ROMANIAN JOURNAL OF INFORMATION SCIENCE AND TECHNOLOGY
Volume **12**, Number 1, 2009, 3–12

Blind Separation for Speech Signals by Minimizing the Contingency Coefficient in the Chi-Square Independence Test

Marina-Anca CIDOTA

University of Bucharest, Faculty of Mathematics and Computer Science,
Str. Academiei no. 14, 010014 Bucharest, Romania

E-mail: cidota@fmi.unibuc.ro

Abstract. In this paper, the problem of blind source separation (BSS) using an independence measure for observations is addressed. Random variables are transformed by their respective estimate distribution functions into uniform random variables, whose independence is evaluated using the contingency coefficient from the chi-square independence test. A new objective function for the adaptive blind source separation algorithm is thus proposed. Simulations illustrate that this algorithm based on minimizing the contingency coefficient leads to reliable results when applied to separate speech signals which are linearly mixed.

Key words: blind source separation, independent component analysis, independence measure, contingency coefficient.

1. Introduction

Blind source separation (BSS) means to extract independent signals from their linear mixtures without knowing how they have been mixed. There are many interesting applications of blind source separation that need reliable solutions. One example from acoustics (known as the "cocktail-party" problem) includes the separation of the speech signals recorded from several microphones where different persons speak simultaneously. BSS has been intensively studied in the last 20 years and many techniques based on elements from information theory have been developed. The principle that lies behind these approaches is called independent component analysis and requires the statistical independence of each component of the output vector. Bell

and Sejnowski used the maximum entropy principle to maximize the joint differential entropy of the output signals, each transformed by a sigmoid function [5]. High order statistics may also be used to derive BSS algorithms, such as the FastICA algorithm that achieves separation by finding the local extrema of a cost function of kurtosis [1].

Other existing BSS criteria are special cases of the Kullback-Leibler divergence and Renyi's mutual entropy [3, 5, 12].

In this paper, a new objective function for the adaptive blind source separation algorithm is proposed. The separation process takes place in two phases: spatial whitening to produce uncorrelated but not independent mixture, followed by a series of Givens rotations in the mixture space to determine the independent components [3, 4]. The uncorrelated mixed signals are transformed by their respective distribution functions into uniform random variables. Then, the parameters of the Givens rotations are found adaptively to minimize the contingency coefficient from the chi-square independence test.

2. Formalized Problem

Let $s(k) = (s_1(k), s_2(k), ..., s_n(k))^T$ be the unknown vector of the source signals which are independent stationary zero-means processes and $x(k) = (x_1(k), x_2(k), ..., x_n(k))^T$ be the observations vector resulted from linear mixtures of sources:

$$x(k) = As(k), k = 0, 1, 2, ...$$

where $A \in R^{n \times n}$ is a non-singular unknown matrix (called the mixing matrix) (see Fig. 1). Blind source separation means to retrieve the original signals from observations $x(k)$ without having prior information about the sources or the mixing matrix, except the hypothesis of independence of the source signals. The demixing model is a linear transform:

$$y(k) = Wx(k), k = 0, 1, 2, ...$$

where $y(k) = (y_1(k), y_2(k), ..., y_n(k))^T$ is the output vector and $W \in R^{n \times n}$ is the demixing matrix that must be found.

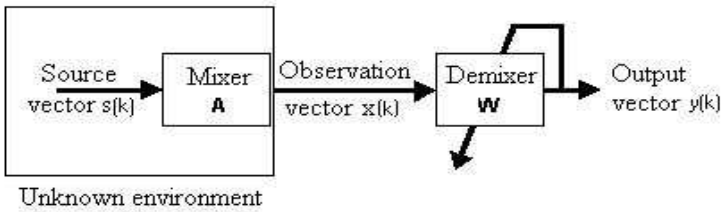


Fig. 1. Block diagram of processor for the BSS problem.

Below we use a well known statistical independence criteria between each component of $y(k)$ to solve the BSS problem.

3. The contingency coefficient

The contingency coefficient in the chi-square independence test is an independence measure for two random variables.

