Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

A novel blind deconvolution algorithm using single frequency bin*

ZHANG Gui-bao†, LI Jia-wen, LI Cong-xin

(*National Die & Mold CAD Engineering Research Center, Shanghai Jiao Tong University, Shanghai 200030, China*)

†E-mail: gbzhangf2@sjtu.edu.cn

Received Nov. 29, 2006; revision accepted Apr. 29, 2007

Abstract: Former frequency-domain blind devolution algorithms need to consider a large number of frequency bins and recover the sources in different orders and with different amplitudes in each frequency bin, so they suffer from permutation and amplitude indeterminacy troubles. Based on sliding discrete Fourier transform, the presented deconvolution algorithm can directly recover time-domain sources from frequency-domain convolutive model using single frequency bin. It only needs to execute blind separation of instantaneous mixture once there are no permutation and amplitude indeterminacy troubles. Compared with former algorithms, the algorithm greatly reduces the computation cost as only one frequency bin is considered. Its good and robust performance is demonstrated by simulations when the signal-to-noise-ratio is high.

Key words: Blind deconvolution, Single frequency bin, Convolutive mixture
doi:10.1631/jzus.2007.A1271 **Document code:** A **CLC number:** TN911

INTRODUCTION

The blind separation of convolutive signals (generally named multichannel blind deconvolution) has received attentions in extensive fields such as wireless communications, speech enhancement, biomedical signal analysis, etc. (Peled *et al*., 2005; Unger and Zeevi, 2006; Throckmorton *et al*., 2007). Its algorithms can be mainly classified as operating in time-domain and operating in frequency domain. The time-domain algorithms include information theory solutions, higher-order cumulant solutions, etc. Because convolution mixture has close relationship with instantaneous mixture, Sun and Scott (2001) pointed out that time-domain blind separation algorithms of instantaneous mixture could be extended to solve blind deconvolution problems. The time-domain deconvolution algorithms have good performances when the filter length of the mixing matrix $\boldsymbol{A}$ is short. With filter length increasing, the algorithm's performances degrade.

The frequency-domain blind deconvolution algorithms transform a convolutive mixture problem into multiple instantaneous mixture problems which can be solved by many algorithms. But traditional frequency-domain algorithms consider a large number of frequency bins and need to execute blind separation of instantaneous mixture many times, which results in permutation and amplitude indeterminacy problems. Several approaches have been proposed to solve the indeterminacy problem: Smaragdis (1998) exploited the frequency coupling relationship between neighboring frequency bins; Kurtia and Saruwatari (2000) used beamforming to estimate the directions of signal arrival (DOA); Parra and Spence (2000) imposed a smooth constraint on the unmixing filter length and applied a projection operator to the filter estimation at each iteration. All the above approaches exploit a large number of frequency bins, therefore increase the computation cost. Based on sliding Fourier transform, Depena *et al*. (2003) exploited two frequency bins to separate convolutive mixture signals. Its computation complexity is lower than the former approaches, but it still needs to solve the permutation and amplitude indeterminacy problems of the two selected frequency bins.

* Project (No. 2005EB040486) supported by the National Torch Program of China

Based on sliding discrete Fourier transform (sliding DFT), the novel blind deconvolution approach proposed in this paper can directly recover the time-domain sources $S(t)$ from their frequency-domain convolutive model and does not suffer from indeterminacy trouble as only one frequency bin is considered. Compared with the former algorithms, the new approach greatly reduces the computation complexity, and has good and robust performance which is demonstrated by simulations of separating audio signals.

FREQUENCY-DOMAIN BLIND SEPATATION MODEL

The general blind source separation problem of convolutive mixture can be described as follows:

$$\boldsymbol{X}(t)=\sum_{k=0}^{K}\boldsymbol{A}(k)\boldsymbol{S}(t-k), \tag{1}$$

where $\boldsymbol{S}(t)=[s_1(t), s_2(t), \ldots, s_n(t)]^{\mathrm{T}}$ (t=1, 2, …, T) is the source signal vector which is real-value, non-gaussian distributed and statistically independent, $\boldsymbol{X}(t)=[x_1(t), x_2(t), \ldots, x_n(t)]^{\mathrm{T}}$ is the observed signal vector, $\boldsymbol{A}(k)$ is the mixing matrix filter, K denotes the maximum filter length.

