AUTOMATIC CALIBRATION FOR 3-D FREE-HAND ULTRASOUND

R.W. Prager, R.N. Rohling,
A.H. Gee and L. Berman

CUED/F-INFENG/TR 303

September 1997

Cambridge University Engineering Department
Trumpington Street
Cambridge CB2 1PZ
England

E-mail: `rwp@eng.cam.ac.uk`, `rnr20@eng.cam.ac.uk`, `ahg@eng.cam.ac.uk`,
`lb@radiol.cam.ac.uk`

Abstract

3-D free-hand ultrasound is a new imaging technique which is rapidly finding clinical applications. A position sensing device is attached to a conventional ultrasound probe, so that as B-scans are acquired they can be labelled with their relative positions and orientations. This allows the B-scans to be inserted into a 3-D voxel array, which can then be visualised using any-plane slicing, volume rendering or surface rendering. A key requirement of all free-hand imaging systems is **calibration**: determining the position and orientation of the B-scan with respect to the position sensor. This is typically a lengthy and tedious process, which may need repeating every time a sensor is mounted on a probe. This paper describes a new calibration technique, which takes only a few minutes to perform and produces more accurate and repeatable results than all previously published alternatives.

1 Introduction

3-D ultrasound is a new imaging modality that has already been recognised as a valuable tool for a variety of clinical applications. Conventional 2-D diagnostic imaging uses a hand-held probe which transmits ultrasound pulses into the body and receives the echoes. The magnitude and timing of the echoes are used to create a 2-D grey-level image (B-scan) of a cross-section of the body in the scan plane. 3-D ultrasound extends this concept so that volumes of intensity data are created from pulse-echo information.

High quality, instantaneous 3-D imaging remains a long term research goal. One promising approach centers around the development of a new type of phased array probe, which sends and receives echos from a 2-D array of elements (instead of the usual 1-D array). Unfortunately, several technical challenges must be overcome before such probes receive clinical acceptance [20]. Alternative approaches, which make use of conventional 2-D ultrasound technology, include the free-hand and swept volume techniques [17, 22]. Instead of taking an instantaneous 3-D snapshot, these techniques construct a 3-D data set from a number of 2-D B-scans acquired in rapid succession.

The swept volume approach uses a special probe to mechanically sweep the plane of the B-scan through a volume of interest. In contrast, the free-hand approach relies on the physician to guide a standard probe over the volume, while a position sensor, attached to the probe, measures the B-scans' relative positions. Each method has its advantages. Swept volume systems are easy to use and produce standardised volumes of densely sampled data without gaps. However, they require the considerable expense of a dedicated machine, and are also limited to a maximum volume dictated by hardware constraints in the probe. Free-hand systems can be used to obtain arbitrary volumes of data, since the motion of the probe is unconstrained. They are also cheaper, requiring only existing, conventional ultrasound systems and relatively inexpensive additional components. However, the physician needs to learn how to move the probe to acquire regular, densely sampled data sets.

The cost and flexibility of free-hand imaging ensures it remains a popular choice, and with recent improvements in the speed of data acquisition, combined with careful scanning practice, the quality and resolution of free-hand data sets is now comparable to those obtained using swept volume systems. For these reasons, research into free-hand systems is very active, and several commercial systems have recently become available.

A 3-D free-hand examination can be broken into three stages: scanning, reconstruction and visualisation: see Figure 1. Before scanning, some sort of position sensor is attached to the probe. This is typically the receiver of an electro-magnetic position sensor, as illustrated in Figure 1. Measurements from the position sensor are used to determine the positions and orientations of the B-scans with respect to the fixed transmitter. In the next stage, the set of acquired B-scans and their relative positions are used to fill a regular voxel array. Finally, this voxel array is visualised using, for example, any-plane slicing, volume rendering or surface rendering (after segmentation).

A key requirement of all free-hand imaging systems is calibration. This involves determining the position and orientation of the B-scans 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 must be added to the sensor measurements to calculate the positions of

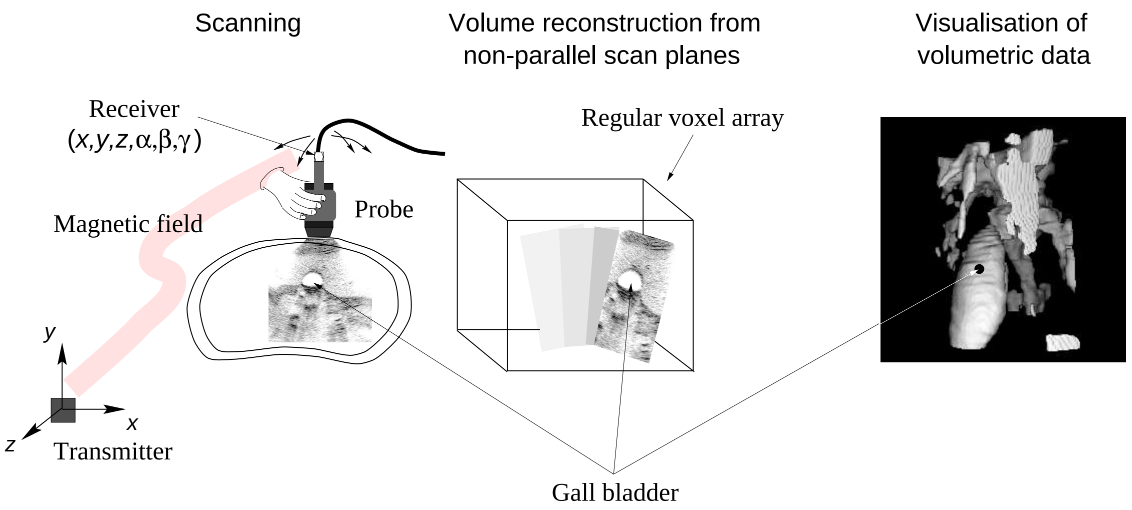


Figure 1: **3-D free-hand ultrasound imaging**. Clinical free-hand imaging involves scanning, reconstruction and visualisation. This example illustrates an examination of a gall bladder.

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. Previous work has described the use of acoustic spark gaps [9], mechanical arms [15], AC magnetic sensors [8, 14], DC magnetic sensors [4, 12], optical sensors [21, 23], and multiple sensors [11]. Only a few, however, have seriously considered the question of accurate calibration.

A rough estimate of the calibration parameters can be obtained by external measurements of the probe and position sensor. However, the origin of the sensor (for example, the centre of the wire coils in an electro-magnetic receiver) and the corner of the B-scan are not well defined with respect to the external cases of the sensor and probe. Furthermore, it is common to allow arbitrary cropping of the B-scans during acquisition, which also affects calibration. For these reasons, calibration must be 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.

Existing calibration techniques are limited in both their accuracy and ease of use. A typical calibration process, which often needs repeating every time a sensor is mounted on a probe, takes several hours for a skilled technician to perform. In this paper we present a novel calibration technique which can be performed in a few minutes and improves on the accuracy of all previously published alternatives. The paper is organised as follows. In Section 2 we describe the reconstruction process and formulate the calibration problem. Section 3 reviews existing calibration procedures and introduces the new technique which is the focus of this paper. We compare the accuracy, repeatability and usability of the techniques in Section 4, and draw conclusions in Section 5. The appendices contain auxiliary material referenced in the main text.

2 Reconstruction

Figure 2 shows the four coordinate systems used for reconstruction. **P** is the coordinate system of the B-scan plane, with an origin in the top left hand corner of the cropped image. The y-axis is in the beam direction, and the x-axis in the lateral direction. The z-axis is in the elevational

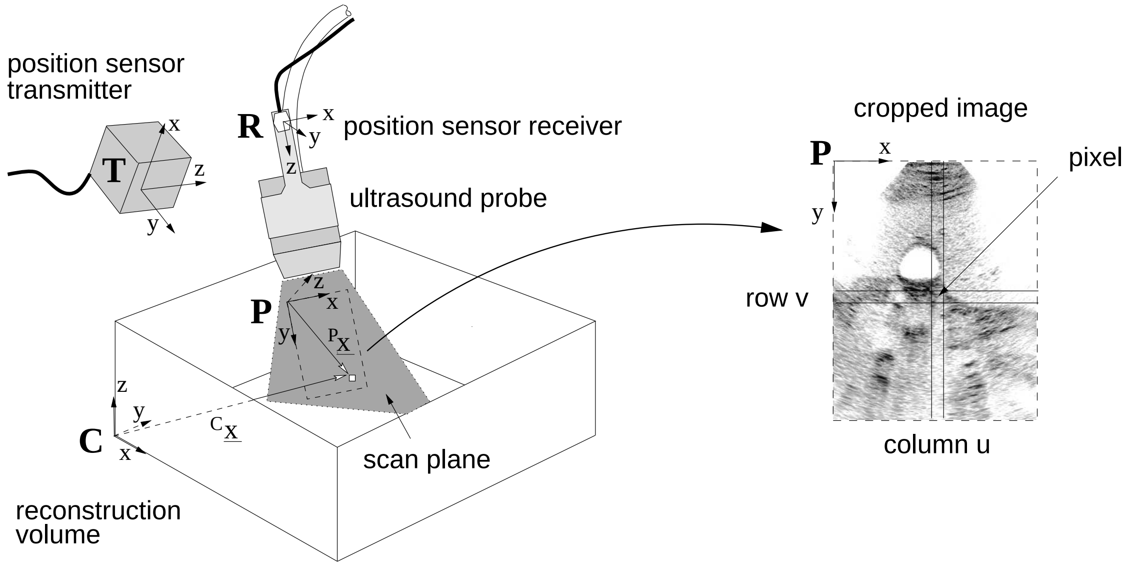


Figure 2: **Coordinate systems**. The reconstruction process uses four coordinate systems. Note that the transmitter is often placed on a cantilevered support above the examination table, to give an unobstructed path to the receiver.

