Echo-cardiographic Segmentation: Via Feature-Space Clustering

Vidhyadhari Gondle
Center for Visual Information Technology
International Institute of Information Technology
Hyderabad 500032, India
Email: vidhyadhari@research.iiit.ac.in

Jayanthi Sivaswamy
Center for Visual Information Technology
International Institute of Information Technology
Hyderabad 500032, India
Email: jsivaswamy@iiit.ac.in

Abstract—Segmentation in echo-cardiographic images is a difficult task due to the presence of speckle noise, low contrast and blurring. We present a novel method based on clustering performed in the feature space. A noise robust image representation is obtained by computing a local feature descriptor at every pixel location. This descriptor is derived using the Radon-Transform to effectively characterise local image context. Next, an unsupervised clustering is performed in the feature space to segment regions in the image. The performance of the proposed method is tested on both synthetic and real images. A comparison against well established feature descriptors is carried out to demonstrate the strengths and applicability of the proposed representation. Overall, the results indicate promise in the strategy of doing segmentation of noisy data in image descriptor space.

I. Introduction

Echo-cardiogram is the sonogram of the heart. It provides a wealth of information like shape and size of the heart valves, velocity of blood, location and the extent of the damaged tissues. It aids an assessment of the functionality and abnormality of heart. Due to its non-invasive and inexpensive nature it is one of the most primary diagnostics conducted during emergency cases. In order to assist the diagnosis, the regions in the echo-cardiographic images need to be segmented appropriately classifying blood region, tissue region and other abnormalities if present. But due to the physics of acquisition, segmentation of echo-cardiographic images is quite difficult. Some of the main problems are: i) presence of speckle noise which occurs due to back-scattering of ultrasound waves by the particles whose size is less than the wavelength of projected waves [1]; ii) blurring of the spatial information in the direction perpendicular to the ultrasound waves [2] and finally, iii) Discontinuities in contours due to the differences in the orientation of the reflecting structure [2]. Hence, any segmentation algorithm aimed for ultrasound images must be robust to speckle noise and should be able to detect the discontinuous edges in low contrast.

Adapting intensity based techniques which are popular in general vision to ultrasound segmentation requires either de-noising via pre-processing or inclusion of the noise model in the algorithm. Pre-processing can remove extra information present in the image like speckle variations in different regions. Insertion of the noise models into the segmentation algorithms makes the design of the algorithm difficult and complex. A variety of geometrical or statistical approaches such as active contours [3] [4] [5] [6] [7] [8] [9] [10], Bayesian framework [11] [4], level-sets [12] [13] [14], neural network [15] etc have been proposed to address above mentioned challenges. In particular, level-set and active contour based techniques require a robust energy formulation to compensate noise factor, a challenging aspect in ultrasound images. However, high computation and time complexity make them inadequate for real-time application. For example, object tracking in a ultrasound image sequence a potential use case of the image segmentation.

Our objective is to devise a solution that can handle noisy data naturally and produce a good segmentation. Hence, we explore transforming the image into an appropriate feature space and segmenting via clustering in that space. The key issue here would be to identify the right feature space that is robust to noise. An existing attempt at clustering based approach for ultrasound segmentation is the spectral clustering technique [16]. In this, ultrasound images are first *pre-processed* using anisotropic filtering and then segmented using clustering. The proposed spectral clustering is based upon N-cut algorithm where the similarity matrix between the pixels is built using intensity. The eigen vectors obtained induce an embedding of pixels in a low dimensional subspace and then a final partitioning is done based upon N-cut. In summary, to compensate noise factor they employ a pre-processing step and a strong clustering algorithm.

In this work, we argue that a similarity measure which is computed in an image descriptor space can provide better robustness to the noise compared to a measure defined solely on point intensity, as in [16]. We propose a local point descriptor derived from the Radon Transform which is capable of capturing shape and intensity information reliably. To assess the strength of the proposed strategy, we do not handle noise explicitly by employing any pre-processing and use a simple clustering scheme unlike the one used in spectral clustering [16].

II. Method

In image segmentation, the main aim is to be able to classify the pixels of the image into a set of non-overlapping regions.