According to the properties of discrete Fourier transform, the result of linear convolution is equal to that of circular convolution, if the length L of Fourier transform is much greater than the filter length K. Thus the time-domain convolutive model (1) can be transformed to a frequency-domain instantaneous mixture model by using an L-point short time Fourier transform (Sun and Scott, 2001; Mitianoudis and Davies, 2003):

$$\boldsymbol{X}(\omega_k,t)=\boldsymbol{A}(\omega_k)\boldsymbol{S}(\omega_k,t),\ k=0, 1, 2, \ldots, L-1, \tag{2}$$

where $\boldsymbol{X}(\omega_k, t)$, $\boldsymbol{S}(\omega_k, t)$ are the frequency-domain transform elements of $\boldsymbol{X}(t)$ and $\boldsymbol{S}(t)$ respectively at the frequency ω_k and the time t, $\boldsymbol{A}(\omega_k)$ is the frequency-domain element of the mixing matrix $\boldsymbol{A}$ at the frequency bin ω_k.

For most of the frequency-domain approaches, the unmixing process can be formulated as follows:

$$\boldsymbol{Y}(\omega_k,t)=\boldsymbol{W}(\omega_k)\boldsymbol{X}(\omega_k,t),\ k=0, 1, 2, \ldots, L-1, \tag{3}$$

where $\boldsymbol{W}(\omega_k)$ represents the unmixing matrix of $\boldsymbol{A}(\omega_k)$. To get the estimation of $\boldsymbol{S}(t)$, former approaches need to execute blind source separation of instantaneous mixture many times, thus suffer from the permutation and amplitude troubles. However, the developed approach in this paper only needs to run blind instantaneous mixture separation once by applying the sliding DFT.

SLIDING DISCRETE FOURIER TRANSFORM

Different from the sliding DFT proposed by Jacobsen and Lyons (2003), Depena *et al.*(2003) proposed a new sliding DFT by applying the L-point discrete Fourier transform with moving windows of size P with their right $L-P$ points padded with zero, that is to say, $\boldsymbol{S}(\omega_k, t)$ of L-point DFT is computed by the new time sequence $\boldsymbol{S}(t)$, $\boldsymbol{S}(t+1)$, $\boldsymbol{S}(t+2)$, …, $\boldsymbol{S}(t+P-1)$, 0, 0, …, 0:

$$\boldsymbol{S}(\omega_k,t)=\sum_{m=0}^{P-1}\boldsymbol{S}(t+m)\exp(-\mathrm{j}\omega_k m),\ \omega_k=2\pi k/L. \tag{4}$$

Therefore, the $\boldsymbol{S}(\omega_k, t+1)$ of L-point DFT at the given frequency bin ω_k and time $t+1$ can be formulated as:

$$\boldsymbol{S}(\omega_k,t+1)=\sum_{m=0}^{P-1}\boldsymbol{S}(t+1+m)\exp(-\mathrm{j}\omega_k m). \tag{5}$$

The following formula can be obtained by Eq.(4)–Eq.(5)·exp(−jω_k):

$$\begin{aligned}&\boldsymbol{S}(\omega_k,t)-\boldsymbol{S}(\omega_k,t+1)\exp(-\mathrm{j}\omega_k)\\&\quad=\boldsymbol{S}(t)-\boldsymbol{S}(t+P)\exp(-\mathrm{j}\omega_k P).\end{aligned} \tag{6}$$

