International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:10, No:1, 2016

A Robust and Efficient Segmentation Method Applied for Cardiac Left Ventricle with Abnormal Shapes

Peifei Zhu, Zisheng Li, Yasuki Kakishita, Mayumi Suzuki, Tomoaki Chono

***Abstract*—**Segmentation of left ventricle (LV) from cardiac ultrasound images provides a quantitative functional analysis of the heart to diagnose disease. Active Shape Model (ASM) is widely used for LV segmentation, but it suffers from the drawback that initialization of the shape model is not sufficiently close to the target, especially when dealing with abnormal shapes in disease. In this work, a two-step framework is improved to achieve a fast and efficient LV segmentation. First, a robust and efficient detection based on Hough forest localizes cardiac feature points. Such feature points are used to predict the initial fitting of the LV shape model. Second, ASM is applied to further fit the LV shape model to the cardiac ultrasound image. With the robust initialization, ASM is able to achieve more accurate segmentation. The performance of the proposed method is evaluated on a dataset of 810 cardiac ultrasound images that are mostly abnormal shapes. This proposed method is compared with several combinations of ASM and existing initialization methods. Our experiment results demonstrate that accuracy of the proposed method for feature point detection for initialization was 40% higher than the existing methods. Moreover, the proposed method significantly reduces the number of necessary ASM fitting loops and thus speeds up the whole segmentation process. Therefore, the proposed method is able to achieve more accurate and efficient segmentation results and is applicable to unusual shapes of heart with cardiac diseases, such as left atrial enlargement.

***Keywords*—**Hough forest, active shape model, segmentation, cardiac left ventricle.

I. Introduction

ECHOCARDIOGRAPHY has been widely used for a number of various cardiac examinations, such as structure and function analysis. Typically, the segmentation of left ventricle (LV) from cardiac ultrasound images provides a quantitative functional analysis of the heart to diagnose diseases. However, the manual LV segmentation can be time consuming, and might only be performed by a clinical specialist. In addition, such segmentation has poor repeatability since examinations are done visually and subjectively by different clinicians. Automating the LV segmentation can solve these issues. It has the potential to speed up cardiac examinations and reduce inter-user variation in the LV segmentation procedure.

Prior art in automatic LV segmentation consists of active contours or snakes [1], deformable models [2], and supervised learning methods [3]. Deformable model based methods, such as Active shape models (ASM) [4], build statistical models of the object shape and iteratively deform the shape to optimally fit a new image. A drawback of shape model base segmentation is that the shape model is not initialized sufficiently close to the target. On the other hand, supervised learning methods use a database of annotated LV images as a training set for learning. These methods have competitive accuracy but suffer from computationally expensive and complex boundary search.

In this work, a two-step framework that combines pattern recognition and a shape model is proposed for robust and efficient cardiac LV segmentation. To solve the initialization problem of ASM, a detector based on Hough forest [5] is first applied to localize cardiac feature points. Such points are able to predict a relatively close fitting of the LV shape model. Two contributions are made for improving accuracy of point detection: 1) a coarse-to-fine strategy is applied with Hough forest detector and 2) artificially generated training data with rotations and scales of the original image is used in the detector training. The performance of the proposed method is evaluated on a dataset of 810 cardiac ultrasound images. This dataset includes many cardiac disease images, such as left atrial enlargement and left ventricular hypertrophy. The experiment results demonstrate that with a robust initialization of the LV shape, the proposed method can achieve more accurate segmentation results and is applicable to unusual shapes of heart with cardiac diseases.

II. Proposed Method

A. A Two- Step Framework for LV Segmentation

Framework of automatic LV segmentation is shown in Fig. 1. For an input cardiac image, three feature points selected as cardiac apex, left mitral annulus and right mitral annulus (marked by red crosses) are detected at the first step. The detection is mainly realized by Hough forest algorithm, which will be introduced in Section II. *B*.

Second, detected feature points are used as reference positions to generate an initialization shape. ASM is then used to adjust each point on the shape model with the input image. The algorithm of ASM will be described in Section II. *C*. A final LV contour is generated after ASM optimization. Since the initialization shape is sufficiently close to the target by using feature point initialization, the ASM optimization process is relatively fast, and segmentation accuracy is largely improved.

