Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Neurocomputing 69 (2006) 2065–2077
NEUROCOMPUTING
www.elsevier.com/locate/neucom

# A filter bank approach to independent component analysis for convolved mixtures

Hyung-Min Park[a,b,*], Chandra Shekhar Dhir[b], Sang-Hoon Oh[c], Soo-Young Lee[a,b]

[a]Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology, Daejeon, 305–701, Republic of Korea
[b]Department of BioSystems, and Brain Science Research Center, Korea Advanced Institute of Science and Technology, Daejeon, 305-701, Republic of Korea
[c]Department of Information Communication Engineering, Mokwon University, Daejeon, 302-729, Republic of Korea

Received 17 November 2004; received in revised form 25 September 2005; accepted 27 September 2005
Available online 8 June 2006
Communicated by E.W. Lang

**Abstract**

We present a filter bank approach to perform independent component analysis (ICA) for convolved mixtures. Input signals are split into subband signals and subsampled. A simplified network performs ICA on the subsampled signals, and finally independent components are synthesized. The proposed approach achieves superior performance than the frequency domain approach and faster convergence with less computational complexity than the time domain approach. Furthermore, it requires shorter unmixing filter length and less computational complexity than other filter bank approaches by designing efficient filter banks. Also, a method is proposed to resolve the permutation and scaling problems of the filter bank approach.


*Keywords:* Independent component analysis; Filter bank; Blind source separation

## 1. Introduction

Independent component analysis (ICA) is a signal processing method to express multivariate data as linear combinations of statistically independent random variables [8,13,15]. With the emerging trend of using higher order statistical methods, ICA plays an important role in numerous applications such as speech enhancement, telecommunications, medical signal processing, and feature extraction [2,6,16,17]. In real-world situations, we often meet with convolved mixtures of acoustic signals where mixing environments have very complex reverberation. Therefore, ICA for acoustic convolved mixtures is a challenging problem that has attracted much interest. Among many approaches to ICA, a simple and biologically plausible adaptive learning algorithm has been proposed with entropy maximization by Bell and Sejnowski [5]. To deal with convolved mixtures, the algorithm has been extended to deconvolution of mixtures in the time domain [25] and the frequency domain [19,22,23].

The time domain approach requires intensive computations with a long reverberation, and it shows slow convergence speed, especially for colored input signals. On the other hand, the frequency domain approach can decrease the computational load, because multiplication in each frequency bin replaces the convolution operation in the time domain. However, performance of the frequency domain approach is limited due to the fact that a long frame size is required to cover a long reverberation, whereas the number of learning data in each frequency bin decreases as the frame size increases [4].

In an effort to overcome these disadvantages of the time domain and the frequency domain approaches, we present a filter bank approach to ICA for convolved mixtures. In this approach, the input signals are split into a number of subbands. Then, each subband signal is decimated and used for ICA. Using an oversampled filter bank that is alias-free and provides near perfect reconstruction, we

*Corresponding author. Department of BioSystems, and Brain Science Research Center, Korea Advanced Institute of Science and Technology, Daejeon, 305-701, Republic of Korea. Tel.: +82 42 869 5351;
fax: +82 42 869 8490.
*E-mail address:* hmpark@neuron.kaist.ac.kr (H.-M. Park).

0925-2312/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2005.09.016

attempt to obtain desired independent components, which have negligible side-effects from the filter bank. Since the ICA algorithm in each subband is basically the same as the time domain approach, the filter bank approach does not have the performance limitation of the frequency domain approach. It also enables us to select the number of filters of the filter bank independently of reverberation. In addition, the decimation of the subband signals provides computational savings and a convergence speed that is much faster than that of the time domain approach [28,30].

Like the frequency domain approach, filter bank approaches also have permutation and scaling indeterminacy, which is general in ICA algorithms. In these approaches, fixing the indeterminacy is very essential and important to provide desired independent components since even one permutation error makes severe performance degradation. Therefore, the indeterminacy should be carefully considered and methods to fix the indeterminacy should be provided according to their own approaches. As a solution to fix the indeterminacy, we propose a modification of the Murata's method [18] by using correlations among envelopes of subband frequency spectra. In addition, we design filter banks whose decimation factor is slightly smaller than the number of subbands. Therefore, the required number of taps in unmixing filters is rather small to span a time range, and the computational complexity to perform ICA is rather little.

The entropy maximization algorithm is employed as the learning rule of ICA networks for fair comparison of our filter bank approach with the conventional time domain and frequency domain approaches. The comparison demonstrates the efficiency of the proposed approach against the time domain approach as well as the frequency domain approach clearly.

The remainder of the paper is organized as follows: Section 2 briefly reviews the time domain and the frequency domain approaches to ICA for convolved mixtures. Section 3 presents our filter bank approach to ICA and compares properties of the approach with those of the other approaches. In Section 4, indeterminacy of ICA is discussed, and a method to solve the indeterminacy for the filter bank approach is also proposed. The proposed method is compared with the other approaches through experiments on blind source separation in Section 5. Finally, some concluding remarks are presented in Section 6.

## 2. Independent component analysis

Let us consider a set of unknown independent components, $\mathbf{s}(n) = [s_1(n), s_2(n), \ldots, s_N(n)]^{\mathrm{T}}$, such that the components $s_i(n)$ are zero-mean and mutually independent. If mixing involves convolution and time-delays, an observation is

$$x_i(n) = \sum_{j=1}^{N} \sum_{m=0}^{L_m-1} a_{ij}(m) s_j(n-m), \tag{1}$$

where $a_{ij}(m)$ denotes a mixing filter coefficient [16].

To obtain the independent components from the observations, a feedback architecture [15,26] can be considered as

$$u_i(n) = \sum_{m=0}^{L_a} w_{ii}(m) x_i(n-m) + \sum_{j=1, j\neq i}^{N} \sum_{m=1}^{L_a} w_{ij}(m) u_j(n-m), \tag{2}$$

where adaptive filters $w_{ij}(m)$ force outputs $u_i(n)$ to reproduce the original independent components $s_i(n)$. Entropy maximization algorithm provides learning rules of the adaptive filter coefficients as follows [25]:

$$\begin{aligned}
&\Delta w_{ii}(0) \propto 1/w_{ii}(0) - \varphi(u_i(n)) x_i(n),\\
&\Delta w_{ii}(m) \propto -\varphi(u_i(n)) x_i(n-m), \quad m \neq 0,\\
&\Delta w_{ij}(m) \propto -\varphi(u_i(n)) u_j(n-m), \quad i \neq j,\\
&\qquad \varphi(u_i(n)) = -\frac{\partial p(u_i(n))/\partial u_i(n)}{p(u_i(n))},
\end{aligned} \tag{3}$$

where $\varphi(\cdot)$ is called a score function and $p(u_i)$ denotes the probability density function (pdf) of $u_i$. With a long reverberation, the time domain approach has a heavy computational load to compute convolution of long filters and updating a large number of filter coefficients. In addition, convergence speed is slow, especially for colored input signals such as speech signals.

Instead of the time domain approach, one can consider the frequency domain approach [22,23], wherein the convolved mixtures can be expressed as

$$\mathbf{x}(f,n) = \mathbf{A}(f)\mathbf{s}(f,n), \quad \forall f. \tag{4}$$

Here $\mathbf{x}(f,n)$ and $\mathbf{s}(f,n)$ are vectors, each of which is a frequency component of the mixtures and the independent components at frequency $f$, respectively. $\mathbf{A}(f)$ denotes a matrix containing elements of the frequency transforms of the mixing filters at frequency $f$. From Eq. (4), one can reason that a convolved mixture can be represented by a set of instantaneous mixtures in the frequency domain. Thus, the independent components can be recovered by applying ICA for the instantaneous mixtures in each frequency bin. Contrary to the time domain approach, the input signals $\mathbf{x}(f,n)$ are complex numbers. In order to deal with complex-valued data, a score function was proposed as [20]

$$\varphi(u_i) = -\frac{\partial p(|u_i|)/\partial |u_i|}{p(|u_i|)} \exp(j \cdot \angle u_i). \tag{5}$$

Applying the natural gradient from Amari et al. [2,7], the entropy maximization algorithm in each frequency bin is

$$\Delta \mathbf{W}(f) \propto [\mathbf{I} - \varphi(\mathbf{u}(f,n))\mathbf{u}^{\mathrm{H}}(f,n)]\mathbf{W}(f), \tag{6}$$

where $\mathbf{W}(f)$ is a matrix which consists of the frequency transforms of the unmixing filters at frequency $f$.