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 3-D matrix **C** of volume elements (voxels). The origin of **C**'s coordinate system is 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 **C**. Each pixel's scan plane 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:

$$^C\underline{\mathbf{x}} = {}^C\mathbf{T}_T \ {}^T\mathbf{T}_R \ {}^R\mathbf{T}_P \ {}^P\underline{\mathbf{x}} \tag{1}$$

where

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

We adopt a standard notation, such 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 coordinate systems has six degrees of freedom: three rotation (α, β, γ) and three translation (x, y, z). There are many ways to parameterise a rotation using three angles: we adopt an x-y-z fixed angles scheme [3]. 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 *fixed* rotation axes are aligned with the first coordinate system. Using this convention, the homogeneous matrix describing the transformation takes the following form [3]:

$$
{}^{J}\mathbf{T}_I(x,y,z,\alpha,\beta,\gamma) =
\begin{pmatrix}
\cos\alpha\cos\beta & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma & x \\
\sin\alpha\cos\beta & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma & y \\
-\sin\beta & \cos\beta\sin\gamma & \cos\beta\cos\gamma & z \\
0 & 0 & 0 & 1
\end{pmatrix} \tag{2}
$$

Each of the transformation matrices in equation (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. We could omit it and align the reconstruction volume with the transmitter, but this might, for instance, place the B-scans a long way from the origin of **C**, resulting in a vast, largely empty voxel array. Or 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. While the scale factors s_x and s_y could be derived from the axis markings on the B-scan, we have decided to include them in the calibration for more accurate and automatic estimates.

Once ${}^{C}\underline{\mathbf{x}}$ has been found for every pixel, the voxels of **C** can be set according to the intensities of the pixels they intersect. The reconstruction algorithm is described in more detail in [18]. After a significant portion of the voxels are filled, the volume can be visualised by a number of methods: in this paper we use surface rendering and any-plane slicing [5].

3 Calibration Phantoms

3.1 Cross-wire phantom

Calibration is performed by scanning a phantom of known geometric dimensions. Equations similar to (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.

A common calibration phantom is the cross-wire phantom [4]. Two intersecting wires are mounted in a water bath, with the transmitter placed at some fixed location with respect to the wires — see Figure 3(a). For calibration purposes, the origin of **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.

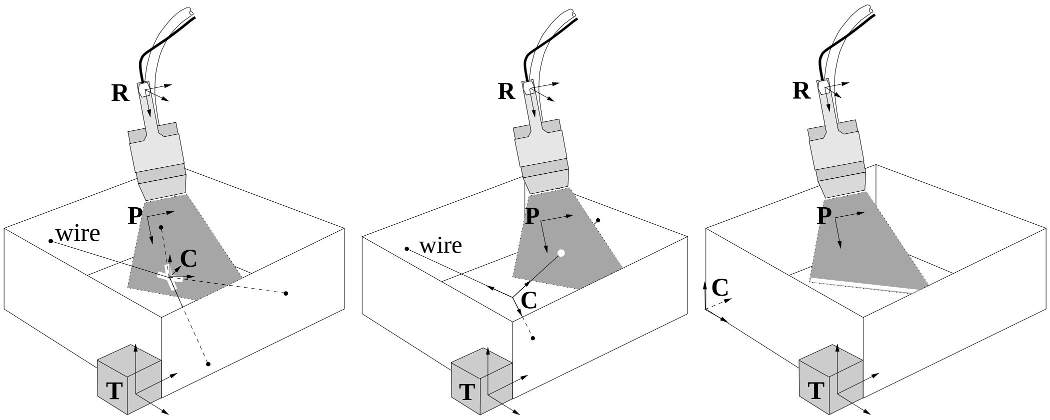
The location where the wires cross is scanned repeatedly from different directions, with each B-scan showing a detectable cross (the reader may wish to skip ahead to Figure 8(a) for a typical example). The pixel at the center of the cross should satisfy:

$$
\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = {}^{C}\mathbf{T}_T\ {}^{T}\mathbf{T}_R\ {}^{R}\mathbf{T}_P \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix} \tag{3}
$$

The first three rows of (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 $3m$:

$$
\mathbf{0} = \mathbf{f}(\theta, \phi) \tag{4}
$$

where θ are the measurements and ϕ are the unknowns. The system can be solved by several iterative methods: we choose the popular and robust Levenberg-Marquardt algorithm [13]. At iteration j, the update $\Delta\phi$ to the current estimate ϕ_j is derived from the first order Taylor expansion



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

Figure 3: **Calibration phantoms.** In (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. In (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. In (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.

of (4):

$$\begin{aligned} \mathbf{0} = \mathbf{f}(\theta, \phi) \quad &\approx \quad \mathbf{f}(\theta, \phi_j) + \frac{\partial \mathbf{f}(\theta, \phi_j)}{\partial \phi}(\phi - \phi_j) \\ \Rightarrow \quad \Delta \mathbf{f} \quad &= \quad \mathbf{J}(\phi - \phi_j) = \mathbf{J}\Delta\phi \end{aligned} \tag{5}$$

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}:

$$\phi_{j+1} = \phi_j + (\mathbf{J}^T\mathbf{J} + \varepsilon\mathbf{I})^{-1}\mathbf{J}^T\Delta\mathbf{f} \tag{6}$$

where ε is a damping term chosen at each step to stabilise convergence. 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.

For calibration, we are only interested in the scale values and ${}^R\mathbf{T}_P$, but we must also solve for ${}^C\mathbf{T}_T$, even though we will subsequently discard these values and adopt an arbitrary, convenient ${}^C\mathbf{T}_T$ for reconstruction, as explained in Section 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 ${}^C\mathbf{T}_T$. However, it is clear from inspection of the geometry that the coordinate system $\mathbf{C}$ can be at any orientation and still satisfy (3). This means that the three orientation angles of ${}^C\mathbf{T}_T$ are not identifiable. These angles are therefore removed from ϕ, and for convenience set to zero in (4). Table 1 lists the identifiable parameters ϕ.

Several papers [11, 12, 21] 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.

Accuracy of calibration with a point object depends on how well the center of the point can be located, as well as the stationarity of the point with respect to the transmitter **T**. Given the

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

Table 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.

quality of ultrasound images, it is common practice to locate the point by hand in each B-scan, making the calibration process tedious and time-consuming.

3.2 Three-wire phantom

Another calibration technique uses a three-wire phantom [1], as shown in Figure 3(b). This is similar to the cross-wire method since it involves scanning wires in a water bath, but it requires that the three wires are accurately mounted in orthogonal directions. The accuracy of calibration therefore depends on the orthogonality, straightness and stationarity of the wires.

We place the coordinate system $\mathbf{C}$ at the origin of the wires, and orient the x,y and z axes along the wires. Each wire is scanned, one at a time, along its length from a variety of directions. The wire appears as a detectable dot in the B-scan — see Figure 8(b). For the wire along the x-axis, the pixel at the center of the wire should satisfy:

$$\begin{pmatrix} x \\ 0 \\ 0 \\ 1 \end{pmatrix} = {}^C\mathbf{T}_T \; {}^T\mathbf{T}_R \; {}^R\mathbf{T}_P \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix} \tag{7}$$

The left hand side of the equation becomes $(0 \ y \ 0 \ 1)^T$ for the y-axis and $(0 \ 0 \ z \ 1)^T$ 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 1. As with the cross-wire phantom, many B-scans are acquired, 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 centered on a crossing point. However, each wire must be scanned separately, keeping track of which wire each B-scan intersects.

3.3 Single-wall phantom

While scanning the cross-wire and three-wire phantoms, we observed that the floor of the phantom produced clear, consistent lines in the B-scans. We therefore removed the wires and performed calibration using the floor of the water bath alone, as shown in Figure 3(c). The accuracy of the calibration will depend on the flatness and stationarity of the floor.

The idea of calibrating a 3D ultrasound system by scanning a wall of a water bath is intuitively attractive. Simple water baths with planar surfaces are readily available, so there is no need to

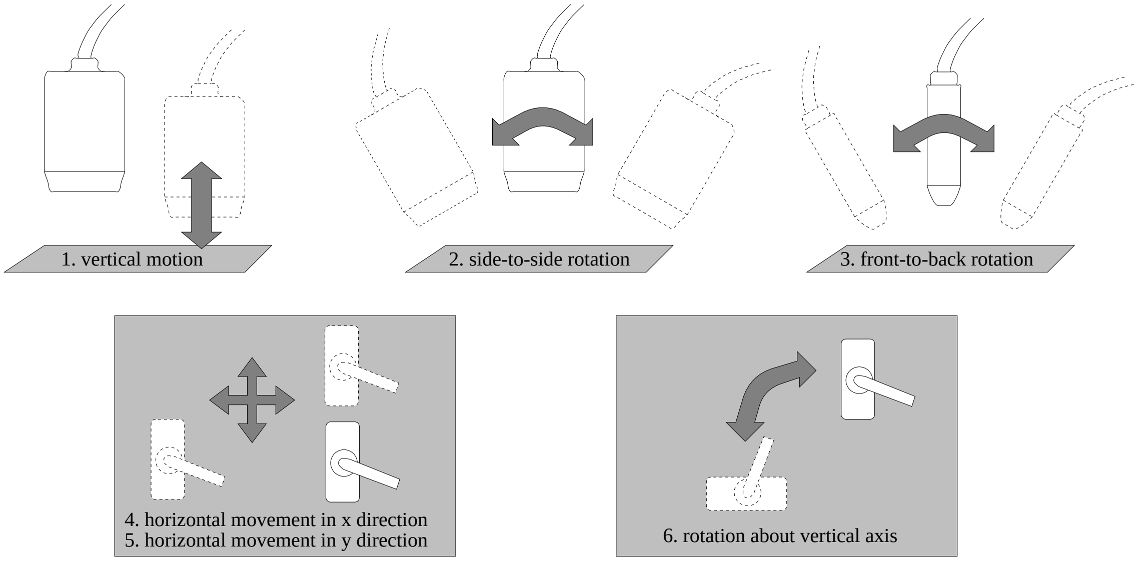


Figure 4: **Minimal sequence of motions for single-wall calibration**. All six degrees of freedom must be exercised to identify the 11 parameters.

construct a special phantom. Furthermore, the plane should show up as a strong, straight line in the image — Figure 8(c) shows a typical example — and this line can be detected *automatically*. While it might be argued that points and crosses can also be detected automatically, there is no doubt that the results would not be as reliable as those for lines. This is because there is much more redundant evidence for the presence of a line in an image: 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 crosses, and practice speaks for itself: everyone calibrating with wire-based phantoms seems to locate the dots by hand. In contrast, we describe a suitably robust, automatic line detection algorithm in Appendix B.

