A Causal Locally Competitive Algorithm for the Sparse Decomposition of Audio Signals

Adam Charles, Abbie Kressner & Christopher Rozell
Georgia Institute of Technology

2011 Digital Signal Processing Workshop
Audio Coding

- Standard coding: Fourier/Wavelet
- Modern processing uses sparsity
- Sparse audio decompositions:

\[ x(t) = \sum_i \sum_m s_i^m \phi_i(t - \tau_i^m) \]

- Make \( s_i^m \) sparse!
Motivation

- Potential applications for sparse inference
  - Audio coding
  - Audio enhancement
  - Hearing aids and cochlear implants
- How can we find $s_i^m$?
Matching Pursuit

- Vector-Matrix form
  \[ x = \Phi a \]

- Algorithm
  - Pick best \( a_i \) at time \( n \)
  - Calculate the residual
  - Repeat

(Mallat and Zhang 1993)
Filter and Threshold

(Smith and Lewicki, 2005)
LCA Structure

Use feedback to sparsify outputs while retaining signal integrity:

(Rozell et. al. 2008)
LCA Dynamics

1) \( \dot{u}_i(t) = \frac{1}{T} \left( \langle x, \phi_i \rangle - u_i(t) - z_i(t) \right) \)
   where \( z_i(t) = \sum_{j \neq i} \langle \phi_i, \phi_j \rangle a_j(t) \)

2) \( a_i(t) = T_\lambda (u_i(t)) \)

![Soft Threshold](image1.png)

![Hard Threshold](image2.png)
Correlations

Figure: Basis Correlation Functions for $\phi_3$, $\phi_4$ and $\phi_6$
Figure: Track correlations through **space & time**
Figure: Recently written coefficients continue inhibiting
1. Read new sample and move sliding window

2. Allow LCA to converge at time $n$,

$$2a) \dot{u}_i(t) = \frac{1}{T} \left( \langle x, \phi_i \rangle - u_i(t) - z_i(t) \right)$$

where $z_i(t) = \sum_k \langle \phi_i, \phi_k \rangle \hat{a}_k + \sum_{j \neq i} \langle \phi_i, \phi_j \rangle a_j(t)$

3. Write last coefficients in the buffer and move all other values back a timestep

$$2b) a_i(t) = T_{\lambda} (u_i(t))$$
Rate Distortion Curve

![Rate Distortion Curve Graph]

- **CLCA - soft**
- **CLCA - hard**
- **FT**
- **MP**

**X-axis:** Signal to Noise Ratio (dB)

**Y-axis:** Number of coefficients
Conclusions

- Sparsity with causality
- Analog system: low power and real-time (50KHz)
- 10ms window: within lip sync tolerance
Introduction

Causal LCA

Results & Conclusions

Spikegrams: Speech Signal

Rate Distortion Curve

Conclusions

**Spikegrams**

![Spikegram Image]

**Matching Pursuit**
- 15 dB SNR
- 17 coefficients

**Filter and threshold**
- 7 dB SNR
- 84 coefficients

**Causal LCA**
- 15 dB SNR
- 162 coefficients