A NEW FAST–CONVERGING METHOD FOR BLIND SOURCE SEPARATION OF SPEECH SIGNALS IN ACOUSTIC ENVIRONMENTS

Kamran Rahbar and James P. Reilly

Electrical & Computer Eng.
McMaster University, 1280 Main St. w., Hamilton, Ontario, Canada L8S 4K1
email: kamran.rahbar@zarlink.com, reillyj@mcmaster.ca

ABSTRACT

In this paper we propose a new frequency domain approach to blind source separation (BSS) of audio signals mixed in a reverberant environment. It is first shown that joint diagonalization of the cross power spectral density matrices of the signals at the output of the mixing system is sufficient to identify the mixing system at each frequency bin up to a scale and permutation ambiguity. The frequency domain joint diagonalization is performed using a new and quickly converging algorithm which uses an alternating least-squares (ALS) optimization method. An efficient diadic algorithm to resolve the frequency dependent permutation ambiguities is presented. The effect of the unknown scaling ambiguities is partially resolved using a novel initialization procedure for the ALS algorithm. The performance of the proposed algorithm is demonstrated by experiments conducted in real reverberant rooms. Audio results are available at
"www.ece.mcmaster.ca/~reilly/kamran/index.htm".

1. INTRODUCTION

In the blind source separation (BSS) problem the objective is to separate multiple sources, mixed through an unknown mixing system(channel), using only the system output data (observed signals) and in particular without using any (or least amount of) information about the sources or the system. The blind source separation problem arises in many fields of studies, including speech processing, data communication, biomedical signal processing, etc. In the audio context where the objective is to separate signals mixed in reverberant rooms, the BSS problem is very difficult, since the mixing system is convolutive with transfer functions of very large order between the sources and the sensors.

In recent years a few blind source separation methods have been proposed that exploit the non-stationarity of the source signals [1–4]. The advantage of using frequency-domain methods is that a time–domain estimation problem with a large number of parameters is decomposed into multiple, independent estimation problems, each with fewer parameters to be estimated. As a result, in general the frequency-domain estimation algorithms have a simpler implementation and better convergence properties over their time–domain counterparts. The main difficulties with frequency–domain blind source separation of convolutive mixtures however is the arbitrary permutation and scaling ambiguities of the estimated frequency response of the un-mixing system at each frequency bin.

Support of the Natural Sciences and Engineering Council of Canada (NSERC) is gratefully acknowledged.
Kamran Rahbar is now with Zarlink Semiconductor, Kanata, ON, Canada.

The method proposed in this paper exploits second-order non-stationarity of the input signals to separate convolved audio sources using frequency domain techniques, using a joint diagonalization procedure. Novel approaches are proposed to deal with the arbitrary permutation and scaling ambiguity problems.

2. PROBLEM STATEMENT

2.1. The Convolutive Mixing Problem

We consider the following N-source J-sensor multi–input multi–output (MIMO) linear model for the received signal for the convolutive mixing problem[1]:

$$\mathbf{x}(t) = [\mathbf{H}(z)]\mathbf{s}(t) + \mathbf{n}(t), \qquad t \in \mathbb{Z} \tag{1}$$

where $\mathbf{x}(t) = (x_1(t), \cdots, x_J(t))^T$ is the vector of observed signals, $\mathbf{s}(t) = (s_1(t), \cdots, s_N(t))^T$ is the vector of sources, $\mathbf{H}(z)$ is the $J \times N$ transfer function of the mixing system and $\mathbf{n}(t) = (n_1(t), \cdots, n_J(t))^T$ is the additive noise vector. We assume $h_{ij}(z)$, the ij_{th} element of $\mathbf{H}(z)$, to be a rational function of z. The objective of the blind source separation algorithm is to estimate the un-mixing filters $\mathbf{W}(z)$ from the observed signals $\mathbf{x}(t)$ such that