Theorem 1. *If X is a continuous random variable with the cumulative distribution function (CDF) F, then the random variable $F(X)$ is distributed uniformly on the interval $[0,1]$ [9].*

If $(U_k, V_k)_{k=1,N}$ is a sample of size N of two uniform variables on the domain $[0,1] \times [0,1]$ divided in $r \times c$ cells, we denote by n_{ij} the frequency of the appearance of (U_k, V_k) in the cell (i,j), where $i = 1, r$ and $j = 1, c$.

Let us consider the following $r \times c$ dimensional table:

$$\begin{array}{cccc|c} n_{11} & n_{12} & \cdots & n_{1c} & n_{1.} \\ n_{21} & n_{22} & \cdots & n_{2c} & n_{2.} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ n_{r1} & n_{r2} & \cdots & n_{rc} & n_{r.} \\ \hline n_{.1} & n_{.2} & \cdots & n_{.c} & N \end{array}$$

where $\sum_{k=1}^{c} n_{ik} = n_{i.}$, $\sum_{k=1}^{r} n_{kj} = n_{.j}$ and $\sum_{i,j} n_{ij} = N$.

If the above table corresponds to a sample of two independent random variables, then the frequency in the cell (i,j) is $n_{i.}n_{.j}/N$. The deviation from independence in the cell (i,j) is measured by:

$$D_{ij} = n_{ij} - n_{i.}n_{.j}/N.$$

The contingency coefficient is defined as:

$$\chi^2 = \sum_{i,j} \frac{D_{ij}^2}{n_{i.}n_{.j}/N} = \sum_{i,j} \frac{(n_{ij} - n_{i.}n_{.j}/N)^2}{n_{i.}n_{.j}/N} = N\left(\sum_{i,j} \frac{n_{ij}^2}{n_{i.}n_{.j}} - 1\right).$$

Theorem 2. *In the independence hypothesis χ^2 is asymptotically distributed as $\chi^2((r-1)(c-1))$ [6].*

So, the hypothesis of independence is accepted if $\chi^2 < \chi^2_{((r-1)(c-1)),\alpha}$, where $\chi^2_{((r-1)(c-1)),\alpha}$ is the $\alpha-$ upper quantile (α has a small value close to zero).

Now, if we consider two random continuous variables s_1 and s_2 with their cumulative distribution functions F_1 and F_2, by transforming s_1 and s_2 by their CDFs, we get two random variables:

$$z_1 = F_1(s_1), z_2 = F_2(s_2).$$

From Theorem 1 we have that $z_1, z_2 \in [0,1]$ and they are uniformly distributed. Also, we know the following Lemma [1]:

Lemma 1. *z_1 and z_2 are independent iff s_1 and s_2 are independent.*

Transforming a sample $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ of size n, of the random variables (s_1, s_2) by their CDFs and computing the corresponding contingency coefficient χ^2, it results from Theorem 2 and Lemma 1 that if $\chi^2 < \chi^2_{((r-1)(c-1)),\alpha}$ then the variables s_1 and s_2 are independent. The parameters of the model are found adaptively to minimize the contingency coefficient.

4. The BSS algorithm based on the minimization of the contingency coefficient

The criterion proposed for the BSS problem is the minimization of the contingency coefficient from the chi-square independence test. The separation process takes place in two phases: spatial whitening to produce uncorrelated but not independent mixture, followed by a series of Givens rotations in the n-dimensional mixture space to determine the independent components [3, 4]. Let $z = As$ be the observations vector. Assuming that s is zero mean, the mixture covariance matrix is $\Sigma = E[zz^T]$. If Λ is the diagonal eigenvalue matrix and Φ is the corresponding orthonormal eigenvector matrix of Σ, then $W = \Lambda^{-1/2}\Phi^T$ is the decorrelation matrix for z. The adaptive part of the model is the rotation matrix that produces the output vector $y = R(\theta)x$ (see Fig. 2).

