

RESEARCH

Open Access

Independent vector analysis using subband and subspace nonlinearity

Yueyue Na^{1*}, Jian Yu¹ and Bianfang Chai^{1,2}

Abstract

Independent vector analysis (IVA) is a recently proposed technique, an application of which is to solve the frequency domain blind source separation problem. Compared with the traditional complex-valued independent component analysis plus permutation correction approach, the largest advantage of IVA is that the permutation problem is directly addressed by IVA rather than resorting to the use of an ad hoc permutation resolving algorithm after a separation of the sources in multiple frequency bands. In this article, two updates for IVA are presented. First, a novel subband construction method is introduced, IVA will be conducted in subbands from high frequency to low frequency rather than in the full frequency band, the fact that the inter-frequency dependencies in subbands are stronger allows a more efficient approach to the permutation problem. Second, to improve robustness and against noise, the IVA nonlinearity is calculated only in the signal subspace, which is defined by the eigenvector associated with the largest eigenvalue of the signal correlation matrix. Different experiments were carried out on a software suite developed by us, and dramatic performance improvements were observed using the proposed methods. Lastly, as an example of real-world application, IVA with the proposed updates was used to separate vibration components from high-speed train noise data.

Keywords: Blind source separation, Independent vector analysis, Subband, Subspace, Nonlinearity

1. Introduction

Blind source separation (BSS) aims at recovering individual source signals from their mixed observations, the word “blind” means that neither the sources nor the mixing environment is known [1]. The applications of BSS techniques include speech enhancement, robust speech recognition, analyzing EEG or fMRI signals, feature extraction, image denoising, etc. [1-3]

Independent component analysis (ICA) [1] is a standard BSS method, which works under the assumption that sources are mutually independent, and the mixing procedure is linear and instantaneous. However, for speech and audio separation problems in real-world acoustic environment, such as the “cocktail party problem”, signals are often mixed in a convolutive manner. One common way to extend the instantaneous ICA to the convolutive model is the so-called frequency domain blind source separation (FDBSS) approach [2-4]. In

FDBSS, observed signals are transformed to time-frequency (T-F) domain via short time Fourier transform (STFT) first, so that the convolutive mixture in the time domain becomes instantaneous mixture in individual frequency bins. Then, complex-valued ICA algorithms [5,6] are used to separate data in each frequency bin independently. Although FDBSS has many advantages, it suffers from the well-known “permutation problem” [7-12]: the separated data must be aligned to make sure that each output signal only contains the data from the same source. After the permutation problem is solved, the inverse STFT is used to reconstruct the sources in the time domain. Many algorithms have been proposed in order to overcome the permutation ambiguity in FDBSS. For speech and audio signals, neighboring frequency bins from the same source are strongly correlated, and algorithms like [8-10] utilize this feature to correct the permutation problem. In many cases, different signals are likely to come from different directions, so direction of arrival patterns, which are hidden in the demixing matrices, can also be used to solve the permutation problem [10-12].

* Correspondence: 08112069@bjtu.edu.cn

¹Department of Computer Science, Beijing Jiaotong University, No. 3 Shang Yuan Cun, Hai Dian District, Beijing, China

Full list of author information is available at the end of the article

In recent years, independent vector analysis (IVA) was developed as an extension of ICA from univariate to multivariate components [13-18], and sources in the IVA model are considered as vectors instead of scalars. When IVA is used to perform source separation in the frequency domain, sources in different frequency bins are optimized together as vectors. IVA utilizes not only the statistical independency among different sources, but also the statistical inner dependency of each source vector in the optimization procedure [16]. Compared with the traditional ICA plus permutation correction approach, the largest advantage of IVA is that the permutation problem is automatically avoided, so there is no need for a postprocessing step after ICA to align sources. Moreover, as pointed out in [17], since the inter-frequency bin dependencies are considered in IVA, the separation results are expected to be better than the results of ICA algorithms which are applied on individual frequency bins alone.

The basic IVA framework was originally introduced in [13,14], this method used Kullback-Leibler (KL) divergence, which is completely equivalent to mutual information of the sources, as an objective function, with a natural gradient-based updating rule to optimize the demixing matrices frequency bin-by-frequency bin. To speed up the optimization procedure, in [15,16] a fast fixed-point IVA (FIVA) algorithm was proposed, which is based on the complex-valued FastICA algorithm [6]. In [19], an online IVA algorithm was proposed for real-time audio separation task, and a two-channel hardware demonstration system was developed. Another IVA algorithm was proposed in [20], as an extension of independent factor analysis [21] to the multivariate case. An expectation-maximization algorithm was used in the optimization procedure, and a Gaussian mixture model (GMM) was used to fit the source prior. Mixing models of noiseless, noisy, and online cases can conveniently be integrated in this framework. The IVA model was also used to solve the joint blind source separation problem for fMRI data in [17], where both gradient based and Newton updating rules are derived. There is a number of other recently proposed IVA algorithms such as IVA incorporating video information [22], non-circular IVA [23], chain clique IVA [24,25], etc.

In this article, we present two improvements for IVA. First, a new subband construction technique is introduced to enhance the inter-frequency bin dependency. Second, in each subband, the nonlinear mapping is calculated in the one-dimensional subspace of the estimated source data to further improve the separation performance. These proposals are integrated in a software suite for BSS research and applied usage, which is publicly available [26].

The remainder of this article is organized as follows: The general framework of the IVA algorithm is briefly introduced in Section 2. Then, in Section 3 we present

the subband policy, and the nonlinear mapping improvement is depicted in Section 4. Computational complexities of the proposed methods are analyzed in Section 5, and some methods are also introduced to reduce the complexity. Different experiments were carried out to show the performance improvements of the proposed methods, the experimental results are reported in Section 6. In addition to the simulated experiments, a real-world application on high-speed train noise component separation is described in Section 7 to demonstrate the usefulness of the proposed methods. At last, we conclude this article in Section 8.

The frequently used notations in this article are listed below for easy reference.

1. Italic lowercase letters denote scalars, boldface italic lowercase letters denote column vectors, and boldface italic uppercase letters denote matrices, e.g., a , \mathbf{a} , and \mathbf{A} .
2. Superscripts * , T , and H denote complex conjugate, matrix and vector transpose, and conjugate transpose, respectively, e.g., $\mathbf{A}^H = (\mathbf{A}^T)^* = (\mathbf{A}^*)^T$.
3. Commas separate values within rows, e.g., $\mathbf{a} = [a_1, a_2]^T$, while semicolons separate rows, e.g., $\mathbf{a} = [a_1; a_2]$, $[\mathbf{a}; \mathbf{b}] = [\mathbf{a}^T, \mathbf{b}^T]^T$.
4. The original source signal, the observed signal, and the separated signal are denoted by the letters s , x , and y , respectively.
5. Mixing matrices and demixing matrices are denoted by the letters \mathbf{A} and \mathbf{W} .
6. Indices m , n , and f denote sensor index, source index, and frequency bin index, respectively. There are M sensors, N sources, and F frequency bins in the IVA model. Indices t and τ represent time domain sample index, and STFT frame index, respectively. Variables with index t indicate time domain data, while variables with indices f and/or τ indicate frequency domain data.

2. Independent vector analysis

2.1. From FDBSS to IVA

In real-world acoustic environment, signals are mixed with each other, as well as their delays, attenuations, and reverberations, i.e., signals are convolutively mixed together. Supposing there are N sources and M sensors ($M \geq N$), the signal captured by sensor m can be modeled as (1) [14], where \star is the convolution operation, $a_{mn}(t)$ is the finite duration impulse response mixing filter from source n to sensor m .

$$x_m(t) = \sum_{n=1}^N a_{mn}(t) \star s_n(t) \quad (1)$$

When STFT is used, if the STFT frame length is sufficiently longer than the mixing filter length [14], the time

domain convolution in (1) can approximately be converted to the frequency domain multiplication in (2), where $s_n^{[f]}(\tau)$, $x_m^{[f]}(\tau)$, and $a_{mn}^{[f]}$ are frequency domain versions of $s_n(t)$, $x_m(t)$, and $a_{mn}(t)$, respectively. For all sources $s^{[f]}(\tau) = [s_1^{[f]}(\tau), \dots, s_N^{[f]}(\tau)]^T$ and sensors $x^{[f]}(\tau) = [x_1^{[f]}(\tau), \dots, x_M^{[f]}(\tau)]^T$, the complete mixing process can be formulated as (3), where $A^{[f]}$ is the mixing matrix for frequency bin f , with $a_{mn}^{[f]}$ as its entries.

$$x_m^{[f]}(\tau) = \sum_{n=1}^N a_{mn}^{[f]} s_n^{[f]}(\tau) \quad (2)$$

$$\mathbf{x}^{[f]}(\tau) = \mathbf{A}^{[f]} \mathbf{s}^{[f]}(\tau) \quad (3)$$

$$\mathbf{y}^{[f]}(\tau) = \mathbf{W}^{[f]} \mathbf{x}^{[f]}(\tau) \quad (4)$$

Since signals are instantaneously mixed in each frequency bin, complex-valued ICA algorithms like [5,6] can be used to separate signals, as depicted in (4), where $\mathbf{W}^{[f]}$ is the demixing matrix for frequency bin f , which is estimated by ICA. FDBSS utilizes (4) to separate signals, an example of 2×2 FDBSS demixing model is shown in Figure 1a. In this example, each horizontal layer is an ICA demixing model in (4) for each frequency bin, and the demixing procedure is carried out in layers independently. Since ICA in different layers may output the separated results in different order, the permutation ambiguity will occur in FDBSS, which is indicated by the different color of $\mathbf{y}^{[f]}$ in Figure 1a. The permutation ambiguity must be carefully addressed by algorithms like [7-12] before the inverse STFT is performed, or else the separation procedure will fail.

