
**Rapid, Easy and Reliable
Calibration for Freehand 3D Ultrasound**

P-W. Hsu, R. W. Prager
A. H. Gee and G. M. Treece

CUED/F-INFENG/TR 534

August 2005

University of Cambridge
Department of Engineering
Trumpington Street
Cambridge CB2 1PZ
United Kingdom

Email: pwh24@cam.ac.uk, rwp/ahg/gmt11@eng.cam.ac.uk

Rapid, Easy and Reliable Calibration for Freehand 3D Ultrasound

Po-Wei Hsu, Richard W. Prager, Andrew H. Gee and Graham M. Treece

University of Cambridge
Department of Engineering
Trumpington Street
Cambridge CB2 1PZ

Abstract

This paper presents improvements to the plane-based technique for calibrating freehand 3D ultrasound systems. The improvements are designed to make it easier for inexperienced users to perform plane-based calibration and to know that they have got a reliable result. In particular, we enable the calibration to be performed using water at room temperature while producing a result that is valid for average soft tissue, and we show how it is possible to provide feedback on the reliability of the calibration using a metric based on the curvature of the calibration criterion function. We present comprehensive results showing that these innovations improve the precision of the calibration and offer useful feedback to the user.

1 Introduction

Freehand three-dimensional (3D) ultrasound [5] is a technique for acquiring ultrasonic data of a 3D volume by measuring the trajectory of the ultrasound probe using a position sensor. It has applications in clinical volume measurement, analysis of complex geometry, surgery planning and radiotherapy planning [7]. As the probe is swept across the volume, the sensor records the position and orientation of the probe with respect to an external world coordinate system. The recorded position, however, is the position of the sensor, rather than the scan plane, as shown in Figure 1. We therefore need to determine the scales in the ultrasound images and the rigid-body transformation from the electrical centre of the position sensor to the corner of the ultrasound scan plane. This is done through a process called probe calibration [12]. Points in the B-scan are first transformed to the sensor's coordinate system by the parameters found by calibration, and then to the world space by the sensor's readings. It is crucial that probe calibration be as accurate as possible, as any errors will propagate through to subsequent stages of image analysis and measurement.

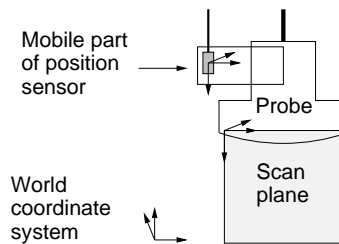


Figure 1: The coordinates associated with the scan plane and mobile part of the position sensor.

The usual technique for 3D ultrasound calibration is to perform a series of scans of an object with known dimensions (a phantom) [13]. The simplest such object is a point fiducial formed by cross wires [4, 10] or a small spherical ball [1, 20] in a water bath. This point is scanned from different directions and segmented in each B-scan. For a given set of calibration parameters, each point in the B-scan can be mapped to the world coordinate system. Since the same point is being scanned, they should be all mapped to the same point in space. This sets up an over-determined system of constraints that can be solved using optimisation techniques.

There are several difficulties in following the point phantom approach. First of all, each point in the B-scan needs to be manually or semi-automatically segmented. This makes the calibration process long and tiresome. Secondly, due to the finite thickness of the ultrasound beam, it is very difficult to align the probe directly at the centre of the point, as the point fiducial may appear in the B-scan even when the probe is not directly pointing at it.

Muratore and Galloway Jr. [15] used a 3D localizer, which is essentially an additional position sensor attached to the point phantom. In general, this requires an additional calibration for the localizer itself, unless it is precisely designed and manufactured. In any case, this means that the position of the phantom is known in space during calibration. This greatly simplifies the optimisation, and allows for closed-form solutions.

A three-dimensional localizer is in general able to locate any point in space, and has been used to locate points on other types of phantoms. Sato et al. [19] used a two-dimensional phantom comprising of three pointed edges. These three points are located in space with a 3D localizer. The planar phantom is then aligned with the B-scan and the three points manually segmented. Since these three points are known in space, the calibration parameters can be solved algebraically in a closed form. This approach has the advantage that only one frame is needed for calibration, but it is nevertheless extremely difficult to align the plane phantom, with the thick beam compounding the difficulty.

The alignment problem is solved by the design of Z-phantoms [3, 12, 16], where wires are connected in a ‘Z’ shape. The intersection of the wires with the B-scan forms a virtual plane phantom, whose location can be deduced from the imaged points. This type of calibration is fast, since only one image is required and does not require accurate alignment with the plane. However, the thick beam is still a problem. The wires do not show up as clear dots, but as smeared ellipses. Locating the centres of such ellipses accurately is difficult and requires human intervention.

Gee et al. [8] designed a high definition mechanical device, where the plane phantom is adjusted with micrometers and is therefore able to be aligned very accurately. The alignment is also aided by the use of wedges, which show up as bright patches in the B-scans, and so a slight misalignment of the phantom can be easily seen from the B-scans. The wedges also serve as the fiducial points on the planar phantom, which are semi-automatically segmented. This mechanical device has been shown to be the highest definition system in the current literature.

In all the aforementioned approaches, the calibration requires a purpose built phantom, some of which need to be very accurately manufactured. Another calibration technique is the single wall approach [17, 18, 23]. All the hardware that is required is a flat plane emersed in a water tank. The flat plane could be as simple as the bottom of the water tank. The Cambridge phantom [17] may be used as a virtual plane to avoid the beam thickness problem and produce strong echoes even when the plane is scanned from an oblique angle. As the plane is scanned from different angles, the images of the plane appear as clear, straight lines in the B-scans. This allows the possibility of automatic segmentation, subject to a few threshold settings prior to calibration. Prager et al. [17] used the RANSAC line detection algorithm [6] for its robustness against outliers. Rousseau et al. [18] used Hough transforms [9] to extract the lines and reject outliers from each B-scan. They went further to assume a smooth varying of the line in neighbouring B-scans in order to reject any falsely detected lines. For a given set of calibration parameters, points on these lines can be transformed to the position sensor’s coordinate system and then to the world coordinate system using the position sensor’s readings. If the calibration is correct, these points will also lie on a plane in world space. This sets up a list of constraints involving eleven parameters. Six of these define the rigid-body transformation from the mobile part of the position sensor (attached to the probe) to the corner of the B-scan image. Three further parameters define the location of the phantom plane in 3D space. These three are only required internally for the process of calibration itself; once calibration is complete they are discarded. Two final parameters define the horizontal and vertical scales in the B-scan image. An iterative optimisation algorithm is then utilized to find the best fit plane by adjusting the calibration parameters, see Figure 2.

