High Resolution Freehand 3D Ultrasound

G.M. Treece, A.H. Gee, R.W. Prager,
C.J.C. Cash and L.H. Berman

CUED/F-INFENG/TR 438

August 2002

Cambridge University Engineering Department
Trumpington Street
Cambridge CB2 1PZ
England

Corresponding e-mail: gmt11@eng.cam.ac.uk

Abstract

This paper describes a high resolution freehand 3D ultrasound system, whose accuracy surpasses that of previously documented systems. Such accuracy is achieved through a series of novel system design and calibration techniques. The accuracy is quantified using a purpose-built, tissue-mimicking phantom, designed to create idealised *in vivo* scanning conditions. The paper includes a thorough discussion of the various different ways of measuring system accuracy and their relative merits, and compares in this context all recently documented freehand 3D ultrasound systems.

Contents

1 Introduction

Advances in the resolution and quality of two-dimensional (2D) ultrasound imaging are increasingly enabling detailed examination of arterial and musculoskeletal anatomy [15, 31]. However, high resolution ultrasound images (B-scans) have a limited field of view, generally sufficient for scanning the cross-section, but not the length, of the anatomy of interest. Three-dimensional (3D) ultrasound can overcome this limitation. Not only does it provide the ability to generate extended images, it also allows the visualisation of complex structures, like ligaments and cartilage or arterial plaque, in a much more intuitive way. A further advantage is that 3D ultrasound, by its very nature, offers much more precise measurement of volume and the relative orientation of structures.

There are many ways to design 3D ultrasound systems [14]. The most appropriate technique for arterial and musculoskeletal anatomy is freehand 3D ultrasound, where the probe is moved by hand, and the resulting sequence of B-scans is located in 3D space by either intrinsic (image-based) or extrinsic (position sensing) means. This is the only technique which gives the clinician complete freedom to guide the probe along the path of the anatomy.

Most recently documented freehand 3D ultrasound systems use either an electromagnetic [4, 13, 26] or an optical [8, 9] position sensor. B-scans are transferred from the ultrasound scanner to an external PC by digitising from the scanner's video output [13, 25, 22], from a video recording [4] or by direct digital transfer [6]. All position sensing techniques, and most image transfer protocols, introduce additional sources of error not present in the original 2D B-scans: typical accuracy of documented 3D systems is of the order of $\pm$2mm. This is significantly worse than the inherent resolution of high frequency B-scan images, which is better than 0.1mm/pixel.

The position sensing and image acquisition techniques are only two of many steps which affect the eventual resolution of the 3D system. A review of the engineering challenges in such systems, together with some practical suggestions for good acquisition protocols, is given in [16]. This review uses our system, Stradx [26], as a worked example. Stradx is a *sequential* freehand 3D ultrasound system. In this paradigm, the (arbitrarily orientated but only gradually varying) sequence of B-scans is preserved, rather than resampled onto a regular voxel array, such that each visualisation or quantification step is calculated directly from the original data. This is an ideal starting point for improving the resolution of 3D systems, since we preserve the 2D resolution for as long as possible.

In this paper, we present recent developments to Stradx which aim to improve the overall system accuracy, so that the 3D resolution can approach that of the original, high resolution 2D B-scans. The system is rigorously assessed by using a purpose-built, tissue-mimicking phantom, allowing us to estimate the actual errors which might be expected when making 3D measurements in real clinical situations. We also consider the practical aspects of using the system and investigate the degradation due to re-mounting the position sensor on the ultrasound probe, and changing the ultrasound machine's depth setting, without carrying out a full re-calibration.

1.1 Analysis of errors in freehand 3D ultrasound

Before considering how we might improve the resolution of a freehand 3D ultrasound system, we need a clear picture of where the errors come from. An overview of the main sources of error is given in Figure 1. These can be grouped into errors in the B-scan images themselves,

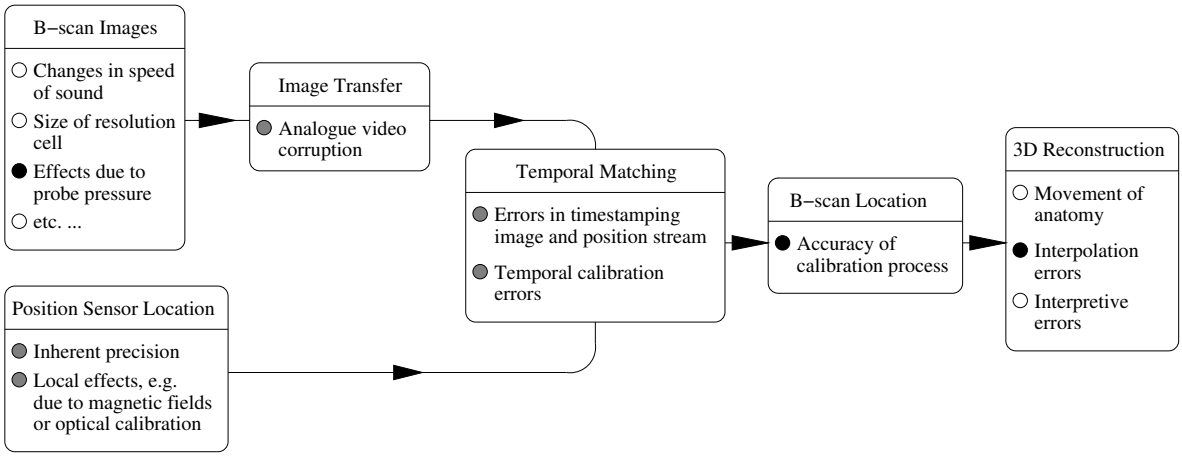


Figure 1: Errors in freehand 3D ultrasound systems. Some of the dominant errors are shown: those with black bullet points have been addressed in previous work, those with mid-grey bullet points are addressed in this paper, those with white bullet points are residual sources of error.

the readings from the position sensor, temporal matching of B-scans and positions, location of the B-scan relative to the position reported by the sensor, and errors in the 3D reconstruction of the B-scans.

Errors in the B-scans themselves are largely determined by the size of the resolution cell, which varies in all three dimensions. Typically the out-of-plane resolution (or beam width) is significantly worse than the in-plane resolution, and varies across the depth of the image, dependent mainly on the out-of-plane focusing. Variation in the speed of sound can also have a significant effect on the beam width: errors of $\pm 5\%$ in sound speed, typical of variations in human tissue, can generate over 200% increases in beam width, as well as affecting the depth scale [1]. For high resolution images, compression of anatomy due to probe pressure can be a large source of error, but this can be reduced by image correlation techniques [30].

Of all the position sensing techniques, optical position sensors are the most accurate, although they require a line of sight between the probe and the camera. Such systems can achieve a spatial accuracy (for the position sensor alone) of up to ± 0.2mm [8]. Electromagnetic position sensors can achieve an accuracy of up to ± 0.5mm in location and $\pm 0.7^\circ$ orientation (again, for the position sensor alone) when optimised for very specific situations and small spatial ranges [3]: in general use, however, they are subject to distortions which impair their accuracy considerably [7].

Errors due to image transfer and temporal calibration (the matching of images to positions) are less widely discussed in the literature: most researchers opt for the practical solution of digitising the analogue video output of the ultrasound machine, at between 10 and 25 frames per second. The images are corrupted by conversion to and from analogue video formats, and the temporal resolution is limited by the low frame rates. A notable exception is the system described in [6], where images are transferred digitally at 150 frames per second, thus giving a temporal resolution of 7ms. Temporal calibration at 25 frames per second can be achieved to within 40ms by looking for sudden changes in the image and position streams [22, 26]. Inaccurate temporal calibration will result in spatial errors with a

magnitude dependent on the speed with which the probe is moved.

Spatial calibration, the estimation of the rigid body transformation between the position sensor's reference frame and the B-scan plane, is one of the most dominant sources of error in freehand systems. The various calibration techniques are compared in [27]. The calibration process involves scanning a known object from a variety of orientations — this can be a single point [20], a set of points [6], a cross-wire [4, 22], a 'z-shape' [9], a real or virtual plane [27] or in fact any known shape [8]. By constraining the 3D reconstruction to match the known geometry of the scanned object, it is possible to derive a system of equations for the eight spatial calibration parameters (six defining the location and orientation of the B-scan relative to the position sensor, and two defining the x and y scales of the B-scan in mm/pixel). The system of equations can either be inverted directly, or more usually optimised iteratively.

Even after the location of each B-scan has been correctly determined, there are still further sources of error. An unspoken assumption in the subsequent 3D reconstruction is that the subject has not moved during the acquisition: any such movements result in a distortion of the 3D data. External movement can be ameliorated by attaching a coordinate reference to the patient [12], and repetitive internal movement (i.e. due to cardiac activity) by the use of an electrocardiogram to gate the acquisition of B-scans [5, 24]. In general, it is best to acquire data within a single breath hold and review it immediately for motion artefacts, so that the scan can be repeated if necessary [16].

Visualising the semi-structured 3D freehand data involves interpolating the data onto some sort of regular pixel or voxel array. Significant interpolation errors can arise as a consequence of the scanning pattern [10] combined with simplistic interpolation schemes, optimised for speed rather than quality [28]. Such errors can be limited by not resampling onto a regular voxel array, as described earlier. This approach also suppresses some interpretive errors, for instance when delineating structures in artefact-ridden out-of-plane reconstructions [2]. Other interpretive errors arise from poor cursor placement when making measurements [18].

1.2 Assessment of system accuracy

There are many ways of assessing the performance of a freehand 3D ultrasound system, and unfortunately there is no agreed standard. More confusingly, results are generally quoted simply as system 'accuracy', despite differences in what was tested, where it was tested and how the results were analysed, which can lead to as much as a factor of three variation in the quoted result. Sometimes, insufficient information is provided to be able to interpret the quoted 'accuracy' at all.

It is therefore necessary to clarify the differences between some of these measures before attempting to compare those systems which are described in the literature and place our system amongst them. This can helpfully be done by asking three questions: "*What part of the system is included in the measurement?*", "*What is it a measurement of?*", and "*How are the measurements analysed?*".

Firstly, "*What part of the system is included in the measurement?*". With reference to Figure 1, when designing a system it might be helpful to know the accuracy of a specific part, for instance the position sensor alone [3], or the spatial calibration alone [27]. Ultimately, however, it is the accuracy of the entire system, in the context in which it will be used, which is relevant to the clinician. *In vivo* accuracy is very difficult to assess, so *in vitro* accuracy is usually reported instead, by scanning a specially designed phantom in a water bath. This excludes some of the B-scan image errors, such as the speed of sound variation

in human tissue, and tissue deformation due to probe pressure. It also excludes some of the 3D reconstruction errors, specifically those due to movement of anatomy, and to some extent those due to interpretation of data (since phantom images are often significantly less complex than *in vivo* images). Clearly, the accuracy of the *entire* system can only be worse than that of its component parts.

Secondly, "*What is it a measurement of?*". There are a variety of possibilities here, in terms of the *quantity* measured, *where* it is measured, and what it is *compared* to. Generally the *quantity* is either the location of a fixed point [4, 8, 9, 22, 27], the distance between points [8, 20, 27] or the volume of a defined object [4, 6]. The location of a point can perhaps be regarded as a more fundamental measure, since volume and to a lesser extent distance are not affected by certain distortions of the 3D data. *Where* the quantity is measured is particularly important if the spatial calibration has been optimised from the same data used to assess the system accuracy (which is, unfortunately, common practice in the literature). If this is the case, then what is being measured is only the calibration residual error [4], but how well this reflects the actual system accuracy is highly dependent on how well conditioned the calibration optimisation is, and how well the calibration scanning pattern represents actual scanning practice. Even a repeated scan of the point on which the calibration was based [27] can be misleading: it is better to assess accuracy based on a completely different set of measurements. Finally, measurements can either be *compared* to 'true' values (known from some other independent source), in which case they reveal system accuracy, or to themselves, in which case they reveal only the precision of the system.

