Blind Speech Separation in Multiple Environments Using a Frequency Oriented PCA Method for Convolutive Mixtures

Y. Benabderrahmane[1], S. A. Selouani[2], and D. O'Shaughnessy[1]

[1] INRS-EMT, 800 de la Gauchetière O, H5A 1K6, Montréal, QC, Canada,
[2]Université de Moncton, campus de Shippagan E8S 1P6 NB, Canada

Abstract

This paper reports the results of a comparative study on blind speech separation (BSS) of two types of convolutive mixtures. The separation criterion is based on Frequency Oriented Principal Components Analysis (FOPCA). This method is compared to two other well-known methods: the Degenerate Unmixing Evaluation Technique (DUET) and Convolutive Fast Independent Component Analysis (C-FICA). The efficiency of FOPCA is exploited to derive a BSS algorithm for the under-determined case (more speakers than microphones). The FOPCA method is objectively compared in terms of signal-to-interference ratio (SIR) and the Perceptual Evaluation of Speech Quality (PESQ) criteria and subjectively by the Mean Opinion Score (MOS). Usually, the conventional algorithms in the frequency domain are subject to permutation problems. On the other hand, the proposed algorithm has the attractive feature that this inconvenience usually arising does not occur.

Index Terms— Blind source separation (BSS), convolutive mixture, speech signals, second order statistics, Frequency Oriented Principal Component Analysis (FOPCA).

1. Introduction

For several years, the separation of sources has been a particularly active research topic. This interest can be explained by the wide spectrum of possible applications, which includes telecommunications, the biomedical field, the separation of speakers (so-called "cocktail party problem"), detection and separation in communication systems for multiple access, etc. Often, it is preferable to extract all sources from the recorded mixtures or to isolate at least a particular source. Most approaches to BSS assume the sources are statistically independent and thus often seek solutions of separation criteria using higher-order statistical information [1] or using only second-order statistical information in cases where the sources have temporal coherency [2], are non-stationary [3], or eventually are cyclo-stationary. Second-order methods and higher-order ones are based on different assumptions. Higher-order methods assume white sources and do not apply to Gaussian signals but the second-order methods assume that the sources are temporally colored and do not have any Gaussian constraint. Many different approaches have been proposed in recent years [4]. One of these methods is the OPCA (Oriented Principal Components Analysis) algorithm that was previously proposed by Diamantaras and Papadimitriou [5] for separating random signals in the instantaneous case, specifically for four multilevel PAM (Pulse Amplitude Modulation) signals filtered by an ARMA (Auto-Regressive Moving Average) coloring filter. Later, it turned out that the instantaneous model does not suit all situations encountered in practice. Therefore, more realistic modeling is the convolutive mixing. Here we propose a new FOPCA to do BSS on a convolutive mixture of speech signals.

2. Problem statement

With a discrete time index t, a set of P source signals $s(t)=(s_1(t), \ldots, s_P(t))$ is received at an array of N sensors. The received signals are denoted $x(t)=(x_1(t), \ldots, x_N(t))$. In many real-world applications the sources are said to be convolutively (or dynamically) mixed. This model introduces the following relation between the n'th mixed signal and the original source signals. The real convolutive mixing process (including delays) can be assumed as:

$$x_m(t) = \sum_{n=1}^{P}\sum_{k=0}^{K-1} a_{mnk} s_n(t-k) \quad (1)$$

The mixed signal is a linear mixture of filtered versions of each of the source signals, and a_{mnk} represents the corresponding mixing filter coefficients. In practice, these coefficients may also change in time, but for simplicity the mixing model is often assumed stationary. In matrix form, the convolutive model can be written as:

$$x(t) = \sum_{k=0}^{K-1} A_k s(t-k), \quad (2)$$

where A_k is a $P \times N$ matrix that contains the k'th filter coefficients. The convolutive mixing process in eq. (2) can be simplified by transforming the mixtures into the frequency domain. The linear convolution in the time domain can be written in the frequency domain as separate multiplications for each frequency:

$$X(f) = A(f)S(f). \quad (3)$$

At each frequency f, $A(f)$ is a complex $P \times N$ matrix, $X(f)$ is complex $P \times 1$ vector, and similarly $S(f)$ is a complex $N \times 1$ vector. The frequency transformation is typically computed using a discrete Fourier transform (DFT) within a time frame of size T starting at time t :

$$X(f,t) = DFT[x(t), \ldots. x(t+T-1)], \quad (4)$$

and correspondingly for $S(f, t)$. Often a windowed discrete Fourier transform is used:

$$X(f,t) = \sum_{\tau=0}^{T-1} w(\tau)x(t+\tau)e^{-j2\pi f\tau/T}, \quad (5)$$

where the window function $w(t)$ is chosen to minimize band-overlap. Using the fast Fourier transform (FFT), convolutions can be implemented efficiently in the discrete Fourier domain. The objective of BSS is to find an estimate, $\hat{s}(t)$, which is a model of the original source signals $s(t)$. It is often sufficient to estimate separation filters W that remove the cross-talk introduced by the mixing process. The goal is not necessarily to recover identical copies of the original sources; each model source signal can be a filtered version of the original source signals,

$$\hat{S}(f,t) = W(f)A(f)S(f,t). \quad (6)$$

The criterion for separation is satisfied if the recovered signals are permuted, and possibly scaled and filtered, versions of the original signals,

$$W(f)A(f) = P(f)D(f), \quad (7)$$

where P is a permutation matrix and $D(f)$ is a diagonal matrix with scaling filters on its diagonal. If one can identify $A(f)$ exactly and choose $W(f)$ to be its inverse, then $D(f)$ is an identity matrix, and one recovers the sources exactly.

A survey of frequency-domain BSS is provided in [6]. An advantage of blind source separation in the frequency domain is that the separation problem can be decomposed into smaller problems for each frequency bin in addition to the significant gains in

28 – 31 August 2011, Florence, Italy

computational efficiency [7]. The convolutive mixture problem is reduced to "instantaneous" mixtures for each frequency. Another problem that arises in the frequency domain is the permutation and scaling ambiguity. For each frequency, the source signals may be estimated with an arbitrary permutation and scaling,

$$\hat{S}(f,t) = P(f)D(f)S(f,t). \tag{8}$$

3. FOPCA Implementation

OPCA corresponds to the generalized eigenvalue decomposition of two covariance matrices in the same way that PCA corresponds to the eigenvalue decomposition of a single covariance matrix [5]. Oriented PCA (OPCA) describes an extension of PCA involving two signals $u(k)$ and $v(k)$. The aim is to identify the so-called oriented principal directions $e_1,...,e_n$ that maximize the signal-to-signal power ratio $E(e_i^T u)^2/E(e_i^T v)^2$ under the orthogonality constraint: $e_i^T R_u e_j=0$, $i\neq j$. The solution of FOPCA, as shown in Fig. 1, is a Generalized Eigenvalue Decomposition (GED) of the matrix pencil $[R_u,R_v]$. Subsequently, we shall relate the BSS problem with the FOPCA analysis of the observed signal x and almost any filtered version of it. There exist $M > 1$ positive time lags $l_1,....,l_M$ with $l_0=0$ such that $\boldsymbol{R}_s(l_M)$ is diagonal and must be nonzero. If we assume a covariance matrix of $\boldsymbol{s(k)}$: $\boldsymbol{R}_s(0)=\boldsymbol{I}$, the 0-lag covariance matrix of $x(k)$ is:

$$R_x(0) = AR_S(0)A^T = AA^T. \tag{9}$$

Now, consider a scalar, linear filter having $h=[h_0,....,h_M]$ (referred to as pre-filter in Figure 1) operating on $X(f,t)$:

$$Y(f,t) = \sum_{m=0}^{M} H_m X(f,t-l_m). \tag{10}$$

The 0-lag covariance matrix of Y is expressed as:

$$R_Y(0) = E\left\{Y(f,t)Y(f,t)^T\right\} = \sum_{p,q}^{M} H_p H_q R_X(l_p - l_q). \tag{11}$$

