

## 10—3 A Knowledge Based Boundary Delineation from Left Ventriculograms

Lei Sui

Department of Bioengineering

Robert M. Haralick

Department of Electrical Engineering

Florence H. Sheehan

Cardiovascular Research and Training Center

University of Washington

Seattle, WA 98195

### 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 multi-stage method to automatically delineate the LV boundary at end diastole and end systole is discussed in this paper. It has a mean absolute boundary error of about 2mm 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 multi-image pixel region classification, a shape regression and a rejection classification. The method was trained and tested on a database of 375 studies whose ED and ES boundary have been manually traced as the ground truth. The cross-validated results presented in this paper shows that the accuracy is close to and slightly above interobserver variability.*

### 1 Introduction

The LVG boundary at end diastole (ED) and ends at end systole (ES) are particularly interested in, because of their importance of assessing the LV function, Tracing the LV boundaries from LVG is a tedious job. Although automated LV boundary delineation (ABD) in LVGs has long been sought [1, 2, 3, 4, 5, 6], no accurate methods have been reported due to the challenges of LVG variation. Moreover, most methods were tested on only a limited number of LVG samples so that no consistent accuracy evaluation concerning LV volume and ejection fraction (EF) was available. Thus, the ABD methods used in clinics performed poorly.

Most previously reported methods have difficulties in handling the large variation in LVGs because the information in LVGs and the human knowledge in-

formation were not effectively utilized. First, most of the methods only processed a single image at a time to detect the boundary [1, 2, 4]. Such methods ignored the helpful LV movement information through time in the cardiac cycle. Second, the edge based methods [2, 3, 4, 5] used the gradient maxima whose locations, however, do not correspond to the boundary positions and the region based methods [1, 6] assumed a Gaussian distribution of the gray scales in LV and background, which is not true. Third, the LV shape information was used only to model the LV boundary [2, 3] rather than the LV region. Fourth, all previous existing algorithms processed all input LVGs indiscriminately. They did not screen out the difficult-to-process cases. On contrast, human experts do utilize movement information in the time domain to resolve boundary fuzziness. They are trained to have anatomical knowledge about the LV and to reject cases with bad image quality. Our method implement these sources of knowledge in the ABD system.

Our method is composed of three stages. The first stage is a pixel region non-parametric classifier. It segments the LV region from the background and obtains a raw LV boundary. The second stage is a shape regression [8] that corrects the systematic error of the region classifier. The last stage is a rejection classifier that flags unreliable results. 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 for expected shape and size.

In the next section, our method is discussed in detail. Some experimental results are presented in Section 3. The conclusion and discussion are in Section 4.

## 2 Method

Figure 1 shows the flow chart of our three-stage ABD system. Because 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.

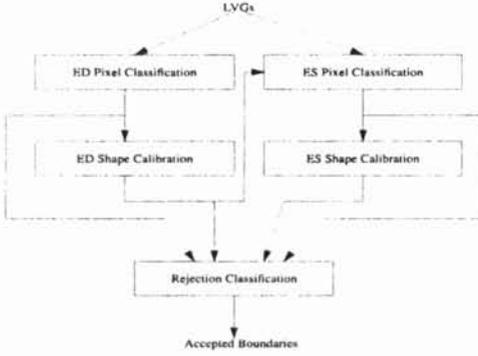


Figure 1: Sequential processing for ED and ES region.

### 2.1 Pixel Region Classification

The Bayesian pixel classifier labels each pixel in the image frame of the 12 dimensional systole images with one of 3 classes, background, ED-not-ES and ES based on the systole gray scale vector, shown in Figure 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 location as the LV border passes over the pixels during contraction.

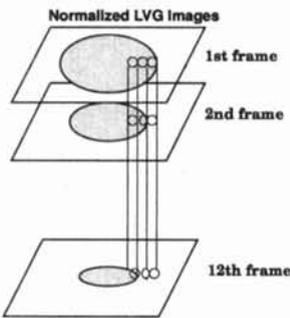


Figure 2: The Gray Scale Vector.

Before classification, the systole images from ED to ES go through pixel size calibration [6], noise filtering [6] and are normalized to 12 frames to adjust for variability in frame rate and heart rate [6]. The pixel size calibration makes the pixel square in millimeter

( $mm$ ) unit. This normalizes the image size and benefits the performance evaluation in  $mm$  unit. Noise filtering is implemented by the gray scale morphological opening and closing [7]. This preserves the location of gray scale boundary transition at the same time filtering out the noise. The heart rate normalization is essential to the classification in that it makes the raw feature vector dimension a constant 12.

The gray scale values over the 12 image frames are then normalized by the cumulative distribution function (CDF) within the region of interest (ROI) in the sequence, where the ROI is the area outside which no part of the LV would occur. 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,  $y_i(u, v)$  is the pixel value at  $(u, v)$  of the  $i^{th}$  image frame after normalization. The normalization is with respect to the entire systole sequence because this keeps the gray scale value change pattern through systole. The CDF normalization expands the gray scale dynamic range and increases the separability among classes.

