

Performance Analysis of Blind Source Separation Using Canonical Correlation

V. Anil Kumar¹ · Ch. V. Rama Rao¹ · Anirban Dutta¹

Received: 14 September 2016 / Revised: 18 April 2017 / Accepted: 20 April 2017 /
Published online: 4 May 2017
© Springer Science+Business Media New York 2017

Abstract Separation of blind source signals from a mixture remains an open issue. Many algorithms have been proposed for blind source separation (BSS) in the literature, but none outperforms the other. Most of the earlier BSS methods were based on the assumption that the sources are independent and non-Gaussian. From the literature, it is observed that speech signals are modelled using Gaussian models. This work focuses on a new approach for BSS in speech processing applications by considering the second-order statistics of the speech signals based on a canonical correlation approach. The performance of the algorithm is analyzed using signal-to-interference ratio, signal-to-distortion ratio, signal-to-artifact ratio and signal-to-noise ratio. Simulation results highlight the better performance of the proposed method as compared to the state of the art approaches like principal component analysis, singular value decomposition and independent component analysis algorithms.

Keywords Blind source separation · Speech signals · Principle component analysis · Singular value decomposition · Independent component analysis · Canonical correlation

✉ V. Anil Kumar
anil2siddu@gmail.com

Ch. V. Rama Rao
chvramaraogec@gmail.com

Anirban Dutta
anirbandutta118@gmail.com

¹ National Institute of Technology Meghalaya, Shillong 793003, India

1 Introduction

In signal and image processing, blind source separation (BSS) is used to reconstruct the source signals from the observed mixtures without the knowledge of either the mixtures or the input source signals [13]. Many applications of BSS have been identified in diverse areas like speech processing, seismic signal processing, data communication, bio-medical signal processing and passive sonar and antenna arrays used in wireless communication [6, 11, 21, 37]. BSS in speech processing is used for determining the mixing or unmixing matrix for the original speech signals from the observed mixtures [18]. The use of BSS has found use in many industrial applications like chemical component identification, infrared optical source separation and analysis of vibrations in rotating machine [7]. Apart from this, BSS is also used to separate ultrasonic signals for non-destructive evaluation (NDE) [36] and to distinguish the toe walking gait from normal gait in idiopathic toe walkers (ITW) [27]. BSS along with entropy rate bound minimization method [5, 25] is used for improving performance of the driver fatigue classification system. The single-channel source separation method based on ICA is used to characterize the different brain networks from an artificial language, speech stream, for learning of sub-serving word and to remove artifact from the source [6, 11]. It is also used to analyze the facial thermal images, such as separation for electromyogram and identifying different gestures [22, 23, 26]. In frequency domain ICA, more stable separation is possible using multiple frequency bins [28]. When the signals are contaminated by additive noise, robustness of ICA can be increased by spreading the noise power and localizing the source energy in the time-frequency domain [10]. The fast ICA algorithm of [37] was used on compressed spectral data for increasing independence of the source signals. ICA-based BSS is used for finding the physical causes in variations of the behavior of the climate. It is also used for improving performance of classification of finger movements for trans-radial amputee subjects along with Icaasso clustering [1, 24]. The performance of the ICA algorithm for noisy observations was improved by considering the time-frequency distribution [10]. ICA algorithm formulated for BSS is based on the assumption that, the independent components have non-Gaussian distributions. Some researchers improved performance of the BSS by considering the temporal autocorrelation of the source signals, but it fails, if such temporal correlation does not exist. ICA algorithm can still separate non-Gaussian sources, even in this situation. However, ICA- and temporal autocorrelation-based approaches for BSS are unable to separate the non-stationary and Gaussian source signals. Speech signals are modelled using the Gaussian model and these signals are non-stationary in nature. In speech processing applications, ICA- and temporal autocorrelation-based approaches are unable to separate the speech signals from its observations. This work introduces a new algorithm for BSS, which is capable of separating the speech signals in speech processing applications by exploiting the second-order statistics of the non-stationary observations based on canonical correlation technique.

Organization of this paper as follows: Sect. 2 deals with source separation with known sources and unknown sources using PCA-, SVD- and ICA-based BSS algorithms. Section 3 deals with the proposed method of canonical correlation-based BSS, performance analysis is explained in Sect. 4, and the conclusion is given in Sect. 5.

2 Source Separation

This section deals with source separation with known sources and without knowing the sources by considering the instantaneous mixing model.

2.1 With Known Sources

Source separation is the process of identifying the original signals from the mixtures. If $s_1(t)$ and $s_2(t)$ are known signals and $x_1(t)$, $x_2(t)$ are observed signals obtained by using instantaneous mixing model then, $x_1(t)$ and $x_2(t)$ can be written as

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t) \quad (1)$$

$$x_2(t) = a_{21}s_1(t) + a_{22}s_2(t) \quad (2)$$

Matrix representation for (1) and (2) is

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} \quad (3)$$

And in vector form (3) will be represented as $\mathbf{x} = \mathbf{A}\mathbf{s}$, where \mathbf{A} is a mixing matrix. Separated sources from \mathbf{x} are given by (4)

$$\begin{bmatrix} \hat{s}_1(t) \\ \hat{s}_2(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} \hat{s}_1(t) \\ \hat{s}_2(t) \end{bmatrix} = \frac{1}{(a_{11}a_{22} - a_{21}a_{12})} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad (5)$$

Let $\det \mathbf{A} = (a_{11}a_{22} - a_{21}a_{12})$ then (5) can be represented by the following equations

$$\hat{s}_1(t) = \frac{a_{22}}{\det \mathbf{A}}x_1(t) - \frac{a_{12}}{\det \mathbf{A}}x_2(t) \quad (6)$$

$$\hat{s}_2(t) = \frac{-a_{21}}{\det \mathbf{A}}x_1(t) + \frac{a_{11}}{\det \mathbf{A}}x_2(t). \quad (7)$$