This paper describes some refinements to make the process of single-wall calibration easier and more reliable for inexperienced users to perform. There are two main problems: (1) the speed of sound in water at room temperature is slower than its speed in average soft tissue. The sound

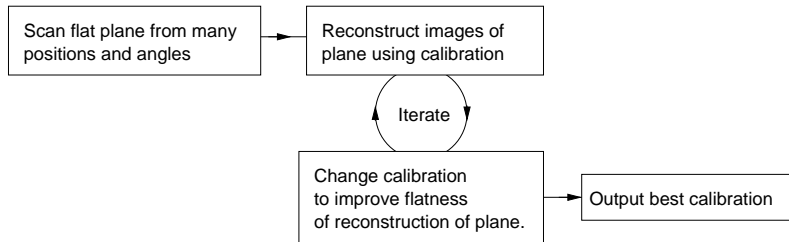


Figure 2: Block diagram of plane-based calibration

speed in water at 20 Celsius is only 1482 m/s [2], while sound travels at 1540 m/s in average soft tissue. For accurate calibration the water bath should be at 48 Celsius which gives a sound speed of 1540 m/s. This is difficult to maintain, and uncomfortable for the operator. If the calibration is performed in cold water without any adjustments, the temperature difference would cause an error of 3.7%. (2) There is currently no feedback to the user to indicate when the plane has been scanned from a sufficiently diverse set of positions and angles to produce a well constrained solution to the resulting non-linear equations. The residue error from the optimisation is not an accurate measure of calibration quality, since it is possible to perform an incorrect calibration with a low residue error. An extreme example would be to repeatedly scan the plane from the same position. In this case, the residue error would be low, but the user has no idea that the calibration is unreliable. As a result, some inexperienced users have recorded under-constrained motion sequences leading to incorrect calibrations. In this paper, we show how it is possible to provide feedback to the user on the reliability of the calibration.

2 Improving the Reliability of Single Wall Calibrations

In this section, we first present the modifications we have made to current single wall calibration systems used by Prager et al. [17] and Treece et al. [22]. We then performed and analyzed a number of calibrations in order to measure the reliability of a calibration using a metric based on the curvature of the criterion function. At the same time, these calibrations were used to compare with the results from Treece et al. [22], in order to verify the effectiveness of the modifications we have made to the single wall calibration protocol.

2.1 Improvement to Single Wall Calibrations

We describe the two modifications we have made to improve current plane-based calibration systems. The first modification is to account for the difference between the speed of sound in water and soft tissue. The second is to determine the B-scan scales explicitly.

2.1.1 Speed of Sound

The speed of sound problem is fixed by allowing the calibration to be performed at room temperature, but requiring the user to provide a measurement of the water temperature. The ultrasound images can then be corrected using one of the well-known polynomial models for the speed of ultrasound in water as a function of temperature [2].

Since sound travels slower in cold water than in average soft tissue, the image of the plane (dotted lines) appears further away than its actual position (solid lines), as shown in Figure 3. Corrections can therefore be made by moving the image of the plane toward the face of the probe.

We used the same edge detection algorithm used by Prager et al. [17]. Vertical samples are taken within the user defined region in the ultrasound image. A likely edge point is where the derivative of the smoothed signal first exceeds some threshold. The edge could then be found by applying the RANSAC line detection algorithm to these likely edge points [6]. For a linear probe,

the correction is a translation of each likely edge point toward the probe face; the axial coordinates v of the points are multiplied by the temperature correction factor $t = \frac{\text{speed in cold water}}{\text{speed in average soft tissue}}$.

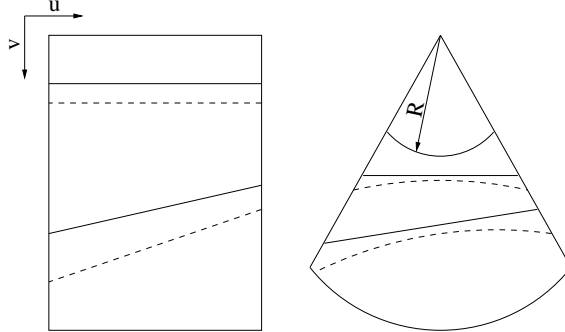


Figure 3: The dotted lines shows the distorted the plane due to slower sound speed in cold water. The solid lines show the corrected plane. The diagram is exaggerated to emphasize the distortion.

In the scenario of a curvilinear probe, the plane appears as a curve rather than a line. We therefore first detect the probe shape automatically [21]. The probe centre is then the intersection of the left and right edges. We further assume that the probe face is part of a circular arc with radius R centred at the probe centre. We then move each point on the curve (shown as a dotted line in Figure 3) towards the probe face, in the direction of the probe centre, by the temperature correction factor. The shifted points will then lie on a straight line, and can be detected using the RANSAC line detection algorithm [6]. Since this correction is mathematically sophisticated and extra steps need to be followed by the user to detect the probe shape reliably, an approximated correction is possible for common abdominal curvilinear probes. Please refer to Appendix A for a detailed discussion.

2.1.2 Explicit Scale Estimation

We increase the robustness of the calibration process by reducing the number of parameters involved. We can establish the image scales at the start of the procedure using a separate protocol, and hence remove them from the optimisation. The result is that the non-linear optimisation only has to determine nine rather than eleven parameters, and this makes the system of equations easier to constrain.

The image scales are determined using the distance measurement tool that is available on most (if not all) clinical ultrasound machines. This tool provides the distance, in centimetres, between two points marked in the B-scan image. A simple graphical user interface has been implemented to enable the user to indicate the point positions and enter their separation. Since most ultrasound machines provide square pixels to a very good approximation, we have further simplified this procedure, with little loss of accuracy, by assuming that the horizontal and vertical image scales are the same. The scale is therefore estimated by $\frac{\text{distance (cm)}}{\sqrt{(\Delta u)^2 + (\Delta v)^2}}$, where Δu and Δv the horizontal and vertical distances, in pixels, between the two marked points on the B-scan.

2.2 Quantification of Calibration Quality

In order to measure the calibration quality and provide feedback to the user, we performed calibrations using three probes on two different ultrasound machines. The first is the Diasus¹ 5–10MHz linear-array probe. The analog radio-frequency (RF) ultrasound data, after receive focusing and time-gain compensation but before log-compression and envelope detection, was digitized using a Gage CompuScope 14100 PCI 14-bit analog to digital converter,² and transferred at 10 frames

¹Dynamic Imaging Ltd., <http://www.dynamicimaging.co.uk/>

²Gage Applied Technologies Inc., <http://www.gage-applied.com/>

per second to a Pentium(R) 4 2.80GHz PC running Microsoft Windows XP. Since we have access to RF data, the image scales could be set at our discretion and the B-scans were displayed at 0.01cm/pixel. We also performed calibrations with the Toshiba³ model SSA-270A/HG 7.5MHZ linear probe, and the 3.75MHz curvilinear probe. The B-scan images were digitized using an inexpensive frame-grabber card and transferred to an 800MHz PC running Linux. Both probes were tracked using an AdapTrax⁴ infrared LED target for the Polaris⁵ optical tracking system.

