A Learning Framework for the Automatic and Accurate Segmentation of Cardiac Tagged MRI Images

Zhen Qian[1], Dimitris N. Metaxas[1], and Leon Axel[2]

[1] Center for Computational Biomedicine Imaging and Modeling (CBIM),
Rutgers University, New Brunswick, New Jersey, USA
[2] Department of Radiology, New York University, New York, USA

Abstract. In this paper we present a fully automatic and accurate segmentation framework for 2D tagged cardiac MR images. This scheme consists of three learning methods: a) an active shape model is implemented to model the heart shape variations, b) an Adaboost learning method is applied to learn confidence-rated boundary criterions from the local appearance features at each landmark point on the shape model, and c) an Adaboost detection technique is used to initialize the segmentation. The set of boundary statistics learned by Adaboost is the weighted combination of all the useful appearance features, and results in more reliable and accurate image forces compared to using only edge or region information. Our experimental results show that given similar imaging techniques, our method can achieve a highly accurate performance without any human interaction.

1 Introduction

Tagged cardiac magnetic resonance imaging(MRI) is a well known technique for non-invasively visualizing the detailed motion of myocardium throughout the heart cycle. This technique has the potential of early diagnosis and quantitative analysis of various kinds of heart diseases and malfunction. However, before it can be used in routine clinical evaluations, an imperative but challenging task is to automatically find the boundaries of the epicardium and the endocardium. (See Figure 1(a-c) for some examples.)

Segmentation in tagged MRI is difficult for several reasons. First, the boundaries are often obscured or corrupted by the nearby tagging lines, which makes the conventional edge-based segmentation method infeasible. Second, tagged MRI tends to increase the intensity contrast between the tagged and un-tagged tissues at the price of lowering the contrast between the myocardium and the blood. At the same time, the intensity of the myocardium and blood vary during the cardiac cycle due to the tagging lines fading in the myocardium and being flushed away in the blood. Third, due to the short acquisition time, the tagged MR images have a relatively high level of noise. These factors make conventional region-based segmentation techniques impractical. The last and the most

Y. Liu, T. Jiang, and C. Zhang (Eds.): CVBIA 2005, LNCS 3765, pp. 93–102, 2005.

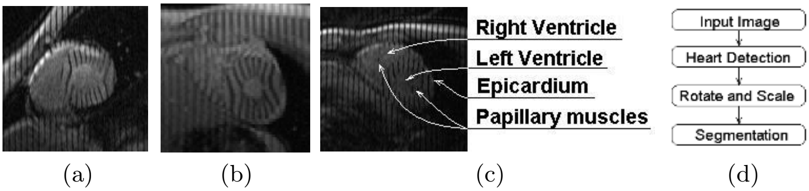


Fig. 1. (a-c) Some examples of tagged cardiac MRI images. The task of segmentation is to find the boundaries of the epicardium and endocardium (including the LV and RV and excluding the papillary muscles.) (d) The framework of our segmentation method.

important reason is that, from the clinicians' point of view, or for the purpose of 3D modeling, *accurate* segmentation based solely on the MR image is usually not possible. For instance, for conventional clinical practice, the endocardial boundary should exclude the papillary muscles for the purpose of easier analysis. However, in the MR images, the papillary muscles are often apparently connected with the endocardium and cannot be separated if only the image information is used. Thus prior shape knowledge is needed to improve the results of automated segmentation.

There have been some efforts to achieve tagged MRI segmentation. In [1], grayscale morphological operations were used to find non-tagged blood-filled regions. Then they used thresholding and active contour methods to find the boundaries. In [2], a learning method with a coupled shape and intensity statistical model was proposed. In [3,4], Gabor filtering was used to remove the tagging lines before the segmentation. These methods work in some cases. However they are still imperfect. In [1], morphological operations are sensitive to image noise, and the active contour method tends to get irregular shapes without a prior shape model. In [2], their intensity statistical model cannot capture the complex local texture features, which leads to inaccurate image forces. And in [3,4], the filtering methods blur the boundaries and decrease the segmentation accuracy.

