TECHNICAL UNIVERSITY OF LIBEREC
Faculty of Mechatronics, Informatics and Interdisciplinary Studies

SKFM 2019
Studentská konference Fakulty mechatroniky, informatiky a mezioborových studií

Online Implementation of Independent Vector Extraction based on Auxiliary function (AuxIVE)

Jakub Janský
jakub.jansky@tul.cz

Jaroslav Čmejla
jaroslav.cmejla@tul.cz

Tomáš Kounovský
tomas.kounovsky@tul.cz

Abstract

Separation of the sources from the mixture is an interesting field of audio processing. This paper is focusing only on the extraction of one source from a mixture. We propose a novel approach, how to achieve this task in online processing. To avoid unpractical gradient method we introduce an auxiliary function technique.

Problem definition

The Short-time Fourier Transform (STFT) model of a determined instantaneous mixture of M sources is

$$\mathbf{X}_{k,\ell} = \mathbf{A}_k \mathbf{S}_{k,\ell} = \mathbf{a}_k s_{k,\ell} + \mathbf{Y}_{k,\ell}, \quad (1)$$

- $\mathbf{S}_{k,\ell}$ is $(M \times 1)$ vectors of the original signals,
- $\mathbf{X}_{k,\ell}$ is $(M \times 1)$ vectors of the mixed signals,
- $\mathbf{A}_k$ denotes $(M \times 1)$ mixing matrix,
- $k = 1, \dots, K$ denote frequency bin,
- $\ell = 1, \dots, L$ denote frame index.

To obtain separated sources from the mixture we are looking for demixing matrix $\mathbf{W}_k$, which fulfils

$$\mathbf{W}_k \mathbf{X}_{k,\ell} = \mathbf{W}_k \mathbf{A}_k \mathbf{S}_{k,\ell} \approx \mathbf{S}_k(\omega, \ell). \quad (2)$$

We are looking only for $s_{k,\ell}$, thus $\mathbf{W}_k$ needs to be partitioned as [1]

$$\mathbf{W}_k = \begin{pmatrix} \mathbf{w}_k^H \\ \mathbf{B}_k \end{pmatrix} = \begin{pmatrix} \overline{\beta_k} & \mathbf{h}_k^H \\ \mathbf{g}_k & -\gamma_k \mathbf{I}_{M-1} \end{pmatrix}, \quad (3)$$

where $\mathbf{w}_k^H$ is mixing vector of $s_{k,\ell}$ and $\mathbf{B}_k$ corresponds to $\mathbf{Y}_{k,\ell}$
$\mathbf{A}_k$ can be also partitioned as $\mathbf{A}_k = \begin{pmatrix} \mathbf{a}_k & \mathbf{D}_k \end{pmatrix}$.
Note that $\mathbf{B}_k \mathbf{a}_k = 0$ and $\mathbf{w}_k^H \mathbf{a}_k = 1$ is mandatory

We can assume that $s_{k,\ell}$ and $\mathbf{Y}_{k,\ell}$ are mutually independent so we can utilise a log-likelihood function

$$\mathcal{L}(\mathbf{w}_k, \mathbf{a}_k) = \log p_s(\mathbf{w}_k^H \mathbf{X}_{k,\ell}) - \log p_z(\mathbf{B}_k \mathbf{X}_{k,\ell}) + \log |\det \mathbf{W}_k|^2. \quad (4)$$

- $p_s(\cdot)$ is unknown $\rightarrow$ replaced by model pdf $f(\cdot)$
- $p_z(\cdot)$ is unknown $\rightarrow$ assumed to be circular Gaussian
- minimisation in both $\mathbf{a}_k$ and $\mathbf{w}_k$ is unpractical and $\mathbf{w}_k^H \mathbf{a}_k = 1$ is weak bond $\rightarrow$ introduce orthogonal constraint $\mathbf{a}_k = \frac{\mathbf{C}_\mathbf{x} \mathbf{w}_k}{\mathbf{w}_k^H \mathbf{C}_\mathbf{x} \mathbf{w}_k}$

All changes leads to contrast function

$$\mathbf{J}(\mathbf{w}_k) = \mathrm{E}\left[f(\mathbf{w}_k^H \mathbf{X}_{k,\ell})\right] - E\left[\mathbf{X}_{k,\ell}^H \mathbf{B}_k^H \hat{\mathbf{C}}_\mathbf{z}^{-1} \mathbf{B}_k \mathbf{X}_{k,\ell}\right] + (d-2)\log|\gamma_k|^2, \quad (5)$$

Can be solved by gradient algorithm [1] $\rightarrow$ hard to find proper gradient step size

Online AuxIVE

To avoid gradient method $\rightarrow$ introducing auxiliary variable $\mathbf{V}_k$. THe minimum is obtained by optimising contrast function in normal and auxiliary variable alternately. The first term of (5) can be modified in terms of auxiliary function as [2]

$$E\left[f(\mathbf{w}_k^H \mathbf{X}_{k,\ell})\right] \le E\left[\frac{f'(r)}{2r} \mathbf{w}_k^H \mathbf{X}_{k,\ell} \mathbf{X}_{k,\ell}^H \mathbf{w}_k\right] = \frac{1}{2}\mathbf{w}_k^H E\left[\frac{f'(r)}{r} \mathbf{X}_{k,\ell} \mathbf{X}_{k,\ell}^H\right] \mathbf{w}_k = \frac{1}{2}\mathbf{w}_k^H \mathbf{V}_k \mathbf{w}_k. \quad (6)$$

