

A Knowledge-Based Boundary Delineation System for Contrast Ventriculograms

Lei Sui, Robert M. Haralick, *Fellow, IEEE*, and Florence H. Sheehan

Abstract—Automated left-ventricle (LV) boundary delineation from contrast ventriculograms has been studied for decades. Unfortunately, no accurate methods have ever been reported. A new knowledge based multistage method to automatically delineate the LV boundary at end diastole (ED) and end systole (ES) is discussed in this paper. It has a mean absolute boundary error or about 2 mm and an associated ejection fraction error of about 6%. The method makes extensive use of knowledge about LV shape and movement. The processing includes a multiimage pixel region classification, shape regression, and rejection classification. The method was trained and cross-validated tested on a database of 375 studies whose ED and ES boundary had been manually traced as the ground truth. The cross-validated results presented in this paper show that the accuracy is close to and slightly above the interobserver variability.

Index Terms—Boundary delineation, boundary regression, image segmentation, left ventricle.

I. INTRODUCTION

LEFT-VENTRICULAR angiography is a routine imaging method in clinics [1]. Dye, which is opaque to X-rays, is injected into the left ventricle (LV) via a catheter. Images are continuously recorded through the cardiac cycle to produce a left ventriculogram (LVG), from which the size and function of the LV can be assessed. These measurements are routinely made from the LV boundaries at end diastole (ED) and end systole (ES).

Tracing the LV boundaries from an LVG is a tedious job. Although automated boundary delineation (ABD) in LVGs has long been sought [2]–[8], [11], [16], [10], [25], it has been difficult to achieve clinically acceptable levels of accuracy. On the other hand, most methods were tested on only a limited number of LVG samples. Therefore, no consistent accuracy evaluation concerning LV volume and ejection fraction (EF) was available.

The LVG challenges lie in several aspects [25]. First, there are several other organs visible in the LVG besides the LV. Usually the diaphragm, spine, ribs, and lung markings can be seen in LVG. Second, the LV has a large variation in shape, direction, and position in the image. It is difficult to set up an LV model. Third, due to the different imaging settings, the variation of the

image gray scale is also large. Last, and the most importantly, the dye is unevenly mixed with blood in the LV. The contrast at the endocardial border is extremely low and the locations of gray-scale gradient maxima do not necessarily correspond to the manually traced border.

This study was performed to enhance LVG ABD accuracy by incorporating human knowledge into the process. For example, human experts utilize movement information in the time domain to resolve boundary fuzziness. They are trained to have anatomical knowledge about LV and to reject the cases with bad image quality. Our method uses these sources of knowledge in the ABD system.

Our method is comprised of three stages. The first stage is a nonparametric pixel region classifier operating on the image sequence. It segments the LV region from the background and obtains a raw LV boundary. The second stage is a shape regression that corrects the systematic error of the region classifier [3], [4]. The last stage is a rejection classifier that flags unreliable results [4]. In order to incorporate anatomical knowledge into the system, the user enters points at ED and ES to identify the endpoints of the aortic valve (AoV) and the LV apex. The three points are used to normalize region prior probabilities for the expected LV shape and size and to constrain the shape regression.

In Section II, some related LV ABD methods are reviewed in a comparative way to illustrate the difference between their methods and ours. In Section III, our method is discussed in detail. Some experimental results are presented in Section IV. A conclusion and discussion are then presented in Section V.

II. RELATED METHODS REVIEW

ABD of LVG has been studied for decades [2]–[8], [11], [16], [10], [25]. Two approaches, the edge-based [3]–[6], [8], [10] and the region-segmentation-based [2], [7], [11], [16] have been applied to solve this problem. The edge-based approach tries to find the boundary location using the pixel gray-scale gradient. The region-based approach tries to segment the LV from the background so that a boundary can be obtained by tracing the outline of the LV region. However, the accuracy of these approaches was limited by their failure to fully utilize the information in the images, either at the low level or the high image analysis level.

The low-level information is the information contained in LVG. This kind of information is carried by the pixel gray scale that is used as the main features of the ABD systems. Due to the complexity of the gray-scale appearance, the edge-based methods have difficulties in linking the boundary location with

Manuscript received July 5, 2000. This work was supported by the Goodman Ltd. Company, Nagoya, Japan.

L. Sui was with the Department of Bioengineering, University of Washington, Seattle, WA 98195 USA. He is now with the Ultrasound Group, Siemens Medical System, Issaquah, WA 98029 USA.

R. M. Haralick is with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195 USA.

F. H. Sheehan is with the Cardiovascular Research and Training Center, School of Medicine, University of Washington, Seattle, WA 98195 USA.

Publisher Item Identifier S 1089-7771(01)02030-1.

the gradient information. Indeed, the spatial locations of gradient maxima do not correspond to the manually traced endocardial border. The low contrast at the boundary region tends to show little significant gradient information. The uneven mixture of the dye with blood inside the LV tends to produce many gray-scale gradient maxima near the LV center. Moreover, most of the previous ABD methods process the pixel gray scales in a single frame [2], [3], [6]. This discards the temporal information contained in the image sequence, which we know helps human experts find the LV boundary at the low-contrast region since human eyes are more sensitive to the change of gray scale than the static gray scale.

Some previous methods also assumed that the class conditional distributions of the pixel gray-scale values in the ventricle and the background are Gaussian [2], [7], [11]. This simplified the problem, but at the risk of misleading the algorithm if the assumption is proven to be wrong. In contrast, a method that is nonparametric and uses the temporal information of the gray scale is more suitable for solving the ABD problem in LVG. That is the pathway our method follows.

The high-level information is the general knowledge about LV anatomy. Some researchers used knowledge about the smoothness of LV boundary as the constraint to confine the solution. Pope [6] used dynamic programming to connect the boundary location candidates obtained from the gradient operation. Figueiredo [7] controlled the smoothness in terms of the energy of the Gibbs distribution, which approximated the boundary probability. However, smoothness alone is not enough. The boundary needs to be shaped like a heart.

Some researchers have tried to model the LV boundary using certain critical points either entered by the user or detected by computer. Clayton [3] used five points: the apex, endpoints of the AoV, and right- and left-most points on the LV boundary. Barrett [4] initially used seven points, adding two points above the AoV to the five points used by Clayton. Later, Barrett [5] guided the algorithms using just four points: the two AoV endpoints, apex, and mitral valve points. Except for the AoV end points and apex, which have an explicit anatomical meaning, the other points were subjectively judged and, thus, might introduce a large variance to the model.

Other researchers trained their algorithms with the hand-traced LV boundaries. Fan's deformable model [9] set up the LV boundary region from the most interior LV boundary to the most exterior LV boundary in the training set. Costa trained his artificial neural network with the chord angle and length between the LV center and contour along the boundary [10]. All the above techniques used edge-based methods. As discussed earlier, however, the edge-based methods are not suitable for the LVG boundary delineation problem. Consequently, we formulated the high-level knowledge in terms of the LV region for our region-based method.

Automatic border delineation for LVGs has been studied for ten years in our laboratory [11], [16], [25]. It became apparent that no single-stage operation can fulfill the mission. Lee [11] formulated the ABD problem with an LV region segmentation in the Bayesian framework. Motivated by the usefulness of the temporal information, he used the gray scale at a pixel location through the systolic cycle where ED is the first image frame

and ES is the last image frame as the feature vector to assign the pixel a class. Each pixel in the image frame belongs to one of three classes: the background, ED-not-ES, and ES. The ES pixels are further constrained to be located within the ED region.¹ Thus, the classifier produced the ED and ES boundaries after segmenting the image into ED and ES regions. At first, this classification was made parametrically by assuming that the class conditional probability of the multidimensional feature vector had a Gaussian distribution. Lee improved upon this by obtaining the class conditional probability from the class conditional Mahalanobis histograms of the feature vector. His prior probability of each class was modeled with two AoV points. A large uncertainty was observed at the apex region. The algorithms were cross validatedly tested on a large database of 375 studies, each of which had been manually traced. He had a mean absolute boundary error (defined in Section III) of 5.2 mm.

Suri [16] improved these results by introducing a boundary regression following the classification. A boundary was treated as a vector by concatenating its vertex coordinates into a vector. Given the hand-traced boundary vectors and classifier boundary vectors, the regression coefficients that best transformed the classifier boundary vectors to the hand-traced boundary vectors were estimated by the least squares criterion. In order to minimize the generalization error, the dimension of the regression vector was reduced by down sampling the boundary vertices. He reduced the mean absolute boundary error to 3.1 mm from its original 5.2 mm. Based on Lee and Suri's formulation, we made a further improvement. Our system is a three-stage operation, whose mean absolute boundary error is about 2 mm. Furthermore, the human interobserver variability in terms of ED volume, ES volume, and EF are just slightly below those of our ABD procedure.

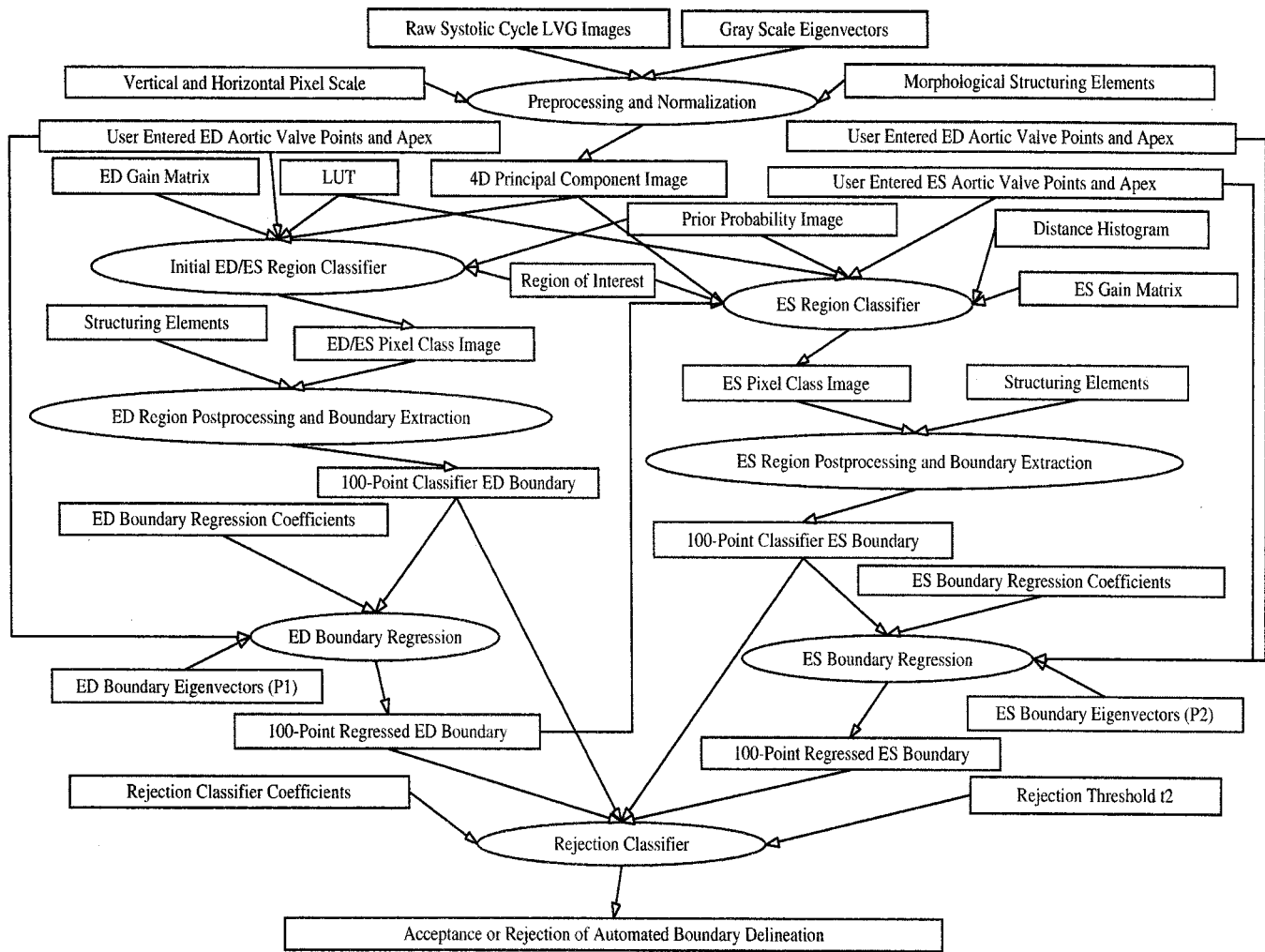
III. METHOD

Our three-stage ABD method is composed of a nonparametric Bayesian classification [11], [25], a shape calibration regression [16], [25], and a rejection classification [17], [25]. Fig. 1 shows the object processing diagram (OPD) of our three-stage ABD system. Since ES involves more variation than ED, the ED result is usually more reliable than the ES result and can help with the ES boundary delineation. Therefore, our processing of the ES boundary is conditioned on the detected ED boundary.