2.2 Blind Source Separation (Unknown Sources)

The method of extracting the input source signals from mixtures without any earlier information of mixing ways or sources is called blind source separation. For single input multiple outputs, ensemble empirical mode decomposition (EEMD) is commonly used [12]. The BSS model for instantaneous mixture [4, 30, 38] is mathematically represented by (8)

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (8)$$

where $\mathbf{x}(t)$ represents the column vector of the observed output signals and $\mathbf{s}(t)$ represents input source signals. The matrix $\mathbf{x}(t)$ is formed by superimposing the columns

of the mixing matrix \mathbf{A} on $\mathbf{s}(t)$ which has independent signals $s_i(t)$, $i = 1, \dots, n$ and it is represented as a linear combination of input source signals given in [2, 17]. For separating independent signals from their mixtures, the following methods have been taken from [31, 33].

2.2.1 Principle Component Analysis

It is a technique that can be used to simplify a data set. By linear transformation, it selects a new coordinate system for the data set such that the greatest variance by projection on the first axis is known as the first principal component and second one on the second axis. Principal component analysis is used to reduce the dimensionality by removing the other principal components.

2.2.2 Singular Value Decomposition

The difference between singular value decomposition (SVD) and PCA is that projection in PCA is not scaled by the singular values. The SVD is decomposed into the product of three matrices U , Σ and V . If $\mathbf{x} = U \Sigma V^T$ where U consists of set of left orthonormal bases, Σ is a diagonal matrix and V is a set of right orthonormal bases.

2.2.3 Independent Component Analysis

Using ICA, a complex data set is decomposed into independent sub-parts. For a given two linear mixtures containing independent source signals where the value of first source signal does not contain any information of second source signal are represented in (8). Separation of the mixed signals using ICA gives needful results based on two assumptions [29, 35].

1. The given source signals should be independent.
2. Source signals should have non-Gaussian distribution values.

In the first step, the observed data are whitened (sphere) [19] which is nothing but removing correlations in the data, i.e., to make the signals uncorrelated [20]. Separated source matrix, \mathbf{y} , will be obtained using $\mathbf{y} = \mathbf{w}\mathbf{x}$. Mathematically, a linear transformation of \mathbf{w} (unmixing matrix) is done such that $E\{\mathbf{y}\mathbf{y}^T\} = I$. This can be obtained by setting, $\mathbf{w} = C^{-1/2}$ having $C = E\{\mathbf{x}\mathbf{x}^T\}$ which is called as correlation matrix of the signal data. Then it can be shown that $E\{\mathbf{y}\mathbf{y}^T\} = E\{\mathbf{w}\mathbf{x}\mathbf{x}^T\mathbf{w}^T\} = C^{-1/2}CC^{-1/2} = I$. A fast ICA algorithm using negative entropy concept given in [9, 32] is explained by following steps.

1. The observed signal \mathbf{x} is centered to make the average or mean value zero

$$\mathbf{x} - \mathbf{x}_m, \quad \mathbf{x}_m = E\{\mathbf{x}\}.$$

2. \mathbf{x} is whitened to maximize non-Gaussian characteristics

$$\mathbf{z} = V\Lambda^{-1/2}V^T\mathbf{x}, \quad V\Lambda V^T = E\{\mathbf{x}\mathbf{x}^T\}.$$

3. The first random vector \mathbf{w} is chosen such that,

$$\|\mathbf{w}\| = 1$$

4. Update \mathbf{w} (to make non-Gaussian)

$$\begin{aligned}\mathbf{w} &= E \left\{ \mathbf{z}^* g \left(\mathbf{w}^T \mathbf{w} \right) \right\} - E \left\{ g' \left(\mathbf{w}^T \mathbf{z} \right) \right\} \mathbf{w}, \\ g(\mathbf{y}) &= \tanh(a_1 \mathbf{y}) \text{ or } \mathbf{y}^* \exp \left(-\mathbf{y}^2 / 2 \right), \quad 1 < a_1 < 2 \\ \mathbf{w} &= \frac{\mathbf{w}}{\|\mathbf{w}\|}\end{aligned}$$

5. If convergence is not taking place, go to step 4.

6. Independent components of \mathbf{s} are indicated by

$$\mathbf{s} = [w_1 w_2 \dots w_n] \mathbf{x}$$

The performance measure of the ICA can be improved by segmenting the speech signal. PCA minimizes the covariance of the data; but ICA lowers the higher-order statistics like kurtosis (or cumulant of fourth order). As a result, it will reduce the mutual information [15] of respective output signal. Particularly, PCA contains orthogonal vectors of high energy contents in the form of the variance, but ICA identifies non-Gaussian signals [8] of independent components.

3 Blind Speech Separation Using Canonical Correlation

The canonical correlation approach was discussed in [16] which was more popular in various signal processing and data analysis applications. The linear relationship between multidimensional data set will be measured using canonical correlation. The canonical correlation property says that correlation variables will not change if the transformations of variable changes. This is an important difference between general correlation and canonical correlation. Mathematical procedure for canonical correlation is illustrated in Fig. 1. Where $s_1(t) = \mathbf{w}_x^H x_1(t)$ and $s_2(t) = \mathbf{w}_y^H x_2(t)$. Here, if we consider $x_1(t) = \mathbf{X}$ and $x_2(t) = \mathbf{Y}$ then the separated source signals can be represented as $s_1(t) = \mathbf{w}_x^H \mathbf{X}$ and $s_2(t) = \mathbf{w}_y^H \mathbf{Y}$ with two demixing vectors \mathbf{w}_x^H and \mathbf{w}_y^H , respectively. Using canonical correlation analysis, $\mathbf{w}_x^H \mathbf{X}$ and $\mathbf{w}_y^H \mathbf{Y}$ are obtained from \mathbf{w}_x and \mathbf{w}_y which has the canonical correlations as columns. The correlation between \mathbf{X} and \mathbf{Y} is given by

$$\rho \left(\mathbf{w}_x^H \mathbf{X}, \mathbf{w}_y^H \mathbf{Y} \right) = \text{Corr} \left(\mathbf{w}_x^H \mathbf{X}, \mathbf{w}_y^H \mathbf{Y} \right) = \frac{\text{Cov} \left(\mathbf{w}_x^H \mathbf{X}, \mathbf{w}_y^H \mathbf{Y} \right)}{\sqrt{\text{Var} \left(\mathbf{w}_x^H \mathbf{X} \right)} \sqrt{\text{Var} \left(\mathbf{w}_y^H \mathbf{Y} \right)}} \quad (9)$$