We performed calibrations including 30, 60, 90 and 120 images of the plane. Every calibration was repeated 10 times. For each repetition, the calibration consisting of 60 images was formed by capturing 30 images of the plane in addition to the 30 images from the previous calibration, and similarly for the calibrations with 90 and 120 frames. It is important for plane-based calibration that the plane is scanned in a non-degenerate sequence to constrain the plane in space [17, 22]. These calibrations all included the required motion sequences, namely: vertical movements, side to side rotations, front to back rotations, translation in the x direction, translation in the y direction and rotations about the z -axis, where x and y are the two axes in the plane phantom, aligned with the sides of the water tank, and z the corresponding axis perpendicularly out of the plane, as shown in Figure 4. The probes and depth settings used in the experiments are listed in Table 1. One focus at approximately half the depth setting was used in each calibration. The scales in the Diasus calibrations were set manually to 0.01cm/pixel, since this is the scale at which the RF data, after log-compression, envelope detection and band-pass filtration, were displayed. The distance measurement tool was utilized to estimate the scales in the Toshiba calibrations. Both the temperature and scale estimation is highly repeatable so that the same readings are obtained in every measurement. Hence there is no variation in the scales in each calibration. In all, a total of: 4 probes \times 4 numbers of plane images \times 10 repetitions = 160 calibrations were acquired.

Machine	Probe Type	Frequency	Depth	B-scan	Scales
Diasus	Linear	5–10MHz	3cm	360 \times 300	Preset
Diasus	Linear	5–10MHz	6cm	360 \times 600	Preset
Toshiba	Linear	7.5MHz	6cm	222 \times 416	Distance Measurement
Toshiba	Curvilinear	3.75MHz	6cm	388 \times 380	Distance Measurement

Table 1: Probes and other settings used in the calibrations.

The quality of each calibration was assessed by measuring its precision and accuracy [22]. Precision is measured by the consistency of the locations of the four corners and the middle of the image when located in 3D space using the calibration parameters. For each of these points we measured its mean location in 3D space, and then computed the average magnitude of its deviation from this point. This measure was called *mean 3D error* by Treece et al. [22]. These five errors were then averaged. Figure 5 shows the averaged error as the number of scans is increased in our calibrations. The decreasing graphs show that more consistent calibrations were achieved at a higher number of scans.

The precision measurements indicate the variation in the image due to spatial calibration errors. This variation could be reduced further by averaging over a larger number of calibrations. Moreover, the repeatability of the calibrations is not an indication of accuracy, as the calibrations could be highly repeatable but biased.

We assess the accuracy of our calibrations by scanning a tissue mimicking phantom consisting of coplanar spheres with 2mm diameter [11]. These spheres lie on a grid with known dimensions (Appendix B), and they show up as dark circles in the B-scans. We scanned the phantom with the Diasus probe at 3cm and 6cm depth, using the three patterns in Figure 6. In order to maximize the visibility of the spheres, six focal points were used in the middle region of the image. The images were recorded at approximately 7–15 frames per second. Although many foci were used during image acquisition, some of the spheres away from the focal points could not be

³Toshiba Corporation, <http://www.toshiba.com/>

⁴Traxtal Technologies, <http://www.traxtal.com/>

⁵Northern Digital Inc., <http://www.ndigital.com/>

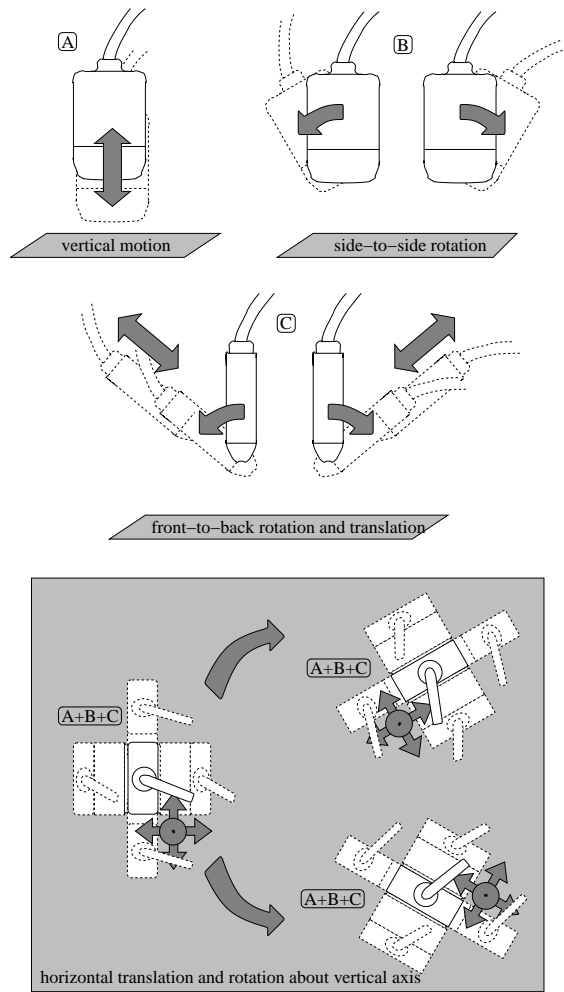


Figure 4: The sequence of probe movement required during calibration to constrain the 11 calibration parameters (before scales estimation). The diagram is taken from Treece et al. [22].

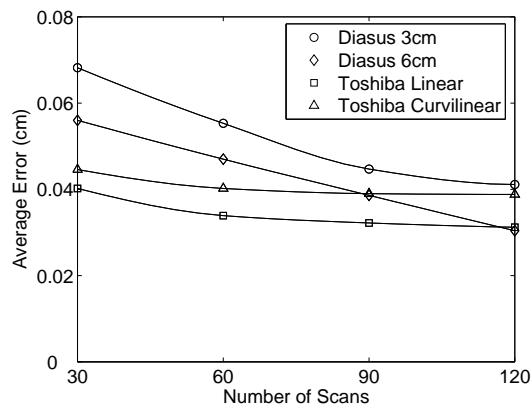


Figure 5: The variation of the four corners and the centre of the cropped B-scan due to spatial calibration for the different types of probes.

seen clearly and segmented properly, and so they were not included in the accuracy assessment. The phantom was scanned freehand, but as slow as possible so that each sphere appears in a number of consecutive images. A typical recording consisted of approximately 400 scans down the 10cm length of the scanning window. The images of the spheres were then semi-automatically segmented, reconstructed in space using the parameters from our calibrations, and fitted to our predefined grid of spheres [22]. We measure the error as the 3D distances between the centres of the reconstructed spheres and their true location in world space.

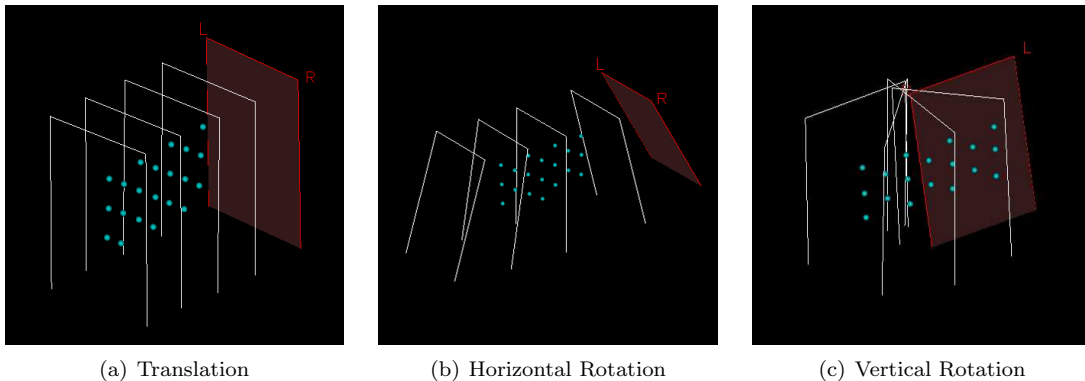


Figure 6: Scanning patterns used when scanning the phantom. Approximately 400 scans were recorded in each case.

