Multiscale Analysis and Diffusion Semigroups With Applications

Karamatou Yacoubou Djima
Advisor: Wojciech Czaja

Norbert Wiener Center
Department of Mathematics
University of Maryland, College Park
http://www.norbertwiener.umd.edu

April 7, 2015

Norbert Wiener Center for Harmonic Analysis and Applications

Outline

Motivation

- Availability of increasingly large data sets.
- Possible useful properties:
  - Intrinsic low-dimensionality,
  - Multiscale behavior.
- General strategy:
  - Representation systems analogous to harmonic analysis tools on $\mathbb{R}^n$.
  - Efficient representations from data-dependent operators.

Data deluge (The Economist)

Wavelets and Multiresolution Analysis (MRA)

- Family of dilations and $\mathbb{Z}^n$-translations of one or several functions.
- Used for approximation of L^2-functions on subsets of $\mathbb{R}^n$ at different resolutions.

Definition (S. Mallat, Y. Meyer, 1986)

A sequence of closed subspaces $\{V_j\}_{j\in\mathbb{Z}}$ of $L^2(\mathbb{R})$ together with a function ϕ is a **multiresolution analysis (MRA)** for $L^2(\mathbb{R})$ if

(i) $\cdots V_{-1} \subset V_0 \subset V_1 \cdots$,

(ii) $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$ and $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$,

(iii) $f \in V_j \iff f(2x) \in V_{j+1}$,

(iv) $f \in V_0 \implies f(x-k) \in V_0$, for all $k \in \mathbb{Z}$,

(v) $\{\phi(x-k)\}_{k\in\mathbb{Z}}$ is an orthonormal basis for V_0.

Wavelets and Multiresolution Analysis (MRA)

- Wavelets spaces W_j are orthogonal complements of V_j in V_{j+1} and

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j.$$

- Advantage: fast pyramidal schemes in numerical computation
- Wavelets perform well in image processing applications

Composite (Directional) Wavelets

- Extensions of traditional wavelets (K. Guo et al, 2006).
- Affine systems of the type

$$\{D_A D_B T_k \psi(x)\}_{x \in \mathbb{R}^n} ,$$

T_k: translation operators, $k \in \mathbb{Z}^n$,
D_A, D_B: dilation operators, $A, B \in GL_n(\mathbb{R})$.

- Examples: Contourlets (M. Do, M. Vetterli, 2002), Curvelets (E. Candes et al. 2003), Shearlets (D. Labate et al. 2005): basis elements with various orientations, elongated shapes with different aspect ratios.

Goal

Construct representations analogous of composite wavelets on graphs and manifolds.

Representation using Frames

Definition (Frame)

A countable family of elements $\{f_k\}_{k=1}^{\infty}$ in a Hilbert space $\mathcal{H}$ is a **frame** for $\mathcal{H}$ if for each $f \in \mathcal{H}$ there exist constants $C_L,\ C_U > 0$ such that

$$C_L \|f\|^2 \leq \sum_{k=1}^{\infty} |\langle f,\ f_k \rangle|^2 \leq C_U \|f\|^2.$$

- Overcomplete set of functions that span an inner product space.
- Generalization of orthonormal bases.
- Redundancy can yield robust representation of vectors or functions.
- No independence and orthogonality restrictions $\Longrightarrow$ varied characteristics that can be custom-made for a problem.

Kernel-Based Methods

- Efficient representations from data-dependent operators.
- Data set $\mathcal{X} = \{x_1, \cdots x_N\}$, $x_i \in \mathbb{R}^D$, D large
- Algorithm
  1) Represent the data as a graph
  2) Design kernel that captures similarity between points on graph
  3) Define graph operator based on kernel
  4) Recover underlying data manifold in terms of most significant eigenvectors of graph operator
- Examples: Kernel PCA (B. Schlkopf et al. 1999), Laplacian Eigenmaps (M. Belkin, P. Niyogi, 2002), Diffusion Maps (R. Coifman, S. Lafon, 2006)...