In this paper, in order to address the difficulties stated above, we propose a novel and fully automatic segmentation method based on three learning frameworks: 1. An active shape model(ASM) is used as the prior heart shape model. 2. A set of confidence-rated local boundary criteria are learned by Adaboost, a popular learning scheme (see Section 2.2), at landmark points of the shape model, using the appearance features in the nearby local regions. These criteria give the probability of the local region's center point being on the boundary, and force their corresponding landmark points to move toward the direction of the highest probability regions. 3. An Adaboost detection method is used to initialize the segmentation's location, orientation and scale. The second component is the most essential contribution of our method. We abandon the usual edge or region-based methods because of the complicated boundary and region appearance in the tagged MRI. It is not feasible to designate one or a few edge or region rules to solve the complicated segmentation task. Instead, we try to use all possible information, such as the edges, the ridges, and the breaking points of tagging lines, to form a *complex rule*. It is apparent that at different locations on the heart boundary, this *complex rule* must be different, and our confidence in the *complex rule* varies too. It is impractical to manually set up each of these *complex rules* and weight their confidence ratings. Therefore, we implement Adaboost to learn a set of rules and confidence ratings at each landmark point on the shape model. The first and the second frameworks are tightly coupled. The shape model deforms under the forces from Framework 2 while controlled and smoothed by Framework 1. To achieve fully automatic segmentation, in Framework 3 the detection method automatically provides an approximate position and size of the heart to initialize the segmentation step. See Figure 1(d) for a complete illustration of the frameworks.

The remainder of this paper is organized as follows: in Section 2, we present the segmentation methodology, including Frameworks 1 and 2. In Section 3, we briefly introduce the heart detection technique of Framework 3. In Section 4 we give some details of our experiments and show some encouraging initial experimental results.

2 Segmentation Based on ASM and Local Appearance Features Learning Using Adaboost

There has been some previous research on ASM segmentation methods based on local features modeling. In [5], a statistical analysis was performed, which used sequential feature forward and backward selection to find the set of optimal local features. In [6], an EM algorithm was used to select Gabor wavelet-based local features. These two methods tried to select a small number of features, which is impractical to represent complicated local textures such as in tagged MRI. In [7], a simple Adaboost learning method was proposed to find the optimal edge features. This method didn't make full use of the local textures, and didn't differentiate each landmark point's confidence level. In our method, similarly using Adaboost, our main contributions are: the ASM deforms based on a more *complex* and robust rule, which is learned from the local appearance, not only of the edges, but also ridges and tagging line breakpoints. In this way we get a better representation of the local appearance of the tagged MRI. At the same time, we derive the confidence rating of each landmark point from their Adaboost testing error rates, and use these confidence ratings to weight the image forces on each landmark point. In this way the global shape is affected more by the *more confident* points and we eliminate the possible error forces generated from the *less confident* points.

2.1 ASM Shape Model

Since the shape of the mid portion of the heart in short axis (SA) images is consistent and topologically fixed (one left ventricle (LV) and one right ventricle (RV)), it is reasonable to implement an active shape model [8] to represent the desired boundary contours.

We acquired two image datasets each, from two normal subjects, using two slightly different imaging techniques. The datasets were acquired in the short axis plane. There are two sets of tagging line orientations (0° and 90°, or −45° and 45°) and slightly different tag spacings. Each dataset included images acquired at phases through systole into early diastole, and at positions along the axis of the LV, from near the apex to near the base, but without topological changes. An expert was asked to segment the epicardium (Epi), the left ventricle (LV) endocardium and the right ventricle (RV) endocardium from the datasets. In total, we obtained 220 sets (each set includes one LV, one RV, and one Epi) of segmented contours to use as the training data.

