Automated 3-D Echocardiography Analysis Compared With Manual Delineations and SPECT MUGA

Gerardo I. Sanchez-Ortiz, Gabriel J. T. Wright*, Nigel Clarke, Jérôme Declerck, Adrian P. Banning, and J. Alison Noble, *Member, IEEE*

Abstract—A major barrier for using 3-D echocardiography for quantitative analysis of heart function in routine clinical practice is the absence of accurate and robust segmentation and tracking methods necessary to make the analysis automatic. In this paper, we present an automated three-dimensional (3-D) echocardiographic acquisition and image-processing methodology for assessment of left ventricular (LV) function. We combine global image information provided by a novel multiscale fuzzy-clustering segmentation algorithm, with local boundaries obtained with phase-based acoustic feature detection. We then use the segmentation results to fit and track the LV endocardial surface using a 3-D continuous transformation. To our knowledge, this is the first report of a completely automated method. The protocol is evaluated in a small clinical case study (nine patients). We compare ejection fractions (EFs) computed with the new approach to those obtained using the standard clinical technique, single-photon emission computed tomography multigated acquisition. Errors on six datasets were found to be within six percentage points. A further two, with poor image quality, improved upon EFs from manually delineated contours, and the last failed due to artifacts in the data. Volume–time curves were derived and the results compared to those from manual segmentation. Improvement over an earlier published version of the method is noted.

Index Terms—Left ventricular (LV) volume, quantitative analysis, three-dimensional (3-D) echocardiography, 3-D functional imaging.

I. INTRODUCTION

A. Motivation

THE LAST few years have seen the emergence of three-dimensional (3-D) echocardiography acquisition systems in the market. Methods of acquisition are improving in terms of spatial and temporal resolution, moving now toward real-time volumetric acquisition. However, interpretation and analysis of the data is more complex and time consuming than for conventional two-dimensional (2-D) echocardiography. As recent research studies have shown—for example, [1]–[3]—automated three-dimensional analysis provides more precise information about the pathophysiology of the heart than conventional analysis of 2-D views ([4]–[6] and references therein) and is of particular help for volume and ejection fraction (EF) calculation.

A good summary of the state-of-the-art in clinical testing of 3-D cardiac image analysis methods can be found in [7]. Only a few papers referenced concern echocardiography [1], [8]–[13], and of these, only [10]–[14] deal with quantitative analysis of function and performance. Even for volume or ejection fraction computations, standard clinical routines like conventional single-photon emission computed tomography (SPECT) and multigated acquisition [15] (MUGA) are costly and require the injection of radioactive isotopes. In the specific case of 3-D echocardiography, automated analysis would seem particularly hard to achieve because of the inherent difficulty in segmenting noisy and contrast-variable ultrasound images. Yet, advantages, such as high temporal resolution, the wide range of cardiac disorders on which it is used as an important indicator, as well as other characteristics such as being a safe, noninvasive, low-cost, portable, and fast procedure, make it ideal for routine clinical practice and an important topic of research.

For these reasons, the goal of this research was to demonstrate the feasibility of obtaining good volume estimates from data derived from a state-of-the-art commercial rotational 3-D echocardiography transducer using a fully automated image-processing solution.

B. Related Work

Previous work at the University of Washington has shown the feasibility of reconstructing a three-dimensional surface of the heart from sparse view freehand echocardiography [2], [10]. However, in that work, a significant amount of interaction was required to obtain a reconstruction of the endocardium, and segmentation was performed manually. Clinically, this is a critical point because manual interaction with 3-D data might be acceptable for research purposes but is too time-consuming and tedious to do on a routine basis.

Recent work from the University of Yale [1] is probably the most advanced heart modeling and analysis method developed to date for echocardiography imagery. In this approach, a 3-D finite-element mesh of the left ventricular (LV) myocardium is

G. I. Sanchez-Ortiz was with the Department of Engineering Science, University of Oxford, OX1 3PJ Oxford, U.K. He is now with the Department of Computing, Imperial College London, SW7 2AZ London, U.K.

*G. J. T. Wright is with the Department of Engineering Science, University of Oxford, Parks Road, OX1 3PJ Oxford, U.K. (e-mail: gabriel@robots.ox.ac.uk).

N. Clarke and A.P. Banning are with the Oxford Radcliffe Hospitals NHS Trust, OX3 9DU Oxford, U.K.

J. Declerck was with the Department of Engineering Science, University of Oxford, OX1 3PJ Oxford, U.K. He is now with Mirada Solutions Ltd., OX2 0JX Oxford, U.K.

J. A. Noble is with the Department of Engineering Science, University of Oxford, OX1 3PJ Oxford, U.K.
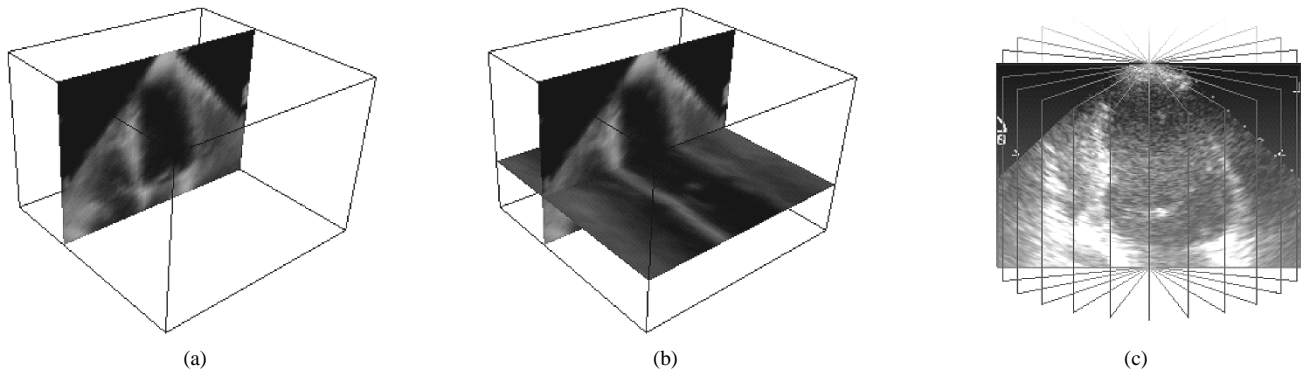
Fig. 1. (a) A single time frame from a 2-D $+T$ sequence. The sequence can be treated as a 3-D image with time as the third dimension; (b) shows additional slices projected through time; (c) shows the position of the 12 slices that form the complete data set.

computed and used to perform myocardial strain analysis. The approach is interesting. However, the analysis depends on the high-quality images acquired from an open-chest acquisition and is obviously not viable for routine clinical practice.