Figure 7 shows the orthographic projection of the sphere centres' distribution. The small grey circles are the actual positions and size of the spheres forming the predefined grid. In order to make the errors more visible, the small grey circles are magnified by a factor of four and shown as larger black circles. The crosses show the centre of each reconstructed sphere with the error magnified by a factor of four. The diagrams in the left column show the reconstruction errors for the Diasus probe at 3cm depth; the errors at 6cm depth are shown in the right column.

Figure 8 shows the average error as a function of the number of scans. The graphs are relatively flat. This is most likely because the dominant errors are not caused by spatial calibration. Possible sources include the optical tracker, temporal misalignments, image capture and segmentation of the spheres. Nevertheless, from the low errors and the even distribution of located spheres in Figure 7, we can deduce that our calibrations are unbiased.

In this paper, we will use precision as the measure for calibration quality. This is because precision is more sensitive to calibration errors than accuracy, since precision measures the variations caused by spatial calibration only, and not a combination of calibration and other sources of errors. We have also verified that this choice is viable by the low accuracy errors in each calibration, confirming that the calibrations are unbiased and accurate.

2.3 Curvature of the Criterion Function at the Solution

In each calibration we obtain a set of constraints involving nine variables, since the B-scan scales have been determined explicitly. These variables are solved using a non-linear optimisation algorithm, and we have ensured that the correct solution is obtained (Appendix C). We wish to measure how well the solution is constrained. As the nine parameters are similarly scaled, with distances in centimetres and angles in radians, we can get an indication of an under-constrained solution using the minimum curvature of the criterion function at the solution. If the minimum curvature is high then the solution is well-constrained. If the minimum curvature is low then one or more parameters will be under-constrained. We can thus spot poor solutions and advise the user to scan the plane from some more angles to introduce further constraints.

We wish to use the minimum curvature at the solution as a measure of the quality of the solution. We therefore measured the minimum curvature of our calibrations by computing the Hessian,

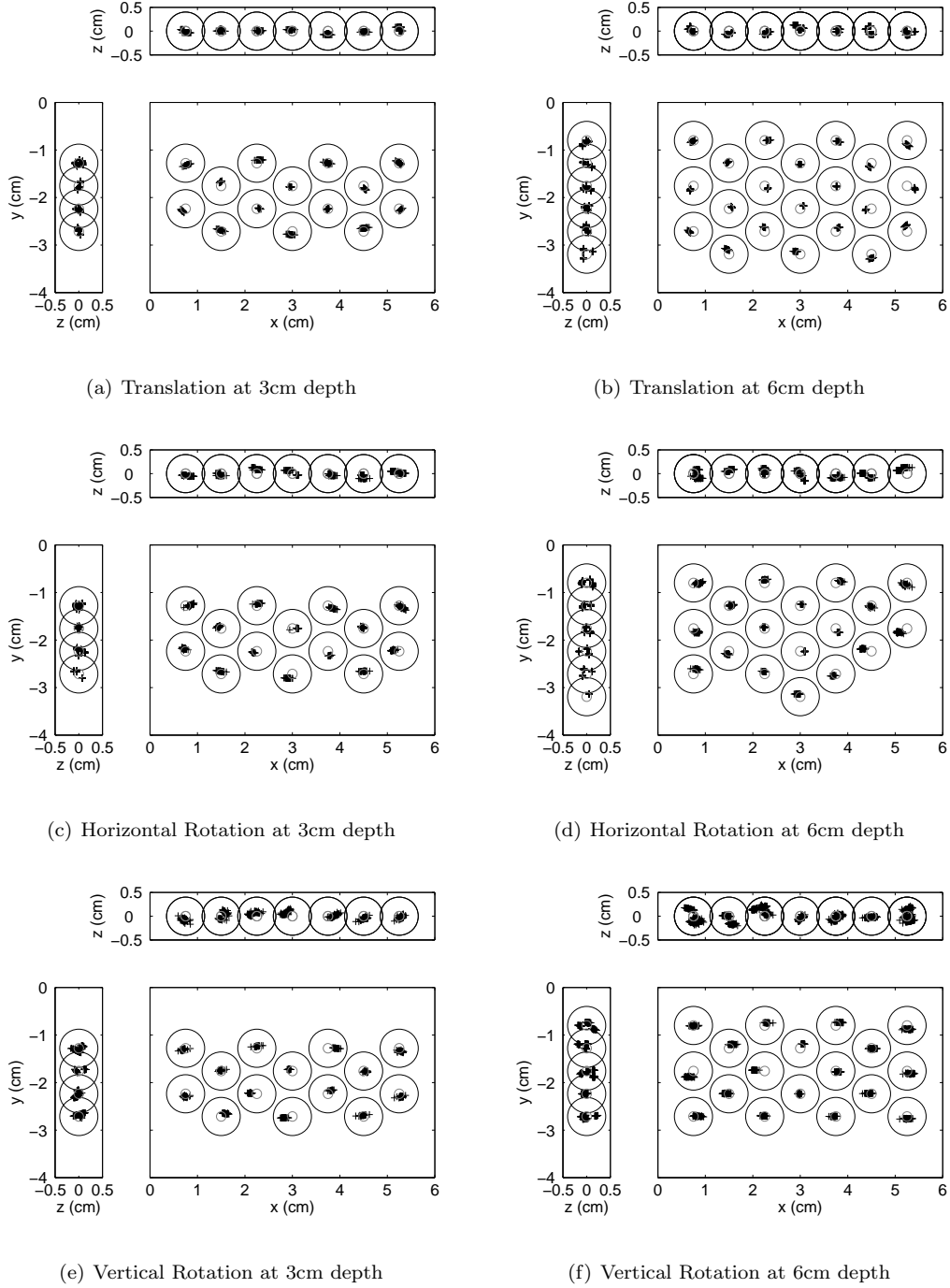


Figure 7: Orthographic projection of the phantom and the reconstructed spheres. The small grey circles shows the true size and location of the spheres in space. The larger black circles shows these spheres magnified by a factor of four. The distribution of the reconstructed spheres are shown as crosses, with the error magnified by a factor of four. The consistent magnification means that the crosses and the large black circles are directly comparable.

