Anatomical Structure Labeling in Apical Four-Chamber View Echocardiogram Images

Yu Cao, Colin B Compas, Hongzhi Wang, Tanveer F Syeda-Mahmood

IBM Research - Almaden, San Jose CA, USA

Abstract

Anatomical structure labeling in echocardiogram images will assist cardiac disease diagnosis by providing a framework for doing geometrical statistics. General labeling algorithms often focus on stationary body structures and do not perform well in echocardiography due to cardiac motion, low signal to noise ratio, and structural deformation caused by diseases. In this paper, we propose a new method for anatomical structure labeling in echocardiography that adopts the structure layout consistency, and works on mid-level primitives (segments). Specifically, the proposed method defines a constellation model, and based on which, a model score function is designed to measure the consistency between a testing candidate configuration and the constellation model. The parameters of the score function are learned through a discriminative training framework. Given a test image and its corresponding multi-level segmentation, we use an MCMC-based algorithm to infer the configuration which best fits the constellation model. We evaluate the proposed method on 50 *images. The qualitative and quantitative results demonstrate the effectiveness of the proposed method.*

1. Introduction

Echocardiography is a popular screening tool for diagnosing cardiac diseases, including septal hypertrophy, valvular insufficiency, aneurysms, etc. Echocardiogram images can be collected from multiple viewpoints. In clinical practice, a popular viewpoint used for diagnosis is the apical four-chamber view, as shown in Fig. 1. The

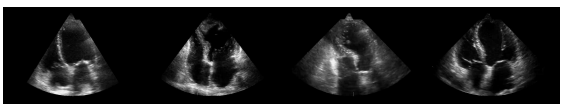

Figure 1. Apical four-chamber echocardiogram images from different subjects.

anatomical structures including heart muscle, chambers, and valves can be clearly identified in the images. Automatic labeling of these structures can help to do statistics across a large number of medical images, which further serves pathological analysis and disease diagnosis.

In medical image processing, there are many region labeling algorithms that have been developed. Generally, these methods can be categorized into two groups [1]: 1) region-based methods; and 2) pixel classification methods. Among the region-based methods, the active shape model [2] and the active appearance model [3] are two popular methods. Both methods rely on a term that measures the consistency of the testing shape with the statistical object model. These methods usually work well when the deformation across subjects is not dramatic. However, this is not true in echocardiography due to cardiac motion and underlying dramatic geometrical deformation caused by cardiac diseases. In the pixel classification category, atlas-based labeling methods are frequently used [4–6]. The atlas-based methods first perform pairwise image registrations between training and target samples, followed by a label propagation from the training samples to the target samples using label fusion, major voting, etc. The atlas-based methods provide good results for many medical imaging modalities, including structural MRI and CT. However, due to low contrast and low signal to noise ratio, the registration is not able to provide contiguous labeling on echocardiogram images. In addition, many previous works focus on segmentation and boundary extraction on a single cardiac structure [7–11], e.g. left ventricle and endocardium. When multiple similar structures are presented in the image at the same time (left atrium and right atrium), it is not clear if these methods are able to correctly capture the different structures.

In this paper, we propose a new anatomical structure labeling method for echocardiogram images, as illustrated in Fig. 2. The proposed method makes usage of the spatial layout consistency between cardiac structures. Specifically, by working on mid-level primitives (segments), the proposed method first defines a constellation model (to characterize the common spatial layout between structures) and a model score function which measures the consistency between the constellation model and a given

testing candidate configuration. Moreover, the parameters of the score function are learned through a discriminative training framework by minimizing the total classification hinge loss. Given a testing image and its corresponding multi-level segmentation, we use an Markov chain Monte Carlo (MCMC)-based algorithm to infer the configuration that best fits the constellation model. In summary, the proposed method makes the following contributions: 1) it utilizes segments rather than pixels as primitives to provide a more compact region labeling results; and 2) the constellation model enforces spatial layout for each anatomical structure and further eliminates the labeling error.

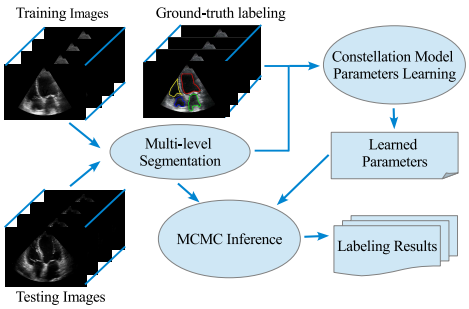

Figure 2. The framework of the proposed method.

The remainder of this paper is organized as following. Section 2 defines the constellation model and formulates the problem. The parameter learning and inference are described in Section 3. Experiment results are provided in Section 4. Section 5 concludes the paper.

