois. Development of the original spline code was done in cooperation with the United States Army Construction Engineering Research Laboratories.

Finally, I am greatly indebted to Churchill College for the generous studentship awarded to me. This studentship, in conjunction with an Overseas Research Students Award, provided the means to support my studies.

Declaration

This dissertation is the result of my own original work and includes nothing done in collaboration with others. It has not been submitted in whole or in part for a degree at any other university. The length is approximately 40,000 words, including footnotes, appendices and references. The following publications were derived from this work:

Journal articles

Rohling, R., Gee, A., and Berman, L. (1998). Automatic registration of 3-D ultrasound images. *Ultrasound in Medicine and Biology*, 24(6):841–854.

Rohling, R., Gee, A., and Berman, L. (1998). Radial basis function interpolation for 3-D ultrasound. *Medical Image Analysis*, submitted.

Prager, R., Rohling, R., Gee, A., and Berman, L. (1998). Rapid calibration for 3-D freehand ultrasound. *Ultrasound in Medicine and Biology*, 24(6):855–869.

Rohling, R., Gee, A., and Berman, L. (1997). 3-D spatial compounding of ultrasound images. *Medical Image Analysis*, 1(3):177–193.

Conference presentations

Rohling, R., Gee, A., and Berman, L. (1998). Automatic registration of 3-D ultrasound images. In *Proceedings of the 1998 IEEE International Conference on Computer Vision*, pages 298–303, Bombay, India.

Rohling, R., Gee, A., and Berman, L. (1997). Spatial compounding of 3-D ultrasound images. In *Proceedings of the 1997 International Conference on Image Processing in Medical Imaging*, pages 519–524, Poultney, Vermont, USA.

Rohling, R., Gee, A., and Berman, L. (1997). Automatic registration of 3-D ultrasound images. In *Proceedings of the First British Medical Image Understanding and Analysis Conference*, pages 37–40, Oxford, UK.

Rohling, R., Gee, A., and Berman, L. (1996). Correcting motion-induced registration errors in 3-D ultrasound images. In *Proceedings of the 1996 British Machine Vision Conference*, pages 645–654, Edinburgh, UK.

Contents

List of Figures

List of Tables

Notation

General Matrix and Vector Notation

$\mathbf{I}$	Identity matrix.
$\mathbf{A}^{\mathrm{T}}$	Transpose of a matrix.
$\mathbf{x}$	Position in three space (x, y, z).
$\mathbf{x}_j$	Position of point (x_j, y_j, z_j).
$\lvert\mathbf{x}\rvert$	Vector norm $\sqrt{x^2 + y^2 + z^2}$.
${}^{I}\underline{\mathbf{x}}$	4×1 position vector $(x, y, z, 1)$ with respect to coordinate system I.
${}^{J}\mathbf{T}_{I}$	4×4 homogeneous transform from coordinate system I to J.

Calculation of Focal Point

F	Distance from the probe face to the focal point.
λ	Wavelength of an ultrasound pulse.
a	Radius of the effective aperture.

Coordinate Systems and Transformations

u	Column index of pixel in B-scan.
v	Row index of pixel in B-scan.
u_{max}	Number of columns of pixels in B-scan.
v_{max}	Number of rows of pixels in B-scan.
s_x	Scale factor to convert the column index of a pixel into mm.
s_y	Scale factor to convert the row index of a pixel into mm.
P	Coordinate system of the cropped B-scan.
R	Coordinate system of the receiver.
T	Coordinate system of the transmitter.
C	Coordinate system of the cuberille.
$\mathbf{P}$	2D array of pixels representing a B-scan.
$\mathbf{C}$	Cuberille.
α	Angle of rotation about the z-axis.

β	Angle of rotation about the y-axis.
γ	Angle of rotation about the x-axis.
${}^C\mathbf{T}_T$	Homogeneous transform from transmitter to cuberille.
${}^T\mathbf{T}_R$	Homogeneous transform from receiver to transmitter.
${}^R\mathbf{T}_P$	Homogeneous transform from B-scan to receiver.
${}^C\mathbf{T}_P$	Combined transformation from B-scan to cuberille.
p_{mn}	Pixel at index (m,n).
c_{ijk}	Voxel at index (i,j,k).
${}^P\underline{\mathbf{x}}$	4×1 vector to a pixel with respect to the coordinate system P.
${}^C\underline{\mathbf{x}}$	4×1 vector to a voxel with respect to the coordinate system C.

Calibration

Φ	Column vector of unknown parameters.
Φ_j	Current estimate of unknown parameters.
$\Delta\Phi$	Update to current estimate of unknown parameters.
m	Number of B-scans.
θ	Combined vector of known parameters, *i.e.* measurements.
$\mathbf{f}(\theta, \Phi)$	Functions extracted from the zero components of the calibration equations.
$\Delta\mathbf{f}(\theta, \Phi_j)$	Error at current estimate of unknown parameters.
ε	Damping term in Levenberg-Marquardt algorithm.
$\mathbf{J}$	Jacobian matrix $\partial\mathbf{f}(\theta, \Phi_j)/\partial\Phi$ at current estimate of unknown parameters.
κ	Condition number.
$\Delta{}^R\underline{\mathbf{x}}$	Change in position, relative to R, of (u_{max},v_{max}) from different calibrations.
$\Delta{}^T\underline{\mathbf{x}}$	Change in position, relative to T, of the crossed wires from different viewpoints.
${}^R\mathbf{T}_P{}^{trial\ i}$	Homogeneous transform between P and R, determined by calibration trial i.
${}^P\underline{\mathbf{x}}^{trial\ i}$	Vector to (u_{max},v_{max}) determined by calibration trial i.
$s_x^{trial\ i}$	Lateral scale factor from calibration trial i.
$s_y^{trial\ i}$	Axial scale factor from calibration trial i.
${}^T\mathbf{T}_R{}^{view\ i}$	The position sensor measurements for view i.
${}^P\underline{\mathbf{x}}^{view\ i}$	The vector to the location of the crossed wires in the B-scan from view i.

Radial Basis Functions

$O(N)$	Order of complexity.
N	Number of pixels summed over all B-scans.
p_j	Pixel located at $\mathbf{x}_j$.
c_j	Voxel value located at $\mathbf{x}_j$
r_{max}	Radius of local neighbourhood.

$S(\mathbf{x})$	Spline function.
$I(s)$	Smoothness function.
$T(\mathbf{x})$	Trend function.
$\mathbf{R}(\mathbf{x}, \mathbf{x}_j)$	Radial basis function, centered about $\mathbf{x}_j$.
a_j	Scalar coefficients of radial basis functions.
r	Distance $\lvert\mathbf{x} - \mathbf{x}_j\rvert$.
ϕ	Tension parameter.
δ_{ij}	Kronecker delta function.
w	Weight (cost) of the spline's smoothness versus deviation from the data points.
K_{max}	Maximum number of data points in segment.
K_{min}	Minimum number of data points in window.
M	Number of artificially removed pixels.
V	Mean error between interpolated values and original voxel values.
μ	Mean of V.
σ	Standard deviation of V.

Spatial Compounding

$\mathbf{T}^{reg}$	Additional homogeneous transform for registration.
${}^{C}\mathbf{T}_P^*$	Updated ${}^{C}\mathbf{T}_P$ that incorporates $\mathbf{T}^{reg}$ for registration.
$\mathbf{L}$	Landmark volume.
l_{ijk}	Element of landmark volume at index (i,j,k).
n	Number of B-scans used for spatial compounding.
n_{high_res}	Effective number of B-scans used for high resolution spatial compounding.
n_{low_res}	Effective number of B-scans used for low resolution spatial compounding.
$\mathbf{b}$	Vector of registration parameters: three translations and three rotations.

Glossary

℄ Centre line, especially the centre line of the ultrasound beam.

2D Two-dimensional.

3D Three-dimensional.

AC Alternating current.

Anechoic Does not produce echoes.

Any-plane slicing The method of viewing 3D data whereby images are created from cross-sections of the data set.

Apodization The method of unequal weighting of the transmit or receive elements' signals to reduce grating lobes in an array transducer.

Artifact Any portion of an ultrasound image that obscures diagnostic information or does not correspond to the true anatomy.

B-scan A 2D cross-sectional image produced by sweeping an ultrasound beam through a plane.

Calibration phantom A device with certain known dimensions that is used to calibrate a 3D freehand ultrasound system.

CT Computed tomography.

Cuberille A regular 3D array.

DC Direct current.

DW Distance-weighted.

DOF Degree(s) of freedom.

Doppler ultrasound The measurement and display of fluid motion using the change in frequency of ultrasound pulses.

ECG Electrocardiogram.

Edgel Edge element.

Fps Frames per second.

Freehand Movement of the ultrasound probe without restriction.

In vivo Within a living organism.

In vitro In an artificial environment outside the living organism, *i.e.* opposite of *in vivo*.

Look direction The direction of an ultrasound beam with respect to an anatomical region of intest.

MR Magnetic resonance.

Organ phantom A tissue-equivalent material that has some characteristics representative of tissues (*e.g.* scattering or attenuation properties).

Orientation The three degrees of freedom that determine the direction or line of action of a vector.

Phantom An artificial object used to generate ultrasound images with some known characteristics.

Pixel Picture element, *i.e.* element of a 2D array that forms an image.

PNN Pixel nearest neighbour.

Position The location of a point in a Cartesian reference system.

RANSAC Random sample consensus.

RBF Radial basis function.

Rendering A computer-generated image showing a depiction of a structure.

Reconstruction The stage of a freehand ultrasound system where the set of acquired B-scans are converted into a regular 3D array of voxels.

RMS Root mean squared.

Scan To perform an ultrasound examination.

Scan plane The plane of a B-scan.

SONAR Sound and navigation ranging.

SNR The signal to noise ratio, calculated by dividing the mean grey level by the standard deviation.

Surface rendering The method of synthesising a view of a surface from a particular viewpoint under selected lighting conditions.

TGC Time-gain compensation.

Volume rendering The method of viewing 3D data whereby an image is created by casting rays through a cuberille.

VNN Voxel nearest neighbour.

Voxel Volume element, *i.e.* element of a 3D array that forms a cuberille.

Increasing knowledge brings increasing power, and great as are the triumphs of the present century, we may well believe that they are but a foretaste of what discovery and invention have yet in store for mankind.

—Lord Rayleigh, **British Association Report**, 1884

Chapter 1

Introduction

1.1 History

Scientists have long dreamed of exploring the inner world of the human body. Yet, before modern medicine, no practical means were available to visualise living anatomy non-invasively. Today, images are routinely made of a wide variety of internal anatomical structures, and one of the most popular methods of imaging is ultrasound.

Using sound is a natural way to describe a scene not visible to the human eye. We do this instinctively when entering a darkened room by listening to the sounds around us. Thomas Hardy described this well in *The Return of the Native* (1878), "Compound utterances address themselves to their senses, and it was possible to view by ear the features of the neighbourhood. Acoustic pictures were returned from the darkened scenery;..." (Cosgrove et al., 1993).

Near the same time, the famous theoretician Lord Rayleigh wrote the influential book *The Theory of Sound* (1877). This work provided the intellectual foundation for the subject of Acoustics. Shortly later, the discovery of the piezo-electric effect by Jean and Pierre Curie in 1880 (Curie and Curie, 1880) provided the key for practical applications of ultrasound imaging.

One of the first applications of ultrasound was the detection of submerged objects. After the sinking of the Titanic in 1912, many new devices were proposed for the detection of icebergs. An initial device that used acoustic signals was patented by M. L. F. Richardson (1912), but was never constructed . It was R. A. Fessenden who staged the first successful demonstration in 1914. He used an electromagnetic moving-coil transducer and detected icebergs up to two miles away (Hunt, 1954). These ideas were further developed in secret by Paul Langevin, a graduate student of Pierre Curie. The main application then was the detection of submarines during the war of 1914 to 1918. Langevin's work on piezo-electric transducers was made possible by the newly developed high frequency amplifiers supplied by the French Army. With these amplifiers he achieved considerable success. This work laid the foundations for subsequent advances in marine SONAR (sound navigation and ranging) (Hackmann, 1984), non-destructive materials testing (Hull and Vernon, 1988) and eventually medical imaging.

Probes in all of these fields share a common technique: a piezoelectric crystal is excited to produce a short ultrasound pulse and the resulting echoes are recorded. The intensity and timing of the echoes gives information about the structures along the direction of the ultrasound beam. If the beam is swept through a plane, a 2D (two-dimensional) image can be created. Because the echo intensities are converted into "brightness" for display, the 2D image is normally called a B-scan.

The first B-scan of human anatomy was published by the Dussik brothers in 1942 (Dussik, 1942). Yet this image is considered to be mainly artifactual, and it was not until after the war of 1939 to 1945 that more serious investigations into medical ultrasound imaging were made. With improvements in display devices and transducers, the first commercial ultrasound machine, the 'Diasonograph', was produced in 1964. Soon after, many more commercial systems were developed with improved ergonomics and image quality.

Since the 1970's, image quality has improved dramatically from one generation of machine to the next. These improvements have been derived mainly from advancements in miniaturisation and electronics. As ultrasound systems continue to improve, the popularity of ultrasound increases. In the past few years, the revenue generated from sales of ultrasound equipment has exceeded that of all other imaging modalities (Cosgrove et al., 1993). The popularity arises from the fact that ultrasound imaging is portable, fast (high frame rate), easy to operate, comfortable for the patient, relatively cheap, and is safe compared to ionising radiation such as X-rays. Moreover, the latest generation of machines can generate images of impressive quality.

With the increasing popularity of ultrasound, the number of applications has expanded. For example surgical, intravascular and needle guidance techniques with 2D ultrasound are now commonplace. 3D (three-dimensional) ultrasound, the subject of this thesis, is also starting to be used in regular clinical practice.

Although the basic principles of 3D ultrasound imaging can be traced back as far as 1956 (Howry et al., 1956), the heavy computational requirements have limited development. With the advent of modern computers, there has been considerable renewed interest in 3D methods. In 1989, Kretztechnik AG (Zipf, Austria) produced the first commercial 3D ultrasound machine, the Combison 330 Voluson. Other commercial 3D systems soon followed.

It will take many clinical trials to fully evaluate the potential of 3D ultrasound, but initial studies have shown that 3D ultrasound offers unique advantages in a variety of clinical applications. These include studies of the fetus (Chervenak et al., 1993; Deng et al., 1996; Steiner et al., 1994), heart (Belohlavek et al., 1993; Ofili and Nanda, 1994; Salustri and Roelandt, 1995), eye (Cusumano et al., 1998; Downey et al., 1996), brain (Rankin et al., 1993), vascular structure (Franceschi et al., 1992), breast (Lalouche et al., 1989; Moskalik et al., 1995), urethra (Ng et al., 1994) and the abdominal organs such as the kidney (Gilja et al., 1995), gall bladder (Fine et al., 1991), and prostate (Basset et al., 1991). More and more clinical applications are appearing as experience with 3D ultrasound increases.

1.2 Advantages of 3D Ultrasound

Clinical trials have shown 3D ultrasound to have a number of advantages over conventional 2D ultrasound. First, images of the surfaces of structures allow the 3D nature of the structure to become fully apparent. For example, the diagnosis of cleft palate in fetal examinations is made easier from 3D renderings of the fetal face, than from 2D cross-sections (Baba and Jurkovic, 1993; Lee et al., 1995).

From 3D data, 2D cross-sectional images can be produced at any angle, including angles impossible to achieve by a standard examination, *e.g.* at a constant depth below the skin. This plane of imaging has been shown to be useful, for example, in monitoring follicular development in infertility treatment (Baba and Jurkovic, 1993).

Measurement of organ volume can also be more accurate. Measuring changes in the volume of an organ over time is often used to monitor the progression of a disease or its response to treatment. Normally, volumes are estimated from equations involving a few linear distance measurements, information about the patient (such as age and sex), and a simple model of the organ shape (Chervenak et al., 1993; McCallum and Brinkley, 1979). With 3D ultrasound, organ volume can be measured directly from the 3D data, giving improved results. This is especially important for organs of non-standard shape and size. Examples where volume or size determination is useful include the fetus (Chervenak et al., 1993), prostrate (Rahmouni et al., 1992), ventricles (Martin et al., 1990; Ofili and Nanda, 1994; Salustri and Roelandt, 1995), kidney (Gilja et al., 1995) and blood vessels (Hodges et al., 1994).

The creation of 3D data sets also aids review of the data after the examination. Instead of producing copies of one or two of the best B-scans, the entire volume can be reviewed at will. The ability to review the 3D data, after the patient is released, can improve time efficiency and reduce costs (Nelson and Elvins, 1993). With standardisation, it can also provide opportunities for tele-medicine (Shahidi et al., 1998). It may also be easier to re-check diagnoses, which has important legal implications (Baba and Jurkovic, 1993).

3D ultrasound may also improve the fetomaternal relationship. Expectant mothers often find it hard to understand 2D images, but readily identify with 3D renderings of the fetal face. A heightened perception of the fetus may have a positive effect on maternal behaviour, *e.g.* improved diet, reduction of smoking and alcohol consumption, and better antenatal attendance (Baba and Jurkovic, 1993).

Finally, education can be enhanced by training medical students on 3D data sets. For example, a student can interactively re-slice a previously acquired 3D data set to simulate a real examination. The student can then be tested on the correct diagnosis (Shahidi et al., 1998).

It must not be overlooked that perhaps the most important advantage is improved ease of understanding of 3D spatial relationships. Mentally reconstructing a complicated 3D structure from a series of cross-sections, as required by 2D ultrasound, can be difficult For this reason and the reasons listed above, the use of 3D ultrasound is likely to become widespread in the future.

1.3 3D Ultrasound Techniques

High quality, rapid 3D imaging remains a long term research goal. One promising approach centres around the development of a new type of probe, which sends and receives echos from a 2D array of elements (instead of the usual one-dimensional array). With parallel transmission and reception, 2D array probes are capable of scanning an entire volume in the same time as a single B-scan. Unfortunately, several technical challenges must be overcome before such probes receive clinical acceptance (Light et al., 1998). Alternative approaches, which make use of conventional 2D ultrasound technology, include the freehand (Sherebrin et al., 1996), mechanical swept volume (Rankin et al., 1993; Steiner et al., 1994) and thick-slice techniques (Baba and Jurkovic, 1993).

Thick-slice volume imaging is a novel technique that produces real-time images of organ surfaces. The basic idea is to attach a cylindrical defocusing lens to a linear array probe so that a volume of tissue is scanned— see Figure 1.1. Beams diverging between 30° to 45° have been used. The ultrasound machine produces images in the same manner as B-scans, so echoes are depicted as emanating from the central plane. If acoustic interfaces produce total reflection of the ultrasound energy, renderings of the surface of the interface are produced. When the probe is rotated axially, the surface renderings also rotate, giving the appearance of changing the viewpoint.

Although the real-time surface renderings can be quite spectacular, the thick-slice technique has some fundamental limitations. First, only organ surfaces that produce nearly total reflections (such as the fetal skin after the third trimester and air-backed structures) can be imaged this way. Even visualisation of the surface renderings of these structures are limited to viewpoints from the same direction as the ultrasound beam; it is impossible to see the back sides of objects. Moreover, it can only generate surface renderings, so cross-sectional slices are unavailable. Although the thick-slice technique has some advantages, the limitations mean it is not likely to replace methods that produce full 3D data sets.

Mechanical swept-volume and freehand techniques construct a 3D data set from a number of 2D B-scans acquired in rapid succession. The mechanical swept volume approach uses a special mechanism inside the probe to sweep the plane of the B-scan through a volume of interest. A hand-held ultrasound probe has six DOF (degrees of freedom) of movement, but the swept-volume method constrains the movement of the probe to one DOF. As the probe is moved with one DOF, the scan plane sweeps through a volume, and a set of B-scans are acquired. The movement of the probe (and therefore the locations of the B-scans) is typically measured by a potentiometer. The size of the swept volume is determined by the area of the image in the B-scan and the range of the DOF. Three different shapes of swept volumes can be produced by scanning with this technique.

The first shape is produced by rotating the probe about an axis in the axial direction (see Figure 1.2 for definitions of the direction names). This method produces a cone-shaped swept volume (assuming a sector-shaped 2D image), as shown in Figure 1.3(a). The second shape is produced by rotating the probe about an axis in the lateral direction.

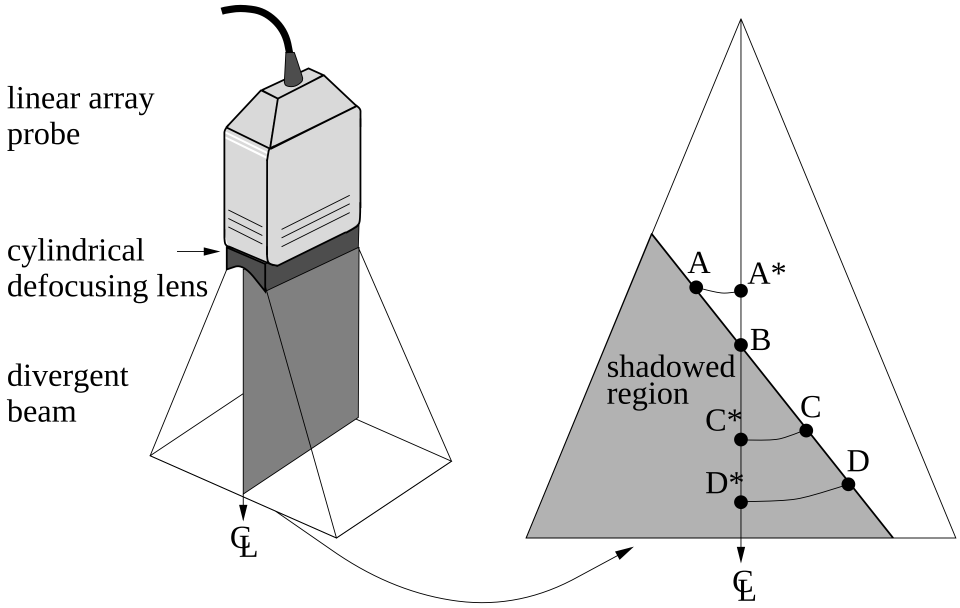


Figure 1.1: **3D thick-slice ultrasound imaging.** A defocusing lens is mounted on the face of a linear array probe to produce a divergent beam profile. If the interface ABCD totally reflects the ultrasound energy, the shadowed region does not contribute to the image. The interface is then projected onto the central image plane as $A^{*}BC^{*}D^{*}$

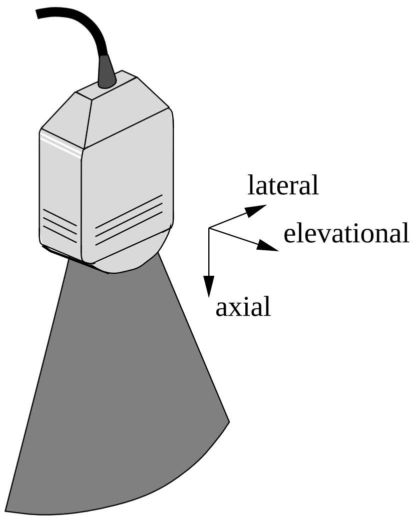
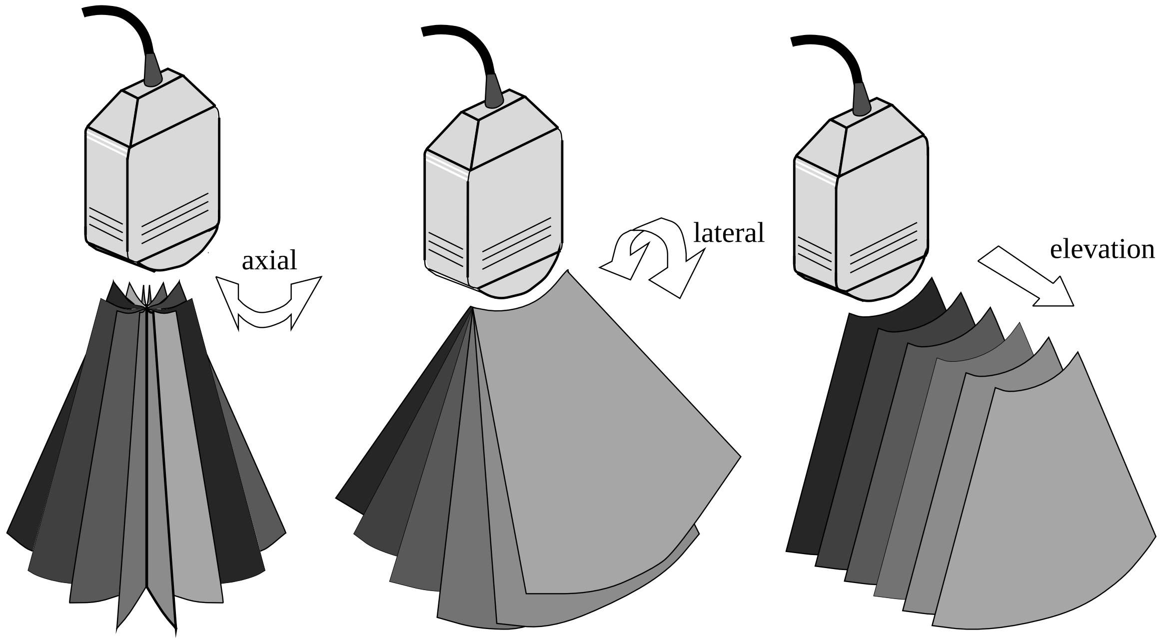


Figure 1.2: **Nomenclature of imaging directions.** The axial and lateral directions define the plane of the B-scan, also called the scan plane. The elevational direction is perpendicular to the scan plane.



(a) Cone (b) Fan (c) Prism

Figure 1.3: **Mechanical sweep motions.** (a) Rotation about an axis in the axial direction produces a cone-shaped volume. (b) Rotation about an axis in the lateral direction produces a fan-shaped volume. (c) Translation along an axis in the elevational direction produces a prism-shaped volume.

This method produces a fan-shaped swept volume (Figure 1.3(b)). Finally, a prismatic swept volume is produced by translating the probe along an axis in the elevational direction (Figure 1.3(c)).

There are advantages and disadvantages of each of these methods. The main advantage of the prismatic volume is the equal interval between each of the 2D image slices. This means the volume of interest is uniformly sampled. The fan-shaped data set has non-uniform inter-slice spacing, with increasing gaps at greater depths. The cone-shaped data set also has non-uniform inter-slice spacing, with the gaps increasing away from the centre line of the cone. The main disadvantage of the cone-shaped and fan-shaped data sets is the presence of large gaps. Acquiring a greater number of B-scans reduces the size of the gaps, but also oversamples other regions and increases acquisition times.

The main advantage of the cone-shaped and fan-shaped data sets is the small acoustic window used for the entire sweep. Acoustic windows on the human body are small and sometimes difficult to obtain. Prismatic data acquisition requires large acoustic windows, which limits the number of possible applications. One exception is 3D intravascular ultrasound, where cross-sectional images are acquired by moving a catheter-probe combination through a blood vessel (Fenster and Downey, 1996; Pasterkamp et al., 1995).

Mechanical swept volume systems are easy to use and produce standardised volumes

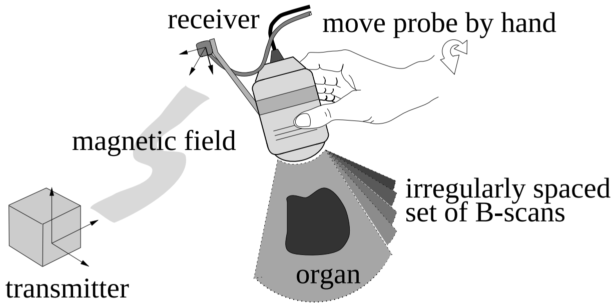


Figure 1.4: **3D freehand ultrasound.** Freehand imaging allows the physician to move the probe freely, so the B-scans can have arbitrary relative locations and may overlap each other. As the probe moves, the B-scans are captured and stored in computer memory. A magnetic position sensor measures the position of a receiver (mounted on the probe) relative to a transmitter (fixed with respect to the patient's body). The relative location of each scan plane is then calculated and the B-scans combined into a single volume of data.

of data. The first commercial 3D system (Kretztechnik Combison 330 Voluson) used the mechanical swept-volume approach, and subsequent developments of the Voluson machine have proven popular. Yet they require the considerable expense of a dedicated machine, and are also limited to a maximum volume dictated by hardware constraints in the probe.

Freehand systems allow arbitrary motion of the probe — see Figure 1.4. The motion of the probe is unrestricted and the six DOF — three translations and three rotations — are measured by a position sensor. With a freehand approach, the scan planes are chosen by the trained physician, for example in a manner consistent with a normal examination. The scan planes may intersect each other so that the same region in space may be sampled from more than one probe direction (also called look direction). Gaps may also remain where no samples are taken.

There are several advantages of a freehand imaging approach over swept-volume approaches. Freehand systems can be used to obtain arbitrary volumes of data, since the motion of the probe is unconstrained. They are also cheaper, requiring only a conventional ultrasound system and relatively inexpensive additional components. Yet the physician needs to learn how to move the probe to acquire densely sampled data sets. The cost and flexibility of freehand imaging ensures it remains a popular choice, and with recent improvements in the speed of data acquisition, combined with careful scanning practice, the quality of freehand data sets is improving. For these reasons, research into freehand systems is active, and several commercial systems have recently become available. 3D freehand ultrasound forms the basis of the research in this thesis.

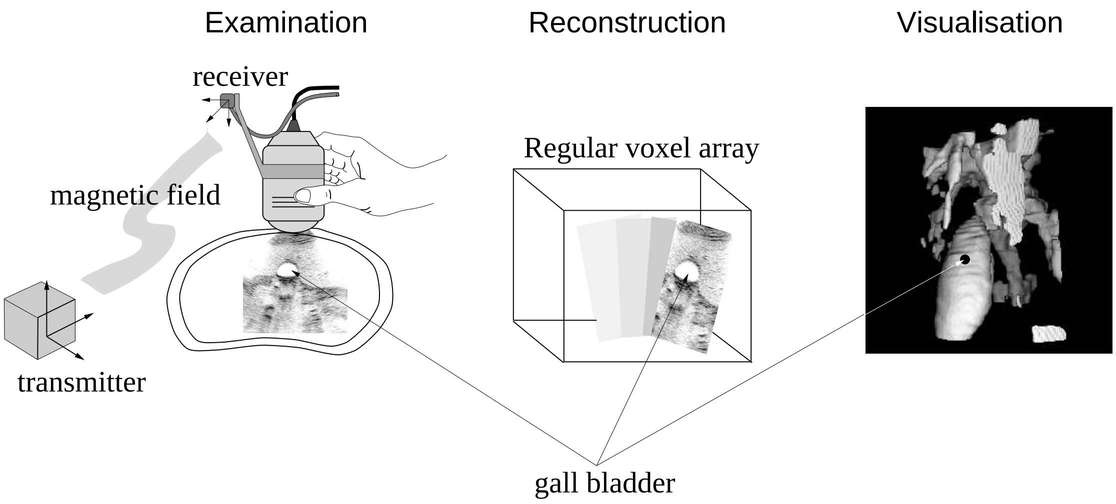


Figure 1.5: **Overview of 3D freehand ultrasound imaging.** Clinical freehand imaging involves an examination, reconstruction and visualisation. This example illustrates a study of a gall bladder.

1.4 The 3D Freehand Technique

The 3D freehand process can be considered in three stages: examination, reconstruction and visualisation — see Figure 1.5. Figure 1.6 shows a photograph of an examination in progress and Figure 1.7 shows the data paths.

In the first step of freehand imaging, a position sensor is attached to the probe. The system is then calibrated to determine the offsets between the position sensor and the corner of the B-scan plane. The offsets are used with the position measurements to calculate the relative positions of each of the acquired B-scans.

The ultrasound machine controls are then adjusted to obtain the best images for the organ of interest. Appendix A gives a brief description of the controls on an ultrasound machine and how they are typically adjusted. During the examination, the physician moves the probe slowly and steadily over the region of interest. The B-scans are digitised by a frame grabber, tagged with the measurements from the position sensor, and stored in computer memory as 2D arrays of pixels (picture elements).

The set of B-scans are then assembled into a 3D data set using the associated position data. This process is called reconstruction. Usually the format of the 3D data set is a scalar array of voxels (volume elements). The 3D scalar array is also known as a cuberille.

In the visualisation stage, portions or all of the reconstructed 3D data are displayed on the computer monitor. The important properties of a good image are clear detail, visualisation of detail at depth, distinguishable tissue types, low noise and consistency of these properties throughout the image (Chervenak et al., 1993). A bad image has no visible diagnostic information. The goal of the visualisation method is to produce good

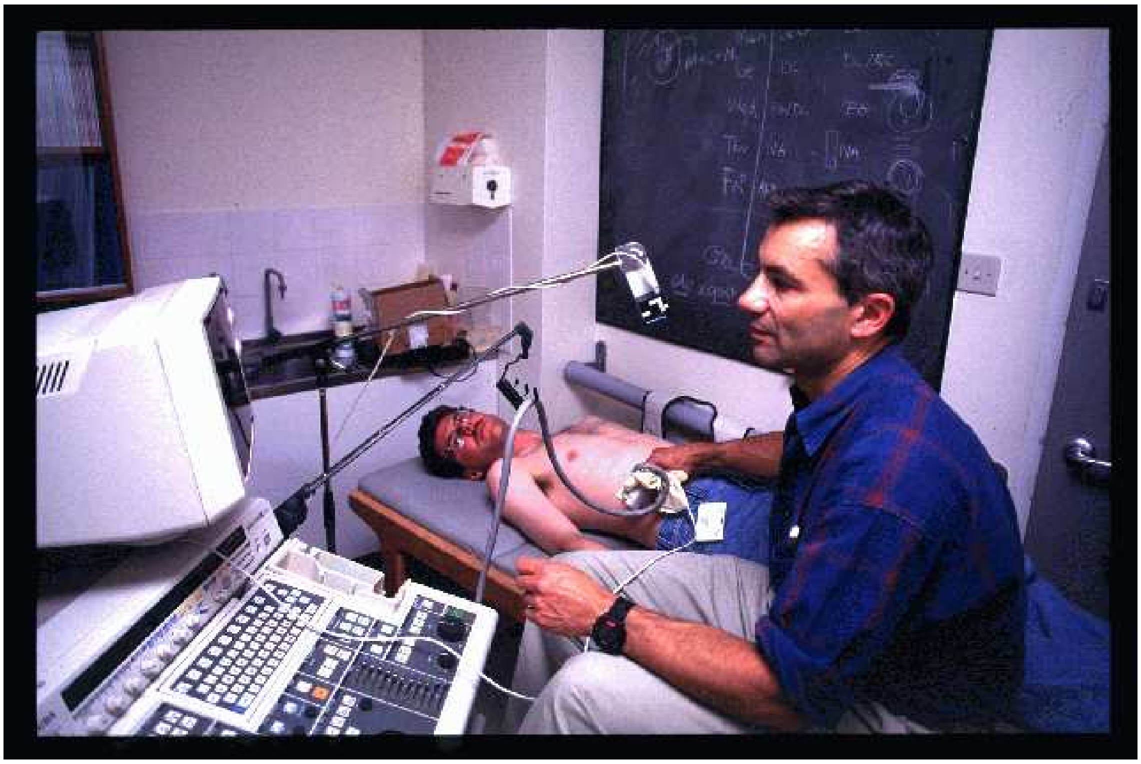


Figure 1.6: **3D freehand ultrasound examination.** The author is examined by Dr. Lol Berman in the Department of Radiology at Addenbrooke's Hospital.

images that accurately represent the anatomy. Several good reviews of the visualisation of 3D medical data sets are available (Shahidi et al., 1998; Soferman et al., 1998; Taylor et al., 1996).

There are many different ways to display 3D data sets. The most conventional is to view slices of the data set in a manner similar to normal 2D scanning (Chervenak et al., 1993; Halliwell et al., 1989; Wells, 1993). The difference is that slices from any look direction may be displayed, regardless of whether any of the original 2D images were taken in that particular look direction. This is called any-plane slicing.

Efficient algorithms for any-plane slicing allow images to be generated at interactive rates, making it easy to find structures of interest. The review process is still 2D in nature, and diagnoses are made in a manner similar to standard B-scan imaging. One difference between any-plane slicing and B-scan imaging is that the 3D data set may have been substantially processed during reconstruction. This may include interpolation and filtering of the data. Therefore, even if a selected look direction coincides with one of the original look directions, the displayed image may not be exactly the same as the original B-scan.

As well as cross-sectional images, it is also useful to create 3D renderings to get a better understanding of the true shape of the organs. Two techniques, surface rendering and volume rendering, achieve this through different means.

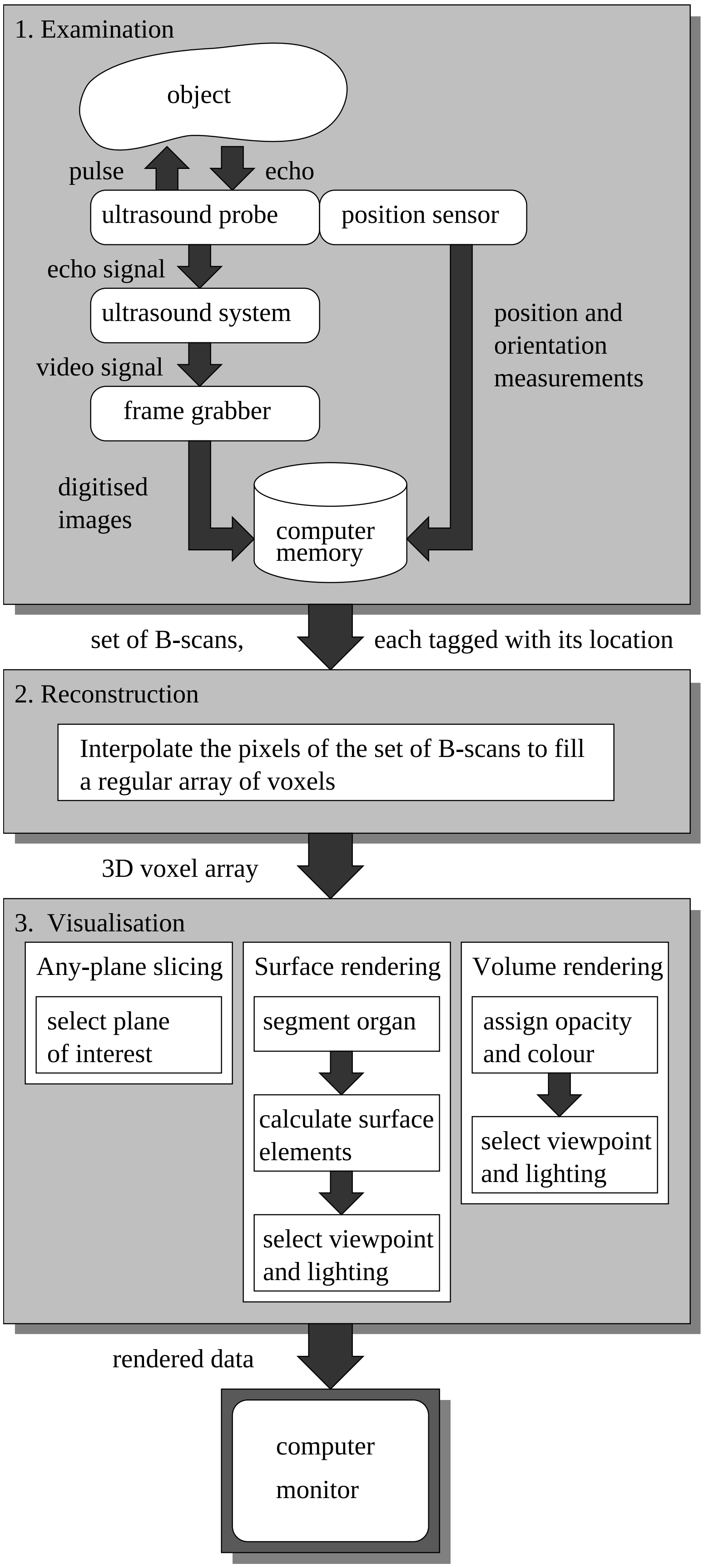


Figure 1.7: **Data paths of 3D freehand ultrasound imaging.**

Surface rendering requires a surface representation of an organ to be derived from the 3D data set. This normally involves classifying the voxels according to tissue type. The result is a binary segmentation of the organ of interest. This can be done by manual outlining, grey-level thresholding or more sophisticated methods (Bouma et al., 1996; Bouma et al., 1997). Next, the segmentation results are used to calculate a set of surface elements. The surface can be depicted as a set of points, line segments, polygons, patches or cubes. A popular method is "Marching Cubes", which constructs a surface of small triangles from the binary segmentation (Lorensen and Cline, 1987). Visualisation then proceeds on the surface representation alone, allowing it to be viewed from chosen viewpoints and under selected lighting conditions.

In this thesis, surface rendering is performed with the 3DViewnix software package (Udupa et al., 1993) using the "semiboundary" approach (Udupa and Dewey, 1991). Here, a binary version of the 3D data set is created by segmenting the object of interest by grey-level thresholding. The semiboundary representation is created from filled (segmented) voxels that are surrounded by at least one empty voxel. Each voxel in the semiboundary also stores the gradient of the data at that point. The semiboundary is then displayed from any chosen viewpoint using the gradient information for shading. Finally the lighting direction, and specular and diffuse reflection coefficients, are adjusted to optimise the image.

The advantages of surface rendering include the relatively short rendering times and reduced storage requirements (Fenster and Downey, 1996). The surfaces look very clear since most obstructing or noisy data are normally removed by segmentation. The drawbacks of using surface fitting algorithms on ultrasound data are the production of false positive and false negative surface pieces, and incorrect handling of small features (Nelson and Elvins, 1993). These drawbacks arise because segmentation, including grey-level thresholding, is particularly difficult with ultrasound data (Sakas and Stefan, 1995; Sakas et al., 1995; Steen and Olstad, 1994)

Volume rendering maps the 3D data set directly into screen space. Parallel ray casting is the most common method of volume rendering (Watt, 1993). First, each voxel is assigned a colour and opacity, normally calculated from simple functions of the original voxel intensity values. Rays are then cast from each screen pixel through the 3D data set. The colour and opacity of each voxel that the ray passes through are blended until the opacities reach unity or the ray exits the rear of the volume of data. The blended opacities and colours are then displayed by the screen pixel. The procedure is repeated for all screen pixels to form an image. Additional visual cues, such as shadowing and reflections, may also be added.

The main advantage of volume rendering is the use of the entire 3D data set for each 2D display. No *a priori* segmentation is required, so surface fitting artifacts are eliminated. It has also been suggested that the averaging process of ray casting reduces the effect of noise on the rendered image (Wells, 1993). The main disadvantage of volume rendering is related to its main advantage: the rendering process is slow because the entire 3D data set is used for each new viewpoint. It is also difficult to see the interior details of a large, noisy data set without removing some of the surrounding data.

Standard surface and volume rendering methods used successfully in MR (magnetic resonance) and CT (computed tomography) imaging have limitations when applied to ultrasound data. Ultrasound images have several unique properties (Sakas et al., 1995): a significant amount of noise; lower dynamic range; blurred boundaries several voxels wide; boundaries with varying grey levels caused by variations in surface curvature and orientation; shadowed regions; and non-uniform spatial resolution. This list demonstrates that visualisation of ultrasound data is a challenging problem. Fetuses are exceptions that are easier to render because they are surrounded by anechoic amniotic fluid. It is relatively simple, therefore, to isolate the fetus from surrounding tissue and noise (Steiner et al., 1994). Yet for most other examinations, the standard surface and volume rendering methods are unsuccessful and new methods are required (Sakas et al., 1995).

In addition to visualisation, the 3D data sets are often subject to further data analysis. Examples of data analysis include size and volume estimation (Chervenak et al., 1993), organ shape modelling (Cootes et al., 1995; Neveu et al., 1994; Taylor et al., 1996) and texture classification (Chervenak et al., 1993; Wells, 1993). All of these techniques can be used for making diagnoses.

This thesis is concerned with improving the performance of 3D freehand ultrasound. Each of the steps of the 3D freehand process involves measurement errors. In order to get the best performance out of the system, it is important to consider the sources of error that affect the quality of the results.

1.5 Sources of Measurement Error

1.5.1 Ideal Imaging Process

Since 3D freehand ultrasound is made up of a set of B-scans, it is helpful to look at the physics of an idealised 2D ultrasound imaging process:

> *A finite length ultrasound pulse is transmitted by the transducer in the probe. The pulse passes uniformly through the material, and the scatterers that make up the material reflect part of the pulse and transmit the rest deeper into the material. The pulses are transmitted and reflected linearly. The amplitude and the timing of the pulses received by the transducer are converted into intensity-distance data. The higher attenuation of pulses returning from deeper locations is compensated for using the exact attenuation coefficient of the material. As the beam is swept through a plane, the set of collected intensity-distance data is used to create a 2D image.*

In reality, the interactions described in the ideal case are non-linear and include additional errors. The pulse consists of a small range of frequencies, instead of a single frequency. The amplitude and frequency properties of the pulse change in a non-linear

Frequency (MHz)	Axial Resolution (mm)	Maximum Imaging Depth (mm)
2	0.77	300
3.5	0.44	170
5	0.31	120
7.5	0.20	80
10	0.15	60

Table 1.1: **Maximum imaging depth and axial resolution.** Approximate values are listed for common frequencies of ultrasound probes. Calculations are based on a two cycle pulse propagating through human tissue (Kremkau, 1993).

manner as it passes through the tissue. For example, the speed of propagation and amplitude attenuation depend on both the type of tissue the ultrasound pulse traverses and the pulse frequency. Both specular and diffuse reflection occur. The ultrasound beam can be refracted or de-focused. Real transducers produce non-uniform beams. In 3D freehand ultrasound, the position sensor and system calibration also contribute errors. Despite these phenomena and errors, images useful for diagnosis are still obtainable by careful selection of the imaging parameters.

The following is a review of the measurement errors and limitations to give an intuitive feel for the types of problems that arise in 3D freehand ultrasound imaging. The categories are not independent nor comprehensive, but cover the most significant effects. The description of these categories demonstrates that the interaction of ultrasound in biological materials is not fully understood and is an active area of research (Evans and Martin, 1988; Evans, 1986; Hill, 1986).

1.5.2 Sonographic Artifacts: Propagation Group

Axial Resolution

Axial resolution is the minimum distance between objects positioned axially that are imaged individually. A first approximation is the spatial pulse length divided by two (Harris et al., 1991; Kremkau, 1993; Wells, 1993) — see Figure 1.8. Typical values are listed in Table 1.1. For a given frequency, the axial resolution varies slightly throughout the image because of changes to the pulse shape as it travels through the medium (Harris et al., 1991; Preston, 1991; Wells, 1993). Despite changes in pulse shape, the first approximation is close to the real axial resolution, and a number of ultrasound systems have demonstrated resolutions similar to the Table 1.1 values (Chervenak et al., 1993).

Table 1.1 also shows that the physician is faced with the trade-off between imaging depth and axial resolution. High frequency pulses give high axial resolution but low imaging depth because of higher attenuation (Chervenak et al., 1993; Kremkau, 1993; Wells, 1993). The physician normally chooses the highest frequency probe capable of

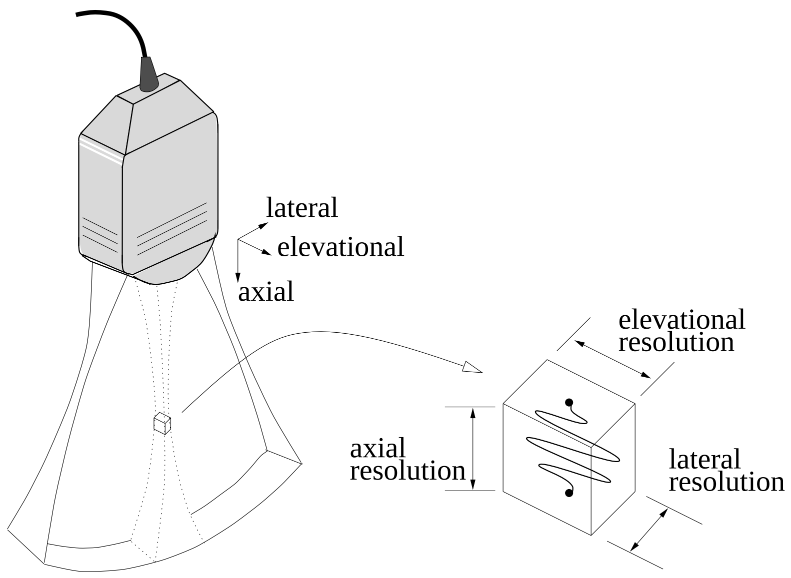


Figure 1.8: **Spatial resolution.** Axial resolution is related to the pulse length. Lateral and elevational resolutions are related to the focusing of the beam.

imaging to the required depth.

