

Echocardiographic Contour Extraction with Local and Global Priors through Boosting and Level Sets

Ayşe Betül Oktay and Yusuf Sinan Akgül

GIT Vision Lab, <http://vision.gyte.edu.tr/>

Department of Computer Engineering, Gebze Institute of Technology,

Gebze, Kocaeli 41400 Turkey

{oktay, akgul}@bilmuh.gyte.edu.tr

Abstract

Extracting endocardium and epicardium from echocardiographic images is a challenging task because of large amounts of noise, signal drop-out, unrelated structures, and unseen wall parts. This paper introduces a new technique that automatically extracts cardiac borders by incorporating local and global priors through boosting and level set methods. The shape-based global prior is incorporated into the system by regularly re-initializing the level set surface under the influence of the expert detected contours. The local priors with image and temporal information are learned through boosting.

The proposed system has many advantages. First, boosting encodes the knowledge about the image information and the temporal cardiac wall motion effectively by using spatiotemporal filters. Second, the local priors can use any features from the images including different filters and intensity profiles. Furthermore, other hard constraints like local shape, texture, distance, etc. can be added to local features effectively. The system is validated on echocardiograms and the results are found to be promising.

1. Introduction

Echocardiographic image segmentation is a challenging task because of ultrasound related problems like signal drop-out, large amounts of speckle noise, unseen side walls, and papillary muscles interfering with the endocardium. These problems cause large variations even between the delineations of human experts (Figure 1). As a result, automatic extraction of the cardiac walls gains importance to provide more accurate segmentations and fast consultation.

In order to address the issues with the echocardiographic image segmentation, domain related information was brought into the segmentation in the form of shape priors, e.g. [2]), where the geometry of the contours is learned

from a number of examples and this information is used during the segmentation process. The inclusion of shape prior information in the echocardiographic segmentation process improved the overall results significantly for the borders with no image information. However, prior shape information is sometimes not enough to recover a novel shape with very low image gradient magnitude. Chen *et al.* [1] suggested using intensity profiles of object contours by training an intensity model from a set of images. A similar intensity and curvature profile information was used earlier by Leventon *et al.* [6]. Although using image or curvature profiles from the training images is very useful in recovering segment borders, it is difficult to use other image and non-image based features with these methods.

This paper presents a new approach for level set based echocardiographic image segmentation that uses what we call global and local priors. Global priors include classical shape priors that impose expected geometric shape related constraints into the segmentation. Our current global prior is based on deformable matching of expert shapes with the current shape during the level set optimization. The local priors impose constraints from domain dependent data that describes the segment borders only locally. Local priors include the image priors and other locally definable features of the segment borders such as local geometric shape, distance from the shape reference point, etc. In order to bring together different local features into the local prior information, we use boosting [4], which is popularly used by the medical imaging community for handling complex structures since the shape, texture, and temporal information can be encoded easily through features.

As one of the boosting features, we use the distance from the heart center. The other features are extracted by a number of Haar-like 3D spatiotemporal and 2D spatial filters on echocardiographic images in polar coordinates. A separate local prior training is performed for a number of angle ranges in the polar space; so that an angle sensitive image

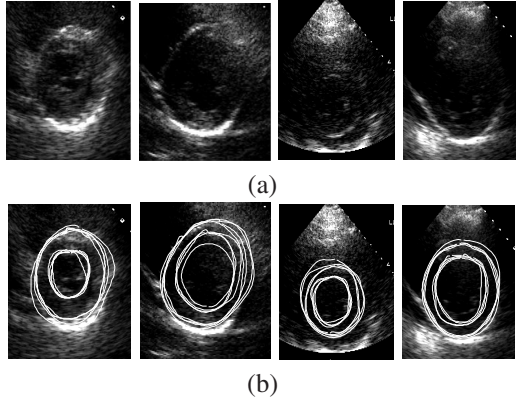


Figure 1. (a) Four sample echocardiographic images; (b) Each image delineated by four different experts.

border information can be learned. The separate training of endocardium and epicardium produces different local priors for inner and outer walls that brings more image information to the segmentation process. Finally, the local and global priors are used under a level set framework to produce the heart wall contours.