2. Constellation Model and Problem Formulation

For apical four-chamber view echocardiogram images, chamber structures, i.e. "Left Ventricle" (LV), "Left Atrium" (LA), "Right Atrium" (RA), and "Right Ventricle" (RV), provide important information for diagnosis of many cardiac diseases. The typical shape and spatial layouts of these four structures are illustrated in Fig. 3(a). A constellation model [12] can be constructed by organizing these structures as shown in Fig. 3(b). The segments, also called "parts", in the constellation model have corresponding anatomical labels. The edges in the graph impose the relative positions and pairwise displacement constraints for the parts. For the constellation model, we can use area, average intensity, intensity variation, location of each part, and pairwise displacement as features. Specifically, when given annotated training images, we can extract the mean statistics of the above features across all the training samples and save them as the template T, which has the following form

$$\begin{aligned} T =& (a_1^T, \ldots, a_P^T, m_1^T, \ldots, m_P^T, v_1^T, \ldots, v_P^T, \\ & x_1^T, \ldots, x_P^T, y_1^T, \ldots, y_P^T, \\ & \triangle x_1^T, \ldots, \triangle x_E^T, \triangle y_1^T, \ldots, \triangle y_E^T), \end{aligned} \tag{1}$$

where P is the number of structures, and E is the number of pairwise edges. a_p^T, m_p^T, v_p^T, x_p^T and y_p^T ($p = 1, \cdots, P$), are mean statistics of area, average intensity, intensity variation, structure locations, respectively, across the training samples. $\triangle x_e^T$ and $\triangle y_e^T$ ($e = 1, \cdots, E$), are mean statistics of pairwise displacement along the constellation model edges.

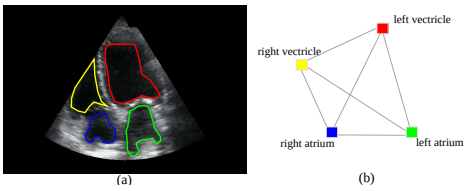

Figure 3. (a) A typical shape and spatial layout of chambers in a four-chamber view. (b) A constellation model consists of 4 chambers.

Given an image I and its multi-level segmentation $S = \{s_n\}_{n=1}^N$, we can define a configuration $Z = \{z_p\}_{p=1}^P$ ($P < N$) as an ordered subset of S, where segment z_p is associated with anatomical label p. Then We can measure how well a configuration Z fits the template T, using the following score function:

$$\begin{aligned} & f(I, S, Z, T) = \\ & -\sum_{p=1}^{P} (d_p * |a_p^Z - a_p^T| + d'_p * (m_p^Z - m_p^T)^2 \\ & \qquad + d''_p * |v_p^Z - v_p^T|) \\ & -\sum_{p=1}^{P} (f_p * (x_p^Z - x_p^T)^2 + f'_p * (y_p^Z - y_p^T)^2) \\ & -\sum_{e=1}^{E} (g_e * (\triangle x_e^Z - \triangle x_e^T)^2 + g'_e * (\triangle y_e^Z - \triangle y_e^T)^2), \end{aligned} \tag{2}$$

where $d_p, d'_p, d''_p, f_p, f'_p, g_e$ and g'_e are linear combination coefficients.

By combining the multiplication coefficients and difference information into two vectors $\vec{\beta}$ and $\Phi(I, S, Z, T)$, respectively, we can rewrite the Eq.(2) as

$$f(I, S, Z, T) = -\vec{\beta} \cdot \Phi(I, S, Z, T). \tag{3}$$

Then the echocardiogram region labeling problem can be formulated as the following optimization problem

$$Z^* = \arg\max_Z \; -\vec{\beta} \cdot \Phi(I, S, Z, T), \tag{4}$$

subject to $\vec{\beta} \geq 0$.

3. Learning and Inference

3.1. Learning Parameters

Inspired by the Deformable Part Model (DPM) [13] framework, we learn the score function parameters $\vec{\beta}$ using a discriminative training framework. Let $D = \{< I_j, S_j, Z_j, Y_j >\}_{j=1}^{J}$ be J labeled configurations with labels $Y_j \in \{-1, +1\}$, and T is the consequent template. Then following the maximum-margin/minimum-loss principle, we can formulate the parameter training problem as

$$\begin{aligned}\vec{\beta}^*(D) =& \arg\min_{\vec{\beta}} \lambda||\vec{\beta}||^2 + \\ & \sum_{j=1}^{J} max(0, 1 - Y_j f(I_j, S_j, Z_j, T)).\end{aligned} \tag{5}$$