Goal (Updated)

Construct frame systems analogous to composite systems with dilations on graphs and manifolds and in the family of kernel-based methods.

Thesis Contribution

- Frame MRA
  - Established sufficient conditions to obtain frame MRA with composite dilations.
  - Constructed an example of an “approximate” MRA.
- Composite Diffusion Frames
  - Diffusion Frames MRA/Wavelet Frames.
  - Diffusion Frames MRA with composite dilations.
- Laplacian eigenmaps applied to retinal images
  - LE for dimension reduction and enhancement of eye anomalies.
  - OMF/VMF for classification methods.

General Idea of Construction

- Define abstract space that encompasses Euclidean spaces, graph, manifolds.
- Define families of operators that encompass many graphs/manifolds operators.
- Define MRA based on eigenfunctions of these operators.
- Construct frames with composite dilations that spans the MRA subspaces.

Spaces of Homogeneous Type

- Quasi-metric d on X with quasi-triangle inequality
$$d(x,\, y) \leq A\left(d(x,\, z) + d(z,\, y)\right),\, \forall\, x,\, y,\, z \in X,\, A > 0.$$
- $B_\delta\left(x\right) = \{y \in X :\, d(x,\, y) < \delta\}$ is open ball of radius δ around x.

Definition (Spaces of Homogeneous Type)

A quasi-metric measure space $(X,\, d,\, \mu)$ with μ, a nonnegative measure, is said to be of **homogeneous type** if for all $x \in X$ and all $\delta > 0$ and there exists a constant $C > 0$ such that
$$\mu\left(B_{2\delta}\left(x\right)\right) \leq C\mu\left(B_\delta\left(x\right)\right).$$

- $\mathbb{R}^n$, with Euclidean metric and Lebesgue measure.
- Finite graphs of bounded degree, with shortest path distance and counting measure.
- Compact Riemannian manifolds of bounded curvature with geodesic metric.

Symmetric Diffusion Semigroups

Definition (Symmetric Diffusion Semigroup – E. M. Stein, 1979)

A family of operators $\{S^t\}_{t\geq 0}$ is a **symmetric diffusion semigroup** on (X,μ) if

(a) Semigroup: $S^0 = I$, $S^{t_1}S^{t_2} = S^{t_1+t_2}$, $\lim\limits_{t\to 0^+} Sf = f \forall f \in L^2(X,\mu)$,

(b) Symmetry: S^t is self-adjoint for all t,

(c) Contraction: $\|S^t\|_p \leq 1$ for $1 \leq p \leq +\infty$,

(d) Positivity: for each smooth $f \geq 0$ in $L^2(X,\mu)$, $S^t f \geq 0$,

(e) Infinitesimal generator: $\{S^t\}_{t\geq 0}$ has negative self-adjoint generator A, so that

$$S^t = e^{At}.$$

Example of Symmetric Diffusion Semigroup

- X: weighted graph (V, E, W).
  - V: vertices, points in X.
  - E: edges, $x \sim y$ if x and y are connected.
  - W: matrix of positive weights w_{xy} if x and y are connected.
- Measure μ

$$\begin{aligned} d_x &= \sum_{x \sim y} w_{xy} \\ \mu(x) &= d_x. \end{aligned}$$

- Form diagonal matrix D with the d_x.
- The normalized Laplacian $L = I - D^{-1/2} W D^{-1/2}$ induces a symmetric diffusion semigroup on $L^2(X, \mu)$.

Compact & Differentiable Symmetric Diffusion Semigroup

Definition (Compact symmetric diffusion semigroup)

A symmetric diffusion semigroup $\{S^t\}_{t\geq 0}$ is called **compact** if S^t is compact for $0 < t < \infty$.

Definition (Differentiability)

Let $\{S^t\}_{t\geq 0}$ be a symmetric diffusion semigroup on $L^2(X,\mu)$. The semigroup $\{S^t\}_{t\geq 0}$ is called **differentiable** for $t > t_0$, if for every $f \in L^2(X,\mu)$, $S^t f$ is differentiable for $t > t_0$. $\{S^t\}_{t\geq 0}$ is called **differentiable** if it is differentiable for every $t > 0$.

