Clustering Based Cardiac Resynchronization Therapy Prediction using Open Source Toolkit PRTools

Heng Huang[1], Li Shen[2], Rong Zhang[1], Fillia Makedon[1], Justin Pearlman[1,3]

[1] Department of Computer Science, Dartmouth College, Hanover, NH 03755
[2] Computer and Information Science Department, Univ. of Mass. Dartmouth, MA 02747
[3] Department of Cardiology, Dartmouth Medical School, Lebanon, NH 03756
hh@cs.dartmouth.edu

Abstract. We propose a novel framework at MICCAI 2005 [1] to predict pacing sites in the left ventricle (LV) of a heart. This framework can be used to assist pacemaker implantation and programming in cardiac resynchronization therapy (CRT) that is a widely adopted therapy for heart failure patients. Hierarchical agglomerative clustering technique is performed to the time series of LV wall thickness to identify pacing site candidates. Meanwhile, pearson correlation coefficients of wall motion series show the dissimilarity between them. These main components of our clustering based prediction framework are implemented by using open source software toolkit PRTools.

1 Introduction

In the United States, heart failure is responsible for almost 1 million hospital admissions and 40,000 deaths annually. Heart failure is the pathophysiological state in which an abnormality of cardiac function is responsible for the failure of the heart to pump blood at a rate for the requirements of the body tissues. As the main problem, the walls of the LV cannot contract synchronously that damaged the heart's pumping action of patient.

De Teresa *et al.* [2] demonstrated that cardiac function can be improved by changing the sequence of the ventricular electrical activation using pacing. They also noted that the LV ejection fraction (an important index for cardiac function) was maximal when wall contractions were simultaneous. Based on the previous sudies, a promising therapeutic option, called cardiac resynchronization therapy (CRT), has been proposed as an alternative treatment in patients with severe, drug-refractory heart failure.

Our paper in MICCAI 2005 [1] proposes an efficient framework to predict the optimal LV pacing sites that should be stimulated by electrical impulses of pacemaker. Hierarchical clustering method is used on a time series of wall thickness measurements. Based on the clustering result, we can find candidate pacing sites with abnormal local motion.

In our framework implementation, open source toolkit PRTools is used for wall motion series similarity measurement, hierarchical clustering and clustering results visualization. PRtools is a powerful MATLAB toolbox for pattern recognition purposes. It can be downloaded from the PRTools website (http://www.ph.tn.tudelft.nl/ bob/PRTOOLS.html). PRTools offer 200 pattern recognition routines and the additional 200 support routines that are a basic set covering largely the area of statistical pattern recognition [3].

The rest of this paper is organized as follows. Section 2 describes the data set and the methods used in this study. Section 3 presents and discusses the results. Section 4 concludes this work.

2 Materials and Methods

2.1 Image Acquisition

Cardiac magnetic resonance imaging (MRI) is used to capture 3D images of a heart during its normal operation in the short-axis orientation. With acquisition timed according to heartbeat frequency, seventeen images can be acquired during each heartbeat. In this work imaging was performed using FIESTA on a 1.5 Tesla scanner with flip angle 20° and interslice gap of 5 mm. The short-axis orientation was operator-determined from four-chamber scout views, optimizing for perpendicularity to the cardiac wall.

The sequences of heart images were produced in the DICOM format with 256×256 pixels size ($260 \times 260\,\mathrm{mm}^2$). Each sequence consists of 17 volume images that together represent one complete heartbeat cycle.

2.2 Left Ventricle Motion Modeling

Recent studies indicate that beneficial effects of CRT are related to improved mechanical synchrony, thereby it can increase the pump function efficiency [4–6]. A successful CRT will synchronize the wall contraction so that LV ejection fraction is maximized. Therefore, the improvement in cardiac performance is highly dependent on the pacing site that changes the sequence of ventricular activation in a manner that translates to an improvement in cardiac performance.

In order to quantify the ventricular mechanical asynchrony or synchrony that can directly help determine optimal treatment, we develop our spatio-temporal model to describe a temporal sequence of wall thickness changing during a heart cycle. It is one of the most sensitive indicators of ventricular dysfunctional contraction and can be used to index the ventricular wall motion.

