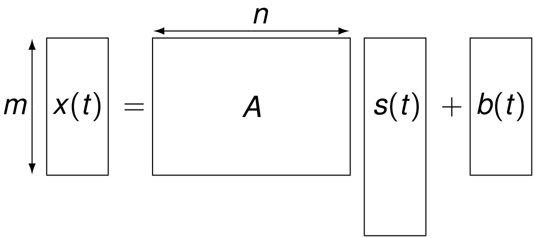
Robust Underdetermined Blind Audio Source Separation of Sparse Signals in the Time-Frequency Domain

S.M. Aziz Sbaï, A. Aïssa-El-Bey and D. Pastor

May 26 2011

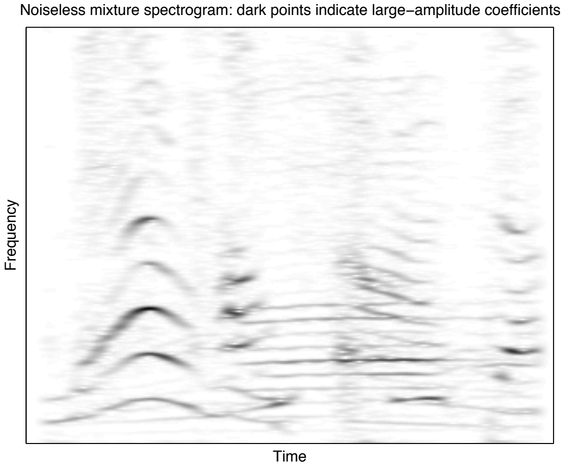
Main Problem

- Underdetermined system of linear equations

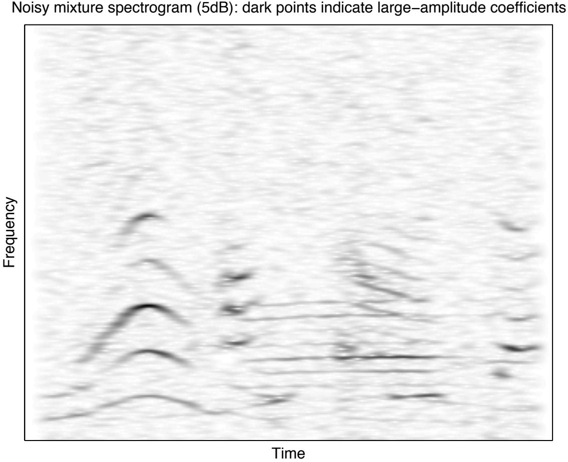


- Instantaneous mixing case with independent AWGN
- The n components of $s(t)$ are sparse in the time frequency domain (in cont. of [Aïssa-El-Bey, IEEE-IT, 2007], [Bofill, Zibulevsky, Elsevier SP, 2001], [Linh-Trung et al, JASP, 2005])
- See survey [Pedersen et al, Springer Press, 2007]

Weak sparseness



Weak sparseness



Weak sparseness

Hypotheses [P., IEEE-IT 2002]

1. Signal components are either present or absent in the transformed domain (Fourier, wavelet, $\ldots$) with a probability of presence $\leqslant 1/2$,
2. When present, signal components are *relatively big* in that their amplitude is above some *minimum amplitude* $\rho > 0$.

Comments

- weak sparseness = constraints that bound our lack of prior knowledge on the signal distributions
- weak sparseness slightly differs from sparsity in compressed sensing

Outline

1. Proposed Algorithm: Robust Underdetermined Blind Source Separation (RUBSS)
2. Experimental results
3. Conclusions and perspectives

Hypotheses

- Any $m \times m$ submatrix of the mixing matrix has full rank, that is , for all $J \subset \{1, 2, \cdots, n\}$ of cardinality less than or equal to m, the column vectors $(A_j)_{j \in J}$ are linearly independent.

$$\begin{array}{cc} & \overset{n}{\longleftrightarrow} \\ m \updownarrow & \begin{pmatrix} A_1 & A_2 & \cdots & A_n \end{pmatrix} \end{array}$$

Hypotheses

- Any $m \times m$ submatrix of the mixing matrix has full rank, that is , for all $J \subset \{1, 2, \cdots, n\}$ of cardinality less than or equal to m, the column vectors $(A_j)_{j \in J}$ are linearly independent.
- The number of active sources at any time-frequency point is strictly less than the number m of sensors.

Time-Frequency model :

$$X(t, f) = A_J S_J(t, f) + B(t, f)$$

J : indexes of the active sources at (t, f)

Procedure (extends [Aïssa-El-Bey et al., IT-SP, 2007])

1. Compute STFT of the mixtures ;

Procedure (extends [Aïssa-El-Bey et al., IT-SP, 2007])

1. Compute STFT of the mixtures ;
2. Estimate the noise standard deviation via the Modified Complex Essential Supremum Estimate (MC-ESE);

[Socheleau, P. and Aïssa-El-Bey, IEEE-AES, 2011] :

[P., CSDA, 2008] (weak sparseness) $\rightarrow$ for coord. # k :

$$\frac{\sum_{t,f} |X_k(t,f)| \mathbb{1}(|X_k(t,f)| \leqslant \sigma_0 \tau)}{\sum_{t,f} \mathbb{1}(|X_k(t,f)| \leqslant \sigma_0 \tau)} - \sigma_0 \frac{\int_0^\tau t^2 e^{-t^2/2} \mathrm{d}t}{1 - e^{-\tau^2/2}} \approx 0$$

