

# Characterizing spatio-temporal patterns for disease discrimination in cardiac echo videos

T. Syeda-Mahmood<sup>1</sup>, F. Wang<sup>1</sup>, D. Beymer<sup>1</sup>, M. London<sup>2</sup>, and R. Reddy<sup>3</sup>

IBM Almaden Research Center, San Jose, CA, USA<sup>1</sup>,  
UCSF Veterans Hospital, San Francisco, CA, USA<sup>2</sup>,  
Mediciti Hospital, Hyderabad, India<sup>3</sup>

**Abstract.** Disease-specific understanding of echocardiographic sequences requires accurate characterization of spatio-temporal motion patterns. In this paper we present a method of automatic extraction and matching of spatio-temporal patterns from cardiac echo videos. Specifically, we extract cardiac regions (chambers and walls) using a variation of multiscale normalized cuts that combines motion estimates from deformable models with image intensity. We then derive spatio-temporal trajectories of region measurements such as wall motion, volume and thickness. The region trajectories are then matched to infer the similarities in disease labels of patients. Validation results on patient data sets collected from many hospitals are presented.

## 1 Introduction

Analyzing the spatio-temporal regional motion patterns of the heart is important for cardiac disease discrimination. This problem is complicated by the heart’s non-rigid motion, involving twists, rotations, and contractions. Furthermore, the poor imaging quality of 2D echo videos due to low contrast, speckle noise, and signal dropouts cause problems in interpretation. Thus, new methods for robust extraction of spatio-temporal regional motion patterns are needed.

In this paper we present a method of automatic extraction and matching of spatio-temporal patterns from cardiac echo videos. Specifically, we extract cardiac regions (chambers and walls) using a variation of multi-scale normalized cuts that combines motion estimates with image intensity. We then extract spatio-temporal trajectories of region measurements such as wall motion, volume, thickness, etc. In order to compare the regional measurements across diseases, the underlying regions must be brought into correspondence. For this, we use the framework of active shape models. The region trajectories are then compared to differentiate disease labels of videos. Validation results on patient data sets collected from many hospitals are presented.

While there is considerable work in cardiac echo region segmentation and tracking [1], not much work exists for automatic disease discrimination. An earlier work on diagnosis validation using video similarity [2] used features from the entire heart region, restricting its use in characterizing region-specific diseases.

The work reported here builds on the considerable literature for optical flow-based methods [3–6], spatial region segmentation methods [7] and deformable

shape models [8] (to perform a non-rigid alignment of shapes). Optical flow-based segmentation works poorly for cardiac regions due to the low quality of the echo videos and non-smooth heart motion. Spatial region segmentation [7], on the other hand, frequently results in under-segmentation due to the poor resolution in the echo video. While regional measurements are frequently used by physicians to diagnose diseases, many of these require human intervention for region annotation. In our approach, we attempt to automate the extraction of many of these regional measurements. Also, while many of the regional measurements are taken at snapshots in time (e.g. end of systole), our regional measurements capture variations within a heart cycle and can serve as a new diagnostic tool for physicians that is complementary to the existing measurements based on AHE guidelines.

## 2 Region extraction

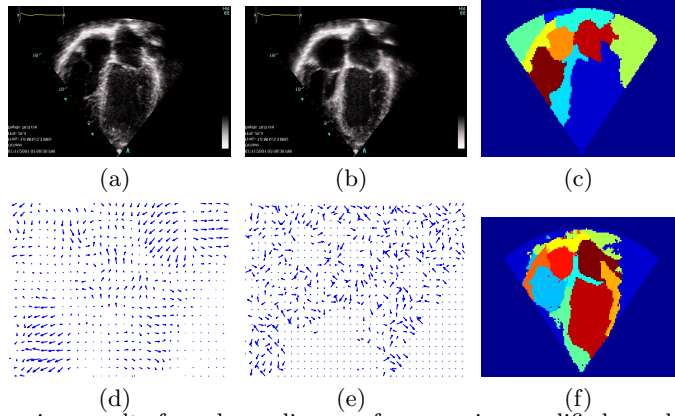
We extract cardiac regions using a variant of multi-scale normalized cuts. Specifically, let  $G = (V, E, W)$  be a graph, with the pixels as graph nodes  $V$ , and edges in  $E$  connecting pairs of neighboring pixels. Let the weight  $W$  associated with each edge be based on the intensity and motion at each pixel. The normalized cut algorithm uses the normalized cut criterion [9] to partition the pixels into self-similar regions. Recently, Coor et. al. [10] improved the efficiency of normalized cuts using a scale-space implementation. We adapt the multiscale normalized cuts to use a weighting function that combines image intensity with motion information so that nearby pixels with similar intensity values and motion are likely to belong to one object. The revised weighting function is given below:

$$W_I(i, j) = e^{-\|X_i - X_j\|/\delta_x - \|I_i - I_j\|/\delta_I - \|V_i - V_j\|/\delta_V} \quad (1)$$

where  $X_i$  denotes pixel location,  $I_i$  and  $V_i$  are image intensity and the motion at pixel location  $i$ .