$$\mathbf{W}(z)\mathbf{H}(z) = \mathbf{\Pi}\mathbf{D}(z) \tag{2}$$

where $\mathbf{\Pi} \in \mathbb{R}^{N \times N}$ is a permutation matrix and $\mathbf{D}(z)$ is a diagonal matrix with diagonal elements which are rational functions of z. Often with frequency-domain methods, the matrix $\mathbf{\Pi}$ is dependent on ω. However, to ensure adequate separation performance, $\mathbf{\Pi}$ must be made independent of frequency. To ensure satisfactory audio output quality, $\boldsymbol{D}(z)|_{z=e^{j\omega}}$ should also at least be approximately constant with frequency.

2.2. Main Assumptions

- **A0**: $J \geq N \geq 2$; i.e, we have at least as many sensors as sources and number of the sources are at least two.
- **A1**: The sources $\mathbf{s}(t)$ are zero mean, second-order quasi-stationary signals and the cross–spectral density matrices of the sources $\mathbf{P}_s(\omega, m)$ are diagonal for all ω and m where ω denotes frequency and m is the epoch index.
- **A2**: The mixing system is modelled by a causal system of the form $\mathbf{H}(z) = [\mathbf{h}_1(z), ..., \mathbf{h}_N(z)]$ and does not change over the entire observation interval.

[1] Here we use the notation $[\mathbf{H}(z)]\mathbf{s}(t)$ to denote the convolution between a system with z-transform $\mathbf{H}(z)$ and source vectors $\mathbf{s}(t)$.

A3: $\mathbf{H}(\omega_k)$, the DFT of $\mathbf{H}(z)$, has full column rank for all ω_k, $k = 0, \ldots, K-1$, $\omega_k = \frac{2\pi k}{K}$. Also, for convenience, we assume $||\mathbf{h}_i(\omega_k)||_2^2 = 1$ where $\mathbf{h}_i(\omega_k)$ is the i_{th} column of $\mathbf{H}(\omega_k)$.

A4: The noise $\mathbf{n}(t)$ is zero mean, *iid* across sensors, with power σ^2. The noise is assumed independent of the sources.

Let $\mathbf{P}_x(\omega, m)$ represent the cross-spectral density matrix of the observed signal at frequency ω and time epoch m. Based on the above assumptions we can write:

$$\mathbf{P}_x(\omega, m) = \mathbf{H}(\omega)\mathbf{P}_s(\omega, m)\mathbf{H}^\dagger(\omega) + \sigma^2\mathbf{I}, \tag{3}$$

where $\mathbf{P}_s(\omega, m)$ is a diagonal matrix which represents the cross-spectral density matrices of the sources at epoch m. In practice we use a discretized version of ω given as $\omega_k = (2\pi k/K)$ where K is the total number of frequency samples. For $J > N$, σ^2 can be estimated from the smallest eigenvalue of the matrix $\mathbf{P}_x(\omega, m)$; so for now we consider the following noise free case:

$$\mathbf{P}_x(\omega_k, m) = \mathbf{H}(\omega_k)\mathbf{P}_s(\omega_k, m)\mathbf{H}^\dagger(\omega_k). \tag{4}$$

3. THE JOINT DIAGONALIZATION PROBLEM

The joint diagonalization problem first introduced by Flurry [5]. The problem is expressed as finding a single matrix that jointly (approximately) diagonalizes the set of matrices $\mathbf{R}_1, \ldots, \mathbf{R}_M$. The motivation behind using joint diagonalization as part of the proposed algorithm can be explained through the following theorem:

Theorem 1 *Consider the set of matrices*

$$\mathcal{R} = \{\mathbf{R}_m \in \mathbb{C}^{J\times J} | \mathbf{R}_m = \mathbf{A}\mathbf{\Lambda}_m\mathbf{A}^\dagger \ \ m = 0, \ldots, M-1\} \tag{5}$$

where $\mathbf{A} \in \mathbb{C}^{J\times N}$ is some full column rank matrix and $\mathbf{\Lambda}_m \in \mathbb{R}^{N\times N}$ are diagonal matrices such that the set of vectors $\boldsymbol{\lambda}_m = \mathrm{diag}\{\mathbf{\Lambda}_m\}$ span $\mathbb{R}^N$. We also assume that $\mathbf{a}_i$, the i_{th} column of $\mathbf{A}$ has unit norm, i.e.; $||\mathbf{a}_i||_2 = 1$. Now if there is a matrix $\mathbf{B} \in \mathbb{C}^{J\times N}$ and diagonal matrices $\tilde{\mathbf{\Lambda}}_m$ such that

$$\mathbf{R}_m = \mathbf{B}\tilde{\mathbf{\Lambda}}_m\mathbf{B}^\dagger \tag{6}$$

and assuming that $\mathbf{B}$ has unit norm columns then $\mathbf{B}$ must be related to $\mathbf{A}$ as

$$\mathbf{B} = \mathbf{A}\mathbf{\Pi}e^{j\mathbf{D}} \tag{7}$$

where $\mathbf{\Pi}$ is a permutation matrix and $\mathbf{D} \in \Re^{N\times N}$ is a diagonal matrix.

Proof: [4] [6].

Based on the above Theorem we can easily see that for the set of matrices $\mathbf{P}_x(\omega_k, m), m = 0, \ldots, M-1$, given by (4), if we find a matrix $\mathbf{B}(\omega_k)$ and diagonal matrices $\mathbf{\Lambda}(\omega_k, m)$ such that $\mathbf{P}_x(\omega_k, m) = \mathbf{B}(\omega_k)\mathbf{\Lambda}(\omega_k, m)\mathbf{B}^\dagger(\omega_k)$ and with the constraint that $||\mathbf{b}_i(\omega_k)||_2^2 = 1$, where $\mathbf{b}_i(\omega_k)$ is the i_{th} column of $\mathbf{B}(\omega_k)$, then we should have

$$\mathbf{B}(\omega_k) = \mathbf{H}(\omega_k)\mathbf{\Pi}(\omega_k)e^{j\mathbf{D}_k} \tag{8}$$

where $\mathbf{D}_k$ are diagonal matrices. In other words, using a joint diagonalization procedure, we can estimate the mixing system $\mathbf{H}(\omega_k)$ up to a frequency dependent permutation $\mathbf{\Pi}(\omega_k)$ and frequency dependent phase ambiguity $e^{j\mathbf{D}_k}$.

In order to attain adequate separation performance, the permutation ambiguity must be resolved. In order to achieve good audio quality at the output, the phase ambiguity problem must be addressed. The permutation issue is addressed in Section 5. A procedure to partially address the phase problem is discussed in Section 4.1.

Having $\mathbf{B}(\omega_k)$ at each frequency bin we can calculate the separating matrix $\mathbf{W}(\omega_k)$ from

$$\mathbf{W}(\omega_k) = \mathbf{B}^+(\omega_k) \tag{9}$$

where $\mathbf{B}^+(\omega_k)$ is the pseudo inverse of matrix $\mathbf{B}(\omega_k)$.

4. THE ALGORITHM

Based on Theorem 1 we can propose the following least-squares based joint diagonalization criterion for the case when a sample estimate of each $\mathbf{P}_x(\omega_k, m)$ is available:

$$\min_{\mathbf{B}(\omega_k),\mathbf{\Lambda}(m)} \sum_{k=0}^{K-1}\sum_{m=0}^{M-1} ||\hat{\mathbf{P}}_x(\omega_k, m) - \mathbf{B}(\omega_k)\mathbf{\Lambda}(\omega_k, m)\mathbf{B}^\dagger(\omega_k)||_F^2, \tag{10}$$

where $\mathbf{B}(\omega)$, which has unit-norm columns, is an estimate of the mixing system $\mathbf{H}(\omega_k)$, $\hat{\mathbf{P}}_x(\omega_k, m)$ is a sample estimate of the observed signal cross spectral density matrix at frequency bin ω_k and time epoch m, and $\mathbf{\Lambda}(\omega_k, m)$ is a diagonal matrix, representing the unknown cross-spectral density matrix of the sources at epoch m.