Mathematical procedure for canonical correlation is as follows

1. Divide the mixture data into two groups $x_1(t) = \mathbf{X}$ and $x_2(t) = \mathbf{Y}$ which are linear mixture of sources from $s_1(t)$ and $s_2(t)$. These are assigned by the following random vectors.

$$\mathbf{X} = [X_1, X_2, \dots, X_p]^T \text{ and } \mathbf{Y} = [Y_1, Y_2, \dots, Y_q]^T$$

2. Find the linear combinations of \mathbf{X} and \mathbf{Y} by following equations.

$$U_1 = \mathbf{a}_1^T \mathbf{X} = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \text{ Where } \mathbf{a} = \mathbf{w}_x^T.$$

$$U_2 = \mathbf{a}_2^T \mathbf{X} = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p$$

$$\vdots$$

$$U_p = \mathbf{a}_p^T \mathbf{X} = a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pp}X_p$$

And

$$V_1 = \mathbf{b}_1^T \mathbf{Y} = b_{11}Y_1 + b_{12}Y_2 + \dots + b_{1q}Y_q \text{ Where } \mathbf{b} = \mathbf{w}_y^T.$$

$$V_2 = \mathbf{b}_2^T \mathbf{Y} = b_{21}Y_1 + b_{22}Y_2 + \dots + b_{2q}Y_q$$

$$\vdots$$

$$V_q = \mathbf{b}_q^T \mathbf{Y} = b_{q1}Y_1 + b_{q2}Y_2 + \dots + b_{qp}Y_q$$

3. The above linear equations U and V will maximize the correlation given by

$$\text{Corr}(U_i, V_i) = \frac{\text{Cov}(U_i, V_i)}{\sqrt{\text{Var}(U_i)}\sqrt{\text{Var}(V_i)}}, \quad i = 1, 2, \dots, n$$

Where $n = \min(p, q)$.

4. The first pair canonical variables is given by

$$\text{Corr}(U_1, V_1) = \rho_1 = \sqrt{\rho_1^2} \text{ where } \rho_1 \text{ is the first canonical correlation.}$$

It is the highest of the eigenvalues, $\rho_1^2 \geq \rho_2^2 \geq \dots \geq \rho_n^2$ which is the eigenvalues of

$$\text{the matrix } \sum_{xx}^{-1/2} \sum_{xy} \sum_{yy}^{-1} \sum_{xy}^T \sum_{xx}^{-1/2}$$

where \sum_{xx} is the variance or covariance of \mathbf{X}

\sum_{yy} is the variance or covariance of \mathbf{Y}

\sum_{xy} is the covariance matrix of the random vector \mathbf{XY}

The second pair canonical variable is given by

$$\text{Corr}(U_2, V_2) = \rho_2 = \sqrt{\rho_2^2} \text{ and so on.}$$

5. Mixing matrix is estimated using singular value decomposition (SVD)

Fig. 1 Mathematical procedure for canonical correlation

In the above equation, it can be observed that if p_1, p_2 scalars are multiplied, then $\rho(p_1\mathbf{X}, p_2\mathbf{Y}) = \rho(\mathbf{X}, \mathbf{Y})$. Hence, the optimization problem is not affected from multiplication by positive scalars. The problem is to find the directions \mathbf{X} and \mathbf{Y} that maximizes the Eq. (9). It is observed that the choice of scaling is arbitrary; therefore, we maximize the Eq. (9) subject to constraint $\mathbf{w}_y^H R_y \mathbf{w}_y = 1$ and $\mathbf{w}_x^H R_x \mathbf{w}_x = 1$

using Lagrangian generalized eigenvalue problem, where $R_x = E[\mathbf{X}\mathbf{X}^H]$ and $R_y = E[\mathbf{Y}\mathbf{Y}^H]$. The minimization function for maximizing the correlation between \mathbf{w}_x and \mathbf{w}_y is given in (10)

$$\Phi(\mathbf{w}_y, \mathbf{w}_x) = E \left[\left| \mathbf{w}_y^H \mathbf{Y} - \mathbf{w}_x^H \mathbf{X} \right|^2 \right] \tag{10}$$

The Eq. (10) can be written as

$$\Phi(\mathbf{w}_y, \mathbf{w}_x) = \text{trace} \left[-\mathbf{w}_y^H R_y \mathbf{w}_y + \mathbf{w}_x^H R_x \mathbf{w}_x - \mathbf{w}_y^H R_y \mathbf{X} \mathbf{w}_x - \mathbf{w}_x^H R_x \mathbf{Y} \mathbf{w}_y \right] \tag{11}$$

To minimize $\Phi(\mathbf{w}_y, \mathbf{w}_x)$, we have to derive component of $\mathbf{w}_y, \mathbf{w}_x$ by using Lagrange operations Λ et Δ which are not discussed in this paper.

