Combination of Graph Theoretic Grouping and Time-Frequency Analysis for Image Segmentation

Rahele Kafieh[a], Hossein Rabbani[*a], Saeed Gazor[b]
[a] Biomed. Eng. Dept., Med. Image & Signal Proc. Research Center, Isfahan Univ. of Med. Sciences, Isfahan, Iran
[b] Electrical & Computer Engineering Department, Queen's University, Kingston, Canada

ABSTRACT

We introduce a nonparametric approach to multiscale segmentation of images using a hierarchical matrix analysis framework called diffusion wavelets. This approach benefits from the advantages of both graph theory and wavelet transform. Till now a broad range of multiscale transforms like wavelets (and other x-lets) have been introduced for image segmentation task. The graph theoretic formulation of grouping is also well-known to deal with this problem. The combination of multiscale transforms and graph based partitioning results in a scale-spectral method exploring through different scales of the image, over a great deal of spectral methods in graph partitioning. The method constructs multiscale basis functions and a series of dilation and orthogonalizations build a hierarchy, automatically. At each level, a set of basis functions is built by applying dyadic powers of a diffusion operator on the bases at the lower level. Two approaches are proposed for multiscale segmentation of images using diffusion wavelets. The first method is based on extended bases functions at each level and designing a competition between the bases value for partitioning. The second approach is defining a new distance for each level and clustering based on such distances.

Keywords: diffusion wavelets, segmentation, multiscale.

1. INTRODUCTION

Image segmentation is the dividing of an image into parts which are loosely connected; but an ideal segmentation is a meaningful organization of objects in a scene, having strong resemblance to partitioning by a human.

Two main approaches have been developed to achieve this goal: central grouping methods and pair wise grouping methods. Central grouping methods such as Gaussian Mixture Models based on EM [1] or K-means clustering, first assume a number of cluster prototypes and assign each pixel to one of them, according to mutual proximity in a defined space. Such methods are computationally efficient, but the underlying assumption behind this strategy, supposing that feature vectors of each cluster have a Gaussian distribution, is the most important drawback for these methods [2]. Moreover, the log likelihood can have more than one local minima and therefore in order to find a correct solution, multiple start points using iterative algorithms should be designed [3].

On the other hand, pair wise grouping methods make use of pair wise affinities between two pixels to measure the degree of similarity. The pixels can then be clustered based on the calculated affinities using methods like spectral methods [3-8], deterministic annealing [9] or stochastic clustering [10]. Pair wise grouping methods has the main disadvantage of high computational complexity as a result of the method's intrinsic requirement of comparing every possible pair of pixels in the image. This problem can be reduced by comparing only a set of points located in a determined distance from a selected pixel. The most striking advantage associated with pair wise grouping methods is that against the central grouping methods, all points in a cluster are not restricted to be located close to a prototype (similar to a Gaussian distribution) and the similarity propagates from a pixel to a neighboring pixel.

Spectral methods for image segmentation are established on calculation of eigenvectors and eigenvalues of an $N \times N$ affinity matrix (W), where N is the number of pixels in an image. In this procedure, a mapping from pixel space to a low dimensional subspace- build by the eigenvectors – makes a new embedding, capable of being meaningfully partitioned by central grouping methods such as Gaussian Mixture Models based on EM [1] or K-means.

One example of spectral graph theory is introduced by Shi and Malik [7] by finding eigenvector corresponding to the second smallest eigenvalue of the normalized laplacian L defined as:

$$L = D^{-1/2}(D-W)D^{-1/2} = I - D^{-1/2}W\,D^{-1/2} \quad (1)$$

where D is the diagonal matrix with entries $D_{ii} = d_i$, equal to sum of W 's i-th row. A two-class solution is provided by thresholding this eigenvector and extension to higher number of groups can be achieved by iterating the algorithm on each partition.

Another example of spectral graph theory was introduced by Ng [3], by finding $\Upsilon_1, \Upsilon_2, \Upsilon_3, \dots \Upsilon_k$, k largest eigenvectors of L (chosen to be orthogonal to each other in the case of repeated eigenvalues) and forming the matrix $\Upsilon = [\Upsilon_1, \Upsilon_2, \Upsilon_3, \dots \Upsilon_k]$ by stacking the eigenvectors in columns. The matrix Λ is then constructed by renormalizing each of Υ's row to have unit length and treating each row of Υ as a point in $\mathbb{R}^k$. A k-means clustering on new space is then applied to find the points pertaining to each of k clusters.