Segmented contours were centered and scaled to a uniform size. Landmark points were placed automatically by finding key points with specific geometric characteristics. As shown in Figure 2(a), the black points are the key points, which were determined by the curvatures and positions along the contours. For instance, $P1$ and $P2$ are the highest curvature points of the RV; $P7$ and $P8$ are on opposite sides of the center axis of the LV. Then, fixed numbers of other points are equally placed in between. In this way, the landmark points were registered to the corresponding locations on the contours. Here, we used 50 points to represent the shape.

For each set of contours, the 50 landmark points (x_i, y_i) were reshaped to form a shape vector $X = (x_1, x_2, ..., x_{50}, y_1, y_2, ..., y_{50})^T$. Then Principal Component Analysis was applied and the modes of shape variation were found. Any heart shape can be approximately modeled by $X = \bar{X} + Pb$, where $\bar{X}$ is the mean shape vector, P is the matrix of shape variations, and b is the vector of shape parameters weighting the shape variations.

After we find the image forces at each landmark point, as in Section 2.2, the active shape model evolves iteratively. In each iteration, the model deforms under the influence of the image forces to a new location; the image forces are then calculated at the new locations before the next iteration.

2.2 Segmentation Via Learning Boundary Criteria Using Adaboost

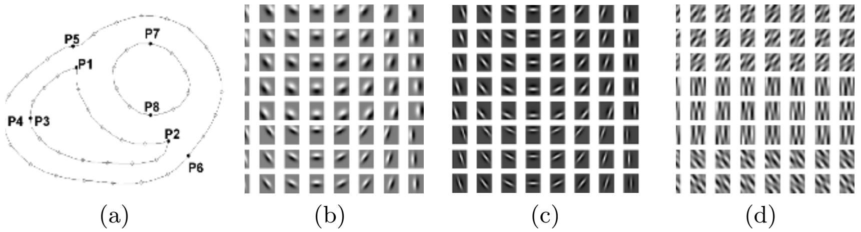
Feature Design. To capture the local appearance characteristics, we designed three different kinds of steerable filters. We use the derivatives of a 2D Gaussian to capture the edges, we use the second order derivatives of a 2D Gaussian to capture the ridges, and we use half-reversed 2D Gabor filters [9] to capture the tagging line breakpoints.

Assume $G = G((x - x_0)\cos(\theta), (y - y_0)\sin(\theta), \sigma_x, \sigma_y)$ is an asymmetric 2D Gaussian, with effective widths σ_x and σ_y, a translation of (x_0, y_0) and a rotation of θ. We set the derivative of G to have the same orientation as G:

$$G' = G_x \cos(\theta) + G_y \sin(\theta) \tag{1}$$

The second derivative of a Gaussian can be approximated as the difference of two Gaussians with different σ. We fix σ_x as the long axis of the 2D Gaussians, and set $\sigma_{y2} > \sigma_{y1}$. Thus:

$$G'' = G(\sigma_{y1}) - G(\sigma_{y2}) \tag{2}$$



(a) (b) (c) (d)

Fig. 2. (a) shows the automatic method used to place the landmark points. (b-d) are the sample sets of feature filters: (b) are the derivatives of Gaussian used for edge detection, (c) are the second derivatives of Gaussian used for ridge detection, and (d) are the half-reversed Gabor filters used for tag line breakpoint detection.

In the previous two equations, we set $x_0 = 0$, and tune y_0, θ, σ_x, σ_y, σ_{y1} and σ_{y2} to generate the desired filters.

The half-reversed 2D Gabor filters are defined as a 2D sine wave multiplied with the 2D derivative of a Gaussian:

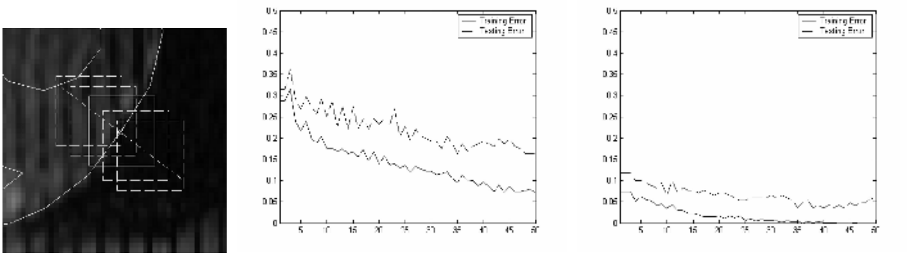
$$F = G'(x, y) \cdot \mathbb{R}\{e^{-j[\phi+2\pi(Ux+Vy)]}\} \tag{3}$$

where G' is the derivative of a 2D Gaussian. U and V are the frequencies of the 2D sine wave, $\psi = arctan(V/U)$ is the orientation angle of the sine wave, and ϕ is the phase shift. We set $x_0 = 0$, $\sigma_x = \sigma_y = \sigma$,$-45° \leq \psi - \theta \leq 45°$, and tune y_0, θ, σ, ϕ, U and V to generate the desired filters.

For a 15x15 sized window, we designed 1840 filters in total. See Figure 2(b-d) for some sample filters.

Adaboost Learning. In the learning section, each training image is scaled proportionally to the scaling of its contours. At each landmark point of the contours, a small window (15x15) around it was cut out as a positive appearance training sample for this particular landmark point. Then along the normal of the contour, on each side of the point, we cut out two 15x15-sized windows as negative appearance training samples for this particular landmark point. Thus for each training image, at a particular landmark point, we got one positive sample and four negative samples (shown in Figure 3(a).) We also randomly selected a few common negative samples outside the heart or inside the blood area, which are suitable for every landmark point. For image contrast consistency, every sample was histogram equalized.

The function of the Adaboost algorithm [10,11] is to classify the positive training samples from the negative ones by selecting a small number of important features from a huge potential feature set and creating a weighted combination of them to use as an accurate strong classifier. During the boosting process, each iteration selects one feature from the total potential features pool, and combines it (with an appropriate weight) with the existing classifier that was obtained in the previous iterations. After many iterations, the weighted combination of the selected important features can become a strong classifier with high accuracy.



(a) (b) (c)

Fig. 3. (a) shows the method of setting the training data. The solid box is the positive sample around the landmark points. The four dashed line boxes along the normal are the negative samples. This way of setting the negative samples is chosen to make the classifier more adaptive to the particular landmark position. (b) and (c) show the training error (solid lines) and testing error (dash lines) of two landmark points versus Adaboost iteration times. (b) is a point on the LV, (c) is a point on the Epi. Note how the training and testing error decrease as Adaboost iterates. Also note the testing error of (b) is higher than (c): we are more confident of landmark point (c)'s classification result.

The output of the strong classifier is the weighted summation of the outputs of each of its each selected features, or, the weak classifiers: $F = \Sigma_t \alpha_t h_t(x)$, where α are the weights of weak classifiers, and h are the outputs of the weak classifiers.

We call F the boundary criterion. When $F > 0$, Adaboost classifies the point as being on the boundary. When $F < 0$, the point is classified as off the boundary. Even when the strong classifier consists of a large number of individual features, Adaboost encounters relatively few overfitting problems [12]. We divided the whole sample set into one training set and one testing set. The function of the testing set is critical. It gives a performance measure and a confidence level that tells us how much we should trust its classification result. Figure 3(b, c) shows the learning error curve versus the boosting iteration numbers at two selected landmark points. Note that every landmark point i has its own α, h and F_i.

Segmentation Based on Confidence Ratings. In the segmentation stage, we first select an initial location and scale, and then overlay the mean shape $\bar{X}$, which is obtained from ASM, onto the task image. In section 3 we describe an automatic initialization method.

