Blind Separation of Underdetermined Mixtures With Additive White and Pink Noises

Ossama S. Alshabrawy[1,3], Aboul ella Hassanien[2,3], W. A. Awad[4], A. A. Salama[4]
[1]Dept. of Mathematics & Computer Science, Faculty of Science, Damietta University, Damietta, Egypt
[2]Faculty of Computers & Information, Cairo University, Cairo, Egypt
[3]Scientific Research Group in Egypt (SRGE), http://www.egyptscience.net
[4]Dept. of Mathematics & Computer Science, Faculty of Science, Port Said University, Egypt

***Abstract*—This paper presents an approach for underdetermined blind source separation in the case of additive Gaussian white noise and pink noise. Likewise, the proposed approach is applicable in the case of separating $I+3$ sources from I mixtures with additive two kinds of noises. This situation is more challenging and suitable to practical real world problems. Moreover, unlike to some conventional approaches, the sparsity conditions are not imposed. Firstly, the mixing matrix is estimated based on an algorithm that combines short time Fourier transform and rough-fuzzy clustering. Then, the mixed signals are normalized and the source signals are recovered using modified Gradient descent Local Hierarchical Alternating Least Squares Algorithm exploiting the mixing matrix obtained from the previous step as an input and initialized by multiplicative algorithm for matrix factorization based on alpha divergence. The experiments and simulation results show that the proposed approach can separate $I+3$ source signals from I mixed signals, and it has superior evaluation performance compared to some conventional approaches.**

***Keywords*-Underdetermined Blind Source Separation; Rough Fuzzy clustering; Short Time Fourier transform; Hierarchical Alternating Least Squares**

I. INTRODUCTION

Blind Source Separation (or Blind Signal Separation, BSS) in combination with information theory, artificial neural networks, and computer science applications has a wide range of applications in the fields of digital communication systems, wireless communications, speech processing, feature extraction, speech processing, medical imaging, water marking, biomedical engineering, and data mining [4]–[8] in the last decade. Blind separation or blindness means that no or very little information is known about the mixing system or the original source signals that need to be extracted [1].

The main goal of BSS algorithms is to estimate or extract original source signals using only the information gathered from observable mixed signals without or with very limited knowledge about the source signals or the mixing system. The approaches developed by researchers in the last two decades can be classified into two methodologies, namely overdetermined BSS and underdetermined BSS, according to the number of source signals and observable mixed signals [12]. If the number of sensors or mixed signals is less than the number of source signals, then the problem is called underdetermined BSS while if the number of sensors or mixed signals is greater than or equal the number of source signals, then the problem is called over-determined BSS. Underdetermined BSS is challenging and is more realistic to most practical situations and real world problems. However, most approaches for BSS rarely involve underdetermined BSS cases. The classical independent component analysis (ICA) approach fails to solve underdetermined BSS problems [11]. Moreover, in many practical problems there are a large number of source signals but a few numbers of sensors that means the underdetermined case. Another major difficulty of ICA is that its ambiguities and that the order, sign, and the variances of the independent components cannot be determined, therefore the mixing matrix and the magnitude of original source signals cannot be estimated [3].

Matrix factorization is very important and unifying topic that has a great deal of attention in signal processing, and linear algebra and has found numerous applications in many other areas [17]. A special case of matrix factorization is Nonnegative Matrix Factorization (NMF) which has the non-negativity constraints. Recently, NMF has been widely applied to many BSS problems. However, the separation results are sensitive to the initialization of parameters. Another major drawback of NMF is that the additive parts by NMF are not necessarily localized consequently, the solution is not unique. Avoiding the subjectivity of choosing parameters, we use general matrix factorization (GMF), which completely relaxes the non-negativity constraints from its factors with matrix factorization multiplicative algorithm as an initialization to the source signals instead of random initial values. GMF is a generalization of the well-known NMF where the NMF is constrained by non-negativity on all its factors, is not necessarily localize, has low convergence and, does not provide a unique solution in some cases without additive constraints and parameters. However, GMF has no constraints of non-negativity and is fast convergent with the ALS method used for initialization and improvement [2].

Most of the conventional BSS approaches assume that the source signals are as statistically independent as possible given the sensors data. Another hypothesis by these approaches is that the mixing matrix is of full column rank.

In many real-world situations, however, this hypothesis is not valid. Consequently, recovering the source signals by multiplying the observable data mixtures by the pseudo inverse of the mixing matrix cannot be used. This makes recovering the source signals a difficult and very challenging task [9]. In practical terms, the overdetermined mixture assumption does not always hold. For instance, in radio communications the probability of receiving more source signals than sensors data or observed mixed signals increases with increase of reception bandwidth, hence it is urgently necessary to solve the problem of underdetermined blind source separation (UBSS) [10].