Spectral Theorem

Proposition (A. Pazy)

Let $\{S^t\}_{t\geq 0}$ be a compact symmetric diffusion semigroup with self-adjoint generator A. Suppose that $\{S^t\}_{t\geq 0}$ is also differentiable.

- *If μ is an eigenvalue of A and ξ_μ is the corresponding eigenvector, then $\lambda^t = e^{\mu t}$ is an eigenvalue of S^t and $\xi_\lambda = \xi_\mu$ is the corresponding eigenvector.*
- *By the spectral theorem, $\{\xi_\lambda\}_{\lambda\in\sigma(S)}$ forms a countable orthonormal basis of $L^2(X,\mu)$ and for $f \in L^2(X,\mu)$, we can write*
$$S^t f = \sum_\lambda \lambda^t \langle f, \xi_\lambda\rangle \xi_\lambda.$$
- *By the contraction property, $\lambda \in [0, 1]$.*

Eigenfunctions as scaling functions

- Let $\{S^t\}_{t\geq 0}$ be a compact symmetric diffusion semigroup.
- View S as a dilation operator with spectrum $\sigma(S)$
- Consider a discretization of $\{S^t\}_{t\geq 0}$ at times t_j.

Definition (Multiresolution spaces V_j^ε)

Let $0 < \varepsilon < 1$ and let $\sigma_{\varepsilon,j}(S) := \{\lambda \in \sigma(S), \lambda^{t_j} \geq \varepsilon\}$. Define

$$\begin{aligned} V_{-1} &= L^2(X,\mu), \\ V_j^\varepsilon &= \langle \{\xi_\lambda : \lambda \in \sigma_{\varepsilon,j}(S)\} \rangle, \; j \geq 0. \end{aligned}$$

Multiresolution Analysis

- The spaces $\{V_j^\varepsilon\}_{j\geq -1}$ form a MRA in the sense that:
  - (i) $V_{j+1}^\varepsilon \subseteq V_j^\varepsilon$ for all $j \geq -1$.
  - (ii) $\{\xi_\lambda : \lambda \in \sigma_{\varepsilon,j}(S)\}$ is an orthonormal basis for V_j^ε.
  - (iii) $V_{-1} = L^2(X,\mu)$.
- For $j \geq -1$, define W_j^ε such that

$$V_j^\varepsilon = V_{j+1}^\varepsilon \oplus W_j^\varepsilon.$$

- $L^2(X,\mu) = \bigoplus_{j\geq -1} W_j^\varepsilon$ is a wavelet decomposition of $L^2(X,\mu)$.

Diffusion Wavelets (R. Coifman, M. Maggioni, 2006)

- Basis functions are typically non-localized.
- The computation of the eigenfunctions is required.
- Given
  - Space of homogeneous type X and bump functions Φ centered on dyadic cubes covering X, and which approximately span V_0^ε,
  - Compact symmetric semigroup $\left\{S^t\right\}_{t\geq 0}$.
- Apply a variant of Gram-Schmidt orthogonalization process to families of the form $\{S^{t_j}\Phi\}$.
- Subspaces obtained approximate $\left\{V_j^\varepsilon\right\}_{j\geq 0}$.
- Diffusion Wavelets: $V_j^\varepsilon = V_{j+1}^\varepsilon \oplus W_j^\varepsilon$.
- High computational cost and unstability for some examples.

* M. Christ proved the existence of dyadic cubes on spaces of homogeneous type

Composite Diffusion Frames

- We use the idea of approximating MRA of eigenfunctions.
- We construct frames instead of orthonormal wavelets and therefore, we will not need an orthogonalization process.
- We use a composition of dilations to obtain more versatile representations.

