

Stradx: real-time acquisition and visualization of freehand three-dimensional ultrasound

Richard W. Prager^{1*}, Andrew Gee¹ and Laurence Berman²

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

²Department of Radiology, Addenbrooke's Hospital, Cambridge CB2 2QQ, UK

Abstract

Conventional freehand three-dimensional (3-D) ultrasound is a multi-stage process. First, the clinician scans the area of interest. Next, the ultrasound data is used to construct a 3-D voxel array, which can then be visualized by, for example, any-plane slicing. The strict separation of data acquisition and visualization disturbs the interactive nature of the ultrasound examination. Furthermore, some systems require the clinician to wait for an unacceptable amount of time while the voxel array is constructed. In this paper, we describe a novel freehand 3-D ultrasound system which allows accurate acquisition of the raw data and immediate visualization of arbitrary slices through the data. Minimal processing separates the acquisition and visualization processes: in particular, at no stage is a voxel array constructed. Instead, the standard graphics hardware found inside most desktop computers is exploited to synthesize arbitrary slices directly from the raw B-scans.

Keywords: 3-D ultrasound imaging, any-plane slicing, panoramic imaging, temporal calibration

Received April 23, 1998; revised July 28, 1998; accepted September 2, 1998

1. INTRODUCTION

Conventional diagnostic ultrasound imaging is performed with a hand-held probe which transmits ultrasound pulses into the body and receives the echoes. The magnitude and timing of the echoes are used to create a two-dimensional (2-D) grey-scale image (B-scan) of a cross-section of the body in the scan plane.

Using a technique called freehand three-dimensional (3-D) ultrasound imaging (Rankin *et al.*, 1993; Steiner *et al.*, 1994; Deng *et al.*, 1996), it is possible to construct 3-D data sets from a series of 2-D B-scans (see Figure 1). A 3-D freehand examination can be broken into three stages: scanning, reconstruction and visualization. Before scanning, some sort of position sensor is attached to the probe. This is typically the receiver of an electromagnetic position sensor (Nelson and Elvins, 1993; Detmer *et al.*, 1994; Hughes *et al.*, 1996; Barry *et al.*, 1997; Leotta *et al.*, 1997), as illustrated

in Figure 1, although alternatives include acoustic spark gaps (King *et al.*, 1991), mechanical arms (Ohbuchi *et al.*, 1992) and optical sensors (State *et al.*, 1994; Trobaugh *et al.*, 1994). Measurements from the position sensor are used to determine the positions and orientations of the B-scans with respect to a fixed datum, usually the transmitter of the electromagnetic position sensor. In the next stage, the set of acquired B-scans and their relative positions are used to fill a regular voxel array. Finally, this voxel array is visualized using, for example, any-plane slicing, volume rendering or surface rendering (after segmentation).

Freehand systems can be used to obtain arbitrary volumes of data, since the motion of the probe is unconstrained. They are also cheap, requiring only existing, conventional ultrasound systems and relatively inexpensive additional components. For these reasons, research into freehand systems is very active (this paper's bibliography provides a representative survey), and several commercial systems have recently become available. To the best of the authors' knowledge, all these systems reconstruct the data on a regular voxel array before attempting any visualization.

*Corresponding author
(e-mail: rwp@eng.cam.ac.uk)

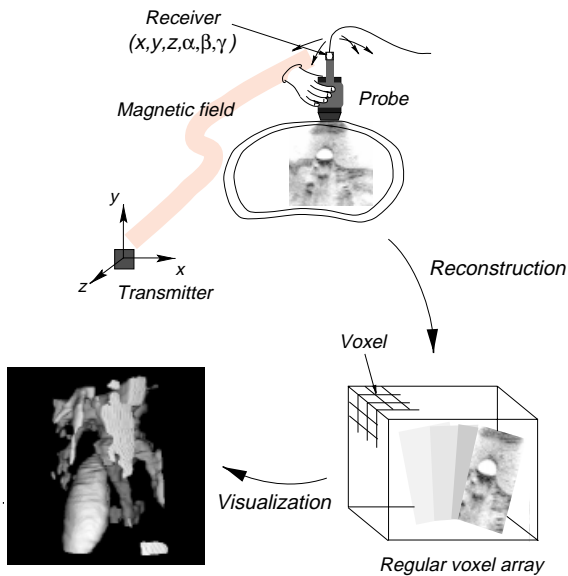


Figure 1. Conventional freehand 3-D ultrasound imaging. This is a multi-stage process, involving scanning, reconstruction and visualization. The figure illustrates an examination of a gall bladder.

There are, however, several problems with the conventional, voxel-based approach. While it is straightforward to construct a voxel array from regularly sampled data, such as the parallel slices produced by CT and MRI scanners, the same cannot be said for freehand 3-D ultrasound. Figure 2 illustrates some of the problems. Depending on the scanning pattern, voxels may be empty or intersected by one or more B-scans. Since empty voxels create artefacts in renderings and confuse automatic segmentation algorithms, they are usually filled-in using some sort of interpolation scheme. The better schemes (Ohbuchi *et al.*, 1992) take into account the finite width of the ultrasound beam (see Figure 3). Voxels intersected by multiple B-scans are often set using a simple average of the contributions from each B-scan: this has the beneficial effect of suppressing speckle noise (Nelson and Elvins, 1993; Moskalik *et al.*, 1995; Hernandez *et al.*, 1996; Hughes *et al.*, 1996; Barry *et al.*, 1997; Rohling *et al.*, 1997, 1998a). Alternatively, a maximum intensity approach can compensate for drop-out and shadowing in the B-scans (Moskalik *et al.*, 1995).

Whatever approach is used to construct the voxel array, it is clear that there are many parameters to set. What size should the voxels be? What should be done with those voxels intersected by multiple B-scans? What should be the size and shape of the interpolation kernel? Another problem is that a lot of filtering takes place during the construction of the voxel array: it is often said that slices through 3-D ultrasound data

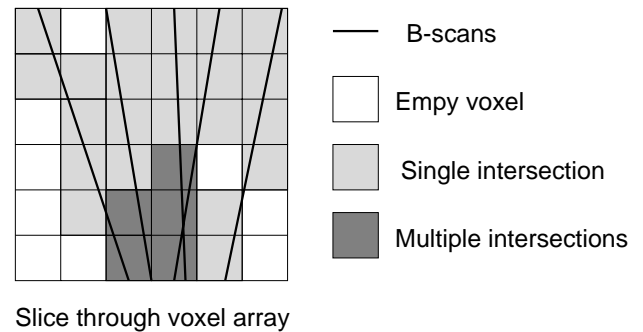


Figure 2. Constructing voxel arrays from freehand 3-D ultrasound data. Depending on the scanning pattern, voxels may be empty or intersected by any number of B-scans.

