Volume Reconstruction from Sparse 3D Ultrasonography

Mark J. Gooding[1,⋆], Stephen Kennedy[2], and J. Alison Noble[1]

[1] Medical Vision Laboratory, University of Oxford, UK
{gooding,noble}@robots.ox.ac.uk

[2] Nuffield Dept. of Obstetrics and Gynaecology, University of Oxford, John Radcliffe Hospital, UK skennedy@molbiol.ox.ac.uk

Abstract. 3D freehand ultrasound has extensive application for organ volume measurement and has been shown to have better reproducibility than estimates of volume made from 2D measurement followed by interpolation to 3D. One key advantage of free-hand ultrasound is that of image compounding, but this advantage is lost in many automated reconstruction systems. A novel method is presented for the automated segmentation and surface reconstruction of organs from sparse 3D ultrasound data. Preliminary results are demonstrated for simulated data, and two cases of in-vivo data; breast ultrasound and imaging of ovarian follicles.

1 Introduction

Ultrasound imaging is used widely in clinical medicine. Its benefits include speed, low cost and the limited exposure risk associated with it. Although imaging in 3D is starting to become more common place, most clinical scanning remains 2D. This has obvious disadvantages when it comes to quantitative analysis, but even used qualitatively, the method is problematic. Clinicians must mentally reconstruct the 3D tissue structure in order to ascertain shape or position of a region of interest. This mental reconstruction is subjective and dependent on the knowledge and experience of the ultrasonographer. A brief outline of 3D scanning techniques is given here, a more detailed review can be found in [1]. There are three main methods used for 3D ultrasonography; free-hand scanning, mechanically driven 3D, and 2D array probes. In this paper we concentrate on the first two techniques. The latter is only starting to be used in particular clinical applications, e.g. cardiology. These probes need further development to have wider applicability. Both free-hand and mechanical 3D ultrasound scanning produce sparse data-sets. Benefit may be derived from using image compounding to reduce noise and artifacts, where image planes intersect [2]. In the case of free-hand scanning this allows multiple views of the same organ, which can be used to circumvent problems associated with acoustic shadowing.

⋆ Mark Gooding is funded by the EPRSC as part of the MIAS-IRC. (GR/N14248)

R.E. Ellis and T.M. Peters (Eds.): MICCAI 2003, LNCS 2879, pp. 416–423, 2003.

There are two main techniques for object reconstruction; those in which segmentation of images is performed prior to object reconstruction and those which perform it after image reconstruction. In the former case, the 2D boundaries are used to guide a meshing algorithm to provide a 3D view of an object of interest. In the latter case, the images are placed into a 3D voxel array and grey-level interpolation is used to fill the gaps. This approach enables generation of images from any viewpoint, together with viewing of a segmentation object.

A review of medical applications of image segmentation and object reconstruction is presented in [1]. Our interest primarily lies in the use of these methods in ovarian follicular volume estimation during assisted reproduction techniques such as in-vitro fertilisation (IVF). A number of studies have shown the benefit of 3D imaging in fertility treatment using manual segmentation of mechanically generated 3D scans to calculate volume, for example[3]. The use of automated methods for object reconstruction has been limited [4,5]; ter Harr-Romeny et al. [5] used mechanical 3D ultrasound to scan ovarian follicles, performing processing on the 3D reconstructed image data. After detecting the follicle centre, the follicle boundary was found using a form of edge detection at multiple scales on lines radiating from the center points. A spherical harmonic surface was fitted to the sparse edge points. In general it appears that most reconstruction methods, with the exception of [5], adopt the approach of segmentation prior to object reconstruction. However, segmentation of individual images prior to reconstruction cannot take advantage of noise suppression through image compounding [2].

We use a Level Set method for object reconstruction. Conventionally a single level set function is used to embed a single object class as in [6]. However we require identification of multiple object classes. In this paper we present a novel variant of the Level Set method [7] which allows for the simultaneous reconstruction of multiple objects from sparse data. Although not limited to these applications, preliminary results are presented for in-vivo data from free-hand 3D breast ultrasound and ovarian scans.