Assumptions for Composite Diffusion Frames

a) A bounded space of homogeneous type $(X,\, d,\, \mu)$.

b) A compact symmetric diffusion semigroup $\{S^t\}_{t\geq 0}$ on $L^2(X, \mu)$ with spectrum $\sigma(S)$ with eigenvalues λ and corresponding eigenvectors ξ_λ. Each λ has multiplicity 1.

c) Another compact symmetric diffusion semigroup $\{T^t\}_{t\geq 0}$ with spectrum $\sigma(T) = \sigma(S)$ and eigenvectors ζ_λ, such that, T and S are **similar**, i.e., for some invertible operator U,

$$\zeta_\lambda = U\xi_\lambda.$$

d) A fixed $\varepsilon \in (0,\, 1)$.

e) A frame $\Phi_0 = \{\phi_{0,k}\}_{k\in\mathcal{K}}$, with frame constants $C_L,\, C_U > 0$ such that $\langle \Phi_0 \rangle = V_0^\varepsilon$.

Construction of Composite Diffusion Frames

Theorem (W. Czaja, K.Y.D.)

Suppose that the assumptions (a)-(e) hold. Consider a discretization of $\{S^t\}_{t\geq 0}$ at times t_j and let V_j^ε be as defined earlier. Then, there exists a sequence of frames $\Phi = \{\Phi_j\}_{j=0,\ldots}$, with the following properties:

(i) $\langle \Phi_j \rangle = V_j^\varepsilon$ *and hence* $\langle \Phi_{j+1} \rangle \subseteq \langle \Phi_j \rangle$,

(ii) Φ_j *is a frame for* $\langle \Phi_j \rangle$.

Architecture of Proof

1. Define $T_{\varepsilon,i}$:

$$T_{\varepsilon,i} f = \sum_{\lambda \in \sigma_{\varepsilon,i}} \lambda^{t_i} \langle f, \zeta_\lambda \rangle \zeta_\lambda,$$

Properties of $T_{\varepsilon,i}$

(i) For each $i \geq 0$, $T_{\varepsilon,i} : V_i^\varepsilon \longrightarrow V_i^\varepsilon$ is a closed range, bounded operator. Moreover, V_i^ε is invariant under $T_{\varepsilon,i}$.

(ii) For each $i, j \geq 0$, and $i < j$, the operator $T_{\varepsilon,i} : V_i^\varepsilon \longrightarrow V_j^\varepsilon$ is a closed range, bounded operator from V_i^ε to V_j^ε.

(iii) For each $i \geq 0$, $T_{\varepsilon,i}$ is self-adjoint, and has a self-adjoint pseudo-inverse $T_{\varepsilon,i}^\dagger$.

(iv) For each $f \in V_i^\varepsilon$, we can write

$$T_{\varepsilon,i} f = T_{\varepsilon,i}^\dagger (T_{\varepsilon,i})^2 f.$$

Architecture of Proof

2. For $i \leq j$, let $\Psi_i = T_{\varepsilon,i}\Phi_0$ and show Ψ_i spans V_j^ε.
    - Use assumptions $\langle \Phi_0 \rangle = V_0^\varepsilon$.
    - Use 1(i) and invertibility of U.
3. For $i < j$, show Ψ is frame for each i
    - Use part 2 to show $f \in V_j^\varepsilon$ implies $f \in V_0^\varepsilon$.
    - For upper bound, use the fact that $T_{\varepsilon,i}$ is a contraction.
    - For lower bound, use the fact that $T_{\varepsilon,i}$ has closed, bounded pseudoinverse. Rewrite any function $f \in \langle \Phi_i \rangle$ in terms of $T_{\varepsilon,i}$ and its pseudoinverse and bound below.

Architecture of Proof

4. For $i < j$, Let $\Phi_j = \bigcup_i S_{\varepsilon,j}\Psi_i = \bigcup_i S_{\varepsilon,j}T_{\varepsilon,i}\Phi_0$, where

$$S_{\varepsilon,j}f = \sum_{\lambda\in\sigma_{\varepsilon,j}(S)} \lambda^{t_j} \langle f, \xi_\lambda\rangle \xi_\lambda.$$

Show Φ_j spans V_j^ε.

