

Harmonic Analysis and Big Data: Data Dependent Representations

Wojciech Czaja

Institute for Mathematics and Its Applications 2015 PI Summer
Graduate Program
College Park, August 5–7, 2015



Introduction

- We have introduced several examples of data dependent representation methods, such as PCA or Laplacian Eigenmaps.
- These methods are well suited for the analysis of complex, noisy, high-dimensional data.
- As a drawback, we are going to encounter increased computational requirements in comparison to fast a priori methods, such as FFT or DWT.
- Also, the nonlinear dimensionality reduction methods such as LE or SE are typically non-invertible.
- We shall show today how to resolve some of these limitations.

Today we shall cover several important topics:

- The computational bottleneck of data dependent methods;
- Approximate preimage for kernel methods;
- Diffusion wavelets;
- Frames for real data applications.
- Data recovery based on LE.

Output normalized DR methods provide us with good approximate models for complex heterogeneous data structures. However, the computation of those approximations comes at a price.

- If D is the ambient dimension, and N is the number of points, time complexity of constructing an adjacency graph is $O(DN^2)$.
- What can we do about D ?
- What can we do about the exponent 2?
- What can we do about N ?
- What can we do about the computational complexity of eigendecomposition?

- 1 Data Compression via Incoherent Random Projections
- 2 Fast Approximate k Nearest Neighbors algorithms
- 3 Quantization Landmarking
- 4 Randomized low-rank SVD decompositions

1. Setting for data compression

- Dataset $\{x_1, x_2, \dots, x_N\}$ in \mathbb{R}^D , sampled from a compact K -dimensional Riemannian manifold
- Assume $\|x_i - x_j\| \leq A$ for all i, j and some $A > 0$
- Let $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_K$ be the first K nonzero eigenvalues computed by LE, assumed simple, with $r = \min_{i,j} |\lambda_i - \lambda_j|$, and let f_j be a normalized eigenvector corresponding to λ_j
- Use a random orthogonal projector Φ to map the points to \mathbb{R}^M . Let \hat{f}_j be the j th eigenvector computed by LE for the projected data set

1. Laplacian Eigenmaps with random projections

Theorem (with A. Halevy)

Fix $0 < \alpha < 1$ and $0 < \rho < 1$. If

$$M \geq \frac{4 - 2 \ln(1/\rho)}{\epsilon^2/200 + \epsilon^3/3000} K \ln(CKD/\epsilon), \text{ where } \epsilon = \frac{r\alpha}{4AN(N-1)},$$

then, with probability at least $1 - \rho$,

$$\|f_j - \hat{f}_j\| < \alpha.$$

The constant C depends on properties of the manifold. Precisely, $C = \frac{1900RV}{\tau^{1/3}}$, where R , V and $1/\tau$ are the geodesic covering regularity, volume, and condition number, respectively.

Similar results for other output normalized DR methods, e.g., for Schroedinger Eigenmaps.

1. Application: Classification of Hyperspectral Data



Table : Comparison of performance on Urban

Method	Time (min)	Accuracy (percent)
LE	15.26	79.05
LERP	11.78	78.44

HYDICE Scene from Copperas Cove, TX, courtesy of NGA

1. Application: Classification of Hyperspectral Data

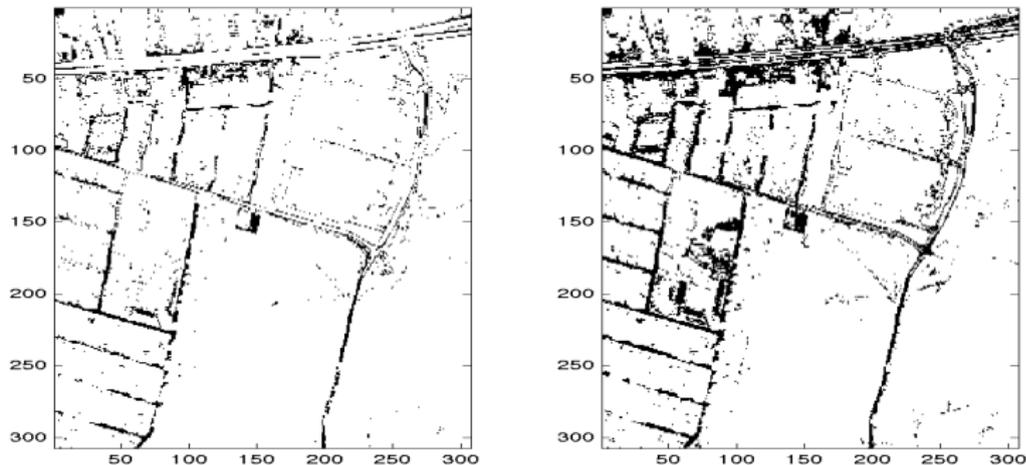
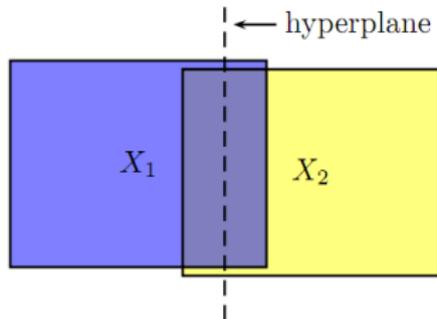