If the coordinate system $\mathbf{C}$ 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:

$$\begin{pmatrix} x \\ y \\ 0 \\ 1 \end{pmatrix} = {}^C\mathbf{T}_T \; {}^T\mathbf{T}_R \; {}^R\mathbf{T}_P \begin{pmatrix} s_x u \\ s_y v \\ 0 \\ 1 \end{pmatrix} \tag{8}$$

The zero component of the equation gives one equation in the unknown parameters ϕ. However, we can write the equation at 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 it's own plane (x, y) do not affect the z-component of ${}^C\underline{\mathbf{x}}$. This leaves 11 identifiable parameters — see Table 1.

The optimisation algorithm needs at least 11 independent equations to solve for these 11 unknowns. While 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 — see Figure 4. As before, the over-determined set of $2m$ equations is solved using the Levenberg-Marquardt algorithm.

It is not strictly necessary to solve for all 11 parameters at once. In fact, it is faster to solve for the angles on their own first, and then solve for the remaining five parameters, holding the angles constant. A final stage, optimising all 11 parameters together, is then used to improve the

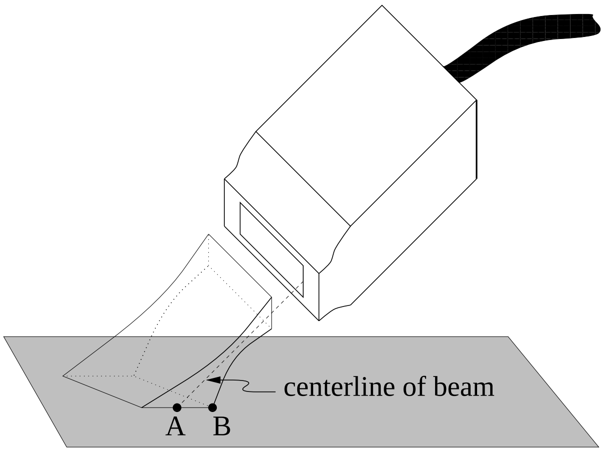


Figure 5: **Beam thickness problem in single-wall calibration**. When the wall is scanned from oblique angles, the finite beam thickness results in a blurred image of the surface. This problem arises because point B is encountered by the ultrasound pulse before point A on the centerline. The echo from point B produces a response in the ultrasound image which does not reflect the true position of the wall.

solution. Details of this three-stage technique are given in Appendix A.

3.4 Cambridge phantom

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 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 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" [16], has been designed to 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 6 and 7.

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 remarkable properties: only the slice required to reflect the centerline of the ultrasound beam exists at each moment in time, and that slice is always oriented towards the ultrasound probe, so the beam is reflected straight back to the probe producing a strong, clear image — see Figure 8(d).

The calibration procedure using the Cambridge phantom may be summarised as follows:

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

2. Immerse the phantom in water, slot the clamp over the bar, and scan the bar with the probe from all possible angles, subject to the constraints imposed by the clamp. The shape of the

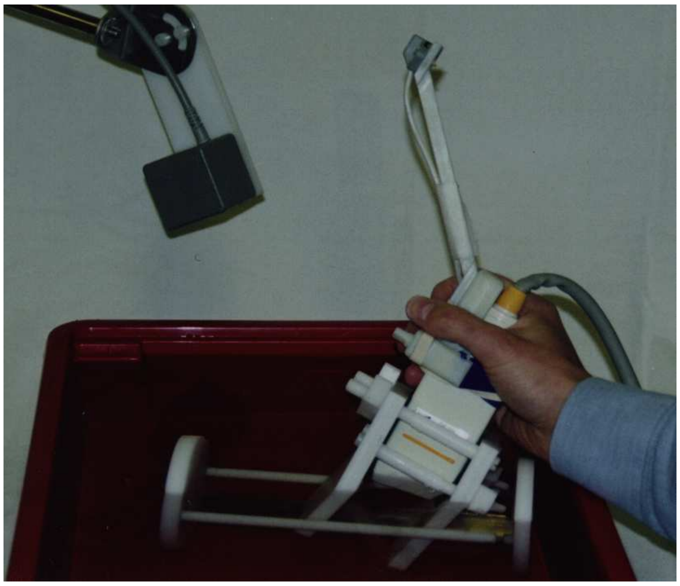


Figure 6: **Calibration using the Cambridge phantom**. A 7 MHz linear phased array probe is fitted into the phantom's clamp, which itself is fitted around the brass bar immersed in the water bath. The position sensor's transmitter is suspended from a cantilevered support above the water bath, while the receiver is attached to the probe at the end of a rigid mount.

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. Because of the clarity of the images and the fact that the phantom produces a line (as opposed to a dot) in the B-scan, it is possible to automatically detect the line in each B-scan — see Appendix B.

4. Once these lines have been located, calibration proceeds in exactly the same manner as described for the single-wall phantom.

Figure 7 shows the minimum sequence of motions required for calibration. The calibration procedure is fully automated and takes less than five minutes to complete, including scanning, line detection, and optimisation. Such speed is not possible with the wire-based techniques, since the dots need to be located by hand in each image.

4 Comparison of Calibration Techniques

4.1 Phantom construction and data acquisition

The cross-wire phantom was constructed by crossing two cotton wires and immersing them in a water bath, as described in [4]. The three-wire phantom comprised 1 mm diameter nylon wires placed under tension between precision drilled holes in the walls of a perspex water bath, as described in [1]. 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 3.4.

The ultrasound machine used in our experiments was a Toshiba model SSA-270A/HG. Calibration experiments were performed on a 7 MHz linear phased array probe. The position sensor was an AC magnetic field receiver-transmitter pair (Polhemus FASTRAK, Polhemus Incorporated, Colchester, Vermont, USA). The receiver was mounted a short distance from the ultrasound probe

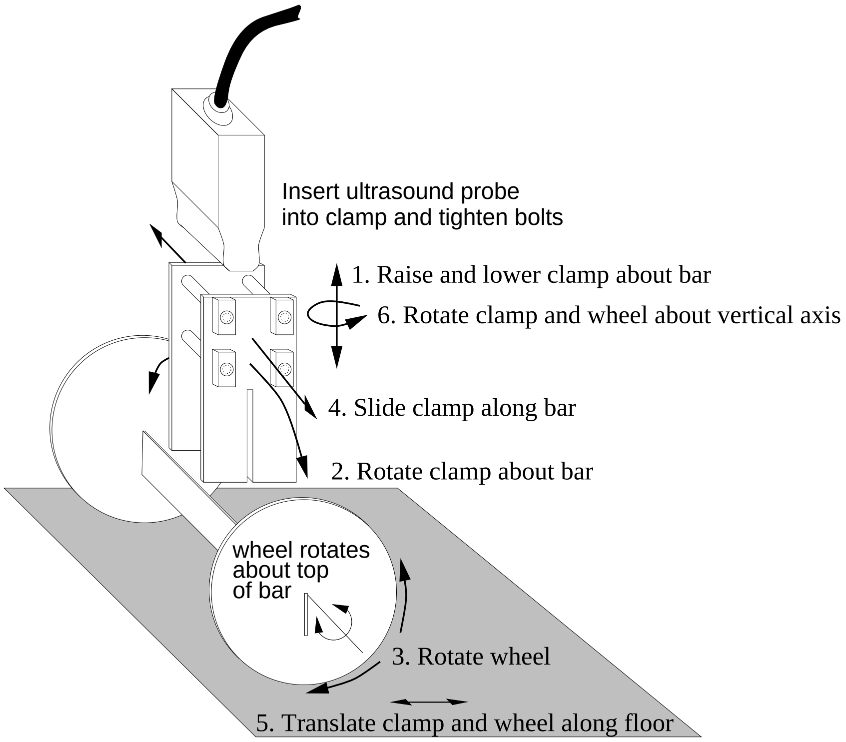


Figure 7: **Cambridge phantom**. Accurate calibration requires a minimal sequence of motions: compare with Figure 4.

at the end of a rigid nylon extension bar. Images from the scanner were digitised using an 8-bit frame grabber, tagged with the position data, and stored in the memory of a Silicon Graphics Indy workstation. While our acquisition software can record at 25 frames/s, we selected a lower frame rate for the calibration experiments, to allow significant motion of the probe between frames.

A depth setting of 40 mm was used throughout. Other imaging controls, such as time-gain compensation, focus, overall gain and acoustic power, were adjusted to give optimal images of each phantom. Figure 8 shows some typical B-scans of the phantoms.