Thirdly, "*How are the measurements analysed?*". As an example, consider a set of measurements of point location, with independent errors in each of the x, y and z dimensions which are normally distributed with zero mean and 1mm standard deviation. The 95% confidence limits in each dimension are approximately twice the standard deviation, i.e. $\pm$2mm. However, the absolute 2D location error (for instance in the x–y plane) is not normally distributed: it follows a Rayleigh distribution. The absolute 3D location error has a more complex distribution still. The *mean absolute 2D error* is approximately 1.25mm, and 95% of the points lie within 1.85 times this, i.e. within 2.3mm of the true location. For the 3D case, the mean error is 1.6mm and the 95% limit is $<$ 2.8mm (approximately 1.75 times the mean). Note that these confidence limits are lower than the pessimistic estimate from simply summing the variances in each dimension, as in [20]. The standard deviation of the 2D and 3D errors is sometimes also quoted: this is a misleading quantity, since these errors are not normally distributed, and can lead to optimistic assessments of system accuracy. For the example above, the standard deviation is approximately 0.7mm, for both the 2D and 3D cases.

A further complication arises from the use of paired analysis [8, 27], where a set of point measurements is analysed by considering the distribution of the *absolute distance between all possible pairs of measurements*. If we continue the example above, this analysis would give a mean 3D error of 2.3mm and 95% confidence limit of $<$ 4mm, whereas we already know that 95% of the values will lie within 2.8mm of the correct location. In effect, the paired analysis measures *relative distance* accuracy, which has twice the variance of the *point location* distribution, since it is a measure of difference.

Although it is not in general possible to compare accuracy results which differ with regard to the first two of these questions, it *is* possible to use the example above to convert between results which are analysed differently, provided sufficient information has been given to determine the nature of the result which has been presented. In the comparison below, results are

converted to the *3D confidence limit*: this is the distance away from the mean location (for precision) or known location (for accuracy) within which 95% of the measured points will lie.

The performance of a freehand 3D ultrasound system is clearly dependent on the type of position sensor and the frequency of the ultrasound. Prager et al. [27] used an electromagnetic position sensor and a 7MHz probe at a 4cm depth setting, Blackall et al. [8] used an optical position sensor and a 10MHz probe, again at a 4cm depth setting, Meairs et al. [22] used an electromagnetic position sensor and a 5-12MHz probe, Legget et al. [20] used an electromagnetic position sensor and a 3MHz probe, and Bouchet et al. [9] used an optical position sensor and a 3.5MHz probe. Systems using optical position sensors and higher frequency probes at lower depth settings will tend to be more accurate.

Since spatial calibration is such an important step in the design of an accurate system, several authors quote the point precision due to the spatial calibration alone. The 3D confidence limits achieved for this value are < 1.2mm [27] and < 2.3mm [8] (both derived from mean paired absolute error). Another frequently quoted value is the point precision of the entire system (as measured by scanning a phantom — note the earlier comments about *in vitro* measurements). 3D confidence limits for this value are < 2.7mm [27], < 1.4mm [8] and < 2.6mm [22] (all derived from mean paired absolute error), < 3.4mm [20] (derived from the sum of the variances in each dimension) and < 2.2mm [9] (derived from the mean absolute error in each dimension). Finally, several authors quote the errors in distances between several points: this leads to a measure of accuracy *within* a particular data set, which is sensitive to any distortion introduced into the data, but not to systematic errors in point location. 3D confidence limits for point location accuracy based on this measure are < 1.9mm [27], < 1.0mm [8] and < 1.1mm [20] (all derived from the standard deviation of the paired signed distance errors).

2 High resolution system

2.1 Physical layout

Figure 2 shows the physical layout of our freehand 3D ultrasound system. B-scans are acquired with a Diasus ultrasound machine[1], using 5-10MHz and 10-22MHz linear array probes, on 2cm, 3cm and 6cm depth settings. 8-bit digital log-compressed data is transferred via ethernet at 25 B-scans per second to an 800MHz PC running Linux. The probe position is measured by a Polaris[2] optical system tracking an AdapTrax[3] target attached to the probe. The camera is mounted on a stable, but highly manoeuvrable stand[4] designed for studio video cameras. Calibration, acquisition, processing and display of the data are performed on the PC by Stradx [26][5], which can also be used in conjunction with a number of other video sources and position sensors.

In order to avoid having to re-calibrate the system each time a different probe is used, the probe mount shown in Figure 3 was designed. The mount can be attached to most ultrasound probes by simply winding a length of Velcro™ tightly around the mount and the probe: the rubber holds the mount firmly in place, and the locating hole ensures that the position sensor

[1] Dynamic Imaging Ltd., http://www.dynamicimaging.co.uk/
[2] Northern Digital Inc., http://www.ndigital.com/
[3] Traxtal Technologies, http://www.traxtal.com/
[4] Unicol, http://www.unicol.co.uk/
[5] http://svr-www.eng.cam.ac.uk/~rwp/stradx/

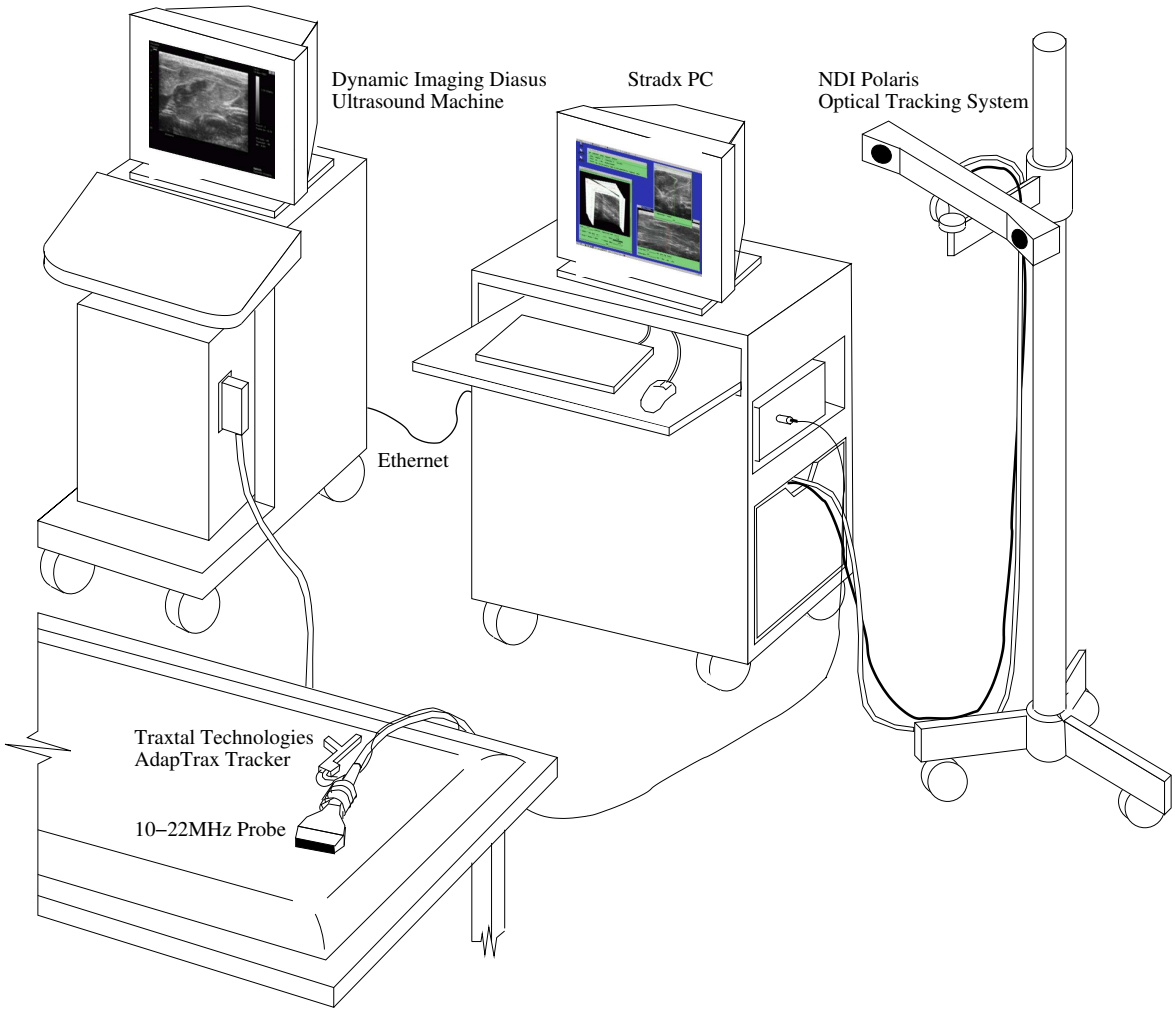


Figure 2: Physical layout of the high resolution freehand 3D ultrasound system. The diagram indicates the approximate size and the required connections between the components, to help give a feel for the system's usability.

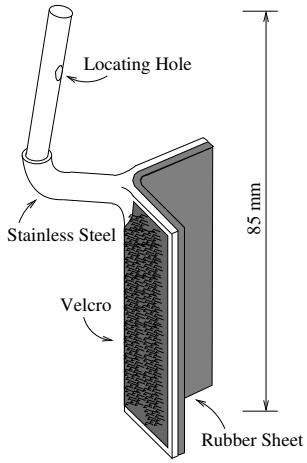


Figure 3: Generic probe mount for the optical tracker. The mount can be quickly attached to almost any type of ultrasound probe, and is sufficiently small and cheap to remain attached between 3D scanning sessions. The optical tracker can be fixed at the same position on the mount to a high degree of precision, such that spatial calibration is only necessary if the mount itself is removed from the probe.

can be fixed to the mount at a known, repeatable position. The mount is small and its presence on the probe does not inhibit normal use. In practice, we have sufficient mounts for all the probes in use, and these mounts remain on the probes between 3D scanning sessions.

2.2 System design

Figure 4 shows the basic design of the part of the system involved with delivering B-scans, tagged with their correct spatial position and orientation, to Stradx's front end visualisation and quantification algorithms (the sequential reslice, manifold and panoramic displays are described in [26], volume rendering extensions in [17], and volume measurement and surface reconstruction in [29]). We are concerned in this paper with the quality of the raw data on which these algorithms are based.

Key components of this system, discussed in the following sections, are the transfer of ultrasound images to the PC, the synchronisation of the PC and ultrasound machine timers, the matching of images to positions and the gating of acquired images. Temporal and spatial calibration of the system will be discussed later in Section 3.