Our employment of boosting as the local prior information tool has several advantages. First, by including 3D spatiotemporal Haar-like filters, our system automatically incorporates temporal heart wall information into the segmentation process without using an explicit 3D segmentation model. This is an important advantage in echocardiography because some parts of the heart walls can be located only by seeing them in motion. Second, the interaction calculations between the heart wall angle and the ultrasound beam angle are handled by boosting training without any explicit formulations. Third, our local prior can use any feature from the images including different filters or intensity profiles. For example, the image prior methods of Chen *et al.* [1] can be incorporated as another feature in our system. Furthermore, our local prior can include non-image related local features such as the geometric relationships between neighboring border points. Finally, the separately trained inner and outer wall information can be conveniently used in level set segmentation to eliminate any ambiguity between endocardium and epicardium.

The rest of the paper is organized as follows: Section 2 includes the level set formulation with our new local and global priors. The experiments are presented in Section 3. Finally, we conclude in Section 4.

2. Segmentation with Level Set Method

We use a variational level set approach [5] to extract the cardiac walls from the echocardiographic images by employing global and local shape priors. Global shape prior is imposed during the re-initialization of the evolving surface.

Local shape priors are learned by Haar-like 3D spatiotemporal and 2D spatial filters from echocardiograms and the pixels are given scores that indicates whether they are an epicardium or endocardium border through Adaboost.

2.1. The Variational Level Set Formulation

For the cardiac wall extraction, let $c_1(t)$ and $c_2(t)$ be two closed curves evolving on the plane \mathbb{R}^2 with time t where $c_1(t)$ is used for extracting the endocardium and $c_2(t)$ is used for extracting the epicardium.

Consider C as the set of points on $c_1(0)$ and $c_2(0)$ and ϕ as a signed distance function. Let \mathbf{x} be the position vector and ϕ be a signed distance function defined as:

$$\phi(\mathbf{x}) = \begin{cases} 0, & \text{if } \mathbf{x} \in C, \\ -d(x), & \text{if } \mathbf{x} \text{ is outside } c_1 \text{ but inside } c_2, \\ d(x), & \text{otherwise.} \end{cases} \quad (1)$$

where d is the shortest Euclidian distance to C from point \mathbf{x} .

According to the variational level set formulation, the 3D surface ϕ is evolved under the influence of the internal energy term $P(\phi)$ and the external energy term $\varepsilon_m(\phi)$. Internal energy term function is:

$$P(\phi) = \int_{\Omega} \frac{1}{2} (|\nabla\phi| - 1)^2 dx dy. \quad (2)$$

Let I^{endo} include the score of each pixel assigned by Adaboost for endocardium. Let I^{epi} be the similar image for epicardium. We will explain how they are formed in Section 2.4. The surface evolves on image I which includes scores through Adaboost. It is defined as:

$$I(\mathbf{x}) = \begin{cases} \frac{1}{1+I^{endo}(x)}, & \text{if } \mathbf{x} \text{ is closer to } c_1, \\ \frac{1}{1+I^{epi}(x)}, & \text{if } \mathbf{x} \text{ is closer to } c_2. \end{cases} \quad (3)$$

The external energy term is used for moving the contours $c_1(t)$ and $c_2(t)$ towards the cardiac walls defined in I^{endo} and I^{epi} , respectively. The external energy term consists of the image length and area of the zero level contour of ϕ :

$$\varepsilon_{I,\lambda,v} = \lambda \int_{\Omega} I \delta(\phi) |\nabla\phi| dx dy + v \int_{\Omega} I H(-\phi) dx dy, \quad (4)$$

where λ and v are weighting parameters, δ is the univariate Dirac function and H is the Heaviside function.

