Whole Heart Region Segmentation from CT and MRI Images Using a Hybrid Model

Feng Chen, Xubo Yang
School of software
Shanghai Jiao Tong University
Shanghai, P.R.China

Benjamin Longhi
Telecommunications Department
INSA Lyon
Lyon, France

Lixu Gu[1]
Med-X Research Institute
Shanghai Jiao Tong University
Shanghai, P.R.China

Jianrong Xu[1]
Radiology Department
Shanghai Ren Ji Hospital
Shanghai, P.R.China

Abstract—**Image segmentation is a very important pre-requisite in medical image analysis, therapy planning and CAD (Computer Aided Diagnose). Heart segmentation is a difficult task for its similarity in gray level with neighboring organs coursed by tissue conglutination and the complexity of anatomical configuration. Some earlier methods focused on the segmentation of parts of heart, e.g. ventricles, which is much easier even though they are time consuming and do not converge in some cases. In this paper, we present a fast and robust hybrid method for the segmentation of the whole heart, including ventricles and myocardium from CT and MRI images. This method consists of two major steps: The first step focuses on the segmentation of different configurations of heart based on classification method; the second step focuses on getting the boundary of the whole heart and separates it from neighboring organs. At the end of this paper, some experiments have been made to verify the efficiency and the validity of our method.**

Keywords-heart segmentation; EM algorithm; Marker Controlled Watershed; Hybrid model

I. INTRODUCTION

Medical image segmentation is a fundamental research topic which is quite popular in the recent years. Many different approaches have been proposed dealing with different kinds of cases. Generally speaking, all segmentation methods can be divided into two major categories. One is boundaries segmentation (pixel differences). And the other one is coherent regions segmentation (pixel similarities) [1, 2]. Some of these approaches are quite famous. For example, the level set method, which was first introduced by Osher and Sethian, is used in capturing moving fronts [3]. Active shape models (ASMs), firstly introduced by Cootes T.F, Taylor J.C etc. [4], is used for segmentation of complex shapes [5,8]. The watershed algorithm, which was first introduced by Serge Beucher and Christian Lantuéjoul, is a classic image processing segmentation algorithm that splits an image into areas, based on the topology of the image [6].

With the increase of living condition in recent years, cardiovascular disease has become the leading direct or contributing cause of non-accidental deaths around the world [7]. Consequently, it is really important to pay more attention to early diagnosis and therapy plan of cardiovascular diseases, in which heart segmentation plays a key role. Due to the heart-beat and its complex anatomical configuration, heart segmentation is a challenging research topic. Some early researches focus on part of heart, e.g. left heart ventricle [8], rather than whole heart, which are easier but not suitable in some diagnosis and therapy cases.

In our method, heart segmentation consists of two steps. One step mainly deals with segmentation of the interior anatomic configuration of heart. The basic idea is to use a classification method based on the EM (Expectation Maximization) algorithm and Gaussian Mixture Model, combined with optimizing processing. The other step focuses on extracting the exterior boundary of whole heart from the neighboring organs in medical images. This step is mainly based on an optimized version of the Marker-Controlled Watershed segmentation with some image optimization methods to improve results.

II. METHODS

A. EM Classification Algorithm

The Expectation Maximization (EM) is an algorithm that can be used to find the maximum-likelihood estimates of parameters in probabilistic models where the model depends on unobserved latent variables, when the samples can be viewed as incomplete data [9]. It is an iterative computation where each iteration consists of an expectation step followed by a maximization step. It is a semi-supervised algorithm since

---

This work is partial sponsored (supported) by the National Nature Science Foundation of China with grant number of 60872103, the National 863 High-Tech Plan with grant number of 2007AA01Z312 and the 973 Research Plan with grant number of 2007CB512700-1, 2006CB504801.