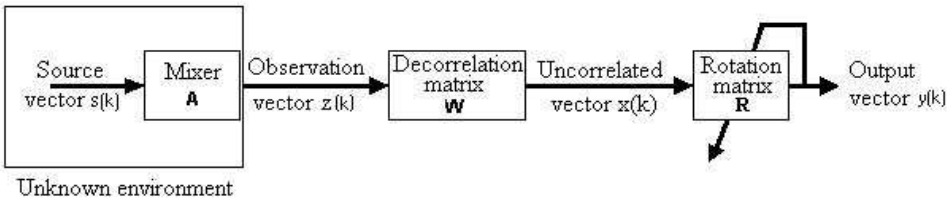


Fig. 2. System block diagram of processor for the BSS problem in two stages: decorrelation and rotation.

The BSS algorithm can be described as follows:

Step 1 - the decorrelation phase
The signals decorrelation matrix is $W = \Lambda^{-1/2}\Phi^T$, where Λ is the diagonal matrix of the eigenvalues and Φ is the orthogonal matrix of the eigenvectors corresponding to the covariance matrix $\Sigma = E[zz^T]$.
Thus we obtain the uncorrelated mixtures:

$$x = Wz.$$

Step 2 - the rotation phase
for each pair $(i, j), 1 \leq i < j \leq n$ find the rotation parameter $\theta_{ij} \in [0, 2\pi]$ which

minimizes the contingency coefficient χ^2_{ij} corresponding to the signals i and j. Minimization is done by a non-gradient optimization technique (for example, the golden section search) because the derivative of the criterion function χ^2_{ij} with respect to the rotation parameter θ_{ij} is not available. The vector x becomes

$$x = R_{ij}(\theta_{ij})x,$$

where the rotation matrix $R_{ij}(\theta_{ij})$ is defined as an identity matrix whose (i,i)th, (i,j)th, (j,i)th and (j,j)th positions are cosθ_{ij}, -sinθ_{ij}, sinθ_{ij} and cosθ_{ij}, respectively.
Go to Step 2 until the convergence of the algorithm is reached.

Step 3 - the output vector $y = x$, STOP.

The convergence of this iterative scheme has not been demonstrated mathematically, but based on many authors'experience [1, 2, 7, 11], as well as by running the algorithm presented above, the convergence is reached in practice. The complexity of the algorithm that minimizes the contingency coefficient is $O(n^2NlogN)$, because the step 2 takes $O(n^2)$, and the rotation matrix $R_{ij}(\theta_{ij})$ (that requires the computation of the CDF for each signal) can be found in $O(NlogN)$ if we sort ascending the values of the signal with quicksort.

5. Simulations

To test the algorithm presented above, we used both Romanian and English speech signals. The Romanian speech signals were taken from [8] and the English speech signals were taken from the TIMIT speech database. The WAV files from these databases were processed with the HTK software [10] to obtain the input files for the BSS algorithm. We have adjusted the signals' amplitude to be in $(0,1)$ and to be zero-mean. The mixture matrix we used is

$$A = \begin{pmatrix} 0.56 & 0.79 & -0.37 \\ -0.75 & 0.65 & 0.86 \\ 0.17 & 0.32 & -0.48 \end{pmatrix}.$$

To evaluate the algorithm, we define a performance measure as [5]

$$P = \sum_{i=1}^{n}\left(\sum_{j=1}^{n}\frac{|q_{ij}|}{\max_k|q_{ik}|} - 1\right) + \sum_{j=1}^{n}\left(\sum_{i=1}^{n}\frac{|q_{ji}|}{\max_k|q_{ki}|} - 1\right),$$

where $Q = (q_{ij})_{i,j=1,n}$ is a n-dimensional square matrix. A monomial matrix is defined as a square matrix having one and only one non-zero element per row and per column. $P = 0$ if Q is a monomial matrix. For the BSS algorithm based on the minimization of the contingency coefficient, the matrix $Q = RWA$. The smaller the value of P, the better the performance of the algorithm. Some results obtained by running this algorithm are presented below.