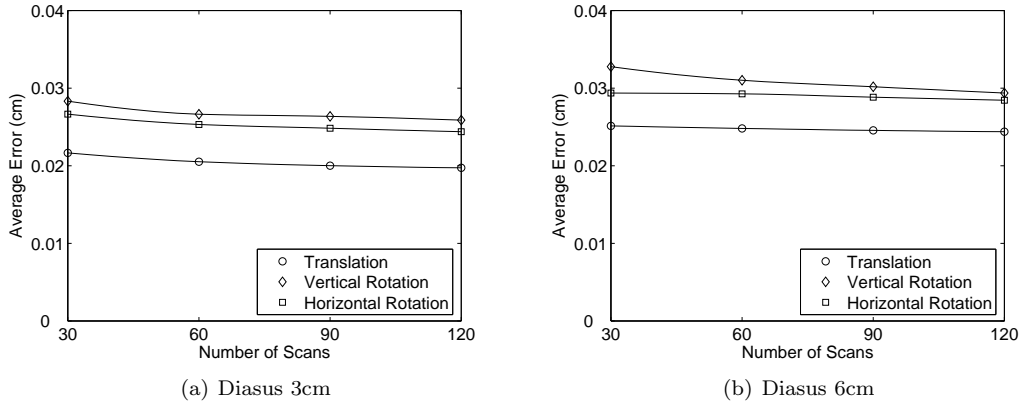


Figure 8: The variation of the system error in terms of distance measurement as the number of scans is increased in the spatial calibrations.

the matrix consisting of all the second derivatives of the objective function. The eigenvalues and eigenvectors of the Hessian are evaluated to give the magnitudes and directions of greatest and least curvature. In all four probes, the eigenvector corresponding to the minimum eigenvalue for each calibration is found to be in the translation direction that defines the phantom in world space. This eigenvalue is expected to be lower than other eigenvalues, because this translation is typically an order of magnitude larger than the other calibration parameters. This means that the error in this direction is also expected to be an order of magnitude higher than other parameters. This does not bother us as the three parameters defining the phantom in world space are subsequently discarded, and do not judge how well a calibration is performed. We therefore use the second minimum eigenvalue, whose corresponding eigenvector is in the direction concerning a combination of the six parameters that define the transformation from the mobile part of the position sensor to the corner of the B-scan image. For brevity, we will use the word *eigenvalue* or *minimum eigenvalue* in this paper while we are in fact referring to the *second minimum eigenvalue* (*minimum eigenvalue of interest*).

Figure 9 shows the graph of the second minimum eigenvalue (minimum eigenvalue of interest) of the Hessian at the calibration solution as the number of scans is increased. A clear linear relationship can be seen. This shows that the minimum curvature at the solution is related to the number of scans of the plane used in the calibration.

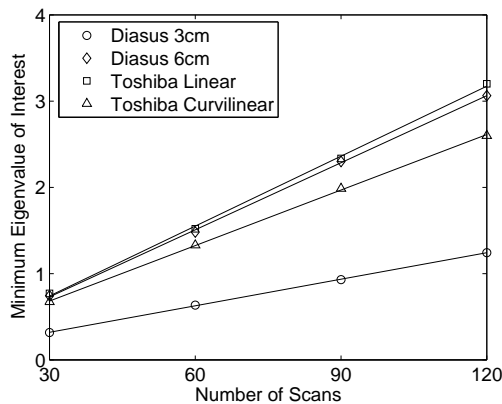


Figure 9: The graphs shows the second minimum eigenvalue (minimum eigenvalue of interest) of the Hessian calculated at the solution of the objective function as the number of scans is increased.

We have shown that the error is related to the number of scans in the calibration in Figure 5. Hence the number of scans is related to both error and eigenvalue. This enables us to plot the error against the eigenvalue, as shown in Figure 10. This shows that we can use the size of the eigenvalue as an indicator for the quality of the solution, and therefore advise the user whether the calibration is well-constrained.

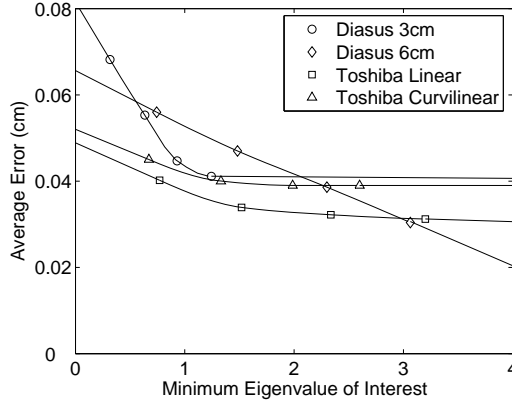


Figure 10: The variation of the four corners and the centre of the B-scan as the calibration becomes more constrained.

2.4 Incomplete Scanning Patterns

We have shown in the previous section that we can use the eigenvalue to determine whether the solution is well-constrained, hence the reliability of the calibration. This assumes that the user has performed all the required scanning patterns in Figure 4. Inexperienced users often neglect some of these motions inadvertently. We performed calibrations consisting of 30 images with one of the motions missing with the Diasus probe at 6cm depth to investigate whether incomplete calibrations do result in an under-constrained solution, reflected by a low eigenvalue.

The minimum eigenvalues for these incomplete calibrations are shown in Figure 11. In the two cases where side to side and front to back rotations are missing, the under-constrained optimisation resulting from incomplete scanning patterns is reflected by a near zero minimum eigenvalue. For the case where vertical movements are missing, the minimum eigenvalue also dropped by more than a factor of two, indicating an insufficiently constrained solution. In the other three cases the eigenvalue did not drop as expected. This suggests that although these motions are absent, the solution is well-constrained. We verify this by performing a series of calibrations without these motions and show that they are just as precise as calibrations with all motions exercised.

We performed an additional 10 calibrations with minimal rotations and translations in the aforementioned directions, but still containing the remaining three types of motions. Figure 12 shows the graph of average error against the number of scans. The graph for good calibrations where all motions are exercised using the same probe (Figure 5) is duplicated here for comparison.

From the figure it could be seen that the apparent paradoxical high eigenvalue for calibrations where rotations about the z -axis and translations in the xy -plane are missing is in fact correct. The average error is similar in magnitude from a low to a high number of scans. This means that these motions are not necessary to sufficiently constrain the optimisation for accurate calibration. This is because we have reduced the number of parameters to be optimized from eleven to nine, hence reducing the degrees of freedom which the calibration needs to constrain. This observation can be confirmed by simulating the calibration process mathematically. We have found that successful calibrations can be simulated excluding the aforementioned motions.

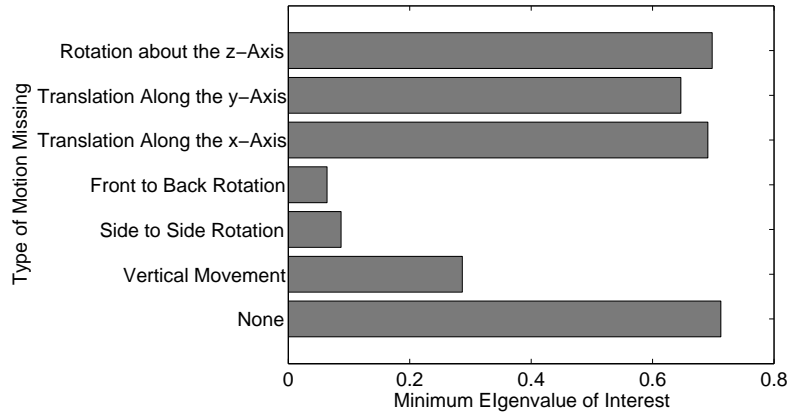


Figure 11: The minimum eigenvalues for the calibrations where a particular motion is missing.