Considering the overlapping consecutive sliding window, the $\boldsymbol{S}(\omega_k, t+1)$ can be simply inferred from $\boldsymbol{S}(\omega_k, t)$. Furthermore, the $\boldsymbol{S}(\omega_k, t)$ and $\boldsymbol{S}(\omega_k, t+1)$ need not be calculated to get the value of $\boldsymbol{S}(\omega_k, t)-\boldsymbol{S}(\omega_k, t+1)\exp(-\mathrm{j}\omega_k)$, because $\boldsymbol{S}(\omega_k, t)-\boldsymbol{S}(\omega_k, t+1)\exp(-\mathrm{j}\omega_k)=\boldsymbol{S}(t)-\boldsymbol{S}(t+P)\exp(-\mathrm{j}\omega_k P)$. Namely, the frequency-domain value $\boldsymbol{S}(\omega_k,t)-\boldsymbol{S}(\omega_k,t+1)\exp(-\mathrm{j}\omega_k)$ can be directly calculated from the time-domain signal $\boldsymbol{S}(t)$ and $\boldsymbol{S}(t+P)$ through one subtraction and one multiplication operation.

DECONVOLUTION USING SINGLE FREQUENCY BIN

According to Eq.(6), the following Eqs.(7) and (8) can be obtained by computing the L-point sliding DFT of the source signal $\boldsymbol{S}(t)$ and the observation signal $\boldsymbol{X}(t)$ at the given frequency bin ω_k and time t:

$$\boldsymbol{S}(\omega_k,t)-\boldsymbol{S}(\omega_k,t+1)\mathrm{e}^{-\mathrm{j}\omega_k}=\boldsymbol{S}(t)-\boldsymbol{S}(t+P)\exp(-\mathrm{j}\omega_k P), \quad (7)$$

$$\boldsymbol{X}(\omega_k,t)-\boldsymbol{X}(\omega_k,t+1)\mathrm{e}^{-\mathrm{j}\omega_k}=\boldsymbol{X}(t)-\boldsymbol{X}(t+P)\exp(-\mathrm{j}\omega_k P). \quad (8)$$

Letting $\boldsymbol{Z}(\omega_k, t)=\boldsymbol{X}(\omega_k, t)-\boldsymbol{X}(\omega_k, t+1)\exp(-\mathrm{j}\omega_k)$ for explicitness, considering model (2), the relationship between the frequency-domain signals $\boldsymbol{Z}(\omega_k, t)$ and time-domain signals $\boldsymbol{X}(t)$ and $\boldsymbol{S}(t)$ is:

$$\boldsymbol{Z}(\omega_k,t)=\boldsymbol{X}(t)-\boldsymbol{X}(t+P)\exp(-\mathrm{j}\omega_k P), \quad (9)$$

and

$$\begin{aligned}\boldsymbol{Z}(\omega_k,t)&=\boldsymbol{A}(\omega_k)[\boldsymbol{S}(\omega_k,t)-\boldsymbol{S}(\omega_k,t+1)\exp(-\mathrm{j}\omega_k)]\\&=\boldsymbol{A}(\omega_k)[\boldsymbol{S}(t)-\boldsymbol{S}(t+P)\exp(-\mathrm{j}\omega_k P)]. \quad (10)\end{aligned}$$

When the inverse of the mixing matrix $\boldsymbol{A}(\omega_k)$ exists, Eq.(10) can be formulated as:

$$\boldsymbol{S}(t)-\boldsymbol{S}(t+P)\exp(-\mathrm{j}\omega_k P)=\boldsymbol{A}^{-1}(\omega_k)\boldsymbol{Z}(\omega_k,t). \quad (11)$$

Regarding $\boldsymbol{Z}(\omega_k ,t)$ as virtual observation signals and $\boldsymbol{S}'(t)=\boldsymbol{S}(t)-\boldsymbol{S}(t+P)\exp(-\mathrm{j}\omega_k P)$ as virtual recovered sources, Eq.(11) implies that $\boldsymbol{S}'(t)$ can be obtained by running instantaneous mixing blind separation only once based on the given frequency bin ω_k, so $\boldsymbol{S}'(t)$ has no permutation and amplitude indeterminacy troubles because $\boldsymbol{S}'(t)$ has relationship with only one frequency bin, which shows a prominent elevation to the former frequency-domain blind deconvolution algorithms.

