A Multiple Active Contour Model for Cardiac Boundary Detection on Echocardiographic Sequences

Vikram Chalana, David T. Linker, *Member, IEEE*, David R. Haynor, and Yongmin Kim,* *Fellow, IEEE*

Abstract—**Tracing of left-ventricular epicardial and endocardial borders on echocardiographic sequences is essential for quantification of cardiac function. We designed a method based on an extension of active contour models to detect both epicardial and endocardial borders on short-axis cardiac sequences spanning the entire cardiac cycle. We validated the results by comparing the computer-generated boundaries to the boundaries manually outlined by four expert observers on 44 clinical data sets. The mean boundary distance between the computer-generated boundaries and the manually outlined boundaries was 2.80 mm ($\sigma = 1.28$ mm) for the epicardium and 3.61 ($\sigma = 1.68$ mm) for the endocardium. These distances were comparable to interobserver distances, which had a mean of 3.79 mm ($\sigma = 1.53$ mm) for epicardial borders and 2.67 mm ($\sigma = 0.88$ mm) for endocardial borders. The correlation coefficient between the areas enclosed by the computer-generated boundaries and the average manually outlined boundaries was 0.95 for epicardium and 0.91 for endocardium. The algorithm is fairly insensitive to the choice of the initial curve. Thus, we have developed an effective and robust algorithm to extract left-ventricular boundaries from echocardiographic sequences.**

I. Introduction

TRACING of the epicardial and endocardial boundaries of the left ventricle (LV) on echocardiographic images is essential for quantification of cardiac function [1]. Tracing the borders on the end-diastolic (ED) and end-systolic (ES) images allows the computation of clinically important measures such as ejection fraction and regional wall thickening. Tracing the borders on a series of images covering the entire cardiac cycle allows quantitative interpretation of LV contraction dynamics [2]. Manual tracing of these borders is a time consuming and tedious task. Moreover, the resulting outlines vary between different observers and may suffer from a subjective bias [3]. Hence, there is a clinical need for automatically detecting these borders.

Several researchers have attempted to identify the LV boundaries on echocardiograms, automatically or semi-automatically, with varying degrees of success. Among the most important reasons for the limited clinical applicability of these methods are the following.

- Most methods rely on using only the image gray scale information to identify the cardiac borders [4]–[7]. Due to poor contrast at the boundaries and low signal-to-noise ratio (SNR), using the image gray scale information alone may not be sufficient and other *a priori* sources of information may be necessary to get useful results [8].
- Most methods have not been evaluated on clinically important data sets. Many researchers have evaluated their algorithms on images of phantoms [6], *in vitro* hearts [9], live animals under nearly ideal laboratory conditions [6], [9], or on highly selected sets of good quality clinical data [5], [7].
- Most methods require some input from the user, either in the form of values of different algorithm parameters, an initial seed point [5], or an initial curve [4], [7]. No research group has studied the influence of user's input on the performance of their algorithms in detecting cardiac boundaries.

Some research groups have addressed these issues in isolation, but no single group has addressed all the issues. For example, Friedland and Adam [10] and Herlin and Ayache [11] have proposed algorithms which allow integration of multiple sources of information, but it is difficult to judge the effectiveness of their methods due to a lack of clinical validation. Perez *et al.* [12] carried out an evaluation of their algorithm on data from 66 patients; however, due to limitations in their algorithm, they were able to obtain satisfactory results on only 72% of the patients.

To address all of the above issues, we designed an algorithm for detecting the LV borders on echocardiographic images which integrates multiple sources of information. Also, we evaluated this algorithm's performance on a representative clinical data set and studied the sensitivity of the algorithm's output to the user-supplied inputs.

We have previously applied the active contour model for detecting cardiac boundaries in static echocardiograms [13]. However, the algorithm performance was limited due to low SNR which is more apparent in static echocardiograms. Expert observers outlining the cardiac borders often use temporal image sequences to track the borders visually across frames [14]. This observation motivated us to design an algorithm

Manuscript received February 8, 1995; revised February 6, 1996. This work was supported, in part, by Siemens Medical Systems, Ultrasound Group. *Asterisk indicates corresponding author.*

V. Chalana is with the Center for Bioeningeering, University of Washington, Seattle, WA 98195 USA.

D. T. Linker is with the Division of Cardiology, University of Washington, Seattle, WA 98195 USA.

D. R. Haynor is with the Department of Radiology, University of Washington, Seattle, WA 98195 USA.

*Y. Kim is with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195 USA (e-mail: kim@ee.washington.edu).

Publisher Item Identifier S 0278-0062(96)04235-8.

that uses temporal domain information. Our algorithm is based on a temporal domain extension of the two-dimensional (2-D) active contour model [15] to find cardiac boundaries on a sequence of ultrasound images. We call this model the multiple active contour model.

Our algorithm detects both epicardial and endocardial borders on mid-papillary short-axis views of the heart. We evaluated the performance of our algorithm on sequences of echocardiographic images from 44 patients selected without regard to image quality. The computer-generated boundaries were compared against boundaries which had been manually outlined by four experienced observers. This allowed us to compare our algorithm to the interobserver variability. We also studied the sensitivity of the algorithm to the specification of the starting curve by varying the starting curve in a systematic way.

In this paper, we first describe our multiple active contour model in relation to the original active contour model of Kass *et al.* [15]. Next, we describe the application of this model to cardiac boundary detection from echocardiograms. Then we discuss the methods used to compare the results to manually outlined borders. Finally, we present and discuss the results of our algorithm.

II. Methods

A. Multiple Active Contour Model

The original active contour model described by Kass *et al.* [15] refers to a set of points, $\mathbf{v}(s) = (x(s), y(s))$, on an image parameterized with respect to the contour length, s. Each possible configuration of the contour has an energy associated with it which is a combination of internal and external energies. Features of interest on the image are detected by defining the internal and external energies of the contour such that the desired contour has the lowest possible total energy. For example, to detect the boundaries of a smooth object on an image, the internal energy is defined to keep the contour smooth and continuous while the external energy is defined to attract the contour toward the boundary of the object. Thus, Kass *et al.* pose the segmentation problem as an optimization problem which is solved using a gradient descent method. This optimization problem requires an initial approximation to the solution and their algorithm was quite sensitive to the selection of this initial curve.

To use temporal information to detect cardiac boundaries from echocardiograms, we extended the original active contour model to three dimensions. Our model is similar in spirit to the simplified three-dimensional (3-D) active contour model proposed by Cohen and Cohen [16]. We call our model the multiple active contour model. We use the multiple active contour model instead of the 2-D tracking approach described by Herlin and Ayache [11] or by Cohen [17] because individual echocardiographic images often contain missing boundaries and the detection errors using a 2-D model may propagate and increase further in the sequence.