To estimate the parameters $\mathbf{B}(\omega_k)$ and $\mathbf{\Lambda}(\omega_k, m)$ we optimize the criterion given by (10) using an alternating least-squares (ALS) approach. The basic idea behind the ALS algorithm is that in the optimization process we divide the parameter space into multiple sets. At each iteration of the algorithm we minimize the criterion with respect to one set conditioned on the previously estimated sets of the parameters. The newly estimated set is then used to update the remaining sets. This process continues until convergence is achieved.

In the interests of brevity, the description of the proposed joint diagonalization algorithm is necessarily brief. Full details are available in [6]. We set $\mathbf{G}(\omega_k) = \mathbf{B}(\omega_k) \odot \mathbf{B}(\omega_k)$, $\mathbf{d}(\omega_k, m) = \mathrm{diag}\{\mathbf{\Lambda}(\omega_k, m))\}$ and $\hat{\mathbf{p}}_x(\omega_k, m) = \mathrm{vec}\{\hat{\mathbf{P}}_x(\omega_k, m)\}$, where $\odot$ represents the Khatri-Rao product. Given the $\mathbf{d}(\omega_k, m)$, the ALS estimate of $\mathbf{b}_i(\omega_k)$ is then obtained as the "maximum" eigenvector of a matrix $\mathbf{Y}_i(\omega_k)$, which is derived from the ith column of $\mathbf{G}(\omega_k)$. Then, given the $\mathbf{b}_i(\omega_k)$, the $\mathbf{d}(\omega_k, m)$ are estimated according to

$$\begin{aligned}\hat{\mathbf{d}}(\omega_k, m) =& \hat{\mathbf{G}}^+(\omega_k)\hat{\mathbf{p}}(\omega_k, m),\\ & m = 0, ..., M-1, \ \ k = 0, \ldots, K-1,\end{aligned} \tag{11}$$

where $\hat{\mathbf{G}}^+(\omega_k)$ is also a matrix derived from $\mathbf{G}(\omega_k)$. This procedure is iterated until convergence is achieved.

4.1. Initialization

A novel, *ad hoc* initialization method which partially alleviates the unknown phase ambiguity problem, and is crucial in obtaining adequate quality of the separated audio signals, is described as follows. The main idea of this initialization procedure, is that first we choose the initial value of $\mathbf{B}(\omega_0)$, the first frequency bin, using an exact closed form joint diagonalization method. We then apply the ALS algorithm to find the final estimate of $\mathbf{B}(\omega_0)$. This final estimate is then used as an initial value for the next adjacent frequency bin, which is $\mathbf{B}(\omega_1)$. The outcome of this frequency bin is also used as an initial value for the next frequency bin and this procedure continues until all the frequency bins have been covered. Note that in this way we need to apply the exact closed form joint diagonalization algorithm (which is more expensive computationally) only for one frequency bin.

As has been demonstrated in our simulation results, this initialization procedure significantly improves the quality of the separated audio signals. An intuitive explanation for this is as follows. We realize that the estimate of $\boldsymbol{B}$ is not unique, because each column $\boldsymbol{b}_i$ is subject to a multiplicative phase ambiguity of the form $e^{j\mathbf{D}_k}$, even though the condition $\|\ \boldsymbol{b}_i\ \|_2 = 1$ in the solution of (10) has been enforced. Fast variation of this phase ambiguity in frequency can cause the resulting time–domain estimate $\hat{\boldsymbol{H}}(t)$ of the channel to be excessively long. By initializing the algorithm in the manner proposed, this phase ambiguity varies smoothly with frequency, therefore creating an $\hat{\boldsymbol{H}}(t)$ which can be of moderate length. As is shown in our simulations, the resulting overall system (channel + inverse) is then more localized in time, resulting in better perceptual performance.