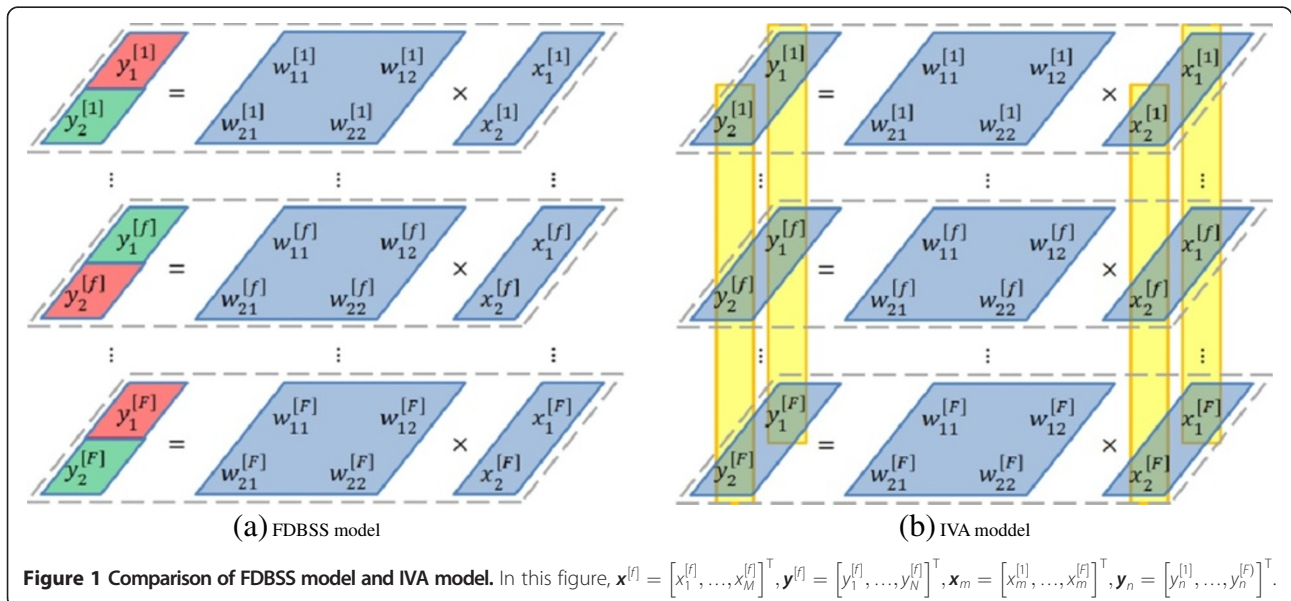
In addition to separate sources in each frequency bin, IVA utilizes inter-frequency bin information to solve the permutation problem in the separation procedure. The IVA model is very similar with the FDBSS model, as shown in Figure 1b. Their difference is that signals are considered as vectors in IVA, i.e. $\mathbf{x}_m = [x_m^{[1]}, \dots, x_m^{[F]}]^T$, $\mathbf{y}_n = [y_n^{[1]}, \dots, y_n^{[F]}]^T$ (vertical bars in Figure 1b), and they will be optimized as multivariate variables, instead of independent scalars like in ICA. The IVA model can also be formulated in a single equation: After data in each layer are concatenated into vectors as: $\mathbf{x} = [\mathbf{x}^{[1]}; \dots; \mathbf{x}^{[F]}]$, $\mathbf{y} = [\mathbf{y}^{[1]}; \dots; \mathbf{y}^{[F]}]$, and \mathbf{W} is a block diagonal matrix with each $\mathbf{W}^{[f]}$ in its diagonal, the demixing procedure can be denoted as: $\mathbf{y} = \mathbf{W}\mathbf{x}$, just as the same expression as ICA.

2.2. IVA objective function

Mutual information $I(\cdot)$ is a natural measure of independence, which is minimized to zero when random variables are mutually independent, and it is often employed as the objective function in ICA. Mutual information can be calculated in the form of KL divergence $\text{KL}(\cdot \| \cdot)$ in (5), where p_y denotes the probability density function (PDF) of a random vector \mathbf{y} , p_{y_n} denotes the n th marginal PDF of \mathbf{y} , and \mathbf{z} is a dummy variable for the integral [16].

$$I(\mathbf{y}) = \text{KL}\left(p_y \| \prod_n p_{y_n}\right) = \int p_y(\mathbf{z}) \log \frac{p_y(\mathbf{z})}{\prod_n p_{y_n}(z_n)} d\mathbf{z} \quad (5)$$

IVA objective function has the similar form as (5); however, each \mathbf{y}_n in IVA is a vector rather than a scalar. The IVA objective function and the corresponding derivations are given in (6) [16,17], where $H(\cdot)$ represents the entropy.



$$\begin{aligned}
 \mathcal{J}_{\text{IVA}} &= \text{KL}(p_{\mathbf{y}} \| \Pi_n p_{\mathbf{y}_n}) \\
 &= \sum_{n=1}^N H(\mathbf{y}_n) - H(\mathbf{y}_1; \dots; \mathbf{y}_N) \\
 &= \sum_{n=1}^N H(\mathbf{y}_n) - H(\mathbf{y}^{[1]}; \dots; \mathbf{y}^{[F]}) \\
 &= \sum_{n=1}^N H(\mathbf{y}_n) - H(\mathbf{W}^{[1]} \mathbf{x}^{[1]}; \dots; \mathbf{W}^{[F]} \mathbf{x}^{[F]}) \quad (6) \\
 &= \sum_{n=1}^N H(\mathbf{y}_n) - H(\mathbf{W} \mathbf{x}) \\
 &= \sum_{n=1}^N H(\mathbf{y}_n) - \sum_{f=1}^F \log |\det(\mathbf{W}^{[f]})| - C
 \end{aligned}$$

In formula (6), the last equation is derived since $H(\mathbf{W}\mathbf{x}) = \log|\det(\mathbf{W})| + H(\mathbf{x})$ holds for a linear invertible transformation \mathbf{W} , and the determinant of the block diagonal matrix $\det(\mathbf{W}) = \prod_{f=1}^F \det(\mathbf{W}^{[f]})$. The term $C = H(\mathbf{x})$ is a constant because the observed signals will not change in the optimization procedure [16,17].

When the observed signals in each frequency bin are centered and whitened ($\mathbf{x} \leftarrow \mathbf{x} - E(\mathbf{x})$ so that $E(\mathbf{x}) = 0$, then $\mathbf{x} \leftarrow \mathbf{V}\mathbf{x}$ so that $E(\mathbf{x}\mathbf{x}^H) = \mathbf{I}$, $E(\cdot)$ for expectation, \mathbf{V} is the whitening matrix), the demixing matrices $\mathbf{W}^{[f]}$ become orthonormal, so the term $\sum_{f=1}^F \log |\det(\mathbf{W}^{[f]})|$ becomes zero. Then, by noting that $H(\mathbf{y}_n) = \sum_f H(y_n^{[f]}) - I(\mathbf{y}_n)$, minimizing the IVA objective function in (6) is equivalent to minimizing (7) [17].

$$\mathcal{J}_{\text{IVA}} = \sum_n \left(\sum_f H(y_n^{[f]}) - I(\mathbf{y}_n) \right) \quad (7)$$

From here we can see that minimization of (7) balances the minimization of the $H(y_n^{[f]})$ term and the maximization of the $I(\mathbf{y}_n)$ term. According to the basic ICA theory, independency is measured by non-Gaussianity, and minimizing $H(y_n^{[f]})$ is equivalent to maximizing the non-Gaussianity, which is responsible for separating data in individual frequency bins. Meanwhile, maximizing $I(\mathbf{y}_n)$ means enhancing the dependency of entries in \mathbf{y}_n , which is responsible for solving the permutation problem. In short, minimizing the IVA objective function can simultaneously separate the data and solve the permutation problem [17].

2.3. Optimization procedures

To minimize the objective function in (6), the entropy of the estimated source vectors must be calculated. Although the actual PDF of each \mathbf{y}_n is unknown, a prior target PDF $\hat{p}(\mathbf{y}_n)$ is often used, so the objective function in (6) can be simplified as in (8) [14].

$$\mathcal{J}_{\text{IVA}} = -\sum_n E(\log \hat{p}(\mathbf{y}_n)) \quad (8)$$

Natural gradient descent and fast fixed-point iteration are two frequently used optimization methods in IVA. In

the natural gradient-based approach [13,14], after differentiating the objective function with respect to the demixing matrices, the updating rule can be formulated as (9)

$$\mathbf{W}^{[f]} \leftarrow \mathbf{W}^{[f]} + \eta \left\{ \mathbf{I} - E \left[\phi^{[f]}(\mathbf{y}^{[f]}) (\mathbf{y}^{[f]})^H \right] \right\} \mathbf{W}^{[f]} \quad (9)$$

In this equation, η is the learning rate, and $\phi^{[f]}(\cdot)$ is a multivariate nonlinear function (also called score function) for frequency bin f . This nonlinear function is highly related to the chosen source prior PDF:

$$\phi^{[f]}(\mathbf{y}_n) = -\frac{\partial \log \hat{p}(\mathbf{y}_n^{[1]}, \dots, \mathbf{y}_n^{[F]})}{\partial \mathbf{y}_n^{[f]}} \quad (10)$$

In [15,16], a FIVA algorithm was proposed. Compared with the natural gradient-based approach, the convergence speed of FIVA is dramatically improved, and there is no need to choose the learning rate manually. After applying a nonlinear mapping G , the FIVA objective function can be transformed from (8) to (11) [15,16]. The corresponding updating rule can be formulated in (12), followed by the symmetric decorrelation scheme in (13). In (12), $(\mathbf{w}_n^{[f]})^H$ represents the n th row of the demixing matrix $\mathbf{W}^{[f]}$. In (13), the inverse square root of a symmetric matrix $\mathbf{W}^{-1/2} = \mathbf{P}\mathbf{D}^{-1/2}\mathbf{P}^H$, and $\mathbf{W} = \mathbf{P}\mathbf{D}\mathbf{P}^H$ is the eigendecomposition of \mathbf{W} .

$$\mathcal{J}_{\text{FIVA}} = \sum_{n=1}^N E \left[G(|\mathbf{y}_n|^2) \right] = \sum_{n=1}^N E \left[G \left(\sum_{f=1}^F |y_n^{[f]}|^2 \right) \right] \quad (11)$$

$$\mathbf{w}_n^{[f]} \leftarrow E \left[G'(|\mathbf{y}_n|^2) + |y_n^{[f]}|^2 G''(|\mathbf{y}_n|^2) \right] \mathbf{w}_n^{[f]} - E \left[(y_n^{[f]})^* G'(|\mathbf{y}_n|^2) \mathbf{x}^{[f]} \right] \quad (12)$$

$$\mathbf{W}^{[f]} \leftarrow \left(\mathbf{W}^{[f]} (\mathbf{W}^{[f]})^H \right)^{-1/2} \mathbf{W}^{[f]} \quad (13)$$

Although the original nonlinearity G used in (11) is also derived from the source prior PDF as: $G(|\mathbf{y}_n|^2) = -\log \hat{p}(\mathbf{y}_n)$ [15,16], nonlinearities in FIVA should be considered as entropy estimators, so, different nonlinearities can also be used, which may not have a direct association with source prior PDF. For example, $G(\cdot) = \sqrt{\cdot}$ and $G(\cdot) = \log(\cdot)$ are two frequently used nonlinear functions.