We first define an active surface, $\mathbf{v}(s,r)$, as a function of two parameters, s and r, $\mathbf{v}(s,r) = (x(s,r), y(s,r), z(s,r))$, for $0 \leq s \leq 1$ and $0 \leq r \leq 1$. The multiple active contour model is a special case of the active surface model in which the z component of the active surface depends only on the second parameter, r. In such a case, the surface is represented as a sequence of planar contours, with the second parameter, r, being the index of the plane. As in the original active contour model, the first parameter, s, is the normalized contour length. We may reparameterize $\mathbf{v}$, by replacing $z(s,r)$ with r, to get

$$\mathbf{v}(s,r) = (x(s,r), y(s,r), r) \tag{1}$$

for $0 \leq s \leq 1$, and $0 \leq r \leq 1$. Each contour in the multiple active contour is discretized into an equal number of points, and the multiple active contour is represented by ordered pair of vectors of x and y coordinates, $(\mathbf{x}, \mathbf{y})$.

The internal energy of the multiple active contour model is expressed as

$$\begin{aligned} E_{\text{int}}(\mathbf{v}) = \int_0^1 \int_0^1 \{&(w_1 \|\mathbf{v}_s(s,r)\|^2) + (w_2 \|\mathbf{v}_{ss}(s,r)\|^2) \\ &+ (w_3 \|\mathbf{v}_r(s,r)\|^2) + (w_4 \|\mathbf{v}_{rr}(s,r)\|^2) \\ &+ (w_5 \|\mathbf{v}_{sr}(s,r)\|^2)\} ds dr \end{aligned} \tag{2}$$

where $\mathbf{v}_s$ and $\mathbf{v}_{ss}$ are first and second derivatives with respect to s, and $\mathbf{v}_r$ and $\mathbf{v}_{rr}$ are the first and second derivatives of the surface with respect to r. The parameters w_1 and w_2 are the weights of the in-plane regularization terms, which can also be interpreted as coefficients of elastic energy and bending energy of a contour [15]. The parameters w_3, w_4, and w_5 are the weights for the various between-plane regularization terms. The parameters w_3 and w_4 are coefficients for the elastic energy and bending energy between corresponding pixels in the third dimension, respectively, and the parameter w_5 is the coefficient for energy of the interaction between in-plane and between-plane constraints. In general, these coefficients are functions of the parameters s and r.

The external energy of our multiple active contour is defined to cause the contour to be attracted to the areas of high gradient and to enforce monotonic motion constraints. For the model to be attracted to large image gradients, the energy term is defined as follows:

$$E_{\text{grad}}(\mathbf{v}) = -w_g \int_0^1 \int_0^1 \{G_\sigma * \|\nabla I(x,y,r)\|^2\} ds dr \tag{3}$$

where ∇I is the 2-D gradient of the image intensity, I, G_σ is a 2-D circular Gaussian with standard deviation σ, and the operator, $*$, represents 2-D convolution. The parameter w_g controls the relative importance of image gradient force in the model. The Gaussian convolution of the image gradient ensures that the contour is attracted to the edges from distances comparable to σ.

For the model to be constrained by the monotonic motion assumption, we define an energy term as follows:

$$\begin{aligned} E_{\text{time}}(\mathbf{v}) = w_t \Bigg\{ &\int_{r \in \text{sys.}} \int_0^1 (\mathbf{v}(s,r) - \mathbf{v}(s, r-\epsilon)) ds dr \\ &+ \int_{r \in \text{dia.}} \int_0^1 (\mathbf{v}(s, r-\epsilon) - \mathbf{v}(s,r)) ds dr \Bigg\} \end{aligned} \tag{4}$$

where $\mathbf{v}(s, r-\epsilon)$ is the curve in the sequence immediately preceding the curve, $\mathbf{v}(s,r)$. This constraint is based on the assumption of monotonic motion of the heart in the diastolic and systolic phases of the cardiac cycle, i.e., in the systolic phase, the heart is monotonically decreasing in size, and in the diastolic phase, it is monotonically increasing. In the systolic phase, E_{time} is minimized when $\mathbf{v}(s,r)$ is smaller than $\mathbf{v}(s, r-\epsilon)$, i.e., the heart is decreasing in size, and in the diastolic phase, E_{time} is minimized when $\mathbf{v}(s,r)$ is larger than $\mathbf{v}(s, r-\epsilon)$, i.e., the heart is increasing in size.

The total external energy of the multiple active contour is given as the sum of these two energy terms

$$E_{\text{ext}} = E_{\text{grad}} + E_{\text{time}}. \tag{5}$$

The active surface, $\mathbf{v}(s,r)$, that minimizes the total energy, $E_{\text{int}} + E_{\text{ext}}$, is found by solving the Euler–Lagrange equation associated with the total energy [15]

$$-\frac{\partial}{\partial s}(w_1\mathbf{v}_s) + \frac{\partial^2}{\partial s^2}(w_2\mathbf{v}_{ss}) - \frac{\partial}{\partial r}(w_3\mathbf{v}_r) + \frac{\partial^2}{\partial r^2}(w_4\mathbf{v}_{rr}) + 2\frac{\partial^2}{\partial s \partial r}(w_5\mathbf{v}_{sr}) = -\nabla E_{\text{ext}}(\mathbf{v}). \tag{6}$$

We discretize this equation by finite differences and after rearranging the terms and separating the equation into its x and y components we get

$$A\mathbf{x} + \mathbf{f}_x = 0 \tag{7}$$

$$A\mathbf{y} + \mathbf{f}_y = 0 \tag{8}$$

where $\mathbf{x}$ and $\mathbf{y}$ are vectors of the x and y coordinates of the multiple active contour model, $\mathbf{f}_x$ and $\mathbf{f}_y$ are the derivatives of the external energy with respect to x and y, and A is the regularizing matrix. The matrix A has a block pentadiagonal structure whose terms are dependent on the parameters, w_1, w_2, w_3, w_4, and w_5.