We estimate motion by treating each successive pairs of intensity image frames as surfaces  $(x, y, I(x, y))$  and finding a deformable surface model that warps one frame into the next as described in the Demons algorithm [11]. The resulting deformation field gives a consistent set of directional velocity vectors, sampling motion densely in both space and time.

Fig.1 illustrates cardiac region segmentation using the modified multilevel normalized cuts combining intensity and motion information. Fig.1a,b show adjacent frames of an echo video. Using the Demons algorithm, the motion field is illustrated in Fig.1d between the frames. In the echo video, the inward and outward pull of the muscles in the different chambers and septal wall can be clearly captured in the resulting deformation field. In comparison, the optical flow field shows haphazard motion inconsistent with the actual observed motion in the echo videos (Fig.1e). Fig. 1c shows the results of using normalized cut-based segmentation based on intensity alone. Finally, Fig.1f shows the normalized cut result using both intensity and motion recovered from deformable registration. As can be seen, the combined use of intensity and motion information has resulted in improved delineation of chamber and septal wall boundaries.



**Fig. 1.** Grouping results for echocardiogram frames using modified graph decomposition algorithm. (See text for details.)

### 3 Spatio-temporal regional measurements

We now discuss the extraction of spatio-temporal region trajectories. Using the deformable surfaces model of motion field, we can establish correspondence between pixels in adjacent fields as described in Section 2. Using this correspondence and the pixel label given by the segmentation step, we obtain the spatio-temporal trajectories of cardiac regions by tracking these regions across a heart beat cycle. To compare echo videos of similar patients, we extract meaningful measurements from the regions based on AHE guidelines. Specifically, we measure the following:

Volumetric trajectory  $V(t)$ : The atrial and ventricular volumes are one of the best diagnostic parameters to be measured for various cardiovascular diseases including myocardial infarction. For example, the change in ventricular volume through the heart beat cycle gives a good indication of the contractile performance, and any impairments can affect the temporal trajectory. We measure this simply as the area of the segmented regions obtained from Section 2 focusing on the cardiac chambers. The temporal trajectories allow the changing volume and hence the 2D + time aspects of the shape to be better modeled.

Translational Displacement  $X(t)$ : The extent of displacement is also a good indicator for many diseases such as hypertrophy, cardiomyopathy, etc. For example, hypokinesia patients often depict smaller septal wall motion. Thus the total extent of displacement as well as the instantaneous displacement of regions such as septal walls within a heart beat cycle can be a good indication of diseases.

Total Motion  $M(t)$ : The displacement feature measures the horizontal component of the full 3D motion of the heart regional volumes. To capture the overall motion of the heart region better, we use the average velocity curve described in [2]. The average velocity curve preserves a common sense of perceived motion per direction and is obtained by averaging the speed and direction of the velocity vectors at each pixel in the region within each frame.

By taking the projection of the average velocity curve along  $x$ ,  $y$ , and  $t$ , we can obtain three additional region features. In particular, the projection onto  $x, y$

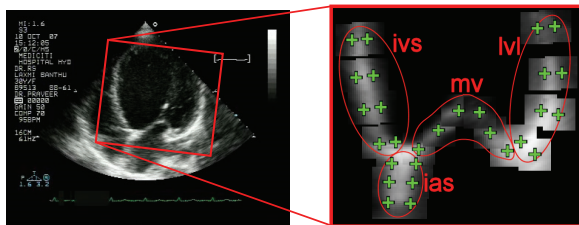
gives the total extent of planar motion of the region and is a good indication of the mechanical performance of the corresponding anatomical region.