The signals shown in Fig. 3 are speech signals spoken by three different persons in Romanian language, each person uttering different words. For these three signals grouped two by two, the contingency coefficients are 114.36, 212.49 and 281.28 respectively. But the $\alpha-$ upper quantile $\chi^2_{81,0.05} = 103.0095$ and by applying the independence test we conclude that the signals do not meet the statistical independence assumption. So, even if the speech signals are spoken by different persons, the independence assumption is not plausible. In spite of the dependence among speech signals, the results obtained with the proposed BSS algorithm are reliable (see Fig. 3 and Table 1).

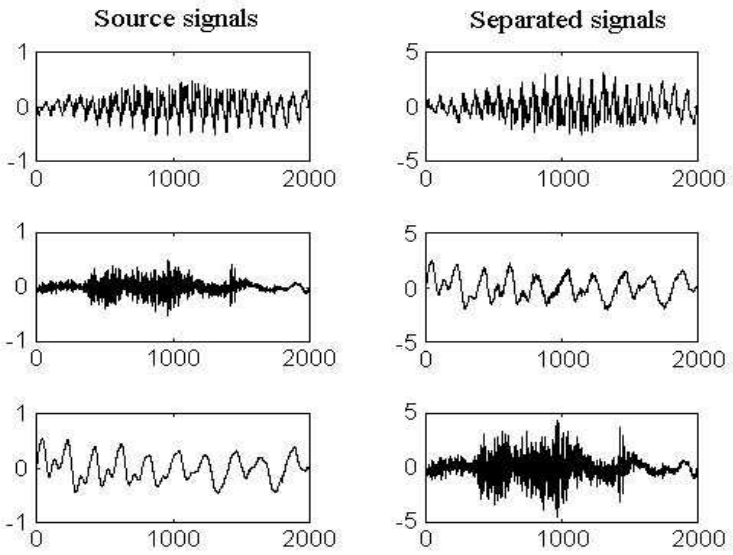


Fig. 3. Source signals and separated signals resulted by applying the algorithm that minimizes the contingency coefficient.

Table 1 shows the product of the mixing and the demixing matrices, so that the connection between the monomial form of this product matrix (column 2) and the performance index (column 3) can be noticed.

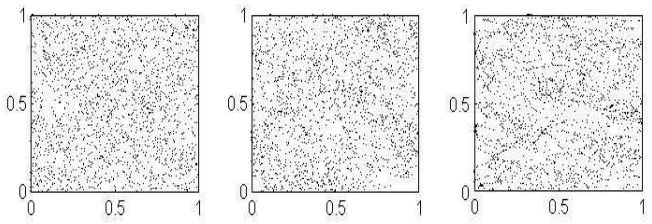
In Figs. 4–7 the same speech signals grouped by twos, transformed into random uniform variables on the interval $[0, 1]$ in different stages of the BSS algorithm are displayed. In Fig. 4 there are the initial signals before mixture, Fig. 5 presents the observed mixed signals, in Fig. 6 there are the signals after the decorrelation phase and the demixed signals are shown in Fig. 7. The domain $[0, 1] \times [0, 1]$ is divided in 100 cells ($r = 10$ and $c = 10$).

To prove the effectiveness and performance of the proposed BSS algorithm based on the minimization of the contingency coefficient, we have also tested this algorithm on other two types of signals, including independent and identically distributed signals uniformly distributed in $[0, 1]$ and exponential signals (with parameter $\lambda = 1$). In each case, we have simulated three signals with the same distribution, we have transformed

them so that to be zero mean and we have linearly mixed them with the same mixture matrix A. The result obtained can be seen in Table 2. Comparing the performance indexes shown in Table 1 and in Table 2, it can be observed, as it was expected, that this BSS algorithm gives better results when the signals are statistically independent from one another.

Table 1. The results obtained by applying the algorithm that minimizes the contingency coefficient

Signal duration	$Q = RWA$	Performance index P
500	$\begin{pmatrix} -0.3222 & 0.3950 & 3.7903 \\ 4.7662 & 0.0601 & 0.2614 \\ -0.1585 & 6.0064 & -0.4258 \end{pmatrix}$	0.7119
1000	$\begin{pmatrix} -0.8218 & 6.9562 & -0.2024 \\ 4.7331 & 0.9303 & -0.5729 \\ -0.6165 & 0.1688 & -4.0575 \end{pmatrix}$	1.3113
2000	$\begin{pmatrix} -5.4271 & 0.3526 & -0.5407 \\ -0.8228 & -0.5853 & 4.2817 \\ -0.0563 & 8.8843 & -0.0475 \end{pmatrix}$	0.9101 (see Fig. 3)



(a) (b) (c)

Fig. 4. The initial signals transformed in random uniform variable on the interval $(0, 1)$ ((a) the signals 1 and 2, (b) the signals 2 and 3, (c) the signals 1 and 3).