2.2.1 Ultrasound image transfer

In almost all other freehand 3D ultrasound systems, B-scans are transferred between the ultrasound machine and the PC in an analogue video format. This is an appealing transfer mechanism because most ultrasound machines have a suitable analogue video output, and video framegrabbers for PCs are cheap and commonplace. However, the unnecessary digital to analogue conversion in the ultrasound machine, followed by an analogue to digital conversion in the framegrabber, introduces unwanted noise into the B-scans. For our high resolution system, we would like to avoid this source of error and instead transfer the images in an

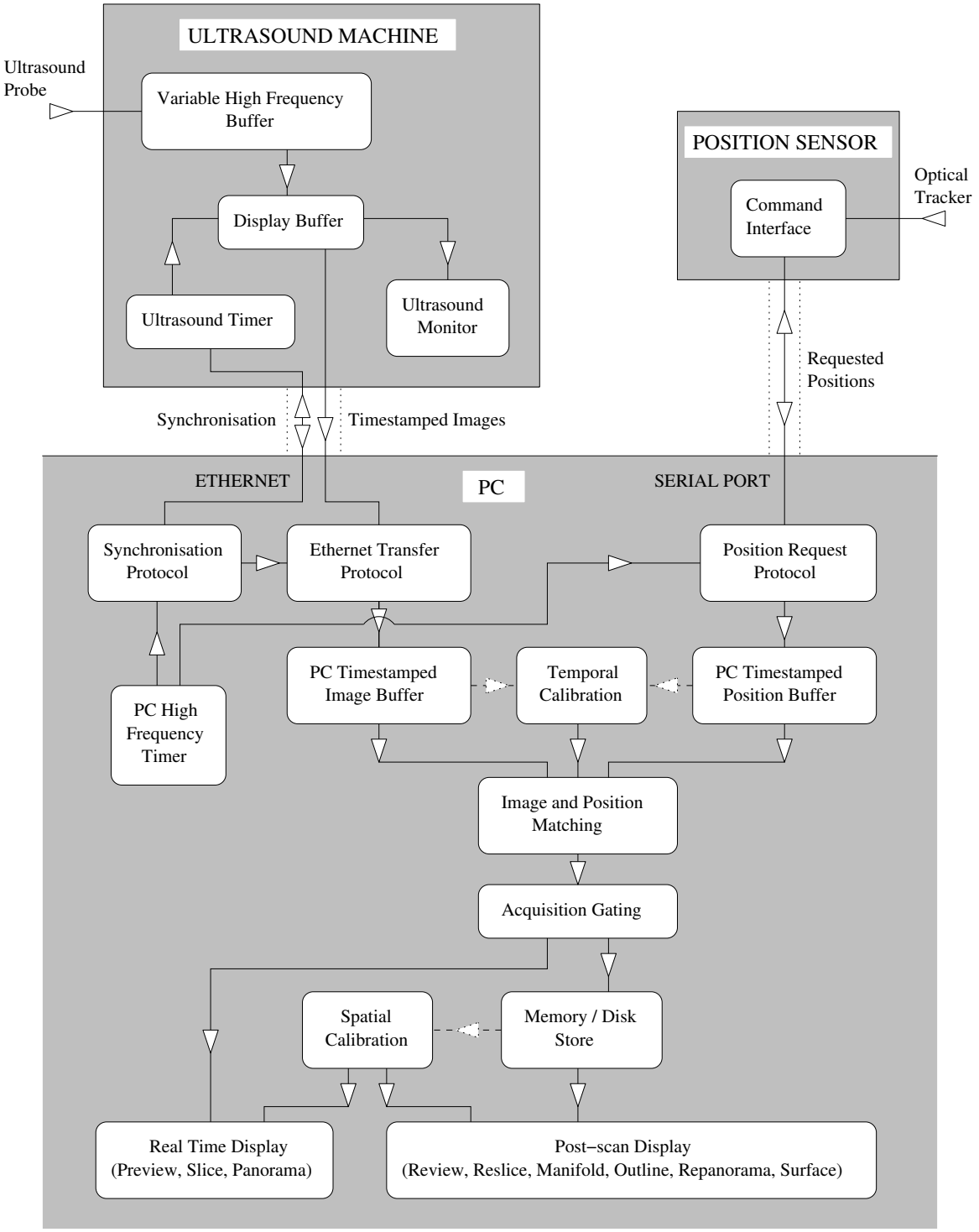


Figure 4: High resolution freehand 3D ultrasound system design. The main hardware and software components of the PC handling the acquisition and display of the B-scan images are shown, together with the relevant parts of the ultrasound system.

uncorrupted digital format.

One solution would be to employ a custom digital link, like the one described in [6]. However, in the interests of cost, flexibility and rapid prototyping, it is preferable to avoid specialised hardware and work instead with standard PC components. The Diasus, like many modern ultrasound machines, is equipped with an off-the-shelf PC front end which includes an ethernet network interface. This raises the possibility of streaming the B-scans across the network: for a switched 100MBit/s link, and assuming one byte per greyscale pixel, it should be possible to transfer large B-scans (up to 724×724 pixels) at the full PAL frame rate of 25Hz.

This solution proved to be both feasible and highly effective. Minimal changes to the Diasus's internal software were required to send each B-scan to the ethernet card at the same time as it was displayed on the screen. A suitable module was added to Stradx to receive the frames from the ethernet instead of the usual video framegrabber. The software for both ends of the link was developed in a few man days, with no need for specialised hardware. Moreover, the resulting system is compatible with any ultrasound machine equipped with a PC front end. Even though the Diasus is a purely greyscale machine, the network transfer mechanism could work equally well with Doppler machines, provided each pixel is coded into a single-byte colour index before transmission[6].

The one complication concerns the timestamping of the B-scans. Each B-scan needs to be tagged with the time at which it was acquired, so that it can later be matched with a position read from the position sensor at the same time. Timestamping the B-scans when they are received by the client PC would be unreliable, since the time of acquisition is not deducible from the time of receipt: there is an unpredictable network latency to consider. So the timestamps must be applied by the server (the ultrasound machine), and sent down the ethernet link with each B-scan. In our case, the timestamps are applied at the point at which B-scans are extracted from the display buffer, which runs at a constant rate of 30Hz. The rate at which the scans are actually acquired is usually higher than this, dependent mainly on the number of focal points in the transmit path. Applying the timestamps at the display buffer rather than the variable rate high frequency buffer introduces a delay which is accounted for by temporal calibration[7].

Sending a suitably high resolution timestamp down the ethernet link requires only a few extra bytes per B-scan, so there is no impact on the ethernet frame rate. But there is an issue with clock synchronisation: the clock in the ultrasound server will not be telling the same time as the clock in the client PC, and will not even run at the same rate: there will be a certain amount of relative clock drift. Since the positions are timestamped by the client PC, we need a way to translate an ultrasound server timestamp into a client PC timestamp, so the B-scan times can be compared directly with the position times.

A simple network time protocol was developed to achieve this. When the client PC first establishes ethernet communication with the ultrasound machine, it requests a single timestamp t_s from the server. When t_s is received, the client immediately takes a local timestamp t_c. The difference $\Delta_1 = t_c - t_s$ can be used to translate server timestamps into client timestamps. However, t_s might have been held up between the server and client by a network delay, causing Δ_1 to be bigger than it should be, so the exchange is repeated ten

[6] 128 grey levels, and 64 shades of red and blue, are all that is needed to represent a colour Doppler ultrasound image.

[7] This delay is itself dependent on the actual rate of acquisition, and will therefore change with the number of focal points.

times and the smallest Δ_1 is retained.

This simple procedure captures the time difference between the two clocks at an instance, but tells us nothing about the relative clock drift. The B-scan timestamps will soon become inaccurate unless we account for this drift. So the next time the client connects to the server (perhaps at the start of a 3D ultrasound recording), the timestamp exchange is repeated, yielding a new difference Δ_2. Since we know when Δ_1 and Δ_2 were measured, we can estimate the relative clock drift and use this to correct subsequent server-to-client timestamp translations. Furthermore, a new value of Δ can be measured at the start of each 3D recording, leading to better estimates of the drift over longer time scales. The end result is accurate B-scan timestamps, applied by the ultrasound machine but reliably translated into client time, ready for matching with timestamped positions from the position sensor.

2.2.2 Image to position matching

Recently acquired B-scans are stored in a circular buffer in the client PC's memory. A similar circular buffer holds recently acquired positions. B-scans are harvested from the circular buffer after a slight delay: that is, we start by processing a recent B-scan, though not the most recent one. This delay is important because we might not yet have a position to match with the most recent B-scan, depending on the latency of the position sensor[8].

The timestamp of the harvested B-scan is compared with the timestamps in the position circular buffer. An accurate position for the B-scan is obtained by linear interpolation of the two position readings on either side of the B-scan timestamp. At this point, the B-scan is tagged with its position and made available to Stradx's visualisation front end. The process continues with the next frame in the B-scan circular buffer: further details can be found in [26].

2.2.3 Acquisition modes

In its default mode, Stradx processes B-scans sequentially, harvesting one frame after another from the B-scan circular buffer. Making use of every B-scan ensures the maximum possible acquisition rate, though this is not always desirable. For example, if the probe is being held stationary, the system will record endless B-scans on top of each other, repeatedly sampling the same slice of space. For this reason, Stradx offers motion-gated acquisition [11]: after position tagging, B-scans are discarded if the probe has not moved by more than a user-defined threshold since the last retained B-scan. Motion-gated acquisition maintains an even spatial sampling density, independent of probe velocity.

A further acquisition mode is stable-gated acquisition. Again, the probe must have moved by a certain amount for a B-scan to be retained, but the probe must also be moving relatively slowly (another user-defined threshold), otherwise the B-scan is discarded. This ensures the acquisition of well spaced B-scans free of motion blur artefacts. Such data sets are particularly useful for spatial calibration, as discussed in Section 3.

[8] The position sensor latency is accounted for by the temporal calibration procedure described in Section 3. The result of this calibration is used to correct the position timestamps before they are stored in the circular buffer.

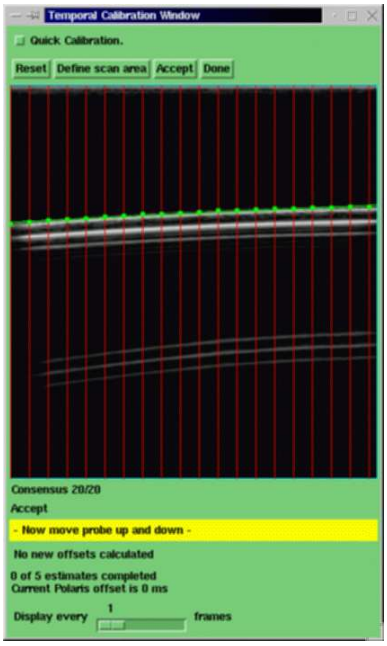


Figure 5: Automatic line detection for temporal calibration. The key to temporal calibration is the ability to provide an automatic, real-time estimate of the location of a horizontal echo in the B-scan, which can subsequently be compared to the position sensor readings.

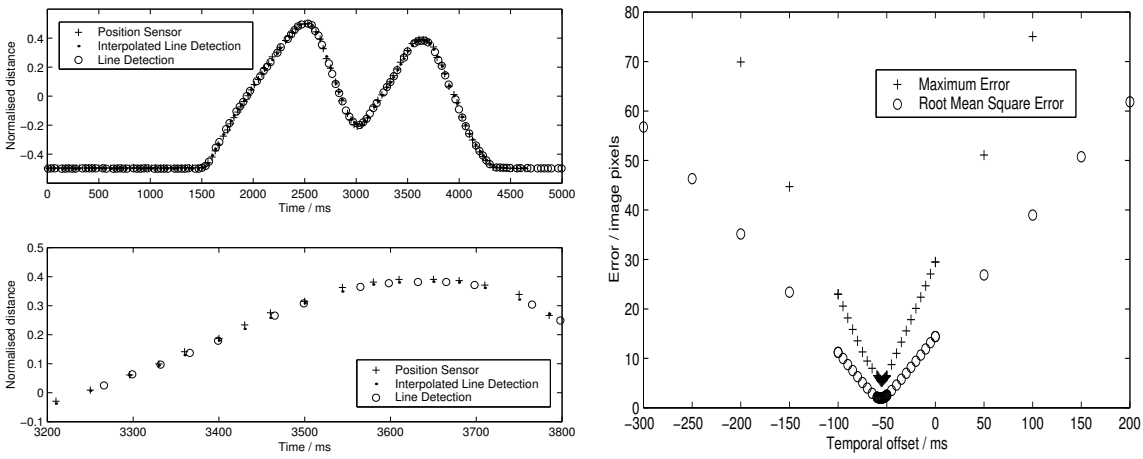
3 Calibration

Temporal calibration is necessary to determine the offset between the position sensor timestamps and the B-scan timestamps. A matching B-scan and position reading will not necessarily have the same timestamp, depending on the latencies of the two data streams. Particularly pertinent is the latency of the position sensor, since the position timestamps are applied not by the position sensor (which has no clock), but by the PC each time it requests a reading from the sensor. There is an unknown delay between the PC making the request and the sensor taking the reading.

Spatial calibration determines the size and location of the invisible B-scan plane relative to the position sensor's reference frame. Since spatial calibration involves the analysis of matched B-scans and positions, it follows that poor temporal calibration will affect the accuracy of the spatial calibration: hence temporal calibration must be performed first. This means that the temporal calibration procedure must be designed with no knowledge of the location or size of the B-scan.

3.1 Temporal calibration

Temporal calibration has previously been achieved by looking for sudden changes in B-scan and position sensor readings, caused, for example, by suddenly jerking the probe away from the skin surface, and matching them up in time [22, 26]. For images acquired at 25Hz, this gives a temporal resolution of only 40ms. It is possible to improve on this value by comparing a stream of position sensor readings with positions derived from the B-scans themselves.



