Correcting Motion-Induced Registration Errors in 3-D Ultrasound Images*

Robert Rohling and Andrew Gee
University of Cambridge, Department of Engineering
Cambridge, England, CB2 1PZ
{rnr20,ahg}@eng.cam.ac.uk

Abstract

3-D ultrasound is a new medical imaging technique which can quickly and non-invasively deliver clinically useful renderings of internal structures. The technique is vulnerable to tissue motion during the scan which can result in an inconsistent 3-D data set. This paper describes an algorithm which corrects for tissue motion and other errors that cause mis-registration during a scan. The resulting 3-D data sets are clearer and of enhanced diagnostic utility. The registration is achieved by applying visual tracking techniques outside their traditional area of application. The algorithm is automatic, incremental and easily incorporated into existing 3-D ultrasound systems.

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 2-D grey-level image (B-scan) of a cross-section of the body in the scan plane. One of the limitations of conventional imaging is the requirement that the physician mentally reconstruct 3-D anatomy given multiple 2-D images. Research is underway to overcome this limitation using **3-D free-hand ultrasound imaging**. In this paradigm, a 3-D position sensor is attached to the probe, so that each B-scan can be labelled with the position and orientation of the scan plane — see Figure 1. Subsequent processing can build up a 3-D description of the imaged anatomy, in much the same manner as is possible using CT or MRI, but with less expensive and invasive technology. Physicians have indicated that there is significant utility in 3-D ultrasound imaging of a variety of anatomical structures, including the fetus [1], vascular structure [2], gall bladder [3], breast [4], kidney [5], and heart [6].

A major shortcoming of 3-D free-hand ultrasound imaging is its vulnerability to registration errors. Typically the largest source of registration error is tissue motion during the scan which can take up to a minute to complete. Significant errors can also arise from the position sensor (including calibration), refraction, propagation speed estimation, and others [7]. The cumulative errors result in mis-registered slices and an inconsistent 3-D data set which is difficult to interpret

*Proc. British Machine Vision Conference 1996, volume 2, pages 645–654, Edinburgh 1996.

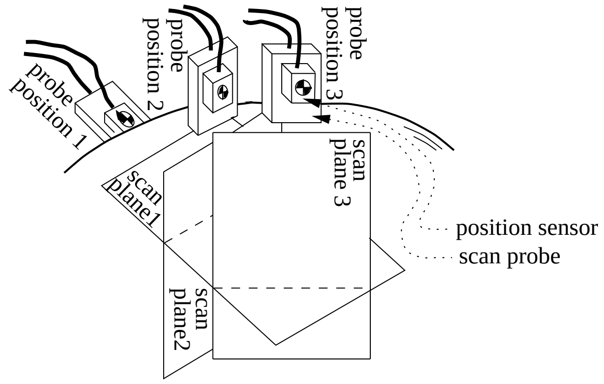


Figure 1: 3-D free-hand ultrasound imaging. Free-hand imaging allows the physician to move the probe as in a normal ultrasound examination. The position sensor measures the position and orientation of each scan plane. Note that the planes may intersect each other.

at best and completely unintelligible at worst. An example of motion-induced registration errors is illustrated in Figures 2 and 3. This paper presents a technique which corrects for tissue motion and several other errors that cause misregistration during a free-hand scan. The resulting 3-D data sets are much clearer and of greatly enhanced diagnostic utility. The registration is achieved by applying established visual tracking techniques outside their traditional area of application.