Lateral Resolution

Lateral resolution is the minimum distance between objects positioned laterally that are imaged individually. 2D images are produced by steering a beam through the scan plane. Most modern probes use an array of transducer elements and steer the beam either by selective activation of a subset of elements or by selective timing of the elements' activations. The lateral resolution may therefore be approximated as the beam width in the scan plane (Kremkau, 1993; Wells, 1993). The beam width is usually taken as the 6 dB width of the amplitude distribution.

The beam shape depends on the type of probe used, the intensity of the ultrasound pulses and the properties of the tissue. The beam width is a function of axial distance and has a minimum cross-sectional area at the focal point — see Figure 1.8. For single focus transducers, the lateral resolution is worse away from the focal region. Probes using multiple focal points help to maximise lateral resolution throughout the image.

For linear array probes, it has been shown (Whittingham, 1991) that the diameter of the focal spot can be calculated from the focal length (F), the wavelength (λ) and the radius of the effective aperture (a):

$$\text{lateral resolution} = \frac{F\lambda}{a}. \tag{1.1}$$

For $F = 50$ mm, $\lambda = 0.44$ mm (3.5 MHz wave in average human tissue) and $a = 10$ mm (the transducer effective radius) the beam width is 2.2 mm. Experiments on a circular disk transducer performed *in vivo* and *in vitro* verify the calculations (Moshfeghi and

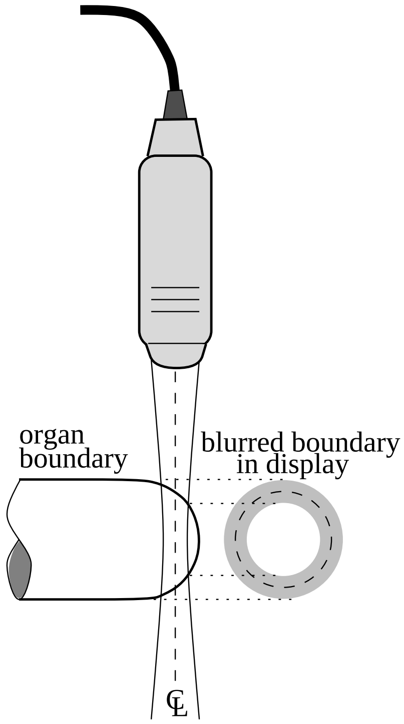


Figure 1.9: **Beam width artifact.** The thin interface of the organ boundary becomes blurred because of the finite width of the beam.

Waag, 1988). Advanced techniques and higher frequency probes produce higher lateral resolutions.

Elevational Resolution

Resolution is also limited in the elevational direction because of the finite beam width in the elevation plane. Alternatively, artifacts are introduced in the B-scan by objects that are off the imaging plane, but still in the beam path (Chervenak et al., 1993; Fish, 1990; Kremkau, 1993). The finite elevation thickness also blurs sharp material boundaries — see Figure 1.9.

The beam's elevational width varies with the type of probe, imaging depth and material properties. On most modern ultrasound machines it is focused by a fixed lens. A study of the elevational beam width of a 3.5 MHz phased array sector probe gives beam widths ranging from 10 mm at a 2 cm depth, to 6 mm at a 12 cm depth, with a minimum of 3 mm at the 8 cm focus (Skolnick, 1991).

Speckle

The grainy appearance of 2D ultrasound images is caused mainly by speckle — see Figure 1.10. The speckle phenomenon results from the constructive-destructive interference of the coherent ultrasound pulses back-scattered from the multiple tiny reflectors that constitute biological materials.

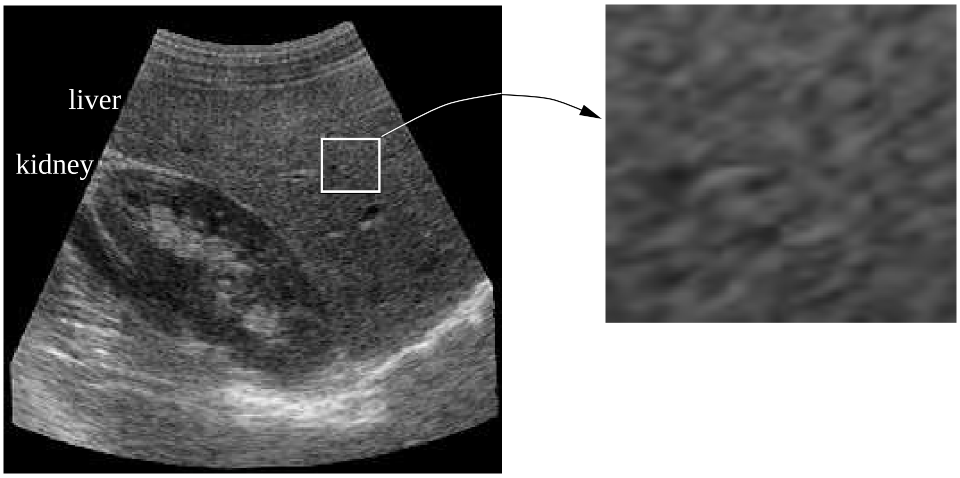


Figure 1.10: **Speckle.** A B-scan of the right kidney and liver is shown on the left. The expanded portion on the right clearly shows the grainy texture of the liver caused by speckle.

In some situations, the speckle pattern provides useful information for the physician, such as for classification of tissue type and identification of organ pathology. But in many other situations, speckle is considered an annoyance and a hindrance to diagnosis. Speckle unfortunately falls into the high sensitivity range of spatial frequency in human vision. The frequency spectrum of speckle is also similar to the imaging system modulation transfer function. The combination of these two effects means speckle can obscure important diagnostic information in ultrasound images (Wells, 1993). For this reason, speckle is removed in some applications to improve diagnosis. For example, filters are often employed in volume rendering of 3D ultrasound data sets (Bamber et al., 1992; Steen and Olstad, 1994).

Reverberation

Reverberation is the phenomenon whereby multiple reflections are produced off the sensor face. Ideally the sensor face should have the same acoustic impedance as skin, so that all ultrasound energy is transmitted, and none reflected. Unfortunately, there is a substantial difference in acoustic impedances (Chervenak et al., 1993; Fish, 1990; Kremkau, 1993) so multiple reflections can occur — see Figure 1.11. Matching layers, such as coupling gel, help to reduce this effect. "Comet tails" (often seen below hard objects such as kidney stones) are also caused by reverberation, but the reverberation occurs within the object.

Refraction

Refraction of the beam changes the beam direction and results in distorted images — see Figure 1.12. Objects are located at the wrong position because the true beam path differs from the estimated path (Chervenak et al., 1993; Fish, 1990; Kremkau, 1993).

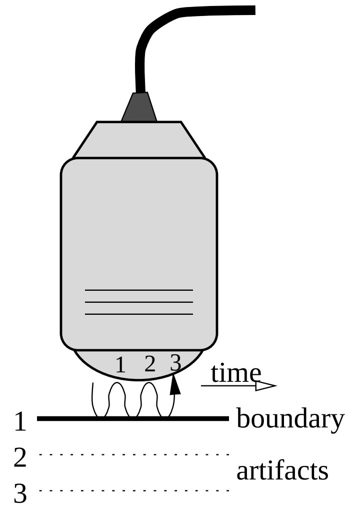


Figure 1.11: **Reverberation artifact.** Reverberation off the sensor face produces multiple echoes.

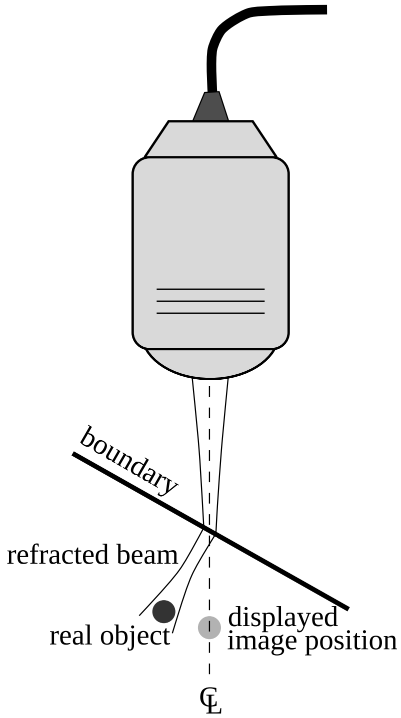


Figure 1.12: **Refraction artifact.** Incorrect placement of the object results from refraction of the beam direction.

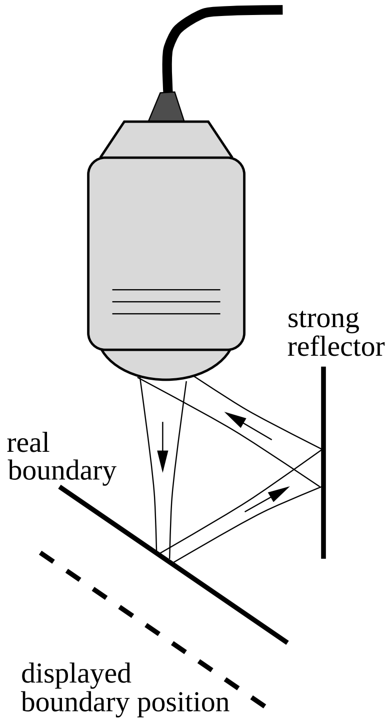


Figure 1.13: **Multi-path artifact.** Incorrect placement of the object results from a lengthening of the beam path through multiple reflections.

Multi-path Reflections

Improper positioning of an object can also result from a difference in the length of the return path from the incident ultrasound path. An example is shown in Figure 1.13. In general, multi-path reflections blur the object boundaries in the image (Fish, 1990; Kremkau, 1993). They can also produce mirror-images of the object, often seen below the uterine wall in fetal images.

Side Lobes and Grating Lobes

Side lobes are minor beams of sound travelling in a direction different from the main beam — see Figure 1.14. Side lobes of sufficient intensity produce echoes from objects off the main beam path (Chervenak et al., 1993; Fish, 1990; Kremkau, 1993).

Grating lobes are lobes produced by array transducer probes (Fish, 1990; Kremkau, 1993). As the number of array elements increases and the gaps between the elements decrease, the size of the lobes decreases (Whittingham, 1991). The lobes also change shape as the beam is steered. Apodization – unequal weighting of the transmit or receive elements' signals – reduces grating lobes but increases the width of the main beam (Wells, 1993). Lobes off the main beam are present to some degree in all ultrasound systems.

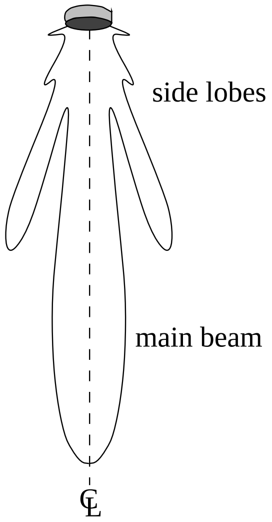


Figure 1.14: **Side lobes of an unfocused disk transducer.** The side lobes of ultrasound can produce artifacts from strong reflectors off the main beam path.

Tissue	Propagation Speed (mm/μs)
Fat	1.44
Brain	1.51
Liver	1.56
Kidney	1.56
Muscle	1.57

Table 1.2: **Propagation speeds of soft tissue *in vivo*.**

Speed Errors

Ultrasound systems typically use a propagation speed of 1.54 mm/μs to calculate the distance to reflectors from the elapsed return time of the echoes (Kremkau, 1993). The value of 1.54 mm/μs is the average propagation speed of ultrasound in soft human tissue. The actual propagation speed varies with the type and temperature of the tissue. Table 1.2 lists the propagation speeds of several types of commonly imaged tissue at normal body temperature. Objects located in materials with a propagation speed greater than 1.54 mm/μs will be incorrectly placed too near the probe and vice versa (Fish, 1990; Kremkau, 1993). The shape of objects within the image will also be distorted — see Figure 1.15

A less significant effect is the dependence of propagation speed on frequency for a given material. For example, the ultrasound propagation speed in the human brain (measured *in vitro* at 37° Celsius) changes from 1.561 mm/μs to 1.566 mm/μs as frequency is increased from 1 to 10 MHz (Hill, 1986).

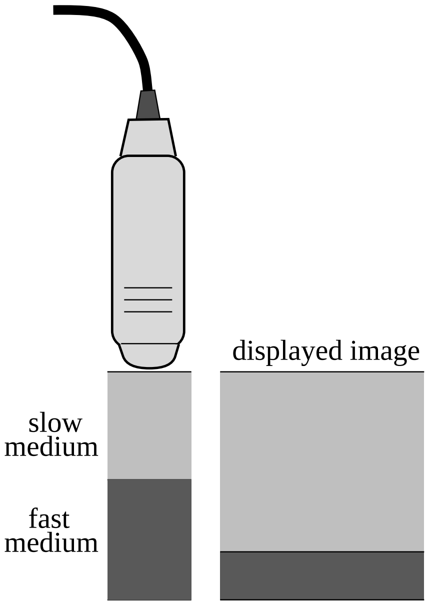


Figure 1.15: **Speed errors.** The difference in the speed of sound in the two types of tissue results in expansion of the slower medium and compression of the faster medium.

Range Ambiguity

The maximum depth that can be imaged unambiguously is determined by the pulse repetition frequency (Kremkau, 1993). In other words, a second pulse should be transmitted only after all echoes are received from the first pulse. Otherwise, echoes from the first pulse will appear as artifacts in the second image. In reality, this is not a problem if the physician chooses the correct probe and pulse repetition frequency for the desired imaging depth. The trade-off that occurs is between range ambiguity and frame rate.

1.5.3 Sonographic Artifacts: Attenuation Group

Contrast Resolution

Contrast resolution is the smallest detectable difference in echo intensity. The contrast resolution perceived by the physician depends on the ultrasound system, the computer monitor and the human visual system. The echo contrast dynamic range is usually compressed to the range most sensitive to the human eye. For example, the ratio of smallest to largest intensities of returning echoes may be 100 dB. A typical display with 256 grey levels has a dynamic range of 24 dB. The dynamic range of human perception of grey levels is approximately 20 dB (Kremkau, 1993).

High contrast resolution also benefits spatial resolution. Individual detection of two closely spaced objects depends on measuring the amplitude of two separate but partially overlapping distributions — see Figure 1.16. Contrast and spatial resolution are also affected by noise and artifacts that obscure or distort objects in the image.

Furthermore, contrast resolution is also dependent on the processing of the echo data.

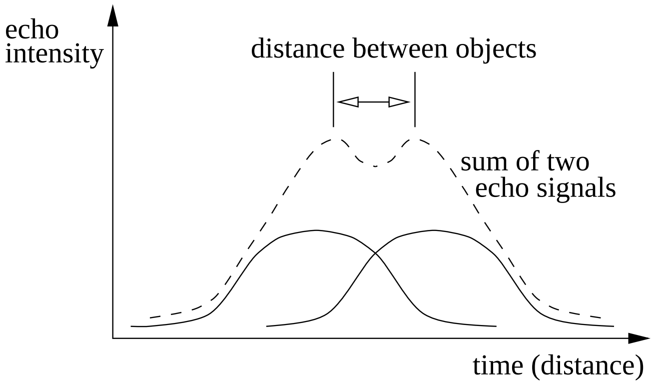


Figure 1.16: **Contrast resolution.** For two echoes to be distinguishable individually, the depression between the two peaks of the summed signals must be detectable — an issue of contrast resolution.

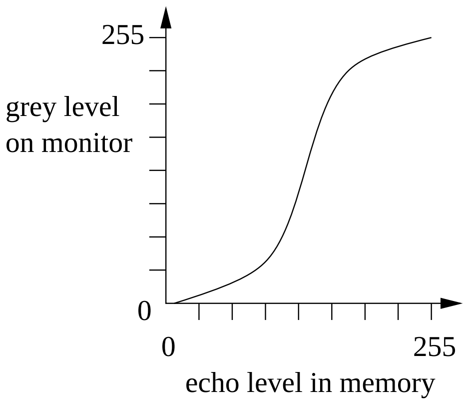


Figure 1.17: **Post-processing curve.** The curve is used to convert between the echo levels in memory and grey levels on the display monitor. A steep slope improves contrast resolution, although a steep slope over one range of echo levels is accompanied by a shallow slope over the remaining ranges.

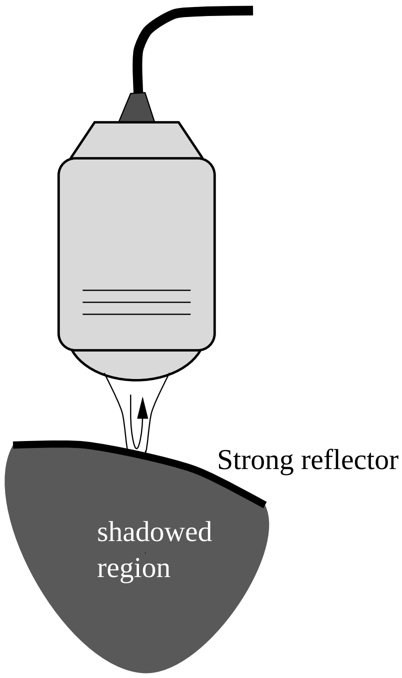


Figure 1.18: **Shadow artifact.** When the transmitted acoustic energy at a boundary approaches zero, no information is obtained from tissue below the boundary. This results in a dark shadowed region.

Some ultrasound systems allow manual control of the post-processing curve that relates echo levels after processing (amplification, demodulation, filtering and compression) to grey levels on the display monitor (Kremkau, 1993). A steep slope of the curve produces very different grey levels from echo levels of near equal amplitude. The physician can improve contrast resolution over a portion of the dynamic range at the expense of the remaining dynamic range — see Figure 1.17. In general, image processing techniques that enhance visibility of features in 2D images also enhance visibility in 3D.

Shadowing

Shadowing occurs when ultrasound pulses are greatly attenuated behind strong reflectors such as bone or lung (Chervenak et al., 1993; Fish, 1990; Kremkau, 1993), as shown in Figure 1.18. The only way to produce image data from the shadowed region is to view it from another look direction that is not blocked by the strong reflector.

Shadows can also be caused when the beam is refracted and de-focused by certain tissue interfaces. If the beam is sufficiently de-focused, then the returning echoes are too weak to be detectable. The image, therefore, exhibits a shadowed region below the interface (Chervenak et al., 1993; Kremkau, 1993).

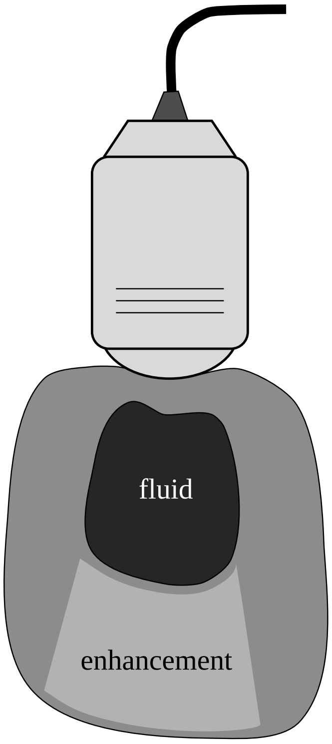


Figure 1.19: **Enhancement artifact.** When the ultrasound pulse passes through an anechoic region, such as a fluid filled cavity, little of the acoustic energy is absorbed. Since the time-gain amplification attempts to compensate for higher attenuation at greater depths, over-amplification occurs below the anechoic region.

Enhancement

Enhancement is the opposite of shadowing. Some tissues or fluids (such as amniotic fluid) are anechoic, so they do not reflect or absorb much ultrasound energy. Regions below anechoic tissues or fluids return echoes of greater than expected amplitude (Chervenak et al., 1993; Kremkau, 1993), as shown in Figure 1.19. Enhancement can also occur because of the inherent higher intensities near the beam focus (Kremkau, 1993). TGC (time-gain compensation) amplifies distant echoes more than proximal echoes to account for the greater attenuation of distant echoes. Manual control of the TGC curve is required because of the variation in attenuation coefficient among different types of materials. The physician typically controls the shape of the curve to reduce the enhancement phenomenon.

Anisotropy

It has been found that some tissues contain anisotropic scatterers that only appear from certain look directions (Moskalik et al., 1995). The physician has been trained to find the look directions that offer high visibility of anatomical features. Freehand imaging has a clear advantage over other 3D imaging methods by allowing the physician to choose these look directions during the examination.

Electronic Noise

Electronic noise is unavoidably introduced during the analogue processing of the echo pulses. The analogue process normally involves amplification, demodulation and compression of the signals before digital conversion. Electronic noise dominates at low intensities and ultimately limits the maximum imaging depth (Wells, 1993).

1.5.4 Position Sensor

Sensor Accuracy and Precision

One way to measure the position and orientation of the probe for 3D freehand imaging is to use a transmitter-receiver system. The two most popular sensors are the Fastrak (Polhemus Incorporated, Colchester, Vermont, USA), and the Flock of Birds (Ascension Technology Corporation, Burlington, Vermont, USA). Although the Fastrak and the Bird sensors both use magnetic fields in their operation, the Fastrak is energised by bursts of sinusoidal current, and the Bird is energised by rectangular pulses of direct current. Thus the Fastrak is an AC (alternating current) system, and the Bird is a quasi-DC (direct current) system. Since the earth's magnetic field contributes to the DC magnetic field, the Bird sensor measures the earth's field and subtracts it from its own signals. The accuracy and precision ratings of the two sensors are listed in Table 1.3.

Both sensors have sufficient accuracy and resolution for 3D freehand ultrasound, but the errors in orientation can be amplified because of the lever-arm effect. The position sensor is attached at a location above the top of the B-scan, so pixels at the bottom of the B-scan are furthest away. A small angle of error in orientation can result in a large error in position for deeper pixels. The position sensor is normally attached to the probe with a small extension bar, as shown in Figure 1.4. The length of the extension bar is therefore kept as short as possible to minimise the effect of orientation errors, while avoiding interference with the probe.

The performance of the Fastrak and Bird sensors is also affected by nearby metal objects and interfering fields (Nixon et al., 1998). In practice, errors are minimised by carefully preparing the examination room before tests. For example, a wooden bed is normally used for examinations, a clear path is kept between the transmitter and receiver, and other field generating devices are placed as far away as practical.

A number of other position sensors have also been used for 3D freehand ultrasound. Spark-gap acoustical sensors were among the first devices to be used, but accuracy can be severely degraded by changes in air pressure and temperature (King et al., 1991). Also, patients did not always tolerate the spark-gap devices near their bodies. Mechanical arms (Ohbuchi et al., 1992) offer high performance and robustness (Rohling et al., 1994), but are relatively expensive and can be obtrusive during an examination. Optical-video systems (State et al., 1994; Trobaugh et al., 1994) can provide exceptional accuracy and precision (Rohling et al., 1994) but are also expensive and require a clear optical path between the receiver and the transmitter. The cost-performance ratio of the magnetic

Sensor	Position		Orientation		Notes
	Accuracy	Resolution	Accuracy	Resolution	
Polhemus Fastrak	0.76 mm	0.38 mm	0.15°	0.025°	(a)
Ascension Flock of Birds	1.78 mm	0.76 mm	0.5°	0.1°	(b)

Table 1.3: **Manufacturers' specifications of Bird and Fastrak position sensors.** Accuracy ratings are expressed as the RMS (root mean squared) errors from the true positions and orientations. Resolution ratings signify the smallest detectable difference in measurements. Notes: (a) specifications at 760 mm separation between the transmitter and receiver; (b) specifications at 305 mm separation.

position sensors is the main reason why the Polhemus Fastrak and the Ascension Bird are the most popular position sensors for 3D freehand ultrasound. The Fastrak position sensor is used for all experiments in this thesis.

Sensor Latency and Acquisition Rate

A moving sensor will provide readings of past positions because of latency: the time between movement of the sensor and the corresponding change in the recorded position. For example, a 4 ms latency (Polhemus manufacturer rating) for a probe moving at 100 mm/s gives a position error of 0.4 mm. Slow motion of the probe during an examination reduces the effect of latency errors. If constant, the latency can be measured and its effect subtracted from the position measurements (Prager et al., 1998a).

Although the position sensor can produce readings at a high rate, the video frame grabber normally limits the acquisition rate. A low acquisition rate requires the physician to move the probe slowly to avoid large gaps between images. So, for a system acquiring 25 fps (frames per second), the probe must be moved at 25 mm/s to get 1 mm spacing between B-scans.

Calibration

The angles and offsets from the position sensor receiver to the corner of the B-scan are used in the reconstruction stage to determine the relative positions of the B-scans. Since calibration determines these angles and offsets, any inaccuracies in calibration are carried over as systematic errors in the reconstruction.

1.5.5 Tissue Movement

Movement of the subject during a prolonged examination results in errors when reconstructing a 3D data set from 2D images. There are few clinical applications where the imaged anatomy is nearly stationary (an exception is the examination of a breast immobilised in a mammographic unit). In many examinations, significant errors arise from heart pulsations, breathing, varying pressure on the skin in contact with the probe and

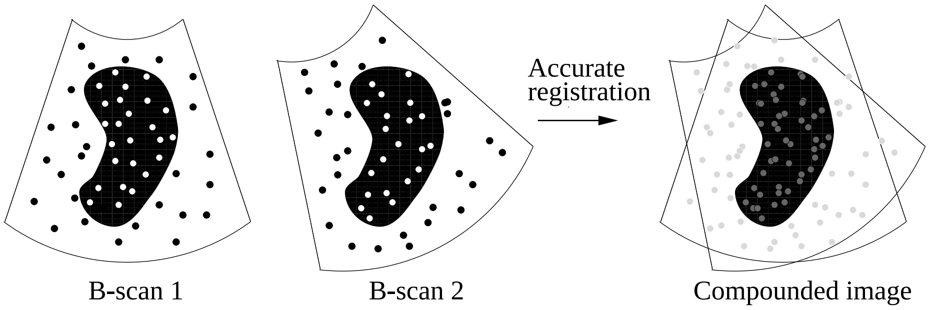


Figure 1.20: **Spatial compounding.** This is a simple illustration of 2D spatial compounding. Two B-scans are taken of the same object, in the same plane, but from different viewpoints. The two B-scans are accurately aligned and then averaged together. The averaging process produces a compounded image with an improved SNR. The principle extends to 3D, where compounding can be performed wherever scan planes intersect.

whole-body movement. Some of the errors involve simple translations and rotations of the subject, but others produce non-rigid tissue deformation. Motion-related errors can be minimised by increasing the acquisition rate, thus reducing the examination time. ECG (electrocardiogram) gating and temporary suspension of breathing can also reduce motion artifacts. Nevertheless, the design of systems that are tolerant to tissue movement remains one of the major challenges in 3D ultrasound.

1.6 Spatial Compounding

Considerable research has been directed towards methods to reduce ultrasound artifacts in an effort to improve diagnosis. Spatial compounding is one method that has been investigated in the past. The underlying principle is to average multiple images of the same region of interest taken from different look directions. The speckle signal, which de-correlates from different look directions, is suppressed by the averaging operation. Conversely, real anatomical features (tissue boundaries, for example) will be observed in the same location from all look directions. Provided the alignment of the scan planes is accurate, the averaging operation will highlight the real anatomical features — see Figure 1.20.

The different viewpoints must be separated by a minimum distance to fully de-correlate the speckle patterns. Translation of forty percent of one transducer aperture diameter was experimentally derived for de-correlation in the beam focus of a linear phased array probe (Trahey et al., 1986b). For fully uncorrelated speckle, it has been established that compounding multiple B-scans lying in the same plane can improve the SNR (signal to noise ratio) by as much as the square root of the number of B-scans used (Burckhardt, 1978; Hernandez et al., 1996; Kerr et al., 1986). SNR is defined here as the ratio of the mean grey level to the standard deviation, for an image with no resolvable structures. This is a commonly used figure of merit for imaging systems (Trahey et al., 1986b). Spatial compounding has also been demonstrated to fill in shadowed regions (He, 1997) and reduce other artifacts (Wells, 1993).

Unfortunately, there is a fundamental trade-off between speckle reduction and spatial resolution: multiple compounding results in a decrease in spatial resolution (Wells, 1993). If structures in one image are not exactly coincident with the same structures in another image, then the compounded image will appear blurred. Even if an accurate position sensor is used to measure the locations of the B-scans, other measurement errors still result in a small amount of misalignment. The main errors that cause misalignment (also called registration errors) include refraction, multi-path reflection, speed errors, sensor errors, latency, calibration and tissue movement. Despite these errors, if the misalignment is small, then the trade-off may be acceptable.

In addition to spatial compounding, there are a number of other techniques for reducing speckle in ultrasound images. These include filtering, frequency compounding, temporal compounding and signal reconstitution.

Filtering is commonly used, and a variety of different filters have been suggested (Crawford et al., 1993; Evans and Nixon, 1995; Sakas et al., 1995). For example, median filters attempt to suppress speckle while retaining sharp edges. More sophisticated filters attempt to adapt themselves for different parts of an image. In this way, the filter is optimised for each type of tissue in the image, instead of using a single filter for the entire image. But finding a filtering method that works well on a wide variety of images is a difficult challenge. Moreover, unlike compounding techniques, filtering does not introduce any new information.

Frequency compounding is similar to spatial compounding, except that different pulse frequencies are used instead of different look directions (Trahey et al., 1986a). Compounding images taken at different frequencies reduces speckle because speckle also de-correlates with a change in frequency. Unfortunately, if the transducer's available bandwidth is divided into a number of partially overlapping spectra for frequency compounding, then each pulse contains a narrower range of frequencies. This limits the ability to create short pulses, so axial resolution is reduced. Because of this trade-off, little success has been achieved with frequency compounding (Wells, 1993).

Temporal compounding is closest to spatial compounding since it directly averages conventional B-scans. It is called temporal compounding because it is implemented with a variable persistence display that averages a number of successive frames (Chervenak et al., 1993; Wells, 1993). The average of several closely spaced B-scans will show a reduction in speckle since the look direction has changed slightly. The main problem with temporal compounding is that the images being averaged do not generally depict the same anatomical region. Yet, if the ultrasound probe is moved slowly, and the anatomy is slowly changing, the anatomical features will be only slightly blurred. In many cases, physicians prefer a small level of display persistence and simply use the speed of the probe to control the level of compounding.

Signal reconstitution refers to techniques that attempt to recover information in the echo signals before it is lost in the image formation process. For example, the speckle artifact can be recognised in the phase domain using the original echo signals (Leeman and Seggie, 1987). This suggests it can be removed before image formation, but signal reconstitution remains a difficult challenge.

In summary, there are a number of alternative techniques for reducing speckle, but many produce images with considerably reduced spatial resolution (Wells, 1993). Spatial compounding is particularly attractive with 3D freehand ultrasound, since the locations of the B-scans are already measured by the position sensor. It can also be performed with a standard freehand system without the need for any internal modifications.

1.7 Motivation and Original Contributions

1.7.1 Reconstruction

As already mentioned, 3D ultrasound imaging has proven diagnostic utility in a number of different clinical applications. This thesis is concerned with making advances in 3D freehand ultrasound imaging in an effort to make better diagnoses. It is sensible to begin with the reconstruction process in order to obtain the highest quality data. First, accurate calibration of the position sensor to the B-scans is required to correctly place each B-scan into the reconstruction volume. Then an interpolation step is needed to fill the voxel array while preserving as much of the image information as possible. These two issues are tackled in Part I.

To start, Chapter 2 looks into the problem of calibration by comparing the various calibration techniques found in the literature. The most important contributions of this chapter are the establishment of a common mathematical framework, and a comprehensive series of comparative tests. The aim of the tests is to quantitatively answer the most important questions of calibration: "What are the relative repeatabilities, precisions and accuracies of each of the different calibration methods?" These tests provide an idea of what levels of calibration errors to expect in the reconstructed data sets. The mathematical framework allows questions on parameter identification, numerical conditioning and the effects of measurement errors to be answered. In this way, the differences in performance of the calibration methods can be explained using the theoretical framework.

Chapter 3 tackles the problem of reconstruction. The main goal is to prove the feasibility of using smooth functions to interpolate the set of irregularly spaced B-scans into a regular 3D voxel array. A spline-based method is adapted to the interpolation problem, and improvements are made to deal with the very large numbers of clustered data points encountered with 3D freehand ultrasound data. Chapter 3 also summarises the traditional reconstruction methods found in the literature. An important contribution is the evaluation of the advantages and disadvantages of the various methods (including the spline method) in a series of quantitative tests. The main questions are "How do the methods deal with gaps and overlapping data?" and "Are any reconstruction artifacts generated?" In other words, can spline interpolation improve on standard methods and produce the highest quality reconstructions?

1.7.2 Spatial Compounding

Spatial compounding is investigated as a method for realising further improvements in the quality of the 3D data sets. Although 2D spatial compounding has been the subject of considerable research in the past, it is not employed on any widely available commercial ultrasound system. Since the emphasis of modern ultrasound machines is on real-time B-scan imaging, there is little time for the calculations required for spatial compounding. Only temporal compounding is common, since it is easy to implement without taking much away from the real-time abilities. But, as mentioned, temporal compounding is a non-optimal way of reducing speckle, since the B-scans from non-coincident planes are averaged.

With 3D freehand ultrasound, it is easy to incorporate spatial compounding by simply examining a region of interest from a variety of look directions. This creates a dense set of overlapping B-scans. The position sensor readings, combined with calibration, ensure that the B-scans are compounded only where they overlap. Unfortunately, small errors in data alignment result in a loss of resolution (blurring). Yet 3D ultrasound brings new challenges, such as 3D visualisation and volume estimation, that demand a certain degree of speckle reduction. This suggests that the trade-off between resolution and speckle reduction in spatial compounding may be acceptable. Plus, because the reconstruction stage of freehand imaging is performed off-line, time is available to align the B-scans more accurately using the features in the images themselves. This process is called registration. The new challenges of 3D ultrasound, combined with advances in image registration and high speed computing, make 3D spatial compounding a promising topic of research.

Part II of this thesis performs an in-depth investigation into spatial compounding from the perspective of 3D freehand imaging. Since little work has been done on 3D spatial compounding, Chapter 4 starts with a statistical study. The statistical theory for the improvement in SNR is developed, and a series of tests are performed *in vitro* to verify the theory. This work is the first to investigate whether 3D visualisation and volume estimation are improved by spatial compounding. Another important contribution is the development of a landmark-based registration technique to correct some of the misalignment. As mentioned, the key to spatial compounding is accurate registration.

In Chapter 5, spatial compounding is extended to examinations performed *in vivo*, to verify its merits on human patients. This study is the first to apply a robust correlation-based registration technique, originally designed for other modalities such as MR and CT, to ultrasound data. The main question is “Can sufficiently accurate registration be achieved with data obtained *in vivo*?” In summary, the main contribution of Part II is the proof-of-concept of 3D spatial compounding.

Part I

Reconstruction

The work which deserves, but I am afraid does not always receive,
the most credit is that in which discovery and explanation go
hand in hand, in which not only are new facts presented,
but their relation to old ones pointed out.

—Lord Rayleigh, **British Association Report**, 1884

Chapter 2

Calibration

2.1 Introduction

A key requirement of all freehand imaging systems is calibration. This involves determining the position and orientation of the B-scan plane with respect to the sensor mounted on the probe. The results of the calibration take the form of six constant offsets, three for position and three for orientation. These offsets are combined with the sensor measurements to calculate the positions of the B-scans during reconstruction. Accurate calibration is essential for a consistent reconstruction that preserves true anatomical shape.

Calibration is required regardless of the type of position sensor used in the examination. A number of different position sensors have been used for freehand systems (see Section 1.5.4), but only a few articles in the literature mention calibration (Barry et al., 1997; Carr, 1996; Detmer et al., 1994; Hughes et al., 1996; Leotta et al., 1997; Nelson and Elvins, 1993; State et al., 1994). Even fewer have seriously investigated the accuracy and precision of calibration.

A rough estimate of the calibration parameters can be obtained by external measurements of the probe and position sensor. Yet the origin of the sensor (for example, the centre of the wire coils in an electromagnetic receiver) and the corner of the B-scan are not well defined with respect to the external cases of the sensor and probe. For these reasons, calibration is preferably performed by imaging a phantom: an artificial object with some known physical properties or dimensions. Measurements from an examination of the phantom, combined with its known physical properties, can be used to determine the six offsets.

In this chapter, all calibration phantoms mentioned in the literature are described using a common mathematical framework. The precision, accuracy and repeatability of each of these techniques are then compared in a series of tests.

2.2 Coordinate Systems

To see how calibration fits into the 3D freehand process, it is necessary to describe the reconstruction equations. Figure 2.1 shows the four coordinate systems used for recon-

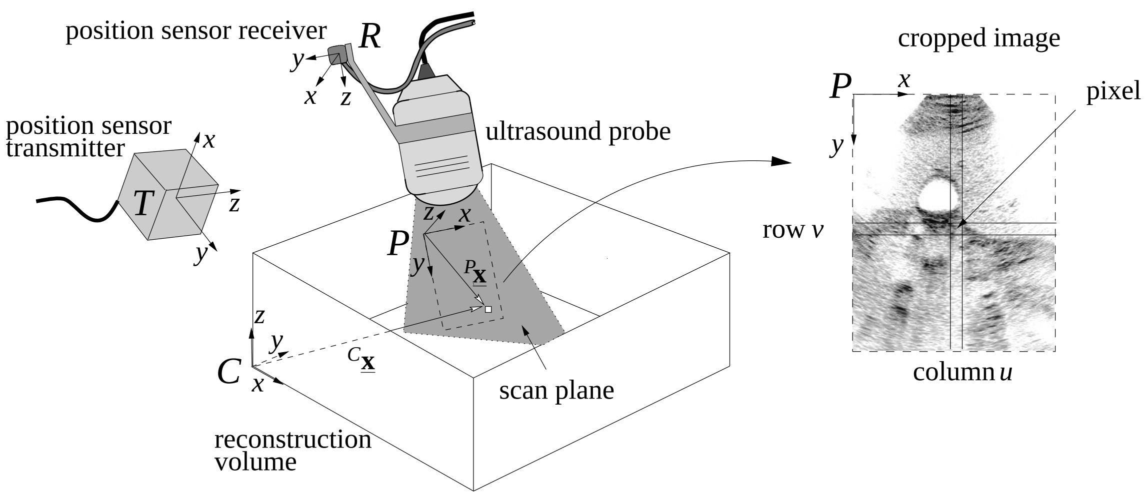


Figure 2.1: **Coordinate systems.** The reconstruction process uses four coordinate systems. The transmitter is often placed on a cantilevered support above the examination table, to give an unobstructed path to the receiver.

struction. P is the coordinate system of the cropped B-scan $\mathbf{P}$, with an origin in the top left hand corner. The y-axis is in the axial direction, and the x-axis in the lateral direction. The z-axis is in the elevational direction, out of the plane of the B-scan. R is the coordinate system of the position sensor's receiver and T is the coordinate system of the transmitter. The reconstruction volume, created from the set of acquired B-scans, takes the form of a 3D voxel array $\mathbf{C}$. The origin of coordinate system C is located at the corner of the reconstruction volume.

During reconstruction, every pixel in every B-scan has to be located with respect to the reconstruction volume. Each pixel's location (${}^{P}\underline{\mathbf{x}}$) is transformed first to the coordinate system of the receiver R, then to the transmitter T and, finally, to the reconstruction volume C. The overall transformation can be expressed as the multiplication of homogeneous transformation matrices:

$$ {}^{C}\underline{\mathbf{x}} = {}^{C}\mathbf{T}_{T} \ {}^{T}\mathbf{T}_{R} \ {}^{R}\mathbf{T}_{P} \ {}^{P}\underline{\mathbf{x}} \tag{2.1} $$

where

$$ {}^{P}\underline{\mathbf{x}} = \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix}. $$

A standard notation is adopted, so that ${}^{J}\mathbf{T}_{I}$ is the transformation from coordinate system I to coordinate system J. u and v are the column and row indices of the pixel in the cropped image, and s_x and s_y are scale factors with units of mm/pixel. ${}^{C}\underline{\mathbf{x}}$ is the pixel's location in the coordinate system C.

A transformation between two three-dimensional coordinate systems has six degrees of freedom: three rotational (α, β, γ) and three translational (x, y, z). There are many ways to parameterise a rotation, including unit quaternions, rotation matrices and a variety of three-angle systems. An x-y-z fixed-angles scheme (Craig, 1989) is chosen here. The rotation between two coordinate systems is effected by first rotating through γ around the x-axis, then through β around the y-axis, and finally through α around the z-axis (the order of rotation is significant). The fixed rotation axes are aligned with the first coordinate system. Using this convention, the homogeneous matrix describing the transformation takes the following form:

$$
{}^{J}\mathbf{T}_{I}(x, y, z, \alpha, \beta, \gamma) =
\begin{pmatrix}
\cos\alpha\cos\beta & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma & x \\
\sin\alpha\cos\beta & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma & y \\
-\sin\beta & \cos\beta\sin\gamma & \cos\beta\cos\gamma & z \\
0 & 0 & 0 & 1
\end{pmatrix}. \quad (2.2)
$$

Each of the transformation matrices in equation (2.1) plays a different role in reconstruction. The most straightforward is ${}^{T}\mathbf{T}_{R}$, which is derived directly from the position sensor readings. In the context of reconstruction, ${}^{C}\mathbf{T}_{T}$ is included largely as a matter of convenience. If omitted, and the reconstruction volume is aligned with the transmitter, this might place the B-scans a long way from the origin of C, resulting in a vast, largely empty voxel array. In addition, the anatomy might appear upside down in C, with the head towards the bottom of C and the feet towards the top: it all depends on the positioning of the transmitter with respect to the patient. Such effects can be corrected by allowing an arbitrary transformation ${}^{C}\mathbf{T}_{T}$, set as required for each examination. That leaves just ${}^{R}\mathbf{T}_{P}$, which needs to be determined by calibration. Although the scale factors s_x and s_y could be derived from the axis markings on the B-scan, these are also included in the calibration for convenience.

Once calibration is performed, equation (2.1) can be used to calculate ${}^{C}\underline{\mathbf{x}}$ for every pixel. Then the voxels of $\mathbf{C}$ can be set according to the intensities of the pixels they intersect. A basic algorithm simply runs through each pixel in every B-scan and fills the nearest voxel with the value of that pixel. Multiple contributions to the same voxel are averaged. This algorithm, called the PNN (pixel nearest neighbour) algorithm, is depicted in Figure 2.2. Other reconstruction algorithms are introduced in Chapter 3.

2.3 Calibration Phantoms

2.3.1 Cross-wire Phantom

Calibration is performed by scanning a phantom of known geometric dimensions. Equations similar in form to equation (2.1) can be written using knowledge of the phantom geometry and the position sensor measurements. These equations are then solved to determine the calibration parameters.

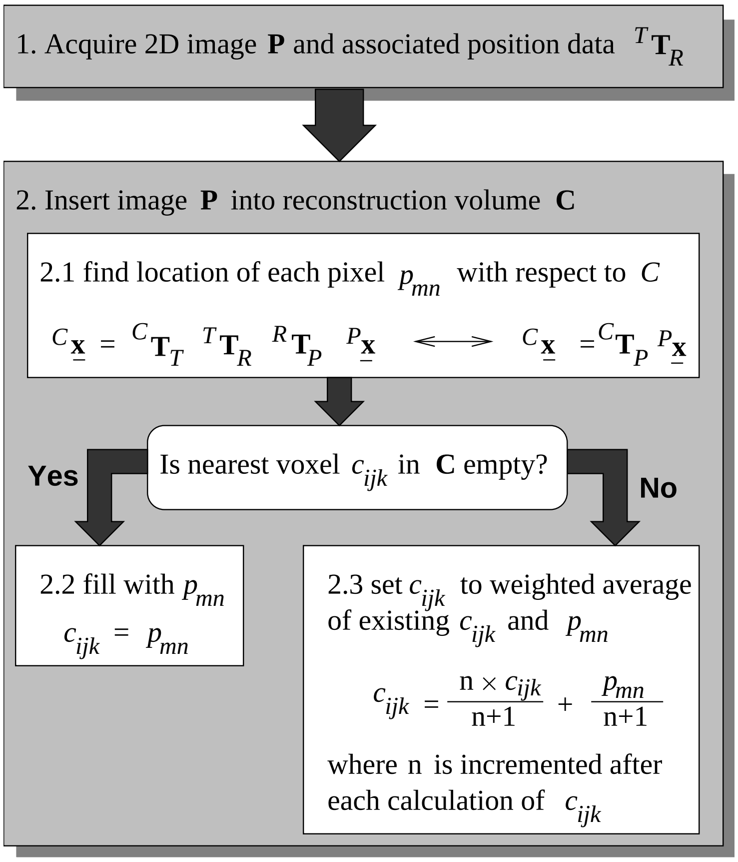
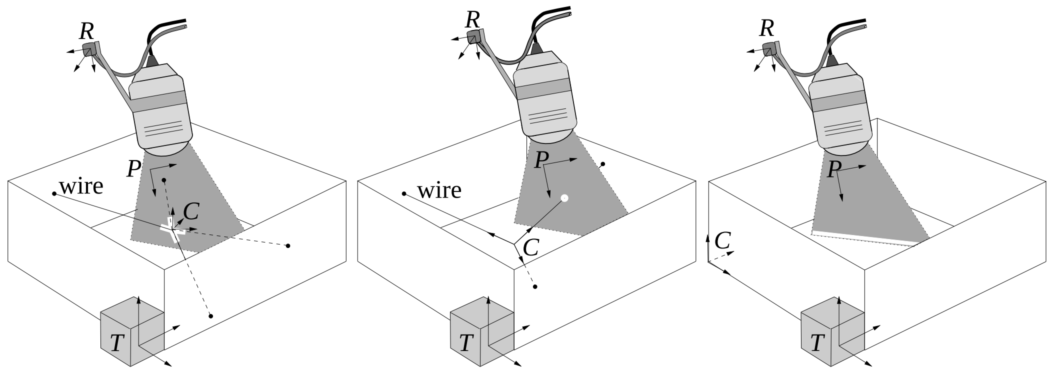


Figure 2.2: **PNN reconstruction algorithm.** Steps 1 and 2 are repeated for all B-scans acquired in an examination. The accumulator n in step 2.3, used to average overlapping data, is also stored at each voxel and set initially to zero. In practice, step 2.2 is omitted since it is equivalent to step 2.3 for n=0.



(a) Cross-wire phantom (b) Three-wire phantom (c) Single-wall phantom

Figure 2.3: **Calibration phantoms.** (a) A crossed wire is scanned in a water bath, producing a white cross in the B-scan. The crossed wires can be scanned from a variety of directions, but the scan plane must intersect the crossing point. The origin of C is placed at the point where the wires cross. (b) Three mutually orthogonal, intersecting wires are mounted in a box. B-scans of each wire exhibit a white dot. The origin of C is located at the point where the wires cross, with the axes running along the three wires. (c) The floor of the water bath is scanned, producing a white line in the B-scan. The origin of C is located on the floor of the water bath, with one axis orthogonal to the plane of the floor.

A common calibration phantom is the cross-wire phantom (Barry et al., 1997; Detmer et al., 1994). Two intersecting wires are mounted in a water bath, with the transmitter placed at some fixed location with respect to the wires — see Figure 2.3(a). For calibration purposes, C is not coincident with the corner of the reconstruction volume, but is placed instead at the intersection of the wires. This simplifies the calibration equations considerably.