Left ventricular spatio-temporal model is proposed in our previous paper[1]. In this model, both endocardium and epicardium of the LV are reconstructed by using the spherical harmonic (SPHARM) method, which was introduced by Brechbühler, Gerig and Kübler [7] for modeling any simply connected 3D object. Our surface alignment algorithm [8] computes a new parameterization for $\mathbf{v}_1$ and $\mathbf{v}_2$ so that the Euclidean distance $D(\mathbf{v}_1, \mathbf{v}_2)$ is minimized.

Combining the SPHARM description and our surface alignment method offers a set of spatio-temporal surface correspondences for our wall motion descriptor. In our experiments, each sampling mesh on one surface has $32 * 32$ nodes and each node has a wall thickness value. The wall motion series we create includes the thickness values for each surface node at each time phase during a heart cycle, from end-diastolic phase to next end-diastolic one. We obtain about n wall motion series, where n varies from 80 to 100 in different experiments. The corresponding points of these n series are uniformly distributed on the LV surfaces. These wall motion series can characterize local contraction behaviors of the LV wall and have a potential to capture the contraction abnormality of a failing heart.

2.3 Similarity measurement

For a pair of (θ, ϕ), the corresponding wall motion series is denoted as $\boldsymbol{w}(\theta, \phi) = \{w_1(\theta, \phi), w_2(\theta, \phi), ..., w_n(\theta, \phi)\}$, where $w_i(\theta, \phi)$ is the wall thickness value of wall motion phase i corresponding to the parametrized point (θ, ϕ) on the epicrdium.

Formlly, given two wall motion series $\boldsymbol{w}(\theta_x, \phi_x)$ and $\boldsymbol{w}(\theta_y, \phi_y)$, we employ the following formula to measure the dissimilarity between them:

$$d_{corr}(\boldsymbol{w}(\theta_x, \phi_x), \boldsymbol{w}(\theta_y, \phi_y)) = 1 - r(\boldsymbol{w}(\theta_x, \phi_x), \boldsymbol{w}(\theta_y, \phi_y)) =$$

$$1 - \left[\sum_{i=1}^{n} \left(\frac{w_i(\theta_x, \phi_x) - w_{mean}(\theta_x, \phi_x)}{\sigma_x} \right) \left(\frac{w_i(\theta_y, \phi_y) - w_{mean}(\theta_y, \phi_y)}{\sigma_y} \right) \right] / n,$$

where $\sigma = \sqrt{(\sum_{i=1}^{n} (w_i(\theta, \phi) - w_{mean}(\theta, \phi))^2)/n}$. $r(\boldsymbol{w}(\theta_x, \phi_x), \boldsymbol{w}(\theta_y, \phi_y))$ is the Pearson correlation coefficient of two wall motion series, $w_{mean}(\theta, \phi)$ is the mean of wall motion series, and σ is the standard deviation of $\boldsymbol{w}(\theta, \phi)$. A Pearson correlation coefficient indicates how the two wall motion series are related and the strength of that relationship. The Pearson correlation coefficient is always between -1 and 1, and we normalize distance function as $d_{corr}/2$ (the result will change from 0 to 1) in our experiments.

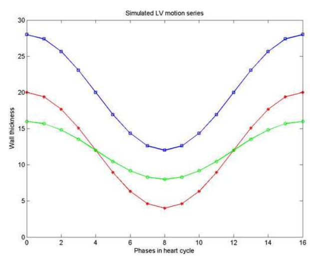

Fig. 1. Example of shift and scaling relationship of wall motion series.

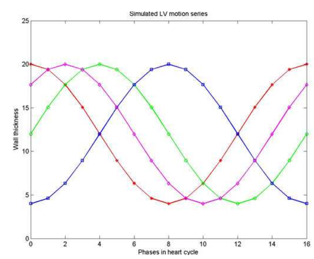

Fig. 2. Example of the contraction delay between wall motion series.