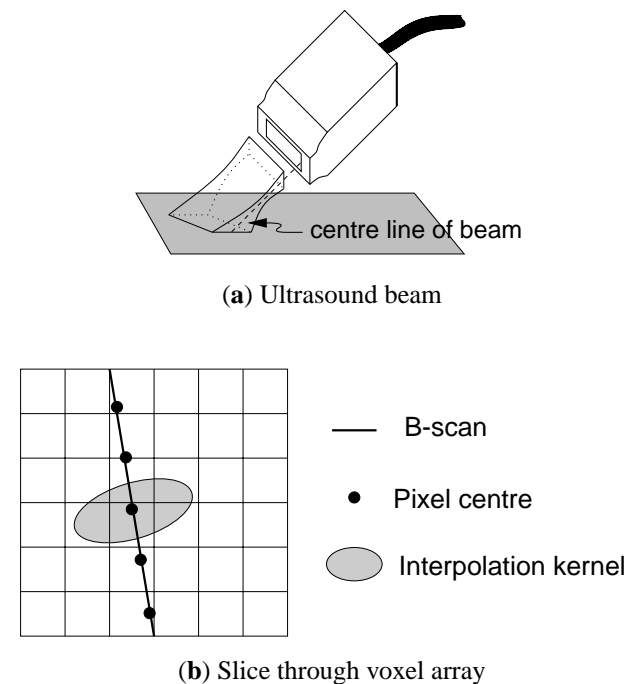


Figure 3. Interpolation for freehand 3-D ultrasound. The ultrasound beam (a) has a finite thickness, so each pixel in the B-scan samples a region of space which extends some distance normal to the B-scan. The interpolation kernel (b) should resemble this sampling function. Each B-scan pixel influences the contents of voxels lying within its interpolation kernel.

sets appear processed beyond recognition. The final problem is the time it takes to construct the voxel array: the better interpolation schemes (which make use of steerable kernels) are computationally expensive. There may be a delay of

several minutes after scanning before the clinician is able to visualize the data.

In this paper, we describe the Stradx^a freehand 3-D ultrasound system, which aims to avoid the pitfalls of a voxel-based approach. The Stradx system deals exclusively with raw 3-D ultrasound data, just B-scan images and positions. At no stage is a voxel array constructed. Great care is taken to label each B-scan with the most accurate position and orientation possible, as described in Section 2. In Section 3, we describe how visualization by panoramic imaging or any-plane slicing is possible immediately following, or even during, the examination, without the need to set any parameters. Finally, we draw some conclusions in Section 4.

2. ACQUIRING 3-D ULTRASOUND DATA

At the heart of the Stradx system lie two asynchronous processes. One process acquires B-scan images from the ultrasound machine as fast as possible (every 40 ms for a PAL video stream) and places them in a circular buffer in the main memory of the computer. The other performs a similar service with the position sensor readings. These are acquired every 30 ms or so, and placed in another circular buffer in the computer's main memory.

Since Stradx does not require a voxel representation of the data, it remains only to match the images with the positions. A simple approach might be to label the most recent image with the most recent position reading. However, greater accuracy can be achieved by interpolating the position between the readings just before and just after each image is acquired. In order to produce the real-time tools in the Stradx system, we would like to perform this temporal matching and interpolation on the fly, as the images and positions are coming in.

2.1. Matching images and positions

To match images and positions, we need to know the time that each image is grabbed, and the time each position reading is taken, as accurately as possible. The hardware we use (Silicon Graphics Indy and O2 workstations) provides a nanosecond time-stamp on each image acquired by the frame-grabbing card. We can record a similar time-stamp when we send a character to the position sensor to request a reading. We have found that there is an approximately constant time offset between the point just after sending the character, and the point when the sensor actually takes its

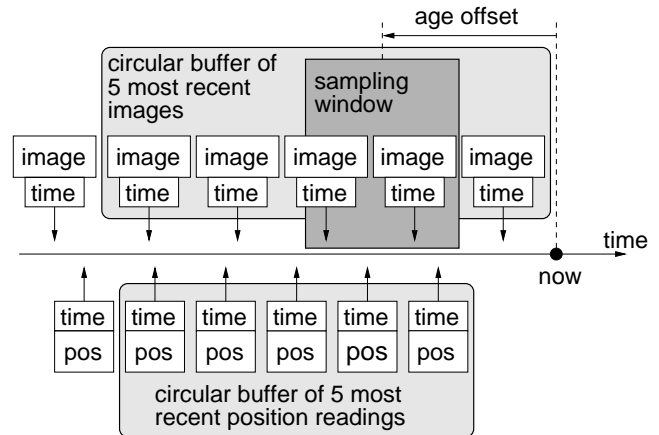


Figure 4. Matching images and positions. The five most recent images and positions are stored in circular buffers (for clarity, the circular buffers have been drawn unwrapped along a straight time line). By selecting images which lie within a certain age range, it is possible to find their positions accurately by linear interpolation between two position readings.

reading. This enables us to time-stamp the position readings as well. The constant time offset is determined using an automatic calibration procedure described in Subsection 2.2.

Now it just remains to linearly interpolate between the relevant pairs of readings to find the position and orientation of each B-scan. This is where we exploit the two circular buffers. Instead of picking the most recent image out of the circular buffer, we select the most recent image that lies within a specific age range (see Figure 4). In this way, there is a good chance that a position reading will be available on either side of it. The age range is calculated to span about one and a half times the image sampling period, so there will always be an eligible image available.

It is hard to predict the correct age offset of the image sampling window, since it depends on the latency of the position sensor. This parameter is therefore determined automatically using an adaptive algorithm. The system maintains a record of the time offset between the current image and the reading at the centre of the position circular buffer, averaged over the last 100 images. If more than five out of the last 100 images are unable to be matched with position readings, then this record is used to correct the age offset of the image sampling window.

Given a suitable age offset, the most recent image in the image sampling window is matched to one of the four possible intervals in the position circular buffer using a binary search. Once the precise position and orientation of the B-scan has been calculated by linear interpolation, the raw position data stream is discarded and the B-scan image, along

^aThe name 'Stradx' derives from an X-windows implementation of an earlier system developed under the UK Engineering and Physical Sciences Research Council's 'Stradivarius' project. The Stradx software, along with full documentation, can be downloaded from <http://svr-www.eng.cam.ac.uk/~rwp/stradx/>.