These equations are solved using an iterative gradient descent algorithm which requires an initial approximation to the multiple curves, $(\mathbf{x}_0, \mathbf{y}_0)$. The iterative solution is of the form

$$(A+\gamma I)\mathbf{x}_t = \gamma\mathbf{x}_{t-1} - \mathbf{f}_x(\mathbf{x}_{t-1}, \mathbf{y}_{t-1}) \tag{9}$$

$$(A+\gamma I)\mathbf{y}_t = \gamma\mathbf{y}_{t-1} - \mathbf{f}_y(\mathbf{x}_{t-1}, \mathbf{y}_{t-1}) \tag{10}$$

where $(\mathbf{x}_t, \mathbf{y}_t)$ is the vector of x and y coordinates of the multiple active curves at a time step t in the iteration and γ is the damping constant for the movement of the contour between iterations ($\frac{1}{\gamma}$ is the length of each time step). The derivatives of the external forces, $\mathbf{f}_x$ and $\mathbf{f}_y$, move the x and y coordinates of the model, and the internal forces provide regularization through the inverse of the positive definite matrix $(A+\gamma I)$. The iterations are terminated when the multiple active contours stop changing between iterations.

We invert the block pentadiagonal matrix, $A+\gamma I$, using a relaxation scheme. The relaxation method is based on splitting the matrix into two parts, one of which is easy to invert. We write

$$A + \gamma I = D - R \tag{11}$$

where D represents the block diagonal part of the matrix and R represents the rest of the matrix. The block diagonal matrix, D, is easily inverted using LU-decomposition in $O(n)$ time [18].

Substituting this decomposition into (9) and (10) and replacing $R\mathbf{x}_t$ and $R\mathbf{y}_t$ by $R\mathbf{x}_{t-1}$ and $R\mathbf{y}_{t-1}$, we get the following update equations

$$\mathbf{x}_t = D^{-1}((R+\gamma I)\mathbf{x}_{t-1} - \mathbf{f}_x(\mathbf{x}_{t-1}, \mathbf{y}_{t-1})) \tag{12}$$

$$\mathbf{y}_t = D^{-1}((R+\gamma I)\mathbf{y}_{t-1} - \mathbf{f}_y(\mathbf{x}_{t-1}, \mathbf{y}_{t-1})). \tag{13}$$

This kind of a semi-implicit solution allows a rapid solution to (7) and (8). The in-plane internal forces are transmitted faster than out-of-plane internal forces because the update equations are implicit with respect to the in-plane forces only. This solution method is different from that proposed by Cohen and Cohen [16]. In their method, the matrix $(A+\gamma I)$ is inverted by using the approximation, $(I + \frac{A}{\gamma})^{-1} \approx (I - \frac{A}{\gamma})$, which is valid only if $\gamma \gg \|A\|$. Such an approximation imposes a constraint on the value of the damping constant γ. Their resulting update equations are totally explicit, and hence, the convergence rate is slow.

In our model, we use the method of backtracking line search from numerical optimization to adaptively change the damping constant γ [18]. This method involves choosing the largest step size which leads to a decrease in the value of the objective function.

B. Epicardial and Endocardial Border Detection

In Fig. 1, we show two sample images and the location of the endocardial and epicardial boundaries as outlined by an experienced observer. We applied the multiple active contour method described previously to detect the epicardial and endocardial borders on echocardiogram sequences. To initialize the algorithm, the user needs to specify a rough epicardial boundary on the ED image.

First, we detected the epicardial borders on the entire image sequence by using the rough ED epicardial border as the initial epicardial border for all images in the sequence. Image gradients were computed using Canny's edge detector [19] with a neighborhood of size 5×5. The variance of the Gaussian kernel used to convolve the gradient image was successively reduced ($\sigma = 10.5, 4.5$, and 1.5 pixels) as the iterations of the algorithm progressed. This achieves a scale-space effect, i.e., initially the model is attracted to coarser edges from far away and as the algorithm progresses, the attraction distance becomes smaller, but the edge definition becomes sharper.

After the epicardial borders had been detected on the entire sequence, we used these borders as the initial curves for endocardial border detection. The epicardial border also helped us define a region-of-interest (ROI) on the images in which to look for the endocardial borders. The same algorithm with slightly different parameter values was used both for epicardial and endocardial border detection. The computer-generated epicardial and endocardial boundaries on the two sample images are also shown in Fig 1. The values of the model parameters used for epicardial and endocardial boundary detection were empirically determined and are listed in Table I.

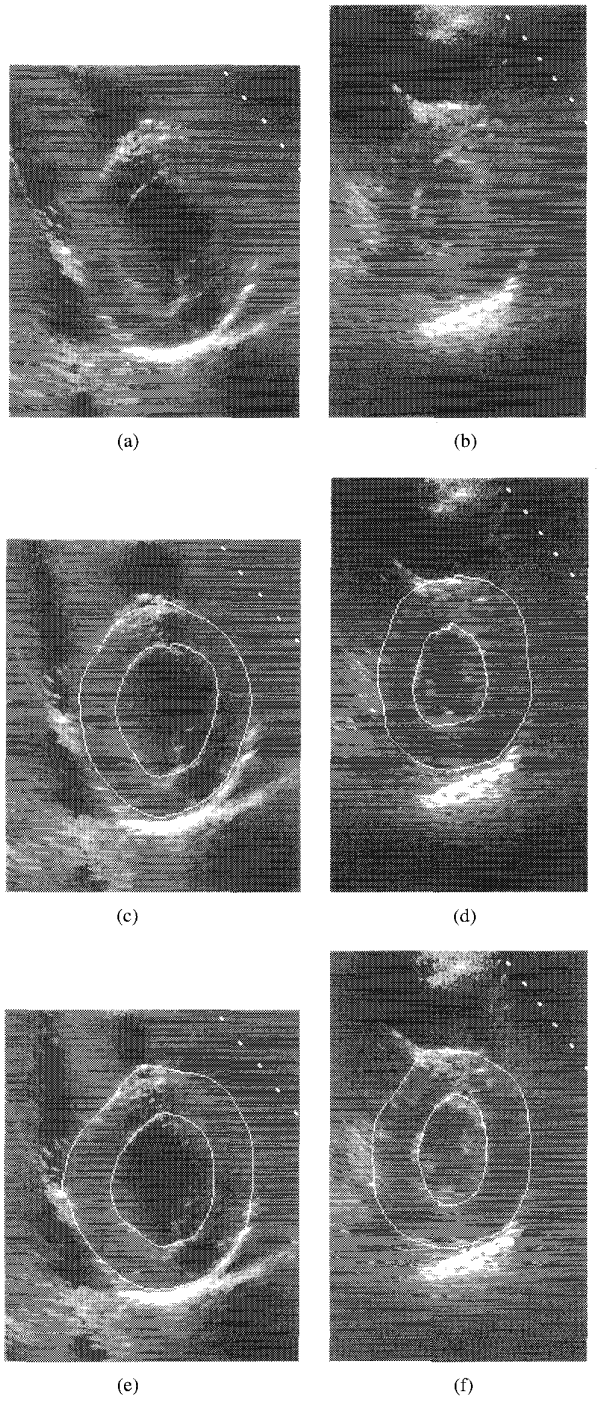

Fig. 1. (a) and (b) Two sample short-axis cardiac images at end diastole, (c) and (d) the same images with hand-outlined epicardial and endocardial boundaries, and (e) and (f) the same images with computer-generated epicardial and endocardial boundaries.