Equation $\boldsymbol{S}'(t)=\boldsymbol{S}(t)-\boldsymbol{S}(t+P)\exp(-\mathrm{j}\omega_k P)$ can be written in another form:

$$\boldsymbol{S}(t+P)=[\boldsymbol{S}(t)-\boldsymbol{S}'(t)]\exp(\mathrm{j}\omega_k P),\ t=1,2,...,T-P, \quad (12)$$

here ω_k and P are given, and $\boldsymbol{S}'(t)$ can be calculated by running instantaneous mixture blind separation from the virtual observation $\boldsymbol{Z}(\omega_k , t)$, therefore, $\boldsymbol{S}(t+P)$ can be calculated iteratively from $\boldsymbol{S}(t)$ if the initial values of $\boldsymbol{S}(1), \boldsymbol{S}(2), \ldots, \boldsymbol{S}(P)$ are known.

$\boldsymbol{S}(1), \boldsymbol{S}(2), \ldots, \boldsymbol{S}(P)$ can be simply assumed to be zero and $\boldsymbol{X}(1), \boldsymbol{X}(2), \ldots, \boldsymbol{X}(P)$ are also set to zero no matter what the value of $\boldsymbol{A}$ is, this assumption will introduce some errors to $\boldsymbol{X}(P+1), \boldsymbol{X}(P+2), \ldots, \boldsymbol{X}(P+K)$ according to model (1), because $\boldsymbol{S}(1), \boldsymbol{S}(2), \ldots, \boldsymbol{S}(P)$ are not zero in fact. However, these errors have little bad effect on the algorithm's performance and can be neglected, because the linear convolution and circular convolution can be regarded as the same if the window width P is greatly larger than the filter length K, i.e., $L>P>>K$.

The algorithm can be summarized as follows:

(1) Parameter setting: set suitable parameters L, P and k;

(2) Initializing: set the values of $\boldsymbol{S}(0), \boldsymbol{S}(1), \ldots, \boldsymbol{S}(P-1)$ to zero, set the values of $\boldsymbol{X}(0), \boldsymbol{X}(1), \ldots, \boldsymbol{X}(P)$ to zero, and $\omega_k=2\pi k/L$;

(3) Filtering: compute $\boldsymbol{Z}(\omega_k, t)$ according to Eq.(9);

(4) Blind separation: regard $\boldsymbol{Z}(\omega_k, t)$ as a virtual observation of an instantaneous mixing system, run blind source separation to get $\boldsymbol{S}'(t)$;

(5) Inverse filtering: compute the source $\boldsymbol{S}(t+P)$ iteratively according to Eq.(12).

SIMULATION

In this section three simulation results are presented to demonstrate the performance of the single frequency bin deconvolution algorithm. The sources used in these simulations are two clear audio voices of a man and a woman with 50000 samples and 16 kHz sample rates. The mixing matrix $\boldsymbol{A}$ is generated randomly as follows:

$$\boldsymbol{A}(z)=\begin{bmatrix}a_{11}(z) & a_{12}(z)\\ a_{21}(z) & a_{22}(z)\end{bmatrix},$$

$$a_{11}(z)=0.9+0.5z^{-1}+0.2z^{-2}+0.2z^{-3}+0.1z^{-4},$$
$$a_{12}(z)=-0.3+0.8z^{-1}+0.4z^{-2}+0.1z^{-3}+0.3z^{-4},$$
$$a_{21}(z)=-1.2-0.7z^{-1}+0.1z^{-2}+0.3z^{-3}+0.2z^{-4},$$
$$a_{22}(z)=-0.2+0.8z^{-1}+0.5z^{-2}+0.1z^{-3}+0.5z^{-4},$$

where $\boldsymbol{A}(z)$ is the z-transform form of $\boldsymbol{A}$. Many instantaneous blind source separation algorithms can be used to recover the sources. Among them, the robust JADE (Joint Approximate Diagonalization of Eigenmatrices) algorithm (Cardoso and Souloumiac, 1993) is selected (downloaded from http://www.tsi.enst.fr/~cardoso/guidesepsou.html) to separate the virtually observed signals $\boldsymbol{Z}(\omega_k)$. Meanwhile, some noises are imported for simulating the real audio environment:

$$\boldsymbol{X}(t)=\sum_{k=0}^{4}\boldsymbol{A}(k)\boldsymbol{S}(t-k)+\boldsymbol{\sigma}. \tag{13}$$

