Spectral clustering with PCANet and DCTNet

Yin Xian¹, Andrew Thompson², Xiaobai Sun³, and Robert Calderbank¹,³,⁴

¹Department of Electrical and Computer Engineering, Duke University, ²Department of Mathematics, Oxford University, ³Department of Computer Science, Duke University, ⁴Department of Mathematics, Duke University

Abstract

- Introduce PCANet [1] and propose the efficient alternative DCTNet for acoustic signal classification.
- Use the eigenfunctions of PCA and DCT as filterbanks for feature extraction;
- Each layer can reveal different time-frequency content of signals;
- Experimental results show that both PCANet and DCTNet outperform scattering transform in classification of DCLDE whale vocalization data.

Road Map

![Diagram of signal and Toeplitz matrix](image)

Optimal circulant preconditioner for Toeplitz

Given a \( n \times n \) Toeplitz matrix \( T \), Chan [3] derived an optimal circulant preconditioner \( C = \arg \min_{C'} \| C' - T \|_F^2 \). Where \( C \) is given by:

\[
C_i = \frac{it_{i-1} + (n - i)t_i}{n}, \quad i = -(n-1), \ldots, 0, \ldots, (n-1).
\]

and \( t_i \) are the entries of the Toeplitz matrix.

The eigenvectors of a circulant matrix are of Fourier bases.

\( C = F \Lambda F^T \)

We can use DCT to approximate the eigenfunctions of the Toeplitz matrix.

Main Reference


Signal and Toeplitz Matrix

A signal of length N, \( x = (x_1, \ldots, x_N) \), construct a Hankel matrix:

\[
X = \begin{bmatrix}
x_1 & x_2 & x_3 & \cdots & x_{N-M+1} \\
x_2 & x_3 & x_4 & \cdots & x_{N-M+2} \\
\vdots & \vdots & \vdots & & \vdots \\
x_M & x_{M+1} & x_{M+2} & \cdots & x_N
\end{bmatrix}
\]

Let \( x_1, \ldots, x_M \), and \( x_{N-M+1}, \ldots, x_N \) be padded zeros, the sample covariance matrix \( XX^T \) is a Toeplitz matrix.

When the autocorrelation of the signal decays fast, the discrete cosine transform (DCT) basis functions [2] approximate well the eigenfunctions of the Toeplitz matrix.

Short time discrete cosine transform

\[
\sum_{n=\infty}^{\infty} x(m-n)w(n) \cos(\omega(m-n))
\]

\[
= \sum_{n=\infty}^{\infty} x(m-n)w(n)[\cos(\omega m) \cos(\omega n) + \sin(\omega m) \sin(\omega n)]
\]

\[
= \cos(\omega m)[(w(m) \cos(\omega n)) \ast x(m)]... + \sin(\omega m)[(w(m) \sin(\omega n)) \ast x(m)]
\]

It can be viewed as the sum of modulation of two band-pass filters.

Results

Data: DCLDE (blue whale and fin whale vocalization);
Approach: PCANet/DCTNet + Laplacian Eigenmap.