[1] Corresponding authers.

some incomplete data is needed as sample points [10]. Compared with entire-supervised and unsupervised algorithm, semi-supervised one has some notable advantages. For example, it doesn't need large amount of labeled data, which is quite a time consuming and boring job in entire-supervised algorithms. On the other hand, semi-supervised algorithms make good use a little basic information such as the number of classes and the approximate mean value of each class, which is easy to get. With the help of these elements the precision and accuracy will be improved a lot, compared with unsupervised algorithms [11].

In our method, medical image is modeled as a Gaussian mixture model based on its gray level. Each Gaussian distribution of the Gaussian mixture corresponds to a class of the segmentation results. We use EM algorithm to find the parameters of the Gaussian mixture model.

Let's recall the definition of the maximum-likelihood estimation problem. We have a density function $p(x \mid \Theta)$ that is governed by the set of parameters Θ. In Gaussian mixture model p is a set of Gaussians and Θ could be the means and covariances. We also have a data set of size N, supposedly drawn from this distribution, i.e., $\chi = \{x_1, \ldots, x_N\}$. That is, we assume these data vectors are independent and identically distributed with distribution p. Therefore, the resulting density for the samples is

$$p(\chi \mid \Theta) = \prod_{i=1}^{N} p(x_i \mid \Theta) = L(\Theta \mid \chi) \tag{1}$$

This function $L(\Theta \mid \chi)$ is called the likelihood of the parameters given the data. The likelihood is thought as a function of the parameters Θ where the data χ is fixed. In the maximum likelihood problem, our goal is to find the Θ which could maximizes L. That is, we wish to find Θ^* where

$$\Theta^* = \arg\max_{\Theta} L(\Theta \mid \chi) \tag{2}$$

Often we maximize $\log(L(\Theta \mid \chi))$ since it is analytically easier.

In our classification method, the gray level histogram of image is viewed as a Gaussian mixture model. Each Gaussian distribution represents a class of anatomical configuration of tissue. Assume that a sample of N vectors (or scalars) $y_1, \ldots, y_N$, where $y \in \Re^l$, is drawn from one of n Gaussian distributions. So the goal of classification problem is deciding which Gaussian distribution every pixel belongs to. Let $z_j \in \{1, 2, \ldots, n\}$ denote which Gaussian y_j is from. The probability that a particular y comes from the ith D-dimensional Gaussian is:

$$\begin{aligned} p(y \mid z = i, \theta) &= \mathrm{N}(\mu_i, \sigma_i) \\ &= (2\pi)^{-D/2} \left|\sigma_i\right|^{-1/2} \exp\left(-\frac{1}{2}(y - \mu_i)^{\mathrm{T}} \sigma_i^{-1} (y - \mu_i)\right) \end{aligned} \tag{3}$$

Then our task is to estimate unknown parameters $\theta = \{\mu_1, \ldots, \mu_n, \sigma_1, \ldots, \sigma_n, P(z = 1), \ldots, P(z = n)\}$. To solve this problem, we use EM algorithm to find the maximum-likelihood estimates of Gaussian mixture parameters. It is achieved by iteratively applying the following two steps:

Expectation step: Calculate the estimation of the unobserved z conditioned on the observation (which Gaussian is used), using the values from the last maximization step:

$$\begin{aligned} p(z_j = i \mid y_j, \theta_t) &= \frac{p(z_j = i, y_j \mid \theta_t)}{p(y_j \mid \theta_t)} \\ &= \frac{p(y_j \mid z_j = i, \theta_t) p(z = i \mid \theta_t)}{\sum_{k=1}^{n} p(y_j \mid z_j = k, \theta_t) p(z = k \mid \theta_t)} \end{aligned} \tag{4}$$

Maximization step: Maximize the expected log likelihood of the joint event:

$$\begin{aligned} Q(\theta) &= E_z\left[\ln \prod_{j=1}^{m} p(y_j, z \mid \theta) \mid y_j\right] \\ &= E_z\left[\sum_{j=1}^{m} \ln p(y_j, z \mid \theta) \mid y_j\right] \\ &= \sum_{j=1}^{m} \sum_{i=1}^{n} p(z_j = i \mid y_j, \theta_t) \ln(p(y_j \mid z_j = i, \theta) p(z_j = i \mid \theta)) \end{aligned} \tag{5}$$