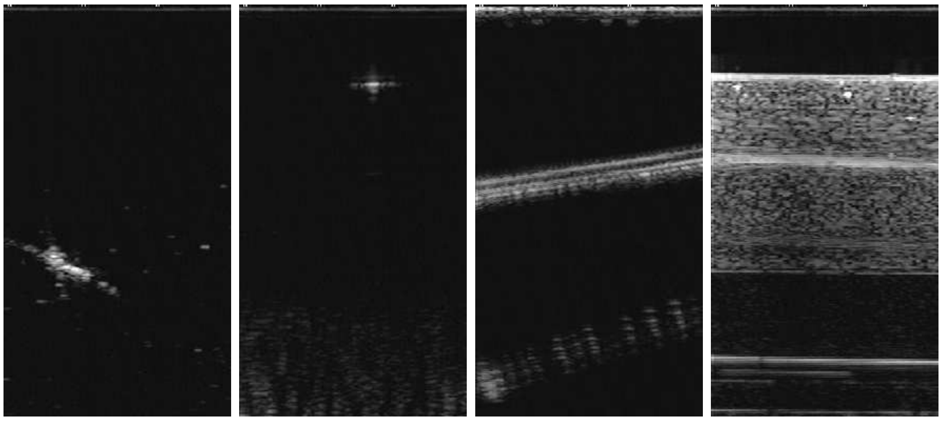
Two calibration experiments were performed using each of the phantoms. A third trial of the Cambridge phantom was carried out, to give additional confidence in the repeatability of the process. For each trial, a large number of B-scans were acquired: any scans which did not reveal a clear image of the phantom were discarded. 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. In contrast, the pixel coordinates for the single-wall and Cambridge phantoms were determined automatically, using the line detection algorithm described in Appendix B.

4.2 Numerical issues

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 free-hand ultrasound calibration — there are guidelines for establishing a well-conditioned problem [7]. 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, a singular value decomposition (SVD) is performed:

$$\mathbf{J} = \mathbf{U}\Sigma\mathbf{V}^T \tag{9}$$



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

Figure 8: **Typical B-scans of calibration phantoms**. In (a), the cross from the two wires is located near the bottom left hand corner of the cropped image. The dot near the top of (b) is one of the wires of the three-wire phantom. In (c), the top line is the floor of the single-wall phantom, the second one down is a reverberation. (d) shows a typical image of the Cambridge phantom: the top of the brass bar is clear and distinct.

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\Sigma = \mathrm{diag}(\mu_1, \ldots, \mu_q)$ is the matrix of ordered singular values. Here $q = 11$ for the cross-wire, single-wall and Cambridge phantoms, while $q = 14$ for the three-wire phantom.

The smallest singular value μ_q will be near zero if there is an unidentifiable parameter. If all μ 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 we have adopted. If $\beta = \pm 90°$, then there exist multiple triples (α, β, γ) that produce the same transformation ${}^{R}\mathbf{T}_P$ [3]. 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, such that β is not close to $90°$.[1]

The second important issue is scaling. Task variable scaling is the process of scaling the errors produced by equation (4) so they contribute equally to the optimisation. Since each of the equations in (4) has the same units and describes a deviation from zero, they are directly comparable. However, 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 (4). This issue can be addressed by parameter scaling, which involves modifying equation (5) as follows:

$$\Delta \mathbf{f} = \mathbf{J}\Delta\phi = \mathbf{J}(\mathbf{H}\mathbf{H}^{-1})\Delta\phi = (\mathbf{J}\mathbf{H})(\mathbf{H}^{-1}\Delta\phi) \tag{10}$$

[1]Note that while degeneracy in the representation of rotation increases κ, it has no bearing on the calibration matrix ${}^{R}\mathbf{T}_P$, and any of the multiple solutions for (α, β, γ) are equally valid. Nevertheless, we avoid this degeneracy to allow meaningful comparison of the condition numbers attained using the various calibration techniques.

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

Table 2: **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 1) the condition number is considerably reduced. Column scaling of the remaining 11 parameters reduces the condition number to below 100.

$\mathbf{H}$ is a scaling matrix chosen such that $\mathbf{JH}$ is better conditioned than $\mathbf{J}$ alone. Equation (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 [10], whereby $\mathbf{H} = \mathrm{diag}(h_1, \ldots, h_q)$ and

$$h_i = \begin{cases} ||j_i||^{-1} & \text{if } ||j_i|| \neq 0 \\ 1 & \text{if } ||j_i|| = 0 \end{cases} \tag{11}$$

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 [19]. Table 2 shows how to establish a well-conditioned calibration problem.

The same process of singular value decomposition, removal of unidentifiable parameters and parameter scaling was performed for each trial, ensuring that all problems were as well conditioned as possible. In order to compare the four calibration techniques without bias, we used the simple Levenberg-Marquardt algorithm (6), and not the faster three-stage technique described in Appendix A, to determine ϕ. With this single stage method, convergence to a solution usually took fewer than 15 iterations.

We ran the Levenberg-Marquardt algorithm from a number of different starting positions for each trial. While local minima were evident when using a small number of scans, increasing the number of scans appeared to solve this problem. Calibration trials with several hundred B-scans *always* converged to a unique solution, independent of the starting position. Working with a large number of B-scans is no disadvantage with the single-wall and Cambridge phantoms, since the lines are detected automatically in the images. Indeed, we show in Appendix D that increasing the number of B-scans helps to improve the calibration accuracy.

There are, however, several different values of ϕ which can achieve the same, well-conditioned global minimum. We call these "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 same calibration, they need to be expressed in a canonical form to allow straightforward comparison of results. A suitable convention is described in Appendix C.[2]

4.3 Results

4.3.1 Cross-wire phantom

A typical set of scans used to image the cross-wire phantom is shown in Figure 9. In the first trial, 251 B-scans were acquired, of which 222 clearly showed the crossed wires. In the second trial, 223

[2]Other researchers parameterise rotation using different schemes, such as raw rotation matrices [4, 12] and unit quaternions. While such schemes avoid singularities and mirror solutions, the redundancy in the representation means that the optimisation has to be constrained. Iterative techniques for solving constrained optimisation problems are characterised by slow convergence and sensitivity to local minima.

	Cross-wire		Three-wire		Single-wall		Cambridge		
	Trial 1	Trial 2	Trial 1	Trial 2	Trial 1	Trial 2	Trial 1	Trial 2	Trial 3
s_x (mm/pixel)	0.148	0.136	0.133	0.132	0.140	0.138	0.137	0.139	0.138
s_y (mm/pixel)	0.138	0.140	0.139	0.144	0.141	0.141	0.138	0.138	0.138
x (mm)	-270.12	-270.25	-269.60	-267.39	-258.86	-261.85	-270.86	-271.22	-270.85
y (mm)	4.73	4.98	5.81	-1.40	4.42	4.39	4.86	5.11	4.79
z (mm)	-64.59	-64.10	-64.51	-65.11	-52.84	-54.33	-65.08	-64.77	-64.59
α (deg)	4.04	-3.15	10.50	12.95	-2.77	-3.21	-0.55	-0.64	-0.84
β (deg)	46.11	43.74	43.02	44.22	44.24	44.17	44.22	44.18	44.22
γ (deg)	-88.64	-92.63	-87.52	-72.88	-88.38	-85.73	-90.37	-90.79	-90.12
B-scans used	222	223	865	694	635	738	697	686	748
B-scans total	251	252	1100	1495	675	775	706	696	757
κ	17.5	16.0	47.5	40.2	447.8	211.1	87.5	68.2	73.9
RMS error (mm)	0.319	0.308	0.515	0.517	0.580	0.591	0.622	0.600	0.585
$\Delta^C \underline{x}_{center}$ (mm)	-	1.74	-	6.07	-	4.08	-	0.24	0.33

Table 3: **Calibration results**. Calibration of a 7 MHz linear probe was performed using the various phantoms. Each solution is given by s_x, s_y and the six parameters of ${}^R\mathbf{T}_P$: x, y, z, α, β, and γ. The RMS error is the root mean square residual of the set of equations (4) at the solution ϕ. κ is the condition number and $\Delta^C \underline{x}_{center}$ is a repeatability measure: see the text for more details. The value of $\Delta^C \underline{x}_{center}$ given for Trial 3 of the Cambridge phantom compares Trial 3 with Trial 1.

of 252 B-scans were acceptable. The results of the two calibration trials are listed in Table 3.

The low condition number of both trials suggests a sufficiently wide range of scan directions. A low root-mean-squared (RMS) error shows that the Polhemus readings were relatively free from noise, and that reliable identification of the pixel near the center of the crossed wires was possible.

We can use the two trials to get a feeling for the repeatability of the technique. We could, for instance, directly compare the solutions between the two trials, but it would be more convenient to derive a single number to measure repeatability. We decided to look at the change in the position of the centre of a B-scan in the reconstruction volume $\mathbf{C}$:

$$\Delta^C \underline{x}_{center} = {}^C\mathbf{T}_T \; {}^T\mathbf{T}_R \; {}^R\mathbf{T}_P^{Trial\ 1} \; {}^P\underline{\mathbf{x}}^{Trial\ 1} - \; {}^C\mathbf{T}_T \; {}^T\mathbf{T}_R \; {}^R\mathbf{T}_P^{Trial\ 2} \; {}^P\underline{\mathbf{x}}^{Trial\ 2} \qquad (12)$$

where

$${}^P\underline{\mathbf{x}}^{Trial\ i} = \begin{pmatrix} s_x^{Trial\ i} \times no.\ columns/2 \\ s_y^{Trial\ i} \times no.\ rows/2 \\ 0 \\ 1 \end{pmatrix}$$