## 4 Disease discrimination from regional trajectories

We now turn to the problem of disease similarity inference by a match of the regional trajectories described above. Due to the variabilities introduced by anatomical differences, viewpoint variations, imaging quality, and segmentation errors, we expect the regional feature trajectories to show considerable variation even for patients within the same disease category. However, we expect to derive higher order statistics from the trajectories that make the comparison meaningful. Furthermore, the trajectories of corresponding regions (e.g. left ventricle to left ventricle) must be compared, so we chose to model the structure of the heart using an active shape model. Echo videos from patients with similar disease are then brought into alignment with the active shape model. The labels associated with the regions in the training stage are then used to infer the region correspondence.

### 4.1 Region correspondence using active shape models

In active shape models [8], the shape of an object is represented by a set of feature points on the object, usually chosen to sample object contours at regular intervals. Given the locations of  $n$  feature points  $f_1, f_2, \dots, f_n$ , the shape vector  $\mathbf{s}$  concatenates the  $(x, y)$  coordinates of all feature points,  $\mathbf{s} = [x_1, y_1, \dots, x_n, y_n]^T$ . For the heart, we have initially focussed on the apical four chamber view (Fig. 2). The feature points are centered around the mitral valve (point group mv), interventricular septum (ivs), interatrial septum (ias), and the lateral wall (lv). We chose not to trace the entire boundary of the left ventricle because the apex is often not visible due to the ultrasound zoom and transducer location. Wall boundaries are annotated on both the inner and outer sections.



**Fig. 2.** Our active shape model contains 28 features in four groups centered about the mitral valve region. The right image shows the mean texture for the 28 feature patches.

Image texture is represented by image patches centered around the feature locations. Given a shape with  $n$  features, the texture vector  $\mathbf{t}$  concatenates the pixels from all the patches into one long vector.

We used a standard implementation of active shape models (ASMs) as described in [8] with a few enhancements. The training images for the active shape

model were obtained from echo video sequences of patients with different diseases and different video frames within a heart beat cycle to capture the range of spatial and temporal variations of the heart structure. Cardiac cycle clips from apical four chamber (A4C) views were manually extracted from sample echo video. Each patient has at least two cycles extracted, with the first used for ASM training, and the remainder used for testing. The extracted cardiac cycles are synchronized on frame 0 using the peak of the R wave as seen in the echo video’s ECG lead.

Unlike most manual initializations for fitting active shape models to new images, we used an automatic initialization method where a distance-to-eigenspace method is first used to generate seed ASM initializations. The seed locations are then refined and evaluated using a coarse-to-fine pyramid approach.

Once we have fit the ASM model to a new image, corresponding regions are determined by simply associating the region label of the projected points of the active shape model with the underlying pixels.

## 4.2 Matching regional trajectories

Once the corresponding regions are identified, we normalize the trajectories to make them invariant to various artifacts of imaging. We normalize for heart rate differences by extracting a single heart cycle from the video starting from the peak of the R-wave as a common reference point (using the ECG signal embedded in the video). We also report all measurements in cm instead of image coordinates, which makes the trajectories invariant to effects of scale (zoom-in or out). This is done by using the 10 cm calibration markers on the left and right sides of the echo sector. Finally, we match trajectories of videos with the same viewpoint to make the comparison meaningful (e.g. 4AC to 4AC). Since the regional measurements used (area, extent of motion, etc.) are not affected by rotation, or translation, and since scale is accounted by keeping the centimeter unit, the trajectories are invariant to similarity transformations as well.

The regional measurements capture different aspects of cardiac regions. Hence, we measure their similarities differently as described below:

**Volumetric trajectories** The distance between two volumetric trajectories  $V_1(t)$  and  $V_2(t)$  is given by

$$D(V_1, V_2) = \sum_i (V_{n1}(i) - V_{n2}(i))^2 \quad (2)$$

where  $V_{n1}(i) = \frac{V_1(i)}{Max(V_1)}$  and  $V_{n2}(i) = \frac{V_2(i)}{Max(V_2)}$ . By normalizing the regional volume, we factor out the effects of absolute volume and focus on trajectory shape. To use absolute volumes, further calibration will be needed.