C. Images

For evaluating the algorithm, we obtained transthoracic ultrasound image sets from 44 patients. The data was collected without any regard to image quality during routine echocardiographic examination at the University of Washington Medical Center's Echocardiology Laboratory using a Hewlett-Packard (HP) Sonos 1500 (Hewlett-Packard, Andover, MA) ultrasound machine with a 2.5-MHz or a 3.5-MHz transducer. The only criteria for inclusion were that the patient was studied on that particular ultrasound machine and that sufficient time was available in the schedule to allow digital data capture. The data was saved digitally (as 8-b images) on a magneto-optical (MO) disk using a built-in feature of the ultrasound machine. A single cycle (as determined by the electrocardiogram) of parasternal short-axis mid-papillary views of the LV was saved for each patient. A typical data set consisted of 25 images. The digital file saved on the ultrasound machine included the calibration information which was used to convert the pixel measurements into quantitative data, e.g., distance in millimeters. The pixel size for the 2.5-MHz transducer was 0.57 mm $\times$ 0.46 mm, and that for the 3.5 MHz transducer was 0.43 mm $\times$ 0.34 mm. The imaging depth was 16 cm for the 2.5-MHz transducer and 12 cm for the 3.5-MHz transducer.

TABLE I
PARAMETER VALUES USED FOR EPICARDIAL AND ENDOCARDIAL BOUNDARY DETECTION

	w_1	w_2	w_3	w_4	w_5	w_g	w_t
Epi.	0.2	0.7	0.01	0.01	0.0	0.5	0.1
Endo.	0.5	0.8	0.01	0.01	0.0	0.5	0.1

The images were transferred to a personal computer where the cardiac borders on these images were manually outlined. Four experienced observers (three sonographers and a cardiology fellow) traced the borders on the ED and ES images on each of the 44 data sets. While outlining the endocardial borders, the papillary muscles were considered as a part of the left-ventricular cavity as recommended by the American Society of Echocardiography [14]. In addition to being used to validate the algorithm, one of the hand-outlined epicardial borders on the ED images was used as the initial curve to the boundary detection algorithm.

Five experienced observers were asked to assess the quality of each of the image sets. Each observer rated an image set as being of good (with a score of 2), average (score: 1), or poor (score: 0) quality based on the clarity of the endocardial and epicardial boundaries. All the observers' scores were added and each image set was assigned a quality score (Q) ranging from zero to ten. Fig. 2 shows a histogram of the quality scores for all the 44 image sets used in this study.

D. Analysis of Algorithm Performance

First, we define a metric to measure the distance, $e(\mathcal{A}, \mathcal{B})$, between two given curves, $\mathcal{A}$ and $\mathcal{B}$. If the two curves are represented as sets of points $\mathcal{A} = \{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n\}$, and $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n\}$, where each $\mathbf{a}_i$ and $\mathbf{b}_i$ is an ordered pair of the x and y coordinates of a point on the curve, we define the distance to the closest point (DCP) for point $\mathbf{a}_i$ on the curve $\mathcal{A}$ as

$$d(\mathbf{a}_i, \mathcal{B}) = \min_j \|b_j - a_i\|. \tag{14}$$

These distances are computed for all the points on the two curves and are averaged to yield the mean absolute distance

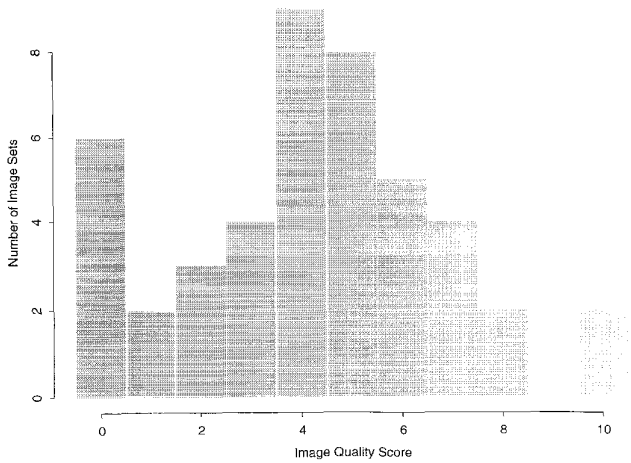

Fig. 2. Histogram of image quality scores (determined by five experienced observers) for the 44 image sets used in the study. The higher the image quality score, the better the definition of the endocardial and epicardial boundaries.

(MAD) between the two curves. Hence

$$e(\mathcal{A},\mathcal{B}) = \frac{1}{2}\left\{\frac{1}{n}\sum_{i=1}^{n} d(\mathbf{a}_i,\mathcal{B}) + \frac{1}{n}\sum_{i=1}^{n} d(\mathbf{b}_i,\mathcal{A})\right\}. \quad (15)$$

We used the average of all the hand-outlined curves as the ground truth curve. The method of computing an average curve is based on establishing corresponding points between the curves to be averaged. To establish correspondence between curves, we used a modification to the iterated conditional points algorithm which is normally used for aligning 3-D shapes [20].

We analyzed our results in three different ways. First, we compared the computer-generated boundaries to each of the hand-outlined boundaries which helped us analyze the computer-generated boundaries with respect to the interobserver variability. Second, we compared the results from the multiple active contour algorithm with different user input curves to evaluate the sensitivity of the algorithm to the initial curve. Finally, we compared the areas enclosed by the computer-generated borders to the areas enclosed by the hand-outlined borders which helped us compare our results to those obtained by other researchers.

For the first set of tests, we computed the distance metric, $e(\cdot,\cdot)$, between all combinations of curves on each image. If the curves outlined by the four observers on image k are denoted by $\mathcal{O}_{k1}, \mathcal{O}_{k2}, \mathcal{O}_{k3}$, and $\mathcal{O}_{k4}$, and the computer-generated curve is denoted by $\mathcal{C}_k$, we computed $e(\mathcal{C}_k, \mathcal{O}_{ki})$ and $e(\mathcal{O}_{ki}, \mathcal{O}_{kj})$ for all $1 \le k \le N$ and $1 \le i, j \le 4$, where $i < j$ and N is the number of images used in the test. For each image and for each observer, we computed the following number

$$r_{ki} = \max_j \{e(\mathcal{O}_{ki}, \mathcal{O}_{kj}) \mid j \neq i\} \quad (16)$$

which is the maximum interobserver boundary distance for observer i on image k. This number gives an upper limit for each observer's curve within which the errors of the computer-generated boundaries should lie. Our test statistic was

$$k_i = \frac{1}{N}\sum_{k=1}^{N}(e(\mathcal{C}_k, \mathcal{O}_{ki}) - r_{ki}) \quad (17)$$

for each observer i. If k_i is less than zero, the distance between the computer boundary and observer i's boundary is less than the maximum distance between the observer i's boundary and other observers' boundaries (the maximum interobserver distance). We computed the 95% confidence interval for k_i and checked to see if its lower limit was less than zero.