Following our own experience with other approaches to 2-D and 3-D echocardiography and ultrasound image analysis (including two-dimensional tracking [16] and 3-D freehand acquisitions [17], [18]), we believe that 3-D rotational probe echocardiography is the most viable approach for immediate clinical use, given the current technology. Hence our approach is based on this type of 3-D ultrasound acquisition.

II. METHOD

A. General

The protocol consists of the following steps.

1) A transthoracic rotating 3-D probe is used to obtain 12 equally spaced coaxial 2-D time sequences (what we call 2-D $+T$ images). See Section II-B.
2) Phase-based boundary detection is applied to each 2-D $+T$ image sequence for endocardial border localization. See Section II-C.
3) In parallel to 2), multiscale fuzzy clustering is applied to each 2-D $+T$ image sequence for robust continuous segmentation of the LV cavity. See Section II-D.
4) A combined filtering and reconstruction algorithm is used to select candidate endocardial boundary points from the complete set of candidate boundary points identified by the phase-based segmentation. This step also combines the 12 2-D $+T$ data sets into one 3-D $+T$ data set. See Section II-E.
5) Fitting and tracking is performed using the 3-D endocardial cluster as an initialization surface and deforming it to 3-D filtered boundary candidates at time frame zero. Then, the resulting surface is used to initialize the surface fit to boundary candidates at the next time frame. Volume and deformation parameters are then estimated from the resulting 3-D endocardial surface sequence. See Section II-F.

On the algorithmic side, new aspects of the work include step 3), where we introduce a simple geometric and dynamic model of the LV as part of the clustering algorithm used to initialize the surface fit, and step 4), the image feature filtering

stage. The other main contribution of this paper is the evaluation of the complete method against MUGA in a small clinical study, which to our knowledge is the first such study completed using a fully automated image-analysis method.

B. Image Acquisition

Digital 3-D echocardiographic data were acquired on an HP SONOS 5500 ultrasound machine (Agilent Technologies, Andover, MA) using a 3–5 MHz rotating transducer and respiration and ECG gating. In this machine, data are stored as 2-D $+T$ images, one for each probe angle [see Fig. 1(b) for an example of a 2-D $+T$ image sequence].

Data were acquired at a frame rate of 24 frames per second, the pixel size being 0.5×0.33 mm^2. Scanning was performed on an apical view, i.e., the probe was located at the apex and roughly aligned with the LV long axis in two orthogonal views prior to the 3-D acquisition. We used sparse data i.e., 12 coaxial planes, one every 15° [see Fig. 1(c)]. In previous work [3], [19], we studied the effect of using dense (60 slices, one every 3°) and sparse (7–12 slices) data sets and concluded that 12 was a reasonable compromise between the speed of image acquisition and the number of images necessary to obtain a reasonably continuous sampling of the LV wall.

It is important to point out that the data acquired using current 3-D technology are of lower quality (in terms of spatial and temporal resolution as well as signal-to-noise ratio) relative to state-of-the-art standard 2-D probes. In our particular case, we did not have a second harmonic imaging option on the machine, which appeared on later versions of this machine. Although the protocol we present is affected by the quality of data, we show later that it is robust and can provide good results on data derived using current 3-D probe technology.

Fourteen patients [12 males, two females, age range 39–71, average age 58 ($\sigma = 10$)] with suspected coronary artery disease and with good acoustic windows when imaged using standard 2-D techniques were invited to return to participate in a study of 3-D echocardiography at the John Radcliffe Hospital, Oxford, U.K. These patients had already undergone a routine MUGA within the previous month (average two weeks). Local ethic approval was obtained for this study. Nine patient datasets were selected from these, which had the best image quality, as assessed by the clinician when looking at the sequence as a whole in 3-D.

C. Phase-Based Image Feature Extraction

A phase-based feature detection method was used to find candidate endocardial border points [6]. Briefly, the idea is to detect endocardial border points according to their phase signature (edge shape) rather than intensity gradient information. The reason for this choice is that the acoustic reflection from the endocardial border varies according to the relative angle between the boundary and the transducer. The net effect is to produce a border with variable intensity contrast around its length. This makes it difficult to detect the endocardial border using an intensity-gradient-based approach. This problem is even more of an issue in 3-D echocardiography, since some 2-D acquisition planes might be nonoptimal because they are determined automatically once the first plane position has been chosen. In the current implementation, we used a single scale version of the feature detection method and feature asymmetry as the feature measure. The method is described in [6], and the performance of the technique is tested in detail on clinical data in [20] within the context of 2-D echocardiographic image tracking.

D. Cluster-Based LV Region Extraction

To reliably perform LV surface fitting and tracking without human intervention, boundary feature points need to be present at reasonable intervals all over the LV surface. Phase-based segmentation provides well-localized features but also detects a large number of nonendocardial boundary points in practice. In order to select only points close to the endocardial surface, we have developed a novel cluster-based method for identifying the LV region. The identified region is used both in the definition of a mask to select good candidate endocardial feature points and to initialize the surface fitting method described in Section II-E and -F, respectively.

Unlike the phase-based method, the fuzzy-clustering-based algorithm does not rely exclusively on the local differential structure of the image but takes into account its global intensity distribution and a model of the LV that incorporates a priori knowledge. In this way, a continuous approximation of the LV cavity boundary is provided even in image regions with low contrast and low signal-to-noise ratio, such as where the LV wall is parallel to the insonification beam. This also diminishes the effect of outliers detected by the phase-based method that correspond to noise and other anatomical structures.

The new method is based on standard fuzzy clustering but works iteratively at different levels of a scale space (the scale space is created using anisotropic diffusion and the fuzzy clustering itself). A model of the objects of interest in the image (in this case the LV) is used as a constraint, i.e., to incorporate useful a priori information. As the new method, which is based on our prior work [21], [22], has not appeared in the literature before, the theory underlying the method is described next. We then discuss how it is applied in our specific application.