**Horizontal displacement:** For the displacement trajectory  $X(t)$ , we measure the total motion of these curves by integrating the speed vector  $|dX(t)/dt|$  over the cardiac cycle. Thus, given two displacement trajectories  $X_1(t), X_2(t)$ , the distance between them is given by

$$D(X_1, X_2) = \left| \int \left| \frac{dX_1(t)}{dt} \right| dt - \int \left| \frac{dX_2(t)}{dt} \right| dt \right|. \quad (3)$$

Total motion: For the total motion trajectory, we measure the area under the curve. The difference between two projected motion trajectories  $M_1(t), M_2(t)$  is

$$D(M_1(t), M_2(t)) = \left| \int \int M_1(x(t), y(t)) dx, dy - \int \int M_2(x(t), y(t)) dx dy \right| \quad (4)$$

## 5 Results

To verify our conjecture that the trajectories of regional features are useful in inferring disease similarity, we experimented with a number of public cardiac echo data sets, namely, the GE Vivid online medical library, the Yale Medical School library, and data collected by cardiologists from various hospitals in India. The data sets depicted over 500 heart beat cycles chosen from over 50 patients with a number of cardiac diseases including cardiomyopathy, hypokinesia, mitral regurgitation, hypertrophy, etc. The echo videos from hospitals came with diagnoses as well as the common regional measurements such as wall thickness and ventricular volumes which were used as ground truth for testing our algorithms.

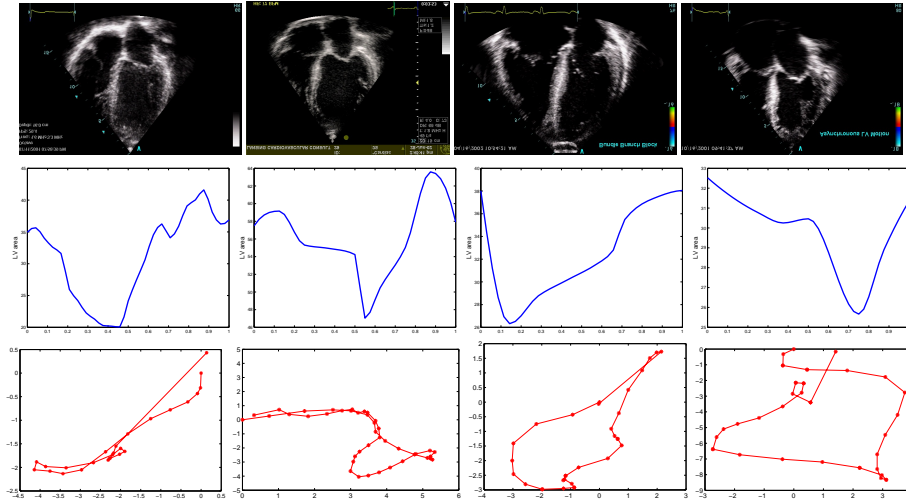
### 5.1 Disease discrimination using regional trajectories

We first illustrate the discrimination of diseases through regional trajectories. Fig. 3 shows the regional trajectories for patients with different disease labels. In the first row of Fig. 3, the two images on the left show videos of normal patients taken from the GE Vivid online library collection, while the third and fourth image show videos of patients diagnosed with asynchronous LV motion and left bundle-branch block respectively. The trajectories of the corresponding left ventricular volumes are shown in the second row of Fig. 3. The  $x$ - $y$  projected motion trajectories of the intra-ventricular septum are shown in last row of Fig. 3. The volumetric trajectories show that the ventricular volume is minimum at roughly the midpoint of the cardiac cycle showing the symmetry of movement, while the diseases cases show asymmetry when the ventricular volume minimum is reached. In the projected motion trajectories, overall displacement of the IVS is much contained as seen by the small area spanned by the trajectories, whereas the diseased cases show a much larger area of total displacement.

The distance measures between each of the regional trajectories for the videos in Fig. 3 are shown below in Eqn. 5. Here the rows and columns correspond to the videos depicted in the first row of Fig. 3 respectively. As can be seen in each case, the normal videos are more similar in the distance metrics used than the abnormal cases.

$$\text{MSE Distance} = \begin{pmatrix} 0 & 1.9728 & 2.4524 & 3.4176 \\ 1.9728 & 0 & 3.3106 & 2.9627 \\ 2.4524 & 3.3106 & 0 & 3.3013 \\ 3.4176 & 2.9627 & 3.3013 & 0 \end{pmatrix} \quad (5)$$

Using the total motion measurement, the total motion values for the 4 echo videos in square millimeter are 0.13, 0.55, 1.79, 1.16 respectively. Thus even using this measure, the normal videos are deemed similar and are discriminable from the abnormal ones.



