DIFFUSION MAP APPROACH TO CLASSIFYING EARLY STAGE CARDIAC DYSFUNCTION

Hsun-Hsien Chang[1], *José M. F. Moura*[1], *Yijen L. Wu*[2], *and Chien Ho*[2,3]

[1]Electrical and Computer Engineering
[2]Pittsburgh NMR Center for Biomedical Research
[3]Department of Biological Sciences
Carnegie Mellon University,
Pittsburgh, PA, USA

ABSTRACT

Magnetic resonance (MR) tagging technology can assist us in determining the motions of the myocardial pixels in a sequence of MR images. This paper presents a semi-supervised algorithm that processes these motion maps and classifies automatically myocardial dysfunctional motions. In distinction with other methods, our algorithm requires that only a few *normal* motions are labeled *a priori*. This is significant because, while *normal* motions can be confidently labeled by a human expert, abnormal motions are very difficult to label with high reliability by an operator.

We use a graph to capture the motion map of the left ventricle. The normalized weighted adjacency matrix of the graph is interpreted as a stochastic matrix. Performing random walks, or diffusion, on the graph determines how similar myocardial motions are. Similar motions on the graph are represented by the diffusion maps framework as closer vectors in a Euclidean space. In the Euclidean space, we adopt eigen-analysis on a small portion of labeled normal motions. The analysis leads to a hyperelliptic surface that classifies the remaining cardiac motions as normal or dysfunctional.

Index Terms— cardiac motion, dysfunction, classification, diffusion maps, spectral graph.

1. INTRODUCTION

With the advent of cardiac magnetic resonance tagging technology, we can estimate dense motion maps of the left ventricle [1]. That is, each pixel of the myocardium is assigned a vector indicating the moving direction and magnitude toward the next cardiac phase. Monitoring the dense cardiac motions, cardiologists can learn where the dysfunctional and abnormal cardiac tissue is.

To determine dysfunctional motions, cardiologists need to scrutinize entire motion maps and then manually determine the regions of abnormalities. This manual work is time consuming. In [2], we proposed a computer assisted method to detect regional heart malfunction based on spectral graph theory [3–5]. The classifier in [2] is semisupervised—initially, a human expert labels as normal or abnormal a small portion of the motions and then the classifier propagates the human prior knowledge to the remaining unlabeled motions. In early stages of heart disease, the abnormal motions are not prominent. As a result, the human expert lacks confidence when labeling dysfunctional motions, while labeling normal motions is usually a reliable task. We propose a new algorithm that requires only prior knowledge on the *normal* motions.

We develop our classifier in the framework of *diffusion maps* [6]. A human expert labels a small portion of normal motions. Then, the algorithm automatically diffuses the normal labels to the remaining undetermined motions. The abnormal motions are in regions where diffusion is hard. After a sufficient number of diffusion steps, the regions with low probability of diffusion are classified as abnormal.

1.1. Algorithm Overview

The algorithm starts by describing the image of the left ventricle with a graph. Each myocardial pixel corresponds to a vertex in the graph. The variations of motions within small neighborhoods are assigned as weights to the graph edges. From graph theory, we associate a weighted adjacency matrix to the graph. We normalize the weighted adjacency matrix, which leads to a stochastic matrix that defines a Markov chain on the graph. Similarity between two arbitrary myocardial motions is determined by how easily they can be reached from one another through random walks, or diffusion, on the graph. The framework of diffusion maps [6] transforms vertices that are easily reached by diffusion from one another on the graph into close vectors in a Euclidean space. In other words, the normal and abnormal motions are mapped to two clusters of vectors in a Euclidean space. Classification in this Euclidean space simplifies the task of classifying motions on the graph. We then train the classifier by performing eigen-analysis on the vectors corresponding to the prelabeled normal motions.

This work was supported by NIH grants, R01EB/AI-00318 and P41EB001977, to the Pittsburgh NMR Center for Biomedical Research, Carnegie Mellon University. Correspondence to H-H Chang (hsunhsien@cmu.edu) or JMF Moura (moura@cmu.edu).

The eigen-analysis leads to a hyperelliptic surface enclosing the labeled normal vectors. The hyperelliptic surface is used to classify the myocardial motions as normal or abnormal.

1.2. Paper Organization

This paper is organized as follows. Section 2 develops in detail the classification method. Section 3 demonstrates the experimental results with real cardiac MRI data. Finally, we conclude the paper in Section 4.