The location where the wires cross is scanned repeatedly from different directions, with each B-scan showing a detectable cross (the reader may wish to skip ahead to Figure 2.8(a) for a typical example). The pixel at the centre of the cross should satisfy:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = {}^C\mathbf{T}_T \, {}^T\mathbf{T}_R \, {}^R\mathbf{T}_P \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix}. \tag{2.3}$$

The first three rows of equation (2.3) give three equations involving the measurements ${}^T\mathbf{T}_R$, u and v, and the unknowns ${}^R\mathbf{T}_P$, ${}^C\mathbf{T}_T$, s_x and s_y. If there are m B-scans, then the equations can be stacked together to produce a system of non-linear homogeneous equations of size $3m$:

$$\mathbf{0} = \mathbf{f}(\theta, \Phi) \tag{2.4}$$

where θ are the measurements and Φ are the unknowns. Given an over-determined set of equations (such that $3m$ is greater than the number of unknowns Φ), the system can be solved by several iterative methods. The popular and robust Levenberg-Marquardt algorithm (More, 1977) is used here. At iteration j, the update $\Delta\Phi$ to the current estimate Φ_j is derived from the first order Taylor expansion of equation (2.4):

$$\begin{aligned} \mathbf{0} = \mathbf{f}(\theta, \Phi) &\approx \mathbf{f}(\theta, \Phi_j) + \frac{\partial \mathbf{f}(\theta, \Phi_j)}{\partial \Phi}(\Phi - \Phi_j) \\ \Rightarrow \Delta \mathbf{f} &= \mathbf{J}(\Phi - \Phi_j) = \mathbf{J}\Delta\Phi \end{aligned} \tag{2.5}$$

where $\Delta\mathbf{f}$ is the error vector $-\mathbf{f}(\theta, \Phi_j)$ and $\mathbf{J}$ is the gradient matrix $\partial\mathbf{f}(\theta, \Phi_j)/\partial\Phi$, also known as the Jacobian. The Levenberg-Marquardt algorithm produces the updated parameter values Φ_{j+1}:

$$\Phi_{j+1} = \Phi_j + (\mathbf{J}^{\mathrm{T}}\mathbf{J} + \varepsilon\mathbf{I})^{-1}\mathbf{J}^{\mathrm{T}}\Delta\mathbf{f} \tag{2.6}$$

where ε is a damping term chosen at each step to stabilise convergence, and $\mathbf{I}$ is the identity matrix. This method becomes the standard iterative least squares algorithm for $\varepsilon = 0$. At each step, $\Delta\mathbf{f}$ and $\mathbf{J}$ are evaluated at the current estimate Φ_j. This process is iterated until the corrections $\Delta\Phi$ are sufficiently small.

Only the scale values and ${}^{R}\mathbf{T}_{P}$ are of interest in calibration, but ${}^{C}\mathbf{T}_{T}$ is also unknown. The parameters of ${}^{C}\mathbf{T}_{T}$ must also be solved for, even though these values will be subsequently discarded and a convenient ${}^{C}\mathbf{T}_{T}$ adopted for reconstruction (as explained in Section 2.2). At first sight, therefore, it appears that Φ must be a 14-element vector, composed of s_x and s_y, the six parameters of ${}^{R}\mathbf{T}_{P}$ and the six parameters of ${}^{C}\mathbf{T}_{T}$. Yet it is clear from inspection of the geometry that the coordinate system C can be at any orientation and still satisfy equation (2.3). This means that the three orientation angles of ${}^{C}\mathbf{T}_{T}$ are not identifiable [1]. These angles are therefore removed from Φ, and, for convenience, set to zero in equation (2.4). Table 2.1 lists the identifiable parameters Φ.

The final requirement for identifiability of the parameters is sufficient range of motion of the probe about the crossed wires. All six DOF must be exercised to produce a set of independent equations. This means the crossed wires must be imaged from a variety of different viewpoints.

Given a sufficiently wide range of viewpoints, accuracy of calibration with a point object depends on how well the centre of the point can be located, as well as the degree to which the point remains fixed in space with respect to the transmitter T. This means care must be taken so that both the transmitter and phantom are rigidly mounted and are not susceptible to mechanical vibrations. Given the quality of ultrasound images, it is common practice to locate the point by hand in each B-scan, making the calibration process fairly time-consuming.

Several papers (Leotta et al., 1995; Leotta et al., 1997; State et al., 1994) have described this calibration method using other point objects (such as a pinhead or a small suspended bead) instead of the crossed wires. The calibration equations, however, remain unchanged.

[1] The unidentifiable parameters can also be discovered by singular value decomposition of the Jacobian matrix — see Appendix B

	Φ	Cross-wire	Three-wire	Single-wall
	s_x	$\surd$	$\surd$	$\surd$
	s_y	$\surd$	$\surd$	$\surd$
	x	$\surd$	$\surd$	$\surd$
	y	$\surd$	$\surd$	$\surd$
${}^R\mathbf{T}_P$	z	$\surd$	$\surd$	$\surd$
	α	$\surd$	$\surd$	$\surd$
	β	$\surd$	$\surd$	$\surd$
	γ	$\surd$	$\surd$	$\surd$
	x	$\surd$	$\surd$	–
	y	$\surd$	$\surd$	–
${}^C\mathbf{T}_T$	z	$\surd$	$\surd$	$\surd$
	α	–	$\surd$	–
	β	–	$\surd$	$\surd$
	γ	–	$\surd$	$\surd$

Table 2.1: **Identifiable parameters.** Some parameters of ${}^C\mathbf{T}_T$, marked '–', are not identifiable by the cross-wire and single-wall methods, so are removed from the vector of parameters Φ.

2.3.2 Three-wire Phantom

Another calibration technique uses a three-wire phantom (Carr, 1996), as shown in Figure 2.3(b). This technique also involves scanning wires in a water bath, but it requires three wires to be accurately mounted in orthogonal directions. The accuracy of calibration depends on the orthogonality and straightness of the wires, as well as the degree to which the wires remain fixed in space with respect to the transmitter.

The coordinate system C is placed at the origin of the wires, and the x, y and z axes oriented along the three wires. Each wire is scanned, one at a time, along its length from a variety of look directions. The wire appears as a detectable dot in the B-scan (for a typical example see Figure 2.8(b)). For the wire along the x-axis, the pixel at the centre of the wire should satisfy:

$$\begin{pmatrix} x \\ 0 \\ 0 \\ 1 \end{pmatrix} = {}^C\mathbf{T}_T \; {}^T\mathbf{T}_R \; {}^R\mathbf{T}_P \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix}. \tag{2.7}$$

The left hand side of the equation becomes $(0\ y\ 0\ 1)^{\mathrm{T}}$ for the y-axis and $(0\ 0\ z\ 1)^{\mathrm{T}}$ for the z-axis. The two zero components of each equation give two equations in the unknowns Φ. All six parameters of ${}^C\mathbf{T}_T$ are identifiable, giving 14 unknown parameters — see Table 2.1. As with the cross-wire phantom, many B-scans are acquired from a variety of viewpoints, ensuring all 6 DOF are exercised. The wires are located by hand in each image, and the over-determined set of $2m$ equations is solved using the Levenberg-Marquardt algorithm.

The advantage of the three-wire method over the cross-wire method is that it is intrinsically easier to scan a length of wire than to keep the B-scan centred about a fixed point. Yet each wire must be scanned separately, keeping track of which wire each B-scan intersects. This means calibration times are almost as long as with the cross-wire technique.

2.3.3 Single-wall Phantom

In the process of designing a new 3D freehand ultrasound system, researchers at the University of Cambridge (Prager et al., 1998a) started to investigate alternative methods of calibration. Their aim was to find a method that was less laborious than the wire-based techniques, so that calibration could be easily performed by a non-specialist. In this way, a conventional ultrasound machine, with almost any type of probe, could be quickly converted into a 3D freehand system.

When scanning the cross-wire and three-wire phantoms, they found that the floor of the phantom produced clear, consistent lines in the B-scans. The wires were removed and calibration was performed using the floor of the water bath alone — see Figure 2.3(c).

Calibration in this way has several advantages. Simple water baths with planar surfaces are readily available, so there is no need to construct a special phantom. Furthermore, the plane should show up as a strong, straight line in the image (Figure 2.8(c) shows a typical example) making it easier to detect than dots and crosses. This is because there is much more redundant evidence for the presence of a line in an image. In other words, it is possible to locate a straight line even when portions of the line are corrupted or missing. The same cannot be said of dots and small crosses. Making the most of this advantage, an automatic line detection algorithm was developed (Prager et al., 1998b).

The equations for single-wall calibration can be expressed in the same form used for the cross-wire and three-wire techniques. If the coordinate system C is defined to lie in the floor of the water bath, with the z-axis orthogonal to the floor, pixels lying on the line in the B-scan should satisfy:

$$\begin{pmatrix} x \\ y \\ 0 \\ 1 \end{pmatrix} = {}^{C}\mathbf{T}_{T}\,{}^{T}\mathbf{T}_{R}\,{}^{R}\mathbf{T}_{P} \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix}. \tag{2.8}$$

The zero component of the equation gives one equation in the unknown parameters $\mathbf{\Phi}$. Yet the equation can be written for two pixels on the line (two points uniquely define a line), giving two equations per B-scan.

From inspection of the geometry, it is evident that several parameters of ${}^{C}\mathbf{T}_{T}$ are not identifiable. The rotation of the floor about the z-axis (α) and the translation of the floor in its own plane (x, y) do not affect the z-component of ${}^{C}\underline{\mathbf{x}}$. This leaves 11 identifiable parameters — see Table 2.1.

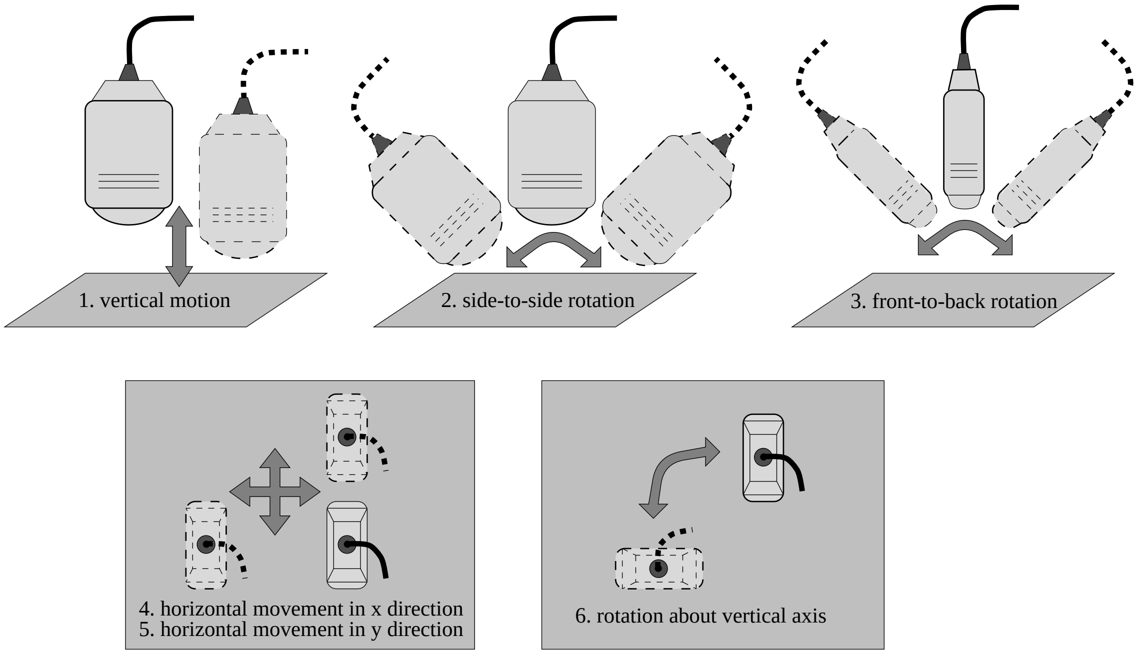


Figure 2.4: **Minimal sequence of motions for single-wall calibration.** All six degrees of freedom must be exercised to identify the 11 parameters.

The optimisation algorithm needs at least 11 independent equations to solve for these 11 unknowns. Although each B-scan can be used to write two equations, the resulting equations will not necessarily be independent unless the probe is moved appropriately while scanning the wall. As with the other methods, the motion of the probe must exercise all six DOF — see Figure 2.4. The over-determined set of $2m$ equations is then solved using again the Levenberg-Marquardt algorithm. The accuracy of the calibration depends on the flatness of the floor, as well as the degree to which the floor remains fixed in space with respect to the transmitter.

2.3.4 Cambridge Phantom

Unfortunately, single-wall calibration is susceptible to problems caused by the width of the ultrasound beam and the nature of specular reflection. When the beam is not normal to the wall, the first echo to return to the probe comes from the edge of the beam closest to the wall — see Figure 2.5. This effect introduces position errors into those B-scans taken at oblique angles. The problem is compounded by the weak echo obtained when the wall is scanned obliquely, since much of the ultrasound energy is specularly reflected away from the probe and the amount of Lambertian reflection is small. Both of these effects conspire to spoil the ultrasound image when the scan plane is not normal to the wall. Unfortunately, such scanning angles are required for effective calibration, as discussed previously.

A new phantom, dubbed the "Cambridge phantom" (Prager, 1997), was designed to

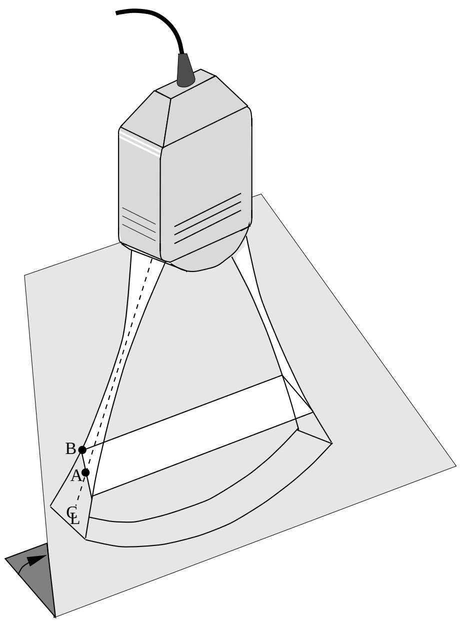


Figure 2.5: **Beam thickness problem with single-wall calibration.** When the wall is scanned from oblique angles, the finite beam thickness results in a thick, blurred line in the image. The thick line arises because point B is encountered by the ultrasound pulse before point A on the centre line. The echo from point B corresponds to the top of the line in the ultrasound image, which does not reflect the true position of the wall.

overcome the difficulties experienced with planar calibration. The phantom consists of two parts: a clamp that fits around the probe, and a thin brass bar mounted between two circular disks — see Figures 2.6 and 2.7. The top of the bar, which is approximately 1.5 mm thick, is squared-off and slightly roughened with emery cloth.

The idea is that the clamp constrains the thin bar to move only in the centre of the ultrasound beam. The bar is attached to the disks in such a way that the upper edge of the bar is aligned with the centre of each disk. This means that as the disks are rolled from side to side, the upper edge of the bar stays at a constant height above the floor. In essence, the Cambridge phantom is similar to the single-wall phantom, but the wall is a virtual plane traced out by the top of the bar. This plane has two important properties. First, only the slice required to reflect the centre of the ultrasound beam exists at each moment in time. Secondly, that slice is usually oriented towards the ultrasound probe, so the beam is reflected straight back to the probe, producing a strong, clear image — see Figure 2.8(d). The exception is when the clamp is rotated about the bar (motion 2 in Figure 2.7). Even with this motion, however, by roughening the surface of the bar slightly, clear images can be obtained for a sufficiently wide range of angles.

The calibration procedure using the Cambridge phantom is summarised as follows:

1. Place the clamp around the ultrasound probe. Adjust it so that the slots through the two sides of the clamp are aligned with the scan plane of the probe. This is

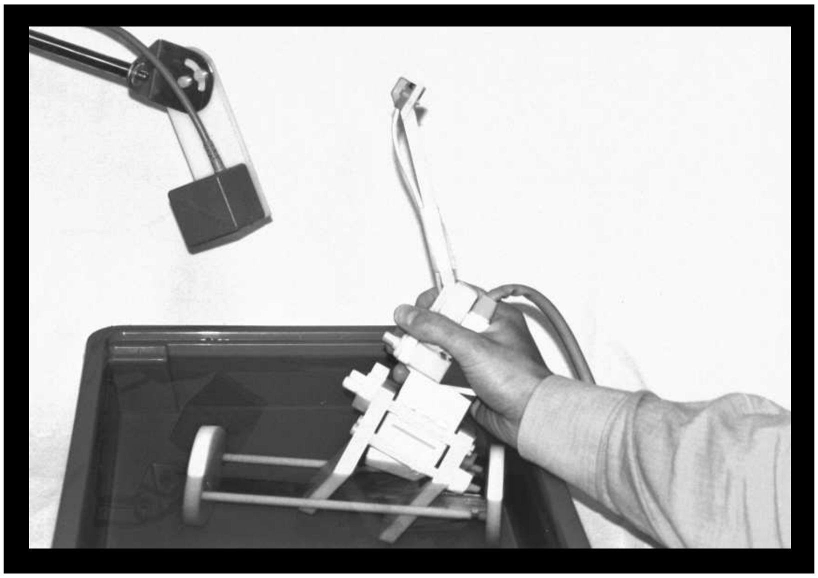


Figure 2.6: **Calibration using the Cambridge phantom.** A 7 MHz linear array probe is fitted into the phantom's clamp, which itself is fitted around the brass bar immersed in the water bath. The position sensor's transmitter is suspended from a cantilevered support above the water bath, and the receiver is attached to the probe at the end of a rigid extension bar.

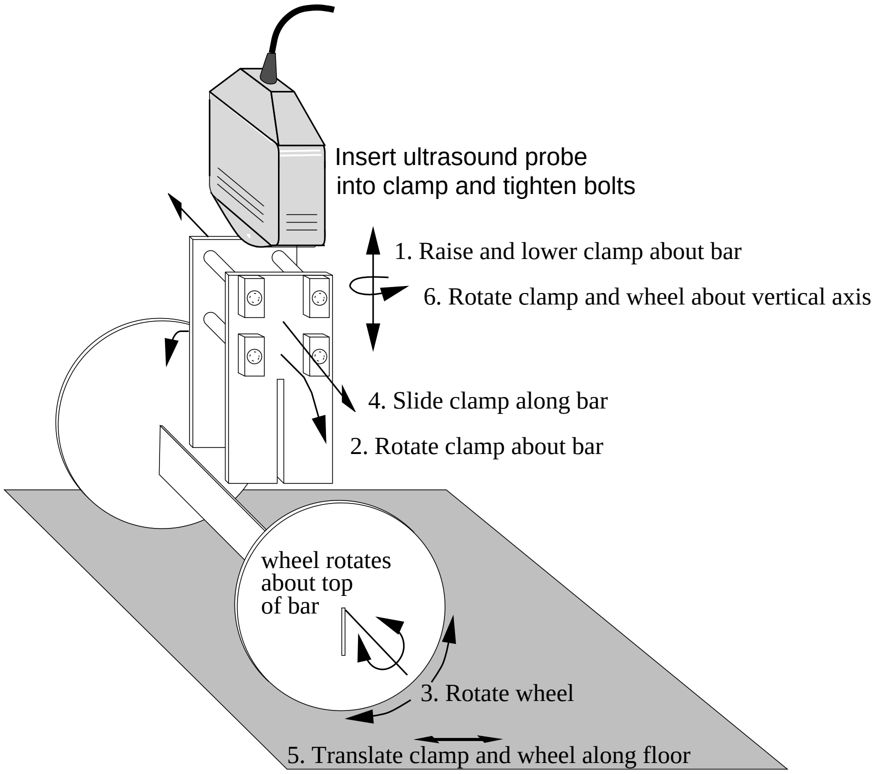


Figure 2.7: **Motions of the Cambridge phantom.** Accurate calibration requires a minimal sequence of motions covering all six degrees of freedom: compare with Figure 2.4.

easily checked by ensuring that a clear image is obtained when the clamp is placed over the bar.

2. Immerse the phantom in water and slot the clamp over the bar. Scan the bar from all possible angles, subject to the constraints imposed by the clamp. Figure 2.4 shows the minimum sequence of motions required for calibration. The shape of the phantom and the way it interacts with the clamp ensures that a clear image of the bar will always be visible in the B-scan.

3. Detect the lines in the image.

4. Solve the same system of equations derived for the single-wall phantom — see equation (2.8).

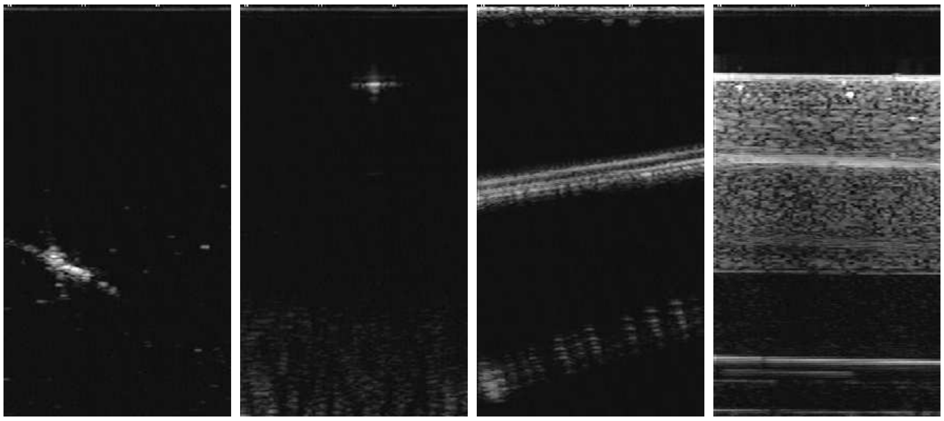
2.4 Comparison of Calibration Techniques

2.4.1 Method

A series of tests was undertaken to evaluate the performance of each of the four calibration methods under typical scanning conditions. Phantoms for three of the calibration methods (three-wire, single-wall, and Cambridge) were obtained directly from the inventors of each method. The cross-wire phantom was constructed by crossing two cotton wires and immersing them in a water bath, as described in (Detmer et al., 1994). The wires were crossed so that they did not form a plane; this ensured crosses were produced in images from all look directions. The three-wire phantom was constructed by placing 1 mm diameter nylon wires under tension between precision drilled holes in the walls of a perspex water bath, as described in (Carr, 1996). The holes were positioned to make the wires orthogonal. The single-wall phantom was no more than the floor of a water bath. The weight of the water, acting on a flat table under the floor of the bath, helped to reduce any error arising from the flexibility of the floor. Construction of the Cambridge phantom is described in Section 2.3.4.

The ultrasound machine used in these experiments was a Toshiba model SSA-270A/HG (Toshiba America Medical Systems, Tustin, California). The Stradx acquisition software (Prager et al., 1998a) was used in conjunction with an 8-bit frame grabber and a Silicon Graphics Indy workstation (Silicon Graphics Incorporated, Mountain View, California). Although the acquisition software can record at 25 fps, a lower frame rate of 5 fps was selected for the calibration experiments to allow significant motion of the probe between frames.

Calibration experiments were performed with a 7 MHz linear array probe, a 40 mm depth range and the imaging controls (such as time-gain compensation, focus, overall gain and acoustic power) adjusted to give optimal images of each phantom. The images were then cropped to 328×409 (width$\times$depth) pixels. Figure 2.8 shows some typical B-scans of the phantoms.



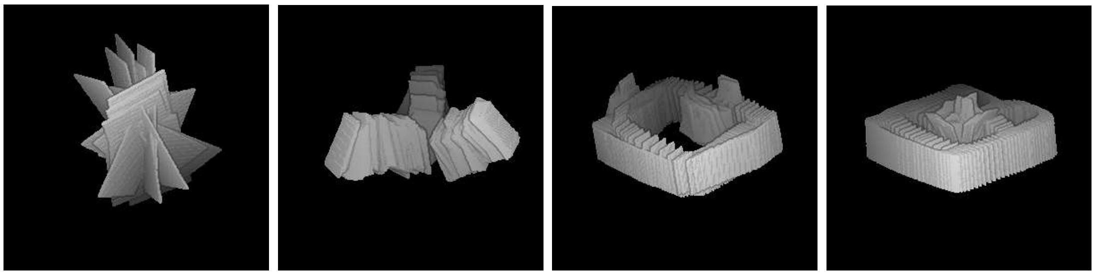
(a) Cross-wire (b) Three-wire (c) Single-wall (d) Cambridge

Figure 2.8: **Typical B-scans of calibration phantoms.** (a) The cross from the two wires is located near the bottom left hand corner of the B-scan. (b) The dot near the top of the B-scan is one of the wires of the three-wire phantom. (c) The top line is the floor of the single-wall phantom, the second one down is a reverberation. (d) The top of the brass bar is clearly visible.

A Polhemus Fastrak position sensor was used for all tests. To minimise interference from the electronics inside the ultrasound probe, the receiver was mounted at the end of a rigid nylon extension bar about 25 cm from the probe face. The mounting was not disturbed during the experiments, so the results of successive calibrations can be directly compared.

Twenty calibration experiments were performed using each of the cross-wire, single-wall and Cambridge phantoms. Practical time constraints limited the evaluation of the three-wire phantom to ten trials. The scanning protocols are described in Figure 2.9 and Table 2.2. Any B-scans that did not reveal a clear image of the phantom were discarded. For the Cambridge phantom, the probe was completely removed from the clamp between each trial, and then re-mounted. For the cross-wire and three-wire phantoms, the pixel coordinates (u, v) needed for the calibration equations were extracted by hand in each B-scan. The pixel coordinates for the single-wall and Cambridge phantoms were determined with an automatic line detection algorithm (Prager et al., 1998b). In all experiments, the Levenberg-Marquardt algorithm (equation (2.6)) was used to determine Φ, taking care to remove unidentifiable parameters and scale the remaining parameters to produce as well-conditioned a problem as possible. Details of the numerical methods are provided in Appendix B.

For all calibration experiments, the water bath was filled with water at room temperature. This means that the estimated scale factors s_x and s_y will reflect the speed of sound in water (1.48 mm/μs at 20° Celsius (Hill, 1986)) and not human tissue (1.54 mm/μs: the average for soft tissue (see Section 1.5.2)). The temperature in the room was ther-

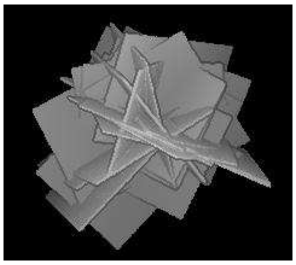


(a) Cross-wire (b) Three-wire (c) Single-wall (d) Cambridge

Figure 2.9: **Scanning protocols.** The outlines of the B-scans are shown for a typical examination of each phantom. Table 2.2 provides details of the range of angles and positions covered by each examination.

	Cross-wire	Three-wire	Single-wall	Cambridge
number of B-scans	~ 53	~ 193	~ 578	~ 530
x translation (mm)	± 20	± 84	± 67	± 108
y translation (mm)	± 20	± 84	± 67	± 71
z translation (mm)	± 12	± 44	± 17	± 17
max. inclination to z-axis	65°	50°	19°	69°
twist of B-scan	$\pm 87^\circ$	$\pm 90^\circ$	$\pm 90^\circ$	$\pm 90^\circ$

Table 2.2: **Scanning protocols.** The table gives approximate values for the ranges of angles and positions covered by the examinations in Figure 2.9. The translations relate to the centre of the probe face, with x and y axes oriented along the floor of the water bath, and the z-axis pointing upwards. These are the same directions as the axes of coordinate system C in Figure 2.1. The angle of inclination to the z-axis defines a cone of look directions, and the twist angle describes rotation of the B-scan around its centre line. Relatively large numbers of B-scans were acquired for the Cambridge and single-wall phantoms, since the images they produce are amenable to automatic detection.



x translation	± 18 mm
y translation	± 20 mm
z translation	± 7 mm
max. inclination to z-axis	67°
twist of B-scan	$\pm 87^\circ$

Figure 2.10: **Scanning pattern used for reconstruction precision experiments.** With this motion, the cross was observed within the full depth range of the B-scan (40 mm). See the caption of Table 2.2 for an explanation of the motion parameters.

mostatically controlled, so the water temperature was stable. Any variations observed in the calibration parameters cannot, therefore, be attributed to a change in the prevailing speed of sound. Of course, calibration before an examination performed *in vivo* should use a fluid in which the speed of sound matches that observed in human tissue.

The **repeatability** of the calibration solutions was assessed by looking at their stability across the multiple trials of each phantom. To this end, the change in the position (relative to the position sensor's receiver) of the bottom right hand corner (u_{max}, v_{max}) of the B-scan was determined under two different calibrations:

$$\Delta^R\underline{\mathbf{x}} = {}^R\mathbf{T}_P{}^{trial\ 1}\ {}^P\underline{\mathbf{x}}^{trial\ 1} - {}^R\mathbf{T}_P{}^{trial\ 2}\ {}^P\underline{\mathbf{x}}^{trial\ 2} \tag{2.9}$$

where

$${}^P\underline{\mathbf{x}}^{trial\ i} = \begin{pmatrix} s_x^{trial\ i} \times u_{max} \\ s_y^{trial\ i} \times v_{max} \\ 0 \\ 1 \end{pmatrix}.$$

This measure involves all six parameters of ${}^R\mathbf{T}_P$ and the two scale values, so $\left|\Delta^R\underline{\mathbf{x}}\right|$ indicates the repeatability with which points in the scan plane can be defined relative to the position sensor's receiver. It is a direct measure of calibration repeatability that does not depend on other sources of error associated with reconstruction. For the Cambridge, cross-wire and single-wall phantoms, 190 values of $\left|\Delta^R\underline{\mathbf{x}}\right|$ were calculated (using all possible pairs drawn from the twenty trials). For the three-wire phantom, 45 values of $\left|\Delta^R\underline{\mathbf{x}}\right|$ were calculated (using all possible pairs drawn from the ten trials).

The **reconstruction precision** was also assessed by scanning the cross-wire phantom immediately after completion of the calibration tests. The cross was observed from 49 different viewpoints, covering a range of angles and depths — see Figure 2.10. The location of the cross was labelled by hand in each B-scan, and then mapped to the reconstruction space using the calibration parameters and the position sensor readings. This was done for the multiple calibration solutions obtained using each of the four calibration phantoms, giving 70 reconstructions in total.

In each reconstruction, the cross appeared as a cloud of 49 points. The reconstruction precision was assessed by looking at the tightness of this cloud. To this end, the change

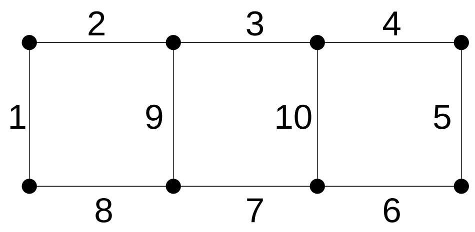
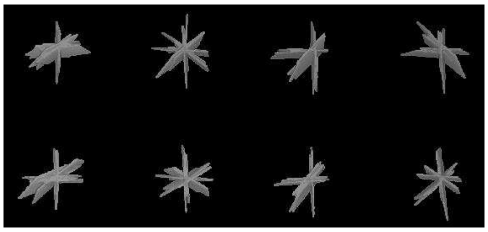


Figure 2.11: **Phantom used for assessing reconstruction accuracy.** Each of the ten inter-pin distances marked on the diagram is about 100 mm.



x translation	± 156 mm
y translation	± 57 mm
z translation	± 14 mm
max. inclination to z-axis	31°
twist of B-scan	$\pm 83^\circ$

Figure 2.12: **Scanning pattern used for reconstruction accuracy experiments.** With this motion, the pinheads were observed within the full depth range of the B-scan (40 mm). See the caption of Table 2.2 for an explanation of the motion parameters.

in the position (relative to the position sensor's transmitter) of the cross observed from two different viewpoints was measured:

$$\Delta^T\underline{\mathbf{x}} = {}^T\mathbf{T}_R{}^{view\ 1}\ {}^R\mathbf{T}_P\ {}^P\underline{\mathbf{x}}^{view\ 1} - {}^T\mathbf{T}_R{}^{view\ 2}\ {}^R\mathbf{T}_P\ {}^P\underline{\mathbf{x}}^{view\ 2} \tag{2.10}$$

where

$$ {}^P\underline{\mathbf{x}}^{view\ i} = \begin{pmatrix} s_x \times u^{view\ i} \\ s_y \times v^{view\ i} \\ 0 \\ 1 \end{pmatrix}.$$

So $\left|\Delta^T\underline{\mathbf{x}}\right|$ indicates the precision with which points can be located in the reconstruction volume. 1176 values of $\left|\Delta^T\underline{\mathbf{x}}\right|$ were calculated (using all possible pairs drawn from the 49 views). The process was repeated for each of the 70 reconstruction volumes.

The **reconstruction accuracy** was also assessed by scanning (immediately after completion of the calibration runs) the phantom shown in Figure 2.11. The phantom consists of eight dressmaker's pins arranged in a plane. Each pin was observed from eight different viewpoints, covering a range of angles and depths, as shown in Figure 2.12. The locations of the pins were labelled by hand in each B-scan, and then mapped to the reconstruction space using the calibration parameters and the position sensor readings. This was done for the multiple calibration solutions obtained using each of the four calibration phantoms, giving 70 reconstructions in total.

To assess the accuracy of the reconstruction, the ten inter-pin distances marked in Figure 2.11 were examined. Each of these distances was measured directly using sophisticated

	Cross-wire	Three-wire	Single-wall	Cambridge
no. trials	20	10	20	20
B-scans used	51/53	106/193	480/578	530/530
RMS error (mm)	0.56	1.04	0.48	0.34

Table 2.3: **Quality of calibration data.** The table lists the number of trials used in the comparative tests, and the proportion of B-scans deemed to be of acceptable quality for inclusion in the calibration. The RMS residuals of the set of equations (2.4) at the solution Φ are also listed. The proportion of acceptable B-scans and the residual errors represent the average across all trials of each phantom.

Calibration Technique	κ
Cross-wire	17.5
Three-wire	47.5
Single-wall	447.8
Cambridge	87.5

Table 2.4: **Condition number.** The condition number (κ) is listed for representative calibrations of each of the four calibration techniques.

metrology equipment to an accuracy in excess of one-hundredth of a millimetre. In each reconstruction, the ten inter-pin distances were estimated 64 times, using all combinations of the eight observations of each pin. The error in each of these measurements was used to assess the reconstruction accuracy. In total, 640 error measurements in each of the 70 reconstructions were obtained.

2.4.2 Results

Table 2.3 shows statistics indicative of the quality of the data being used to estimate the calibration parameters. For all four phantoms, a large proportion of the acquired B-scans revealed the calibration feature (wire or plane) clearly enough that the feature could be reliably labelled and used to generate some of the calibration equations (2.4). In general, the Cambridge phantom produced the highest quality images.

For all phantoms, the Levenberg-Marquardt algorithm (equation (2.6)) converged to a stable solution in around 15 iterations. No local minima were encountered: the algorithm repeatedly converged to a unique solution, independent of starting position (though note the discussion on "mirror" solutions in Appendix B). After the Levenberg-Marquardt algorithm had converged, the RMS (root mean square) residual of the system of equations (2.4) was evaluated at the solution Φ. The residuals are listed in Table 2.3. Assuming Φ is the globally optimal solution, the residual errors give an indication of the self-consistency of the calibration equations. Perfectly consistent equations should produce an RMS error of zero.

	Cross-wire	Three-wire	Single-wall	Cambridge
no. observations	190	45	190	190
mean (mm)	1.47	5.37	3.27	0.92
max (mm)	3.37	12.09	10.07	2.32

Table 2.5: **Calibration repeatability.** The quantity tabulated is $\left|\Delta^R \underline{\mathbf{x}}\right|$ for the bottom right hand corner of the B-scan, where $\left|\Delta^R \underline{\mathbf{x}}\right|$ is defined in equation (2.9). $\left|\Delta^R \underline{\mathbf{x}}\right|$ was calculated for each pair of calibration solutions drawn from the multiple trials of each phantom.

	Cross-wire	Three-wire	Single-wall	Cambridge
no. observations	23520	11760	23520	23520
mean (mm)	1.65	2.67	3.43	2.17
max (mm)	4.78	10.31	11.67	7.13

Table 2.6: **Reconstruction precision.** The quantity tabulated is $\left|\Delta^T \underline{\mathbf{x}}\right|$ for a point target (the cross in the cross-wire phantom), where $\left|\Delta^T \underline{\mathbf{x}}\right|$ is defined in equation (2.10). $\left|\Delta^T \underline{\mathbf{x}}\right|$ was calculated for each pair of views drawn from the 49 views of the cross, giving a total of 1176 values per reconstruction.

To determine the numerical conditioning of the four calibration methods, condition numbers were calculated for representative examinations — see Table 2.4. A low condition number corresponds to good identifiability of parameters, as described in Appendix B. The condition numbers were less than 100 for the Cambridge, cross-wire and three-wire phantoms, indicating good identifiability. Only the single-wall phantom produced a slightly ill-conditioned set of equations. This can be attributed to the limited range of scanning motions that resulted in clear images of the wall. In other words, it was not possible to observe the wall from a sufficiently large range of angles to produce a well-conditioned problem.

Table 2.5 shows the results of the calibration repeatability experiments described in Section 2.4.1. These results indicate to what extent repeated trials with the same phantom produce the same calibration parameters.

Table 2.6 shows the results of the reconstruction precision experiments described in Section 2.4.1. These results indicate how precisely the position of a point can be defined in a 3D reconstruction volume. They reflect the size of the cloud obtained when imaging a point from different directions, but say nothing about reconstruction accuracy (whether the cloud is in the right place). Table 2.7 shows the results of the reconstruction accuracy experiments with the pinhead phantom described in Section 2.4.1. These results indicate whether distances calculated between points in the reconstruction volume correspond to the true distances.

	Cross-wire	Three-wire	Single-wall	Cambridge
no. observations	12800	6400	12800	12800
mean (mm)	0.04	−0.15	0.14	0.23
std. dev. (mm)	1.12	2.18	1.63	1.33
min (mm)	−3.98	−8.10	−6.53	−5.62
max (mm)	4.31	7.83	5.72	4.46

Table 2.7: **Reconstruction accuracy.** The quantity tabulated is the error in each of the ten inter-pin distances shown in Figure 2.11. The true inter-pin distances were all approximately 100 mm. As explained in the text, 640 error measurements were obtained in each reconstruction.

2.4.3 Discussion

Performance

There are many ways to assess the performance of a calibration procedure. We might start by asking, "If I do another calibration tomorrow, shall I get the same result?" The figures in Table 2.5 help us answer this question. The calibration parameters define the position and orientation of the scan plane, with respect to the position sensor's receiver. These tests have measured how a supposedly fixed point on that plane (the bottom right hand corner of the B-scan) moves around under repeated calibrations with the same phantom. The Cambridge phantom appears to provide the most stable calibration parameters. Of course, we have no idea whether these calibration parameters are correct: they're just stable. Yet it is clear that stability is a prerequisite for a "perfect" calibration technique.

To decide whether the calibrations are correct, we might again ask a question: "Given a calibration, if I look at a point from different viewpoints, will it always map to precisely the same point in the reconstruction volume?" The results in Table 2.6 relate to this question. The Cambridge and cross-wire calibrations appear to produce the most precise reconstructions. If we look at a point from two significantly different directions, we can expect a distance of about 2 mm between the two reconstructed positions of the point. With reference to the results in Table 2.5, this suggests that the Cambridge calibrations are somewhat biased. They are tightly bunched together but not evenly spread around the optimal solution. The cross-wire calibrations are more variable but more equally disposed around the optimal solution.

We have yet to establish full confidence in the calibrations. We know that fixed points are reconstructed as, more or less, points, but we do not yet know whether these points are in the right place. So we ask one, final question: "If I scan something that is 100 mm long in the real world, will it come out as 100 mm long in the reconstruction?" The results in Table 2.7 help us answer this question. The cross-wire calibration is remarkably accurate in this respect. There is hardly any bias in the measured lengths, and the standard deviation of the measured distance is just greater than 1 mm: about 1% of the length being measured. The Cambridge phantom is also good: it has a slight tendency to produce over-estimates of length (by about 0.25 mm on average), but can still measure lengths to within about 1.5% of their true values.

The calibrations produced by the single-wall and three-wire phantoms are not as good. There is significant uncertainty in the calibration parameters (Table 2.5), which leads to relatively poor point reconstruction precision (Table 2.6). The length accuracy figures (Table 2.7) do not reflect these errors to the same extent, since the measurements are not sensitive to point reconstruction errors out of the plane of the pinheads.

In Section 2.3.4, two of the main sources of error in single-wall calibrations were identified. The first concerns identifiability of the calibration parameters. Because only a limited amount of energy is reflected back towards the probe, scanning can only be performed from a limited range of viewing angles. The limited range leads to relatively poor conditioning of the calibration problem. The poor reflection at oblique angles can be improved by roughening the floor of the water bath, thus producing better images of the plane from a wider range of angles. The other source of calibration error is the effect of the finite beam width. Simulations were performed to investigate the effects of beam width, and the results are presented in Appendix C. The simulations support the idea that the finite beam width is responsible for a large proportion of the errors. Unfortunately, roughening the floor of the water bath will do nothing to reduce beam width effects.

The performance of the three-wire phantom is disappointing. In Table 2.3 we see that the calibration equations are not as self-consistent as with the other techniques. This suggests inaccuracies in the phantom construction: the calibration equations are based on mutually orthogonal wires, and yet the wires are not, in reality, mutually orthogonal. The three-wire phantom was relatively difficult to construct, requiring great precision in the positioning of the wires.

Other researchers (Detmer et al., 1994; King et al., 1991; Leotta et al., 1995; Leotta et al., 1997) have used similar measures of reconstruction accuracy and precision, but these should be compared with caution to those reported here. Measured reconstruction performance depends critically on the range of scanning angles and positions, as well as the type of position sensor and the quality of the ultrasound equipment. The results presented in Tables 2.6 and 2.7 are meaningful only in the context of examinations using similar ultrasound equipment, position sensing devices, and examinations that cover a similar range of angles and positions to those described in Figures 2.10 and 2.12. In range-of-motion studies, it was observed that many common clinical examinations [2] fall within the envelope of motions described in Figures 2.10 and 2.12. Thus, Tables 2.6 and 2.7 provide an upper bound on the reconstruction accuracy and precision one can expect in such clinical examinations. It is an upper bound, since reconstructions of data obtained *in vivo* are also susceptible to a whole range of other errors, including motion artifacts.

[2] A number of routine clinical examinations of the left kidney, right kidney, liver, Achilles tendon, rotator cuff, aorta and thyroid gland were monitored.

Ease of use

Ease of use is another important measure by which a calibration technique should be assessed. It is rather difficult to quantify ergonomics, but calibration with the single-wall phantom is certainly the easiest and fastest. The probe can be moved about freely, and B-scans acquired continuously until a sufficiently wide range of motions are covered. The automatic detection of lines then helps to greatly reduce the time required to determine the calibration parameters.

The Cambridge phantom is equally easy to use, once the probe is inserted into the clamp. It is sometimes difficult to accurately align the scan plane with the bar of the phantom, so care must be taken to achieve good results. Since the alignment is only performed once before the calibration, the extra time required to get good alignment is relatively small, but experience is still required to get repeatable results. Once inserted, it is easy to adopt a particular series of movements to ensure all six DOF are exercised. In this way, the examinations can be performed rapidly, and with automatic line detection, the calibration results are readily obtained.

Unlike the Cambridge phantom, the cross-wire technique does not require the probe to be accurately mounted on the apparatus before the calibration runs, yet the crossed wires must be placed in the plane of each acquired B-scan. This requires the operator to simultaneously look at the crossed wires on the ultrasound monitor, hold the probe steady in the right location, and start and stop data acquisition. This process makes cross-wire calibration the most time consuming technique.

The three-wire phantom can be scanned more rapidly than the crossed-wire phantom, since the entire length of the wires can be imaged with continuous data acquisition. Yet each of the wires must be imaged one at a time, so the examination takes longer than the Cambridge or single-wall techniques. Some care must still be taken with alignment of the probe, since the best images (depicting a small circular dot) are obtained when the B-scan is perpendicular to the wire. This must be traded off against the requirement of exercising all six DOF.

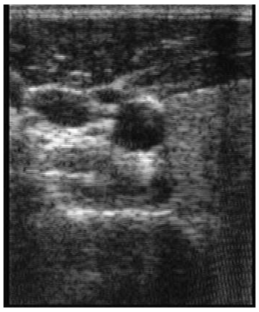
It may be possible to create software to assist in labelling images from the cross-wire and three-wire techniques to speed up calibration. Determination of which feature in the image corresponds to the wire (noise in the image occasionally obscures the wire) may be assisted by tracking features in a series of B-scans. Fully automatic detection, however, is not likely to be feasible.

To illustrate the total time required for calibration, reconstruction and visualisation, a representative 3D freehand examination was performed. Calibration was performed using the Cambridge phantom, a human carotid artery was examined by a physician, and any-plane slicing and surface rendering were used to display the results. The elapsed time between mounting the position sensor on the probe and viewing the results was 15 minutes. Figure 2.13 shows the results.

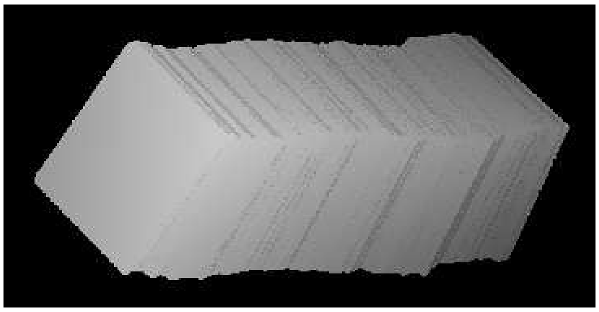
To achieve the best possible results, calibration was performed with suitably warm water (approximately 40° Celsius), so that the speed of sound matches that observed in human tissue. The data acquisition was not gated to an ECG, so the pulsative motion of the artery results in small ripples in the reconstruction. Other reconstruction errors were caused by variation in the contact pressure between the probe and the skin, position sensor errors, refraction, and the other errors listed in Section 1.5. These errors cannot be addressed by calibration. Despite these limitations, fairly clear and consistent 3D reconstructions are obtained.

2.5 Conclusions

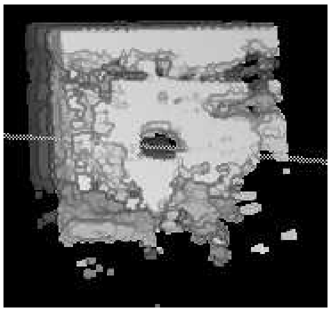
This chapter has addressed the issue of freehand 3D ultrasound calibration. The existing techniques were reviewed and their performance assessed under representative scanning conditions. The cross-wire technique produced impressive results, but took the longest to perform, typically several hours. The three-wire technique was slightly easier, but construction of the phantom required much greater precision, and its performance was worse. The single-wall phantom was the easiest to use and the simplest to construct. Unfortunately, the effects of the finite beam width and limitations on the range of motion severely degraded its performance. These two problems are alleviated by the Cambridge phantom, which is based on the same principle of using a single flat wall as the standard for calibration. Combined with automatic line detection, the Cambridge technique was easy, rapid and produced good results. The main disadvantages of the Cambridge technique are the need to precisely machine the apparatus and the time required to correctly clamp the ultrasound probe to it.



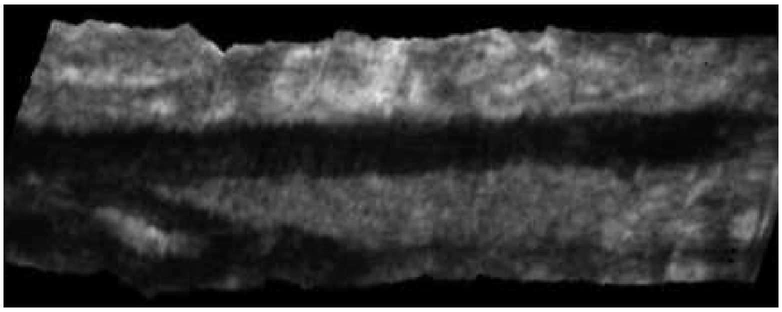
(a) B-scan



(b) Outlines of B-scans



(c) Surface rendering



(d) Cross-sectional slice

Figure 2.13: **3D reconstruction of the carotid artery.** Reconstruction is performed with the PNN algorithm after calibration with the Cambridge phantom. In (a), a B-scan of the carotid artery is shown. The artery is the dark circle just above the centre of the image. In (b), the outlines of the B-scans are shown. The reconstruction volume has dimensions of $277 \times 242 \times 299$ voxels, with a voxel dimension of 0.3 mm. In (c), an axial view of the artery is shown. A surface rendering is produced with a narrow grey-level threshold, shown from a viewpoint looking down the artery. In (d), a re-slice of the voxel array is shown. A grey-level image is obtained by slicing the voxel array at the location shown in (c). A 90 mm length of the carotid artery is visible: such a view cannot be obtained using standard 2D ultrasound.

