2015 DCLDE Conference

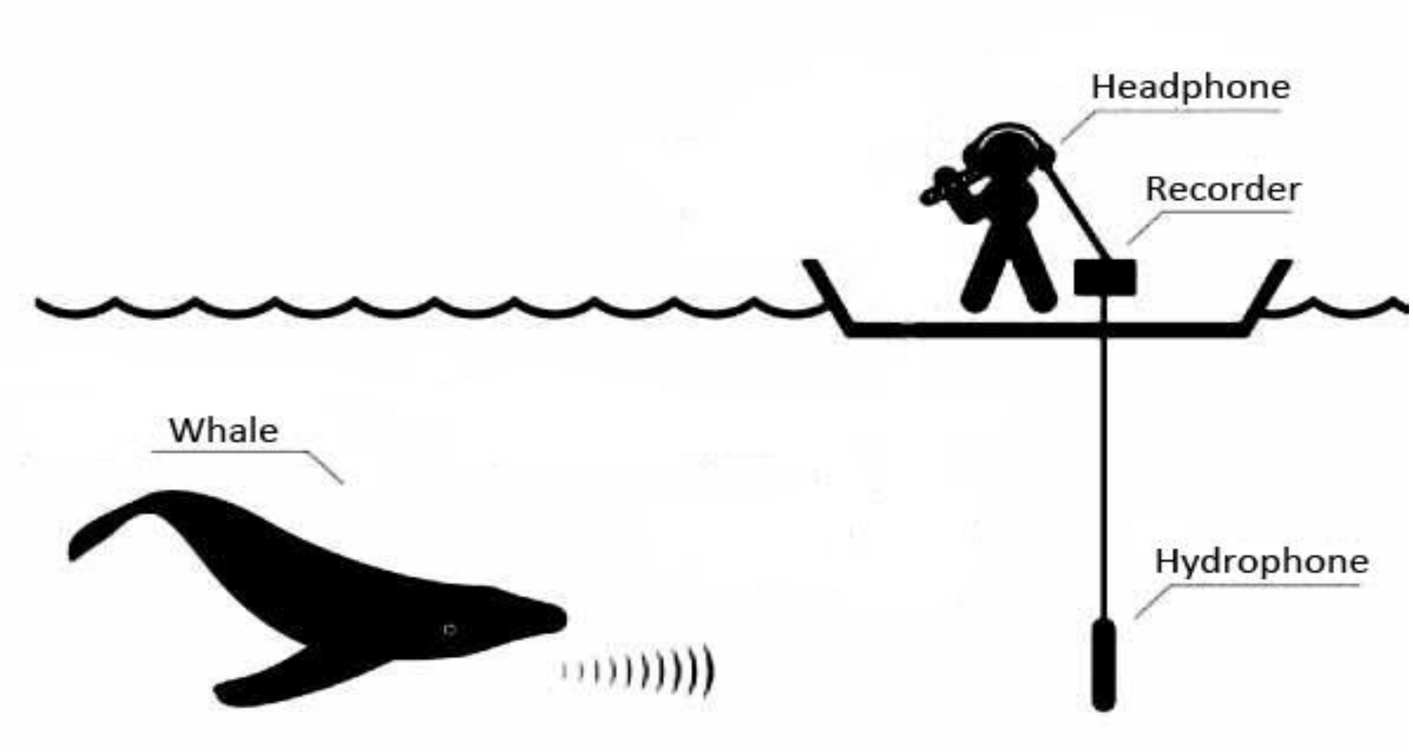
Intrinsic Structure Study on Whale Vocalizations

Yin Xian[1], Xiaobai Sun[2], Yuan Zhang[3], Wenjing Liao[3]
Doug Nowacek[1,4], Loren Nolte[1], Robert Calderbank[1,2,3]

[1]Department of Electrical and Computer Engineering, Duke University
[2]Department of Computer Science, Duke University
[3]Department of Mathematics, Duke University
[4]Duke Marine Lab, Duke University

Problem Statement

Goal: classify the whale signal from the hydrophone.

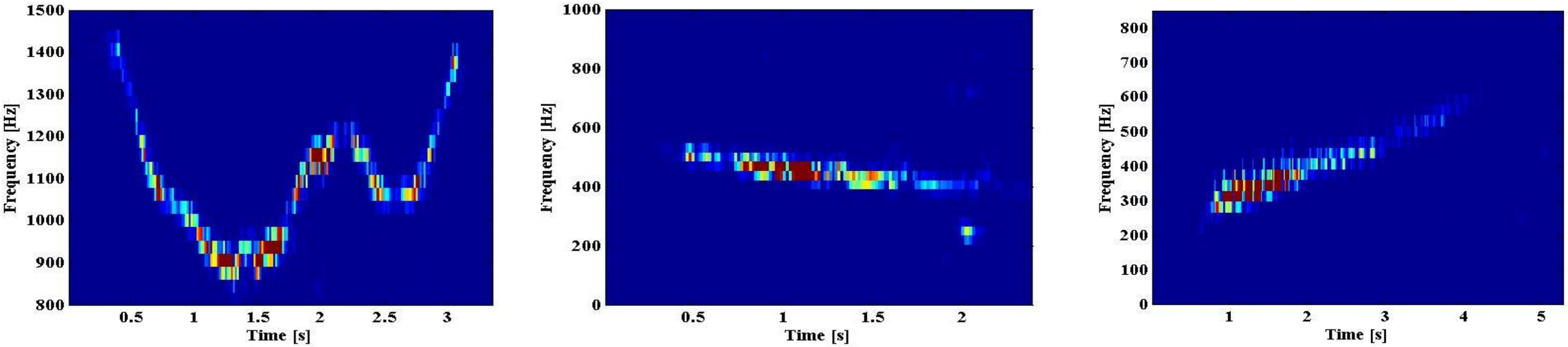


- Passive acoustic;
- Challenge: variation of whale vocalizations, background noise

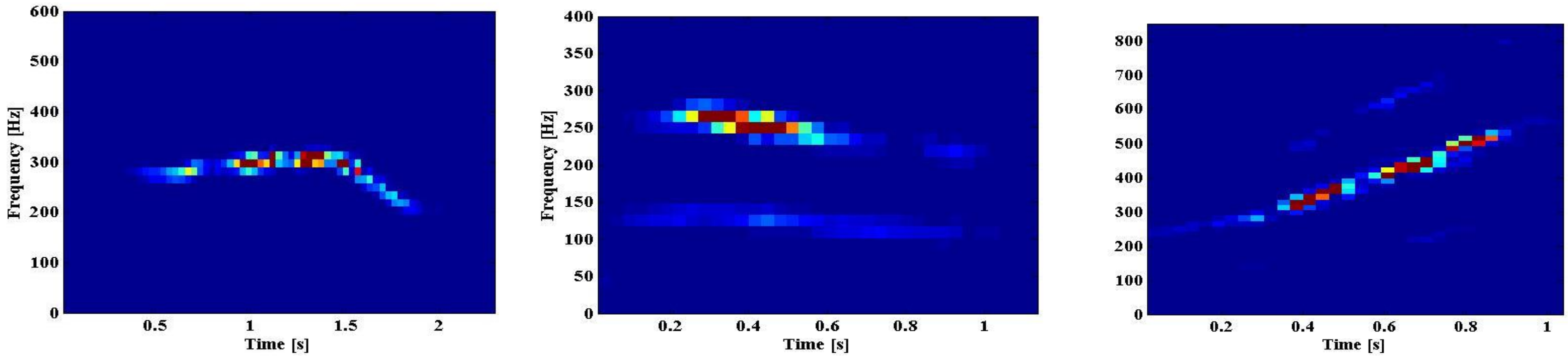
Duke University School of Engineering

Variation of Whale Vocalizations

Bowhead whale calls [1]



Humpback whale calls [1]



[1] Mobysound data. http://www.mobysound.org/mysticetes.html.

Overview

- Many whale vocalizations frequency modulated and can be modeled as polynomial phase signals[2,3].

$$s(t) = A\cos(2\pi(\sum_{m=0}^{N} a_m t^m))$$

- The intrinsic dimension can be described and estimated by the number of polynomial phase parameters.
- Use low dimension representation for the signals and classify them.

[2] I. R. Urazghildiiev, and C. W. Clark. "Acoustic detection of North Atlantic right whale contact calls using the generalized likelihood ratio test," J. Acoust. Soc. Am. 120, 1956-1963 (2006).

[3] M. D. Beecher. "Spectrographic analysis of animal vocalizations: implications of the "uncertainty principle"," Bioacoustics 1, 187-208 (1988).

Dimension reduction

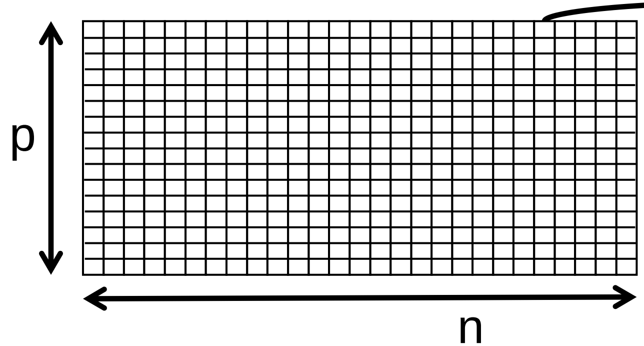
Linear methods:

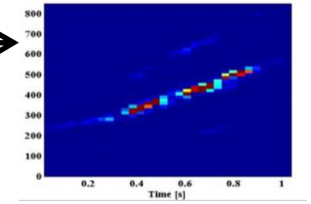
- PCA
- MDS (Multidimensional Scaling)

Non-linear methods:

- Laplacian-Eigenmap
- Isomap

Road map





X

Linear model | Non-linear model

PCA: $\frac{1}{n-1}\tilde{X}\tilde{X}^T$

MDS: $\tilde{X}^T\tilde{X}$

....

