Machine Vision and Applications
DOI 10.1007/s00138-012-0458-9

SPECIAL ISSUE PAPER

Extraction of left ventricle borders with local and global priors from echocardiograms

Ayse Betul Oktay · Yusuf Sinan Akgul

Received: 31 December 2011 / Revised: 25 June 2012 / Accepted: 17 September 2012

Abstract This paper presents a novel technique for the extraction of the left ventricle borders from echocardiograms with prior information. Although the literature includes many successful prior based methods, priors that include both image and non-image related features are rare for the contour extraction. We classify these features as local and global priors where the local priors refer to the locally definable features of the target borders and global priors refer to the geometric shape properties. The local priors, which include image, motion, and local shape information, are learned with AdaBoost. The scores produced by AdaBoost for the target images are combined with the global shape prior under a level set framework. The main contributions of this paper are to learn different types of local features efficiently with machine learning and to combine these features with the geometric shape information for the contour extraction task. The system is validated on the real echocardiograms and synthetic images. The results indicate that using local and global priors together produces better extraction results and the contours extracted by the proposed system are in accord with the expert delineated borders.

Keywords Contour extraction · Active contours · AdaBoost · Level set · Echocardiogram · Prior

A. B. Oktay (✉)
Department of Computer Engineering, Istanbul Medeniyet University, D100 Karayolu No 6/1, Goztepe, Kadikoy, Istanbul, Turkey
e-mail: aboktay@medeniyet.edu.tr

Y. S. Akgul
GIT Vision Lab, Department of Computer Engineering, Gebze Institute of Technology, Gebze, Kocaeli 41400, Turkey
e-mail: akgul@bilmuh.gyte.edu.tr
URL:http://vision.gyte.edu.tr/

1 Introduction

Extraction of the inner and outer cardiac borders (endocardium and epicardium) from echocardiographic images produces crucial information about the cardiac functions and cardiac morphology like the size and thickness of the cardiac wall, the volume of chambers, the ejection fraction (blood pumping capacity), etc. In the current clinical practice, the cardiac borders are generally delineated manually by sonographers [15]. However, the manual delineation process is user-dependent, time consuming, ineffective, and results in variations between experts (Fig. 1). As a result, automation of the cardiac wall extraction is considered important for increasing the accuracy of measurements and speeding up the cardiac assessment.

Conventional contour extraction techniques (such as [14,18,20]) cannot be directly employed for the automatic extraction of the cardiac borders because of ultrasound modality related problems. The literature partially addresses these problems by employing different regularization strategies through the incorporation of domain related information into the extraction process in the form of shape priors [5,7,26,31]. The overall results improve significantly for the borders with poor image information. However, the shape prior is sometimes not adequate to recover a shape with very low image gradient magnitude. Chen et al. [6] suggest using intensity profiles of object contours besides the shape prior by training an intensity model from a set of images. A similar intensity and curvature profile information is used by Leventon et al. [17]. Bresson et al. [3] propose a variational model that uses a shape model learned with Principle Component Analysis, image gradient, and region information in its energy functional. Although using such priors are very useful for the extraction of object boundaries, it is difficult to use other image and non-image based

Published online: 20 October 2012

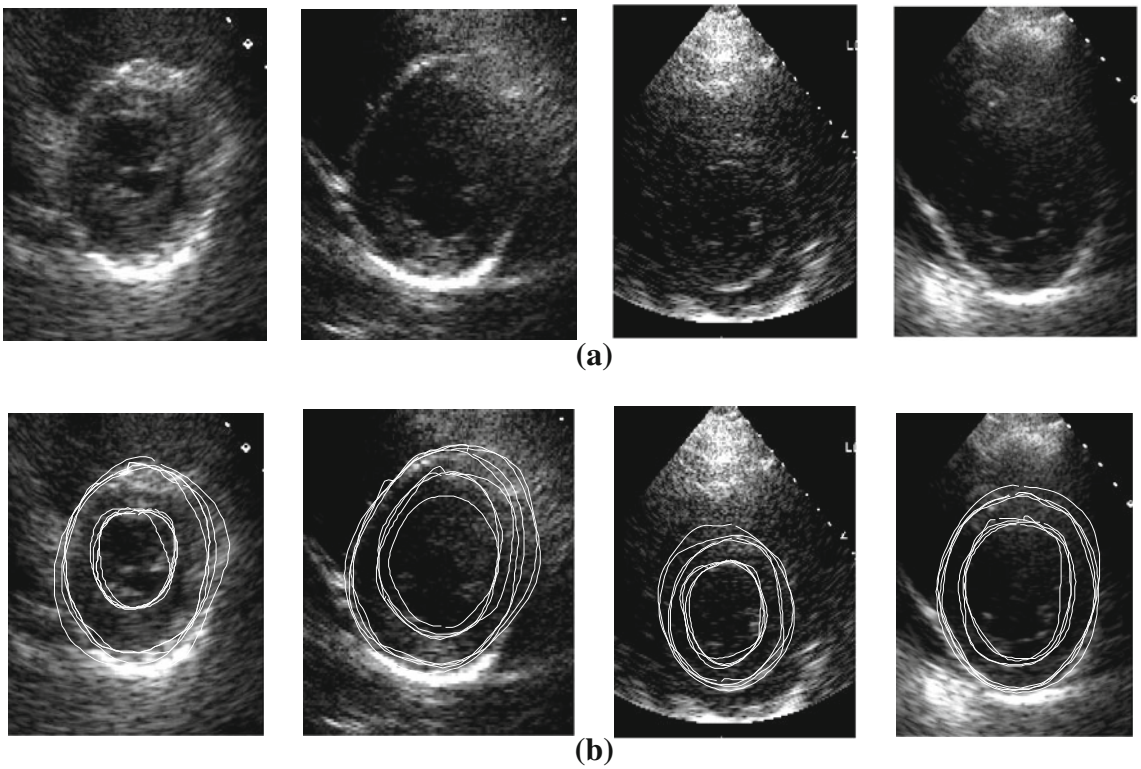

Fig. 1 **a** Four sample echocardiographic images of the left-ventricular short-axis transthoracic views. Note that the left and right parts of the borders have minimal image gradient information due to signal dropout problem. **b** The same images delineated by four different experts

features (e.g. motion and distance to a reference point) together as the prior information with the contour extraction methods.