The frequency domain approach can decrease the computational load because the convolution operation in the time domain are substituted with multiplication in each frequency bin. However, a long frame size is required to cover a long reverberation. To maintain computational efficiency and obtain data that are not significantly overlapped with those in adjacent frames, the frame shift has to increase as the frame size increases. Then, the number of data in each frequency bin decreases, causing a shortage of data to learn the unmixing matrices and to measure the independence of outputs. Therefore, the performance of the frequency domain approach is limited because of the above-mentioned trade-off between covering a long reverberation and sufficient amount of learning data [4].

## 3. A filter bank approach to ICA

In proposing a filter bank approach to ICA, we first consider filter banks. Many researchers have studied adaptive filtering in subbands generally with least-mean-square (LMS) type algorithms [10,30,31]. If the input signals are decomposed by critically sampled filter banks, cross adaptive filters between adjacent bands are required to compensate for the distortion caused by aliasing [10], or spectral gaps are required in order to avoid aliasing [31]. However, the cross adaptive filters introduce additional adaptive parameters and may induce slow convergence speed with poor performance. On the other hand, the spectral gaps distort reconstructed signals.

With oversampled filter banks, in which the decimation factor $M$ is smaller than the number $K$ of analysis filters, aliasing can be neglected with each filter having a high stopband attenuation. The oversampled filter banks make it possible to perform adaptive filtering without requiring cross adaptive filters or distorting reconstructed signals. The oversampled filter banks can be efficiently and systematically constructed by complex modulation from a real-valued low-pass prototype filter [12,29,30]. In the filter bank, analysis filters $h(k,n)$ are obtained by a generalized discrete Fourier transform (GDFT) [9],

$$h(k,n) = e^{j(2\pi/K)(k+1/2)(n-(L_q-1)/2)} \cdot q(n),$$
$$k = 0,1,\ldots,K-1, \;\; n = 0,1,\ldots,L_q-1, \tag{7}$$

where $L_q$ is the length of the prototype filter $q(n)$. Complex-conjugate and time-reversed versions of the analysis filters,

$$f(k,n) = \tilde{h}(k,n) = h^*(k, L_q - n - 1), \tag{8}$$

are selected for synthesis filters. The analysis and synthesis filters can be derived from one prototype filter $q(n)$, and the prototype filter can be designed by iterative least-squares algorithm with a cost function that considers reconstructiveness and stopband attenuation [12]. In addition, we can implement the filter bank efficiently by employing polyphase representation of the analysis and synthesis filters [29,30].

When we perform ICA in the oversampled filter bank, adaptive parameters in each subband can be adjusted without any information of the other subbands because of the negligible aliasing of the filter bank [27,29,30]. Thus, the filter bank approach is appropriate for parallel processing. Fig. 1 shows a $2 \times 2$ network for the oversampled filter bank approach to ICA. The input signals, which are mixtures of unknown independent components, are decomposed into subband signals by analysis filters. Then, each subband signal is subsampled by a factor $M$. Although the input signals are split into subband signals, each subband still covers a somewhat broad frequency band. Moreover, when the subband signal is subsampled, the decimation factor $M$ is usually much smaller than the reverberation length in mixing environments. For blind source separation on acoustic mixtures as an example, the decimation factor and the reverberation length are normally ten or several tens and hundreds or thousands, respectively. Therefore, one should regard the subsampled signals as convolved mixtures whose reverberation length decreases by a factor $M$, and a typical ICA algorithm for convolved mixtures can be used to obtain independent components from the subsampled signals in each subband. Here, the unmixing filter length is much shorter than that of the full-band time domain approach. Each output signal from the ICA network is expanded, and the original independent components can be reconstructed from the subband output signals through synthesis filters after fixing permutation and scaling.

As a method to perform ICA for convolved mixtures in each subband, we may apply the time domain approach to ICA in Section 2. With the oversampled filter bank above, subband signals are complex-valued data. Therefore, the learning rules of the adaptive filter coefficients are changed to deal with complex-valued data, and the polar-coordinate

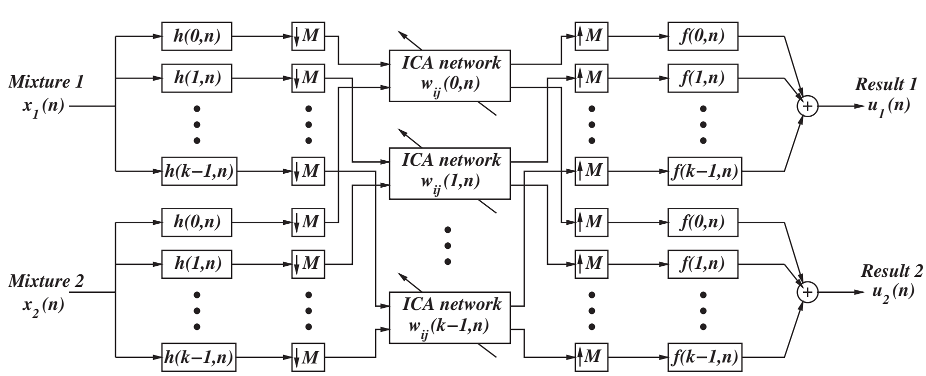


Fig. 1. A $2 \times 2$ network for the oversampled filter bank approach to ICA.

based score function Eq. (5) is used in each subband. Using a feedback network in each subband and considering complex-valued data, the learning rules are

$$\begin{aligned} \Delta w_{ii}(k,0) &\propto 1/w_{ii}^*(k,0) - \varphi(u_i(k,n))x_i^*(k,n), \\ \Delta w_{ii}(k,m) &\propto -\varphi(u_i(k,n))x_i^*(k,n-m), \quad m \neq 0, \\ \Delta w_{ij}(k,m) &\propto -\varphi(u_i(k,n))u_j^*(k,n-m), \quad i \neq j, \end{aligned} \tag{9}$$

where $w_{ij}(k,m)$, $u_i(k,n)$, and $x_i(k,n)$ represent adaptive filter coefficients, estimated independent components, and input signals in the $k$th subband, respectively.

In summary, the procedure for the filter bank approach will be described as follows:

*Step* 1: Split each mixture $x_i(n)$ into subband signals $\hat{x}_i(k,n)$ by analysis filters $h(k,m)$ to give

$$\hat{x}_i(k,n) = \sum_{m=0}^{L_q-1} h(k,m)x_i(n-m), \quad i = 1,2,\ldots,N, \tag{10}$$

where $k$ denotes the subband index.

*Step* 2: Decimate(subsample) each subband signal to make subband input signal for the ICA network by

$$x_i(k,n) = \hat{x}_i(k,nM), \quad n = 0,1,2,\ldots, \tag{11}$$

where $M$ denotes the decimation factor. One sample is obtained at the decimated rate for each subband signal.

*Step* 3: At the subsampled rate, collect subband input signals $x_i(k,n)$ for the corresponding subband (here, the $k$th subband), and compute subband output signals $u_i(k,n)$ for the ICA network. Using a feedback architecture for the ICA network in each subband,

$$\begin{aligned} u_i(k,n) = &\sum_{m=0}^{L_{sa}} w_{ii}(k,m)x_i(k,n-m) \\ &+ \sum_{j=1, j\neq i}^{N} \sum_{m=1}^{L_{sa}} w_{ij}(k,m)u_j(k,n-m), \\ &i = 1,2,\ldots,N, \end{aligned} \tag{12}$$

where $w_{ij}(k,m)$ denotes adaptive filter coefficients and $L_{sa}$ is the adaptive filter length.

*Step* 4: At the subsampled rate, update the adaptive filter coefficients for the ICA network in each subband. For the feedback architecture represented by Eq. (12), use Eq. (9) as learning rules.

*Step* 5: Fix permutation and scaling of the subband output signals $u_i(k,n)$. (For more detail, refer to the next section.)

*Step* 6: Expand each subband output signal by

$$\hat{u}_i(k,n) = \begin{cases} u_i(k,l) & \text{if } n = lM, l = 0,1,2,\ldots, \\ 0 & \text{otherwise,} \end{cases} \tag{13}$$

where $M$ denotes the decimation factor.

*Step* 7: Reconstruct desired independent components $u_i(n)$ from the expanded subband output signals $\hat{u}_i(k,n)$ using synthesis filters $f(k,m)$. That is,

$$u_i(n) = \sum_{k=0}^{K-1} \sum_{m=0}^{L_q-1} f(k,m)u_i(k,n-m), \quad i = 1,2,\ldots,N. \tag{14}$$

In the filter bank approach, the mixtures at each subband are considered as convolved mixtures, and the decimation factor can be chosen by designing an appropriate filter bank. If one uses a decimation factor which does not degrade the performance, the filter bank approach to ICA does not have the performance limitation of the frequency domain approach. In addition, the ICA network in each subband can have a much simpler structure (composed of much shorter filters) than that for the full-band time domain approach because the input subband signals are decimated by a factor $M$ and approximately one $M$th filter length in each subband is sufficient to span the corresponding time ranges of the time domain approach. Since the ICA network in each subband processes signals at the subsampled rate and the input signals of the network are also decimated, computational complexity is considerably reduced for long-length adaptive filters.