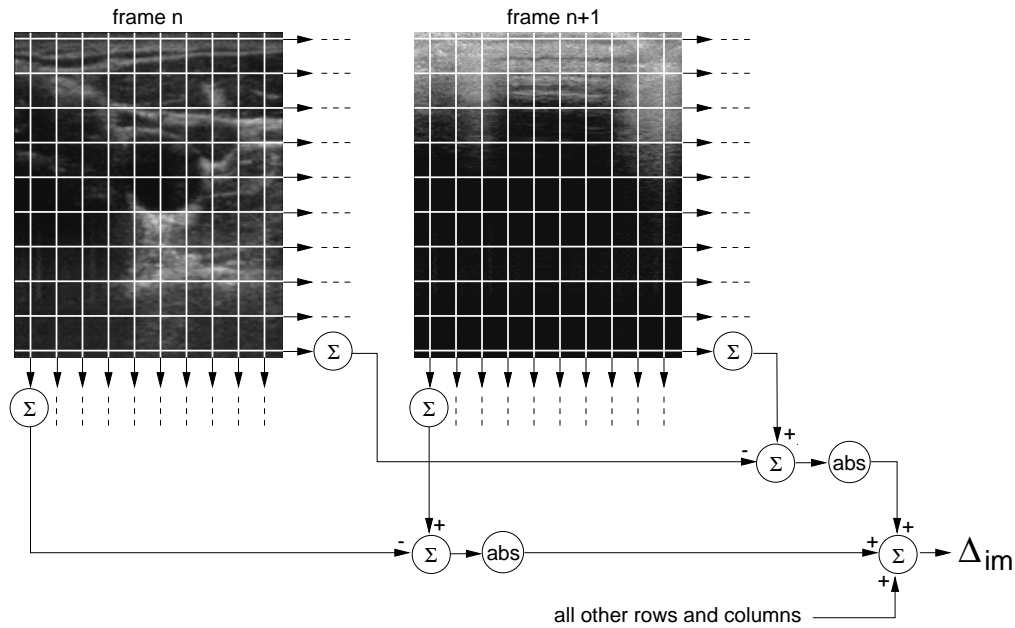


Figure 5. Detecting sudden changes in the image. The difference measure Δ_{im} is robust to speckle noise, quick to compute and sensitive to changes in all parts of the image, as long as the image contains some sort of large-scale structure.

with its position and orientation, is provided for use by the rest of the Stradx system. Care must be taken when interpolating between two orientation angles, since they wrap around at $\pm 180^\circ$.

2.2. Temporal and spatial calibration

There are two calibration processes which are central to the accuracy of the Stradx system. The first of these involves estimating the latencies in the position sensing device and the ultrasound machine. As explained in Subsection 2.1, images are automatically time-stamped by the frame-grabbing card. However, we do not know whether there are significant latencies in the ultrasound machine and the frame-grabbing hardware, which might cause the image to be time-stamped some time after it is actually acquired by the ultrasound probe. Similarly, there is an unknown latency in the position sensor readings: we take a time-stamp whenever we request a reading, although the delivered reading generally comes from a buffer inside the position sensing device and might already be fairly old. The temporal calibration procedure involves estimating the difference between the image and position latencies^b. This difference is then added to the position sensor

time-stamps, so that the image and position time-stamps can be meaningfully compared.

The temporal calibration is performed by applying a step input simultaneously to both the image and position streams. In practice, this is achieved by holding the probe still against the skin and then suddenly jerking it away, producing step changes in both the image content and position readings. Any discrepancy in the timings of the observed disturbances can be used to adjust the offset added to the position sensor time-stamps.

The motion of the probe is monitored by looking at the amount it has translated between successive position readings (the orientation angles are ignored). The calibration software asks the user to hold the probe steady for a period of 2.5 s, where ‘steady’ means that the probe should not move by more than 1 mm between successive position readings. When the user has steadied the probe for the required time, the system primes itself, asking the user to jerk the probe suddenly. The subsequent motion is accepted if the probe moves by more than 5 cm in 100 ms. The system records the time of the step, which is deemed to take place mid-way between the two position readings which first differed by > 1 mm^c.

^bSimilar temporal calibration procedures are employed in augmented reality systems, where user-input devices, tracking devices and imaging devices need to be registered temporally with the user’s view of the surroundings (Jacobs *et al.*, 1997).

^cThe various motion thresholds are compatible with the average user’s ability to hold a probe steady and then jerk it suddenly. The user is provided with essential feedback as to when the system considers the position sensor to be stationary and in motion.

While this is happening, the system is also examining the incoming image stream. It compares successive images and records a measure of their difference in a circular buffer. The difference measure must be robust to speckle noise, quick to compute and sensitive to changes in all parts of the image. To this end, Stradx sums the pixel intensities along ten evenly spaced rows and columns spanning the image. These sums are subtracted from their values in the previous frame, and the absolute values are accumulated to give the difference measure Δ_{im} (see Figure 5). Sudden changes in the image produce significant jumps in Δ_{im} .

The time of the step is estimated from the image stream as follows. Immediately following the jerk, the system looks in the circular buffer of difference measures and picks out those which were recorded while the probe was being held steady. From these, the system notes the maximum steady-state value of Δ_{im} . Care must be taken at this stage, since the image time-stamps cannot be reliably compared with the position time-stamps: there is an unknown latency which we are currently trying to estimate. So when deciding which images were acquired with a steady probe, the system allows for some predetermined maximum latency, which Stradx generously considers to be 0.5 s. After the calibration process is primed (again, allowing for up to 0.5 s latency), the system looks for the first Δ_{im} that exceeds the maximum steady-state value by 10%. The step is deemed to take place mid-way between the two images that produced this increased Δ_{im} .

The constant offset added to the position sensor time-stamps can now be adjusted by looking at the delay separating the image and position disturbances. Note, however, that if images are acquired every t seconds, and positions every T seconds, then the offset can be estimated only to an accuracy of $\pm(T + t)/2$ seconds. For this reason, it is advisable to perform several temporal calibrations to establish the range of results, and then accept a value that lies at the centre of this range.

The second calibration process is a spatial one. The position sensor indicates the position of the small receiver, mounted on the probe, with respect to the fixed transmitter. What we really want to know is the position of each B-scan pixel with respect to the transmitter. The calibration process involves deducing the six degree-of-freedom rigid-body transformation between the position sensor receiver and the corner of the B-scan plane, and also the x and y dimensions of the pixels (in mm/pixel). This is a fairly complex process which is usually performed by scanning and reconstructing some known object, and then using the discrepancy between the reconstructed shape and the known shape to re-estimate the calibration parameters (Detmer *et al.*, 1994; State *et al.*, 1994; Leotta *et al.*, 1995, 1997, Barry *et al.*, 1997). Stradx uses a novel, state-of-the-art spatial calibration