1) Theory: Briefly, we use a generalization of Bezdek's fuzzy c-means clustering algorithm [23] to provide a fuzzy partition of a possibly multivalued image $\mathbf{F} = \mathbf{F}(\mathbf{p})$ in which every pixel $\mathbf{p}$ of a P-dimensional image domain D is classified according to the attribute vectors specified by the feature space function $\mathbf{S}$. In general, the feature space $\mathbf{S}$ could be a function of position and values of the multidimensional image and of any other higher level interpretations of it (texture, for instance), or simply be a simple scalar function, as in the common case of classification based exclusively on image intensity.

Based on information provided by the feature space function $\mathbf{S}$, the algorithm assigns each pixel a membership grade to each of the Q regions Q_q in which the image is assumed to be divided. The membership function $\mathbf{U}(\mathbf{p})$ specifies how strongly $\mathbf{p}$ is correlated with each region Q_q of the image partition. In fact, the membership function $\mathbf{U}(\mathbf{p}) = (U_1(\mathbf{p}), \ldots, U_q(\mathbf{p}), \ldots, U_Q(\mathbf{p}))$, where $\mathbf{U} : \Re^P \rightarrow \Re^Q$ and $U_q : \Re^P \rightarrow \Re^+$, is computed using the feature space function $\mathbf{S}$ and is therefore, in a strict sense, also a function of $\mathbf{F}$. However, since its final result is to assign a vector value to each pixel $\mathbf{p}$ of the image domain, we write it as a function exclusively of $\mathbf{p}$.

To create a fuzzy partition, $\mathbf{U}$ must satisfy the following constraints:

$$\sum_{q \in \{1, \ldots, Q\}} U_q(\mathbf{p}) = 1 \quad \forall \mathbf{p} \in D$$

$$\sum_{\mathbf{p} \in D} U_q(\mathbf{p}) > 0 \quad \forall q \in \{1, \ldots, Q\}$$

$$U_q(\mathbf{p}) \in [0,1] \quad \forall \mathbf{p} \in D, \forall q \in \{1, \ldots, Q\}. \quad (1)$$

A set of membership functions that satisfy these constraints is given by

$$U_q(\mathbf{p}) = \left(\sum_{j=1}^{Q} \left(\frac{d_q(\mathbf{p})}{d_j(\mathbf{p})} \right)^{2/m-1} \right)^{-1}$$
$$\forall \mathbf{p} \in D, \forall q \in \{1, \ldots, Q\} \quad (2)$$

where $d_q(\mathbf{p}) = \|\mathbf{S}(\mathbf{F}, \mathbf{p}) - \mathbf{K}_q\|$ is the distance in the feature space representation between the pixel $\mathbf{p}$ and the center $\mathbf{K}_q$ of the region (or cluster) Q_q and m is a real number greater than one that controls the overlap between different fuzzy regions. The center of each cluster is computed using the expression

$$\mathbf{K}_q = \frac{\sum_{\mathbf{p} \in D} (U_q(\mathbf{p}))^m \mathbf{S}(\mathbf{F}, \mathbf{p})}{\sum_{\mathbf{p} \in D} (U_q(\mathbf{p}))^m} \quad \forall q \in \{1, \ldots, Q\}. \quad (3)$$

The fuzzy c-means algorithm tries to group the pixel attributes by using an iterative process to approach a local minima of the function

$$J_m(\mathbf{U}, \mathbf{K}) = \sum_{\mathbf{p} \in D} \sum_{q \in \{1, \ldots, Q\}} (U_q(\mathbf{p}))^m (d_q)^2. \quad (4)$$

Equation (4) measures the similarity between the attribute vectors and the cluster centers of each region. J_m is minimized by iterating (4) and monitoring the norm of $\mathbf{U}$ until its change between iteration goes below a prespecified threshold value (the cluster centers and membership functions are recomputed every iteration). We use $m = 2$, and the algorithm is initialized distributing the cluster centers $\mathbf{K}_q$ homogeneously over the feature space (by dividing the dynamic range of each of the feature space axes into equal-sized segments over the identity

line). To measure the probability that two neighboring pixels belong to the same cluster, we need to define a similarity measure $\mathbf{E}$ between the fuzzy cluster membership functions of two given pixels. We use the maximum component difference with an extra component [24], as defined by

$$
\begin{aligned}
\mathbf{E} =&\, \mathbf{E}\left(\mathbf{U}(\mathbf{p}_1), \mathbf{U}(\mathbf{p}_2)\right) \\
=&\, 1 - \left[\, \left|U_{q_{(\max P_1)}}(\mathbf{p}_1) - U_{q_{(\max P_1)}}(\mathbf{p}_2)\right| \right. \\
&\, + \left|U_{q_{(\max P_2)}}(\mathbf{p}_1) - U_{q_{(\max P_2)}}(\mathbf{p}_2)\right| + 1 \\
&\, \left. - \frac{\delta^{q_{(\max P_1)}, q_{(\max P_2)}}}{\left(3 - \delta^{q_{(\max P_1)}, q_{(\max P_2)}}\right)}\,\right]
\end{aligned}
\tag{5}
$$

where $q = q_{(\max P_n)}$ is the cluster membership index that determines the largest cluster membership of $\mathbf{p}_n$ (i.e., maximizes the function $U_q(\mathbf{p}_n)$) and $\delta^{i,j}$ is the Kronecker delta function.

2) The Spatiotemporal Model: To guide the clustering algorithm without overconstraining the segmentation, we introduce model-based knowledge in the form of a simple approximate geometric and dynamic model of the LV.

Specifically, we use a general a priori model derived from the following simple approximations:

1) an approximately sinusoidal contraction of the LV during a heart beat, which represents the *dynamics* of the model;
2) an approximately ellipsoidal shape of the LV cavity as the 3-D *geometric* model of the heart.

The *geometric* parameters of the model are those related to the ellipsoidal shape of the LV cavity. These, together with the dynamic part of the model, can be written as the *transformation of coordinates*

$$
\mathbf{p}' = \mathbf{T}(\mathbf{p})
$$

where $\mathbf{T}$ is defined by

$$
\begin{aligned}
x' =&\, A(t)\left[(x - o_{x\phi})\cos(\theta_\phi) - (y - o_{y\phi})\sin(\theta_\phi)\right] \\
y' =&\, A(t)\left[(x - o_{x\phi})\sin(\theta_\phi) + (y - o_{y\phi})\cos(\theta_\phi)\right]e_{\mathrm{ell}} \\
t' =&\, t\left(1 + k_1\left|\frac{d}{dt}A(t)\right|\right).
\end{aligned}
\tag{6}
$$

Here $A(t)$ describes the sinusoidal heart contraction where the sampling points in time (t) are fixed by the frame rate of the ultrasound machine. $o_{x\phi}$, $o_{y\phi}$, and θ_ϕ are projections onto the particular slice of the ellipsoid center and the angle of its main axis. e_{ell} is the eccentricity of the projected ellipsoid and k_1 is a scaling constant.