(a) Image and position sensor signals

(b) Signal correlation against temporal offset

Figure 6: Temporal calibration of the B-scan image stream to the position stream. (a) For a given temporal offset, the position signal derived from the images is interpolated to give values at the same time points as each of the position estimates from the position sensor. (b) The (interpolated) image and position sensor position streams can then be correlated at various temporal offsets. A multi-resolution approach is used to speed up the search for the minimum error between the streams.

We have previously described an algorithm which can automatically detect horizontal lines in a B-scan, as in Figure 5 [27]. Such lines can easily be generated by scanning the bottom of a bath full of water. The line detection algorithm can run sufficiently fast to give a real-time estimate (i.e. at the PAL frame rate of 25Hz) of the location of the base of the water bath in the B-scan. Note that this measurement only gives the variation in height of the probe above the (horizontal) surface, and cannot detect horizontal movements. In order to be able to compare this with the position sensor readings, we must ensure that the probe is only moved in the vertical direction.

At this stage, without a spatial calibration, neither the B-scan scale nor the orientation of the B-scan with respect to the position sensor coordinate system are known: we do not know how far the probe has travelled, nor which direction is up. However, assuming the probe was moved vertically, 3D position sensor readings can be converted to vertical distances by taking the scalar product of each reading with the direction of maximal movement over the whole calibration sequence. The sign can be deduced by comparison with the distances derived from the B-scan. Finally, both the B-scan and position sensor distances must be normalised to lie in the range -0.5 to 0.5, to allow for the unknown scale. In order to prevent this normalisation from badly conditioning the problem, only calibrations where the probe has moved a certain distance (as measured in the B-scan) are accepted for further processing.

A temporal calibration therefore proceeds as follows. After one second with the probe held steady, four further seconds of B-scans are acquired, during which the probe is moved up and down above the base of the water bath. After normalisation, this results in a stream of distance measurements from both the position sensor and the B-scan, as in Figure 6(a)

(shown at the correct temporal offset, rather than that at which they were acquired). These measurements are at different rates: approximately every 40ms for the images, and 30ms for the positions. In order to correlate the two streams, distances derived from the images are interpolated with a cubic function to give values which coincide (for a given temporal offset) with those from the position sensor.

The temporal offset is found which gives the minimum root mean square error between the distances derived from the B-scans and those derived from the position sensor. The correlation function is generally well behaved with respect to the temporal offset, as can be seen in Figure 6(b), so a simple multi-resolution search can be used to find the optimum offset. The maximum error for a given offset is also calculated, and the result is disregarded if this is greater than 10 pixels (as measured in the B-scan): this gives some protection against calibrations where the probe was not moved with sufficient care.

The whole process is guided by prompts from the user interface as to when to hold the probe steady, when to move it, and whether the movement was sufficient and the eventual correlation good enough. Acceptable results can either be used individually, or averaged to improve the precision. With a little practice, ten or more calibrations can be performed within one minute, leading to a temporal resolution which is significantly better than that of the original image and position streams.

3.2 Spatial calibration

Spatial calibration of the high resolution system is performed using the Cambridge phantom [27], which essentially offers calibration on a flat plane. The plane appears as lines in the B-scans: by constraining the 3D reconstruction of these lines to be co-planar, the eight calibration parameters can be estimated. Moreover, since the lines can be detected automatically in the B-scans, as illustrated in Figure 5, the calibration process is extremely rapid. However, the accuracy of this approach is limited by beam width effects: with the probe held at an oblique angle to the plane, reflections from the edge of the ultrasound beam can cause the plane to appear in the wrong position in the B-scan. The Cambridge phantom overcomes this problem by generating a 'virtual' plane, with reflections coming only from the centre of the ultrasound beam. Full details of the phantom and its use are described in [27]. Here we describe some refinements which increase the accuracy of the spatial calibration.

Firstly, the improved temporal calibration process described in Section 3.1, combined with the stable gated acquisition mode described in Section 2.2.3, significantly reduces temporal errors in the calibration process: the result is more accurate labelling of each calibration B-scan with its position. In addition, the high quality of the digitally-transferred images, as described in Section 2.2.1, improves the accuracy with which the line detection process can be performed. We also aim to keep the water fairly hot during the calibration: tap water at 50°C has the same speed of sound, 1540ms^{-1}, as is normally assumed for average human tissue. These steps improve the conditioning of the equations which must be solved to give the spatial calibration parameters, and ensure the suitability of the resulting parameters for normal clinical use.

Secondly, we have slightly adapted the pattern in which the probe is moved during calibration, in order to constrain the calibration parameters reliably. This pattern, shown in Figure 7, has been derived experimentally. Motion C is particularly important to distinguish between the z (out of plane) translation, and rotation of the B-scan about a vertical axis.

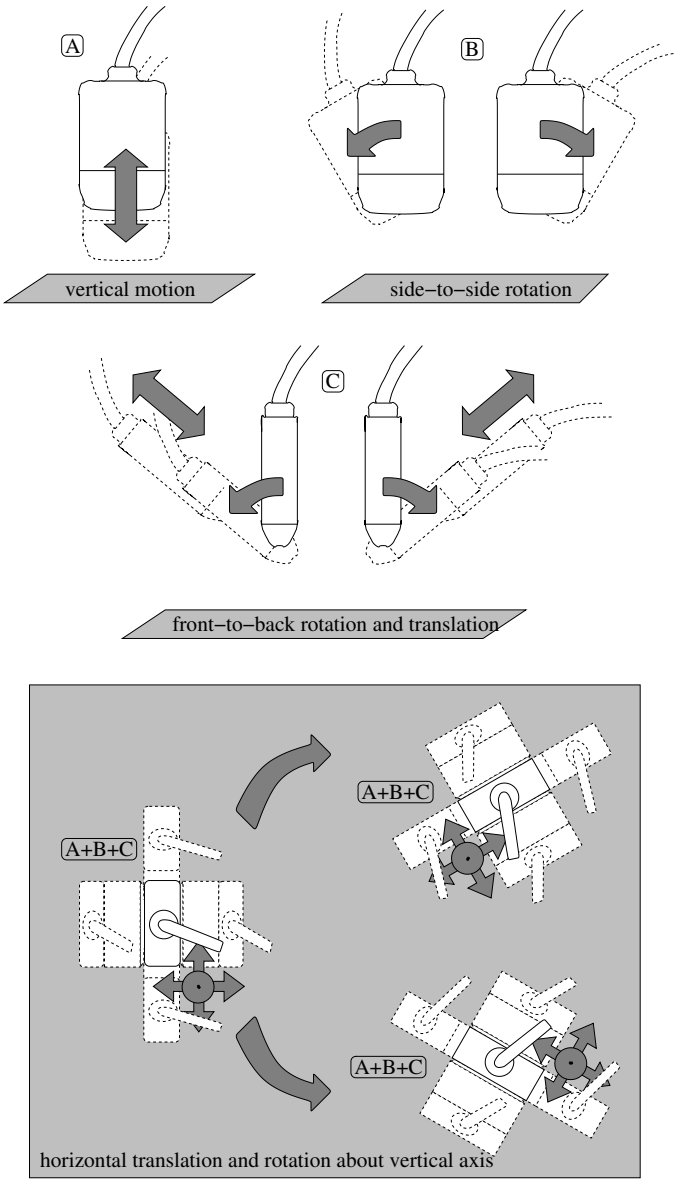


Figure 7: Probe movement during spatial calibration. The sequence of movements A, B and C is repeated at three locations and orientations on the plane being imaged. This is necessary to define both the spatial calibration parameters and the location of the calibration plane.

3.3 Depth change re-calibration

While spatial calibration can be performed in around ten minutes using this technique, it must be repeated whenever the relative locations of the position sensor and the ultrasound probe, *or* the size or location of the B-scan in the video frame, change. This latter restriction implies that pan, zoom and depth changes require a new set of spatial calibration parameters and hence must be avoided during a scanning session: a constraint which is contrary to normal clinical practice. However, such changes do not affect the rotational calibration parameters — they only affect the translation $(t_x,\ t_y,\ t_z)$ and the scale $(s_x,\ s_y)$. Indeed, changes to this subset of parameters are dependent on only three independent variables: the translation in the plane of the B-scan $(b_x,\ b_y)$, and the zoom z.

In certain situations, b_x, b_y and z can be estimated directly from the ultrasound image without the need for a full spatial re-calibration. As long as the top and some of both sides of the B-scan data can be seen, we can estimate the location of the central top point of the B-scan $(c_x,\ c_y)$ relative to the top left hand corner of the image, and the B-scan width (w). This estimation is trivial for linear probes. For convex probes, the point $(c_x,\ c_y)$ is the centre of curvature, w is the width at the top of the B-scan, and an additional parameter r is required to define the radius of curvature [30].

If these probe parameters were $(c_{xo},\ c_{yo},\ w_o)$ at the time the spatial calibration was performed, and are measured as $(c_{xn},\ c_{yn},\ w_n)$ after a change to the zoom or depth settings, then

$$z = \frac{w_o}{w_n} \tag{1}$$

$$\Delta b_x = c_{xo}s_{xo} - c_{xn}s_{xn}, \quad \Delta b_y = c_{yo}s_{yo} - c_{yn}s_{yn} \tag{2}$$

where $(\Delta b_x,\ \Delta b_y)$ is the change in offset within the plane of the B-scan. Changes in the actual calibration parameters $(t_x,\ t_y,\ t_z,\ s_x,\ s_y)$ can be derived from these values[9].

This means that for most depth changes, and some pan and zoom changes, the spatial calibration can be updated by analysing a single image from the ultrasound machine. Furthermore, this analysis (i.e. the detection of the top and sides of the B-scan data) can be carried out by automatic image processing algorithms [30].

4 Experimental method

4.1 Ultrasound phantom

In order to establish the performance of the system described in Section 2, a highly accurate, tissue-mimicking ultrasound phantom was required. Phantoms consisting of very thin nylon wires embedded in a tissue-mimicking material are often used to determine the resolution of 2D ultrasound machines. However, these phantoms are only designed to be scanned from one insonification angle. To test a freehand 3D ultrasound system, we need a target which can be scanned from multiple angles, such as a small, non-echogenic sphere. Phantoms containing such targets are used to determine lesion detectability for 2D ultrasound systems. If we know the exact location of each sphere, then we can compare the known location with the apparent location from each of a sequence of freehand scans of the phantom, and therefore deduce system accuracy in a (nearly) clinical setting.

[9] The spatial calibration translation $(t_x,\ t_y,\ t_z)$ is in the coordinate frame of the position sensor, *not* the B-scan.

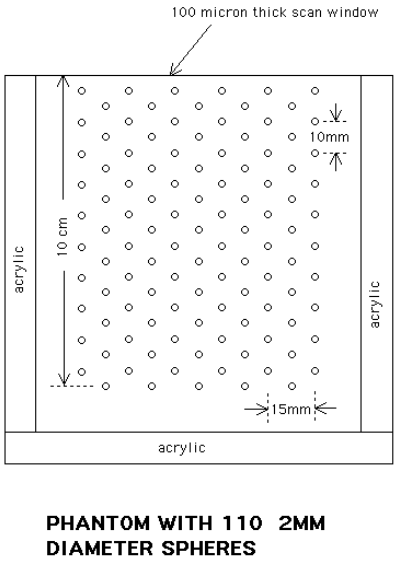
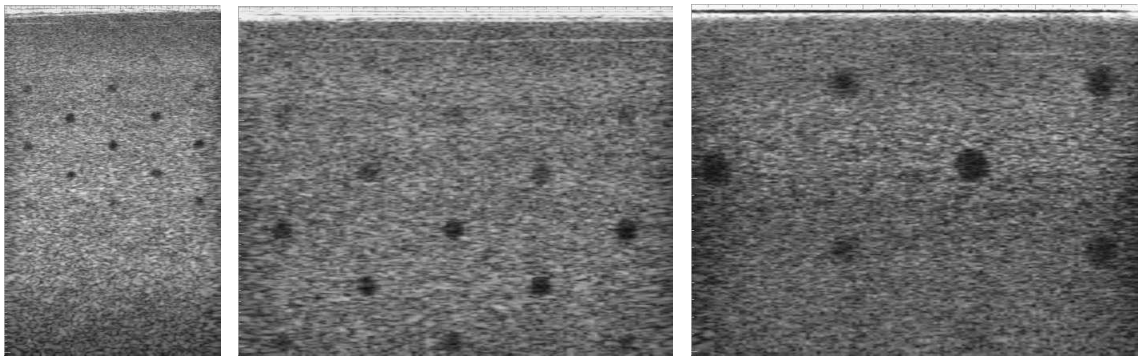


Figure 8: Schematic of the ultrasound phantom. The phantom is made from a tissue-mimicking substance containing a planar array of 110 precisely located, non-echogenic, 2mm diameter spheres [19, 21].