Here we have the constraint $\sum_{i=1}^{n} p(z = i \mid \theta) = 1$. We add a Lagrange multiplier and expand the probability density function (1), so we get

$$\begin{aligned} G(\theta) = &-\frac{D}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\sigma_i\right| \\ &-\frac{1}{2}(y_j - \mu_i)^T \sigma_i^{-1} (y_j - \mu_i) + \ln p(z = i \mid \theta) \end{aligned} \tag{6}$$

$$\begin{aligned} L(\theta) = &\left(\sum_{j=1}^{m} \sum_{i=1}^{n} p(z_j = i \mid y_j, \theta_t) G(\theta)\right) \\ &-\lambda\left(\sum_{i=1}^{n} p(z = i \mid \theta) - 1\right) \end{aligned} \tag{7}$$

To find the new estimate θ_{t+1}, we look for a maximum where

$$\frac{\partial L(\theta)}{\partial \theta} = 0 \tag{8}$$

Solve (8) for $\mu_i, \sigma_i, p(z = 1 \mid \theta)$ and λ respectively. So we got the following calculation results:

$$\begin{aligned} u_i &= \frac{\sum_{j=1}^{m} p(z_j = i \mid y_j, \theta_t) y_j}{\sum_{j=1}^{m} p(z_j = i \mid y_j, \theta_t)} \\ \sigma_i &= \frac{\sum_{j=1}^{m} p(z_j = i \mid y_j, \theta_t)(y_j - \mu_i)(y_j - \mu_i)^{\mathrm{T}}}{\sum_{j=1}^{m} p(z_j = i \mid y_j, \theta_t)} \\ p(z = i \mid \theta) &= \frac{1}{\lambda} \sum_{j=1}^{m} p(z_j = i \mid y_j, \theta_t) \\ \lambda &= \sum_{i=1}^{n} \sum_{j=1}^{m} p(z_j = i \mid y_j, \theta_t) \end{aligned} \tag{9}$$

These estimates now become our θ_{t+1}, which will be used in the next estimation step iteration. The EM iteration will be initialized with the parameters from the sample points and will not stop until parameters variations get minor between two iterations. When it converged, every pixel will be classified into a class which it belongs the most, that is to say with the global maximum-likelihood. Then we'll get the most precise classification result.

B. Unreliable Classified Pixels Processing

In our classification method, each pixel is classified into a class where the probability is the highest. At this point we may encounter the unreliable classification problem, this problem happens when the highest probability matching is still too low. As we can see in Figure 1, the red part of the Gaussian mixture model is unreliable part due to the gray level there is near the boundary of two classes.

To solve this problem in our segmentation method, we involve the principle of spatial from computer science. It's an empirically observed phenomenon that has substantial intuitive appeal and numerous practical implications [12]. The basic idea here is based on the fact that all the pixels within a relative close locality belong to same organ or anatomical configuration.

When dealing with an unreliable pixel, we find the closest high reliable pixel as its classification. That is

$$z = \begin{cases} \arg\max\limits_{i} p(x \mid z_i) & \text{when } \max(p(x \mid z_i)) > T \\ \arg\max\limits_{i} p(c \mid z_i) & \text{when } \max(p(x \mid z_i)) < T \end{cases} \tag{10}$$

Here T is the threshold parameter to divide reliable and unreliable, x is a pixel to classify, and c is the closest reliable pixel to x.

Noise is an annoying problem in medical image segmentation. Especially in our classification method based on gray level, noise will lead to incorrect segmentation. Let's consider a simple case, as shown in Fig.2(a). In this figure there are only two classes, one is black and the other is white. The desired segmentation result should be two completely separated rectangles, as shown in Fig.2(b). However there're some white noise points in black part, and vice versa. Obviously, if black and white are chosen as the initial parameter of EM, we won't get our desired result.