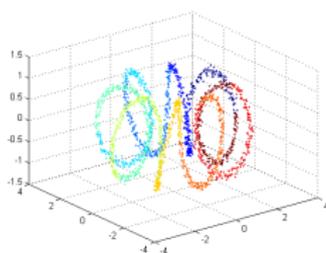
Figure : Urban class 2 (secondary road): left - LE, right - LERP

2. Fast Approximate k Nearest Neighbors

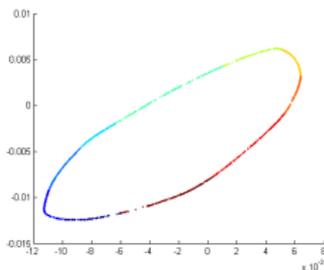
- There are many approximate nearest neighbor algorithms, e.g., Locality-sensitive Hashing (P. Indyk), Best Bin First (D. Lowe), or Clustered Point Sets Search (D. Mount). We present the Divide and Conquer method of Chen, Fang, and Saad
- Divide the set of points into two overlapping subsets using spectral bisection based on the Lanczos algorithm
- Once the size of a subset is less than a threshold r , compute using brute-force.
- If a point belongs to more than one subset, its nearest neighbors are selected from the neighbors in each of the subsets.



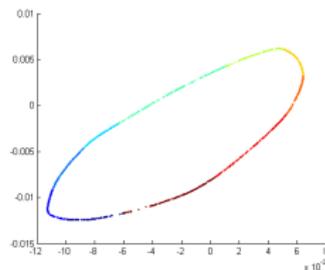
2. Numerical Experiments: Synthetic Data



(a) Helix



(b) Exact



(c) Approximate

Figure : Mapping a one-dimensional helix embedded in \mathbb{R}^3 . In the above example the exponent used is approx. 1.16 (depends on the size of overlap).

3. Landmarking

- Random sampling.
- Deterministic Landmarking via quantization techniques
 - PCM or Sigma-Delta quantization from analog-to-digital conversion
 - Vector quantization
- Hierarchical clustering with cluster centroids
- Supervoxel technology

- W. Sun, A. Halevy, J. J. Benedetto, W. Czaja, W. Li, C. Liu, B. Shi, and R. Wang, "Nonlinear dimensionality reduction via the enh-ltsa method for hyperspectral image classification," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 2, pp. 375 - 388, 2014.
- W. Sun, A. Halevy, J. J. Benedetto, W. Czaja, W. Li, C. Liu, B. Shi, and H. Wu, "UL-isomap based dimensionality reduction for hyperspectral imagery classification," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 89, no. 3, pp. 25 - 36, 2014.

4. Approximate SVD

- Rokhlin, Szlam and Tygert introduced a randomized, approximate SVD algorithm (ASVD) that works well when matrix is low rank.
- ASVD algorithm constructs an approximate SVD $U\Sigma V^T$ of a real $m \times n$ matrix A that satisfies

$$\|A - U\Sigma V^T\| \leq Cm^{1/(4i+2)}\sigma_{k+1}$$

with high probability.

- Since the normalized graph Laplacian is symmetric, we can use the ASVD algorithm to obtain the r largest eigenvalues. However, for the purposes of Laplacian Eigenmaps (LE), we need the smallest eigenvalues. This can be accomplished in the following manner. Let A be the normalized graph Laplacian. Use ASVD with $r = 1$ to compute the largest eigenvalue, σ_1 . Next, use ASVD again to find the largest r' eigenvalues of $2\sigma_1 I - A$. The smallest r' eigenvalues of A are given by subtracting the results of the previous step from $2\sigma_1$.

4. Robust Principle Component Analysis

- Consider PCA of data, with a fraction of the entries grossly corrupted due to, e.g., sensor malfunction on some measurements or random pixels occluded by irrelevant data.
- E. Candès (Stanford) introduced a version of PCA that eliminates such gross corruption via compressive sensing.
- Algorithm relies on using Singular Value Decomposition (SVD) which is computationally too expensive.
- It can be combined with V. Rokhlin's (Yale) randomized, approximate SVD algorithm that works well when matrix is low rank.