When the IVA updating rules in (9) and (12) are compared with the corresponding updating rules in conventional InfomaxICA [5] and complex-valued FastICA [6], one can find that they have nearly the same expressions, the only difference is the improvement from univariate nonlinearities to multivariate nonlinearities. It means

that multivariate nonlinearity is very important for IVA algorithms, choosing proper nonlinearities will improve the source separation performance.

3. Subband IVA

3.1. Clique-based approach

To illustrate the idea of the subband approach, the spectrogram of an estimated source signal is shown in Figure 2a as an example, and its corresponding correlation matrix is visualized in Figure 2b. The correlation matrix Σ_n of the estimated source y_n is calculated upon the amplitude of its T-F data by (14) [27]. According to the demixing model in (4), since the observed data $x^{[f]}$ are centered and whitened, and the demixing matrix $W^{[f]}$ is orthonormal because of (13), data in Figure 2a are kept zero mean and unit variance in each frequency bin (the scaling ambiguity [2] is still unsolved), and the correlation matrix in Figure 2b is symmetric and positive semidefinite, with all entries belong to [0,1] and with all ones in its diagonal.

$$\Sigma_n = E \left\{ \left[|y_n^{[1]}|, \dots, |y_n^{[F]}| \right]^T \left[|y_n^{[1]}|, \dots, |y_n^{[F]}| \right] \right\} \quad (14)$$

From Figure 2 we can see that, when time domain signals are transformed to T-F domain by STFT, neighboring frequency bins usually have strong correlation; however, two frequency bins may be weakly correlated if they are far apart from each other [10]. Original IVA algorithms [13-16] treat all frequency bins as a whole, and although this approach is easy to implement and its computational time is relatively short, weakly correlated frequency bins will degrade the separation performance.

Some policies have been proposed to compensate the weakly correlated frequency bins in IVA. For example, a chain clique model was introduced in [24], and its update in [25] with variable clique size. Neighboring frequency bins are treated as a “clique”, i.e., a fully connected subgraph, to increase the inter-frequency bin dependency, and consecutive cliques are chained together by proper overlapping, as depicted in Figure 3a. For speech and music signals, harmonic structures usually exist in the signal T-F data (can also be observed in Figure 2), and neighboring frequencies may still be weakly correlated because of the harmonic structure. In order to choose strongly correlated frequency bins for IVA, a novel clique construction method was proposed in [29] to utilize the harmonic-dependent property of speech and music signals. The frequently used source prior PDF of the preceding clique-based approaches can be summarized in (15), where C is the number of cliques, C_k means all frequency bins in clique k . The corresponding nonlinear mapping can be derived according to (10).

$$p(y_n) \propto \exp \left(-\sum_{k=1}^C \sqrt{\sum_{f \in C_k} |y_n^{[f]}|^2} \right) \quad (15)$$

3.2. Subband-based approach

In addition to the local dependency property of the T-F data, from Figure 2 we can also conclude another property that the high-frequency part of the speech T-F data usually has stronger correlation than the low-frequency part. As visualized in Figure 2a, the harmonic structure of the speech data is weak in the high-frequency part, so, the corresponding frequency bins are highly correlated in a relatively large neighborhood, as visualized in

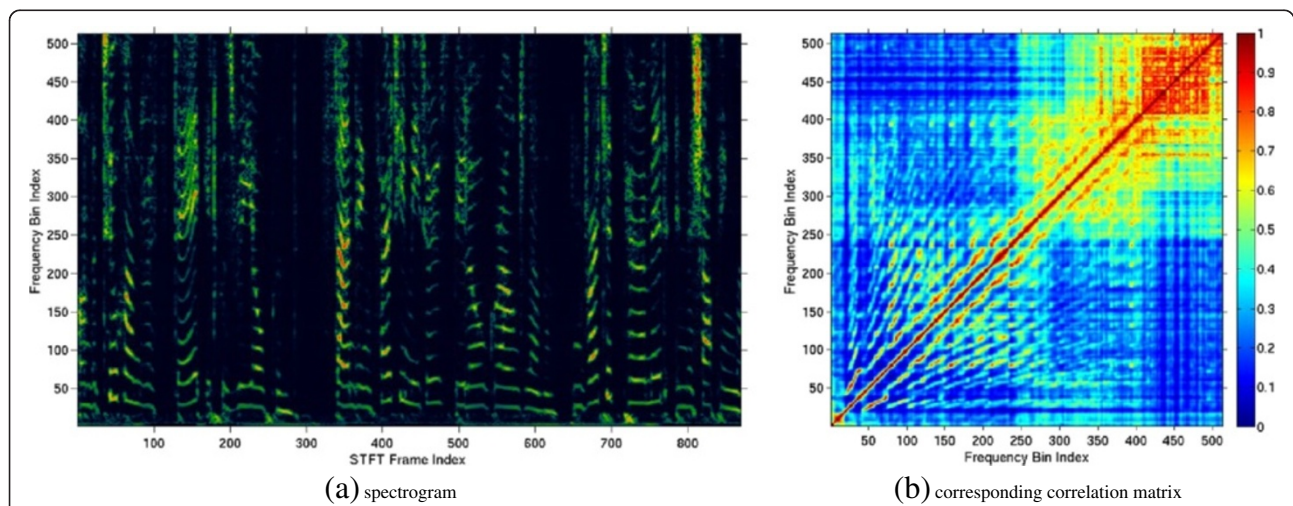


Figure 2 Separated signal's spectrogram and corresponding correlation matrix. In this example, dataset from [28] was used, and the proposed subband subspace IVA algorithm was used for separation, STFT frame size, STFT frame overlap, and FFT size were set to 512, 7/8, and 1024. In Figure 2a, brighter color means larger magnitude. In Figure 2b, the correlation matrix is reversed upside down for visualization.

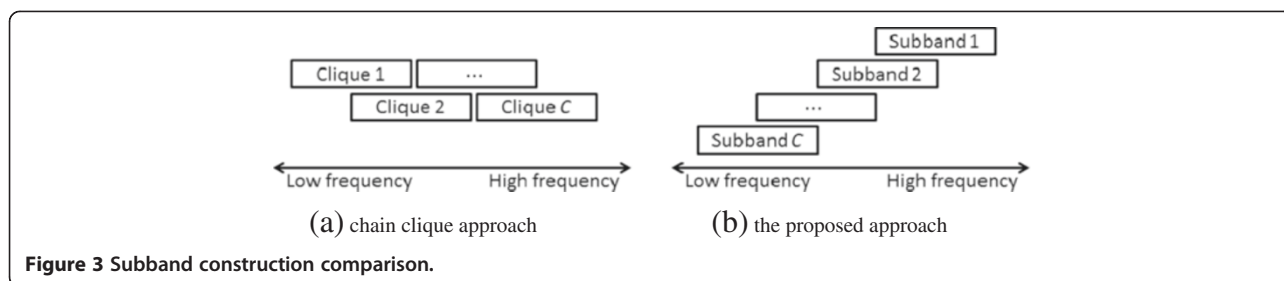


Figure 3 Subband construction comparison.

Figure 2b. When the frequency bin index goes down, the harmonic structure becomes clearer, and the neighborhood of high correlation is decreased. According to these two properties, our subbands are constructed as depicted in Figure 3b. The difference between the clique-based approach and the proposed approach is that although the full frequency band is divided into cliques in the former approach, the updating procedure is still performed in the full frequency band, so the terminology “clique” is used. On the other hand, in the proposed approach, IVA is carried out in individual subbands from high frequency to low frequency, the underlying nonlinearity is only a function of the multivariate data within subbands, and the updating procedure is also performed in current subband, so, the terminology “subband” is used in the proposed approach.

In each subband (except the first one), the input of the IVA algorithm can be considered as partially separated data, as depicted in Figure 4. Since two consecutive subbands are overlapped with each other, data in the overlapped part have already been handled by the IVA performed in previous subband. Usually, most frequency bins in the overlapped part will converge, only left a few (possibly none) unseparated frequency bins. So, IVA in current subband only needs to handle the unconverged frequency bins, which are coming from the newly added data, and inherited from the previous subband. Sometimes, permutation errors may occur in the overlapped part of current subband, which are caused by previous IVA. However, since current IVA utilizes the information

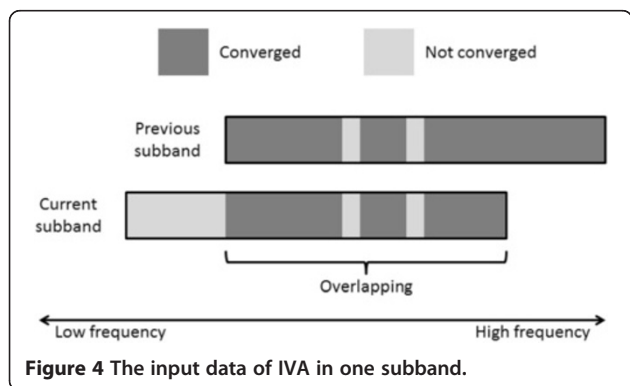


Figure 4 The input data of IVA in one subband.

of the entire subband to perform the separation, the algorithm is robust against permutation errors if the number of the misaligned frequency bins is not too large, i.e., old permutation errors will not easily bring new errors and propagate to lower subbands.