At a selected landmark point i on the shape model, we select several equally spaced points along the normal of the contour on both sides of i, and use their F values to examine the corresponding windows centered on these points. In [12], a logistic function was suggested to estimate the relative boundary probabilities:

$$Pr(y = +1|x) = \frac{e^{F(x)}}{e^{F(x)} + e^{-F(x)}} \tag{4}$$

We find a point j whose test window has the highest probability of being on the heart boundary. Thus an image force $\boldsymbol{f}$ should push the current landmark point

i toward j. Recall that, as discussed in the previous subsection, Adaboost gives the errors of the testing data e_i. We define the confidence rating as:

$$c_i = \ln \frac{1}{e_i}; \tag{5}$$

Intuitively, when c_i is big, we trust its classification and increase the image force $\boldsymbol{f}$, and conversely. Thus, we define the image force at landmark point i as:

$$\boldsymbol{f} = \mu \cdot \frac{[\boldsymbol{x}(j) - \boldsymbol{x}(i)] \cdot c(i)}{||\boldsymbol{x}(j) - \boldsymbol{x}(i)||_2} \tag{6}$$

where μ is a scale as a small step size.

The detailed algorithm to update the parameters of the ASM model with the image force $\boldsymbol{f}$ can be found in [8].

3 Heart Detection Based on Adaboost Learning

The heart detection algorithm used is influenced by the Adaboost face detection algorithm developed by Paul Viola and Michael Jones [13]. The reason we adapt a face detection method is that these two problems are closely related. Often, there are marked variations between different face images, which come from different facial appearance, lighting, expression, etc. In heart detection, we have the similar challenges: the heart images have different tag patterns, shape, position, phase, etc.

We use the same Haar wavelet features as in [13]. The training data contained 297 manually cropped heart images and 459 randomly selected non-heart images. The testing data consisted of 41 heart images and 321 non-heart images. These data were resized to 24x24 pixels and contrast equalized. Adaboost training gave a strong classifier by combining 50 weak features. For an input task image, the detection method searched every square window over the image, and found a window with the highest probability as the final detection. If we rotate the task

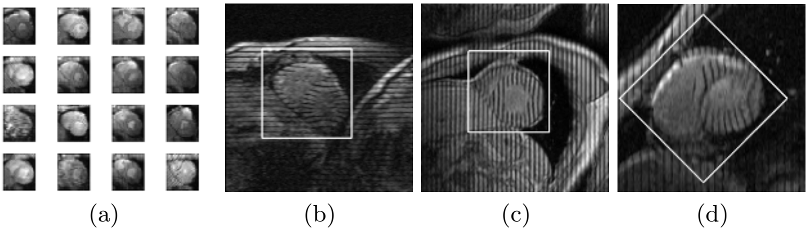


Fig. 4. (a) shows a few samples of the training data. (b), (c) and (d) are three detection results. For image (d), the image was rotated by a set of discrete angles before the detection, and the final detection is of the highest probability among all the discrete angles tested.

image by a set of discrete angles before the detection procedure, and compare the probabilities across the discrete angles, we are also able to detect hearts in rotated images (see Figure 4).

4 Representative Experimental Results and Validation

We applied our segmentation method to three data sets, one from the same subject and with the same imaging settings as the training data (but excluding the training data), and the other two novel data sets from two different subjects and with slightly different imaging settings. Respectively, the three data sets each contained 32, 48 and 96 tagged MRI images, with different phases, positions and tagging orientations. Each task image was rotated and scaled to contain a 80x80-pixel-sized chest-on-top heart, using the detection method before the segmentation. Each segmentation took 30 iterations to converge. Our experiment was coded in Matlab 6.5 and run on a PC with dual Xeon 3.0G CPUs and 2G memory. The whole learning process took about 20 hours. The segmentation process of one heart took 120 seconds on average. See Figure 5 for representative results.