Since $\mathbf{T}$ is only a function of the coordinates and not of $\mathbf{F}$, we write the feature space function $\mathbf{S}$ as the composition of functions $(\mathbf{S}' \circ \mathbf{T})$

$$
\mathbf{S}(\mathbf{F}(\mathbf{p}), \mathbf{p}) = (\mathbf{S}' \circ \mathbf{T})(\mathbf{F}(\mathbf{p}), \mathbf{p}) = \mathbf{S}'(\mathbf{F}(\mathbf{p}'), \mathbf{p}')
$$

where $\mathbf{S}'$ can further be decomposed into two components $\mathbf{S}' = (S_1', S_2')$ such that

$$
\begin{aligned}
S_1' =&\, S_1'(\mathbf{F}(\mathbf{p}')) = w_1 I(\mathbf{p}') = w_1 I(x', y', t') \\
S_2' =&\, S_2'(\mathbf{p}') = w_2 R(\mathbf{p}') = w_2 (x'^2 + y'^2)^{1/2}
\end{aligned}
$$

with the weights for each function defined as

$$
w_1 = \alpha/I_{\max}, \quad w_2 = (1 - \alpha)/R_{\max}, \quad \alpha \in [1].
$$

Here $I_{\max}$ and $R_{\max}$ are the maximum intensity and radial position computed from each image and used to normalize the feature space and α used to weight intensity versus geometry.

Although varying α can improve the segmentation results, we found that a value of $\alpha = 0.5$ gave good results in general, and we used this value in the results reported here. In our study, we set the ellipse minor axis to be two-thirds the size of the major axis ($e_{\mathrm{ell}} = 2/3$). The orientation and center of the ellipse were set from the image coordinate system: the major axis in the vertical direction ($\theta_\phi = 0$) was aligned with the ultrasound probe as using a standard 3-D echocardiography imaging protocol for the apical view; one tries to center the LV in the field of view. The center (x_o, y_o, z_o) was set to the origin of coordinates, first placed in the geometric center of the image and thereafter computed from the clustering results as the center of the cluster corresponding to the LV.

3) Model-Based Clustering: A simple and computationally efficient way to introduce a spatiotemporal model into the fuzzy-clustering equations is via the feature space. The image attributes of the feature space used are intensity and position in a cylindrical coordinate system, which is a natural choice for the 2-D + T LV long-axis images.

The LV model is introduced using $\mathbf{T}$ to deform the feature space, stretching the isocontours of the 2-D + T space to embrace ellipses of different sizes in the 2-D space dimensions and sinusoidal lines instead of straight lines in the time dimension. The origin of the cylindrical coordinate system is first placed in the center of the image, and then the position is refined by computing the center of mass of the LV cluster. The cluster corresponding to the LV cavity is automatically identified as the one with the lowest intensity and position closest to the center of the image. In case some pixels belonging to this cluster are scattered around the image (which rarely happens after the image has been smoothed, as described in the next section), the largest connected component is computed to select only the cluster points belonging to the LV cavity.

There are two parameters in the clustering method: the number of clusters and the geometry-intensity weight. Experimental evaluation has shown that the algorithm is insensitive to either parameter. For all nine datasets (108 2-D + T images), using either five or six clusters and $\alpha = 0.5$ gave satisfactory results. We found that small variations to these values slightly changed how close the cluster approximated the LV cavity, but did not, for instance, fuse the cavity and the background clusters. In our application, a cluster that approximates the LV cavity is sufficient. If a precise identification of the LV is required, these parameters could be useful for weighting the belief in the geometry versus intensity in images with poor contrast or boundary definition.

4) Model-Based Multiscale Fuzzy Clustering: To overcome the problematic effect of intensity fluctuations of the noisy ultrasound images, the clustering process is performed at different levels of resolution in a scale space of the image [25].

The scale space is generated using a knowledge-based anisotropic diffusion (KBAD) algorithm [22], where the conductance term of the diffusion process is a tensor and an explicit function of the position $\mathbf{p}$ (in 2-D + T in our image domain), the image intensity, and its gradient, i.e., $\mathbf{C} = \mathbf{C}(\mathbf{p}, I, \nabla I)$.

This scheme also introduces a probabilistic measure of the image intensity distribution [26].

The combined diffusion-clustering algorithm is implemented as follows. The first clustering is done after some iterations of the diffusion. The clustering is then updated at regular intervals during the diffusion process (i.e., at different levels of the scale space), and the initially coarse segmentation of the image is gradually improved until it converges to a meaningful region partition, as the smoothing action of the diffusion process reduces noise. The computational expense of repeating the clustering at different scales of resolution is not high because energy minimization is faster in the lower dynamic range of the smoothed image. Since we are only interested in obtaining an approximate clustering segmentation, the process is performed on subsampled images (reduced by a factor of eight), making processing time shorter (a 2-D $+ T$ image can be processed in under 30 s on an SGI O2 workstation).

E. Filtering

A geometric filtering method is used to extract a reduced set of candidate boundary points from the complete set of feature asymmetry boundary points to be used in the surface fitting method discussed in Section II-F. Geometric constraints are used to discard points that clearly do not belong to the endocardial surface—for example, spurious responses in the LV cavity, features on the epicardium, or right ventricle features.

We assume that the left ventricle forms a closed surface, i.e., ignore the effect of the mitral valve. We also assume that the LV cavity appears mainly dark, while the myocardium appears mostly light (this follows from the appearance of ultrasound images). A filter based on both the position of a boundary point and the direction of its normal is defined as follows.