Speed up of Robust PCA (with A. Cloninger and G. Warnell)

Under certain assumptions on corrupted entries, Rokhlin's randomized SVD algorithm is used to speed up Candès PCA by **several orders of magnitude without loss of precision.**

Preimage Problem

- How to map data back from feature space to input space?
- The pre-image of $\psi \in \mathbb{R}^m$ is a point $x \in \mathbb{R}^d$ such that $\phi(x) = \psi$. Because x may not necessarily exist, this problem is ill-defined. Thus, instead, we look for $x \in \mathbb{R}^d$ such that $\phi(x)$ is “as close as possible” to ψ .
- Optimality criteria for “closeness” include, e.g.,

$$x = \arg \min_{x \in \mathbb{R}^d} \|\phi(x) - \psi\|_2, \quad (\text{Distance})$$

or

$$x = \arg \max_{x \in \mathbb{R}^d} \left\langle \frac{\phi(x)}{\|\phi(x)\|}, \frac{\psi}{\|\psi\|} \right\rangle. \quad (\text{Collinearity})$$

- Preimage problem introduced by Mika, Schoelkopf, and Kwok.

- Bakir et al. (2004) exploit labeling available for the unknown pre-image map, combined with a kernel regression technique.

G. Bakör, A. Zien, and K. Tsuda, Learning to find graph pre-images, in Pattern Recognition. Springer, 2004, pp. 253–261.

G. H. Bakör, J. Weston, and B. Schoelkopf, Learning to find pre-images, Advances in neural information processing systems, vol. 16, no. 7, pp. 449–456, 2004.

- F. Segonne et al. (2007-09) use diffusion maps as their embedding technique and seek to find a pre-image for the purposes of learning shape priors.

P. Etyngier, F. Segonne, and R. Keriven, Shape priors using manifold learning techniques, in International Conference on Computer Vision. IEEE, 2007, pp. 1–8.

- A non-iterative solution to the pre-image problem proposed by P. Honeine and C. Richard (2009), which improves the computational complexity of the original algorithms.

P. Honeine and C. Richard, Solving the pre-image problem in kernel machines: A direct method, in IEEE International Workshop on Machine Learning for Signal Processing. IEEE, 2009, pp. 1–6.

- Zheng et al. (2010) use weakly supervised penalty functions, in conjunction with the optimization function, to improve the pre-image learning process.

W.-S. Zheng, J. Lai, and P. C. Yuen, Penalized preimage learning in kernel principal component analysis, IEEE Transactions on Neural Networks, vol. 21, no. 4, pp. 551–570, 2010.

- R. Talmon, D. Kushnir, R. Coifman, I. Cohen, and S. Gannot (2012), use diffusion kernels with interpolation to learn the pre-image map.

R. Talmon, D. Kushnir, R. Coifman, I. Cohen, and S. Gannot, Parametrization of linear systems using diffusion kernels, IEEE Transactions on Signal Processing, vol. 60, no. 3, pp. 1159–1173, 2012.

- P. Arias, G. Randall, and G. Sapiro (2005) introduced Nystroem extension for Kernel PCA pre-image.

P. Arias, G. Randall, and G. Sapiro, Connecting the out-of-sample and pre-image problems in kernel methods, 2007 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8, Jun. 2007.

D. Kushnir, A. Haddad, and R. R. Coifman, Anisotropic diffusion on sub-manifolds with application to earth structure classification, Applied and Computational Harmonic Analysis, 2012.

Nystroem Extension for Kernel Methods

Given $x \neq x_1, \dots, x_N$, the Nystroem extension is an approximation of $\phi(x)$:

$$\hat{\phi}(x) = \sum_{i=1}^N K_{x,i} \phi(x_i),$$

where $K_{x,i}$ is a normalized extension of the original kernel:

$$K_{x,i} = \frac{k(x, x_i)}{\sqrt{\sum_{j=1}^N k(x, x_j) \sum_{j=1}^N k(x_i, x_j)}}.$$

Letting $K_x := [K_{x,i}]_{i=1, \dots, N}$, we have $\hat{\phi}(x) = V^* K_x$, for $V = [\phi_1, \dots, \phi_m]$.
Also, $\|K_x\|_0 = \#\mathcal{N}(x)$ for a nearest neighbor Gaussian kernel.

Preimage Problem with Nystroem Extension

Sapiro et al. proposed to modify the pre-image problem by using the Nystroem extension. Let $\mathcal{E} \in \mathbb{R}^{m \times N}$ be the Nystroem extension for a general dimension reduction scheme. Create a new objective function by solving:

$$\arg \min_{x \in \mathbb{R}^D} \|\mathcal{E}K_x - \psi\|_2.$$

They note that, due to normalization and since ψ is constant, minimizing $\|\mathcal{E}K_x - \psi\|_2$ is equivalent to maximizing the collinearity:

$$\|a - b\|_2^2 = 1 + \langle b, b \rangle - 2\langle a, b \rangle.$$

However, they also point out that for a generic approximation of a pre-image, one needs to assume that ψ is normalized, as well.