Et harum scientarum porta et clavis est Mathematica.
'Mathematics is the door and the key to the sciences.'
—Roger Bacon, **Opus Majus**, 1267

Chapter 3

RBF Interpolation

3.1 Introduction

While the B-scans acquired in a freehand examination can be reviewed individually, reconstructing them into a regular array makes it possible to use conventional 3D visualisation and data analysis tools. These tools include any-plane slicing, volume rendering, surface rendering, segmentation and registration procedures. The reconstruction step is important because any loss of image quality, or the introduction of artifacts, must be avoided. For example, a small loss in image quality during reconstruction can make a barely detectable pathology in a B-scan become undetectable in the reconstructed data. This may result in a misdiagnosis.

In the previous chapter, the pixel nearest neighbour (PNN) reconstruction method was described. This popular reconstruction method, although fast and easy to implement, may not be ideal. A new reconstruction technique that aims to improve on existing techniques is proposed in this chapter and compared with the standard methods, including the PNN method. This is done by performing a series of comparative tests to evaluate each method's ability to fill gaps and average overlapping data.

3.2 Reconstruction Methods

3.2.1 Overview

Since the motion of the probe in freehand imaging is controlled by the physician, the B-scans can be at any relative position and orientation. This means the pixels lie at irregular locations in the array of voxels. The reconstruction problem is therefore classified as unstructured, or scattered, data interpolation [1] (Nielson, 1993).

A survey of the literature reveals a number of different methods for reconstruction of 3D freehand ultrasound data sets. Most of these methods are rather simple because they were designed to minimise the time and memory required for reconstruction. This is because physicians want to visualise the 3D data sets immediately after acquisition, so

[1] Since the pixels are grouped onto planes, the data are actually best described as semi-structured.

the reconstruction should take ideally only a few seconds. Although almost all freehand systems perform reconstruction at some stage, details of the reconstruction method are often unpublished since the method is considered *ad hoc*. Nevertheless, the methods that have been published can be classified into the following categories: voxel nearest neighbour interpolation, pixel nearest neighbour interpolation and distance-weighted interpolation.

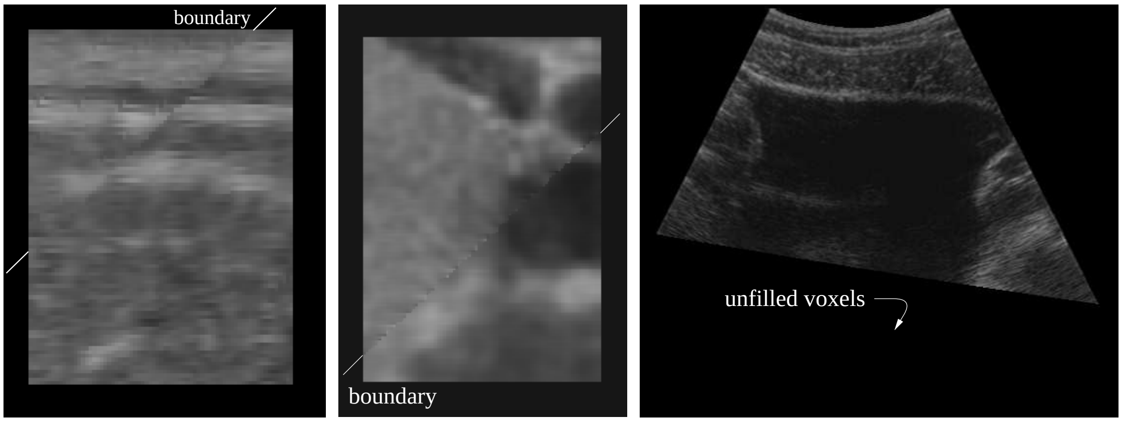
3.2.2 Voxel Nearest Neighbour Interpolation

The concept of VNN (voxel nearest neighbour) interpolation is easy to understand: each voxel is assigned the value of the nearest pixel (Sherebrin et al., 1996). There are no parameters to set. A naive implementation traverses the array one voxel at a time and finds the value of the nearest pixel, but this is computationally inefficient. Using the fact that the nearest pixel lies along a line normal to the nearest B-scan greatly speeds up the reconstruction, making it perhaps the fastest of all reconstruction methods. Moreover, a new 3D ultrasound imaging system, developed at the University of Cambridge (Prager et al., 1998a), can rapidly produce slices of the set of B-scans without reconstructing an entire voxel array. By using the voxel nearest neighbour interpolation method with dedicated graphics hardware, slices can be generated at interactive rates.

Although this reconstruction method has the advantage of avoiding gaps in the voxel array, reconstruction artifacts can be observed in slices through the voxel array. When a slice plane intersects several of the original B-scans, we can consider the interpolated image as a collage of projections from the intersected B-scans. Registration errors, including tissue motion and sensor errors, contribute to slight misalignment of the B-scans. This results in mismatches among the neighbouring pieces of the collage. The lines of intersection between the pieces then become visible — see Figure 3.1(a).

3.2.3 Pixel Nearest Neighbour Interpolation

PNN (pixel nearest neighbour interpolation) is perhaps the most popular reconstruction method (Fine et al., 1991; Hottier and Collet Billon, 1990; McCann et al., 1988; Nelson and Pretorius, 1997), and was used in Chapter 2. The basic algorithm consists of two stages. In the first stage, the algorithm simply runs through each pixel in every B-scan and fills the nearest voxel with the value of that pixel. Multiple contributions to the same voxel are usually averaged, although recording only the maximum value has also been suggested (Nelson and Pretorius, 1997). The parameters to set at this stage are the weights on the multiple contributions. The PNN algorithm in Section 2.2 described this first stage of the algorithm, and used simple averaging of multiple contributions to the same voxel. Yet gaps can occur in the voxel array when the voxel size is small compared to the distance between the acquired B-scans. In practice, this situation is inevitable with voxel arrays of similar resolution to the B-scans. The second stage fills these remaining gaps in the voxel array. A variety of methods have been used, including averaging of filled voxels in a local neighbourhood (McCann et al., 1988; Nelson and Pretorius, 1997) and interpolating between the two closest non-empty voxels in the transverse direction to the



(a) VNN (b) PNN (c) DW

Figure 3.1: **Reconstruction artifacts.** In (a), voxel nearest neighbour interpolation is used to reconstruct an examination of a neck muscle. A small slice of the voxel array is shown. The slight misalignment between the projections of the two nearest B-scans arises from registration errors. In (b), pixel nearest neighbour interpolation is used to reconstruct an examination of the thyroid. The boundary between the voxels filled by the first bin-filling stage and voxels filled by the second hole-filling stage is evident. In (c), distance-weighted interpolation is used to reconstruct an examination of the bladder. The slice is truncated because the neighbourhood for selecting points is set too small, leaving some voxels unfilled.

B-scans (Hottier and Collet Billon, 1990). Other publications (Fine et al., 1991) do not describe the method for filling gaps, so it is assumed that they adopt similar approaches or choose the voxel size large enough to avoid gaps. The parameters to set at this stage are the weights of the nearby voxels used to fill the gaps.

In summary, this method can be considered as a two stage process: "bin-filling" and "hole-filling". The first stage of bin-filling the voxel arrays with pixels is fast, but the second stage may take longer, depending on the particular method chosen. Unfortunately, artifacts can be generated by this two stage process. For example, voxels calculated by local averaging in the second hole-filling stage can appear smoothed. Thus a slice passing through regions of both first stage and second stage filled voxels may show the boundary between the highly detailed bin-filled voxels and the smoothed hole-filled voxels — see Figure 3.1(b).

3.2.4 Distance-weighted Interpolation

Like the voxel nearest neighbour interpolation method, DW (distance-weighted) interpolation proceeds voxel by voxel. Yet instead of using the nearest pixel, each voxel value is assigned the weighed average of some set of pixels from nearby B-scans. The parameters to choose are the weight function and the size and shape of the neighbourhood.

The simplest approach is to consider a fixed spherical neighbourhood of radius r_{max}, centred about each voxel (Barry et al., 1997). All pixels in this neighbourhood are weighted by the inverse distance to the voxel and then averaged. This is similar to Shepard's interpolation method (Shepard, 1968). Unfortunately, it has the disadvantage of not reproducing any of the local shape properties implied by the data, because the resulting interpolant typically has local extrema at the data sites (Nielson, 1993). It also requires r_{max} to be set prior to reconstruction. If r_{max} is set too small, gaps may result — see Figure 3.1(c). Yet if r_{max} is set too large, the voxel array will appear highly smoothed, since the effect of inverse distance weighting can be quickly overwhelmed by the much larger number of data points falling into the larger local neighbourhood. Nevertheless, with dense B-scans and a small value of r_{max}, excellent results are claimed (Barry et al., 1997).

Another distance-weighted interpolation method was designed to be incremental, so that the reconstruction volume could be viewed during acquisition of the B-scans (Ohbuchi et al., 1992). This means that filtering and hole-filling cannot wait until the end of the examination. To reduce the chance of gaps remaining after all B-scans are acquired, each B-scan fills a "thick slab" of voxels surrounding it. The voxels in this thick slab are calculated by convolving the 2D B-scans with a truncated 3D Gaussian kernel. To mimic the spatial resolution of the ultrasound beam, the kernel is made wide perpendicular to the B-scan and narrow within the B-scan plane. The convolution acts as both a way of filling gaps and a filter to smooth the data. Multiple contributions to the same voxel can occur when the thick slabs overlap, but the Gaussian shape of the kernel ensures that the multiple contributions are averaged with larger weights on the contributions from closer

B-scans [2]. In summary, each voxel value represents a weighted average of nearby B-scans, with an asymmetric Gaussian weight function attached to each B-scan. Although this method has the advantage of being incremental, the trade-off is that gaps may still occur. If the truncated Gaussian weight function is smaller than the distance between B-scans, some voxels will remain unfilled.

Another distance-weighed method that uses a non-uniform neighbourhood is based on grouping pairs of B-scans (Trobaugh et al., 1994). In this method, each voxel is filled only by the nearest two B-scans that fall on either side of it. For each of the two B-scans, lines are drawn perpendicular to the B-scan plane, passing though the voxel. Bi-linear interpolation within each B-scan determines the contributing pixel value from that B-scan. The voxel is then set to the inverse distance-weighted average of the two contributing pixel values. This method has the advantage of retaining the resolution of each B-scan in the voxel array and avoiding gaps. Although it is clear which pairs of B-scans to use with simple motions, such as linear sweeps, it is not clear which pairs to use with more arbitrarily located B-scans. It has been suggested that the pairs be chosen time sequentially (Carr, 1996), but this may not be optimal for all types of scanning motions.

3.2.5 Radial Basis Function Interpolation

It is perhaps intuitively obvious that reconstruction should be attempted using a traditional functional interpolation method. In other words, choose a particular function, such as a polynomial, determine the polynomial coefficients to make it pass through the pixel values, then evaluate at regular intervals to produce a voxel array. Yet there have been no previously published attempts at functional interpolation of 3D freehand ultrasound data, since there are severe computational demands to overcome. For example, simple volume splines (Nielson, 1993) require solving a set of N linear equations, where N is the number of pixels summed over all B-scans. This requires $O(N^3)$ calculations. For modern computer workstations, N cannot be much larger than 500 before this becomes prohibitively slow and numerical inaccuracies arise (Nielson, 1993). This makes a typical ultrasound reconstruction problem of several million pixels infeasible.

There is a great deal of literature on interpolation, spanning a variety of disciplines, and progress has recently been made in the area of tackling very large data sets. Reviews of some of the popular methods are described in (Foley, 1987; Franke, 1982; Nielson, 1993). To help select an appropriate method for solving the 3D ultrasound reconstruction problem, the requirements are listed below:

- the method must interpolate scattered trivariate data;
- the method must be fast, *i.e.* $O(N)$ complexity and non-iterative;
- the interpolating function must be smooth;

[2]The authors also include a weight related to the "age" of each of the multiple contributions, so that new overlapping B-scans overwrite old B-scan values.

- both interpolation and approximation must be possible;
- large overshoots must be avoided.

The smoothness requirement arises from the assumption that the input data (the set of B-scans) is smooth. Although the underlying anatomy is not smooth, it is measured by a finite width ultrasound beam with a smooth intensity profile. This is tantamount to convolution of the anatomical function with the beam profile, creating a smoothed image. Signal processing, such as filtering of the echo signals, smooths the data even more.

The approximation requirement arises from the existence of measurement errors. These errors include tissue motion (such as breathing, pulsative and whole body motions), position sensor errors and calibration errors. This means it may be desirable to change the interpolating function into an approximating function that passes close to, but not exactly through, the data points [3].

The final requirement of eliminating overshoots refers to the desire to have the range of interpolated voxel values in the same [0, 255] grey-level range as the B-scans. This requirement is related to the approximation requirement, since an approximating function can often reduce overshoots compared with interpolating functions.

After surveying the recent advancements in trivariate interpolation of large data sets, a method was discovered that ideally suits these requirements. This method was recently developed by researchers at the University of Illinois for interpolation of multivariate geographical data sets (Mitášová et al., 1995). They dubbed the method "completely regularized splines with tension". The Illinois method is summarised as follows.

Consider a set of pixel values $p_j, j = 1, \ldots, N$ that are located at the positions $\mathbf{x}_j$, where $\mathbf{x}_j = (x_j, y_j, z_j)$ is expressed with respect to the voxel array. The basic idea is to find a spline $S(\mathbf{x})$ that passes as close as possible to the data points and is as smooth as possible. These two requirements can be combined together in such a way that $S(\mathbf{x})$ fulfills

$$\sum_{j=1}^{N} |p_j - S(\mathbf{x}_j)|^2 + wI(s) = \text{minimum}. \tag{3.1}$$

The first component is the deviation of the spline from the data points, and the second is a smoothness function $I(s)$. The weight w determines the relative cost of the two components.

The solution can be expressed (Talmi and Gilat, 1977) as

$$S(\mathbf{x}) = T(\mathbf{x}) + \sum_{j=1}^{N} a_j \mathbf{R}(\mathbf{x}, \mathbf{x}_j) \tag{3.2}$$

[3]Strictly speaking, the term interpolation should only refer to functions that pass exactly through the data points. For simplicity, hereafter interpolation shall refer to either interpolating or approximating functions.

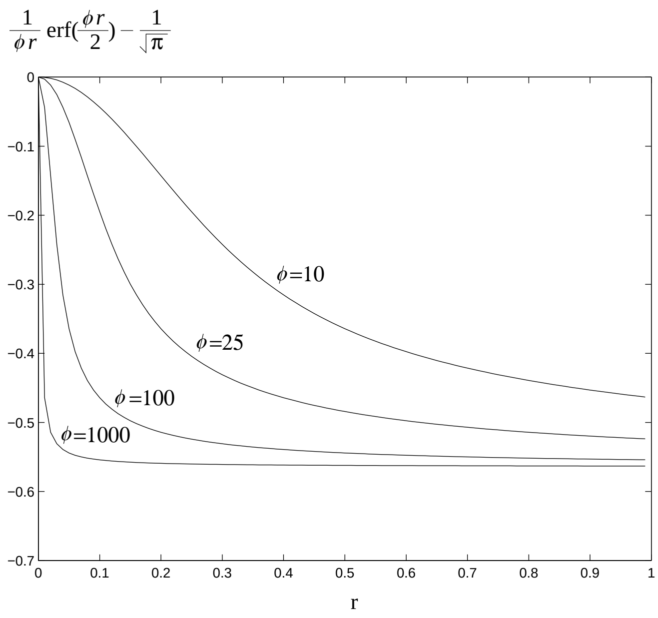


Figure 3.2: **Radial basis function.** Increasing the tension reduces the range of influence of the radial basis function.

where $T(\mathbf{x})$ is the trend function, a_j are scalar coefficients, and $\mathbf{R}(\mathbf{x}, \mathbf{x}_j)$ is an RBF (radial basis function) whose form depends on the choice of $I(s)$. In other words, the interpolation function is a combination of a trend function (typically a polynomial) and a set of simple radial functions placed at each data point. The radial basis functions are weighted so that their sum, when added to the trend function, passes close to the values of the given data points.

For the 2D case, if $I(s)$ is chosen to minimise the cost of the second (partial) derivatives only, then the familiar thin plate spline results. If the same $I(s)$ is used for the 3D case, the first derivatives of the RBF become divergent at the data points (Mitáš and Mitášová, 1988). By carefully choosing a more general $I(s)$, it is possible to obtain a simple analytic expression for the RBF with regular derivatives of all orders everywhere (Mitášová and Mitáš, 1993). This choice results in $T(\mathbf{x}) = a_0$, a constant, and

$$\mathbf{R}(\mathbf{x}, \mathbf{x}_j) = \frac{\phi^3}{4\pi}\left[\frac{1}{\phi r}\mathrm{erf}\left(\frac{\phi r}{2}\right) - \frac{1}{\sqrt{\pi}}\right] \tag{3.3}$$

where $r = |\mathbf{x} - \mathbf{x}_j|$ is the distance from $\mathbf{x}$ to $\mathbf{x}_j$, and erf is the error function (Abramowitz and Stegun, 1965). The parameter ϕ is a generalised tension parameter, and it controls the distance over which the point influences the resulting hypersurface. The multiplicative constant $\phi^3/4\pi$ can be omitted, since it can be combined with the coefficients a_j. A plot of the RBF is shown in Figure 3.2.

The spline coefficients can then be found by solving the set of linear equations

$$a_0 + \sum_{j=1}^{N} a_j \left[\mathbf{R}(\mathbf{x}, \mathbf{x}_j) + \delta_{ij} w\right] = p_i \qquad i = 1 \ldots m \tag{3.4}$$

$$\sum_{j=1}^{N} a_j = 0 \tag{3.5}$$

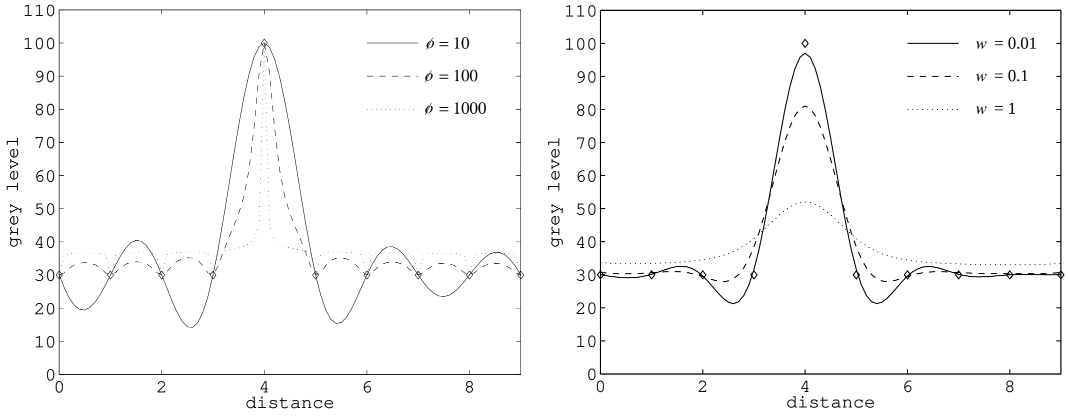
where δ_{ij} is the Kronecker delta function.

There are therefore two parameters that can be adjusted to tune the nature of the interpolant: ϕ controls the tension, and w controls the level of approximation. The goal of tuning the interpolant is to find the optimal balance between the requirements of obtaining small deviations from the data points and avoiding overshoots. As Figure 3.3(a) shows, a high level of tension limits the distance at which each point influences the overall interpolant. Yet the RBF appears spiky and does not replicate the overall shape of the data. Low tension results in overshoot. Figure 3.3(b) shows that the overshoot from low tension interpolation may be effectively controlled by allowing a small amount of approximation. In fact, it will be shown that the combination of low tension and a small level of approximation works well with ultrasound data.

As previously mentioned, the RBF interpolant cannot be calculated using all the data points of an ultrasound examination at once. In order to localise the volume splines, the scattered input data must be divided into manageable **segments**. The basic idea is that the interpolating function in a local region is not influenced by data at some sufficiently distant point. The voxel array is therefore divided into many small, non-overlapping rectangular segments. Individual interpolating functions are then calculated for each segment until all of the voxel array is covered.

To ensure smooth connections among the RBF's of neighbouring segments, overlapping **windows** are used. This means that a window is established around each segment under consideration in such a way that it encompasses not only all the data points in the segment but also a sufficient number of neighbouring data points. All data points in the window are then used to calculate the RBF's for that segment. Since the windows overlap each other, the RBF for each segment will closely (but not necessarily exactly) match the neighbouring RBF's. With a reasonable amount of overlap, the differences will be negligible compared to the quantisation of the grey levels in the 8-bit voxel array.

Segments of fixed size, each containing fewer than a maximum set number of data points (K_{max}), were initially proposed in (Mitášová and Mitáš, 1993). Windows were expanded from $3{\times}3{\times}3$ surrounding segments to $5{\times}5{\times}5$ and so on, until a sufficient number (K_{min}) of neighbouring points were included. In later work, the segmentation method was further developed using an oct-tree representation of the voxel array to improve the ability to interpolate data with a heterogeneous spatial distribution (Mitášová et al., 1995). This means the segments were continuously divided into eight subsegments until each contained no more than K_{max} points. In this way, segments of variable size were created to account



(a) Interpolation

(b) Approximation

Figure 3.3: **1D example of radial basis function interpolation and approximation.** 10 sample data points obtained from an impulse function are shown as diamonds. In (a) w is set to zero so the RBF exactly passes through the data points. The average of the data points sets the trend part of the interpolant to a value of 37. With high tension, the interpolation quickly returns to the trend between data points. With low tension, overshoots appear. Increasing the tension can be thought of as changing the nature of the interpolant from a stiff plate to a membrane. In (b) the tension level ϕ was set to a value of 20. Increasing the level of approximation reduces the amount of overshoot but can result is large deviations from the data points. A small level of approximation produces a reasonable trade-off.

for the clustering of data points. The window around each segment expanded equally in all directions until it encompassed at least K_{min} points.

If the data are only mildly heterogeneous, this windowing method works well. Unfortunately, the problem of 3D ultrasound reconstruction involves highly clustered data; all of the input data lie within the planes of the B-scans. For this reason, a more flexible windowing method is developed here.

The basic problem with expanding a window until a fixed number of data points is encompassed, is that all the data points may lie on only one side of the segment. This can create problems with continuity of the interpolating function between segments. A new windowing method is proposed instead, where the window around the segment is expanded until data points are found in all directions around the segment. Examples of the conventional and improved windowing methods are shown in Figure 3.4.

In the 3D case, the window is expanded in all directions at first, but each of the six faces of the window stops expanding only when a sufficient number of data points fall within the region defined by that direction — see Figure 3.5. In practice, expansion of the window in a given direction also stops when it reaches the extents of the voxel array.

The Illinois software was originally designed to work with 1000 to 10000 data points, so a number of practical changes are necessary. In order to accommodate several million data points from an ultrasound examination, points in the oct-tree are stored using recursive dynamic memory allocation. The oct-tree segmentation is also modified to allow segments to be divided into four (quad-tree) or two (binary tree) subsegments when one dimension of the segment is only one voxel wide. This improves the ability to divide the volume into manageable segments. Yet some very large segments can still remain. For example, a single empty segment between two nearly parallel B-scans can extend the full length and width of the B-scans. The window around this segment, expanding one voxel at a time, will suddenly encompass an entire neighbouring B-scan in a single step. This produces an intractable number of data points, so these large, empty segments are chopped into smaller segments too.

In summary, a localised trivariate spline with $O(N)$ complexity is optimised for the freehand reconstruction task. The method is efficient enough to handle very large data sets and flexible enough to handle highly heterogeneous data. The smooth interpolating function has regular derivatives of all orders, and has good accuracy compared with other interpolation methods (Mitášová and Mitáš, 1993). It also has a variable tension and can be tuned between interpolation and approximation.

3.3 Comparison of Reconstruction Methods

3.3.1 Method

Tests

A Toshiba Powervision 7000 ultrasound machine was used with a Polhemus Fastrak position sensor mounted on the probe. To minimise interference from the electronics inside

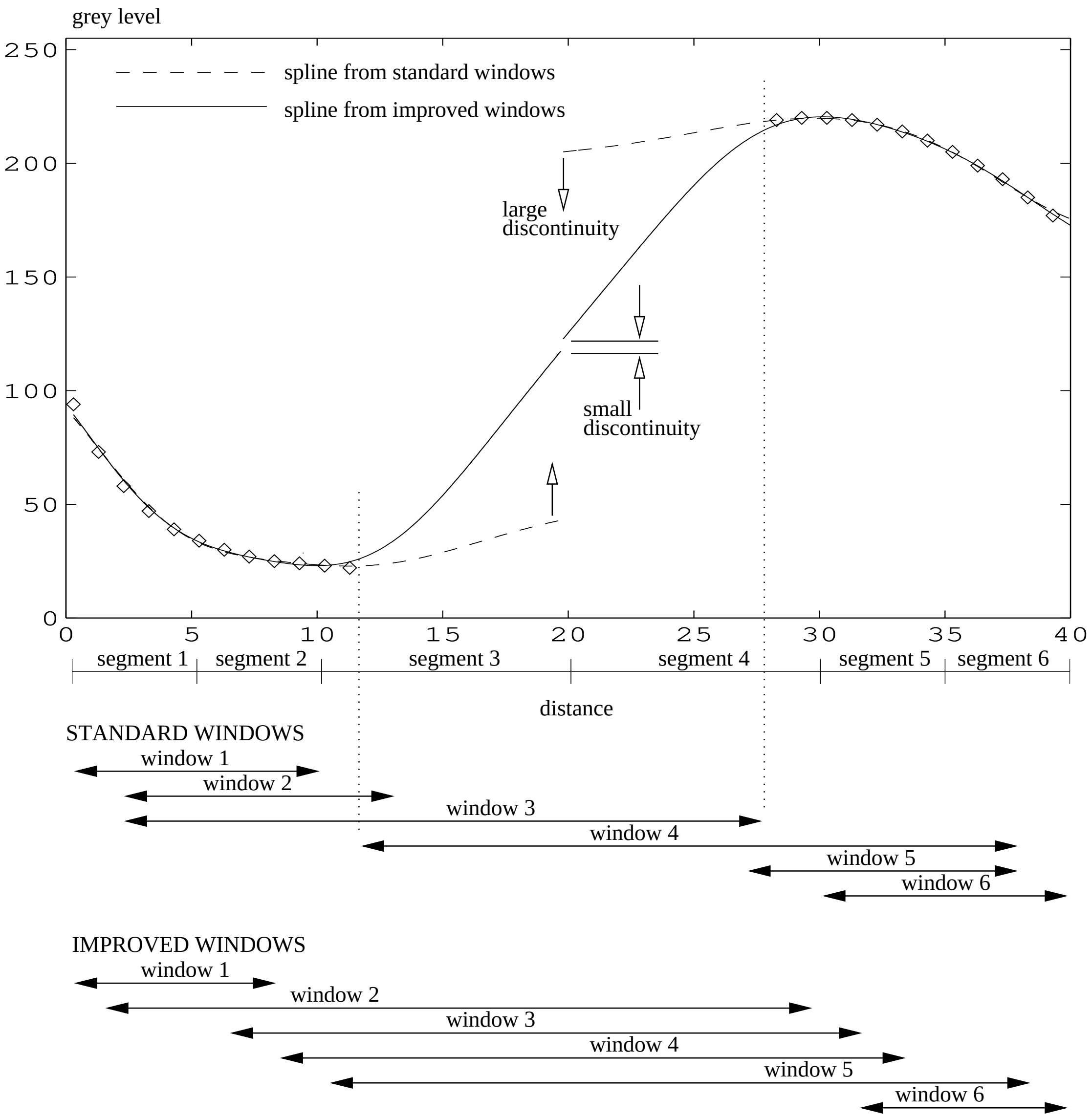


Figure 3.4: **Example of improved windowing method.** This contrived 1D example illustrates the potential problem of the standard windowing method. Segments are created by subdividing the length of interest in a binary tree so that each segment contains no more than 5 (K_{max}) data points. The RBF's are calculated (with tension $\phi = 3$, and smoothing $w = 0.01$) for each segment using the data points that fall in the surrounding window. The standard windowing method expands in both directions from the segment in single increments until 10 (K_{min}) data points in total are obtained. The standard windows 3 and 4 only contain data on one side of the segment. This creates a large discontinuity in the interpolating function between segments 3 and 4. The new windowing method expands in both directions until at least 4 points are found on each side of the segment, or the limits of the data set are encountered. This way, the interpolants for segments 3 and 4 contain data from both sides of the segments and therefore meet much closer together.

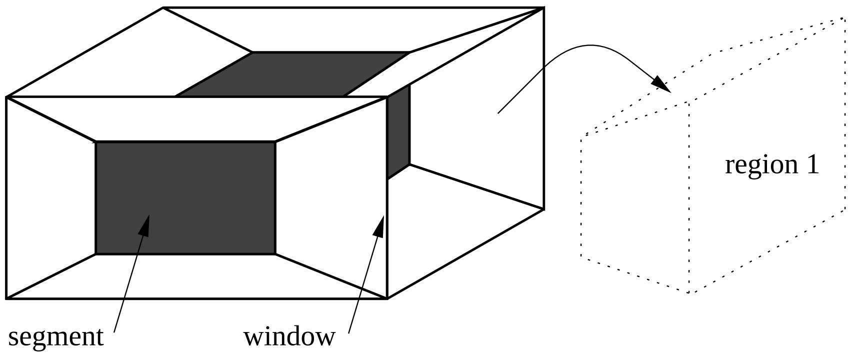


Figure 3.5: **Segment and surrounding window.** The segment is the shaded volume surrounded by the window. The window must encompass data points in each of the six regions around the segment, one of which is highlighted. This is done by expanding each of the six faces of the window until each region contains a minimum number of data points. Determination of which region a point falls into is calculated efficiently using the cross products of the vectors to the corners of the planes separating the regions.

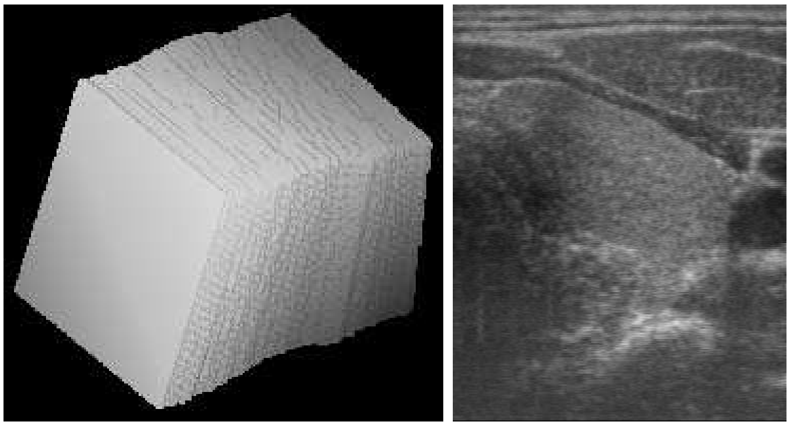
the ultrasound probe, the receiver was mounted at the end of a rigid nylon extension bar about 25 cm from the probe face. Calibration was performed with the Cambridge phantom (Prager et al., 1998b) using suitably warm water (approximately 40° Celsius) to match the speed of sound in human tissue. The Stradx acquisition software (Prager et al., 1998a) was used in conjunction with an 8-bit frame grabber and a Silicon Graphics Indy workstation. The images were matched to the position sensor readings and recorded at a rate of 25 fps.

Two examinations were performed *in vivo* by a physician on a healthy human subject. First, an examination of a thyroid gland was performed with a 7 MHz linear array probe. Each B-scan was cropped to 328×409 pixels, with a pixel size of 0.087 mm. A single sweep of the organ with a slow and steady motion resulted in a dense set of 219 nearly parallel B-scans. Figure 3.6 shows the outlines of the set of B-scans and a typical B-scan.

The second examination was performed with a fan-shaped sweep over the extents of the bladder. A 3.75 MHz convex curvilinear array probe was used, and a total of 468 B-scans were acquired. Each B-scan was cropped to 480×413 pixels, with a pixel size of 0.34 mm. Because the probe produces sector-shaped B-scans, the acquired images were masked so that only the ultrasound intensity data were used in each of the interpolation methods. Figure 3.7 shows the outlines of the set of masked B-scans and a typical B-scan.

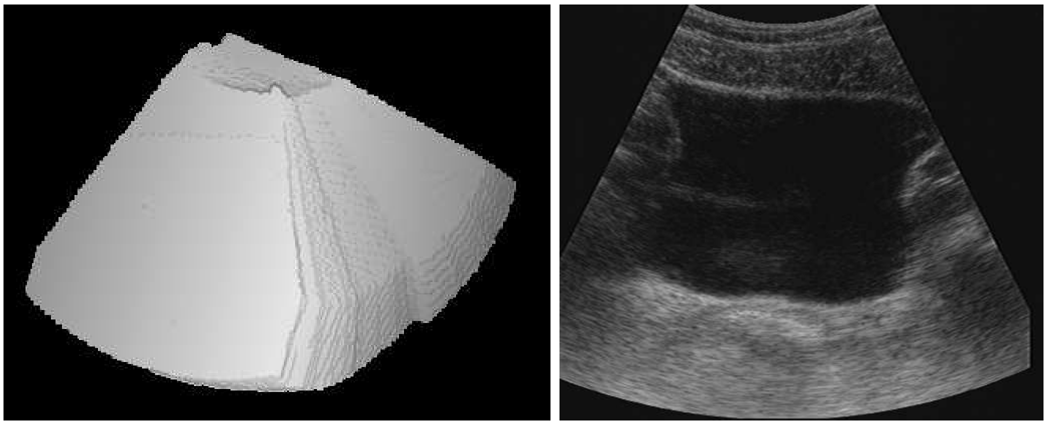
These two examinations were chosen because they allow tests of the reconstruction methods with different organs, probes, depth settings, ultrasound machine settings and types of probe motion.

The tests were designed to evaluate the ability of the reconstruction methods to interpolate the ultrasound data and fill in gaps. Since the true underlying anatomical function is unknown, the different methods were tested by artificially removing data from the two examinations. The four different reconstruction methods were then evaluated on their



(a) Outlines (b) B-scan

Figure 3.6: **Thyroid examination.** The outlines of the B-scans are shown in (a) with a typical B-scan shown in (b).



(a) Outlines (b) B-scan

Figure 3.7: **Bladder examination.** The outlines of the masked B-scans are shown in (a) with a typical B-scan shown in (b).

ability to predict the intensity values at the locations where the data were removed. In other words, a good reconstruction method will interpolate the removed data points with values very near to the data that was originally there.

First, a B-scan near the middle of the sweep was selected. The voxel array (with voxels equal in size to the pixels) was aligned exactly with this B-scan so that pixels fell exactly onto voxels. A percentage of the pixels was then removed randomly from the B-scan, creating gaps of various sizes. The rest of the pixels and all other B-scans in the reconstruction were used in the interpolation to fill in all voxels in the voxel array. The values of the removed (original) pixels could then be compared to the values of the voxels aligned with them. The voxel array stored the interpolation results as floating point numbers to avoid the influence of quantisation. The average of the difference between the interpolated and the original data over all missing data points was then calculated by

$$V = \frac{1}{M} \sum_{i=1}^{M} |p_i - c_i| \tag{3.6}$$

where p_i is the original pixel that was removed from the reconstruction, c_i is the interpolated value of the voxel aligned with p_i and M is the number of removed pixels. A low value of V indicates a good ability to interpolate over the gaps.

The tests were performed with eight different percentages of removed data: 0%, 25%, 50%, 75%, 100%, 300%, 500% and 700%. For the 25%, 50%, 75% and 100% tests, pixels were removed only from the selected B-scan n. The 300% test removed all of the pixels of B-scan n and all of B-scans n–1 and n+1. The 500% and 700% tests further removed B-scans $n \pm 2$ and $n \pm 3$. The 0% test was also included, since a reconstruction method may not exactly replicate the original data points. For example, a functional approximation method will miss the data points. For the 0% test alone, V was calculated over all pixels of the selected B-scan.

Each of the eight tests were repeated for ten different B-scans to give mean and variance estimates of V. In this way, 80 voxel arrays were created for each of the four reconstruction methods, producing 320 voxel arrays for the thyroid examination. A further 320 voxel arrays were created in the same way for the bladder examination.

Implementation of Reconstruction Methods

The VNN interpolation method was implemented by traversing the voxel array and filling each voxel with the value of the nearest pixel. The PNN interpolation method was implemented in two steps. The first step assigned each pixel to the nearest voxel in the array. Multiple contributions to a single voxel were averaged together. The second step filled the remaining gaps. Empty voxels were filled by taking the average of the filled voxels in a $3 \times 3 \times 3$ neighbourhood. The remaining unfilled voxels were then filled by averaging originally filled voxels in a $5 \times 5 \times 5$ neighbourhood and so on, until all voxels were filled. This is similar to the method described in (Nelson and Pretorius, 1997).

The high quality reconstructions reported in (Barry et al., 1997) motivated the implementation of the inverse distance-weighted method with a spherically shaped local neighbourhood. The one parameter to choose in this method was the radius, r_{max}, of the neighbourhood. If r_{max} was set to cover the largest gaps in the 700% tests, it would have been much too large for the tests with smaller gaps and excessive smoothing would have occurred. The radius r_{max} must be set to reasonable but not arbitrary values for each of the tests.

In (Barry et al., 1997), a 7 MHz linear array probe was used to examine the carotid artery, a similar examination to the thyroid gland. The authors suggested two r_{max} values: 0.25 mm and 0.5 mm. In the examination of the thyroid gland, the distances between the centres of the B-scans ranged from 0.09 mm to 0.52 mm, with a mean of 0.32 mm. Using a radius r_{max} equal to 0.25 mm would not have been large enough to fill in all gaps, so 0.5 mm was used in the 0%, 25%, 50%, 75% and 100% tests [4]. For the 300%, 500% and 700% tests, r_{max} must be increased to cover all gaps. It was sensible to increase r_{max} by the mean (0.32 mm) of the B-scan spacing, so r_{max} was set to 0.82 mm (0.5 mm + 0.32 mm), 1.14 mm (0.5 mm + 2×0.32 mm), and 1.46 mm (0.5mm + 3×0.32 mm) for the 300%, 500% and 700% tests respectively.

To be consistent, the values of r_{max} for the bladder reconstruction were set in a similar fashion. The radius r_{max} was set to the maximum distance between B-scans for the 0%, 25%, 50%, 75% and 100% tests, then increased by the mean distance between B-scans for the 300%, 500% and 700% tests. Since the examination used a fan-shaped sweep, the B-scan spacing was measured at the bottom centre of the B-scan, where the gaps were larger. The B-scan spacing ranged from 0.41 mm to 0.80 mm, with a mean of 0.60 mm. The value of r_{max} was therefore set to 0.80 mm for the 0%, 25%, 50%, 75% and 100% tests, and 1.40 mm (0.80 mm + 0.60 mm), 2.00 mm (0.80 mm + 2×0.60 mm) and 2.60 mm (0.80 mm + 3×0.60 mm) for the 300%, 500% and 700% tests respectively.

The RBF method was implemented using the oct-tree segmentation and improved windowing technique described in Section 3.2.5. K_{max} was set to 30 data points, and each region of the window was required to contain at least 5 data points.

One final practical change was made to the RBF method. The removal of data in the 300%, 500% and 700% tests of the thyroid examination produced very large gaps in the reconstruction volume. As mentioned in Section 3.2.5, large empty segments that result from large gaps in the voxel array are chopped into smaller segments. Yet even with smaller segments, the surrounding window still occasionally encompassed a very large number of data points when the gaps were large. To resolve this occasional problem in the 300%, 500% and 700% tests, the window was expanded asymmetrically — see Figure 3.8. This was done simply to speed up the calculations, since such a large overlap was not required to get smooth transitions among the segments. As a reassurance for smooth connections among segments, at least 20 points per region were required for these 300%, 500% and

[4] It appears that setting r_{max} to half of the maximum centre distance between B-scans would have filled the entire volume without gaps. But because the scan planes were not exactly parallel, larger gaps than 0.52 mm existed. In fact, the value of 0.5 mm was just sufficient to cover all gaps.

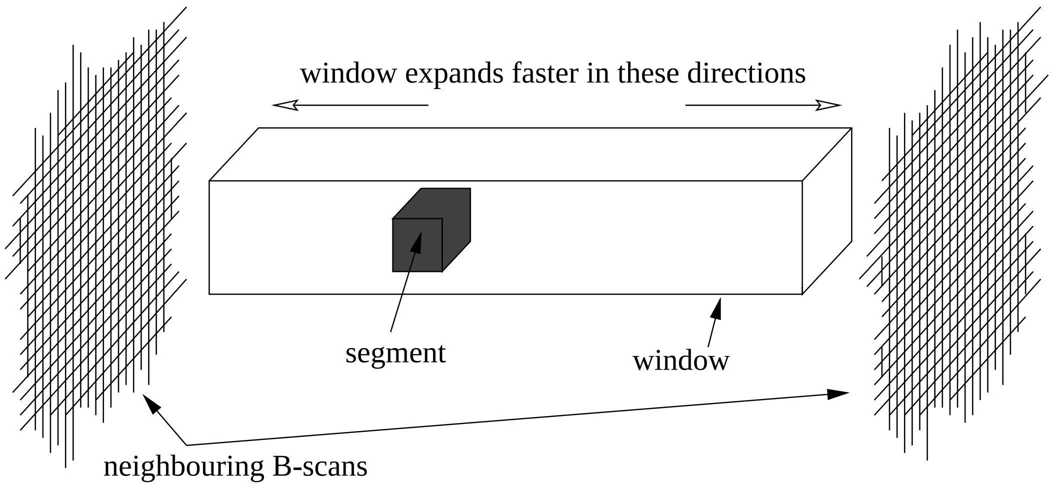


Figure 3.8: **Asymmetric expansion of window.** Instead of expanding all six faces by the same increment at each step (and counting the data points that fall in each region), the two faces parallel to the selected B-scan were expanded with a larger increment. This means the window expands faster towards the nearest data (neighbouring B-scans), ensuring that the faces for these directions are still relatively small when the window encounters data.

700% tests, compared with the minimum of 5 points per region for all other tests.

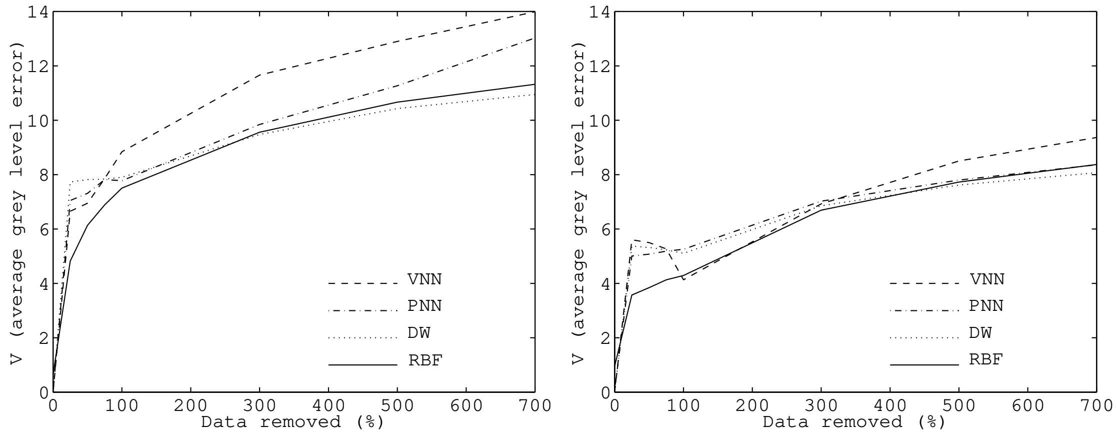
The tension and approximation parameters were tuned manually by viewing a slice of the voxel array. A fairly low tension ϕ combined with a small amount of data smoothing w resulted in a good trade-off between minimising overshoot and passing near the data points. This, in turn, produced high quality slice images. For the thyroid examination, the values $\phi = 17$ and $w = 0.01$ were used. The bladder examination used $\phi = 25$ and $w = 0.1$. These values fall within the range of values typically used for geographic data interpolation (Mitášová et al., 1995).

3.3.2 Results

The overall trends of the test results are illustrated in Figure 3.9 and tabulated in Tables 3.1 and 3.2. To give an indication of the distributions of V over the 10 trials, box-and-whisker plots are shown in Figures 3.10 and 3.11. A small but representative set of images of the interpolated data is shown in Figures 3.12, 3.13, 3.14 and 3.15.

Voxel Nearest Neighbour Interpolation

For both examinations, V is zero at 0% data removal since the voxels are set to their nearest neighbours: the original pixels of the B-scan. At 25%, 50% and 75%, the nearest neighbours of the voxels are mainly the remaining pixels of the selected B-scan. Therefore, the resulting interpolated images appear as a patchwork of irregularly shaped pieces and relatively large values of V result.



(a) Thyroid

(b) Bladder

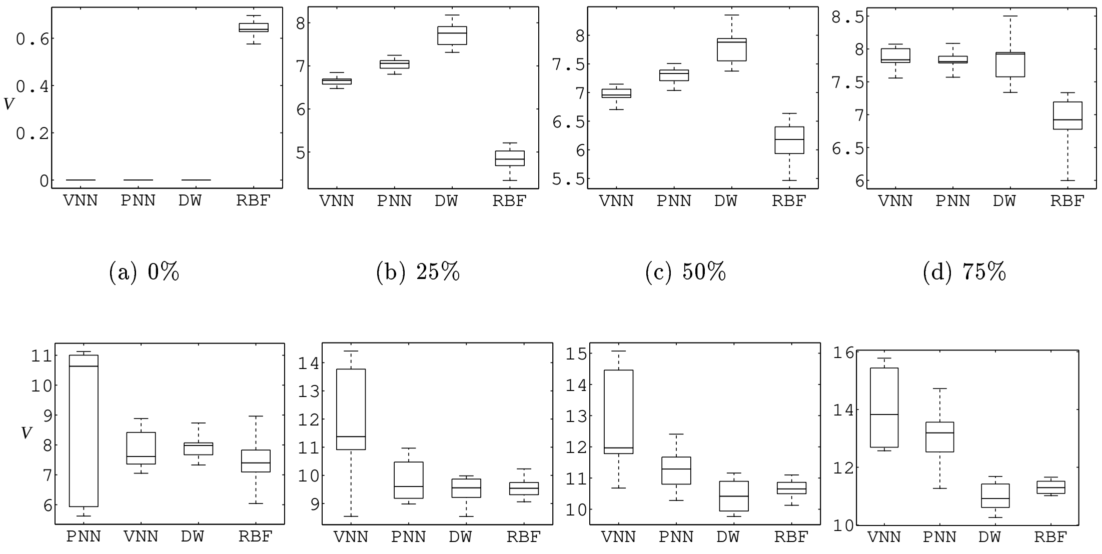
Figure 3.9: **Average interpolation error V for the thyroid and bladder data sets.** For each of the four reconstruction methods, V is calculated for various percentages of removed data.

Test	VNN		PNN		DW		RBF	
	μ	σ	μ	σ	μ	σ	μ	σ
0%	$0.00^\star$	0.00	$0.00^\star$	0.00	$0.00^\star$	0.00	0.64	0.03
25%	$6.65^\dagger$	0.12	$7.05^\dagger$	0.12	$7.73^\dagger$	0.27	4.82	0.25
50%	$6.95^\dagger$	0.14	$7.31^\dagger$	0.13	$7.82^\dagger$	0.30	6.14	0.33
75%	$7.86^\dagger$	0.16	$7.82^\dagger$	0.13	$7.84^\dagger$	0.35	6.89	0.39
100%	8.84	2.60	$7.78^\dagger$	0.64	7.90	0.42	7.51	0.78
300%	$11.66^\dagger$	2.00	9.85	0.75	9.48	0.47	9.56	0.34
500%	$12.90^\dagger$	1.65	$11.27^\dagger$	0.73	10.43	0.53	10.67	0.29
700%	$13.99^\dagger$	1.32	$13.03^\dagger$	1.00	$10.95^\star$	0.52	11.32	0.23

Table 3.1: **Interpolation error V for the thyroid examination.** μ is the mean of V and σ is the standard deviation. $\dagger$ means that the assertion $\mu > \mu_{RBF}$ is statistically significant for a confidence level of 0.05. $\star$ means that the assertion $\mu < \mu_{RBF}$ is statistically significant for a confidence level of 0.05. The assertions are tested with the paired-sample t-test statistical method (Hogg and Ledolter, 1987).