B. Hough Forest Based Feature Point Detection

Hough forest provides a way to map from image patches to anatomical locations. It is a combination of random forest and Hough transform. Hough forest is a set of decision trees learned

Peifei Zhu, Zisheng li, Yasuki Kakishita and Mayumi Suzuki are with the Media Systems Research Department, Hitachi, Ltd., Reaserch & Development Group, Japan (phone: +81-42-323-1111, e-mail: peifei.zhu.ww@hitachi.com).
Tomoaki Chono is with Hitachi Aloka Medical, Ltd.

from the training data. Training and testing process are described below.

Training: Each tree T of Hough forest is constructed based on a set of patches $\{P_i = (I_i, c_i, \mathbf{d}_i)\}$, where I_i is the appearance of the patch, c_i is the class label, and $\mathbf{d}_i$ is the offset from patch center to the object center. For each leaf node L, C_L which is the proportion between object patches and background patches and the list $\mathbf{D}_L = \{\mathbf{d}_i\}$ of the offset vectors are stored. Training data and leaf information of Hough forest is shown in Fig. 2. A cardiac ultrasound image is manually divided into background sets (negative samples) and object sets (positive samples), as shown in (a). The structure of Hough forest, where the input is a subset of patches from images, is shown in (b). Each tree starts from the root and during regression, each node receives a set of patches for binary test and splitting. The information recorded in leaves is shown in (c). It consists of the proportion C_L and the list of the offset $\mathbf{D}_L$. Blue crosses are the object center and red crosses are the offset vectors. For example, $C_L = 0.5$ means half the patches are background and half are object in this leaf. Leaves of the tree thus form a discriminative codebook.

Based on input sets of patches, the Hough Tree is constructed recursively, starting from the root. Random tests is selected based on how well they separate the input set of patches. A key point of Hough forest is how the binary test is evaluated. To conduct an optimal test, the uncertainties in both the class labels and the offset vectors should decrease towards the leaves. A set of patches is defined as $A = \{P_i = (\mathbf{I}_i, c_i, \mathbf{d}_i)\}$, and class label uncertainty $U_1(A)$ and offset uncertainty $U_2(A)$ are defined as:

$$U_1(A) = -|A| \cdot \sum p(c|A) \ \ln(p(c|A)), \tag{1}$$

$$U_2(A) = \sum_i (\mathbf{d}_i - \mathbf{d}_A)^2, \ \text{ when } c_i = 1, \tag{2}$$

where $|A|$ is the number of patches, $p(c|A)$ is the proportion of patches with label c in set A , and $\mathbf{d}_i$ is the mean offset vector over all object patches. The node construction steps are as follows. Given a training set of patches, a pool of pixel tests $\{t^k\}$ is generated by uniformly choosing one feature channel and two pixel locations inside a patch. The randomized decision is made as to whether the node should minimize the class-label uncertainty or the offset uncertainty. The process can be represented as:

$$\operatorname{argmin}\left(U_*(\{P_i|t^k(\mathbf{I}_i) = 0\}) + U_*(\{P_i|t^k(\mathbf{I}_i) = 1\})\right), \tag{3}$$

where * is either class label uncertainty or offset uncertainty.

Testing: Testing can be divided into regression and voting steps. The regression process is as follows. 1) For each pixel location $\mathbf{p}$, a patch is extracted and starts regression from the root; 2) when passing each node, this patch is sorted into the left or right child node in accordance with the binary test. All pixels in the image go through the forest simultaneously in the first and second steps until they reach the leaves. During the voting process, the information stored in leaves is used to cast the probabilistic Hough votes to the location of the object center. Leaf information consists of proportion C_L and offset vectors $\mathbf{D}_L$, so that $C_L/\mathbf{D}_L$ is defined as a weight value for a vote. Each pixel in leaves carries a location $\mathbf{p}$, and it votes to all locations $\{\mathbf{p} - \mathbf{d} | \mathbf{d} \in \mathbf{D}_L\}$ with a weight value $C_L/\mathbf{D}_L$. After all votes from each pixel have been summed up, Hough image $V(x)$ can be obtained by applying a two dimensional Gaussian-filter.