Preimage Problem with Nystroem Extension

Since our objective is now in the form of a linear least squares problem, the optimal K_x can be obtained by, e.g., calculating the Penrose-Moore pseudo-inverse:

$$\hat{K}_x = \mathcal{E}^\dagger \psi.$$

The map \mathcal{E} in our construction is unitary, so we can define the pseudo inverse as, e.g., the left inverse:

$$\mathcal{E}^\dagger = (\mathcal{E}^* \mathcal{E})^{-1} \mathcal{E}^*.$$

In the special case of kernel PCA, $K_{x,j} = e^{-\frac{\|x-x_j\|_2^2}{2\sigma^2}}$. This yields

$$\|x - x_j\|_2^2 = -2\sigma^2 \log(\hat{K}_{x,j}).$$

Finding x now reduces to a localization problem that is solved using MDS and the distances between x and $\mathcal{N}(x)$.

Preimage Problem for LE

- With A. Cloninger and T. Doster (2014).
- Laplacian Eigenmaps does not have a simple, closed form solution that relates $K_{x,i}$ to $\|x - x_i\|_2$, as each $K_{x,i}$ is a function of $\|x - x_l\|_2$ for all $x_l \in \mathcal{N}(x)$.
- Utilize the sparsity of LE construction, $\|K_x\|_0 = c \ll n$, to solve for $\|x - x_i\|_2^2$.
- Our approach:
 - Use pseudo-inverse or L^1 regularization to obtain estimates for K_x .
 - Modify the K_x term to correct for noise.
 - Use Newton's method or a constrained optimization to solve the resulting system of equations.
 - This solution leads to a localization problem.
 - Solve the localization problem using MDS.

Digit denoising

Digit	$\epsilon^2 = .2$		$\epsilon^2 = .4$		$\epsilon^2 = .6$	
	kPCA	LE w/ L_1	kPCA	LE w/ L_1	kPCA	LE w/ L_1
0	4.28	5.38	4.08	5.09	4.20	4.75
1	5.37	4.85	5.02	4.64	5.13	4.45
2	4.27	5.12	3.92	4.76	3.68	4.54
3	4.17	4.73	4.02	4.30	3.94	4.32
4	3.66	4.65	3.66	4.32	3.37	3.72
5	3.54	4.63	3.48	4.39	3.35	3.99
6	4.20	5.41	3.98	5.07	3.99	4.85
7	4.33	4.85	4.35	4.50	4.04	3.90
8	3.68	4.44	3.33	4.16	3.55	4.10
9	3.97	4.72	3.76	4.25	3.67	4.01
Avg.	4.15	4.88	3.96	4.55	3.89	4.26

Table : SNR for MNIST Dataset Comparing Pre-image Algorithms for kernel PCA, and Laplacian Eigenmaps with L_1 Regularization. White Gaussian noise added at several intensities, and the SNR is calculated over 10 samples of each digit by means of Liu, Tanaka and Okutomi algorithm.

Diffusion maps and diffusion wavelets

- Coifman and Lafon introduced the **diffusion distance** - a new and efficient metric measuring distances in graphs and point clouds:

$$D_t^2(x, y) = \sum_z \frac{(\rho(z, t|x) - \rho(z, t|y))^2}{\phi_0(z)},$$

where each $\rho(z, t|x)$ is the probability of transition in time t from z to x , and $\phi_0(z)$ is the stationary eigenvector of the probability transition matrix.

R. R. Coifman and S. Lafon, Diffusion maps, Applied and computational harmonic analysis 21 (2006) no. 1, pp. 5–30.

- Coifman and Maggioni used the diffusion distances to construct a new notion of wavelet representations of graphs, **diffusion wavelets**.

R. R. Coifman and M. Maggioni, "Diffusion wavelets", ACHA , 2006, Vol. 21(1), 53–94.

- Singer and Wu generalized this concept to **vector diffusion maps** (VDM).

A. Singer and H. Wu, Vector diffusion maps and the connection Laplacian, Communications on Pure and Applied Mathematics vol., 65 (2012) no. 8, pp. 1067–1144.

- Bremer, Coifman, Maggioni, and Szlam showed how to construct **diffusion wavelet packets**, generalizing the classical construction of wavelet packets of Coifman and Wickerhauser.

J. C. Bremer, R. R. Coifman, M. Maggioni, and A. D. Szlam, Diffusion wavelet packets, Applied and Computational Harmonic Analysis 21 (2006) no. 1, pp. 95–112.

Given is $\{\delta_n\}_{n=1}^N$, the standard orthonormal basis for $V_0 = \mathbb{R}^N$.

$$V_0 = \text{span} \{\delta_n : n = 1, \dots, N\}$$

\downarrow
 K dilation

$$V_1 = \text{span} \{K(\delta_n) : n = 1, \dots, N\}$$

$\{\Phi_n = K(\delta_n)\}_{n=1}^N$ is a frame for $V_1 = \mathbb{K}$.

We want to compute the eigenbasis $\{e_n\}_{n=1}^r$ for K . Recall that

$$V_1 = \mathbb{K} = \text{span} \{K(\delta_n) : n = 1, \dots, N\} \quad \text{and} \quad r = \dim V_1.$$

Hence, $\{e_n\}_{n=1}^r$ is an orthonormal basis for K .