Test	VNN		PNN		DW		RBF	
	μ	σ	μ	σ	μ	σ	μ	σ
0%	0.00*	0.00	0.00*	0.00	0.00*	0.00	0.96	0.03
25%	5.60†	0.39	5.01†	0.22	5.37†	0.09	3.57	0.25
50%	5.50†	0.40	5.08†	0.25	5.32†	0.09	3.85	0.31
75%	5.27†	0.50	5.19†	0.35	5.24†	0.10	4.13	0.40
100%	4.13	0.38	5.25†	0.40	5.11†	0.14	4.29	0.37
300%	6.92	0.40	7.03†	0.15	6.85†	0.12	6.69	0.19
500%	8.50†	0.23	7.80†	0.14	7.62*	0.11	7.73	0.16
700%	9.37†	0.26	8.36	0.18	8.07*	0.09	8.37	0.16

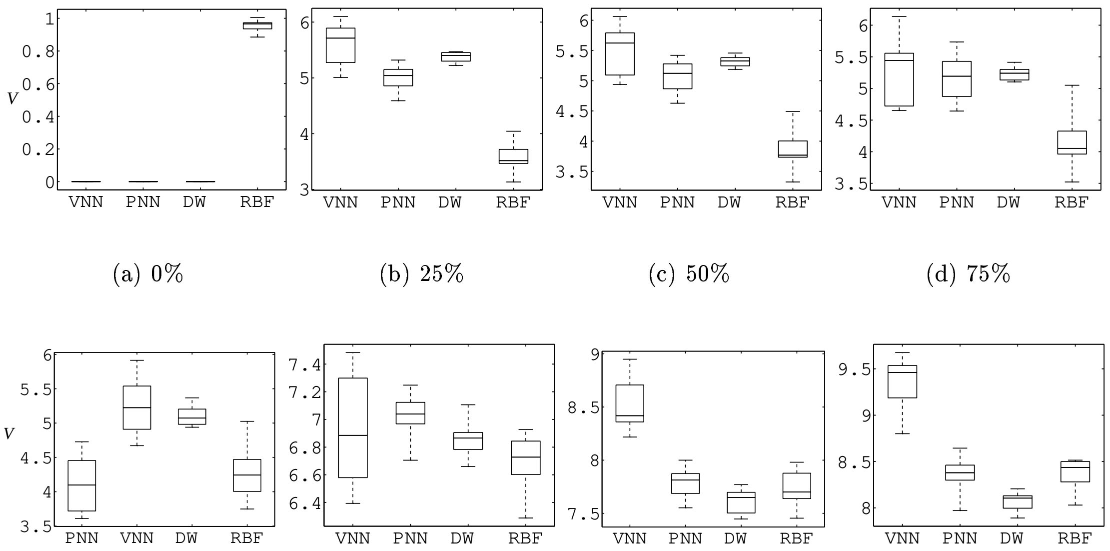
Table 3.2: **Interpolation error V for the bladder examination.** See the caption of Table 3.1 for an explanation of the tabulated terms and symbols.



(a) 0% (b) 25% (c) 50% (d) 75%

(e) 100% (f) 300% (g) 500% (h) 700%

Figure 3.10: **Interpolation error V for the thyroid examination.** The box-and-whisker plots show the three quartiles and the extremes of V calculated over the 10 trials at each level of removed data. The median (50th percentile) is the line inside the box of the 25th and 75th percentiles of V. The whiskers show the minimum and maximum values.



(a) 0% (b) 25% (c) 50% (d) 75%

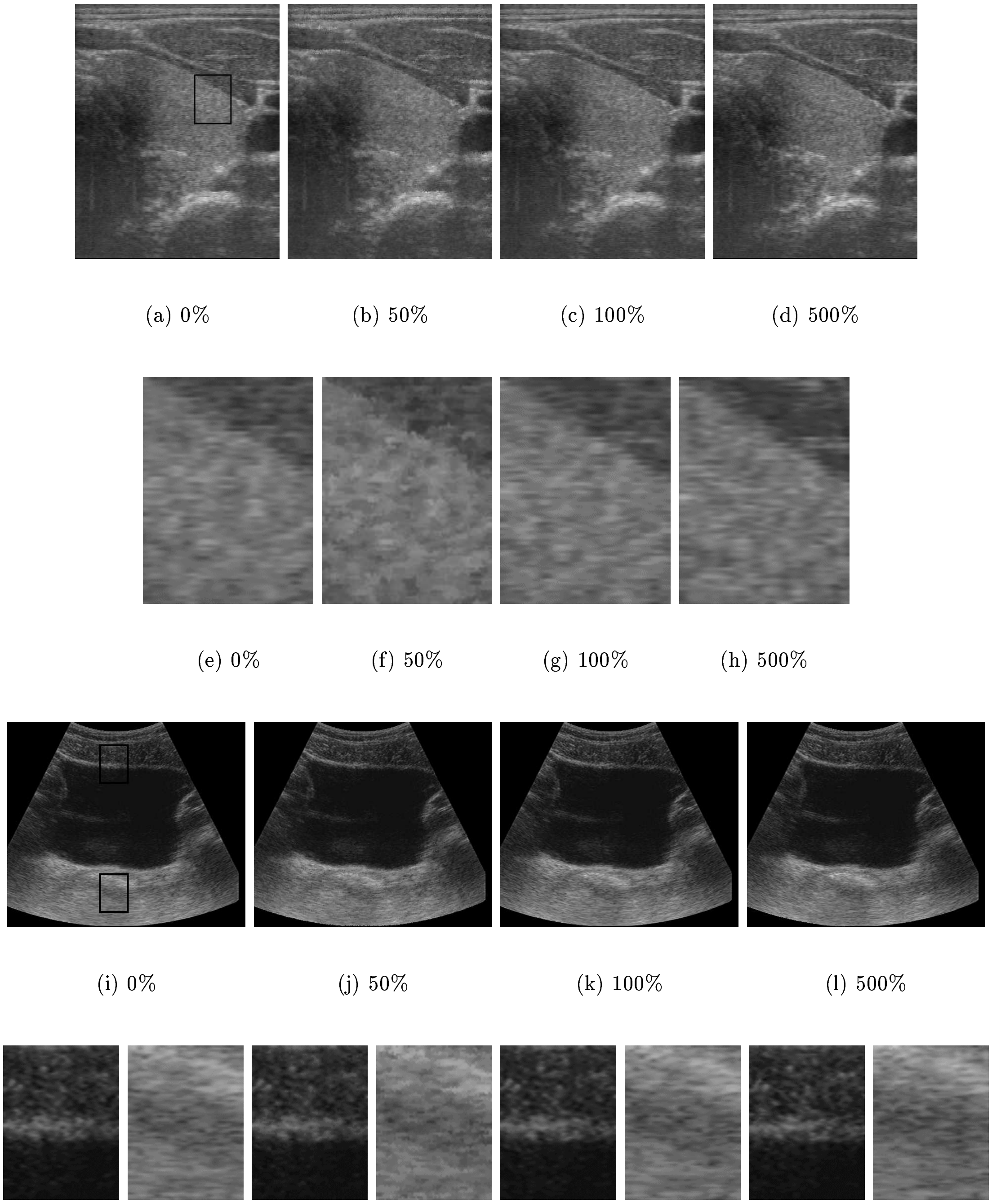
(e) 100% (f) 300% (g) 500% (h) 700%

Figure 3.11: **Interpolation error V for the bladder examination.** See the caption of Figure 3.10 for an explanation of the box-and-whisker plots.

For the 100% to 700% tests, the interpolated image is formed from the projection of the pixels from the nearest B-scan, so the patchwork appearance disappears. For the thyroid examination, the variance of V is large because some of the projections of neighbouring B-scans closely match the original B-scan, but others do not. This depends on both the level of registration error and the similarity of the neighbouring B-scans (a slowly changing anatomy results in similar neighbouring B-scans). The values of V are generally greater than those observed with the other three reconstruction methods.

For the bladder examination, the 100% to 700% tests show a smaller variance of V than the thyroid examination, but the variance is still large compared with the other methods. The mean drops at 100%, however, before increasing again for the higher percentage tests. This may be explained by comparing the level of registration errors to the voxel size. If registration errors, such as position sensor error, are of similar absolute magnitudes for the two examinations, the errors will have a smaller effect on the bladder tests. This is because the voxel size is larger in the bladder tests than in the thyroid tests, so the effect of the errors is reduced. Yet, as more and more B-scans are removed, the relative effect of registration errors increases, thereby making V increase again.

In general, however, the results look sharp and detailed for the 100% to 700% tests, since the projected data exhibit no blurring. Moreover, the join lines between the portions of the projected data are not discernible for any tests of either examination. This means that the registration errors are mainly small and the images vary slowly from one B-scan to the next.



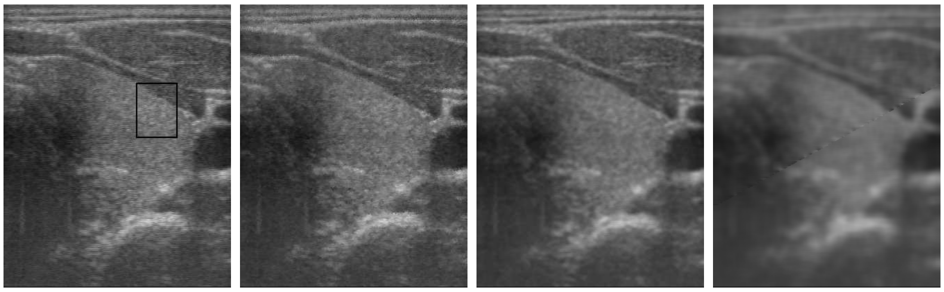
(a) 0% (b) 50% (c) 100% (d) 500%

(e) 0% (f) 50% (g) 100% (h) 500%

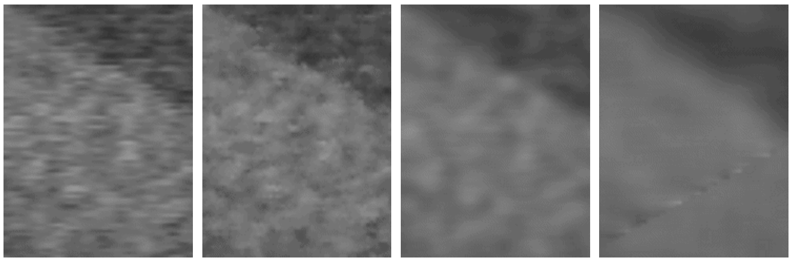
(i) 0% (j) 50% (k) 100% (l) 500%

(m) 0% (n) 0% (o) 50% (p) 50% (q) 100% (r) 100% (s) 500% (t) 500%

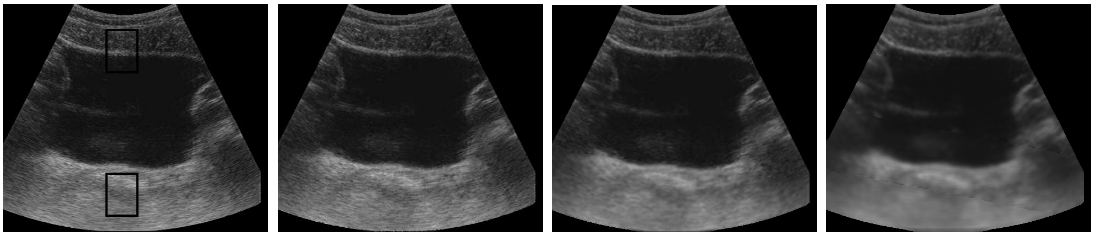
Figure 3.12: **Typical B-scans interpolated by the voxel nearest neighbour method.** All images are shown for a particular B-scan with various percentages of removed data. Images (a) to (d) are from the thyroid examination, with expanded views of the area indicated in image (a) shown in images (e) to (h). Images (i) to (l) are from the bladder examination, with expanded images of the top and bottom areas indicated in image (i) shown in images (m) to (t).



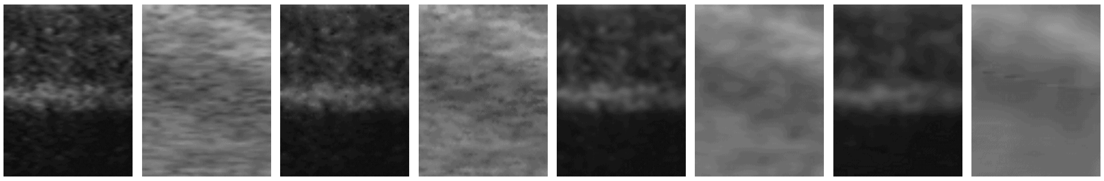
(a) 0% (b) 50% (c) 100% (d) 500%



(e) 0% (f) 50% (g) 100% (h) 500%



(i) 0% (j) 50% (k) 100% (l) 500%



(m) 0% (n) 0% (o) 50% (p) 50% (q) 100% (r) 100% (s) 500% (t) 500%

Figure 3.13: **Typical B-scans interpolated by the pixel nearest neighbour method.** See the caption of Figure 3.12 for an explanation of the images.

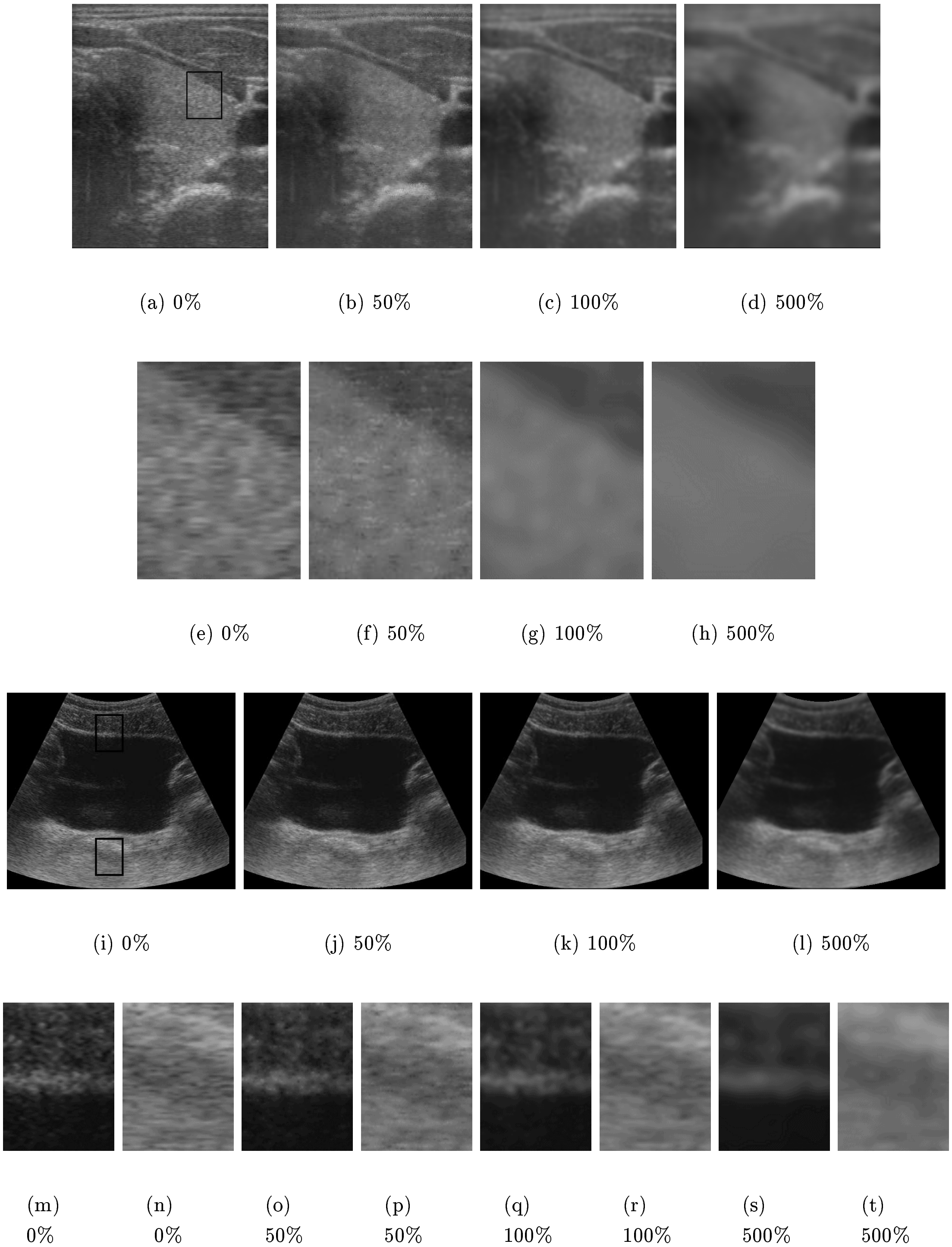
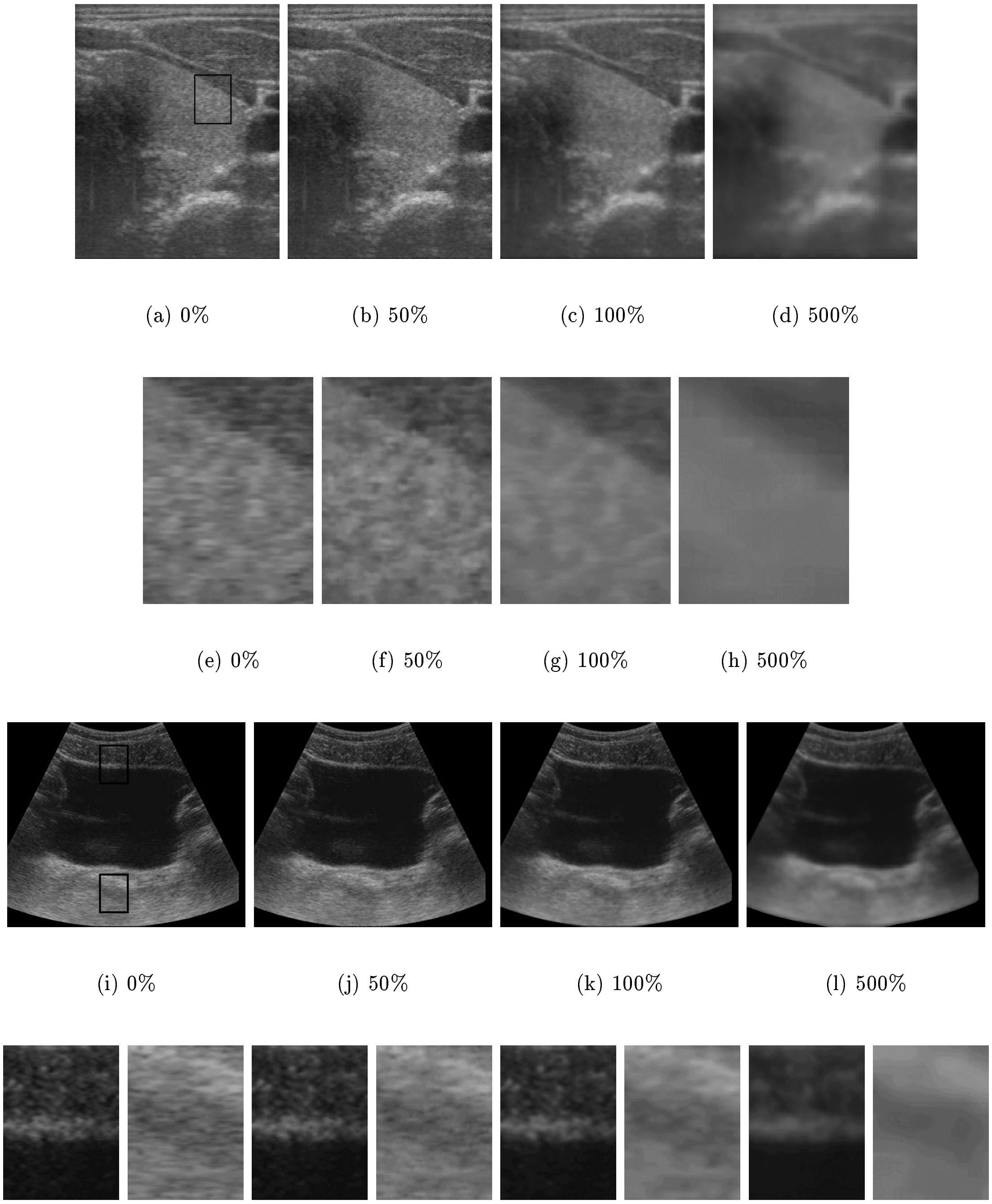


Figure 3.14: **Typical B-scans interpolated by the distance-weighted method.** See the caption of Figure 3.12 for an explanation of the images.



(a) 0% (b) 50% (c) 100% (d) 500%

(e) 0% (f) 50% (g) 100% (h) 500%

(i) 0% (j) 50% (k) 100% (l) 500%

(m) 0% (n) 0% (o) 50% (p) 50% (q) 100% (r) 100% (s) 500% (t) 500%

Figure 3.15: **Typical B-scans interpolated by the radial basis function method.** See the caption of Figure 3.12 for an explanation of the images.

Pixel Nearest Neighbour Interpolation

For both examinations, V is zero at 0% data removal since the nearest neighbours to the voxels are the original pixels. At 25%, 50% and 75%, the gaps are filled mainly with an average of the remaining pixels in the original B-scan. The interpolated image appears as a patchwork again, and relatively large values of V result.

For the thyroid examination, the mean of V increases progressively for the 500% and 700% tests in contrast to the relatively good performance in the bladder examination. This can be explained by noticing that the gaps in the bladder data set are much smaller at the top than the bottom, which limits the blurring effect of the hole-filling stage to the large gaps at the bottom of the images. With lower levels of blurring in the bladder tests, lower values of V result.

Unfortunately, the images exhibit significant reconstruction artifacts in both examinations, especially for the 500% and 700% tests — see Figure 3.13(d) and (h). A visible boundary exists between portions that are filled, for example, by averages of a $7 \times 7 \times 7$ neighbourhood, and portions filled by a neighbourhood of $9 \times 9 \times 9$, because they involve different levels of smoothing. The boundaries between the voxels filled by the first bin-fill stage and second hole-fill stage are also visible.

Distance-Weighted Interpolation

For both examinations, V is zero for the 0% tests since the weights on the original pixels that fall exactly on the voxels approach infinity. At 25%, 50% and 75%, the images consist of voxels filled by the original data (weighted by infinity), along with voxels in the gaps that are calculated from a weighted average of many neighbouring pixels. The resulting interpolated data set is therefore a combination of the remaining pixels from the original B-scan and smoothed data in the gaps.

Since the radius r_{max} of the neighbourhood is tailored for each test of 100% and greater (as explained in Section 3.3.1), the values of V for the thyroid examination are relatively small. Yet for the 100% test on the bladder examination, the mean of V is considerably higher than the RBF method, since r_{max} is too large at the top of the image where the data are dense (but just large enough to fill the gaps at the bottom), resulting in excessive smoothing at the top. For 300% to 700%, the values of V improve, since the removal of data offsets the increase in r_{max}.

The images show progressive blurring as the percentage of removed data increases, but no other reconstruction artifacts are generated. For the thyroid examination, the level of blurring is uniform. For the bladder examination, blurring is slightly greater near the top of the images.

Radial Basis Function Interpolation

Because an approximating function (which does not pass exactly through the original pixel values) is used, V is greater than zero for the 0% test. For both examinations, however, the mean of V at 0% is less than 1.0.

At 25%, 50% and 75%, the mean of V is considerably lower than the other methods and the resulting interpolated data appears the most detailed and least artificial. This is because the interpolation utilises both the remaining pixel data of the original B-scan and the neighbouring B-scans. The low values of V demonstrate the ability of a functional method to use the general shape of the underlying anatomical data to interpolate across the gaps.

Yet at percentages of 100% and greater, the RBF is not always significantly better than the other methods. In particular, the DW method matches it in the thyroid examination. In the bladder examination, the VNN method matches it for the 100% and 300% tests, the PNN method matches it for the 700% tests, and the DW method even slightly betters it for the 500% and 700% tests.

One of the reasons the RBF loses some of its advantage is that the underlying shape of the anatomical data is lost when the gaps become too large. Another problem is that the RBF approaches the trend term of the interpolation function in the largest gaps. This is because the tension needs to be set high enough so that no overshoots appear within the B-scan. Yet in the gaps, a higher tension results in a faster approach of the interpolation function towards the trend.

In general, however, the RBF method shows no artifacts in the interpolated data. The join lines between segments are not visible, since the discontinuities are smaller than the grey-level quantisation. Also, no overshoots are generated: the range of the interpolated data matches the range of grey levels in the original B-scans.

For the thyroid examination, the interpolated images become progressively blurred as the percentage of removed data increases. The level of blurring is uniform throughout the image. For the bladder examination, however, the denser data at the top of the images allow the RBF to retain a high level of detail. The blurring then progressively and smoothly increases toward the bottom of the images where the gaps are larger. The level of blurring therefore reflects the level of uncertainty in estimation of the underlying function. In comparison, the DW method shows counter-intuitively the greatest blurring near the top of the bladder images.

3.3.3 Discussion

The RBF method appears to work well; low values of V are produced and no artifacts are generated. It must be noted, however, that the computational demands are much greater than for the other three methods. The thyroid and bladder data sets can be reconstructed by the VNN, PNN and DW methods in a few minutes, but the RBF method requires several hours. Yet the oct-tree segmentation means the RBF method is amenable to parallel processing. The segments can be interpolated easily in parallel because the memory storing the pixel values does not change (eliminating read errors). Each process also writes to a different segment in the voxel array (eliminating write errors). Since many modern ultrasound machines already have the capacity for parallel processing (the Toshiba

Powervision 7000 used for these examinations contains more than 60 Pentium processors), a practical implementation of the RBF method is not infeasible.

The greatest advantage of the RBF method occurs with small percentages of removed data. This is because the underlying shape of the data can be used to interpolate across the small gaps and reduce the effect of outliers. This is why the RBF method performs especially well with small percentages of removed data in the tightly spaced bladder examination.

The higher percentage tests are representative of what would happen with faster hand motion during the examination or a slower rate of acquisition by the computer. When the gaps become too large, the RBF method loses some of its advantage because the shape of the underlying data is eventually lost. Nevertheless, with few exceptions, the RBF method performs at least as well as the other methods. It is possible that the performance of the RBF method can be improved by choosing a different combination of tension ϕ and smoothing w. But any advantage at low percentages of data removal results in a trade-off at high percentages. Moreover, the potential improvements are small. The results are mainly unchanged as long as the tension ϕ is in the range of [10, 25] and the smoothing w is in the range [0.01, 0.1].

A much greater potential improvement to the RBF lies with the introduction of anisotropic tension. It is possible for the tension to be different in different directions. This is a result of the invariance of the RBF functions to translations and rotations, but not scale (Mitášová and Mitáš, 1993). This means a change in the scale of the dimensions is equivalent to a change in the tension parameter. By changing the individual scales of the three axes, the tension is changed for each of these three directions. The tension should ideally be high within the individual B-scans to avoid overshoots, and low in the direction orthogonal to the B-scans to fill the gaps between them. This would also be likely to reduce the blurring in the gaps between B-scans that is evident in these tests.

While the RBF method produces the lowest values of V, it must be admitted that the VNN method produces the images with greatest clarity. Since the registration errors are small, the nearest neighbour projections appear to be seamless. But since both examinations of the thyroid and bladder are of organs with a slowly varying shape, the projections from different B-scans are closely aligned. If more complex shapes are scanned, or the registration errors increase, the seams will become more visible. Moreover, interpolated images in large gaps (which depict projections of distant B-scans) appear sharp, but are strictly incorrect since the anatomy is featured in the wrong place.

In summary, all four methods perform well on both examinations and produce reasonable images when the gaps are kept at realistic levels. It is also true that all of the methods could be improved, or tailored to a particular type of examination. Yet the point of this chapter is not to declare a winner among the different reconstruction methods. Rather, it is to investigate typical reconstruction techniques on typical examinations and determine whether a functional interpolation method offers any advantages.

3.4 Conclusions

The RBF reconstruction method was shown to perform at least as well as traditional reconstruction methods. The RBF method performed particularly well with small gaps, but lost some of its advantage when larger gaps were present. This was because the RBF method alone made use of the underlying shape of the data to interpolate across the gaps.

The RBF method was capable of simultaneously reconstructing areas with densely overlapping B-scans and areas with large gaps, while making a smooth transition between them. The resulting quality of the interpolated data was good, with no visible reconstruction artifacts produced in any of the tests. Unfortunately, the RBF method was the slowest reconstruction method and is not yet practical in its current form. Yet it is easily parallelised and given more efficient coding techniques a practical implementation is feasible.

Many opportunities also exist to exploit the unique properties of the RBF method. For example, derivatives can be calculated directly from the RBF's. Accurate derivatives are often required in applications such as visualisation, registration and segmentation. A functional representation can also be useful for data compression and filtering. Also, since an approximating function in general misses the data points, the distance it misses them by can be considered the predictive error. A large predictive error may be indicative of image misalignment, so determining which regions have large predictive errors can be useful for investigations into registration errors.

Part II

Spatial Compounding

Nessuna humana investigazione si pio dimandara vera scienzia
s'essa non passa per le matematiche dimonstrazione.
'No human investigation can be called real science
if it cannot be demonstrated mathematically.'
—Leonardo da Vinci, **Treatise on Painting**, 1500

Chapter 4

Spatial Compounding *In Vitro*

4.1 Introduction

Part I of this thesis focused on reconstructing the best possible 3D data sets from a set of 2D B-scans. Yet even with accurate system calibration and sophisticated interpolation, problems remain. In particular, visualisation and volume estimation are inhibited by artifacts including speckle. As mentioned in Section 1.6, spatial compounding is an effective approach for speckle reduction, and there has been considerable research in the past on 2D spatial compounding. This chapter takes a fresh look at spatial compounding from the perspective of 3D freehand ultrasound. Every freehand system has to deal with compounding in some manner, since it is almost inevitable that the scan planes will intersect. Deliberate, extensive compounding is proposed here, with the aim of producing high quality 3D data sets which lend themselves to automatic segmentation for visualisation and volume measurement.

The key to effective spatial compounding is to achieve a sufficiently high registration accuracy. The accumulation of measurement errors, listed in Section 1.5, means that the position sensor readings alone will not be sufficient. It is therefore necessary to improve the registration using image-based techniques.

Little work on the registration of 3D ultrasound data sets is evident in the literature. In one exception (Moskalik et al., 1995), two separate data sets of the breast were retrospectively registered using manual landmark matching. This constitutes a labour intensive solution to a specific registration problem. Instead, an automatic, incremental registration algorithm is proposed here, for use with generic freehand ultrasound imaging.

Similarly, little work on 3D spatial compounding has been performed. In some early cardiovascular work, excised hearts were scanned in a water bath (Belohlavek et al., 1993; McCann et al., 1988). 2D cross-sectional images were created by averaging a number of B-scans taken in the same plane but from different look directions. Although a set of these compounded cross-sectional images were stacked to form a 3D data set, the spatial compounding was strictly 2D. Since the tests were performed *in vitro*, registration errors from tissue movement (and to a smaller extent, changes in the speed of sound in surrounding tissue) were reduced. This meant that image alignment was accurate and high quality images were produced.

In another article (He, 1997), spatial compounding was used for the specific application of imaging an entire limb. Again, the limb was submerged in water. Cross-sectional images were created by averaging 36 B-scans acquired while rotating the transducer in 10° increments around the limb. The limb was imaged therefore from a 360° range of viewpoints within a single plane. Although a set of these cross-sectional images was then used to create a 3D data set, the spatial compounding was still only 2D. The author's main goal was to circumvent the problem of shadowing behind bones, but he also noted a decrease in speckle noise and an increase in SNR. Blurring was said to be a result of limb motion, and research on realignment methods was suggested, but not implemented.

One article did use a freehand system to perform spatial compounding in 3D, and improved visualisation and segmentation were reported (Barry et al., 1997). The authors used an inverse-distance-weighted interpolation method with a relatively large local neighbourhood to reconstruct the 3D data sets. As described in Section 3.2.4 and demonstrated in Section 3.3.2, this method of reconstruction can blur the data, making it difficult to separate the effects of reconstruction from spatial compounding. Alignment of the B-scans was based on the position sensor alone, so again, no image based registration was performed. Another brief article (Nelson and Pretorius, 1994) cited the improvements that are possible by 3D compounding, but it simply stated the need for accurate registration without providing further detail.

The goal of this chapter is to perform an in-depth study of spatial compounding with three major objectives in mind. The first is to demonstrate how accurate 3D registration can be achieved using a technique that takes sensor-based measurements of B-scan positions and applies small adjustments to align anatomical landmarks in the reconstructed volume. The second objective is to demonstrate how spatial compounding, coupled with accurate registration, can improve the SNR of the reconstructed data. The final objective is to develop a statistical theory of 3D spatial compounding and establish agreement between the predicted and observed improvements in SNR. These objectives are met by first studying an artificial object (phantom) *in vitro*. The phantom study allows the 3D reconstruction process to be evaluated and verified before proceeding to an examination of a human patient *in vivo*. This chapter describes the study performed *in vitro*, and the next chapter describes the study performed *in vivo*.

4.2 Reconstruction With Registration

4.2.1 Overview

To acquire high quality, spatially compounded data sets, any registration errors must be corrected before compounding. (The reader may wish to skip ahead to Figure 4.9 to get an indication of the level of registration errors.) These errors are corrected by adding an image registration step to the reconstruction algorithm.

A variety of image registration techniques are available (Brown, 1992; Maintz and Viergever, 1998), but many are specialised for particular applications or are computationally expensive. Since the artificial object studied here produces images with well defined boundaries, registration is attempted using landmarks.

Landmarks are anatomical features which are prominent in the B-scan images. When a newly acquired B-scan is compounded into a filled (or partially filled) voxel array, any landmarks in the B-scan should align with existing landmarks in the voxel array. Registration errors can be corrected by searching for corresponding landmarks and repositioning the new B-scan plane so that the landmarks are brought closer together.

Registration must be performed with respect to a reliable baseline. If all the B-scans are repositioned, it is possible to construct a voxel array where all landmarks are in perfect alignment, but the reconstruction bears little resemblance to the underlying anatomy (for instance, the voxel array could contain a sheared image of the anatomy). For this reason, the ultrasound examination commences with a quick pass over the region of interest, filling most of the voxels with little overlap of the B-scans. No attempt is made to register these initial B-scans: they act as the baseline. Any small registration errors that occur at this stage are not corrected, but since the true shape of the anatomy is unknown, these initial B-scans provide the best estimate of the true shape. Subsequent passes over the region of interest, from different look directions, are compounded with the baseline data in the voxel array. The inevitable registration errors are automatically corrected by landmark alignment as each new B-scan is acquired.

The basic PNN algorithm without the hole-filling stage is chosen for reconstruction because it offers a simple basis for the study of the SNR statistics: direct averaging of multiple contributions to the same voxel is performed, and no parameters need to be to specified by the user. With low resolution voxel arrays, no gaps are likely to remain after all of the overlapping data are compounded, but with high resolution arrays (small voxels), small gaps may remain. Instead of filling in the gaps in high resolution reconstructions, the second stage of the PNN algorithm is omitted so that the statistics are independent of the type of hole-filling method. The calculations are then made only on the filled voxels. In this way, the SNR statistics are kept as generic as possible.

Figure 4.1 describes the PNN reconstruction algorithm with the addition of landmark-based registration. The main difference between this algorithm and the algorithm without registration (Figure 2.2) is that the image transformation ${}^C\mathbf{T}_P$ is replaced with the optimal transformation ${}^C\mathbf{T}^*_P$ that registers the landmarks in the image $\mathbf{P}$ with those already present in the voxel array $\mathbf{C}$. The transformation ${}^C\mathbf{T}^*_P$ may be considered as the combination of the original ${}^C\mathbf{T}_P$ and an additional transformation $\mathbf{T}^{reg}$:

$$ {}^C\underline{\mathbf{x}} \;=\; {}^C\mathbf{T}_P\;{}^P\underline{\mathbf{x}} \qquad \longrightarrow \qquad {}^C\underline{\mathbf{x}} \;=\; {}^C\mathbf{T}^*_P\;{}^P\underline{\mathbf{x}} \;=\; \mathbf{T}^{reg}\;{}^C\mathbf{T}_P\;{}^P\underline{\mathbf{x}}. \tag{4.1} $$

In this study, $\mathbf{T}^{reg}$ is constrained to a 6 DOF rigid body transformation. Although some non-rigid registration errors are undoubtably present, it is hoped that the majority of registration errors can be corrected with rigid body registration.

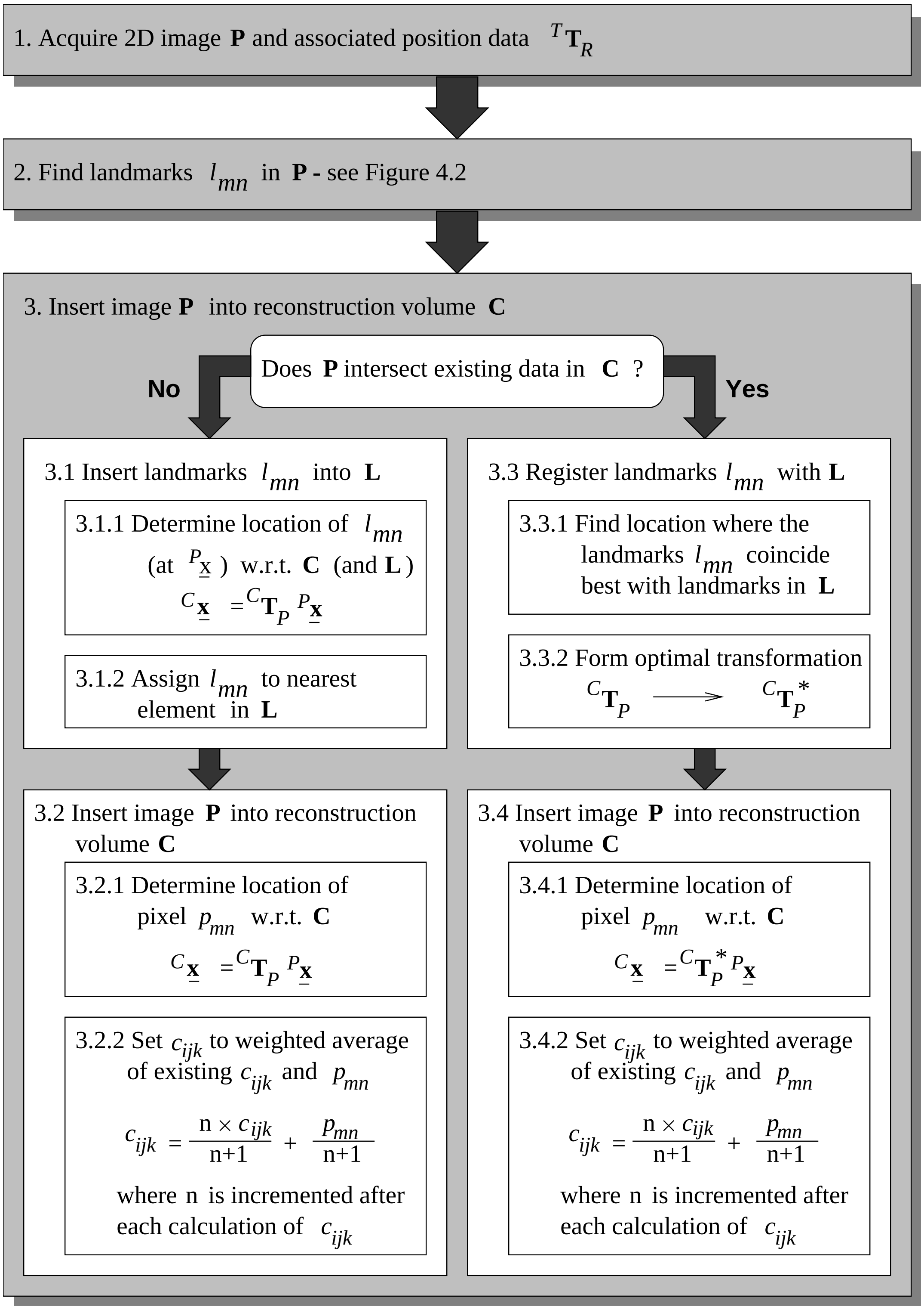
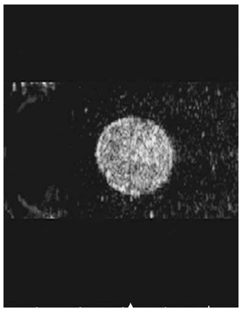


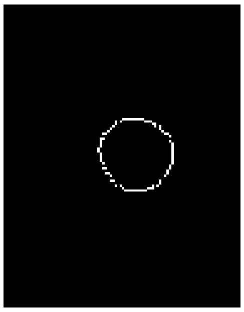
Figure 4.1: **PNN reconstruction with registration.** Steps 1 to 3 are repeated for all B-scans acquired in an examination.

Parameter	Organ Phantom
Canny kernel size (mm)	5
Landmark search radius (mm)	6
Consensus set threshold (%)	40

Table 4.1: **Registration algorithm parameters.** The size of the kernel used in the Canny convolution represents the best trade-off between the accuracy of boundary localisation and sensitivity to speckle noise. The radius of the landmark search volume is related to the expected levels of registration error. The consensus set threshold is specified as a proportion of the total number of landmarks in each B-scan. Its value is based on an estimated ratio of the number of true landmarks to the total number of landmarks in a typical B-scan.



(a) High resolution image



(b) Landmarks

Figure 4.2: **Landmark detection.** (a) shows the original high resolution image of the organ phantom. Notice the speckling throughout the image. (b) depicts the edgels extracted from image (a) after chaining and pruning.

4.2.2 Detection of Landmarks

The landmarks used in this study are edgels (edge elements) automatically extracted by the Canny edge detection algorithm (Canny, 1986). The edgels are produced at the resolution corresponding to the voxel size chosen for the reconstruction, so the resolution of the B-scans is reduced correspondingly before edge detection. The kernel size in the Canny convolution is chosen large enough to minimise the effect of speckle, but small enough to give good location accuracy. The size of the kernel is given in Table 4.1.

The edgel set is further pruned by chaining neighbouring edgels together and eliminating short chains. Chains with fewer than three elements are eliminated because speckle typically produces one or two element chains, while structural boundaries produce longer chains. Deleting short chains therefore reduces the number of false landmarks. Figure 4.2 shows an example of the detected landmarks.

4.2.3 Selection of Correspondences

As the B-scans are acquired, landmarks are stored in a 3D binary array $\mathbf{L}$ which is aligned with $\mathbf{C}$. The landmark array $\mathbf{L}$ is binary because only the presence of each edgel is stored; the orientation and strength are discarded since they are not required for alignment. When a B-scan intersects non-empty (previously filled) voxels in $\mathbf{C}$, ${}^{C}\mathbf{T}^{*}_{P}$ is determined by finding correspondences between landmarks in the new B-scan and landmarks in $\mathbf{L}$.

A minimum number of intersections between pixels in the scan plane and non-empty voxels in $\mathbf{C}$ is needed for accurate registration. In the phantom study, landmark registration is attempted only when more than 25% of the pixels in the B-scan intersect previously filled voxels in $\mathbf{C}$ [1].

Potential correspondences between landmarks in the B-scan and landmarks in $\mathbf{L}$ are found by searching a spherical volume in $\mathbf{L}$ for each landmark in the B-scan. Since the B-scans are acquired in rapid succession, registration errors are expected to vary slowly from one B-scan to the next. This observation is used to place the spherical search volume at an appropriate location in the voxel array, and also limit its size — see Figure 4.3. The registration algorithm effectively tracks the registration error, which is far more efficient than performing an unconstrained search for ${}^{C}\mathbf{T}^{*}_{P}$ for each B-scan. The size of the search volume is given in Table 4.1.

4.2.4 Determination of Transformation Matrix and Compounding

The set of correspondences produces more constraints than are required to determine ${}^{C}\mathbf{T}^{*}_{P}$. For example, each B-scan of the phantom generates approximately 1200 candidate correspondences, but only three are needed to determine ${}^{C}\mathbf{T}^{*}_{P}$. A least squares estimation of ${}^{C}\mathbf{T}^{*}_{P}$ is inappropriate, since many of the candidate correspondences are outliers. Instead, the RANSAC (random sample consensus) regression technique (Fischler and Bolles, 1981) is used to determine ${}^{C}\mathbf{T}^{*}_{P}$ and reject the erroneous correspondences. RANSAC is sufficiently robust to tolerate a significant proportion of erroneous correspondences. The steps of the RANSAC regression are listed in Figure 4.4.

${}^{C}\mathbf{T}^{*}_{P}$ is used to transform the new B-scan image $\mathbf{P}$ into the coordinates of the voxel array $\mathbf{C}$, then the pixels of $\mathbf{P}$ are added to $\mathbf{C}$. Where the B-scan intersects previously filled voxels, weighted averaging is used to compound the new image with the existing data (see step 3.4.2 in Figure 4.1).

4.3 Method

4.3.1 Acquisition System

The acquisition system consisted of a Toshiba model SSA-270A/HG ultrasound machine, a 7 MHz linear array probe, and a Polhemus Fastrak position sensor mounted on the probe.

[1] In practice, after the first sweep in the ultrasound examination, all subsequent B-scans intersect a sufficient amount of baseline data to perform registration.

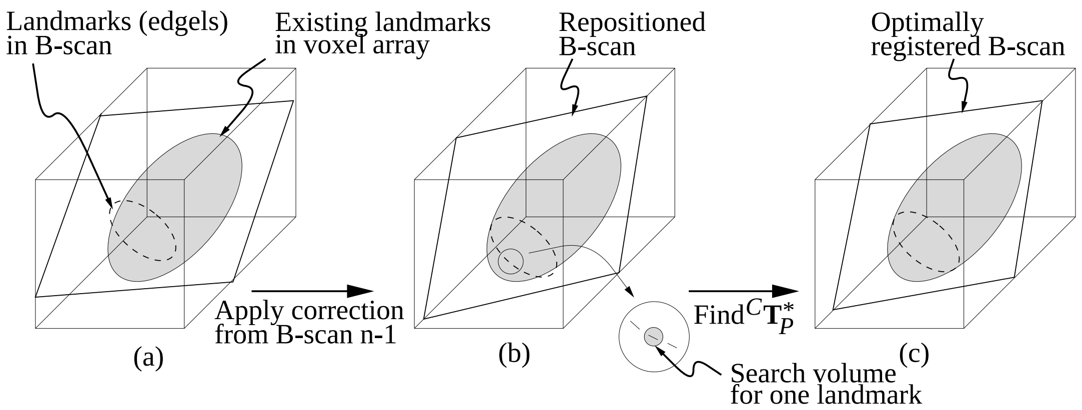


Figure 4.3: **Landmark-based registration.** In (a), B-scan n is inserted into the voxel array at the location indicated by the position sensor. Significant registration errors are evident (exaggerated here for clarity). After applying the rigid body correction found from registering B-scan $n-1$, only small residual errors remain (b). These are corrected by landmark-based registration. A small spherical search volume is defined around each repositioned landmark (edgel) in $\mathbf{P}$: any landmarks in $\mathbf{L}$ found within this volume are marked as candidate correspondences for the landmark in $\mathbf{P}$. RANSAC regression is used to find the optimal rigid body transformation ${}^{C}\mathbf{T}_{P}^{*}$ which aligns as many of the corresponding landmarks as possible — see Figure 4.4. Finally, the B-scan is inserted into its optimal position in the voxel array using ${}^{C}\mathbf{T}_{P}^{*}$ (c).