Let us assume that we have $L_a$ adaptive filter coefficients for a branch of the time domain approach. Assuming that there are $N$ mixtures convolved from $N$ independent components and direct filters are forced to scaling factors, the number of adaptive filter coefficients is approximately $(N^2-N)L_a$. Then, $2(N^2-N)L_a$ number of multiplications are required to compute outputs of the ICA network and update the filter coefficients whenever a sample has been obtained in an observation. On the other hand, the number of required filter coefficients of a branch $L_{sa}$ is approximately $L_a/M$ in each subband for the filter bank approach with a decimation factor $M$. For real input signals, we need the lower half subbands of a uniform GDFT filter bank because of symmetry. With a $K$-channel oversampled filter bank, the required number $N_a$ of multiplications per sampling interval for the observations is

$$N_a \approx \frac{1}{M} 4 \frac{K}{2} 2(N^2-N)\frac{L_a}{M}. \tag{15}$$

The factor $1/M$ and 4 are respectively multiplied because the filter bank approach processes signals at the subsampled rate in each subband and one complex multiplication equals four real multiplications. Therefore, the number of multiplications for the filter bank approach is approximately $2K/M^2$ times as large as that for the time domain approach. In addition, computational savings will be maximized for the same number of subbands $K$ when the decimation factor $M$ approaches to $K$ with negligible aliasing of the filter banks.

Additionally, in this approach the number of subbands can be chosen regardless of reverberation. Also, the decimation operation omits several adjacent samples, so that correlation between adjacent samples in a sub sampled signal is much smaller than that in the signal before decimation. Therefore, this approach improves convergence of the adaptive filter coefficients of the ICA network because subband input signals might be more whitened by decimation and the number of the adaptive filter coefficients in each subband is much smaller than that of the time domain approach.

As some related works to date, several papers have been reported on blind source separation in subbands. Some researchers have proposed use of overcomplete subband representation and orthogonal filter banks [11,24]. Although they generated oversampled signals using some operations which were different from the Fourier transform, the methods basically regarded the subband signals as instantaneous mixtures rather than convolved mixtures. Therefore, they can be regarded as modifications of the frequency domain approach in this paper. As such, they still have the basic drawbacks of the frequency domain approach, although they might mitigate the problems. These methods require a lot of subbands (frequency bins in this paper) in order to cover a long reverberation. In this case, too many subbands worsen the performance limitation problem because of shortage of data in subbands. They also make the permutation problem severe because of the great number of subbands and inaccurate estimation of envelopes in subbands.

Other authors have attempted to separate convolved mixtures in each subband [3,14]. However, they did not consider the permutation indeterminacy or mitigated it by using null beamformers as the initial value of the ICA network. Note that this method cannot function as a substantial solution for permutation correction which is very important to recover desired independent components [3,14]. On the contrary, in the next section, we will provide an algorithm which is fit for the permutation problem of the filter bank approach. In addition, the separation algorithm in the paper authored by Araki et al. does not converge at all for the long unmixing filters without initialization using null beamformers, and the decimation factor of the filter bank was one-fourth of the number of subbands [3]. Since the decimation factor is rather small comparing with the number of subbands, it increases unmixing filter length to span a time range and the computational load regardless of separation methods in subbands. Regarding filter bank structure, the Araki's structure required additional computation for single side-band modulation and demodulation. Furthermore, in the GDFT filter bank structure of this paper, the computational burden for analyzing and synthesizing signals can be reduced by using special relation among filters which is GDFT modulation in addition to polyphase representation [29]. In the Huang's method, the used filter banks suffered from large aliasing distortion or did not assure perfect reconstruction of output signals [14]. The alias-free property and the perfect reconstruction are very essential requirements in order to use the filter banks without any side-effects because they limit the performance of the methods primarily apart from separation capability in subbands. They reported that the increased aliasing distortions significantly deteriorated the performance with more than 4 subbands [14].

## 4. Discussion on ICA indeterminacy

For convolutive mixtures which are mixed from temporally correlated independent components, there exists indeterminacy of outputs up to permutation and arbitrary filtering. Entropy maximization attempts to make the outputs temporally whitened, which may degrade outputs in many applications such as separation of natural signals. Whitening the recovered outputs can be avoided by forcing direct filters $w_{ii}(k,m)$ to scaling factors [25].

The filter bank approach has an ICA network in each subband whose filter coefficients can be adapted independently of the other subbands. Thus, the filter bank approach has the same permutation and scaling problems as the frequency domain approach. Assuming that the independent components have time-varying statistical properties, we propose a modification of the Murata's method to fix permutation and scaling as follows:

*Step* 1: Normalize estimated independent components $u_i(k,n)$ in each subband by the corresponding scaling factors $w_{ii}(k,d_i)$ to give

$$v_i(k,n) = \frac{u_i(k,n)}{w_{ii}(k,d_i)}, \quad i = 1,2,\ldots,N, \tag{16}$$

where $d_i$ is an appropriate delay for unmixing systems.

*Step* 2: Compute envelopes of frequency spectra by

$$\xi[v_i(k,n)] = \frac{1}{2T+1}\sum_{n'=n-T}^{n+T} |v_i(k,n')|, \tag{17}$$

where $T$ is a positive constant.

*Step* 3: Find a subband $k_1$ which shows the smallest similarity. That is,

$$k_1 = \arg\min_k \sum_{i\neq j} r\{\xi[v_i(k,n)], \xi[v_j(k,n)]\}. \tag{18}$$

Here, $r$ denotes a normalized correlation expressed as

$$r\{\alpha(n),\beta(n)\} = \frac{1/L_n\sum_n \alpha(n)\beta(n)}{\sqrt{1/L_n\sum_n \alpha^2(n)1/L_n\sum_n \beta^2(n)}}, \tag{19}$$

where $L_n$ denotes the length of $\alpha(n)$ and $\beta(n)$.

*Step* 4: For $k_1$, assign estimated independent components to specific outputs by

$$u_i'(k_1,n) = v_i(k_1,n), \quad i = 1,2,\ldots,N. \tag{20}$$

*Step* 5: For $k = \{k_1+1, k_1+2, \ldots, K-1\}$, find a permutation $\sigma(i)$ which maximizes the normalized correlation between the envelope of the $k$th subband and a weighted average from the envelopes of the previous subbands with a forgetting factor $\mu$, and assign the permutation to outputs. That is,

$$\sigma(i) = \arg\max_{\sigma(i)} \sum_{i=1}^{N} r\{\xi[v_{\sigma(i)}(k,n)], \Xi[u_i'(k,n)]\}, \tag{21}$$

where

$$\Xi[u_i'(k,n)] = \begin{cases} \xi[u_i'(k_1,n)] & \text{if } k = k_1+1, \\ \mu\Xi[u_i'(k-1,n)] + (1-\mu) \\ \quad \times \xi[u_i'(k-1,n)] & \text{if } k > k_1+1, \end{cases} \tag{22}$$

and

$$u_i'(k,n) = v_{\sigma(i)}(k,n), \quad i = 1,2,\ldots,N. \tag{23}$$

*Step* 6: For $k = \{k_1-1, k_1-2, \ldots, 0\}$, repeat step 5 except

$$\Xi[u_i'(k,n)] = \begin{cases} \xi[u_i'(k_1,n)] & \text{if } k = k_1-1, \\ \mu\Xi[u_i'(k+1,n)] + (1-\mu) \\ \quad \times \xi[u_i'(k+1,n)] & \text{if } k < k_1-1. \end{cases} \tag{24}$$

In this method, step 1 removes the ambiguity of scaling for estimated independent components of the ICA network in each subband of the filter bank approach. When the original Murata's method is used in the frequency domain approach [18], each estimated independent component is multiplied by the inversed unmixing matrix in each frequency bin in order to avoid any ambiguity of scaling as follows:

$$\mathbf{v}(f,n;i) = \mathbf{W}(f)^{-1}[0 \ \cdots \ 0 \ u_i(f,n) \ 0 \ \cdots \ 0]^{\mathrm{T}}, \tag{25}$$

where $u_i(f,n)$ denotes the $i$th element of $\mathbf{u}(f,n)$. However, as we described in step 1, the ambiguity of scaling in the filter bank approach, is removed by normalizing the estimated independent components with the corresponding scaling factors used as the direct filters instead of using the decomposition of the frequency spectra in Eq. (25). Therefore, the vector $\mathbf{v}(f,n;i)$ in Eq. (25) corresponds to a scalar value $v_i(k,n)$ with subband index $k$, and for computing the envelopes of frequency spectra, step 2 has a different formulation from the original Murata's method for the frequency domain approach

$$\xi[\mathbf{v}(f,n;i)] = \frac{1}{2T+1}\sum_{n'=n-T}^{n+T}\sum_{j=1}^{N}|v_j(f,n';i)|, \quad i = 1,2,\ldots,N, \tag{26}$$

where $v_j(f,n;i)$ denotes the $j$th element of $\mathbf{v}(f,n;i)$.