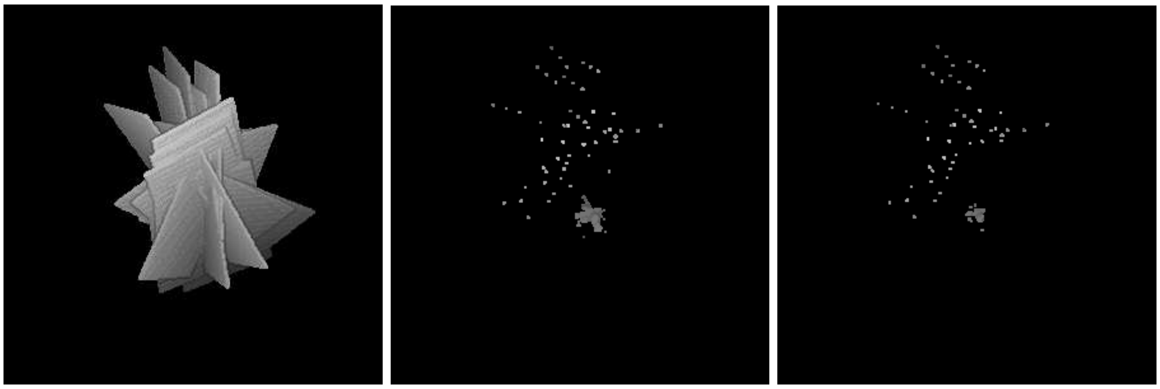
and ${}^C\mathbf{T}_T$ and ${}^T\mathbf{T}_R$ are arbitrary transformation matrices[3]. This measure involves all six parameters of ${}^R\mathbf{T}_P$ and the two scale values, but not the parameters of ${}^C\mathbf{T}_T$ that are included in ϕ but are not required for reconstruction. So $\Delta^C \underline{x}_{center}$ reflects how accurately we can locate a typical B-scan pixel in the reconstruction volume. The cross wire method achieved a small $\Delta^C \underline{x}_{center}$ of 1.74 mm, indicating good repeatability.

4.3.2 Three-wire phantom

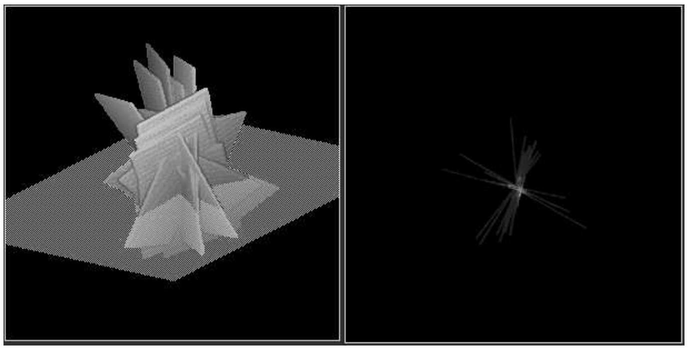
Each of the three wires were scanned in succession, as shown in Figure 10. In Trial 1, 865 out of 1100 B-scans were deemed to be of sufficiently high quality for use in the calibration procedure. In Trial 2, 694 out of 1495 B-scans were accepted. Many scans had to be discarded because the wires do not show up well when viewed from oblique angles. Nevertheless, more than enough good images were available for calibration. The results of the two calibration trials are listed in Table 3.

The low condition numbers suggest that the B-scans were acquired from a sufficient variety of scanning angles. The low RMS error also shows relatively noise-free Polhemus readings and reliable

[3]The particular values of ${}^C\mathbf{T}_T$ and ${}^T\mathbf{T}_R$ have no effect on $\Delta^C \underline{x}_{center}$.



(a) B-scans (b) Cross-wire (c) Smaller cross-wire



(d) Re-slice of voxel array

Figure 9: **Trial 1 of the cross-wire phantom**. (a) shows a surface rendering of the outlines of the B-scans, reconstructed using the calibration solution ϕ. The large variety of scanning angles is evident. (b) shows a similar surface rendering, except that the voxel array has been thresholded to a range around the intensity of the wires. The wires form a small cross near the centre of the image. The remaining cloud of noise is caused by speckle in the B-scans. A tighter range of thresholding is used in (c), revealing a smaller portion of the cross. The right part of (d) shows a slice through the voxel array at the location shown on the left. The small size of the reconstructed cross is evident, suggesting that calibration errors are small.

	Cambridge
s_x (mm/pixel)	0.143
s_y (mm/pixel)	0.136
x (mm)	74.08
y (mm)	-1.04
z (mm)	22.24
α (deg)	125.77
β (deg)	-87.93
γ (deg)	143.05
B-scans used	810
B-scans total	887
κ	50.8
RMS error (mm)	0.531

Table 4: **Calibration results**. Calibration of a 3.75 MHz curvilinear probe was performed using the Cambridge phantom.

pixel identification. However, the repeatability measure $\Delta^C \underline{x}_{center}$ is high at 6.07 mm. One possible explanation is that constructing an accurate three-wire phantom is relatively difficult. In contrast, the cross-wire phantom does not require straight or orthogonal wires.

4.3.3 Single-wall phantom

A large portion of the phantom floor was scanned, as shown in Figure 11. In Trial 1, the line corresponding to the wall was prominent in 635 out of 675 B-scans. In Trial 2, 738 out of 775 B-scans were acceptable. Table 3 lists the results of the two trials.

The main point to note is that the condition numbers are relatively high. This technique is the only one not to achieve a condition number below 100, reflecting the intrinsic problem with the phantom. As discussed in Section 3.4, images of the floor are clearest when scanned from directly above, but become progressively less clear when scanned from an angle. This restricts the range of allowable scanning angles and makes the calibration problem slightly ill-conditioned. The low RMS error suggests relatively noise-free Polhemus readings and reliable line detection. However, the repeatability measure $\Delta^C \underline{x}_{center}$ is high at 4.08 mm. As we show in Appendix D, the poor repeatability is a result of the finite beam width.

4.3.4 Cambridge phantom

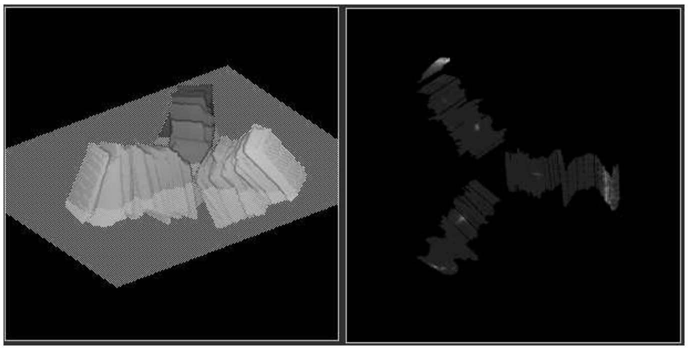
The phantom was moved around the floor and imaged from a wide range of angles, as shown in Figure 12. In Trial 1, the top of the brass bar was clearly detectable in 697 out of 706 B-scans. In Trial 2, 686 out of 696 were acceptable, and 748 out of 757 were used in Trial 3. The low number of discarded B-scans reflects the high quality of the phantom images.

Table 3 lists the calibration results and shows that the condition numbers were lower than with the single-wall method. This is because a wide range motion was possible without compromising the quality of the B-scans. The RMS errors are similar to those obtained with the other techniques, but the repeatability is greatly improved. In fact, $\Delta^C \underline{x}_{center}$ is better than the next most repeatable method, the cross-wire method, by more than a factor of five. The improvement in performance over the single-wall phantom is in line with theory, as discussed in Appendix D.

The Cambridge phantom was also found to be the easiest to use. Once the probe was attached to the clamp, the probe could be moved freely without much worry about obtaining clear images. Furthermore, the entire calibration process was complete a matter of minutes after attaching the position sensor to the probe. To give additional confidence in its performance, the Cambridge phantom was used to calibrate a 3.75 MHz convex, curvilinear array probe. The results, which are given in Table 4, reveal an equally low RMS error and condition number.

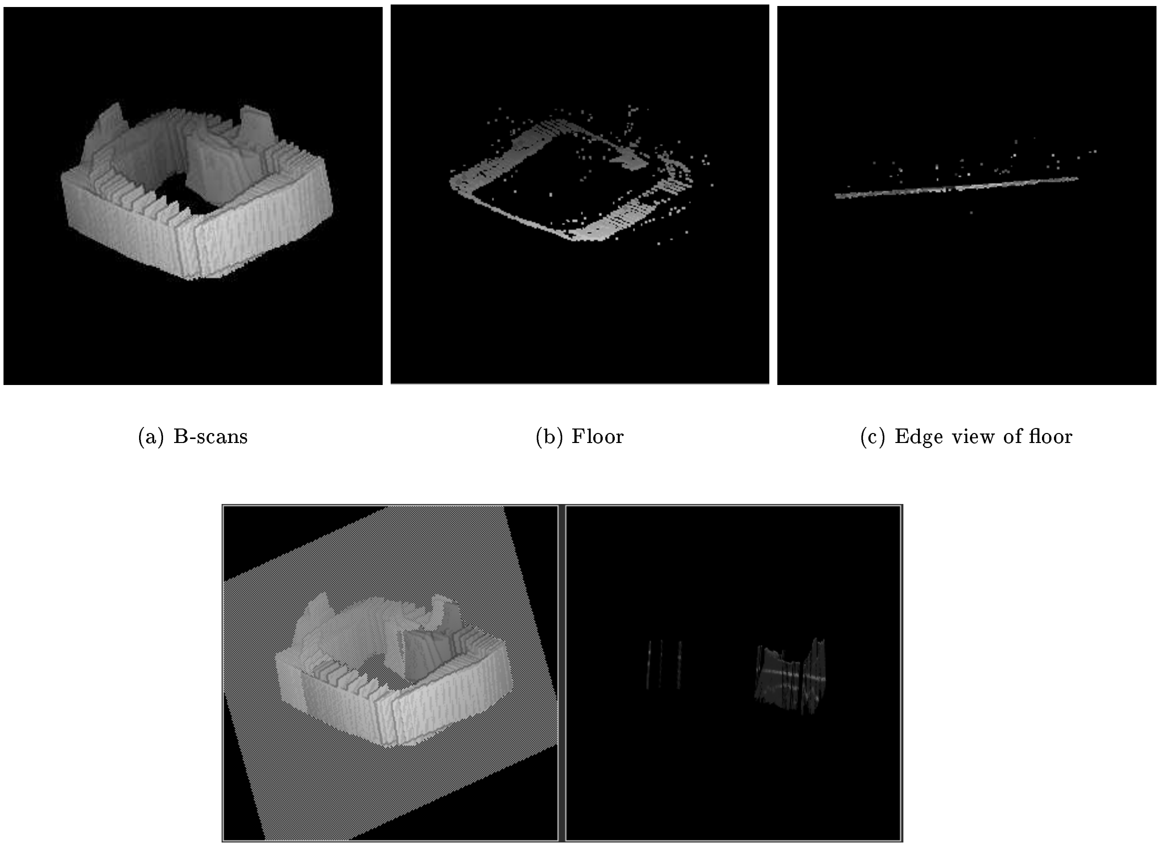
Finally, an *in-vivo* ultrasound examination was performed on a healthy human subject, using a calibration obtained with the Cambridge phantom. Figure 13 shows a 3-D scan of the carotid