For validation, we used the manual segmentation contours as the ground truth for the first and second data sets. For the third data set, because we don't have independent manual contours, we used cross validation, since we know that at the same position and phase, the heart shapes in the vertical-tagged and horizontal-tagged images should be similar. We denote the ground truth contours as T and our segmentation contours as S. We defined the average error distance as $\bar{D}_{error} = mean_{s_i \in S}(min||T - s_i||_2)$. Similarly the cross distance is defined as $\bar{D}_{cross} = mean_{s_i^{vertical} \in S^{vertical}}(min||S^{horizontal} - s_i^{vertical}||_2)$. In a 80x80 pixel-sized heart, the average error distances between the automatically segmented contours and the contours manually segmented by the expert for the first data set were: $\bar{D}_{error}(LV) = 1.12$ pixels, $\bar{D}_{error}(RV) = 1.11$ pixels, $\bar{D}_{error}(Epi) = 0.98$ pixels. For the second data set, $\bar{D}_{error}(LV) = 1.74$ pixels, $\bar{D}_{error}(RV) = 2.05$ pixels, $\bar{D}_{error}(Epi) = 1.33$ pixels. In the third dataset, the cross distances are: $\bar{D}_{cross}(LV) = 2.39$ pixels, $\bar{D}_{cross}(RV) = 1.40$ pixels, $\bar{D}_{cross}(Epi) = 1.94$ pixels. The larger distance in the cross validation arises in part from underlying misregistration between the (separately acquired) horizontal and vertical images. Thus, the true discrepancy due to the segmentation should be smaller. From the above quantitative results, we find that for a normal-sized adult human heart, the accuracy of our segmentation method achieves an average error distance of less than 2mm. The cross validation results of the third data set suggest that our method is very robust as well.

5 Discussion

In this paper, we have proposed a learning scheme for fully automatic and accurate segmentation of cardiac tagged MRI data. The framework has three steps. In the first step we learn an ASM shape model as the prior shape constraint. Second, we learn a confidence-rated complex boundary criterion from the local