- Note that $S_{\varepsilon,j}$ has same properties as $T_{\varepsilon,i}$.
- Since Ψ_i spans V_j^ε, by 2(i), $S_{\varepsilon,j}\Psi_i$ spans V_j^ε.
- Since Φ_j is finite union of $S_{\varepsilon,j}\Psi_i$, Φ_j spans Vj.
- Since Φ_j is finite union of frames, it is a frame.

Potential Application

- Diffusion Shearlets (W. Czaja, K.Y.D., Upcoming)
  - Suppose that we have all assumptions of composite dilation theorem.
  - We are given an invertible, shear matrix U.
  - Form a family of operator T_s, "similar" to S with

$$\zeta_\lambda = U^s \xi_\lambda .$$

  - By composite dilation theorem, for $s < \infty$, $j \geq 0$,

$$\bigcup_i S_{\varepsilon,j} T_{\varepsilon,i,s} \Phi_0$$

    is a frame for V_j.

Age-related Macular Degeneration (AMD)

- Leading cause of blindness in elderly patients in industrialized nations.
- Earliest observable sign of retinal pigment epithelium (RPE) dysfunction, which causes AMD, is accumulation of irregularly shaped, color fundus deposits called drusen.

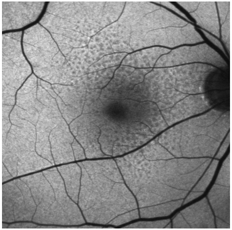
Age-related Macular Degeneration (AMD)

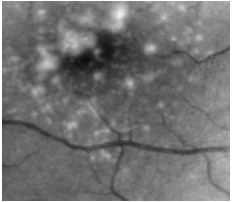
Growing interest in automated, analytic tools for:

- Early diagnosis.
- Tracking progression over time.
- Testing effectiveness of new treatment methods.

Common procedure:

- Retinal imaging.
- Feature extraction using segmentation/dimension reduction methods.
- Classification of features.



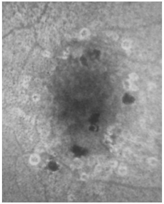


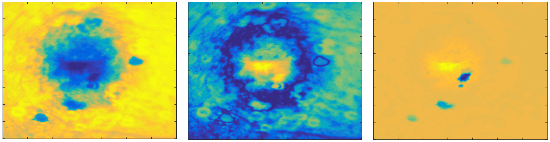
Retinal image (from NIH): drusen appear as bright spots.

LE-VMF

- **Laplacian Eigenmaps (LE)** is a nonlinear dimensionality reduction algorithm with locality preserving properties, which represent the data in the form of eigenimages, some of which accentuate the visibility of anomalies.
- **Vectorized Matched Filtering (VMF)** is a matched-filtering based algorithm that classifies anomalies across significant eigenimages simultaneously.

Laplacian Eigenimages





Eigenimages enhance appearance of structures in retinal images, pointing to possibility of distinct anomaly spectral signatures. In eigenimage 1 and 2, two classes of anomalies. In eigenimage 3, structures with large dark centers surrounded by a thin contour with white fluffy appearance.

Matched Filtering

- Used to detect the presence of specific elements in an unknown signal.
- A series of filters (templates) are correlated with an unknown signal.
- Templates yielding "high" correlation coefficients are retained as components of the signal.
- In 2-D, consider a filter or template image T of size $p \times q$. Compute the normalized cross-correlation NCC of the templates and each image, yielding a response matrix. For a pixel $I(x, y)$ in the image I

$$NCC(x, y) = \frac{\sum_i^p \sum_j^q I(x+i, y+j)T(i, j)}{\sqrt{\sum_i^p \sum_j^q |I(x+i, y+j)|^2}\sqrt{\sum_i^p \sum_j^q |T(i, j)|^2}}.$$

Vectorized Matched Filtering

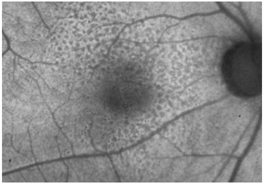
Vectorized Matched Filtering “averages” the response matrices accross the data:

- Given some templates, perform matched filtering for each image obtained via PCA or LE.
- Take the average of the absolute value of NCC matrices across all images for an individual.
- Apply a threshold to keep more significant correlations in the average response matrix. Set of all other correlations to 0.
- Identify anomalies for each individual using the average correlation matrix after threshold.