The incorporation of global information is achieved by stopping the surface deformations and re-initializing the surface under the influence of shape knowledge at regular intervals.

2.2. Incorporating Global Information

In order to incorporate global prior information into the level set deformation, instead of adding a new energy term to the level set functional, we prefer to re-initialize the surface according to the most similar expert contours which is explained in our previous work [8] in detail.

We follow a piecewise uniform scaling approach in selecting the most similar expert contours. Let c_1 and c_2 be the zero level contours and c_1^e, c_2^e be the expert delineated contours expressed in polar coordinates (in $\theta - r$ space), where $e = 1 \dots m$ and m is the number of experts. First, we transform each expert contour to the zero level contour and then find the best matching expert contour as follows:

- Divide contours c_1, c_2, c_1^e and c_2^e into θ ranges.
- For each range, calculate a local scaling factor between c and c^e by dividing average r values of c to average r values of c^e of that θ range (Figure 2-a).
- For each range, transform the expert contour c^e to c by multiplying the r positions of the contour with the local scaling factor (Figure 2-b). The transformed expert contour is called c_t^e .
- Transform other expert contours as defined above.
- The expert contour c_t^e that has the minimum absolute difference is selected as the most similar contour.

The best matching contours for endocardium and epicardium are found separately as explained above. The surface that contains the global shape information is constructed by embedding the best matching transformed contours (endocardium and epicardium) into the zero level of the surface according to the Equation 1. The constructed surface is re-initialized on the image and thus the global shape information is incorporated by re-initialization.

2.3. Local Priors

Our local prior information includes locally definable features around the cardiac borders. In order to capture local features better, we choose to transform the rectangular echocardiographic images to polar coordinates using the heart center as the polar origin (Figure 3). First, any translation difference between different images is minimized because in the polar coordinates, the heart centers of all images are registered to origin. As a result, using local heart wall distance to polar origin becomes a valuable feature that can be used as a local prior information. Second, in polar coordinates, the heart wall orientations are similar (all almost horizontal). We can use a smaller number of filter orientations to capture the contour orientations, which would be very useful in decreasing the number of Haar-like filters

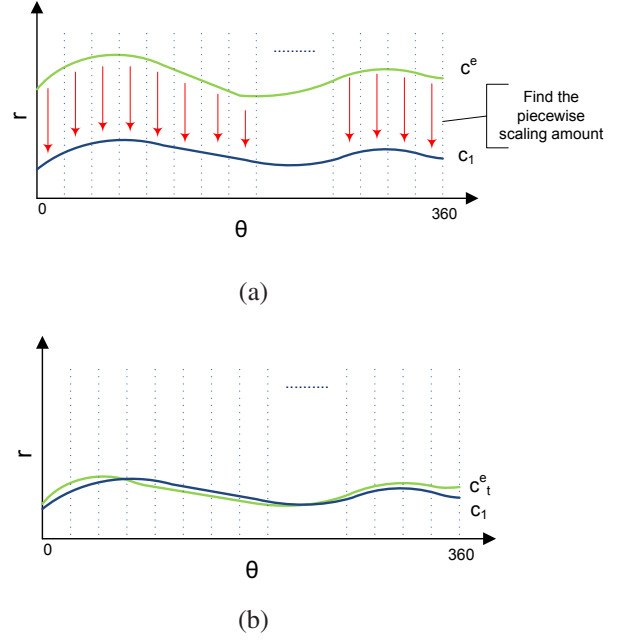


Figure 2. (a) The zero level and expert contour before transformation; (b) Local scaling factor is calculated for each θ range and the expert contour is transformed to zero level contour.

in boosting. Similar polar coordinate techniques on medical images are used in the literature, *e.g.* [9]. The downside of using polar coordinates is to choose the heart center before the extraction of heart walls. Although it can be estimated automatically as in [11], we give the heart centers manually during the system run.