With the definition of similarity given by

$$\mathrm{sim}(f) \equiv \sum_{i\neq j} r\{\xi[\mathbf{v}(f,n;i)], \xi[\mathbf{v}(f,n;j)]\}, \tag{27}$$

the original Murata's method sorts frequency bins in order of weakness of similarity among independent components, so that

$$\mathrm{sim}(f_1) \leqslant \mathrm{sim}(f_2) \leqslant \cdots \leqslant \mathrm{sim}(f_F). \tag{28}$$

Here $F$ is the number of frequency bins. For the frequency bin $f_1$ which has the smallest correlation, its independent components are assigned to specific outputs $\mathbf{u}'(f_1,n;i) = \mathbf{v}(f_1,n;i)$. Then, for the frequency bins $\{f_l, l = 2,3,\ldots,F\}$ sorted in the increasing order of the correlation, the independent components are assigned to the outputs that have more correlation between the envelopes of the frequency bins. That is,

$$\mathbf{u}'(f_l,n;i) = \mathbf{v}(f_l,n;\sigma(i)), \tag{29}$$

where the permutation is given as

$$\sigma(i) = \arg\max_{\sigma(i)} \sum_{i=1}^{N} r\left\{\xi[\mathbf{v}(f_l,n;\sigma(i))], \sum_{j=1}^{l-1}\xi[\mathbf{u}'(f_j,n;i)]\right\}. \tag{30}$$

The process is repeated in turns till all the frequency bins are covered.

However, if the next frequency bin is far from the previous frequency bins, the envelopes of the frequency spectra may be very different even though the frequency spectra are obtained from the same signal. Therefore, we do not use this order for the subbands to fix the permutation in the filter bank approach except a subband which has the minimum similarity, and we just find the subband by Eq. (18) in step 3. After assignment by Eq. (20) has been done for the $k_1$th subband, we perform the assignments for the subbands adjacent to the previous subbands in step 5 and 6 instead of the subband which has the next smallest similarity. In addition, when independent components are assigned to outputs, we use the normalized correlation between the envelope of a subband and the weighted average from the envelopes of the previous subbands by Eqs. (22) and (24). This will emphasize the envelopes of close subbands. In this way, we may get more desirable results than those obtained by using Eq. (30) since the envelopes from close subbands will be more similar than those from distant subbands.

In the frequency domain approach, we need many frequency bins to cover a long reverberation, e.g. thousands of bins for a normal room reverberation. In this case, the frame size and the frame shift are also very large and the number of data in each frequency bin becomes quite small. Thus, the envelope of each frequency spectrum cannot be accurately estimated to fix the permutation. However, in the filter bank approach, we can determine the bandwidths of subbands regardless of reverberation. Therefore, we can resolve the permutation problem quite easily in the filter bank approach because each subband can have a sufficiently broad band so as to exactly and minutely estimate the envelope of the frequency spectrum.

## 5. Experimental results

### 5.1. Experiments on blind source separation for simulated mixtures

We have performed experiments on blind source separation for simulated mixtures with the proposed filter bank approach. We have used two streams of speech as the independent source signals. Each signal had 5 s length at 16 kHz sampling rate. To construct a $2 \times 2$ mixing system, impulse responses were generated by the image method, which simulates acoustics between two points in a rectangular room [1]. Fig. 2 shows a virtual room to simulate the impulse responses from 2 speaker points to 2 microphone points, and Fig. 3 shows the resulting impulse responses. All reflection coefficients were 0.6, respectively.

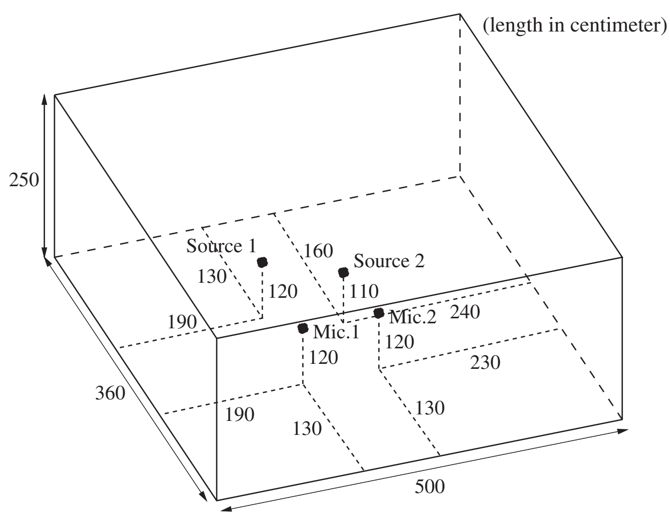


Fig. 2. Virtual room to simulate impulse responses from 2 speaker points to 2 microphone points.

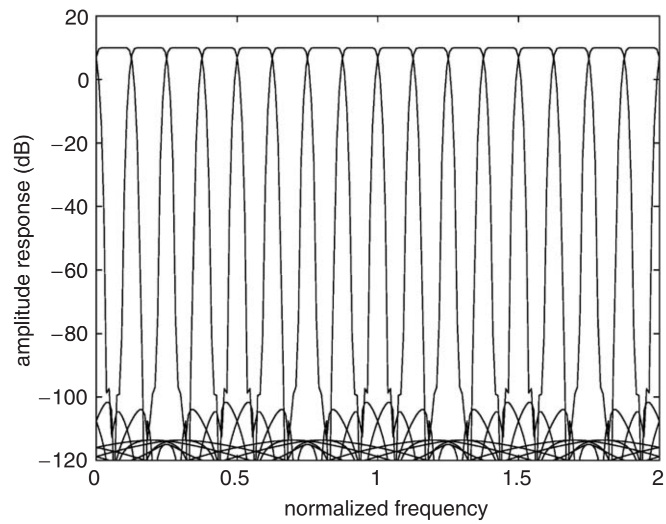


Fig. 4. Frequency response of analysis filters of a uniform sixteen-channel oversampled filter bank.