The advantages of this subband-based approach are: first, the separated data in each subband can be used as a kind of heuristic information by IVA, which makes the unseparated data converge towards the separated part, and this is useful for solving the permutation problem. Second, since subbands with stronger correlations can be handled by IVA easier, the separation is performed from high frequency to low frequency, which is according to the difficulty increasing order. Plenty of heuristic information will simplify the separation in low-frequency subbands. Third, only a few unseparated frequency bins will participate in the IVA iteration, so, subbands overlapping will not increase the complexity too much. As the conclusion to this section, the pseudo code of the complete subband IVA algorithm is given as Algorithm 1. In Algorithm 1, a subband is converged if all of its frequency bins are converged, and the convergence criteria for a frequency bin is measured by the average cosine of the corresponding rows of the demixing matrices in two consecutive iterations $\mathbf{W}_0^{[l]}$ and $\mathbf{W}_0^{[l+1]}$, which can be compared according to (16), where $\text{tr}(\cdot)$ calculates the trace of a matrix, $\text{abs}(\cdot)$ is entry to entry absolute value of a matrix.

$$|1 - \text{tr} \left\{ \text{abs} \left[\mathbf{W}^{[l]} \left(\mathbf{W}_0^{[l]} \right)^H \right] \right\} / N| \leq \varepsilon \quad (16)$$

In our experiments, the maximum iteration times was set to 500, and the convergence threshold ε was set to 1×10^{-10} .

4. Subspace nonlinearity

4.1. IVA nonlinearity

The usage of the multivariate nonlinearities is the key reason why IVA algorithms can avoid the permutation problem. As the nonlinear function is closely related to the source prior PDF, different PDFs are designed. The multivariate spherically symmetric Laplace (SSL) distribution proposed in [14] is probably the most widely

Algorithm 1 The subband IVA algorithm

Input: centered and whitened sensor data in time-frequency domain

Output: $\mathbf{W}^{[1]}, \dots, \mathbf{W}^{[F]}$; demixing matrices

1. Initialize the convergence tag $converged[f]=0$ for every frequency bin.
2. Construct subbands according to Fig 3b.
3. Call the subroutine “subband_iva(subband)” on each subband in order.
4. Return $\mathbf{W}^{[1]}, \dots, \mathbf{W}^{[F]}$.

```

subband_iva(subband):
while(true) {
    if (all frequency bins in the subband are converged or exceeds the maximum number of iterations) break;
    for ( $f \in$  current subband) {
        if ( $converged[f]=1$ ) continue;
        update  $\mathbf{W}^{[f]}$  according to (9) or (12);
        if ( $\mathbf{W}^{[f]}$  is converged)  $converged[f]=1$ ;
    }
    increase iteration count by 1;
}
    
```

used source prior PDF in IVA. The formulation of SSL is given in (17), this PDF is based on the observation that speech and audio signals in frequency domain usually exhibit super Gaussian distribution, and their real and imaginary parts are uncorrelated [14]. In SSL, although different dimensions are uncorrelated, they are not independent, IVA uses this property to correct the permutation problem.

$$p(\mathbf{y}) \propto \exp\left(-\sqrt{\sum_f |y^{[f]}|^2}\right) \quad (17)$$

In source separation problems, different sources may have different distributions. In [20], a mixture of Gaussians is used to model the source prior, as shown in (18), where k is the index of Gaussians, π_k is the mixture weight of the k th Gaussian, σ_{kf} is the standard deviation for Gaussian k frequency bin f . Since the GMM can approximate almost any distribution with appropriate parameters [30], this approach is more flexible than (17). However, as the tradeoff, the GMM must be trained before or during the separation procedure.

$$p(\mathbf{y}) \propto \sum_k \pi_k \prod_f \frac{1}{\sqrt{2\pi}\sigma_{kf}} \exp\left[-\frac{(y^{[f]})^2}{2\sigma_{kf}^2}\right] \quad (18)$$

In [17], the multivariate Gaussian distribution in (19) is used as the source prior for fMRI data, where Σ is the correlation matrix of the estimated source. Unlike the first two PDFs which are spherically symmetric, second-order correlations are considered and modeled in this

approach. In [17], Σ is learned from the estimated source data in every iteration round.

$$p(\mathbf{y}) \propto \exp\left(-\frac{1}{2}\mathbf{y}^T \Sigma^{-1} \mathbf{y}\right) \quad (19)$$

Given a source prior PDF, the nonlinearity for natural gradient IVA can be derived as (10), while the nonlinearity for FIVA can be derived as: $G(|\mathbf{y}_n|^2) = -\log p(\mathbf{y}_n)$, however, different nonlinear functions G can also be used. In this section, we mainly consider the FIVA algorithm, as it is easy to extend the complex-valued FastICA to FIVA [15,16], and the convergence speed of FIVA is high. Moreover, nonlinear functions G in FIVA are real valued univariate functions, which are easy to differentiate.

4.2. The proposed approach

After the nonlinear function G is chosen for the FIVA algorithm, its first- and second-order derivatives can be derived, and the algorithm can be updated according to (12) and (13). For example, the nonlinear function $G(\cdot) = \log(\cdot)$ is used in all our experiments in Sections 6 and 7, then, $G'(|\mathbf{y}_n|^2) = 1/|\mathbf{y}_n|^2$, and $G''(|\mathbf{y}_n|^2) = -1/(|\mathbf{y}_n|^2)^2$ can be substituted in (12) for the original FIVA iteration.

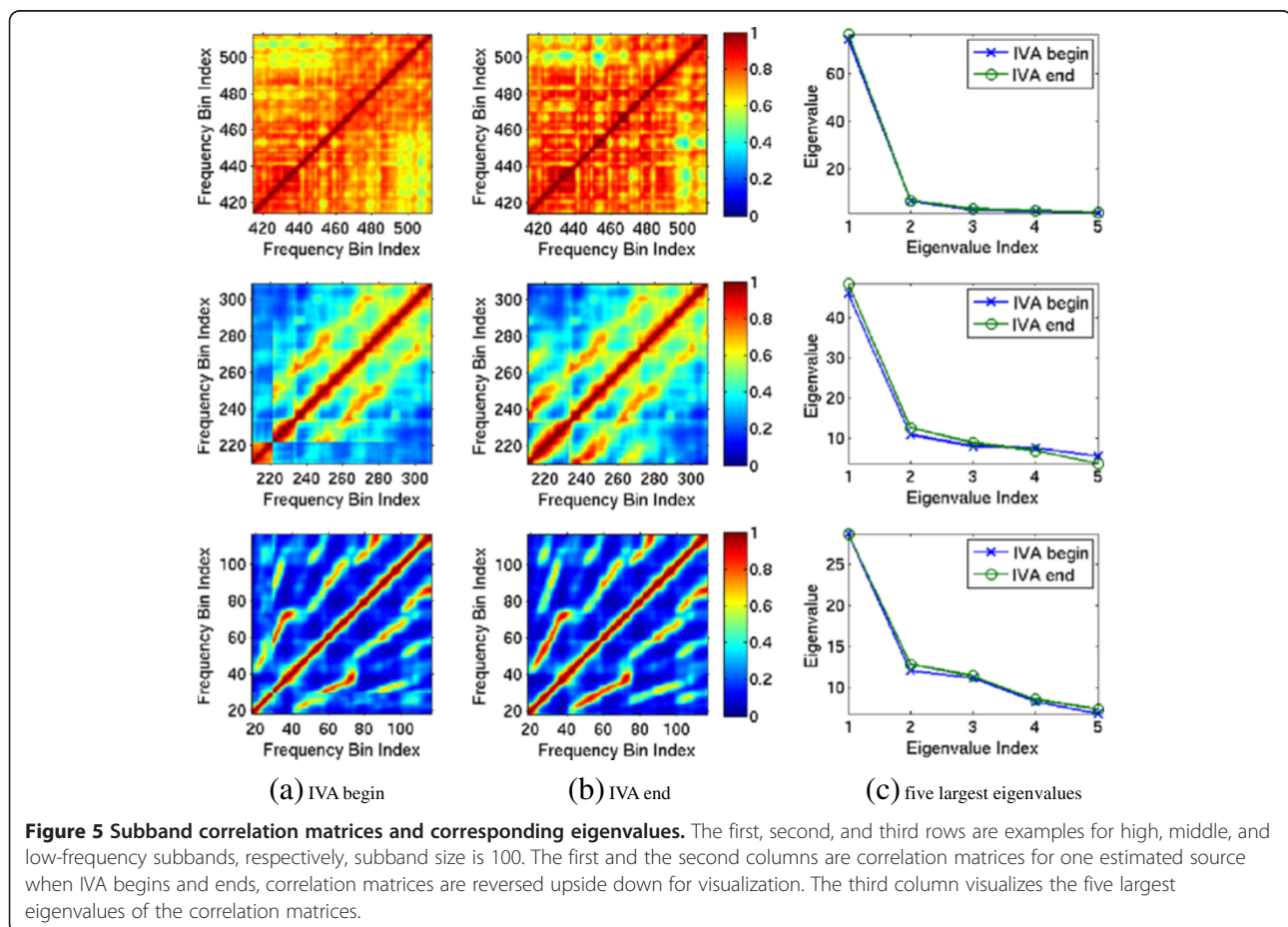
In (12), the nonlinearity is calculated from the squared norm of the estimated source data as: $|\mathbf{y}_n|^2 = \sum_{f=1}^F |y_n^{[f]}|^2$ in every iteration step, where $|y_n^{[f]}|$ is also called the signal envelope for frequency bin f . It is well known that signal envelopes of the same source are highly correlated

in neighboring frequency bins, and this feature is often used to solve the permutation problem [8-10]. When the subband technique proposed in the previous section is used, stronger correlations are expected to be observed because of the local dependency property of the data. Figure 5 is an example of the estimated source correlation matrices Σ_n in different subbands, which are calculated according to (14). We still use Σ_n to denote the correlation matrix of source n , and $y_n = [y_n^{[1]}, \dots, y_n^{[F]}]^T$ for the corresponding data in the current subband without any ambiguity since all nonlinearities are calculated within subbands. From Figure 5, we can see that frequency bins in subbands are strongly correlated with each other.