The motivation of this paper is to separate sparse, and super and sub-Gaussian signals in the underdetermined case with additive noise such as Gaussian white noise and pink noise without imposing any sparsity conditions. Another motivation of this paper and to increase the performance of the separation in case of noisy mixtures. The rest of the paper is organized as follows. Section II formulates the problem. In Section III, we present the details of the proposed approach. In section IV, we show the analysis of typical experiments and the results obtained by different BSS methods, where the simulation results show the effectiveness and high performance of the proposed algorithm. Finally, a short conclusion and future work are presented in Section V.

II. Problem formulation

The problem considered in this paper is an underdetermined instantaneous BSS with additive white Gaussian background noise and pink noise, which can be mathematically formulated as follows: Assume that for I unobservable components $X(t) = tr[X_1(t), X_2(t), ..., X_J(t)]$, where J is the number of source signals, and $X(t)$ is a zero-mean vector. The available sensor vector $Y(t) = tr[Y_1(t), Y_2(t), ..., Y_I(t)]$, where I is the number of sensors and tr is the transpose of the vector, is given by

$$Y(t) = AX(t) + E(t). \tag{1}$$

Here $A \in R^{I\times J}$ is unobservable matrix, and the rank of A is I. $X \in R^{J\times T}$, $Y \in R^{I\times T}$, $t = 0, ..., T-1$ are the sampling instant time points. Moreover, The vector $E(t)$ represents the noise which will be in this paper two kinds of noise throughout the experiments.

III. Proposed UBSS algorithm

In this section, the proposed approach is presented starting with estimating the mixing matrix knowing only the observable mixtures matrix which contains noise. Also, a method for GMF gradient descent based update rules initialized with matrix factorization multiplicative algorithm based on alpha divergence is introduced.

A. Mixing matrix estimation based on short time Fourier transform and rough fuzzy clustering

Conventional algorithms estimate the mixing matrix based on clustering algorithms such as the k-means algorithm require that the source signals to be very sparse in the time domain and this is unavailable in many practical real world problems. Other algorithms are based on an assumption that there exist many TF points of single source occupancy (SSO), or require that there exists at least one small region in the TF plain with only a single source and such a TF region must exist for each source. All aforementioned approaches require that for each source there exist many TF points of SSO. However, single source detection (SSD) requires that there exists at least one TF point of SSO and is hence less restrictive than the other approaches [13]. The short time Fourier transform (STFT) of the i^{th} observed signal is defined by the following equation:

$$Y_i^{Fourier}(t, r) = \sum_{l=0}^{\infty} h(l - t)X_i(l)e^{-jrl} \tag{2}$$

at frame t and frequency bin r where $h(l)$ is a window sequence. In equation (2), $i = 1, 2, , I$; $t = 0, 1, , T - 1$ are the sampling points over the time domain and $r = 0, 1, , T - 1$ are the sampling points over the frequency domain. The SSD is based on the ratio of the TF transforms and finds a set of TF points where a single source is active for each source. Therefore, for a given $\varepsilon > 0$ the set that represents the detected points can be obtained by the following equation:

$$\chi_F = \{(t, r)| \left\| Im[\frac{Y^{Fo}(t, r)}{Y_1^{Fo}(t, r)}] \right\|_F < \varepsilon, \; Y_1^{Fo}(t, r) \neq 0\} \tag{3}$$

where, Y^{Fo} represents the matrix of $Y^{Fourier}$ obtained from Eq. (2), Im[.] denotes the imaginary part. We can choose any of the mixture instead of Y_1.

After clustering, the ith column vector of A, denoted as $\hat{a}_i$, is estimated as

$$\hat{a}_i = \frac{1}{|\chi_{C_i}|} \sum_{(t,r)\in\chi_{C_i}} Re[Y^{Fo}(t, r)] \tag{4}$$

Here, χ_{C_i} represents the number of TF points in cluster C_i for $i = 1, 2, , J$.

B. The initialization technique for source signlas estimation

The source signals estimation algorithm will be initialized using multiplicative alpha GMF algorithm. This initialization technique will help to improve the results and obtain a better separation performance. The cost function for alpha divergence is outlined in the following equation:

$$COST^{\alpha}(Y||AX) = \frac{1}{\alpha(\alpha - 1)} \sum_{it} \left([Y]_{it}[AX]_{it}^{1-\alpha} - \alpha Y_{it} + (\alpha - 1)[AX]_{it} \right) \tag{5}$$

The final alpha multiplicative learning algorithm is expressed by the following update rules:

$$(x_{jt})_{new} = (x_{jt})_{old}\left(\frac{\sum_{i=1}^{I} a_{ij}(y_{it}/[AX]_{it})^{\alpha}}{\sum_{i=1}^{I} a_i j}\right)^{\frac{1}{\alpha}},$$
$$(a_{ij})_{new} = (a_{ij})_{old}\left(\frac{\sum_{t=1}^{T}(y_{it}/[AX]_{it})^{\alpha} x_{jt}}{\sum_{t=1}^{T} x_j t}\right)^{\frac{1}{\alpha}}; \alpha \neq 0 \quad (6)$$

But here we will make initialization to only the source signals, so we will exploit only the first part of the equation.