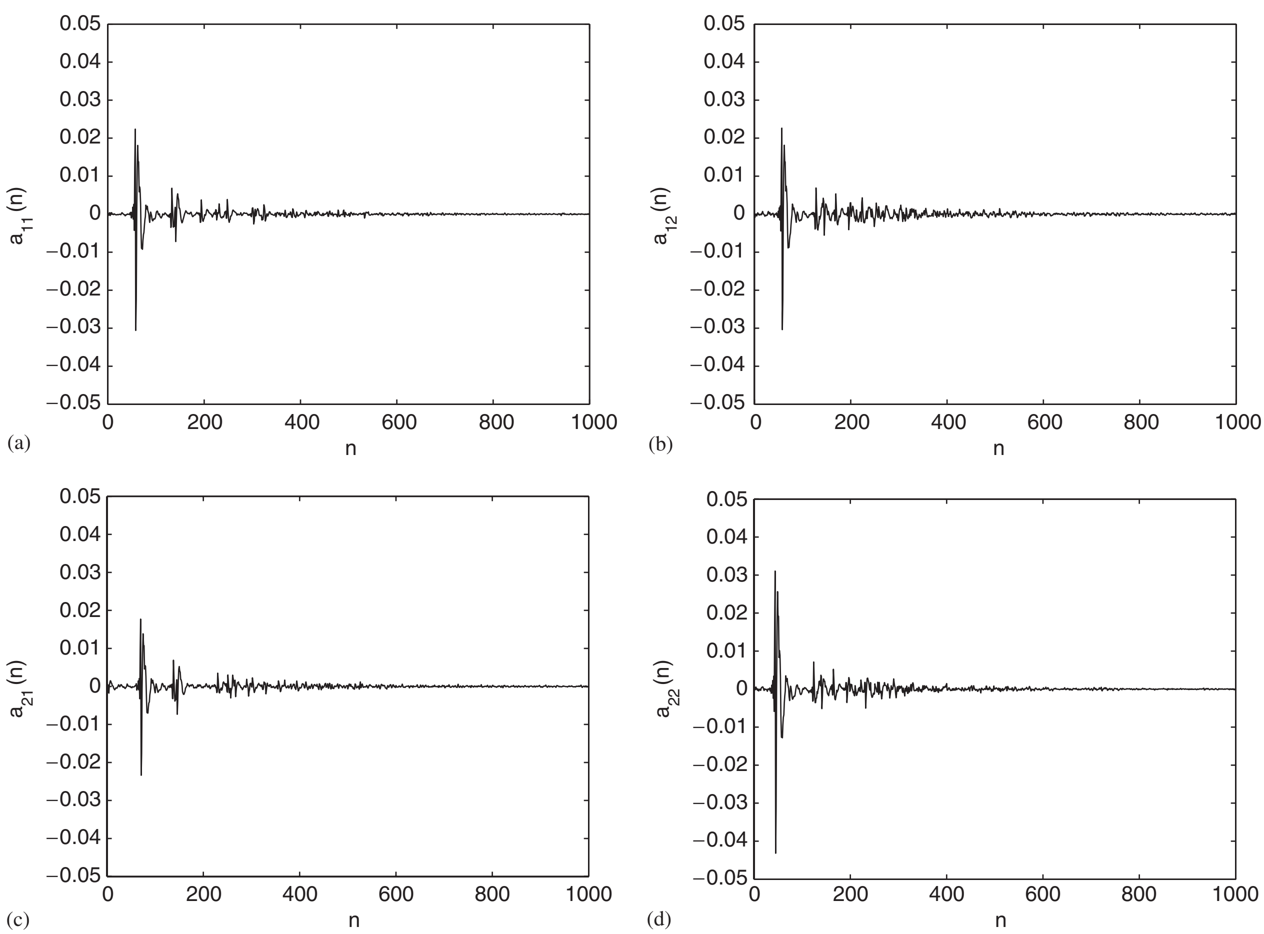


Fig. 3. Impulse responses of the mixing system for experiments on blind source separation: (a) $a_{11}$; (b)$a_{12}$; (c)$a_{21}$; (d)$a_{22}$.

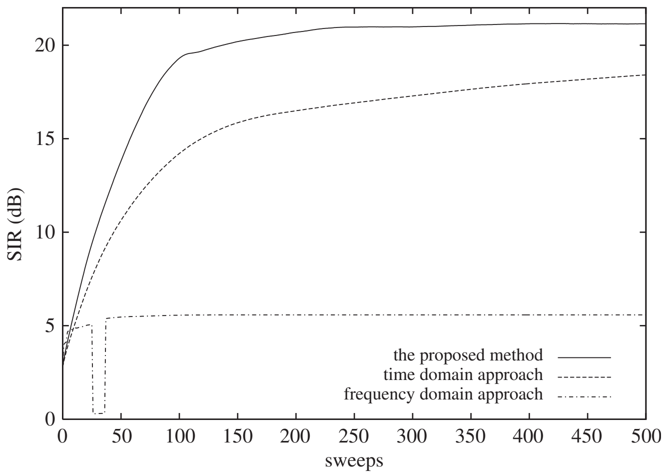


Fig. 5. Learning curves of the three approaches to blind source separation.

Fig. 4 shows the frequency response of analysis filters of a uniform sixteen-channel oversampled filter bank using the GDFT. The filter bank was designed for alias-free decimation by a factor $M = 10$, and it was constructed from a prototype filter with 220 taps by iterative least-squares algorithm with a cost function which considers reconstructiveness and stopband attenuation. The ratio of the decimation factor to the number of subbands was larger than that in the paper authored by Araki et al. [3]. As mentioned in Section 3, the efficient filter bank reduced the required filter length in the ICA network and the computational burden.

With the oversampled filter bank above, we have performed experiments on the proposed filter bank approach to blind source separation. For the separation system in each subband, we have used a feedback network in which the number of taps of each filter was $\lceil \frac{2048}{M} \rceil$. Based on the fact that speech signal approximately follows Laplacian distribution, $\mathrm{sgn}(|u_i|)\exp(j \cdot \angle u_i)$ was used as the score function $\varphi(u_i)$. Fig. 5 shows a learning curve of the proposed filter bank approach. For comparison, learning curves of the time domain and the frequency domain approaches are also displayed. Experimental results were compared in terms of signal-to-interference ratio (SIR). For a $2 \times 2$ mixing/unmixing system, the SIR is defined as a ratio of the signal power to the interference power at the outputs [21],

$$\mathrm{SIR(d}B) = \frac{1}{2}\left| 10 \log\left( \frac{\langle (u_{1,s_1}(n))^2 \rangle}{\langle (u_{1,s_2}(n))^2 \rangle} \cdot \frac{\langle (u_{2,s_2}(n))^2 \rangle}{\langle (u_{2,s_1}(n))^2 \rangle} \right) \right|. \tag{31}$$

In Eq. (31), $u_{i,s_j}(n)$ denotes the $i$th output of the cascaded mixing/unmixing system when only $s_j(n)$ is active. We have used a feedback network in which each filter length was 2048 taps for the time domain approach. In the frequency domain approach, the frame size was 2048, and the frame shift was an eighth of the frame size. The SIRs of the frequency domain approach were much smaller than those of the other two approaches. This supports the previous argument that the frequency domain approach has a performance limitation which arises from the conflict between long reverberation covering and sufficient learning data. Abruption in the SIRs of the frequency domain approach is caused by wrong permutation correction because the envelopes of estimated independent components are very similar in the beginning part. In addition, the learning curves show that the filter bank approach had much faster convergence speed than the time domain approach since subband signals were decimated in each subband for the filter bank approach.

We also compared the proposed method with the Araki's method [3] for the same blind source separation problem. For the Araki's method, we have used the same number of subbands as our method, whereas the decimation factor was 4 since they used a fourth of the number of subbands for the decimation factor [3]. For the separation network of the Araki's method, the number of taps of each filter was $\lceil \frac{2048}{4} \rceil$. As in the Araki's paper [3], we initialized the filters as the null beamformers for $\pm 60^\circ$. Fig. 6 shows the learning curves of the methods with the SIR at the outputs. The Araki's method shows rather small SIRs. For the cases where the initialization fits a problem (direction of the sources is close to that of the initial null beamformer), the Araki's method also shows a satisfactory performance. For this mixing system, however, the direction of the sources is very different from the initialization.

Fig. 7(a) shows envelopes of estimated independent components according to Eq. (17) for the lower half 8 subbands in the proposed method. In each subband, an independent component shows a quite different envelope from the other of the same subband, which indicates that source separation has been successfully achieved. In order to quantify the ease of fixing the permutation, each bar in Fig. 8(a) represents the normalized correlation between the envelope of a subband and the weighted average of the

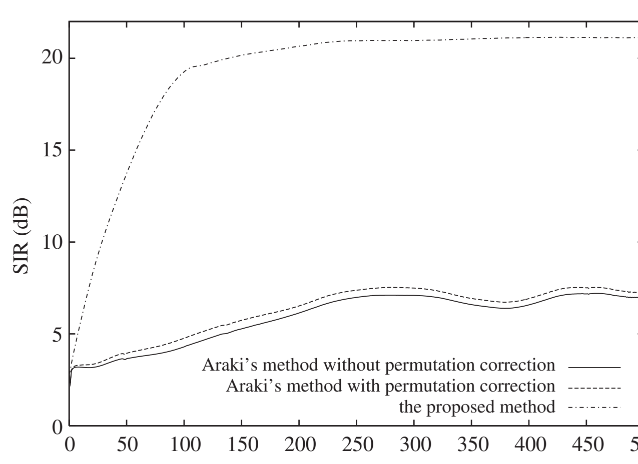


Fig. 6. Learning curves of the proposed method and the Araki's method for blind source separation.

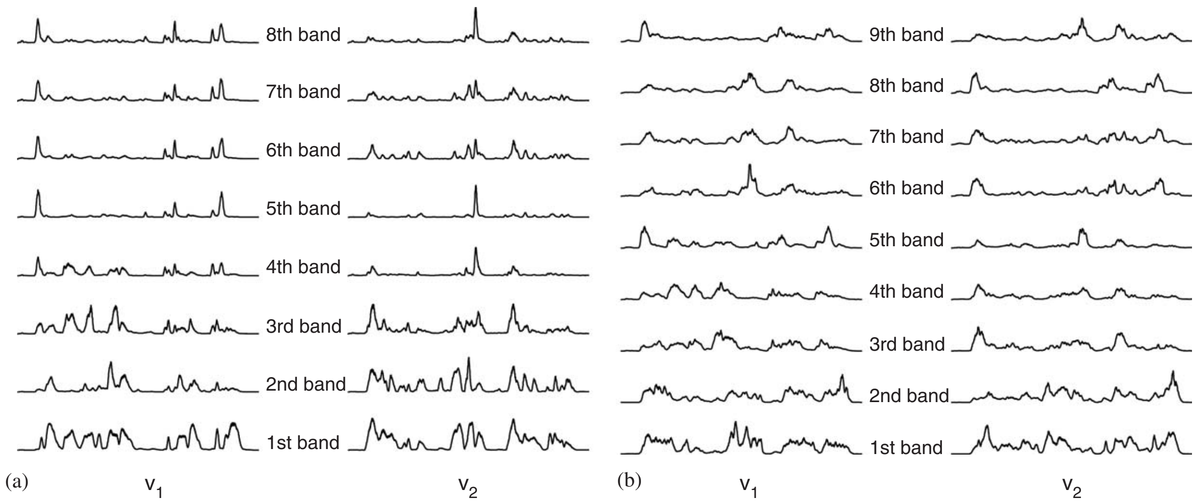


Fig. 7. Envelopes of estimated independent components: (a) the proposed method; (b) the Araki's method.