A. Pixel Region Classification

The Bayesian pixel classifier labels each pixel in the image frame of the normalized systolic LVG with one of three classes: background (Class 1), ED-not-ES (Class 2), and ES (Class 3), which is based on the systole gray scale vector. This is shown in Fig. 2. The underlying concept is that background pixels retain their intensities throughout systole, ES pixels increase their intensities throughout systole, and ED-not-ES pixels demonstrate a characteristic pattern of gray values over time depending on their location as the LV border passes over the pixels during LV contraction.

¹It is not always true. Since the possibility that the ES pixels stay outside the ED region is small, that assumption was made to simplify the problem and avoid unreliable probability estimation.



Online LVG Boundary Delineation

Fig. 1. OPD of a three-stage ABD system.

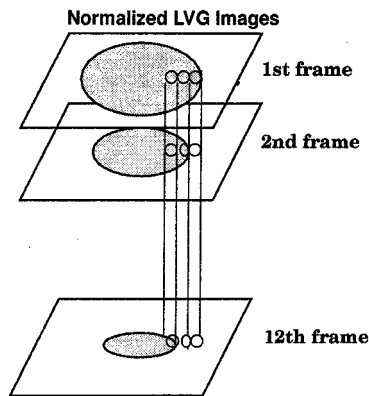


Fig. 2. Gray-scale vector.

1) *Online Processing:* Before classification, the systole images from ED to ES go through pixel size calibration [11], noise filtering [11], and are normalized to 12 frames to adjust for variability in frame rate and heart rate [11]. The pixel size calibration makes the pixel square in millimeter unit. This not only nor-

mizes the image size, but also benefits the performance evaluation in millimeter unit. Noise filtering is implemented by the gray-scale morphological opening and closing [15]. This is similar to median filtering, but in addition, it preserves the location of the gray-scale boundary transition, and at the same time, filters out the noise. The heart-rate normalization is essential to the classification in that it makes the raw feature vector dimension a constant 12.

Only the pixels within the region of interest (ROI) are processed. The ROI is the area outside in which no part of the LV would occur. It was obtained by overlapping all the ground truth (GT) class region images together. The largest LV region resulting from the overlapping was morphologically dilated by a 7×7 square structuring element. The dilation ensures the mask can be generalized to the entire LVG population. The ROI is shown in Fig. 3.

The gray-scale values of the 12 image frames are normalized by the cumulative distribution function (CDF) within the ROI in the sequence. That is, each pixel's gray-scale value is transformed to a fraction between 0–1, where the fraction means the

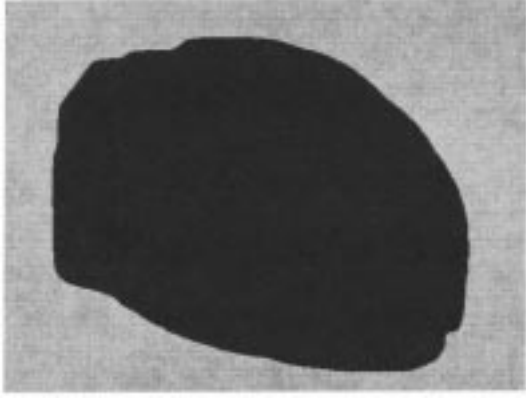


Fig. 3. ROI displayed as the black hole.

fraction of pixels within the ROI on all 12 image frames that have a gray-scale value less than or equal to it. Hence,

$$y_i(u, v) = \sum_{j=0}^{g_i(u, v)} \frac{n_j}{N} \quad (1)$$

where $i \in 1, \dots, 12$ is the image frame index, (u, v) is the pixel location in the image frame, n_j is the number of pixels whose gray scale is j in the ROI on the 12 image frames, N is the total number of pixels in the ROI of the 12 image frames, $g_i(u, v)$ is the original pixel gray scale at (u, v) of the i th image frame, and $y_i(u, v)$ is the pixel value at (u, v) of the i th image frame after normalization. The normalization is done with respect to the entire systole sequence rather than a single image frame because this keeps the gray-scale value change pattern through systole.

The 12 CDF normalized pixel values at each pixel location form a vector $Y = (y_1, \dots, y_{12})$. This vector is projected onto the eigenvectors of the second moment matrix of the sample pixel value vectors $\mathbf{Y} = \sum_{j=1}^M Y_j^T Y_j$, where M is the total number of training vectors. In our data set of 375 cases, $M = 40, 056, 750$. The eigenvectors are ordered so that their eigenvalues are descending. The first four principal components (PCs) of Y , $X = (x_1, \dots, x_4)$, are used as the feature vector for the classifier.

The pixel classification is based on a Bayesian rule maximizing the expected gain. Given the gain matrix $\mathbf{G} = \{g(c, c') | c, c' \in \{1, 2, 3\}\}$, where 1 denotes the background class, 2 denotes the ED-not-ES class, 3 denotes the ES class, and $g(c, c')$ denotes the gain of assigning to class c while the true class is c' , the expected gain of classifying X at (u, v) to class c is

$$E(g(c|X)) = \sum_{c'=1}^3 g(c, c') p(c'|X, (u, v))$$

where $p(c|X, (u, v))$ is the posterior probability of class c given the feature vector X at location (u, v) . The Bayesian rule classifies X at (u, v) to class c if the assignment to class c yields the expected maximal gain

$$E(g(c|X)) = \max_{c'} E(g(c'|X)) \quad (2)$$

where $c' \in \{1, 2, 3\}$.

For the ED pixel classification

$$p(c|X, (u, v)) \propto p(X|c) p(c|(u, v)) \quad (3)$$

where $p(c|(u, v))$ is the prior probability of class $c \in \{1, 2, 3\}$ at (u, v) , and $p(X|c)$ is the class conditional probability of X given class c , which is assumed to be independent of (u, v) .

For the ES pixel classification, the posterior probability $p(c|X, (u, v))$ depends on the previously delineated ED boundary and is written as

$$p(c|X, d, (u, v)) \propto p(X|c) p(c|(u, v)) p(d|c, (u, v)) \quad (4)$$

where $p(d|c, (u, v))$ is the probability of the shortest distance d between (u, v) and the automatically delineated ED boundary given class c .

After classification, the pixels of the same class are grouped together. For the ED classification, the candidate ED regions are composed of the pixels of classes 2 and 3. The largest candidate region is selected as the true ED region. The ED region is further refined by removing the class 3 pixels outside the class 2 pixels. This position checking enforces the class category assumption that the ES pixels are not located outside the ED region. The binary morphological opening and closing are used to smooth the region boundary. After that, the raw ED boundary is traced from the ED region and represented by 100 evenly spaced points. The same postprocessing is applied to the ES classification in which the ES region is composed of class 3 pixels.

2) *Classifier Training*: The classifier training includes estimating the class conditional probability $p(X|c)$, prior probability $p(c|(u, v))$, and class conditional distance probability $p(d|c, (u, v))$.

Reducing the feature vector dimension from 12 to 4 makes it possible to estimate the class conditional probabilities $p(X|c)$ nonparametrically. A four-dimensional (4-D) lookup table (LUT) is constructed by dividing the 4-D space into nonuniformly sized hyperrectangular bins and estimating the probability in each hyperrectangle by simply counting. Compared to the decision tree or the artificial neural network, the LUT method has the advantage of being simple and fast. To determine the class conditional probability of a feature vector, the feature vector is quantized to make an address to access the LUT. The probability is looked up from the addressed entry in the table.

Given a number of samples of X and the number m of bins of the LUT, the number of bins b_i for the i th dimension is allocated according to the marginal entropy h_i on that dimension so that $m = \prod_{i=1}^4 b_i$ and the marginal information loss is minimal. According to Shannon's theory [13], the entropy reflects the information amount. The total bits available for an m bin LUT to describe the marginal information is $\log_2 m$. The number of bits assigned to axis i is proportional to the i th marginal entropy

$$h_i = \sum_{j=1}^{l_i} p_j \log p_j \quad (5)$$

$$\lambda_i = \frac{h_i}{\sum_{j=1}^4 h_j} \log_2 m \quad (6)$$

where l_i is the number of digitizing level for x_i , p_j is the probability of x_i at level j , and $i = 1, \dots, 4$. The number of bins on axis i is

$$b_i = \lceil 2^{l_i} \rceil \quad (7)$$

where $\lceil \cdot \rceil$ denotes the closest integer. Each marginal distribution of x_i is quantized into b_i bins with equal probability quantization, which minimizes the marginal information loss.

After each training sample has been put into one of the bins in the LUT, the class mixture samples are smoothed by a $k - NN$ kernel [14]. As a result, each bin contains three class conditional probabilities estimated by frequency counting of how many samples of each class are in the bin.

The LUT size m and smoothing parameter k determine the generalization capability of the LUT. The optimal m and k are obtained by searching the flat minima of the cross-validated bidirectional Kullback–Liebler distance between the observed LUT and smoothed LUT

$$L(p(X), \hat{p}(X))|_{k,m} = \sum_{i=1}^m (p_i(X) - \hat{p}_i(X)) \log \frac{p_i(X)}{\hat{p}_i(X)} \quad (8)$$

where $p_i(X)$ is the observed class mixture probability at bin i and $\hat{p}_i(X)$ is the class mixture probability estimated by $k - NN$ by leaving out the samples at bin i . The flat minima of $L(p(X), \hat{p}(X))|_{k,m}$ suggests a good generalization parameter region [26].

The class prior probabilities at each pixel location constitute the knowledge about the LV shape, size, and position. A set of 375 GT class region images generated from the hand-traced boundaries are aligned by a transformation to a common LV region by a least squares fitting of their AoV angles and long axes. The LV AoV angle and long axis are defined by the AoV end points and apex, shown in Fig. 4. Those aligned GT class region images are stacked one on the top of the other. The probability of each class at each pixel location is estimated by frequency counting from the aligned GT class region images and smoothing the frequency count images with a 7×7 square box filter.

The prior probability image alignment needs the user to enter three points for ED and three points for ES. The alignment based on the ED three points defines the geometric transformation f_n that aligns the prior probability images for ED-not-ES and background with image sequence for patient n . The user entered three points of ES defines the geometric transformation g_n that aligns the prior probability image for ES with the image sequence for patient n .

Let: 1) B_n be the background GT class region image for patient study n ; 2) D_n be the ED (ED-not-ES) GT class region image for patient study n ; 3) S_n be the ES GT class region image for patient study n ; 4) PB be the prior probability image for background; 5) PD be the probability image for ED-not-ES; and 6) PS be the probability image for ES

$$NB(r, c) = \# \{ n | B_n(i, j) = 1, (i, j) = f_n(r, c) \} \quad (9)$$

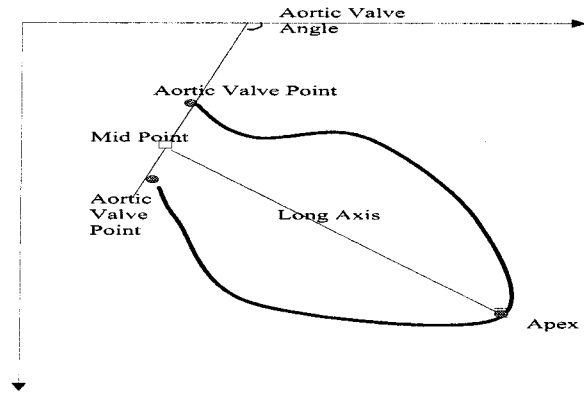


Fig. 4. Three-user input points determines the AoV angle, AoV middle point, and apex, which define the alignment transform.

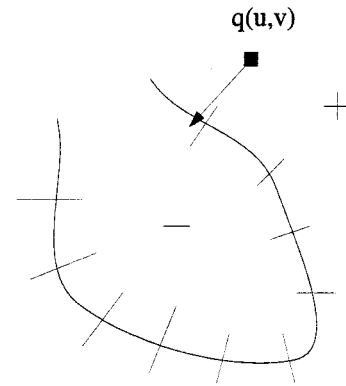


Fig. 5. Distance between the pixel and the delineated ED boundary.

$$ND(r, c) = \# \{ n | D_n(i, j) = 1, (i, j) = f_n(r, c) \} \quad (10)$$

$$NS(r, c) = \# \{ n | S_n(i, j) = 1, (i, j) = g_n(r, c) \} \quad (11)$$

where “#” is the counting operator. The prior probabilities at (r, c) are then

$$PB(r, c) = \frac{NB(r, c)}{NB(r, c) + ND(r, c) + NS(r, c)} \quad (12)$$

$$PD(r, c) = \frac{ND(r, c)}{NB(r, c) + ND(r, c) + NS(r, c)} \quad (13)$$

$$PS(r, c) = \frac{NS(r, c)}{NB(r, c) + ND(r, c) + NS(r, c)}. \quad (14)$$

Since the alignment uses an AoV angle and long axis defined by the three points with the anatomical meanings, the alignment has a good estimation reliability.