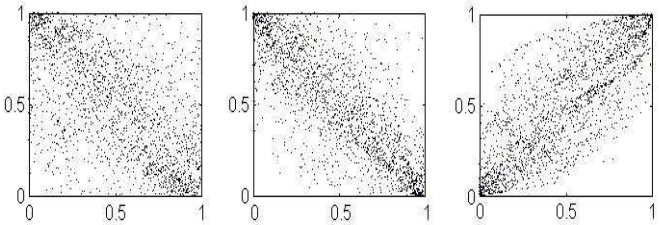


Fig. 5. The observed (mixed) signals.

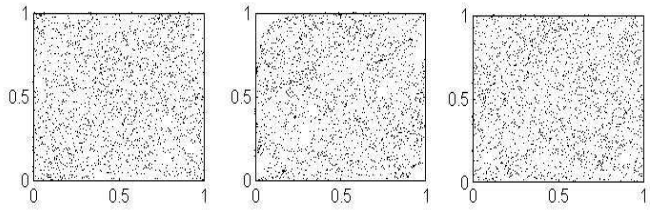


Fig. 6. The decorrelated signals.

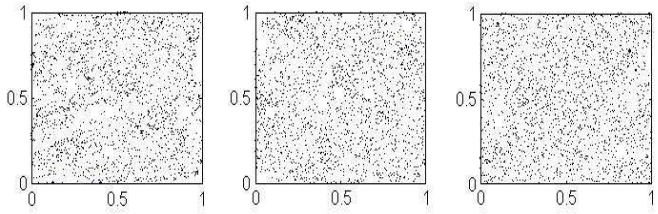


Fig. 7. The demixed signals.

Table 2. The results obtained by applying the algorithm that minimizes the contingency coefficient on independent uniform and exponential distributed signals

Signal duration	Performance index P for the uniform distributed signals	Performance index P for the exponential distributed signals
500	1.2980	0.5587
1000	0.4915	0.3735
2000	0.2115	0.1905

We have also linearly mixed two independent signals: a speech signal and a white Gaussian noise (see Fig. 8). The performance index obtained was $P = 0.1196$, which means that the separation of the signals was very good.

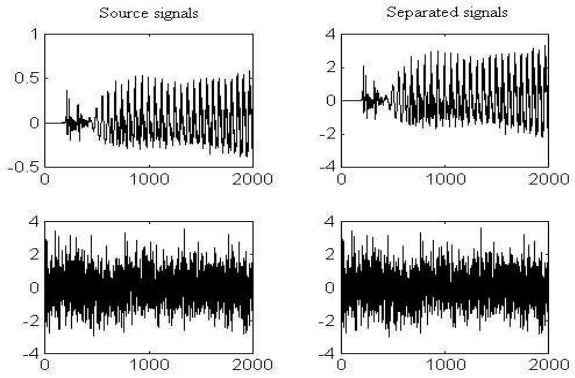


Fig. 8. A speech signal and a white Gaussian noise.

In the last simulation we present, we have tested another BSS algorithm based on the same scheme: decorrelation followed by a series of Givens rotations in the mixture space. The parameters of this model are found adaptively to minimize the Renyi's mutual entropy [3]. In Table 3 we present the results obtained by testing this algorithm on the speech signals showed in Fig. 3. Comparing the results from Table 1 and Table 3, it can be noticed that the algorithm that minimizes the contingency coefficient achieved a better performance than the algorithm that is based on minimization of the Renyi's mutual entropy. But testing these two algorithms using other speech signals, we can see that in some cases the results are better for the algorithm that minimizes the Renyi's mutual entropy (Table 4). In conclusion, there are speech signals for which the algorithm that minimizes the contingency coefficient gives better results and there are speech signals that are better separated by applying the algorithm that minimizes the Renyi's mutual entropy.