In Eq. 5, the second term $max(0, 1 - Y_j f(I_j, S_j, Z_j, T))$ is the standard hinge loss, which is convex but non-differentiable. However, it has a subgradient with respect to $\vec{\beta}$, in form of

$$\nabla f_{\vec{\beta}}(I_j, S_j, Z_j, T) = \begin{cases} -\Phi(I_j, S_j, Z_j, T) \cdot Y_j & \text{if } f(I_j, S_j, Z_j, T) \cdot Y_j < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Therefore, to minimize the total loss, we use a gradient descent search algorithm to find the optimal parameter $\vec{\beta}$ using

$$\vec{\beta}_{t+1} = \vec{\beta}_t - \eta_t \nabla f_{\vec{\beta}}(I_j, S_j, Z_j, T), \tag{7}$$

where η_t is the learning ratio.

3.2. Inference

We define a binary matrix M of size $N \times (P + 1)$ to represent the selection of segments from S. Each row of M corresponds to a segment in S, and column 1 to column P of M correspond to P anatomical structures (defined in Fig. 3). The last column is called dummy column. Considering $M_{(n,p)}, n = 1, \ldots, N, p = 1, \ldots, P$, if $M_{(n,p)} = 1$, then the segment s_n is selected to be anatomical structure p in the configuration. For the segments that are not selected in the configuration, we set the dummy column of the corresponding row to be "1", namely if $M_{(n,\{1,\ldots,P\})} = 0$, then we would have $M_{(n,P+1)} = 1$. Consequently, the posterior of a configuration can be defined as

$$\pi(M \mid I, S, T) \propto sigmoid(-\vec{\beta} \cdot \Phi(I, S, Z(M), T)/\sigma), \tag{8}$$

where $sigmoid(x) = \frac{1}{1+e^{-x}}$, and σ is a normalizing parameter. Inferring the best configuration can be formulated as

$$M^* = \arg\max_{M} \pi(M \mid I, S, T). \tag{9}$$

We solve the inference problem Eq.(9) using the MCMC method. Specifically, we adopt the Metropolis-Hastings sampling [14], which is summarized in Algorithm-1.

Algorithm 1 Metropolis-Hastings Sampling

1: At each iteration t:
2: Sample $M' \sim q(M'|M^{(t)})$.
3: Accept it with probability
$\alpha(M^{(t)}, M') = \min\{1, \frac{\pi(M'|I,S,T)q(M^{(t)}|M')}{\pi(M^{(t)}|I,S,T)q(M'|M^{(t)})}\}$
4: **if** accepted **then**
5: $M^{(t+1)} = M'$
6: **else**
7: $M^{(t+1)} = M^{(t)}$
8: **end if**

In the algorithm, $M^{(t)}$ is the solution matrix at iteration t. $\pi(M'|I, S, T)$ is the posterior density for M' given I , S and T. $q(M^{(t)}|M')$ is the proposal density for moving from M' to $M^{(t)}$. Similarly, $q(M'|M^{(t)})$ is the proposal density from moving $M^{(t)}$ to M'. When proposing M' from $M^{(t)}$, we choose an anatomical structure with probability of $1/P$. Then we change the segment associated with this anatomical structure from the original one to a new one with probability $1/(1 + dist_{new})$, where the $dist_{new}$ measures the Euclidean distance between the new proposed segment and the corresponding location in the template. Similarly, the probability to move back is $1/(1 + dist_{old})$. This gives the transition ratio

$$\frac{q(M^{(t)}|M')}{q(M'|M^{(t)})} = \frac{\frac{1}{P} \cdot \frac{1}{1+dist_{old}}}{\frac{1}{P} \cdot \frac{1}{1+dist_{new}}} = \frac{1 + dist_{new}}{1 + dist_{old}}. \tag{10}$$

4. Experiments

The test dataset consists of 50 images. These images are collected from 50 subjects, one image per subject, with 14 having normal cardiac anatomy and 36 being abnormal. The images present structures from minor to severely deformed due to cardiac diseases, including amyloidosis, ventricular/atrial hypertrophy, and rheumatic heart disease. For the quantitative evaluation, we construct the ground-truth by having an expert label anatomical structures from each of the images using the LabelMe [15] annotation tool. The proposed method works on mid-level primitives (segments) for both parameter learning and inference. To get these mid-level primitives, we use multi-level segmentation [16] to provide candidate anatomical segments from coarse to fine levels, as shown in Fig. 4. We mea-

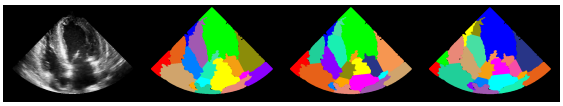

Figure 4. An illustration for multi-level segmentation from coarse to fine.