Given the delineated ED boundary, the shortest distance between a pixel location and the boundary can be computed. The distance is signed, as shown in Fig. 5. The pixel inside the ED boundary has a minus sign, outside has a plus sign, and on the boundary is zero. The ED boundary is divided into several segments since the wall movement along the ED boundary is not even. The class conditional distance histograms are then setup for each segment, given the pixel class and the corresponding segment on the ED boundary. The class conditional probabilities are estimated from the histograms.

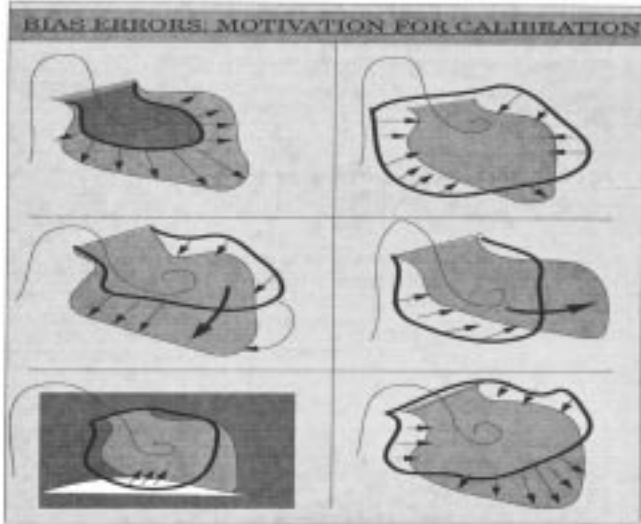


Fig. 6. Bias errors produced by the pixel classification system. (top) Classifier boundary inside and outside the ideal boundary. (middle) Classifier boundary toward the anterior and inferior wall. (bottom) Classifier boundary overlaps the diaphragm. (bottom, right) Classifier boundary outside the ideal boundary in top two-third region and underestimated in the bottom one-third region.

B. Shape Regression

It was observed that the classifier has some systematic shape bias errors. Fig. 6 shows some examples of the shape bias. In order to normalize out those errors, we use a shape regression after the classifier. Each LV boundary is taken as a two-hundred-dimensional vector by concatenating its vertex coordinates into a vector. The boundary vector is then approximated by representing it in the first few dimensions of the LV shape space defined by the eigenvectors of the second moment matrix of the sample GT boundary vectors. The reason for compressing data by PC analysis is to sufficiently reduce the number of coefficients the regression has to estimate so that the regression is able to generalize rather than memorize the training data.

These first few PCs contain most of the shape information instead of using the information in a deformable model, as in [24], to reconstruct the target. We adopt a regression, which is more time efficient. The idea of the regression is to transform the coordinates of the raw classifier boundary vector to a place in the LV shape space where the coordinates are more like an LV boundary. The regression vector is composed of the projected coordinates of the raw boundary vector on the eigenvectors of the second moment matrix of the sample GT boundary vectors. Furthermore, the regression vector is augmented by the user-entered AoV endpoints and apex, which put more constraints on the regression results with the size and shape indicated by the three points. The regressed boundary coordinates are the linear combinations of the coordinates in the shape space and the quadratically augmentation terms of the user entered points.

We will now describe the online process in detail.

1) *Online Regression:* Given: 1) a raw boundary vector $Y = (r_1, c_1, \dots, r_{100}, c_{100})$ from the classifier; 2) two AoV points and the apex coordinates entered by the user; 3) a matrix \mathbf{A} , whose columns are the full set of the eigenvectors of the sample second moment matrix of the GT boundary space; and 4) a re-

gression coefficient matrix $\mathbf{C}_{(P_2+t) \times P_3}$ determined offline by training, the regressed boundary is

$$\hat{B} = ([Y \mathbf{A}_{P_2}; T_t] \mathbf{C} \mathbf{A}'_{P_3}) \quad (15)$$

where \hat{B} is the calibrated boundary, \mathbf{A}_{P_2} and \mathbf{A}_{P_3} are matrices of the first P_2 and P_3 eigenvectors in \mathbf{A} , the prime denotes the transpose, and T_t contains t augmentation terms coming from the three-user input points. For the ED boundary regression, T_t is the full quadratic terms of the coordinates of the user-entered three ED points. Those consist of a constant, 6 first order items and 21 second order items. Thus, t is equal to 28. For the ES boundary regression, T_t is the full quadratic terms of the three user-entered ES points, and t' partial quadratic terms of the three ED points. Thus, t is equal to $(28 + t')$.

2) *Regression Training:* The regression coefficient estimation has a potential generalization problem due to the two-hundred-dimensional boundary vector space. In order to reduce the degrees of freedom, the number of PCs P_2 and P_3 involved in the regression from the raw boundary and the GT, respectively, need to be as small as possible, while maintaining good performance. Their sizes are determined by an optimization procedure.

Let $\mathbf{Z} = (Z_1, Z_2, \dots, Z_n)'$ be a matrix of n GT boundaries, in which Z_i is the i th boundary; let $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)'$ be a matrix of n classifier output boundaries, in which Y_i corresponds to Z_i , and let \mathbf{A}_{P_2} and \mathbf{A}_{P_3} be two subsets of \mathbf{A} , which are composed of the first P_2 and P_3 eigenvectors in \mathbf{A} , respectively. The coefficient matrix \mathbf{C} is estimated offline to minimize

$$\|\mathbf{W} - \mathbf{V}\mathbf{C}\| \quad (16)$$

where $\mathbf{V} = ([Y \mathbf{A}_{P_2}; \mathbf{T}])$, $\mathbf{W} = \mathbf{Z} \mathbf{A}_{P_3}$, and \mathbf{T} is the augmentation matrix. The standard least squares solution for (16) estimates a $(P_2 + t) \times P_3$ matrix \mathbf{C}

$$\mathbf{C} = (\mathbf{V}'\mathbf{V})^{-1} \mathbf{V}'\mathbf{W} \quad (17)$$

where t is the number of the augmentation terms.

Motivated by the knowledge that the ED boundary contains information about the ES boundary, some of the quadratic terms of the three ED points are used in the ES boundary regression. To determine the partial quadratic augmentation from the ED points for the ES regression, let $\mathbf{R} = (R_1, \dots, R_i, \dots, R_n)'$, where R_i is the i th 1×27 row vector including the six row and column coordinates of the i th three ED points and their 21 quadratic combinations. Let $\mathbf{Z} = (Z_1, \dots, Z_n)'$ be the ES GT boundary vector matrix. R_i corresponds to Z_i . A 27×200 coefficient matrix \mathbf{Q} is estimated by minimizing

$$\|\mathbf{Z} - \mathbf{R}\mathbf{Q}\|. \quad (18)$$

The i th row of \mathbf{Q} , $(q_{i1}, \dots, q_{i200})$, $i = 1, \dots, 27$, contains 200 coefficients of the corresponding i th ED augmentation term for the 200 ES boundary coordinates, respectively. The term that corresponds to a large mean of the 200 coefficients $(1/200) \sum_{j=1}^{200} q_{ij}$ is of a big influence on the ES boundary coordinates and selected as the augmentation term for the ES boundary regression.

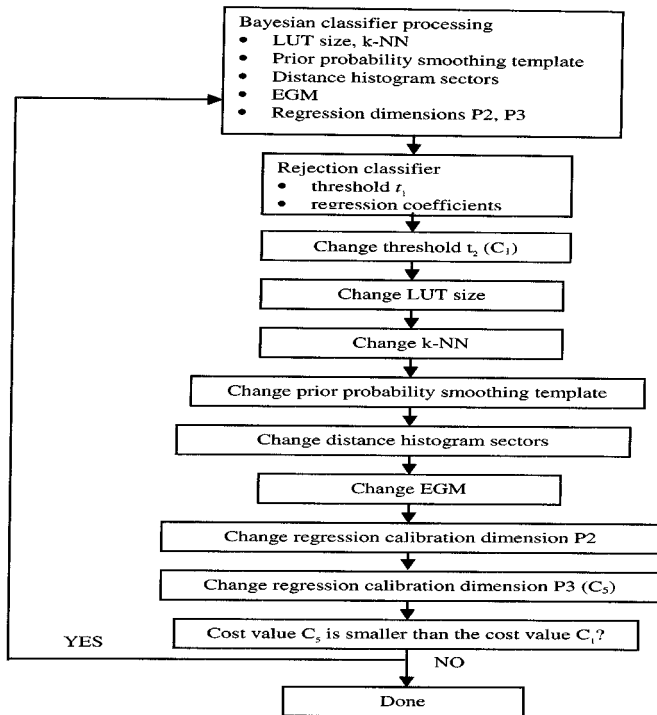


Fig. 7. Flowchart for the systematic optimization to minimize the cost.

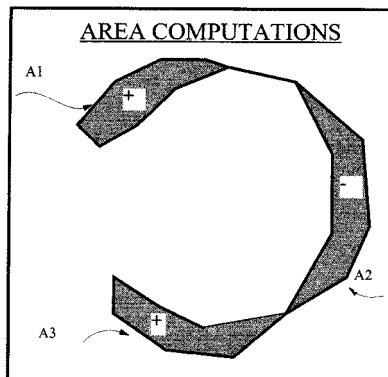


Fig. 8. Hamming distance between the area enclosed by two boundaries. It is equal to the sum of the shaded areas in the graph.

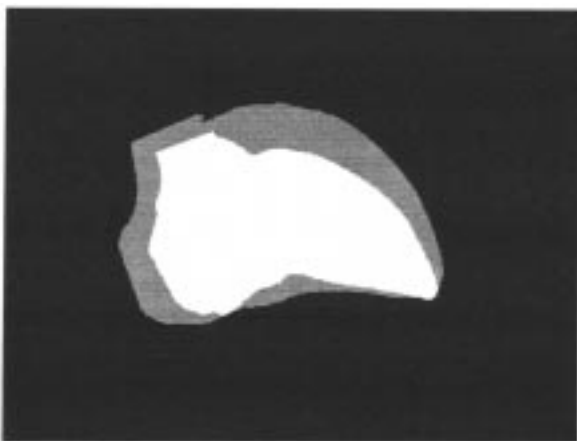


Fig. 9. Example of the GT class image. (bright): ES class (class 3). (gray): ED-not-ES class (class 2). (dark): background class (class 1).

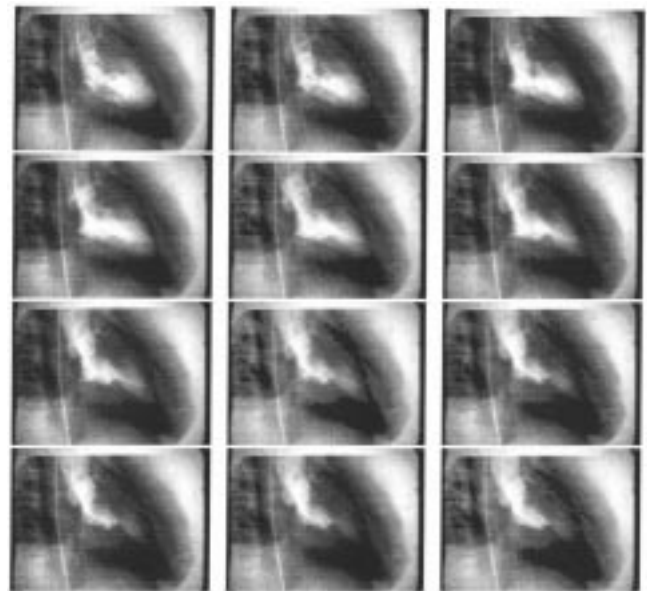


Fig. 10. Frame rate and square pixel normalized 12-frame systolic cycle images. The gray scales are before the CDF normalization. The upper left-hand side is the ED frame, the bottom right-hand side is the ES frame.