Table 3. The results obtained by applying the algorithm that minimizes the Renyi's mutual entropy using the speech signals showed in Fig. 3

Signal duration	Performance index P
500	6.28
1000	4.26
2000	4.04

Table 4. The results obtained by applying the algorithm that minimizes the contingency coefficient (column 2) and by applying the algorithm that minimizes the Renyi's mutual entropy (column 3)

Signal duration	Performance index P (contingency coefficient)	Performance index P (Renyi's mutual entropy)
500	4.71	5.41
1000	4.42	0.54
2000	0.93	2.04

6. Conclusions

In this paper we present a new BSS algorithm based on the minimization of the contingency coefficient and the results obtained by testing it on different types of signals. By applying the independence test for the speech signals we used, we conclude that they are not statistically independent from one another, although the signals were spoken by different persons uttering different words. Even if the BSS algorithm uses the independence assumption, the results obtained for demixing speech signals are reliable. For the other types of signals we tested (independent and identically

distributed signals), the BSS algorithm achieved a better performance because these signals were generated independently. It is clear that in real-life BSS problems, the independence hypothesis is not always acceptable, and therefore it is necessary to develop new BSS techniques that relax the statistical independence assumption of the mixed signals.

References

[1] CHEN Y., *Blind Separation Using Convex Functions*, IEEE Transactions on Signal Processing, **53** , 6, 2005, pp. 2027–2035.

[2] COMON P., *Independent Component Analysis: A new concept?*, Signal Processing, **36**, 1994, pp. 287–314.

[3] ERDOGMUS D., HILD K. E., PRINCIPE J. C., *Blind source separation using Renyi's α-marginal entropies*, Neurocomputing, **49**, 2002, pp. 25–38.

[4] ERDOGMUS D., RAO Y. N., PRINCIPE J. C., ZHAO J., HILD K. E., *Simultaneous Extraction of Principal Components Using Givens Rotations and Output Variances*, Proceedings of ICASSP'02, **1**, Orlando, Florida, 2002, pp. 1069–1072.

[5] HAYKIN S., *Neural Networks-A Comprehensive Foundation*, Prentice Hall, Upper Saddle River, New Jersey, 1999.

[6] MIHOC GH., CRAIU V., *Tratat de statistică matematică, Volumul 2. Verificarea ipotezelor statistice*, Editura Academiei Republicii Socialiste Romania, 1977, Bucureşti.

[7] ORSAK G. C., DOUGLAS S. C., *Code-length-based universal extraction for blind signal separation*, Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., Istanbul, Turkey, 2000, pp. 3422–3425.

[8] TEODORESCU H. N., FERARU M., TRANDABĂŢ D., ZBANCIOC M., LUCA R., VERBUŢĂ A., HNATIUC M., GANEA R., VORONEANU O., PISTOL L., *Proiectul Sunetele Limbii Române*, Retrieved February, 2009, from http://www.etc.tuiasi.ro/sibm/romanian_spoken_language/ro/ton_neutru.htm.

[9] VĂDUVA I., *Modele de simulare − Note de curs*, Editura Universitaţii din Bucureşti, 2004.

[10] YOUNG S., EVERMANN G., GALES M., HAIN T., KERSHAW D., MOORE G., ODELL J., OLLASON D., POVEY D., VALTCHEV V., WOODLAND P., *The HTK Book*, Retrieved February, 2005, from http://htk.eng.cam.ac.uk/docs/docs.shtml.

[11] ZARZOSO V., NANDI A. K., *Adaptive blind source separation for virtually any source probability density function*, IEEE Trans. Signal Process., **48**, 2, 2000, pp. 477–488.

[12] ZHANG L., CICHOCKI A., AMARI A., *Self Adaptive Blind Source Separation Based on Activation Function Adaptation*, IEEE Transactions on Neural Networks, **15**, 2, 2004.