....

Isomap:

$$K = \begin{cases} ||x_i - x_j||^2 & i,j \text{ neighbors} \\ \min_{path} \sum ||x_i - x_j||, & \text{otherwise} \end{cases}$$

Laplacian-Eigenmap: $L = \exp(D, t)$

$G = U\Sigma U^T$, then spectral embedding to k dimensions

↓

Classifier

Duke University School of Engineering

PCA

- Denote the data by $X = [x_1, \dots, x_n] \in R^{p\times n}$,
- Covariance matrix: $\Sigma_n = \frac{1}{n-1}\tilde{X}\tilde{X}^T$, where $\tilde{X} = X - \frac{1}{n}Xee^T$
- The eigenvalue decomposition: $\Sigma_n = U\Lambda U^T$
- Choose the top k eigenvalues and the corresponding eigenvectors for Σ_n, and compute $\bar{Y}_k = U_k(\Lambda_k)^{-1/2}$

- The PCA compute the top k right singular vectors for $\tilde{X}$.

[4] H. Hotelling. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24(4):17–441,498–520 (1933).

MDS

- Denote the data by $X = [x_1, \ldots, x_n] \in R^{p \times n}$,
- The distance matrix: $D_{ij} = {d_{ij}}^2 = ||x_i - x_j||^2$.
- Compute $B = -\frac{1}{2} HDH^T$, where $H = I - e\mathrm{e}^T/n$, the centering matrix.
- Compute eigenvalue decomposition $B = U\Lambda U^T$.
- Choose top k nonzero eigenvalue and corresponding eigenvector for B , $\bar{X}_k = U_k(\Lambda_k)^{-1/2}$

[5] J. B. Kruskal, and M. Wish. Multidimensional scaling. Vol. 11. Sage (1978).

MDS

- Denote the data by $X = [x_1, \ldots, x_n] \in R^{p \times n}$,
- The distance matrix: $D_{ij} = {d_{ij}}^2 = ||x_i - x_j||^2$.
- Compute $B = -\frac{1}{2} HDH^T$, where $H = I - ee^T/n$, the centering matrix.
- Compute eigenvalue decomposition $B = U\Lambda U^T$.
- Choose top k nonzero eigenvalue and corresponding eigenvector for B , $\bar{X}_k = U_k(\Lambda_k)^{-1/2}$
- The relationship between B and the covariance matrix via $\tilde{X}$:

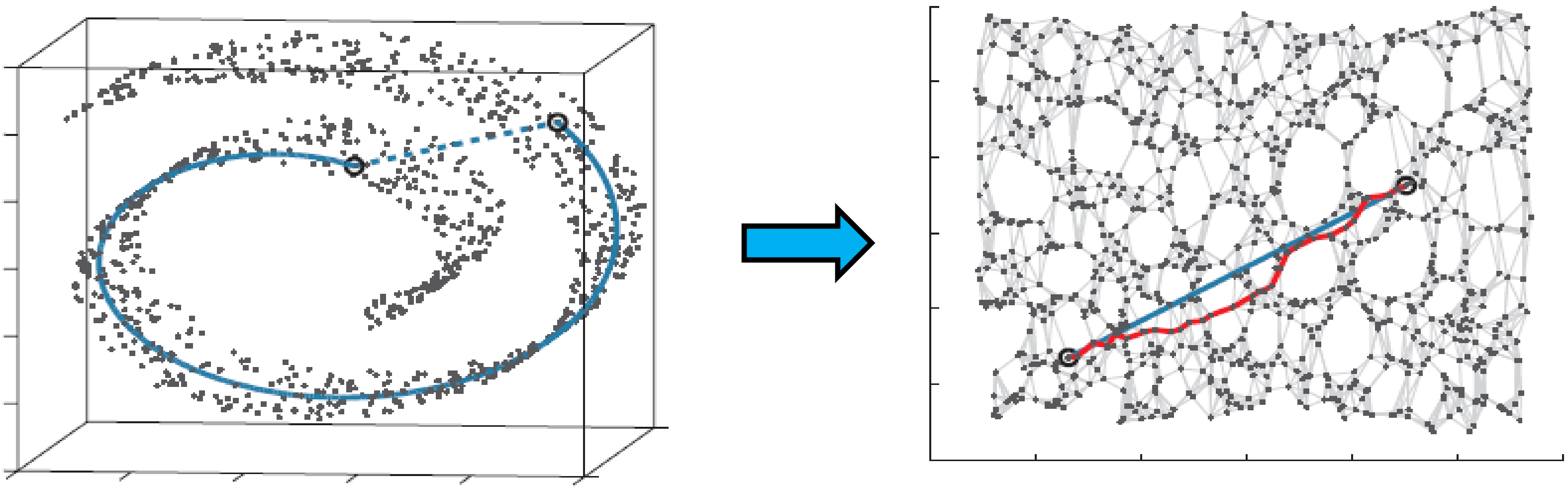
$$B = \tilde{X}^T \tilde{X}, \qquad \Sigma_n = \frac{1}{n-1} \tilde{X} \tilde{X}^T$$

so the MDS is actually compute the top left singular vectors of $\tilde{X}$.

[5] J. B. Kruskal, and M. Wish. Multidimensional scaling. Vol. 11. Sage (1978).

Idea of Isomap

The “Swiss roll” data set, illustrating Isomap exploits geodesic paths for nonlinear dimensionality reduction.



- For two arbitrary points on a nonlinear manifold, their Euclidean distance in the high-dimensional input space may not accurately reflect their intrinsic similarity, as measured by geodesic distance along the low-dimensional manifold.
- The two-dimensional embedding recovered by Isomap, which best preserves the shortest path distances in the neighborhood graph.

[6] J. B. Tenenbaum et al. "A global geometric framework for nonlinear dimensionality reduction." Science 290, 2319-2323 (2000).

Isomap

- Construct a neighborhood graph G=(X, E, D), based on k nearest neighborhood, or ε-neighborhood.
- Compute D,

$$D_{ij} = \begin{cases} ||x_i - x_j||^2 & \text{if } i \text{ and } j \text{ are neighbors} \\ \min\limits_{\{t_1,\ldots,t_k\} \text{ is a path between } i,j} (||x_{t_1} - x_i|| + \cdots + ||x_{t_k} - x_j||), & \text{otherwise} \end{cases}$$

- Compute $K = -\frac{1}{2}HDH^T$, where $H = I - ee^T/n$ is the centering matrix.
- Compute eigenvalue decomposition $K = U\Lambda U^T$.
- Choose the top k eigenvalues and eigenvectors and compute $\bar{X}_k = U_k(\Lambda_k)^{-\frac{1}{2}}$.

[6] J. B. Tenenbaum et al. "A global geometric framework for nonlinear dimensionality reduction." Science 290, 2319-2323 (2000).

Laplacian-Eigenmap

- Construct a neighborhood graph G=(X, E, W), based on k nearest neighborhood, or ε-neighborhood.
- Choose the weight:

$$w_{ij} = \begin{cases} e^{-\frac{||x_i - x_j||^2}{t}}, & i, j \ connected \\ 0, & otherwise \end{cases}$$

- Eigenmap:
  – Construct Laplacian matrix L=D-W, where D=diag($\sum_{j \in N_i} w_{ij}$)
  – Compute eigenvalues and eigenvectors:

$$L\boldsymbol{f} = \lambda D\boldsymbol{f}$$

$\boldsymbol{f} = [f_0, \ldots, f_k]$ corresponds to D = diag($\lambda_1, \ldots, \lambda_k$), $\lambda_i <= \lambda_{i+1}$

- Leave out the eigenvector f_0.
- The m dimensional embedding with $(f_1, \ldots, f_m)$.

[7] M. Belkin, and P. Niyogi. "Laplacian Eigenmaps for dimensionality reduction and data representation." Neural computation, 15, 1373-1396, (2003).

Laplacian-Eigenmap

- Construct a neighborhood graph G=(X, E, W), based on k nearest neighborhood, or ε-neighborhood.
- Choose the weight:

$$w_{ij} = \begin{cases} e^{-\frac{||x_i - x_j||^2}{t}}, & i, j \ connected \\ 0, & otherwise \end{cases}$$

- Eigenmap:
  - Construct Laplacian matrix L=D-W, where D=diag($\sum_{j \in N_i} w_{ij}$)
  - Compute eigenvalues and eigenvectors:

$$L\boldsymbol{f} = \lambda D\boldsymbol{f}$$

$\boldsymbol{f} = [f_0, \dots, f_k]$ corresponds to D = diag($\lambda_1, \dots, \lambda_k$), $\lambda_i <= \lambda_{i+1}$