2. METHODOLOGY

2.1. Graph Representation of Motion Maps

Let a set $I = \{1, 2, \cdots, N\}$ index the myocardial pixels, where N is the number of myocardial pixels. A myocardial pixel i has coordinate $\mathbf{x}_i = [x_i, y_i]^T$ and motion $\mathbf{u}_i = [u_i, v_i]^T$. We describe the myocardial motion map by a graph $\mathcal{G}(V, E)$, which consists of a set V of vertices and a set E of edges linking the vertices. In the graph $\mathcal{G}(V, E)$, each vertex i is a pixel; it is linked to itself and to its four nearest neighboring pixels. An edge linking vertices i and j is denoted by e_{ij}.

To capture the motion map by a graph, we use weights W_{ij} to record the similarities between pairs of motions $\mathbf{u}_i, \mathbf{u}_j$ if there are edges e_{ij}. Due to the errors in the motion estimation step, we treat the motions $\mathbf{u}_i$ as random vectors and use the Mahalanobis distance ρ_{ij}, see [7], to compute W_{ij}. The Mahalanobis distance is

$$\rho_{ij} = \sqrt{(\mathbf{u}_i - \mathbf{u}_j)^T \Sigma_{ij}^{-1} (\mathbf{u}_i - \mathbf{u}_j)}\,, \tag{1}$$

where Σ_{ij} is the covariance matrix between $\mathbf{u}_i$ and $\mathbf{u}_j$. In the graph representation of the motion map, we assign larger weights to edges linking pixels with shorter Mahalanobis distances. We compute these weights using the Gaussian kernel; that is, the weight W_{ij} between a pair of myocardial pixels i, j is

$$W_{ij} = \begin{cases} \exp\left(-\frac{\rho_{ij}^2}{\epsilon}\right), & \text{if } i \text{ and } j \text{ are linked,} \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The graph $\mathcal{G}$ can now be represented by the *weighted adjacency matrix* $\mathbf{W}$ whose entries are W_{ij}.

We normalize the weighted adjacency matrix $\mathbf{W}$ by its row sums to obtain a new matrix $\mathbf{P}$, namely

$$P_{ij} = \frac{W_{ij}}{\sum_k W_{ik}}\,. \tag{3}$$

The matrix $\mathbf{P}$ is used to specify the transition probabilities between vertices in the graph and, as such, describes a Markov chain. The entry P_{ij} is the one-step transition from vertex i to vertex j. The multiple-step transition probabilities are the entries of powers of P; i.e., let $\mathbf{P}^m$ be the m-th iterate of $\mathbf{P}$, the entry $(P^m)_{ij}$ is the probability of going from vertex i to vertex j in m steps.

The i-th row $(P^m)_{i-}$ of $\mathbf{P}^m$ collects the transition probabilities of starting at vertex i and ending at each of the other vertices. When two rows, say $(P^m)_{q-}$ and $(P^m)_{r-}$, are equal, the vertices q and r will behave as identical random walks in subsequent steps. Hence, the vertices q and r have similar characteristics and can be determined as belonging to the same class. However, to build a classifier from this simple idea, we need to consider two issues. First, there is a significant computational cost in computing powers of $\mathbf{P}$. Although $\mathbf{P}$ inherits the sparse structure from the weighted adjacency matrix $\mathbf{W}$, the matrix $\mathbf{P}^m$ will be a full matrix when $m \gg 1$. Second, we need a good measure to determine the similarity between two probability distributions $(P^m)_{q-}$ and $(P^m)_{r-}$. To handle these issues, we adopt the diffusion maps framework [6].

2.2. Diffusion Maps

The eigen-decomposition of the matrix $\mathbf{P}$ gives a set of left eigenvectors and a set of right eigenvectors:

$$\{\phi^{(j)}\}_{j=0}^{N-1} \quad \text{and} \quad \{\psi^{(j)}\}_{j=0}^{N-1}\,, \tag{4}$$

respectively. The left and right eigenvectors correspond to the same set of eigenvalues $\lambda_0 > \lambda_1 \geq \cdots \geq \lambda_{N-1}$, where $\lambda_0 = 1$. The components $\phi_i^{(0)}$ of the zeroth left eigenvector $\phi^{(0)}$ are

$$\phi_i^{(0)} = \frac{\sum_j W_{ij}}{\sum_{i,j} W_{ij}}\,, \tag{5}$$

and the zeroth right eigenvector $\psi^{(0)} = \mathbf{1}$ is a constant vector with all components being 1. The left eigenvectors are equivalent to the stationary distributions on the graph. The right eigenvectors are the dual of the left eigenvectors, and they satisfy the inner product