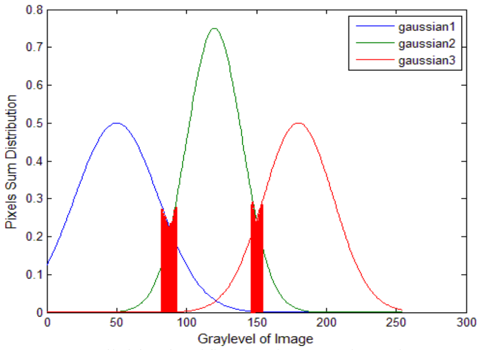
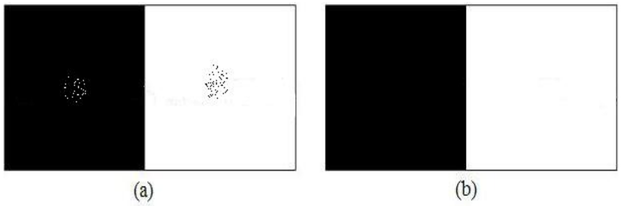


Figure 1. Unreliable classification in Gaussian mixture model



(a) (b)

Figure 2. (a) A sample image with simple noises (b) the desired classification result

To solve this problem, we involve the connected component labeling algorithm. In our method, we labeled the connected area using the algorithm presented in [13] based on the 8-neighbouring pixels. Any classification with too small connected area is considered as unreliable. Then we use the same processing method in previous part of this section to deal with the unreliable pixels.

C. Marker-Controlled Watershed

Usually the Watershed segmentation is used to segment the different elements of images, in the heart medical image context to get the different part of heart. However our researches tend to show that the watershed segmentation can gives pretty good results in heart boundary segmentation. Indeed by modifying the segmentation function using the results of 2 markers, one for the foreground and one for the background, good results can be obtained in heart segmentation. This segmentation is usually called Marker-Controlled Watershed Segmentation (MCWS) which is presented more deeply in [6] and [14]. In our case the MCWS is composed of 6 main parts:

1. Image preprocessing. This step mainly deals with noise and image optimization.
2. Segmentation function processing. By computing the image gradient magnitude.
3. Foreground markers computing. This process is obtained by using an "opening & closing by reconstruction" to get the main image elements which will be represented as blobs of pixels in the inside of each part of the heart image.
4. Background markers computing. These are pixels that are not part of any object.
5. Final segmentation using a modified segmentation function so that it only has minima at the foreground and background marker locations.
6. Compute the final boundary of the heart out of the segmentation results.

Dealing with the image quality and noise is really an important step in the watershed algorithm. Indeed, depending on the image attributes such as gamma or contrast, different results quality can occurs. Also noise can really lower the accuracy of our results. In our method we also use a Median filter which has an interesting propriety which is to keep the edges of our image [15]. This is exactly what we want since we are looking forward to getting the edges of the heart.

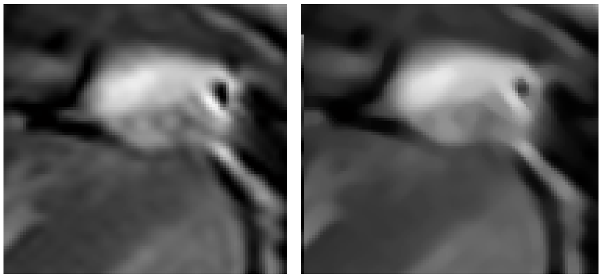


Figure 3. On the left a sample image with simple noises and on the right the same image after median filtering (strong filtering)

As you can see in Fig. 3, the median filtering here allows us to get remove some noisy part of the image and also clarify the boundary of heart. Thanks to this the variation of the gradient magnitude computing are clearer and enhance our final results.