For the second set of tests, we computed the boundary distance measure, $e(\cdot,\cdot)$, between the computer-generated boundary and the average manually outlined curve for different input curves to the algorithm. The results for these different user inputs were compared using a nonparametric multiple comparison procedure [21].

For the third test, we performed Pearson's correlation test of the areas enclosed by the computer-generated borders with the areas enclosed by the average manually outlined borders. The correlation coefficient so determined helped us compare our results to the results obtained by other researchers.

E. Sensitivity to the Initial Curve

Since our algorithm requires the user to trace a rough initial curve, we investigated the sensitivity of the algorithm to the initial curve specification. This sensitivity study was performed by perturbing the initial curve in three different ways. First, we tested the sensitivity of the algorithm to different observers using initial curves outlined by different observers. Second, we tested the sensitivity of the algorithm to the size of the initial curve, using scaled versions of the initial curve as inputs to the algorithm. Third, we tested the sensitivity of the algorithm to the resolution of the initial curve, using subsampled versions of the initial curves as inputs to the algorithm.

The initial curve was scaled by first converting the curve into polar coordinates by choosing the centroid of the curve as the origin and then adding a constant offset, c, to the radius of each point on the curve. We used ten different c values, both positive and negative, so that we got both "deflated" and "inflated" versions of the initial curve. Multiple resolution versions of the initial curve were generated in the following way: First, the points constituting the curve were subsampled by two, i.e., one in every two points was selected. There are two possible ways of choosing the points constituting the curve. Next, the curve was subsampled by three, i.e., one in every three points was selected; there are three ways of choosing such a curve. Next, the curve was subsampled by four, i.e., one in every four points was selected; there are four ways of choosing such a curve. This subsampling scheme produced a total of ten versions of the curve (including the original curve).

III. Results

We chose one of the manually outlined ED epicardial borders as the starting curve to the multiple active contour algorithm. To detect the epicardial and endocardial borders on

TABLE II
COMPUTER-TO-OBSERVER DISTANCES (MEAN OVER ALL IMAGES) VERSUS MEAN MAXIMUM INTEROBSERVER DISTANCES FOR EPICARDIAL BOUNDARIES (ALL DISTANCES IN MILLIMETERS). CO = MEAN COMPUTER TO OBSERVER DISTANCE, IO = MEAN MAXIMUM INTEROBSERVER DISTANCE, AND P= PERCENTAGE OF CASES WITHIN THE INTEROBSERVER RANGE

Observer	CO	IO	CO - IO	95% CI	P (%)
1	2.91	5.31	-2.40	(-2.98, -1.81)	89.7%
2	3.36	5.18	-1.82	(-2.40, -1.83)	80.5%
3	3.29	5.17	-1.88	(-2.48, -1.28)	81.6%
4	4.55	5.80	-1.25	(-1.57, -0.94)	86.2%

all images in a sequence, our program required approximately 15 s of CPU time per image (image size 200 × 300 × 8 b) on a Sun Sparcstation 20 Model 61 (Sun Microsystems, Mountain View, CA) workstation (SPECint = 98.2, SPECfp = 107.2).

A. Comparing Results to Interobserver Variability

For the epicardial borders, the mean distance between the computer-generated boundaries and the manual outlines was 2.80 mm ($\sigma = 1.28$ mm). For good quality images ($Q \geq 6$), the mean distance decreased to 2.24 mm ($\sigma = 0.93$ mm). The mean distance for the poor and average quality data sets was 3.04 mm ($\sigma = 1.34$ mm). For comparison, the mean interobserver distance was 3.79 mm ($\sigma = 1.53$ mm) for all images. One observer, observer 4, appeared to be using a somewhat different criterion for the identification of epicardial borders than the other observers. Excluding observer 4's hand-outlined boundaries, the mean interobserver distance decreased to 2.42 mm ($\sigma = 0.66$ mm). We can see that the average interobserver distance is about the same as the distance between the computer-generated boundary and the averaged manual outline.

Table II shows the mean differences and the 95% confidence interval of the mean difference between $e(\mathcal{C}_k, \mathcal{O}_{ki})$ (the computer to observer distance) and r_{ki} (maximum interobserver distance) for each observer, i. Not only the lower limit, but also the upper limit of the 95% confidence interval was less than zero for all the observers. Table II also shows the percentage of image sets for which $e(\mathcal{C}_k, \mathcal{O}_{ki})$ was less than r_{ki}.

For the endocardial borders, the mean distances of the computer-generated boundaries to the average manual outlines was 3.61 mm ($\sigma = 1.68$ mm). For the good quality images, the mean distance was 3.29 mm ($\sigma = 1.61$ mm), and for average and poor quality images the mean distance was 3.75 mm ($\sigma = 1.70$ mm). For comparison, the mean interobserver distance was 2.67 mm ($\sigma = 0.88$ mm). The computer to observer distances are larger for endocardial borders than epicardial borders.

Table III shows the mean difference and the 95% confidence interval of the mean difference between $e(\mathcal{C}_k, \mathcal{O}_{ki})$ and r_{ki} for each observer, i. The lower limit of the confidence interval was less than zero for two of the four observers. For the other two observers, the lower limit was greater than zero.