The 12 CDF normalized pixel values at each pixel location form a vector,  $Y = (y_1, \dots, y_{12})$ . The vector is projected onto the eigenvectors of the second moment matrix of the sample pixel value vectors,  $\mathbf{Y} = \sum_{j=1}^M Y_j Y_j^T$ , 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 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 expect gain. Given the gain matrix  $\mathbf{G} = \{g(c, c') | c, c' \in \{1, 2, 3\}\}$ , where 1 means the background class, 2 means the ED-not-ES class, 3 means the ES class, and  $g(c, c')$  is 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

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

The Bayesian rule classifies  $X$  at  $(u, v)$  to class  $c$  if the assignment to class  $c$  yields the maximal gain.

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

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

$p(c|X, (u, v))$  is the posterior probability of class  $c$  given the feature vector  $X$  at location  $(u, v)$ . 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  $(u, v)$ ,  $p(X|c)$  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 delineated ED boundary given class  $c$ . This is the way that the information from the delineated ED boundary is introduced to the ES boundary detector.

After the classification, the pixels of the same class are grouped together. The largest area consisting of Class 2 and 3 pixels is selected as the ED region and the ES region is the largest area composed of Class 3 pixels. The binary morphological opening and closing are used to smooth the region boundary. After that, the raw ED or ES boundary is traced from the ED or ES region and represented by 100 evenly spaced points.

The classifier training includes estimating the class conditional probability  $p(X|c)$ , the prior probability  $p(c|(u, v))$  and the 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  $p(X|c)$  non-parametrically from a 4D look-up table (LUT). The LUT divides the 4D space into non-uniformly sized hyperrectangular bins and estimating the probability in each hyperrectangle by simply counting. 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. The quantization and bin size allocation depend on the marginal entropy on each dimension and are generalized by the cross validated Kullback-Liebler distance and  $k$  nearest neighbors ( $k - NN$ ).

The class prior probabilities at each pixel location are the knowledge about the LV shape, size and position. A set of ground truth class region images generated from the hand traced boundaries are aligned to a common LV region by fitting their AoV angles

and long axes with the least squares. The LV AoV angle and long axis are defined by the AoV end points and apex. This alignment efficiently keeps the LV region information. The probability of each class at each pixel location is estimated by frequency counting from the aligned ground truth class region images and smoothing with a  $7 \times 7$  square template. By this way, we formulate the high level LV anatomical knowledge in terms of LV region.

Given the delineated ED boundary, the shortest distance between a pixel location and the boundary can be computed. The distance is signed. 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 correspondent segment on the ED boundary. The class conditional distance probabilities are estimated from the histograms.

## 2.2 Shape Regression

It is observed that the classifiers have some systematic errors which makes the raw classifier boundaries out of the LV like shape. In order to calibrate those errors, a shape regression is designed. An LV boundary is taken as a 200 dimensional vector by concatenating its vertex coordinates into a vector. The boundary vector has its coordinates in the LV shape space defined by the eigenvectors of the second moment matrix of the sample ground truth boundary vectors. The regression is to transform the 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 ground truth boundary vectors and 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 3 points. As a result, the regressed boundary coordinates are the linear combinations of the coordinates in the shape space and the quadratic terms of the user entered points.

Given

- A raw boundary vector,  $Y = (r_1, c_1, \dots, r_{100}, c_{100})$ , from the classifier
- Two AoV points and apex coordinates entered by the user

- A matrix  $\mathbf{A}$ , whose columns are the full set of the eigenvectors of the sample second moment matrix of the ground truth boundary space
- A regression coefficient matrix  $\mathbf{C}_{(P_2+t) \times P_3}$  determined offline by training

The regressed boundary is

$$\hat{B} = ([Y \mathbf{A}_{P_2}] : T) \mathbf{C} \mathbf{A}'_{P_3} \quad (5)$$

where  $\hat{B}$  is the calibrated boundary,  $T$  is the augmentation terms coming from the 3 user input points,  $\mathbf{A}_{P_2}$  and  $\mathbf{A}_{P_3}$  are matrices of the first  $P_2$  and  $P_3$  eigenvectors in  $\mathbf{A}$ , the prime denotes the transpose. For the ED boundary, the regression vector is augmented with the full quadratic terms of the three user entered ED points which form 28 terms. For the ES boundary, the regression vector is augmented with the full quadratic terms of the three user entered ES points and some partial quadratic terms of the three ED points.

The regression coefficient estimation is implemented as a least squares fitting given the ground truth boundary vectors and the classifier boundary vectors. The numbers of principal components involved in the regression,  $P_2$  and  $P_3$ , are optimized to avoid the potential generalization problem. The partial quadratic terms which represent the information ED boundary contains for the ES boundary is also determined by the least squares fitting.