Image Preparation

Start with blue, green, and yellow AF images of the human retina.

- **Registration & Alignment.** We determine the overlap between images and produce a common coordinate system. After these processes, the images have same size and uniform overlap.
- **Vessel Mask.** We create a mask for the images to remove contribution from the blood vessels, as they are otherwise selected as anomaly by the detection algorithm.



Anomaly Detection

- **Feature Enhancement.** Perform LE on X.
- **Template Creation.** Rely on the following properties of anomalies:
  - circular or ellipsoid shapes, with different orientations, whose size can vary significantly across patients.
  - centers appear darker compared to other retinal surfaces.
- **Vectorized Matched Filtering.**
- **Thresholding.** Apply a threshold coefficient r to eliminate high correlation responses due to some background elements.
- **Detection.** We use the average response matrix after thresholding to identify the anomalies.

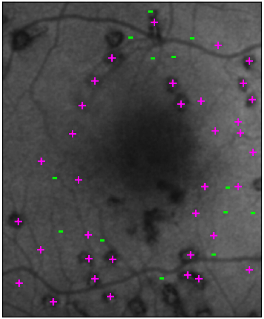
Evaluation of Method

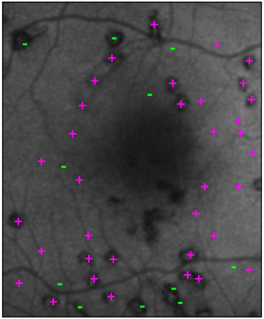
- Compare PCA and LE for feature enhancement.
  - Linear transformation that creates a new coordinate system for the data such that greatest variance by some projection of the data comes to lie on the first coordinate, second greatest variance on the second coordinate and so on.
  - Compute matrix $Q = XX^T$.
  - Perform eigendecomposition of Q:

    $$Q = P\Lambda P^T,$$

    where, Λ: matrix of eigenvalues λ_i, $i = 1, \ldots D$, P: matrix of orthonormal eigenvectors vi_i, $i = 1, \ldots D$.
  - Use eigenvectors as PCA eigenimages.
- Test method on various patients.
- Test method on nosiy data.

Comparison of Detections





Comparison of PCA and LE as anomaly enhancing schemes. For each patient, in both images the plus markers are the common detections. In the left image, the minuses mark anomalies detected by PCA-VMF and not detected by LE-VMF and vice versa in the right image.

Numerical comparison

The table below shows the number of correct detections, false positive (false detections) and true negatives (missed detections). We also give the rate of correct detections.

Table: Performances of VMF applied to PCA images versus LE images.

Type of Detections	PCA-VMF	LE-VMF
Correct	16	21
True Negative	6	1
False Positive	22	11
Rate of Correct anomaly Detection	73 %	95 %

Conclusion

- Summary
  - Composite Diffusion Frames as new method in harmonic analysis toolbox for graphs and manifolds.
  - Laplacian-Eigenmaps and Vectorized Matched Filtering for analysis of retinal images.
- Future
  - Diffusion Shearlets and application to vessel detection.
  - Retinal Imaging using various diffusion methods including composite diffusion frames.

Main References

1 K. Guo, D. Labate, D. W.-Q Lim, G. Weiss, E. Wilson, *with composite dilations and their MRA properties*, 2006.

2. R. Coifman and S. Lafon, *Diffusion Maps*. Applied and Computational Harmonic Analysis, 2006.

3. R. Coifman and M. Maggioni. *Diffusion Wavelets*. Applied and Computational Harmonic Analysis, 2006.

4. A. Pazy. *Semigroups of Linear Operators And Applications*, 1983.

5. M. Belkin and P. Niyogi P. *Laplacian eigenmaps for dimensionality reduction and data representation*. Neural Computation, 2003.

Thank you!