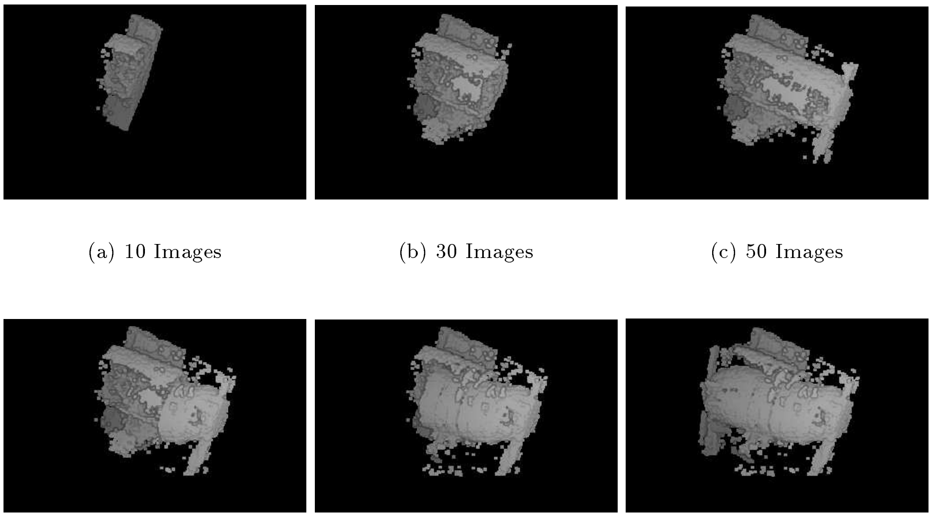
Little work on the registration of 3-D ultrasound data sets is evident in the literature. In one exception [4], two separate data sets were *retrospectively* registered using *manual* landmark matching. This constitutes a one-off, labour intensive solution to a specific registration problem. Here we propose an *automatic*, *incremental* registration algorithm for use with generic free-hand ultrasound imaging.

2 System Overview

The system comprises a Toshiba model SSA-270A/HG ultrasound scanner with a 3.75 MHz convex curvilinear array probe. The position and orientation of each acquired scan plane, relative to a fixed transmitter, are measured by an AC magnetic field receiver (Polhemus FASTRAK) mounted on the probe. B-scan images from the scanner are recorded by an 8-bit frame grabber at 5 frames per second and stored with the position data in the memory of a Sun SparcStation 10 workstation.

Each image is represented as a 2-D array ($\mathbf{P}$) of intensity values (p_{mn}). A 3-D scalar array ($\mathbf{C}$) of voxels c_{ijk} is chosen as the volumetric representation of the set of images. The voxel size, chosen according to the tradeoff between memory and resolution, is typically 7 times the size of the p_{mn} elements.

Before the start of the examination, the c_{ijk} voxels are all zero. As each image is acquired, each voxel c_{ijk} that is intersected by the scan plane is set to the average value of the p_{mn} elements which intersect the voxel. When an image intersects



(a) 10 Images (b) 30 Images (c) 50 Images

(d) 90 Images (e) 110 Images (f) 125 Images

Figure 2: Reconstruction of an organ phantom *without* correction for organ motion. A set of 125 B-scans was obtained by free-hand scanning. The images were acquired by scanning first from left to right, then right to left. These figures show the incremental filling of the reconstruction volume. Motion-induced registration errors are evident: the surface of the phantom constructed after the left-to-right scan (a to c) does not match the surface constructed by including the following right-to-left scan (d to f). In particular, the top of the reconstruction of all slices in (f) shows multiple surfaces created from the overlapping, mis-registered images. The rendering program attempts to generate surfaces from voxel intensities above a certain threshold. Portions of the surface of the phantom are missing in (c) because the threshold was chosen to minimise speckle and improve clarity of the image. The surfaces in all images are calculated with the same threshold.

voxels whose values have previously been set, a registration procedure is used to align the image to the existing data in $\mathbf{C}$ — see Section 3. The registered image is then added to $\mathbf{C}$ using a form of compounding.

Initial tests were performed on an organ phantom which consists of a water-filled latex balloon submerged in a warm water bath. To simulate breathing motion, the phantom was slowly moved half-way through each examination to a new location one half the width of the phantom away. *In-vivo* examinations were then performed on the gall bladder of a healthy human subject. A breath-hold was maintained but small subject movement during the examination was inevitable.

After a portion of $\mathbf{C}$ is filled it can be displayed by several methods including volume ray-tracing, surface rendering and any-plane slicing [8, 9]. Surface rendering is used in Figures 2 and 5, and any-plane slicing in Figures 3, 6 and 7.