5. RESOLVING PERMUTATIONS

A problem with the cost function in (10) is that it is insensitive to permutations of the columns of $\mathbf{B}(\omega_k)$. More specifically if $\mathbf{B}_{opt}(\omega_k)$ is an optimum solution to (10) then $\mathbf{B}_{opt}(\omega_k)\mathbf{\Pi}_k$, where $\mathbf{\Pi}_k$ is an arbitrary permutation matrix for each ω_k, will also be a optimum solution. Since in general $\mathbf{\Pi}_k$ can vary with frequency, poor overall separation performance will result.

In this section we suggest a novel solution for solving the permutation problem which exploits the cross-frequency correlation between diagonal values of $\mathbf{\Lambda}(\omega_k, m)$ and $\mathbf{\Lambda}(\omega_{k+1}, m)$ given in (10). Notice that $\mathbf{\Lambda}(\omega_k, m)$ can be considered an estimate of the sources' cross-power spectral density at epoch m. When the sources are speech signals, the temporal trajectories of the power spectral density of speech, known as spectrum modulation of speech, are correlated across the frequency spectrum. Using this correlation we can correct wrong permutations, as discussed for the two source case. This same principle was also used in, e.g., [7]. The proposed method for resolving permutations can be extended to an arbitrary number of sources.

Assume that $\mathbf{\Lambda}(\omega_k, m), m = 0, \ldots, M-1$, represents the estimated cross-spectral density of the two sources at frequency bin ω_k. We want to adjust the permutation at frequency ω_j such that it has the same permutation as frequency bin ω_k. To do so we first calculate the cross frequency correlation between the diagonal elements of $\mathbf{\Lambda}(\omega_j, m)$ and $\mathbf{\Lambda}(\omega_k, m)$ using following measure:

$$\rho_{qp}(\omega_k, \omega_j) = \frac{\sum_{m=0}^{M-1} \lambda_q(\omega_k, m)\lambda_p(\omega_j, m)}{\sqrt{\sum_{m=0}^{M-1} \lambda_q^2(\omega_k, m)}\sqrt{\sum_{m=0}^{M-1} \lambda_p^2(\omega_j, m)}} \quad (12)$$

where $\rho_{qp}(\omega_k, \omega_j)$ represents the cross frequency correlation between $\lambda_q(\omega_k, m)$, the q_{th} diagonal element of $\mathbf{\Lambda}(\omega_k, m)$, and $\lambda_p(\omega_j, m)$, the p_{th} diagonal element of $\mathbf{\Lambda}(\omega_j, m)$. To determine whether frequency bins ω_k and ω_j have the same permutation, we consider the quantity

$$r = \frac{\rho_{11}(\omega_k, \omega_j) + \rho_{22}(\omega_k, \omega_j)}{\rho_{12}(\omega_k, \omega_j) + \rho_{21}(\omega_k, \omega_j)}. \quad (13)$$

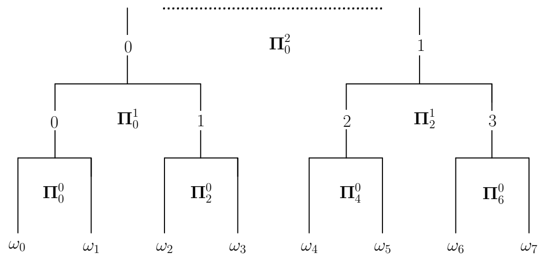

Figure 1: An example of the diadic permutation sorting algorithm for the case when the total number of frequency bins is eight.

If $r < 1$, the permutation at one of the frequency bins ω_k or ω_j is changed. The extension to the general case where number of sources can be greater than two is discussed in [6].