Now, we have two equations:

$$R_{xy} \mathbf{w}_y = R_x \mathbf{w}_x \Lambda \tag{12}$$

$$R_{yx} \mathbf{w}_x = R_y \mathbf{w}_y \Delta \tag{13}$$

After multiplying (12) by $R_{yx}^{-1} R_{yx}$ we get as

$$\begin{aligned} R_{xy} \mathbf{w}_y &= R_x R_{yx}^{-1} R_{yx} \mathbf{w}_x \Lambda \\ &= R_x R_{yx}^{-1} R_y \mathbf{w}_y \Delta \Lambda \end{aligned} \tag{14}$$

If \mathbf{w}_x is required, we can have the dual equation. Assume $T = \Delta \Lambda$, then

$$R_y^{-1} R_{yx} R_x^{-1} R_{xy} \mathbf{w}_y = \mathbf{w}_y T \tag{15}$$

After multiplying $R_y^{1/2}$ to both sides of (15), we have the equation as

$$R_y^{-1/2} R_{yx} R_x^{-1/2} R_x^{-1/2} R_{xy} R_y^{1/2} R_y^{1/2} \mathbf{w}_y = R_y^{1/2} \mathbf{w}_y T \tag{16}$$

If $R_y^{-1/2} R_{yx} R_x^{-1/2} = D$, and $R_y^{1/2} \mathbf{w}_y = \mathbf{w}_y^*$ then Eq. (16) can be written as

$$D D^H \mathbf{w}_y^* = \mathbf{w}_y^* T \tag{17}$$

From (17) it is easy to find the eigenvectors and eigenvalues of D by choosing the L eigenvectors of \mathbf{w}_x according to L higher eigenvalues. A matrix \mathbf{w}_y can be written as

$$\mathbf{w}_y = R_y^{1/2} \mathbf{w}_x \tag{18}$$

Application For evaluating the canonical correlation-based BSS, we consider the instantaneous mixing model. Here, two speech signals are considered and these are mixed with following mixing matrix.

$$\mathbf{A} \approx \begin{bmatrix} 0.2588 & 0.7071 \\ 0.9659 & 0.7071 \end{bmatrix}$$

Using the canonical correlation approach the estimated unmixing matrix is obtained as

$$\hat{\mathbf{A}} \approx \begin{bmatrix} 0.8045 & 0.5939 \\ -0.5939 & 0.8045 \end{bmatrix}$$

Using the estimated unmixing matrix, source speech signals are separated and performance of this approach is analyzed and discussed in Sect. 4.

4 Results and Discussion

The simulation is carried out using MATLAB simulator. For analysis, two clean speech signals are taken from a NOIZEUS database [14], each with duration of 2 s and sampling frequency of 16,000 Hz. The performance is also analyzed on same speech signals in the airport noise environment. Considered source speech signals are shown in Fig. 2. The observed signals from mixture are shown in Fig. 3.

The performance of the proposed canonical correlation-based BSS is compared with PCA-, SVD- and ICA-based algorithms. The following performance measures are used for the comparison:

1. *Signal-to-interference parameter (SIR)* It is defined as the logarithmic ratio of normalized target value to the normalized interference value in the signal. The

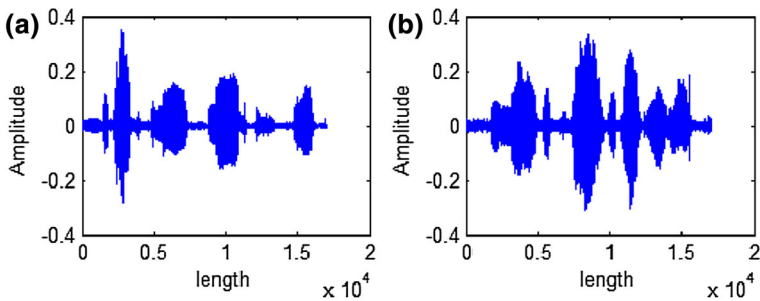


Fig. 2 a Clean speech signal 1, b clean speech signal 2

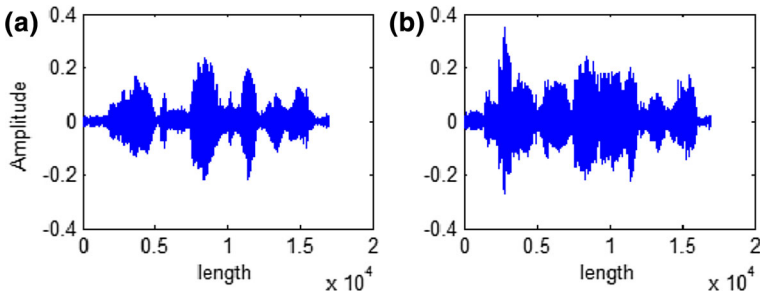


Fig. 3 a Observed signal 1, b observed signal 2

Table 1 The performance measures with respect to separated source 1

S. no.	Methods →	performance parameter ↓	Principal component analysis (PCA)	Singular value decomposition (SVD)	Independent component analysis (ICA)	Independent component analysis using frames	Canonical correlation-based BSS
1	SIR (dB)	Clean speech signal	2.34	2.47	-0.52	10.76	33.03
		Airport Environment	2.41	2.51	-1.52	10.84	7.54
			2.66	2.78	-0.91	11.01	33.59
			2.38	2.50	-0.82	10.78	39.52
			2.31	2.45	-0.47	10.75	25.29
2	SDR (dB)	Clean speech signal	2.34	2.47	-2.65	10.76	33.03
		Airport environment	2.41	2.51	-4.01	10.84	7.54
			2.66	2.78	-3.53	11.01	33.59
			2.38	2.50	-3.37	10.78	39.52
			2.31	2.45	-2.87	10.75	25.29
3	SAR (dB)	Clean speech signal	232.76	235.25	4.72	52.26	233.96
		Airport environment	255.85	256.35	3.42	51.75	260.22
			254.25	251.12	3.39	51.86	252.91
			247.97	248.98	3.59	52.14	247.41
			249.60	251.84	4.10	52.17	248.43
4	SNR (dB)	Clean speech signal	-28.47	7.77	-26.66	-2.53	11.52
		Airport environment	-25.68	7.75	-25.22	-2.53	14.31
			-27.44	7.74	-26.62	-2.68	12.55
			-28.08	7.80	-26.87	-2.56	11.91
			-28.33	7.74	-26.87	-2.51	11.66