Both the materials and the manufacturing process exist to make such a phantom, consisting of a grid of 2mm diameter spheres which are located to an accuracy of ±0.1mm. A cubic phantom, with sides approximately 10cm and spheres arranged as in Figure 8, was designed and manufactured specifically for this purpose by Prof. Madsen at the University of Wisconsin. The manufacturing process is described in [19] and the tissue-mimicking material used for the non-echogenic spheres in [21].The phantom has a water bath at its top so that it can be scanned from varying directions.

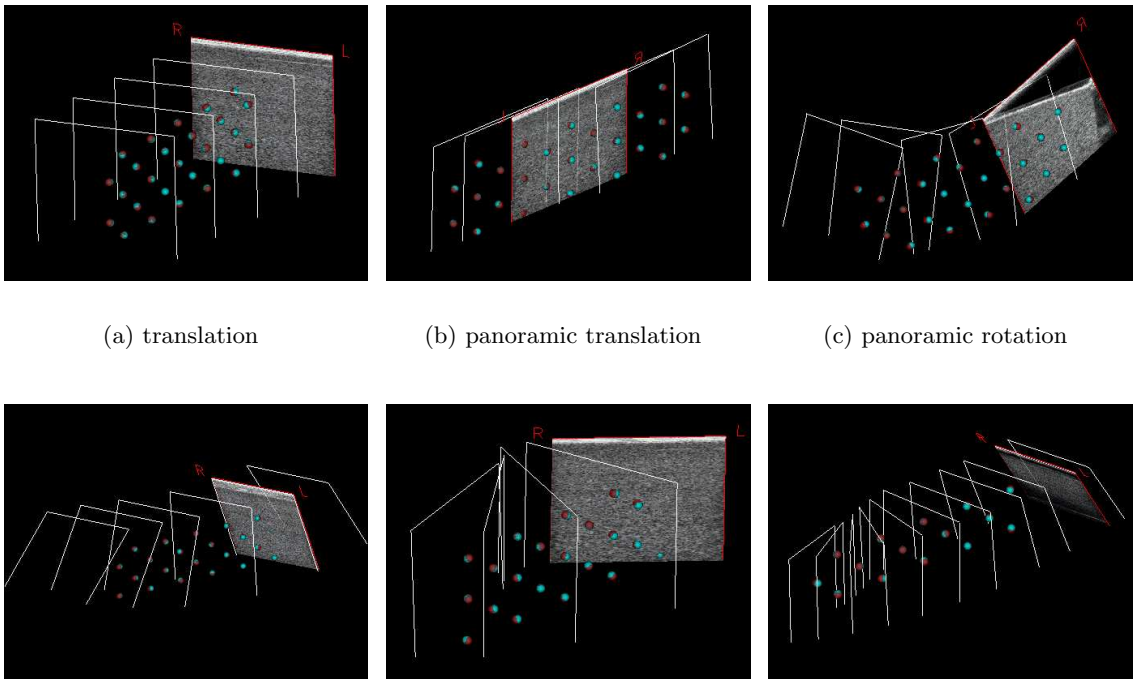
Figure 9 shows some typical B-scans of this phantom, at different frequencies and depth settings. The propagation speed and attenuation coefficient for the background material are 1540ms^{-1} and 0.55dBcm^{-1}MHz^{-1} respectively, and for the spheres are 1541ms^{-1} and 0.53dBcm^{-1}MHz^{-1}, measured at 22°C and 8MHz. Close matching of these values ensures that the spheres generate only minimal distortion to the underlying image. Nevertheless, the backscatter of the sphere material is 40dB lower than that of the background, hence the spheres are clearly visible as black circles. Note that the spheres are not necessarily visible across the entire depth of the image: this is due to the beam thickness being greater than 2mm at very shallow or very deep locations. Scans were performed using four or five focal points, arranged across the mid-region of each B-scan, in order to maximise the visibility of the spheres without reducing the frame update rate below 25Hz.

For each test, the phantom was scanned with five patterns, shown in Figures 10(a) to (e). These represent the five basic motions of the probe during freehand 3D ultrasound acquisition. Patterns (b) and (c) actually generate extended-field-of-view data — although not strictly 3D, this sort of data comes 'for free' with freehand 3D systems [16]. A combined motion, as in Figure 10(f), was used to test the repeatability of mounting the position sensor on the probe.



(a) 6cm (b) 3cm (c) 2cm

Figure 9: Typical scans of the ultrasound phantom at varying depth settings. (a) and (b) were scanned using a 5-10MHz probe, and (c) using a 10-22MHz probe.



(a) translation (b) panoramic translation (c) panoramic rotation

(d) horizontal rotation (e) vertical rotation (f) combined

Figure 10: Scanning patterns. A few sample B-scans, displayed as white 'goal posts', are shown to demonstrate the scanning pattern: typically four to five hundred B-scans were acquired in each sequence.

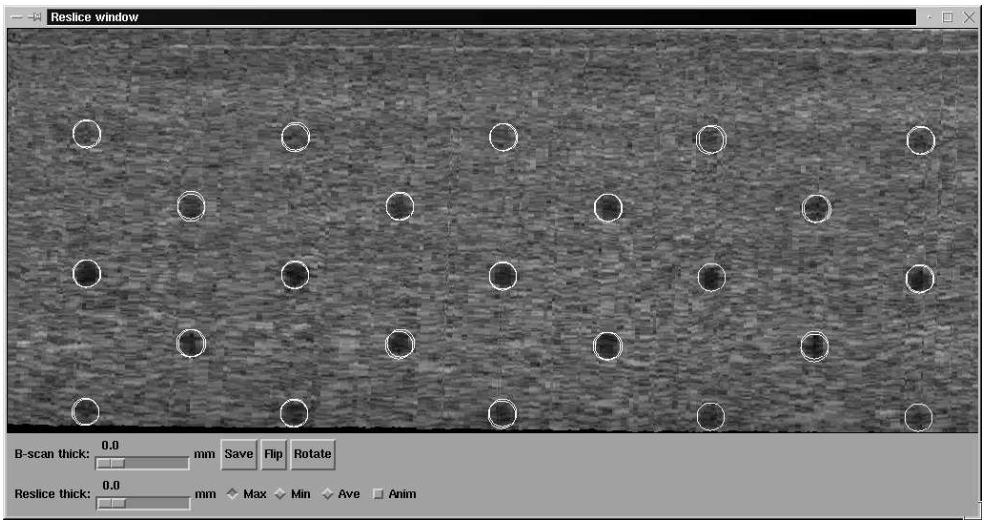


Figure 11: Reslice through the phantom data set. Two white circles are shown for each sphere. One circle shows the location of each sphere as derived from the B-scan data. The other is a regular array of spheres with relative locations as in Figure 8, whose global location and orientation has been optimised to match as closely as possible the set of data-derived spheres.

4.2 Automatic detection of spheres

In order to be able to compare the location of a sphere as determined from a freehand 3D data set with its true value, we need an accurate method of locating the centre of a sphere contained in a set of sequential B-scans. The number of measurements required (over 5000) prohibited the use of entirely manual sphere location. Hence, an automatic method was used to refine the location of each sphere centre, given an initial approximate location from the user clicking in a B-scan. The same approximate location could be used to initialise the algorithm for all sets of spatial calibration parameters, thus significantly reducing the number of manual interactions required.

For each candidate centre point on each B-scan within the sphere diameter (2mm) of the initial click, the B-scan was multiplied locally with a 2D mask, radius 1mm, of the form $m(x, y) = (1 - \sqrt{x^2 + y^2}), (0 \leq m \leq 1)$. The point of minimum summed response to this mask was considered to represent the centre of the intersection of the sphere in any given B-scan. The 3D location of the sphere centre was then calculated by assuming that the scans were approximately parallel in the locality of the sphere, and interpolating between the centres in the three B-scans with the minimum responses.

Although the geometry of the phantom is known, the exact location of each sphere in the position sensor coordinate system is not. Sphere locations derived from the ultrasound data were therefore compared to a regular grid of spheres, spaced as in Figure 8 and *globally* aligned to the measured sphere centres using a non-linear optimisation algorithm (Levenberg-Marquardt [23]). A typical result of this process is shown in Figure 11.

There are several important points to note regarding this method of analysis:

- Only the alignment of the grid of spheres was optimised, *not* the scale parameters.

Table 1: Accuracy of automatic sphere detection. The mean error and 95% confidence intervals (±95%) are given for sphere detection accuracy, estimated from ten sphere location measurements.

Frequency (depth)	**mean error (mm)** 3D	95% **confidence limits (mm)** x	y	z	3D
5-10MHz (3cm)	0.16	±0.19	±0.19	±0.29	< 0.35

- The grid was optimised to all the visible sphere centres (i.e. all those which could be detected by the algorithm described above): there were typically 20 or more in each scan.
- By definition, such analysis will result in a total mean location error in each of the x, y and z dimensions of zero.
- The variance of the error will be slightly less than the population variance. Six parameters (the location and orientation of the plane containing the spheres) were derived from the data, therefore we must multiply the mean square error by $\frac{3n}{3n-6}$ (n being the number of points) to give an unbiased estimate of the variance. Normally, we would only need to multiply by $\frac{n}{n-1}$ to allow for the use of the mean value in each dimension.
- Given the adjustment above, the variance of the error gives an unbiased estimate of the location accuracy. It does *not* reveal any systematic errors in locating a feature with respect to the position sensor coordinate system (which is what might be required for an ultrasound-guided biopsy, for instance). However, it *does* reveal the accuracy of locating a feature with respect to another feature within the data set, which is the important quantity for diagnostic medical imaging. Distance and volume measurement accuracy can also be derived from this result.

5 Results

Wherever confidence limits are given in the following analysis, these are derived from unbiased estimates of the population statistics, allowing for the quantity of measured data and the number of parameters derived from this data during the calculation. In each case, the coordinate system is aligned with the phantom, such that x is along the rows of spheres, y down the columns and z out of the plane of the spheres.

Three probe frequencies and depth settings were used in the following experiments. The 10-22MHz probe was the highest frequency probe available, and 2cm was the most shallow depth setting. The 5-10MHz probe was also tested at depth settings of 3cm and 6cm; these collectively cover the working range for musculoskeletal and arterial scans. The B-scan pixel size was approximately 0.05mm (2cm depth), 0.07mm (3cm depth) and 0.14mm (6cm depth).

5.1 Automatic sphere detection

Before investigating the freehand system accuracy, it would be wise to ascertain the maximum attainable accuracy, given the 2D ultrasound machine, the tissue-mimicking phantom and the

Table 2: Variation of spatial calibration parameter values. 20 spatial calibrations were performed for each probe frequency and depth setting. The mean, root mean square error (rms) and 95% confidence intervals (±95%) are given for each of the estimated spatial calibration translations (t_x, t_y, t_z), rotations (α, β, γ) and scale factors (s_x, s_y).