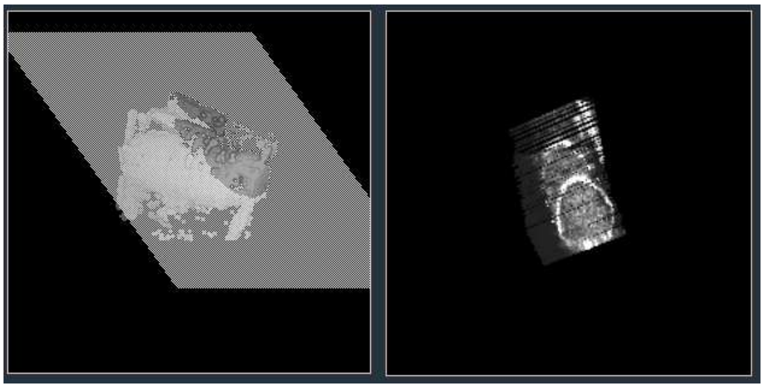


Figure 3: Cross-section of a reconstruction, *without* registration, of an organ phantom. The surface reconstruction (left image) is the same as in Figure 2(f). The cross-section clearly shows the two mis-aligned ellipses which correspond to the boundary of the phantom. The top ellipse corresponds to images taken during the left-to-right sweep of the scan and the bottom ellipse to the right-to-left sweep, after a small amount of motion of the phantom. The stratified look of the cross-section results from selecting a slice at an angle to the original B-scan images.

3 Registration Algorithm

3.1 Overview

Traditionally, registration problems in medical imaging are viewed as unconstrained searches for the transformation which brings one image (or set of images) into optimal alignment with another [10]. The search space is typically very large, resulting in a computationally expensive registration algorithm. With 3-D freehand ultrasound imaging, we can take advantage of the fact that individual images are acquired in rapid succession, and any registration error should vary smoothly from one frame to the next. This observation raises the possibility of performing efficient registration within a visual tracking framework.

$\mathbf{T}^*$, the optimal transformation that registers the image $\mathbf{P}$ with the existing data in $\mathbf{C}$, is constrained to compensate only for physically realistic registration errors. In general, it may be necessary to compensate for tissue motion, inaccurate measurement of the scan planes' positions, and imaging errors such as refraction and propagation speed estimation. In this study, however, the largest source of error was tissue motion during the scan. Registration by a rigid-body transformation can correct for simple tissue motion and scan plane position measurement errors. The imaging errors of refraction and speed estimation can only be partly corrected by rigid-body transformations. Imaging errors are small compared to motion and scan plane position errors if a small angular range of probe directions is used. For example, the phantom study corrected errors of up to 12 mm, and the *in-vivo* study up to 7 mm. For comparison, an example of speed estimation errors is readily calculated. The scanner uses a propagation speed estimate of

1540 m/s in the production of B-scans. If a B-scan is taken through 10 mm of fat (at 1440 m/s) followed by 90 mm of muscle (at 1570 m/s), object positions are mis-calculated by up to 1 mm. Warm water has a propagation speed near 1540 m/s so speed errors are negligible in the phantom tests. Different geometries and tissue types will produce errors that are still likely smaller than motion-induced errors. $\mathbf{T}^*$ is therefore constrained to a rigid 6 degree-of-freedom transformation.

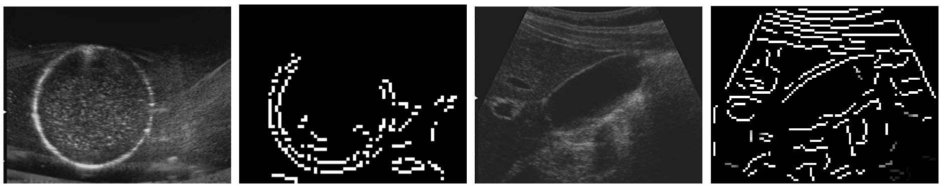
Smoothly varying registration errors result in smoothly varying parameters of $\mathbf{T}^*$, hence the suitability of a tracking framework. The tracking framework works best when objects are completely scanned from one extent to the other, then rescanned from different directions without lifting the probe from the subject. This is consistent with normal ultrasound examinations of internal organs.