Diffusion maps are more recent example of spectral graph theory introduced by Coifman [11]. The method is based on finding m (not essentially equal to k (the number of clusters)) largest eigenvectors and corresponding eigenvalues of L. The diffusion map embeds each node i = 1, …, N into an m −dimensional Euclidean space (equation 2) named diffusion coordinates, Euclidean distance of which is proved to be diffusion distance between the nodes. Any kind of clustering like k-means may be applied in this new space. The application of diffusion maps in image segmentation is proposed by Kafieh[12] and more details on construction of affinity matrix and coarse grained approach [13] for k-means clustering and selection of a correct number of clusters is discussed in more detail in [12]. It is also important to know that Diffusion maps are mathematically proved to have a performance similar to Fourier transform on graphs and this approach provides the motivation to define different version of Fourier transform (like short time Fourier transform and wavelets) on graphs.

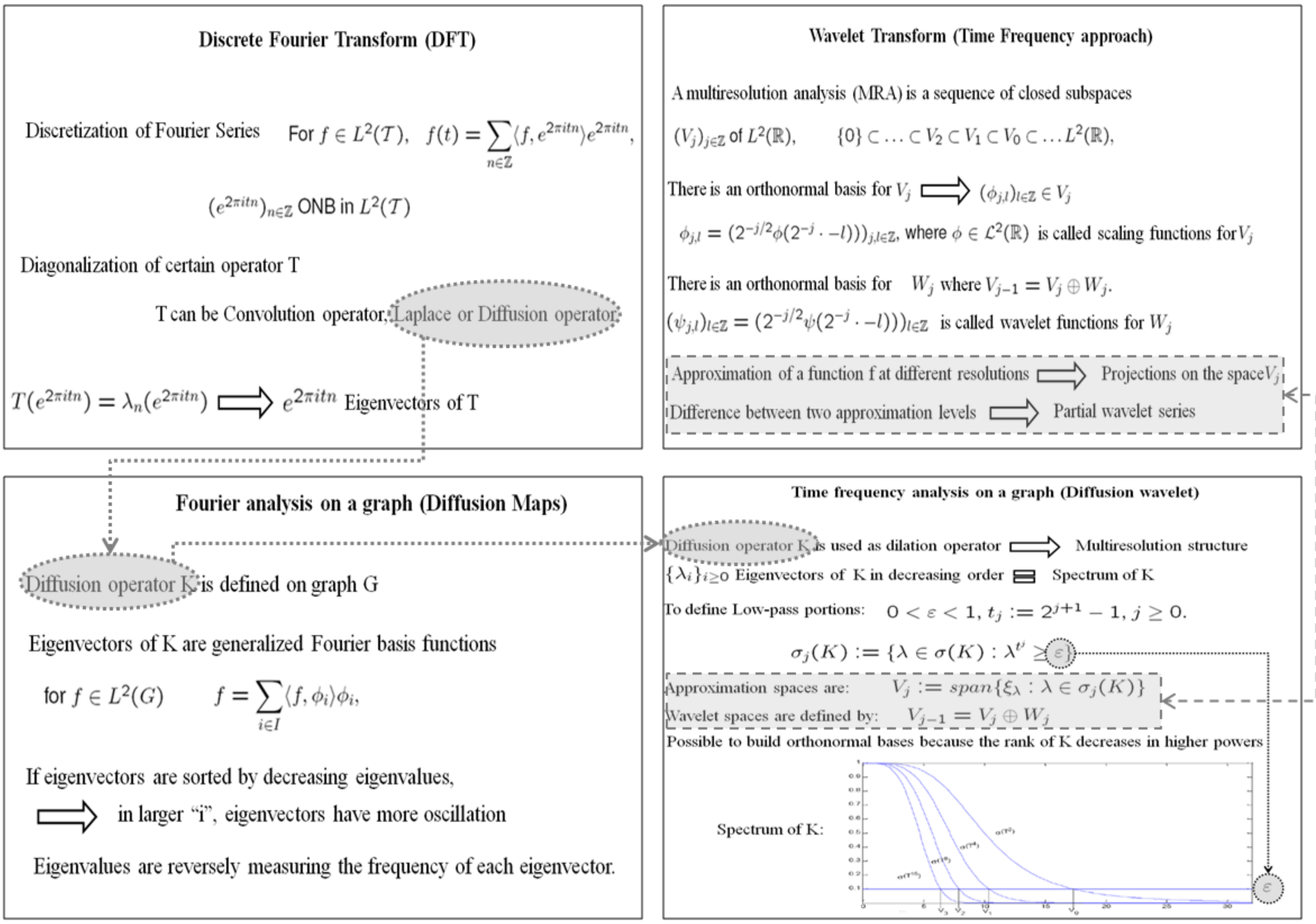

Figure 1. Schematic correlation between Fourier and wavelet transforms in continues spaces and diffusion maps and wavelets in discrete spaces.

$$i \rightarrow \Psi(i) = \begin{pmatrix} \lambda_1\psi_1(\mathrm{i}) \\ \lambda_2\psi_2(\mathrm{i}) \\ \vdots \\ \lambda_m\psi_m(\mathrm{i}) \end{pmatrix} \qquad (2)$$

The mentioned methods based on spectral graph theory are only dealing with spectral properties of the graph; however, as discussed above, with assuming diffusion maps as Fourier transform on graphs, the scale-spectral theories on graphs sound appealing. Diffusion wavelets introduced by Coifman [14] can be considered as a perfect specimen of such the scale-spectral theories, application of which in image segmentation is discussed in this paper. Figure 1 shows a schematic correlation between Fourier and wavelet transforms in continues spaces and diffusion maps and wavelets in discrete spaces.

2. DIFFUSION WAVELETS

In this paper, we present a new diffusion model based approach that finds multiscale embeddings of pixels in an image. Our approach is based on work of Coifman [14] in harmonic analysis. Harmonic analysis is commonly referred to Fourier analysis in continuous spaces, but wavelet methods can also classified as new versions of harmonic analysis, dealing with both temporal and spatial properties of the signals. Coifman [14] introduced diffusion wavelets as a recent extension of wavelets to perform on discrete spaces like graphs. In this new approach, the basis functions are constructed in each level and despite traditional wavelets, they are not predefined [15].