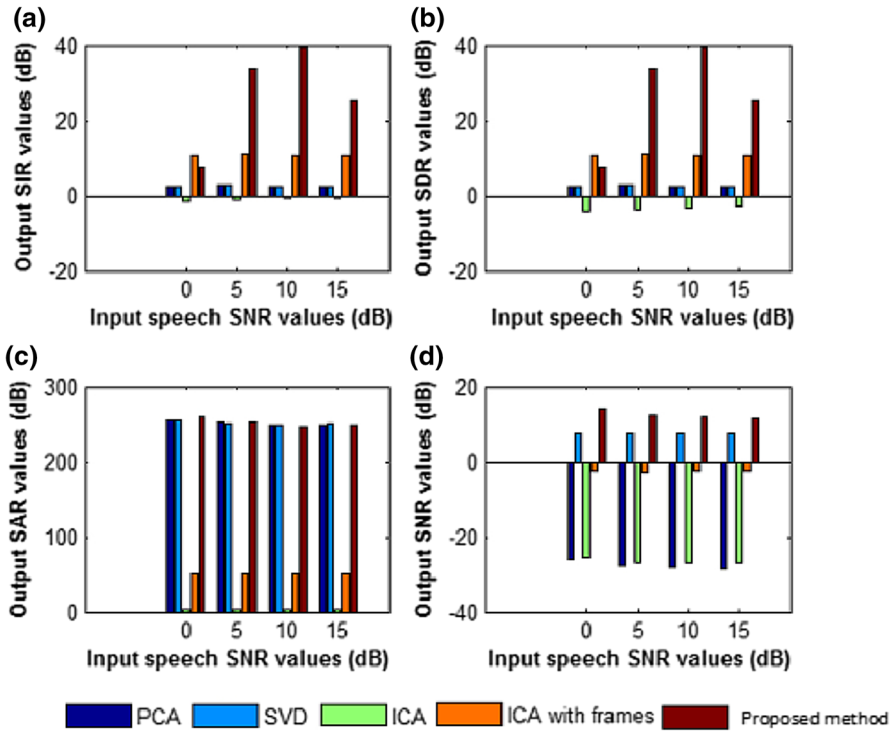


Fig. 4 The performance measures with respect to separated source 1 using PCA-, SVD-, ICA-, ICA with frames-based BSS and proposed method, **a** output SIR values, **b** output SDR values, **c** output SAR values, **d** output SNR values

high value of the SIR shows the best performance in separation [34]. The value of SIR is obtained using (19)

$$SIR = 10 \log_{10} \frac{\|s_{\text{target}}\|^2}{\|e_{\text{interference}}\|^2} \text{ (dB)} \tag{19}$$

2. *Signal-to-distortion ratio* (SDR) It is defined as the logarithmic ratio of the normalized target value to the normalized sum of the interference, noise and artifact values in the signal. The SDR is valid global performance measure used for measuring performance of BSS algorithms [34]. Mathematically, it is represented by

$$SDR = 10 \log_{10} \frac{\|s_{\text{target}}\|^2}{\|e_{\text{interference}} + e_{\text{noise}} + e_{\text{artifact}}\|^2} \text{ (dB)} \tag{20}$$

3. *Signal-to-artifact ratio* (SAR) It is defined as the logarithmic ratio of the normalized sum of the target, interference and noise value to the normalized interference value in the signal. The high value of the SAR is needed for better performance in

Table 2 The performance measures with respect to separated source 2

S. no.	Methods	→ performance parameter ↓	Principal component analysis (PCA)	Singular value decomposition (SVD)	Independent component analysis (ICA)	Independent component analysis using frames	Canonical correlation -based BSS
1	SIR (dB)	Clean speech signal	2.51	4.55	0.69	0.75	38.27
		Airport environment	2.70	4.20	1.85	0.76	7.89
			2.69	4.57	1.38	0.67	33.73
2	SDR (dB)	Clean speech signal	2.46	4.41	1.05	0.82	36.65
		Airport environment	2.57	4.68	0.59	0.71	28.06
			2.51	4.55	-1.70	0.75	38.27
3	SAR (dB)	Clean speech signal	2.70	4.20	-1.47	0.76	7.89
		Airport environment	2.69	4.57	-1.82	0.67	33.73
			2.46	4.41	-1.94	0.82	36.65
4	SNR (dB)	Clean speech signal	2.57	4.68	-2.05	0.71	28.06
		Airport environment	236.94	241.34	4.70	52.13	248.03
			256.86	258.66	3.42	51.58	254.73
5	SAR (dB)	Clean speech signal	258.01	258.21	3.37	51.75	262.84
		Airport environment	245.88	246.16	3.58	52.02	248.85
			251.40	249.97	4.07	52.08	252.60
6	SNR (dB)	Clean speech signal	-26.47	-1.79	-24.67	2.46	13.52
		Airport environment	-23.64	-1.75	-23.17	2.48	16.35
			-25.23	-1.75	-24.41	2.45	14.76
7	SNR (dB)	Clean speech signal	-26.09	-1.82	-24.83	2.45	13.90
		Airport environment	-26.34	-1.77	-24.88	2.47	13.65