The machine learning methods learn different types of features for detecting the target structures. AdaBoost [12] is one of the popular machine learning methods used heavily by the medical imaging community for the detection of anatomical structures. Carneiro et al. [4] detect the fetal anatomies and Feng et al. [11] detect fetal faces from ultrasound images through a probabilistic boosting tree. Georgescu et al. [13] use AdaBoost for extracting the endocardial borders of the left ventricle with structure detection. AdaBoost is also used for locating the vessel boundaries [27] and tracking the cardiac borders [28].

In this paper, we present a novel approach for the automatic extraction of the cardiac borders with AdaBoost and active contours. We first detect the candidate cardiac borders by learning local information with AdaBoost and then extract the final borders by incorporating global shape information. We classify the incorporated knowledge as local and global priors because they include different types of information about the target structure. The global prior refers to the classical shape prior [5,7] that incorporates the geometric shape requirements into the extraction process. Local priors include locally definable features of the target borders such as image information, time related properties, local geometric shape, distance from a shape reference point, etc. Since there are different types of local features, it is hard to learn and combine them with classical regularization techniques. In this study, different local features are all trained and learned via AdaBoost easily without using any regularization techniques. In the testing phase, score images that include the local information are produced. We employed the global prior incorporation method of [25] in bringing the local and global priors together. Oktay and Akgul [25] incorporates the global prior by re-initializing the best matching expert contour on the score image at regular intervals. Preliminary work on this study was presented in [23,24]. We extend [23,24] by utilizing the motion information more efficiently and by providing a more complete system validation.

The proposed method is tested on short axis view echocardiograms of end-systole and end-diastole phases to extract endocardium and epicardium. It is hard to evaluate the accuracy of the system and effectiveness of the filters with real echocardiograms due to the lack of proper ground truth. Therefore, experiments on the synthetic images are also performed to compare our method with the methods of the literature and to show the robustness of our method under various controlled noise levels.

The rest of this paper is organized as follows. The local priors and the AdaBoost scoring process are described in Sect. 2. Section 3 includes the method for the incorporation of the global prior. The experiments are presented in Sect. 4. Finally, we provide concluding remarks in Sect. 5.

2 Local priors

Incorporation of local priors is necessary when the target object has different characteristics at different spatiotemporal image locations. We extract these local, spatial, and motion features after converting the echocardiographic images into polar coordinates (Fig. 2). After training the extracted features with AdaBoost, each pixel of the target image is given two scores to be used with global priors. The architecture of AdaBoost scoring system is shown in Fig. 3.

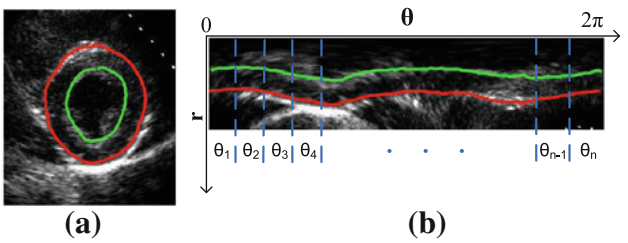

Fig. 2 **a** An echocardiographic image where the epicardium is in *red color* and the endocardium is in *green color*. **b** The same image converted to polar coordinates in (θ, r) space and divided into angle ranges from θ_1 to θ_n (color figure online)

2.1 Extracting local features

We convert the echocardiographic images into polar coordinates [16] using the linear interpolation technique. The left ventricle center is chosen as the center of the polar coordinates. Although the left ventricle center can be estimated automatically as in [30], we use the centers of expert contours during the AdaBoost training and manually extract the centers during the testing. In the echocardiographic images, the left ventricle is usually in the middle of the image and the midpoint of the image is near to the polar center. So, in our system the polar center does not affect the system performance very much. However, for the images in which target organ is not in the center of the image, the performance is dependent on the polar center.

The employment of polar coordinates (Fig. 2) has advantages for the local feature extraction. First, the translational differences between different images are minimized because the left ventricle centers of all images are registered to the origin in the polar coordinates. As a result, the distance from the heart wall to the polar origin becomes a valuable feature that can be used as a local prior. Second, the heart wall orientations are similar (all almost horizontal) in the polar coordinates. This allows us to use a smaller number of filter orientations to capture the contour orientations.

In the echocardiograms, the upper and lower parts of the cardiac wall have higher image gradient magnitudes than the side walls because of the signal dropout. In order to address this modality related problem, we divide the polar images into a number of overlapping angle ranges θ_i where $i = 1 \ldots n$ and n is the number of ranges (Fig. 2b). The features in each θ_i range are extracted and trained separately so that different

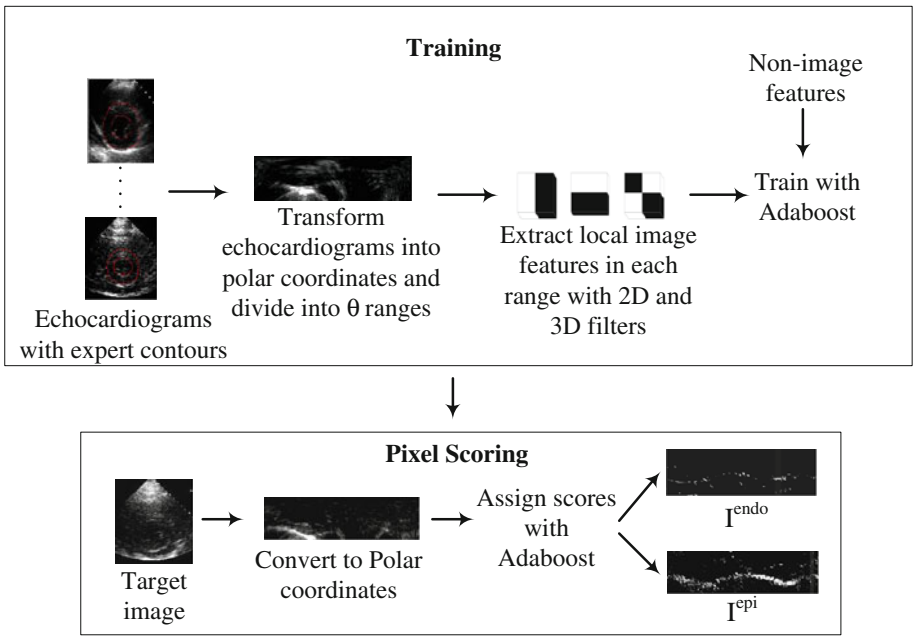

Fig. 3 The local feature extraction and scoring process with AdaBoost

characteristics of the cardiac borders including the signal dropout rates at specific positions can be handled appropriately.