2 Reconstruction Method

The Level Set approach is a powerful tool which finds application in many fields including medical image segmentation and object reconstruction [7]. The essence of the approach is to define a boundary implicitly as the zero level set of a higher dimensional function, for example a curve (1D) is represented by the zero level set ($\phi = 0$) of a surface, ϕ (2D). The advantage of this representation is that complex topology and surface evolution, for example curve merging, can be handled in an elegant manner. A full explanation of the method can be found in [7]. The main equation solved by the method is:

$$\phi_t + F|\nabla\phi| = 0 \tag{1}$$

where the function, $\phi(t)$, is evolved over time using a speed function, F, such that the zero level set, $\phi = 0$, at time $T = \infty$ is the optimal solution for the application of interest; in our case, the segmentation and reconstruction of sparse ultrasound data. Equation 1 may be numerically minimised by defining the iterative update equation:

$$\phi_{new} = \phi_{old} - \Delta T F|\nabla\phi| \tag{2}$$

where ΔT is a small time step. A speed function, F, must be defined for the application of interest. A method for reconstructing an object from sparse known edge points was presented in [6], where F was defined as:

$$F = \nabla d.\frac{\nabla\phi}{|\nabla\phi|} + \frac{d}{p}\nabla.\frac{\nabla\phi}{|\nabla\phi|} \tag{3}$$

Here d is the distance to the nearest edge point and p is a weighting factor controlling the smoothness of the solution. In this case the speed function finds the weighted minimal surface with respect to the edge points. Although such a method could be used to fit a surface to 2D segmentations, our aim is to segment sparse 3D images after reconstruction. To this end we propose a new speed function as follows:

$$F = \alpha F_{surf} + \beta F_{image} + \gamma F_{reg} \tag{4}$$

where F_{surf} is the surface reconstruction term in Equation 3, F_{image} is a segmentation term and F_{reg} is a regularisation term; in this instance proportional to the level set curvature $\nabla.\frac{\nabla\phi}{|\nabla\phi|}$. The purpose of this last term is to keep the segmentation result smooth. The parameters α, β and γ are application specific and must be determined empirically. Our method is as follows: first the free-hand data is reconstructed as a volume image. Then the level set is evolved using information from the volume image to guide both the segmentation and reconstruction. The distance to the edge point required for Equation 3 is calculated at each iteration from the current positions where the zero-level set intersects the image data.

A relatively simple segmentation term, F_{image} is used in the work. Given a prior segmentation, whether by initialization or as a result of a previous iteration, each region is labelled with a class, c, such that $c(\mathbf{x})$ is the current class at point $\mathbf{x}$ within the volume image. Associated with each class is a non-parametric probability density function (PDF) derived from the intensity of the points contained within the class. We then define $p_{c(\mathbf{x})}(v)$ as the probability that intensity value v belongs to class $c(\mathbf{x})$, which is found from the PDF of $c(\mathbf{x})$. The intensity value at point $\mathbf{x}$ is given by the mean intensity within a neighbourhood, $N_2(\mathbf{x})$, around that point. For a particular point, $\mathbf{x}$, we consider the probability of membership to the region to which it currently belongs and to any region within a neighbourhood, $N_1(\mathbf{x})$ around that point. The size of N_1 is on the scale of the level set function discretisation, whilst N_2 is chosen to get a good approximation to the local data. F_{image} is set to the difference in the probability of membership between the current class and the most probable neighbouring class. For non-boundary pixels where all points within $N_1(\mathbf{x})$ are the same class, or for areas where there is no data within $N_2(\mathbf{x})$, F_{image} is set to zero. This results in the segmentation term, F_{image}, having a value between -1 and 1, with the sign chosen such that the region is extended if it is more probable that the point belongs to the class that it is already in and shrinks if the probability of belonging to a neighbouring class is higher. This can be expressed as:

$$F_{image}(\mathbf{x}) = \max_{\forall \mathbf{u} \in N_1(\mathbf{x}) | \mathbf{u} \neq \mathbf{x}, c(\mathbf{u}) \neq c(\mathbf{x})} \left[p_{c(\mathbf{u})}(\mu(N_2(\mathbf{x}))) - p_{c(\mathbf{x})}(\mu(N_2(\mathbf{x}))) \right] \quad (5)$$

where $\mu()$ is the mean value over a neighbourhood. For all $\mathbf{x}$ where $N_2(\mathbf{x}) = \emptyset$ or where $\forall \mathbf{u} \in N_1(\mathbf{x}), c(\mathbf{u}) = c(\mathbf{x})$;

$$F_{image}(\mathbf{x}) = 0 \quad (6)$$

2.1 Implementation of the Object Reconstruction

The implementation of the level set method is done in a similar way to [6], but with two important modifications.