For each volume corresponding to an instance in time, an approximation to the center-of-gravity and major axis of the LV cavity is known from the fuzzy-clustering algorithm (Section II-D). Now, consider a point $\mathbf{x_i}$ that is a member of the set of candidate boundary points, and consider two lines

$$\mathbf{g}(\kappa) = \mathbf{x_i} + \kappa \mathbf{n_i} \tag{7}$$
$$\mathbf{h}(\lambda) = \mathbf{x_0} + \lambda \mathbf{n_0} \tag{8}$$

where κ and λ are scalars. In (7), $\mathbf{g}(\kappa)$ is the line passing through the point $\mathbf{x_i}$ in the direction of the normal of the point $\mathbf{n_i}$. In (8), the line $\mathbf{h}(\lambda)$ is the major axis of the ventricle, defined by the center of the LV $\mathbf{x_0}$ and the direction of its major axis $\mathbf{n_0}$.

Using simple vector geometry, the points of closest approach on the two lines $\mathbf{G} = \mathbf{x_i} + \kappa_G \mathbf{n_i}$ and $\mathbf{H} = \mathbf{x_0} + \lambda_H \mathbf{n_0}$ are defined by particular values of the scalar constants $\kappa = \kappa_G$ and $\lambda = \lambda_H$, respectively, where

$$\kappa_G = \mathbf{n_0} \cdot \frac{(\mathbf{x_0} - \mathbf{x_i}) \wedge (\mathbf{n_0} \wedge \mathbf{n_i})}{\|\mathbf{n_0} \wedge \mathbf{n_i}\|^2} \tag{9}$$

$$\lambda_H = \mathbf{n_i} \cdot \frac{(\mathbf{x_0} - \mathbf{x_i}) \wedge (\mathbf{n_0} \wedge \mathbf{n_i})}{\|\mathbf{n_0} \wedge \mathbf{n_i}\|^2}. \tag{10}$$

Here, $\wedge$ denotes the cross product. The values of κ_G and λ_H are used to filter out image feature points in the following way. First, points with normals pointing into the left ventricle are selected. Secondly, from this subset, points that result in the point $\mathbf{H}$ lying

within the ventricle are selected. Thirdly, points are filtered out based on their distance from the major axis by placing upper and lower bounds on acceptable values of κ_G. Mathematically, these three constraints can be encapsulated as

$$D_{\min} \leq \kappa_G \leq D_{\max} \tag{11}$$
$$|\lambda_H| \leq \Lambda \tag{12}$$

where Λ is related to the length of the ventricle, while $D_{\min}$ and $D_{\max}$ are related to the width of the ventricle. These values are estimated from the LV cluster for a given time frame and vary over the cardiac cycle, depending on the length and width of the cluster; hence the region of search contracts with the cluster.

A different filtering system is employed at the apex and base of the ventricle (the end zones). This is based on experimental experience, which showed that too many points were being filtered out in these areas. To solve this, we define a cone aligned with the major axis with its apex at the center-of-gravity of the LV cavity and a user-defined cone-angle θ. Within this cone image, features are accepted if either they satisfy the above criterion [(11) and (12)] or if the following two conditions are met. Let $\mathbf{XC}$ be the vector from the image feature point to the center-of-gravity and $\widehat{\mathbf{XC}}$ be the normalized version of this vector. A point is assumed to be in an end zone if $\widehat{\mathbf{XC}} \cdot \mathbf{n_0} \geq \cos(\theta)$ (recall that $\mathbf{n_0}$ is the major axis direction vector). Then it follows that

$$D_{\min} \leq \|\mathbf{XC}\| \leq D_{\max} \tag{13}$$
$$\widehat{\mathbf{XC} \cdot \mathbf{n_i}} \geq \cos\theta. \tag{14}$$

F. Surface Fitting and Motion Tracking

The filtered phase-based feature points are reconstructed into 3-D for each instance in time (i.e., slices are placed coaxially). In addition, at each time frame, the LV cluster boundary is reconstructed in 3-D in the same manner and a mesh surface fit to the cluster boundary points. The end-diastolic cluster-based surface is used as the surface initialization for fitting a surface to the 3-D set of phase-based feature points of the first volume. The LV cavity shape is then sequentially estimated by using the surface fit of the previous frame to initialize the fit in the current frame using a variant of the method described in [27]. Namely, at each time frame, the distance between a point on the model mesh surface $\mathbf{M_i}$ and a data point $\mathbf{X_i}$ is computed. This is calculated as

$$d_{\mathbf{M_i}} = \min_{\forall \mathbf{X_j}} \left(d_{\mathbf{M_i X_j}} \right)$$

where

$$d_{\mathbf{M_i X_j}} = \|\mathbf{M_i} - \mathbf{X_j}\|^2 + \beta \|\mathbf{n_{M_i}} - \mathbf{n_{X_j}}\|^2.$$

($\beta = 0.5$ was used in the examples shown in this paper.) The surface S is then deformed—either globally, by translation, rigid, or affine; or locally, using a B-spline transformation to minimize the cost function $\sum_{\forall i} d_{\mathbf{M_i X_i}}$ subject to certain constraints. Computation of the local B-spline transformation is iterated until convergence is achieved. Details of the subset selection and minimization process are outlined in [27].
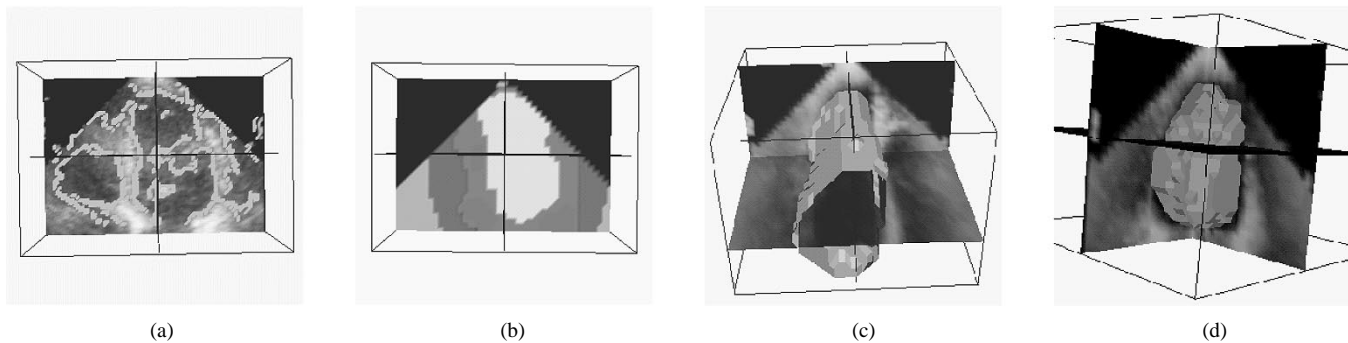
Fig. 2.   (a) Example of low-quality ultrasound image with phase-based edges overlaid. (b) Clustering results of the same image. (c) and (d) show the cluster superimposed on the original image in 2-D $+$ T and in 3-D.
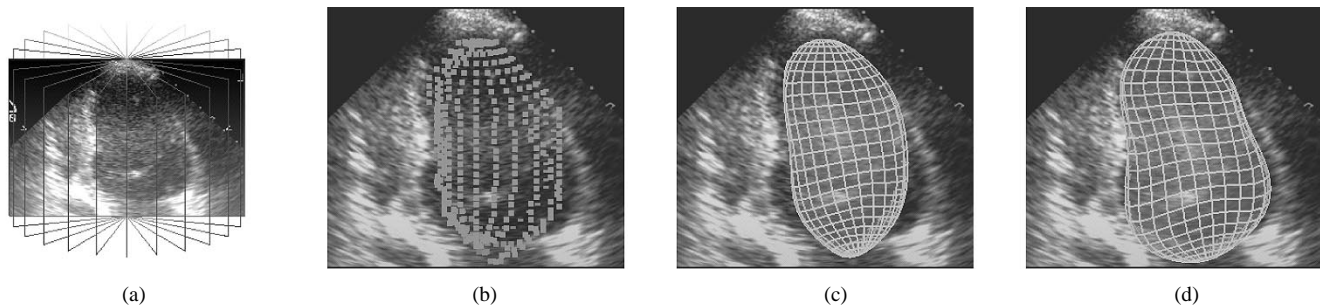