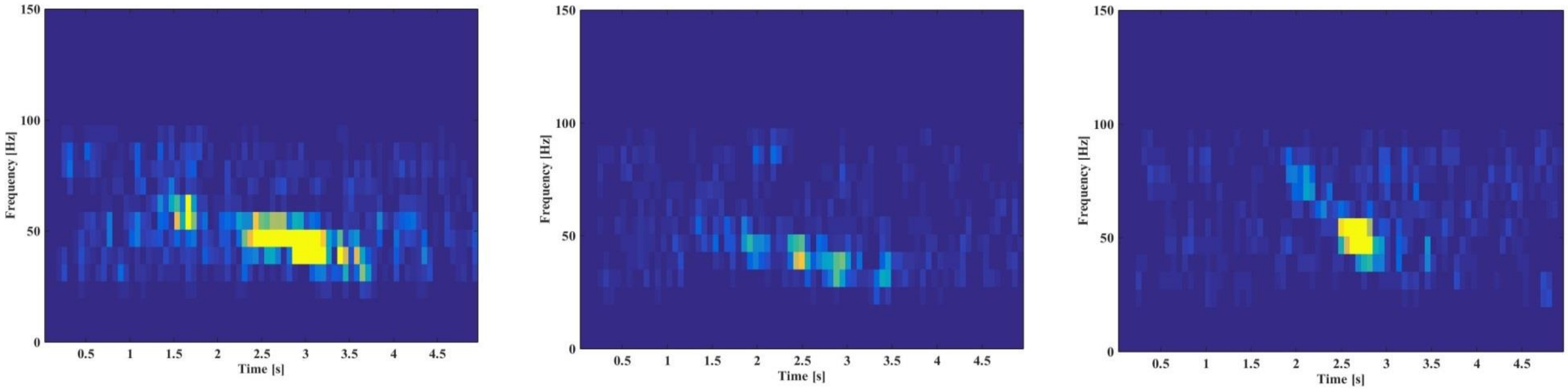
- Leave out the eigenvector f_0.
- The m dimensional embedding with ($f_1, \dots, f_m$).
- For **normalized Laplacian-Eigenmap**, we compute Φ:

$$D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}\Phi = \lambda\Phi$$

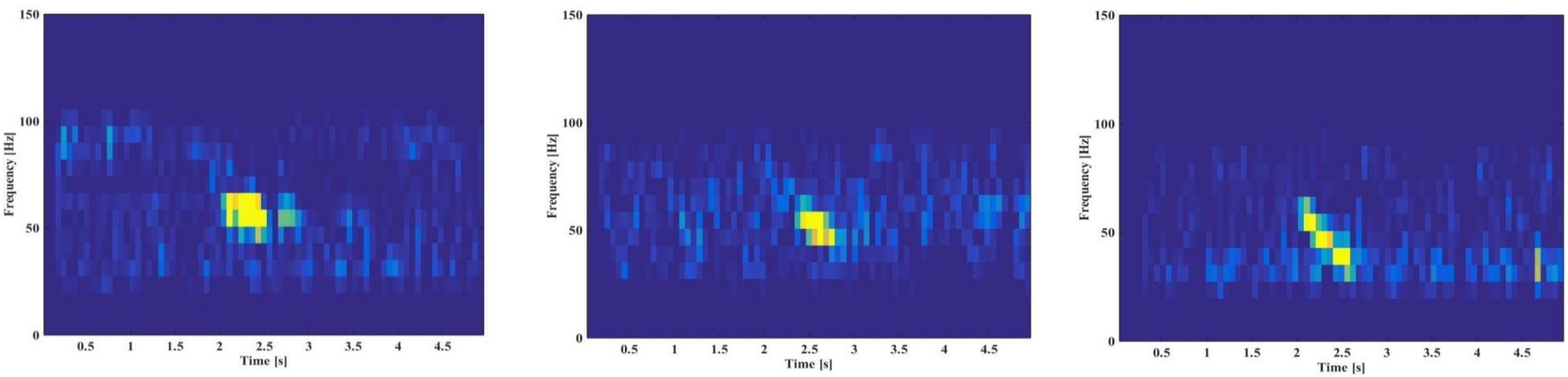
Duke University School of Engineering

DCLDE 2015 Data

Blue whale (# 851)



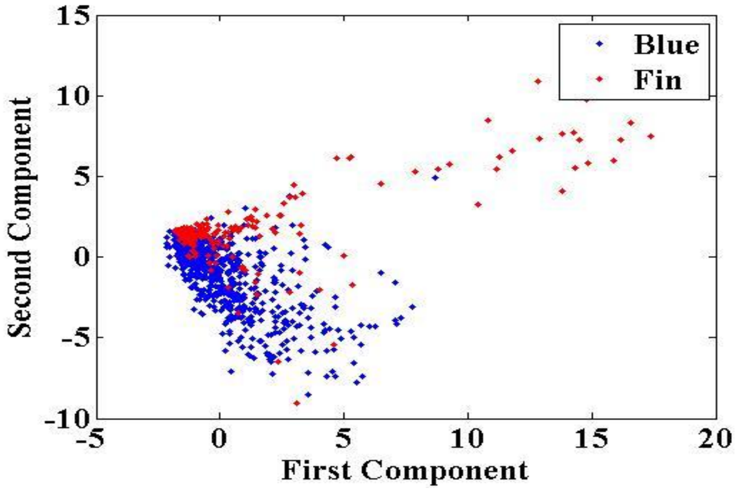
Fin whale (# 244)



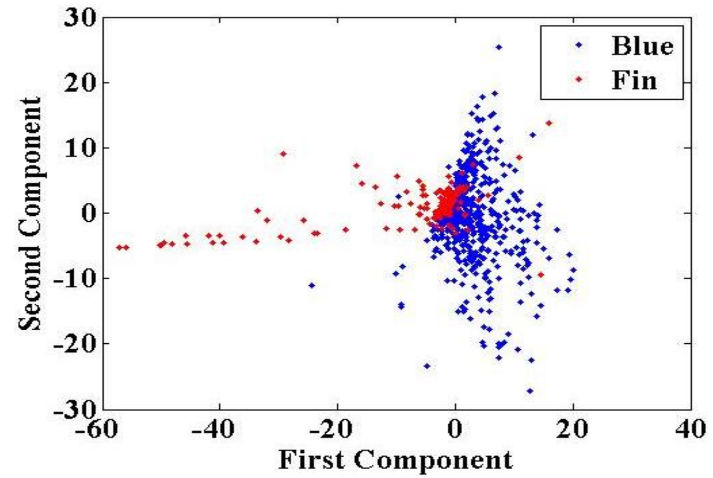
[8] DCLDE conference data. http://www.cetus.ucsd.edu/dclde/dataset.html.