From Eq. (1) it follows that:

$$R_X(l_m) = AR_S(l_m)A^T. \tag{12}$$

So,

$$R_Y(0) = ADA^T, \tag{13}$$

$$D = \sum_{p,q=0}^{M} H_p H_q R_S(l_p - l_q). \tag{14}$$

Provided that A is square and invertible we can write:

$$R_Y(0)A^{-T} = AD = AA^T A^{-T} D = R_X(0)A^{-T} D. \tag{15}$$

Eq. (15) expresses a Generalized Eigenvalue Decomposition problem for the matrix pencil $[R_Y(0), R_X(0)]$. This is equivalent to the FOPCA problem for the pair of signals $[Y(f,t), X(f,t)]$. The generalized eigenvalues for this problem are the diagonal elements of D. The columns of the matrix A^{-T} are the generalized eigenvectors. The eigenvectors are unique up to a permutation and scale provided that the eigenvalues are distinct (this is true in general). In this case, for any generalized eigenmatrix W we have $W = A^{-T}P$ with P being a scaled permutation matrix; each row and each column contains exactly one non-zero element. Then the sources can be estimated as:

$$\hat{S}(f,t) = W^T X(f,t), \tag{16}$$

which can be written as:

$$\hat{S}(f,t) = P^T A^{-1} AS(f,t) = P^T S(f,t), \tag{17}$$

where $\hat{S}(f,t)=[\hat{S}_1(f,t), \hat{S}_2(f,t)]^T$ is the estimated source signal vector and $W(f)$ represents an unmixing matrix at frequency bin f. The unmixing matrix $W(f)$ is determined so that $\hat{S}_1(f,t)$ and $\hat{S}_2(f,t)$ become mutually uncorrelated. Then we apply the IFFT of $\hat{S}(f,t)$ for recovering the estimated signals in the time domain.

$$\hat{s}(t) = IFFT\left(\hat{S}(f,t)\right). \tag{18}$$

For the special case $M=2$, the pre-filter mentioned in Figure 1 is expressed as:

$$h = [h_0, h_1, h_2] = [1, \alpha, \beta], \tag{19}$$

where α and β are parameters to be fixed. These parameters are optimized by re-formulating the D matrix of eq. (15) as the following [5]:

$$D = (1+\alpha^2+\beta^2)I + 2\alpha R_S(l_\alpha) + 2\beta R_S(l_\beta) + 2\alpha\beta . R_S(l_\alpha - l_\beta). \tag{20}$$

Note that the optimality criterion of the pre-filter is related to the eigenvalue spread [5]. The maximization criterion used to find α and β is given by:

$$J(\alpha,\beta) = \min_i \left[\min_{j\neq i} \frac{(d_i - d_j)^2}{\max_k d_k^2} \right], \tag{21}$$

where $d_{i,j}$ represent the diagonal elements of D. In our experiments, the pre-filter order of 3 was chosen. The search of the optimal filter is transformed into the search for the filter that spreads the eigenvalues as much as possible [5]. The search is exhaustive and is performed for values of α and b varying within a given interval of h ($\forall\ \alpha,\ \beta\in[h_{min}, h_{max}]$). In the experiments we fixed $h_{min}=-5$, $h_{max}=5$, while the increasing step was 0.2. These values were chosen empirically. The window length was of 512. This length gave the best results when the signals are nearly stationary.

4. Experiments and results

To evaluate our approach in the convolutive case, we compared it with the well-known C-FICA and DUET techniques, and use two types of mixture, the first mixture uses the HRTF (Head Related Transfer Function) filters [8] and the second one represents an anechoic mixture.