$$\langle \phi^{(j)}, \psi^{(k)} \rangle = \delta_{jk} \tag{6}$$

and the relation

$$\psi_\ell^{(j)} = \frac{\phi_\ell^{(j)}}{\phi_\ell^{(0)}}\,. \tag{7}$$

In the diffusion maps framework, the eigenvectors are normalized with respect to $\phi^{(0)}$:

$$\|\phi^{(j)}\|^2 = \sum_\ell \frac{{\phi_\ell^{(j)}}^2}{\phi_\ell^{(0)}} = 1 \tag{8}$$

and

$$\|\psi^{(j)}\|^2 = \sum_\ell {\psi_\ell^{(j)}}^2 \phi_\ell^{(0)} = 1\,. \tag{9}$$

Coifman *et al.* [6] define the m-step diffusion distance $\xi_m(q,r)$ between vertices q and r, or between $(P^m)_{q-}$ and $(P^m)_{r-}$, as

$$\xi_m(q,r) = \sqrt{\sum_{k=1}^{N} \frac{\left[(P^m)_{qk} - (P^m)_{rk}\right]^2}{\phi_k^{(0)}}} \,. \tag{10}$$

The spectral representation of $(P^m)_{ab}$ is, see [6],

$$(P^m)_{ab} = \sum_{j=0}^{N-1} \lambda_j^m \psi_a^{(j)} \phi_b^{(j)} \,. \tag{11}$$

The substitution of equation (11) into the diffusion distance (10) results in

$$\xi_m(q,r) = \sqrt{\sum_{j=1}^{N-1} \lambda_j^{2m} \left[\psi_q^{(j)} - \psi_r^{(j)}\right]^2} \,. \tag{12}$$

Because the eigenvalues are decreasing and the high order eigenvalues λ_j^{2m} vanish asymptotically fast with m, we can further approximate the diffusion distance

$$\xi_m(q,r) \approx \sqrt{\sum_{j=1}^{p_m} \lambda_j^{2m} \left[\psi_q^{(j)} - \psi_r^{(j)}\right]^2} \tag{13}$$

by choosing the first p_m eigen components.

An important implication of equation (13) is that there is a diffusion map $\mathcal{D}_m$ that maps a vertex q on the graph $\mathcal{G}$ to a vector $\mathbf{s}_q$ in the Euclidean space $\mathbb{R}^{p_m}$, i.e.,

$$\mathcal{D}_m : q \mapsto \mathbf{s}_q = \left[\lambda_1^m \psi_q^{(1)}, \lambda_2^m \psi_q^{(2)}, \cdots, \lambda_{(p_m)}^m \psi_q^{(p_m)}\right]^T \,. \tag{14}$$

The insight of the diffusion map (14) is the following. When computing the Euclidean distance between $\mathbf{s}_q$ and $\mathbf{s}_r$, we obtain the same formulation as the diffusion distance in equation (13). In other words, the diffusion map $\mathcal{D}_m$ is an isometry. Our goal of comparing the similarities between two probability distributions $(P^m)_{q-}$ and $(P^m)_{r-}$ under the m-step diffusion reduces to the evaluation of the Euclidean distance between $\mathbf{s}_q$ and $\mathbf{s}_r$. This simplifies significantly the task of designing the classifier since it is much simpler to design it in the Euclidean space $\mathbb{R}^{p_m}$ where we can easily find a surface that acts as a classification function for normality and abnormality.

For practical implementation, we need to determine the number m of diffusion steps and the number p_m of eigenvectors to be used in the Euclidean representation (14). The number m should be large enough to reduce the number p_m of eigenvectors needed, because, as mentioned earlier, high order eigenvalues, λ_j^{2m} for $j > p_m$ in equation (13) vanish asymptotically. On the other hand, we can not set m too large. Note that the zeroth eigenvalue and eigenvectors do not play a role in the Euclidean representation (14). Since $\lambda_j < 1$ for $j = 1, \cdots, N-1$, a too large m makes all Euclidean representations $\mathbf{s}_q$ collapse to zero vectors that are useless for classification. We choose empirically the value of m. Given m, the value of p_m can be determined through a predefined threshold τ_λ, which should be a small number close to zero. We choose the first p_m eigenvectors determined by

$$\lambda_{p_m}^{2m} \geq \tau_\lambda \quad \text{and} \quad \lambda_{p_m+1}^{2m} < \tau_\lambda \,. \tag{15}$$