Mapping Data to two dimensions

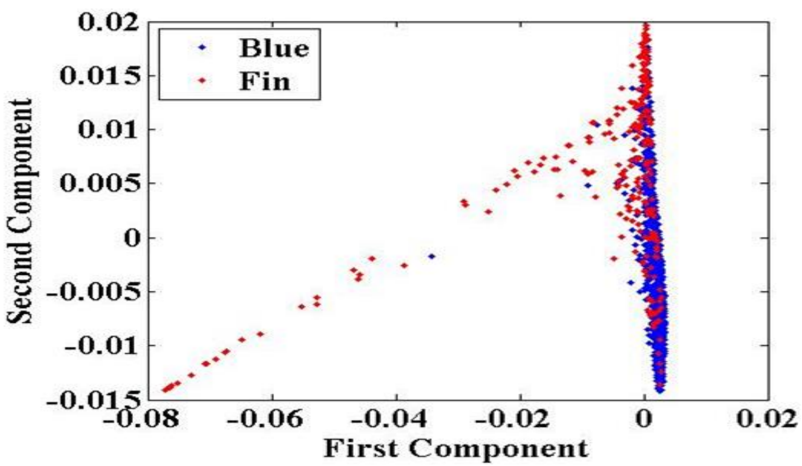
PCA



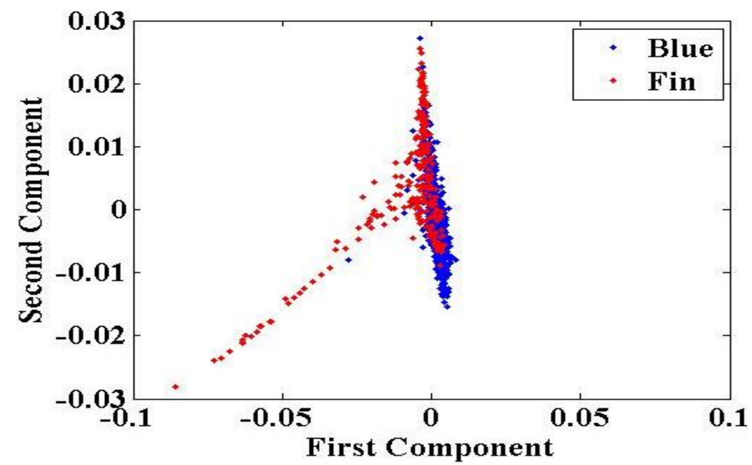
Isomap



Laplacian Eigenmap



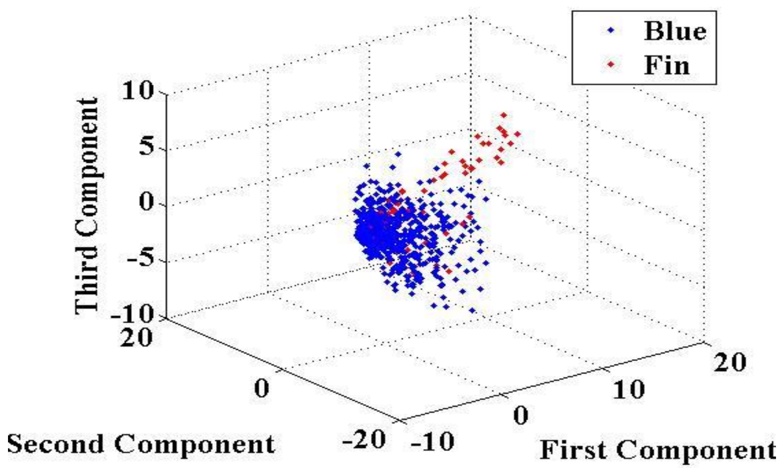
Normalized Laplacian Eigenmap



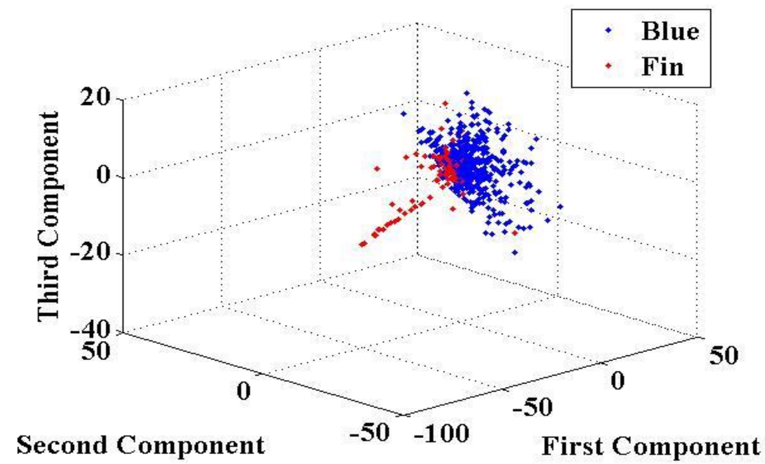
Duke University School of Engineering

Mapping Data to three dimensions

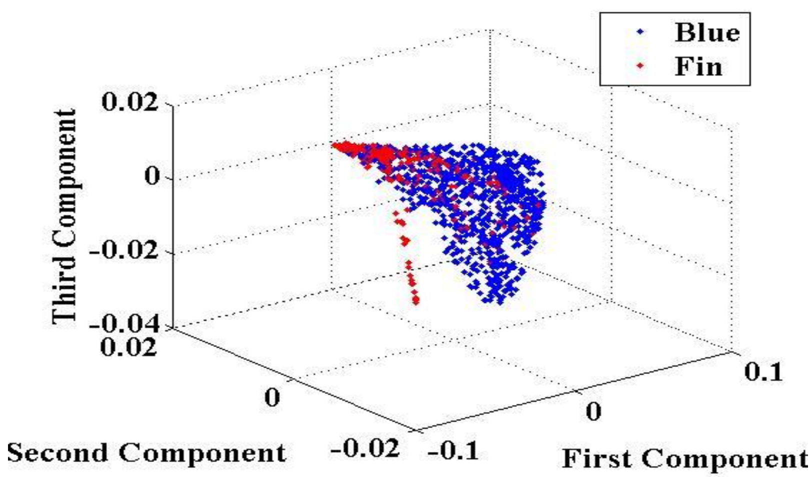
PCA



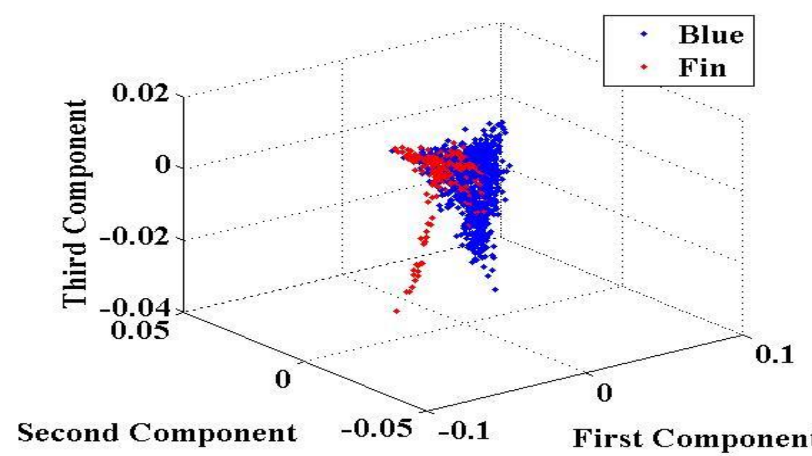
Isomap



Laplacian Eigenmap



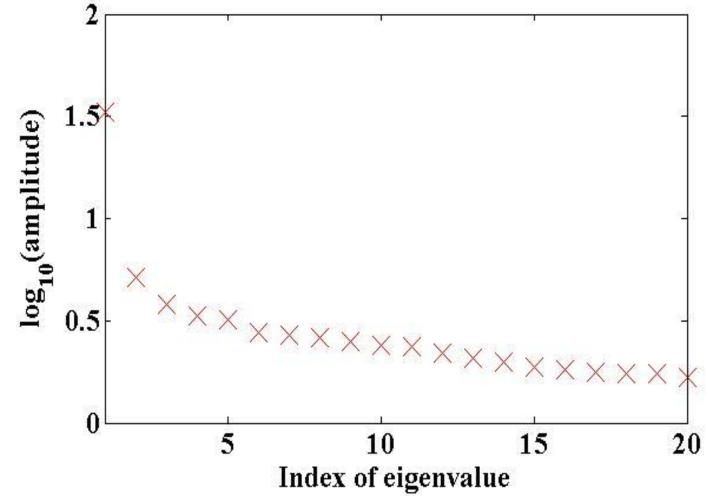
Normalized Laplacian Eigenmap



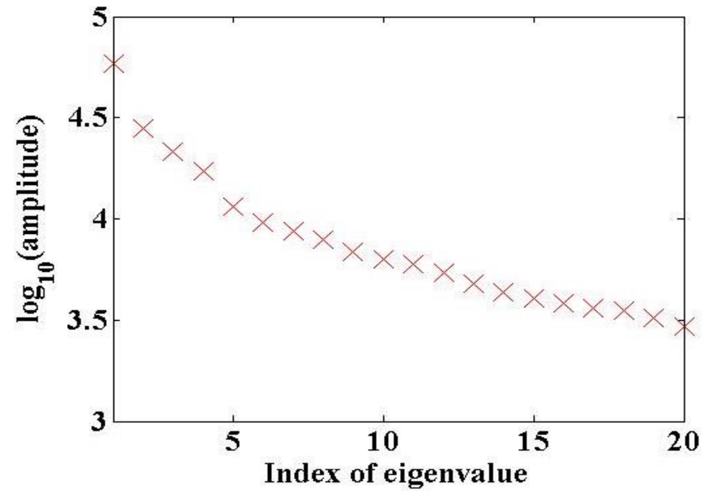
Duke University School of Engineering

Eigenvalues (energy) distributions

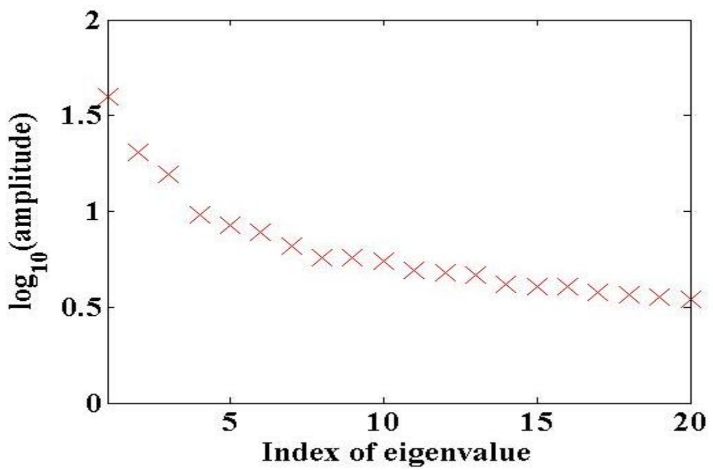
PCA



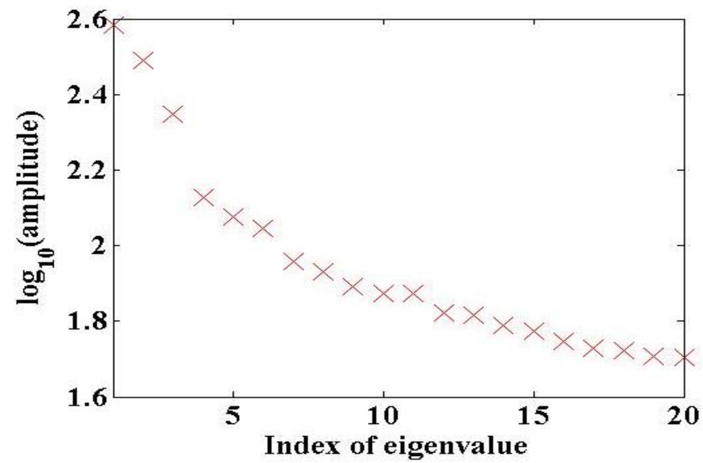
Isomap



Laplacian Eigenmap



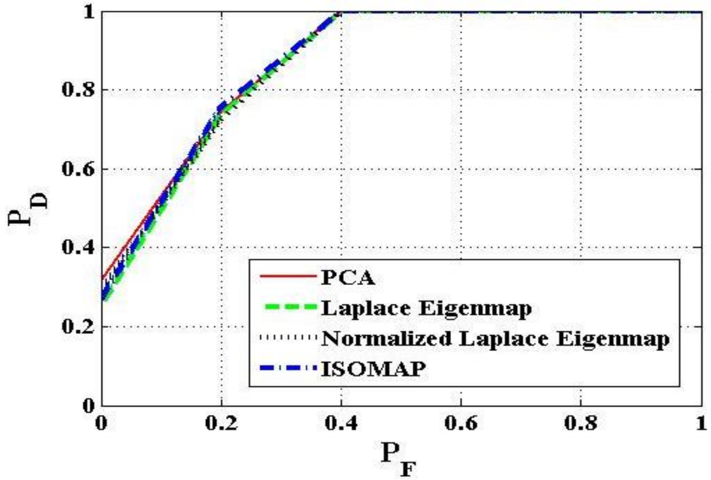
Normalized Laplacian Eigenmap



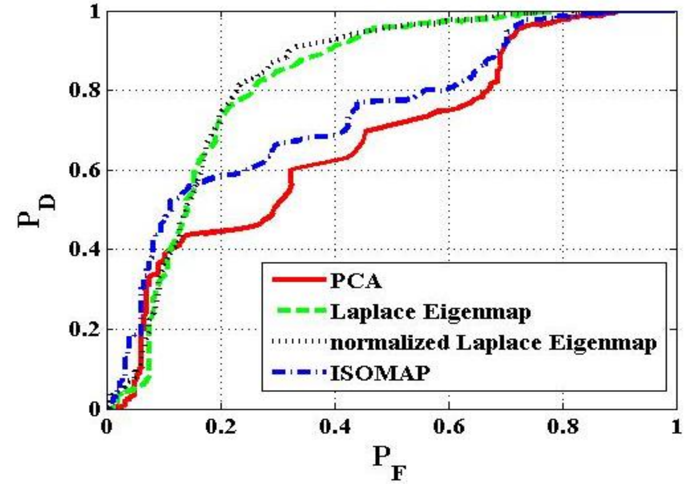
Duke University School of Engineering

ROCs comparisons (2D)