2.1 Construction of the affinity and normalized Laplacian matrixes

We focus on gray-level images (which can be simply extended to colored ones). In order to apply the diffusion wavelets to an image, graph nodes must be associated with the image pixels. To reduce the complexity, we select 5×5 pixel boxes as graph nodes and a kernel is defined as:

$$w(x,y) = \exp - \left(\left(\frac{d^2(x,y)}{2{\sigma_{geo}}^2} \right) + \left(\frac{d^2(g(x),g(y))}{2{\sigma_{gray}}^2} \right) \right) \qquad (3)$$

where x, y indicate the centroids of selected 5×5 boxes (we call them pixel nodes), $g(.)$ is the mean gray level of each box, and σ_{geo} and σ_{gray} point out the scale factor (calculated as 0.15 times the range of $d(x,y)$ and $d(g(x),g(y))$, respectively). The affinity matrix (W) is a $N \times N$ matrix, where N is the number of pixels in the image.

The normalized pixel-pixel matrix T is built by:

$$T = D^{-1/2} W D^{-1/2}, \qquad (4)$$

and the normalized laplacian matrix L is also constructed by:

$$L = I - T \qquad (5)$$

which is the same as the matrix formulated in (1).

Spectral graph theory is usually looking for eigenvectors of T, while multiscale diffusion analysis is trying to find scaling functions of T using diffusion wavelets [14]. It is similar to projecting data to lower dimensional spaces by using scaling functions, while keeping the large scale information [15]. The important property of the method is its ability to detect the geometric structure of data at different scales to provide a multiscale embedding. The main assumption on this method is that T is local, i.e. it has a small support and that high powers of T have low dimensional rank. Figure 1 (the section titled "Diffusion wavelet") describes how the spectral powers of T relate with the multiscale eigen-space decomposition. Here the rank of T decreases when increasing the powers of T. From the analyst's perspective, high powers are smooth functions with small gradient, hence they are compressible, leading to data reduction. Table 1 shows a comparison between continuous and discrete transforms both in multiscale and single-scale modalities.