Fig. 3.   Steps in surface fitting: (a) 3-D dataset with 12 planes; (b) 3-D reconstructed boundary of the LV cavity cluster; (c) cluster's ellipsoid of inertia deformed onto the cluster; and (d) surface from (c) deformed onto phase-based image feature points.

G. Illustration of Processing Steps

Fig. 2 illustrates steps in the clustering algorithm. Fig. 2(a) shows a typical phase-based feature detection result on a low-quality image. Fig. 2(b) shows the resulting clusters within one frame of one slice; the large central cluster corresponds to the left ventricle cavity. The boundary of this cluster was extracted in each slice over all time frames and smoothed to give a continuous 2-D$+T$ contour [Fig. 2(c)]. It was then combined with the contours from other slices to create a three-dimensional contour at each time frame [Fig. 2(d)].

Fig. 3 illustrates surface fitting. The distribution of image planes in 3-D is shown in Fig. 3(a). Boundary points on a cluster are shown reconstructed in 3-D in Fig. 3(b). The surface fit to these points, i.e., the fitting surface initialization, is shown in Fig. 3(c). Fig. 3(d) shows the result after deforming the surface shown in (c) to the phase-based image feature points.

H. Comparing User-Guided and Fully Automatic Processing

Three processing protocols were compared on each of the nine data sets.

1) Manual segmentation performed by an expert.
2) Automatic tracking using an ellipsoid initial surface manually placed by an expert with the help of a 3-D graphic interface and without geometric filtering of image features. Below, we refer to this method as the user-guided method.
3) The new method including geometric filtering of image features and automatic surface fitting, as described in Section II-A–II-F. We refer to this method as the fully automatic processing method below.

TABLE I
AVERAGE AND STANDARD DEVIATION OF ERROR IN VOLUMES (IN cm^3)
OVER ALL PATIENTS AND TIME FRAMES (202 VOLUMES IN TOTAL)
WITH RESPECT TO MANUAL DELINEATIONS

Method	μ	(σ)	RMSE
User-guided	37.82	(37.92)	53.49
Fully Auto. Processing	6.63	(37.07)	37.57

Volume-time plots and EFs were calculated for each method. Errors in volumes were calculated between the manually derived surfaces and the user-guided method and the fully automatic processing method, respectively.

III. RESULTS AND DISCUSSION

Table I compares the average and standard deviation of errors in volume over all data sets. It can be seen that compared to the user-guided method, the fully automatic processing method gives a closer match to contours drawn by a clinician, although the standard deviation (σ) in the difference between the volumes is still high.

Fig. 4(a) shows the volume–time curve obtained by each method on a typical data set. Note that in this case, the fully automatic processing method and manually tracked volumes give similar results, showing that fitting based on automatic cluster initialization can give as good a result as manual segmentation. However, the overestimation observed here is not a general result. Fig. 4(b) shows, for all volumes, a volume scatter plot of the volumes from both automatic methods plotted against the volume from manually delineated contours. The R^2 values are 0.3898 for the user-guided method and 0.3163 for the fully automatic processing method, i.e., both low.
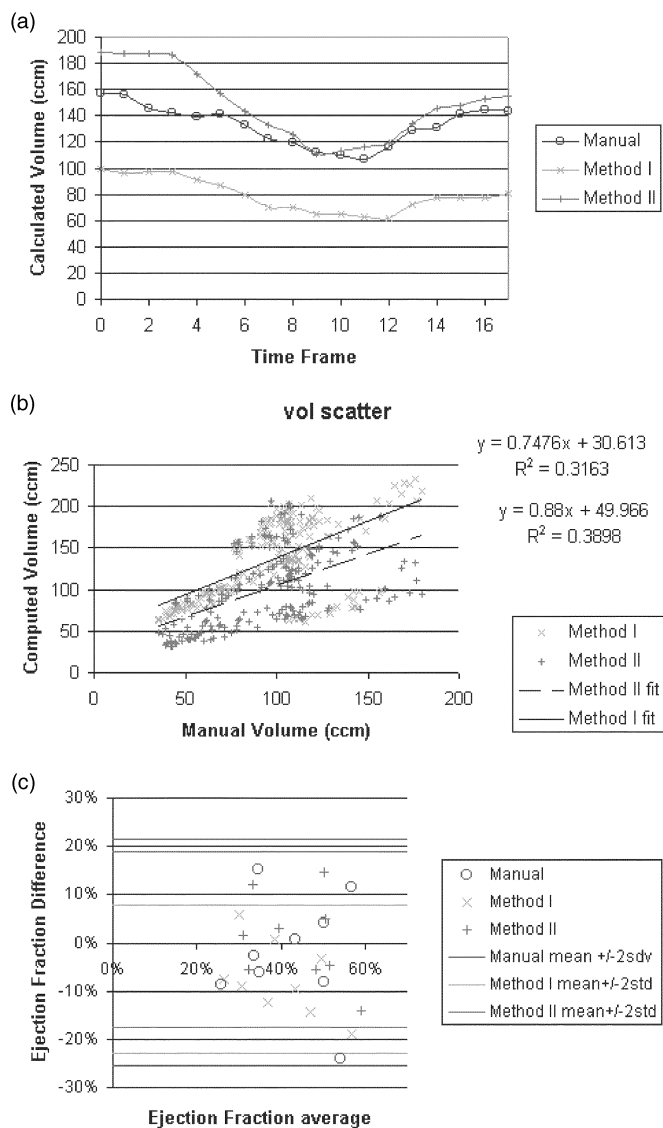
Fig. 4.   (a) A typical volume–time curve computed from the fitted surface. (b) A scatter plot for all volumes, comparing automatically computed volumes to those from manually drawn delineations. (c) A Bland–Altman plot of the automatically calculated and MUGA EFs.