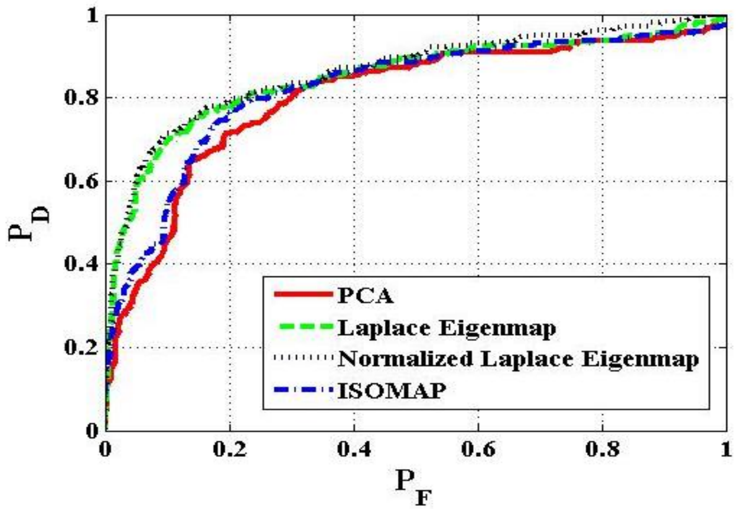
KNN



Naïve Bayes



Logistic regression

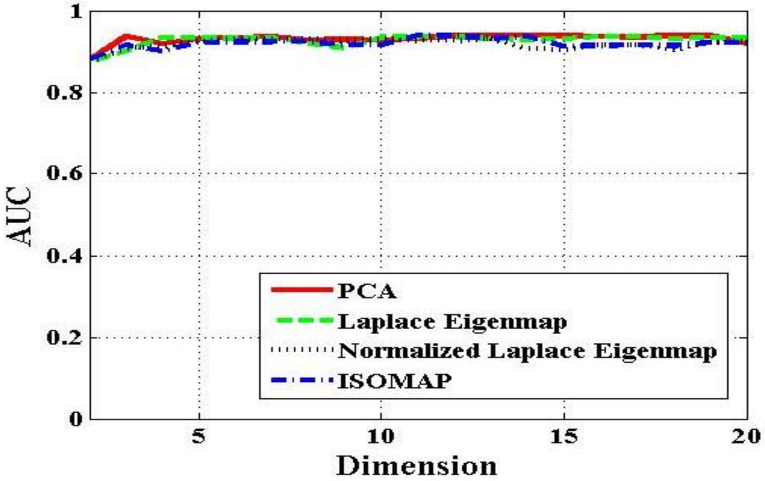


- We use 5-folds cross validation to generate the ROC. That is, 681 blue whale sounds and 196 fin whale sounds for training, and 170 blue whale sounds and 48 fin whale sounds for testing.
- We use k=7 for Isomap, and k=7, t=1 for Laplacian Eigenmap.

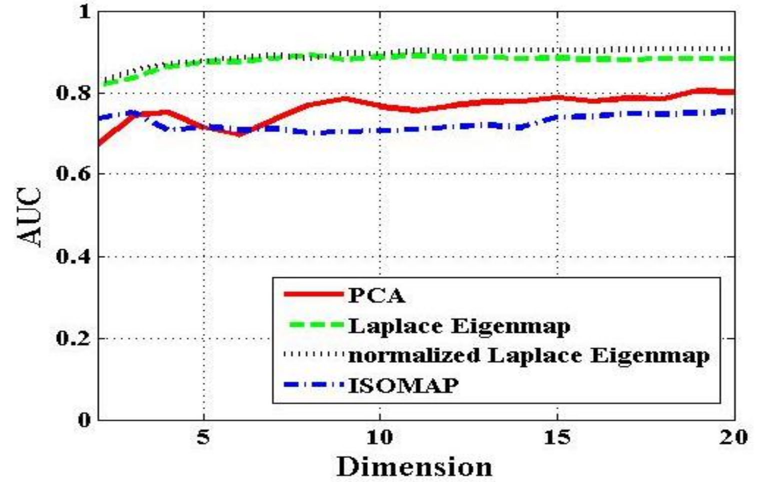
Duke University School of Engineering

AUCs comparisons

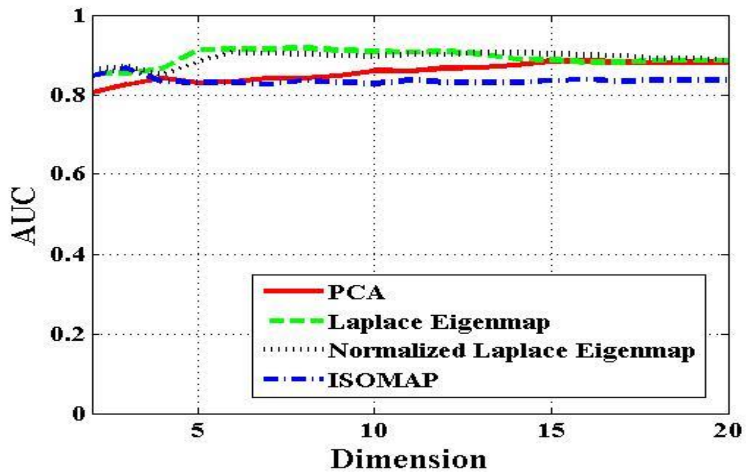
KNN



Naïve Bayes

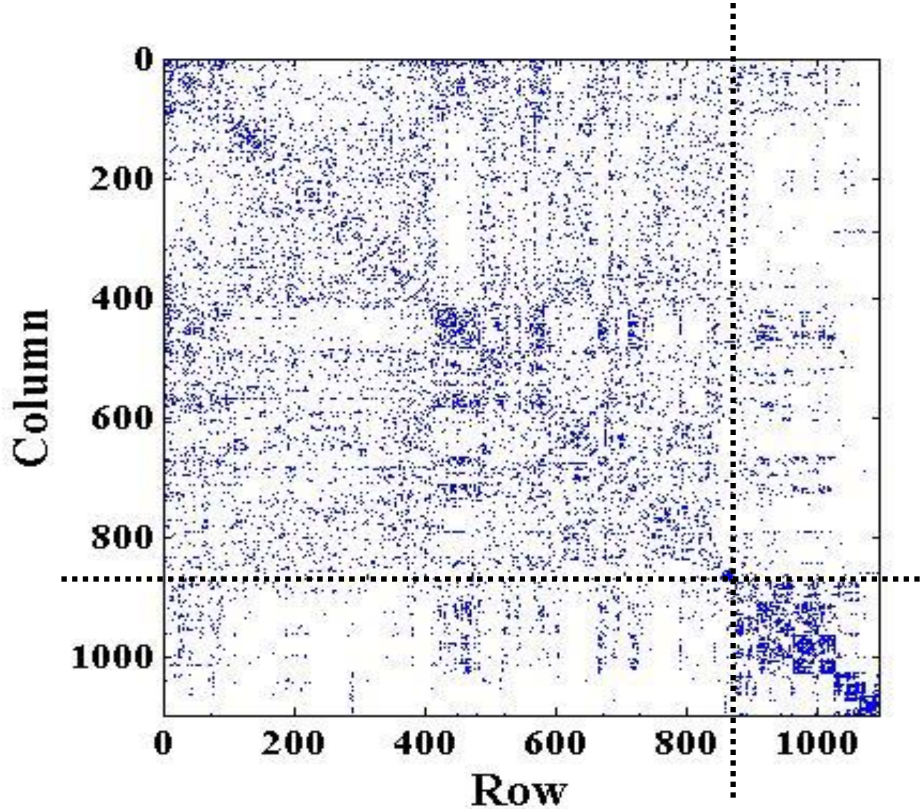


Logistic regression



Duke University School of Engineering

Plots of adjacency matrix of Laplacian-Eigenmap



By systematic spectral re-ordering, the blue whale data and fin whale data are well separated in the adjacency matrix (we use 851 blue whale data, and 244 fin whale data).

Summary

- Efficient classification of the whale vocalizations from low-dimensional intrinsic structure.
- The intrinsic dimension of whale vocalizations can be recovered from the eigenvalues energy distribution.
- The nonlinear dimensional reduction methods work better with data of nonlinear structure.

Future topics

- Further develop efficient manifold mappings for more complex whale vocalizations, and other acoustic signals.
- Apply optimization methods to enhance computational efficiency for nonlinear dimensional mappings.