Evaluating the performance of blind source separation or blind deconvolution, two kinds of performance indexes based on mixing matrix and signal distortion can be exploited (Peled *et al*., 2005). In this paper, a correlation coefficient C is introduced as a performance index, which is based on signal distortion to evaluate the algorithm's capability quantificationally:

$$C(\boldsymbol{s}_1,\boldsymbol{s}_2)=\frac{\mathrm{cov}(\boldsymbol{s}_1,\boldsymbol{s}_2)}{\sqrt{\mathrm{cov}(\boldsymbol{s}_1,\boldsymbol{s}_1)\cdot\mathrm{cov}(\boldsymbol{s}_2,\boldsymbol{s}_2)}}, \tag{14}$$

where cov() is the covariance function, $\mathrm{cov}(\boldsymbol{s}_1, \boldsymbol{s}_2)=E\{[\boldsymbol{s}_1-E(\boldsymbol{s}_1)]^{\mathrm{T}}[\boldsymbol{s}_2-E(\boldsymbol{s}_2)]\}$, $E(\cdot)$ is the mean operator. $C(\boldsymbol{s}_1, \boldsymbol{s}_2)=0$ means that $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ are uncorrelated, and the signals correlation increases as $C(\boldsymbol{s}_1, \boldsymbol{s}_2)$ approaches unity, the signals become fully correlated as C becomes unity. In our simulations the correlation coefficients between actually observed signal $\boldsymbol{X}$ and source $\boldsymbol{S}$ are $C(\boldsymbol{x}_1, \boldsymbol{s}_2)=0.63$, $C(\boldsymbol{x}_2, \boldsymbol{s}_2)=0.54$ respectively. They are very low, which means that $\boldsymbol{X}$ and $\boldsymbol{S}$ are uncorrelated.

In the first experiment, letting $P=64$, $L=1\,024$, $SNR=30$ dB, and randomly selecting the frequency bin $k=5$ to run JADE for virtual observation of $\boldsymbol{Z}(\omega_k)$. After separation, the correlation coefficients between the recovered signals and the sources increase from 0.63 to 0.98 and from 0.54 to 0.96 respectively. As shown in Fig.1, the sources are well recovered. This result demonstrates that convolutive signals can be separated using single frequency bin. More importantly, the permutation indeterminacy and amplitude indeterminacy will not affect the final separation performance because of the use of a single frequency bin. On the contrary, the previous frequency-domain algorithms used several frequency bins and run blind separation of instantaneous mixture several times which lead to permutation and amplitude problems.

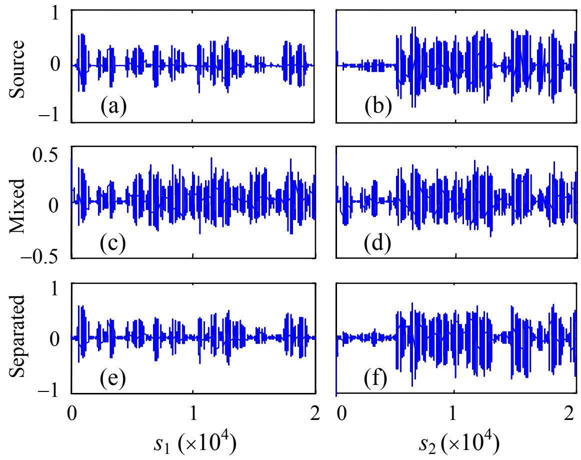

Fig.1 The time-domain waveforms of blind deconvolution at P=64, L=1 024. (a) Signal-1; (b) Signal-2; (c) First mix of signals; (d) Second mix of signals; (e) First separated source; (f) Second separated source