The incorporation of temporal information is crucial for detecting the cardiac borders [1,21]. The sonographers utilize the temporal information in the manual extraction of the cardiac borders by viewing the echocardiograms as movies [15]. We incorporate the temporal information using local features in the form of 3D Haar-like filters [19], hence our system avoids any explicit spatiotemporal deformable models which are complicated and expensive to minimize. This solution allows us to conveniently use our local prior enforcement method in imposing spatiotemporal constraints.

Three types of rectangular Haar-like filter sets D, E, and F are used for the image feature extraction (Fig. 4). For each filter type, the feature value is calculated by subtracting the total intensity value of the darker area from the total intensity value of the lighter area [29].

The 2D filter set D extracts the spatial appearance information. The 3D spatiotemporal filter set E, which is the stacked version of type D, is used to extract motion and appearance information in the time dimension. Similarly, the filters in set F are stacked versions of the type D filters, but while stacking these filters, each filter is shifted depending on the anticipated heart wall motion. The F filters are shifted up (down) for the neighboring frames of the end-systole (end-diastole). For the cardiac images in polar (θ, r) space, the border positions have the smallest r value in the end-systole and the largest in the end-diastole phase. The borders of the neighboring frames of the end-systole phase have r values which are typically 1 to 3 pixels greater than the r value of the end-systole. In order to register the contour motion of the end-systole, the F filters shift up for each neighboring frame 1, 2, or 3 pixels. Similarly for the end-diastole, the F filters shift down for each neighboring frame 1, 2, or 3 pixels. Since we do not know the exact amount of border motion between the frames, we try different shift amounts in order to capture the cardiac motion. These filters describe motion information better and capture motion patterns effectively.

The spatiotemporal filter sets E and F are very useful in addressing the ultrasound noise since they use the motion information of neighboring frames. We use various versions of the filters with different sizes (from the size of $2 \times 2 \times 1$ to $20 \times 20 \times 5$), gray level inverses, and rectangular shapes. The filters are efficiently calculated by using the integral images [10,29]. The calculation of E and F filters with integral volumes are the same as [19].

In order to train the local prior model, we extract a number of features f_m where m is the feature type. The features are calculated separately for each possible (θ-B) pair where θ is the angle range and $B \in \{\text{epi}, \text{endo}\}$ is the border type (epicardium or endocardium).

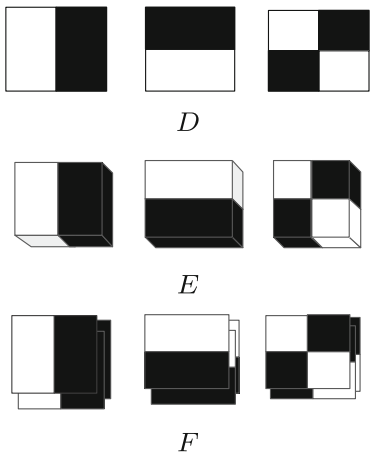

Fig. 4 Three sets of filters used for feature extraction. The set D contains only spatial filters whereas the sets E and F contain spatiotemporal filters

The first feature type $f_1^{(\theta,B)}(p)$ is the distance of the pixel p from the center of the left ventricle:

$$f_1^{(\theta,B)}(p) = R(p), \tag{1}$$

where $R(p)$ returns the r position of pixel p in polar space.

The other feature types are all image related:

$$f_k^{(\theta,B)}(p) \in \{D(p)\} \cup \{E(p)\} \cup \{F_i(p)\}, \tag{2}$$

where $\{D(p)\}$ and $\{E(p)\}$ are the sets of all D and E filters applied to pixel p, respectively. F filters are different from the sets D and E in that they have to specify shift amounts for the neighboring frames. Since we do not know the amount of shift in the cardiac movies, for a given F filter and pixel p, we measure the responses of all possible shift values and choose $F_i(p)$ that gives the maximum value. The classical AdaBoost assigns a weight to each feature and sums them linearly which is used in our handling of D and E filter sets. However in our system, not all F filters are assigned weights. Instead, we only assign weights for a special subset of F filters (maximum $F_i(p)$) which makes our approach non-linear. This is a novel method that we use to effectively address the unknown contour movement values of borders between the cardiac frames.

2.2 Training and scoring process

Consider a pair (p, c) for a pixel p and class $c \in \{wall, non-wall\}$ as the classification label of p. We define p_f as the set of features extracted with Eqs. 1 and 2 for the pixel p. AdaBoost uses weak simple classifiers $h^{\theta,B}(p_f)$ and assigns weights α inversely proportional to the accuracy of the classification. Different from the classical AdaBoost, instead of using the discrete class value $\text{sign}(H(p))$, we employ $H(p)$ directly as the score value in our local prior estimations. The final strong score value is built by

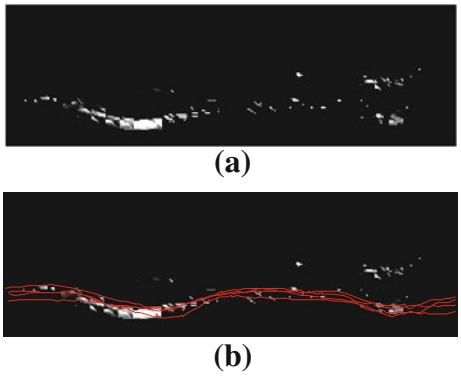

Fig. 5 **a** The result of the pixel scoring process for the epicardium of the echocardiogram of Fig. 2 which is built by combining the score values of each angle range. **b** Same score image with 4 expert detected contours

$$H^{\theta,B}(p) = \sum_{t=1}^{T} \alpha_t h_t^{\theta,B}(p_f), \tag{3}$$

where T is the total number of weak classifiers.

For training, the polar converted echocardiographic images are divided into $n = 36$ θ ranges where $n = 36$ is selected empirically. The expert contour positions are used as positive examples and 10 % of the other positions in the same angle range are used as negative examples. We train and test only the end-systole and the end-diastole frames in the cardiac cycle for which the expert delineations are available. Each (θ, B) pair is trained separately because they include different types of local knowledge, i.e., for each (θ, B) tuple we train a separate classifier $H^{\theta,B}$.