Diffusion wavelets solution: Use a modified Gram- Schmidt orthogonalization algorithm on the frame $\{\Phi_n = K(\delta_n)\}_{n=1}^N$ to obtain an orthonormal basis $\{\Psi_n\}_{n=1}^r$ for \mathbb{K} .

Applications of diffusion maps in remote sensing

- Coifman and Hirn (2013) proposed to use diffusion maps and their embeddings for change detection in hyperspectral imagery over time.
- Diffusion geometry principles via tree structures with evolving graphs were proposed by Lee and Maggioni (2012).
- Evolving Riemannian manifold with changing diffusion processes were studied by H. Abdallah (2010).
- Shape analysis via heat kernels was proposed by F. Mémoli (2011).

- *Another solution:* Use the *frame minimization algorithm* on \mathbb{K} to obtain a FUNTF $\{\psi_n\}_{n=1}^r$. Since $r = \dim \mathbb{K}$, $\{\psi_n\}_{n=1}^r$ is an orthonormal basis for \mathbb{K} .
- *Additional advantage:* A FUNTF generalizes the concept of an orthonormal basis. We can obtain FUNTFs with s elements for any $s > r$. The truly redundant FUNTFs can be optimized for sparsity.

Motivation for frames originating in remote sensing

- Different classes of interest may not be orthogonal to each other; however, they may be captured by different frame elements. It is plausible that classes may correspond to elements in a frame but not elements in a basis.
- A *frame* generalizes the concept of an orthonormal basis. Frame elements are non-orthogonal.
- Frames provide over-complete data decompositions, often useful for numerical stability and noise reduction.

HYDICE Copperas Cove data set

- HYDICE Copperas Cove, TX, HSI dataset with 23 different classes.
- If the 23 classes were to correspond roughly to orthogonal subspaces, then one cannot achieve effective dimension reduction less than dimension $d = 23$.
- However, we could have a frame with 23 elements in a space of reduced dimension $d < 23$.



Figure : HYDICE Copperas Cove, TX — <http://www.tec.army.mil/Hypercube/>

Spectral signatures of selected classes

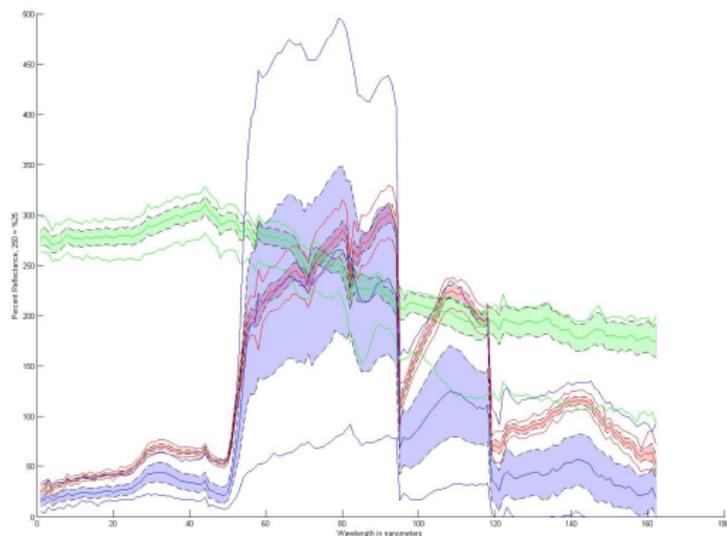


Figure : Spectral signatures of Walmart, pasture, and trees are not orthogonal to each other.

Comparison of spectral signatures

We focus on **Walmart**, **pasture**, and **trees**.

Angles between the mean signature of each class:

- **Walmart** and **pasture**: 36 degrees.
- **Walmart** and **trees**: 42 degrees.
- **Pasture** and **trees**: 14 degrees.

Maximum angles between classes:

- **Walmart** and **pasture**: 40 degrees.
- **Walmart** and **trees**: 49 degrees.
- **Pasture** and **trees**: 28 degrees.