Essentially, it is proposed to find $\mathbf{T}^*$ using *landmarks* [10]: anatomical features which can be reliably detected in all frames. Landmarks generally lie on surface boundaries which are distributed smoothly in 3-D space. Thus, when a newly acquired image $\mathbf{P}$ is inserted into $\mathbf{C}$, its landmarks should lie close to landmarks detected in other frames. Registration errors are corrected by searching for corresponding landmarks and applying a transformation to the new plane so that the landmarks are brought closer together. The landmarks used in this study are edge elements (edgels) automatically extracted by the Canny edge detection algorithm [11] from a thresholded B-scan. The edgels are produced at a lower resolution corresponding to the voxel size in the reconstruction. The edgel set is further pruned by chaining neighbouring edgels together and eliminating chains with fewer than three edgels. This reduces the number of edgels produced by speckle, an artifact common to all ultrasound images. Figure 4 shows typical B-scans of the phantom and gall bladder along with the edgels that are produced by the Canny edge detector. The set of edgels produced in this manner is sufficient for high contrast phantom and gall bladder reconstructions but more sophisticated landmark extraction techniques could be used for other *in-vivo* studies.

As the B-scans are acquired, edgels are stored in a 3-D vector array $\mathbf{L}$ which is aligned with $\mathbf{C}$. The data in the first image, and likely the next several images, will not intersect any existing data in the initially empty array $\mathbf{C}$. When a scan plane intersects non-empty voxels in $\mathbf{C}$, $\mathbf{T}^*$ is determined by finding correspondences between edgels in the new image and edgels in $\mathbf{L}$.

3.2 Selection of correspondences

Potential correspondences between edgels in the B-scan and edgels in $\mathbf{L}$ are found by searching an sub-volume in $\mathbf{L}$ around each edgel in the image. The search for correspondences must be restricted if the registration is to proceed rapidly. To this end, a *motion model* is required to describe the expected change in registration error from one frame to the next. Then, given a registration error in one frame, the search for correspondences in the next can be greatly constrained. A simple zero-velocity motion model, with a small uncertainty in position based on expected maximum motion, is sufficient to account for the motion-induced errors in the phantom tests. The motion model is restricted further to account for the smaller level of errors in the *in-vivo* examinations.



(a) Phantom (b) Edgels (c) Gall Bladder (d) Edgels

Figure 4: Automatic landmark detection. Images (a) and (c) show B-scans of the organ phantom and gall bladder respectively. Notice the speckling throughout the images and other artifacts. The intensity of the boundaries of both the phantom and gall bladder are also non-uniform and contain gaps. Images (b) and (d) show the edges detected by the Canny algorithm after pruning short chains of edgels. As well as the phantom and gall bladder boundaries, the detector also finds edges which do not correspond to real physical structures. These can generate false correspondences which must be tolerated by the registration algorithm.

3.3 Determination of transformation matrix

The set of correspondences produces more constraints than are required to determine $\mathbf{T}^*$. A least squares estimation of $\mathbf{T}^*$ is inappropriate because many of the candidate correspondences are outliers. Instead, the RANSAC regression technique [12] is used to determine $\mathbf{T}^*$ and reject the erroneous correspondences:

1. Choose three pairs of corresponding edgels at random from the full set of correspondences.

2. Reject these correspondences if they are not consistent with a rigid body transformation. Also reject these correspondences if the edgels in $\mathbf{P}$ are too close together or co-linear (otherwise the calculation of $\mathbf{T}$ in step 3 is ill-conditioned).

3. Calculate a linear affine transformation $\mathbf{T}$ which brings the three edgels in $\mathbf{P}$ into precise registration with the corresponding edgels in $\mathbf{L}$. The pruning in step 2 ensures that $\mathbf{T}$ represents a rigid body transformation. Transform all the remaining edgels in $\mathbf{P}$ by $\mathbf{T}$.

4. Count how many of the transformed edgels in $\mathbf{P}$ register with their counterparts in $\mathbf{L}$. Those edgels that do register contribute to the *consensus set* for $\mathbf{T}$. The remaining edgels are deemed outliers.

5. Repeat from 1 until a $\mathbf{T}$ is found with a consensus set larger than a preset threshold: this $\mathbf{T}$ becomes the estimate of the optimal transformation $\mathbf{T}^*$.