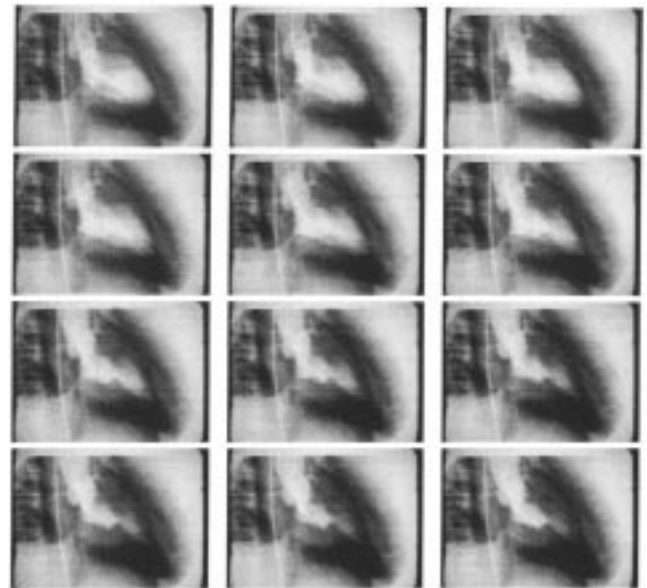


Fig. 11. Frame rate and square pixel normalized 12-frame systolic cycle images. The gray scales are CDF normalized. The upper left-hand side is the ED frame, the bottom right-hand side is ES frame.

C. Rejection Classifier

1) *Online Rejection*: From the raw and regressed boundaries, two sets of parameters can be computed: the ED and ES volumes, and the EF, the ED, and ES areas. In addition, the mean absolute ED boundary difference between classifier and regression, and the mean absolute ES boundary difference between classifier and regression can also be obtained. Those measurements are defined at the conclusion of this section. The rejection classifier takes those parameters as the rejection feature vector. The difference between the two sets of the parameters indicates, more or less, the unreliability of the border detection. Given the classification weights $Y_{12 \times 1}$ on the rejection feature vector

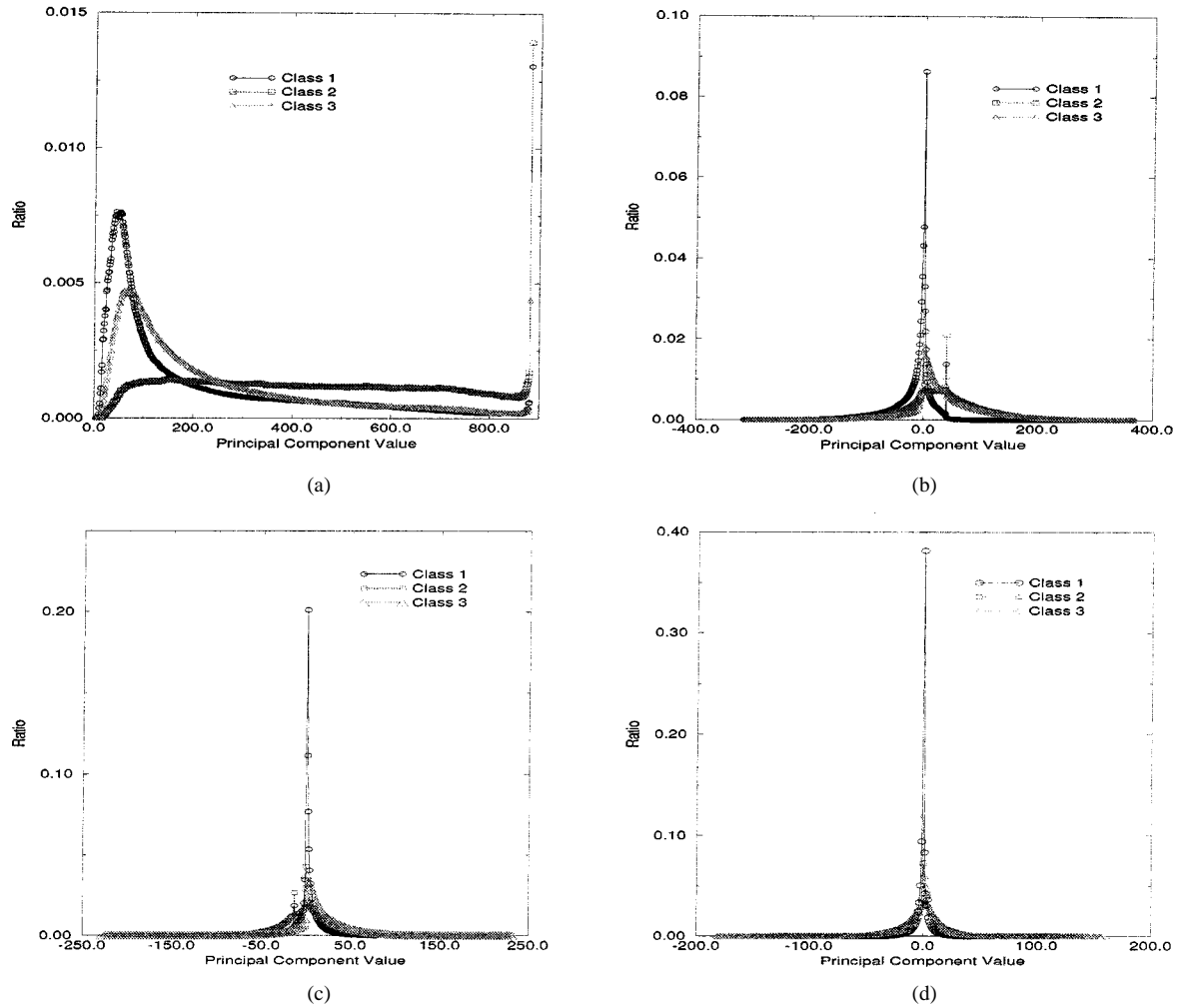


Fig. 12. PC histograms before the gray-scale CDF normalization. (a) Histograms of the first PCs of the three classes. (b) Histograms of the second PCs of the three classes. (c) Histograms of the third PCs of the three classes. (d) Histograms of the fourth PCs of the three classes. Class 1 is the background class, class 2 is the ED-but-not-ES class, class 3 is the ES class.

$U_{1 \times 12}$ and a decision threshold t_2 , a study is rejected if the difference is greater than the threshold $UY > t_2$.

2) *Rejection Training*: Suppose n training cases are available. Let E be an $n \times 1$ objective vector whose elements are one or zero. One indicates that the corresponding training case associated with its component position has an ED volume, ES volume or EF error greater than an acceptable threshold. Zero indicates that its error is acceptable. The fraction of one components in the rejection objective vector is roughly 5% of the training cases such that the same percentage of the worst cases of the ED volume, ES volume, and EF, respectively, are flagged. Let V be an $n \times 12$ matrix, each of whose rows is a rejection classifier vector for its corresponding study. The classifier coefficients Y , an 12×1 vector, can be estimated by minimizing

$$\|E - VY\|. \quad (19)$$

The standard least squares solution for (19) is

$$Y = (V'V)^{-1}V'E. \quad (20)$$

A rejection decision is made on the study whose feature vector is V when $VY > t_2$. When $VY \leq t_2$, the study results are not

rejected. Here, t_2 is the decision threshold, which is obtained by minimizing a cost function. The cost function is defined as

$$\text{cost} = \sum_{i=1}^{n-n'} e_i + \sum_{j=1}^{n'} w \quad (21)$$

where n is the total number of the training cases, n' is the number of the rejected cases, e_i is the true error of the i th accepted case, and w is the predefined cost for a rejected case. The second term of (21) encourages the rejection of studies whose error is greater than w , and punishes the rejection of studies with smaller errors than w .

D. System Optimization

The whole system is optimized after the estimation of each tuning parameter is in a suitable range. The optimization is to minimize the cross-validated EF error cost after the rejection. The tuned training parameters include the LUT size, the LUT smoothing parameter k -NN, the prior probability smoothing parameter, the class conditional distance histogram sectors, which divides the delineated ED boundary, the gain matrices (EGMs), the regression dimensions $P2$, $P3$, and the rejection threshold t_2 . Fig. 7 shows the optimization flowchart.

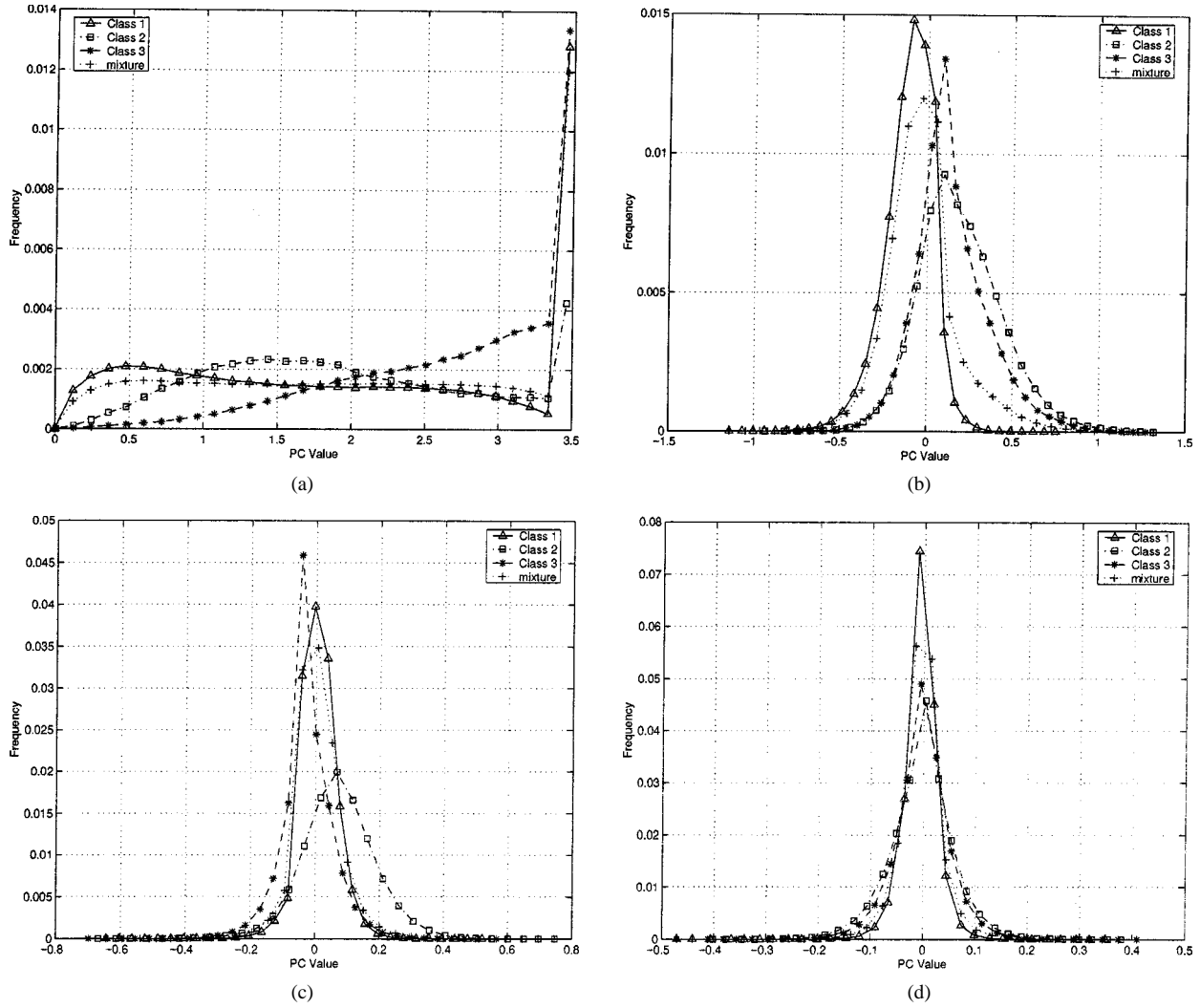


Fig. 13. PC histograms after the gray-scale CDF normalization. (a) Histograms of the first PCs of the three classes and the mixed class. (b) Histograms of the second PCs of the three classes and the mixed class. (c) Histograms of the third PCs of the three classes and the mixed class. (d) Histograms of the fourth PCs of the three classes and the mixed class.