Procedure (extends [Aïssa-El-Bey et al., IT-SP, 2007])

1. Compute STFT of the mixtures ;
2. Estimate the noise standard deviation via the Modified Complex Essential Supremum Estimate (MC-ESE);
3. Reject the time frequency points that correspond to observations of noise alone ;

Thresholding test:

$$\mathcal{T}(X(t,f)) = \begin{cases} 1 & \text{if } \|X(t,f)\|_\infty \geq \widehat{\sigma_0}\sqrt{-2\ln \mathcal{P}_{FA}} \\ 0 & \text{if } \|X(t,f)\|_\infty < \widehat{\sigma_0}\sqrt{-2\ln \mathcal{P}_{FA}} \end{cases}$$

$$u = (u_1, \cdots, u_m)^T \quad \|u\|_\infty = \max_i\{|u_i|\}$$

Procedure (extends [Aïssa-El-Bey et al., IT-SP, 2007])

1. Compute STFT of the mixtures ;
2. Estimate the noise standard deviation via the Modified Complex Essential Supremum Estimate (MC-ESE);
3. Reject the time frequency points that correspond to observations of noise alone ;
4. Estimate STFT of the sources at the accepted points ;
   - Identification of active sources

Identification:

- Compute $P_{\mathcal{J}} = I_m - A_{\mathcal{J}}(A_{\mathcal{J}}^H A_{\mathcal{J}})^{-1} A_{\mathcal{J}}^H$ where $\#(\mathcal{J}) < m$.
- Compute $\mathcal{K} = \arg\min_{\mathcal{J}} \{\|P_{\mathcal{J}} X(t,f)\|\}$.

Procedure (extends [Aïssa-El-Bey et al., IT-SP, 2007])

1. Compute STFT of the mixtures ;
2. Estimate the noise standard deviation via the Modified Complex Essential Supremum Estimate (MC-ESE);
3. Reject the time frequency points that correspond to observations of noise alone ;
4. Estimate STFT of the sources at the accepted points ;
   - Identification of active sources
   - Linear estimation

Linear estimation (Wiener filtering):

$$\hat{S}_{\mathcal{J}}(t,f) \approx \widehat{\mathbf{R}}_{\mathcal{J}} A_{\mathcal{J}}^{H} (A_{\mathcal{J}} \widehat{\mathbf{R}}_{\mathcal{J}} A_{\mathcal{J}}^{H} + \widehat{\sigma_0}^2 \mathbf{I}_m)^{-1} X(t,f)$$

$\widehat{\mathbf{R}}_{\mathcal{J}}$: empirical correlation matrix of S_J

Procedure (extends [Aïssa-El-Bey et al., IT-SP, 2007])

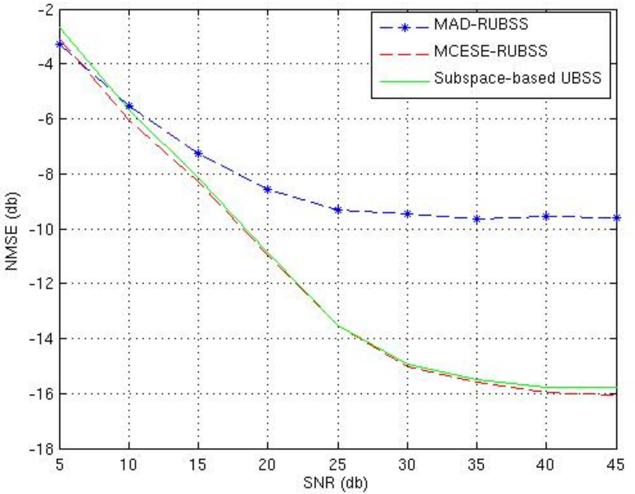
1. Compute STFT of the mixtures ;
2. Estimate the noise standard deviation via the Modified Complex Essential Supremum Estimate (MC-ESE);
3. Reject the time frequency points that correspond to observations of noise alone ;
4. Estimate STFT of the sources at the accepted points ;
   - Identification of active sources
   - Linear estimation
5. Inverse STFT to recover the sources in the time domain.

Procedure

1. Compute STFT of the mixtures ;
2. Estimate the noise standard deviation with MC-ESE ;
3. Reject the time frequency points that correspond to observations of noise alone ;
4. Estimate STFT of the sources at the accepted points ;
   - Identification of active sources
   - Linear estimation
5. Inverse STFT to recover the sources in the time domain.

[Our contribution] [Aïssa-El-Bey et al., IT-SP, 2007]

Source separation performance: NMSE vs SNR ($P_{FA} = 10^{-3}$)



Source separation performance: Audio Results

Mixtures

- Mixture 1
- Mixture 2
- Mixture 3

Original signals

- Sound 1
- Sound 2
- Sound 3
- Sound 4

Estimated signals

- Estimated sound 1
- Estimated sound 2
- Estimated sound 3
- Estimated sound 4

Conclusions and perspectives

Conclusions

- The role of sparseness
- Only one parameter
- No prior knowledge on the exact nature of the sources

Perspectives

- Computational cost of MC-ESE $\Longrightarrow$ DATE algorithm ;
- False alarm rate dependence $\Longrightarrow$ Alternative approaches in robust and non parametric detection;
- The convolutive mixing case