(a) B-scans (b) Wires (c) View along x-axis



(d) Re-slice of voxel array

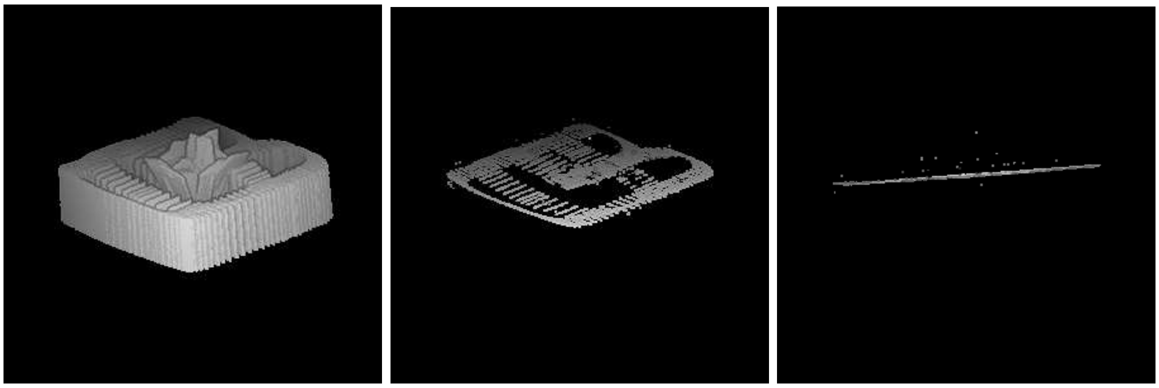
Figure 10: **Trial 1 of the three-wire phantom**. (a) shows a surface rendering of the outlines of the B-scans, reconstructed using the calibration solution ϕ. A long portion of each wire was scanned. (b) shows a similar surface rendering, except that the voxel array has been thresholded to a range around the intensity of the wires. The three wires are clearly visible, as are some flat surfaces of the water bath that happened to fall into the B-scans. The remaining cloud of noise is caused by speckle in the B-scans. (c) is identical to (b), except that the reconstruction is viewed along the x wire. The y and z wires are close to orthogonal. The right part of (d) shows a slice through the voxel array at the location shown on the left. The three wires appear as sharp dots, suggesting that calibration errors are small.



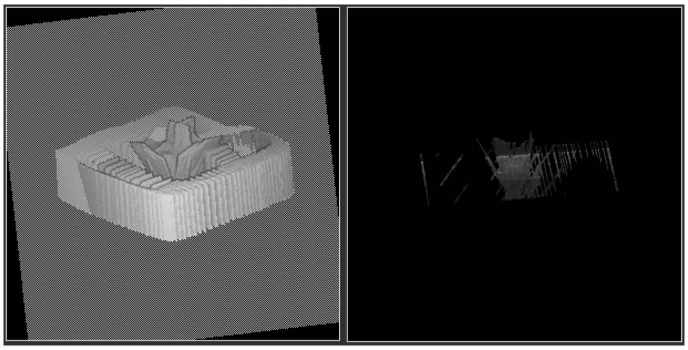
(a) B-scans (b) Floor (c) Edge view of floor

(d) Re-slice of voxel array

Figure 11: **Trial 1 of the single-wall phantom**. (a) shows a surface rendering of the outlines of the B-scans, reconstructed using the calibration solution ϕ. The wide range of motion is evident. (b) shows a similar surface rendering, except that the voxel array has been thresholded to a range around the intensity of the floor. The scanned parts of the floor are clearly visible. The remaining cloud of noise is caused by speckle in the B-scans. (c) is identical to (b), except that the reconstruction is viewed along the edge of the floor, revealing its flatness. The right part of (d) shows a slice through the voxel array at the location shown on the left. Again, the flatness of the floor is evident.

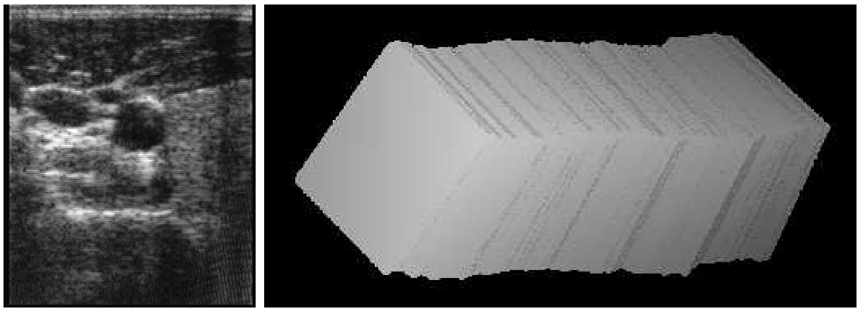


(a) B-scans (b) Plane (c) Edge view of plane

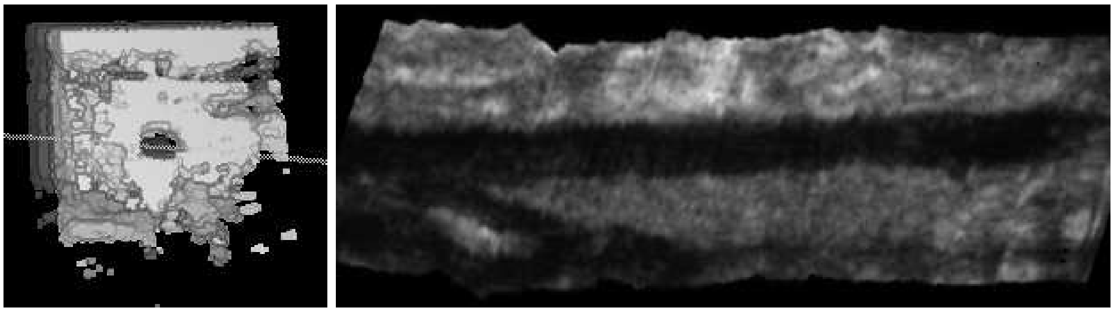


(d) Re-slice of voxel array

Figure 12: **Trial 1 of the Cambridge phantom**. (a) shows a surface rendering of the outlines of the B-scans, reconstructed using the calibration solution ϕ. The wide range of motion and scanning angles is evident. (b) shows a similar surface rendering, except that the voxel array has been thresholded to a range around the intensity of the brass bar. The virtual plane swept out by the top of the bar is clearly visible. The remaining cloud of noise is caused by speckle in the B-scans. (c) is identical to (b), except that the reconstruction is viewed along the edge of the virtual plane, revealing its flatness. The right part of (d) shows a slice through the voxel array at the location shown on the left. Again, the flatness of the plane is evident.



(a) B-scan (b) Outline of B-scans



(c) Axial view of vein (d) Re-slice of voxel array

Figure 13: **Carotid artery**. A B-scan of the carotid artery is shown in (a). The vein is the dark circle just above the center of the image. (b) is a surface rendering of the outlines of the B-scans, reconstructed using a calibration obtained with the Cambridge phantom. The reconstruction volume has dimensions of $277 \times 242 \times 299$ voxels, with a voxel dimension of 0.3 mm. (c) shows a surface rendering, produced with a narrow grey level threshold, looking down the vein. (d) is a grey level image 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 2-D ultrasound.

artery. The vein was scanned down from the base of the jaw for 90 mm of its length. Such a large volume cannot be acquired using mechanical swept volume systems.

The calibration appears to be accurate, with no discontinuities in the reconstruction. The data acquisition was not gated to an ECG, so the pulsatile motion of the vein results in a slight blurring of the data. Other reconstruction errors are caused by refraction of the ultrasound beam and changes in the speed of sound as the pulse travels through different kinds of tissue [24]. These errors cannot be addressed by calibration. Despite these limitations, remarkably clear and consistent 3-D ultrasound images are obtained.

4.4 Discussion

A search of the literature reveals a large number of papers on free-hand 3-D ultrasound, though very few pay much attention to calibration. In [21], for example, the authors describe the use of a 4 mm bead and the cross-wire calibration method, but give no numerical results. Another paper [9] describes a phantom with multiple pin elements to estimate system accuracy. The authors use an acoustic position sensor and claim a standard deviation of 0.66 mm to 1.26 mm for point location accuracy. However, a very restricted set of scanning angles was used. It follows that these results are probably rather optimistic.

The first paper to comprehensively investigate the accuracy of calibration was [4]. The authors used a 7.5 MHz probe, a DC magnetic field sensor (Flock of Birds, Ascension Technology Corporation, Burlington, Vermont, USA), and the cross-wire calibration method. Using a depth setting of 60 mm, four trials of 30 $\pm$ 2 B-scans were performed. The scale parameters were not estimated as part of the calibration, but the authors added a manual correction for the speed of sound in water. The calibration parameters were calculated as the average of the results from the four trials. The RMS error of the solution ranged from 2.1 mm to 3.5 mm, depending on the distance between the receiver and transmitter.