The ejection fractions calculated using the three methods were also compared with those obtained from a clinical MUGA study performed on the same patients. Table II summarizes these results where the data sets, ranked by image quality, are good (1–6,9) and medium (7,8). The results show that the fully automatic processing method has a lower mean error (0.70%) than both manual contours (−1.98%) and the user-guided method (−7.61%). The standard deviation (9.14%) in the error for the fully automatic processing method is slightly higher than for the user-guided method (7.68%), but with a small sample set, it is difficult to draw strong conclusions from this. For data sets of good quality (1–6), the fully automatic processing method calculates ejection fractions within six percentage points of the MUGA value.[1] The image quality of data sets 7 and 8 was not as good, and manual delineation

[1]The literature [15] states that five percentage points is the clinical reproducablity of MUGA.

was more difficult. Hence, the estimates of EF for both manual and automatic methods were not as good. Data set 9 was rated a good quality dataset. This allowed for a good manual segmentation and a good estimate of ejection fraction by the manual protocol. However, the papillary muscles were very evident in this data set, which confused the automatic methods, leading to a poor volume estimate.

The mean and standard deviation for the good quality cases (i.e., omitting data sets 7, 8, and 9) are shown in the last row of Table II. These show that on good quality images, the fully automatic processing method provides a more reliable estimate of EF than the user-guided method and a similarly good estimate to the manual protocol.

Fig. 4(c) shows a Bland–Altman plot of calculated ejection fractions from manual contours and the user-guided and fully automatic processing methods when compared to ejection fractions from MUGA. Note that the manual method has the largest 2σ range and shows a slightly negative mean ($\mu = -1.98$). The user-guided method shows a smaller 2σ range and a definite negative bias ($\mu = -7.61$). On the other hand, the fully automatic processing method has a lower 2σ range than the manual result and a low positive bias ($\mu = 0.70$). Although the small number of data sets analyzed means that strong statistically based conclusions cannot be drawn from this part of the work, the results indicate that the fully automatic processing method is consistently comparable in performance to the manual and user-guided methods.

IV. CONCLUSION AND FUTURE WORK

In this paper, we have described an automatic method of tracking endocardial boundaries from 3-D transthoracic ultrasound data. We first used multiscale clustering to provide an approximate segmentation of the LV cavity. This estimated shape was used to initialize a 3-D fitting and tracking algorithm, which tracked candidate boundaries derived from a feature asymmetry-based segmentation method followed by a geometric filtering postprocessing step. The combination of these three techniques overcomes the sensitivity that tracking has to the initial shape by using local and global information from the 3-D image sequences.

On good quality data sets, we obtained similar tracking results for a manually traced surface and the fully automatic method. For these data sets, ejection fractions were within six percentage points of the value obtained from MUGA. Although our results are preliminary and further validation is required, from this part of the work, we draw the conclusion that the new method is suitable for clinical research use. Performance was less good on low-quality data, which shows that further work is needed to improve the methods for routine clinical use.

There are a number of ways in which the current volume estimation method could be improved. The low quality of transthoracic image sequences today is a major issue. We have been investigating image-processing-based methods of enhancing the quality of $2\text{-D} + T$ images to improve feature detection [28]. Future work will also be based on using MRI as the gold standard, which is more accurate than SPECT MUGA. This was not available at our institution at the time of this research. Finally,

TABLE II
EJECTION FRACTIONS COMPUTED FOR THE NINE PATIENTS USING MANUAL CONTOURS,
THE USER-GUIDED METHOD, AND THE FULLY AUTOMATIC PROCESSING METHOD

PATIENT	MUGA	Ejection Fraction as a percentage (Difference to MUGA (%))					
		Manual		User-guided		Fully automatic processing	
1	38.00	31.96	(-6.04)	38.62	(0.62)	40.88	(2.88)
2	35.00	32.28	(-2.72)	26.07	(-8.93)	29.55	(-5.45)
3	51.00	62.48	(11.48)	47.87	(-3.13)	45.46	(-5.54)
4	48.00	52.16	(4.16)	38.48	(-9.52)	52.97	(4.97)
5	30.00	21.46	(-8.54)	22.54	(-7.46)	31.64	(1.64)
6	54.00	46.02	(-7.98)	39.51	(-14.49)	49.19	(-4.81)
7	27.00	42.17	(15.17)	32.78	(5.78)	39.10	(12.10)
8	66.00	42.07	(-23.93)	47.04	(-18.96)	52.00	(-14.00)
9	43.00	43.60	(0.60)	30.64	(-12.36)	57.54	(14.54)
All	Mean (Std)	-1.98	(11.71)	-7.61	(7.68)	0.70	(9.14)
Best 6	Mean (Std)	1.61	(7.92)	-7.15	(5.28)	1.05	(4.74)

the approach we have developed could be changed with minor modification to work on real-time 3-D echocardiography.

REFERENCES

[1] X. Papadimetris, A. Sinusas, D. Dione, and J. Duncan, "3D cardiac deformation from ultrasound images," in *MICCAI*, Cambridge, U.K., October 1999, pp. 420–429.

[2] F. Sheehan, E. Bolson, R. Martin, G. Basein, and J. McDonald, "Quantitative three-dimensional echocardiography: Methodology, validation and clinical applications," in *MICCAI*, Boston, MA, Oct. 1998, pp. 102–109.