First, we subsample the 3D image into a voxel array of the same resolution as the level set function voxel array, with the mean intensity being used in any voxel with more than a single pixel falling in it. In such an arrangement we may consider the neighbourhood, N_2, of a point as being the voxel in which it falls, with calculations of level set speed only occurring at each voxel centre near the boundary. N_1 is defined as the 27-voxel neighbourhood of each voxel centre. In principle, the reconstructed image can be kept in the form of a position-intensity pair, where the position is not quantised to a voxel array but is in "real space". Such a scheme is used in [6], however once the distance field d is calculated for each point within the Level Set voxel array, the raw data can be discarded. In our method the raw data cannot be discarded since the intensity at each position is needed for the F_{image} term, and d is recalculated at each iteration. Since our data sets are very large (of the order 10^6 points), the memory requirements to store the information make such an approach unfeasible so we adopt the voxel based representation.

Second, Level Set segmentation methods exist which operate by evolving multiple coupled surfaces in parallel, requiring N[8], or at best $logN$[9], embedded functions for N classes. In [10] a method is presented for embedding N classes in a single level set function, which although slow is memory efficient. For 3D applications, memory becomes more constrained than for 2D image analysis and as a result a modification of the implementation in [10] has been developed as follows. Multiple classes evolution is achieved by storing a class label for each voxel. When the sign of ϕ changes for a particular voxel, its label either becomes that of the background class, for $\phi > 0$, or the same as the object that it is touching. If two different object classes come within 2 voxels of each other both have the speed set $F = -1$ such that they will be driven apart again, as this prevents problems of class assignment occurring on the boundary between the object classes. Once the regions are "driven back", the class with the highest true speed value is the first to move back into the gap and the two regions compete in this way. This method varies from [10], by allowing for non-binary speed functions, storing of the class labels, and preventing region merging. The initial seeding is performed manually and merging is prevented because, in this application, neighbouring objects share the same statistical appearance.

3 Experimental Analysis

Examples of applying the method to both simulated and in-vivo data are given in the section. In-vivo results are shown for two clinical ultrasound applications: breast mass detection and fertility treatment.

3.1 Simulated Data Sets

In this experiment the simulated data consists of a spherical object of radius 20 voxels. Simulated scans of this object were made, such that for each plane the regions corresponding to the sphere would have intensity values in the range from 60 to 120, uniformly distributed, while the background has intensities from 10 to 240, uniformly distributed. The sampled intensities from each plane were quantised to a voxel array. Each voxel on a plane had between 30 and 60 intensity values assigned to it for simulated compounding. Simulations were made with 4 scan patterns; linear sweep across the x-axis at 2 and 5 voxel spacing, and rotational about the x-axis at $\frac{\pi}{24}$ and $\frac{\pi}{12}$ radian spacing. Two spherical initialisations were used, centred at $(50, 50, 50)$, with radii of 15 and 25 voxels.

Table 1 shows the volume error for each of the simulated data test. All volume estimates fall within an equivalent of 1 voxel change of radius. The linear scan measure shows larger error for the smaller initialisation as the method cannot extend to unconnected scan planes. Closer spacing of planes, for both linear and rotational scans, gives greater accuracy as expected.

Table 1. Volume results for simulated data compared with true volume.

Scan type and spacing	initial radius (voxels)	estimated volume (voxels)	error (%)
linear, 5 voxels	15	29819	-10.65
linear, 5 voxels	25	32173	-3.58
linear, 2 voxels	15	29861	-10.52
linear, 2 voxels	25	33493	0.37
rotation, $\frac{\pi}{12}$ rads	15	30573	-8.38
rotation, $\frac{\pi}{12}$ rads	25	32417	-2.86
rotation, $\frac{\pi}{24}$ rads	15	32406	-2.89
rotation, $\frac{\pi}{24}$ rads	25	33338	-0.10

3.2 In-vivo Scanning

Breast data: The breast ultrasound data consist of 174 B-mode images recorded at approximately 25Hz using a linear sweep across a cyst. The images were scanned using an AuIdea4 (Esaote) and an LA13 7.5Mhz linear array probe. The positions were recorded by a Polaris Hybrid optical tracker (Northern Digital Inc). No quantitative measurements of the cyst volume were available.