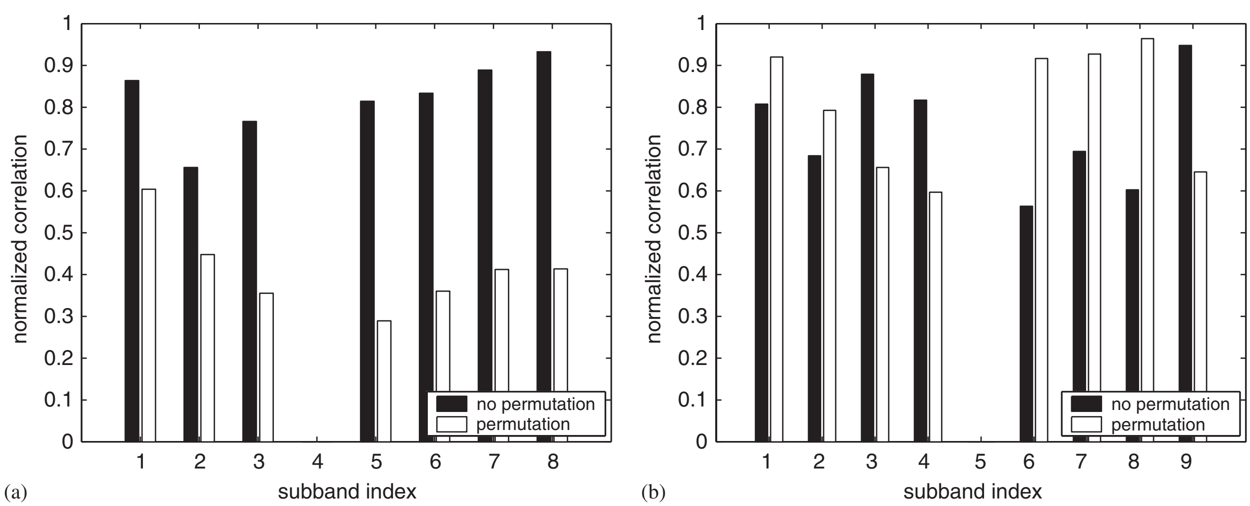


Fig. 8. Normalized correlations between the envelopes of estimated independent components ("No permutation" means that $\sigma(1) = 1$ and $\sigma(2) = 2$, whereas "permutation" means that $\sigma(1) = 2$ and $\sigma(2) = 1$): (a) the proposed method; (b) the Araki's method.

envelopes of the previous subbands expressed as

$$\frac{1}{2}\sum_{i=1}^{2} r\{\xi[v_{\sigma(i)}(k,n)], \Xi[u_i'(k,n)]\}, \tag{32}$$

where $\Xi[u_i'(k,n)]$ can be obtained from Eqs. (22) and (24). In this source separation problem, the 4th subband showed the smallest similarity. For the other subbands, differences of the normalized correlations were greater than 0.2. The filter bank approach provided quite large differences of the normalized correlations, and we confirmed that the method successfully resolved the permutation problem.

For the Araki's method, we have also displayed the envelopes of estimated independent components for the lower 9 subbands in Fig. 7(b) and the normalized correlations between the envelopes in Fig. 8(b). In order to cover frequency range from 0 to $\pi$, the proposed method used 8 subbands whose center frequencies were $(\pi/16) \times (2k+1), k = 0, 1, \ldots, 7$. On the other hand, the Araki's method used 9 subbands since center frequencies were $(\pi/16) \times 2k, k = 0, 1, \ldots, 8$. In Fig. 7(b), several subbands provided quite similar envelopes to each other. Therefore, one can conclude that the method is in difficulty for separation in the subbands. Furthermore, it should be noted that there are several subbands where the "permutation" has larger normalized correlation than the "no permutation". If the subbands do not perform the permutation correction as in [3], the resulting SIR may be lower than the SIR with the permutation correction as in Fig. 6. Although difference between the two SIRs is small, real speech quality is quite different. Most energy of speech is concentrated in low frequency, but human auditory system is sensitive to not only low frequency but also high frequency.

### 5.2. Experiments on blind source separation for real recorded mixtures

The proposed filter bank approach has also been applied to separate real acoustic mixtures recorded with two microphones in an office room. Mixtures of 5 s length at 16 kHz sampling rate were recorded where two streams of speech signals were used as source signals. Speakers and microphones were placed as shown in Fig. 9. In addition to the recordings of mixtures while the two speakers are active simultaneously, we also recorded signals while only one of the speaker is active. The recordings made while only one speaker is active correspond to the terms $\sum_{m=0}^{L_m-1} a_{ij}(m)s_j(n-m)$ in Eq. (1), and the signals are denoted as $x_{ij}(n)$. From these signals, we can approximately compute $u_{j,s_i}(n)$ using the estimated ICA network so as to estimate the SIR according to Eq. (31).

When recordings of acoustic signals are made in a real environment, there are many noise components such as background noise, measurement noise, and nonlinearity of equipment. Therefore, the overall output signal $u_i(n)$ has not only desired signal and interference but also noise components. To consider these noise components in a performance measure, we have used signal-to-noise ratio (SNR) which can be expressed as

$$\mathrm{SNR(d}B) = \frac{1}{2} \times 10 \log\left(\frac{\langle(\tau_1(n))^2\rangle}{\langle(v_1(n))^2\rangle} \cdot \frac{\langle(\tau_2(n))^2\rangle}{\langle(v_2(n))^2\rangle}\right), \tag{33}$$

where $\tau_i(n)$ and $v_i(n)$ denote the desired signal and noise component of $u_i(n)$ $(i = 1, 2)$, respectively. Since the normalized independent components after fixing the permutation, $u_i'(k, n)$, are used for synthesizing the overall

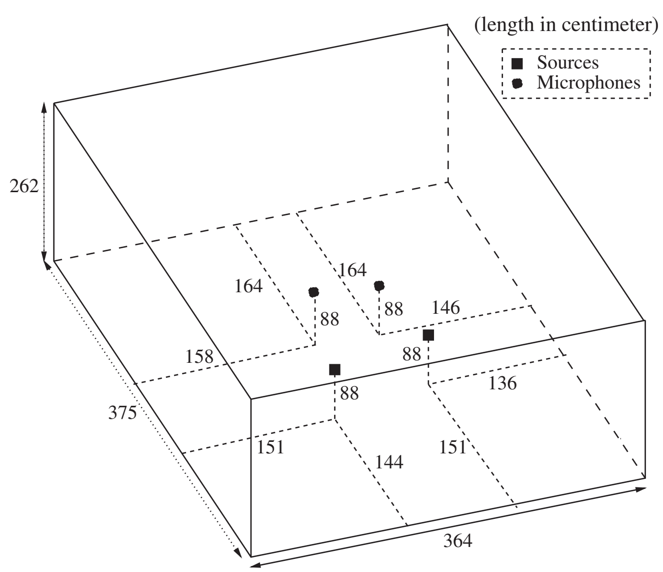


Fig. 9. Recording arrangement for 2 speakers and 2 microphones.

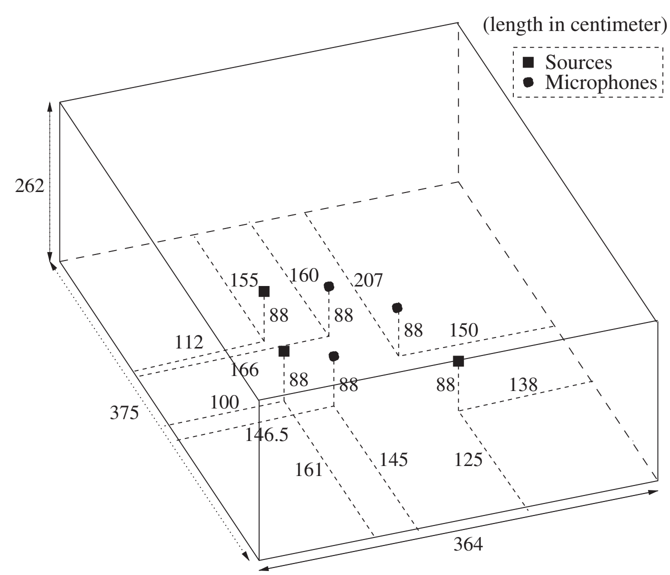


Fig. 11. Recording arrangement for 3 speakers and 3 microphones.