E. Measurements Used in the Rejection Classifier

From an LV boundary $\{(r_1, c_1), (r_2, c_2), \dots, (r_{100}, c_{100})\}$ in millimeter unit, the perimeter of the LV is

$$\rho = \sum_{i=1}^{100} \sqrt{(r_i - r_{i+1})^2 + (c_i - c_{i+1})^2} \quad (22)$$

where $(r_{101}, c_{101}) = (r_1, c_1)$. The area of the LV is

$$A = \frac{1}{2} \sum_{i=1}^{100} |c_{i+1}r_i - c_i r_{i+1}| \quad (23)$$

where $(r_{101}, c_{101}) = (r_1, c_1)$. The LV volume can be estimated by [27]

$$V = 0.81 \times \frac{8A^2}{3\pi l} + 1.9 \quad (24)$$

where l is the maximum chamber length. The EF is estimated as the ratio of the volume difference between ED and ES over the ED volume

$$EF = \frac{V_{ED} - V_{ES}}{V_{ED}}. \quad (25)$$

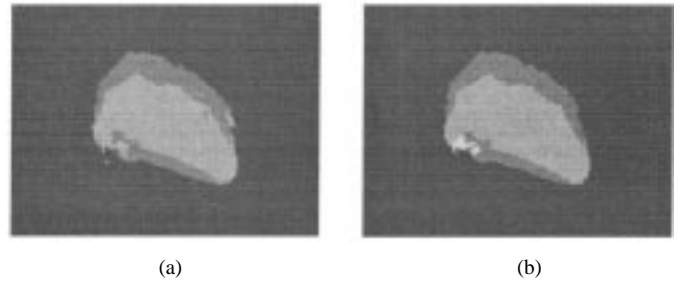


Fig. 14. Example of the class region image after the ED classifier and the result after the postprocessing. The dark gray is the background region, the mid gray is the ED-not-ES region, and the bright gray is the ES region. The white in the right-hand-side image are the rejected ES pixels after position checking.

The mean absolute ED volume deviation, ES volume deviation, and EF deviation between the ABD results and hand-traced results are used for the ABD system performance evaluations.

Besides those volume-related measurements, the mean absolute boundary error between two boundaries is used in the development of the ABD system. Let A_1 and A_2 be the areas en-

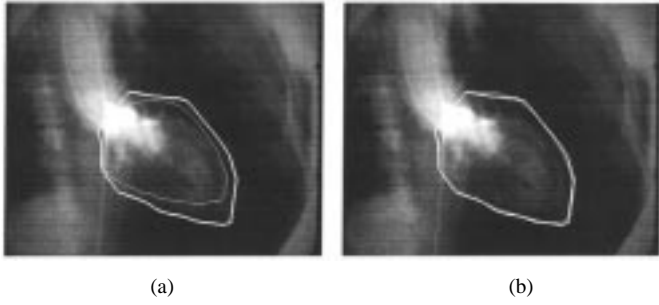


Fig. 15. ED classifier (thin) and regression boundary (thin) compared with its GT boundary (thick). (a) ED frame: GT and classifier. (b) ED frame: GT and regression.

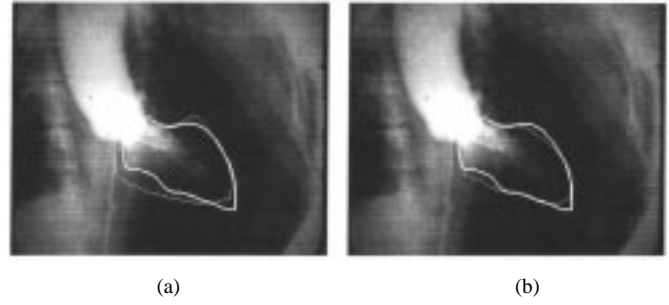


Fig. 19. ES classifier (thin) and regression boundary (thin) compared with its GT boundary (thick). (a) ES frame: GT and classifier. (b) ES frame: GT and regression.

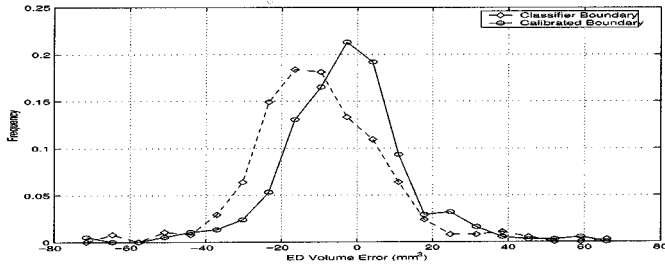


Fig. 16. Histograms of the signed ED volume errors of the classifier outputs and calibrated results. Cross-validation groups $K = 5$. The number of studies is 375.

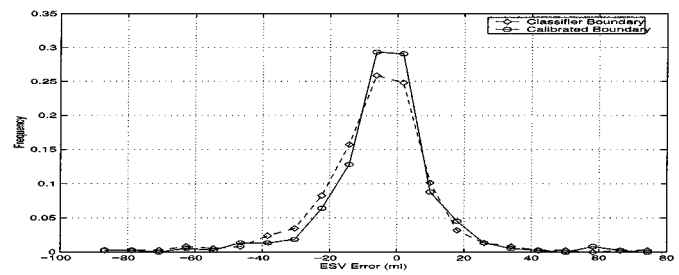


Fig. 20. Histograms of the signed ES volume errors of the classifier outputs and calibrated results. Cross-validation groups $K = 5$. The number of studies is 375.

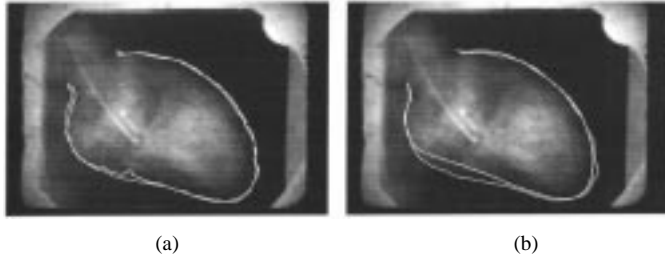


Fig. 17. (a) ED gray-scale image overlaid with the GT boundary (gray) and the classifier boundary (bright). (b) ED gray-scale image overlaid with the GT boundary (gray) and the regression boundary (bright).

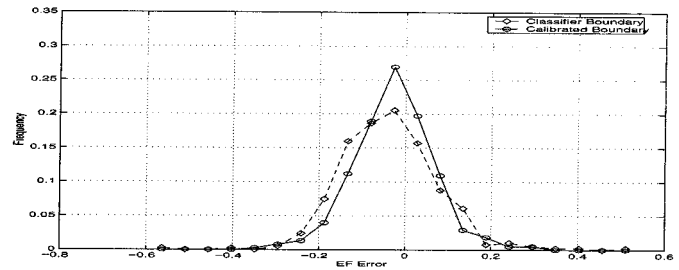


Fig. 21. Histograms of the signed EF errors of the classifier outputs and calibrated results. Cross-validation groups $K = 5$. The number of studies is 375.

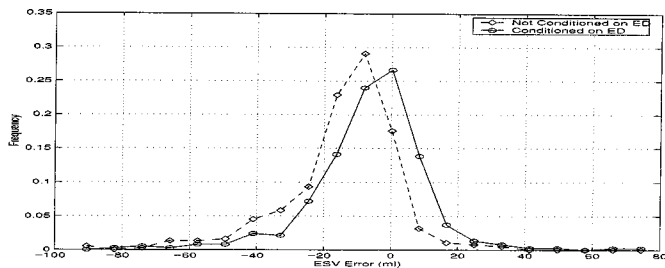


Fig. 18. Histograms of the signed ES volume errors of the classifiers not conditioned and conditioned on the ED results. Cross validation groups $K = 5$.

closed by two LV boundaries, the Hamming distance between the two areas is

$$h(A_1, A_2) = |A_1 \cap \overline{A_2}| + |\overline{A_1} \cap A_2| \quad (26)$$

An example of the Hamming distance is shown in Fig. 8. Let ρ_1 and ρ_2 be the perimeter of the two boundaries. The mean absolute boundary error is defined by

$$e_b = \frac{h(A_1, A_2)}{\frac{\rho_1 + \rho_2}{2}}. \quad (27)$$

TABLE I
ES VOLUME AND EF ERRORS OF THE AUGMENTED REGRESSION WITH AND WITHOUT THE DEPENDENCE ON ED

Error Type	Without ED Augmentation	With ED Augmentation
Mean of the Abs. ES Volume Error	12.4ml	10.4ml
Variance of the Abs. ES Volume Error	150	97

The boundary error reflects the shape deviation between the ABD results and the hand-traced results.

IV. EXPERIMENTS AND RESULTS

A. Data Preparation

The system was trained and tested with cross validation on 375 LVGs. Of these, 210 were of 35 patients studied serially over one year following acute myocardial infarction [28]. The remaining LVGs were from diagnostic cardiac catheterizations.

All LVGs were recorded on 35-mm cine film and included images of a calibration figure of known dimension. The images in the selected cardiac cycle were digitized using a frame grabber

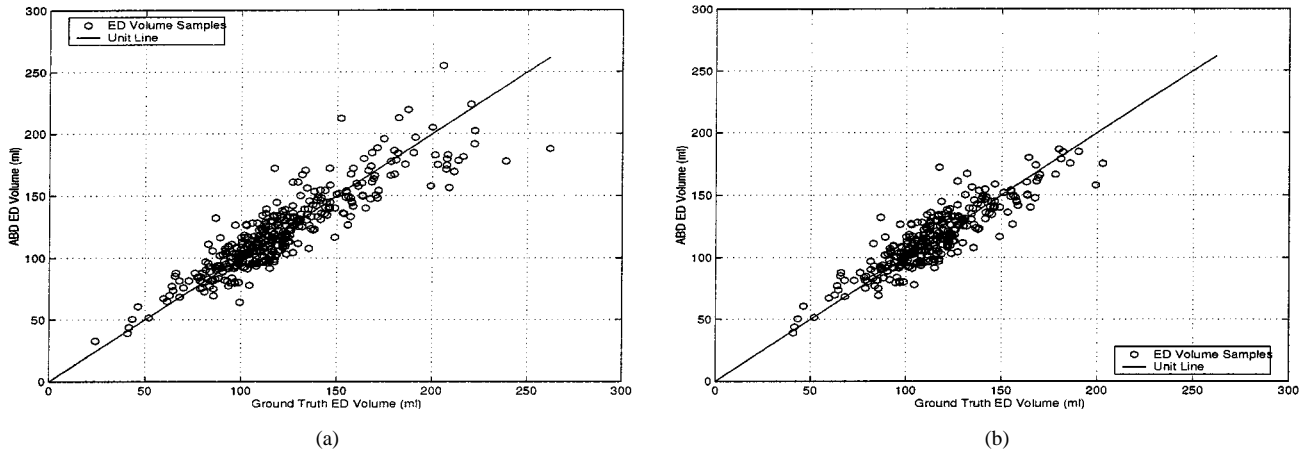


Fig. 22. (a) ED volume scatter plot before rejection. (b) ED volume scatter plot after rejection. The line is the identity line in both figures.

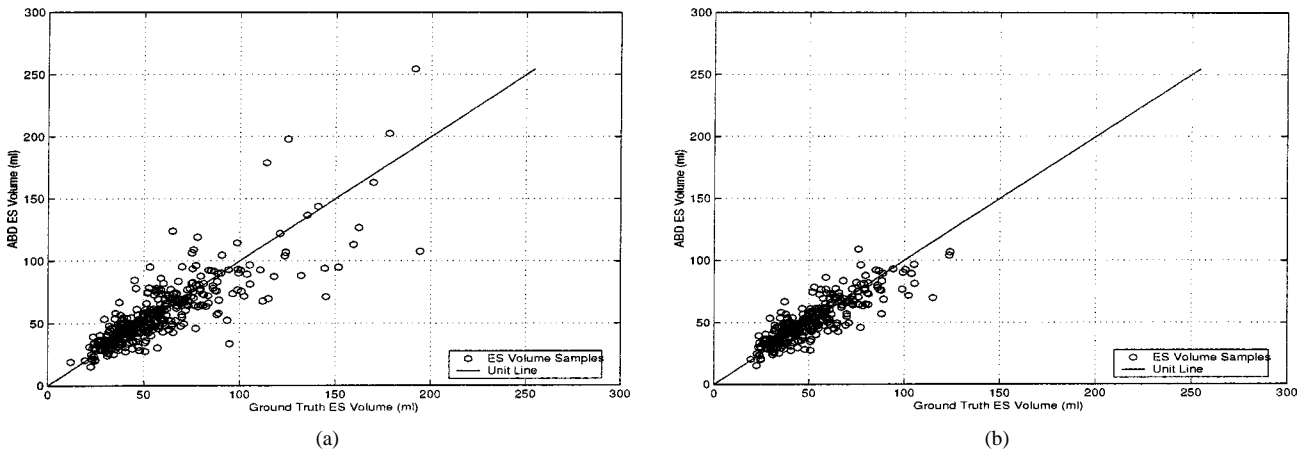


Fig. 23. (a) ES volume scatter plot before rejection. (b) ES volume scatter plot after rejection. The line is the identity line in both figures.