The cardiac borders have different image gradient characteristics at different θ positions in the polar images (Figure 3) because the ultrasound modality produces higher image gradient magnitudes for the tissue-tissue interfaces with normals parallel to ultrasound beams. If the tissue-tissue interface normal is perpendicular to the ultrasound beams, then only minimal image gradient magnitudes are produced for those regions. In order to capture this modality related information in our local priors, we divided the polar images into θ ranges. The local priors in each θ range are trained separately so that different image characteristics for each range can be extracted.

Including temporal information in echocardiographic contour extraction process is known to produce better segmentation results. As a result, there are many techniques that model and incorporate temporal information as prior knowledge in the literature [7]. Without complicating the overall system, we choose to use 3D Haar-like filters to learn about the intensity motion around the cardiac borders. This solution allows us to use 2D global priors with the added benefit of 3D temporal intensity motion information through the local priors. A similar technique is used

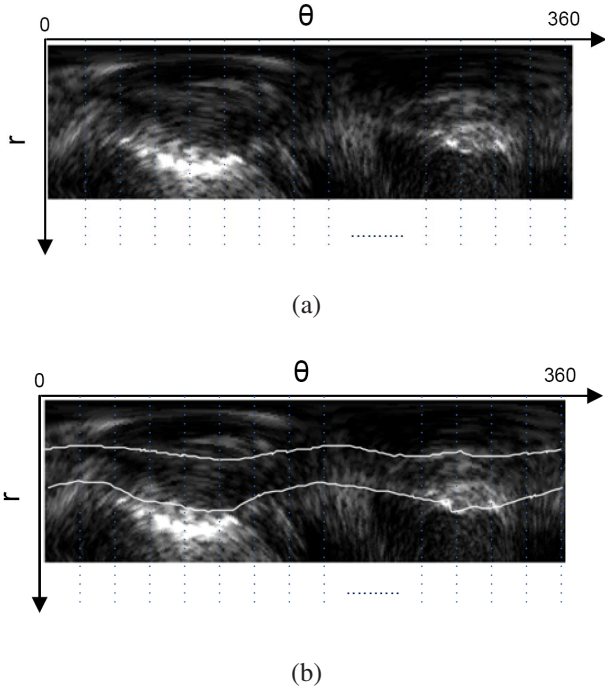


Figure 3. (a) Echocardiographic image converted to polar coordinates; (b) Same image with expert detected contour in polar coordinates.

by Viola and Jones [10] that integrates intensity with motion information for detecting pedestrians by using differences between pairs of sequential images in time. In our system, we used at most 5 sequential frames of cardiac cycles around end-systole and end-diastole frames. The method includes not only the data information from the end-systole/end-diastole frame, but also the temporal/motion information from the sequential frames and this makes the segmentation results more robust.

Rectangular 2D and 3D spatiotemporal Haar-like filters F_k^n are used for feature extraction, where $n = 1, 3, 5$ shows the number of frames used and $k = 1 \dots 12$ defines the type of the filter (Figure 4). The filters in Figure 4(a) extract the image and temporal information directly in sequential frames. These filters are applied to 1, 3 and 5 sequential frames around end-systole and end-diastole. In order to capture the cardiac wall motion, the filters in Figure 4(b) are used for registering the frames by shifting the filters for the preceding and succeeding frames 1 pixel up for end-systole and 1 pixel down for end-diastole.

We use various versions of these filters with different spatial sizes, gray level inverses, and aspect ratios. The filters describe the image properties well and they are efficiently calculated by using the integral images[3]. The boosting algorithm automatically selects the filters that classify the data better and ineffective filters are eliminated.