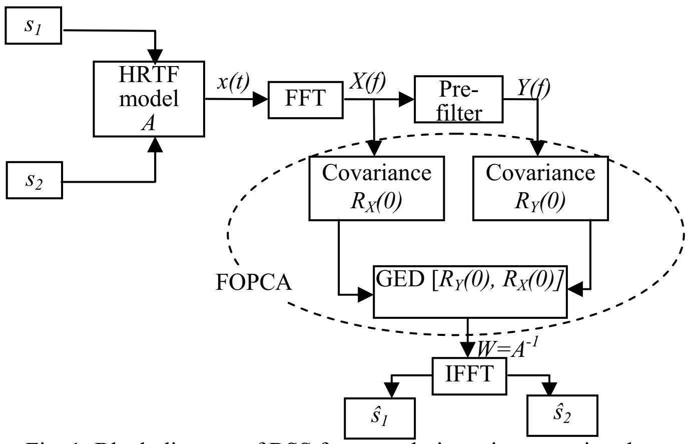

Fig. 1: Block diagram of BSS for convolutive mixtures using the FOPCA method

Through this comparison, we aim to demonstrate the effectiveness of the proposed separation technique based on the FOPCA method.

- **C-FICA** algorithm (Convolutive extension of Fast-ICA) [9] is a time-domain fast fixed-point algorithm that realizes BSS of convolutive mixtures. It is based on a convolutive sphering process (or spatio-temporal sphering) that lets the use of the classical Fast-ICA updates extract iteratively the innovation processes of the sources in a deflation procedure.

- **DUET** (Degenerate Unmixing and Estimation Technique) is a method that applies when sources are W-disjoint orthogonal, i.e., when the time-frequency representations of any two signals in the mixtures are disjoint sets. The method uses an online algorithm to do gradient search for the mixing parameters and simultaneously construct binary time-frequency masks that are used to partition one of the mixtures to recover the original source signals [10].

4-1 Case of two-by-two sources

In the first experiment, the observation data dimension is $n=2$ and speech is from the TIMIT database with typical durations of six seconds. We have also used a set of six other signals, from different speakers (three females and three males) reading sentences lasting approximately 30 seconds. We named them s_1, s_3, s_5 for female speech signals and s_2, s_4, s_6 for male speech signals. The discussion is mainly focused on two signals s_1 and s_2 seen in figure 2 for sake of clarity. The same approach was tried with the other signals s_3, s_4, s_5 and s_6 and the same conclusions were obtained and presented in table 5. These speech signals were collected by Dimitri Nion, & al [11]. The sampling frequency of the signals is 16 kHz. We mixed in convolution two speech signals: $s_1(n)$ and $s_2(n)$, respectively pronounced by a man and a woman. In the experiment for the HRTF model, a dummy head with two microphones (one in each ear) was used instead of the microphone array. This kind of recording was used to investigate how effective the BSS is during a more natural configuration of the sources. We selected impulse responses associated with source positions defined by different angles in relation to the dummy head. For the anechoic model, the mixtures contained the two sources with relative amplitudes (1.1, 0.9) and sample delays (-2, 2). The reverberation time of the environment is not discussed in this work with the same parameters indicated in [10]. In tables 1 and 2 we show results for signals mixed by HRTF convolution with the positions defined by 30 and -80 degrees and in anechoic conditions, and in table 3 the results when the HRTF impulse responses associated with source positions defined by different angles were selected. The angles were chosen randomly.

4-1-1 Evaluation of the Separation Quality

a- To measure speech quality, we use a very reliable method: the PESQ (Perceptual Evaluation of Speech Quality). This method is normalized in ITU-T recommendation P.862 [12] and is generally used to evaluate speech enhancement systems [13]. PESQ returns a score from 0.5 to 4.5.