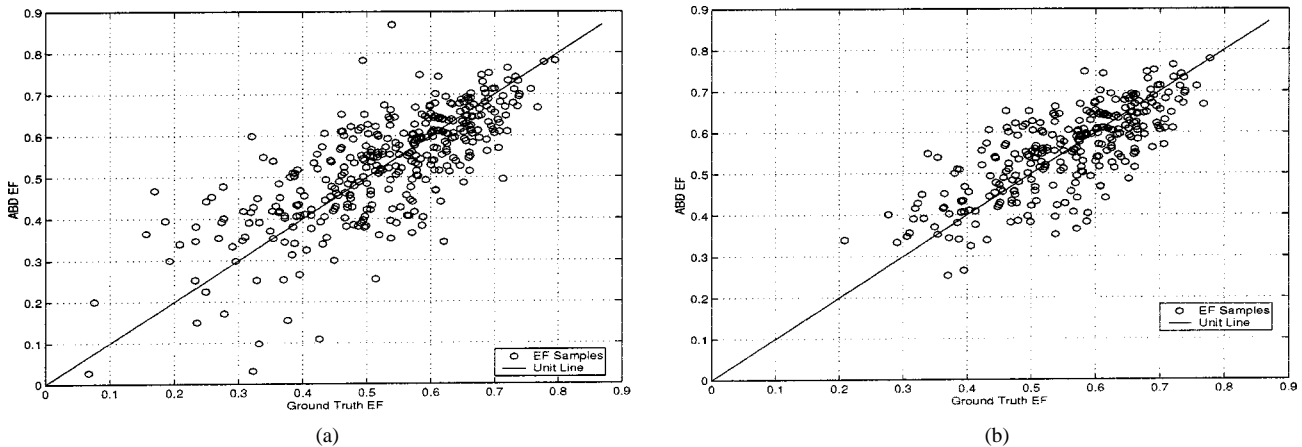


Fig. 24. (a) EF scatter plot before rejection. (b) EF scatter plot after rejection. The line is the identity line in both figures.

(Imaging Technology, Woburn, MA) seated in a personal computer (Intel) at a resolution of 512×480 . Each study of the 375 cases had its ED and ES boundary traced by a human expert. Those hand-traced boundaries were used as the GT for the ABD performance evaluation. From the hand-traced boundary, the GT class region image was generated by filling in the area enclosed by the boundaries. It gave each pixel in the image a GT class. Fig. 9 shows an example of the GT class region image.

The 375 studies were randomly divided into five groups. The following experiments, except stated otherwise, were trained on four groups and tested on the remaining. They were repeated for five times until every study was involved in the testing.

B. Online Processing and Results

A raw systolic image sequence is input into the system. It goes through the preprocessing, the ED region classification, the ED

TABLE II
MEAN ABSOLUTE ERRORS OF THE ABD SYSTEM WITH AND WITHOUT REJECTION

Error Type	Without Rejection	With Rejection
Mean Absolute ED Volume Error	11.6ml	10.5ml
Mean Absolute ES Volume Error	10.4ml	8.2ml
Mean Absolute EF Error	6.9%	6.0%
ED Border Error	1.7mm	1.6mm
ES Border Error	2.3mm	2.2mm
Mean Absolute ED Volume Error	-0.1ml	0.7ml
Mean Absolute ES Volume Error	-0.2ml	-0.3ml
Mean Absolute EF Error	0.1%	0.3%

TABLE III
INTEROBSERVER VARIABILITY IN MEASURING LV VOLUME AND EF

Reference	EDV	ESV	EF	Notes
[18]	20ml	10ml	5%	mean absolute difference
[19]	10.8 ± 1.9ml	5.7 ± 0.8ml	3.9% ± 0.5%	mean absolute difference ± SE
[20]	10ml	6ml	4%	SEE
[21]	3.0 ± 11.8ml	2.6 ± 4.9ml	1.9% ± 4.1%	mean signed difference ± SD
[27]	8.3ml	7.3ml	4ml	SEE
[22]	7.5ml	4.2ml	3.1ml	SEE
[23]			5.1% ± 4.8%	mean signed difference

TABLE IV
LINE FITTING USING GROUND TRUTH AS THE INDEPENDENT VARIABLE AFTER REJECTION FOR THE 375 TRAINING CASES

Measurement	a	γ	SEE
ED Volume	1	0.76	12.6 ml
ES Volume	0.97	0.67	10.1 ml
EF	0.99	0.47	7.5%

shape regression, the ES region classification, and the ES shape regression. Finally, the ED and ES classifier boundaries, the ED, and ES regression boundaries are input to the rejection classifier. It labels the result as accepted or rejected.

1) *Preprocessing*: After the pixel size calibration, noise filtering, and heart rate normalization, the raw systolic LVGs are normalized to have 12 frames. Fig. 10 shows an example of a 12-frame systolic image sequence. The pixel gray scales are then normalized by its CDF within the sequence. Fig. 11 shows the result in which the CDF values were quantized to 256 gray-scale levels between 0–255. The CDF gray-scale normalization effect is obvious since the LV region became more uniform and the contrast at the border region increased. As a matter of fact, the CDF normalization expanded the gray-scale dynamic range so that the separability of the first four PCs among different classes increased. The results are shown in Figs. 12 and 13. Fig. 13 also suggested that four PCs were enough for the classification since the histograms of different classes began to merge in the fourth PC.

2) *ED Region Classification*: The first four PCs of a gray-scale vector at a pixel location is input to the ED region classifier, together with the ED decision gain matrix, a 0.7-M class conditional probability LUT ($M = 2^{20}$) and the prior probability images aligned to the user entered points. The output of the ED region classification is a pixel class image in which each pixel value is a class label. A typical output from the ED classifier is shown in Fig. 14(a). The output is noisy. There are holes in the regions. Some ES class pixels are outside the ED-not-ES class pixels and a curl could be seen near the mitral valve. Fig. 14(b) shows the results after postprocessing. The classifier boundary is ready to be traced.

TABLE V
LINE FITTING USING A BD AS THE INDEPENDENT VARIABLE AFTER REJECTION FOR THE 375 TRAINING CASES

Measurement	a	γ	SEE
ED Volume	0.99	0.77	12.5 ml
ES Volume	0.99	0.72	10.2 ml
EF	0.99	0.53	7.5%

TABLE VI
MEAN ABSOLUTE BOUNDARY ERRORS OF LEE'S CLASSIFIER AND OUR CLASSIFIER

Error Type	Lee's Classifier	Our Classifier
Average ED and ES mean Boundary Error	3.4mm	2.2mm

TABLE VII
BOUNDARY ERRORS OF THE CLASSIFIER OUTPUT AND CALIBRATED RESULT WITH LEE'S CLASSIFIER

Frame	ED and ES Error Statistics	
	Classifier	Regression
Mean of the ED Error	3.1 (mm)	2.1 (mm)
Mean of the ES Error	3.7 (mm)	2.6 (mm)
Variance of the ED Error	3.0	1.2
Variance of the ES Error	3.4	1.3

TABLE VIII
ENTROPIES

Dimension	1	2	3	4
Entropy	18.496	16.445	15.036	14.090

TABLE IX
LUT SIZES ($M = 2^{20}$) FOR THE CLASSIFIER TRAINING

LUT Size	Bins			
	Dimension 1	Dimension 2	Dimension 3	Dimension 4
0.24M	37	24	19	15
0.49M	45	29	22	18
0.75M	50	33	24	20
1.2M	58	37	27	22
1.7M	64	40	29	24

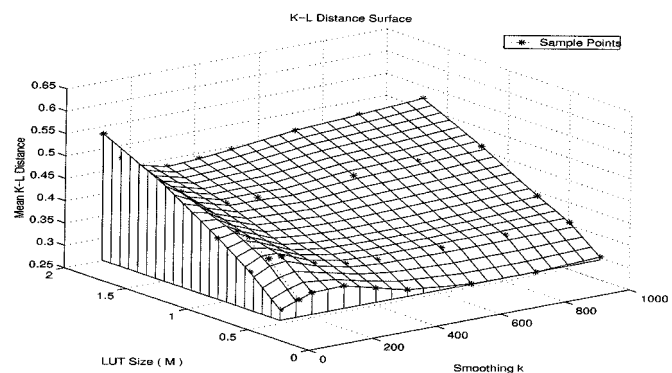


Fig. 25. 3-D mesh surface of the K - L distance (8) obtained with 19974966 training samples.

3) *ED Shape Regression*: Given the raw classifier ED boundary, the user enters three ED points and the regression coefficients, the ED shape regression makes the boundary more heart shaped. Fig. 15 shows an example of the shape regression. The ED volume error histogram taken over 375 studies is shown in Fig. 16.

The shape regression improved the ED volume accuracy by correcting the bias of the classifier results. However, it did not

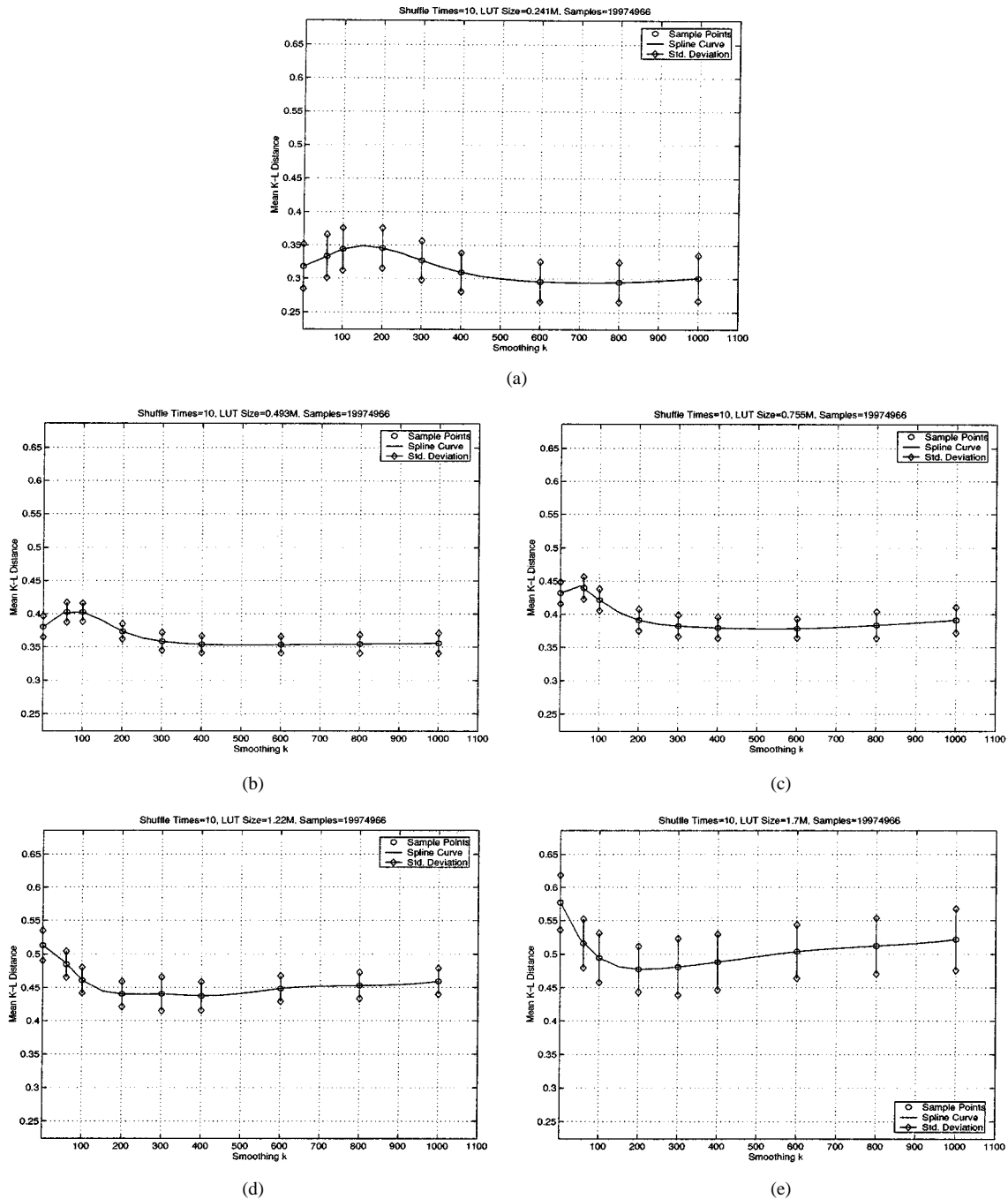


Fig. 26. Projecting the sample curves on the 3-D surface in Fig. 25 onto the axis of smoothing kernel size k . (a)–(e) Graphs of 0.24-, 0.49-, 0.75-, 1.2-, and 1.7-M LUT. They were constructed from 19974966 samples.

improve the boundary accuracy. The ED boundary errors before and after the regression were both 1.7 mm. That is because the boundary error measurement is more critical than the volume measurement. The small errors along the ED boundary can balance the volume error, but worsen the boundary error.