In this distance function, $\boldsymbol{w}_{mean}(\theta, \phi)$ is used to remove the shift difference. In Fig. 1, the square-curve (each curve represents a simulated wall motion series) and star-curve should be quite similar, since one of them can be shifted up vertically to superpose with the other one. Similarly, σ is used to normalize the wall motion series when we calculate the similarity score between them. In Fig. 1, values of star-curve are roughly 1.6 times those of circle-curve. Since we are interested in the wall thickness change instead of the actual thickness value, these two motion series are also similiar. For example, the mid-epicardium should have a larger wall thickness value and contract more acutely than apex-epicardium. But if they contract and dilate synchronously, their wall motion should be treated as the same.

In Fig. 2, star-curve represents a normal wall motion series. The wall motion series represend by diamond-curve activate a little late, and their correlation value is 0.7265.

As the activity delay increases, the Pearson correlation values between the other motion series and the normal one desrease to 0 (circle-curve) and -1 (square-curve, it is perfectly divergent). Since the Pearson correlation coefficient is sensitive to direction of change (increasing or decreasing), it is reasonable to use it to measure the similarity between wall motion series. The Pearson correlation coefficient is always between -1 and 1, and we normalize distance function as $d_{corr}/2$ (the result will change from 0 to 1) in our experiments.

2.4 Hierarchical clustering

We apply hierarchical clustering algorithm [9] to group similar wall motion series together. It is a bottom-up clustering method where clusters can have sub-clusters. For any set of n objects, hierarchical clustering starts with every single object in a single cluster. Then, in each successive iteration, it merges the closest pair of clusters by satisfying their proximity information criteria, until all of the data are in one cluster.

In our case, the objects are the wall motion series of sampled points on epicardium, and the proximity criteria is defined by the distance described in between pairs of wall motion series. The algorithm is sketched as follows:

1. Each wall motion series is assigned to a separate cluster.
2. All pair-wise distances between clusters are calculated and stored into a distance matrix.
3. **repeat**
   – The pair clusters with the closest distance is found
   – Those two clusters are merged into one cluster
   – The distances between the new groups and the remaining groups are computed to get a new distance matrix
4. **until** All of the wall motion series are clustered into a single group.

In Step 3, we employ the average-link approach: the distance between two clusterings is defined as the average of distances between all pairs of wall motion series, where each pair is made up of wall motion series from each group. Thus, the distance matrix can be updated using the following formula:

$$d(r, p+q) = \frac{n_p}{n_p + n_q} d(r, p) + \frac{n_q}{n_p + n_q} d(r, q)$$

where p and q are merged into one new cluster, and n_p and n_q are the numbers of wall motion series in group p and q respectively.

The hierarchical clustering process usually stops after performing $n - 1$ iterations in Step 3, and results in a dendrogram. This dendrogram is a binary tree (see Fig. 3 for an example) in which each data point corresponds to a leaf nodes, and distance from the root to a subtree indicates the similarity of subtrees – highly similar nodes or subtrees have joining points that are farther from the root.

3 Results

We have implemented our pacing site prediction framework using Matlab 6.5. The open source platform toolkit PRTools, a MATLAB based toolbox for pattern recognition, is used as the implementation for wall motion series similarity measurement, hierarchical clustering and clustering results visualization.

To show the effectiveness of this framework, we use cardiac MRI data from 20 patients in our experiments, where half of them have heart failure problems. These experiments are conducted on a PC with a 2.40GHz CPU and 512 MB main memory. Note that the patients are diagnosed by specialized physicians, and these diagnostic results are used to validate our results in the experiments.