2.2 Diffusion wavelets algorithm

Some notations should be defined for this algorithm: $[\mathrm{T}]_{\phi_a}^{\phi_b}$ is used to indicate the matrix representing the linear operator T with respect to the basis ϕ_a in the domain (row space or input space) and ϕ_b in the range (column space). A set of vectors ϕ_a represented on a basis ϕ_b will be written in matrix form $[\phi_a]_{\phi_b}$ (basis ϕ_b written on the basis ϕ_a[14].The subscripts in ϕ_a and ϕ_b denote the scale.

A pseudo-code is presented for construction of a diffusion wavelet tree:

$\{\phi_j, \psi_j\} = DWT\left([\mathrm{T}]_{\phi_0}^{\phi_0}, \phi_0 = I, QR, J, \varepsilon\right)$,

$//\phi_j$: Scaling basis functions at scale j.

$//\psi_j$: Wavelet basis functions at scale j.

//QR : A function computing a sparse QR decomposition.

$//J$: Maximum number of steps to compute.

$//\varepsilon$: Precision.

For $j = 0\ to\ J-1$ {

$\left([\phi_{j+1}]_{\phi_j}, [\mathrm{T}^{2^j}]_{\phi_j}^{\phi_{j+1}}\right) \leftarrow QR\left([\mathrm{T}^{2^j}]_{\phi_j}^{\phi_j}, \varepsilon\right)$;

$[\mathrm{T}^{2^{j+1}}]_{\phi_{j+1}}^{\phi_{j+1}} = \left([\mathrm{T}^{2^j}]_{\phi_j}^{\phi_{j+1}}\right)\left([\mathrm{T}^{2^j}]_{\phi_j}^{\phi_{j+1}}\right)^*$;

$[\psi_j]_{\phi_j} \leftarrow QR\left(\mathrm{I} < \phi_j > -[\phi_{j+1}]_{\phi_j}[\phi_{j+1}]_{\phi_j}^{\ *}, \varepsilon\right)$; }

A multiresolution decomposition of the functions on the graph is a family of nested subspaces $V_0 \supseteq V_1 \supseteq V_2 \supseteq \ldots V_j \supseteq \cdots$ spanned by orthogonal bases of diffusion scaling function ϕ_j. If T^t is an operator on functions on the graph G, then the subspace V_j is defined as the numerical range up to the precision ε of $T^{2^{j+1}-1}$ and the scaling functions are smooth bump functions with some oscillations, at a scale roughly 2^{j+1}. The orthogonal complement of subspace V_{j+1} into V_j is called W_j and is spanned by a family of orthogonal diffusion wavelets ψ_j , which are smooth and localized oscillatory functions at the same scale.

The DWT procedure can be described in two steps [15]:

1. Generating Diffusion Models:

$$\{\phi_j, \psi_j\} = DWT(T, I, QR, J, \varepsilon), \quad (6)$$

- I is an identity matrix; J is the max step number; ε is the desired presicion, QR is the sparse QR decomposition [14].
- ϕ_j: diffusion scaling function at level j. ψ_j: wavelet function at level j.

2. Computing the extended basis functions:

- $[\phi_j]_{\phi_0}$, the representation of the basis functions at level j in the original space, is computed as follows:

$$[\phi_j]_{\phi_0} = [\phi_j]_{\phi_{j-1}}[\phi_{j-1}]_{\phi_{j-2}} \cdots [\phi_1]_{\phi_0}[\phi_0]_{\phi_0}. \quad (7)$$

- $[\phi_j]_{\phi_0}$ is an $N \times N_j$ matrix.

At any scale like i, N_i basis functions with length of l_i exist, $[\mathrm{T}]_{\phi_a}^{\phi_b}$ is a $N_b \times l_a$ and $[\phi_b]_{\phi_a}$ is a $l_a \times N_b$ matrix. $[\phi_j]_{\phi_{j-1}}$ plays the role of mapping between data at coarse and fine levels and the basis function at level j can be represented in terms of basis functions at the lower levels.

Table 2. A comparison between some continuous and discrete transforms both in multiscale and single-scale modalities.