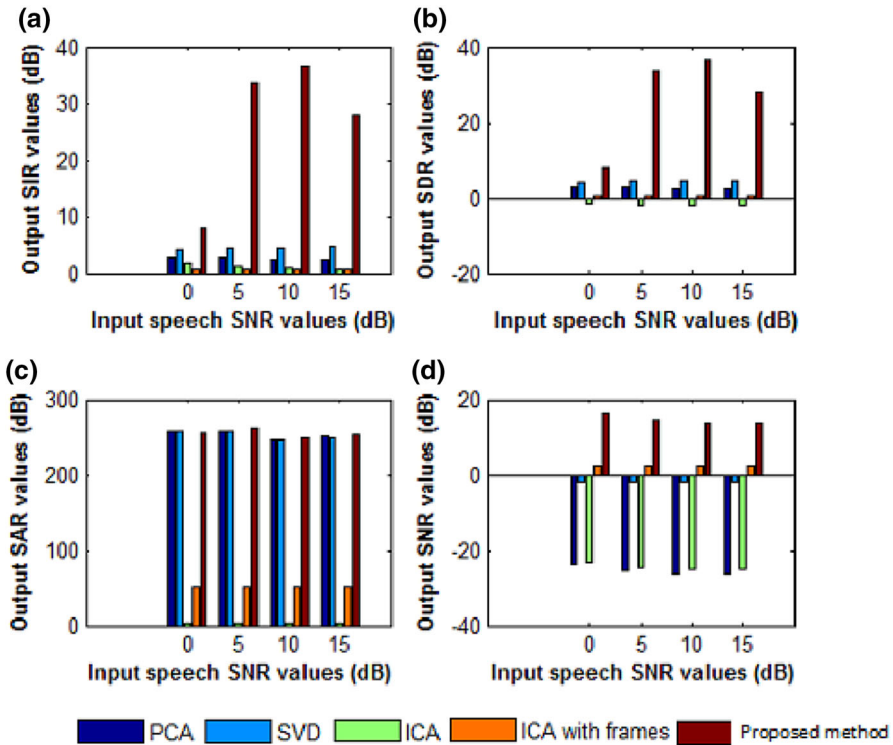


Fig. 5 The performance measures with respect to separated source 2 using PCA-, SVD-, ICA-, ICA with frames-based BSS and proposed method, **a** output SIR values, **b** output SDR values, **c** output SAR values, **d** output SNR values

BSS [34]. Mathematically, it is represented by

$$SAR = 10 \log_{10} \frac{\|s_{\text{target}} + e_{\text{interference}} + e_{\text{noise}}\|^2}{\|e_{\text{artifact}}\|^2} \text{ (dB)} \tag{21}$$

4. *Signal-to-noise ratio* (SNR) The signal-to-noise ratio is given by

$$SNR = 10 \log_{10} \frac{\frac{1}{N} \sum_{t=1}^N s_i(t)^2}{\frac{1}{N} \sum_{t=1}^N [s_i(t) - \hat{s}_i(t)]^2} \text{ (dB)} \tag{22}$$

The SNR value is useful for knowing rejection of the sensor noise in BSS [34].

The performance of the proposed BSS method with respect to the source 1 is tabulated in Table 1 and graphically represented in Fig. 4. From Table 1, it is observed that for higher input SNR values, the proposed method gives higher SIR, SDR and output SNR values, i.e., about 14, 14 and 5 dB, respectively, compared to the existing algorithms.

The performance measures with respect to the separated source 2 are tabulated in Table 2 and graphically illustrated in Fig. 5. As per Table 2, it is clear that for higher

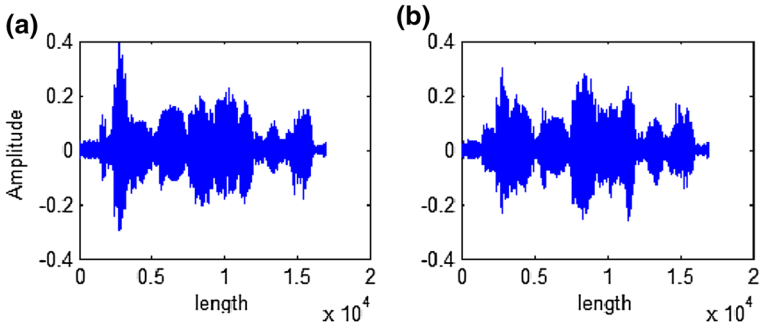


Fig. 6 Using PCA-based BSS, **a** separated speech signal 1, **b** separated speech signal 2

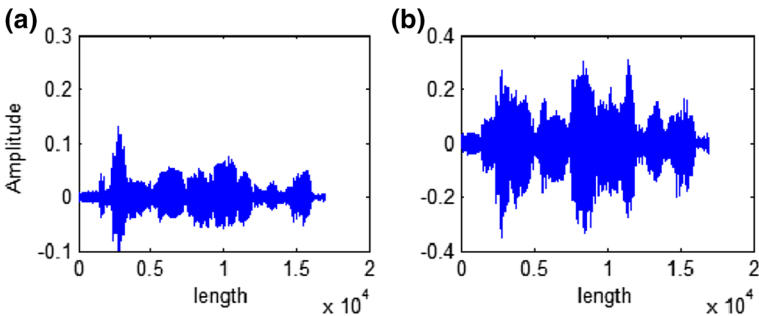


Fig. 7 Using SVD-based BSS, **a** separated speech signal 1, **b** separated speech signal 2

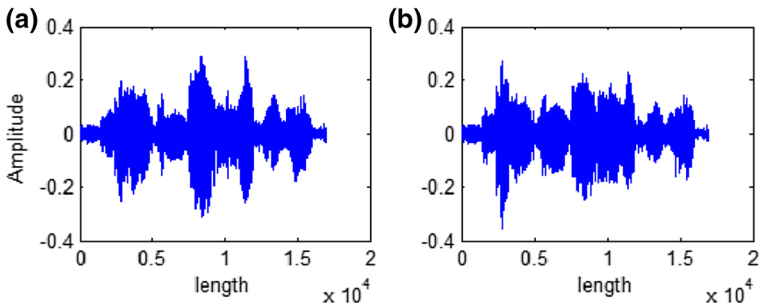


Fig. 8 Using ICA-based BSS, **a** separated speech signal 1, **b** separated speech signal 2

input SNR values for considered noise environment, the proposed method gives higher SIR, SDR and output SNR values, i.e., about 23, 23 and 10 dB, respectively, compared to the existing algorithms.