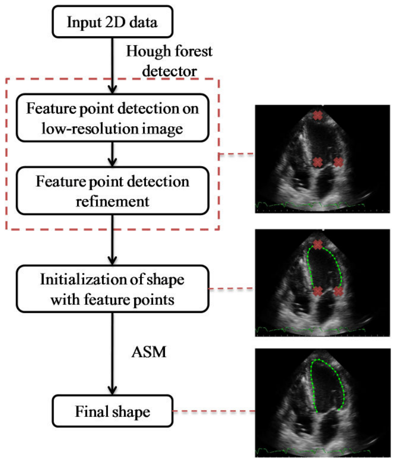

Fig. 1 Framework of automatic cardiac LV segmentation

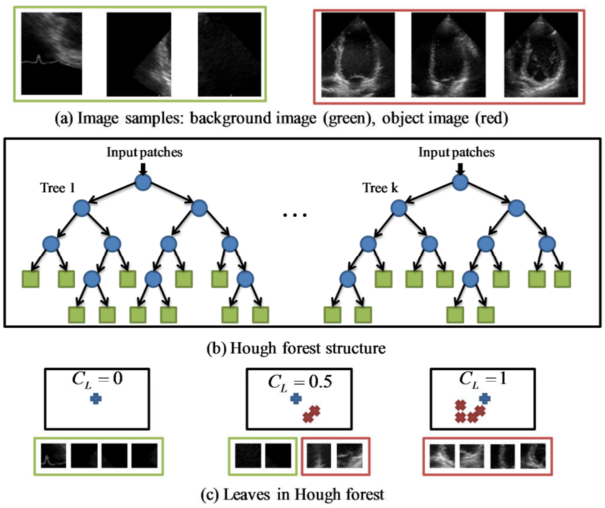

Fig. 2 Training data and Forest structure

C. Active Shape Model Based Segmentation

A shape model is formed by a set of landmarks that represent a distinguishable point present in the image. ASM [4] is a widely used shape model. It is constructed by two sub-models: the profile model and the shape model.

The profile models are used to locate the approximate position of each landmark by template matching. The classical ASM forms a fixed-length normalized gradient vector (called the profile) by sampling the image along a line orthogonal to the shape boundary at the landmark. During training on a manually landmarked model, at each landmark the mean profile vector is $\bar{\mathbf{g}}$ and the profile covariance matrix is $\mathbf{S_g}$. During searching, the landmark is updated to a new location where profile $\mathbf{g}$ has the shortest Mahalanobis distance from the mean profile along the orthogonal line.

$$\text{Mahalanobis Distance} = (\mathbf{g}-\bar{\mathbf{g}})^{\mathrm{T}}\mathbf{S}_{\mathbf{g}}^{-1}(\mathbf{g}-\bar{\mathbf{g}}). \quad (4)$$

The shape model specifies allowable constellations of landmarks. It generates a shape $\hat{\mathbf{x}}$ with

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \boldsymbol{\varphi}\mathbf{b}, \quad (5)$$

where, $\bar{\mathbf{x}}$ is the mean shape, $\mathbf{b}$ is a parameter vector, and $\boldsymbol{\varphi}$ is a matrix of selected eigenvectors of the covariance matrix $\mathbf{S}_{\mathbf{s}}$ of the points of the aligned training shapes. Using a standard principal components approach, we model as much variation in the training set as we want by ordering the eigenvalues of $\mathbf{S}_{\mathbf{s}}$ and keeping an appropriate number of the corresponding eigenvectors in $\boldsymbol{\varphi}$.

D. Contributions to Improve Accuracy of Point Detection

To improve the accuracy of feature point detection, this work makes two main contributions: 1) a coarse-to-fine strategy is applied with Hough forest detector; 2) artificially generated training data with rotations and scales of original image is used in the detector training.

A coarse-to-fine strategy: Feature points are detected serially through a multi-scale hierarchical search. The whole image is first used to provide an estimate of the region of interest that is then refined by only using local information. The process is shown in Fig. 3. The whole process consists of two parts: coarse level and fine level. Hough forest coarse detector and a fine detector need to be trained before testing. Coarse detector is trained using low-resolution images that are down-sampled from original images. Positive patches are chosen from a bounding box region around truth-ground, and negative patches are chosen from a whole image except the positive region. The fine detector is trained on a high-resolution image (original image). Since the candidate region has been largely narrowed, regions for sampling positive patches and negative patches are correspondingly set smaller. During the testing step, first, the input image is down-sampled, and a coarse position is localized using Hough forest coarse detector prepared in the training. In coarse level detection, every pixel in a low-resolution image provides a vote (Hough voting) to a potential target location. Second, in the refinement step, a high-resolution image is used to provide a more accurate detection. Only pixels in the neighborhood of the coarse position (green cross and green bounding box is the voting region) are used to predict the existence of the object. By applying a coarse-to-fine strategy, the searching region has been largely cut down so running time is successfully shortened. Moreover, refinement searching that only uses the region closest to the target may provide a more accurate result than using the information of the whole image.