Though the regression improved the overall volume performance, it could bring some errors to a classifier boundary, Fig. 17 shows an example. Here, the regression smoothed out the correct classifier boundary at the inferior wall.

4) *ES Region Classification*: The ES region classifier is conditioned on the ED result. In addition to the inputs to the ED region classifier, the automatically delineated ED boundary

and the class conditional distance histograms are input into the ES region classifier. After a similar postprocessing to that of the ED region classification, the raw ES classifier boundary is obtained.

The improvement of the ES classification conditioned on the ED result is shown in Fig. 18. The conditioned ES classifier tremendously reduced the system bias that the one-layer classification had.

5) *ES Shape Regression*: The ES shape regression utilizes the user-entered three ED points as well as three ES points. An example of the ES regression results is shown in Fig. 19. The ES volume and EF error histograms are shown in Figs. 20 and 21.

TABLE X
SEGMENT SLOPE IN FIG. 26(a)–(e)

LUT Size	Slope of the Segment between							
	1–60	60–100	100–200	200–300	300–400	400–600	600–800	800–1000
0.24M	0.0003	0.0003	0.0000	-0.0002	-0.0002	-0.0001	-0.0000	0.0000
0.49M	0.0004	-0.0000	-0.0003	-0.0002	-0.0000	-0.0000	0.0000	0.0000
0.75M	0.0001	-0.0005	-0.0003	-0.0001	-0.0000	-0.0000	0.0000	0.0000
1.2M	-0.0005	-0.0006	-0.0002	0.0000	-0.0000	0.0001	0.0000	0.0000
1.7M	-0.0010	-0.0005	-0.0002	0.0000	0.0001	0.0001	0.0000	0.0000

Since the conditioned ES classifier reduced the classifier result bias, the ES volume improvement was not as big as the ED volume improvement. However, the regression had an apparent improvement on the EF result. That is because the regression makes the ES boundaries more heart shaped, as shown by the reduction in the ES border error from 2.7 to 2.3 mm.

With the dependence of the ES boundary regression on the three ED points, the calibrated ES results were improved from the non-ED-augmented regression. The comparison is shown in Table I.

6) *Rejection*: The cross-validated classifier ED and ES boundaries and regressed ED and ES boundaries are input to the rejection classifier. Fig. 22 shows the scatter plots of the ED volume before and after the rejection. The ABD algorithm consistently underestimated large LVs whose ED volumes were bigger than 200 mL. Fig. 23 shows the scatter plots of the ES volume before and after the rejection. LVs with a large ED or ES volume were rejected by this classifier. Due to the errors of the large LVs, the EF scatter plot, shown in Fig. 24(a), had some bad performance for small EF values before the rejection. As shown in Fig. 24(b), these were later rejected by the rejection classifier.

Table II shows the mean absolute and signed volume and EF deviations before and after the rejection. The rejection did not introduce much bias to the final results as the magnitudes of the mean signed differences were barely changed before and after the rejection. Nevertheless, the performance, in terms of the absolute difference, on ED volume, ES volume, and EF were all improved, apparently as a result of the rejection classifier.

C. Performance Evaluations

1) *Comparison with Human Interobserver Variability*: Due to the lack of a gold standard for measuring LV volume *in vivo*, manual tracing of the LVG has been used as the accuracy standard. However, manual tracing is subject to interobserver variability. Our performance evaluation compared the difference between the ABD and hand-traced results with the magnitude of human interobserver variability. Table III shows some published results on human interobserver variability.

Comparing the results after rejection in Table II with the first two rows in Table III, our mean absolute ED volume difference is comparable to human interobserver variability. Our mean absolute ES volume difference is between the published interobserver variability and our mean absolute EF difference is slightly higher than the published data.

The last three rows of Table II shows our signed difference of ED volume, ES volume, and EF between the ABD results and hand-trained results. Compared with the fourth and last row

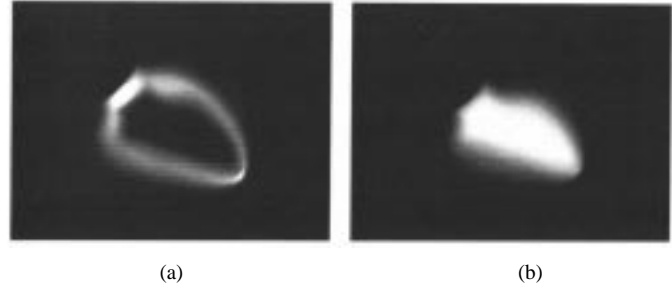


Fig. 27. (a) General prior probability image of ED-not-ES class. (b) General prior probability image of ES class.

of Table III, our cross-validated results had a smaller signed difference than the published data about human interobserver variability. That might be because the published data were tested on a small sample of LVG, about 18 to 40 studies varying from case to case, and a small sample test tends to have a bias. Our results were tested on a 375 case database, thus, it is unlikely to show the bias.

Furthermore, we least square fit the scatter plots in Figs. 22–24 with $y = ax$ in order to assure ourselves that the coefficient a has a value near one. We define the coefficient of determination γ and standard error of the estimate (SEE) by

$$\gamma = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

$$\text{SEE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

where y_i is the i th true y sample, \hat{y}_i is the i th estimated y with $y = ax$, and n is the number of samples, $\bar{y} = (1/n) \sum_{i=1}^n y_i$.

Tables IV and V show the fitted slope a , the coefficient of determination γ , and SEE between the GT and ABD results after the rejection. Compared with Table III, our results after the rejection were slightly above human interobserver variability.

2) *Comparison with Lee's Result*: Table VI shows the boundary errors of Lee's classifier and our classifier. Lee's classifier had a one-layer structure for both ED and ES. It did not include the CDF gray-scale normalization. Lee used Mahalanobis distance histograms to look up the class conditional probabilities of the feature vector and generated the prior probabilities with only two AoV points. We used the raw results from our ED classifier² to compare with Lee's result on the

²Our ED classifier can produce the ES boundary by tracing the ES region composed of only the ES pixels.

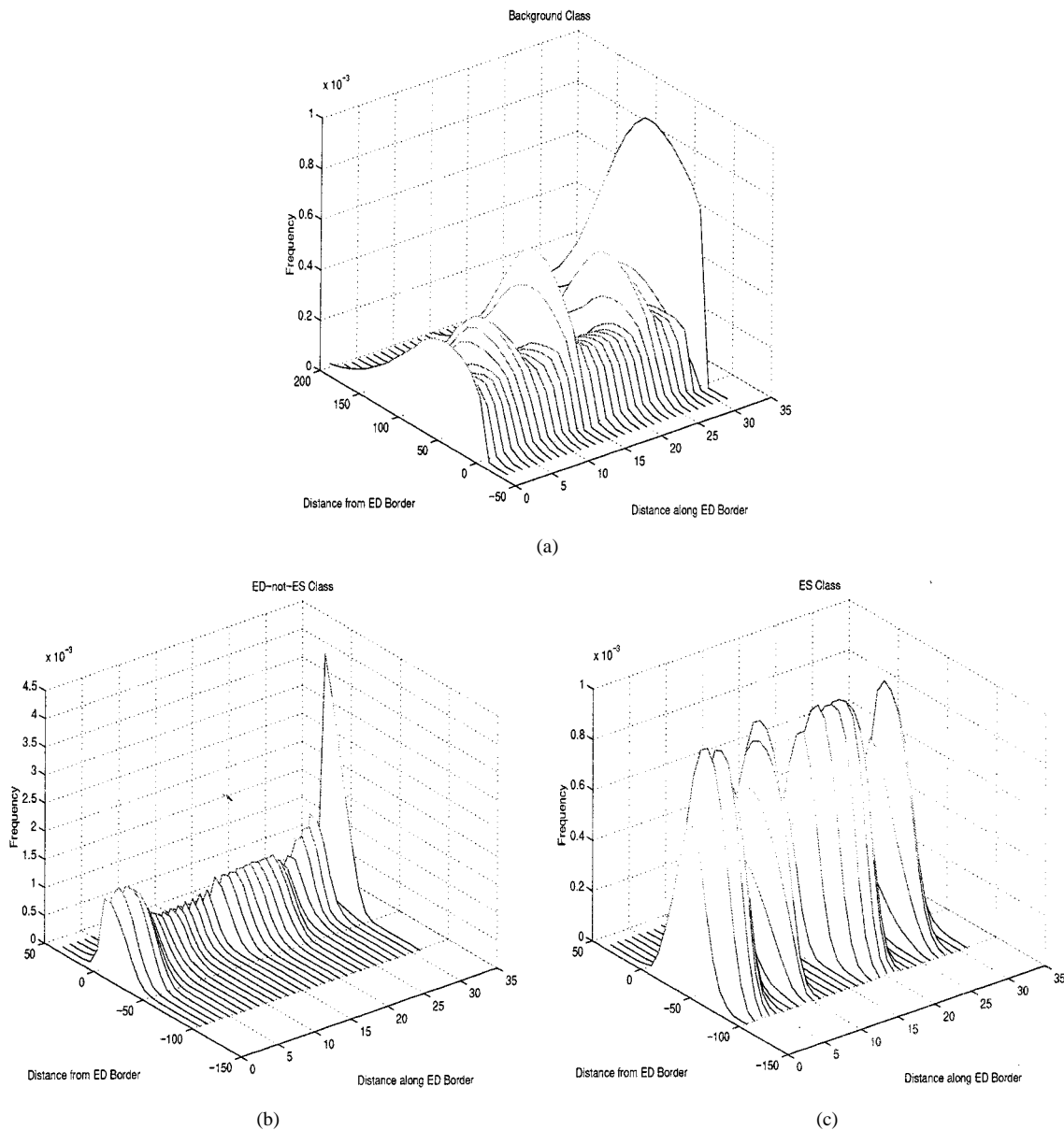


Fig. 28. Class conditional histograms of the signed distance between the pixels in the image frame and the ED boundary. (a) Background class. (b) ED-not-ES class. (c) ES class.

same data. The comparison shows that our border performance was better by 1.2 mm.

3) *Comparison with Suri's Result:* The shape regression was applied to Lee's classifier [11]. The cross validation was the leave-one-out method. The performance was measured by the boundary error, as defined in the previous section. Some statistics of the errors are shown in Table VII. Using the shape regression scheme, the average mean boundary error of ED and ES frames improved from 3.4 to 2.4 mm. This outperformed Suri's method [16], whose average ED and ES error was about 2.7 mm on the same data.

D. Offline Training

Some offline classification training results are briefly discussed here. They are the class conditional probability LUT size

m and smoothing kernel size k , the class prior probabilities, the class conditional histograms of the distance between a pixel, and the delineated ED boundary.

1) *LUT Size and Smoothing:* There are 40 056 750 sample pairs of a 4-D feature vector and its GT class in the ROI of 375 cases. The discrete entropy of each dimension is shown in Table VIII. Table IX lists the number of bins on each dimension given the LUT size m .

From the 40 056 750 samples, 19 974 966 points were randomly picked out to construct a series of LUTs in Table IX, smoothed with $k \in \{1, 60, 100, 200, 300, 400, 600, 800, 1000\}$. The experiment was repeated ten times with the shuffled samples. Fig. 25 shows the three-dimensional (3-D) surface of the mean bidirectional K - L distance (8) of the mixture distribution. As seen in this figure, there is a flat region on the K - L distance surface. That is the place where the distribution is not sensitive

to the LUT size and the smoothing kernel size so that a generalization can be obtained.

Fig. 26 shows the graphs of projecting the 3-D surface in Fig. 25 onto the smoothing kernel size axis k . As seen from those graphs, the LUT of 0.24 M has a big hump at the small k s. The samples in the neighboring bins have a negative influence on the cross-validated bin. The nearest neighbors do not have similarity to the cross-validated samples. That could be due to that too many samples are quantized into a bin. Moreover, the minimal K - L distance obtained at $k = 700$ is not much different from the one at $k = 1$. Its variance among the shuffles at each sample point is quite big. On the other hand, as the LUT size became larger, in the graph of 1.7-M LUT, for instance, the minima are more like a valley than a flat region, together with a big variance. That implies that the observed data are likely to be sparse in the LUT. Table X shows the slopes of the segments in each graph of Fig. 26. The flattest graphs are the ones of the 0.49- and 0.75-M LUT. Also, the shuffle variances in 0.49- and 0.75-M LUTs are relatively small, which implies a good generalization. The suitable k is between 200–800.

Fig. 27 shows the prior probability images of the ED-not-ES class and ES class, in which the bright areas are the high-probability regions. They were generated by aligning all the GT boundaries with their AoV angles and long axes, and estimated by frequency counting. That alignment makes them less sensitive to the exact coordinates location of the user-entered points.