In a follow-up paper [11] results of two more calibration techniques were reported. The first used two Bird receivers, instead of the usual one, to give redundant measurements. The second employed a miniature version of the Bird receiver. The authors used a variable 3-5 MHz phased array sector scan probe, a 1.5 mm brass bead phantom and the cross-wire method. No depth settings were mentioned. Citing refinements in the calibration procedure, the authors claimed a 1.9 mm RMS error with a single Bird receiver. The use of two receivers gives little improvement, and the miniature receiver gives a slightly worse RMS error of 2.5 mm. In this paper, the authors admit that more accurate calibration is needed for imaging smaller structures, such as blood vessels.

Recently, another comprehensive calibration paper [12] was published, featuring an upgraded version of the miniature Bird sensor. The authors used a variable 3-5 MHz probe, a 1.5 mm brass bead and the cross-wire calibration method. Six sets of 32 B-scans were acquired from different directions and combined to give a total of 192 B-scans. Using a variety of depth settings, the RMS error was 2.4 mm. But at the lowest depth setting of 40 mm, a single set of 16 B-scans and limited probe motions gave an RMS error of 0.5 mm. As a measure of accuracy, eight beads, precisely arranged in a grid, were scanned. The error in estimating the distance between beads was 0.06 $\pm$ 0.68 mm.

In a recently published thesis [1], three-wire calibration was performed with a 5 MHz linear phased array probe and a 120 mm depth setting. Any B-scans which contributed a large residual error to equation (4) were removed, leaving at most 30 B-scans per calibration and an RMS error of 0.5 mm. Using all acquired B-scans would result in a significantly larger RMS error.

Care must be taken when comparing results of previous calibration papers. Calibration accuracy depends on the resolution of the B-scans (determined by probe construction and pulse frequency), B-scan depth settings, the position sensor, construction of the phantom and the electromagnetic environment. Moreover, in [4, 11, 12] RMS error is not calculated in exactly the same way as in this paper.

Nevertheless, no previous paper has claimed RMS errors lower than those reported in this paper. We have used a large number of B-scans covering a wide range of scanning angles, and have not attempted to improve the RMS error by selective removal of B-scans. In fact, when scanning the calibration phantoms, we have used a wider range of motions than those expected in a typical *in-vivo* examination. Point position errors in *in-vivo* reconstructions will therefore be even more accurate than the errors in calibration. From our results, we are satisfied that we have optimised the performance of all the phantoms studied, and that the comparisons we have made are therefore meaningful.

5 Conclusions

In this paper we have comprehensively addressed the issue of free-hand 3-D ultrasound calibration. We have reviewed the common techniques and optimised their performance to set an upper bound on their capabilities. Both the cross-wire and three-wire techniques suffer from relatively poor repeatability and, more significantly, extremely long calibration times, typically four to five hours. This has motivated us to examine calibration using a flat plane, since the line it produces in the B-scan can be detected reliably and automatically, greatly speeding up the calibration process. Problems with beam width effects led to the design of the Cambridge phantom, which ingeniously suppresses these effects and guarantees remarkably clear images of the virtual plane. The performance of the Cambridge phantom was significantly better than any of the other methods, while calibration was comfortably performed in a matter of minutes.

		$\phi_{stage\ 1}$	$\phi_{stage\ 2}$	$\phi_{stage\ 3}$
	s_x	s_x/s_y	-	$\surd$
	s_y	-	$\surd$	$\surd$
	x	-	$\surd$	$\surd$
	y	-	$\surd$	$\surd$
${}^R\mathbf{T}_P$	z	-	$\surd$	$\surd$
	α	$\surd$	-	$\surd$
	β	$\surd$	-	$\surd$
	γ	$\surd$	-	$\surd$
	x	-	-	-
	y	-	-	-
${}^C\mathbf{T}_T$	z	-	$\surd$	$\surd$
	α	-	-	-
	β	$\surd$	-	$\surd$
	γ	$\surd$	-	$\surd$

Table 5: **Identified parameters**. The table shows which parameters are identified at each stage of the three-stage optimisation process.

Acknowledgements

Jonathan Carr designed the three-wire phantom and also helped create some of the illustrations. Barbara Levienaise-Obadia performed the initial coding for the line detection algorithm. Robert Rohling is supported by Churchill College and an ORS award. The free-hand acquisition system was developed with the help of Patrick Gosling, and the 3-D renderings were produced using the 3DViewnix visualisation package.

A Three-stage optimisation

A three-stage optimisation process has been developed for the single-wall and Cambridge phantoms, offering faster convergence than the standard, single-stage optimisation. The idea is to solve for the angles first and then, having fixed the angles, solve for the distances. A final, global optimisation is then used to fine-tune the solution.

The phantom appears in each B-scan as a straight line. This line defines two points on the phantom's plane. To solve for the angles alone, we subtract the image positions of these two points to obtain a vector lying in the plane. The direction of this vector must have a zero z-component. As we are now considering only the direction of the vector, we can ignore all the translation elements of the transformation matrices. The resulting optimisation problem has much the same structure as before, except that 3×3 matrices are used throughout, and only the ratio s_x/s_y is required. We therefore solve first for the parameters listed under $\phi_{stage\ 1}$ in Table 5.

In the second stage, the angles are treated as fixed constants. The distances are found by solving the same equations as in the single-stage optimisation, only with fewer unknowns in $\phi_{stage\ 2}$.

A small error in an angle can significantly affect the solution found for the lengths. It is therefore beneficial to run a brief optimisation of all 11 parameters in $\phi_{stage\ 3}$. This converges very quickly and improves the accuracy of the final solution.

B Automatic detection of lines in ultrasound images

One of the principle attractions of the single-wall phantom is the relative ease with which the plane can be reliably and automatically detected in the ultrasound images. This is even more true of the Cambridge phantom, since the virtual plane traced out by the bar shows up as a strong echo, irrespective of the scanning angle. Automatic detection of the plane results in a calibration process which is both rapid and easy to perform.

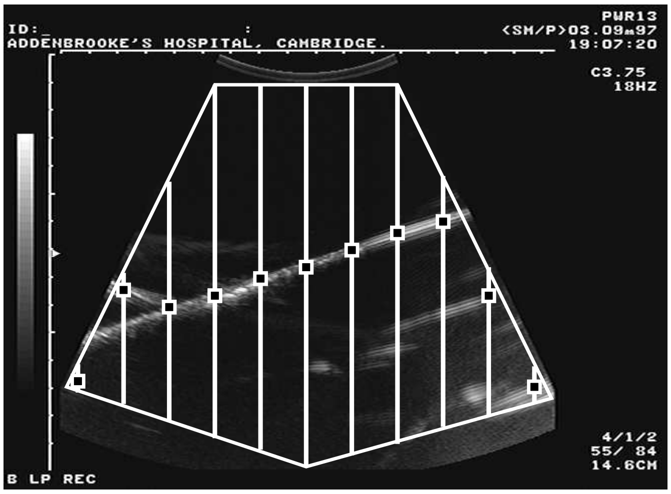


Figure 14: **Edge detection.** Vertical samples are taken within a user-specified region of interest in the ultrasound image. For each vertical sample, a likely edge point is marked at the location nearest the top of the image where the derivative of the smoothed signal exceeds some threshold. Many, but not all, of these points will lie on the echo from the plane.

Since the plane appears as a straight line in the B-scans, measuring its position and orientation comes down to a fairly standard edge detection procedure. However, given the low signal-to-noise ratio of the images (especially with the single-wall phantom), it is necessary to choose a robust edge fitting algorithm which incorporates some outlier rejection.

The approach we adopt is essentially a simplified version of the scheme described in [2]. n vertical lines are sampled at regular intervals from within a region of interest in the ultrasound image. The intensities $I(y)$ along each line are smoothed by convolution with a 1D Gaussian function:

$$s(y) = g_\sigma(y) * I(y)$$

where

$$g_\sigma(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{y^2}{2\sigma^2}\right)$$

The smoothed intensities $s(y)$ are then differentiated. The derivative theorem of convolution allows us to calculate $s'(y)$ from $I(y)$ using just one convolution:

$$s'(y) = g'_\sigma(y) * I(y)$$

Peaks in $s'(y)$ correspond to prominent dark-to-light transitions in the intensities. Since we are looking for the top of the first bright bar in the image, we search for the first maximum of $s'(y)$ that exceeds some threshold T. If no maximum of $s'(y)$ exceeds T, we settle for the biggest local maximum of $s'(y)$. Repeating this process for each of the sampled vertical lines, we obtain a list of points (x_i, y_i), $i \in \{1 \ldots n\}$, corresponding to the first strong dark-to-light edge on each line — see Figure 14.

In the absence of noise, these points would lie on a straight line corresponding to the strong echo from the plane. However, given the quality of ultrasound images, it is likely that some of the points will be located some distance from the plane's actual position. The presence of outliers means that a least squares fit of a straight line to the points (x_i, y_i) is inappropriate. Instead, Random Sample Consensus (RANSAC) [6] is used to fit the line to the data.