Data augmentation: The effectiveness of the supervised learning method depends on the size and the richness of the training data set. Cardiac images of diseases, such as left atrial enlargement, usually have extremely small LV. To deal with such cases, as many abnormal shapes as possible should also be included in training data. However, the number of abnormal image is relatively limited. As a solution, manually generated training data with the rotation and scale of the original image is used in the training step. The angle and size range of LV in the dataset is analyzed first. Cardiac LV tends to be inside the angle range of -25~25 degrees and scale range of 0.8~1.2 times. Based on this analysis, 10 manual patterns are generated with a randomly chosen angle between -25~25 degrees and a randomly chosen scale between 0.8~1.2 times for each original image. All generated patterns and original images are used in training detectors.

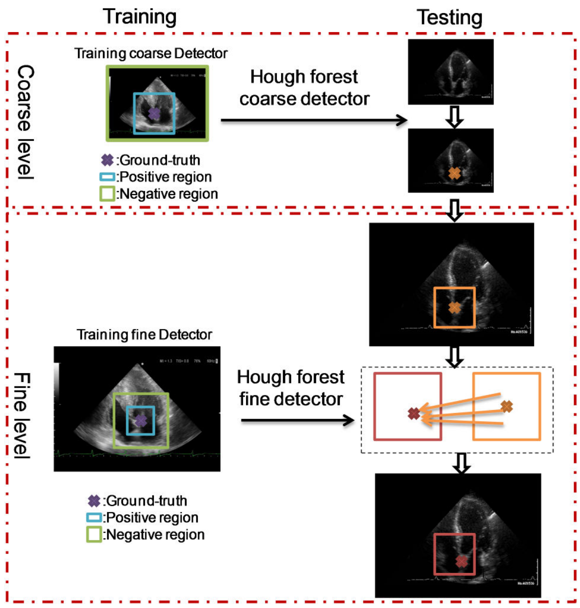

Fig. 3 A coarse-to-fine strategy for feature point detection

III. Experiments

This section describes experiments of feature point detection and cardiac LV segmentation. For LV segmentation, the endocardial border of the LV at the end-diastole phases is the target. As shape model based segmentation is used, 30 points on the border are uniformly set for constructing a shape model. The dataset used in experiments contains 810 samples of 2D cardiac ultrasound images. This dataset includes many images of cardiac disease, such as left atrial enlargement and left ventricular hypertrophy. All images are apical four chamber view at the end-diastole. For experiments, 296 images are for training and the remaining 514 images are used as unseen data for testing. The study was approved by the ethics committee of Hitachi group headquarters.

Three kinds of methods have been used to compare LV segmentation accuracy. All methods use ASM algorithm but with different initializations. Feature points detected in this work are: apex, left mitral annulus and right mitral annulus. They are either detected by an AdaBoost based [6] algorithm, Hough Forest algorithm (proposed method), or manually annotated by users. Detection results of feature points are first measured and shown in Table I. The proposed method has successfully reduced distance error by about 40.5%. As for the number of wrong detections, if mean distance of a sample is larger than 20 mm, it is counted as a wrong detection. AdaBoost based method yields a large number of wrong detections, which means it is not robust to abnormal shapes. Images with diseased hearts, especially left atrial enlargement, are usually unsuccessfully detected. On the other hand, unlike AdaBoost using a linear combination of weak learners, Hough Forest is constructed by a multitude of decision trees and is based on local feature mapping, which is more robust to general changes of target shape. Detection results show proposed method reduces wrong detection by 70% and is able to deal with cardiac disease images.

Detected feature points are used to generate an initialization shape for ASM based segmentation. To measure the accuracy of the segmentation, three metrics are applied: mean distance, precision and recall. Mean distance is mean error of point distance between ground-truth and each feature point from the shape model. Precision and recall are particularly attractive as measures of segmentation quality because they are not biased in favor of over- or under-segmented images. Precision measures the percentage of intersection area in the automatic segmentation that corresponds to a ground-truth area and is sensitive to over-segmentation. Recall measures the percentage of intersection area that corresponds to whole automatic segmented area and is sensitive to under-segmentation.