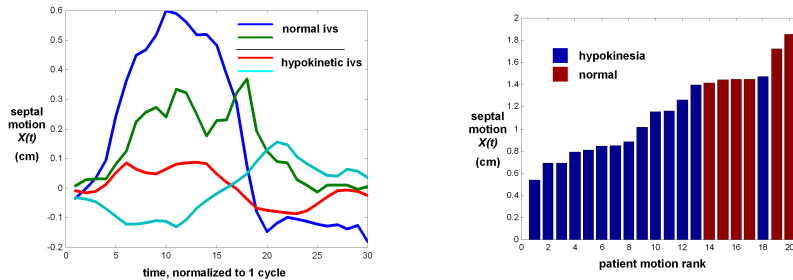
**Fig. 3.** Feature plots of four sample videos from GE Vivid online medical library. (see text for details).

## 5.2 Disease similarity using septal motion

To further evaluate the validity of regional trajectories, we conducted a focused study on one of the regional measurements, namely, the displacement of the septal wall, as an indicator for disease. As the left ventricle (LV) expands and contracts, the septal wall shows a cyclic motion. In the case of motion abnormalities such as hypokinesia, the septal motion is considerably reduced.

To perform this study, we used a collection of echocardiogram videos from 20 cardiology patients from a hospital in India. The cardiologists capture the echo video in their normal clinical setting, and echo findings and disease diagnoses are included as text reports. Thus, we are able to divide the patients by any cardiac disease of interest – in this case, there were 14 patients diagnosed with hypokinesia in the septal wall, and 6 patients with normal motion of the septum.

Fig. 4, left, shows four example displacement curves  $X(t)$ , two for patients with a normal septum, and two for patients with a hypokinetic septum. The time axis in these plots have been normalized to one heart cycle. Qualitatively from the plots of  $X(t)$ , it is easy to see that the normal cases have much more motion displacement than the hypokinetic cases. Finally, we evaluated the effectiveness of the displacement measure for discriminating between patients with and without hypokinesia. Fig. 4, right, shows the result of ranking echo videos based on the displacement metric of Equation 3. As can be seen, the displacement trajectory can be used to discriminate the hypokinesia patients from the normal patients. The one outlier (patient 18) is due to a tracking error during trajectory formation when the lateral walls go out of view, biasing the septum tracking and increasing the septal motion estimate.



**Fig. 4.** Left: The motion displacement  $X(t)$  for four patients, 2 normal ones and 2 with ivs hypokinesia; Right: the patients sorted by their septal motion.

## 6 Conclusions

In this paper, we presented novel methods of automatic extraction and matching of spatio-temporal patterns from cardiac echo videos. The region trajectories are then matched to infer the similarities in disease labels of patients. Validation results on patient data sets collected from many hospitals were presented.

## References

1. Dydenko, I., Jamal, F., Bernard, O., D’Hooge, J., Magnin, I.E., Friboulet, D.: A level set framework with a shape and motion prior for segmentation and region tracking in echocardiography. *Medical Image Analysis* **10**(2) (2006) 162–177
2. Syeda-Mahmood, T.F., Yang, J.: Characterizing normal and abnormal cardiac echo motion patterns. In: *Computers in Cardiology*. Volume 33. (2006) 725–728
3. Gheissari, N., Bab-Hadiashar, A.: Motion analysis: Model selection and motion segmentation. *ICIAP’03* **00** (2003) 442–447
4. Cremers, D., Soatto, S.: Motion competition: A variational framework for piecewise parametric motion segmentation. *IJCV’05* **62**(3) (2005) 249–265
5. Tagare, H.: Shape-based nonrigid correspondence with application to heart motion analysis. *IEEE Trans. Med. Imaging* **18**(7) (1999) 570–579
6. Papademetris, X., Sinusas, A.J., Dione, D.P., Duncan, J.S.: 3d cardiac deformation from ultrasound images. In: *MICCAI’99*, London, UK (1999) 420–429
7. Zerhouni, E., Parish, D., Rogers, W., Yang, A., Shapiro, E.: Human heart: Tagging with mr imaginga method for noninvasive assessment of myocardial motion. *Radiology* **169**(1) (1988) 59–63
8. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models their training and application. *Comput. Vis. Image Underst.* **61**(1) (1995) 38–59
9. Shi, J., Malik, J.: Motion segmentation and tracking using normalized cuts. In: *ICCV’98*, Washington, DC, USA (1998) 1154
10. Cour, T., Benezit, F., Shi, J.: Spectral segmentation with multiscale graph decomposition. In: *CVPR’05*, Washington, DC, USA (2005) 1124–1131
11. Thirion, J.P.: Image matching as a diffusion process: an analogy with maxwell’s demons. *Medical Image Analysis* **2**(3) (1998) 243–260