This strong correlation can be seen more clearly after performing eigendecomposition on Σ_n , as an example, the five largest eigenvalues of the correlation matrices in Figure 5a,b are shown in Figure 5c. An animation for the complete iteration procedure of the proposed subband subspace approach is also provided as Additional file 1 to illustrate the data property in subbands. Several conclusions can be made from Figure 5 and Additional file 1: First, we can see that the first

eigenvalue is much larger than the other eigenvalues. This phenomenon is due to the strong dependency of the T-F data, it implies that data in a subband are almost distributed in a one-dimensional subspace, which is spanned by the dominant eigenvector. Second, from high-frequency subband to low-frequency subband, the T-F data become sparser, and the dominant eigenvalue become smaller; however, the large eigengap between the first and the second eigenvalues still can be observed. Third, only small changes can be observed between the eigenvalues before and after IVA, since only a small part of the correlation matrix is updated because of subband overlapping. After IVA, the dominant eigenvalue in high-frequency subbands will increase a little, as the frequency bin correlation is enhanced, while in low-frequency subbands, the dominant eigenvalue will decrease a little, as the separated data become sparser.

Since data samples are mainly distributed in a one-dimensional subspace, it is better to incorporate this property in the nonlinearity calculation, and this yields the proposed method: instead of calculating the nonlinearity in the original space, we calculate the nonlinearity in the scaled dominant subspace of the input data. Supposing λ_{n1} and ν_{n1} are the dominant



eigenvalue and the corresponding eigenvector of Σ_n , then data are first projected and scaled into the dominant subspace as (20), and the updates are performed as (21). The projected data are also likely super Gaussian distributed in the subspace, many samples are projected around zero, so the scaling parameter α is introduced in (20) to stretch the data. In our experiments, subspaces spanned by more eigenvectors were also tried, and we found that performance improvement was not guaranteed when more eigenvectors were used, so only the dominant eigenvalue and eigenvector are used in the proposed approach to keep computational cost low. In addition, different values of α from 1 to very large scales were also tested, and we found that the separation performance was not sensitive to this parameter, so we fix $\alpha = \sqrt{\lambda_{n1}}$.

$$q = \alpha \mathbf{v}_{n1}^T \left[|y_n^{[1]}|, \dots, |y_n^{[F]}| \right]^T \quad (20)$$

$$\mathbf{w}_n^{[f]} \leftarrow E \left[G'(q^2) + |y_n^{[f]}|^2 G''(q^2) \right] \mathbf{w}_n^{[f]} - E \left[\left(y_n^{[f]} \right)^* G'(q^2) \mathbf{x}^{[f]} \right] \quad (21)$$

When the proposed nonlinearity is compared with the nonlinearities which are derived from (17)-(19), we can find that the proposed nonlinearity is based on the method in (17), as they can share the same nonlinear function G ; however, the proposed nonlinearity is calculated in the dominant subspace. This can be explained as a kind of denoising operation like principal component analysis [27]: since data in trivial component directions usually result from noise or inaccuracy, discarding these parts should improve data purity. Because different sources have different dominant subspaces, the proposed nonlinearities for different sources are also different, this property is the same as equations (18) and (19). At last, after each iteration round, the dominant eigenvectors

should be updated, as the estimated source data are refreshed, this is the same as the situation in (19).

5. Computational complexity analysis

Some improvements can be made to reduce the computational complexity of the proposed methods. In order to calculate the subspace nonlinearity, the correlation matrix Σ_n for each source must be calculated first, estimating large correlation matrices in every iteration step is computationally expensive. However, data in a subband are partially converged (except the first subband), the correlation between two converged frequency bins will not change any more. This means that only those entries corresponding to the unseparated frequency bins in the correlation matrices need to be updated. As the algorithm goes on, more and more frequency bins will converge, so fewer and fewer updates are required. When the algorithm moves from one subband to the next, the subcorrelation matrix of the overlapped part can directly be copied to the corresponding position of the new correlation matrix. In addition, only half of the entries need to be updated since correlation matrices are symmetric. The preceding procedures for correlation matrix calculation are depicted in Figure 6.

After the correlation matrix is constructed, performing full eigendecomposition is unnecessary since only the dominant eigenvalue and the corresponding eigenvector are needed in subspace nonlinearity calculation. Instead, the “power method” [31] can be used for this purpose, whose main operation is matrix-vector multiplication, with the time complexity of $O(F^2)$. In one subband, the order of correlation matrices F is not large, so this operation is efficient. Because of the symmetric and positive semidefinite property, the dominant eigenvalue and eigenvector of Σ_n are guaranteed to be found, moreover, the convergence speed of the power method is proportional to the gap between the dominant and the second

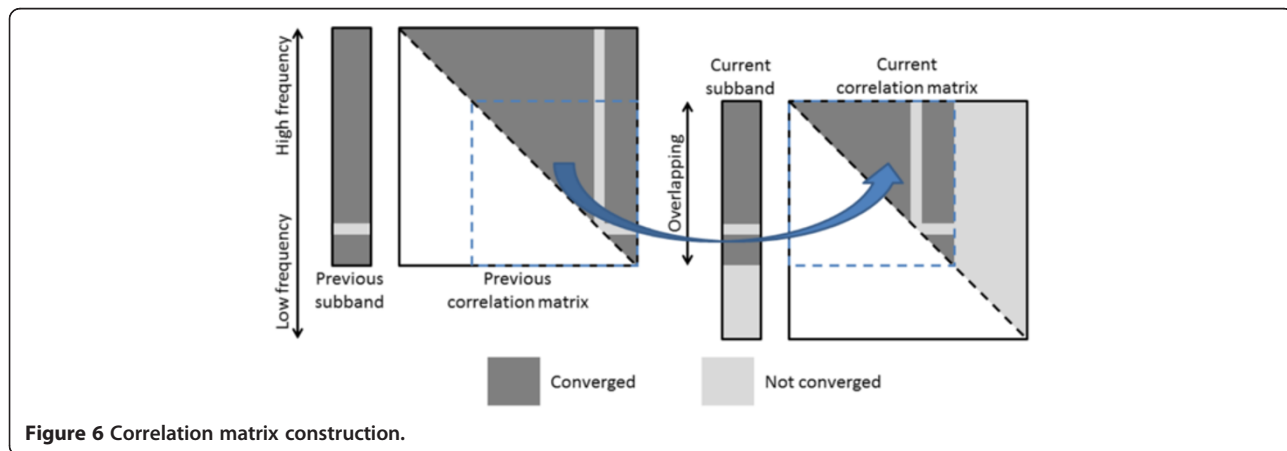


Figure 6 Correlation matrix construction.

eigenvalue [31], this gap typically is very large, as can be seen for example in Figure 5c, which makes the algorithm converge very fast.

As the number of sources, signal length, or FFT block size increase, more data are processed by the separation algorithm; therefore, the spatial complexity becomes a problem. Here, the spatial complexity means the amount of main memory space consumed by the algorithm, for FDBSS problems, the memory is mainly taken up by the observed and the estimated T-F data. Sometimes, it is suitable to store the unused T-F data into disk to decrease the spatial complexity, and load them into memory when they are requested by the separation algorithm. At the first glance, the traditional ICA plus permutation correction approach has the lowest spatial complexity, since only the amount of data for one frequency bin is needed at a time in ICA. However, the permutation problem still exists, and some permutation algorithms need to load all the separated data into memory to correct the permutation, so, the total spatial complexity is still high. For IVA algorithms like [14,15], all sensor data should be loaded since the inter-frequency bin information is required in the separation procedure. Although the clique-based approaches, like [24,25,29], calculate nonlinearities within cliques, the updating procedures are still performed in the full frequency band, so the spatial complexity remains the same. In the proposed subband and subspace approach, all calculations are performed in subbands, only the amount of data for one subband needs to be loaded into memory, so the spatial complexity is reduced.

6. Experiments

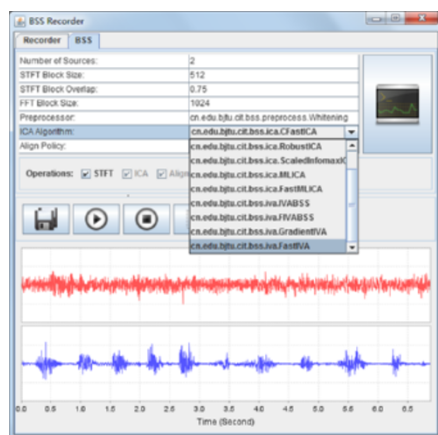
All experiments were carried out on a software suite for BSS research and application purpose (see Figure 7). This platform is developed in Java, some frequently used

source separation algorithms and permutation algorithms have already been implemented in the system, and modules for virtual mixing environment generation and performance evaluation are also provided. Moreover, well-designed interfaces enable new ICA, IVA, permutation algorithms, and also other new features to be conveniently integrated in the platform. The source code is available for public, please refer [26] for more information.

Different algorithms were compared in the following experiments, including the complex-valued FastICA algorithm [6] (ICA), natural gradient-based IVA [14] (IVA), fast fixed-point IVA [15] (FIVA), IVA with the chain clique approach [24] (IVA-C), with 100 clique size and 1/2 overlapping; the proposed subband approach (FIVA-S), with 100 subband size and 7/8 overlapping; the proposed subband with the subspace nonlinearity approach (FIVA-SS). Separation results were also postprocessed by the permutation algorithm proposed in [10] (with the -P suffix) for comparison. Separation performance was evaluated in terms of signal-to-interference ratio (SIR) improvement [32].

6.1. Instantaneous mixture

The first experiment tests the permutation overcoming ability of IVA algorithms, for simplicity, instantaneous mixtures are performed so that we can focus our attention on the permutation problem. The dataset from [28] was used, which includes two male and two female speeches for 7 s, the source sampling rate is 8000 Hz. In this experiment, 2×2 mixtures were performed, all $C_4^2 = 6$ different combinations of 4 sources were mixed by 25 randomly generated mixing matrices, i.e., 150 groups of mixed signals were tested for each algorithm. The STFT frame size was set to 512, with 3/4 overlapping, and the FFT block size was set to 1024. The mean value and the standard deviation of SIR improvements are shown in Figure 8. In this experiment, the nonlinearity



(a) GUI screenshot



(b) sound capture device

Figure 7 The BSS platform.