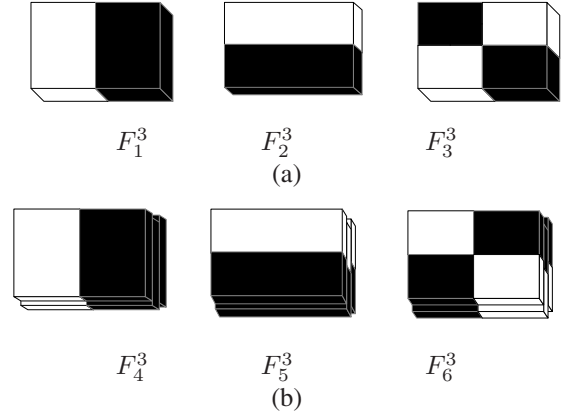


Figure 4. Example rectangular filters used for feature extraction. The total intensity value of darker area is subtracted from the total intensity value of the lighter area. Also gray level inverse of these filters are used and totally there are 12 types of filters. (a) Three types of filters for extracting image features; (b) The other three types of filters created by shifting the neighbor frames by 1 pixel to capture the cardiac wall motion.

In order to train the boosting based local prior model, we choose a number of features f_m , where m is the feature type. Each feature is calculated for each possible θ range, border type B (epicardium or endocardium), and cardiac phase Ph (end-systole or end-diastole).

The first feature type used is the distance of the wall point (p) from the center of the heart.

$$f_1(\theta, B, Ph) = r(p), \quad (5)$$

where r returns the R position of point p in polar space.

The other feature types are all image related.

$$f_l(\Theta, B, Ph) = F_k^n, \quad (6)$$

where $k = 1 \dots 12$, and $n = 1, 3, 5$. Throughout the experiments, we used 36 θ ranges each of which covers 10 degrees.

2.4. Training and Classification through Boosting

The main idea of boosting is to use a number of weak classifiers $h_t^{\theta, B, Ph}(x)$ to obtain a strong classifier $H^{\theta, B, Ph}(x)$ for each possible (θ, B, Ph) value. We used Adaboost, which is one of the most popular boosting methods, to train and classify the features that contains local priors extracted from echocardiograms. Different from the classical AdaBoost, instead of using the discrete class value $sign(H(x))$, we employed $H(x)$ directly in our local prior estimations.

Let the pixels in the echocardiograms be the feature points $\{x_i, c_i\}$ where $c_i = \{1, -1\}$ are the classes that

show cardiac border and non-border parts. Adaboost uses weak simple classifiers $h_t^{\theta,B,Ph}(x)$ and assigns weights α_t inversely proportional to the accuracy of classification. The final strong classifier is built by

$$H^{\theta,B,Ph}(x) = \sum_{t=1}^T \alpha_t h_t^{\theta,B,Ph}(x), \quad (7)$$

where T is the total number of weak classifiers. Note that each $h_t^{\theta,B,Ph}$ corresponds to the feature set f_l defined in Equations 5 and 6.

For training, the polar transformed echocardiographic image is divided into 36 θ ranges. The features are extracted as explained in the previous section and for each (θ, B, Ph) tuple, we trained a separate classifier $H^{\theta,B,Ph}$. In other words, there are 144 different $H^{\theta,B,Ph}(x)$ classifiers in our system. For a given θ range, only the image pixels in that range of the image are used. Similarly, for a given cardiac phase Ph , only the pixels from the frame of that phase are used in the training. A given pixel x of the polar image in angle range θ and phase Ph is used in the training of $H^{\theta,epi,Ph}$ and $H^{\theta,endo,Ph}$. The expert contour positions are used as positive examples and the other positions in the same angle range and cardiac phase are used as negative examples for the training.

During the run time, the images are transformed to the polar coordinates and features are extracted. The classification value of each pixel, which is in the range $[-1, 1]$, is obtained by using both $H^{\theta,epi,Ph}$ and $H^{\theta,endo,Ph}$. Values closer to 1 indicate that the pixel is likely a cardiac border pixel. We formed two new images using these values and these images are converted back to the rectangular representation. The image produced by the classification values of $H^{\theta,epi,Ph}$ is called I^{epi} and the image produced by the classification values of $H^{\theta,endo,Ph}$ is called I^{endo} . These two images are used by the external term of the level set optimization defined by Equation 3.