Table 1. Comparison of PESQ for the C-FICA, DUET and the FOPCA methods

Mixing Model	**HRTF**		**Anechoic**	
PESQ / Methods	Female speech	Male speech	Female speech	Male speech
C-FICA	2.60	2.16	2.74	2.26
DUET	2.98	2.95	3.23	2.69
FOPCA	4.42	4.17	3.25	3.74

b- The received signal quality is typically measured by the SIR (Signal-to-Interference Ratio). A practical way to calculate the SIR ratio is to perform a linear projection or prediction of the sources extracted on the original signals, and differences between the original signals and these projections become the error signals. The signals projected (predicted) and the error signals can then be used for a calculation of the SIR [14]. In our experiments, the SIR measures are implemented within a MATLAB toolbox named *BSS_EVAL,* distributed online under the GNU Public License [15]. The input Signal-to-Interference-Ratio (SIR_{in}) before separation was -4.66 dB for the male speech and -1.75 dB for the female speech. In Table 2, SIR_{out} represents the output Signal-to-Interference-Ratio after separation. As shown in Table 2, the ratio SIR_{out} of the FOPCA method is larger than the SIR_{out} of the C-FICA and the DUET approaches. Taking as an example the HRTF mixing model, the improvement in the SIR_{out} ratio for C-FICA was 14.36 dB and 9.91 dB respectively for male and female speech. It was 14.60 dB and 11.75 dB for DUET and by 45.01 dB and 31.34 dB against the FOPCA approach. We also note the efficiency of the FOPCA method when the HRTF impulse responses associated with source positions defined by different angles were selected. The angles were chosen randomly as shown in Table 3. Note that for angles that are very close on the same side of the ear, such as the angle of -180 degrees with 10 degrees or -10 degrees with the angle -60 degrees, the separation of the two signals does not occur. The very good results were found for speech signals with long duration, collected by Dimitri Nion & al [11] as we can see in Table 5. In this case, the gain reached a value of 50 dB between SIR_{out} and SIR_{in}, which indicates very promising results. This approach is effective, as can be seen in the time domain, where we note that the original signals and estimated signals by FOPCA are very close (Fig. 2-(a) and (c)). A common subjective benchmark for quantifying the performance of the speech is the Mean Opinion Score (MOS) because voice quality in general is subjective to listeners. In our experiments, MOS tests were conducted with a group of ten listeners. The results for the three methods are summarised in table 4, and we can note that the MOS of FOPCA is the best one.

Table 2: Comparison of SIR (in dB) for the C-FICA, DUET and FOPCA methods

Mixing Models	HRTF SIR_{out}		Anechoic SIR_{out}	
Methods	Male speech	Female speech	Male speech	Female speech
C-FICA	9.07	8.16	10.80	7.10
DUET	9.94	10	10.67	3.10
FOPCA	40.35	29.59	24.23	33.28

Table 3: Comparison of SIR for the C-FICA, DUET and FOPCA methods with different angles for HRTF impulse responses

Angles (degrees)	**FOPCA**	**C-FICA**	**DUET**
10 -10	29.60 40.35	9.07 8.16	-1.63 2.07
20 -120	29.60 40.35	9.07 8.16	-1.63 2.07
10 -180	-16.17 3.04	-255 -257	-7.53 2.52
-20 50	29.60 40.40	9.07 8.16	-1.63 2.07
-30 80	29.60 40.40	9.07 8.16	-1.63 2.07
-10 60	29.60 40.40	9.07 8.16	-1.63 2.07
20 -50	29.60 40.35	9.07 8.16	-1.63 2.07
-10 -60	-21.00 3.03	-255 -257	-2.78 -2.50

Table 4. Comparison of MOS for the C-FICA, DUET and FOPCA methods.

Mixing	HRTF	Anechoic
Methods	MOS	
C-FICA	2.62	2.66
DUET	2.31	3.82
FOPCA	4.36	4.36

Table 5: Results of SIR_{out} and PESQ for speech signals collected by Dimitri Nion & al: FOPCA method with HRTF mixing model

Speech signals	s_1 and s_2	s_3 and s_4	s_5 and s_6	s_2 and s_3
SIR (dB)	48.70 48.20	40.82 41.16	32.54 31.56	38.88 39.87
PESQ	4.48 4.28	4.07 4.19	3.75dB 3.75dB	4.05 4.11

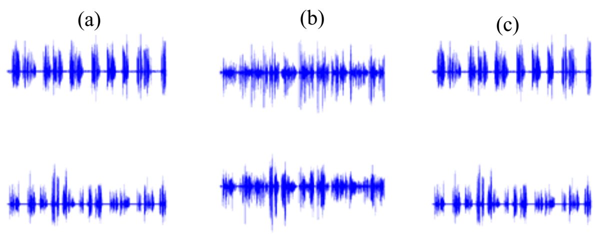

Figure 2: (a) Original speech signals: male s_2 and female s_1 (b) Speech signals mixed by convolution using the HRTF model, (c) Estimated signals by the FOPCA approach: male s_2 and female s_1.