In the second experiment, the influence of P and L on the performance is analyzed. The SNR is fixed to be 30 dB, P is set to be 16, 32, 64, 128, 256 and L is set to be 512, 1 024, 2 048, 4 096, 8 192 respectively. Thus (P, L) has 5×5=25 possible groups: from (16, 512), (16, 1 024) till to (256, 8 192). At each (P, L), the presented deconvolution approach is executed 50 times to get its mean performance. The result is shown in Fig.2, from which it can be concluded that P and L have little effect on the algorithm's performance. Of course, the precondition $L>P>>K$ must be satisfied.

In the last experiment, P is fixed to 64 and L is fixed to 1 024, the frequency bin k is changed from 1 to 60 to analyze the algorithm's robustness in different noise environments. The presented deconvolution algorithm is executed 50 times at each k to get its mean performance. The result is shown in Fig.3, which shows that k has little effect on the algorithm's performance when SNR is high, but the performance degrades when SNR falls. The correlation coefficients between the recovered signals and the sources are all larger than those between the mixing signals and the sources when $SNR>10$ dB. It can be concluded that the presented algorithm can improve the signals' correlation when $SNR>10$ dB. Fig.3 also shows that k influences the algorithm's performance periodically. This characteristic comes from the periodicity of $\boldsymbol{Z}(\omega_k, t)$. In this experiment, $P=64$, $L=1024$, and $\omega_k=2\pi k/L$, so $\omega_k P=2\pi k/16$, and the period of $\omega_k P$ is 16.

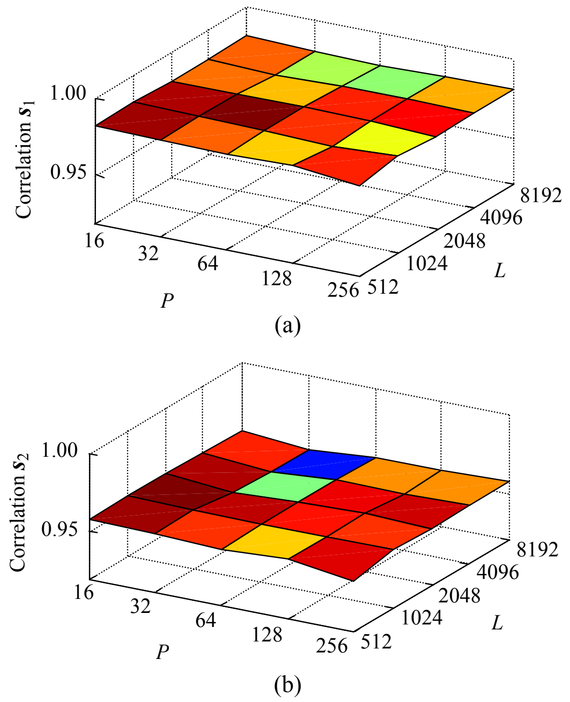

Fig.2 Performance index when changing the sliding window width *P* and the number of points *L*