B. Sensitivity to the Initial Curve

First, we tested the influence of different observers in specifying the initial curve. Starting out with initial curves

TABLE III
COMPUTER-TO-OBSERVER DISTANCES (MEAN OVER ALL IMAGES) VERSUS MEAN MAXIMUM INTEROBSERVER DISTANCES FOR ENDOCARDIAL BOUNDARIES (ALL DISTANCES IN MILLIMETERS)

Observer	CO	CO	CO - IO	95% CI	P (%)
1	3.46	3.42	0.03	(-0.37, 0.43)	55.2%
2	3.92	3.26	0.66	(0.21, 1.10)	40.2%
3	3.84	3.37	0.47	(-0.01, 0.97)	43.7%
4	4.31	3.62	0.68	(0.28, 1.08)	35.6%

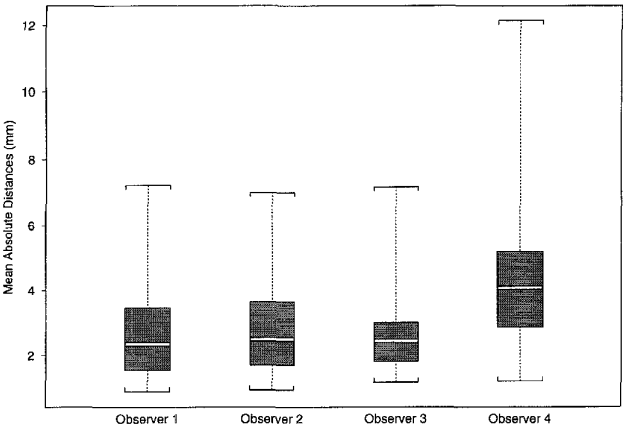

Fig. 3. Boxplots of distances of the computer-generated boundaries to the average manual outline for epicardial boundaries on starting the algorithm with initial curves drawn by different observers. The box represents the middle half of the data, the whiskers extend to the extreme values and the white line inside the box represents the median.

specified by four different observers, we computed the distances of the resulting computer-generated boundaries to the average manual outlines. Fig. 3 shows the boxplots for these distances. We compared the means of these distances using a nonparametric multiple comparison test and found that there was no significant difference in the means of these distances for the first three observers. The fourth observer, as we saw in the last section, outlined borders which were significantly different from the other observers.

Next, we tested the effect of scaling the initial curve. We chose one of the observer curves at random and scaled it with various scaling factors as described in the last section. The scale factor of -4 implied that the perturbed curve was approximately 3.5 mm inside all points on the initial curve. Similarly, for a scale factor of $+4$, the perturbed curve was approximately 3.5 mm outside all the points on the initial curve. Fig. 4 shows the boxplots of the distances of the resulting computer boundaries to the average manual outlines on starting out with scaled versions of the initial curve. There was no significant difference in the means of these distances for scale factors from -4 to $+4$.

We also tested the effect of subsampling the initial curve. We chose one of the observer curves at random and subsampled it by three different factors as described before. A typical hand-drawn curve had 20 points; thus, subsampling it by two implied reducing down to ten points and subsampling it by four implied reducing it down to five points. Fig. 5 shows the boxplots of the distances of the resulting computer-generated

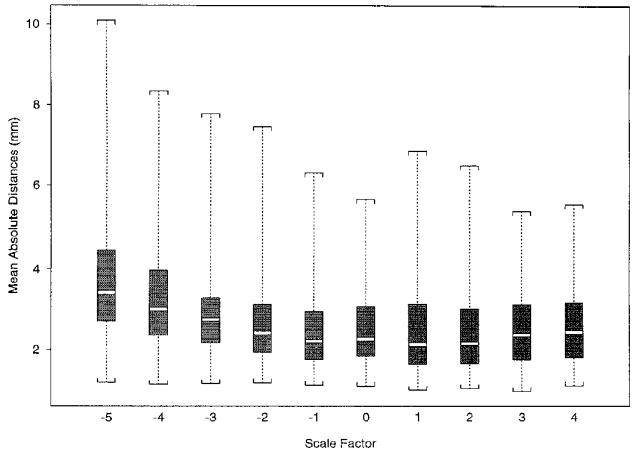

Fig. 4. Boxplots of distances of the computer-generated boundaries to the average manual outline for epicardial boundaries on starting the algorithm with scaled versions of the initial curve.

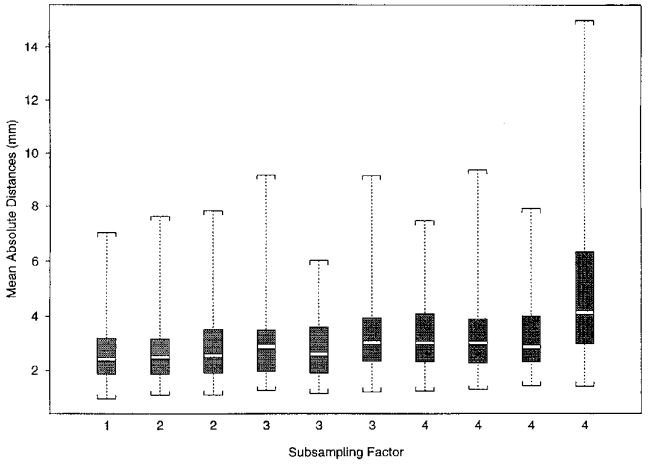

Fig. 5. The distances of the computer-generated boundaries to the average manual outline for epicardial boundaries on starting the algorithm with subsampled versions of the initial curve. For subsampling by two there are two possible choices, for subsampling by three there are three possible choices and for subsampling by four there are four possible choices.

boundaries to the average manual outlines on starting out with these different subsampled versions of the initial curve. There was no significant difference in these distances up to a subsampling factor of three.

C. Comparison of Areas

To compare the areas enclosed by computer-generated boundaries to the areas enclosed by the manually outlined boundaries, we computed the correlation between the area enclosed by the computer boundaries and the area enclosed by the average manual outline. We found a correlation coefficient of 0.95 for the epicardial boundaries. Fig. 6(a) shows a scatter plot of these areas along with the line of unit slope. Next, we divided our data set into good ($Q \geq 6$), average ($3 < Q < 6$), and poor ($Q \leq 3$) quality data and computed the correlation coefficients for these subsets of data.

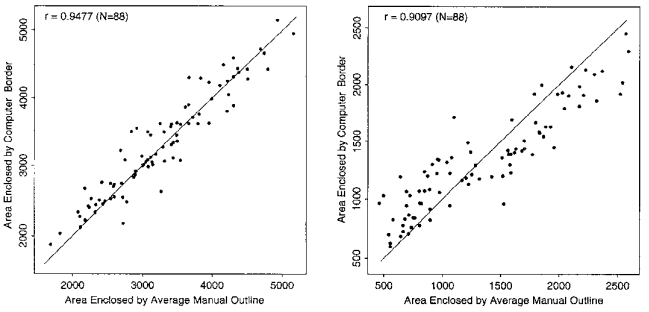

Fig. 6. Scatter plots of the areas enclosed by the computer-generated boundaries against the areas enclosed by the average manual outline for (a) *epicardial* boundaries and (b) *endocardial* boundaries. The areas are in mm^2. The line in each plot is the unity line.