4-2 Under-determined case

In this experiment we consider the under-determined case and we illustrate the performance of the FOPCA algorithm in comparison with the DUET and C-FICA methods. The speech sources were convolved with impulse responses associated with source positions in relation to the dummy head: HRTF. In this case we have three speakers and two microphones. We cannot have more than two microphones because the model of the dummy head has only two ears.

Table 6: Comparison of SIR for the C-FICA, DUET and FOPCA methods with different angles for HRTF impulse responses: under-determined case

	SIR_{out} (dB)		
Angles (degree)	FOPCA	C-FICA	DUET
30	**29**	2.32	-0.38
-80	2,54	14.75	-0.92
10	-2,6	-2.04	-0.85
20	**29**	2.32	-0.38
-50	2,54	14.75	-0.92
40	-2,6	-2.04	-0.85
10	-0.28	5.36	-1.57
-50	**39,26**	0.34	-2.50
-10	-23,14	-9.35	0.83
80	-3	1.20	-0.07
70	3	-6.81	-0.30
-30	**19,6**	10.06	-0.10
60	-2.68	1.20	-0.07
90	3	-6.81	-0.30
-70	**19,5**	10.06	-0.10

As shown in tables 6 and 7, the ratio SIR_{out} of the FOPCA method is larger than the SIR_{out} of the C-FICA and the DUET algorithms. This is true for the two types of mixing: the HRTF and anechoic models. Table 8 gives the PESQ comparison for the three methods. We note that in the case of the FOPCA approach, the results are higher compared to the other two methods, particularly in the case of the mixture with the HRTF model. For the anechoic mixture, the result is better compared to the C-FICA algorithm, but comparable with respect to the DUET method. For this last one we have good separation only for one signal, but for FOPCA we are able to separate the two signals.

Table 7: Comparison of SIR for the C-FICA, DUET and FOPCA methods using an anechoic model: under-determined case

SIR_{out} (dB)		
FOPCA	**CFICA**	**DUET**
-3.4	3.24	13
14	3.89	-0.73
4.27	2.64	1.39

Table 8: Comparison of PESQ for the C-FICA, DUET and FOPCA methods using anechoic model: under-determined case

PESQ		
FOPCA	**CFICA**	**DUET**
1.95	1.87	2.76
1.84	1.89	1.54
2.36	2.03	1.54

5. Conclusion

We have presented a blind speech separation technique of convolutive mixtures using a frequency oriented principal component analysis method. The proposed separation technique of mixed observations into source estimates is effective, as shown in the time domain. Subjective evaluation is performed through listening to the estimated signals before and after mixing and after separation was used. The results are very satisfactory; we obtained a very good separation. We tested the method with other speech signals from the TIMIT database and the results were similar. Tests were also conducted with speech signals with duration of 30s, and the results were very satisfactory not only for the PESQ but also for the SIR as we can see in Table 5. However, to the contrary of the well-known algorithms in the frequency domain, the proposed algorithm has the advantage that the inconvenient permutation problem usually arising in frequency domain methods does not occur [16]. These results confirm the efficiency of the FOPCA method in convolutive mixtures [17] [18]. This approach was previously used for the first time, in the separation of speech signals in an instantaneous mixing case in the time domain [18]. We are continuing our research efforts by implementing FOPCA, applying it in a mobile communication framework. For a 3-source scenario, the SIR_{out} for all three sources is listed in Table 6. One SIR_{out} is extremely high, but the other two SIRs are fairly low. This indicates that only the source with high SIR_{out} is well separated; the other two signals are still mixed. We know that there are some applications where the extraction of a single source is desired and the purpose of a blind source separation is to extract at least one signal source

6. References

[1] J.-F. Cardoso, "Source separation using higher order moments," in Proceedings IEEE ICASSP, Glasgow, U.K., 1989, vol. 4, pp. 2109–2112.