For the scoring process, each pixel of the angle range θ is assigned two score values with $H^{\theta,\text{epi}}$ and $H^{\theta,\text{endo}}$ by Eq. 3. These two scores are used to form two polar images I^{epi} and I^{endo}. An example I^{epi} image is shown in Fig. 5. These two images are converted back to the rectangular representation which are linearly normalized to the [0–255] range and used by the active contour energy functional that will be defined in Sect. 3.

3 Global priors with level set method

Although the local priors incorporate valuable information into the contour extraction process, we still need to impose more global constraints such as geometric shape. Such information would guide the contours towards better positions where there is no sufficient local data. For example, the sections of the score image in Fig. 5, which do not carry any local information, can only be extracted with global information.

The score images I^{epi} and I^{endo} are formed only with the local information from the target images. We use the prior shape incorporation method of [25] which is based on incorporating the shape prior during re-initialization of the level set surface. Note that, at this step our goal is to bring the local and global priors together rather than proposing a new global prior incorporation method. In addition, it is possible to use other contour extraction methods like active shape models [9] and active appearance models [8] with the score images.

The global prior incorporation method is based on re-initializing the level set surface formed with the most similar expert contour [22,25]. We prefer to use this technique because it does not require any training and it is fast.

For the cardiac wall extraction, let $c_1(t)$ and $c_2(t)$ be two closed curves evolving on the plane $\Re^2$ with time t for extracting the endocardium and epicardium, respectively.

Consider C as the set of points on $c_1(t)$ and $c_2(t)$. Let $\mathbf{x} \in \Re^2$ be the position vector and ϕ be a signed distance function defined as:

$$\phi(\mathbf{x}) = \begin{cases} 0, & \text{if } \mathbf{x} \in C, \\ -d(\mathbf{x}), & \text{if } \mathbf{x} \text{ is outside } c_1 \text{ but inside } c_2, \\ d(\mathbf{x}), & \text{otherwise,} \end{cases} \tag{4}$$

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

I^{endo} and I^{epi} are the normalized rectangular images that include the score of each pixel assigned by AdaBoost. The surface ϕ evolves on image I which is defined as

$$I(\mathbf{x}) = \begin{cases} \frac{1}{1+I^{\text{endo}}(x)}, & \text{if } \mathbf{x} \text{ is closer to } c_1, \\ \frac{1}{1+I^{\text{epi}}(x)}, & \text{if } \mathbf{x} \text{ is closer to } c_2. \end{cases} \tag{5}$$

The 3D surface ϕ evolves under the influence of internal and external energy terms according to the variational level set formulation [18]. The incorporation of global information is achieved by stopping the surface deformations and re-initializing the surface under the influence of the shape prior at regular intervals. During the re-initialization, the expert contours are warped onto the evolving contour and the best matching contour is found. The best matching contours for the endocardium and epicardium are found separately. The surface that contains the global shape information is constructed by embedding the best matching warped expert contours (endocardium and epicardium) into the zero level of the surface with Eq. 4. The constructed surface is re-initialized on the image $I(\mathbf{x})$ and thus the global shape information is incorporated by the re-initialization. [25] explains the details of the global shape incorporation method.

4 Experimental results

The system is tested and validated using real echocardiographic and synthetic images. Experiments on the echocardiograms are performed to show the effectiveness of the method and its practical applicability. Experiments on synthetic images are performed to demonstrate the robustness of the local priors under varying controlled levels of noise.

4.1 Echocardiographic images

The proposed method is validated on the echocardiographic images of the left-ventricular short-axis transthoracic views. The data set includes sequences of frames during the cardiac cycle of 20 different people. Each cardiac cycle includes one end-systole frame and one end-diastole frame whose endocardial and epicardial borders are delineated by 4 different experts (Fig. 1). The frames between the end-systole and end-diastole phases exist in the dataset. However, they are not annotated by the experts in medical practice. There are 20 end-systole and 20 end-diastole frames for 20 different subjects. Each end-systole and end-diastole frame is annotated by 4 different experts which makes a total of 160 annotations for 40 images. Each endocardial and epicardial contour is represented by 100 points.

We performed a non-overlapping subset of leave-five-out cross validation experiment for 20 systole and 20 diastole frames separately. In each cross validation sub-experiment, we used 15 cardiac images (each delineated by 4 experts) for training and the remaining 5 cardiac images for testing. The total number of features extracted for each pixel is 1081. We used the expert detected endocardium and epicardium contours in the training set as the global shape prior.

We do not have a proper ground truth data available for the validation of the system. Therefore, the delineations of 4 different experts are used for the evaluation of our extraction results. The expert delineations are also compared with each other to obtain the variations between the experts. The average pixel differences are calculated between expert-expert and expert-system contours. Given two contours to be compared, $C_x = \{x_1, \ldots, x_n\}$ and $C_y = \{y_1, \ldots, y_m\}$, the Chamfer distance $E(C_x, C_y)$ [2] is calculated by

$$E(C_x, C_y) = \frac{\sum_{i=1}^{n} \mathrm{Dist}(x_i, C_y) + \sum_{j=1}^{m} \mathrm{Dist}(y_j, C_x)}{n+m}, \tag{6}$$

where $\mathrm{Dist}(z, C)$ is the minimum Euclidean distance between the point z and the contour C.

We run our method on 40 echocardiograms in the dataset (20 systole and 20 diastole images). The average of expert-expert differences (mean $\pm$ standard deviation) are shown in Table 1. The smallest average pixel difference between the experts is 3.10 pixels (Expert 1 and Expert 3) and the largest is 6.72 pixels (Expert 1 and Expert 4). Our system results are between 3.54 and 5.10 pixels for endocardium and epicardium.

We also run the method of Bresson et al. [3] on both echocardiograms and score images to make comparisons with our system. We choose this method because it is an enhanced combination of successful prior incorporation methods [7] and [17]. The main limitation of the method of [3] is segmenting only one object at a time so we only extract the epicardium with the method of [3]. The results generated with the method of [3] are shown in Table 1. The Chamfer distances on the score images are between 4.68 and 6.46 for the epicardium. We also directly run the method of [3] on the original echocardiograms to extract 40 epicardial borders using the same experimental setup. The results of this step are also listed in Table 1. We immediately observe that the method of [3] performs worse on echocardiograms than on score images because the score images include more relevant local information. Note that, a small pixel difference between two contours only shows the similarity between them. As a result, a smaller pixel number should not always be interpreted as a better result. Instead, the numbers of Table 1 are provided to show the statistical indistinguishability of the contours produced by the system and experts. Three sample images from the dataset and our extraction results are shown in Fig. 6.