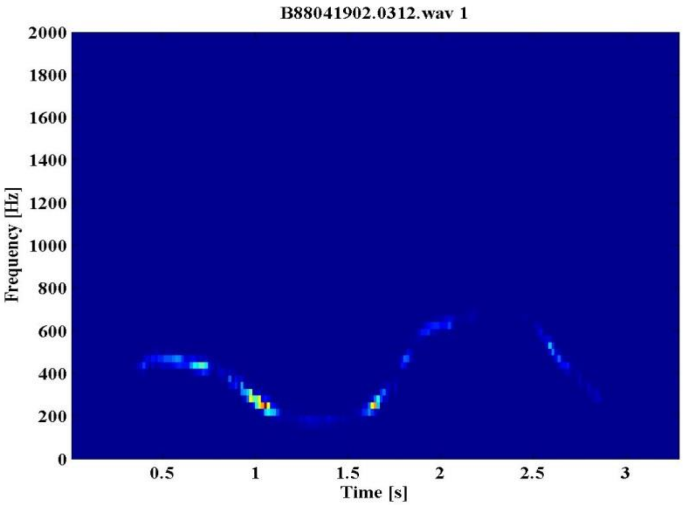
Reference

- [1] Mobysound data. http://www.mobysound.org/mysticetes.html.
- [2] I. R. Urazghildiiev, and C. W. Clark. “Acoustic detection of North Atlantic right whale contact calls using the generalized likelihood ratio test,” J. Acoust. Soc. Am. 120, 1956-1963 (2006).
- [3] M. D. Beecher. “Spectrographic analysis of animal vocalizations: implications of the “uncertainty principle”,” Bioacoustics 1, 187-208 (1988).
- [4] H. Hotelling. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24(4):17–441,498–520 (1933).
- [5] J. B. Kruskal, and M. Wish. Multidimensional scaling. Vol. 11. Sage (1978).
- [6] J. B. Tenenbaum, V. D. Silva, and J. C. Langford. "A global geometric framework for nonlinear dimensionality reduction." Science 290, 2319-2323 (2000).
- [7] M. Belkin, and P. Niyogi. “Laplacian Eigenmaps for dimensionality reduction and data representation.” Neural computation, 15, 1373-1396, (2003).
- [8] DCLDE conference data. http://www.cetus.ucsd.edu/dclde/dataset.html.
- [9] Y. Xian, X. Sun, Y. Zhang, W. Liao, D. Nowacek, L. Nolte, and R. Calderbank. “Intrinsic structure study of whale vocalization.” J. Acoust. Soc. Am. (in preparation)

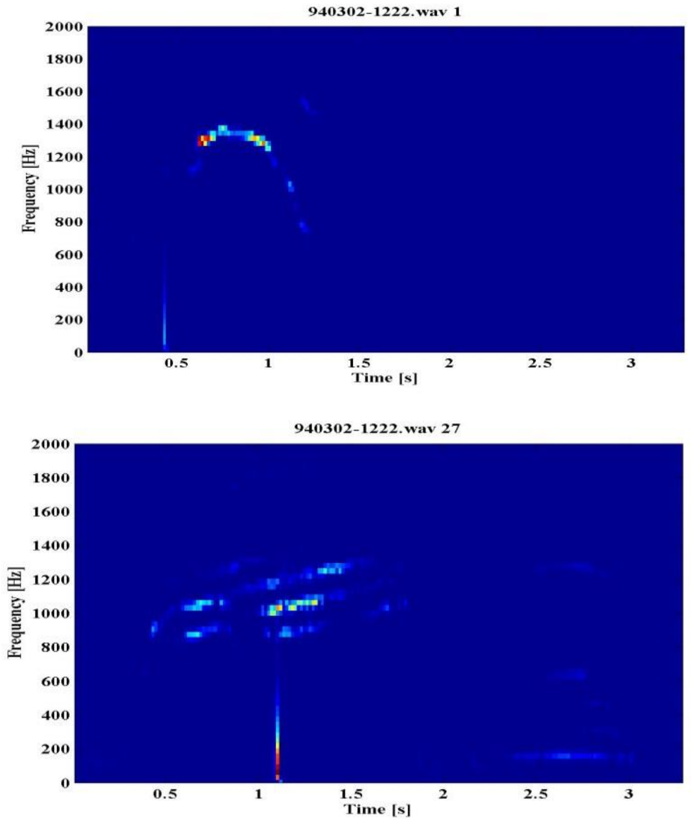
Backup slides

Mobysound data

Bowhead whale (#446)



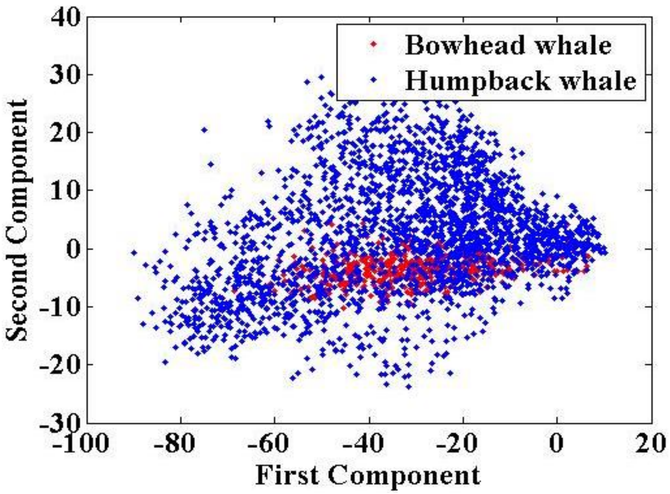
Humpback whale (#2310)



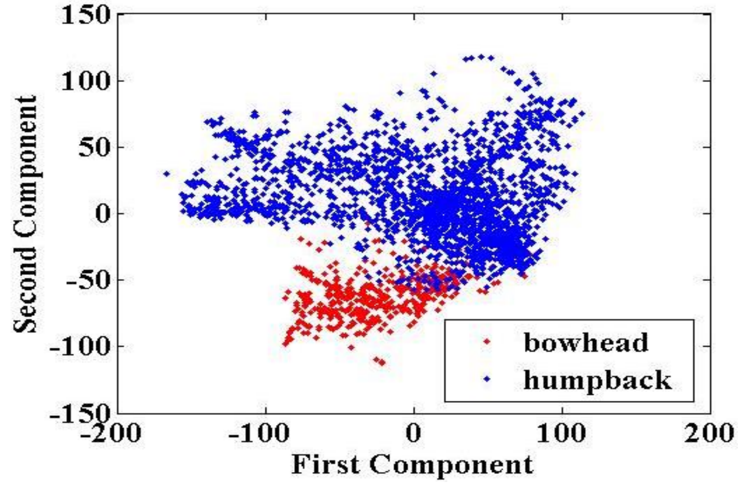
[1] Mobysound data. http://www.mobysound.org/mysticetes.html.