[2] J. Basak and S. Amari, "Blind separation of uniformly distributed signals: A general approach," IEEE Trans. Neural Networks, vol. 10, pp. 1173–1185, Sept. 1999.

[3] D.T. Pham, "Blind separation of instantaneous mixture of sources via an independent component analysis," IEEE Trans. Signal Processing, vol. 44, pp. 2768–2779, 1996.

[4] A. Hyvärinen, J. Karhunen, and E. Oja, "Independent Component Analysis", John Wiley, NY, 2001.

[5] K. I. Diamantaras, Th. Papadimitriou, "Oriented PCA and Blind Signal Separation", 4th Int. Symp. On Independent Component Analysis & BSS, 609-613, Nara, April 2003.

[6] M. S. Pedersen, J. Larsen, U. Kjems, and L.C. Parra, "A Survey of Convolutive Blind Source Separation Methods", Springer Handbook on Speech Processing and Speech Communication, 2007

[7] Makino, S., Sawada, H., Mukai, R., and Araki, S. ,"BSS of convolutive mixtures of speech in frequency domain," IEICE Trans. Fund., vol. E88-A, no. 7, 1640–1655, 2005.

[8] B. Gardner and K. Martin,"HRTF Measurements of KEMAR Dummy-Head Microphone", J. Acoust. Soc. Amer., Vol. 97, 3907-3908, 1995. See http://sound.media.mit.edu/ica-bench/.

[9] J. Thomas, Y. Deville and S. Hosseini, "Time-Domain Fast Fixed-Point Algorithms for Convolutive ICA", IEEE Sig. Proc. Letters, 13, No. 4, 228-231, April 2006.

[10] Makino, S., Lee, T.W., Sawada, H., "Blind Speech Separation", Signals & Comm. Techn., Springer, 2007.

[11] D. Nion, K.N. Mokios, N.D. Sidiropoulos, and A. Potamianos, "Batch and Adaptive PARAFAC-Based Blind Separation of Convolutive Speech Mixtures", IEEE Trans. Audio, Speech and Language Processing, Vol. 18, No. 6, 1193-1207, 2010.

[12] ITU, "Perceptual evaluation of speech quality (PESQ)", ITU-T Recommendation 862, 2000.

[13] P Loizou, 2007. "Speech Enhancement: Theory and Practice". CRC Press LLC, Boca Raton, FL, 2007.

[14] E. Vincent, R. Gribonval, and C. Févotte, "Performance Measurement in Blind Audio Source Separation" IEEE Trans. Audio, Speech, Lang. Proc., Vol. 14, No. 4, 2006.

[15] C. Févotte, R. Gribonval, and E. Vincent, "BSS_EVAL toolbox user guide," IRISA, Rennes, France, Tech. Rep. 1706, 2005. (On-line: http://www.irisa.fr/metiss/bss_eval).

[16] M. Kawamoto and Y. Inouye, "Blind separation of multiple convolved colored signals using second-order statistics" In *Proc. of ICA'03*, Nara, Japan, April 2003.

[17] Y. Benabderrahmane, S.A. Selouani and D. O'Shaughnessy, "Oriented PCA method for Blind Speech Separation of Convolutive Mixtures", Interspeech 2010, September 26-30, 2010, Makuhari, Japan.

[18] Y. Benabderrahmane, S.A. Selouani, D. O'Shaughnessy, and H. Hamam, "A Comparative Study of Blind Speech Separation using Subspace Methods and Higher Order Statistics", Lecture Notes in Computer Science, Springer eds., pp. 117-124, 2009.