In the online scenario:

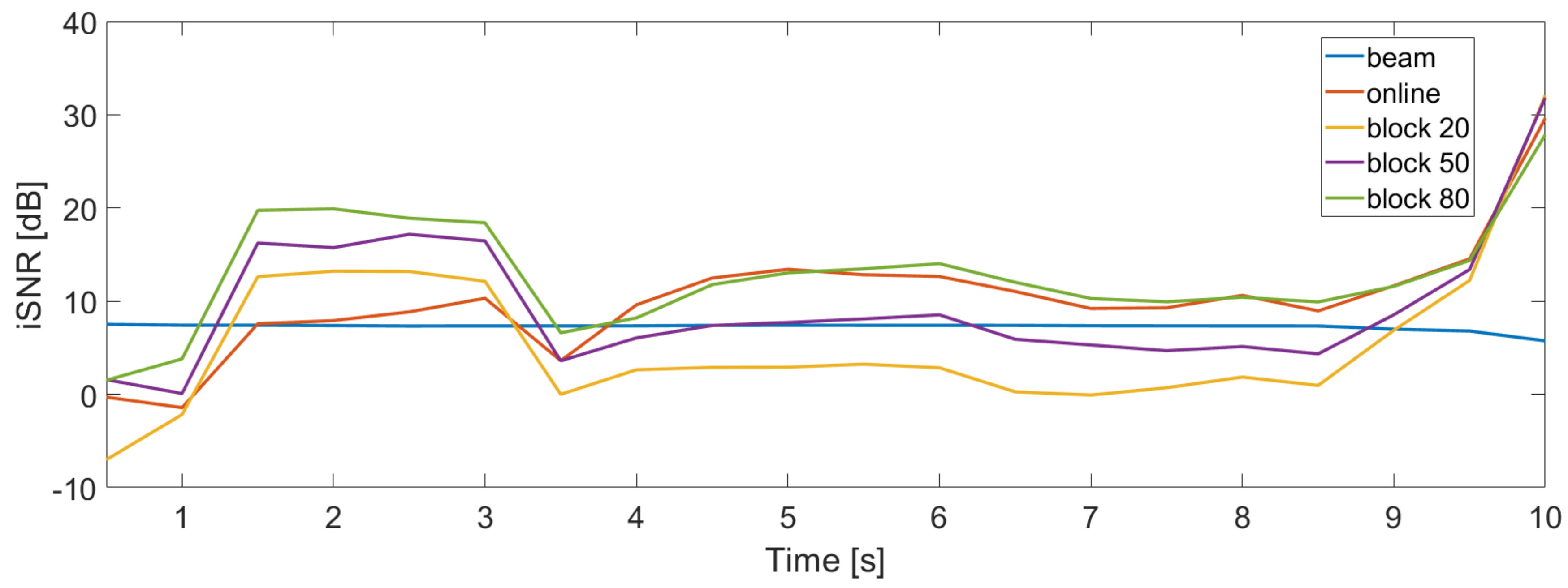
- source can change position $\rightarrow$ separation only over few frames
- demixing vector $\mathbf{w}_k$ can extremely change $\rightarrow$ past computation is included with forgetting factor α

By deriving modified (5) in $\mathbf{V}_k$ and $\mathbf{w}_k$ $\rightarrow$ sufficient update rules

1. $\hat{\mathbf{C}}_{\mathbf{x},b} = \alpha \hat{\mathbf{C}}_{\mathbf{x},b-1} + (1-\alpha) E_b\left[\mathbf{X}_{k,b}^H \mathbf{X}_{k,b}\right]$
2. $\mathbf{a}_k = \frac{\hat{\mathbf{C}}_{\mathbf{x},b} \mathbf{w}_k}{\mathbf{w}_k^H \hat{\mathbf{C}}_{\mathbf{x},b} \mathbf{w}_k}$
3. $r_b = \|\mathbf{w}_k^H \mathbf{X}_{k,b}\|_2$
4. $\mathbf{V}_{k,b} = \alpha \mathbf{V}_{k,b-1} + (1-\alpha) E_b\left[\frac{f(r_b)'}{r_b} \mathbf{X}_{k,b}^H \mathbf{X}_{k,b}\right]$
5. $\mathbf{w}_k^H = \mathbf{a}_k^H \mathbf{V}_{k,b}^{-1}$

Experimental setup

- 5 microphones mixture of female speaker (1m in front of microphones) and white noise (1m in 70 degree angle)
- different frame size $b = [1; 20; 50; 80]$



Conclusion

The online version of the Independent Vector Analysis based on Auxiliary function is presented in this paper. In the experimental part, we compare online proposed a method with a batch online method with variable frame block size. The online method has shown a promising result. However, more research is needed in this area.

References

[1] P. Tichavský Z. Koldovský.
Gradient algorithms for complex non-gaussian independent component/vector extraction, question of convergence.
In *IEEE Transactions on Signal Processing*, volume 67, pages 1050–1064, 2019.

[2] N. Ono.
Stable and fast update rules for independent vector analysis based on auxiliary function technique.
In *2011 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 189–192, 2011.