4.2 Synthetic images

The proposed contour extraction system uses both local and global priors to extract the cardiac borders. We need to analyze the behavior of the local priors separately, for which we view our pixel scoring process as a classification system. We apply standard classification performance tests, e.g. ROC curves. Note that we could not perform the same experiment with the real images as we do not have the proper ground truth data for them. In addition, this experiment provides a test bed environment that allows us to evaluate the effects of the filter sets D, E, and F.

Several sequences of synthetic images are created with various amounts of noise of type gaussian, salt and pepper, and speckle. The images are 100 by 100 pixels in size and have 50 gray level contrast. The synthetic sequences include a ring shaped object that grows and shrinks like the left ventricle. The sequences with different growing and shrinking speeds are included in the training set for a more realistic experiment.

We used 15 different synthetic movies with different noise levels (Table 2) for the training set. For testing, we scored three different synthetic sequences with different noise levels (Table 2) to evaluate the effectiveness of the scoring process. Figure 7a shows sample frames from the test sequences. Figure 7b shows the score images for the outer contour which corresponds to I^{epi}. For the score images, the brighter intensities represent higher scores (likely border positions) and darker intensities represent lower scores (likely non-border positions). In order to obtain a standard classifier that classifies a given pixel into a border or a non-border pixel, we applied a threshold to the score images (results of Eq. 3). ROC curves (Fig. 8) are formed by changing this threshold from $-\infty$ to ∞. As expected, the least noisy image has a

Table 1 The Chamfer distances (mean ± standard deviation) for 40 test cases of epicardium and endocardium

	Endocardium distances				Epicardium distances					
	Exp 2	Exp 3	Exp 4	Our method on scores	Exp 2	Exp 3	Exp 4	Our method on scores	The method of on [3] scores	The method of [3] on scores
Exp 1	3.46 ± 1.08	3.10 ± 0.86	4.36 ± 1.64	3.54 ± 0.69	3.36 ± 0.89	3.29 ± 0.7	6.72 ± 2.64	4.69 ± 0.92	4.68 ± 1.57	6.66 ± 2.03
Exp 2		3.23 ± 1.06	4.23 ± 1.44	3.99 ± 1.00		3.37 ± 1.05	6.64 ± 2.86	5.02 ± 1.27	5.11 ± 1.60	7.00 ± 1.72
Exp 3			4.17 ± 1.65	4.16 ± 0.72			6.63 ± 2.80	4.52 ± 1.01	4.73 ± 1.41	6.67 ± 1.72
Exp 4				4.59 ± 1.32				5.10 ± 1.64	6.46 ± 2.11	7.01 ± 1.92

Exp is used as the abbreviation for the word expert

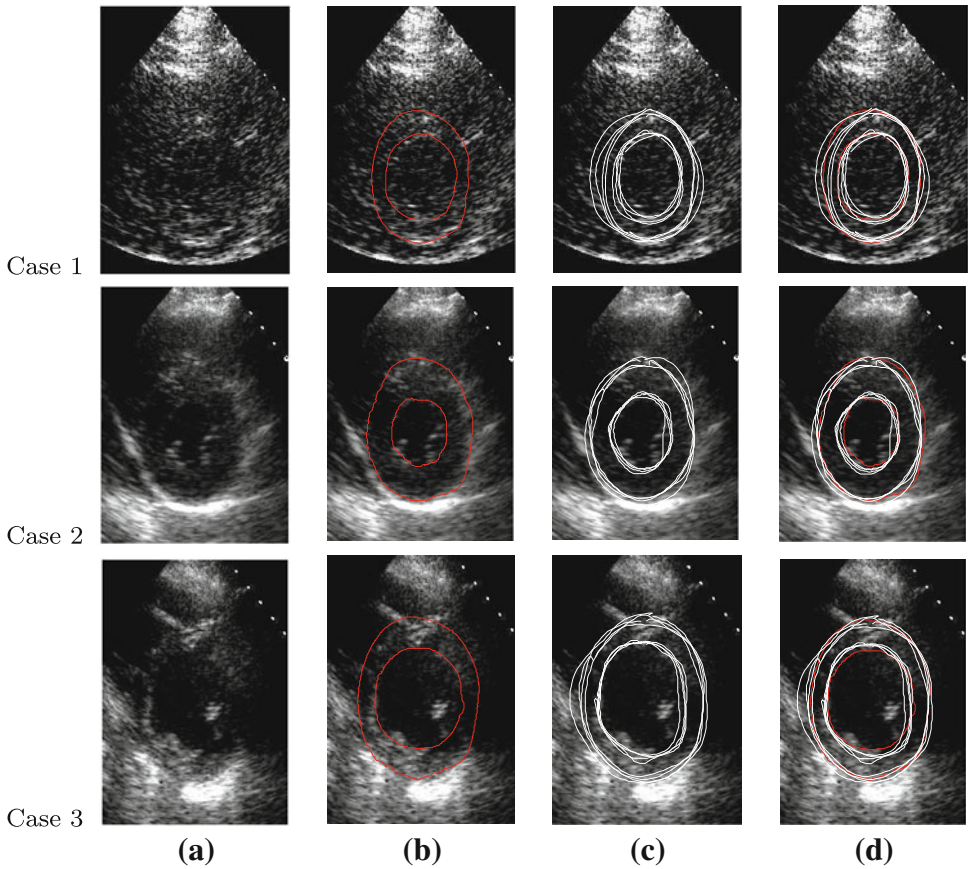

Fig. 6 **a** Three echocardiographic images; **b** epicardium and endocardium contours extracted with our method; **c** contours delineated by four different experts; **d** expert detected and system detected contours

Table 2 Noise levels of synthetic sequences in Fig. 7

Noise type	Training sequences (Noise range)	Test seq. 1	Test seq. 2	Test seq. 3
Gaussian (σ^2)	0.005–0.25	0.04	0.17	0.2
Salt and pepper (%)	0.5–25	4	17	20
Speckle (σ^2)	0.005–0.25	0.04	0.17	0.2

higher classification rate and the noisiest image has a lower classification rate. The analysis of the score images and ROC curves shows that borders of the objects are detectable even under high amounts of noise.