Note that the above algorithm calculates the permutation matrix $\mathbf{\Pi}_k$ between two frequency bins ω_k and ω_j. To obtain a uniform permutation across the whole frequency spectrum, we need to apply the above algorithm repeatedly to all pairs of frequency bins. One way of doing this is to adjust the permutation between adjacent frequency bins in a sequential order where, for example, starting from frequency bin ω_0 we adjust the permutation of each bin relative to it's previous bin. This approach, although simple, has a major drawback as explained as follows. Consider the situation where an error is made in estimating the correct permutation matrix for frequency bin ω_k. In this case, all frequency bins placed after ω_k will receive a different permutation than the ones placed before ω_k. In the worst case scenario we will have half of the frequency bins with one permutation and the other half with a different permutation, which will result in very poor or no separation. To prevent such a catastrophic situation we propose following hierarchical sorting scheme to sort the permutations across all frequency bins. The method is described with the aid of Figure 1, which shows the case for $K = 8$ frequency bins.

This scheme first applies the above test to adjacent pairs of frequency bins; i.e, $[0, 1], [2, 3], \ldots, [K-2, K-1]$. The permutations between the pairs are adjusted as necessary. The data from each pair is then combined, and the process moves up one level in the hierarchy as shown in Figure 1. Permutations between adjacent new pairs are again adjusted, and so on. This process iterates until the last (root) pair is complete[2]. The process is described in detail in [6].

The success of the proposed diadic sorting algorithm can be explained as follows. As the process moves up the hierarchy, the statistics comprising the numerator and denominator of (13) accumulate more data samples. Thus, the probability distribution governing these statistics decreases in variance, with the result that the probability of error in deciding the correct permutation diminishes as the hierarchy is ascended. Thus, we can say that the dominant error mechanism is at the lower levels of the hierarchy, where the impact of a permutation error is not catastrophic. Simulation results verifying the effectiveness of this procedure are given in [6].

[2]It is assumed the number of frequency bins is a power of 2.

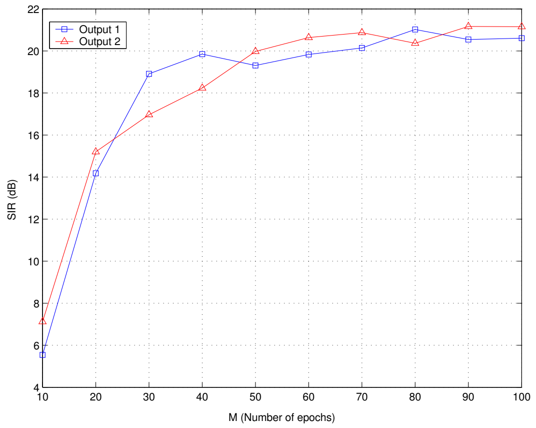

Figure 2: Results of separation performance for recordings in an office room: $\text{SIR}_y i$ versus M, the number of epochs, for $K = 4096$, $i = 1, 2$.

6. REAL ROOM EXPERIMENTS

The experiments were conducted in a typical office room of dimension 3.3m by 4.75m, using $N = 2$ loudspeakers for the sources and $J = 4$ omnidirectional microphones for the sensors. The two sources were created by catenating independent multiple speech segments from the TIMIT speech database. The speech signals then were played simultaneously through the two speakers with approximately the same sound volume.

To quantify the separation performance, we measure the signal–to–interference ratio at each microphone (i.e., the input to the algorithm) and at each output of the algorithm. To achieve this, using the same setup, white noise signals were played through each speaker one at a time (i.e., only one source was active at each time). Let $\hat{\sigma}_x^2(x_i, s_j) = \sum_{t=0}^{T-1} x_i^2(t)$ represent the power of the recorded signal at the i_{th} microphone when only speaker j is active and all other speakers (sources) are inactive. The signal–to–interference ratio of the recorded signal at the i_{th} microphone can be estimated using

$$SIR_x(i) = \frac{\max_j \hat{\sigma}_x^2(x_i, s_j)}{\sum_{j=1}^{N} \hat{\sigma}_x^2(x_i, s_j) - \max_j \hat{\sigma}_x^2(x_i, s_j)}. \quad (14)$$

Using above formula, we can also measure $\text{SIR}_y(i)$, the SIR of the ith output of the separating network, by substituting $\hat{\sigma}_x^2(x_i, s_j)$ with $\hat{\sigma}_y^2(y_i, s_j)$, the power of the signal at the i_{th} output when only source j is active.