The consensus threshold is calculated as the percentage of the edgels in $\mathbf{P}$ that register with edgels in $\mathbf{L}$. For the phantom tests the threshold is set at 25%.

The *in-vivo* study uses 12% to reflect the higher number of false edgels. The technique described above is sufficiently robust to tolerate a significant proportion of erroneous correspondences.

$\mathbf{T}^*$ is used to transform the new image $\mathbf{P}$ into the coordinates of $\mathbf{C}$. The elements of the registered image can now be added to $\mathbf{C}$. Where the registered image intersects already-filled voxels, averaging is used to compound the new image with the existing data in $\mathbf{C}$. Speckle patterns become uncorrelated if a region is viewed from different look directions so speckle (and other artifacts) can be significantly reduced by averaging [9, 13]. Compounding images taken from different look directions theoretically increases the signal-to-noise ratio in proportion to the square root of the number of compounded images [9, 14]. Yet, accurate registration is a prerequisite for maintaining image quality.

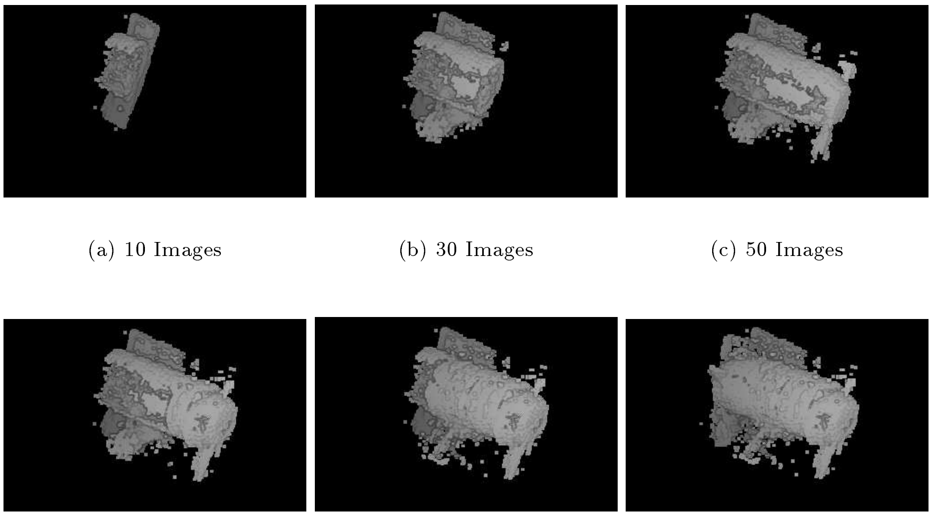
4 Results, Conclusions and Future Work

The registration algorithm yields a more consistent reconstruction of a 3-D data set from multiple B-scans than reconstructions based on position sensor data alone. Figures 5 and 6 show the reconstruction, with registration, of the phantom which was allowed to move during the scan. Figure 7 demonstrates the need and benefits of registration for *in-vivo* examinations. Clearly, motion-induced errors must be corrected to achieve clear 3-D renderings. Furthermore, compounding intersecting registered images results in a reduction of the speckle noise and can fill in regions of the reconstruction that were missed by a single sweep of the anatomical region of interest. The ultrasound examinations in this paper contain a limited number of intersecting images so only a small amount of speckle reduction is observed. Work is underway on a new system to acquire B-scans at 25 frames per second which will allow a higher level of compounding.

Success of the registration technique has shown a sensitivity to the parameters values such as the size of the correspondence search space and the consensus threshold. Essentially, for each new type of examination, the series of B-scans must be previewed to determine the level of false edgels and size of the maximum registration error. Once an optimal set of parameter values is determined for a particular type of examination (the gall bladder for example) the reconstruction proceeds automatically by registering each new B-scan as it is acquired.

The robustness of the algorithm could be improved by employing more sophisticated motion models tuned to specific motion patterns (eg. breathing). The motion models can be learned from examples using established system identification techniques [15]. For difficult data sets, where the detection of landmarks is unreliable, it may be necessary to employ a tracking algorithm which is capable of maintaining several motion hypotheses simultaneously [16].