[3] G. I. Sanchez-Ortiz, J. A. Noble, G. J. T. Wright, J. Feldmar, and M. Mulet-Parada, "Automated LV motion analysis from 3D echocardiography," presented at the MIUA, Oxford, U.K., July 19–20, 1999.

[4] A. Giachetti, "On-line analysis of echocardiographic image sequences," *Med. Image Anal.*, vol. 2, no. 3, pp. 261–284, 1998.

[5] G. Jacob, A. Noble, M. Mulet-Parada, and A. Blake, "Evaluating a robust contour tracker on echocardiographic sequences," *Med. Image Anal.*, vol. 3, no. 1, pp. 63–76, 1999.

[6] M. Mulet-Parada and J. A. Noble, "2d+t boundary detection in echocardiography," *Med. Image Anal.*, vol. 4, no. 1, pp. 21–30, 2000.

[7] A. F. Frangi, W. J. Niessen, and M. A. Viergever, "Three-dimensional modeling for functional analysis of cardiac images: A review," *IEEE Trans. Med. Imag.*, vol. 20, pp. 2–25, Jan. 2001.

[8] T. Gustavsson, R. Pascher, and K. Caidahl, "Model based dynamic 3D reconstruction and display of the left ventricle from 2D cross-sectional echocardiograms," *Comput. Med. Graph.*, vol. 17, no. 4–5, pp. 273–278, July 1993.

[9] A. S. Gobal, D. L. King, L. M. Katz, Boxt, D. L. King Jr., and M. Y. Shao, "Three-dimensional echocardiographic volume computation by polyhedral surface reconstruction: In virto validation and comparison to magnetic resonance imaging," *J. Amer. Soc. Echocardiogr.*, vol. 5, no. 2, pp. 115–124, Apr. 1992.

[10] M. E. Legget, D. F. Leottta, 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 magnetic field position and orientation sensing system," *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 495–504, Apr. 1998.

[11] G. Coppini, R. Poli, and G. Valli, "Recovery of the 3-D shape of the left ventricle from echocardiographic images," *IEEE Trans. Med. Imag.*, vol. 14, pp. 301–317, June 1995.

[12] E. A. Geiser, M. Ariet, D. A. Conetta, S. M. Lupkiewicz, L. G. Christie, and C. R. Conti, "Dynamic three-dimensional echocardiographic reconstruction of the intact human left ventricle: Technique and initial observations in patients," *Amer. Heart J.*, vol. 103, pp. 1056–1065, Sept. 1982.

[13] Y.-H. Tseng, J.-N. Hwang, and F. H. Sheehan, "3-d heart modeling and motion estimation based on continuous distance transform neural networks and affine transform," *J. VLSI Signal Process. Syst. Signal, Imag., Video Technol.*, vol. 18, no. 3, pp. 207–218, Apr. 1998.

[14] E. A. Geiser, S. M. Lupkiewicz, L. G. Christie, M. Ariet, D. A. Conetta, and C. R. Conti, "A framework for three-dimensional time-varying reconstruction of the human left ventricle: Sources of error and estimation of their magnitude," *Comput. Biomed. Res.*, vol. 13, no. 3, pp. 225–241, June 1980.

[15] F. J. T. H. Wackers, H. J. Berger, D. E. Johnstone, L. Goldman, R. A. Reduto, A. Gottschalk, and B. L. Zaret, "Multiple gated cardiac blood pool imaging for left ventricular ejection fraction: Validation of the technique and estimation of variability," *Amer. J. Cardiol.*, vol. 43, pp. 1159–1166, June 1979.

[16] G. Jacob, J. A. Noble, A. D. Kellion, and A. P. Banning, "Quantitative regional analysis of myocardial wall motion," *Ultrasound Med. Biol.*, vol. 27, no. 6, pp. 773–784, 2001.

[17] G. Xiao, J. M. Brady, J. A. Noble, M. Burcher, and R. English, "Non rigid registration on 3D free-hand ultrasound imaging of the breast," *IEEE Trans. Med. Imag.*, vol. 21, pp. 405–412, Apr. 2002.

[18] X. Ye, J. A. Noble, and D. Atkinson, "3D freehand echocardiography for automatic left ventricle reconstruction and analysis based on multiple acoustic windows," *IEEE Trans. Med. Imag.*, vol. 21, pp. 1051–1058, Sept. 2002.

[19] G. I. Sanchez-Ortiz, J. Declerck, J. A. Noble, and M. Mulet-Parada, "Automating 3D echocardiographic image analysis," in *Proc. MICCAI*, Lecture Notes in Computer Science, Pittsburgh, PA, Oct. 2000, pp. 687–696.

[20] M. Mulet-Parada, "Intensity independent feature extraction and tracking in echoc ardiographic sequences," Ph.D. dissertation, Univ. Oxford, Oxford, U.K., 2000.

[21] G. I. Sanchez-Ortiz, "Fuzzy clustering driven aniotropic diffusion: Enhancement and segmentation of cardiac mr images," in *IEEE Nuclear Science Symp. Medical Imaging Conf.*, vol. 3, Toronto, ON, Canada, Nov. 1998, pp. 1873–1875.

[22] G. I. Sanchez-Ortiz, D. Rueckert, and P. Burger, "Knowledge-based tensor anisotropic diffusion of cardiac MR images," *Med. Image Anal.*, vol. 3, no. 1, pp. 77–101, 1999.

[23] J. C. Bezdek and P. F. Castelaz, "Prototype classification and feature selection with fuzzy sets," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-7, pp. 87–92, 1977.

[24] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[25] B. M. ter Haar Romeny, "Computational Imaging and Vision," in *Geometry-Driven Diffusion in Computer Vision*, B. M. ter Haar Romeny, Ed. Norwell, MA: Kluwer Academic, 1994.

[26] S. R. Arridge and A. Simmons, "Application of multi-spectral probabilistic diffusion to dual echo MRI," presented at the Medical Image Understanding and Analysis (MIUA'98), Leeds, U.K., 1998.

[27] J. Declerck, J. Feldmar, M. L. Goris, and F. Betting, "Automatic registration and alignment on a template of cardiac stress and rest reoriented SPECT images," *IEEE Trans. Med. Imag.*, vol. 16, no. 1, pp. 1–11, 1997.

[28] D. Boukerroui, J. A. Noble, and J. M. Brady, "Feature enhancement in low quality images with application to echocardiography," *Ultrasound Med. Biol.*, vol. 27, no. 12, pp. 1583–1594, 2001.