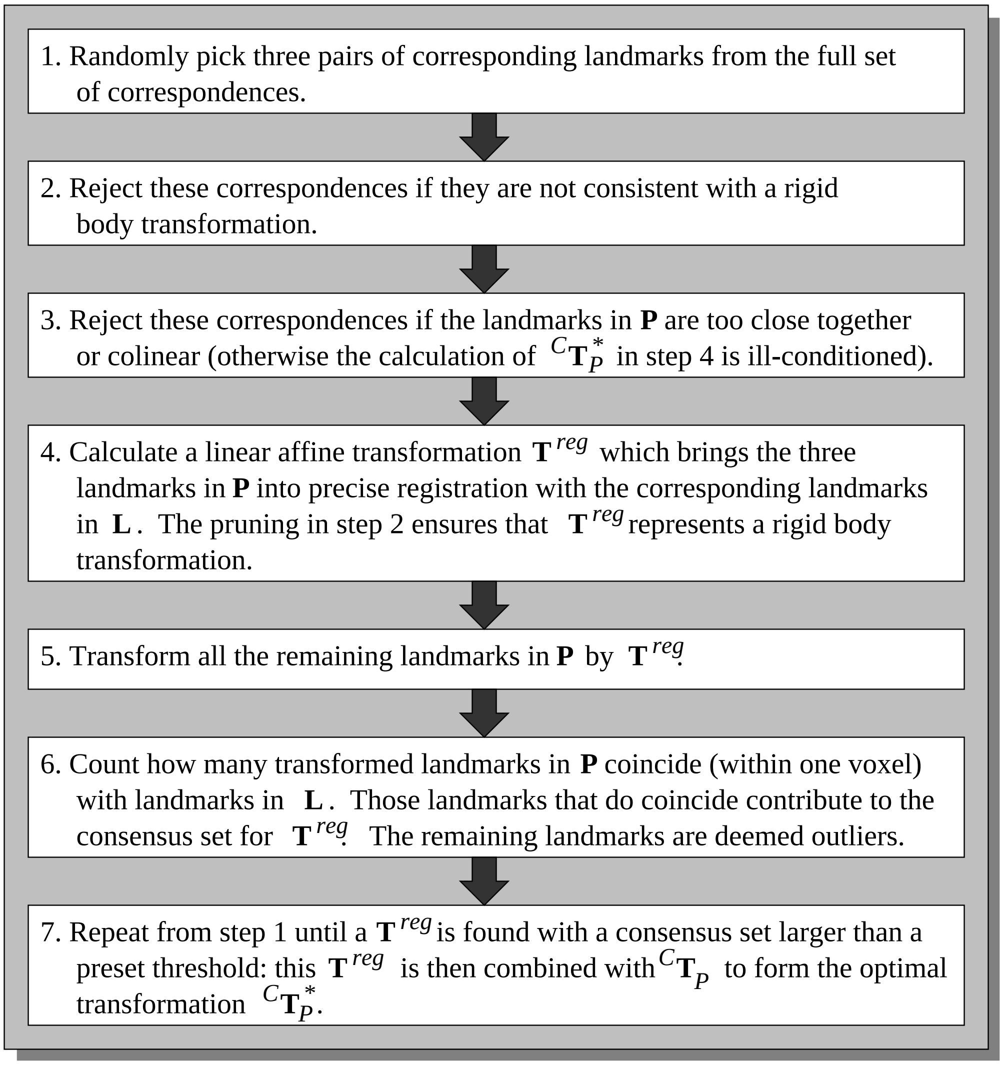


Figure 4.4: **RANSAC regression for determining ${}^{C}\mathbf{T}_{P}^{*}$.** The consensus set threshold is given in Table 4.1.

To minimise interference from the electronics inside the ultrasound probe, the receiver was mounted at the end of a rigid nylon extension bar about 25 cm from the probe face. Images were acquired with an 8-bit frame grabber and a Sun SparcStation 10 workstation (Sun Microsystems, Inc., Palo Alto, California). A shallow depth setting produced images with a pixel size of 0.14 mm. The images were then recorded at a rate of 5 fps, and later cropped off-line to 220×380 pixels.

Because this study was performed before the work on calibration, none of the calibration phantoms were yet available. The calibration parameters were therefore obtained in a simpler manner. The offsets to the receiver were measured with calipers, and the scale values were calculated from the graduations of the axes displayed in the uncropped B-scans. The phantom was mounted in a bath of water at a temperature of approximately 40° Celsius so that the speed of ultrasound in the water approximated the speed in human tissue (the speed assumed by the ultrasound machine).

Performing calibration with calipers is not as accurate as calibration with the phantoms described in Chapter 2. It may appear, therefore, that the registration errors arise solely from miscalibration, but the study in Chapter 5 used the more accurate three-wire calibration method and still encountered significant registration errors. This implies that the errors cannot be attributed only to miscalibration and, even with accurate calibration, image-based registration is required.

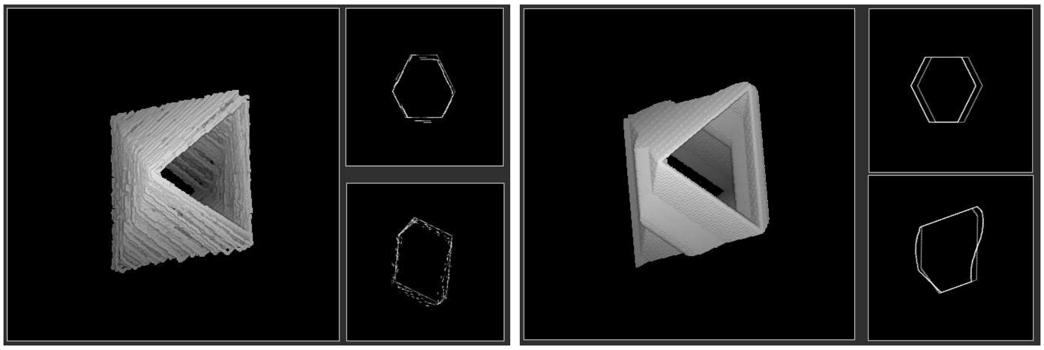
4.3.2 Simulations

Simulations were first performed on a computer generated octahedron. The octahedron shape was chosen because the boundary of the octahedron in each cross-section is unique within a local neighbourhood. In other words, each cross-section makes a perfect fit with the octohedron in only one nearby location, making it easy to verify the results of registration.

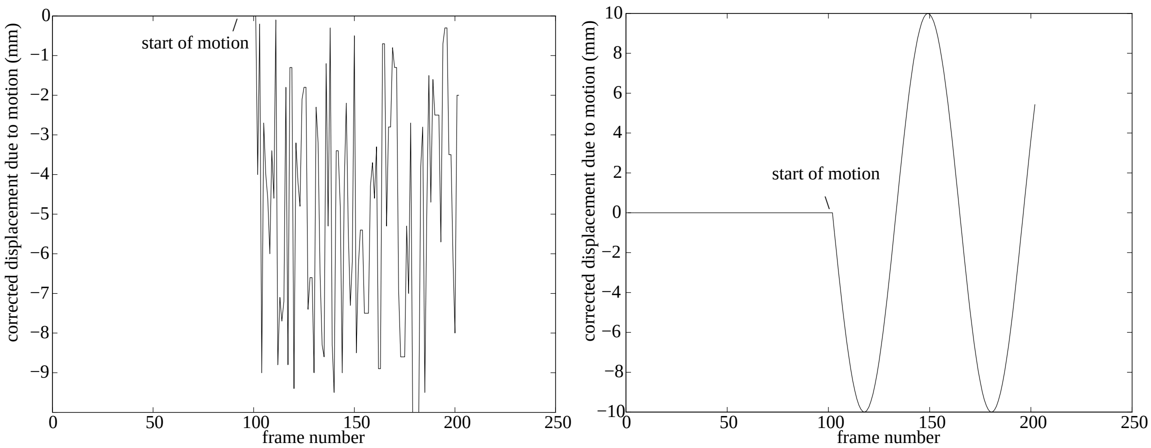
A simulated examination of the octohedron test object was performed from left-to-right, then right-to-left. Errors were added to the right-to-left portion of the examination to simulate motion of the object. Figure 4.5 shows that the algorithm can compensate for different types of motion errors and track the motion accurately. Errors with only one DOF are shown in Figure 4.5 for clarity, but errors in all 6 DOF were also tracked and corrected.

4.3.3 Experiments

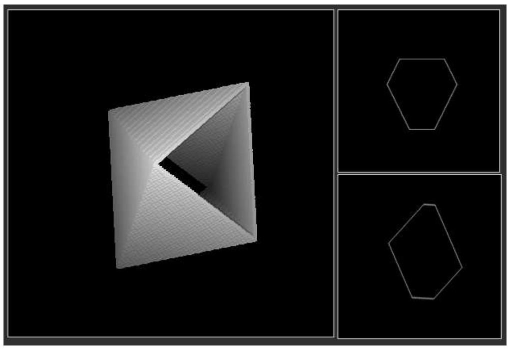
Tests were performed on a phantom comprising an egg-shaped latex balloon filled with a combination of water, ultrasound coupling gel and talcum powder. This type of phantom was used because B-scans of its cross-section produce images of a uniformly speckled interior, a sharp boundary, and a uniformly speckled exterior with a lower mean grey level — see Figure 4.6. Grey-level statistics could therefore be measured in two regions with almost homogeneous statistical features, and the effect of spatial compounding on the SNR easily observed.



(a) Random motion errors (b) Sinusoidal motion errors

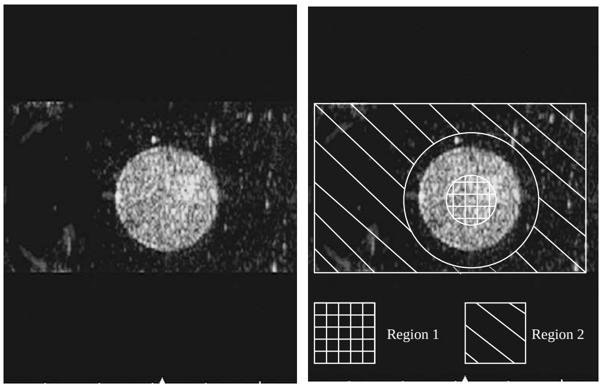


(c) Corrected motion errors (d) Corrected motion errors



(e) Octahedron reconstructed with motion correction

Figure 4.5: **Registration of simulated data.** (a) and (b) show surface renderings of the reconstruction (without registration) and two typical cross-sections. Small random motion errors were added to (a), and small sinusoidal motion errors were added to (b). (c) and (d) show the motion errors in (a) and (b) that were corrected by the registration algorithm. (e) shows the result of reconstruction with registration.



(a) Example B-scan (b) Regions

Figure 4.6: **Regions of the phantom B-scans.** To investigate the effect of compounding on the image statistics, two regions are defined: region 1 and region 2. Region 1 lies completely inside the latex balloon and region 2 on the outside. Similar volumetric regions are defined for the 3D reconstruction volume.

The phantom was scanned in a continuous series of sweeps from one end of the balloon to the other, producing a large number of overlapping B-scans. Each sweep was carried out with the probe at a look direction slightly displaced from the previous sweep. About four and a half sweeps were performed, producing 400 B-scans in total. The first 100 B-scans corresponded to the first complete sweep, and therefore did not overlap each other.

The PNN algorithm was used to reconstruct volumes from the 400 B-scans at both high and low resolution. The high resolution reconstructions contained voxels with the same size as the pixels in the B-scans (0.14mm). The low resolution reconstructions contained voxels seven times the pixel size (1mm). While the high resolution reconstructions preserved the full resolution of the B-scans, they required a large amount of memory (181 MBytes). Plus, since the PNN method was used without the hole-filling stage, many unfilled voxels remained in the gaps between B-scans. Low resolution reconstructions (0.5 MBytes with 1 mm voxels) were generated in significantly less time and contained no gaps. For each resolution, two reconstruction volumes were created, one with registration (using the PNN reconstruction algorithm with registration in Figure 4.1), the other without (using the standard PNN reconstruction algorithm in Figure 2.2). Four volumes were therefore reconstructed in total.

Registration was, in fact, only performed at low resolution, with the results applied to the high resolution reconstructions. Registration at high resolution was infeasible because of the large number of unfilled voxels and the considerable memory requirements of the reconstruction algorithm.

To investigate the improvement in SNR with spatial compounding, the grey-level statistics were analysed in each of the four reconstruction volumes at various stages of reconstruction: after 100 B-scans (initial sweep, no compounding), then 150, 200, 250, 300, 350

No. of	Region 1			Region 2		
B-scans	mean	std. dev.	SNR	mean	std. dev.	SNR
100	103.76	18.37	5.65	33.68	5.76	5.85
150	102.45	18.02	5.68	33.60	5.49	6.12
200	103.09	17.65	5.84	33.39	5.14	6.50
250	103.33	17.03	6.07	32.23	4.77	6.97
300	102.80	16.55	6.21	33.17	4.66	7.12
350	102.30	16.00	6.39	33.07	4.45	7.43
400	101.41	15.74	6.44	33.01	4.34	7.61

Table 4.2: **High resolution reconstruction *without* registration.**

and finally 400 B-scans (heavy compounding). For each case, volumetric regions 1 and 2 were segmented, and then the mean and standard deviation (and therefore the SNR — the ratio of the mean to the standard deviation) were calculated for filled voxels in each region. The unfilled voxels were not included in the calculations. Segmentation was performed by manually constructing volumetric masks for regions 1 and 2. Once each mask was created, it was used for all seven stages of reconstruction (100 to 400 B-scans). In this way, changes in the observed SNR cannot be attributed to changes in segmentation.

4.4 Results

4.4.1 Signal to Noise Ratio

It is immediately apparent that the SNR increases with the amount of compounding — see Tables 4.2, 4.3, 4.4, and 4.5. Furthermore, the SNR improves almost identically for both the registered and unregistered cases. This is because regions 1 and 2 do not include the area where the phantom boundary is blurred by the registration errors.

The improvement in SNR can be predicted by statistical theory. Previous papers (Burckhardt, 1978; Hernandez et al., 1996; Kerr et al., 1986) have demonstrated a $\sqrt{n}$ improvement in SNR for 2D compounding of n uncorrelated B-scans [2]. This theory is not directly applicable to the 3D case, since the voxels are not all compounded the same number of times. The arbitrary positions and orientations of the B-scans result in some voxels being intersected more than others. Furthermore, the statistical theory for low resolution compounding must account for both the compounding caused by the intersections of multiple B-scans as well as the reduction in resolution (when the voxel size is chosen larger than the pixel size). For these reasons, a full theory of 3D spatial compounding is developed and presented in Appendix D.

The measured SNR of both the low and high resolution registered reconstructions is

[2] This assumes that the speckle is uncorrelated across the B-scans. Compounding B-scans from very similar look directions will result in a smaller improvement in the SNR because the speckle will be partially correlated.

No. of	Region 1			Region 2		
B-scans	mean	std. dev.	SNR	mean	std. dev.	SNR
100	103.76	18.37	5.65	33.68	5.76	5.85
150	102.45	18.02	5.68	33.60	5.49	6.12
200	103.09	17.65	5.84	33.39	5.14	6.50
250	102.99	16.77	6.14	33.24	4.77	6.97
300	103.28	16.19	6.38	33.16	4.62	7.18
350	103.37	15.73	6.57	33.06	4.39	7.53
400	103.24	15.35	6.72	32.99	4.26	7.74

Table 4.3: **High resolution reconstruction *with* registration.**

No. of	Region 1			Region 2			Vol.
B-scans	mean	std. dev.	SNR	mean	std. dev.	SNR	(ml)
100	100.39	11.07	9.07	33.22	3.78	8.79	7.20
150	100.09	10.63	9.42	33.15	3.41	9.72	7.17
200	99.94	9.77	10.23	33.07	3.07	10.77	7.19
250	100.05	9.27	10.79	33.00	2.88	11.46	7.27
300	99.93	8.67	11.53	32.97	2.86	11.53	7.31
350	99.71	8.16	12.22	32.93	2.82	11.68	7.43
400	99.44	7.81	12.73	32.90	2.80	11.75	7.56

Table 4.4: **Low resolution reconstruction *without* registration.** The volume of the phantom, as estimated by semi-automatic segmentation, increases with the number of B-scans. This is because the registration errors blur the boundary of the phantom, affecting the semi-automatic segmentation.

No. of	Region 1			Region 2			Vol.
B-scans	mean	std. dev.	SNR	mean	std. dev.	SNR	(ml)
100	100.39	11.07	9.07	33.22	3.78	8.79	7.20
150	100.09	10.63	9.42	33.15	3.41	9.72	7.16
200	99.94	9.77	10.23	33.07	3.07	10.77	7.19
250	99.91	9.09	10.99	33.00	2.84	11.62	7.20
300	100.02	8.13	12.30	32.96	2.78	11.86	7.14
350	100.06	7.76	12.89	32.91	2.68	12.28	7.20
400	100.17	7.27	13.78	32.86	2.57	12.79	7.18

Table 4.5: **Low resolution reconstruction *with* registration.** The volume of the phantom, as estimated by semi-automatic segmentation, remains almost constant with higher levels of compounding.

plotted against the theoretical increase in SNR in Figure 4.7. As the theory predicts, the SNR increases with increasing levels of compounding. Yet there are a number of differences between the idealised theory and the actual empirical study. The first is that no attempt was made to obtain completely uncorrelated B-scans for different sweeps. As mentioned in Section 1.6, the look directions for different sweeps must be separated by more than forty percent of the transducer width to obtain completely uncorrelated speckle patterns (Trahey et al., 1986b). To verify this, the first few sweeps were deliberately performed from similar look directions, saving large variations in the look direction for the last few sweeps. The results in Figure 4.7 confirm the theory: the SNR improves most dramatically for the last few sweeps, when the speckle patterns are less correlated. For this reason, the last four data points in Figure 4.7 are highlighted for comparison with the predicted values (derived assuming completely uncorrelated speckle).

The grey-level variations in the B-scans are also not caused entirely by speckle. The echo amplitudes reflected by a homogeneous medium have an expected Rayleigh distribution when there is a large number of scatterers per resolution cell. The SNR of a Rayleigh distribution is a constant, equal to 1.91 (Burckhardt, 1978). The logarithmic compression of the echo amplitude in the formation of B-scans changes the Rayleigh distribution to an approximately Gaussian distribution, and the SNR changes to 7.64 (Crawford et al., 1993; Hernandez et al., 1996; Thijssen et al., 1988). The measured SNR of a typical B-scan in this study is 5.6 for region 1 and 5.4 for region 2. The lower SNR means that the observed grey-level variations are greater than the variations produced by speckle. If some of the variation is a result of real physical structure (such as clumping of the talcum powder suspension), then it will be correlated across different sweeps and cannot be reduced by compounding. Despite these differences from the assumptions underpinning the statistical theory, the slopes of the curves in Figure 4.7 are in remarkably close agreement with the theoretical predictions.

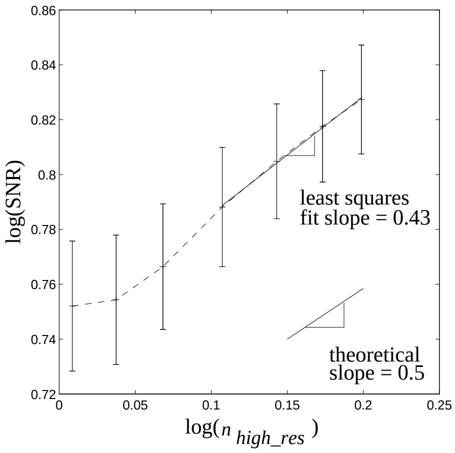
4.4.2 Visualisation

Cross-sectional slices through the high resolution reconstructions are shown in Figure 4.8, both with and without registration. It is evident that the registration errors are large enough to significantly distort the reconstruction, but are dramatically reduced by landmark-based registration. The effect of compounding is not pronounced at high resolution, because each voxel is intersected only a small number of times.

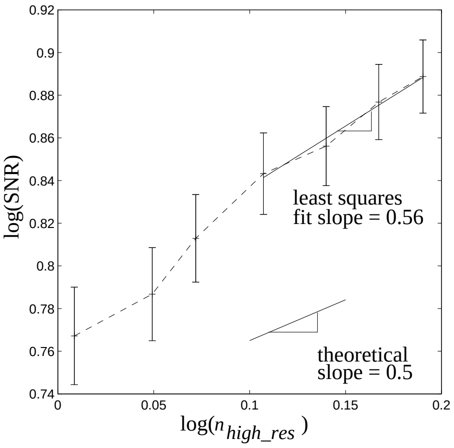
Figures 4.9 and 4.10 show the effects of registration and compounding at low resolution. In both figures the speckle (and other artifacts) are greatly reduced by compounding, but the shape of the phantom in the unregistered case departs significantly from the original shape. The registered reconstruction maintains the original shape.

Surface rendering is performed with the semiboundary approach described in Section 1.4, using segmentation by grey-level thresholding. Figure 4.11 shows how compounding improves a surface rendering of the reconstruction volume.

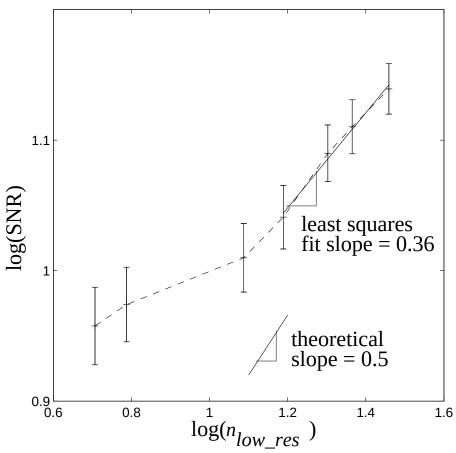
The grey-level histograms of regions 1 and 2, shown in Figure 4.12, change as the level



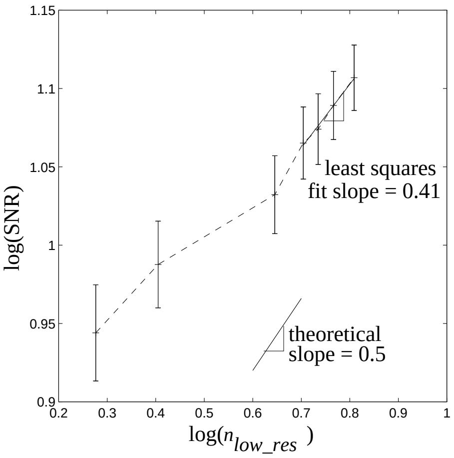
(a) Region 1 at high resolution



(b) Region 2 at high resolution

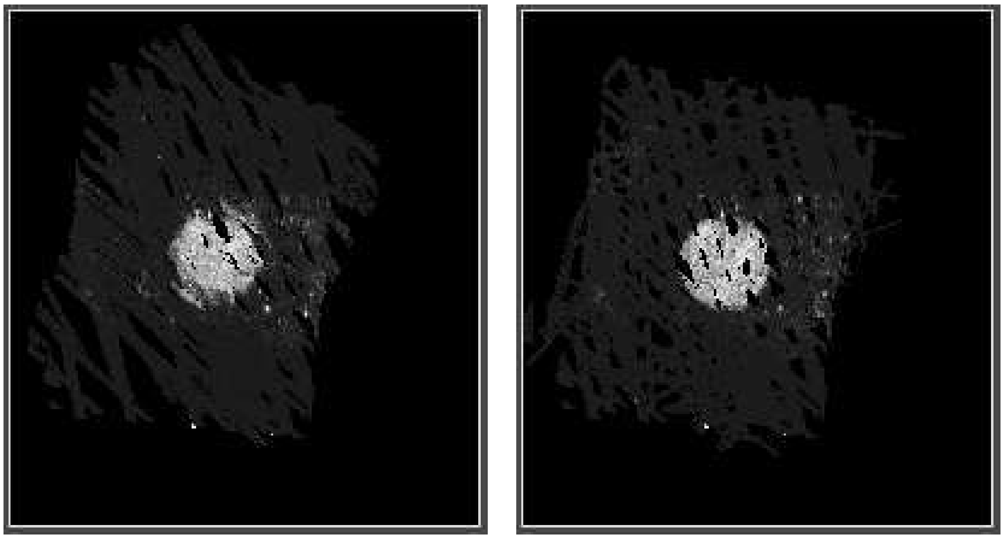


(c) Region 1 at low resolution



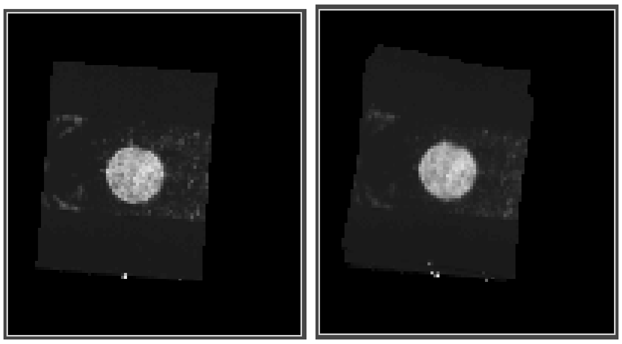
(d) Region 2 at low resolution

Figure 4.7: **Effect of compounding on SNR for registered reconstructions.** Both high resolution (a, b) and low resolution (c, d) reconstructions show an increase in the SNR as the level of compounding increases. Each '+' data point represents the SNR calculated at 100, 150, 200, 250, 300, 350, and 400 B-scans. The error bars were estimated by measuring the RMS variation in the SNR obtained using slightly different volumetric masks. The x-axis indices, n_{high_res} and n_{low_res}, are defined in Appendix D. A log-log plot is used because a slope of 0.5 indicates agreement with the theoretical increase in SNR as $n^{0.5}$. As explained in Section 4.4.1, only the final four data points in each plot can be meaningfully compared to the theory. In all cases, these data points are in remarkably good agreement with the theory.

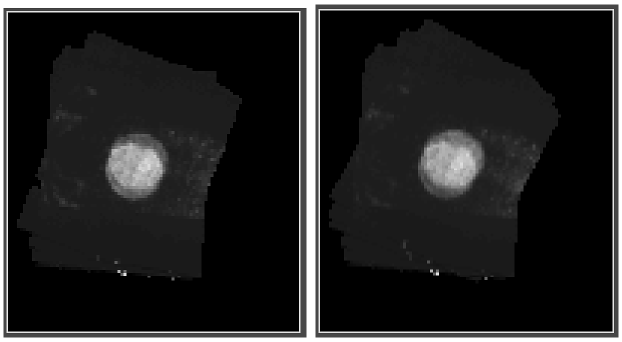


(a) Unregistered (b) Registered

Figure 4.8: **Slices of high resolution reconstructions.** A slice of the reconstruction volume is shown for both the unregistered and registered cases. All 400 B-scans are used in both reconstructions, but gaps still remain. In (a), the registration errors substantially distort the circular cross-section of the phantom. The circular shape is restored in (b) by landmark registration.

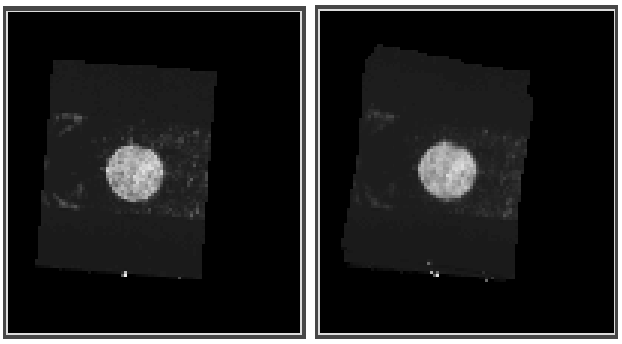


(a) 100 B-scans (b) 200 B-scans

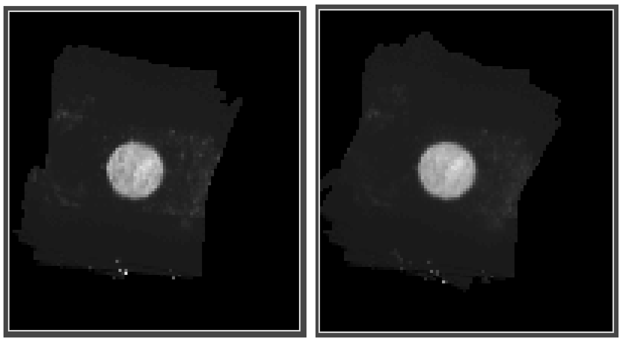


(c) 300 B-scans (d) 400 B-scans

Figure 4.9: **Slices of low resolution reconstructions *without* registration.** All slices are taken at the same location in the reconstruction volume. Speckle is reduced by compounding both inside and outside the object, but registration errors result in substantial blurring of the object boundary.

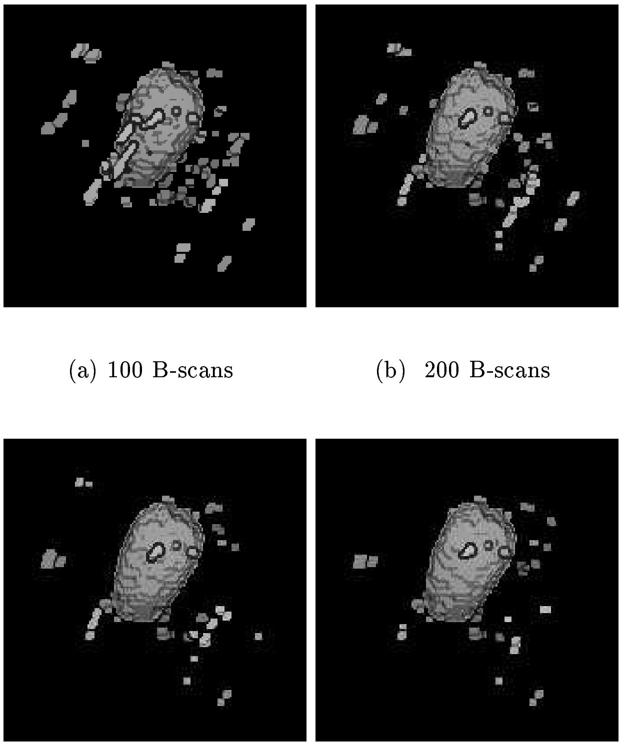


(a) 100 B-scans (b) 200 B-scans



(c) 300 B-scans (d) 400 B-scans

Figure 4.10: **Slices of low resolution reconstructions *with* registration.** All slices are taken at the same location in the reconstruction volume. Speckle is reduced by compounding both inside and outside the object and blurring of the object boundary is minimal.



(a) 100 B-scans (b) 200 B-scans

(c) 300 B-scans (d) 400 B-scans

Figure 4.11: **Surface renderings of low resolution reconstructions *with* registration.** The 3D surface renderings are produced by grey-level thresholding. The amount of speckle outside the object is reduced with increasing levels of compounding. (a) shows the reconstruction after a single sweep. The object is less obscured by speckle when the reconstruction volume is heavily compounded in (d).

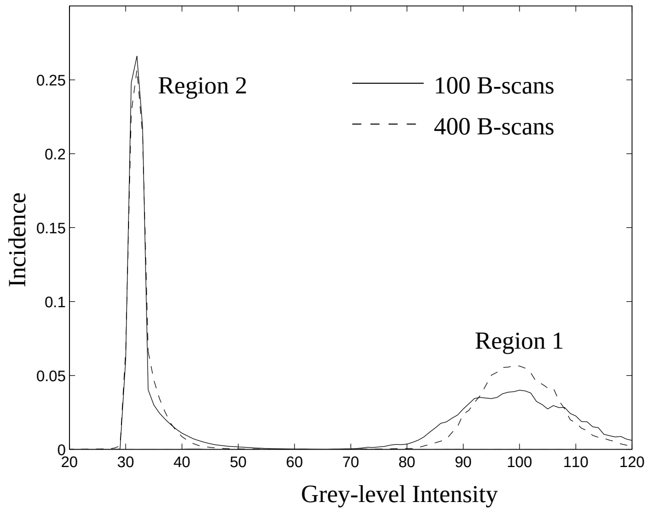


Figure 4.12: **Histogram of grey levels in the reconstruction volume *with* registration.** The two regions are described by individual distributions. As compounding increases, the mean grey level remains constant but the standard deviation decreases. This effect is particularly visible for the distribution of region 1.

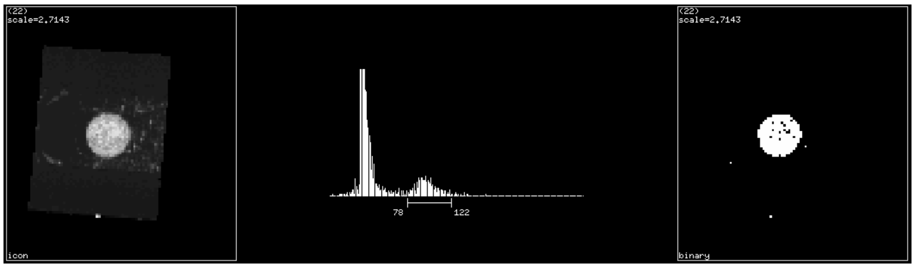
of compounding increases. As expected, the histograms become narrower with higher levels of compounding. If the histograms of the different regions are non-overlapping, then fully automatic segmentation can be achieved by thresholding the reconstruction at the appropriate grey level. Figure 4.13 shows several examples of fully automatic segmentation by grey-level thresholding. One of the motivations for improving SNR is to allow more accurate automatic segmentation. It is evident that compounding with registration improves the accuracy of automatic grey-level segmentation.

4.4.3 Volume Estimation

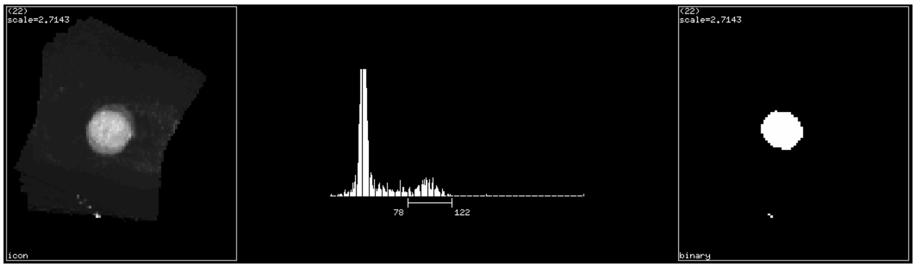
A measure is required to quantify the effect of registration errors on the reconstruction. The volume of the phantom was chosen as the measure for these experiments, since volume is often sought after by physicians when scanning internal organs (Chervenak et al., 1993). It can also be compared to the real volume of the phantom, which was measured with a graduated cylinder to be 7.0 ml $\pm$ 0.2 ml.

Volumes were estimated by "live-wire" segmentation (Barrett and Mortensen, 1997) of slices through the reconstructed volume. Live-wire segmentation is a powerful tool for extracting boundaries in noisy images. It offers a good compromise between accuracy and amount of user intervention. The technique involves laying an active wire (on a slice by slice basis) that is attracted automatically to the object's boundary. The operator assists the live wire by depositing small sections at a time near the boundary, so that the wire does not enclose nearby speckle. For these tests, four sections of wire were sufficient to enclose the phantom cross-section accurately.

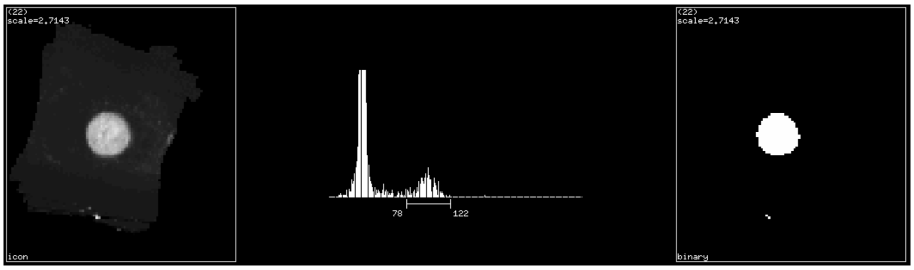
The semi-automatic live-wire technique was used instead of grey-level threshold seg-



(a) Reconstruction using 100 B-scans. The area of the thresholded image on the right is 243.9 mm^2. Live-wire semi-automatic segmentation of the image produces an area estimate of 256.5 mm^2. The difference arises mainly from speckle in the interior that falls outside the threshold range.



(b) Unregistered reconstruction using 400 B-scans. The area of the thresholded image on the right is 217.7 mm^2. The area is lower than in (a) because the blurred regions near the phantom boundary fall outside the threshold range.



(c) Registered reconstruction using 400 B-scans. The area of the thresholded image on the right is 250.9 mm^2. This is close to the area calculated by live-wire segmentation in (a).

Figure 4.13: **Threshold segmentation of low resolution reconstructions.** Corresponding slices through the reconstruction volumes are shown on the left of each figure. For each slice, the histogram of the grey levels is shown in the middle. The images on the right show the regions segmented by thresholding the grey levels falling into the range [78, 122]. The grey-level segmentation in (c) is more accurate than in (a) or (b). The compounding in (c) ensures that the entire object falls within the threshold range and the registration retains the original circular shape of the cross-section.

mentation, since thresholding is incapable of completely separating the phantom from the surrounding noise — see Figure 4.11. Some user intervention is required to avoid incorporating the surrounding noise and, therefore, over-estimating the phantom's volume.

The volume of the phantom was calculated by live-wire segmentation at seven levels of compounding: after 100, 150, 200, 250, 300, 350, and 400 B-scans. The resulting volume estimates are given in Tables 4.4 and 4.5 for the low resolution reconstructions. Live-wire segmentation is not feasible at high resolution, where there are too many gaps around the boundary.

Without registration, the blurring of the phantom boundary results in an increase in the segmented volume (range = [7.17, 7.56]) that does not reflect the true volume. Conversely, the volume calculated from the registered reconstructions is stable (range = [7.14, 7.20]). The increase in volume by blurring is less dramatic than expected from looking at slices through the reconstruction (Figure 4.9). This is because the live wire is attracted to the location with the highest intensity gradient, which was sometimes near the true boundary and not the edge of the blurred region. It is fair to say that considerably worse volume estimates could be expected under less favourable circumstances.

4.5 Conclusions

This chapter has shown how spatial compounding can improve the SNR of 3D ultrasound reconstructions. The improvement in SNR agreed with the theoretical predictions for both the high and low resolution cases. The inevitable registration errors were corrected using an automatic, incremental landmark-based registration algorithm. The resulting high quality 3D reconstructions were particularly well suited to automatic segmentation for visualisation and volume measurement.

At present, the reconstruction algorithm only performs registration with rigid body transformations. This study showed that the majority of registration errors *in vitro* could be corrected this way. Although the current reconstruction algorithm was sufficiently robust to tolerate small non-rigid errors, these errors remained in the final compounded reconstruction. Yet the level of blurring that resulted from these residual errors was small. The next step is to investigate spatial compounding *in vivo*.

The whole is more than the sum of the parts.
—Aristotle, **Metaphysica**, 330 B.C.

Chapter 5

Spatial Compounding *In vivo*

5.1 Introduction

The previous chapter has demonstrated the benefits of 3D spatial compounding on an artificial object. In this chapter, spatial compounding is applied to an examination of a human patient. The gall bladder is chosen for these tests because it is easy to identify and small enough to be encompassed in a single sweep. The main goals are still improvements in visualisation and volume estimation.

To reiterate the discussion in Chapter 1 on measurement errors, there are a number of properties of ultrasound data that make visualisation and volumetric data analysis difficult (Sakas and Stefan, 1995). Ultrasound images often have:

- significant noise and speckle (speckle arises from the constructive-destructive interference of the coherent ultrasound pulses);
- lower dynamic range than MR and CT images;
- blurred boundaries around anatomical features;
- boundaries with varying intensities caused by changes in surface curvature and orientation;
- partially or completely shadowed surfaces from objects nearer the probe;
- variable resolution through the volume, dependent on both the spacing between the B-scans and the location within a B-scan.

Several of the items in the list above are apparent in B-scans of the gall bladder — see Figure 5.1.

In response to these problems, considerable effort has gone into the development of new segmentation and visualisation techniques for ultrasound data. Although some success has been achieved for special cases such as fetal imaging (Sakas and Stefan, 1995), fast and fully automatic generic techniques remain elusive. In this chapter, the aim is to exploit

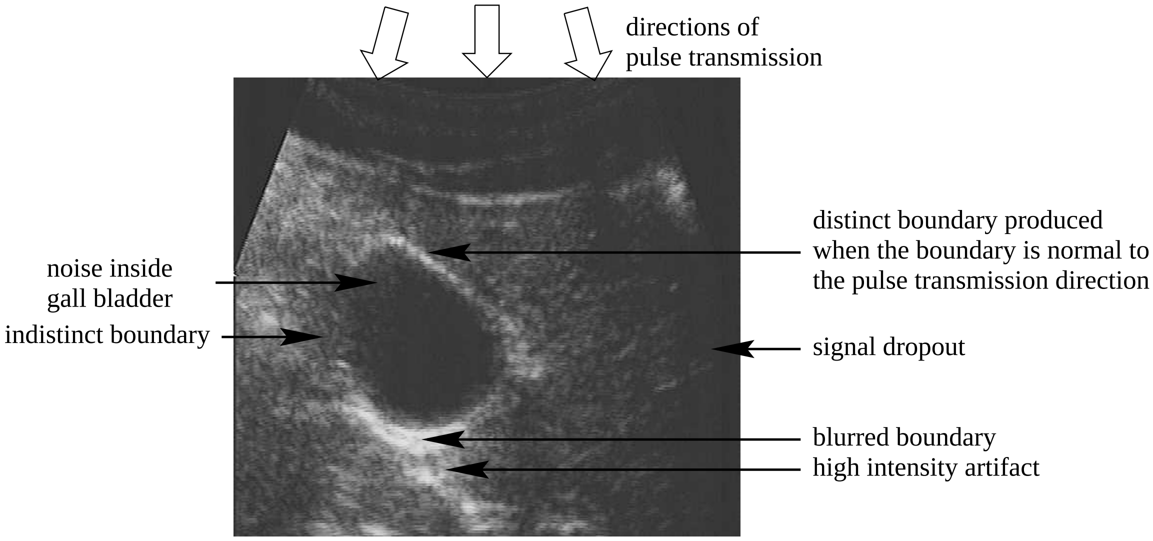


Figure 5.1: **Typical B-scan of a human gall bladder.** The gall bladder is the dark circular region just left of centre. Most of the properties of ultrasound that inhibit automatic analysis are present in this image. In particular, the speckle phenomenon is visible as a granular texture throughout the image. The intensity of the gall bladder boundary varies around its length. Sections of the boundary that are perpendicular to the pulse transmission direction produce high intensity echoes, and sections that are parallel produce low intensity echoes.

spatial compounding to improve the quality of the 3D ultrasound data, so that existing visualisation and analysis techniques are more successful.

The previous chapter has already demonstrated that the key to spatial compounding is accurate registration of image features. This chapter introduces a new, more robust registration technique for experiments on a human patient. These experiments aim to demonstrate that image-based registration is feasible for data acquired *in vivo*.

5.2 Reconstruction With Registration

5.2.1 Overview

In a typical 3D ultrasound examination, the physician scans smoothly and continuously a region of interest from one extent to the other. Each sweep of the region can be used to reconstruct a 3D data set. The look direction of the probe (*i.e.* location of the acoustic window) is changed slightly between each sweep so that different views of the region are obtained.

As in the previous chapter, spatial compounding is performed by accurately registering the subsequent sweeps to the first baseline sweep. The results of the previous chapter also demonstrated that the majority of registration errors *in vitro* could be corrected by a six DOF rigid body registration. It is hoped that rigid body registration will also correct the majority of registration errors *in vivo*. Yet instead of registering 2D B-scans to a 3D baseline data set, here entire 3D data sets are reconstructed (again using the standard PNN reconstruction method without hole-filling — see Section 2.2) and registered to a baseline 3D data set. In this way, the registration problem is simplified, because only a small number of registration procedures (the number of sweeps) are required, which reduces the computational expense. It is assumed that the majority of registration errors arise when the look direction is changed between sweeps, leaving relatively small errors within the sweeps. The registration process is illustrated in Figure 5.2.

There has been considerable research into the automatic registration of CT and MR images, but very few attempts to apply these techniques to 3D ultrasound. Comprehensive surveys of the various registration techniques can be found in (Taylor et al., 1996; van den Elsen et al., 1993). A number of the more popular approaches were evaluated on 3D ultrasound data, with varying degrees of success.

The first class of techniques that were tested are those based on landmark matching, as used for the phantom study of Chapter 4. The simplest of these is assisted point-based landmark matching. In this paradigm, the user identifies corresponding points in the two data sets, and then the transformation that brings one set of points into close alignment with the other is computed. The problem with this approach is the difficulty in finding easily distinguished points on smooth organ boundaries. Curves and surfaces can be used instead, and are extracted by either manual or automatic methods. Given suitable visual feedback, it is possible to perform the alignment of the curves or surfaces by eye. Automatic surface-surface matching (Levine et al., 1988) and automatic curve-curve

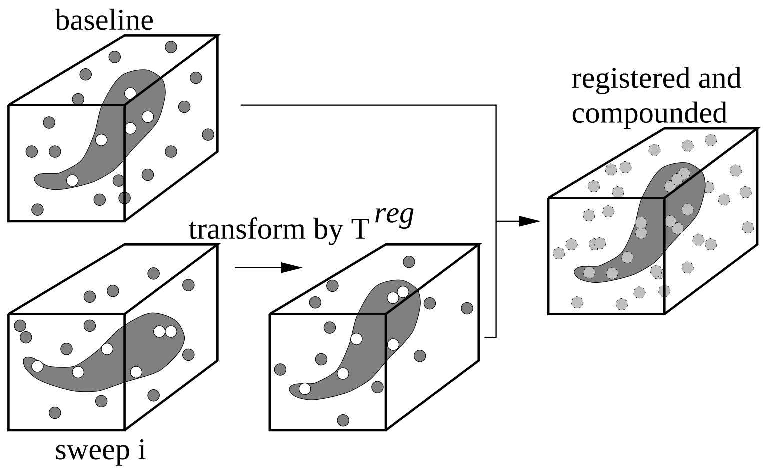


Figure 5.2: **Registration and spatial compounding of noisy 3D data.** The first sweep is called the baseline since all other sweeps are registered to it. $\mathbf{T}^{reg}$ is the homogeneous transformation matrix describing the rigid body registration. Transformation of sweep i by $\mathbf{T}^{reg}$ aligns it with the baseline data so that compounding can be performed with minimal loss of spatial detail. Since sweep i is reconstructed from B-scans taken from different look directions to the baseline B-scans, the noise is uncorrelated. Compounding the two registered data sets therefore reduces the level of noise.

matching (Thirion et al., 1992) are also feasible.

The main problems found with all of these landmark-based methods are accuracy and repeatability. Fully automatic extraction of curves and surfaces from ultrasound data is very difficult, as shown previously when segmenting the organ phantom in Chapter 4. Some manual intervention is inevitably required (the need to use the semi-automatic live-wire segmentation technique for the phantom study in Section 4.4.3 is a good example). With examinations performed *in vivo*, the data sets are even more difficult to segment, and the accuracy and repeatability of operator-assisted segmentation decreases. Human error also means that alignment by eye is not sufficiently accurate to produce sharp, compounded data.

The second broad class of registration techniques makes use of some sort of correlation measure. In other words, no user interaction, explicit segmentation or landmark identification is required: the measure is obtained directly from the original grey-level data. Such techniques have been found to be robust for multimodal (MR to CT) registration (Maintz et al., 1996; van den Elsen et al., 1995). High accuracy can be achieved at acceptable computational expense through a coarse-to-fine search, and it is possible to register 3D data sets that are only partially overlapping. Registering two 3D ultrasound data sets has much in common with registering an MR data set to a CT one. In both cases, the data sets to be aligned can appear substantially different, yet share some common features. For this reason, correlation-based registration of each sweep onto the baseline is chosen as the basis for this investigation.

5.2.2 Correlation-based Registration

Two volumes of intensity data can be registered by searching the transformation parameter space for a peak in the correlation function. In the current application, only rigid body transformations are considered, so the parameter space is six-dimensional (three translations and three rotations). For two volumes $\mathbf{L}_1$ and $\mathbf{L}_2$, the correlation coefficient is:

$$\mathrm{corr}(\mathbf{b}) = \sum_{(x,y,z)\in\mathbf{L}_1} \mathbf{L}_1(x,y,z)\mathbf{L}_2(\mathbf{T}^{reg}(x,y,z)) \tag{5.1}$$

where $\mathbf{b}$ is the vector of the registration parameters, and $\mathbf{T}^{reg}$ is the rigid body transformation matrix corresponding to $\mathbf{b}$.