In summary, the registration technique described in this paper demonstrates the ability to correct motion-induced errors and improve the quality of the 3-D reconstructions. 3-D ultrasound data is normally difficult to segment and visualise. 3-D data sets reconstructed from many overlapping, but accurately registered, B-scans is a very promising technique for improving the data quality and subsequent segmentation and visualisation.



(a) 10 Images (b) 30 Images (c) 50 Images

(d) 90 Images (e) 110 Images (f) 125 Images

Figure 5: Reconstruction of an organ phantom *with* correction for organ motion. The use of the registration algorithm during reconstruction results in a better alignment of boundary contours — compare with Figure 2. The surfaces in all images are calculated at the same threshold, equal to the value used in Figure 2. The compounding of both the left-to-right and right-to-left scanned images is particularly useful for reducing the uncorrelated speckle noise and filling in surface patches missed by the single left-to-right sweep. With 1 mm^3 voxels, 46% of the filled voxels are intersected by 1 B-scan, 34 % are intersected by 2, 10% by 3, and 10% by 4 or more. Further compounding will further reduce the speckle noise.

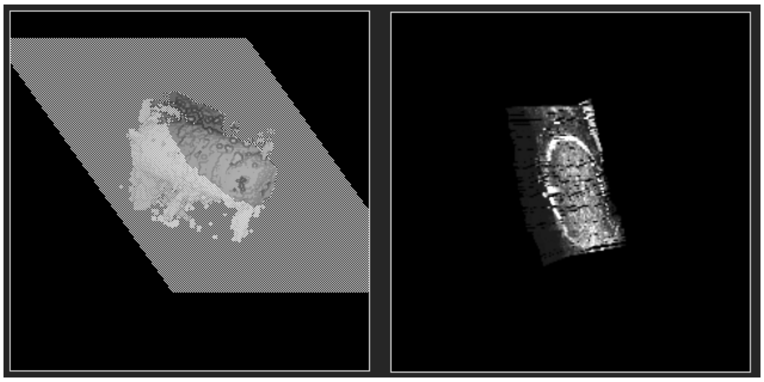
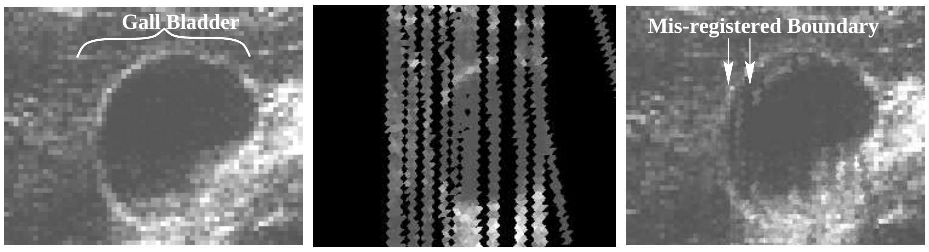


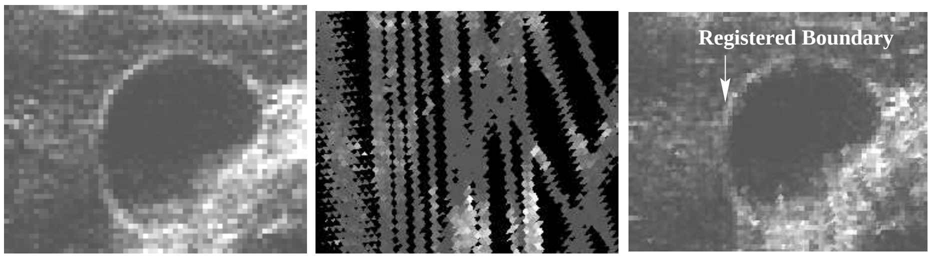
Figure 6: Cross-section of a reconstruction, *with* registration, of an organ phantom. The surface reconstruction (left image) is the same as Figure 5(f). The cross-section is at the same location as in Figure 3. The two ellipses have been aligned.