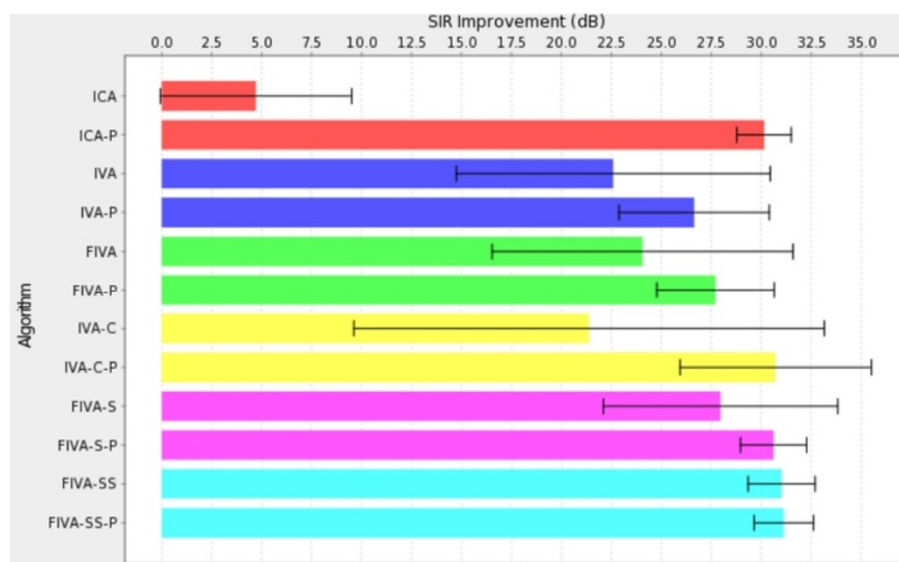


Figure 8 Performance of instantaneous mixture.

derived from the Gaussian distribution in (19) was also tested, which utilizes the second-order statistics between frequency bins; however, the algorithm failed to separate the speech sources. It means that although this non-linearity works well in the joint BSS problems for fMRI data in [17], only second-order correlation may still not adequate to separate mixed speeches by IVA.

Several conclusions can be deduced from this experiment. First, performances of the natural gradient IVA (IVA) and the fast fixed-point IVA (FIVA) are much higher than the ICA without permutation correction (ICA); this means that IVA algorithms can dramatically alleviate the permutation ambiguity. However, the permutation problem is still not perfectly solved by IVA, as their performances can be further improved after permutation correction (IVA-P and FIVA-P). Second, large standard deviations of IVA algorithms (IVA, FIVA, IVA-C) mean that they are not stable enough, their depermutation ability are highly related to the input data and the mixing environment. Third, even postprocessed by the permutation algorithm, performances of the full frequency band IVA (IVA-P, FIVA-P) are still not as good as ICA with permutation correction (ICA-P), a similar result was also observed in [16]. A possible explanation to this phenomenon is that uncorrelated frequency bins in the full frequency band will degrade the IVA performance. When subband techniques are used (IVA-C-P, FIVA-S-P), separation performances become comparable with the traditional ICA plus permutation correction approach, as inter-frequency bin dependencies become stronger in subbands or cliques. Fourth, when the proposed subband approach is used alone (FIVA-S), the average performance is already higher than

other compared IVA algorithms (IVA, FIVA, IVA-C). When the subspace nonlinearity is further used (FIVA-SS), both the separation performance and the stability are improved to a comparable level with the ICA-P approach; however, no permutation algorithm is followed. When postprocessed by the permutation algorithm (FIVA-SS-P), only a tiny performance improvement was observed, the marginal gain in performance achieved by FIVA-SS-P relative to FIVA-SS indicates that the additional complexity of a postprocessing source alignment algorithm is not required for FIVA-SS.

6.2. Convolutional mixture

In this experiment, a virtual room was established similar to [33], the mixing environment configurations are shown in Figure 9, and the mixing filters from each source to each sensor were generated by the image

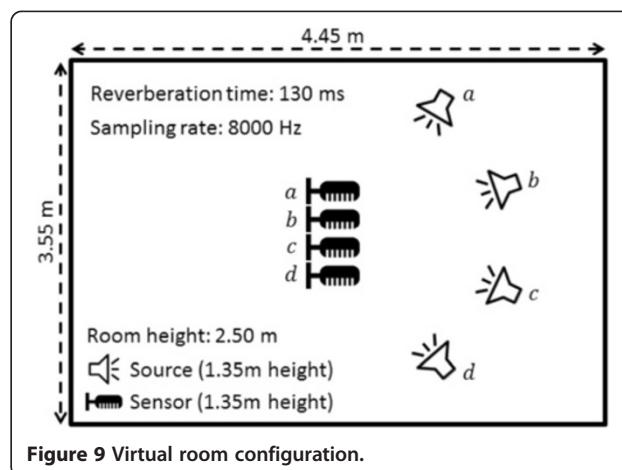


Figure 9 Virtual room configuration.

Table 1 Performance in convolutive mixture

Algorithm	Type of mixtures (source and sensor combinations)		
	2 × 2 mixture (a, d)	3 × 3 mixture (a, b, d)	4 × 4 mixture (a, b, c, d)
ICA-P	19.66	20.06	22.59
IVA	4.39	5.13	10.01
FIVA	0.10	10.91	8.63
IVA-C	4.98	6.04	5.81
FIVA-SS	19.22	18.77	21.18

method [34,35]. The dataset in [28] was also used in this experiment, the STFT frame size was set to 1024, with 7/8 overlapping, and the FFT block size was set to 2048 for all algorithms.

SIR improvements of this experiment are shown in Table 1. Since long STFT size and FFT size are required in long reverberation environment, the difficulty of the permutation problem is also increased. The proposed FIVA-SS approach is only a little inferior to the traditional ICA-P approach; however, FIVA-SS outperforms other IVA algorithms considered in this experiment. The proposed approach was also tested on real-world audio signals that were recorded by the BSS platform and the sound capture device in Figure 7, some separation examples can be found at [26].

7. IVA application: high-speed train noise component separation

High-speed train noise level is an important factor with respect to passenger comfort and life quality of residents along the railway. Determining how to attenuate the noise level is an important research direction that train designers care about [36,37]. Studies show that train noise is a kind of mixed signal which is made up of train body vibration, rolling noise, aerodynamic noise, device noise, etc. [37]. Separating individual noise components

from the overall observations will provide some guide to train noise reduction design. Since the noise component is of interest, here we use “noise signal” or “noise component” to distinguish it from the common use of “noise” for undesired interference.

Since BSS and ICA have many successful applications in speech and audio separation tasks, a natural choice is to use these techniques to perform train noise component separation; however, noise signals from mechanical vibration are very different from speech. Figure 10 shows two spectrogram examples of these signals, compared with the speech spectrogram, we can see from Figure 10 that noise signal T-F data are more stationary but not as sparse as speech, so, non-Gaussianity in each frequency bin is not strong. All these characteristics make individual noise components more difficult to be separated by ICA [38]. Moreover, neighboring frequency bins in noise spectrogram are weakly correlated, which increases the difficulty of the permutation problem.

Since IVA utilizes inter-frequency dependencies in the separation and can avoid the permutation problem, we expect that IVA has better performance in noise component separation tasks than the traditional ICA-P approach.

7.1. Simulation experiment

The SPIB noise dataset [39] was used in the simulation, including (1) destroyer engine room noise, (2) factory noise, (3) tank noise, (4) military vehicle noise. The first 8 s of signals were used, the sampling rate was 19.98 kHz. Mixing filters were randomly generated by concatenating different all-pass filters. The STFT frame size was set to 512, with 3/4 overlapping, and the FFT block size was set to 1024. Experimental results are given in Table 2.

In this simulation experiment, we observed that when the ICA-P approach was used, ICA on many frequency bins failed to converge; however, when IVA was used,

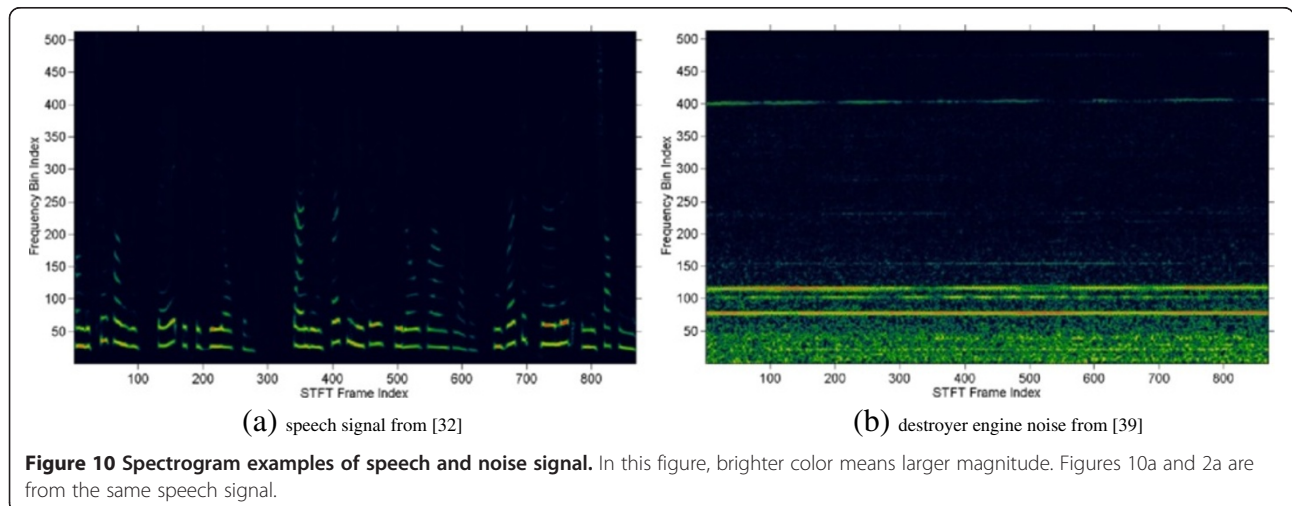


Table 2 SIR improvement of noise component separation

Dataset	Algorithm	Mixing filter length (number of samples)		
		1	256	512
(1), (2)	ICA-P	8.33	9.69	9.10
	FIVA	16.70	14.32	9.50
	FIVA-SS	18.16	14.53	14.00
(2), (3)	ICA-P	12.01	8.54	8.71
	FIVA	13.62	8.10	11.58
	IVA-SS	18.68	11.58	16.09
(4), (3)	ICA-P	7.69	8.27	2.44
	FIVA	9.82	7.25	6.52
	FIVA-SS	14.39	10.71	7.74

better convergence ability was observed. This suggests that for noise signals, inter-frequency bin information can improve the separation. From Table 2, we can see that the separation performances of all algorithms in this experiment are much lower than the performances in the previous section, this is due to the stationary and strong Gaussianity of noise signals make it difficult to separate these components. However, compared with the traditional ICA-P approach, the performances of the IVA algorithms are greatly improved, so IVA is a better choice for noise component separation tasks.