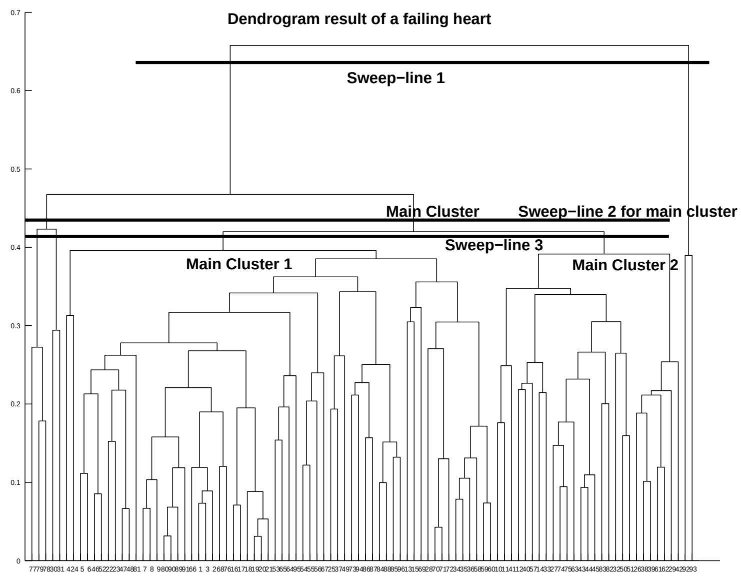

Fig. 3. Dendrogram result of a failing heart. The x label represents the number of wall motion series. The y label corresponds to the distance between clusters. The dendrogram is cut into clusters by the "sweep-line 3".[1]

Our primary purpose for building a cluster hierarchy is to structure and present wall motion series at different levels of abstraction. Using a dendrogram, researchers and technicians can easily know the dissimilarity between subclusters that represent certain parts on the epicardium.

We move the horizontal sweep-line from top to bottom in the dendrogram result (for example, the "sweep-line 1" in Fig. 3) to get the abnormal clusters (small clusters) that have a large dissimilarity to the main cluster. Note that the pacemaker system uses electrical impulses to adjust the sites whose contraction characteristics are considerably different from other sites'. Thus, hierarchical clustering results can help us to find these location candidates for installing the pacing leads.

The physician will test the pacing lead on candidate pacing sites according to the suggested site ordering until they find a suitable region for fixing the tip of pacing lead. If the list is empty and a suitable site isn't found, we will continue to select a lower value sweep-line in the dendrogram result, for example, the “sweep-line 2” and “sweep-line 3” in Fig. 3.

Because the candidates list includes locations with notable asynchronous contraction and timing delay, the optimal resynchronization therapy can be obtained after adding electrical pulse into these candidates. These sites are potentially good candidates to implant the pacemaker for a more efficient CRT.

4 Conclusions

During our framework implementation, the major advantages of using the PRTools open source software are that it provides a lot of pattern recongnition routines that make the success of short term project more feasible and helps user to visualize their datasets and results conveniently.

References

1. Huang, H., Shen, L., Zhang, R., Makedon, F., Hettleman, B., Pearlman, J.: A Prediction Framework for Cardiac Resynchronization Therapy via 4D Cardiac Motion Analysis. International Conf. on Medical Image Computing and Computer Assisted Intervention. (2005)
2. De Teresa, E., Chamorro, J.L., Pulpon, L.A.: An even more physiologic pacing: changing the sequence of activation. Proceedings of the VIIth World Symposium on Cardiac Pacing. (1984) 395–400
3. Duin, R.P.W.: PRTools Version 3.0, A Matlab Toolbox for Pattern Recognition. http://www.ph.tn.tudelft.nl/ bob/PRTOOLS.html, 2000.
4. Bader, H., Garrigue, S., Lafitte, S., *et al.*: Intra-LV electromechanical asynchrony. A new independent predictor of severe cardiac events in heart failure patients. J Am Coll Cardiol. **43(2)** (2004) 248–256
5. Bordachar, P., Garrigue, S., Lafitte, S., *et al.*: Interventricular and intra-LV electromechanical delays in right ventricular paced patients with heart failure: implications for upgrading to biventricular stimulation. Heart. **89(12)** (2003) 1401–1405
6. Schreuder, J.J., Steendijk, P., van der Veen, F.H., *et al.*: Acute and short-term effects of partial left ventriculectomy in dilated cardiomyopathy: assessment by pressure-volume loops. J Am Coll Cardiol. **36(7)** (2000) 2104–2114
7. Brechbühler, Ch., Gerig, G., and Kübler, O.: Parametrization of closed surfaces for 3D shape description. Computer Vision and Image Understanding. **61(2)** (1995) 154–170
8. Huang, H., Shen, L., Zhang, R., Makedon, F., Hettleman, B., Pearlman, J.: Surface Alignment of 3D Spherical Harmonic Models: Application to Cardiac MRI Analysis. International Conf. on Medical Image Computing and Computer Assisted Intervention. (2005)
9. Alpaydin, E.: Introduction to Machine Learning. The MIT Press, (2004)