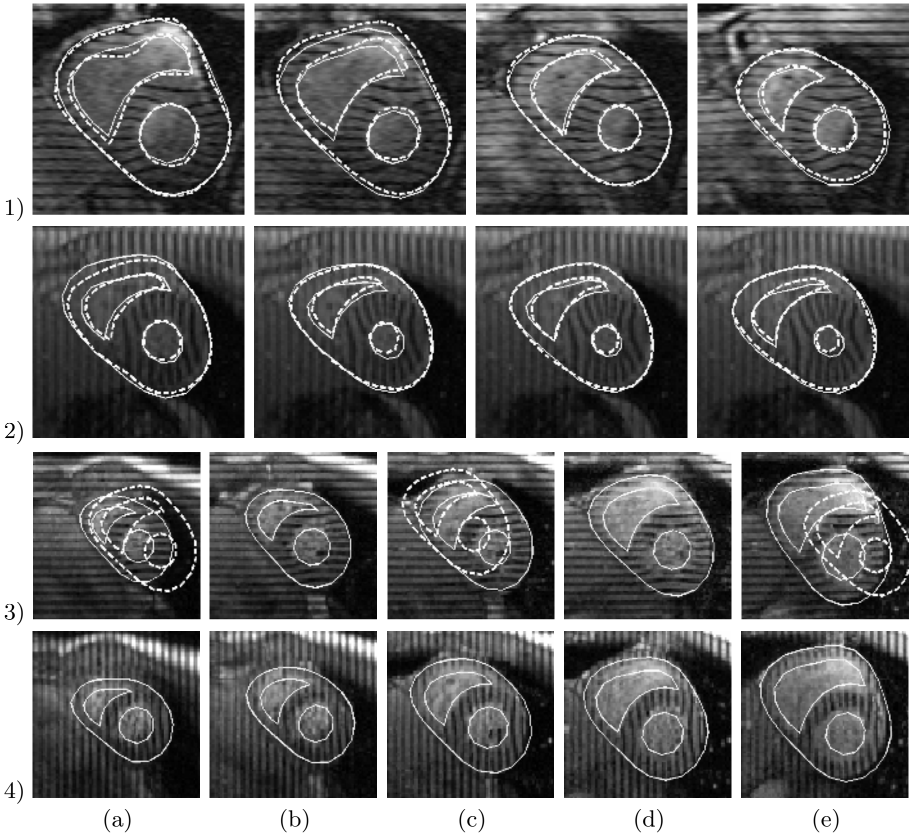


Fig. 5. The first and second rows of images come from the the first and second dataset, respectively. For better representation, the images in the first row vary in position and remain at the same phase, while the images in the second row vary in phase but remain at the same position. The solid contours are from our automatic segmentation method; the dashed contours are manual. Notice that the papillary muscles in LV are excluded from the endocardium. The third and fourth rows are from the third dataset. Manual contours are not available for this dataset, so we compare our segmentation results between the the horizontal and vertical tagged images that are at same position and phase. Qualitatively, the contours are quite consistent, allowing for possible misregistration between the nominally corresponding image sets. In (3a), (3c) and (3e) the dashed contours are testing examples of poor initializations, while the final contours are solid. Although the initialization is fay away from the target, the shape model moves and converges well to the target.

appearance features to use to direct the detected contour to move under the influence of image forces. Third, we also learn a classifier to detect the heart. This learning approach achieves higher accuracy and robustness than other previously available methods. Since our method is entirely based on learning, the way of choosing the training data is critical. We find that if the segmentation method is applied to images at phases or positions that are not represented in the training data, the segmentation process tends to get stuck in local minima. Thus the training data need to be of sufficient size and range to cover all possible variations that may be encountered in practice.

An interesting property of our method is that it is not very sensitive to the initialization conditions. As shown in Figure 5, even if the initial contours are far away from the target position, it can still eventually converge to the right position after a few iterations. This property makes automatic initialization feasible. The detection method gives only a rough approximation of the heart's location and size, but it is good enough for our segmentation purposes.

References

1. Montillo, A., Metaxas, D., Axel, L.: Automated segmentation of the left and right ventricles in 4d cardiac spamm images. In: Medical Imaging Computing and Computer-Assisted Intervention. (2002) 620–633
2. Huang, X., Li, Z., Metaxas, D.N.: Learning coupled prior shape and appearance models for segmentation. In: MICCAI (1). (2004) 60–69
3. Qian, Z., Huang, X., Metaxas, D., Axel, L.: Robust segmentation of 4d cardiac mri-tagged images via spatio-temporal propagation. In: Proceedings of SPIE Medical Imaging: Physiology, Function, and Structure from Medical Images. Volume 5746. (2005) 580–591
4. Metaxas, D., Chen, T., Huang, X., Axel, L.: Cardiac segmentation from mri-tagged and ct images. In: 8th WSEAS International Conf. on Computers, special session on Imaging and Image Processing of Dynamic Processes in biology and medicine. (2004)
5. Ginneken, B.V., Frangi, A.F., Staal, J.J., et al: Active shape model segmentation with optimal features. IEEE Trans. on Medical Imaging **21** (2002)
6. Jiao, F., Li, S., Shum, H., Schuurmans, D.: Face alignment using statistical models and wavelet features. In: IEEE Conf. on CVPR. Volume 1. (2003) 321–327
7. Li, S., Zhu, L., Jiang, T.: Active shape model segmentation using local edge structures and adaboost. In: Medical Imaging Augmented Reality. (2004)
8. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models - their training and application. Computer Vision and Image Understanding **61** (1995) 38–59
9. Daugman, J.: Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. Journal of the Optical Society of America A **2** (1985) 1160–1169
10. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. In: EuroCOLT '95: Proceedings of the Second European Conference on Computational Learning Theory. (1995) 23–37
11. Schapire, R.E.: The boosting approach to machine learning: An overview. In: In MSRI Workshop on Nonlinear Estimation and Classification. (2002)
12. R. E. Schapire, Y. Freund, P.B., Lee, W.S.: Boosting the margin: a new explanation for the effectiveness of voting methods. Annals of Statistics **26** (1998) 1651–1686
13. Viola, P., Jones, M.: Robust real-time object detection. Second International Workshop on Statistical and Computational Theories of Vision - Modeling, Learning, And Sampling. Vancouer, Canada, July 13 (2001)