After evaluating the score images with ROC curves, we run our method on the score images to produce the final results (Fig. 7c, Table 3A). In order to see the contribution of our global prior to the end result, we run a popular level

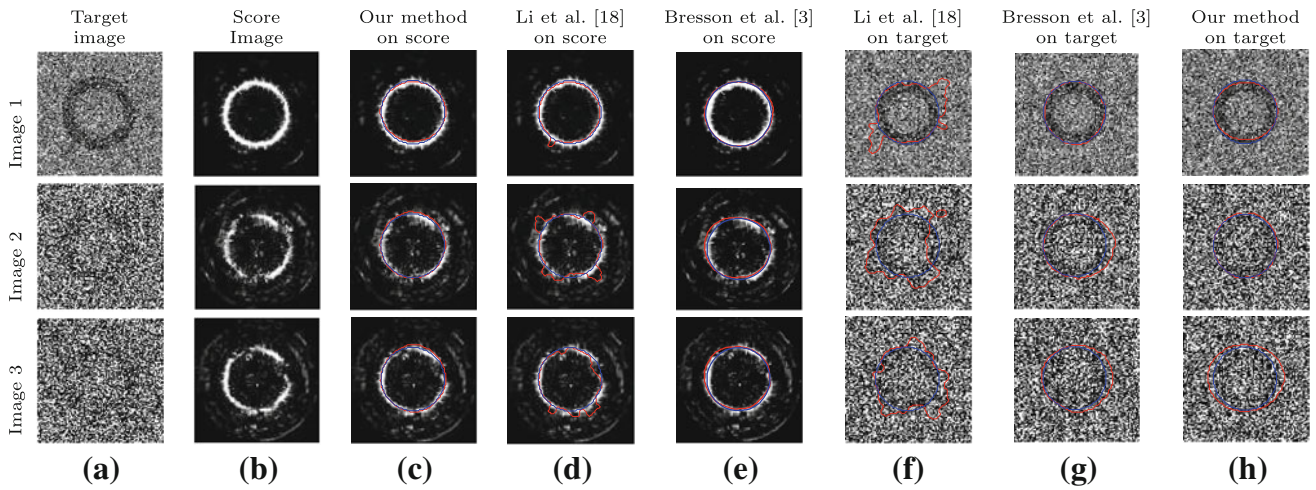

Fig. 7 **a** Three target images corrupted with the noise levels in Table 2 and **b** their score images. **c** The extraction results of our method, **d** those of the method of Li et al. [18], and **e** those of the method of Bresson et al. [3] on score images. **f** The extraction results of the method of Li et al. [18], **g** those of the method of Bresson et al. [3], and **h** those of our method on target images. The *red contours* are the automatically detected contours and the blue contours are the ground truth. The numerical differences are presented in Table 3 (Best viewed electronically) (color figure online)

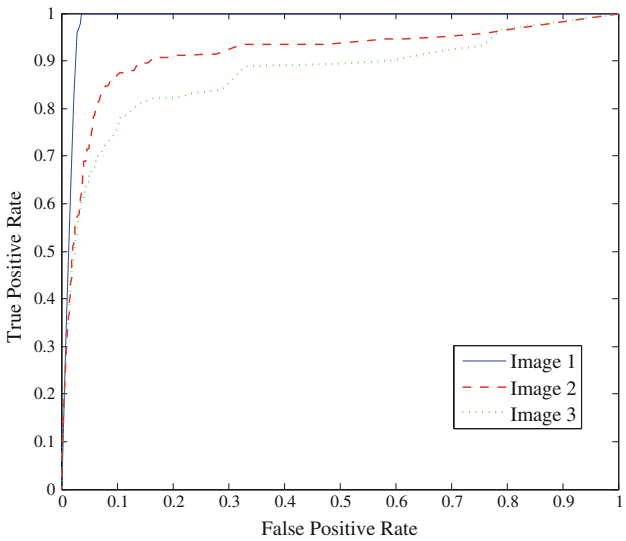

Fig. 8 ROC curves of the synthetic images in Fig. 7a created by applying a threshold from $-\infty$ to ∞

set method [18] using the score images in its external energy without any global prior enforcement (Fig. 7d, Table 3B). Since the local priors cannot extract all the contour information from noisy images, the level set method without global information cannot locate ring borders as good as our method.

We also run the method of Bresson et al. [3] on score images (Fig. 7e, Table 3C). For creating the shape model, we used 10 different ellipses with different orientations and aspect ratios. Next, we run the level set method of [18] without any modifications on the original synthetic images (Fig. 7f, Table 3D). Then, we run method of Bresson et al. [3] (Fig. 8g, Table 3E) on synthetic images. Finally, we run our method that employs only global information on the original target images (Fig. 7h, Table 3F).

The visual and numerical comparisons with similar existing methods indicate that our employment of both global and local priors produce successful results. We argue that the performance enhancement is caused by two main factors. First, our method employs both local and global priors under the same framework, which is an important advantage. Second, our local priors incorporate time related and non-image related information into the extraction process without leaving the 2D deformable contour framework. This advantage brings a considerable amount of relevant information into the contour extraction process without complicating the overall system.

It is interesting to note that we cannot conclude what type of prior is more important by comparing Table 3 columns A, B, and F. Our intuition says that local and global priors become important depending on the available image information at hand and hence we cannot generally say that one of them is more reliable than the other.

The experiments above are performed for the outer contour because the method of Bresson et al. [3] does not work with concentric contours. We also show the contour extraction results of our method for double concentric contours (Fig. 9). Figure 7g shows the pixel differences for inner and outer contours. The visual and numerical results state that our method successfully extracts the inner and outer contours simultaneously.

We also performed another experiment in which the features extracted with the filter sets D, E, and F are trained and scored separately. Figure 10 shows the ROC curves indicating the performance of each filter for the images in Fig. 7a.