The correlation coefficients were 0.98, 0.97, and 0.83 for the good-, average-, and poor-quality images, respectively.

Similarly, the correlation coefficient was computed for the resulting endocardial borders. For the complete data set, the correlation coefficient was 0.91. The correlation coefficients were 0.93, 0.92, and 0.88, for the good-, average-, and poor-quality images, respectively. The scatter plot for the areas enclosed by the endocardial borders are shown in Fig. 6(b).

IV. Discussion

The multiple active contour algorithm performs well in detecting both epicardial and endocardial borders based on the comparison of the resulting boundaries to the boundaries manually outlined by experienced observers. The differences between the computer-generated boundaries and the hand-outlined boundaries were shown to be comparable to the interobserver differences in manual outlining. The epicardial boundaries are detected more accurately than the endocardial boundaries. This algorithm needs a user input in the form of an initial curve representing the ED epicardial border on a single image in the sequence. We studied the effect of varying the initial curve and showed that the detected boundaries do not vary much when different observers outlined the starting curve. Scaled versions of the initial curve did not result in very different output curves. Also, subsampled versions of the initial curve, up to a subsampling factor of three, did not affect the resulting curves much. From this observation, we can conclude that the user can draw a rough starting curve on one image in the sequence, and the algorithm can detect the boundaries on the entire image sequence with small errors. As expected, the results of the algorithm were better when we considered the good-quality images only; however, for good-quality images, the interobserver variability is also low. The algorithm also performed well on average- and poor-quality images.

Three aspects of this study are not addressed by most other studies. First, we validated our results on a routine clinical data set that an average echocardiology section of a hospital generates. Second, we generated both endocardial and epicardial borders for an entire sequence of images over one cardiac cycle. Third, the endocardial boundaries generated

by our algorithm included the papillary muscles as a part of the LV cavity as recommended by the American Society of Echocardiography [14], whereas most other studies have concentrated on detecting the blood-tissue interface excluding the papillary muscles. Anatomically, the papillary muscles are not a part of the myocardium; therefore, these muscles should not be included as a part of the LV myocardium.

The work of Detmer *et al.* [5] is the only other published work on echocardiographic border detection that uses boundary-distance measures to compare computer-generated boundaries to hand-traced ones. All other researchers have compared areas enclosed by the boundaries. The mean boundary distance to manual outlines computed by Detmer *et al.* for endocardial borders was 3.32 mm ($\sigma = 1.84$ mm). The mean boundary distances for epicardial and endocardial borders due to our algorithm are 2.80 mm ($\sigma = 1.28$ mm) and 3.61 mm ($\sigma = 1.68$ mm), respectively. Although their method of measuring the boundary distances is different from ours, the distances measured are similar to each other. Detmer's study, however, used fairly high-quality transesophageal images. Moreover, their algorithm generated endocardial boundaries which excluded the papillary muscles.

Geiser *et al.* [7] obtained a correlation coefficient of 0.98 when comparing the computer-generated areas to the manually outlined areas. In computing the correlation coefficient, they used highly dependent data, generating 1340 observations from only 30 echocardiogram sequences. Moreover, their algorithm relied heavily on user interaction. These two factors probably contributed to such a high correlation. Zwehl *et al.* [9] validated their algorithm on images from the canine heart. Their algorithm also computed the area of the blood in the LV, thus excluding the papillary muscles. They obtained a correlation of 0.92 for images obtained under nearly ideal laboratory conditions. Two recent papers have studied the online automated boundary detection (ABD) feature available on HP systems [12], [22]. The ABD system detects the blood-tissue interface and hence excludes the papillary muscle from the LV cavity. Vandenberg *et al.* [22] obtained a correlation coefficient of 0.92 for 119 observations with very-good-quality images. Their data were also highly dependent because they generated 119 observations from only eight patients. Perez *et al.* [12] obtained satisfactory results in only 72% of the 66 patients considered. For these 72% of the patients, their correlation coefficient was 0.91.

Thus, we obtained better results than most published algorithms who have validated their algorithms on a relatively large number of clinical data sets. We believe that this is because of two reasons: First, our algorithm uses time-domain information, which mimics the behavior of expert observers outlining these boundaries. Second, our algorithm uses multiple sources of information, such as edge strength, local smoothness, and monotonic motion information.

Other methods have been designed for echocardiographic image segmentation which use models similar in spirit to the active contour models, allowing the integration of multiple sources of information. A method proposed by Friedland and Adam [10] uses a similar approach; however, their method defines the cost function along radial lines. Hence, it is limited and not easily extensible to include other sources of information. Another method has been proposed by Herlin and Ayache [11] which is very similar to our model; however, their algorithm operates on prescan converted images. Also, they use simplistic models about the shape of the LV cavity, which may not always be valid. Due to a lack of clinical validation, it is hard to judge how well their algorithm would perform on real data.

Our method has certain limitations which may benefit from further work. First, the use of multiple active contour model assumes that the images in the sequence are registered with respect to each other, i.e., the transducer or the patient does not move during the scanning process. Moreover, this model does not use any temporal gradient information. Hence, in future, a true 3-D active surface model may be used for cardiac boundary detection. Second, our method uses only the gradient information from the image. Various other kinds of information, such as textural parameters or region statistics, may also be derived from the image and used to improve the performance of endocardial boundary detection. We have recently developed a methodology to integrate both region and edge information into active contours [23]. Future work could include extending this integration to a multiple active contour model or a 3-D active surface model. Third, our algorithm requires user input in the form of an initial boundary. This makes the algorithm results still dependent on the user input, although we have shown that this dependence is small for our data set. This issue is related to the larger issue of whether any medical image analysis procedure can, or should, be made completely automatic. Most automated image analysis procedures are based on certain assumptions about shapes, sizes, or textures which may be violated in abnormal cases. The large variability of shapes, sizes, and textures observed in medical images may make the task of designing fully automated image analysis algorithms with comparable performance to manual image analysis elusive. However, future extensions of this work may focus on trying to reduce the degree of user interaction required for our algorithm.

The method described here can be extended so as to handle other views of the heart, such as the four-chamber view, or the long-axis view. Handling other views of the heart poses additional problems because the cardiac boundaries may not form closed contours any more. The use of prescan converted images may lead to better segmentation results because scan conversion result in differing resolution for different parts of the image with higher resolution near the transducer and lower resolution farther away from it. Finally, this algorithm can be extended to detect other important tissue boundaries in ultrasound images such as the outlines of the fetal skull [24].