In order to illustrate the significance of the proposed method, its performance is also compared in terms of time domain waveforms with other BSS algorithms. The time domain waveforms of the separated sources using various BSS algorithms are shown in Figs. 6, 7, 8, 9 and 10.

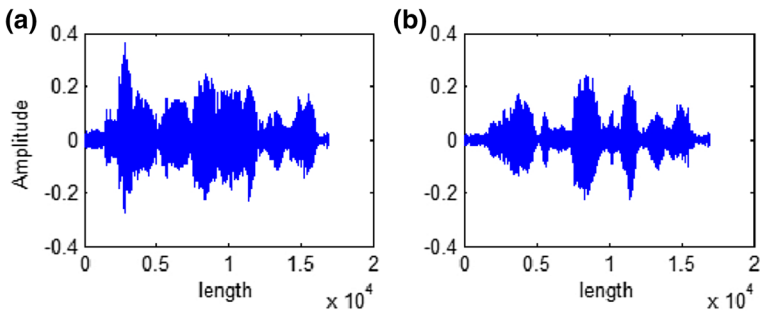


Fig. 9 Using ICA with frames-based BSS, **a** separated speech signal 1, **b** separated speech signal 2

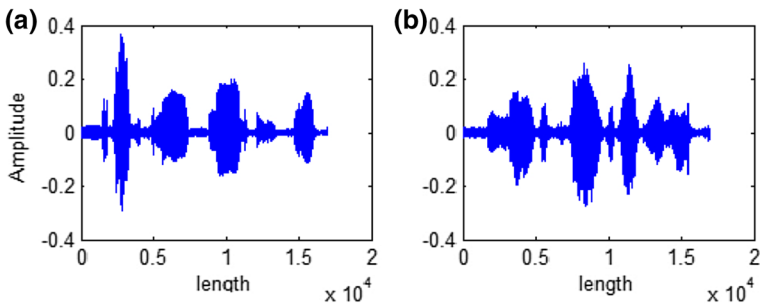


Fig. 10 Using proposed method, **a** separated speech signal 1, **b** separated speech signal 2

5 Conclusion

In this work, for evaluating the performance of the proposed approach for BSS, two speech signals are considered as sources and these signals are mixed using the instantaneous mixing model. From the literature, it was known that speech signals are non-stationary and Gaussian in nature. For separating the speech signals from the observations, we considered the second-order statistics based on canonical correlation. The performance of the proposed method is compared in terms of SIR, SDR, SAR and SNR with conventional BSS approaches like PCA, SVD and ICA. From the simulation results, it was observed that the canonical correlation-based approach is giving better results as compared to PCA-, SVD- and ICA-based BSS. The obtained results highlighted that canonical correlation-based approach is effective to eliminate the interference of the signals in separation. Based on the simulation results, it can be suggested that the proposed method for BSS is suitable for audio/speech processing applications. In further research, performance of the proposed method can be improved by considering direction of the arrival of the signals.

References

1. F. Aires, W.B. Rossow, A. Chédin, Rotation of EOFs by the independent component analysis: toward a solution of the mixing problem in the decomposition of geophysical time series. *J. Atmos. Sci.* **72**(9), 111–123 (2002)