	Space (continuous : C, Discrete : D)	parameters	Introduced in	By
Fourier Transform	C	Frequency	1948	G. Campbell;R. Foster
Cosine Transform	C	Duration-Translation	1974	N. Ahmed;T. Natarajan and K. R. Rao,
Wavelet Transform	C	Scale-Translation	1988	I. Daubechies
Wavelet packet Transform	C	Scale-Translation- Frequency	1992	R.R. Coifman
Geometrical X-lets (wedgelet)	C	Scale-Translation- Rotation	1997	DL Donoho
Complex Wavelet Transform	C	Scale-Translation- Rotation	1999	N. G. Kingsbury
Geometrical X-lets (curvelet)	C	Scale-Translation- Rotation	2000	E. Candès and D. Donoho
Geometrical X-lets (bandlet)	C	Scale-Translation- Rotation	2004	Le Pennec; E., Mallat
Geometrical X-lets (contourlet)	C	Scale-Translation- Rotation	2005	M. N. Do and M. Vetterli,
Diffusion Maps	D	Spectral	2006	R.R. Coifman and S. Lafon
Diffusion wavelet	D	Scale-Spectral	2006	R. R. Coifman and M. Maggioni

3. APPLICATION OF DIFFUSION WAVELETS IN IMAGE SEGMENTATION

Two approaches are proposed in this paper for image segmentation. The first method is based on extended bases functions at each level and designing a competition between the bases value for partitioning. The second approach is defining a new distance for each level and clustering based on such distances. These two methodologies are discussed below.

3.1 Image segmentation based on extended bases functions at each level

As discussed above, in diffusion maps, the diffusion distances were used for feature representation. We also use this approach for feature representation based on the extended bases functions, at level j. Let us consider extended bases functions ϕ_j at level j as

$$\phi_j = \begin{pmatrix} \phi_{j1}(x_1) & \dots & \phi_{jm}(x_1) \\ \vdots & \ddots & \vdots \\ \phi_{j1}(x_n) & \dots & \phi_{jm}(x_n) \end{pmatrix} \tag{8}$$

Where, n is the number of pixel nodes in image and m is the number of extended bases functions at level j. For each arbitrary level j, the number of extended bases functions (m) will differ and the final level will have one extended bases functions ($m = 1$). For instance, in a image with n =800 pixel nodes, we may have $m = 7$ at level $j = 10$. Therefore, ϕ_j will have a size of 800×7 at level $j = 10$. We can consider that the best number of clusters at level $j = 10$ is $m = 7$ and we can design a competition between the bases value for partitioning. For this purpose, x_i (the i-th pixel node) belongs to cluster F which is obtained by:

$$F = \mathrm{argmax}_{f=1:(m=7)} \left(\phi_{10f}(x_i)\right), \tag{9}$$

To clear up, we find the maximum value in each row of ϕ_j and the column, at which the maximum is localized, reveals the correct cluster to which the pixel node corresponding to that row is belonged. To determine the level j at which the extended bases functions ϕ_j should be calculated, we determine the best number of clusters for our image and find level j with nearest number of extended bases functions (m) to the determined cluster number.

3.2 Image segmentation based on defining a new distance for each level

The second approach is based on defining new distances between each pair of pixel nodes and applying a central grouping method like k-means according to these new distances. This approach is similar to diffusion maps and the most important difference is new distance functions pertaining to each level and consequently, leading to different partitioning results at each level. The distance between pixel nodes x_i and x_k at level j is defined based on extended bases functions (ϕ_j) at each level:

$$D_{\phi_j}(x_i, x_k) = \sqrt{\sum_{l=1}^{m} \left(\phi_{jl}(x_i) - \phi_{jl}(x_k)\right)^2}. \tag{10}$$

We may consider the whole extended bases distance matrix at level j, which is given by the following:

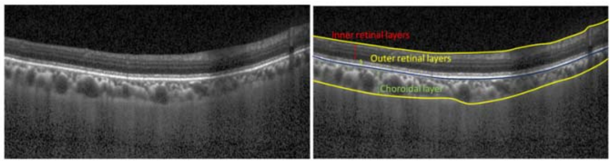

Figure 1 An example of EDI OCT segmentation by diffusion wavelets.

$$D_{\phi_j} = \begin{pmatrix} D_{\phi_j}(x_1, x_1) & \dots & D_{\phi_j}(x_1, x_n) \\ \vdots & \ddots & \vdots \\ D_{\phi_j}(x_n, x_1) & \dots & D_{\phi_j}(x_n, x_n) \end{pmatrix} \tag{11}$$

Now, we can run a method like k-means clustering between the pixel nodes and the only difference with traditional k-means is using distance as calculated in D_{ϕ_j} instead of simple Euclidean or mohanalobis distances.