V. Conclusion

Cardiac boundary detection from ultrasound images remains an important clinical problem. We have designed an effective and robust algorithm for both endocardial and epicardial boundary detection on short-axis sequences of echocardiograms. We tested the algorithm on a representative clinical data set and compared our results to boundaries manually outlined

by four experienced observers. The algorithm performs well and gives results which are comparable to the interobserver variability in outlining. Also, performance of the algorithm is not greatly affected by variation in the user input to the algorithm. Our algorithm gives better results than most other published methods that have been tested on clinical data sets. The image sets used for this study, including the hand-outlined boundaries, are available to other researchers in this field; please contact the authors for more information.

ACKNOWLEDGMENT

The authors would like to thank B. Schwaegler, C. Kraft, C. Miyake, M. Fujioka, and M. Leggett of University of Washington Medical Center's Echocardiology Laboratory for collecting the data and for hand tracing the cardiac boundaries for this study. They also thank P. Sampson of the Department of Statistics, C. Ritchie of the Department of Electrical Engineering, and D. Liu and D. Gustafson of Siemens Medical Systems for their valuable inputs. They also thank the reviewers for their helpful comments and suggestions.

REFERENCES

[1] E. D. Folland, A. F. Parisi, P. F. Moyniham, D. R. Jones, C. L. Feldman, and D. E. Tow, "Assessment of left-ventricular ejection fraction and volumes by real-time two-dimensional echocardiography: A comparison of cineangiographic and radionuclide techniques," *Circ.*, vol. 60, pp. 760–766, 1979.

[2] S. J. Goldberg, "Analysis and interpretation of thickening and thinning phases of left-ventricular wall dynamics," *Ultrasound Med. Biol.*, vol. 10, pp. 797–802, 1984.

[3] D. A. Conetta, E. A. Geiser, L. H. Olivier, A. B. Miller, and R. Conti, "Reproducibility of left-ventricular area and volume measurements using a computer endocardial edge detection algorithm in normal subjects," *Amer. J. Cardiol.*, vol. 56, pp. 947–952, 1985.

[4] H. E. Melton, S. M. Collins, and D. J. Skorton, "Automatic real-time endocardial edge detection in two-dimensional echocardiography," *Ultrason. Imag.*, vol. 5, pp. 300–307, 1983.

[5] P. R. Detmer, G. Bashein, and R. W. Martin, "Matched filter identification of left-ventricular endocardial borders in transesophageal echocardiograms," *IEEE Trans. Med. Imag.*, vol. 9, pp. 396–404, 1990.

[6] C. H. Chu, E. J. Delp, and A. J. Buda, "Detecting left-ventricular endocardial and epicardial boundaries by digital two-dimensional echocardiography," *IEEE Trans. Med. Imag.*, vol. 7, pp. 81–90, 1988.

[7] E. A. Geiser, D. A. Conetta, M. C. Limacher, V. O. Stockton, L. H. Olivier, and B. Jones, "A second-generation computer-based edge detection algorithm for short-axis two-dimensional echocardiographic images: Accuracy and improvement in interobserver variability," *J. Amer. Soc. Echocardiogr.*, vol. 3, pp. 79–90, 1990.

[8] D. Marr and E. Hildreth, "Theory of edge detection," in *Proc. Royal Soc., London, Ser. B*, 1980, vol. 207, pp. 187–217.

[9] W. Zwehl, R. Levy, E. Garcia, R. Haendchen, W. Childs, S. Corday, S. Meerbaum, and E. Corday, "Validation of a computerized edge detection algorithm for quantitative two-dimensional echocardiography," *Circ.*, vol. 68, pp. 1127–1135, 1983.

[10] N. Friedland and D. Adam, "Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated annealing," *IEEE Trans. Med. Imag.*, vol. 8, pp. 344–353, 1989.

[11] I. L. Herlin and N. Ayache, "Feature extraction and analysis methods for sequences of ultrasound images," *Image Vision Comput.*, vol. 10, pp. 673–682, 1992.

[12] J. E. Perez, A. D. Waggoner, B. Barzilai, H. E. Melton, J. G. Miller, and B. E. Sobel, "On-line assessment of ventricular function by automatic boundary detection and ultrasonic backscatter imaging," *J. Amer. College Cardiol.*, vol. 19, pp. 313–320, 1992.

[13] V. Chalana, D. R. Haynor, and Y. Kim, "Left-ventricular boundary detection from short-axis echocardiograms: The use of active contour models," in *Proc. SPIE Conf. Med. Imag.*, M. H. Loew, Ed., 1994, vol. 2167, pp. 786–798.

[14] N. Schiller, P. Shah, M. Crawford, A. DeMaria, R. Devereux, H. Feigenbaum, H. Gutgesell, N. Reichek, D. Sahn, I. Schnittger, N. Silverman, and A. Tajik, "Recommendations for quantitation of the left ventricle by two-dimensional echocardiography: American society of echocardiography committee in standards subcommittee," *J. Amer. Soc. Echocardiogr.*, vol. 2, pp. 358–367, 1989.

[15] M. Kass, A. Witkin, and D. Terzopolous, "Snakes: Active contour models," *Int. J. Comput. Vision*, vol. 1, pp. 321–331, 1988.

[16] L. D. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-D and 3-D images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1131–1147, 1993.

[17] L. D. Cohen, "Note on active contour models and balloons," *CGVIP: Image Understanding*, vol. 53, pp. 211–218, 1991.

[18] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.

[19] J. F. Canny, "Finding edges and lines in images," M.S. thesis, MIT, Cambridge, MA, 1983.

[20] P. D. Sampson, F. L. Bookstein, F. H. Sheehan, and E. L. Bolson, "Eigenshape analysis of left ventricular outlines from contrast ventriculograms," in *Advances in Morphometrics, Proceedings of NATO Advanced Study Institute*, L. Marcus, M. Corti, A. Loy, G. Naylor, and D. Slice, Eds. New York: Plenum, 1995.

[21] W. W. Daniel, *Applied Nonparametric Statistics*. Boston, MA: Houghton Mifflin, 1978.

[22] B. F. Vandenberg, L. S. Ruth, P. Stuhlmuller, H. E. Melton, and D. J. Skorton, "Estimation of left-ventricular cavity area with an on-line, semi-automated echocardiographic edge detection system," *Circ.*, vol. 86, pp. 159–166, 1992.

[23] V. Chalana, W. C. Costa, and Y. Kim, "Integrating region growing and edge detection using regularization," in *Proc. SPIE Conf. Med. Imag.* M. H. Loew, Ed., 1995, vol. 2434, pp. 262–271.

[24] V. Chalana, T. C. Winter, D. R. Cyr, D. R. Haynor, and Y. Kim, "Automatic fetal size measurements from ultrasound images," submitted to *Academic Radiol.*