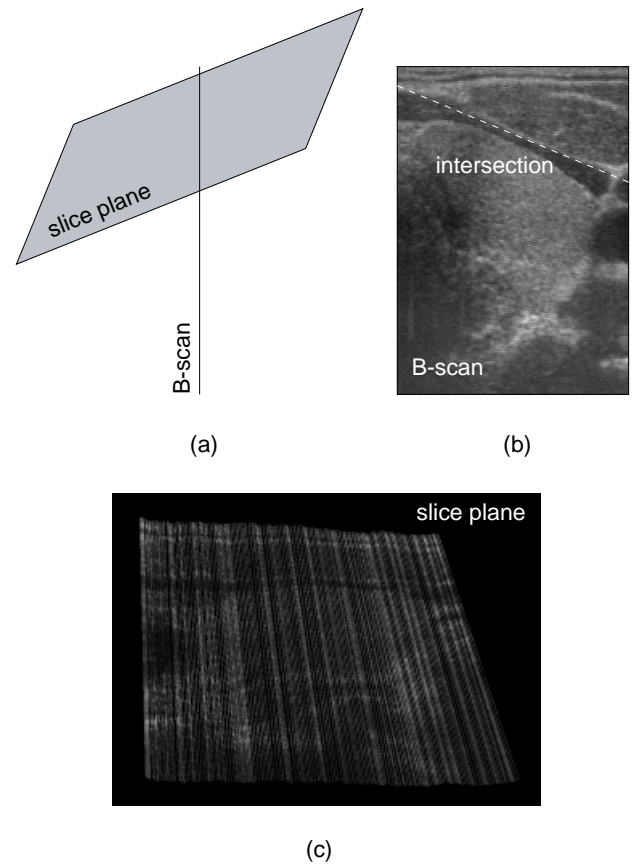


Figure 6. Reslicing thin B-scans. The slice plane intersects each B-scan along a line, as shown in (a) and (b). Intensities can be extracted from the B-scans along these lines and drawn on the slice plane. However, this produces an unsatisfactory, fragmented image, as shown in (c).

technique which we describe in full elsewhere (Prager *et al.*, 1998).

3. VISUALIZING 3-D ULTRASOUND DATA

Stradx provides two visualization tools, both of which are available immediately after the data has been acquired, without having to wait for a voxel array to be constructed. In fact, both visualization tools can be used even at the same time as the data is being acquired. The first tool, any-plane slicing, displays the ultrasound data on any plane selected by the user through the 3-D data set. The second tool, panoramic imaging, allows many B-scans to be stitched together to form a seamless, wide-angle composite.

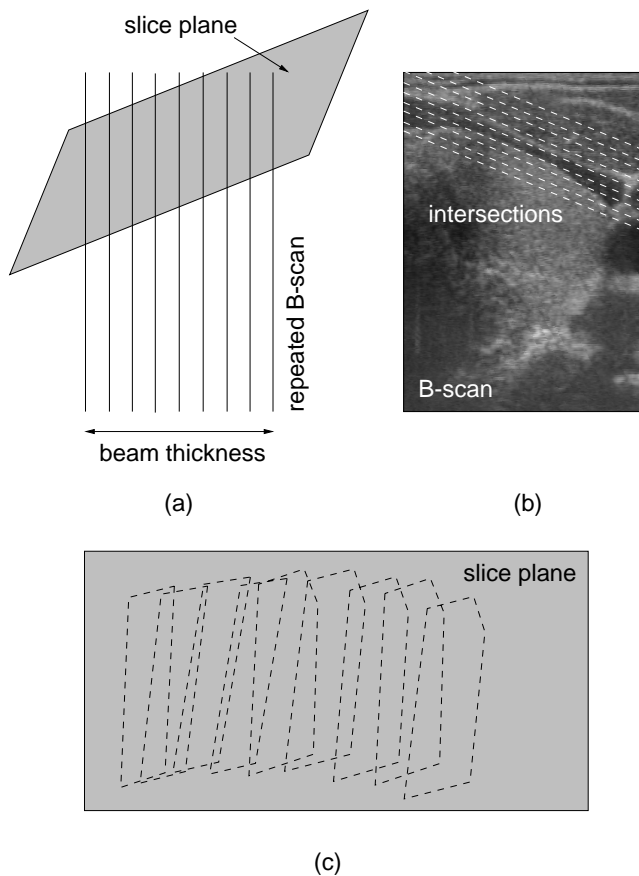


Figure 7. Reslicing thick B-scans. In practice, the ultrasound beam is not well focused out of the plane of the B-scan. In effect, the B-scan can be imagined repeated in space over the extent of the beam thickness, as shown in (a). The intersection of each B-scan with the slice plane is now a polygon, as shown in (b). Intensities can be extracted from within the polygons and drawn on the slice plane. This typically produces a set of filled, overlapping polygons, as shown in (c).

3.1. Any-plane slicing

Any-plane slicing of regular voxel arrays is straightforward and fast. The slice plane is selected interactively, and then the voxels which straddle the slice plane are rapidly picked out of the array (using integer arithmetic) and sent to the display. Any-plane slicing of scattered B-scan data is not so straightforward. A naive reslicing algorithm might extract intensity values along the lines of intersection of each B-scan with the slice plane, and paint these intensities onto the slice plane (see Figure 6). The problem with this simple approach is that the resulting slice comprises a set of fragmented line segments, as shown in Figure 6c.

To improve the quality of the slice we clearly need to interpolate between the line segments. Figure 7 illustrates an

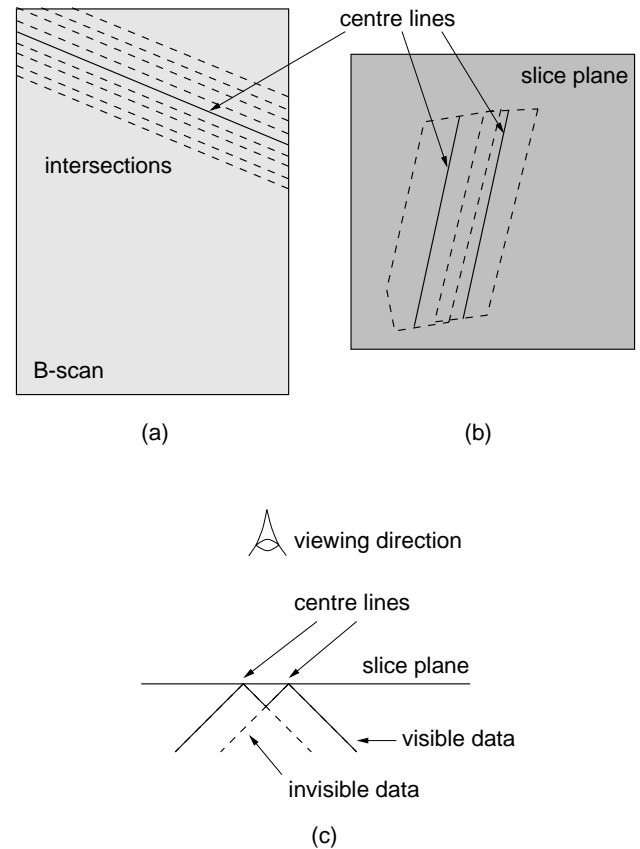


Figure 8. Correct treatment of overlapping polygons. Where polygons overlap, the slice should show data arising from the centre of the ultrasound beam, and not from the edges. This can be accomplished by 'folding' each polygon around its centre line to form a wedge, such that the centre line is close to the viewer and the edges further away. If the graphics system is instructed to render the collection of wedges with hidden surface removal, then the slice will reveal only those portions of the polygons that are as close as possible to their centre lines.

interpolation scheme which takes into account the anisotropic resolution of the 3-D ultrasound data. Since the ultrasound beam has a finite thickness, perhaps as much as 10 mm for low-frequency probes (Skolnick, 1991), we should think not of a single, thin B-scan in space, but a continuum of repeated B-scans smeared across the width of the ultrasound beam^d. The intersection of the slice plane with the smeared B-scan is now a polygon. Intensity values can be extracted within this polygon and drawn at the appropriate position on the slice plane. Repeating this process for all B-scans tiles the slice

^dFor simplicity, we assume that the width of the ultrasound beam is constant, although in reality the width varies with distance from the probe face (Skolnick, 1991) (see Figure 3a).