Frame vs ONB representations in HSI

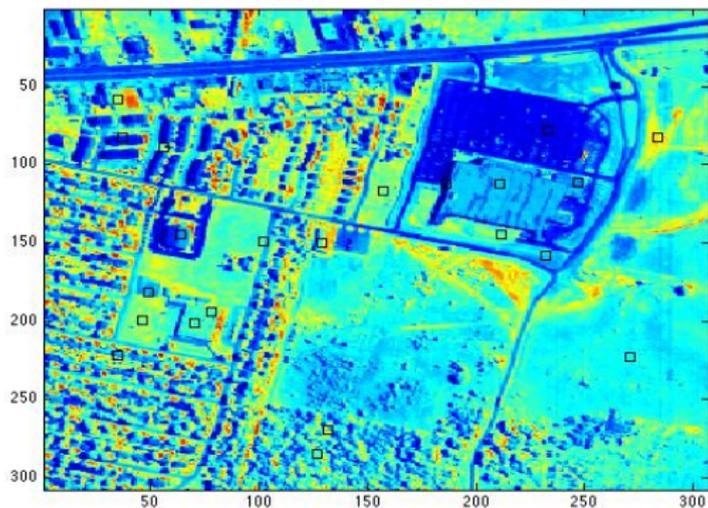


Figure : Frame vectors

ONB representations in HSI

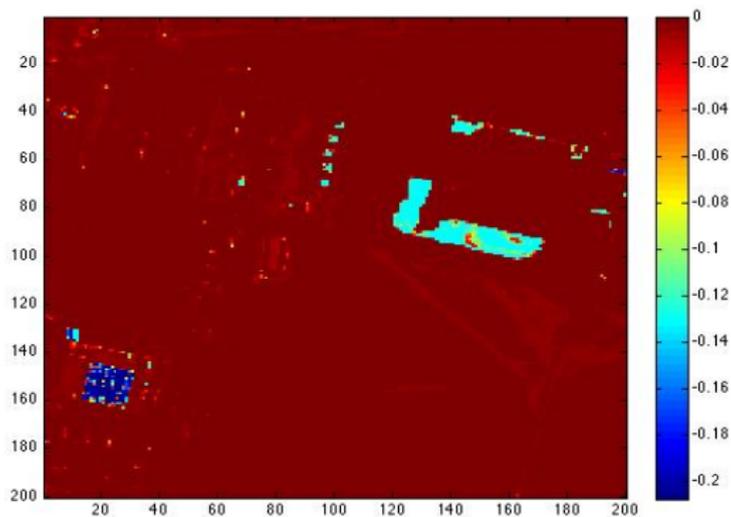
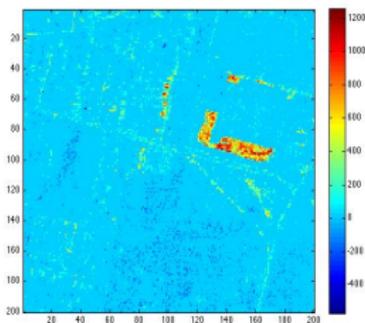
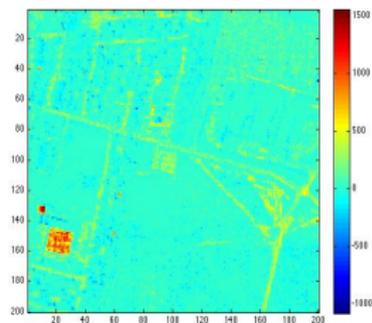


Figure : An element of an orthonormal Laplacian eigenmap basis

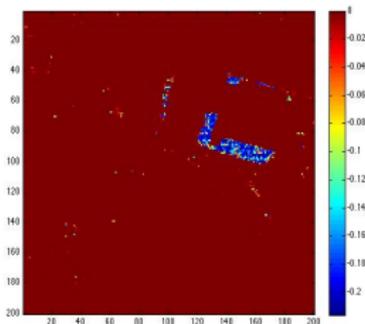
Frame representations in HSI



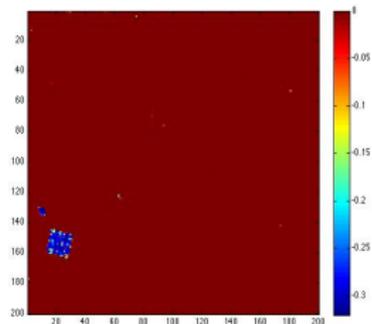
(a) Frame coefficients



(b) Frame coefficients



(c) LE on frame coeffs



(d) LE on frame coeffs

Frame coefficients provide a more suitable representation of the classes of objects which may be spectrally close to each other.

Frame optimization problems

Define the $W(\ell^1, \ell^\infty)$ Wiener amalgam penalty term

$$p_k = \sum_{m=1}^N |\langle y_m, \Psi_k \rangle|, \text{ for all } k = 1, \dots, s.$$

Our frame-theoretic method has two steps:

Step 1: Choose $q < s$ and find a pseudo-FUNTF $\Psi = \{\Psi_k\}_{k=1}^s$ by solving the following minimization problem:

$$\Psi = \arg \min_{\tilde{\Psi}} TFP(\tilde{\Psi}) + \frac{s^2}{d} \sum_{k=q+1}^s p_k. \quad (1)$$

- (1) is solved using a gradient descent method. The method is initialized with a $d \times s$ matrix with entries that are uniformly distributed on $[0, 1]$.