Table 3 The quantitative pixel differences of extracted contours and ground truth in Fig. 7

Image	A	B	C	D	E	F	G-double concentric contours	
	Our method on score	Li et al. [18] on score	Bresson et al. [3] on score	Li et al. [18] on target	Bresson et al. [3] on target	Our method on target	Our method Outer contour	Our method Inner Contour
1	0.8	1.03	0.72	2.81	0.87	0.98	0.87	0.81
2	0.95	1.74	0.84	4.22	1.66	1.23	0.89	0.88
3	0.86	1.65	0.85	4.03	1.48	1.73	0.98	0.89

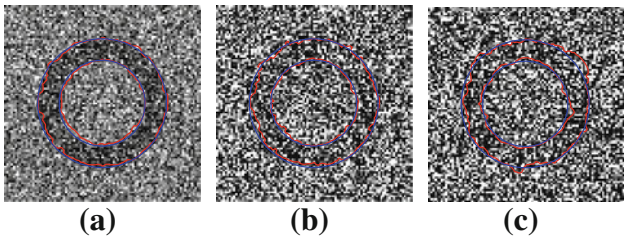

Fig. 9 The extraction results of our method with double concentric contours on **a** Test sequence 1, **b** Test sequence 2, and **c** Test sequence 3 (Best viewed electronically)

The performance of the features learned with 3D spatiotemporal filters E and F are higher than the performance of the spatial filter set D for the noisy images. This experiment shows that more accurate scoring results are produced for the noisy images by using spatiotemporal filters that use more than one frame.

The conventional level set method uses the image gradient magnitudes in its external energy. In contrast, our system uses the score images in the external energy of the level set function. In order to compare our external energy with the conventional level set external energy, we formed a classifier by applying a threshold to the image gradient magnitudes. The pixels with magnitudes higher than the threshold value are considered as border pixels. An ROC curve for the gradient magnitude based classifier is formed by changing this threshold value from 0 to ∞. We manually tuned the Gaussian smoothing values of the gradient calculations for a good classification. Figure 10 compares the obtained ROC curves with our system results. The analysis of this experiment indicates that our scoring process carries considerably more border related information into the contour extraction process.

5 Conclusions

In this paper, we proposed a novel system for extracting the cardiac borders from echocardiographic images by imposing prior knowledge. In order to incorporate the prior information into the extraction process in a more systematic manner, we handle the local and global priors separately. The global prior includes the constraints that are related to the whole image and object shape. Although we used the shape as the only global prior information in our system, it is possible to use other types of global priors such as constraints on the expected region area, length of the contours, and the number of certain features like corners or holes.

The local image features and the motion information, which are mostly extracted with the Haar-like filters, are trained and scored with AdaBoost. Some other non-image related local features such as wall distance to a shape reference point are also included in our system as local priors.

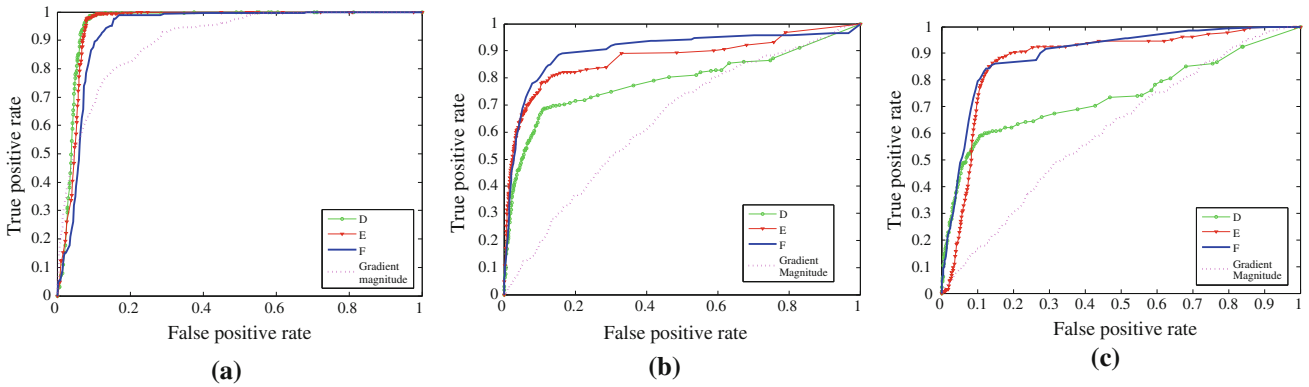

Fig. 10 The ROC curves of each filter type learned individually and the curves when gradient threshold process is applied to **a** Image 1, **b** Image 2, **c** Image 3 in Fig. 7 (Best viewed electronically)

By combining these local features under a single machine learning method, our system can capture motion, image, and other types of local information without employing complicated models.

Our method has several advantages. First, the motion related temporal heart wall information is incorporated easily into the extraction process with 3D spatiotemporal Haar-like filters without using an explicit 3D contour model. This is an important advantage in echocardiography, because some parts of the cardiac borders can be only located with motion information. Second, the division of echocardiographic images into angle ranges provides an angle specific learning mechanism. This advantage largely addresses the signal dropout problem of the ultrasound modality. Third, our local prior can use any features from the images including different filters or intensity profiles. Moreover, our local prior can include non-image related local features such as the geometric relationships between the neighboring border points. Finally, the separation of global and local priors results in a more systematic and modular approach that can integrate different types of knowledge into the extraction process.

The experiments on synthetic and real echocardiograms indicated that the system successfully extracts the borders and the results are not distinguishable from the expert contours. Our system can be easily extended for other types of medical image contour extraction tasks by identifying relevant global and local features.

Acknowledgments The authors would like to thank University of Washington's Echocardiographic Laboratory for the echocardiographic images. Ayse Betul Oktay completed this work while she was in the GIT Vision Laboratory.