Frequency (depth)		t_x	t_y (mm)	t_z	α	β (°)	γ	s_x (mm/pixel)	s_y
5-10MHz (6cm)	mean	122.16	-80.47	-2.43	-111.17	-6.14	-1.11	0.1383	0.1382
	rms	0.13	0.11	0.25	0.03	0.25	0.35	0.0005	0.0003
	±95%	±0.27	±0.21	±0.50	±0.05	±0.51	±0.71	±0.0010	±0.0007
5-10MHz (3cm)	mean	122.19	-80.04	-1.73	-111.32	-6.30	-0.88	0.0693	0.0699
	rms	0.14	0.18	1.06	0.06	0.39	0.64	0.0004	0.0004
	±95%	±0.28	±0.36	±2.13	±0.12	±0.77	±1.28	±0.0007	±0.0007
10-22MHz (2cm)	mean	100.69	-68.46	0.76	-107.52	-3.89	0.29	0.0461	0.0461
	rms	0.10	0.12	0.20	0.04	0.18	0.62	0.0003	0.0002
	±95%	±0.20	±0.23	±0.39	±0.08	±0.35	±1.25	±0.0007	±0.0003

sphere detection process described in Section 4.2. The ultrasound phantom was scanned by mounting the 5-10MHz probe in a rigid clamp attached to the movable section of a precision vice. With the B-scan plane parallel to the plane containing the array of spheres, the probe was moved in a direction orthogonal to this plane, and the movement of the vice jaw (and hence the probe) measured with a micrometer. B-scans were recorded every 0.002″ (0.0508mm) such that each sphere was spanned by approximately 40 B-scans.

The location of all the detectable spheres was measured and compared to the true geometric arrangement, as given in Figure 8. Results are shown in Table 1. The 3D confidence limit of < 0.35mm includes errors in the formation of each B-scan, the location of the spheres in the phantom, and the sphere detection process. The greater error in the z (out of plane) dimension is due to the beam width, and possibly also the construction of the phantom.

5.2 Spatial and temporal calibration precision

The precision of the temporal and spatial calibration procedures, described in Section 3, was analysed by comparing the results of repeated calibrations. Forty temporal calibrations were performed, with the same ultrasound machine settings, giving a root mean square deviation from the mean value of 4.9ms, and a 95% confidence limit of ±10.0ms. This is clearly much better than the period between ultrasound frames (40ms) or between position sensor readings (30ms). The precision is improved further by using the average of ten such calibrations: this was done before each of the spatial calibrations in the following experiments.

The precision of the spatial calibration was tested by repeating the process outlined in Section 3.2 twenty times for each probe frequency and depth setting. The position sensor remained attached to the probe throughout these experiments, but the probe was re-attached to the Cambridge phantom each time. The temperature of the water bath was held approximately constant by replacing some of the water between calibrations.

Table 2 shows the mean, root mean square variation and 95% confidence limits for the eight spatial calibration parameters. The translations (t_x, t_y, t_z) and rotations (α, β, γ) are defined in the coordinate system of the position sensor, *not* the B-scan plane. Due to the

Table 3: Variation in B-scan location due to the spatial calibration. The table shows the effect of the spatial calibration parameter variation in Table 2 on the location of the centre and corners of the B-scan. The mean 3D error and 95% confidence intervals are given in each case.

Frequency (depth)		**mean error (mm)** 3D	95% **confidence limits (mm)** x	y	z	3D
5-10MHz (6cm)	centre	0.44	±0.37	±0.18	±0.93	< 0.78
	top left	0.29	±0.27	±0.21	±0.50	< 0.50
	top right	0.48	±0.25	±0.30	±0.95	< 0.84
	bottom left	0.54	±0.52	±0.18	±1.08	< 0.94
	bottom right	0.67	±0.50	±0.34	±1.42	< 1.18
5-10MHz (3cm)	centre	1.03	±0.32	±0.30	±2.27	< 1.80
	top left	0.94	±0.28	±0.36	±2.13	< 1.65
	top right	1.01	±0.24	±0.35	±2.29	< 1.77
	bottom left	1.07	±0.48	±0.28	±2.35	< 1.87
	bottom right	1.14	±0.38	±0.36	±2.45	< 1.99
10-22MHz (2cm)	centre	0.21	±0.20	±0.14	±0.45	< 0.36
	top left	0.21	±0.20	±0.23	±0.39	< 0.37
	top right	0.24	±0.24	±0.15	±0.47	< 0.42
	bottom left	0.25	±0.22	±0.20	±0.53	< 0.44
	bottom right	0.26	±0.20	±0.18	±0.60	< 0.46

orientation of the position sensor on the probe, t_x and t_y were nevertheless approximately in the plane of the B-scan, and t_z orthogonal to it. α represented rotation about an axis approximately normal to the B-scan plane, β about an approximately vertical axis, and γ about an approximately horizontal axis. The scale factors (s_x, s_y) are defined within the B-scan plane.

It is immediately apparent from the results that the parameters in the plane of the B-scan (t_x, t_y, α, s_x, s_y) are much better determined than those out of the plane. In the worst case, the in-plane parameters are defined to within ±0.36mm, ±0.12° and ±0.001mm/pixel, whereas those out of the plane are only defined to within ±2.13mm and ±1.28°. This is a consequence of the ultrasound beam width, which, away from the foci, can exceed 4mm for the 10-22MHz probe and 8mm for the 5-10MHz probe. Use of the Cambridge phantom, as described earlier, limits the effect of beam width on the spatial calibration — which is why the out of plane accuracy is significantly better than the width of the ultrasound beam — but it does not suppress it entirely.

Of particular note in this respect is the combination of the out of plane translation t_z and rotation about a horizontal axis γ. These parameters can combine to generate movements of the B-scan plane which have only a small effect on the location of the calibration plane within the B-scan. Only with motion C in Figure 7 is the image of the plane sensitive to combined changes in these two parameters, which is why this motion is an essential part of the calibration protocol. This is a particular problem for B-scans which are shallow, but have large beam widths, which may explain the results for the 5-10MHz probe on a 3cm depth setting, where the out of plane parameters are particularly poorly determined compared with those in the B-scan plane.

Table 3 shows the effect of this parameter variation on the location of the corner points

Table 4: Measurement accuracy. Summary results are shown for each of the graphs in Figures 12, 13 and 14. Unbiased confidence limits are calculated given the number of spheres in each case. x, y and z are measured in the plane containing the phantom spheres, as for the graphical results. 2D confidence limits are given for panoramic scans and 3D confidence limits otherwise.

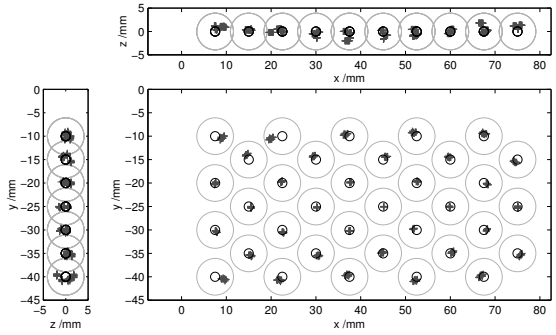
Frequency		**mean error (mm)**		95% **confidence limits (mm)**				
(depth)		2D	3D	x	y	z	2D	3D
5-10MHz	Fig. 12(a)	-	0.27	± 0.41	± 0.25	± 0.45	-	< 0.51
(6cm)	Fig. 12(b)	0.19	-	± 0.39	± 0.35	-	< 0.38	-
	Fig. 12(c)	0.24	-	± 0.40	± 0.38	-	< 0.49	-
	Fig. 12(d)	-	0.32	± 0.54	± 0.31	± 0.40	-	< 0.60
	Fig. 12(e)	-	0.35	± 0.64	± 0.26	± 0.42	-	< 0.65
5-10MHz	Fig. 13(a)	-	0.19	± 0.33	± 0.21	± 0.25	-	< 0.37
(3cm)	Fig. 13(b)	0.14	-	± 0.26	± 0.21	-	< 0.29	-
	Fig. 13(c)	0.21	-	± 0.41	± 0.27	-	< 0.42	-
	Fig. 13(d)	-	0.32	± 0.62	± 0.26	± 0.35	-	< 0.61
	Fig. 13(e)	-	0.33	± 0.64	± 0.21	± 0.45	-	< 0.62
10-22MHz	Fig. 14(a)	-	0.22	± 0.22	± 0.31	± 0.37	-	< 0.44
(2cm)	Fig. 14(b)	0.17	-	± 0.29	± 0.32	-	< 0.37	-
	Fig. 14(c)	0.22	-	± 0.43	± 0.36	-	< 0.49	-
	Fig. 14(d)	-	0.25	± 0.39	± 0.35	± 0.32	-	< 0.50
	Fig. 14(e)	-	0.22	± 0.27	± 0.36	± 0.30	-	< 0.46

and centre of the B-scan. The coordinate system in this table is aligned to the B-scan plane. As expected, the in-plane precision is significantly better than the out of plane (z) precision. The bottom right of the B-scan, being furthest from the position sensor, is also slightly less well constrained. Nevertheless, an (x, y, z) precision at the centre of the B-scan of (± 0.20, ± 0.14, ± 0.45) is achieved for the highest resolution case, giving a 3D confidence limit better than < 0.5mm.

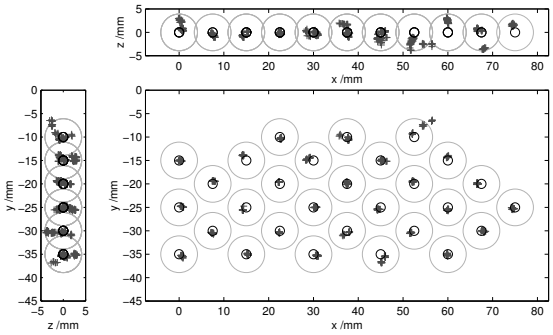
5.3 System accuracy

To assess the accuracy of the entire system for diagnostic ultrasound examinations, the phantom described in Section 4.1 was scanned in each of the scanning patterns of Figures 9(a) to (e). For each probe frequency and depth setting, the five sweeps were recorded after the first ten spatial calibrations. Sphere detection was then performed for all visible spheres in each of the five sweeps, with each of the 20 spatial calibration parameter sets.

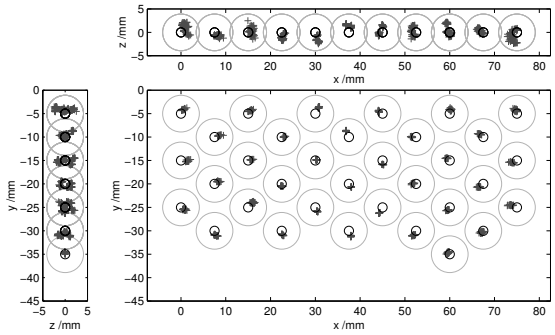
The results are presented in graphical form in Figures 12, 13 and 14, and summarised in Table 4. Each graph shows an orthographic projection of the distribution of (visible) sphere centres, so that the relative errors in each dimension, and across the depth of the B-scan, can be clearly seen. In order to make the location errors more visible, they have been magnified by a factor of four. The small bold circles (representing the spheres) are drawn to scale, whereas the grey circles are also magnified by a factor of four for comparison with the location errors. Note that the scanning patterns in (b) and (c) are panoramic, so we do not expect to be able to estimate the z location from these scans with a great degree of accuracy. The z projections are given in these graphs for completeness only, and 2D rather than 3D confidence limits are