Frame optimization problems

Step 2: Solve for the minimum ℓ^1 norm coefficients of the frame Ψ :

$$\forall m = 1, \dots, N, \hat{a}_m = \arg \min_{\tilde{a}} \|\tilde{a}\|_1 \quad \text{such that} \quad y_m = \Psi \tilde{a}.$$

- ℓ^1 minimization is used instead of ℓ^0 minimization.
- Set $a_m = \hat{a}_m^+$.
- Select ℓ significant frame elements, $\Psi_{k_1}, \dots, \Psi_{k_\ell}$.
- The new coordinates for x_m are $(a_m[k_1], \dots, a_m[k_\ell])$.

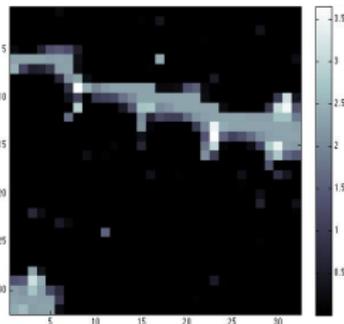
J. J. Benedetto, W. Czaja, and M. Ehler, Frame potential classification algorithm for retinal data, Springer Proceedings Series: Intern. Fed. for Medical and Biological Engineering, 26th Southern Biomedical Engineering Conference (2010).

J. J. Benedetto, W. Czaja, J. C. Flake, and M. Hirn, Frame based kernel methods for automatic classification in hyperspectral data, in: Proceedings of the IEEE 2009 International Geoscience and Remote Sensing Symposium, vol. 4, pp. 697–700, 2009.

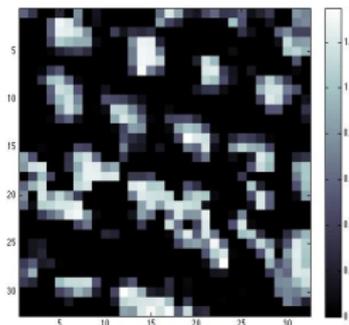
Frame coefficients



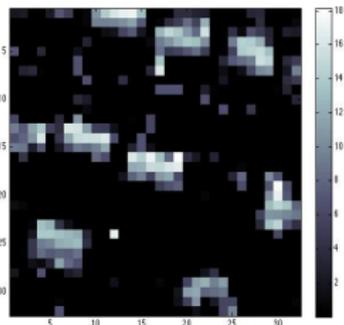
(a) Original



(b) Road coefficients

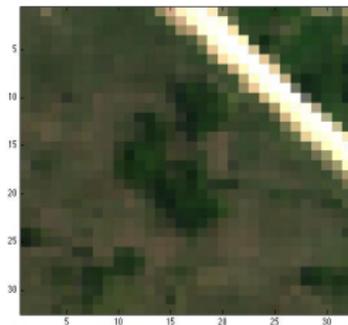


(c) Tree coefficients

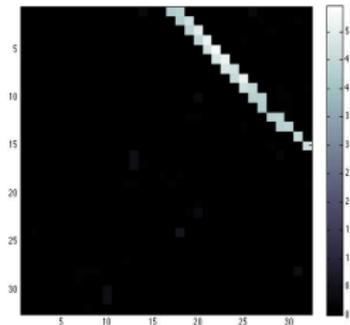


(d) White house coefficients

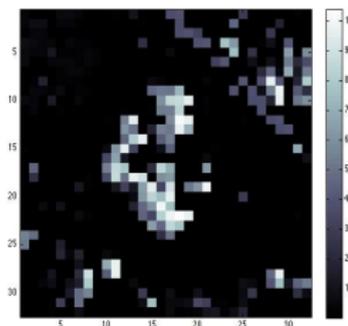
Frame coefficients



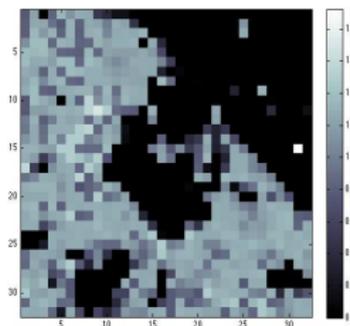
(a) Original



(b) Road coefficients



(c) Tree coefficients



(d) Dirt/grass coefficients

Frames in remote sensing

- Specifically constructed frames with built-in features have been utilized in remote sensing data processing (E. H. Bosch, A. Gonz'alez, J. Vivas, and G. Easley)
- 2D tight frames that provide a new way to analyze, visualize, and process data at multiple scales and directions was proposed by Bosch et al.
- Olshausen and his collaborators used learned dictionaries, which are frames, in their work on improving the performance of supervised classification algorithms for HSI data.
- Z. Xing, M. Zhou, A. Castrodad, G. Sapiro and L. Carin constructed dictionary learning algorithms for noisy and incomplete Hyperspectral images.

Feature Space Rotation for LE

Our final example deals with data recovery which exploits heterogeneous data structures. We begin with a brief description of the feature space rotations.

- Feature space rotation for LE is a modification from rotation for diffusion maps, introduced by Coifman and Hirn (2013).