The pixels which are similar are grouped together and those pixels which are dissimilar are divided into groups. The aspect used for comparison must be chosen meaningfully according to the image on which it is to be applied. In case of ultrasound images we cannot use the intensity to differentiate the regions because of high level of noise.

As mentioned earlier, there are several problems that make ultrasound image segmentation difficult: 1) Speckle noise caused because back-scattering of the particles whose size is smaller than the wavelength of the projected waves. The noise depends upon the number of particles present in that particular tissue. 2)During the acquisition of the image, the reflected waves are captured and then converted into image. All the reflecting structures may not reflect the sound waves in the same direction due to difference in orientation hence the intensity of the waves is not same for similar reflecting structures, hence there will be discontinuities in the edges in the ultrasound images. 3)There is a blurring effect along the direction perpendicular to that of the waves.

In order to segment such images, use of information about the local context of a pixel might be a good alternative. The local context represents the type of speckle noise present in the images and it is also robust to the blurring effect present in the image. We propose to derive a representation for the noisy image by locally transforming every pixel and constructing a feature descriptor for the pixel in this space.The choice of a particular feature descriptor depends upon the requirements like noise robustness, stability with respect to small distortions and invariance to geometric transformations.We choose the Radon transform and derive a descriptor in the Radon space.

The Radon Transform is an integral transform and some successful attempts have been made to derive representations from this transform to capture (largely binary or single grey scale) shape information [17]. However, it has not been applied in noisy grey scale images to derive a general local descriptor. The integral or projection operation underlying the transform should be useful in handling noise.

A. Feature Descriptor

The Radon transform T_{Rf} of a given continuous function f(x,y) is a projection along a line oriented at angle θ. It is given as

$$R^f (\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) \, \delta (x \cos \theta + y \sin \theta - \rho) \, dx \, dy, \tag{1}$$

where $\delta(.)$ is a Dirac delta-function. The range of the arguments are generally $\theta \in [0, \pi]$ and $\rho \in [-\infty, \infty]$. Note that the line is expressed in parametric form: $\rho = x \cos \theta + y \sin \theta$. The angle θ is also called the view angle.

In binary images, the Radon transform gives the length of intersection of the line with the object, which makes it a good representation of shape in binary images. In gray-scale images, it is the line integral of the intensities. Thus, a rise in the intensity results in a high response. This helps it to differentiate between bright regions and dark regions which can be used in segmentation. The radon transform undergoes

a shift in ρ under object translation and a shift in variable θ for object rotation. Consider a feature-vector derived from the Radon transform by integrating the squared values of the Radon transform for every view θ as follows.

$$d(\theta) = \int_{\rho=-\infty}^{\rho=\infty} R^2 (\rho, \theta) \, d\rho \tag{2}$$

The squaring operation boosts the high values of Radon transform. Since an object translation affects only the ρ variable, the feature descriptor $d(\theta)$ is unaffected. Object rotation by some angle θ_0 results in a translation of d by θ_0. Scaling affects only the magnitude of the feature-descriptor. Hence, this descriptor provides a good representation of the shape and context of a pixel.

Next we extend this idea to the discrete case. The line integral becomes a ray sum (projection). Let I[i,j] be the given image. Now consider a patch of size $n \times n$ around a pixel at location p. The Radon transform of this patch, denoted as $R[m, \alpha]$ is an array of size $(\sqrt{2})n \times K$ where K is the number of views. The descriptor d_p is constructed as follows. Next, the smoothed array is projected as follows

$$d_p[\alpha] = R[\alpha] = \sum_{m=1}^{m=(\sqrt{2})n} R^2[m, \alpha] \tag{3}$$

where d_p is a K-long feature vector.

An edge oriented at angle α_0 will yield a descriptor which has a maximum at α_0. Hence the descriptor forms a good representation of the shape or context present around the pixel. Since it is derived by summing the intensities along a particular direction, the descriptor is insensitive to zero mean noise. The descriptor is also resistant to blurring effect present in the image due to the projection operation (along m). Blurring spreads the intensity values but the information present in the actual image and the blurred image are same. Hence, there is no need for pre-processing the image.