3. Experimental Results

The proposed method is tested on echocardiographic images of the left-ventricular (LV) short-axis transthoracic views during the cardiac cycle. The data set includes a sequence of frames during one cardiac cycle including an end-systole and end-diastole image. The endocardium and epicardium of the images are delineated by 4 different experts and each wall is represented by 100 landmark points.

For the training set, we used 20 frames for systole delineated by 4 experts and 20 sequential frames before/after systole images. The systole and diastole sequences are trained and tested separately and the number of samples used for the diastole are the same as systole. The total number of features extracted for each pixel is 1022. We used 8 expert

detected endocardium and epicardium contours for incorporating global shape prior.

The system is tested on 12 end-systole and end-diastole images. The experts' delineations are used for evaluating our segmentation. The delineations of experts are also compared with each other to obtain the variation between experts. The average pixel differences are calculated between expert-expert and expert-automatic contours. The differences are calculated for 12 test cases and we found that the differences between our system and the experts are very similar to expert-expert differences. Visual inspection of the results also verifies this finding. It is promising that our system is within the inter-expert variation numbers.

4. Conclusions

We introduced a new level set based system for echocardiographic cardiac wall segmentation from short axis images. Like many examples in the literature, we employed prior knowledge to address the common problems of medical image analysis. In order to incorporate the prior information into the segmentation process in a more systematic manner, we defined global and local prior information. The global prior information produces problem constraints that can be imposed in a top down methodology. Although we used expected segmentation shapes as the only global prior information in our system, it is possible to use other types of global priors such as hard constraints on the expected segmentation area, the number of certain features like corners or holes. Our enforcement of global prior information is based on repetitive re-initialization of the level set surface with the desired global features.

The local prior information incorporates local geometric and image based constraints into the system. The prior information about how the heart walls would look on ultrasound images is captured using a boosting based classifier that uses many spatiotemporal filters. Imposing local prior information through boosting makes it possible to include non-image related constraints such as local contour distance to a shape reference point. Another advantage of using boosting is the inclusion of 3D spatiotemporal information without using complicated models.

We applied the above general ideas to echocardiographic heart wall segmentation process with promising results. The advantages of local prior information enforcement made it possible to handle the worst quality images with a good performance.

For future work, we plan to include an image profile element in our local feature set. We will also employ local geometric information in the form of contour curvatures for other medical imaging applications.

| | Expert 2 | Expert 3 | Expert 4 | Auto | Expert 2 | Expert 3 | Expert 4 | Auto |
|----------|----------|----------|----------|------|----------|----------|----------|------|
| Expert 1 | 3.68 | 3.47 | 3.92 | 2,87 | 3.45 | 3.61 | 8.36 | 5.39 |
| Expert 2 | | 3.32 | 4.49 | 4.33 | | 3.29 | 8.21 | 5.94 |
| Expert 3 | | | 4.93 | 4.13 | | | 8.01 | 5.05 |
| Expert 4 | | | | 4.70 | | | | 6.23 |

Table 1. First part in the table shows endocardium pixel differences and the second part shows pixel differences for epicardium.