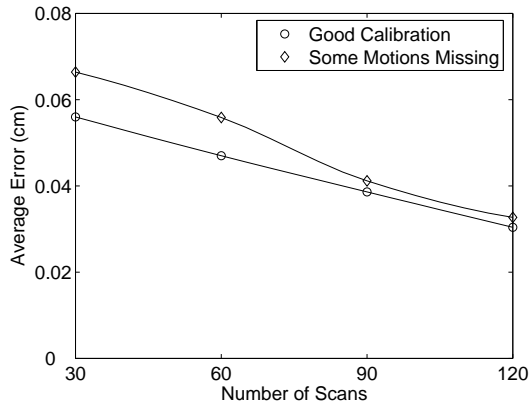


Figure 12: Comparison between calibrations with minimal rotation about the z -axis and minimal translation in the xy -plane, and calibrations where all six types of motion are exercised.

2.5 Precision and Accuracy

We verify the effectiveness of the improvements we have made to single wall calibration by comparing our calibrations based on 120 images with the results from Treece et al. [22], where the highest definition system using plane-based calibrations was reported. Figure 13 shows that the calibrations in this paper are more precise than those achieved by Treece et al. [22]. The error for the Diasus 3cm probe has been reduced by more than a half, and the error for the 6cm depth setting has been reduced by more than a third.

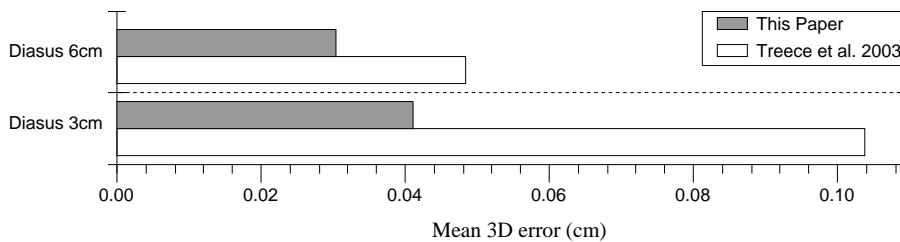


Figure 13: Comparison of the errors due to spatial calibration alone achieved in this paper and in Treece et al. [22]. The errors shown are the mean 3D errors at the five corners.

The probe motion during calibration is restricted when the depth is set low at 3cm. This can result in a weakly constrained optimisation. By removing the scale factors from this optimisation we reduce the dimensionality of the search and greatly improve the precision of the result.

The higher errors in the calibrations performed by Treece et al. [22] are most likely caused by the variation in temperature during calibration. The hot 48 Celsius water will cool down rapidly in the standard room environment. By correcting the temperature differences, calibrations in cold water is possible.

Figure 14 shows the comparison of our system accuracy with the results from Treece et al. [22]. It shows that our adjustments have improved the system accuracy slightly.

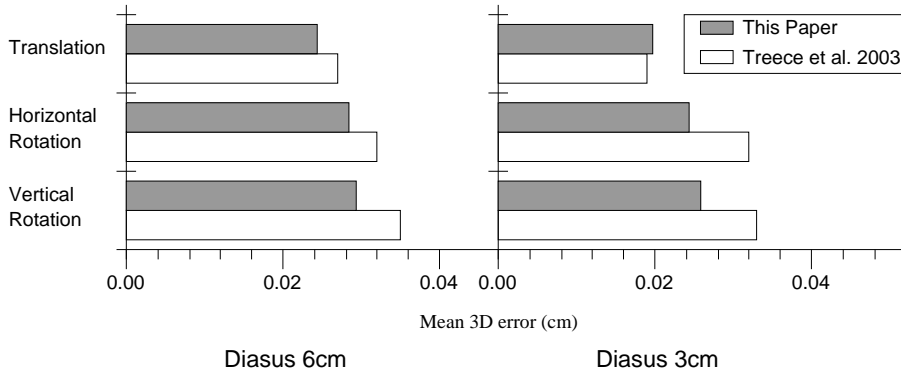


Figure 14: Comparison of the errors from distance measurements of the whole freehand 3D ultrasound system achieved in this paper and in Treece et al. [22]. The errors shown are the mean 3D distance errors between the centres of the segmented spheres and their real locations in space.

3 Conclusion

We have proposed some changes to the plane-based 3D ultrasound calibration protocol that increase its reliability and ease of use, while maintaining its state-of-the-art performance. The compensation for the temperature of the water and the simple technique for explicit entry of scale values are both easy to use. In particular, fewer degrees of freedom need to be constrained by spatial calibration, and so rotations about the z -axis and translations in the xy -plane are no longer needed to achieve a good calibration.

We have also produced a foolproof indicator of the calibration quality by calculating the eigenvalue of the Hessian at the solution of the criterion function. We have ensured that the global minimum is reached by the non-linear optimisation (Appendix C). Figure 10 shows the variation of the B-scan corners due to spatial calibration as a function of the minimum eigenvalue. The high accuracy of the calibrations shown in Figure 8 confirms that these calibrations are unbiased. Hence we may use the eigenvalue as an indicator for the calibration quality. Figure 11 further confirms that if a motion necessary for calibration is missing, which is a common scenario for inexperienced users, the incomplete calibration is indicated by a low eigenvalue. The eigenvalue criterion has been evaluated and shown to be effective on three different probes and at two different depth settings.

A The Curvilinear Probe

It is clear that the temperature correction routine is more sophisticated for the curvilinear probe than for the linear probe. This includes the detection of the probe shape and curvature in addition to the mathematical adjustments, which is an extra step for the user during calibration. We

therefore investigate what is the expected error should we use the temperature correction routine for a linear probe to approximate the exact correction.

Figure 15 (a) shows an exaggerated diagram when the plane is scanned in such a way that the plane appears as a horizontal line, or in fact a slight curve. The exact correction is shown with arrows pointing in the direction of the shift, the result is a horizontal straight line, shown as a dashed line in the diagram. This is the true position of the plane in space. The approximated correction is shown as arrows ending with a circle. It is clear from the diagram that the maximum error occurs at the middle of the image, where the point is over corrected. If this central point appeared at a depth d in the B-scan, and t represents the temperature correction factor, then the point is corrected to position td . For an exact correction, the point should be corrected to $t(d - c) + c$. The overcorrection is thus the difference $(1 - t)c$. The line detection algorithm is likely to estimate the line to be around the middle of the two extreme points as the best fit line (shown as a dashed line), and so the expected error e should be about half the overcorrection error, i.e. $e \approx \frac{1}{2}(1 - t)c$. Since $t = 0.963$ for water at 20 Celsius, and c is less than 1cm for common abdominal curvilinear probes, this error ($e \approx 0.01-0.02\text{cm}$) is negligible for such probes.