sure the overlap with the ground-truth annotation using the Dice coefficient. Since the proposed method is based on a discriminative training framework, we randomly split the whole dataset, using 30 (60%) images as training and 20 (40%) as testing. Then we run 5 times cross validation and report the averaged performance. Some qualitative results are shown in Fig. 5. From the figure, we can observe that the proposed method can provide compact labeling results. Quantitative results for labeling accuracy are shown

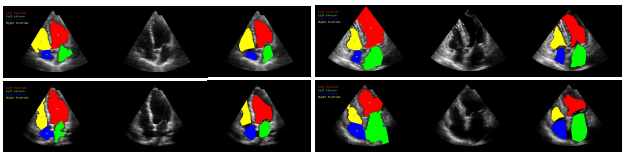

Figure 5. Qualitative results: Left Ventricle (Red), Left Atrium (Green), Right Atrium (Blue) and Right Ventricle (Yellow). In each subfigure, from left to right: result, original image and ground-truth.

in Table 1.

	Left Ventricle	Left Atrium	Right Atrium	Right Ventricle
Run #1	0.762 ± 0.102	0.677 ± 0.257	0.736 ± 0.268	0.638 ± 0.231
Run #2	0.639 ± 0.257	0.66 ± 0.246	0.613 ± 0.357	0.568 ± 0.287
Run #3	0.706 ± 0.228	0.687 ± 0.255	0.656 ± 0.294	0.528 ± 0.349
Run #4	0.736 ± 0.184	0.677 ± 0.288	0.691 ± 0.248	0.47 ± 0.286
Run #5	0.668 ± 0.238	0.662 ± 0.281	0.694 ± 0.287	0.506 ± 0.327
Overall	0.702	0.673	0.678	0.542

Table 1. Dice coefficient $\pm$ standard deviation for each chamber structure for five runs.

5. Conclusion

We have investigated the problem of labeling chamber structures in apical four-chamber view echocardiogram images. The proposed method adopts the spatial layout consistency between structures. By working on mid-level primitives (segments), the method defines a constellation model and then learns the model parameters through a discriminative training framework. The labeling inference is done using an MCMC-based algorithm. The qualitative and quantitative results demonstrate the effectiveness of the proposed method.

References

[1] Withey D, Koles Z. Medical image segmentation: Methods and software. In Proceedings of NFSI and ICFBI. 2007; 140–143.

[2] Cootes T, Taylor C, Copper D, Graham J. Active shape models – their training and application. CVIU 1995;38–59.

[3] Cootes T, Edwards G, Taylor C. Active appearance models. PAMI 2001;681–685.

[4] Wang H, Suh J, Das S, Pluta J, Craige C, Yushkevich P. Multi-atlas segmentation with joint label fusion. PAMI 2013;611–623.

[5] Heckemann R, Hajnal J, Aljabar P, Rueckert D, Hammers A. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. NeuroImage 2006;115–126.

[6] Warfield S, Zou K, Wells W. Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. IEEE Trans Machine Intelligence 2004;903–921.

[7] Carneiro G, Nascimento J, , Freitas A. A threshold selection method from gray–level histograms. TIP 2012;968–982.

[8] Carneiro G, Nascimento J, Freitas A. Robust left ventricle segmentation from ultrasound data using deep neural networks and efficient search methods. In ISFNM. 2010; 1085–1088.

[9] Paragios N, Jolly MP, Taron M, Ramaraj R. Active shape models and segmentation of the left ventricle in echocardiography. Scale Space and PDE Methods in Computer Vision 2005;131–142.

[10] Syeda-Mahmood T, Wang Q, McNeillie P, Beymer D, Compas C. Discriminating normal and abnormal left ventricular shapes in four-chamber view 2D echocardiography. In ISBI. 2014; 401–404.

[11] Guo Y, Wang Y. Automatic endocardium extraction for echocardiogram. In BMEI. 2011; 155–159.

[12] Weber M, Welling M, Perona P. Unsupervised learning of models for recognition. In ECCV. 2000; 18–32.

[13] Felzenszwalb P, McAllester D, Ramanan D. A discriminatively trained, multiscale, deformable part model. In CVPR. 2008; 1–8.

[14] Gilks W, Richardson S, Spiegelhalter D. Markov Chain Monte Carlo in Practice. Chapman and Hall/CRC, 1995.

[15] Russell B, Torralba A, Murphy K, Freeman W. LabelMe: A database and web–based tool for image annotation. International Journal of Computer Vision 2008;157–173.

[16] Cao Y, McNeillie P, Syeda-Mahmood T. Segmentation of anatomical structures in four-chamber view echocardiogram images. In ICPR. 2014; .

Address for correspondence:

Yu Cao
650 Harry Rd, San Jose, CA 95120, USA
caoy@us.ibm.com