To perform the separation, we first divided the recorded signal into multiple time segments (epochs), where we chose the size of each epoch to be 10000 data samples long. For each epoch we calculated the cross spectral density matrices according to the method described in [6], with the number of FFT-points K equal to 4096, and with the overlap-percentage equal to 80%. Figure 2 shows the output $\text{SIR}_y(i)$ versus M, the number of epochs, for $i = 1, 2$. As can be seen for $M \geq 60$, an average SIR of more than 20dB is reached for each output. By comparing the SIRs before and after applying the separating algorithm, it can be seen that the output SIRs have been improved by roughly 19 to 20dB. In this experiment, on average, at each frequency bin the ALS algorithm converged within 30 to 40 iterations.

To demonstrate the effect of the proposed initialization procedure on the impulse response of the overall network, results in [6] show the impulse response of the separating matrix with and without using the proposed initialization method. It is shown that the proposed initialization procedure significantly reduces the length of the overall response (channel + inverse). This is known to result in improved perceptual performance of the processed speech.

The performance of the proposed method has been compared to that of previous convolutive BSS approaches from the literature in [6]. It is demonstrated that, for the same experiment, the proposed method significantly outperforms previous methods, both in perceptual quality and separation performance.

Listening results are available at
"www.ece.mcmaster.ca/~reilly/kamram/index.htm".

7. CONCLUSIONS

In this paper we discussed a new algorithm for blind source separation of convolved non-stationary sources. It was proved that the set of the observed signals' cross spectral density matrices evaluated over different time segments is sufficient to recover the sources up to a frequency dependent scaling and permutation ambiguity, using a joint diagonalization procedure. A two stage frequency-domain algorithm to estimate the un-mixing filters was proposed. A novel procedure for resolving the frequency-dependent permutation ambiguities, and a novel initialization method which significantly improves the perceptual quality of the separated audio signals was also proposed. The performance of the new algorithm using real room experiments was demonstrated.

8. REFERENCES

[1] L. Parra and C. Spence, "Convolutive blind separation of non-stationary sources," *IEEE Transactions on Speech and Audio Processing*, vol. 8, pp. 320–327, May. 2000.

[2] D. T. Pham and J. Cardoso, "Blind source separation of instantaneous mixtures of nonstationary sources," *IEEE Transactions on Signal Processing*, vol. 49, pp. 1837–1848, Sept. 2001.

[3] K. Abed_Meraim, Y. Xiang, J. H. Manton, and Y. Hua, "Blind Source Separation Using Second-Order Cyclostationary Statistics," *IEEE Transactions on Signal Processing*, vol. 49, pp. 694–701, April 2001.

[4] K. Rahbar, J. Reilly, and J. H. Manton, "Blind identification of MIMO FIR systems driven by quasi-stationary sources using second order statistics: A frequency domain approach," *to appear, IEEE Transactions on Signal Processing*, 2003.

[5] B. Flury and W. Gautschi, "An algorithm for the simultaneous orthogonal transformation of several positive definite symmetric matrices to nearly orthogonal form," *Siam J. of Sci. Stat. Comp.*, vol. 7, pp. 169–184, 1986.

[6] K. Rahbar and J. Reilly, "A new frequency domain method for blind source separation of convolutive audio mixtures," *submitted to IEEE Transactions on Speech and Audio Processing; also available at www.ece.mcmaster.ca/~reilly*, Dec. 2002.

[7] N. Mitianoudis and M. Davies, "New fixed-point ica algorithms for convolved mixtures," in *International Workshop on Independent Component Analysis and Signal Separation*, pp. 633–638, Dec 2001.