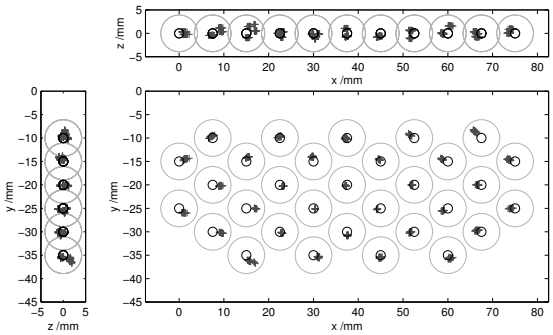
(a) translation



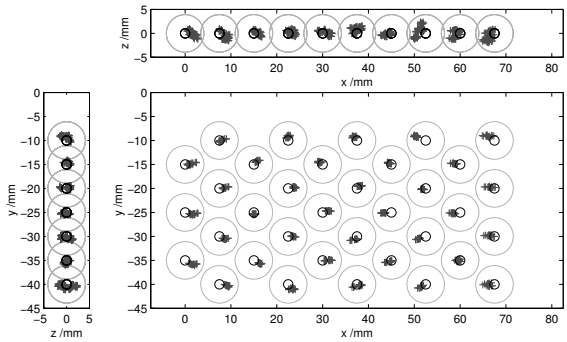
(b) panoramic translation



(c) panoramic rotation

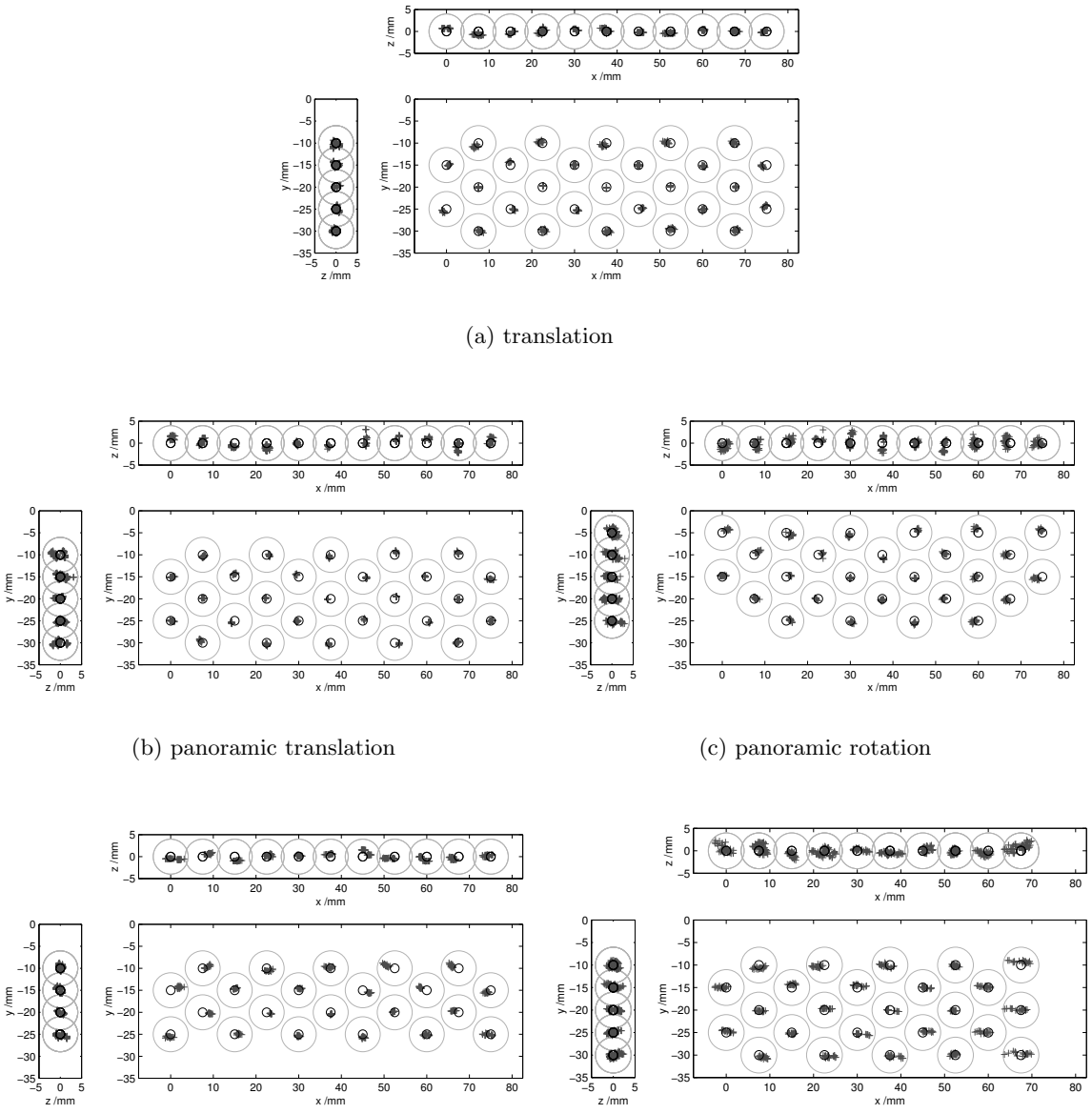


(d) horizontal rotation



(e) vertical rotation

Figure 12: System accuracy: 5-10MHz probe, 6cm depth. (a) to (e) show the distribution of sphere locations for each of the 20 spatial calibrations in Table 2 applied to each of the scanning patterns in Figures 10(a) to (e). Bold circles show the actual sphere location and size; the location errors and grey circles have been *magnified by a factor of four* to show the distribution more clearly.



(a) translation

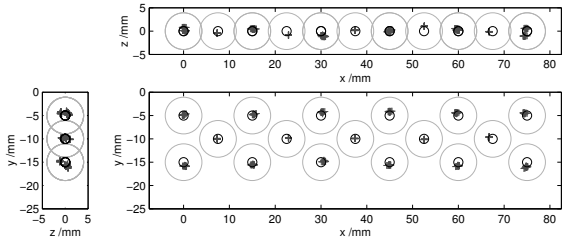
(b) panoramic translation

(c) panoramic rotation

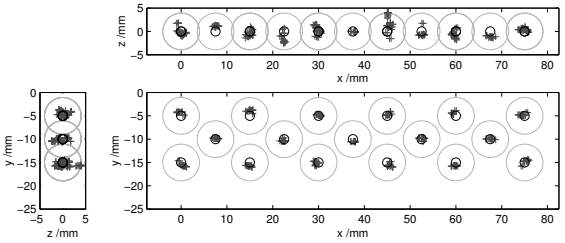
(d) horizontal rotation

(e) vertical rotation

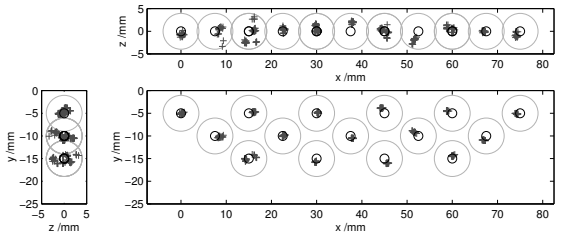
Figure 13: System accuracy: 5-10MHz probe, 3cm depth. (a) to (e) are as in Figure 12, and once again the location errors have been *magnified by a factor of four*.



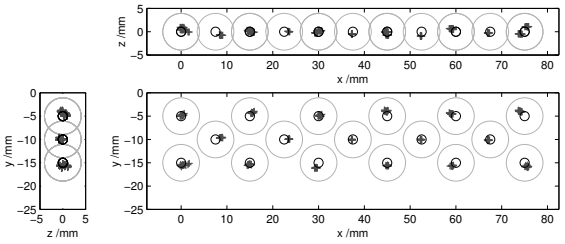
(a) translation



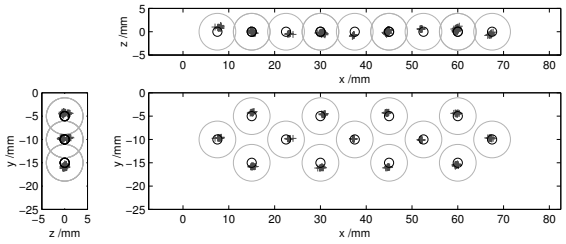
(b) panoramic translation



(c) panoramic rotation

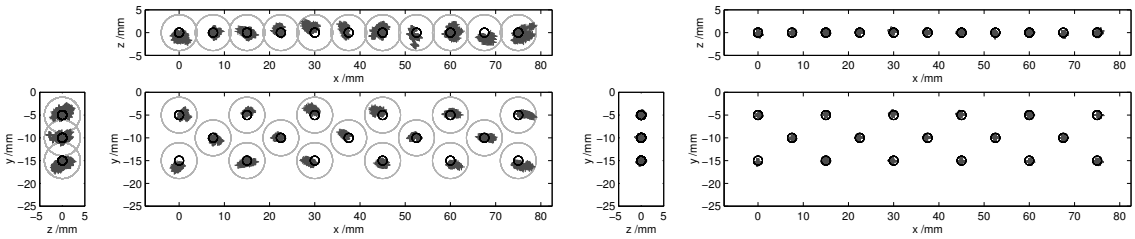


(d) horizontal rotation



(e) vertical rotation

Figure 14: System accuracy: 10-22MHz probe, 2cm depth. (a) to (e) are as in Figure 12, and once again the location errors have been *magnified by a factor of four*.



(a) Errors magnified by four

(b) Actual size

Figure 15: Probe mounting repeatability: 10-22MHz probe, 2cm depth. (a) and (b) show the distribution of location errors for 20 spatial calibrations as in Figure 2, each applied to ten scans using the pattern in Figure 10(f). The optical tracker was re-mounted between each scan.

Table 5: Probe mounting repeatability. Summary results are shown for the graphs in Figure 15.

Frequency (depth)		**mean error (mm)** 3D	95% **confidence limits (mm)** x	y	z	3D
10-22MHz (2cm)	Fig. 15	0.35	±0.53	±0.41	±0.53	< 0.69

quoted in these cases.

As expected, the worst case 3D confidence limit improves with increasing probe resolution and decreasing depth setting, so that with the 10-22MHz probe, the point location accuracy within one scan is < 0.5mm for any scanning pattern. This is similar to the precision of the spatial calibration, which implies that the spatial calibration is relatively unbiased. In all cases, scanning pattern (a) produced the most accurate results: in this orientation, the out of plane parameters of the spatial calibration have only a very limited effect on the geometry of the reconstructed data set. An error in t_z, for instance, would tend to move the entire data set bodily, which is irrelevant for this analysis. However, this is not the case for the rotations in scanning patterns (d) and (e). In a similar fashion, the panorama involving simple translation (b) generated more accurate results than that involving rotation (c).

In nearly all cases, the location errors were well within the radii of the spheres, particularly in the central region of the B-scan where the beam was focused. The variation of accuracy with probe settings is, however, somewhat less than might have been expected from the three-fold reduction in depth. This implies that the dominant errors are those *not* related to probe frequency and depth setting: possible culprits include the accuracy of the position sensor, and temporal effects in the acquisition path.

Table 6: Accuracy of depth change re-calibration. Results are as in Table 4, however the spatial calibrations used for the 3cm depth setting were derived from those calculated for the 6cm depth setting, and *vice versa*.

Frequency (depth)		**mean error (mm)** 2D	3D	95% **confidence limits (mm)** x	y	z	2D	3D
5-10MHz (6cm)	Fig. 10(a)	-	0.30	±0.47	±0.36	±0.45	-	< 0.56
	Fig. 10(b)	0.16	-	±0.29	±0.28	-	< 0.33	-
	Fig. 10(c)	0.19	-	±0.30	±0.34	-	< 0.38	-
	Fig. 10(d)	-	0.29	±0.47	±0.26	±0.41	-	< 0.55
	Fig. 10(e)	-	0.39	±0.71	±0.27	±0.57	-	< 0.72
5-10MHz (3cm)	Fig. 10(a)	-	0.21	±0.29	±0.26	±0.29	-	< 0.39
	Fig. 10(b)	0.18	-	±0.28	±0.31	-	< 0.37	-
	Fig. 10(c)	0.27	-	±0.50	±0.40	-	< 0.56	-
	Fig. 10(d)	-	0.37	±0.70	±0.37	±0.35	-	< 0.69
	Fig. 10(e)	-	0.32	±0.56	±0.32	±0.42	-	< 0.61

5.4 Repeatability of position sensor mounting

The scanning pattern of Figure 10(f) was used to assess the effect of re-attaching the position sensor to the probe mount shown in Figure 3. Ten scans were recorded using the highest resolution probe and lowest depth setting, removing and replacing the position sensor between each scan. These scans were then analysed as before, using each of the 20 spatial calibrations calculated for this probe frequency and depth setting.