References

1. Akgul, Y.S., Kambhamettu, C.: A coarse-to-fine deformable contour optimization framework. IEEE Trans. Pattern Anal. Mach. Intell. **25**, 174–186 (2003)
2. Barrow, H.G., Tenenbaum, J.M., Bolles, R.C., Wolf, H.C.: Parametric correspondence and chamfer matching: two new techniques for image matching. In: Proceedings of 5th International Conference on Artificial Intelligence, pp. 659–663. Morgan Kaufmann, San Francisco (1977)
3. Bresson, X., Vandergheynst, P., Thiran, J.P.: A variational model for object segmentation using boundary information and shape prior driven by the Mumford-Shah functional. Int. J. Comput. Vis. **68**(2), 145–162 (2006)
4. Carneiro, G., Georgescu, B., Good, S., Comaniciu, D.: Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree. IEEE Trans. Med. Imag. **27**(9), 1342–1355 (2008)
5. Chan, T., Zhu, W.: Level set based shape prior segmentation. In: IEEE Conf. Comput. Vis. Pattern Recognit. (2005)
6. Chen, Y., Huang, F., Tagare, H.D., Rao, M.: A coupled minimization problem for medical image segmentation with priors. Int. J. Comput. Vis. **71**(3), 259–272 (2007)
7. Chen, Y., Tagare, H.D., Thiruvenkadam, S., Huang, F., Wilson, D., Gopinath, K.S., Briggs, R.W., Geiser, E.A.: Using prior shapes in geometric active contours in a variational framework. Int. J. Comput. Vis. **50**(3), 315–328 (2002)
8. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. **23**(6), 681–685 (2001)
9. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models-their training and application. Comput. Vis. Image Underst. **61**(1), 38–59 (1995)
10. Crow, F.C.: Summed-area tables for texture mapping. In: Proc. of Conf. Comp. Grap. and Inter. Tech., pp. 207–212, New York, NY, USA (1984)
11. Feng, S., Zhou, S., Good, S., Comaniciu, D.: Automatic fetal face detection from ultrasound volumes via learning 3d and 2d information. In: IEEE Conf. Comput. Vis. Pattern Recognit., pp. 2488–2495 (2009)
12. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. Int. J. Comput. Syst. Sci. **55**(1), 119–139 (1997)
13. Georgescu, B., Zhou, X., Comaniciu, D., Gupta, A.: Database-guided segmentation of anatomical structures with complex appearance. In: IEEE Conf. Comput. Vis. Pattern Recognit., vol. 2, pp. 429–436 (2005)
14. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: active contour models. Int. J. Comput. Vis. **1**(4), 321–331 (1988)
15. Kumar, R., Wang, F., Beymer, D., Syeda Mahmood, T.: Echocardiogram view classification using edge filtered scale-invariant motion features. In: IEEE Conf. Comput. Vis. Pattern Recognit., pp. 723–730 (2009)
16. Lee, H., Codella, N.C.F., Cham, M.D., Weinsaft, J.W., Wang, Y.: Automatic left ventricle segmentation using iterative thresholding and an active contour model with adaptation on short-axis cardiac MRI. IEEE Trans. Biomed. Eng. **57**(4), 905–913 (2010)
17. Leventon, M., Faugeras, O., Grimson, E., Wells, W.: Level set based segmentation with intensity and curvature priors. In: Proceedings of the IEEE Work. on Math. Methods in Biomed. Imag. Analysis (2000)
18. Li, C., Xu, C., Gui, C., Fox, M.D.: Level set evolution without re-initialization: a new variational formulation.In: IEEE Conf. Comput. Vis. Pattern Recognit., vol. 1, pp. 430–436 (2005)
19. Liu, Y., Chen, X., Yao, H., Cui, X., Liu, C., Gao, W.: Contour-motion feature (cmf): a spacetime approach for robust pedestrian detection. Pattern Recogn. Lett. **30**, 148–156 (2009)
20. Malladi, R., Sethian, J.A., Vemuri, B.C.: Shape modeling with front propagation: a level set approach. IEEE Trans. Pattern Anal. Mach. Intell. **17**(2), 158–175 (1995)
21. Noble, J., Boukerroui, D.: Ultrasound image segmentation: a survey. IEEE Trans. Med. Imag. **25**(8), 987–1010 (2006)
22. Oktay, A.B., Akgul, Y.S.: A novel level set based echocardiographic contour extraction method with prior knowledge. In: British Mach. Vis. Conf., vol. 2, pp. 975–84 (2008)
23. Oktay, A.B., Akgul, Y.S.: Echocardiographic contour extraction with local and global priors through boosting and level sets. In: Proceedings of the IEEE CVPR Work. on Math. Methods in Biomed. Imag. Anal., pp. 46–51 (2009)
24. Oktay, A.B., Akgul, Y.S.: Locating cardiac walls with adaboost and level sets (in turkish). In: IEEE 18. Signal Processing and Communications Applications Conference, pp. 503–506. Turkey (2010)
25. Oktay, A.B., Akgul, Y.S.: A modular framework for shape and image prior based anatomical structure contour extraction. J. Med. Biol. Eng. **31**(6) (2011)
26. Paragios, N.: Shape-based segmentation and tracking in cardiac image analysis. IEEE Trans. Med. Imag., pp. 402–407 (2003)
27. Pujol, O., Rosales, M., Radeva, P., Nofrerias-fernandez, E.: Intravascular ultrasound images vessel characterization using

adaboost. Func. Imag. Model. of the Heart: LNCS, pp. 242–251 (2003)
28. Qian, Z., Metaxas, D.N., Axel, L.: Boosting and nonparametric based tracking of tagged mri cardiac boundaries. In: Int. Soc. Conf. Series Med. Image Comput. Comput-Assist. Intervention, pp. 636–644 (2006)
29. Viola, P., Jones, M.J.: Robust real-time face detection. Int. J. Comput. Vis. **57**(2), 137–154 (2004)
30. Wilson, D.C., Geiser, E.A.: Automatic center point determination in two-dimensional short-axis echocardiographic images. Pattern Recogn. **25**(9), 893–900 (1992)
31. Yan, P., Kassim, A.A.: Medical image segmentation using minimal path deformable models with implicit shape priors. IEEE Trans. Inf. Technol. Biomed. **10**(4), 677–684 (2006)

Author Biographies

Ayse Betul Oktay received her BS and MS degrees in computer engineering from Fatih University, Turkey, in 2004 and 2006, respectively. She received her PhD from the Gebze Institute of Technology in 2011. She joined Department of Computer Engineering, Istanbul Medeniyet University as assistant professor in 2011. Her main research interests include medical image analysis, computer vision, and machine learning.

Yusuf Sinan Akgul received the BS degree in computer engineering from Middle East Technical University, Ankara, Turkey in 1992. He received the MS and PhD degrees in computer science from University of Delaware in 1995 and 2000, respectively. He worked for Cognex Corporation, Natick, MA between 2000 and 2005 as a senior vision engineer. He is currently with the Department of Computer Engineering, Gebze Institute of Technology, Turkey. His research interests include application of computer vision in medical image analysis, video analysis, object recognition, 3D analysis, and industrial inspection.