2.3. Classification

The diffusion map $\mathcal{D}_m$ introduced in (14) transforms the graph $\mathcal{G}$ of the myocardial motions into a set of points in the Euclidean space $\mathbb{R}^{p_m}$. Each motion $\mathbf{u}_q$ is now represented by a vector $\mathbf{s}_q \in \mathbb{R}^{p_m}$. We can perform the classification task on the diffusion map representations. Equation (14) explains that similar cardiac motions on the graph $\mathcal{G}$ have similar vector representations in $\mathbb{R}^{p_m}$. In other words, different classes of myocardial motions should be mapped farther away in $\mathbb{R}^{p_m}$. The task of classification is to find in $\mathbb{R}^{p_m}$ one subspace corresponding to the normal cardiac motions and a second subspace corresponding to the dysfunctional motions.

In $\mathbb{R}^{p_m}$, we derive the subspace of normal motions based on the prior knowledge provided by a human expert. Let the first c motions $\mathbf{u}_1, \cdots, \mathbf{u}_c$ be labeled as normal. Hence, the vectors $\mathbf{s}_1, \cdots, \mathbf{s}_c$ are known as normal. We perform eigenanalysis on these vectors and obtain the eigenvalues $\{a_i\}_{i=1}^{p_m}$ and eigenvectors $\{\mathbf{e}_i\}_{i=1}^{p_m}$. Described in terms of the eigenvectors $\mathbf{e}_i$, each vector $\mathbf{s}_q$ has a new representation $\tilde{\mathbf{s}}_q$,

$$\tilde{\mathbf{s}}_q = \mathbf{E}^T (\mathbf{s}_q - \bar{\mathbf{s}}_{\mathrm{L}}) \,, \tag{16}$$

where

$$\mathbf{E} = [\mathbf{e}_1, \cdots, \mathbf{e}_{p_m}] \tag{17}$$

is a matrix collecting the eigenvectors $\mathbf{e}_i$ and

$$\bar{\mathbf{s}}_{\mathrm{L}} = \frac{1}{c} \sum_{i}^{c} \mathbf{s}_i \tag{18}$$

is the mean of the labeled vectors $\mathbf{s}_1, \cdots, \mathbf{s}_c$. Step (16) means that we translate the origin of the Euclidean space to $\bar{\mathbf{s}}_{\mathrm{L}}$ and then rotate the space to align with the eigen basis $\mathbf{E}$.

Let $\tilde{z}_1, \cdots, \tilde{z}_{p_m}$ denote the Cartesian coordinates in the eigen basis system. We now define the classification function as

$$h(\tilde{\mathbf{z}}) = \frac{\tilde{z}_1^2}{a_1} + \frac{\tilde{z}_2^2}{a_2} + \cdots + \frac{\tilde{z}_{p_m}^2}{a_{p_m}} \,. \tag{19}$$

The relation $h(\tilde{\mathbf{z}}) \leq \tau_c$ defines a hyperellipsoid that encloses all the labeled vectors $\tilde{\mathbf{s}}_1, \cdots, \tilde{\mathbf{s}}_c$, where τ_c is a parameter scaling the size of the hyperellipsoid; namely,

$$\forall i = 1, \cdots, c, \quad h(\tilde{\mathbf{s}}_i) = \frac{\tilde{s}_{i,1}^2}{a_1} + \frac{\tilde{s}_{i,2}^2}{a_2} + \cdots + \frac{\tilde{s}_{i,p_m}^2}{a_{p_m}} \leq \tau_c \,. \tag{20}$$

The relation $h(\tilde{\mathbf{z}}) = \tau_c$ defines a hyperelliptic surface. We treat this hyperelliptic surface as the classification boundary between the regions of normal and dysfunctional motions. The classification rule is

$$\forall i \in I, \quad \begin{cases} \mathbf{u}_i \text{ is normal}, & \text{if } h(\tilde{\mathbf{s}}_i) \leq \tau_c \\ \mathbf{u}_i \text{ is dysfunctional}, & \text{if } h(\tilde{\mathbf{s}}_i) > \tau_c \end{cases} . \quad (21)$$

3. EXPERIMENTAL RESULTS

This section details the classification algorithm and presents classification results for abnormal cardiac motions in real MRI data. We use MATLAB® to implement the classifier.

Algorithm Parameters. There are various parameters used by the algorithm. They are chosen as follows.