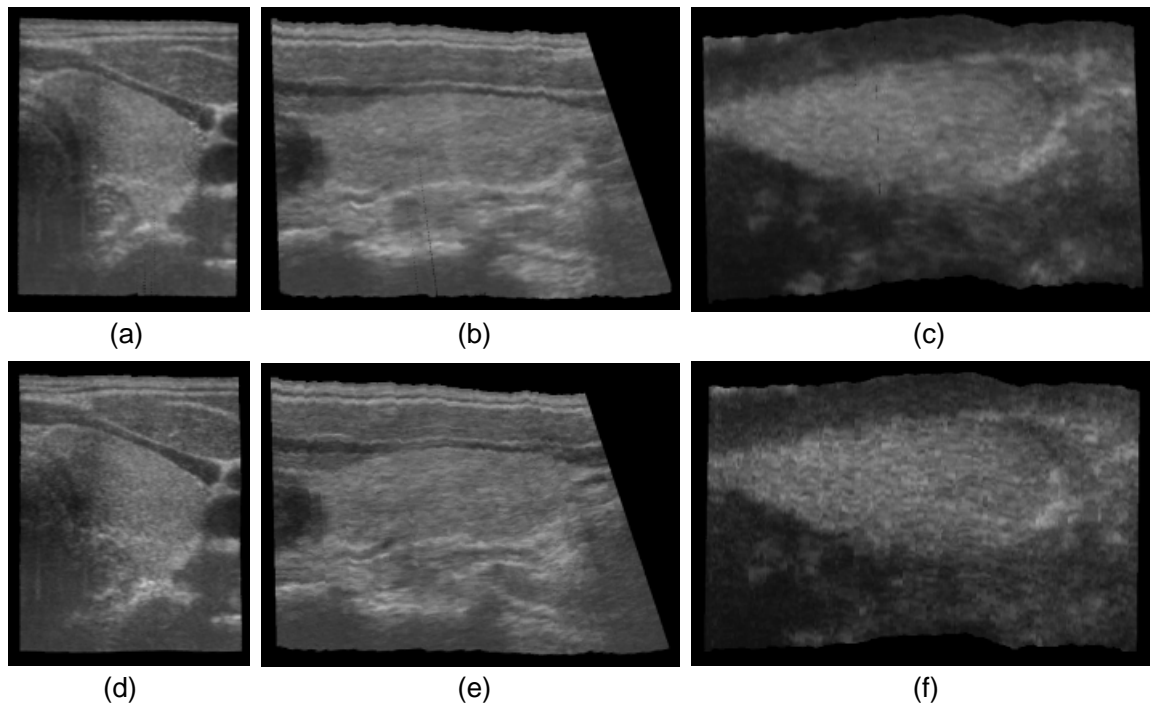


Figure 9. Any-plane slicing of a human thyroid gland. The figure shows results of a voxel-based (a)–(c) and voxel-less (d)–(f) approach. Slices (a) and (d) are virtually parallel to the B-scans. Slices (b) and (e) are orthogonal to these, but the skin is still at the top: similar B-scans could have been obtained by rotating the probe through 90° . Slices (c) and (f) are located a constant depth below the skin, and could not have been obtained using conventional 2-D ultrasound.

plane with a set of overlapping polygons filled with intensity data.

It is only left to decide which intensity to display where two or more polygons overlap on the slice plane. Intuitively, we would like to see data gathered along the centre line of the ultrasound beam, and not the periphery (see Figure 3a). This can be achieved by mapping the intensity data not onto a flat polygon in the slice plane, but a wedge, such that the orthographic projection of the wedge onto the slice plane is the same as the original polygon (see Figure 8). To an observer viewing the slice plane from above, a single wedge looks exactly the same as the flat polygon. However, intensities associated with the centre of the ultrasound beam are rendered close to the viewer, while those from the extremities of the beam are further away. If the graphics system is instructed to render the wedges using orthographic projection with hidden surface removal, then the slice will reveal only those portions of the polygons that are as close as possible to their centre lines.

Note that the voxel-less reslicing technique is effectively free of parameters. We say ‘effectively’, since it is necessary to specify the beam thickness. However, the appearance of

the reslice image is entirely insensitive to this parameter, provided it is set high enough to fill all the gaps in the slice plane. In practice, we tend to acquire B-scans separated by about 0.5 mm, so we just set the thickness parameter to 0.5 mm and leave it there.

Central to the success of this approach is the need to develop a fast implementation. To this end, we exploit standard graphics accelerator hardware, as found in many of today’s desktop computers. The wedges can be rendered using one of two techniques. The first, 2-D texture mapping, defines each wedge as two polygons and projects the appropriate grey-level pattern from the B-scan onto the polygons. This is the preferred technique with hardware that offers flexible, accelerated texture mapping. Alternatively, we can extract adjacent rows of pixels from the B-scans, as in Figure 8a, and place each pixel at its correct position on the wedge. These points can then be rendered as the vertices of a dense mesh of triangles, using Gouraud shading to fill in the gaps.

Given suitable (but not necessarily extravagant) graphics hardware, the slice can be rendered very rapidly. For instance, using the Gouraud technique, several hundred B-scans can be resliced on a Silicon Graphics O2 workstation in

about 1 s. Moreover, once the slice plane has been defined, the slice can be displayed in real-time as the ultrasound data is acquired, providing a powerful interactive tool for the clinician. This is possible because the interpolation strategy requires information from only one B-scan at a time, in any order.

Orthogonal slices rendered using the Gouraud scheme are shown in Figure 9d–f. For comparison, Figure 9a–c show similar slices through a regular voxel array constructed from the same data. The voxels were cubic, with the same dimension as a pixel in the original B-scans. Gaps in the voxel array were filled by averaging intensities in a $5 \times 5 \times 5$ local neighbourhood. Even with this simple interpolation scheme, the voxel array took a few minutes to construct on a powerful workstation. Close inspection of the two sets of slices shows that the voxel-less scheme can produce grey-scale texture which is more faithful to the texture in the original B-scans, since there is no averaging of the raw intensity data. Another example of a voxel-less reslice can be found in Figure 10, while a quantitative evaluation of the various reconstruction and reslicing algorithms is presented in Rohling *et al.* (1998b).