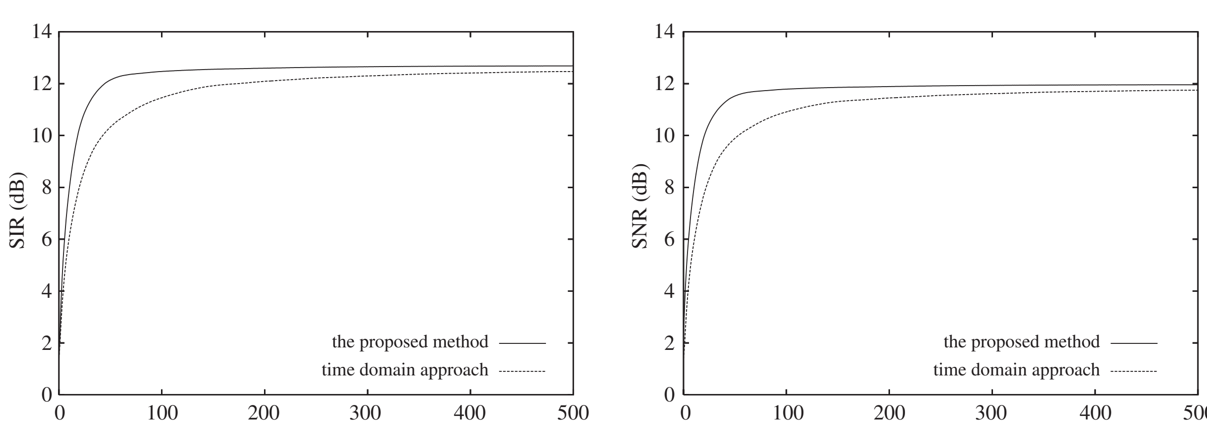


Fig. 10. Learning curves of blind source separation for mixtures recorded from a 2 × 2 mixing environment: (a) SIR; (b) SNR.

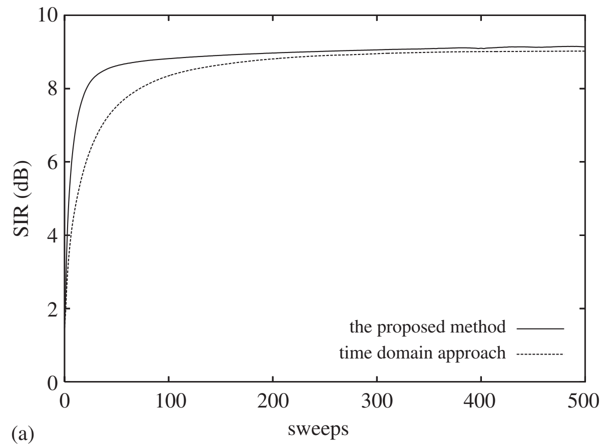


Fig. 12. Learning curves of blind source separation for mixtures recorded from a $3 \times 3$ mixing environment: (a) SIR; (b) SNR.

output signals, desired signal and noise components at the outputs can be approximately calculated as follows:

If $u_1(n)$ and $u_2(n)$, respectively, estimate $s_1(n)$ and $s_2(n)$, then

$$\begin{aligned}
\tau_1(n) &\approx u_{1,s_1}(n),\\
v_1(n) &\approx u_1(n) - u_{1,s_1}(n),\\
\tau_2(n) &\approx u_{2,s_2}(n),\\
v_2(n) &\approx u_2(n) - u_{2,s_2}(n).
\end{aligned} \tag{34}$$

As denoted before, $u_i(n)$ is the $i$th output total signal whereas $u_{i,s_j}(n)$ is the $i$th output signal of the cascaded mixing/unmixing system when only $s_j(n)$ is active. On the other hand, if $u_1(n)$ and $u_2(n)$, respectively, estimate $s_2(n)$ and $s_1(n)$, then

$$\begin{aligned}
\tau_1(n) &\approx u_{1,s_2}(n),\\
v_1(n) &\approx u_1(n) - u_{1,s_2}(n),\\
\tau_2(n) &\approx u_{2,s_1}(n),\\
v_2(n) &\approx u_2(n) - u_{2,s_1}(n).
\end{aligned} \tag{35}$$

Fig. 10 displays the SIR and SNR curves of outputs for mixtures recorded from the microphones of Fig. 9. Setup for separation was the same as in the simulated mixing. We skipped the frequency domain approach since it showed much worse performance than the other approaches in the previous experiments.

Regardless of the used methods, the final SIRs were smaller than those for the simulated mixtures in Fig. 5. That may be not only because the real mixing environment could be much more difficult to separate signals than the simulated mixing, but also because we used different signals for separation and measuring the SIRs. Moreover, it may give an important reason that recorded signals contain artifacts, such as measurement noise and nonlinearity of equipment, which obstruct exact estimation of adaptive filter coefficients. Above all, note that the learning curves of the proposed filter bank approach showed faster convergence than those of the time domain approach like in the previous subsection in terms of both the SIRs and SNRs. The difference between SIR and SNR values is due to many noise components as we mentioned in SNR definition.

We repeated the blind source separation experiment on three recorded mixtures to estimate three source signals. Fig. 11 describes position of speakers and microphones. We performed this experiment in the same way as the previous one except that a $3 \times 3$ mixing/unmixing system has been treated. Therefore, SIR and SNR had to be extended as

$$\mathrm{SIR}(\mathrm{d}B) = \max_{\{\sigma(j),j=1,2,3\}} \frac{1}{3} \times \sum_{j=1}^{3} 10 \log\left(\frac{\langle (u_{j,s_{\sigma(j)}}(n))^2\rangle}{\left\langle \left(\sum_{i=1,i\neq\sigma(j)}^{3} u_{j,s_i}(n)\right)^2\right\rangle}\right) \tag{36}$$

and

$$\mathrm{SNR}(\mathrm{d}B) \approx \max_{\{\sigma(j),j=1,2,3\}} \frac{1}{3} \times \sum_{j=1}^{3} 10 \log\left(\frac{\langle (u_{j,s_\sigma(j)}(n))^2\rangle}{\langle (u_j(n) - u_{j,s_\sigma(j)}(n))^2\rangle}\right), \tag{37}$$

where $\{\sigma(j), j = 1, 2, 3\}$ denotes a permutation of indices so that six values were compared to choose the maximum for the SIR and the SNR.[1]

Learning curves of the blind source separation problem are presented in Fig. 12. The measured performance was worse than that in the previous experiment since we have recorded more mixtures in a more complex mixing environment. Except the absolute measured values, the results were consistent with those of the problem to

[1]Note that there are six different permutations for three indices $\{1, 2, 3\}$.

estimate sources from two recorded mixtures. Since a subband covers somewhat broad band, the SIRs as well as the SNRs would be severely degraded with wrong permutation even in one subband. Therefore, the results certified that the ICA indeterminacy for the filter bank approach has been successfully resolved.

## 6. Conclusions

In this paper, we proposed a filter bank approach to perform ICA for convolved mixtures. The filter bank approach achieved a much better performance than the frequency domain approach and faster convergence speed with less computational complexity than the time domain approach. By proposing a modification of the Murata's method, we resolved the permutation and scaling problems of the filter bank approach and showed its successful working through simulations. Furthermore, the designed filter bank required shorter unmixing filter length and less computational load than those in other filter bank approaches.

## Acknowledgements

This work was supported as the Brain Neuroinformatics Research Program sponsored by Korean Ministry of Science and Technology.

## References

[1] J.B. Allen, D.A. Berkley, Image method for efficiently simulating small-room acoustics, J. Acoust. Soc. Am. 65 (4) (1979) 943–950.

[2] S. Amari, A. Cichocki, H.H. Yang, A new learning algorithm for blind source separation, in: Advances in Neural Information Processing Systems, vol. 8, MIT Press, Cambridge, MA, 1996, pp. 757–763.

[3] S. Araki, S. Makino, R. Aichner, T. Nishikawa, H. Saruwatari, Subband based blind source separation with appropriate processing for each frequency band, in: Proceedings of International Conference on ICA and BSS, 2003, pp. 499–504.

[4] S. Araki, S. Makino, R. Mukai, T. Nishikawa, H. Saruwatari, Fundamental limitation of frequency domain blind source separation for convolved mixtures of speech, in: Proceedings of International Conference on ICA and BSS, 2001, pp. 132–137.