R. R. Coifman and M. Hirn, Diffusion maps for changing data, Applied and Computational Harmonic Analysis, vol. 36 (2014) no. 1, pp. 79–107.

- For LE, define weight function as $a(x, y) = \frac{1}{2} \left(\delta_{x,y} - \frac{k(x,y)}{\sqrt{m(x)m(y)}} \right)$, where $m(x) = \sum_y k(x, y)$.
- LE distance for two different kernels, k_α and k_β , is defined as

$$D(x_\alpha, y_\beta) = \|a_\alpha(x, \cdot) - a_\beta(y, \cdot)\|_2.$$

- Can show $D(x_\alpha, y_\beta) = \|\Phi_\alpha(x) - \mathcal{O}_{\beta \rightarrow \alpha} \Phi_\beta(y)\|_{\ell^2}$, where

$$(\mathcal{O}_{\beta \rightarrow \alpha} x)_i = \sum_j x_j \langle \phi_\alpha^{(i)}, \phi_\beta^{(j)} \rangle_{L^2(X)}, \quad x \in \ell_\beta^2.$$

HSI and LIDAR Data: Gulfport

With A. Cloninger and T. Doster (2014) we proposed to use LE preimage for a data recovery application in heterogeneous setting.

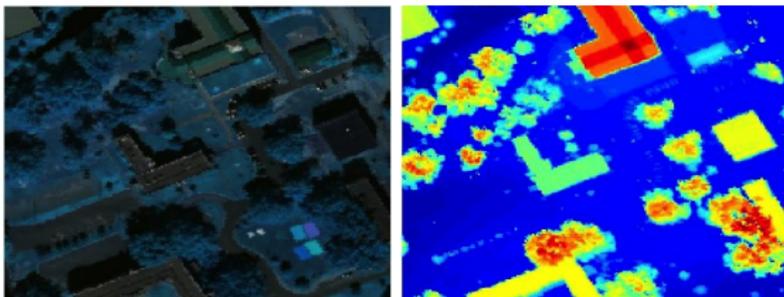


Figure : Lidar and HSI data were acquired concurrently and coregistered using Optech Inc. Gemini Airborne Topographic LIDAR Mapper (ALTM) system, and ITRES Inc. hyperspectral Compact Airborne Spectrographic Imager (CASI-1500). Out of the original 72 HSI bands, 58 bands were selected for higher signal to noise ratio. Courtesy of Paul Gader, University of Florida, and Alina Zare, University of Missouri.

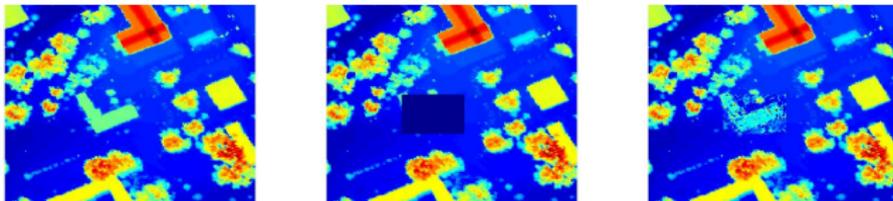


Figure : LIDAR Heat Map, Occluded region, and Reconstruction from HSI

A. Cloninger, Exploiting Data Dependent Structure for Improving Sensor Acquisition and Integration, Ph.D. Thesis, University of Maryland College Park, 2014

T. Doster, Harmonic Analysis Inspired Data Fusion for Applications in Remote Sensing, Ph.D. Thesis, University of Maryland College Park, 2014

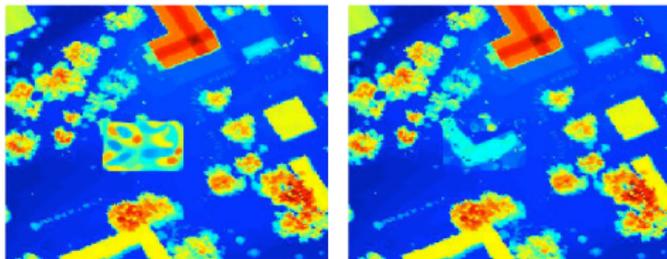


Figure : Euler Elastica Inpainting directly (left) and in the new representation (right)

J. Shen, S. H. Kang, and T. Chan, Euler's Elastica and Curvature-Based Inpainting, SIAM J. Appl. Math., 63 (2003) no. 2, pp. 564–592.

Summary and Conclusions

- Over these 3 days of lectures we have covered a wide range of methods in harmonic analysis that can be potentially applied in Big Data applications.
- Two major classes of techniques we have introduced were: multiscale methods and data-dependent kernel methods.
- There is also a long list of topics in harmonic analysis which can be applied in these applications, but which we have **not** had an opportunity to cover in this short tutorial.
- The major points which we emphasized are that harmonic analysis is one of the best approaches to deal with big data, but there is still a lot to be done before HA techniques are widely applicable in true big data settings.