7.2. Train noise component separation

In this application, train noise signals were collected by four sound pressure sensors *a-d* in Figure 11, the corresponding train speed was 380 km/h, the sampling rate was 65,536 Hz. The low-frequency part of the signals is of more interest to train designers, so data are low pass filtered and down sampled to 1024 Hz before separation to reduce the data size.

Train noise is highly related to mechanical vibration, and different physical devices have different intrinsic frequencies. Although the original train noise signal exhibits greatly uncertainty and randomness, just like “noise”, the

underlying intrinsic frequencies can be revealed by calculating the signal’s autocorrelation sequence [36,40]. Figure 12 shows a sensor signal’s autocorrelation and the corresponding spectrum, we can see that two components are mixed in the observed data, component 1 is related to the train body vibration, and component 2 is related to the nearby device. Signals collected by other sensors in Figure 11b have similar characteristics like Figure 12, while their ratios are different.

After the proposed FIVA-SS approach was used, two of the total four output signals’ autocorrelations and spectrums are given in Figure 13. Unlike Figure 12, two dominant frequencies are individually shown by two separated signals, so we can infer that train noise components were separated in this experiment.

8. Conclusion

With the help of the multivariate nonlinear mapping, IVA is able to solve the permutation problem in the source separation procedure, so there is no need for an extra permutation correction algorithm as the postprocessing step. In this article, the subband and the subspace nonlinearity approaches are proposed as two improvements for IVA. Because of the local dependency property of the T-F data, performing IVA in subbands will gain more inter-frequency dependency than in the full frequency band, and this is useful for overcoming the permutation ambiguity. In each subband, highly correlated T-F data are likely distributed in the one-dimensional subspace of the original data space, so, IVA nonlinearity is calculated in the dominant subspace to meet the actual data distribution and for the sake of denoising. A platform was developed in Java for FDBSS research and real-world application purpose, and all our experiments were carried out on this platform. Experimental results show that the separation performance and the algorithm stability are greatly improved by the

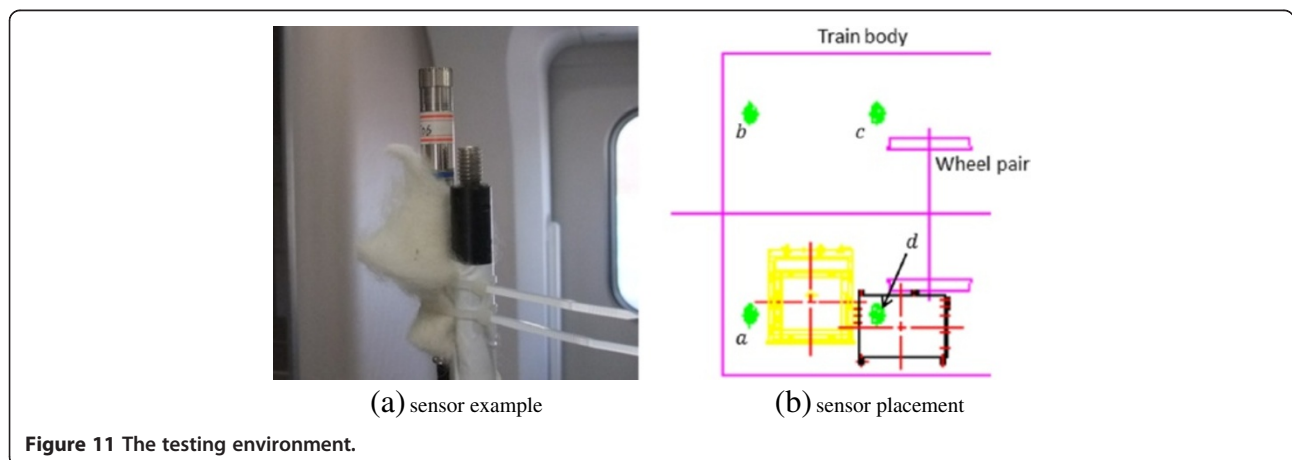
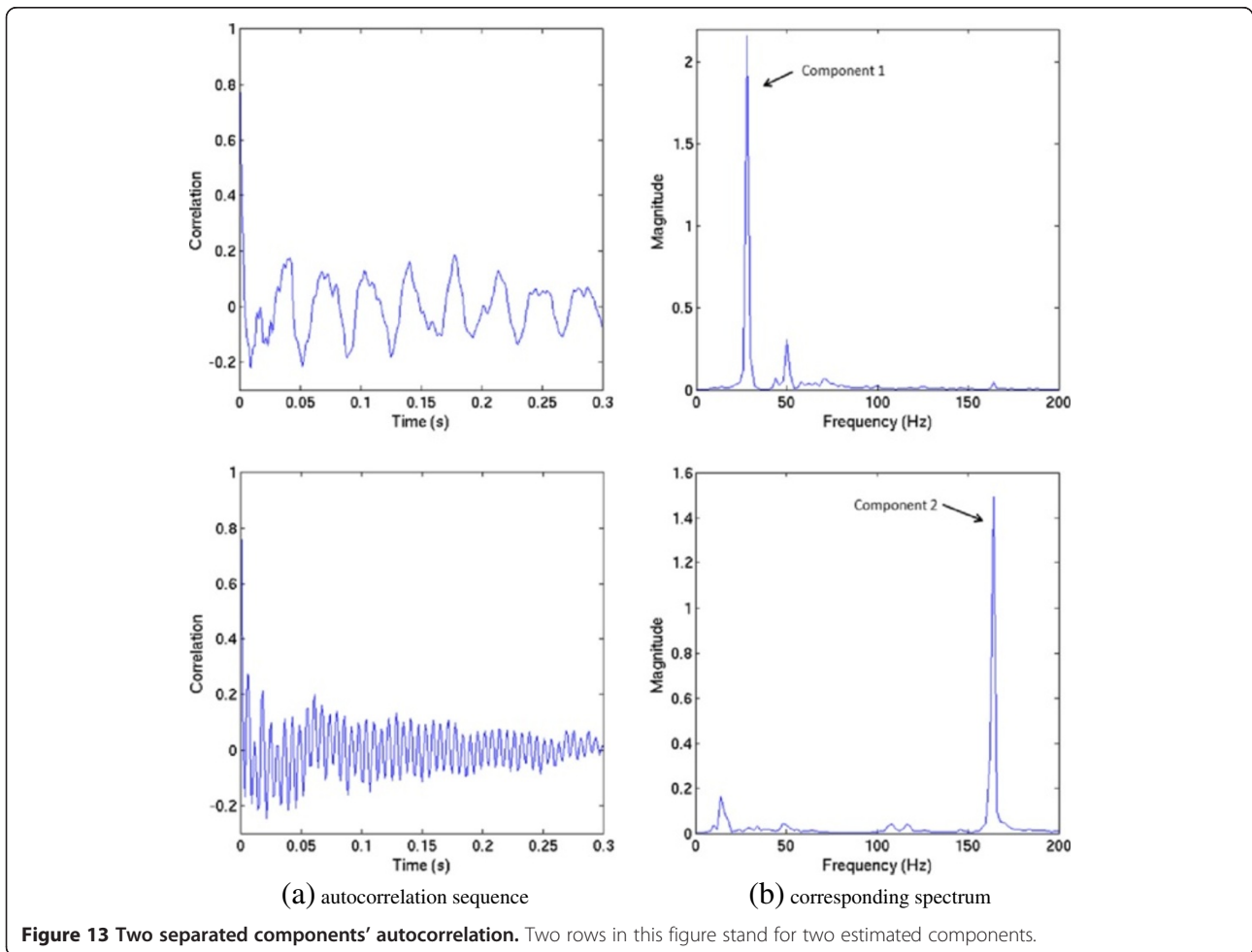
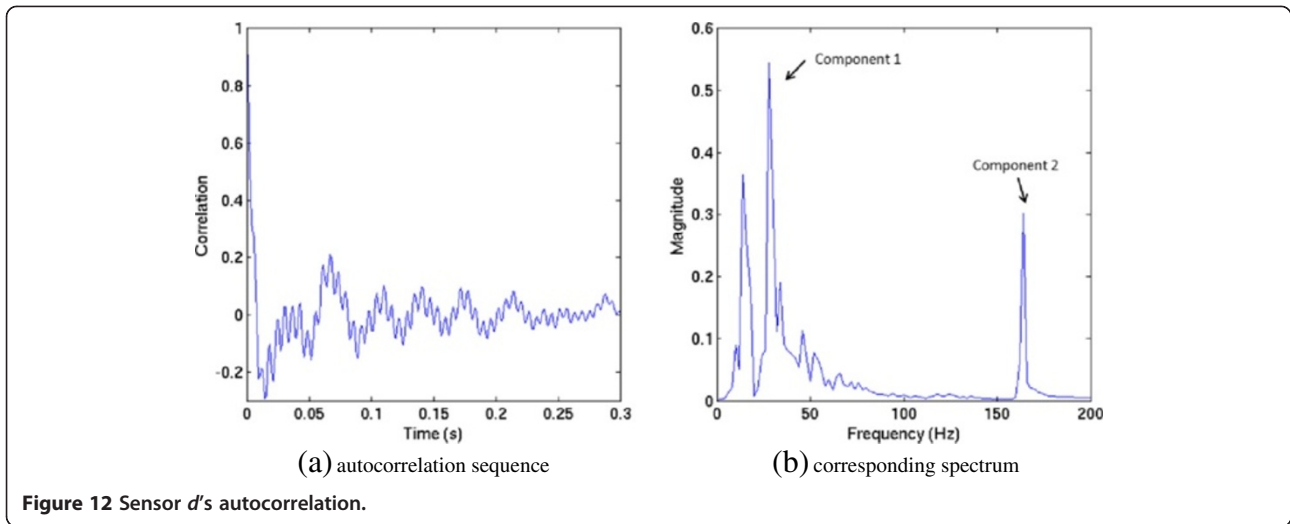


Figure 11 The testing environment.



proposed methods. Lastly, as an example of real-world application, the FIVA algorithm with the proposed updates was used to separate vibration components from high-speed train noise data.