[5] A.J. Bell, T.J. Sejnowski, An information-maximisation approach to blind separation and blind deconvolution, Neural Comput. 7 (1995) 1129–1159.

[6] A.J. Bell, T.J. Sejnowski, Edges are the independent components of natural scenes, in: Advances in Neural Information Processing Systems, vol. 9, MIT Press, Cambridge, MA, 1997, pp. 831–837.

[7] J.-F. Cardoso, B. Laheld, Equivariant adaptive source separation, IEEE Trans. Signal Process. 44 (12) (1996) 3017–3030.

[8] P. Comon, Independent component analysis, a new concept?, Signal Process. 36 (1994) 287–314.

[9] R.E. Crochiere, L.R. Rabiner, Multirate Digital Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[10] A. Gilloire, M. Vetterli, Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation, IEEE Trans. Signal Process. 40 (8) (1992) 1862–1875.

[11] N. Grbic, X.-J. Tao, S.E. Nordholm, I. Claesson, Blind signal separation using overcomplete subband representation, IEEE Trans. Speech Audio Process. 9 (5) (2001) 524–533.

[12] M. Harteneck, S. Weiß, R.W. Stewart, Design of near perfect reconstruction oversampled filter banks for subband adaptive filters, IEEE Trans. Circuits Syst. II 46 (1999) 1081–1085.

[13] S. Haykin (Ed.), Unsupervised Adaptive Filtering, Wiley, New York, 2000.

[14] J. Huang, K.-C. Yen, Y. Zhao, Subband-based adaptive decorrelation filtering for co-channel speech separation, IEEE Trans. Speech Audio Process. 8 (4) (2000) 402–406.

[15] T.-W. Lee, Independent Component Analysis, Kluwer Academic Publishers, Boston, MA, 1998.

[16] T.-W. Lee, A.J. Bell, R.H. Lambert, Blind separation of delayed and convolved sources, in: Advances in Neural Information Processing Systems, vol. 9, MIT Press, Cambridge, MA, 1997, pp. 758–764.

[17] J.-H. Lee, T.-W. Lee, H.-Y. Jung, S.-Y. Lee, On the efficient speech feature extraction based on independent component analysis, Neural Process. Lett. 15 (3) (2002) 235–245.

[18] N. Murata, S. Ikeda, A. Ziehe, An approach to blind source separation based on temporal structure of speech signals, Neurocomputing 41 (2001) 1–24.

[19] H.-M. Park, H.-Y. Jung, T.-W. Lee, S.-Y. Lee, Subband-based blind signal separation for noisy speech recognition, Electron. Lett. 35 (23) (1999) 2011–2012.

[20] H. Sawada, R. Mukai, S. Araki, S. Makino, A polar-coordinate based activation function for frequency domain blind source separation, in: Proceedings of the International Conference on ICA and BSS, 2001, pp. 663–668.

[21] D. Schobben, K. Torkkola, P. Smaragdis, Evaluation of blind signal separation methods, in: Proceedings of the International Conference on ICA and BSS, 1999, pp. 261–266.

[22] P. Smaragdis, Information theoretic approaches to source separation, Master's Thesis, MIT Media Arts and Science Department, MIT, Boston, June 1997.

[23] P. Smaragdis, Blind separation of convolved mixtures in the frequency domain, Neurocomputing 22 (1998) 21–34.

[24] M. Stanacevic, M. Cohen, G. Cauwenberghs, Blind separation of linear convolutive mixtures using orthogonal filter banks, in: Proceedings of the International Conference on ICA and BSS, 2001, pp. 260–265.

[25] K. Torkkola, Blind separation of convolved sources based on information maximization, in: Proceedings of the IEEE International Workshop on NNSP, 1996, pp. 423–432.

[26] K. Torkkola, Blind separation for audio signals—are we there yet?, in: Proceedings of the International Conference on ICA and BSS, 1999, pp. 239–244.

[27] S. Weiß, On adaptive filtering on oversampled subbands, Ph.D. Thesis, Signal Processing Division, University of Strathclyde, Glasgow, May 1998.

[28] S. Weiß, S.R. Dooley, R.W. Stewart, A.K. Nandi, Adaptive equalization in oversampled subbands, Electron. Lett. 34 (15) (1998) 1452–1453.

[29] S. Weiß, L. Lampe, R.W. Stewart, Efficient implementations of complex and real valued filter banks for comparative subband processing with an application to adaptive filtering, in: Proceedings of the First International Symposium on Communication Systems and Digital Signal Processing, 1998, pp. 32–35.

[30] S. Weiß, A. Stenger, R.W. Stewart, R. Rabenstein, Steady-state performance limitations of subband adaptive filters, IEEE Trans. Signal Process. 49 (9) (2001) 1982–1991.

[31] Y. Yamada, H. Ochi, H. Kiya, A subband adaptive filter allowing maximally decimation, IEEE J. Select. Areas Commun. 12 (9) (1994) 1548–1552.

**Hyung-Min Park** received the B.S., M.S., and Ph.D. degrees in Electrical Engineering and Computer Science from Korea Advanced Institute of Science and Technology, Daejeon, Korea, in 1997, 1999, and 2003, respectively. From 2003 to early 2005, he was a Post-Doc. at the Department of Biosystems, Korea Advanced Institute of Science and Technology. Currently, he is working in the Language Technologies Institute, Carnegie Mellon University, as a Post-Doc. His current research interests include the theory and applications of independent component analysis, blind source separation, adaptive noise canceling, and binaural or multi-microphone processing for noise-robust speech recognition.

**Chandra Shekhar Dhir** received B.Tech. in Electrical Engineering from Indian Institute of Technology, Madras, India in 2002. Before joining Department of Biosystems, KAIST as an internship trainee, he served Research and Development section of Jindal Vijayanagar Steel Ltd, India as Jr. Manager. Currently, he is pursuing M.S. program at Department of Biosystems, KAIST and his current research interests include signal processing, information theoretic approaches to blind source separation, and source localization.

**Sang-Hoon Oh** received his B.S. and M.S. degrees in Electronics Engineering from Pusan National University in 1986 and 1988, respectively. He received his Ph.D. degree in Electrical Engineering from Korea Advanced Institute of Science and Technology in 1999. From 1988 to 1989, he worked for the LG semiconductor, Ltd., Korea. From 1990 to 1998, he was a senior researcher in Electronics and Telecommunications Research Institute (ETRI), Korea. From 1999 to 2000, he was with Brain Science Research Center, KAIST. In 2000, he was with Brain Science Institute, RIKEN in Japan. In 2001, he was an R&D manager of Extell Technology Corporation, Korea. Since 2002, he has been with the Department of Information Communication Engineering, Mokwon University, Daejon, Korea. His research interests are supervised/unsupervised learning for intelligent information processing, speech processing, and pattern recognition.

**Soo-Young Lee** received B.S., M.S., and Ph.D. degrees from Seoul National University in 1975, Korea Advanced Institute of Science in 1977, and Polytechnic Institute of New York in 1984, respectively. From 1977 to 1980 he worked for the Taihan Engineering Co., Seoul, Korea. From 1982 to 1985 he also worked for General Physics Corporation at Columbia, MD, USA. In early 1986 he joined the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, as an Assistant Professor and now is a Full Professor at the Department of BioSystems and also department of Electrical Engineering and Computer Science. In 1997 he established Brain Science Research Center, which is the main research organization for the Korean Brain Neuroinformatics Research Program. The research program is one of the Korean Brain Research Promotion Initiatives sponsored by Korean Ministry of Science and Technology from 1998 to 2008, and currently about 35 Ph.D. researchers have joined the research program from many Korean universities. He is a Past-President of Asia-Pacific Neural Network Assembly, and has contributed to International Conference on Neural Information Processing as Conference Chair (2000), Conference Vice Co-Chair (2003), and Program Co-Chair (1994, 2002). Dr. Lee is the Editor-in-Chief of the newly established online/offline journal with double-blind review process, Neural Information processing—Letters and Reviews, and is on Editorial Board for 2 international journals, i.e., Neural Processing Letters and Neurocomputing. He received Leadership Award and Presidential Award from International Neural Network Society in 1994 and 2001, respectively, and APPNA Service Award in 2004. His research interests have resided in artificial brain, the human-like intelligent systems based on biological information processing mechanism in our brain. He has worked on the auditory models from the cochlea to the auditory cortex for noisy speech processing, information-theoretic binaural processing models for sound localization and speech enhancement, the unsupervised pro-active developmental models of human knowledge with multi-modal man–machine interactions, and the top–down selective attention models for superimposed pattern recognitions. His research scope covers the mathematical models, neuromorphic chips, and real-world applications. Also, he had recently extended his research into brain-computer interfaces with simultaneous fMRI and EEG measurements.