2. J. Antoni, Blind separation of vibration component: principles and demonstrations. *Mech. Syst. Signal Process.* **19**(6), 1166–1180 (2005)
3. D. Barroso, P. Ripollés, J.M. Pallarés, B. Mohammadi, T.F. Münte, A.C. Bachoud-Lévi, A. Rodriguez-Fornells, R. Diego-Balaguer, Multiple brain networks underpinning word learning from fluent speech revealed by independent component analysis. *NeuroImage* **110**, 182–193 (2015)
4. A. Belouchrani, G. Moeness, Blind source separation based on time-frequency signal representations. *IEEE Trans. Signal Process.* **46**(11), 2888–2897 (1998)
5. R. Chai, G.R. Naik, T.N. Nguyen, S.H. Ling, Y. Tran, A. Craig, H. Nguyen, Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system. *IEEE J. Biomed. Health Inf.* (2016). doi:[10.1109/JBHI.2016.2532354](https://doi.org/10.1109/JBHI.2016.2532354)
6. P. Comon, C. Jutten, J. Herault, Blind separation of sources. Part II: problem statement. *Signal Process.* **24**(1), 11–20 (1991)
7. Y. Deville, Towards industrial applications of blind source separation and independent component analysis. Laboratoire d' Acoustique, de Metrologie, d' Instrumentation (LAMI) (Universite paul sabtier, France)
8. D. Erdogmus, E.K. Hild, C. J. Principe, L. Vielva, Blind separation of uncorrelated sources via principal component analysis of observations for a symmetric mixing matrix, in *11th European Signal Processing Conference*, pp. 1–4 (2002)
9. C. Ghita, R. Doru Raichu, B. Pantelimon, Implementation of the fast ICA algorithm in sound source separation, in *9th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, pp. 19–22 (2015)
10. J. Guo, Y. Deng, *A Time-Frequency Algorithm for Noisy ICA. Geo-Informatics Resource Management and Sustainable Ecosystem* (Springer, Berlin, 2015)
11. Y. Guo, G.R. Naik, H. Nguyen, Single channel blind source separation based local mean decomposition for biomedical applications, in *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 6812–6815 (2013)
12. Y. Guo, S. Huang, Y. Li, G.R. Naik, Edge effect elimination in single-mixture blind source separation. *Circuits Syst. Signal Process.* **32**(5), 2317–2334 (2013)
13. S. Haykin, *Unsupervised Adaptive Filtering*, vol 1 (Newyork, Chichester, Weinheim, Brisbane, Singapore, Toronto, 2000)
14. <http://www.utdallas.edu/~loizou/speech/oizeus/>
15. M.G. Jafari, W. Wang, J.A. Chambers, T. Hoya, A. Cichocki, Sequential blind source separation based exclusively on second order statistics developed for a class of periodic signals. *IEEE Trans. Signal Process.* **54**(3), 1028–1040 (2006)
16. K. Juha, T. Hao, J. Ylipaavalniemi, A generalized canonical correlation analysis based method for blind source separation from related data sets, in *the 2012 International Joint Conference on Neural Networks (IJCNN)*. IEEE (2012)
17. G. Kerschen, F. Poncelet, J.C. Golinval, Physical interpretation of independent component analysis in structural dynamics. *Mech. Syst. Signal Process.* **21**(4), 1561–1575 (2007)
18. C.T. Le, S. Moghaddammia, C. Kupferschmidt, T. Kaiser, Performance evaluation blind source separation algorithms base on source signal statistics in convolutive mixtures, in *International Conference on Communications and Electronics*, pp. 268–272 (2010)
19. Z. Li, J. An, L. Sun, M. Yang, A blind source separation algorithm based on whitening and non-linear decorrelation, in *Second International Conference on Computer Modelling and Simulation*, vol 1, pp. 443–447 (2010)
20. X. Lin Li, T. Adali, Independent component analysis by entropy bound minimisation. *IEEE Trans. Signal Process.* **58**(10), 5151–5164 (2010)
21. A. Mansour, A. A-Falou, Performance indices for real-world applications, in *14th European Signal Processing Conference*, pp. 1–5 (2006)
22. G.R. Naik, K.D. Kumar, Estimation of independent and dependent components of non-invasive EMG using fast ICA: validation in recognising complex gestures. *Comput. Methods Biomech. Biomed. Eng.* **14**(12), 1105–1111 (2011)
23. G.R. Naik, K.D. Kumar, M. Palaniswami, Signal processing evaluation of myoelectric sensor placement in low-level gestures: sensitivity analysis using independent component analysis. *Expert Syst.* **131**(1), 91–99 (2012)

24. G.R. Naik, A.H. Al-Timemy, H.T. Nguyen, Transradial amputee gesture classification using an optimal number of sEMG sensors: an approach using ICA clustering. *IEEE Trans. Neural Syst. Rehabil. Eng.* **24**(8), 837–846 (2016)
25. G.R. Naik, S.E. Selvan, H.T. Nguyen, Single-channel EMG classification with ensemble-empirical-mode-decomposition-based ICA for diagnosing neuromuscular disorders. *IEEE Trans. Neural Syst. Rehabil. Eng.* **24**(7), 734–743 (2016)
26. R. Okamoto, S. Bando, A. Nozawa, Blind signal processing of facial thermal images based on independent component analysis. *IEE J. Trans. Electron. Inf. Syst.* **136**(8), 1142–1148 (2016)
27. G. Pendharkar, G.R. Naik, H.T. Nguyen, Using blind source separation on accelerometry data to analyze and distinguish the toe walking gait from normal gait in ITW children. *Biomed. Signal Process. Control* **13**, 41–49 (2014)
28. D. Persia, E. Leandro, D.H. Milone, Using multiple frequency bins for stabilization of FD-ICA algorithms. *Signal Process.* **119**, 162–168 (2015)
29. M. Plumbly, Algorithm for nonnegative independent component analysis. *IEEE Trans. Neural Netw.* **14**(3), 534–543 (2003)
30. V.G. Reju, S.N. Koh, I.Y. Soon, An algorithm for mixing matrix estimation in instantaneous source separation. *Signal Process.* **89**(9), 1762–1773 (2009)
31. B. Rivert, L. Girin, C. Jutten, Mixing audio–visual speech processing and blind source separation for the extraction of speech signals from convolutive mixtures. *IEEE Trans. Audio Speech Lang. Process.* **15**(1), 96–108 (2006)
32. T. Shijiea, C. Hang, Blind source separation of underwater acoustic signal by use of negentropy-based fast ICA algorithm, in *IEEE International Conference on Computational Intelligence & Communication Technology*, pp. 608–611 (2015)
33. P.M. Syskind, W. DeLiang, J. Larsen, U. Kjems, Two-microphone separation of speech mixtures. *IEEE Trans. Neural Netw.* **19**(3), 475–492 (2008)
34. E. Vincent, R. Gribonval, C. Févotte, Performance measurement in blind audio source separation. *IEEE Trans. Audio Speech Lang. Process.* **14**(4), 1462–1469 (2006)
35. J. Wang, C. Chang, Independent component analysis based dimensionality reduction with applications in hyper-spectral image analysis. *IEEE Trans. Geosci. Remote Sens.* **44**(6), 1586–1600 (2006)
36. Y. Wu, T.T. Guo, B. Zhang, L. Guo, J. Wu, L. Zhao, Applications of blind source separation in ultrasonic NDE, in *Second International Workshop on Education Technology and Computer Science*, vol 3, pp. 115–118 (2010)
37. W. Ying, G. Tian-tai, J. Jie-wei, Applications of blind sources separation in plant leaves classification, in *10th World Congress on Intelligent Control and Automation*, pp. 4174–4179 (2012)
38. L. Zou, X. Chen, Z. Jane, Underdetermined joint blind source separation for two datasets based on tensor decomposition. *IEEE Signal Process.* **23**(5), 673–677 (2016)