3.2. Panoramic ultrasound

The field of view of standard B-scans is very limited. Stradx offers a panoramic imaging facility which extends this field of view by any amount, allowing long structures to be displayed in a single image. Raw data for a panoramic image is acquired by translating the probe in the plane of the B-scans (see Figure 11). The acquired B-scans are then accurately registered and stitched together to produce a seamless composite. This can be performed in real-time, so the clinician sees the panorama being constructed as the probe is moved over the region of interest.

Recent interest in this type of imaging has been triggered by the release of the Siemens Sonoline Elegra ultrasound machine, which offers a panoramic imaging facility called ‘SieScape’. The Siemens machine achieves accurate registration by correlating consecutive images in real-time. This requires special-purpose, dedicated hardware that contributes to the high cost of the machine. We achieve a similar result using information provided by the position sensing device. Consider B-scans n and $n + 1$. Since we know the 3-D position and orientation of both scans, it is straightforward to calculate the rigid-body transformation \mathbf{T} between the local coordinate systems of the two images. We can then map pixels in B-scan $n + 1$ into B-scan n ’s coordinate system, and display both images on the same axes.

There is one slight complication. Unless the probe is moved perfectly in the plane of the B-scans, \mathbf{T} will involve not only a within-plane translation and rotation, but also

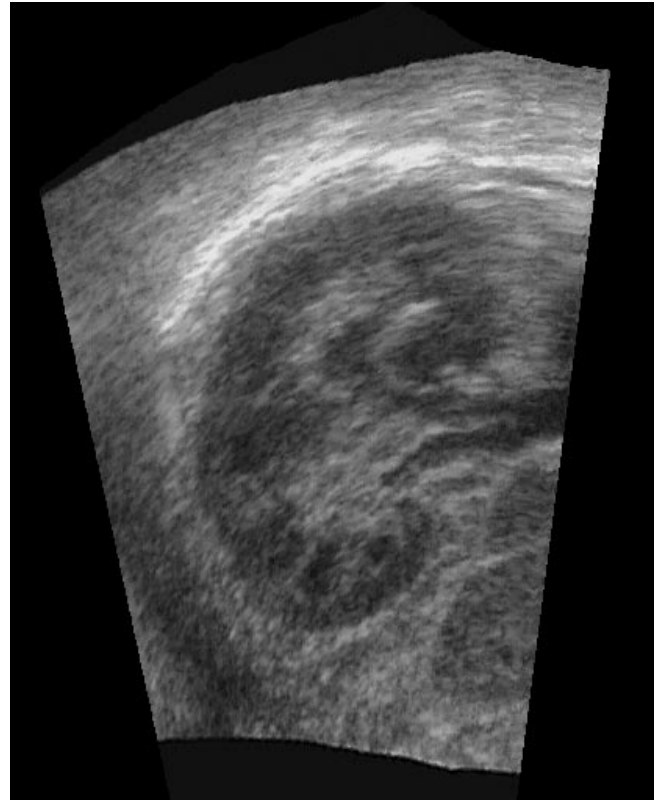


Figure 10. Any-plane slicing of a human kidney. The figure shows a coronal view of a kidney, generated from B-scans acquired in the longitudinal direction. This textbook view could not have been obtained using conventional 2-D ultrasound. The original B-scans were acquired at the full video frame rate using a fan-shaped scanning pattern. The B-scans were self-intersecting, though this is not a problem for the reslicing algorithm.

an out-of-plane transformation. We need to ‘flatten’ this transformation to warp the surface traced out by the B-scans onto a flat plane. This can be achieved as follows. We use \mathbf{T} to transform the centre line of B-scan $n + 1$ into B-scan n ’s coordinate system. The centre and endpoints of the centre line will not generally have zero z -coordinates in the new coordinate system. We flatten the transformation by setting these coordinates to zero. We then calculate the angle θ and the displacement $(\delta x, \delta y)$, such that a within-plane rotation θ followed by a translation $(\delta x, \delta y)$ maps the centre of B-scan $n + 1$ onto its projected position in B-scan n ’s coordinate system (see Figure 12). We can easily construct a planar Euclidean transformation matrix \mathbf{T}' from θ and $(\delta x, \delta y)$. The panoramic image is then constructed using \mathbf{T}' to map pixels in B-scan $n + 1$ into B-scan n ’s coordinate system.

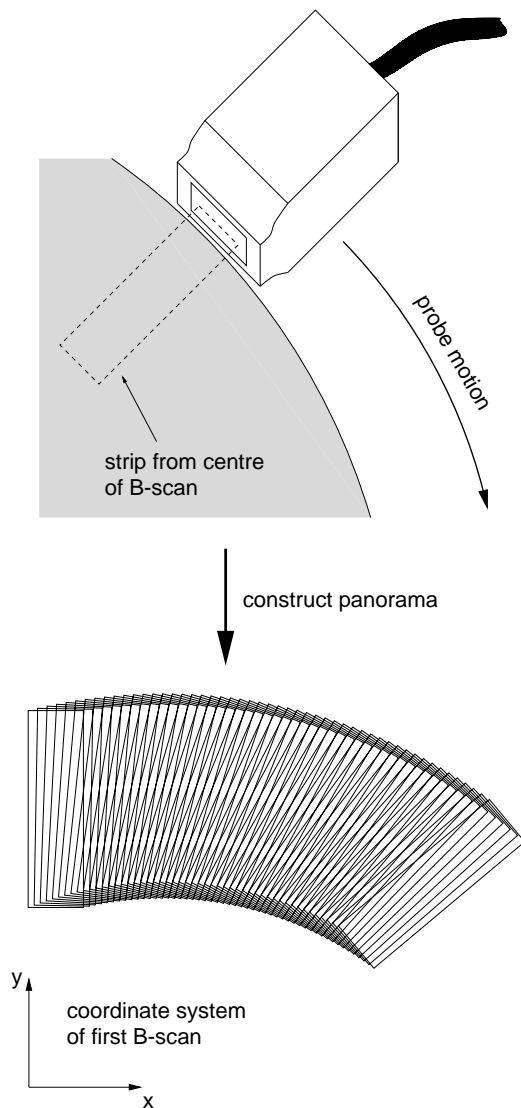


Figure 11. Panoramic ultrasound. The probe is moved roughly in the plane of the B-scans over the region of interest. Thin strips from the centre of the B-scans are used to construct a panoramic image in real-time.

Of course, we do not need to render every B-scan in its entirety to construct the panoramic image. Consecutive B-scans will share a large overlap region, and it suffices to use only a narrow strip from the centre of each B-scan. The panorama is constructed in the coordinate system of the first B-scan. \mathbf{T}' is calculated for each pair of consecutive B-scans, and the frame-to-frame transformations are concatenated to refer the current B-scan back to the coordinate system of the first B-scan.