(a) Original Image

(b) Existing Images at Measured Location of (a)

(c) Superposition of (a) and (b)



(d) Original Image: identical to (a)

(e) Existing Images at Registered Location

(f) Superposition of (d) and (e)

Figure 7: Registration of an *in-vivo* transverse and longitudinal gall bladder examination. The bottom row shows the elimination of a double boundary, shown in the top row, that is produced by mis-registered images. This figure is produced with 1 mm^3 voxels to enhance clarity, but registration was performed with $2 \times 2 \times 2$ mm voxels. With $2 \times 2 \times 2$ mm voxels, 35% of filled voxels are intersected by 1 B-scan, 33% are intersected by 2, 16% by 3, and 16% by 4 or more.

Acknowledgements

The authors thank Richard Prager and Chris Dance for many useful discussions. Ultrasound scanning was done with the help of Laurence Berman at the University of Cambridge Department of Radiology. The 3-D renderings were produced using 3DViewnix. Robert Rohling is supported by Churchill College and an ORS award.

References

[1] Chervenak, F A, Isaacson, G C, Campbell, S, *Ultrasound in Obstetrics and Gynecology*, Little, Brown and Company, Boston, MA, USA, 1993.

[2] Franseschi, D, Bondi, J A, Rubin, J R, "A New Approach for Three-Dimensional Reconstruction of Arterial Ultrasonography", J. Vasc. Surgery,

15, 5, May 1992, pp 800–805.

[3] Fine, D, Perring, S, Herbetko, J, Hacking, C N, Fleming, J S, Dewbury, K C, "Three-Dimensional (3D) Ultrasound Imaging of the Gallbladder and Dilated Biliary Tree: Reconstruction from Real-Time B-Scans", The British Journal of Radiology, **64**, 1991, pp 1056–1057.

[4] Moskalik, A, Carson, P L, Meyer, C R, Fowlkes, J B, Rubin, J M, Roubidoux, M A, "Registration of Three-Dimensional Compound Ultrasound Scans of the Breast for Refraction and Motion Correction", Ultrasound in Medicine and Biology, **21**, 6, 1995, pp 769–778.

[5] Gilja, O H, Smievoll, A I, Thune, N, Matre, K, Hausken, T, Odegaard, S, Berstad, A, "In Vivo Comparison of 3D Ultrasonography and Magnetic Resonance Imaging in Volume Estimation of Human Kidneys", Ultrasound in Medicine and Biology, **21**, 1, 1995, pp 25–32.

[6] Salustri, A, Roelandt, J R T C, "Ultrasonic Three-Dimensional Reconstruction of the Heart", Ultrasound in Medicine and Biology, **21**, 3, 1995, pp 281–293.

[7] Rohling, R N, Gee, A H, *Issues in 3-D Free-hand Medical Ultrasound Imaging*, CUED/F-INFENG/TR 246, Cambridge Univ. Dept. of Eng., January 1996. available by ftp: svr-ftp.eng.cam.ac.uk/reports/rohling_tr246.ps.Z.

[8] Fuchs, H, Levoy, M, Pizer, S M, "Interactive Visualization of 3D Medical Data", Computer, **22**, 8, August 1989, pp 46–51.

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

[10] Brown, L G, "A Survey of Image Tegistration Techniques", ACM Computing Surveys, **24**, 4, December 1992, pp 325–376.

[11] Canny, J, "A Computational Approach to Edge Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, **8**, 6, 1986, pp 679–698.

[12] Fischler, M A, Bolles, R C, "Random Sample Consensus: a Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Communications of the ACM, **24**, 6, 1981, pp 381–395.

[13] Nelson, T R, Pretorius, D H, "3D Ultrasound Image Quality Improvement Using Spatial Compounding and 3D Filtering", Med.Phys.,**21**,6, 1994, pp 998.

[14] Trahey, G E, Smith, S W, von Ramm, O T, "Speckle Pattern Correlation with Lateral Aperture Translation: Experimental Results and Implications for Spatial Compounding", IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, **UFFC-33**, 3, 1986, pp 257–264.

[15] Blake, A, Isard, M, Reynard, D, "Learning to Track the Visual Motion of Contours", Artificial Intelligence, **78**, 1995, pp 101–133.

[16] Isard, M, Blake, A, "Contour Tracking by Stochastic Propagation of Conditional Density", Proc. 4th Europ. Conf. on Comp. Vis., **1**, 1996, pp 343–356.