All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

D. External Boundary Delimitation

Using the final result of the last part we have the segmentation of the full heart. The point of this part is to make an external detection of the watershed fragmented result by computing the 4-connected neighbor of the external heart using a modified version of the «Boundary Following Algorithm».

Indeed due to the configuration of the heart we had to change a little bit the BFA in order to correctly keep the edges of the heart even when there is multiple boundary connection. The specificity of the last part giving results of other image elements such as bones forced us to use a specific search system to get the Starting Point of the Boundary Following Algorithm of the heart depending of the type of image.

III. EXPERIMENTS

We cooperated with a clinical doctor who simply clicked several points in medical images, representing different classes for the segmentation respectively. These points were used as sample points in EM algorithm. In the experiments, there were 4 classes chosen which denoted ventricles, myocardium, bones and other tissues respectively, and 10 sample points for each class. So the time cost of manual operation was only a few seconds for each image. We used cardiac MRI datasets (Philips Intera Achieva, 256*256 images) and CT datasets (GE LightSpeed VCT, 512 * 512 images) in our experiments. Part of our experiment results are showed in Figure 4. The first row shows the original images, the second row shows the results after applying EM, and the last row shows the final segmentation results with exterior boundaries. The parameters used in EM, which were acquired by clicking several points in the images as we mentioned before are normalized and listed in Table 1. All the experiments were taken on a 1.8GHz Pentium dual core E2160 with 2GB DDR2 memory.

Table 1. EM parameters in experiment

Parameters		**Experiment Images**		
		Image1	**Image2**	**Image3**
Class1	μ	0.43451	0.469804	0.649804
	σ	0.003311	0.003536	0.003660
Class2	μ	0.196078	0.228627	0.465098
	σ	0.465098	0.001153	0.001375
Class3	μ	0.932549	0.871373	0.988235
	σ	0.003020	0.004533	0.000474
Class4	μ	0.018039	0.042353	0.116078
	σ	0.000578	0.002183	0.009547

We compared our results with manual segmentation and other methods by calculating the area differences of ventricles, and maximum closest distance of whole heart boundaries between them. Table 2 shows the results. We can see that the traditional segmentation methods encountered over-segmentation problem when dealing with whole heart boundary due to tissue conglutination and edge unclarity, while our method works pretty good with segmenting both interior anatomical configuration and exterior edge of heart.

The novel hybrid model presented in this paper shows a fast and robust whole heart segmentation, together with ventricle and myocardium. With this method we can get the whole heart boundary which is quite difficult when using traditional segmentation methods. Besides that the anatomical configuration such as ventricle and myocardium are segmented precisely too. Therefore it's applicable in a wide range of computer aided diagnosis and therapy plan.

Future works include the application of a heart shape model to improve the heart segmentation quality, and the extension into the 3-D segmentation, which is useful in application such as surgery navigation and virtual surgery training systems.

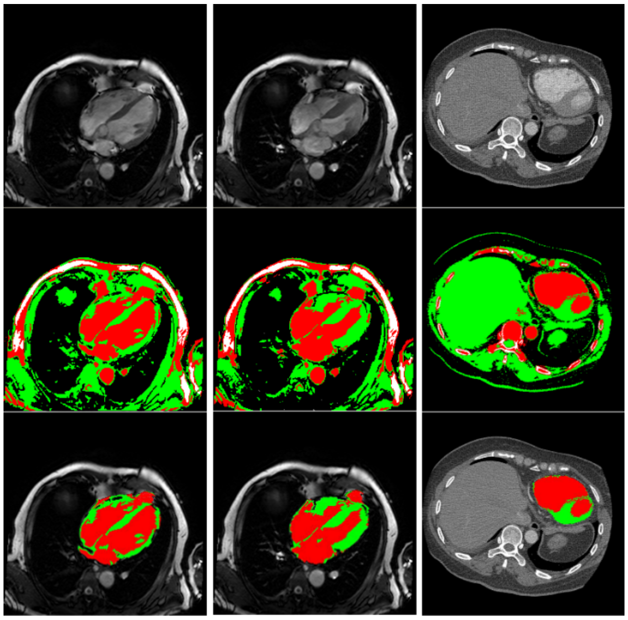


Image1(MRI) Image2(MRI) Image3(CT)