Figure 1 shows the segmentation and surface fit of the breast cyst. Visually, both the segmentation and object reconstruction appear good. A shift in the surface of the cyst can be observed. This was caused by variation in the contact force between the probe and the breast, resulting in variable compression of the cyst and the breast tissue. This is a significant source of error and must be addressed before quantitative measurements can be made [11].

Follicular data: In this experiment the data consists of scans from 2 patients undergoing IVF treatment. Each set contains 180 B-mode images of an ovary recorded at approximately 12Hz using a rotational motion. Ovary 1 contained one follicle, Ovary 2 contained three follicles. The images were scanned using a Powervision 6000 (Toshiba Medical Systems) and a 7.5MHz transvaginal probe. Positions were recorded by a Faro Arm (Faro Technologies Inc). Linear measurements were made of the follicle during 2D scanning. Each follicle was aspirated as part of the normal IVF treatment, shortly after scanning, allowing the associated volume to be recorded.

Consenting patients were scanned at the John Radcliffe Hospital, Oxford, U.K. Ethics committee approval had been granted for both acquistions.

Figure 2A shows the reconstruction. Although the reconstruction appears good, Table 2 shows that the method underestimates the aspirated volume in 3 out of 4 cases. Despite this underestimate the reconstructed volume is of a similar accuracy to the volume of a sphere calculated from clinical measurements. The mean measurement of two diameters is currently used by clinical staff as an indicator of follicle size. The re-sliced compounded image (Fig. 2B) reveals that compounding leads to lower image quality as a result of misplaced images. Patient breathing and motion have an effect on the resulting segmentation and hence the accuracy of the measurements, particularly for the second ovary.

Table 2. Measurements of follicle volume compared to aspirated volume

Ovary/ follicle	mean diameter measured in 2D US (mm)	estimated volume from 2D US (ml)	estimated volume from 3D US (ml)	aspirated [true] volume (ml)	error in estimate from 2D US (%)	error in estimate from 3D US (%)
1/i	21	4.9	6.21	7.0	-30	-11
2/i	22	5.6	2.91	5.5	+1.8	-47
2/ii	22	5.6	4.70	7.0	-20	-32
2/iii	9	0.4	1.57	1.0	-60	+57

4 Discussion and Conclusion

This paper has presented a novel method for the 3D volume reconstruction from sparse 3D (ultrasound) scans. Initial experimental results are encouraging despite the simple segmentation model, with reconstruction of artificial data

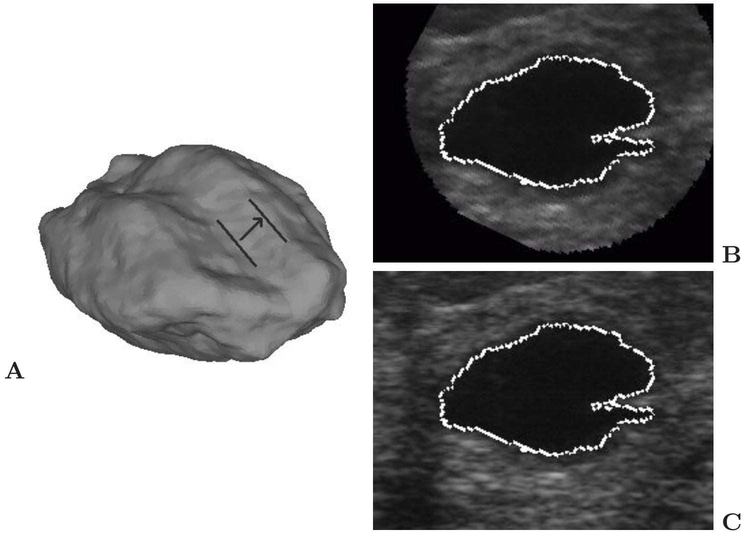

Fig. 1. A shows the 3D shape of the breast cyst when reconstructed in 3D. The shift in the surface, as indicated by the arrow, is as a result of breast deformation under different probe contact pressure with the breast. **B** shows the segmentation overlaid on the compounded image for a particular plane. **C** shows the same segmentation overlaid on the original image from that plane.