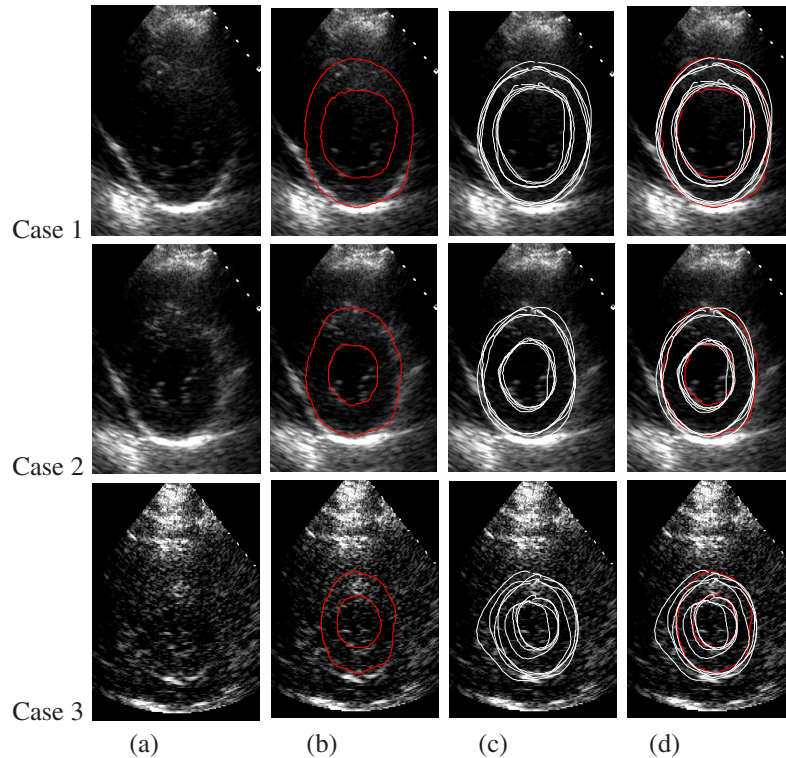


Figure 5. (a) Three echocardiographic images; (b) Automatically extracted contours by our system; (c) The epicardium and endocardium delineated by four different experts; (d) White contours are the expert delineated contours, red contour is automatically detected.

References

- [1] Y. Chen, F. Huang, H. D. Tagare, and M. Rao. A coupled minimization problem for medical image segmentation with priors. *Int. J. Comput. Vision*, 71(3):259–272, 2007.
- [2] Y. Chen, H. D. Tagare, S. Thiruvankadam, F. Huang, D. Wilson, K. S. Gopinath, R. W. Briggs, and E. A. Geiser. Using prior shapes in geometric active contours in a variational framework. *Int. J. Comput. Vision*, 50(3):315–328, 2002.
- [3] F. C. Crow. Summed-area tables for texture mapping. In *Proc. of Conf. Comp. Grap. and Inter. Tech.*, pages 207–212, New York, NY, USA, 1984.
- [4] Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Int. J. of Comput. and Sys. Sciences*, 55(1):119–139, 1997.
- [5] C. Li, C. Xu, C. Gui, and M. D. Fox. Level set evolution without re-initialization: a new variational formulation. In *IEEE Conf. Comput. Vis. Pattern Recognit.*, volume 1, pages 430–436 vol. 1, 2005.
- [6] E. G. M. Leventon, O. Faugeras and W. Wells. Level set based segmentation with intensity and curvature priors. In *Proc. of the IEEE Work. on Math. Methods in Biomed. Imag. Analysis*, page 4, Washington, DC, USA, 2000. IEEE Computer Society.
- [7] J. Noble and D. Boukerroui. Ultrasound image segmentation: A survey. *IEEE Trans. Med. Imaging*, 25(8):987–1010, August 2006.
- [8] A. B. Oktay and Y. S. Akgul. A novel level set based echocardiographic contour extraction method with prior knowledge. In *British Mach. Vis. Conf.*, 2008.
- [9] M. Teles de Figueiredo and J. Leita. Bayesian estimation of ventricular contours in angiographic images. *IEEE Trans. Med. Imag.*, 11(3):416–429, Sep 1992.
- [10] P. Viola, M. Jones, and D. Snow. Detecting pedestrians using patterns of motion and appearance. In *Int. Conf. Comput. Vis.*, pages 734–741, 2003.
- [11] D. Wilson and E. Geiser. Automatic center point determination in two-dimensional short-axis echocardiographic images. *Pattern Recog.*, 25(9):893–900, September 1992.