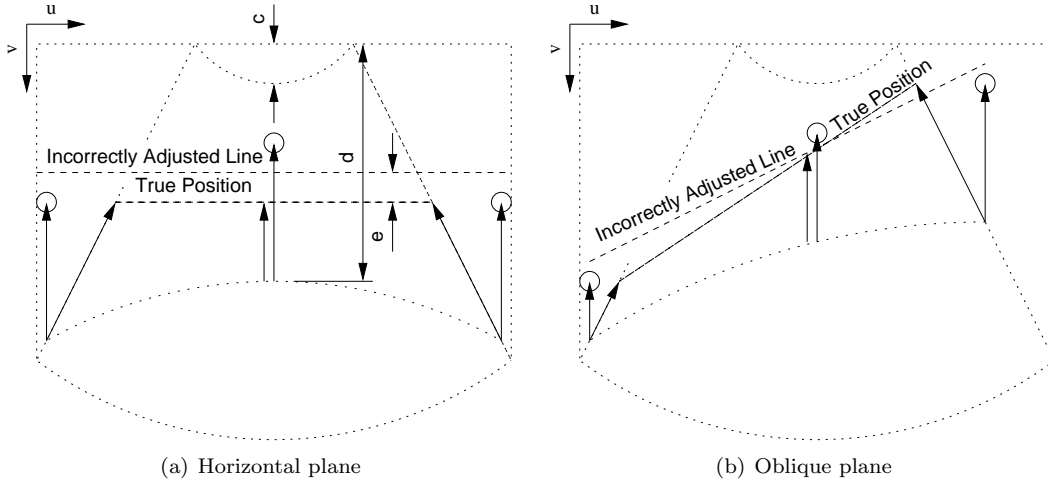


Figure 15: The difference between the corrected planes when applying the correct and approximated correction routines to the displayed plane. The approximated correction is shown with arrows ending with a circle. It is assumed that a best fit line will be fitted to the non-linear corrected points. Both diagrams are exaggerated to show the correction error more clearly.

Figure 15 (b) shows an exaggerated diagram when the plane is scanned from an oblique direction. The difference between the two correction routines in this case is mainly a rotational error, rather than a distance error. The incorrectly corrected line will appear not as steep as the real position of the plane, while the distance at the centre will remain roughly at the same place as before. This error is difficult to quantize, as it is dependent on the depth, curvature and width of the probe, slant of the plane, and even the thresholds in the line segmentation algorithm. As an estimate of the expected error, we compute the error using the dimensions of our Toshiba curvilinear probe, as shown in Figure 16. The plane is scanned at maximum gradient at 17cm depth. In this case, the error is 0.6° , and in practise the error will be less than this value. Since this type of motion only occurs in approximately 20% of the calibration, its impact on the calibration error is limited.

We verify that using the simple temperature correction is a good approximation to the exact correction by performing an additional 10 calibrations with the Toshiba curvilinear 3.75MHz probe using the simple temperature correction routine, and compare this result with the previous calibrations with the same probe, but with the exact temperature correction routine applied. Since the probe shape no longer needs to be detected, it is therefore not necessary to have the sides of the probe visible within the images. This allows the B-scans to be displayed at a wide field of

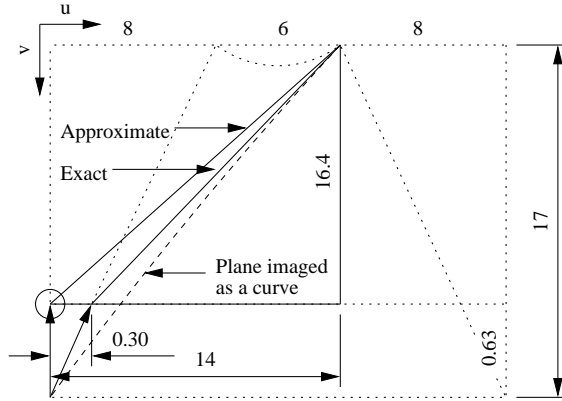


Figure 16: An estimate of the error caused by temperature correction differences. The dimensions are shown in centimetres for our Toshiba curvilinear probe.

view, the more common setting for the curvilinear probe in clinical environments. The difference in settings for the two sets of calibrations is shown in Table 2.

Correction	Cropped B-scan	Refresh Rate
Exact	388×380	36Hz
Approximate	496×416	18Hz

Table 2: Setting differences for the Toshiba curvilinear 3.75MHz probe used for determining the effect of the different temperature correction routines. The depth is set at 6cm.

The repeatability of the curvilinear probe shape detection is shown in Table 3, with the values in pixels. It is clear that the centre and radius of the probe are accurate to within 2 pixels (0.3mm).

	Centre		Radius	
	x	y	2D	
Mean	193.1	-489.5	(193.1; -489.5)	
RMS	0.115	1.120	1.126	1.105
95%	± 0.237	± 2.314	± 2.344	± 2.283

Table 3: Repeatability of the estimation for curvilinear probe centre position and radius. The unit of the values in the table is pixels. The scales in each B-scan are determined explicitly to be 0.0143cm/pixel.

Figure 17 shows the mean 3D errors of the five corners for the curvilinear probe when the two temperature correction routine is applied. The similarity of the precisions showed in the two graphs shows that there is not a need for accurate temperature correction. Although the precision is slightly better when an approximation to the temperature correction routing is used, it is not statistically significant given the spread of errors in the five corners. The difference is also small when compared with the error achieved by Treece et al. [22] and in this paper as shown in Figure 5. An exact temperature correction routine for common abdominal curvilinear probes is therefore not necessary to achieve high precision.

B Verification of the sphere distances

Although the location of each sphere was precisely manufactured, the phantom has shrunk over the years by losing water. This has caused the spheres to be closer than the manufactured dimensions, typically near the scanning window at the top away from the rigid walls. We first measured

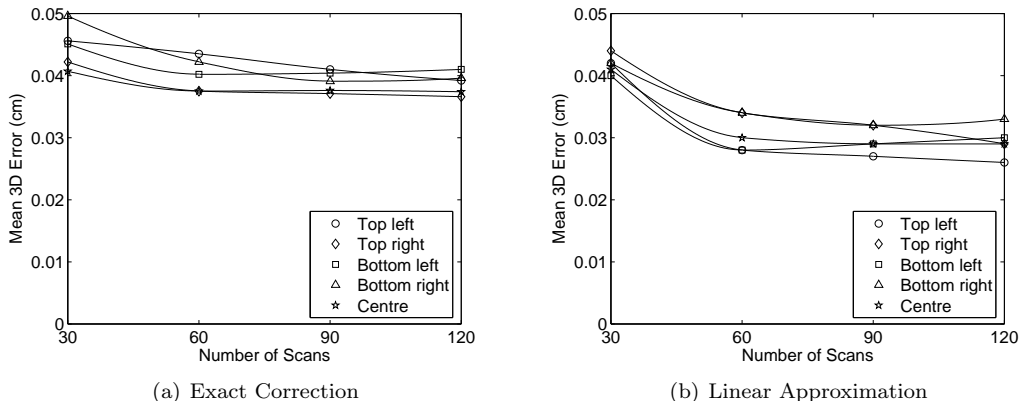


Figure 17: The effect of temperature correction routines on the consistencies of the five B-scan corners.

the distances between the spheres. This was done by capturing multiple ultrasound images in the direction of the coplanar spheres with the Diasus 5–10MHz probe with multiple foci down its depth. The horizontal and vertical distances between the spheres were measured using the distance measuring tool provided by the ultrasound machine in a single image. The measurements are shown in Table 4. The vertical distances between spheres have shrunk by approximately 5%, while the horizontal distances remained roughly unchanged. We modelled the locations of the spheres to lie on a rectangular grid, with the measured vertical distance and the manufactured horizontal distance. We have kept the manufactured horizontal distance as we do not anticipate any shrinking or expanding in the horizontal direction as the phantom is surrounded by rigid walls on its sides. The very slight difference in the measured distances is within measurement error.