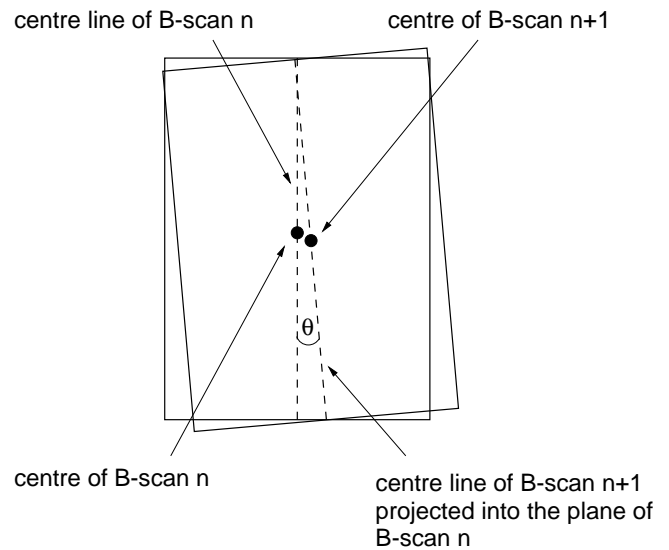


Figure 12. Registering consecutive B-scans. The position sensor readings are used to transfer the centre line of B-scan $n+1$ to B-scan n 's coordinate system. In general, the centre line will lie outside the plane of B-scan n , so its centre and endpoints are projected onto the plane by simply setting their z -coordinates to zero. It is then straightforward to measure the angle θ and deduce the translation $(\delta x, \delta y)$, such that a within-plane rotation θ followed by a translation $(\delta x, \delta y)$ maps the centre of B-scan $n+1$ onto its projected position in B-scan n 's coordinate system.

Where adjacent strips overlap, we would like to see pixels arising from the centre of the strips and not the edges. This can be achieved using the wedge-based technique developed for any-plane slicing (see Figure 8). Each wedge is folded about its centre line before being painted onto the screen: hidden surface removal sees to the rest.

Note that this technique is effectively free of parameters. It is necessary to specify the width of each strip, but this can be determined automatically, using \mathbf{T}' to calculate the minimum width such that there are no gaps between adjacent strips. It is the wedge-based technique that really determines the parts of each strip that are displayed, so using slightly too large a strip width only affects the rendering speed, not the image quality.

Using the Gouraud technique, the panorama is easily rendered on a Silicon Graphics Indy workstation at the video frame rate (25 strips per second), at the same time as the clinician guides the probe over the area of interest. Typical panoramic images constructed using this technique can be found in Figure 13. The figure illustrates some of the problems of scanning non-rigid structures. The *in vivo* data acquisition was not gated to an ECG, so the pulsatile motion of the femoral artery results in small ripples in the panorama.

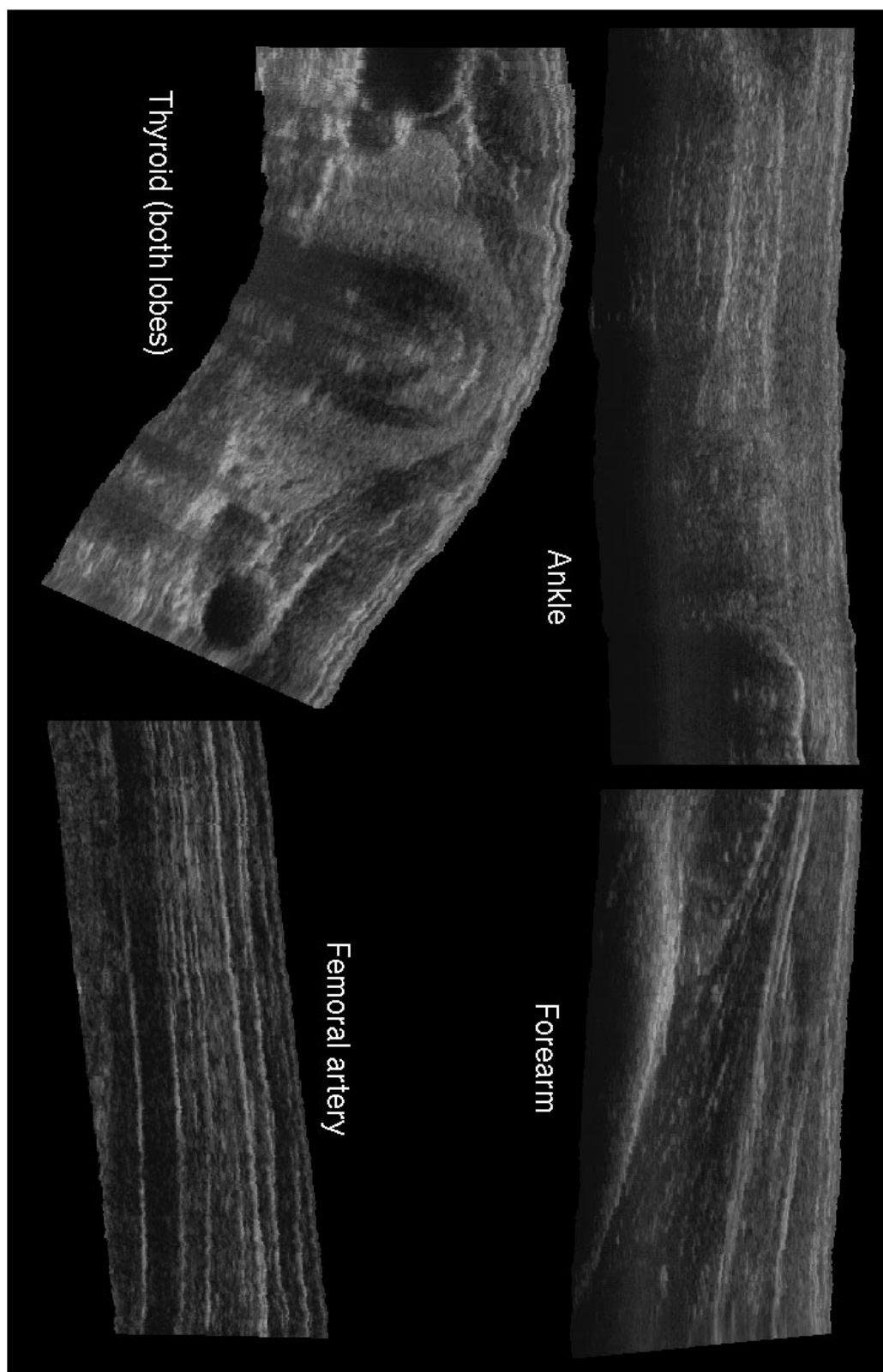


Figure 13. Typical real-time panoramic images.