In our implementation, each row of $R[m, \alpha]$ was smoothed with a 5-length kernel before deriving the descriptor. The aforementioned descriptor was computed for every pixel in the given image to obtain a representation of the image in a feature space.

B. Clustering

Since every pixel is associated with a feature vector, pixels belonging to similar regions (tissue type in ultrasound image) should cluster in the feature space. Thus, we classify these pixels based upon their corresponding feature vectors in the feature space. We use K-means clustering to achieve this classification. K-means clustering is one of the simplest and efficient ways of clustering when compared to other clustering methods like Agglomerative, Fuzzy C-Means and Mixture of Gaussians. In the present case, the target number of classes is found manually based on domain knowledge. Clustering results in each feature vector being assigned with a label. These labels are mapped back to the pixels of the image to get the segmented image.

III. Other descriptors of interest

Since the proposed scheme is one of clustering in a feature space, it is natural to ask if the choice of descriptor is correct. We consider other popular feature descriptors and wish to compare the segmentation performance across the descriptors. The descriptors that are considered are both gradient based and non-gradient based ones. In the former category are the histogram of oriented gradients used recently in non rigid registration [18] and the popular Daisy feature descriptor [19]. In the latter category is geometric blur [20] which is also good at representing shapes.

A. Geometric Blur

Geometric blur [20] is defined as the sum of blurred images in the polar coordinate space, where the blurring is done with a set of bounded transforms. Consequently, it is robust to geometric distortions. Details of the descriptor can be found in [20].

B. Histogram of Oriented Gradients

This descriptor is based on the assumption that the context of a pixel can be described by the distribution of gradients or edge directions. The descriptor is formed by computing the gradient orientation of the image and binning them. This descriptor [18] is similar to SIFT in a dense overlapping grid.

C. DAISY feature descriptor

This feature descriptor [19] is inspired from SIFT but is computationally faster. For a given image, to compute the DAISY descriptor, the gradient of the image in 8 different orientation is calculated. These gradient maps in 8 different orientations are then convolved with a Gaussian kernel with standard deviation.

It is said to be robust to noise due to the Gaussian smoothing step.

IV. Experiments

The proposed scheme for segmenting ultrasound images was implemented with the following settings. In the Radon transform computation of image patches, we chose m = 10 and K = 37. The number of clusters was chosen as per the test data. The segmentation performance was compared using our descriptor plus 3 other descriptors listed above. The performance was evaluated both qualitatively and quantitatively on two types of data described below.

A. Synthetic Data

A synthetic image was generated to have a dark circle on a white background with the intensity of the circle varying inversely according to a Gaussian function i.e., least at the centre of the image and it increases as we move away from the centre of the circle. Four such images are generated with decreasing radii. Since our main goal is to test the algorithm on noisy images, the images were corrupted with speckle noise generated using a Rayleigh distribution which is generally used to model speckle in ultrasound images. Different noise statistics was used for the circle and its background to model the ultrasound image. Specifically, the noise level was set to be high for the background to model non-blood regions in an ultrasound and low noise noise level within the circle which models the blood region. Three different noise levels were studied. The segmentation algorithm was used to classify the pixels into 2 classes, namely, pixels belonging to the dark circle or the background.

B. Echo-cardiographic Data

Echocardiographic data is generally acquired in a video form. There are about 41 frames in each video. But the cycle of heart repeats after every 22 images. Hence, we evaluated the experiments only on 22 frames of echocardiographic data. These images were segmented to get blood and tissue regions.

C. Evaluation Measures

The obtained segmentation results were evaluated against human marked boundaries. We use following validation measures to report method's performance:

1) Rate of misclassified pixels: The ratio of no. of misclassified pixels to the total number of pixels in target foreground region.

$$\text{Rate} = \frac{\text{No. of misclassified pixels}}{\text{Total No. of pixels present in the region}}. \quad (4)$$

2) Dice Coefficient: It gives the measure of overlap between two regions. If X is the segmented region and Y is the region present in ground truth then the measure is given as

$$Dice\ Coefficient = \frac{2 \times |\ X \cup Y\ |}{|\ X \cap Y\ |}. \quad (5)$$

V. Results

In our implementation, we use following parameter settings. In the Radon transform computation on image patches, we chose m = 10 and K = 37. The number of clusters chosen varied according to the underlying data. The descriptors other than ours, were computed using binaries provided by the respective authors.