Statistical and visual results of all methods are shown in Table II and Fig. 4. All methods use ASM based segmentation with different initializations. Three feature points are detected by the AdaBoost algorithm, Hough Forest algorithm, and manually. Results show that the proposed method is able to reduce mean distance by about 40% as well as improve precision and recall by about 3.3% and 8.3%, compared with the AdaBoost based method. The AdaBoost based method is more easily to yields under-segmented results. In fact, the cardiac apex is relatively close to the image border so that the outer side of the boundary might suffer from the lack of feature information. AdaBoost is not robust enough to cope with such incompletion, so the boundary around the apex tends to be close to inner side. On the other hand, Hough forest constructed by a multitude of trees tends to be more robust. In addition, Hough forest uses patch based features for classification so that local information, such as features around the apex, might just only slightly influence the whole detection.

Segmentation results show this trend in Fig. 4. Segmentation results of AdaBoost tend to be smaller than those of the ground-truth area, while the proposed method shows the modification. Results also demonstrate that the proposed method yields segmentation results similar to those of manual feature point's segmentation. Moreover, experiments also show the detection time of feature points by the proposed method is about 100 ms per image. The total time of the segmentation is about 200 ms. All experiments were run on an Intel core i7 3.6GHz computer with 16GB RAM. Future work using parallel processing of Hough forest may be able to realize a near real-time detection. With a robust initialization, necessary ASM fitting loops are reduced about 20% compared to AdaBoost based method. Therefore, the whole segmentation process becomes more efficient.

TABLE I
DETECTION RESULTS OF FEATURE POINTS

	Mean error (mm)	Wrong detection number
AdaBoost	8.08	78
Proposed method	4.81	23

TABLE II
SEGMENTATION RESULTS WITH DIFFERENT INITIALIZATION METHODS

	Mean error (mm)	Precision (%)	Recall (%)
Manual points+ASM	2.65	92.8	93.5
AdaBoost+ASM	7.60	85.1	81.1
Proposed method	4.55	88.4	89.4

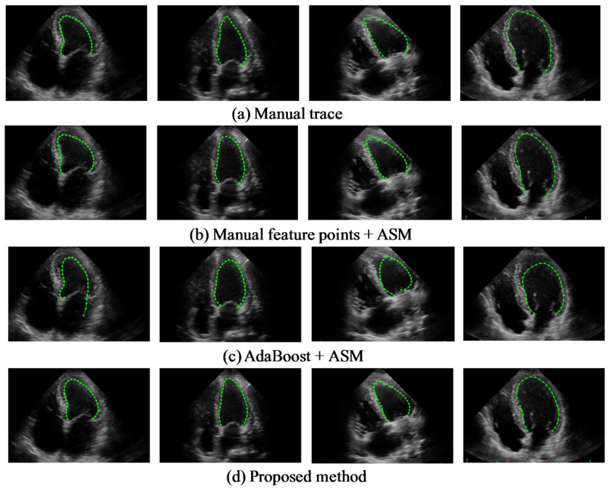

Fig. 4 Examples of segmentation result for abnormal shapes

IV. CONCLUSIONS

A robust and efficient segmentation method is proposed for cardiac LV segmentation. Feature points are detected first to provide an initialization fitting of the shape model. ASM is then applied to further deform the shape to optimally fit input images. Experiments implemented on 2D cardiac show our method is applicable to diseased hearts, and achieves faster and more accurate segmentation results than the existing method.

ACKNOWLEDGMENT

The authors thank Hitachi Aloka Medical, Ltd. for providing cardiac images and ground-truth annotation data for the evaluation experiments of this work.

REFERENCES

[1] C. Corsi and G. Saracino, "Left ventricular volume estimation for real-time three-dimensional echocardiography," *IEEE Trans. Med. Imag.*, vol. 21, no. 9, pp. 1202–1208, Sep. 2002.
[2] T. McInerney, "Deformable models in medical image analysis: a survey," *Med. Image Anal.*, vol. 1, no. 2, pp. 91–108, 1996.
[3] B. Georgescu, "Databased-guided segmentation of anatomical structures with complex appearance," *Proc. Conf. CVPR,* pp. 429–436, 2005.
[4] T. Cootes, "Active shape models—Their training and application," *Comput. Vis. Image Understand.*, vol. 61, no. 1, pp. 38–59, Jan. 1995.

[5] J. Gall, "Hough forests for object detection, tracking, and action recognition," *IEEE Trans. on PAMI.*, vol. 33, no. 11, pp. 2188–2202, 2011.
[6] P. Viola and M. Jones, “Robust real-time face detection,” *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.