Mapping to two dimensions

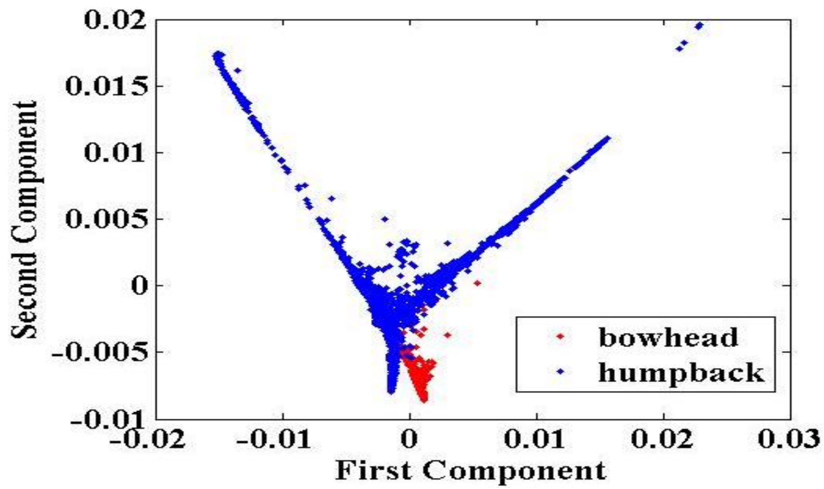
PCA



Isomap



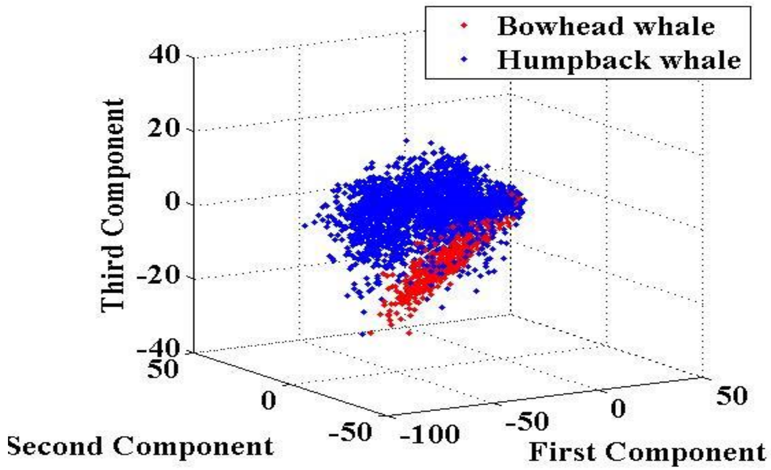
Laplacian Eigenmap



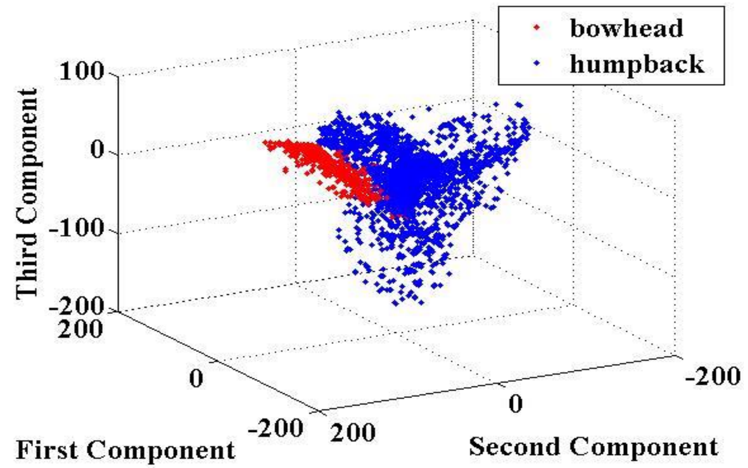
Duke University School of Engineering

Mapping to three dimensions

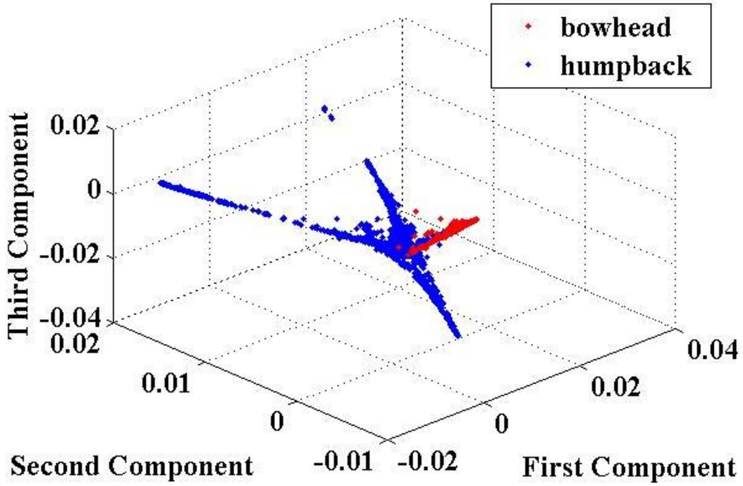
PCA



Isomap



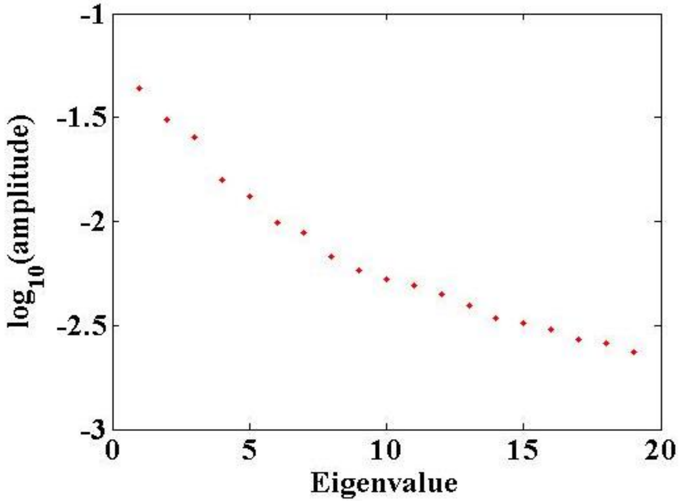
Laplacian-Eigenmap



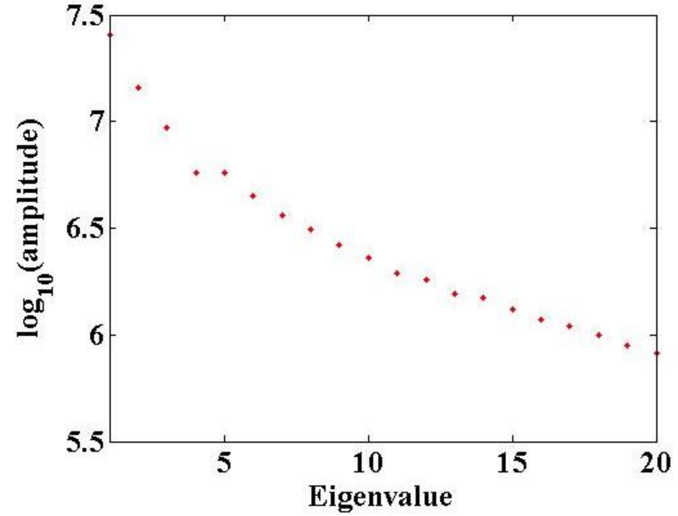
Duke University School of Engineering

Eigenvalues energy distributions

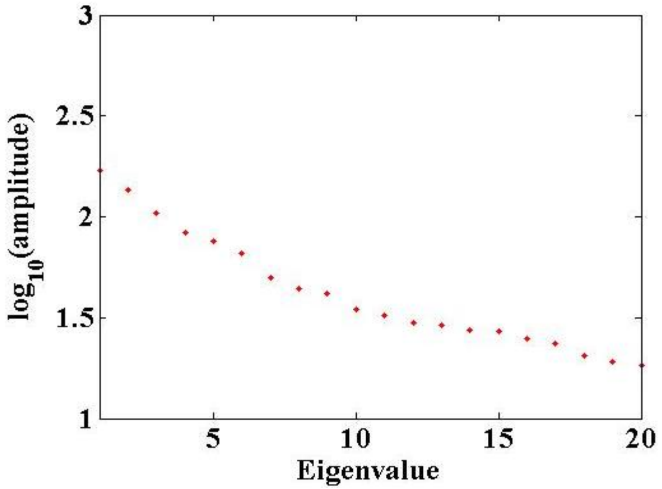
PCA



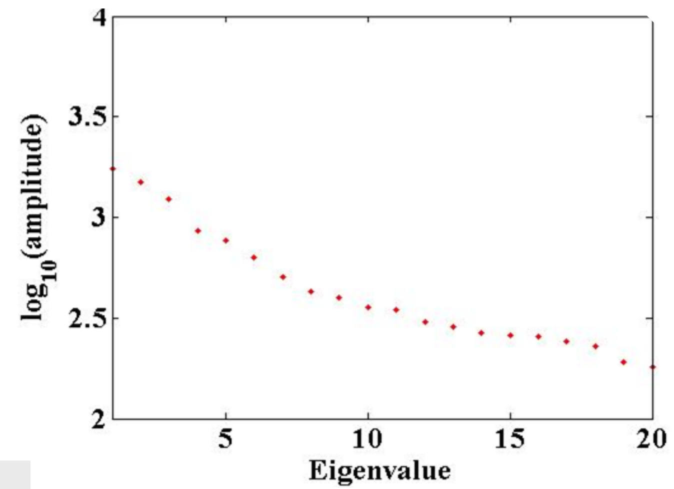
Isomap



Laplacian Eigenmap



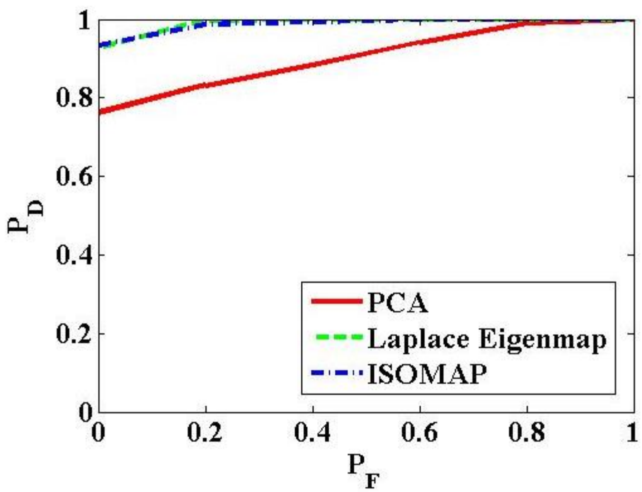
Normalized Laplacian Eigenmap



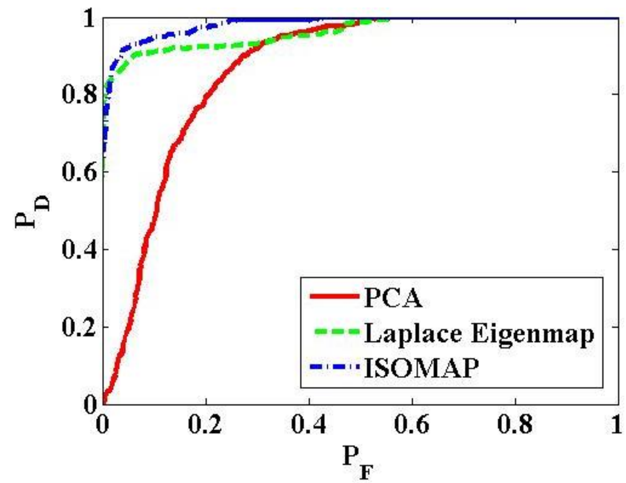
Duke University School of Engineering

ROCs comparisons (2D)