Graphical results are shown in Figure 15 for the combination of 10 scans and 20 calibrations, and summarised in Table 5. The accuracy of the system is degraded only slightly by re-mounting the position sensor, from a 3D confidence limit of < 0.50mm to < 0.69mm.

5.5 Accuracy of depth change re-calibration

The system accuracy experiments in Section 5.3 were repeated for the 5-10MHz probe on a 3cm and 6cm depth setting, but in each case using the spatial calibration from the alternate depth setting, re-adjusted using the algorithm of Section 3.3. Table 6 shows the results in the same format as presented in Table 4. System accuracy is only slightly compromised by the use of this fast re-calibration technique. For the 6cm depth setting, the 3D confidence limit is degraded from < 0.65mm to < 0.72mm, and for the 3cm setting, from < 0.62mm to < 0.69mm.

5.6 Summary

Figure 16 shows a comparison of the results for the highest resolution system (10-22MHz probe and 2cm depth setting) and the other systems cited in Section 1.2. The distance measurement accuracy for the system presented in this paper is approximately $\sqrt{2}$ times the point location accuracy, since it is a measure of difference between two identical distributions.

As has been previously explained, the other systems' accuracies were assessed in slightly different circumstances and quoted in different forms. While every effort has been made to convert each of them to a comparable form, the values should be regarded as indications of

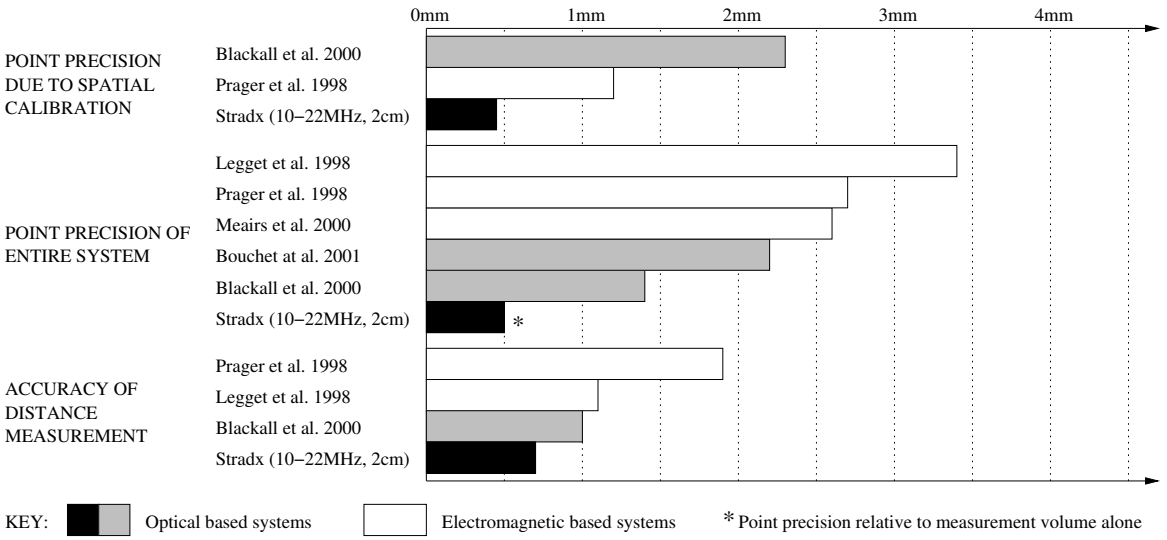


Figure 16: Comparison of the 10-22MHz Stradx system with other cited freehand 3D ultrasound systems. The bar chart shows the 3D confidence limits for various parameters of the systems. In most cases, this parameter has been estimated from alternative quoted values, as described in Section 1.2. As a result, the comparison must be treated with some caution.

system accuracy only.

6 Conclusions

Our system can be used to locate points within a freehand 3D data set to an accuracy of $<$ 0.50mm, using a 10-22MHz probe on a 2cm depth setting. The accuracy with which distances can be measured within a data set is approximately $\pm$0.7mm. This accuracy can be achieved by using the temporal and spatial calibrations outlined in this paper, and subsequently leaving the probe settings and position sensor mounting unchanged. It is valid for all practical freehand scanning patterns. The necessary calibration procedures can be performed in only 10 to 15 minutes from mounting the position sensor on the probe and connecting the PC to the ultrasound machine.

System accuracy for probes covering a wide range of frequencies and depth settings appropriate for musculoskeletal and arterial scans is within $<$ 0.65mm. Faster spatial calibrations are possible if only the depth setting or certain pan and zoom settings have changed, in which case the system accuracy is degraded only slightly, to $<$ 0.72mm. If spatial calibration is only performed each time the probe mount is attached to the probe (rather than each time the position sensor is attached to the mount), then the highest resolution system accuracy drops slightly from $<$ 0.5mm to $<$ 0.69mm.

The achievable resolution of this system is significantly better than that of any other such system presented in the literature. This is still the case if the position sensor is remounted or the spatial calibration recalculated based on depth changes alone: both of which are options that improve the system's usability. This performance owes much to the quality of the position sensor and the ultrasound machine, but also to the careful design and calibration

of the freehand acquisition system.

Acknowledgements

This work was carried out under EPSRC grant GR/N21062. Dynamic Imaging Ltd. provided a modified ultrasound machine to enable digital data acquisition.

References

[1] M. E. Anderson, M. S. McKeag, and G. E. Trahey. The impact of sound speed errors on medical ultrasound imaging. *Journal of the Acoustical Society of America*, 107(6):3540–3548, June 2000.

[2] J. E. Bailey, R. O. Bude, and T. Tuthill. US artifacts: effects on out-of-plane US images reconstructed from three-dimensional data sets. *Radiology*, 218:592–597, 2001.

[3] D. C. Barratt, A. H. Davies, A. D. Hughes, S. A. Thom, and K. N. Humphries. Optimisation and evaluation of an electromagnetic tracking device for high-accuracy three-dimensional ultrasound imaging of the carotid arteries. *Ultrasound in Medicine and Biology*, 27(7):957–968, 2001.

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

[5] M. Belohlavek, D. A. Foley, J. B. Seward, and J. F. Greenleaf. 3D echocardiography: Reconstruction algorithm and diagnostic performance of resulting images. *Proceedings of SPIE*, 2359:680–692, 1994.

[6] S. Berg, H. Torp, D. Martens, E. Steen, S. Samstad, I. Høivik, and B. Olstad. Dynamic three-dimensional freehand echocardiography using raw digital ultrasound data. *Ultrasound in Medicine and Biology*, 25(5):745–753, June 1999.

[7] W. Birkfellner, F. Watzinger, F. Wanschitz, G. Enislidis, C. Kollmann, D. Rafolt, R. Nowotny, R. Ewers, and H. Bergmann. Systematic distortions in magnetic position digitizers. *Medical Physics*, 25(11):2242–2248, November 1998.

[8] J. M. Blackall, D. Rueckert, Jr. C. R. Maurer, G. P. Penney, D. L. G. Hill, and D. J. Hawkes. An image registration approach to automated calibration for freehand 3D ultrasound. In *Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2000*, volume 1935 of *Lecture Notes in Computer Science*, pages 462–471, 2000.

[9] L. G. Bouchet, S. L. Meeks, G. Goodchild, F. J. Bova, J. M. Buatti, and W. A. Friedman. Calibration of three-dimensional ultrasound images for image-guided radiation therapy. *Physics in Medicine and Biology*, 46:559–577, 2001.

[10] H. N. Cardinal, J. D. Gill, and A. Fenster. Analysis of geometrical distortion and statistical variance in length, area, and volume in a linearly scanned 3-D ultrasound image. *IEEE Transactions on Medical Imaging*, 19(6):632–651, June 2000.

[11] J. Carr, R. W. Prager, G. M. Treece, A. H. Gee, M. Fynes, and L. Berman. Design of a clinical freehand 3D ultrasound system. In *Medical Imaging 2000 (Ultrasonic Imaging and Signal Processing)*, volume 3982 of *Proceedings of SPIE*, pages 14–25, San Diego, California, February 2000.

[12] M. L. Chuang, M. G. Hibberd, R. A. Beaudin, M. G. Mooney, M. F. Riley, J. T. Fearnside, and P. S. Douglas. Patient motion compensation during transthoracic 3-D echocardiography. *Ultrasound in Medicine and Biology*, 27(2):203–209, 2001.

[13] W. S. Edwards, C. Deforge, and Y. Kim. Interactive three-dimensional ultrasound using a programmable multimedia processor. *International Journal of Imaging Systems and Technology*, 9(6):442–454, 1998.

[14] A. Fenster, D. B. Downey, and H. N. Cardinal. Three-dimensional ultrasound imaging. *Physics in Medicine and Biology*, 46:R67–R99, 2001.

[15] B. D. Fornage, E. N. Atkinson, L. F. Nock, and P. H. Jones. US with extended field of view: phantom-tested accuracy of distance measurements. *Radiology*, 214(2):579–584, 2000.

[16] A. H. Gee, R. W. Prager, G. M. Treece, and L. Berman. Engineering a freehand 3D ultrasound system. *In Press for Pattern Recognition Letters*, 2002.

[17] A. H. Gee, R. W. Prager, G. M. Treece, and L. Berman. Narrow-band volume rendering for freehand 3D ultrasound. *Computers and Graphics*, 26(3):463–476, 2002.

[18] A. Goldstein. Errors in ultrasound digital image distance measurements. *Ultrasound in Medicine and Biology*, 26(7):1125–1132, 2000.

[19] J. M. Kofler Jr. and E. L. Madsen. Improved method for determining resolution zones in ultrasound phantoms with spherical simulated lesions. *Ultrasound in Medicine and Biology*, 27(12):1667–1676, 2001.

[20] M. E. Legget, D. F. Leotta, E. L. Bolson, J. A. McDonald, R. W. Martin, X.-N. Li, C. M. Otto, and F. H. Sheehan. System for quantitative three-dimensional echocardiography of the left ventricle based on a magnetic-field position and orientation sensing system. *IEEE Transactions on Biomedical Engineering*, 45(4):494–504, April 1998.

[21] E. L. Madsen, G. R. Frank, and F. Dong. Liquid or solid ultrasonically tissue-mimicking materials with very low scatter. *Ultrasound in Medicine and Biology*, 24(4):535–542, 1998.

[22] S. Meairs, J. Beyer, and M. Hennerici. Reconstruction and visualization of irregularly sampled three- and four-dimensional ultrasound data for cerebrovascular applications. *Ultrasound in Medicine and Biology*, 26(2):263–272, February 2000.

[23] 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.

[24] C. Palombo, M. Kozakova, C. Morizzo, F. Andreuccetti, A. Tondini, P. Palchetti, G. Mirra, G. Parenti, and N. G. Pandian. Ultrafast three-dimensional ultrasound — application to carotid artery imaging. *Stroke*, 29:1631–1637, August 1998.

[25] R. Prager, A. Gee, G. Treece, and L. Berman. Freehand 3D ultrasound without voxels: volume measurement and visualisation using the Stradx system. *Ultrasonics*, 40(1–8):109–115, 2002.

[26] R. W. Prager, A. H. Gee, and L. Berman. Stradx: real-time acquisition and visualisation of freehand 3D ultrasound. *Medical Image Analysis*, 3(2):129–140, 1999.

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

[28] R. N. Rohling, A. H. Gee, and L. Berman. A comparison of freehand three-dimensional ultrasound reconstruction techniques. *Medical Image Analysis*, 3(4):339–359, 1999.

[29] G. M. Treece, R. W. Prager, A. H. Gee, and L. Berman. Fast surface and volume estimation from non-parallel cross-sections, for freehand 3-D ultrasound. *Medical Image Analysis*, 3(2):141–173, 1999.

[30] G. M. Treece, R. W. Prager, A. H. Gee, and L. Berman. Correction of probe pressure artifacts in freehand 3D ultrasound. *In Press for Medical Image Analysis*, 2002.

[31] S. Wang, R. K. Chem, E. Cardinal, and K.-H. Cho. Joint sonography. *Radiologic Clinics of North America*, 37(4):653–668, July 1999.