RANSAC works by fitting a straight line through every pair of points (x_j, y_j) and (x_k, y_k) in the set (x_i, y_i), $i \in \{1 \ldots n\}$. This generates ${}^nC_2$ lines. Each line is scored according to the size of its *consensus set*: a point (x_i, y_i) contributes to the consensus set if it is no more than a distance

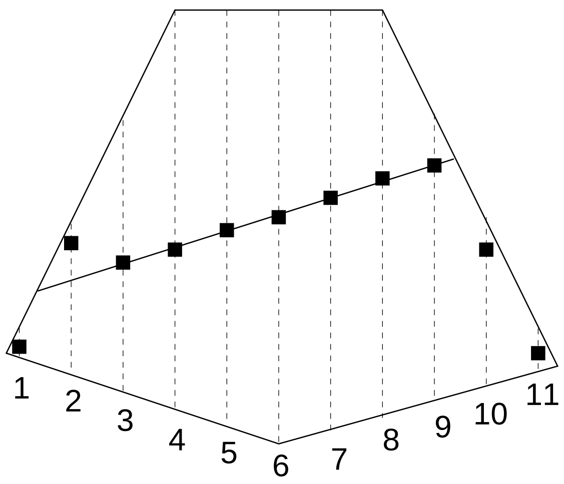


Figure 15: **Line fitting by RANSAC.** The line through points 3 and 9 has a consensus set of size 7, since it is consistent with points 3–9. The maximum size of this line's consensus set is 8, since points 1, 10 and 11 could not possibly lie on the line. The line would be accepted if $p < 7/8$.

d above or below the line. The line with the largest consensus set is retained as the most likely position of the plane in the ultrasound image. Note how one or two outliers in the set (x_i, y_i) will have little effect on the position of the fitted line.

Finally, a check is made that the size of the consensus set is plausible before the line is entered into the calibration data. If the line cannot be clearly detected, then it is better to simply discard that frame than to use unreliable data for calibration. The line is considered accurate if the size of the consensus set is at least a fraction p of the maximum consensus m. Note that m does not necessarily equal n: if the region of interest is not rectangular, then some of the vertical scans may not intersect the fitted line — see Figure 15.

The algorithm is not particularly sensitive to the parameters n, σ, T, d and p. These are set by the user through a simple graphical user interface — see Figure 16. The user can visibly check the performance of the algorithm on a few randomly selected frames (and adjust the parameters if necessary) before running the algorithm on the entire data set. The graphical interface also allows the user to draw the region of interest over that portion of the image that contains ultrasound data. The entire process of setting the region of interest, checking the parameters and extracting the lines typically takes less than a minute for a data set containing several hundred frames.

C Conventions for mirror solutions

There are several distinct sets of angles and scales that produce the same calibration: we call these "mirror" solutions. In order to compare solutions, it is convenient to adopt a canonical form for the angles and scales. We require that:

- s_x and s_y are positive.
- α and γ are in the range $\pm\pi$.
- β is in the range $\pm\pi/2$.

These constraints can be enforced using the following procedure:

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.

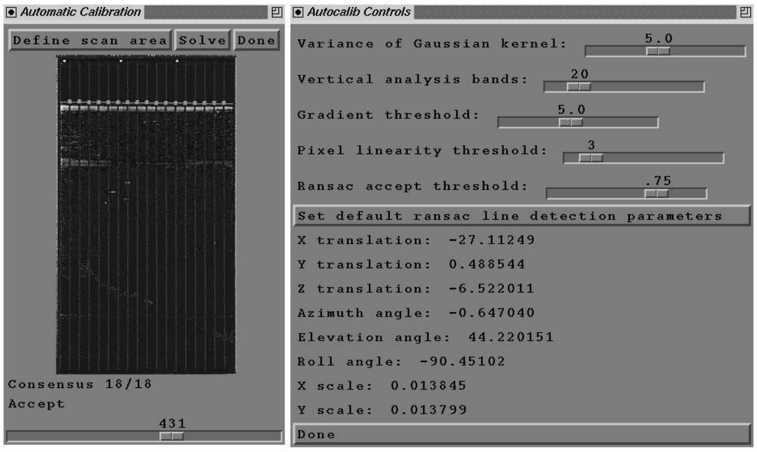


Figure 16: **Graphical user interface.** The left hand window displays the current B-scan and the results of the line detection process. The slider at the bottom of this window allows rapid selection of arbitrary B-scans, with real-time updating of the detected line. The right hand window includes sliders to set the various line detection parameters. The results of the calibration are displayed in the lower part of this window.

3. If $s_y < 0$ then $\gamma \leftarrow \gamma + \pi$ and $s_y \leftarrow -s_y$.
4. If $s_x < 0$ then $\alpha \leftarrow \alpha + \pi$, $\beta \leftarrow -\beta$, $\gamma \leftarrow \pi - \gamma$ and $s_x \leftarrow -s_x$.
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 $\gamma \leftarrow -\gamma$.

D Simulation of beam-width effects

By simulating the effects of the finite beam thickness, we can understand why the Cambridge phantom performs much better than the single-wall phantom. 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 [25]:

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

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

$$\lambda = \frac{\nu}{f} = \frac{1.45 \times 10^3\,\mathrm{mm/sec}}{7 \times 10^6\,\mathrm{sec}^{-1}} = 0.2\,\mathrm{mm}$$

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 in the first trial 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. We used the calibration parameters found in Trial 1 of the Cambridge phantom. Solving equation (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 6.

If now a non-zero beam thickness is introduced, we can calculate the expected change in position of the lines in the B-scans. Solving the calibration equations will result in a different ϕ. For the

	Ideal				Noisy	
Beam Width	0 mm	1 mm	2 mm	3 mm	0 mm	2 mm
s_x (mm/pixel)	0.137	0.137	0.137	0.137	0.137	0.136
s_y (mm/pixel)	0.138	0.138	0.138	0.138	0.138	0.138
x (mm)	-270.86	-269.37	-267.88	-266.39	-271.07	-267.74
y (mm)	4.86	4.87	4.87	4.88	4.84	4.49
z (mm)	-65.08	-63.62	-62.15	-60.69	-65.05	-62.21
α (deg)	-0.55	-0.55	-0.54	-0.54	-0.60	-0.43
β (deg)	44.22	44.22	44.22	44.22	44.26	44.17
γ (deg)	-90.37	-90.42	-90.47	-90.53	-90.63	-89.72
B-scans used	697	697	697	697	697	697
B-scans total	697	697	697	697	697	697
κ	87.5	87.15	86.85	86.57	87.48	86.86
RMS error (mm)	0.000	0.020	0.039	0.059	0.354	0.360
$\Delta^C \underline{x}_{center}$ (mm)	-	2.09	4.18	6.28	0.366	4.27

Table 6: **Simulated 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.

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 6. We observe that the RMS error and condition number are relatively insensitive to beam thickness but, compared with the Cambridge phantom (see Table 3), repeatability is greatly reduced, even for beam widths as low as 1 mm. So the Cambridge phantom's superior repeatability is due to 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.

We can confidently neglect some of these errors. Any latency is likely to be small, since the probe was moved slowly while acquiring the calibration scans. 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 0.8 mm in distance and 0.15° in orientation were added to the position measurements: these are in line with the Polhemus manufacturer's specifications. A reasonable estimate of the image's axial resolution is one wavelength λ. 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 (see Appendix B). Combining this with the resolution error, we added uniform random noise of amplitude 0.9 mm to the end points of the lines in the B-scans.

The simulations were repeated with the noisy measurements. The results are listed in Table 6. The RMS error increased significantly to levels similar to those found in real trials, giving confidence that the noise estimates are reasonable. However, the noise had 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 our decision to calibrate using many more than the minimal number of B-scans.

References

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

[2] J. C. Clarke, S. Carlsson, and A. Zisserman. Detecting and tracking linear features efficiently. In *Proceedings of the British Machine Vision Conference*, pages 415–424, 1996.

[3] J. J. Craig. *Introduction to Robotics: Mechanics and Control.* Addison-Wesley, Reading, Massachusetts, 1989.

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

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

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

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

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

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

[10] C. L. Lawson and R. J. Hanson. *Solving Least Squares Problems.* Prentice-Hall, Englewood Cliffs, New Jersey, 1974.

[11] D. F. Leotta, P. R. Detmer, O. H. Gilja, J. M. Jong, R. W. Martin, J. F. Primozich, K. W. Beach, and D. E. Strandness. Three-dimensional ultrasound imaging using multiple magnetic tracking systems and miniature magnetic sensors. In *Proc. IEEE Ultrasonics Symposium, 1995*, pages 1415–1418, 1995.

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

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

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

[15] R. Ohbuchi, D. Chen, and H. Fuchs. Incremental volume reconstruction and rendering for 3D ultrasound imaging. *Proceedings of SPIE — The International Society for Optical Engineering*, 1808:312–323, 1992.

[16] R. W. Prager. Ultrasound machine calibration. UK patent application no. 9716994.0 filed on 11 August 1997.

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

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

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

[20] S. W. Smith, R. E. Davidsen, C. D. Emery, R. L. Goldberg, and E. D. Light. Update on 2-D array transducers for medical ultrasound. In *Proc. IEEE Ultrasonics Symposium, 1995*, pages 1273–1278, 1995.

[21] A. State, D. T. Chen, C. Tector, A. Brandt, H. Chen, R. Ohbuchi, M. Bajura, and H. Fuchs. Case study: Observing a volume rendered fetus within a pregnant patient. In *Proc. IEEE Visualization, 1994*, pages 364–368, 1994.

[22] H. Steiner, A. Staudach, D. Spitzer, and H. Schaffer. Three-dimensional ultrasound in obstetrics and gynaecology: technique, possibilities and limitations. *Human Reproduction*, 9(9):1773–1778, 1994.

[23] J. W. Trobaugh, D. J. Trobaugh, and W. D. Richard. Three-dimensional imaging with stereotactic ultrasonography. *Computerized Medical Imaging and Graphics*, 18(5):315–323, 1994.

[24] P. N. T. Wells. *Advances in Ultrasound Techniques and Instrumentation*. Churchill Livingstone Inc., New York, 1993.

[25] T. A. Whittingham. Resolution and information limitations from transducer arrays. *Physics in Medicine and Biology*, 36(11):1503–1514, 1991.