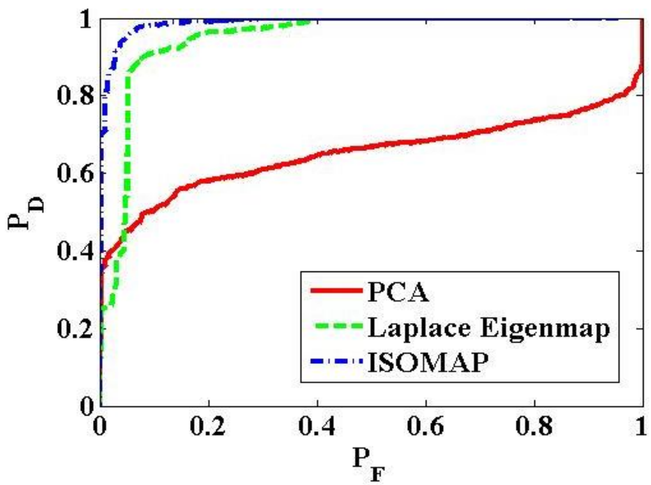
KNN



Naïve Bayes



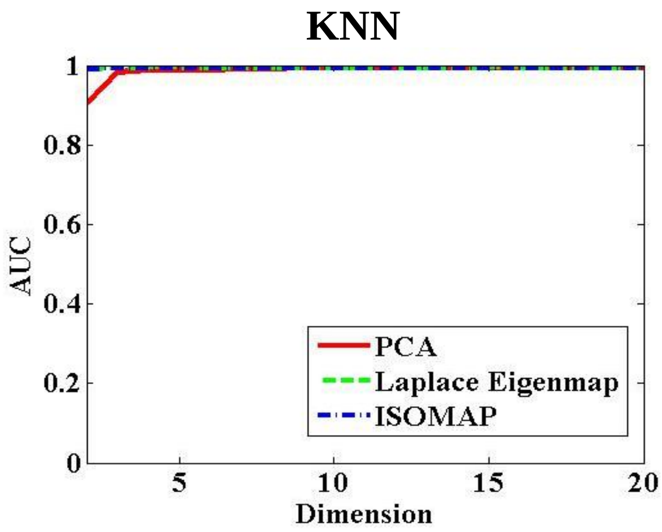
Logistic regression

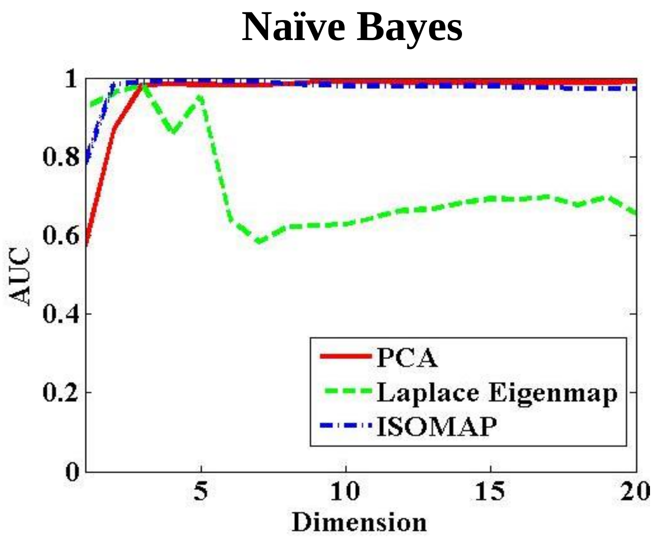


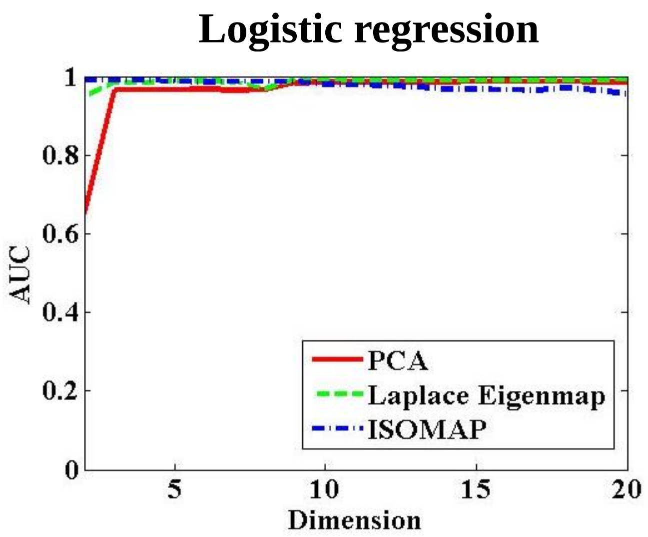
- We use 5-folds cross validation to generate the ROC. That is, 358 bowhead whale sounds and 1848 humpback whale sounds for training, and 88 bowhead whale sounds and 462 humpback whale sounds for testing.
- We use k=7 for Isomap, and k=7, t=1 for Laplacian Eigenmap

Duke University School of Engineering

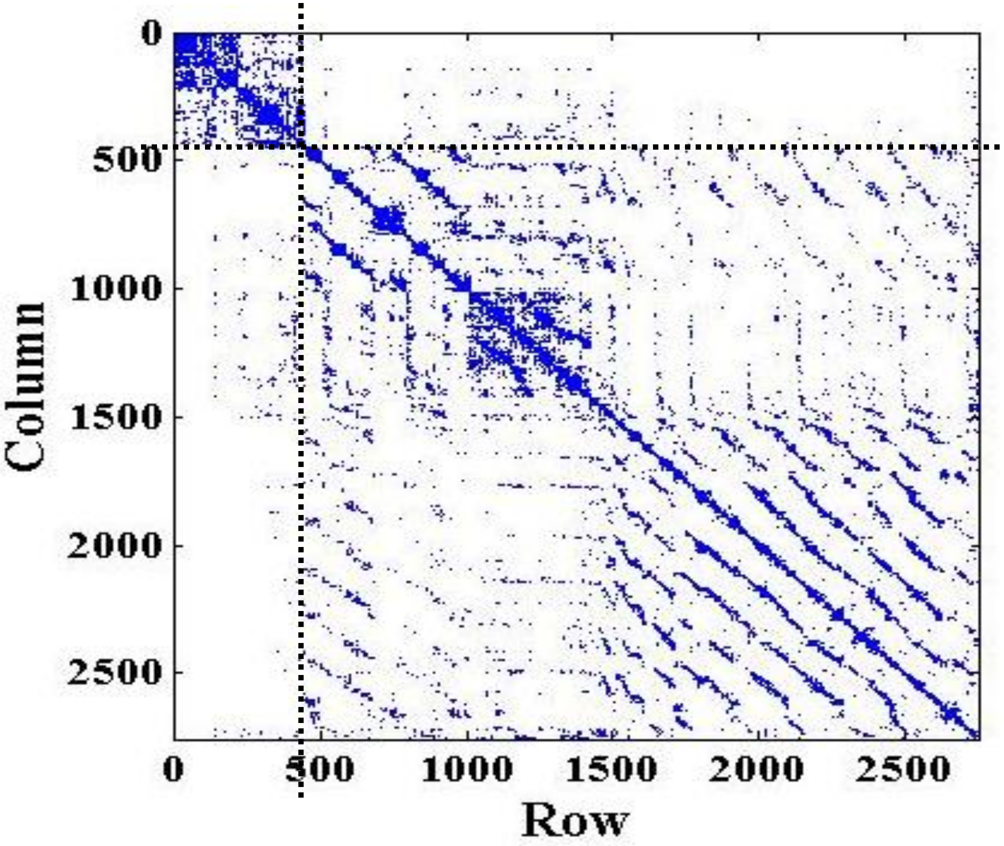
AUCs Comparisons







Plots of adjacency matrix W of Laplacian-Eigenmap



By systematic spectral re-ordering, the bowhead whale data and humpback whale data are well separated in the adjacency matrix (we use 446 bowhead whale data, and 2310 humpback whale data).

Local Linear Embedding

- Construct a neighborhood graph G=(V, E, W),
  - where V is the vertex $\{x_i: i = 1, \dots, n\}$; E is the edge $\{(i, j)$: if j is a neighbor of $i\}$, k-nearest neighbors, ε-neighbors; W is the Euclidean distance: $d(x_i, x_j)$
- Local fitting:
  - Compute the weights $\min_{\sum_{j \in N_i} w_{ij} = 1} ||x_i - \sum_{j \in N_i} w_{ij}(x_j - x_i)||^2$
  - Solve the equation according to Lagrange multipler method:

$$\min_{w_{ij}} \frac{1}{2}||x_i - \sum_{j \in N_i} w_{ij}(x_j - x_i)||^2 + \lambda(1 - \sum_{j \in N_i} w_{ij})$$

Let $w_i = [w_{ij_1}, \dots, w_{ij_k}]^T \in R^k$, $\bar{X}_i = [x_{j_1} - x_i, \dots, x_{j_k} - x_i]$, the local covariance matrix $C_{jk}^{(i)} = < x_j - x_i, x_k - x_i >$, we have:

$$w_i = C_i^\dagger(\bar{X}_i^T x_i + \lambda 1)$$

$$\lambda = \frac{1}{1^T C_i^\dagger 1}(1 - 1^T C_{i=1}^\dagger \bar{X}_i^T x_i)$$

Duke University School of Engineering

Local Linear Embedding

- Global alignment
  - Define a n by n weight matrix W:

$$W = \begin{cases} w_{ij}, & j \in N_i \\ 0, & otherwise \end{cases}$$

  - Compute the global embedding matrix Y:

$$\min_Y \sum_i ||y_i - \sum_{j=1}^{n} W_{ij} y_j||^2 = trace(Y(I-W)^T(I-W)Y^T)$$

That is construct a semi-definite matrix $B = (I - W)^T(I - W)$ and find the d+1 smallest eigenvectors of B: $v_0, v_1, \ldots, v_d$, and the corresponding eigenvalues: $\lambda_0, \ldots, \lambda_d$, drop the smallest eigenvector which is the constant vector, we have:

Y=$\left[\frac{v_1}{\sqrt{\lambda_1}}, \ldots, \frac{v_d}{\sqrt{\lambda_d}}\right]$.

- Advantage of LLE
  - Neighbor graph: k nearest neighbors is of *O(kn)*.
  - W is sparse;
  - $B = (I - W)^T(I - W)$ is positive semi-definite.