Other reconstruction errors are caused by variation in the contact pressure between the probe and the skin (this is particularly evident in the irregular skin line at the top of the thyroid image), refraction of the ultrasound beam, inaccuracy of the position sensor^e and changes in the speed of sound as the pulse travels through different kinds of tissue. Despite these limitations, fairly clear and consistent panoramas are obtained.

4. CONCLUSIONS

We have described a freehand 3-D ultrasound system which allows immediate visualization of the data without constructing a voxel array. Great care is taken to determine the positions and orientations of the B-scans as accurately as possible. Any-plane slicing and panoramic imaging are accomplished with minimal processing of the raw data, with the aim of preserving the fine detail present in the B-scans.

In the future, we aim to extend the system to handle colour Doppler data as well as grey-scale images. We are also developing novel, semi-automatic techniques for segmenting the B-scans and reconstructing the 3-D shape of the segmented object from the resulting multi-planar contours (Treece *et al.*, 1998).

ACKNOWLEDGEMENTS

Stradx was developed from an earlier data acquisition system written by Patrick Gosling. Figure 1 was provided by Jonathan Carr. The reslices through the voxel array in Figure 9a–c were produced by Robert Rohling using the 3DViewnix visualization package. Finally, thanks are due to the anonymous referees, whose constructive suggestions helped to improve the original version of this paper.

REFERENCES

- Barry, C. D., Allot, C. P., John, N. W., Mellor, P. M., Arundel, P. A., Thomson, D. S. and Waterton, J. C. (1997) Three-dimensional freehand ultrasound: image reconstruction and volume analysis. *Ultrasound Med. Biol.*, 23, 1209–1224.
- Deng, J., Gardener, J. E., Rodeck, C. H. and Lees, W. R. (1996) Fetal echocardiography in 3-dimensions and 4-dimensions. *Ultrasound Med. Biol.*, 22, 979–986.
- Detmer, P. R., Bashein, G., Hodges, T., Beach, K. W., Filer, E. P., Burns, D. H. and Strandness Jr, D. E. (1994) 3D ultrasonic image feature localization based on magnetic scanhead tracking: *in vitro* calibration and validation. *Ultrasound Med. Biol.*, 20, 923–936.
- Hernandez, A., Basset, O., Chirossel, P. and Gimenez, G. (1996) Spatial compounding in ultrasonic imaging using an articulated scan arm. *Ultrasound Med. Biol.*, 22, 229–238.
- Hughes, S. W., D'Arcy, T. J., Maxwell, D. J., Chiu, W., Milner, A., Saunders, J. E. and Sheppard, R. J. (1996) Volume estimation from multiplanar 2D ultrasound images using a remote electromagnetic position and orientation sensor. *Ultrasound Med. Biol.*, 22, 561–572.
- Jacobs, M. C., Livingston, M. A. and State, A. (1997) Managing latency in complex augmented reality systems. In *Proc. 1997 Symp. on Interactive 3D Graphics*, Providence, RI, pp. 49–54.
- King, D. L., King Jr., D. L. and Shao, M. Y. (1991) Evaluation of *in vitro* measurement accuracy of a three-dimensional ultrasound scanner. *J. Ultrasound Med.*, 10, 77–82.
- Leotta, D. F., Detmer, P. R., Gilja, O. H., Jong, J. M., Martin, R. W., Primozich, J. F., Beach, K. W. and Strandness, D. E. (1995) Three-dimensional ultrasound imaging using multiple magnetic tracking systems and miniature magnetic sensors. In *Proc. IEEE Ultrasonics Symp.*, pp. 1415–1418.
- Leotta, D. F., Detmer, P. R. and Martin, R. W. (1997) Performance of a miniature magnetic position sensor for three-dimensional ultrasound imaging. *Ultrasound Med. Biol.*, 24, 597–609.
- Moskalik, A., Carson, P. L., Meyer, C. R., Fowlkes, J. B., Rubin, J. M. and Roubidoux, M. A. (1995) Registration of three-dimensional compound ultrasound scans of the breast for refraction and motion correction. *Ultrasound Med. Biol.*, 21, 769–778.
- Nelson, T. R. and Elvins, T. T. (1993) Visualization of 3D ultrasound data. *IEEE Comput. Graphics Applic.*, 50–57.
- Ohbuchi, R., Chen, D. and Fuchs, H. (1992) Incremental volume reconstruction and rendering for 3D ultrasound imaging. In Robb, R. A. (ed.) *Proc. Visualization in Biomedical Computing, SPIE*, 1808, 312–323. International Society of Optical Engineering, Bellingham, WA.
- Prager, R. W., Rohling, R. N., Gee, A. H. and Berman, L. (1998) Rapid calibration for 3-D freehand ultrasound. *Ultrasound Med. Biol.*, 24, 855–869.
- Rankin, R. N., Fenster, A., Downey, D. B., Munk, P. L., Levin, M. F. and Vellet, A. D. (1993) Three-dimensional sonographic reconstruction: techniques and diagnostic applications. *Am. J. Roentgenol.*, 161, 695–702.
- Rohling, R. N., Gee, A. H. and Berman, L. (1997) Three-dimensional spatial compounding of ultrasound images. *Med. Image Anal.*, 1, 177–193.
- Rohling, R. N., Gee, A. H. and Berman, L. (1998a) Automatic registration of 3-D ultrasound images. *Ultrasound Med. Biol.*, 24, 841–854.
- Rohling, R. N., Gee, A. H. and Berman, L. (1998b) *Radial Basis Function Interpolation for 3-D Ultrasound*. Technical Report CUED/F-INFENG/TR 327, Cambridge University Department of Engineering.
- Skolnick, M. L. (1991) Estimation of ultrasound beam width in the elevation (section thickness) plane. *Radiology*, 180, 286–288.

^eThis is particularly problematic with electromagnetic position sensors. Great care needs to be taken to keep the sensor away from metallic objects, which can severely affect the accuracy of the position readings.

- State, A., Chen, D. T., Tector, C., Brandt, A., Chen, H., Ohbuchi, R., Bajura, M. and Fuchs, H. (1994) Case study: observing a volume rendered fetus within a pregnant patient. In *Proc. IEEE Visualization*, pp. 364–368. Washington, DC.
- Steiner, H., Staudach, A., Spitzer, D. and Schaffer, H. (1994) Three-dimensional ultrasound in obstetrics and gynaecology: technique, possibilities and limitations. *Human Reproduction*, 9, 1773–1778.
- Treece, G. M., Prager, R. W., Gee, A. H. and Berman, L. (1998) *Fast Surface and Volume Estimation from Non-parallel Cross-sections, for Freehand 3-D Ultrasound*. Technical Report CUED/F-INFENG/TR 326, Cambridge University Department of Engineering.
- Trobaugh, J. W., Trobaugh, D. J. and Richard, W. D. (1994) Three-dimensional imaging with stereotactic ultrasonography. *Comput. Med. Imag. Graphics*, 18, 315–323.