Figure 4. Experiments results

Table 2. Segmentation results evaluation

Segmentation Methods	Area difference of ventricle	Max. Closest Distance
Our method	6.4%	4.45mm
Level Set [3]	7.2%	n.a(1)
Watershed [6]	10.6%	n.a(1)

(1) No acceptable result acquired by this method due to seriously over-segmentation problem

ACKNOWLEDGMENT

This work is partial supported by the National Nature Science Foundation of China with grant number of 60872103, the National 863 High-Tech Plan with grant number of 2007AA01Z312 and the 973 Research Plan with grant number of 2007CB512700-1, 2006CB504801.

REFERENCES

[1] Bryan S. Morse, Lecture 18: Segmentation (Region Based), Brigham Young University, 1998-2000

[2] Suri. J.S, Liu. K, Singh. S, Laxminarayana. S.N, Zeng. X & Reden. L, “Shape Recovery Algorithms Using Level Sets in 2D/3D Medical Imagery: A State-of-the-Art Review”, IEEE Trans. Inf. Technol. Biomed, 2001, vol. 6, pp. 8-28

[3] S. Osher, J.A. Sethian, “Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations”, J. Comp. Phys., 1988, 79: pp. 12-49

[4] Cootes T.F, Taylor J.C, Cooper D.H, Graham J, “Active Shape Models-Their training and application”, Computer Vision and Image Underst, 1995, 61(1):pp.1-10

[5] Frangi A.F, Niessen W.J, Viergever M.A, “Three Dimensional Modeling for Functional Analysis of Cardiac Images: A Review”, IEEE Trans Med Imag, 2001, 20(1)

[6] Serge Beucher and Christian Lantuéjoul, “Use of watersheds in contour detection”, International workshop on image processing, real-time edge and motion detection, 1979

[7] The World Health Report, World Health Organization (WHO), 2008[Online], available: http://www.who.int

[8] T. Schwarz, T. Hermann, I Wolf, and H.P. Meinzer, “3D Heart Segmentation and Volumetry Using Deformable Shape Models”, Computers in Cardiology, 2007, 34: pp. 741-744

[9] J.A. Bilmes, “A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models”, International Computer Science Institute Berkeley CA 94704 and Computer Science Division Department of Electrical Engineering and Computer Science U.C. Berkeley, TR-97-021, April 1998

[10] A. Dempster, N. Laird, and D. Rubin, “Maximum Likelihood from Incomplete Data via the EM Algorithm”, J. Royal Statist. Soc. Ser. B., 39, 1997

[11] V. Castelli, T. Cover, “On the Exponential Value of Labeled Sample”, Pattern Recognition Letters, 1995, vol: 16, issue: 1, pp. 105-111

[12] Aho, Lam, Sethi, and Ullman, “Compilers: Principles, Techniques & Tools” 2nd ed., Pearson Education, 2007

[13] J.M. Park, C.G. Looney, H.C. Chen, “Fast Connected Component Labeling Algorithm Using a Divide and Conquer Technique”, TR-2004-04, the Univ. of Alabama, 2000

[14] Beucher, S. and Meyer, F., The morphological approach to segmentation: the watershed transformation. In: Dougherty, E. (Ed.), Mathematical Morphology in Image Processing, Marcel Dekker, New York. pp. 433-481, 1992.

[15] Conrad Bovik A., Huang T. S., Munson D. C. JR, “The effect of median filtering on edge estimation and detection”. In: IEEE transactions on pattern analysis and machine intelligence , IEEE Computer Society, Los Alamitos, CA, ETATS-UNIS 1987