- The true covariance matrices Σ_{ij} used in equation (1) are not known, so we have to estimate them from the data. We collect two 3×3 windows of motions centered at $\mathbf{u}_i$ and $\mathbf{u}_j$, respectively, to derive the covariance matrices Σ_{ij}.
- To determine the edge weights in equation (2), we set ϵ to 0.1 as suggested by Laflon [8].
- The number m of diffusion steps, see equation (14), is crucial in the framework of diffusion maps. The literature has not reported how to choose this parameter. We let m be the diameter of the graph. The graph diameter is the least number of steps needed for the two farthest vertices in the graph [4]. This choice of m allows all vertices to have a chance of reaching every other vertices in the graph in m steps.
- In the diffusion map representation, the number p_m of eigenvectors is chosen through a threshold τ_λ that bounds the vanishing eigenvalues λ_i^{2m}, see equation (15). We set $\tau_\lambda = 0.1$.
- Finally, the classification boundary between normal and abnormal motions is the hyperelliptic surface $h(\tilde{\mathbf{z}}) = \tau_c$. Since the labeled normal motions must satisfy $h(\tilde{\mathbf{z}}) \leq \tau_c$, see equation (20), we set the threshold τ_c to

$$\tau_c = \max\{h(\tilde{\mathbf{s}}_1), h(\tilde{\mathbf{s}}_2), \cdots, h(\tilde{\mathbf{s}}_c)\} . \quad (22)$$

Application to Cardiac MRI Data. We apply the motion estimation algorithm [1] to a set of tagged MRI sequences and obtain dense motion maps. Figure 1 shows the motion map of a heart at the end-diastole. In this figure, the motions inside the two squares are labeled normal motions by an expert. We run the classifier on the motion map and obtain the results in Figure 2, where the dots denote the classified abnormal motions. To evaluate the results, a human expert manually classified the dysfunctional regions before running the algorithm.

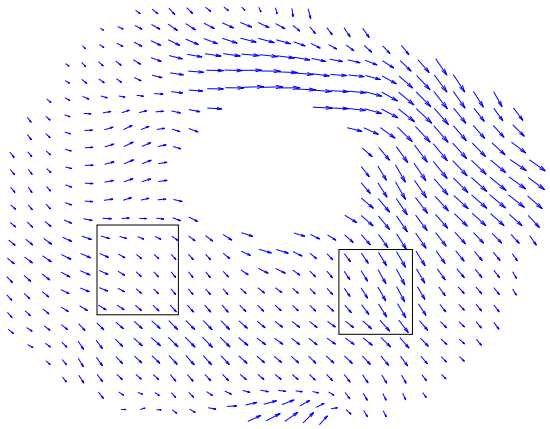


Fig. 1. A map of dense cardiac motions.

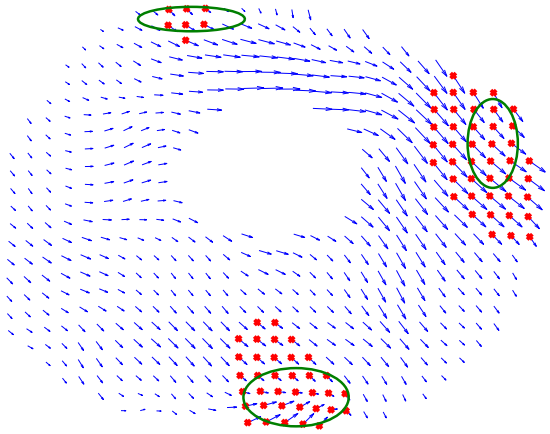
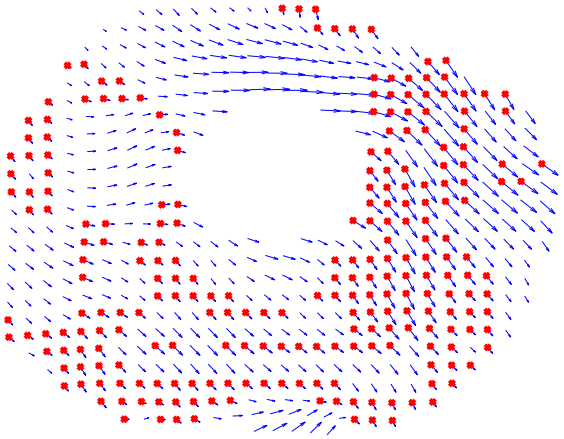


Fig. 2. Classification of abnormal motions using the classifier developed in this paper.