Probe	Measured Distances		Original Distances	
	Horizontal	Vertical	Horizontal	Vertical
Diasus 3cm	3.021 ± 0.0087	0.947 ± 0.012	3 ± 0.01	1 ± 0.01
Diasus 6cm	3.005 ± 0.0181	1.914 ± 0.013	3 ± 0.01	2 ± 0.01

Table 4: The measured and original manufactured distances between two or three spheres, depending on the number of spheres that could be seen in each B-scan. All distances are in centimeters, displayed in the form mean \pm standard deviation.

C Solution of the Criterion Function

We have utilized the non-linear Levenberg-Marquardt optimisation algorithm [14] to solve for the solution to the system of non-linear equations involving nine parameters. The calibration parameters are solved in three stages: the angles are first solved, then the distances, and then a final global optimisation to fine tune the solution [17]. As is the case for all optimisation algorithms, a starting point is necessary. This is chosen at random, but within a reasonable estimate of what the parameters can be. This means that the solution given by the optimisation algorithm is dependent on the starting point. Hence, we need to ensure that the optimisation converges to the global minimum, rather than any other local minimum. Based on trial runs of the optimisation, the convergence to a non global minimum occurs less than 50% of the time. We therefore repeat the optimisation 50 times at different starting points and accept the solution with the minimum error. This process takes less than a second on our 800MHz PC. The probability that the optimisation still converges to the incorrect minimum is 1 in 10^{15} .

References

- [1] D. V. Amin, T. Kanade, B. Jaramaz, A. M. DiGioia III, C. Nikou, R. S. LaBarca, and J. E. Moody Jr. Calibration method for determining the physical location of the ultrasound image plane. In *Proceedings of the fourth International Conference on Medical Image Computing and Computer-Assisted Intervention, Lecture Notes in Computer Science*, volume 2208, pages 940–947. Springer-Verlag, 2001.
- [2] N. Bilaniuk and G. S. K. Wong. Speed of sound in pure water as a function of temperature. *The Journal of the Acoustical Society of America*, 93(3):1609–1612, 1993.
- [3] R. M. Comeau, A. Fenster, and T. M. Peters. Integrated MR and ultrasound imaging for improved image guidance in neurosurgery. In *Proceedings of SPIE — The International Society for Optical Engineering, Medical Imaging 1998: Image Processing*, volume 3338, pages 747–754, 1998.
- [4] P. R. Detmer, G. Bashein, T. Hodges, K. W. Beach, E. P. Filer, D. H. Burns, and D. E. Stradness Jr. 3D ultrasonic image feature localization based on magnetic scanhead tracking: in vitro calibration and validation. *Ultrasound in Medicine & Biology*, 20(9):923–936, 1994.
- [5] A. Fenster, D. B. Downey, and H. N. Cardinal. Three-dimensional ultrasound imaging. *Physics in Medicine and Biology*, 46:R67–R99, 2001.
- [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, 1981.
- [7] A. H. Gee, R. W. Prager, G. H. Treece, and L. H. Berman. Engineering a freehand 3D ultrasound system. *Pattern Recognition Letters*, 24:757–777, 2003.
- [8] A. H. Gee, N. E. Houghton, G. M. Treece, and R. W. Prager. A mechanical instrument for 3D ultrasound probe calibration. *Ultrasound in Medicine & Biology*, 31(4):505–518, 2005.
- [9] P. V. C. Hough. Machine analysis bubble chamber pictures. In *International Conference on High Energy Accelerators and Instrumentation*, pages 554–556. CERN, 1959.
- [10] Q. H. Huang, Y. P. Zheng, M. H. Lu, and Z. R. Chi. Development of a portable 3D ultrasound imaging system for musculoskeletal tissues. *Ultrasonics*, 43:153–163, 2005.
- [11] J. M. Kofler Jr. and E. L. Madsen. Improved method for determining resolution zones in ultrasound phantoms with spherical simulated lesions. *Ultrasound in Medicine & Biology*, 27(12):1667–1676, 2001.
- [12] F. Lindseth, G. A. Tangen, T. Langø, and J. Bang. Probe calibration for freehand 3-D ultrasound. *Ultrasound in Medicine & Biology*, 29(11):1607–1623, 2003.
- [13] L. Mercier, T. Langø, F. Lindseth, and D. L. Collins. A review of calibration techniques for freehand 3-D ultrasound systems. *Ultrasound in Medicine & Biology*, 31(4):449–471, 2005.
- [14] J. J. More. The Levenberg-Marquardt algorithm: Implementation and theory. In *Numerical Analysis, Lecture Notes in Mathematics*, volume 630, pages 105–116. Springer-Verlag, 1977.
- [15] D. M. Muratore and R. L. Galloway Jr. Beam calibration without a phantom for creating a 3-D freehand ultrasound system. *Ultrasound in Medicine & Biology*, 27(11):1557–1566, 2001.
- [16] N. Pagoulatos, D. R. Haynor, and Y. Kim. A fast calibration method for 3-D tracking of ultrasound images using a spatial localizer. *Ultrasound in Medicine & Biology*, 27(9):1219–1229, 2001.

- [17] R. W. Prager, R. N. Rohling, A. H. Gee, and L. Berman. Rapid calibration for 3-D freehand ultrasound. *Ultrasound in Medicine & Biology*, 24(6):855–869, 1998.
- [18] F. Rousseau, P. Hellier, and C. Barillot. Confhustus: A robust and fully automatic calibration method for 3D freehand ultrasound. *Medical Image Analysis*, 9(1):25–38, 2005.
- [19] Y. Sato, M. Nakamoto, Y. Tamaki, T. Sasama, I. Sakita, Y. Nakajima, M. Monden, and S. Tamura. Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization. *IEEE Transactions on Medical Imaging*, 17(5):681–693, 1998.
- [20] 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 *Proceedings of the Conference on Visualization '94, IEEE Visualization*, pages 364–368, California, 1994. IEEE Computer Society Press.
- [21] G. M. Treece, R. W. Prager, A. H. Gee, and L. Berman. Correction of probe pressure artifacts in freehand 3D ultrasound. *Medical Image Analysis*, 6:199–214, 2002.
- [22] G. M. Treece, A. H. Gee, R. W. Prager, C. J. C. Cash, and L. H. Berman. High-definition freehand 3-D ultrasound. *Ultrasound in Medicine & Biology*, 29(4):529–546, 2003.
- [23] J. Varandas, P. Baptista, J. Santos, R. Martins, and J. Dias. VOLUS — a visualization system for 3D ultrasound data. *Ultrasonics*, 42:689–694, 2004.