3.3 *A sample application in segmentation of OCT*

The proposed methods are tested on both manually produced images and on OCT scans of normal subjects acquired using a Heidelberg scanner. The method shows high ability in segmentation of easy-doughnut and difficult-doughnut cases. The results seem to be promising, particularly due to scale dependency of segmentation results. The number of segmented clusters in each scale is different from other scales and in the case of knowing the correct number of classes, a particular scale can be chosen. An example of application in Enhanced Depth Imaging (EDI) OCTs is shown in Figure 2. The retinal structure containing inner, outer and choroidal layers of retina can be segmented apart from the background by applying diffusion wavelet in a scale that produces 3 clusters (yellow curves in Figure 2). The algorithm can be proceeded with a dynamic programming step to detect RPE-choroid interface (blue curve in Figure 2), so the retinal layers and coroidal layer can be segmented correctly.

4. RELATION TO PRIOR WORK

The prior works in relation to this work can be categorized to two classes. The first class is about applications of diffusion wavelets (combinations of graph based and time-frequency methods). No particular report is even published in application of this combinatory method in image segmentation, while this paper presents two novel strategies to be used in segmentation of images using this method.
On the other hand, the second class of the related prior works relates to algorithms on OCT image analysis which makes use of both graph based methods and time-frequency schemes in separate works; however, there is no report of joint application of these methodologies in any previous publication and this paper is the first from this point of view.

REFERENCES

[1] C. M. Bishop, *Neural networks for pattern recognition*: Oxford university press, 1995.

[2] C. Fowlkes, S. Belongie, F. Chung, and J. Malik, "Spectral grouping using the Nystrom method," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 26, pp. 214-225, 2004.

[3] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," *Advances in neural information processing systems,* vol. 2, pp. 849-856, 2002.

[4] P. Perona and W. Freeman, "A factorization approach to grouping," in *Computer Vision—ECCV'98*, ed: Springer, 1998, pp. 655-670.

[5] S. Sarkar and K. L. Boyer, "Quantitative measures of change based on feature organization: Eigenvalues and eigenvectors," in *Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on*, 1996, pp. 478-483.

[6] G. L. Scott and H. C. Longuet-Higgins, "An algorithm for associating the features of two images," *Proceedings of the Royal Society of London. Series B: Biological Sciences,* vol. 244, pp. 21-26, 1991.

[7] J. Shi and J. Malik, "Normalized cuts and image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 22, pp. 888-905, 2000.

[8] Y. Weiss, "Segmentation using eigenvectors: a unifying view," in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, 1999, pp. 975-982.

[9] J. Puzicha, T. Hofmann, and J. M. Buhmann, "Non-parametric similarity measures for unsupervised texture segmentation and image retrieval," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, 1997, pp. 267-272.

[10] Y. Gdalyahu, D. Weinshall, and M. Werman, "Stochastic image segmentation by typical cuts," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, 1999.

[11] R. R. Coifman and S. Lafon, "Diffusion maps," *Applied and computational harmonic analysis,* vol. 21, pp. 5-30, 2006.

[12] R. Kafieh, H. Rabbani, M. D. Abramoff, and M. Sonka, "Intra-Retinal Layer Segmentation of 3D Optical Coherence Tomography Using Coarse Grained Diffusion Map," *Medical Image Analysis,* 2013.

[13] S. Lafon and A. B. Lee, "Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 28, pp. 1393-1403, 2006.

[14] R. R. Coifman and M. Maggioni, "Diffusion wavelets," *Applied and Computational Harmonic Analysis,* vol. 21, pp. 53-94, 2006.

[15] C. Wang and S. Mahadevan, "Multiscale analysis of document corpora based on diffusion models," in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2009, pp. 1592-1597.