Fig. 28 shows the class conditional histograms of the distance between a pixel and the delineated ED boundary. They provides the third term on the left-hand side of (4).

V. CONCLUSION

Our results indicate that the ABD process presented in this paper is able to detect and delineate the endocardial contour of the LV from contrast ventriculograms with an accuracy comparable to the magnitude of human interobserver variability. The success of this process is due to the integration of knowledge concerning human cardiac anatomy and physiology with the imaging data.

The region classifier embodies knowledge concerning the expected regional movement of the ventricular wall during systole. With it we sought to emulate the human observer's practice of reviewing wall motion through the cardiac cycle to help define the endocardial contour. The region classifier also utilizes the shape and size knowledge contained by the user-entered ED and ES points to adjust the prior probability images so that the classifier can focus on the area where the LV is likely to be.

The regression calibration embodies knowledge concerning the expected shape of the LV endocardium. Just as human observers require training to recognize heart contours, we sought to provide this to the ABD process. The shape analysis was not only performed on each image's candidate border, but also between ED and ES image frames. The latter captures the expectation that the ES border will bear some resemblance to the ED border.

The third component is the method for rejecting studies whose image sequence are likely to produce suspicious or unphysiological borders. Just as clinical ventriculograms are rejected for manual tracing if there is poor contrast quality, we sought a method to warn the user of these problems in the ABD process.

Finally, the large number of training studies and the normalization before pixel region classification helped, in large part, to ensure that the methodology is able to generalize rather than memorize the training data.

Future work includes improving the ABD performance on the large-volume LV cases. Due the insufficient representation of those large LVs in the database, the ABD has a systematic underestimation of the volume for those cases. A possible research direction is to set up an independent shape regression calibration for those cases where initial ABD shows them to be large-volume cases.

ACKNOWLEDGMENT

The authors want to thank the staff at the Cardiovascular Research and Training Center, School of Medicine, University of Washington, Seattle, for organizing the LVG data and the hand-traced boundaries. The authors are thankful for the patience and generosity of the Goodman Ltd. Company, Nagoya, Japan.

REFERENCES

- [1] F. H. Sheehan and J. W. Kennedy, "Ventriculography," in *Diagnostic and Therapeutic Cardiac Catheterization*, 3rd ed. Baltimore, MD: Williams & Wilkins, 1998, pp. 247–263.
- [2] C. K. Chow and T. Kaneko, "Automatic boundary detection of the left ventricle from cineangiograms," *Comput. Biomed. Res.*, vol. 5, pp. 388–410, 1972.
- [3] P. D. Clayton, L. D. Harris, S. R. Rumel, and H. R. Warner, "Left ventricular videometry," *Comput. Biomed. Res.*, vol. 7, pp. 369–379, 1974.
- [4] W. A. Barrett, P. D. Clayton, and H. R. Warner, "Determination of left ventricular contours: A probabilistic algorithm derived from angiographic images," *Comput. Biomed. Res.*, vol. 13, pp. 522–548, 1980.
- [5] W. Barrett, H. Eisenberg, B. Arnold, and P. Scheibe, "High-speed processing of digital intravenous angiographic images for enhancement of coronary bypass grafts and quantitation of left ventricular function," *Comput. Cardiol.*, pp. 101–104, 1982.
- [6] D. L. Pope, D. L. Parker, P. D. Clayton, and D. E. Gustafson, "Left ventricular border recognition using a dynamic search algorithm," *Radiology*, vol. 155, pp. 513–518, 1985.
- [7] M. T. de Figueiredo and J. M. N. Leitao, "Bayesian estimation of ventricular contours in angiographic images," *IEEE Trans. Med. Imag.*, vol. 11, pp. 416–429, Mar. 1992.
- [8] P. V. D. Zwet, G. Koning, and J. Reiber, "Left ventricular contour detection: A fully automated approach," *Comput. Cardiol.*, pp. 359–362, 1992.
- [9] N. Fan, C. C. Huang, B. G. Denys, and P. S. Reddy, "Improved automatic left ventricular border detection in cineangiograms using dynamic spider model," *Comput. Cardiol.*, pp. 363–366, 1992.
- [10] C. Costa and L. Moura, "Automatic detection of LV contours in nuclear medicine using geometrical information and a neural net," *Comput. Cardiol.*, pp. 557–560, 1995.
- [11] C. K. Lee, "Automated boundary tracing using temporal information," Ph.D. dissertation, Dept. Elect. Eng., Univ. Washington, Seattle, WA, 1994.
- [12] J. S. Suri and R. M. Haralick, "Systematic error correction in automatically produced boundaries in low contrast ventriculograms," in *Int. Pattern Recog. Conf.*, 1996, pp. 361–365.

- [13] C. E. Shannon, *A Mathematical Theory of Communication*. Urbana, IL: Univ. Illinois Press, 1949.
- [14] D. O. Loftsgaarden and C. P. Quesenberry, "A nonparametric estimate of a multivariate density function," *Annu. Math. Statist.*, vol. 36, pp. 1049–1051, 1965.
- [15] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*. Reading, MA: Addison-Wesley, 1992, vol. 1 and 2.
- [16] J. S. Suri, "Contour generation of motion objects in fuzzy and low contrast multiple image frames: An application in cardiological imaging," Ph.D. dissertation, Dept. Elect. Eng., Univ. Washington, Seattle, WA, 1997.
- [17] A. Guo, "Rejection analysis in automatic left ventricle boundary tracing system," Masters thesis, Dept. Elect. Eng., Univ. Washington, Seattle, WA, 2000.
- [18] P. F. Cohn, J. A. Levine, G. A. Bergeron, and R. Gorlin, "Reproducibility of the angiographic left ventricular ejection fraction in patients with coronary artery disease," *Amer. Heart J.*, vol. 88, pp. 713–720, 1974.
- [19] B. R. Chaitman, H. DeMots, J. D. Bristow, J. Rosch, and S. H. Rahimtoola, "Objective and subjective analysis of left ventricular angiograms," *Circulation*, vol. 52, pp. 420–425, 1975.
- [20] W. J. Rogers, L. R. Smith, W. P. Hood, Jr., J. A. Mantle, C. E. Rackley, and R. O. Russell, Jr., "Effect of filming projection and interobserver variability on angiographic biplane left ventricular volume determination," *Circulation*, vol. 59, pp. 96–104, 1979.
- [21] H. Sigel, W. Nechwatal, P. Kress, and M. Stauch, "Interobserver and intermethod variation in evaluation of regional wall motion of the left ventricle," *Cardiovasc. Intervention. Radiol.*, vol. 6, pp. 14–19, 1983.
- [22] F. H. Sheehan, D. K. Stewart, H. T. Dodge, S. Mitten, E. L. Bolson, and B. G. Brown, "Variability in the measurement of regional ventricular wall motion from contrast angiograms," *Circulation*, vol. 68, pp. 550–559, 1983.
- [23] G. Koning, M. van den Brand, I. Zorn, G. Loois, and J. H. C. Reiber, "Usefulness of digital angiography in the assessment of left ventricular ejection fraction," *Catheter Cardiovas. Diagnos.*, vol. 21, pp. 185–194, 1990.
- [24] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models—Their training and application," *Comput. Vis. Image Understanding*, vol. 61, no. 1, pp. 38–59, 1995.
- [25] L. Sui, "Automated left ventriculogram boundary delineation," Ph.D. dissertation, Dept. Bioeng., Univ. Washington, Seattle, WA, 2000.
- [26] S. Hochreiter and J. Schmidhuber, "Flat minima," *Neural Comput.*, vol. 9, no. 1, pp. 1–42, 1997.
- [27] H. T. Dodge, F. H. Sheehan, and D. K. Stewart, "Estimation of ventricular volume, fractional ejected volumes, stroke volume, and quantitation of regurgitant flow," in *Angio-Cardiology Current Status and Future Development*. Berlin, Germany: Springer-Verlag, 1986, pp. 99–108.
- [28] H. Hosokawa, F. H. Sheehan, and T. Suzuki, "Measurement of post systolic shortening to assess viability and predict recovery of left ventricular function following acute myocardial infarction," *J. Amer. Coll. Cardiol.*, vol. 35, no. 7, pp. 1842–1849, 2000.

Lei Sui received the Bachelor's degree in electrical engineering and the Master's degree in computer engineering from Shanghai Jiaotong University, Shanghai, China, and the Ph.D. degree in bioengineering from the University of Washington, Seattle.

He is currently a Senior Scientist in the Ultrasound Group, Siemens Medical System, Issaquah, WA. His research interest includes image processing, computer vision, and pattern recognition.

Robert M. Haralick (F'84) was born in Brooklyn, NY, on September 30, 1943. He received the B.A. degree in mathematics, the B.S. degree in electrical engineering, the M.S. degree in electrical engineering, and the Ph.D. degree from the University of Kansas, Lawrence, in 1964, 1966, 1967, and 1969, respectively.

In 1969, he joined the faculty of the Electrical Engineering Department, University of Kansas, where he was a Professor from 1975 to 1978. In 1979, he joined the Electrical Engineering Department, Virginia Polytechnic Institute and State University, Blacksburg, where he was a Professor and Director of the Spatial Data Analysis Laboratory. From 1984 to 1986, he was Vice President of Research at Machine Vision International, Ann Arbor, MI. From 1986 to 2000, when he retired, held the Boeing Clairmont Egvedt Professorship in the Department of Electrical Engineering, University of Washington, Seattle. He has made a series of contributions in computer vision. In the high-level vision area, he has been involved with inferring 3-D geometry from one or more perspective projection views. He has also identified a variety of vision problems, which are special cases of the consistent labeling problem. His papers on consistent labeling, arrangements, relation homomorphism, matching, and tree search translate some specific computer vision problems to the more general combinatorial consistent labeling problem and then discuss the theory of the look-ahead operators that speed up the tree search. This gives a framework for the control structure required in high-level vision problems. He has also extended the forward-checking tree search technique to propositional logic. In the low-level and mid-level areas, he has been involved with image texture analysis using spatial gray-tone co-occurrence texture features. These features have been used with success on biological cell images, X-ray images, satellite images, aerial images, and many other kinds of images taken at small and large scales. In the feature detection area, he has developed the facet model for image processing, which states that many low-level image processing operations can be interpreted relative to what the processing does to the estimated underlying gray-tone intensity surface of which the given image is a sampled noisy version. The facet papers develop techniques for edge detection, line detection, noise removal, peak and pit detection, as well as a variety of other topographic gray tone surface features. His involvement in shape analysis and extraction uses the techniques of mathematical morphology. He has developed the morphological sampling theorem, which establishes a sound shape/size basis for the focus of attention mechanisms, which can process image data in a multiresolution mode, thereby making some of the image feature extraction processes execute more efficiently. He has also developed recursive morphological algorithms for the computation of opening and closing transforms. The recursive algorithms permit all possible sized openings or closings for a given structuring element to be computed in constant time per pixel. In the area of document image understanding, he has been responsible for the development of comprehensive ground-truthed data bases consisting of over 1600 document image pages, most in English and some 200 pages in Japanese. The databases are issued in CD ROMs and are being used all around the world by people who are developing character recognition methodologies and techniques for document image structural decomposition. He has developed algorithms for document image skew angle estimation, zone delineation, and word and text line bounding box delineation. Most recently, he has helped influence the computer vision community to be more sensitive to the needs of computer vision performance characterization and covariance propagation. He was a computer vision area editor for *Communications of the Association for Computing Machinery* and an associate editor for *Computer Vision, Graphics and Image Processing*. He has served on the Editorial Board of *Real Time Imaging* and the *Journal of Electronic Imaging*. He has authored or co-authored over 500 papers.

Prof. Haralick is a Fellow of the International Association for Pattern Recognition (IAPR). In August 1998, he completed a two-year term as the president of the International Association for Pattern Recognition. He has served on the Editorial Board of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE and THE IEEE TRANSACTIONS ON IMAGE PROCESSING AND PATTERN RECOGNITION.

Florence H. Sheehan received the B.S. degree from the Massachusetts Institute of Technology, Cambridge, and the M.D. degree from the University of Chicago.

Following her medicine residency at the Medical College of Virginia, she was a Clinical Associate at the National Heart, Lung, and Blood Institute, National Institutes of Health (NIH), Bethesda, MD. She completed her cardiology fellowship at the University of Washington, Seattle, and is currently a Research Professor in Medicine there. Her research has been oriented toward the development and validation of methods for quantitative analysis of cardiac images, and the application of these methods to clinical and experimental research. Over the past decade, she has been involved with 3-D echocardiography and image processing.