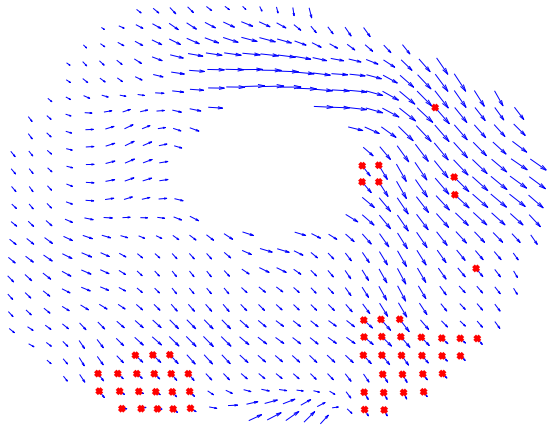
The manually determined abnormal regions are the three ellipses shown in Figure 2. The evaluation demonstrates that our classified dysfunctional regions are consistent with the manually classified results.

Comparisons with Other Approaches. A heuristic automatic method is to perform PCA on each motion vector. The PCA method for classifying a motion as abnormal is as follows: collect a window of neighboring motions, compute the covariance matrix to find out the principal motion in this neighborhood, and then determine the motion as abnormal if it deviates appreciably from the principal motion. The classification results using PCA are shown in Figure 3(a), where the classified abnormalities are randomly dispersed over the heart and are inconsistent with the manual ground truth. The main problem with the PCA method is that it ignores the myocardial motions' global geometry.

A second method to which we compare our algorithm is the classifier based on the spectral graph approach [2]. The



(a) PCA method.



(b) Semi-supervised spectral graph method when fed with normal motions only [2].

Fig. 3. Classification of abnormal motions using other approaches.

spectral graph based method relies on prior knowledge in both normal and abnormal motions. To compare fairly with our diffusion based approach, we feed the spectral graph algorithm with prelabeled normal motions only, leading to the classification results shown in Figure 3(b). We can easily see that the spectral graph approach misclassifies three regions of abnormalities.

4. SUMMARY AND CONCLUSIONS

This paper develops a semi-supervised algorithm to classify dysfunctional myocardial motions. Given a cardiac motion map, it is an easy task to determine prior normal motions. However, it is usually much harder to classify with high confidence abnormal motions. We develop in this paper a classifier that requires a small subset of normal motions.

We let a human expert label a small portion of the normal motions. The classifier automatically propagates this information to other unlabeled motions. The algorithm starts by describing the motion map with a weighted graph. Normalizing the weighted adjacency matrix of the graph results in a stochastic matrix. Associated with the stochastic matrix is a diffusion on the graph. The diffusion maps framework transforms the graph structure to a Euclidean space. In the Euclidean space, eigen-analysis on the labeled normal motions leads to a hyperelliptic surface that partitions the normal and abnormal cardiac motions. The experimental results show that the classifier in this paper performs better than other approaches: one based on PCA and the other based on spectral graph theory (when this latter classifier is given only labeled normal motions).

5. REFERENCES

[1] H.-H. Chang, J. M. F. Moura, Y. L. Wu, K. Sato, and C. Ho, "Reconstruction of 3-D dense cardiac motion from tagged MR sequences," in *Proceedings of IEEE International Symposium on Biomedical Imaging*, Arlington, VA, April 2004, pp. 880–883.

[2] H.-H. Chang, J. M. F. Moura, Y. L. Wu, and C. Ho, "Early detection of rejection in cardiac MRI: A spectral graph approach," in *Proceedings of IEEE International Symposium on Biomedical Imaging*, Arlington, VA, April 2006, pp. 113–116.

[3] M. Belkin and P. Niyogi, "Semi-supervised learning on Riemannian manifolds," *Machine Learning*, vol. 56, no. 1-3, pp. 209–239, June 2004.

[4] F. R. K. Chung, *Spectral Graph Theory*, vol. 92 of *CBMS Regional Conference Series in Mathematics*, American Mathematical Society, 1997.

[5] H.-H. Chang, J. M. F. Moura, Y. L. Wu, and C. Ho, "Automatic detection of regional heart rejection in USPIO-enhanced MRI," submitted to *IEEE Transactions on Medical Imaging*.

[6] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker, "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 21, pp. 7426–7431, May 2005.

[7] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, John Wiley & Sons, New York, NY, second edition, 2001.

[8] S. S. Lafon, *Diffusion Maps and Geometric Harmonics*, Ph.D. thesis, Yale University, New Haven, CT, May 2004.