### 2.3 Rejection Classifier

From the raw and calibrated boundaries, two sets of parameters including the ED and ES volumes, EF, ED and ES areas can be computed. Also the ED boundary difference and ES boundary difference between the classifier and regression results can be obtained. The rejection classifier takes those parameters as the rejection vector vector. The difference between the two sets of the parameters indicates, more or less, the unreliability of the border detection. Given the classification weights  $Y_{7 \times 1}$  on the rejection feature vector  $U_{1 \times 7}$  and a decision threshold  $t_2$ , a study is rejected if the difference is greater than the threshold,

The rejection classifier is trained to reject roughly 5% of the worst cases in terms of ED volume, ES volume and EF. The training is the least squares fitting based and optimized by minimizing the rejection cost.

### 2.4 Performance Evaluation and System Optimization

Given a LV boundary, the LV volume can be computed [12]. The ejection fraction (EF) can be computed if the ED and ES volume are available. The mean absolute ED volume deviation, ES volume deviation and EF deviation between the ABD results and

hand traced results and their standard error estimations (SEE) are used for the ABD system performance evaluations. Beside those, the boundary error between two boundaries is defined as

$$e_b = \frac{h(A_1, A_2)}{\frac{\rho_1 + \rho_2}{2}} \quad (6)$$

where  $h(A_1, A_2)$  is the Hamming distance between the two boundaries,  $\rho_1$  and  $\rho_2$  are the perimeter of the two boundaries respectively.

The whole system is optimized to minimize the cost after the rejection. The tuned training parameters include the LUT size, the LUT smoothing parameter k-NN, the prior probability smoothing parameter, the distance histogram sectors, the gain matrices (EGMs), the regression dimensions  $P_2$ ,  $P_3$ , and the rejection threshold  $t_2$ .

## 3 Experiment and Results

The system was trained and tested with cross validation on 375 LVGs. 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 ground truth for the ABD performance evaluation. From the hand traced boundary, the ground truth class region image was generated by filling the area enclosed by the boundaries. It gave each pixel in the image a ground truth class.

The 375 studies were divided into 5 groups. The experiments were trained on 4 groups and tested on the remaining. They were repeated for 5 times until every study was involved in the testing.

Figure 3 shows an example of the ED and ES ABD results. Figure 4 shows the scatter plots of the ED volume, ES volume and EF after the rejection.

Table 1: Table of the ABD system performance

	Mean Absolute Error
ED Volume	10.48ml
ES Volume	8.17ml
EF	5.96%
ED Border	1.55mm
ES Border	2.15mm

Table 1 shows that the deviations between the ABD results and the ground truth. The average of the mean absolute boundary error of ED and ES was about 1.85mm. It outperformed Lee [6] and Suri's [8] results where they were about 3.4mm and 2.7mm respectively. It also shows the volume related deviations. Compared with Table 2, the results after the

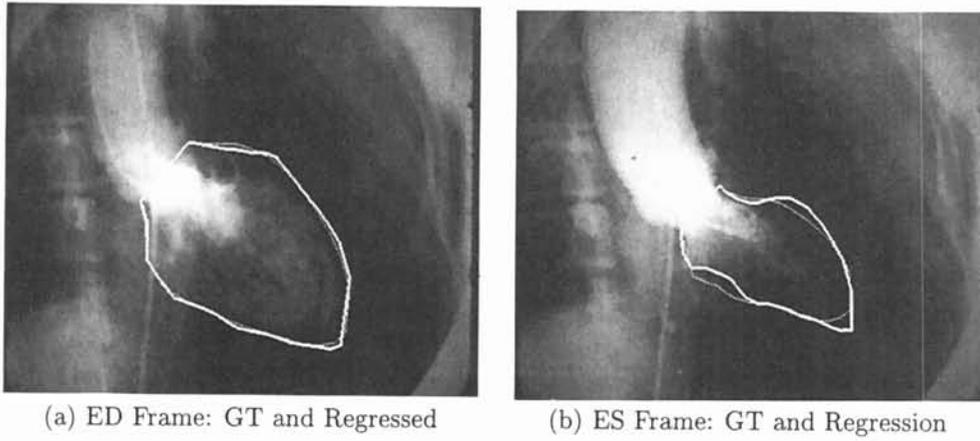


Figure 3: An example of the ABD results (thin) compared with their ground truth (GT) boundaries (thick)