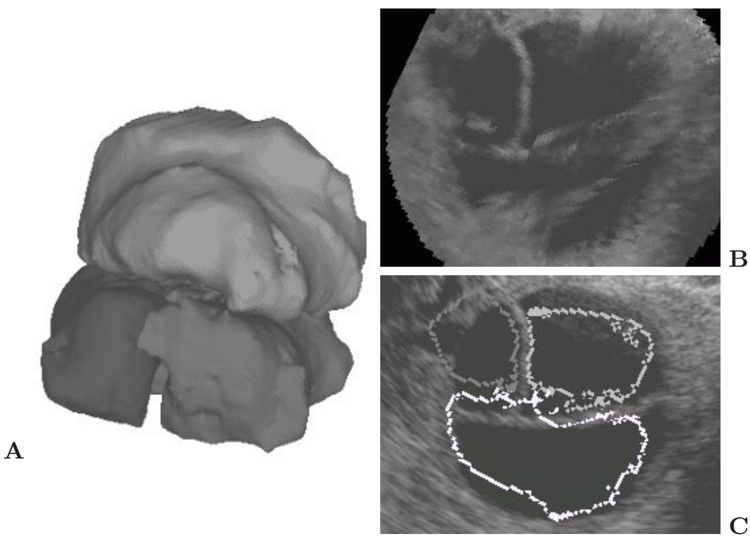

Fig. 2. A shows the shape of the follicles when the ovary is reconstructed in 3D. **B** shows the compounded image for a particular plane. **C** shows the same segmentation overlaid on the original image from that plane. Compounding can be seen to be making image quality and the resulting segmentation worse. This effect is a result of patient motion and breathing.

falling within 1 voxel radius of the true volume. The preliminary results on in-vivo scans are encouraging, showing plausible segmentation results. The resulting volume estimates are disappointing as a result of patient motion, but have similar error range to 2D clinical measurement.

Two particular problems need addressing in future work: first, problems with the data acquisition process, for example patient motion and deformation due to probe contact force, need consideration. These are not problems of the algorithm *per se*, but do affect the accuracy of the resulting segmentation and volume estimation. Second, a feature of the segmentation term is that compounding gives better separation for classes with different mean values. However segmentation will fail for classes with identical, or close, means. This can be addressed by using a different measure to calculate class membership for each voxel. A more sophisticated segmentation term could prevent underestimation resulting from multiple classes falling in a single voxel on the class boundaries.

Acknowledgement. We wish to thank the staff of the Oxford Fertility Unit, particularly Debbie Barber, for their help in data acquisition. We are grateful to Toshiba Medical Systems for the loan of the Powervision for this research.

References

1. A. Fenster and D.B. Downey. 3-d ultrasound imaging: A review. *IEEE Engineering in medicine and biology*, 15(6):41–51, 1996.
2. R.N. Rohling, A.H. Gee, and L. Berman. Three-dimensional spatial compounding of ultrasound images. *MIA*, 1(3):177–193, 1997.
3. A. Kyei-Mensah et al. Transvaginal three-dimensional ultrasound: accuracy of follicular volume measurements. *Fertility and Sterility*, 65(2):371–376, 1996.
4. F.G. Balen et al. 3-dimensional reconstruction of ultrasound images of the uterine cavity. *The British Journal of Radiology*, 66:588–591, 1993.
5. B.M. ter Haar Romeny et al. Computer assisted human follicle analysis for fertility prospects with 3d ultrasound. *IPMI*, pages 56–69, 1999.
6. H.K. Zhao et al. Implicit and non-parametric shape reconstruction from unorganized points using variational level set method. *Computer Vision and Image Understanding*, 80:295–319, 2000.
7. J.A. Sethian. *Level set methods and fast marching methods.* CUP, 2nd edition, 1999.
8. H.K. Zhao et al. A variational level set approach to multiphase motion. *Journal of Computational Physics*, 127:179–195, 1996.
9. L.A. Vese and T. Chan. A multiphase level set framework for image segmentation using the mumford-shah model. *IJCV*, 50(3):271–293, 2002.
10. T. Kadir and J.M. Brady. Unsupervised non-parametric region segmentation using level sets. To appear: ICCV, 2003.
11. M. Burcher, J.A. Noble, and L. Han. Deformation correction in ultrasound images using contact force measurements. *Proc. IEEE MMBIA*, pages 63–70, 2001.