It remains to decide what features ($\mathbf{L}$) of the data sets to correlate. Since different ultrasound images of the same anatomy can appear substantially different (as explained in Section 1.6), simple correlation of the raw intensity data is not likely to succeed. Again, multimodal MR-CT registration provides the inspiration. Raw MR and CT data sets cannot be compared directly, yet they share some common features that can be usefully correlated. For example, differential operators such as "edge" and "ridge" detectors can transform MR and CT scans of the skull into comparable data sets (Maintz et al., 1996).

In a comparative study of multimodal 3D image registration (Maintz et al., 1996), correlation of the magnitude of the 3D gradient was found to give the best performance. The 3D gradient magnitude is invariant under the group of orthogonal transformations (translation, rotation, and reflection), and tends to produce better results than Laplacian and ridgeness operators (Maintz et al., 1996). For this reason, registration of the ultrasound sweeps was attempted by correlating the 3D gradient magnitudes of the data sets. The gradients were calculated by convolution of the 3D data set with the derivative of a 3D Gaussian kernel. When applied to ultrasound data, the gradient magnitude operator produces a transformed image with local maxima near organ boundaries. The gradients were calculated efficiently using a separable, recursive algorithm (Monga et al., 1991).

The main drawback of all correlation-based registration techniques is the significant computational expense of the search over the six-dimensional parameter space. Yet considerable savings are possible using a multi-resolution approach (van den Elsen et al., 1995). A multi-resolution pyramid is constructed with the original data set at the base. Each higher level is produced at half the resolution of the previous level. The pyramid continues until the largest structures are no longer discernible. The basic idea is that wide range searches can be performed efficiently near the top of the pyramid and narrow but accurate searches at the bottom. Results from the searches at each level provide initial search locations at lower levels until the bottom of the pyramid is reached.

Since the position sensor readings provide a fairly good initial guess at the correct registration, only two pyramid levels were required for the gall bladder registration. The final reconstructions of the gall bladder used 1 mm voxels, so searches started with a resolution of 2 mm.

The correlation coefficient was sufficiently well behaved to allow two searches, with different step sizes, at each level of the pyramid. So two searches were performed at the 2 mm resolution, the first with large steps to cover a large part of the parameter search space, the second with smaller steps around the first correlation peak. The maximum of the 2 mm search was then used as the centre of the search at the 1 mm resolution, where again two searches were performed. Because the initial guess is close to the optimal registration, local minima were not encountered.

$\mathbf{T}^{reg}$ was constructed from the parameter values corresponding to the maximum correlation coefficient. In the same manner as in the previous chapter, the optimal transformation ${}^{C}\mathbf{T}_{P}^{*}$ was formed from the combination of the original ${}^{C}\mathbf{T}_{P}$ and $\mathbf{T}^{reg}$:

$$ {}^{C}\underline{\mathbf{x}} = {}^{C}\mathbf{T}_{P}\, {}^{P}\underline{\mathbf{x}} \quad \longrightarrow \quad {}^{C}\underline{\mathbf{x}} = {}^{C}\mathbf{T}_{P}^{*}\, {}^{P}\underline{\mathbf{x}} = \mathbf{T}^{reg}\, {}^{C}\mathbf{T}_{P}\, {}^{P}\underline{\mathbf{x}}. \tag{5.2} $$

To compound the registered sweep i with the baseline data, the position of each B-scan of sweep i was calculated using ${}^{C}\mathbf{T}_{P}^{*}$ then added to $\mathbf{C}$. (see step 3.4.2 in Figure 4.1). An overview of the compounded reconstruction algorithm, using the correlation-based registration method, is given in Figure 5.3.

5.3 Method

5.3.1 Acquisition System

The acquisition system consisted of a Toshiba model SSA-270A/HG ultrasound machine, a 3.75 MHz convex curvilinear array probe, and a Polhemus Fastrak position sensor mounted on the probe. Calibration was performed with the three-wire method described in Section 2.3.2 [1], using water at approximately 40° Celsius to match the speed of sound in human tissue. Images were acquired with an 8-bit frame grabber and a Silicon Graphics Indy workstation. Each B-scan was cropped to 490×380 pixels, with a pixel size of 0.19 mm. The cropped images were then matched to the position sensor readings and recorded at a rate of 25 fps.

5.3.2 Simulations

Simulations were performed on the same octohedron object described in Section 4.3.2. Instead of registering individual B-scans, here an entire octohedron was purposefully misregistered with respect to a baseline octohedron. The simple shell structure of the octohedron produced clearly defined gradients, so registration was easily performed by the correlation-based registration technique — see Figure 5.4.

[1] Since these spatial compounding experiments were performed just as the work on calibration started, the alternative calibration methods were not yet available.

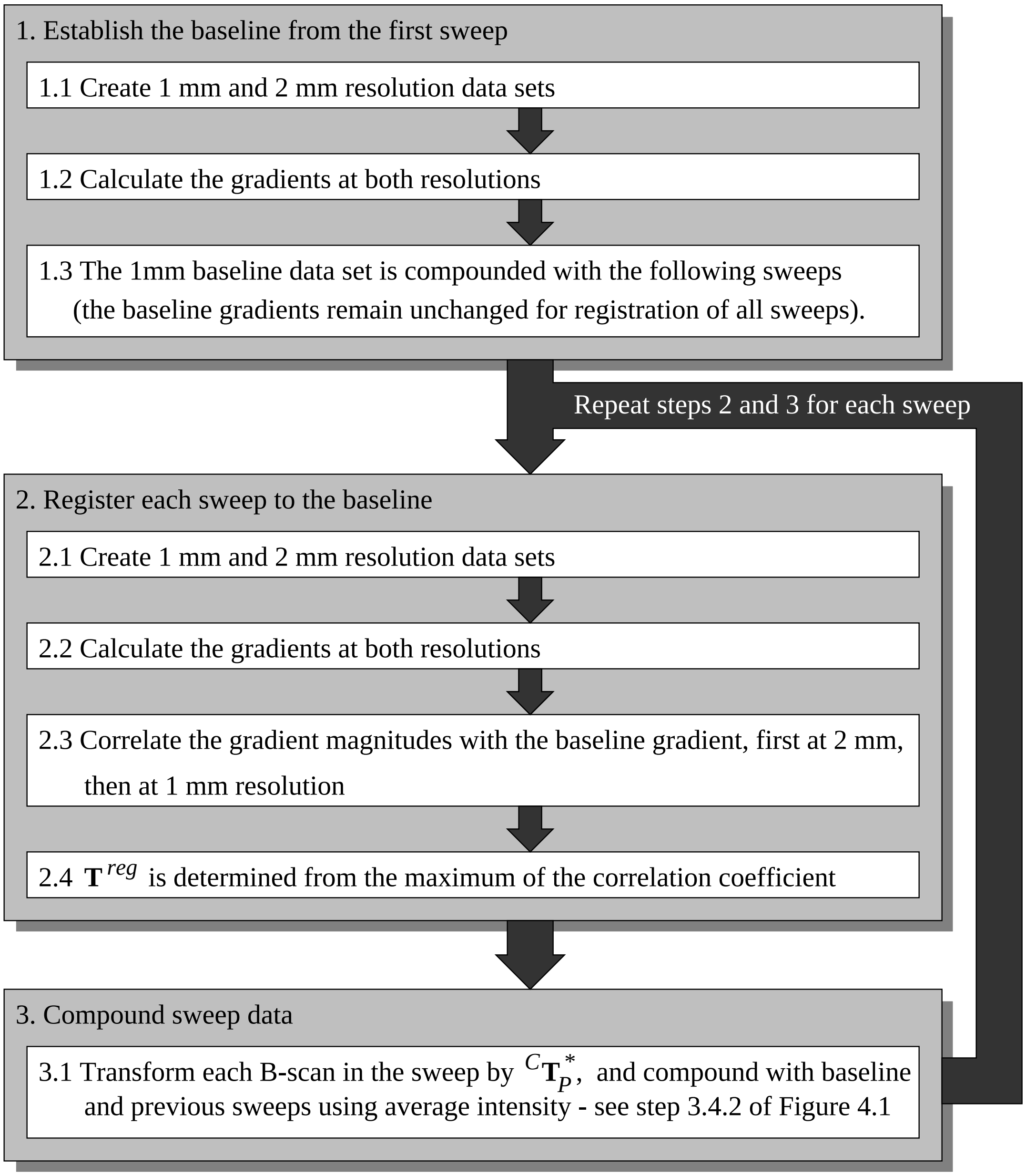
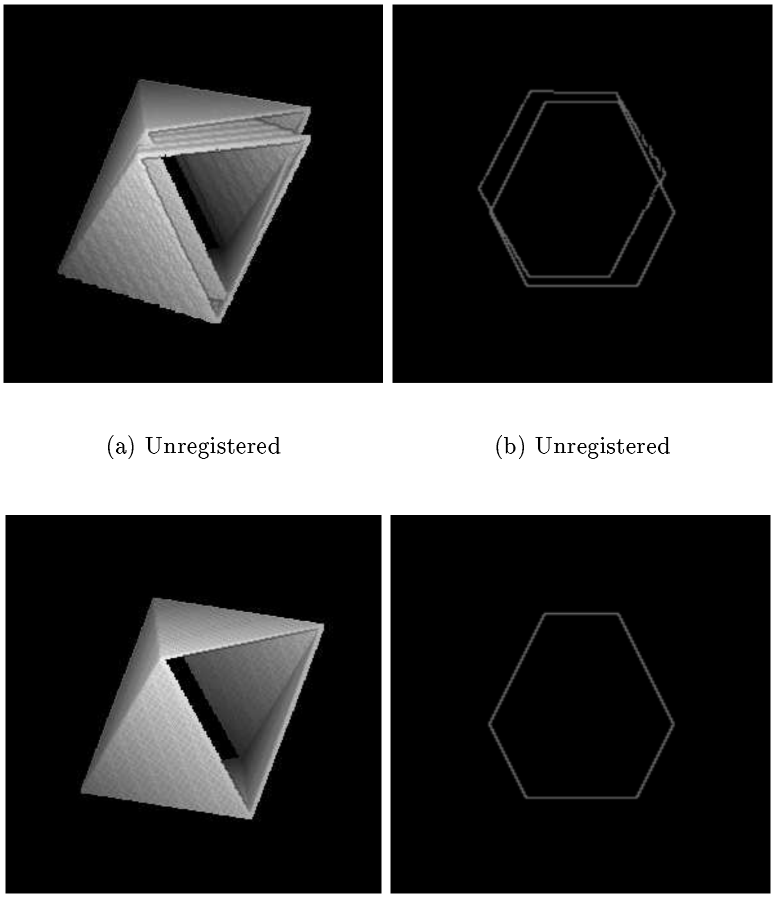


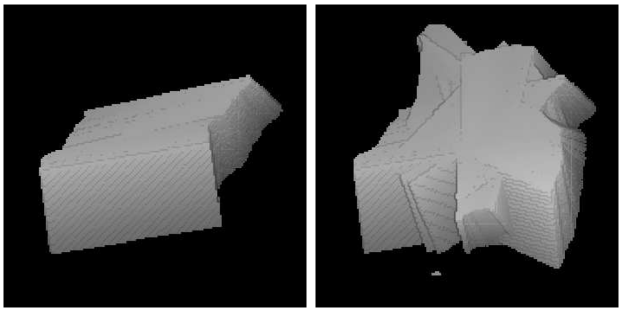
Figure 5.3: **Reconstruction with registration.**



(a) Unregistered (b) Unregistered

(c) Registered (d) Registered

Figure 5.4: **Registration of simulated data.** A surface rendering of the compounded octohedra before registration is shown in (a), and a cross-sectional slice in (b). The misalignment is clearly visible. A surface rendering of the compounded octohedra after registration is shown in (c), and a cross-sectional slice in (d). The registration has resulted in accurate alignment of the two octohedra.



(a) Single sweep (b) Compounded

Figure 5.5: **Outlines of B-scans.** The volume of the reconstruction filled by the baseline sweep 0 is shown in (a). The volume filled by sweeps 0 to 5 is shown in (b). Notice how the volume is enlarged by compounding, since the individual sweeps only partially overlap each other.

5.3.3 Experiments

The gall bladder of a healthy human subject was examined in this study. The gall bladder can be scanned from a number of different look directions, each giving a slightly different view. The variety of look directions used in this study is apparent in Figure 5.5.

The subject was requested to lie motionless on a bed and to breathe lightly to minimise motion-induced errors during scanning. Approximately 100 B-scans were required to cover the gall bladder from tip to tail with sufficient resolution to fill a data set with 1 mm voxels without gaps. Since the acquisition system recorded B-scans at 25 fps, 600 B-scans could be acquired in only 24 seconds. This allowed six sweeps of the organ during that time. Since all six sweeps were acquired in a single examination, the B-scans were subsequently grouped into coherent sweeps by manually reviewing the B-scans in the order of acquisition.

After grouping the B-scans into sweeps, a voxel array was reconstructed for each sweep using the basic PNN algorithm (Figure 2.2). At this stage it is possible to appreciate the registration errors by viewing corresponding slices through the reconstructions — see Figure 5.6. Compounding the sweeps using the position measurements alone gave unsatisfactory results, as can be seen in Figure 5.7. Clearly, image-based registration is essential for accurate spatial compounding.

The first stage of registration was the calculation of the gradients. These were calculated at both 2 mm and 1 mm resolutions — see Figure 5.8. The scale of the gradient operator (kernel standard deviation $\cong$ 2 mm) was chosen to minimise the effect of speckle, yet retain the features of the anatomical structures. The correlation calculations were performed on a Silicon Graphics Indigo 2 Impact 10000 workstation, taking 3.6 hours per sweep. This included calculation of the gradients and the two searches at both resolutions. While this is undoubtably computationally expensive, the times are comparable with those accepted for CT and MR registration (van den Elsen et al., 1995), and the procedure can

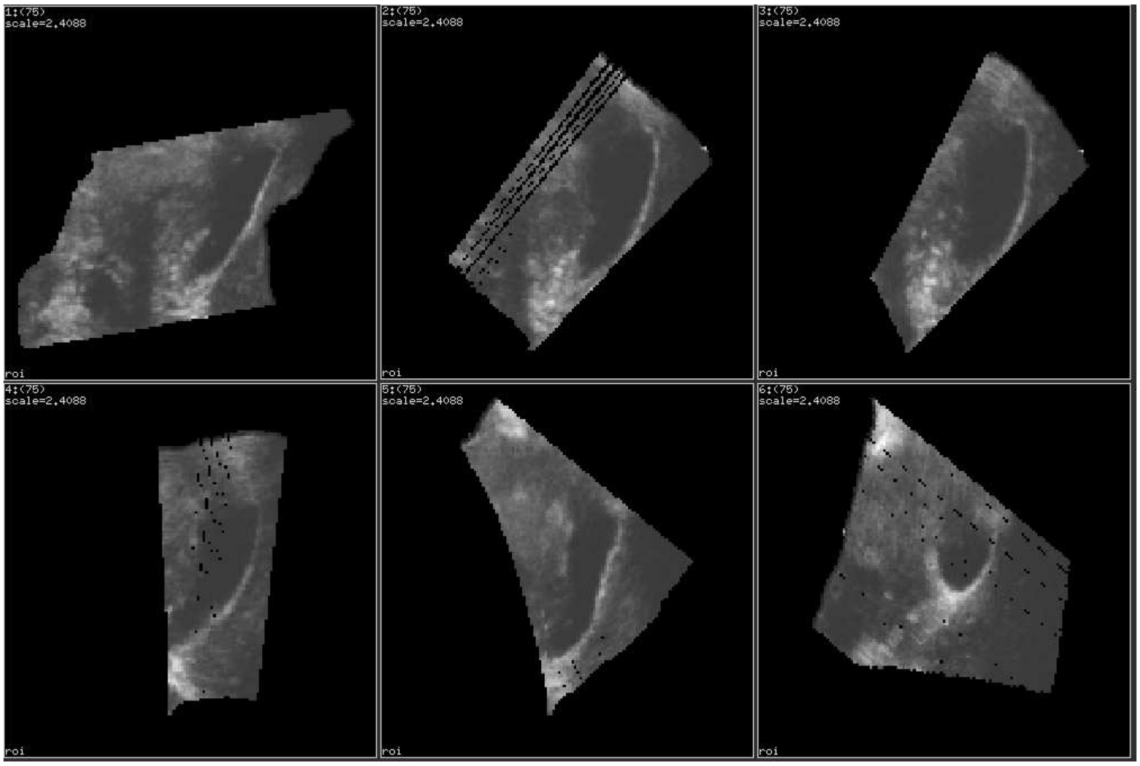


Figure 5.6: **Slices of reconstructions of individual sweeps *without* registration.** Sweeps 0 to 5 are shown in two rows, starting from the top left. All slices are taken at the same location in the reconstruction, identical to the locations used in Figures 5.7 and 5.9. The data have not undergone any hole-filling, so gaps remain in some reconstructions where the B-scans are widely spaced. The cross-sections of the gall bladder appear differently in each sweep, indicating significant registration errors.

Sweep	Registration Error					
	x (mm)	y (mm)	z (mm)	α (degrees)	β (degrees)	γ (degrees)
1	-2.0	9.0	0.0	-6.3	-1.7	0.6
2	-3.0	16.0	-4.8	-10.9	-3.4	2.3
3	1.0	-2.0	-9.0	-5.2	-8.6	-2.3
4	0.0	8.7	-16.0	-2.8	-6.9	4.6
5	10.0	-12.0	-10.0	6.9	-8.0	-3.4

Table 5.1: **Registration errors.** The parameters of $\mathbf{T}^{reg}$ follow the same fixed angle representation described in Section 2.2: translation is expressed as x, y, z and rotation about the z, y and x axes as α, β and γ respectively.

be readily parallelised. The registration results for each sweep are listed in Table 5.1.

5.4 Results

Slices of the registered and compounded data set are shown in Figure 5.9. The borders of the gall bladder are all well aligned compared with the unregistered case in Figure 5.7. The level of speckle noise is also reduced compared with the single sweep reconstructions of Figure 5.6. Furthermore, the organ boundaries of the registered and compounded reconstruction have better contrast and are more uniform than in the single sweep reconstructions (this could not have been achieved by filtering). While all of these improvements are benefits of spatial compounding, it has to be admitted that, on close inspection, a slight loss of spatial detail is apparent. The residual registration errors that are not corrected (such as non-rigid motion of the organ and local refraction of the ultrasound beam) result in a slight blurring of small features. The main benefits of spatial compounding are therefore improved segmentation and visualisation of larger structures.

In order to objectively compare the quality of the compounded reconstruction with the single sweep reconstructions, automatic segmentation by intensity thresholding was attempted. Figures 5.10 and 5.11 show how spatial compounding improves the ability to segment by thresholding. Essentially, spatial compounding has the ability to fill in regions that are not well defined in some sweeps, but are clearer in others. Note that this improvement is not achievable by filtering, which introduces no new information.

The reduction of speckle noise by spatial compounding is difficult to see in the reconstruction slices. It is more clearly evident in 3D surface renderings of original (unprocessed) 3D data sets. Figure 5.12 shows a certain amount of speckle reduction by spatial compounding. The compounding allows a significant fraction of the gall bladder surface to be viewed without any filtering.

Further improvements are possible with a small amount of data processing before segmentation and rendering. A typical ultrasound visualisation procedure involves filtering

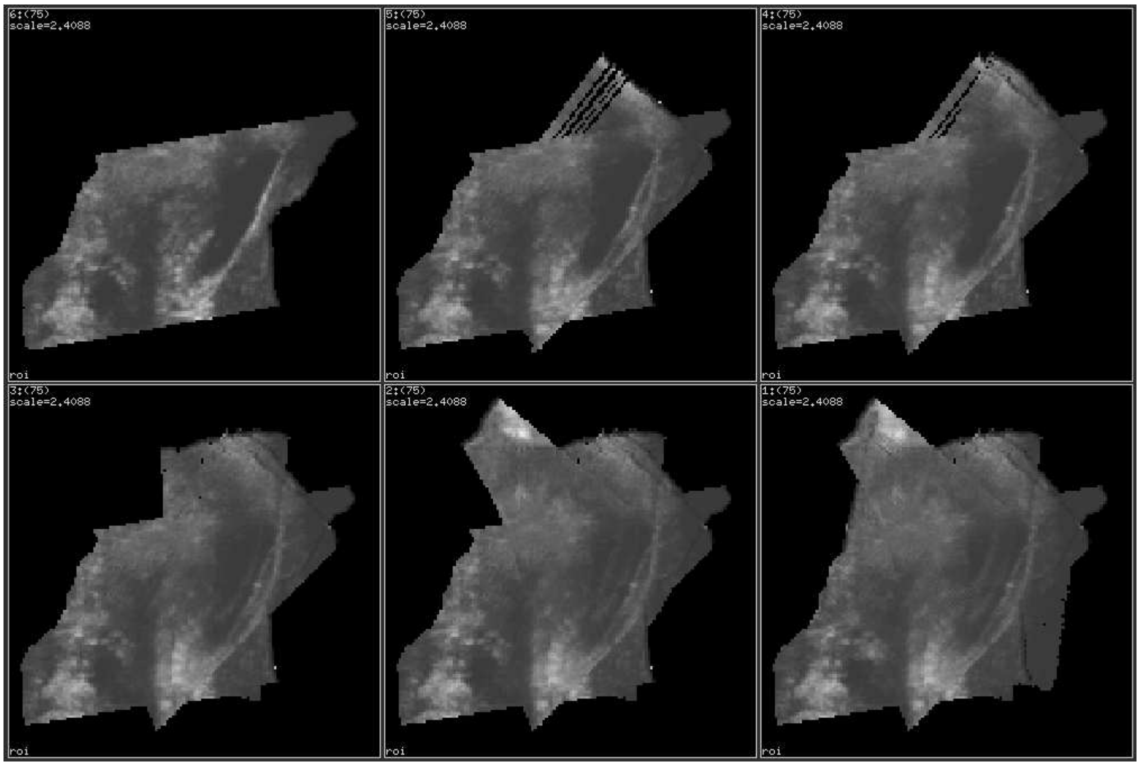
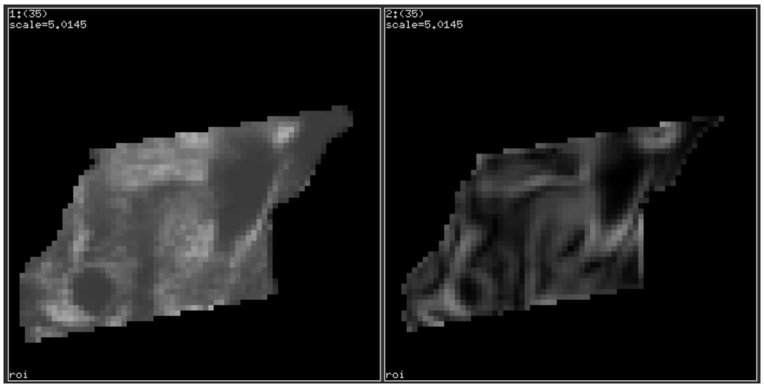
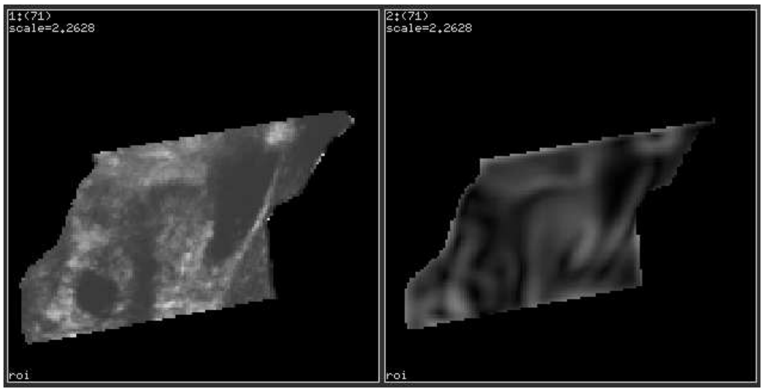


Figure 5.7: **Slices of compounded reconstructions *without* registration.** Slices of reconstructions compounded incrementally with sweeps 0 to 5 are shown in two rows starting from the top left. All slices are taken at the same location in the reconstruction volume, identical to the locations used in Figures 5.6 and 5.9. As more and more sweeps are compounded together using the position sensor data alone, the gall bladder becomes more blurred and eventually indistinguishable. Image-based registration is therefore required for accurate spatial compounding.



(a) 2 mm resolution



(b) 1 mm resolution

Figure 5.8: **Differential geometric features used for registration.** The left images are slices of the original data and the right are the magnitudes of the 3D gradient. The features used for correlation at 2 mm resolution are shown in (a), and the features used for correlation at 1 mm resolution are shown in (b).

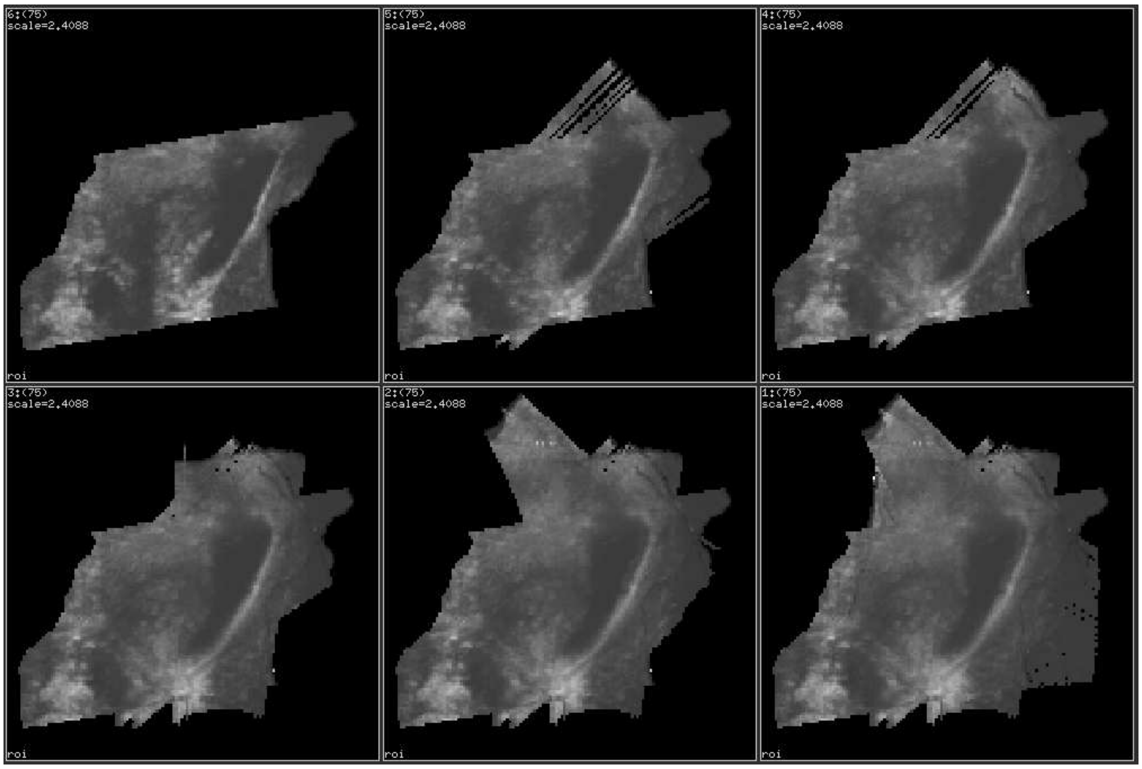


Figure 5.9: **Slices of compounded reconstructions *with* registration.** Slices of reconstructions compounded incrementally with sweeps 0 to 5 are shown in two rows, starting from the top left. All slices are taken at the same location in the reconstruction volume, identical to the locations used in Figures 5.6 and 5.7. The registration has aligned the organ boundaries in the individual sweeps, so the fully compounded data set retains the original organ shape. Compounding has also reduced speckle noise and made the gall bladder more distinct.

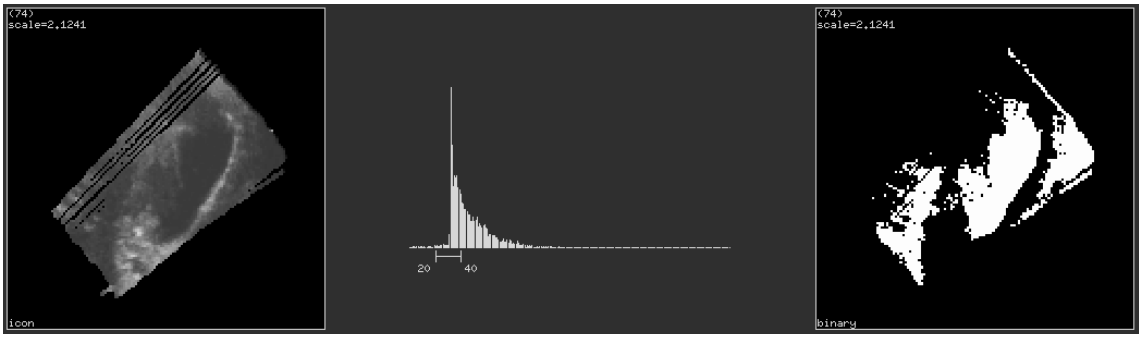


Figure 5.10: **Threshold segmentation of reconstruction of a single sweep.** A slice of the reconstruction from sweep 1 is shown on the left, identical to the location used in Figure 5.11. The slice's grey-level histogram is shown in the middle. The data set has not undergone any processing. The segmentation by thresholding the intensities within the range [20, 40] is shown on the right. The problem with segmentation of ultrasound data by thresholding is that the organ is not easily isolated from the surrounding noise and artifacts. In this case, the image has a dark region from signal drop-out near the gall bladder, which affects the segmentation. A narrower threshold range reduces the volume outside the gall bladder that falls within the range, but also creates gaps within the gall bladder. The range of [20, 40] represents the best trade-off between the two throughout the reconstruction.

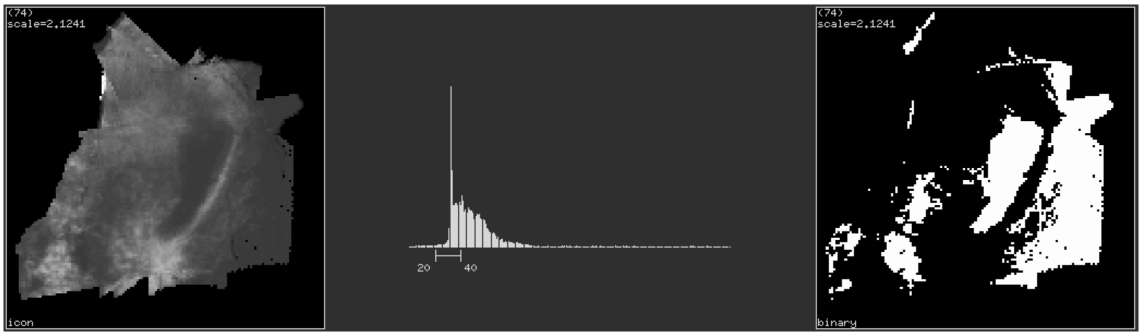
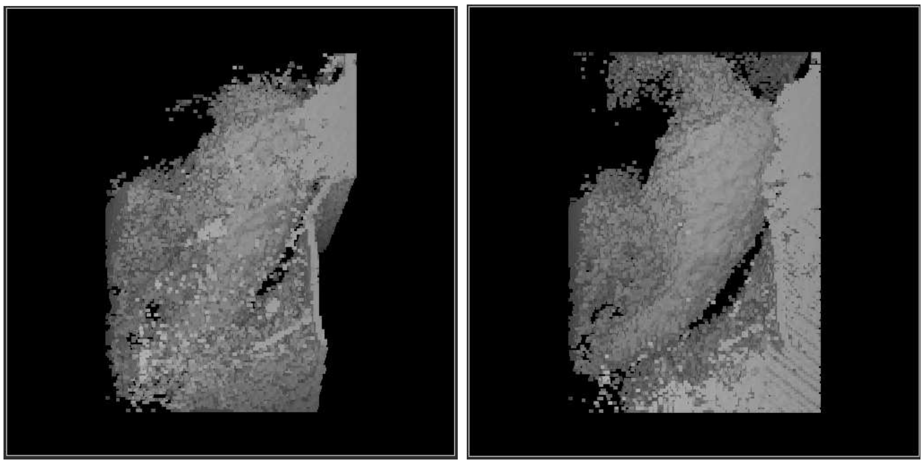


Figure 5.11: **Threshold segmentation of compounded reconstruction.** The segmentation has been performed in an identical manner to Figure 5.10, yet is more accurate.



(a) Single sweep (b) Compounded

Figure 5.12: **3D segmentation of unprocessed data.** The renderings were created using the semiboundary approach described in Section 1.4. Surfaces were extracted by thresholding, similar to the method illustrated in Figures 5.10 and 5.11. No other data processing was performed. Spatial compounding clearly suppresses the noise, making the surface easier to detect.

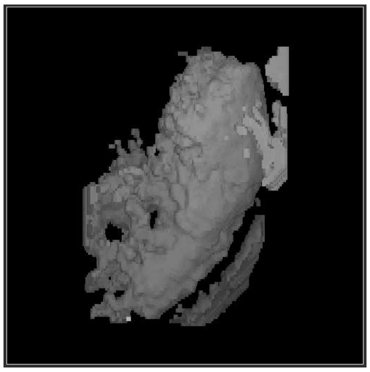
the data before surface detection is attempted (Nelson and Elvins, 1993; Sakas et al., 1995). To investigate this, the single sweep reconstructions are processed in the following manner. First, any holes in the reconstructed 3D data sets are filled with the median value of filled voxels in a local neighbourhood. The data set is then filtered with a Gaussian kernel to remove speckle. Finally, segmentation is performed by thresholding the processed data. Figure 5.13 shows that applying these processing steps to single sweep reconstructions does improve visualisation, but problems still remain. The main problem is that ill-defined boundaries (like the ones in Figure 5.10) are not improved by filtering. The segmentation procedure does not isolate the gall bladder from the surrounding noise and artifacts in single sweep reconstructions. Yet performing the same data processing before segmentation of the registered and compounded data does isolate the gall bladder, as shown in Figure 5.14. Removal of the unwanted, disconnected regions becomes a trivial task, with the results shown in Figure 5.15.

Volume estimation follows a similar trend to surface visualisation, since both rely upon accurate segmentation. Since the true volume of the gall bladder is unknown, live-wire semi-automatic segmentation (Barrett and Mortensen, 1997) of the baseline sweep is chosen as a “gold standard”. See Section 4.4.3 for a description of the live-wire segmentation technique.

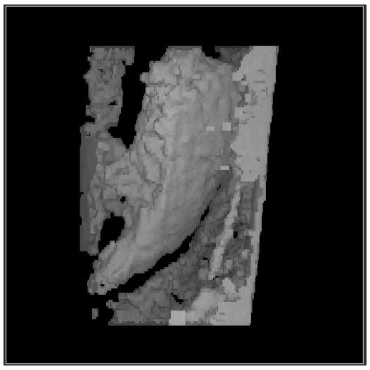
In addition to the live-wire estimate, volumes for individual sweeps and the compounded reconstruction are calculated using the same process (hole-filling, filtering and thresholding) used for the visualisation results in Figures 5.13 and 5.14. In order to cal-

(a) Sweep 0

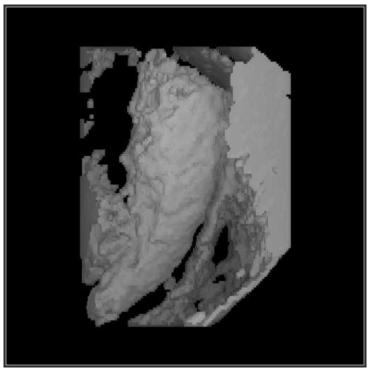
(b) Sweep 1



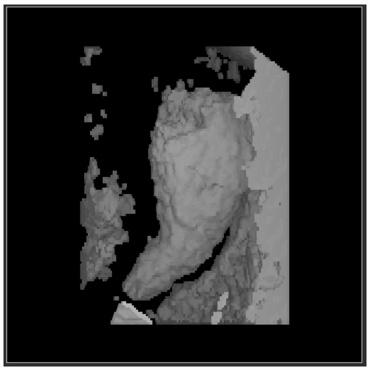
(c) Sweep 2



(d) Sweep 3



(e) Sweep 4



(f) Sweep 5

Figure 5.13: **3D segmentation of individual sweeps after data processing.** All images are produced with identical hole-filling methods, filter sizes and threshold ranges. Compared with Figure 5.12 (a), a considerable amount of noise has been removed by processing. In each sweep, however, the gall bladder is not disconnected from the surrounding noise and artifacts.

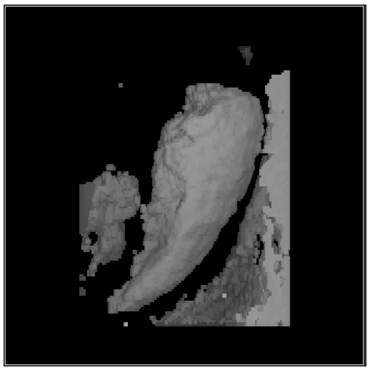


Figure 5.14: **3D segmentation of compounded reconstruction after data processing.** This image is produced in the same manner as in Figure 5.13. Spatial compounding improves the organ segmentation. Compared to the segmentation of the individual sweeps, the compounded data set is the only one in which the gall bladder is completely disconnected from surrounding noise and artifacts. Complete isolation of the gall bladder is now easily performed — see Figure 5.15.

Figure 5.15: **3D segmentation of high resolution compounded reconstruction.** This image is produced from a high resolution reconstruction compounded with sweeps 0 to 5. The data set is first hole filled, then filtered, threshold segmented and manually trimmed of extraneous regions.

Sweep	Volume (ml)
0	13.494
1	16.662
2	15.758
3	14.945
4	15.117
5	14.166
compounded	13.039
live-wire	12.507

Table 5.2: **Volume of gall bladder.** The compounded data set provides the volume estimate closest to the "gold standard" live-wire segmentation.

culate only the gall bladder volume, any segmented part of the data that is not connected to the gall bladder is removed manually. The volume is then calculated by summing the remaining filled voxels that contain the segmented gall bladder.

The calculated volume of the compounded data is closer to the live-wire estimate than any of the individual sweeps — see Table 5.2. The volumes for each of the individual sweeps are higher than the live-wire volume because the segmentation of the gall bladder typically exceeds the gall bladder boundary (as evident in Figure 5.10). This again shows that segmentation of spatially compounded data sets gives better results than single sweeps.

5.5 Conclusions

The benefits of spatial compounding were demonstrated in an examination of a human gall bladder. The level of registration errors was significant, so image-based registration was required. The registration was performed in a fully automatic manner, eliminating user subjectivity. Registration was based on the correlation of features obtained by calculation of the 3D gradient, so neither explicit segmentation nor landmark identification was required. Implementation of the registration technique showed it to be robust and accurate. The efficiency of the algorithm was also improved by using a multi-resolution search for the correlation peak.

Rigid body registration was shown to correct a sufficiently large proportion of the registration errors. Since these errors were known to be small, registration by correlation was feasible by searching a small portion of the six-dimensional parameter space. The resulting registered and compounded 3D data sets offered improved segmentation for better visualisation and volume estimation. Application of conventional visualisation techniques to the compounded data produced dramatically clear and accurate images of the gall bladder. Although visualisation of larger structures was improved, some slight blurring was produced by the residual registration errors. The small loss of detail means spatial compounding has greater advantage in visualisation of larger structures.

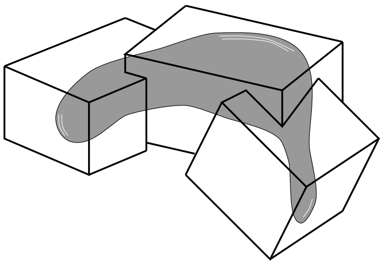


Figure 5.16: **Extended volume of interest.** Several small, partially overlapping volumes can be combined to encompass a larger organ.

Registration of ultrasound data can also be used for purposes other than spatial compounding. For example, the correlation-based registration technique is capable of aligning data sets that are only partially overlapping. An examination volume can therefore be extended so that a large organ is reconstructed from a number of smaller, overlapping examinations — see Figure 5.16.

Felix qui potuit rerum cognoscere caucus.
'Happy is he who gets to know the reasons for things'
[Motto of Churchill College, Cambridge.]
—Virgil, **Georgics II**, 29 B.C.

Chapter 6

Summary and Conclusions

6.1 Calibration

Four different calibration techniques were reviewed and put into a common mathematical framework. Although each technique uses a different type of phantom, they were all expressed with the same basic set of equations. These equations were used to show which parameters are identifiable by each calibration technique. Calculation of the condition number from the equations quantified the identifiability of the parameters, and helped to verify that a sufficiently large range of probe motions was used. Having established a well-conditioned problem, the four calibration techniques were compared in a series of tests. These tests measured the repeatability of calibration, the precision of point reconstruction and the accuracy of distance measurements.

The cross-wire technique demonstrated good repeatability, point precision and distance accuracy. Unfortunately, the technique was the most difficult to use: careful, steady positioning of the probe was required, and the resulting images of the wires needed to be labelled by hand. These factors contributed to the time required for cross-wire calibration — the longest of all techniques.

The three-wire technique was slightly easier to use, but also took a long time to calibrate. Unfortunately, the performance was relatively poor in all three tests. The main problems with the three-wire technique were that the wires in the image still needed to be labelled by hand, and the phantom was difficult to construct accurately.

The single-wall phantom, in contrast, was especially simple to construct, and easy to use. With lines instead of dots to detect in the images, labelling was considerably easier. In fact, automatic line detection was possible. The main problems with the single-wall technique were the limited range of motion and the effect of the beam width. The limited range of motion made identification of the calibration parameters more difficult, as shown by the relatively large condition number of the Jacobian. Substantial errors were also introduced by the beam width at oblique angles of incidence to the phantom wall, as verified in a series of simulations. Overall performance was therefore relatively poor.

The newest calibration phantom, the Cambridge phantom, was shown to use the same principles as the single-wall technique, but the phantom design eliminated the deleterious effect of beam width. The quality of the images was also better over a wider range of angles of incidence. Good performance of the Cambridge technique was demonstrated in all three tests, and was comparable to the cross-wire technique. Yet with automatic line detection, and greater flexibility when scanning, the Cambridge technique was considerably easier to use than the cross-wire technique. The main disadvantages were the need to construct the apparatus of the Cambridge phantom with a high degree of precision, and the the time required to correctly clamp the ultrasound probe to the phantom.

In summary, both the cross-wire and Cambridge calibration techniques produced good results. A large proportion of the remaining errors came from the position sensor itself. The Polhemus Fastrak sensor used in these experiments is rated at about 1 mm accuracy, and additional errors arise from the influence of nearby metal objects and interfering fields. This is consistent with the level of errors found in the comparisons of the calibration techniques. As more accurate sensors are adopted for freehand ultrasound, a point will soon be reached where other errors, such as refraction and tissue movement, will dominate.

6.2 RBF Interpolation

The RBF method was shown to be at least as good as the traditional 3D freehand reconstruction methods for two representative examinations. The RBF method performed particularly well with small gaps, but lost some of its advantage when larger gaps were present. This is because the RBF method alone made use of the underlying shape of the data to interpolate across the gaps. Yet with large gaps, this shape could no longer be followed and the RBF produced smoothed averages of the nearby B-scans.

The RBF method was capable of reconstructing areas with densely overlapping B-scans and areas with large gaps, while making a smooth transition between them. The resulting quality of the interpolated data was good, with no visible reconstruction artifacts produced in any of the tests. Unfortunately, the RBF method was the slowest reconstruction method and is not yet practical in its current form. Yet it is easily parallelised and, given more efficient coding techniques, a practical implementation is feasible.

There is considerable scope for future development of the RBF method. First, the tension should be adjusted to accommodate the nature of the freehand ultrasound data. Since the pixel data all lie in planes, and are usually separated by gaps larger than the pixel size, the tension should be higher within the B-scan and lower perpendicular to the B-scan. It is also possible to introduce weights on the pixels (Mitášová et al., 1995) when, for example, estimates of registration error or signal fall-out are available. Finally, it is possible to find the optimal combination of tension and smoothing by statistical techniques such as cross-validation (Mitášová et al., 1995).

Since the RBF method alone uses a functional approach, it also offers a number of unique opportunities. For example, derivatives of any order can be computed directly from the interpolating function. This may be useful for visualisation, registration and segmentation. Also, because the function usually misses the data points (when a non-zero smoothing parameter is used), the distance from the interpolating function to the data points can be considered the predictive error. Determining which regions have a high predictive error can be useful for investigations into registration errors.

In summary, the comparisons show the importance of interpolation. The traditional methods were designed to accommodate limited computational resources, but artifacts can be introduced and information lost. With increasingly powerful modern workstations, more sophisticated interpolation methods are feasible, and should be used.

6.3 Spatial Compounding

The main advantages of 3D spatial compounding were shown to be increased SNR, improved visualisation and more accurate volume estimation. These improvements came from a reduction in speckle and other artifacts, as well as more uniform organ boundaries. The main disadvantage of 3D spatial compounding, a loss in image detail, was caused by slight blurring of the features. The key to minimising blurring was accurate registration of the images before compounding. Registration using image features was possible and accurate alignment was achieved with only a small loss of resolution.

The benefits of spatial compounding are not new. Compound scanning was popular in the past (before the widespread arrival of real-time machines), and the quality of the 2D images was remarkably good. With the advent of real-time imaging, there was no time to compound data taken from different look directions, so spatial compounding became a lost art. But now, 3D ultrasound brings new requirements that are not encountered in 2D ultrasound. Two important requirements are 3D visualisation and volume estimation, and spatial compounding is especially useful for meeting these requirements. Here the slight loss of detail is less important than the improvement in boundary uniformity and the reduction of speckle. In light of these trade-offs, it makes sense to use heavily compounded data for visualisation and volume estimation, and uncompounded data for any-plane slicing, where greater detail is desired.

The dramatic improvement in computing power in recent years has had a profound effect on the development of modern ultrasound machines. The latest generation of machines use parallel processing and dedicated digital signal processing hardware. It is now possible to perform rapid feature recognition as the images are created. Since registration depends on accurate alignment of features, improvements in feature recognition result in improvement in the accuracy of registration. This, in turn, decreases the level of blurring in spatial compounding. This trend has led to renewed interest in spatial compounding (Whittingham, 1997).

6.4 Future Directions

Although the first commercial 3D ultrasound machine by Kretztechnik used a mechanical swept volume approach, several commercial 3D freehand systems have recently become available. Research on both techniques is active, and image quality has improved over the past few years. Unfortunately, both techniques require examination times of 5 to 30 seconds, and reconstruction is performed off-line. Since conventional 2D B-scan imaging is real-time, 3D ultrasound is currently used alongside 2D ultrasound, not as a replacement. When 3D ultrasound achieves real-time data acquisition rates (more than 30 volumes per second), it will replace 2D ultrasound, because slices of 3D volumes can be used instead of B-scans. Real-time 3D ultrasound has been the subject of research for many years, and eventually new techniques (such as 2D array transducers) will attain a level of performance acceptable for clinical use.

Future 3D ultrasound techniques will still encounter the same problems found today: images will be dependent on the look direction, and artifacts will be present. The artifacts will still include speckle, shadowing, refraction, speed of sound estimates, among others, because the physics of the interaction between coherent sound pulses and human tissue remains unchanged. Spatial compounding will also still be a useful way of reducing these artifacts, and system calibration, image registration and interpolation will continue to be the main issues. The questions of how to align data and how to cope with the combination of overlapping data and large gaps will remain active subjects of investigation.

The trends in other medical imaging modalities, such as CT and MR, suggest what lies ahead for ultrasound. In particular, new registration techniques that take an information theoretic approach have been developed (Viola and Wells, 1997; Wells et al., 1996). These may offer improved accuracy and robustness for registration of ultrasound data. As feature detection and registration techniques improve, registration by image warping may also become feasible (Brown, 1992; Maintz and Viergever, 1998).

RBF's have also been used with other imaging modalities. In this thesis, knowledge of the basic acquisition process was used to model the data with smooth functions. In the future, knowledge of the point spread function (the axial, lateral and elevational dimensions of an ultrasound pulse) may also benefit interpolation.