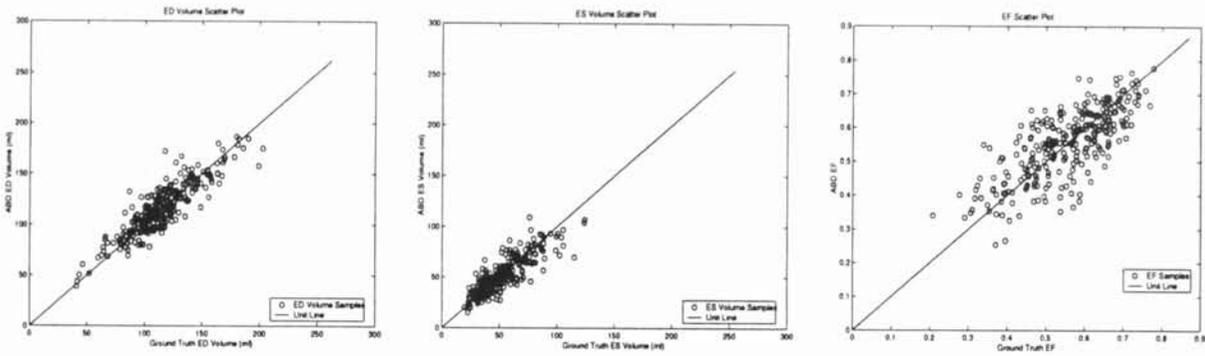


Figure 4: **Left:** The ED volume scatter plot after rejection. **Middle:** The ES volume scatter plot after rejection. **Right:** The EF scatter plot after rejection.

Table 2: Interobserver variability in terms of mean absolute deviation

Reference	EDV	ESV	EF
[9]	20ml	10ml	5%
[10]	10.8ml	5.7ml	3.9%

rejection were slightly above human interobserver variability. Since the ABD was tested on a large database, the performance is expected to be extendible.

#### 4 Conclusion

The results of the present report indicate that the ABD process presented is able to detect and delineate the endocardial contour of the left ventricle 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.

The 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 regression 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 image frames. The latter captures the expectation that the ES border will bear some resemblance to the ED border.

The third component is our code for rejecting studies whose images produced suspicious 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 with the ABD process.

The large number of training studies and the normalization before pixel region classification helped in large part to ensure that the component processes were generalized.

#### 5 Acknowledgment

This project is supported by Goodman Ltd. Co., Nagoya, Japan.

#### References

- [1] C. K. Chow, T. Kaneko, "Automatic Boundary Detection of the Left Ventricle from Cineangiograms", *Computers and Biomedical Research*, Vol. 5, p388-410, 1972.
- [2] P. D. Clayton, L. D. Harris, S. R. Rumel, H. R. Warner, "Left Ventricular Videometry", *Computers and Biomedical Research*, Vol. 7, p369-379, 1974.
- [3] W. Barrett, H. Eisenberg, B. Arnold, P. Scheibe, "High-Speed Processing of Digital Intravenous Angiocardiographic Images for Enhancement of Coronary Bypass Grafts and Quantitation of Left Ventricular Function", *Computers in Cardiology*, p101-104, 1982.
- [4] D. L. Pope, D. L. Parker, P. D. Clayton, D. E. Gustafson, "Left Ventricular Border Recognition Using a Dynamic Search Algorithm", *Radiology*, Vol. 155, p513-518, 1985.
- [5] P. V. D. Zwet, G. Koning, J. Reiber, "Left Ventricular Contour Detection: A Fully Automated Approach", *Computers in Cardiology*, p359-362, 1992.
- [6] C. K. Lee, "Automated Boundary Tracing Using Temporal Information", *Ph.D Thesis*, Department of Electrical Engineering, University of Washington, Seattle, 1994.
- [7] R.M. Haralick, Linda G. Shapiro, "Computer and Robot Vision", Vol. 1 and 2, Addison-Wesley Pub. Co., 1992.
- [8] J. S. Suri, "Contour Generation of Motion Objects in Fuzzy and Low Contrast Multiple Image Frames: An Application in Cardiological Imaging", *Ph.D Thesis*, Department of Electrical Engineering, University of Washington, Seattle, 1997.
- [9] P. F. Cohn, J. A. Levine, G. A. Bergeron, R. Gorlin, "Reproducibility of the angiographic left ventricular ejection fraction in patients with coronary artery disease", *Am Heart J* 88:713-720 1974.
- [10] B. R. Chaitman, H. DeMots, J. D. Bristow, J. Rosch, S. H. Rahimtoola, "Objective and subjective analysis of left ventricular angiograms", *Circulation* 52:420-425 1975.
- [11] S. Hochreiter, J. Schmidhuber, "Flat Minima", *Neural Computation*, 9(1), 1-42, 1997.
- [12] H. T. Dodge, F. H. Sheehan, D. K. Stewart, "Estimation of Ventricular Volume, Fractional Ejected Volumes, Stroke Volume, and Quantitation of Regurgitant Flow", *Angiocardiology Current Status and Future Development*, p99-108, Springer-Verlag, 1986.