First, we report experiment results for synthetic data. Figure 1 shows the segmentation results obtained on synthetic data corrupted with noise. The noise levels vary columnwise from low (top row) to high (bottom row). The results of clustering in different feature spaces are shown columnwise: From left to right these are for Histogram based (HC), Daisy based (DC), Geometric blur (GC) and FSC (ours) respectively. HC based clustering results are poor. The results of DC does not capture the desired circle boundary, it gives shifted boundaries but in high level noise it over segments in some cases. Comparatively, GC results are quite poor. The FSC results though a bit under-segmented are consistent across noise levels.

Table I shows the average rate of mis-classified pixels and the Dice coefficient obtained with different descriptors. The results of HC have been omitted as they were very poor. From this table, it can be seen that the overall performance of FSC
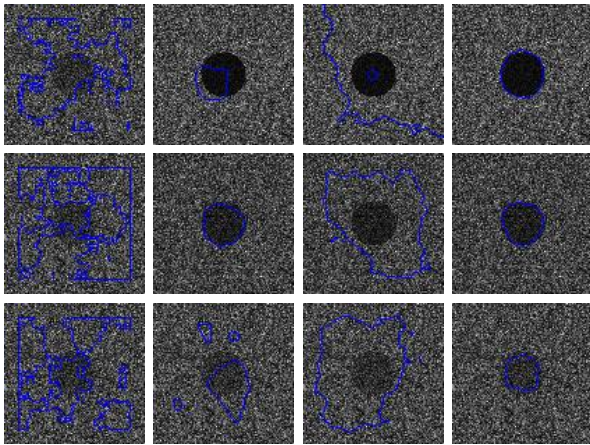
Fig. 1. Results obtained on synthetic images corrupted with speckle noise by a) first column: HC, b) second column: DC, c) third column: GC and d) fourth column: FSC.

Feature	Rate of Misclassification	Dice Coefficient
DC	0.16	0.51
GC	0.62	0.27
FSC	**0.004**	**0.87**

TABLE I
AVERAGE RATE OF MISCLASSIFIED PIXELS OBTAINED BY DIFFERENT DESCRIPTORS

is the best in terms of misclassification and Dice coefficient. In the above evaluations, it can be seen that FSC performance is consistent across different noise levels. This consistency makes it a good choice for echocardiographic segmentation. Next, we report experiments performed on the real echocardiographic data.

The segmentation results obtained on two echocardiographic image frames are shown in fig. 2. The red colour indicates region of interest (blood region) detected by proposed method (second column). In the last column, we show our results as follows: segmented blood region is indicated by white regions and the GT is visible as thin black contour and the non-blood tissue region is indicated by copying the original grey scale image. It can be seen that FSC is able to extract the regions of interest and distinguish it from the background regions quite successfully as over segmented regions are few.

Figure 3 shows the FSC results and results of other descriptors on the first image frame shown in the first row in fig. 2. It can be observed that FSC results are the most accurate. GC results appear to be the second best. However, the non-blood regions (not red) is very poorly handled in that it is inconsistent with the type of tissue characteristics seen in the original image. Inexplicably, DC failed completely.

Next, we quantitatively assess the performance of different methods. Only the GC and FSC have been considered for this purpose as the results of the rest of the methods are unacceptable. Table II shows the obtained values for FSC and GC, respectively. It can be seen that on an average FSC
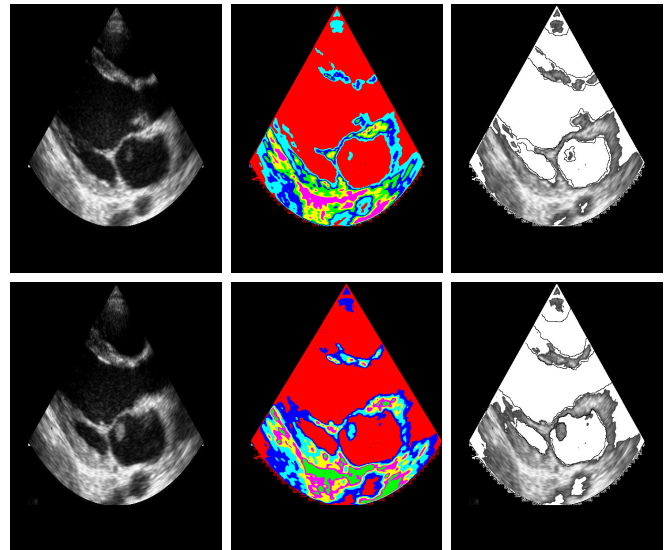