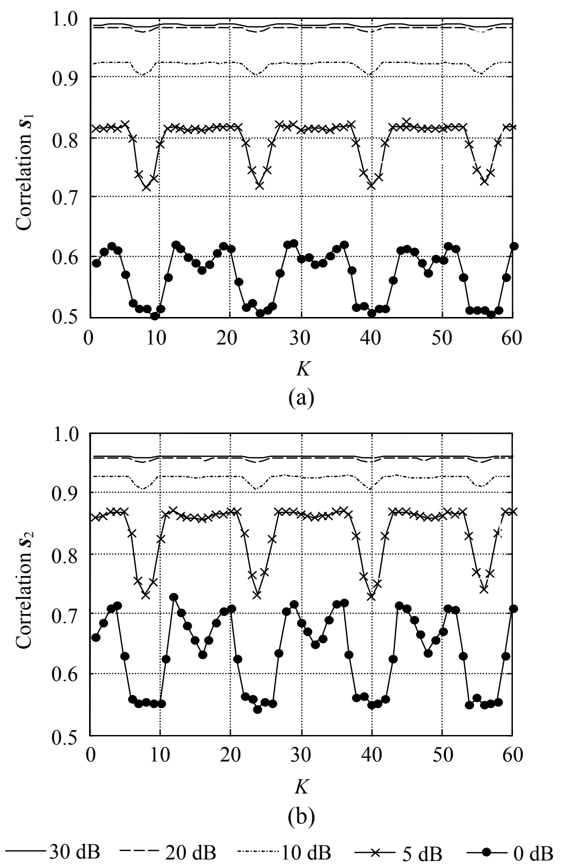

Fig.3 Performance index in different *SNR* when changing the frequency bin *k*

According to Eq.(9), the period of $\boldsymbol{Z}(\omega_k, t)$ should be 16 too. Naturally, the results of blind separation of $\boldsymbol{Z}(\omega_k, t)$ will change periodically along with the change of k.

It is concluded in this section that the parameters L, P and k have limited effect on the presented algorithm's performances when *SNR* is high.

CONCLUSION

By investigating the relationship between time-domain signals and frequency-domain signals, a sliding-Fourier-transform-based blind deconvolution algorithm is presented. It can directly recover the time-domain sources from frequency-domain convolution model after blind separation. It does not suffer from permutation and amplitude indeterminacy problems, which are great troubles for traditional frequency blind deconvolution algorithms. The presented approach greatly reduces the computation complexity and increases the performance robustness. Simulations demonstrated that its performance is stable when changing the algorithm's three parameters: the frequency bin k, the sliding window width P and the number of points L of the Fourier transform.

References

Cardoso, J.F., Souloumiac, A., 1993. Blind beamforming for non-Gaussian signals. *IEE Proc-F*, **140**:362-370.

Depena, A., Serviere, C., Castedo, L., 2003. Inversion of the sliding Fourier transform using only two frequency bins and its application to source separation. *Signal Processing*, **83**(2):453-457. [doi:10.1016/S0165-1684(02)00424-3]

Jacobsen, E., Lyons, R., 2003. The sliding DFT. *IEEE Signal Processing Magazine*, **20**(2):74-80. [doi:10.1109/MSP.2003.1184347]

Kurtia, S., Saruwatari, H., 2000. Evaluation of Blind Separation Method Using Directivity Patter under Reverberant Conditions. ICASSP 2000. Istanbul, Turkey, p.3140-3143.

Mitianoudis, N., Davies, M.E., 2003. Audio source separation convolutive mixtures. *IEEE Trans. on Speech and Audio Processing*, **11**(5):489-497. [doi:10.1109/TSA.2003.815820]

Parra, L., Spence, C., 2000. Convolutive blind source separation of non-stationary sources. *IEEE Trans. on Speech and Audio Processing*, **8**(3):320-327. [doi:10.1109/89.841214]

Peled, R., Braun, S., Zacksenhouse, M., 2005. A blind deconvolution separation of multiple sources, with application to bearing diagnostics. *Mech. Syst. Signal Processing*, **19**(6):1181-1195. [doi:10.1016/j.ymssp.2005.08.019]

Smaragdis, P., 1998. Blind separation of convolved mixtures in the frequency domain. *Neurocomputing*, **22**:21-34. [doi:10.1016/S0925-2312(98)00047-2]

Sun, X., Scott, C.D., 2001. A Natural Gradient Convolutive Blind Source Separation Algorithm for Speech Mixtures. ICA 2001. San Diego, California, USA, p.59-63.

Throckmorton, C.S., Tantum, S.L., Tan, Y.Y., Collins, L.M., 2007. Independent component analysis for UXO detection in highly cluttered environments. *J. Appl. Geophys.*, **61**:304-317. [doi:10.1016/j.jappgeo.2006.05.007]

Unger, H., Zeevi, Y.Y., 2006. Blind separation of spatio-temporal synfire sources and visualization of neural cliques. *Neurocomputing*, **69**:1475-1484. [doi:10.1016/j.neucom.2005.12.024]