Ideally, the interpolation should make use of all available information from the acquisition system. The complex interaction of the ultrasound pulse with living tissue produces a wealth of information, as evidenced by the complex signals of the returning echoes. Some information is discarded by the ultrasound machine, but efforts are being made to understand and utilise this information (Whittingham, 1997). Progress is being made mainly with the help of new digital techniques. Digital representation of the original, unprocessed echo signals means the phase and frequency information of the echoes can now be analysed, whereas it was previously lost during amplification, demodulation, filtering and compression. The interpolation process would also certainly benefit by incorporating as much of this information as possible.

Many conventional ultrasound machines are also capable of producing Doppler ultrasound images. In Doppler ultrasound, pulse-echo information can be used to calculate the speed of fluids passing through blood vessels, providing important vascular flow information (Kremkau, 1993). While Doppler ultrasound can provide information no other imaging modality can, it suffers from considerable signal drop-out and noise. The same comments on spatial compounding and interpolation also apply to Doppler ultrasound data, making it another promising direction for future research.

Appendix A

Ultrasound Machine Controls

The quality of the B-scans is highly dependent on the settings of the ultrasound machine. Some of the parameters that a physician must adjust to suit each examination are listed in Table A.1. Normally the physician adjusts the controls to obtain the best looking B-scans for the organ of interest, then keeps them constant during the freehand examination. This helps to produce a consistent 3D data set.

Although the large number of controls means that an inevitable amount of trial and error is involved in obtaining the best possible image, a basic procedure can be followed. The following procedure is compiled from lists of suggested clinical practice (Cosgrove et al., 1993; Chervenak et al., 1993):

> A probe with a suitable frequency and focal point must first be selected. High frequency probes and short focal points are used for superficial structures and lower frequency probes with deeper focal points for deep lying structures. The probe is then placed over the region of interest using a sufficient amount of coupling gel to ensure a good interface with the skin's surface. The first adjustments are usually made to the TGC and overall amplifier gain. Both of these adjustments control the level of amplification of the echo signals immediately after they are received by the transducer. TGC differs from overall gain adjustment by allowing the user to create an arbitrary gain curve. The shape of the gain curve is controlled by a number of sliders, each corresponding to gain at a particular depth.
>
> A quick scan of the region is made and the TGC and overall gain are adjusted to put the image intensities in the range most easily perceived, and achieve an even balance in the near and far fields. It is also suggested that the pulse amplitude (output power) be set as low as possible. This can be done by setting a high overall gain, then decreasing the pulse amplitude until the desired image intensities are achieved.
>
> Next, the correct pan and zoom settings are made to enlarge the region of interest to the extents of the display monitor. The focal point (or points for multiple focus machines) is set to the centre of the region of interest for

probe used	focus parameters
time-gain compensation	scan line density
overall gain	scan angular range
pulse amplitude	image post-processing
pan and zoom	display persistence
echo enhancement	display monitor controls

Table A.1: **Parameters than can be controlled by the physician during an examination.**

optimal resolution. Other controls such as echo enhancement are also adjusted for best image quality.

Finally the post-processing curve is adjusted to obtain the desired contrast resolution or enhance outlines of anatomical structures. Other optional controls that may enhance the image are the display persistence [1] (for frame averaging and some speckle reduction) and frame rate. On some machines, the frame rate is chosen as a trade-off between capturing organ motion (such as the heart) and resolution, since a slower sweep of the beam gives higher scan line density.

The contrast, brightness and focusing of the display monitor should also be checked. These are often set to obtain uniformity of images among the different machines in a clinic, so comparisons can be made more easily. These controls do not have an impact on the video signal recorded by the frame grabber of a 3D freehand ultrasound system.

[1] Display persistence was not used in any of the examinations in this thesis.

Appendix B

Numerical Issues in Calibration

The accuracy of calibration depends to some extent on the numerical methods employed. In the literature of robot kinematic calibration (a problem very similar to freehand ultrasound calibration) there are guidelines for establishing a well-conditioned problem (Hollerbach and Wampler, 1996). The two key issues are rank determination and scaling.

To ensure that all the parameters of Φ are identifiable, the Jacobian matrix $\mathbf{J}$ is evaluated (near the solution) to check that it is of full rank. To do this, SVD (singular value decomposition) is performed:

$$\mathbf{J} = \mathbf{U}\Sigma\mathbf{V}^{\mathrm{T}}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\Sigma = \mathrm{diag}(\rho_1, \ldots, \rho_q)$ is the matrix of ordered singular values. Here, $q = 11$ for the cross-wire, single-wall and Cambridge phantoms, and $q = 14$ for the three-wire phantom.

The smallest singular value ρ_q will be near zero if there is an unidentifiable parameter. If all ρ are non-zero, then the Jacobian is said to have full rank. A convenient measure of identifiability is the condition number κ, which is simply the ratio of the largest to smallest singular values. A large condition number indicates that the problem is ill-conditioned, and at least one parameter cannot be accurately identified.

The condition number may be large for one of three main reasons. The first is that a parameter may not be intrinsically identifiable using the particular calibration technique. For example, when using the single-wall or Cambridge phantoms, three parameters of ${}^{C}\mathbf{T}_{T}$ are not identifiable because they do not affect the height or orientation of the plane. Removing these unidentifiable parameters from Φ reduces κ. The second reason is that the motion of the probe may not have covered all the necessary degrees of freedom to identify the parameters in Φ. Extending the range of scanning directions will help to reduce κ. The third reason concerns degeneracy in the fixed-axis representation of rotation. If $\beta = \pm 90°$, then there exist multiple triples (α, β, γ) that produce the same transformation ${}^{R}\mathbf{T}_{P}$ (Craig, 1989). In effect, one of the angles becomes unidentifiable. This situation can be avoided by mounting the position sensor at a skew angle to the ultrasound probe, so that β is not close to $90°$.[1]

[1]Note that although degeneracy in the representation of rotation increases κ, it has no bearing on the

	κ
14 parameters	6.97×10^{18} ($\rightarrow\infty$)
11 identifiable parameters	4867
11 scaled parameters	87.5

Table B.1: **Condition numbers for calibration with the Cambridge phantom.** The condition number approaches infinity when Φ includes the full set of 14 parameters. By removing the three unidentifiable parameters x, y, and α (see Table 2.1), the condition number is considerably reduced. Column scaling of the remaining 11 parameters reduces the condition number to below 100.

The second important issue is scaling. Task variable scaling is the process of scaling the errors produced by equation (2.4) so they contribute equally to the optimisation. Since each of the equations has the same units and describes a deviation from zero, they are directly comparable. Yet this is not true for the parameters to be identified. The different parameters have different units (mm, radians, mm/pixel) and different effects on the residual of equation (2.4). This issue can be addressed by parameter scaling, which involves modifying equation (2.5) as follows:

$$\Delta\mathbf{f} = \mathbf{J}\Delta\Phi = \mathbf{J}(\mathbf{H}\mathbf{H}^{-1})\Delta\Phi = (\mathbf{J}\mathbf{H})(\mathbf{H}^{-1}\Delta\Phi).$$

$\mathbf{H}$ is a scaling matrix chosen in such a way that $\mathbf{JH}$ is better conditioned than $\mathbf{J}$ alone. Equation (2.6) is then evaluated using $\mathbf{JH}$ and $\mathbf{H}^{-1}\Phi$ in place of $\mathbf{J}$ and Φ. After convergence, Φ is determined by unscaling the solution $\mathbf{H}^{-1}\Phi$.

The simplest way of selecting $\mathbf{H}$ is by column scaling (Lawson and Hanson, 1974), whereby $\mathbf{H} = \mathrm{diag}(h_1, \ldots, h_q)$ and

$$h_i = \begin{cases} ||j_i||^{-1} & \text{if } ||j_i|| \neq 0 \\ 1 & \text{if } ||j_i|| = 0 \end{cases}.$$

j_i is the ith column of $\mathbf{J}$, evaluated at a reasonable guess of the solution. This scaling ensures that each parameter has the same effect on $\Delta\mathbf{f}$.

The combination of removing unidentifiable parameters and scaling the remaining parameters should reduce κ to an acceptable level. For robot kinematic calibration, a common rule of thumb is that the condition number should be less than 100 for accurate parameter identification (Schröer, 1993). Table B.1 shows how to establish a well-conditioned calibration problem.

There are, however, several different values of Φ which can achieve the same, well-conditioned global minimum. These are called "mirror" solutions. This is not the same as the singularity problem discussed earlier, where an infinite number of solutions exist and the problem is ill-conditioned. A mirror solution can be constructed, for example, by adding a multiple of 2π to any of the angles. Even though mirror solutions represent the

calibration matrix ${}^{R}\mathbf{T}_{P}$, and any of the multiple solutions for (α, β, γ) are equally valid.

same calibration, they need to be expressed in a canonical form to allow straightforward comparisons of results[2]. The requirements are as follows:

- s_x and s_y are positive;
- α and γ are in the range $[-\pi, \pi]$;
- β is in the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

These constraints can be enforced using the following procedure (Prager et al., 1998b):

1. Limit all the angles to $\pm\pi$ by adding or subtracting 2π.
2. If β is outside $\pm\pi/2$, add (or subtract) π to correct it, and add π to both α and γ.
3. If $s_y < 0$ then $\gamma \rightarrow \gamma + \pi$ and $s_y \rightarrow -s_y$.
4. If $s_x < 0$ then $\alpha \rightarrow \alpha + \pi$, $\beta \rightarrow -\beta$, $\gamma \rightarrow \pi - \gamma$ and $s_x \rightarrow -s_x$.
5. Check that α and γ are still within $\pm\pi$: if not, repeat step 1.

Note the order of the corrections for negative scales. If both scales require correction, the composite transformation for γ must be $\gamma \rightarrow -\gamma$.

[2] Other researchers parameterise rotation using different schemes, such as raw rotation matrices (Detmer et al., 1994; Leotta et al., 1997) and unit quaternions. Although such schemes avoid singularities and mirror solutions, the redundancy in the representation means that the optimisation has to be constrained.

Appendix C

Calibration Simulations

By simulating the effects of the finite beam thickness, we can understand why the Cambridge phantom performs better than the single-wall phantom. As mentioned in Section 1.5.2, the width of the beam (w) at the focal point can be estimated from the focal length (F), wavelength (λ) and the radius (a) of the effective aperture (equation (1.1)). For a Toshiba 7 MHz linear array probe the beam width is:

$$w = \frac{F\lambda}{a} = \frac{20\,\text{mm} \times 0.2\,\text{mm}}{7\,\text{mm}} = 0.6\,\text{mm}.$$

F and a are obtained from the manufacturer's technical specifications, and λ is calculated from the speed of sound in water ν and the frequency f:

$$\lambda = \frac{\nu}{f} = \frac{1.48 \times 10^3\,\text{mm/sec}}{7 \times 10^6\,\text{sec}^{-1}} = 0.2\,\text{mm}.$$

The width of the beam increases away from the focal point, so 0.6 mm is a lower bound.

The simulations were performed using the same set of probe motions as one of the trials of the Cambridge phantom. This allows the simulated and observed results to be directly compared. For each probe position, given a set of calibration parameters ϕ, it is straightforward to calculate where the scan plane intersects the floor, and hence the position of the line in the B-scan. The parameters used in the simulations are the ones that resulted from the actual calibration of the Cambridge phantom. Solving equation (2.4) using the simulated set of B-scan lines and the real position sensor readings returns the same ϕ used to create the simulation. These values are listed in the left hand column of Table C.1.

If a non-zero beam thickness is introduced, the expected change in position of the lines in the B-scans can be calculated. Solving the calibration equations will result in a different ϕ. For the purposes of simulation, the beam was approximated as having a constant thickness. Since the lower bound on the beam width is 0.6 mm, simulations were performed using several widths in the range 1 mm to 3 mm: the results are listed in Table C.1.

We observe that the RMS error and condition number are relatively insensitive to beam thickness. But compared with the Cambridge phantom (see Table 2.5), calibration repeatability $\left|\Delta^R \underline{\mathbf{x}}\right|$ is greatly reduced, even for beam widths as low as 1 mm. So the Cambridge phantom's superior repeatability is a result of its suppression of beam width effects.

Beam thickness accounts for only some of the errors in real calibration. Errors also arise from random noise and latency in the position sensor measurements, the influence of nearby metal objects on the position sensor's performance, flexibility of the mount used to attach the receiver to the probe, tolerances of the phantom's construction, B-scan image resolution and detection of the lines in the B-scans.

Some of these errors can be confidently neglected. The latency of the position sensor was measured and subtracted from the readings by the Stradx acquisition software (Prager et al., 1998a). The sensitivity of the position sensor to metal has been addressed by mounting the receiver some distance from the probe and removing all nearby metal objects. The receiver mount was rigid and the phantoms were constructed to small tolerances. The effect of finite lateral image resolution is negligible when detecting a near horizontal line.

This leaves random position sensor noise, the accuracy of line detection and axial image resolution. The extent of these errors can be estimated and incorporated into the simulations. Uniform random fluctuations of 0.76 mm in distance and 0.15° in orientation were added to the position measurements: these are taken from the manufacturer's specifications of the Polhemus Fastrak sensor. A reasonable estimate of the image's axial resolution is one wavelength λ (0.2 mm). The accuracy of line detection was estimated as the standard deviation of the Gaussian kernel (0.7 mm) used to smooth the image before edge detection (Prager et al., 1998b). Combining this with the resolution error, a uniform random noise of amplitude 0.9 mm was added to the end points of the lines in the B-scans.

The simulations were then repeated with the noisy measurements. The results are listed in Table C.1. The RMS error increases significantly to levels similar to those found in real trials, giving confidence that the noise estimates are reasonable. Yet the noise has little effect on the condition number or repeatability. This is because the effect of random errors is reduced when a large number of equations is used to determine ϕ. This justifies the decision to calibrate using many more than the minimal number of B-scans.

	Ideal				Noisy			
Beam Width	0 mm	1 mm	2 mm	3 mm	0 mm	1 mm	2 mm	3 mm
s_x (mm/pixel)	0.137	0.137	0.137	0.137	0.136	0.137	0.137	0.138
s_y (mm/pixel)	0.138	0.138	0.138	0.138	0.138	0.138	0.138	0.138
x (mm)	-270.86	-269.37	-267.88	-266.39	-270.45	-268.83	-267.54	-266.22
y (mm)	4.86	4.87	4.87	4.88	4.48	4.44	5.04	4.91
z (mm)	-65.08	-63.62	-62.15	-60.69	-65.04	-63.54	-61.81	-60.68
α (deg.)	-0.55	-0.55	-0.54	-0.54	-0.69	-0.67	-0.56	-0.65
β (deg.)	44.22	44.22	44.22	44.22	44.27	44.29	44.23	44.19
γ (deg.)	-90.37	-90.42	-90.47	-90.53	-89.38	-89.15	-90.97	-90.60
B-scans used	699	699	699	699	699	699	699	699
κ	87.5	87.15	86.85	86.57	87.50	87.26	86.92	86.58
RMS error (mm)	0.000	0.020	0.039	0.059	0.346	0.360	0.363	0.368
$\lvert\Delta^R\underline{\mathbf{x}}\rvert$ (mm)	-	2.09	4.18	6.28	0.819	2.68	4.74	6.54

Table C.1: **Simulations of single-wall calibration.** Only beam thickness errors were simulated to obtain the “ideal” calibration results. For the “noisy” results, position sensor errors, accuracy of line detection and finite image resolution were also considered. The repeatability is measured relative to the ideal calibration with zero beam width.

Appendix D

Statistical Theory of 3D Spatial Compounding

Simple statistical theory can be used to predict the increase in SNR from 3D spatial compounding. In this appendix two distinct cases are considered: 3D high resolution compounding and 3D low resolution compounding. First, however, some simple results from statistical theory are reviewed.

(1) Linear functions of random variables (Hogg and Ledolter, 1987): Consider the random variable Y as a linear function of *independent* random variables X_i: $Y = \sum_{i=1}^{n} a_i X_i$ where $\{a_1 \ldots a_n\}$ are constant coefficients. Then the expectation (or mean) μ_Y of Y, expressed in terms of the expectation of X_i (μ_{X_i}), is $\mu_Y = \sum_{i=1}^{n} a_i \mu_{X_i}$. The variance of Y (σ_Y^2), in terms of the variance of X_i ($\sigma_{X_i}^2$), is $\sigma_Y^2 = \sum_{i=1}^{n} a_i^2 \sigma_{X_i}^2$.

(2) Sets of samples of random variables: Consider Z as a set of samples Y_j drawn randomly from a number of independent distributions: $Z = \{Y_1, \ldots, Y_j, \ldots, Y_m\}$. The expected mean of the samples in Z, expressed in terms of the expectation (μ_{Y_j}) of the distribution from which each Y_j is drawn, is $\mu_Z = \frac{1}{m} \sum_{j=1}^{m} \mu_{Y_j}$. Provided all μ_{Y_j} are equal, the expected variance of the samples in Z, expressed in terms of the variances ($\sigma_{Y_j}^2$) of the distributions from which each Y_j is drawn, is $\sigma_Z^2 = \frac{1}{m} \sum_{j=1}^{m} \sigma_{Y_j}^2$. These equations are readily derived from the definitions of expectancy and variance.

D.1 3D High Resolution Compounding

The simplest 3D compounding case occurs when the voxel size in the reconstruction is the same as the pixel size of the B-scan. This means that each voxel is intersected no more than once per B-scan (discounting the voxels that can occasionally contain two pixels if intersected obliquely). To predict the improvement in SNR, two key assumptions are made:

Assumption 1: The subject of the ultrasound examination exhibits no resolvable structure, so any grey-level variations in the B-scans are caused entirely by speckle.

Assumption 2: Separate B-scans are taken from look directions spaced sufficiently far apart, so that the speckle across B-scans is uncorrelated.

Now consider Z to be the 3D array of voxels Y_j, each compounded from pixels X_i of the individual B-scans. Each X_i comes from a distribution with mean μ_0 and variance σ_0^2. The X_i averaged for a particular Y_j are independent, because the speckle is uncorrelated across B-scans. There are m voxels in the set Z, and voxel j is intersected n_j times. The maximum possible value of n_j is the total number of B-scans used in the reconstruction.

$$Z = \{Y_1, \ldots, Y_j, \ldots, Y_m\} \qquad \text{where} \qquad Y_j = \sum_{i=1}^{n_j} \frac{1}{n_j} X_i.$$

From (1),

$$\mu_{Y_j} = \sum_{i=1}^{n_j} \frac{1}{n_j} \mu_0 = \mu_0 \qquad \text{and} \qquad \sigma_{Y_j}^2 = \sum_{i=1}^{n_j} \frac{\sigma_0^2}{n_j^2} = \frac{\sigma_0^2}{n_j}.$$

From (2),

$$\mu_Z = \frac{1}{m} \sum_{j=1}^{m} \mu_0 = \mu_0 \qquad \text{and} \qquad \sigma_Z^2 = \frac{1}{m} \sum_{j=1}^{m} \frac{\sigma_0^2}{n_j}.$$

Combining these results gives

$$\frac{\mathrm{SNR}(Z)}{\mathrm{SNR}(Z_0)} = \frac{\mu_Z / \sigma_Z}{\mu_{Z_0} / \sigma_{Z_0}} = \sqrt{\frac{1}{\frac{1}{m} \sum_{j=1}^{m} \frac{1}{n_j}}}.$$

To abbreviate references to this result in the main text, let the quantity n_{high_res} be defined as follows:

$$n_{high\_res} = \frac{1}{\frac{1}{m} \sum_{j=1}^{m} \frac{1}{n_j}} \quad \Rightarrow \quad \frac{\mathrm{SNR}(Z)}{\mathrm{SNR}(Z_0)} = \sqrt{n_{high\_res}}.$$

Note that the $\sqrt{n}$ result of the 2D case can be derived by setting $n_j = n$ for all j.

D.2 3D Low Resolution Compounding

Many 3D reconstructions involve cuberilles with voxels considerably larger than the pixels in the B-scans. Again, the set of voxels in the reconstruction can be written as

$$Z = \{Y_1, \ldots, Y_j, \ldots, Y_m\} \qquad \text{where} \qquad Y_j = \sum_{i=1}^{n_j} \frac{1}{n_j} X_i.$$

but in this case the X_i are *not* all independent, since some are from the same neighbourhood in the same B-scan (from the reduction in resolution). This becomes clear if each voxel Y_j is considered as follows:

$$Y_j = \sum_{k=1}^{k_j} \sum_{p=1}^{p_{j,k}} \frac{1}{n_j} X_{p,k}$$

where k_j is the number of B-scans that intersect voxel j, and $p_{j,k}$ is the number of pixels that intersect voxel j for B-scan k: *i.e.* $\sum_{k=1}^{k_j} p_{j,k} = n_j$. The individual pixels $X_{p,k}$ are labelled according to B-scan k and pixel p. Rearranging gives

$$Y_j = \sum_{k=1}^{k_j} \frac{p_{j,k}}{n_j}\,\Theta_{k,j} \qquad \text{where} \qquad \Theta_{k,j} = \sum_{p=1}^{p_{j,k}} \frac{1}{p_{j,k}} X_{p,k}.$$

$\Theta_{k,j}$ is the mean of the pixels that intersect voxel j for B-scan k. The expected mean (μ_Θ) of $\Theta_{k,j}$ can be determined using the result for the expectation in (1), which is still valid for non-independent $X_{p,k}$ (Hogg and Ledolter, 1987):

$$\mu_\Theta = \sum_{p=1}^{p_{j,k}} \frac{1}{p_{j,k}} \mu_0 = \mu_0.$$

Since the size of the speckle is larger than one pixel, the grey-level value of a pixel is not independent of its close neighbours. Thus the expected variance (σ_Θ^2) of $\Theta_{k,j}$ cannot be determined using the result for the variance in (1), which is not valid if the $X_{p,k}$ are not independent. The variance of the sum of several neighbouring pixels *does* decrease as the number of summed pixels increases, but $\sigma_\Theta^2 \gg \sigma_o^2/p_{j,k}$. σ_Θ^2 depends on the spatial structure of the speckle and can be expressed as follows:

$$\sigma_\Theta^2 = \sigma_0^2 b^2(p_{j,k})$$

where $b(p_{j,k})$ is estimated empirically from the B-scans themselves. Returning now to the expression for Y_j, the summation over k *can* be considered as a sum of independent variables, since each element in the sum comes from a different B-scan. Hence

$$\mu_{Y_j} = \sum_{k=1}^{k_j} \frac{p_{j,k}}{n_j}\mu_0 = \mu_0 \quad \text{and} \quad \sigma_{Y_j}^2 = \sum_{k=1}^{k_j} \left(\frac{p_{j,k}}{n_j}\right)^2 \sigma_0^2 b^2(p_{j,k}).$$

From (2),

$$\mu_Z = \frac{1}{m}\sum_{j=1}^{m} \mu_0 = \mu_0 \quad \text{and} \quad \sigma_Z^2 = \frac{1}{m}\sum_{j=1}^{m}\sum_{k=1}^{k_j} \left(\frac{p_{j,k}}{n_j}\right)^2 \sigma_0^2 b^2(p_{j,k}).$$

Combining these results gives

$$\frac{\mathrm{SNR}(Z)}{\mathrm{SNR}(Z_0)} = \frac{\mu_Z/\sigma_Z}{\mu_{Z_0}/\sigma_{Z_0}} = \sqrt{\frac{1}{\frac{1}{m}\sum_{j=1}^{m}\sum_{k=1}^{k_j}\left(\frac{p_{j,k}}{n_j}\right)^2 b^2(p_{j,k})}}.$$

Again, to abbreviate references to this result in the main text, let us define the quantity n_{low_res} as follows:

$$n_{low\_res} = \frac{1}{\frac{1}{m}\sum_{j=1}^{m}\sum_{k=1}^{k_j}\left(\frac{p_{j,k}}{n_j}\right)^2 b^2(p_{j,k})} \quad \Rightarrow \quad \frac{\mathrm{SNR}(Z)}{\mathrm{SNR}(Z_0)} = \sqrt{n_{low\_res}}.$$

Note that the result of the high resolution case can be derived by setting $p_{j,k} = 1$, $b^2(p_{j,k}) = 1$ (because $\sigma_\Theta^2 = \sigma_0^2$), and $k_j = n_j$.

Bibliography

Abramowitz, M. and Stegun, I. (1965). *Handbook of Mathematical Functions.* Wiley, Dover, NY, USA.

Baba, K. and Jurkovic, D. (1993). *Three-dimensional ultrasound in obstetrics and gynecology.* Parthenon Publishing Group Ltd., New York, NY, USA.

Bamber, J. C., Eckersley, R. J., Hubregtse, P., Bush, N. L., Bell, D. S., and Crawford, D. C. (1992). Data processing for 3-D ultrasound visualization of tumour anatomy and blood flow. In Robb, R. A., editor, *Proceedings of Visualization in Biomedical Computing*, SPIE 1808, pages 651–663. International Society of Optical Engineering, Bellingham, WA, USA.

Barrett, W. and Mortensen, E. (1997). Interactive live-wire boundary extraction. *Medical Image Analysis*, 1(4):331–341.

Barry, C., Allott, C., John, N., Mellor, P., Arundel, P., Thomson, D., and Waterton, J. (1997). Three-dimensional freehand ultrasound: image reconstruction and volume analysis. *Ultrasound in Medicine and Biology*, 23(8):1209–1224.

Basset, O., Gimenez, G., Mestas, J. L., Cathignol, D., and Devonec, M. (1991). Volume measurement by ultrasonic transverse or sagittal cross-sectional scanning. *Ultrasound in Medicine and Biology*, 17(3):291–296.

Belohlavek, M., Foley, D. A., Gerber, T. C., Kinter, T. M., Greenleaf, J. F., and Seward, J. B. (1993). Three- and four-dimensional cardiovascular ultrasound imaging: a new era for echocardiography. *Mayo Clinic Proceedings*, 68(3):221–240.

Bouma, C. J., Niessen, W. J., Zuiderveld, K. J., Gussenhoven, E. J., and Viergever, M. A. (1996). Evaluation of segmentation algorithms for intravascular ultrasound images. In Höhne, K. H. and Kikinis, R., editors, *Proceedings of Visualization in Biomedical Computing*, Lecture Notes in Computer Science 1131, pages 203–212. Springer-Verlag, Heidelberg, Germany.

Bouma, C. J., Niessen, W. J., Zuiderveld, K. J., Gussenhoven, E. J., and Viergever, M. A. (1997). Automated lumen definition from 30 MHz intravascular ultrasound images. *Medical Image Analysis*, 1(4):363–377.

Brown, L. G. (1992). A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376.

Burckhardt, C. (1978). Speckle in ultrasound B-mode scans. *IEEE Transactions on Sonics and Ultrasonics*, 25(1):1–6.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698.

Carr, J. (1996). *Surface Reconstruction in 3D Medical Imaging.* Ph.D. Thesis, University of Canterbury, Christchurch, New Zealand.

Chervenak, F., Isaacson, G., and Campbell, S. (1993). *Ultrasound in Obstetrics and Gynecology.* Little, Brown and Company, Boston, MA, USA.

Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. (1995). Active shape models — their training and application. *Computer Vision and Image Understanding*, 61(1):38–59.

Cosgrove, D., Meire, H., and Dewbury, K. (1993). *Abdominal and General Ultrasound.* Churchill Livingstone, Edinburgh, UK.

Craig, J. J. (1989). *Introduction to Robotics: Mechanics and Control.* Addison-Wesley, Reading, MA, USA.

Crawford, D. C., Bell, D. S., and Bamber, J. C. (1993). Compensation for the signal processing characteristics of ultrasound B-mode scanners in adaptive speckle reduction. *Ultrasound in Medicine and Biology*, 19(6):469–485.

Curie, J. and Curie, P. (1880). Développement par pression, de l'électricité polaire dans les cristaux hémièdres á faces inclinées. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, 91:294–294.

Cusumano, A., Jackson, D., Silverman, R. H., Reinstein, D. Z., Rondeau, M. J., Ursea, R., Daly, S. M., and Lloyd, H. O. (1998). Three-dimensional ultrasound imaging. *Opthalmology*, 105(2):300–306.

Deng, J., Gardener, J. E., Rodeck, C. H., and Lees, W. R. (1996). Fetal echocardiography in 3-dimensions and 4-dimensions. *Ultrasound in Medicine and Biology*, 22(8):979–986.

Detmer, P. R., Bashein, G., Hodges, T., Beach, K. W., Filer, E. P., Burns, D. H., and Strandness Jr., D. (1994). 3D ultrasonic image feature localization based on magnetic scanhead tracking: in vitro calibration and validation. *Ultrasound in Medicine and Biology*, 20(9):923–936.

Downey, D., Nicolle, D., Levin, M., and Fenster, A. (1996). Three-dimensional ultrasound imaging of the eye. *Eye*, 10:75–81.

Dussik, K. T. (1942). Über die Möglichkeit hochfrequente mechanische Schwingungen als diagnostisches Hilfsmittel zu verwerten. *Neurologische Psychiatrie*, 174:153–168.

Evans, A. N. and Nixon, M. S. (1995). Mode filtering to reduce ultrasound speckle for feature extraction. *IEE Proceedings of Vision, Image and Signal Processing*, 142(2):87–94.

Evans, D. H. and Martin, K. (1988). *Physics in Medical Ultrasound II.* Institute of Physical Sciences in Medicine, London, UK.

Evans, J. A. (1986). *Physics in Medical Ultrasound.* Institute of Physical Sciences in Medicine, London, UK.

Fenster, A. and Downey, D. B. (1996). 3-D ultrasound imaging — a review. *IEEE Engineering in Medicine and Biology Magazine*, 15(6):41–51.

Fine, D., Perring, S., Herbetko, J., Hacking, C. N., Fleming, J. S., and Dewbury, K. C. (1991). Three-dimensional (3D) ultrasound imaging of the gallbladder and dilated biliary tree: reconstruction from real-time B-scans. *British Journal of Radiology*, 64:1056–1057.

Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395.

Fish, P. (1990). *Physics and Instrumentation of Diagnostic Medical Ultrasound.* John Wiley and Sons Ltd., Chichester, UK.

Foley, T. (1987). Interpolation and approximation of 3-D and 4-D scattered data. *Computers and Mathematics with Applications*, 13(8):711–740.

Franceschi, D., Bondi, J. A., and Rubin, J. R. (1992). A new approach for three-dimensional reconstruction of arterial ultrasonography. *Journal of Vascular Surgery*, 15(5):800–805.

Franke, R. (1982). Scattered data interpolation: tests of some methods. *Mathematics of Computation*, 38(157):181–200.

Gilja, O. H., Smievoll, A. I., Thune, N., Matre, K., Hausken, T., Odegaard, S., and Berstad, A. (1995). In vivo comparison of 3D ultrasonography and magnetic resonance imaging in volume estimation of human kidneys. *Ultrasound in Medicine and Biology*, 21(1):25–32.

Hackmann, W. (1984). *Seek and Strike: Sonar anti-submarine warfare and the Royal Navy 1914-1954.* Her Majesty's Stationary Office, London, UK.

Halliwell, M., Key, H., Jenkins, D., Jackson, P. C., and Wells, P. N. T. (1989). New scans from old: digital reformatting of ultrasound images. *The British Journal of Radiology*, 62(741):824–829.

Hardy, T. (1878). *The Return of the Native.* Osgood and McIlvaine, London, UK.

Harris, R. A., Follett, D. H., Halliwell, M., and Wells, P. N. T. (1991). Ultimate limits in ultrasonic imaging resolution. *Ultrasound in Medicine and Biology*, 17(6):547–558.

He, P. (1997). Spatial compounding in 3D imaging of limbs. *Ultrasonic Imaging*, 19(4):251–265.

Hernandez, A., Basset, O., Chirossel, P., and Gimenez, G. (1996). Spatial compounding in ultrasonic imaging using an articulated scan arm. *Ultrasound in Medicine and Biology*, 22(2):229–238.

Hill, C. (1986). *Physical Principles of Medical Ultrasonics.* John Wiley and Sons, New York, NY, USA.

Hodges, T. C., Detmer, P. R., Burns, D. H., Beach, K. W., and Strandness Jr., D. (1994). Ultrasonic three-dimensional reconstruction: in vitro and in vivo volume and area measurement. *Ultrasound in Medicine and Biology*, 20(8):719–729.

Hogg, R. and Ledolter, J. (1987). *Engineering Statistics.* Macmillan Publishing Company, New York, NY, USA.

Hollerbach, J. M. and Wampler, C. W. (1996). The calibration index and taxonomy for robot kinematic calibration methods. *International Journal of Robotics Research*, 15(6):573–591.

Hottier, F. and Collet Billon, A. (1990). 3D echography: status and perspective. In Höhne, K., Fuchs, H., and Pizer, S., editors, *3D Imaging in Medicine: Algorithms, Systems, Applications*, NATO ASI Series, pages 21–41. Springer-Verlag, Berlin, Germany.

Howry, D. H., Posakony, G. J., Cushman, C. R., and Holmes, J. H. (1956). Three-dimensional and stereoscopic observation of body structures by ultrasound. *Journal of Applied Physiology*, 9:304–306.

Hughes, S. W., D'Arcy, T. J., Maxwell, D. J., Chiu, W., Milner, A., Saunders, J. E., and Sheppard, R. J. (1996). Volume estimation from multiplanar 2D ultrasound images using a remote electromagnetic position and orientation sensor. *Ultrasound in Medicine and Biology*, 22(5):561–572.

Hull, B. and Vernon, J. (1988). *Non-destructive Testing.* Macmillan Education, Basingstoke, UK.

Hunt, F. V. (1954). *Electroacoustics: the analysis of transduction, and its historical background.* Harvard University Press, Boston, MA, USA.

Kerr, A., Patterson, M., Foster, F., and Hunt, J. (1986). Speckle reduction in pulse echo imaging using phase insensitive and phase sensitive signal processing techniques. *Ultrasonic Imaging*, 8:11–28.

King, D. L., King Jr., D. L., and Shao, M. Y. (1991). Evaluation of in vitro measurement accuracy of a three-dimensional ultrasound scanner. *Journal of Ultrasound in Medicine*, 10:77–82.

Kremkau, F. (1993). *Diagnostic Ultrasound: Principles and Instruments.* W.B. Saunders Company, Philadelphia, PA, USA.

Lalouche, R. C., Bickmore, D., Tessler, F., Mankovich, N. J., Huang, H. K., and Kangarloo, H. (1989). Three-dimensional reconstruction of ultrasound images. In *Proceedings of Medical Imaging III*, pages 450–457, Newport Beach, CA, USA.

Lawson, C. L. and Hanson, R. J. (1974). *Solving Least Squares Problems.* Prentice-Hall, Englewood Cliffs, NJ, USA.

Lee, A., Deutinger, J., and Bernaschek, G. (1995). Three dimensional ultrasound: abnormalities of the fetal face in surface and volume rendering mode. *British Journal of Obstetrics and Gynaecology*, 102:302–306.

Leeman, S. and Seggie, D. (1987). Speckle reduction via phase. In Ferrari, L., editor, *Proceedings of the International Symposium on Pattern Recognition and Acoustical Imaging*, SPIE 768, pages 173–177. International Society of Optical Engineering, Bellingham, WA, USA.

Leotta, D. F., Detmer, P. R., Gilja, O. H., Jong, J. M., Martin, R. W., Primozich, J. F., Beach, K. W., and Strandness, D. E. (1995). Three-dimensional ultrasound imaging using multiple magnetic tracking systems and miniature magnetic sensors. In *Proceedings of the IEEE Ultrasonics Symposium*, pages 1415–1418, Seattle, WA, USA.

Leotta, D. F., Detmer, P. R., and Martin, R. W. (1997). Performance of a miniature magnetic position sensor for three-dimensional ultrasound imaging. *Ultrasound in Medicine and Biology*, 24(4):597–609.

Levine, D. N., Pelizzari, C. A., Chen, G. T. Y., Chen, C. T., and Cooper, M. D. (1988). Retrospective geometric correlation of MR, CT and PET images. *Radiology*, 169(3):817–823.

Light, E. D., Davidsen, R. E., Fiering, J. O., Hruschka, T. A., and Smith, S. W. (1998). Progress in two-dimensional arrays for real-time volumetric imaging. *Ultrasonic Imaging*, 20(1):1–15.

Lorensen, W. E. and Cline, H. (1987). Marching cubes: a high resolution 3D surface construction method. *Computer Graphics*, 21(4):163–169.

Maintz, J. B. A., van den Elsen, P. A., and Viergever, M. A. (1996). Comparison of edge-based and ridge-based registration of CT and MR brain images. *Medical Image Analysis*, 1(2):151–161.

Maintz, J. B. A. and Viergever, M. A. (1998). A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36.

Martin, R. W., Bashein, G., Detmer, P. R., and Moritz, W. E. (1990). Ventricular volume measurement from a multiplanar transesophageal ultrasonic imaging system: an in vitro study. *IEEE Transactions on Biomedical Engineering*, 37(5):442–449.

McCallum, W. D. and Brinkley, J. F. (1979). Estimation of fetal weight from ultrasonic measurements. *American Journal of Obstetrics and Gynecology*, 133(2):195–200.

McCann, H., Sharp, J., Kinter, T., McEwan, C., Barillot, C., and Greenleaf, J. (1988). Multidimensional ultrasonic imaging for cardiology. *Proceedings of the IEEE*, 76(9):1063–1073.

Mitáš, L. and Mitášová, H. (1988). General variational approach to the interpolation problem. *Computers and Mathematics with Applications*, 16(12):983–992.

Mitášová, H. and Mitáš, L. (1993). Interpolation by regularized splines with tension: I. theory and implementation. *Mathematical Geology*, 25(6):641–655.

Mitášová, H., Mitáš, L., Brown, W., Gerdes, D., Kosinovsky, I., and Baker, T. (1995). Modelling spatially and temporally distributed phenomena: new methods and tools for GRASS GIS. *International Journal of Geographical Information Systems*, 9(4):433–446.

Monga, O., Deriche, R., Malandain, G., and Cocquerez, J. P. (1991). Recursive filtering and edge tracking: two primary tools for 3-D edge detection. *Image and Vision Computing*, 9(4):203–214.

More, J. J. (1977). The Levenberg-Marquardt algorithm: implementation and theory. In Watson, A., editor, *Numerical Analysis*, Lecture Notes in Mathematics 630, pages 105–116. Springer-Verlag, Heidelberg, Germany.

Moshfeghi, M. and Waag, R. C. (1988). In vivo and in vitro ultrasound beam distortion measurements of a large aperture and a conventional aperture focussed transducer. *Ultrasound in Medicine and Biology*, 14(5):415–428.

Moskalik, A., Carson, P. L., Meyer, C. R., Fowlkes, J. B., Rubin, J. M., and Roubidoux, M. A. (1995). Registration of three-dimensional compound ultrasound scans of the breast for refraction and motion correction. *Ultrasound in Medicine and Biology*, 21(6):769–778.

Nelson, T. and Pretorius, D. (1994). 3D ultrasound image quality improvement using spatial compounding and 3D filtering. *Medical Physics*, 21(6):998–998.

Nelson, T. and Pretorius, D. (1997). Interactive acquisition, analysis and visualization of sonographic volume data. *International Journal of Imaging Systems and Technology*, 8:26–37.

Nelson, T. R. and Elvins, T. T. (1993). Visualization of 3D ultrasound data. *IEEE Computer Graphics and Applications*, 13(6):50–57.

Neveu, M., Faudot, D., and Derdouri, B. (1994). Recovery of 3D deformable models from echocardiographic images. In Bookstein, F. L., Duncan, J. S., Lange, N., and Wilson, D. C., editors, *Proceedings of Mathematical Methods in Medical Imaging III*, SPIE 2299, pages 367–376. International Society of Optical Engineers, Bellingham, WA, USA.

Ng, K. J., Gardener, J. E., Rickards, D., Lees, W. R., and Milroy, E. J. G. (1994). Three-dimensional imaging of the prostatic urethra — an exciting new tool. *British Journal of Urology*, 74(5):604–608.

Nielson, G. (1993). Scattered data modeling. *IEEE Computer Graphics and Applications*, 13(1):60–70.

Nixon, M. A., McCallum, B. C., Fright, W. R., and Price, N. B. (1998). The effects of metals and interfering fields on electromagnetic trackers. *Presence*, 7(2):204–218.

Ofili, E. O. and Nanda, N. C. (1994). Three-dimensional and four-dimensional echocardiography. *Ultrasound in Medicine and Biology*, 20(8):669–675.

Ohbuchi, R., Chen, D., and Fuchs, H. (1992). Incremental volume reconstruction and rendering for 3D ultrasound imaging. In Robb, R. A., editor, *Proceedings of Visualization in Biomedical Computing*, SPIE 1808, pages 312–323. International Society of Optical Engineering, Bellingham, WA, USA.

Pasterkamp, G., Borst, C., Moulaert, A. F. S. R., Bouma, C. J., van Dijk, D., Kluytmans, M., and ter Haar Romeny, B. M. (1995). Intravascular ultrasound image subtraction: a contrast enhancing technique to facilitate automatic three-dimensional visualization of the arterial lumen. *Ultrasound in Medicine and Biology*, 21(7):913–918.

Prager, R. W. (1997). Ultrasound machine calibration. UK patent application no. 9716994.0.

Prager, R. W., Gee, A. H., and Berman, L. (1998a). 3D ultrasound without voxels. In *Proceedings of Medical Image Understanding and Analysis*, pages 93–96, Leeds, UK.

Prager, R. W., Rohling, R. N., Gee, A. H., and Berman, L. (1998b). Rapid calibration for 3-D freehand ultrasound. *Ultrasound in Medicine and Biology*, 24(6):855–869.

Preston, R., editor (1991). *Output Measurements for Medical Ultrasound*. Springer-Verlag, Berlin, Germany.

Rahmouni, A., Yang, A., Tempany, C. M. C., Frenkel, T., Epstein, J., Walsh, P., Leichner, P. K., Ricci, C., and Zerhouni, E. (1992). Accuracy of in-vivo assessment of prostatic volume by MRI and transrectal ultrasonography. *Journal of Computer Assisted Tomography*, 16(6):935–940.

Rankin, R. N., Fenster, A., Downey, D. B., Munk, P. L., Levin, M. F., and Vellet, A. D. (1993). Three-dimensional sonographic reconstruction: techniques and diagnostic applications. *American Journal of Roentgenology*, 161(4):695–702.

Rayleigh, 3rd Baron (Lord) (1877). *Theory of Sound.* Macmillan, London, UK.

Richardson, M. L. F. (1912). Apparatus for warning a ship at sea of its nearness to large objects wholly or partly under water. UK patent no. 1125.

Rohling, R., Peters, T., Hollerbach, J., and Munger, P. (1994). Comparison of relative accuracy between an optical and a mechanical position tracker for image guided neurosurgery. *Journal of Imaged Guided Surgery*, 1(1):30–34.

Sakas, G., Schreyer, L. A., and Grimm, M. (1995). Preprocessing and volume rendering of 3D ultrasonic data. *IEEE Computer Graphics and Applications*, 15(4):47–54.

Sakas, G. and Stefan, W. (1995). Extracting surfaces from fuzzy 3D-ultrasound data. In *Proceedings of SIGGRAPH*, pages 465–474, Los Angeles, CA, USA.

Salustri, A. and Roelandt, J. R. T. C. (1995). Ultrasonic three-dimensional reconstruction of the heart. *Ultrasound in Medicine and Biology*, 21(3):281–293.

Schröer, K. (1993). Theory of kinematic modelling and numerical procedures for robot calibration. In Bernhardt, R. and Albright, S. L., editors, *Robot Calibration*, pages 157–196. Chapman and Hall, London, UK.

Shahidi, R., Tombropoulos, R., and Grzeszczuk, R. P. (1998). Clinical applications of three-dimensional rendering of medical data sets. *Proceedings of the IEEE*, 86(3):555–568.

Shepard, D. (1968). A two-dimensional interpolation function for irregularly spaced data. In *Proceedings of the 23rd National Conference of the ACM*, pages 517–523. ACM Press, New York, NY, USA.

Sherebrin, S., Fenster, A., Rankin, R., and Spence, D. (1996). Freehand three-dimensional ultrasound: implementation and applications. In van Metter, R. L. and Beutel, J., editors, *Proceedings of Medical Imaging 1996*, SPIE 2708, pages 296–303. International Society of Optical Engineering, Bellingham, WA, USA.

Skolnick, M. L. (1991). Estimation of ultrasound beam width in the elevation (section thickness) plane. *Radiology*, 180(1):286–288.

Soferman, Z., Blythe, D., and John, N. W. (1998). Advanced graphics behind medical virtual reality: evolution of algorithms, hardware, and software interfaces. *Proceedings of the IEEE*, 86(3):531–554.

State, A., Chen, D. T., Tector, C., Brandt, A., Chen, H., Ohbuchi, R., Bajura, M., and Fuchs, H. (1994). Case study: observing a volume rendered fetus within a pregnant patient. In *Proceedings of IEEE Visualization, 1994*, pages 364–368, Washington, DC, USA.

Steen, E. and Olstad, B. (1994). Volume rendering of 3D medical ultrasound data using direct feature mapping. *IEEE Transactions on Medical Imaging*, 13(3):517–525.

Steiner, H., Staudach, A., Spitzer, D., and Schaffer, H. (1994). Three-dimensional ultrasound in obstetrics and gynaecology: technique, possibilities and limitations. *Human Reproduction*, 9(9):1773–1778.

Talmi, A. and Gilat, G. (1977). Method for smooth approximation of data. *Journal of Computational Physics*, 23:93–123.

Taylor, R. H., Lavallee, S., Burdea, G. C., and Mosges, R. (1996). *Computer-Integrated Surgery*. MIT Press, Cambridge, MA, USA.

Thijssen, J., Oosterveld, B., and Wagner, R. (1988). Gray level transforms and lesion detectability in echographic images. *Ultrasonic Imaging*, 10:171–195.

Thirion, J. P., Monga, O., Benayoun, S., Gueziec, A., and Ayache, N. (1992). Automatic registration of 3D images using surface curvature. In *Proceedings of Mathematical Methods in Medical Imaging*, pages 206–216, San Diego, CA, USA.

Trahey, G. E. ., Allison, J. W., Smith, S. W., and von Ramm, O. T. (1986a). A quantitative approach to speckle reduction via frequency compounding. *Ultrasonic Imaging*, 8:151–164.

Trahey, G. E., Smith, S. W., and von Ramm, O. T. (1986b). Speckle pattern correlation with lateral aperture translation: experimental results and implications for spatial compounding. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 33(3):257–264.

Trobaugh, J. W., Trobaugh, D. J., and Richard, W. D. (1994). Three-dimensional imaging with stereotactic ultrasonography. *Computerized Medical Imaging and Graphics*, 18(5):315–323.

Udupa, J., Goncalves, R., Iyer, K., Narendula, S., Odhner, D., Samarasekera, S., and Sharma, S. (1993). 3DVIEWNIX: an open, transportable software system for the visualization and analysis of multidimensional, multimodality, multiparametric images. In Kim, Y. M., editor, *Proceedings of Medical Imaging 1993*, SPIE 1897, pages 47–58. International Society of Optical Engineering, Bellingham, WA, USA.

Udupa, J. K. and Dewey, O. (1991). Fast visualization, manipulation, and analysis of binary volumetric objects. *IEEE Computer Graphics and Applications*, 11(6):53–62.

van den Elsen, P., Pol, E. J. D., and Viergever, M. (1993). Medical image matching: a review with classification. *IEEE Engineering in Medicine and Biology*, 12(1):26–39.

van den Elsen, P. A., Maintz, J. B. A., Pol, E. J. D., and Viergever, M. A. (1995). Automatic registration of CT and MR brain images using correlation of geometrical features. *IEEE Transactions on Medical Imaging*, 14(2):384–396.

Viola, P. and Wells, W. M. (1997). Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154.

Watt, A. H. (1993). *3D computer graphics*. Addison-Wesley, Wokingham, UK.

Wells, P. N. T. (1993). *Advances in Ultrasound Techniques and Instrumentation*. Churchill Livingstone Inc., New York, NY, USA.

Wells, W. M., Viola, P., Atsumi, H., Nakajima, S., and Kikinis, R. (1996). Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1(1):35–51.

Whittingham, T. A. (1991). Resolution and information limitations from transducer arrays. *Physics in Medicine and Biology*, 36(11):1503–1514.

Whittingham, T. A. (1997). New and future developments in ultrasonic imaging. *The British Journal of Radiology*, 70(SISI):S119–S132.