Fig. 2. Two segmentation results on echocardiographic obtained by FSC. Each row shows a) original image b)segmented result and c) segmented result overlaid on ground truth (shown as thin contours).

Image	Rate Of Misclassification		Dice Coefficient	
	FSC	GC	FSC	GC
01	0.0065	0.35	0.95	0.78
05	0.0019	0.43	0.94	0.75
10	0	0.67	0.94	0.70
15	0.0032	0.51	0.89	0.66
20	0	0.48	0.86	0.68
Mean	**0.0023**	0.49	**0.91**	0.72

TABLE II
PERFORMANCE OBTAINED ON ECHOCARDIOGRAPHIC IMAGES

gives low mis-classification rate (accurate segmentation), and high value of dice's coefficient (similarity with the desired region) compared to the GC descriptor. A good performance on real echocardiographic data is mainly due to the robustness (in term of consistency) reported by the proposed descriptor on synthetic data. Hence from the qualitative and quantitative results, FSC shows consistency in all the evaluation measures.

VI. CONCLUSIONS

In this paper, a new segmentation algorithm has been presented for echo-cardiographic images based upon clustering of feature-space representation of the image. A feature descriptor is proposed for this purpose and computed for a patch around every pixel. This is derived from the Radon Transform. Test results from synthetic data demonstrate that the feature space clustering is most effective with the proposed descriptor as it provides robustness to different noise levels. This performance was also found to hold for real images results. The obtained results on ultrasound images were assessed by a medical expert who found them to be qualitatively good. Any segmentation errors made by the algorithm were found to be more due to lack of inclusion of domain knowledge which a medical expert employs while marking. Example of this is in the top part of the image corresponding to myocardial muscle contour. Future
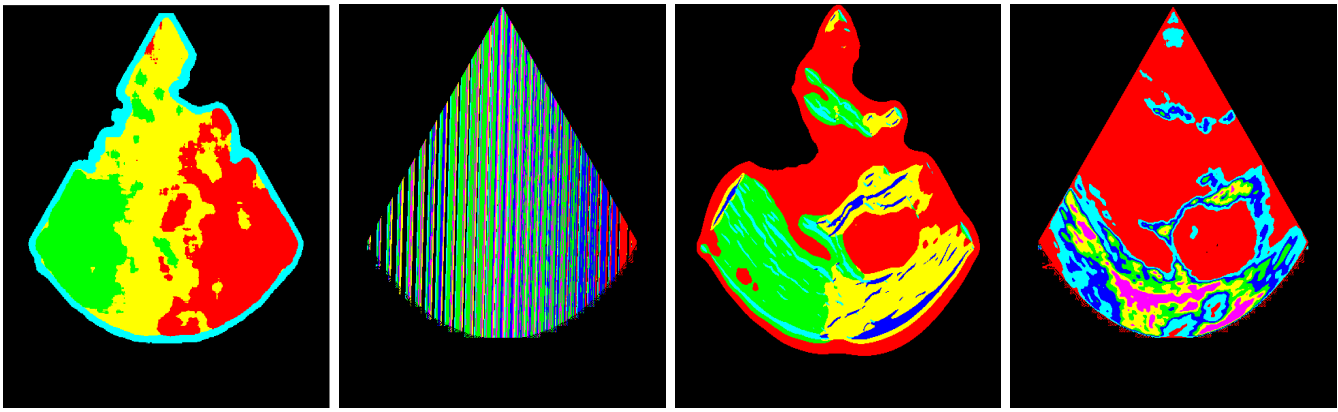
Fig. 3. Segmented Echocardiographic images using HC,DC,GC and FSC

efforts will be towards post processing of the segmentation results on the basis of such domain knowledge.