Additional file

Additional file 1: The animation for the subband subspace IVA iteration procedure in Section 4.2. Description: This animation visualizes the variation of the subband correlation matrix and its five largest eigenvalues as the algorithm proceeds, the experimental configuration is the same as Section 6.1.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions. This research was supported by the National Natural Science Foundation of China, Grant no. 61033013, the National Natural Science Foundation of China, Grant no. 81230086, and the Fundamental Research Funds for the Central Universities, Grant no. 2012YJS027.

Author details

¹Department of Computer Science, Beijing Jiaotong University, No. 3 Shang Yuan Cun, Hai Dian District, Beijing, China. ²Department of Information Engineering, Shijiazhuang University of Economics, Shijiazhuang, Hebei 050031, China.

Received: 22 October 2012 Accepted: 5 March 2013

Published: 10 April 2013

References

1. A Hyvarinen, E Oja, Independent component analysis: algorithms and applications. *Neural Netw.* **13**(4–5), 411–430 (2000)
2. S Makino, H Sawada, R Mukai, S Araki, Blind source separation of convolutive mixtures of speech in frequency domain. *IEICE Trans. Fund. Electron.* **E88-A**(7), 1640–1655 (2005). doi:10.1093/ietfec/e88-a.7.1640
3. MS Pedersen, J Larsen, U Kjems, LC Parra, *A Survey of Convolutive Blind Source Separation Methods. Springer Handbook on Speech Processing and Speech Communication* (Springer, New York, 2007)
4. P Smaragdakis, Blind separation of convolved mixtures in the frequency domain. *Neurocomputing* **22**(1–3), 21–34 (1998). doi:10.1016/S0925-2312(98)00047-2
5. V Calhoun, T Adali, Complex infomax: convergence and approximation of infomax with complex nonlinearities. *J. VLSI Signal Process.* **44**, 173–190 (2006). doi:10.1007/s11265-006-7514-5
6. E Bingham, A Hyvarinen, A Fast, Fixed-point algorithm for independent component analysis of complex valued signals. *Int. J. Neural Syst.* **10**(1), 1–8 (2000)
7. H Sawada, R Mukai, S Araki, S Makino, A robust and precise method for solving the permutation problem of frequency-domain blind source separation. *IEEE Trans. Speech Audio Process.* **12**(5), 530–538 (2004). doi:10.1109/TSA.2004.832994
8. H Sawada, S Araki, S Makino, Measuring dependence of bin-wise separated signals for permutation alignment in frequency-domain BSS, in *IEEE International Symposium on Circuits and Systems* (, New Orleans, LA, 2007), pp. 3247–3250. doi:10.1109/ISCAS.2007.378164. 27–30 May
9. L Wang, H-P Ding, F-L Yin, A region-growing permutation alignment approach in frequency-domain blind source separation of speech mixtures. *IEEE Trans. Audio Speech* **19**(3), 549–557 (2011). doi:10.1109/TASL.2010.2052244
10. Y-Y Na, J Yu, Kernel and spectral methods for solving the permutation problem in frequency domain BSS, in *The 2012 International Joint Conference on Neural Networks* (, Brisbane, QLD, 2012), pp. 1–8. doi:10.1109/IJCNN.2012.6252698. 10–15 June
11. H Sawada, S Araki, R Mukai, S Makino, Grouping separated frequency components by estimating propagation model parameters in frequency-domain blind source separation. *IEEE Trans. Audio Speech* **15**(5), 1592–1604 (2007). doi:10.1109/TASL.2007.899218
12. T-T Ngo, S-H Nam, An expectation-maximization method for the permutation problem in frequency-domain blind source separation, in *International Conference on Acoustics, Speech and Signal Processing* (, Dallas, TX, 2010), pp. 17–20. doi:10.1109/ICASSP.2010.5496274. 14–19 March 2010
13. A Hiroe, Solution of permutation problem in frequency domain ICA, using multivariate probability density functions, in *International Conference on Independent Component Analysis and Blind Signal Separation*, vol. 3889 (, Charleston, SC, USA, 2006), pp. 601–608. doi:10.1007/11679363_75. March 5–8
14. T Kim, HT Attias, S-Y Lee, T-W Lee, Blind source separation exploiting higher-order frequency dependencies. *IEEE Trans. Audio Speech* **15**(1), 70–79 (2007). doi:10.1109/TASL.2006.872618
15. I Lee, T Kim, T-W Lee, Fast fixed-point independent vector analysis algorithms for convolutive blind source separation. *Signal Processing* **87**(8), 1859–1871 (2007). doi:10.1016/j.sigpro.2007.01.010
16. I Lee, T Kim, T-W Lee, Independent vector analysis for convolutive blind speech separation. *Blind Speech*, 169–192 (2007). doi:10.1007/978-1-4020-6479-1_6. Sep
17. M Anderson, T Adali, X-L Li, Joint blind source separation with multivariate Gaussian model: algorithms and performance analysis. *IEEE Trans. Signal Process.* **60**(4), 1672–1683 (2012). doi:10.1109/TSP.2011.2181836
18. T Itahashi, K Matsuoka, Stability of independent vector analysis. *Signal Process.* **92**(8), 1809–1820 (2012). doi:10.1016/j.sigpro.2011.11.008
19. T Kim, Real-time independent vector analysis for convolutive blind source separation. *IEEE Trans. Circuits Syst.* **57**(7), 1431–1438 (2010). doi:10.1109/TCSI.2010.2048777
20. J Hao, I Lee, T-W Lee, TJ Sejnowski, Independent vector analysis for source separation using a mixture of Gaussians prior. *Neural Comput.* **22**(6), 1646–1673 (2010). doi:10.1162/neco.2010.11-08-906
21. H Attias, Independent factor analysis. *Neural Comput.* **11**, 803–851 (1999)
22. Y-F Liang, SM Naqvi, JA Chambers, Audio video based fast fixed-point independent vector analysis for multisource separation in a room environment. *EURASIP J. Adv. Signal Process.* **183**(2012) (2012). doi:10.1186/1687-6180-2012-183
23. H-F Zhang, L-P Li, W-C Li, Independent vector analysis for convolutive blind noncircular source separation. *Signal Process.* **92**(9), 2275–2283 (2012). doi:10.1016/j.sigpro.2012.02.020
24. G-J Jang, I Lee, T-W Lee, Independent vector analysis using non-spherical joint densities for the separation of speech signals, in *International Conference on Acoustics, Speech and Signal Processing*, vol. 2 (, Honolulu, HI), pp. II-629–II-632. doi:10.1109/ICASSP.2007.366314. 15–20 April 2007
25. I Lee, G-J Jang, Independent vector analysis based on overlapped cliques of variable with for frequency-domain blind signal separation. *EURASIP J. Adv. Signal Process.* **113**(2012) (2012). doi:10.1186/1687-6180-2012-113
26. *The BSS platform (Y-Y Na, 2012)*. http://211.71.76.45/bss/, Accessed 19 Oct 2012
27. A Hyvarinen, J Karhunen, E Oja, *Independent Component Analysis* (Publishing House of Electronics Industry, Beijing, 2007)
28. *Hiroshi Sawada's dataset (H Sawada, R Mukai, S Araki, S Makino, 2003)*. http://www.kecl.ntt.co.jp/icl/signal/sawada/demo/bss2to4/index.html, Accessed 19 Oct 2012
29. C-H Choi, W Chang, S-Y Lee, Blind source separation of speech and music signals using harmonic frequency dependent independent vector analysis. *Electron. Lett.* **48**(2), 124–125 (2012). doi:10.1049/el.2011.3215
30. CM Bishop, *Neural Networks for Pattern Recognition* (Oxford University Press, New York, 1995)
31. Q-Y Li, N-C Wang, D-Y Yi, *Numerical Analysis* (Tsinghua University Press, Beijing, 2008), pp. 245–251 (in Chinese)
32. MZ Ikram, DR Morgan, Permutation inconsistency in blind speech separation: investigation and solutions. *IEEE Trans. Speech Audio Process.* **13**(1), 1–13 (2005). doi:10.1109/TSA.2004.834441
33. H Sawada, R Mukai, S Araki, S Makino, Convolutive blind source separation for more than two sources in the frequency domain, in *International Conference on Acoustics, Speech, and Signal Processing*, vol. 3 (Montreal, Quebec, Canada, 2004), pp. iii-885–iii-888. doi:10.1109/ICASSP.2004.1326687. 17–21 May
34. JB Allen, DA Berkley, Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Am.* **65**(4), 943–950 (1979). doi:10.1121/1.382599
35. *Room impulse response generator (E Habets, 2010)*. http://home.tiscali.nl/ehabets/rir_generator.html, Accessed 19 Oct 2012

36. Y-Y Na, J Yu, C Xie, High speed train transmission noise and structural noise separation. *J. Comput. Res. Dev.* in press
37. S-G Zhang, Noise mechanism, sound source localization and noise control of 350 km·h⁻¹ high-speed train. *China Railway Sci.* **30**(1), 86–90 (2009) (in Chinese)
38. A Masnadi-Shirazi, W-Y Zhang, BD Rao, Glimpsing IVA: a framework for overcomplete/complete/undercomplete convolutive source separation. *IEEE Audio Speech Lang. Process.* **18**(7), 1841–1855 (2010). doi:10.1109/TASL.2010.2052609
39. *SBIP noise dataset*, 1995. <http://graphics.stanford.edu/~jwshin/signal.html>. Accessed 18 May 2013
40. JG Proakis, DG Manolakis, *Digital Signal Processing Principles, Algorithms, and Applications*, 4th edn. (Publishing House of Electronics Industry, Beijing, 2010), pp. 116–129

doi:10.1186/1687-6180-2013-74

Cite this article as: Na et al.: Independent vector analysis using subband and subspace nonlinearity. *EURASIP Journal on Advances in Signal Processing* 2013 **2013**:74.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