REFERENCES

[1] O. Bernard, J. D'Hooge, and D. Friboulet, "Statistical modeling of the radio-frequency signal in echocardiographic images based on generalized gaussian distribution," in *ISBI*, 2006, pp. 153–156.

[2] G. Jacob, J. A. Noble, C. P. Behrenbruch, A. D. Kelion, and A. P. Banning, "A shape-space based approach to tracking myocardial borders and quantifying regional left ventricular function applied in echocardiography," *IEEE Trans. Med. Imaging*, vol. 21, no. 3, pp. 226–238, 2002.

[3] A. Mishra, P. K. Dutta, and M. K. Ghosh, "A ga based approach for boundary detection of left ventricle with echocardiographic image sequences," *Image Vision Comput.*, vol. 21, no. 11, pp. 967–976, 2003.

[4] M. Mignotte, J. Meunier, and J.-C. Tardif, "Endocardial boundary e timation and tracking in echocardiographic images using deformable template and markov random fields," *Pattern Anal. Appl.*, vol. 4, no. 4, pp. 256–271, 2001.

[5] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, 1998.

[6] E. Brandt, L. Wigström, and B. Wranne, "Segmentation of echocardiographic image sequences using spatio-temporal information," in *MIC-CAI*, 1999, pp. 410–419.

[7] A. A. Amini, T. E. Weymouth, and R. Jain, "Using dynamic programming for solving variational problems in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 9, pp. 855–867, 1990.

[8] J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber, "Automatic segmentation of echocardiographic sequences by active appearance motion models," *IEEE Trans. Med. Imaging*, vol. 21, no. 11, pp. 1374–1383, 2002.

[9] M. Ma, M. van Stralen, J. H. C. Reiber, J. G. Bosch, and B. P. F. Lelieveldt, "Left ventricle segmentation from contrast enhanced fast rotating ultrasound images using three dimensional active shape models," in *FIMH*, 2009, pp. 295–302.

[10] X. Papademetris, A. J. Sinusas, D. P. Dione, and J. S. Duncan, "3d cardiac deformation from ultrasound images," in *MICCAI*, 1999, pp. 420–429.

[11] D. Boukerroui, A. Baskurt, J. A. Noble, and O. Basset, "Segmentation of ultrasound images-multiresolution 2d and 3d algorithm based on global and local statistics," *Pattern Recognition Letters*, vol. 24, no. 4-5, pp. 779–790, 2003.

[12] J. Yan and T. Zhuang, "Applying improved fast marching method to endocardial boundary detection in echocardiographic images," *Pattern Recognition Letters*, vol. 24, no. 15, pp. 2777–2784, 2003.

[13] N. Lin, W. Yu, and J. S. Duncan, "Combinative multi-scale level set framework for echocardiographic image segmentation," in *MICCAI (1)*, 2002, pp. 682–689.

[14] Y. Chen, H. Tagare, S. R. Thiruvenkadam, F. Huang, D. C. Wilson, K. S. Gopinath, R. W. Briggs, and E. A. Geiser, "Using prior shapes in geometric active contours in a variational framework," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 315–328, 2002.

[15] E. J. H. Wu, M. L. Andrade, D. E. Nicolosi, and S. C. Pontes, "Artificial neural network: border detection in echocardiography," 2008.

[16] N. Archip, R. Rohling, P. Cooperberg, H. Tahmasebpour, and S. K. Warfield, "Spectral clustering algorithms for ultrasound image segmentation," in *MICCAI (2)*, 2005, pp. 862–869.

[17] S. Tabbone, L. Wendling, and J.-P. Salmon, "A new shape descriptor defined on the radon transform," *Computer Vision and Image Understanding*, vol. 102, no. 1, pp. 42–51, 2006.

[18] S. Liao and A. C. S. Chung, "Non-rigid image registration with uniform spherical structure patterns," in *IPMI*, 2009, pp. 163–175.

[19] E. Tola, V. Lepetit, and P. Fua, "A fast local descriptor for dense matching," in *CVPR*, 2008.

[20] A. C. Berg and J. Malik, "Geometric blur for template matching," in *CVPR (1)*, 2001, pp. 607–614.