C. Modified gradient descent local Hierarchical alternating least squares

This section will introduce a quick overview on the analysis and derivation of Hierarchical Alternating Least Squares (HALS). Hierarchical Alternating Least Squares method is suitable for large-scale NMF problems, and it can be applied also for sparse non-negative coding or representation [14]. HALS algorithm can be derived by choosing exploiting a set of local cost functions such as Alpha- and Beta-divergences, and the squared Euclidean distance. Then perform consecutive or simultaneous minimization of these local cost functions. For example, using gradient descent or some nonlinear transformations. The family of HALS algorithms can not only do better for the over-determined case of BSS, but they can also solve underdetermined BSS case under some simple conditions. Especially for the multi-layer technique [15], the extensive experiments and simulation results show the superior performance and validity of the family of HALS algorithms. HALS is used here in this paper in a modified version by relaxing the non-negativity constraints and depending on the gradient descent algorithm.

Denote $A = [a1, a2, ..., a_j]$ and $S = X^T = [s1, s2, ..., s_j]$ to express the squared Euclidean cost function as

$$J[a_1, .., a_j, s1, .., s_j] = \frac{1}{2}\|Y - AS^T\|_f^2 = \frac{1}{2}\|Y - \sum_{j=1}^{J} a_j s_j^T\|_{Fro}^2 \quad (7)$$

where, *Fro* refers to the Frobenious norm. The main idea is to define the residues followed by minimizing the set of local cost functions alternatively with respect to the parameters a_i and s_j. The residues can be obtained as:

$$Y^(j) = Y - \sum_{p \neq J} a_p s_p^T = Y - AS^T + a_j s_P^T = E + a_j s_j^T (j = [1, 2, ..., J]) \quad (8)$$

Then the alternative minimization of the set of cost functions can be obtained by [16]

$$Cost_{Fro}^{(j)}(Y^{(j)}||a_j s_j^T) = \frac{1}{2}\|Y(j) - a_j s_j^T\|_{Fro}^2, \quad Forj = 1, 2, ..., J \quad (9)$$

The optimality conditions for the set of cost functions (9) can be defined as

$$a_j \otimes \nabla_{a_j} Cost_{Fro}^{(j)}(Y^(j)||a_j s_j^T) = 0 \quad (10)$$

$$s_j \otimes \nabla_{s_j} Cost_{Fro}^{(j)}(Y^(j)||a_j s_j^T) = 0 \quad (11)$$

The gradients of the local cost functions in Eq. (9) are computed with respect to the unknown vectors a_j and s_j to obtain the critical or stationary points with the assumption that the other vectors are fixed by the following equation:

$$\nabla_{a_j} Cost_{Fro}^{(j)}(Y^(j)||a_j s_j^T) = \frac{\partial Cost_{Fro}^{(j)}(Y^(j)\|a_j s_j^T)}{\partial a_j} = a_j s_j^T s_j - Y^(j) s_j \quad (12)$$

$$\nabla_{s_j} Cost_{Fro}^{(}j)(Y^(j)||a_j s_j^T) = \frac{\partial Cost_{Fro}^{(j)}(Y^(j)||a_j s_j^T)}{\partial s_j} = a_j^T a_j s_j - Y^{(j)T} a_j \quad (13)$$

Without resorting to any non-negativity constraints on the entries of vectors a_j and s_j $\forall j$, the critical points can be obtained by the following simple update rules:

$$s_j \leftarrow \frac{1}{a_j^T a_j}(Y^{(j)T} a_j) = \frac{1}{a_j^T a_j} Y^{(j)T} a_j. \quad (14)$$

$$a_j \leftarrow \frac{1}{s_j^T s_j}(Y^(j) s_j) = \frac{1}{s_j^T s_j} Y^{(j)} s_j, (j = 1, 2, ..., J) \quad (15)$$

The algorithm of Modified gradient descent local Hierarchical alternating least squares is stated below:

Algorithm 1 Modified gradient descent local Hierarchical alternating least squares algorithm

Input: The observable matrix $Y_{\pm}$, the number of source signals J

Output: The source signals $X = S^T \in R_{\pm}^{J \times T}$ such that the cost function in Eq. 9 is minimized

Intitialize the source signals $X = S^T$ by multiplicative alpha GMF algorithm by Eq. 6;
set $E = Y - AS^T$;
repeat
  for k=1 **to** J **do**
    $Y^{(k)} \leftarrow E + a_k s_k^T$
    $s_k \leftarrow Y^{(k)T} a_k$
    $E \leftarrow Y^{(k)} - a_k s_k^T$
  end for
until a stopping criterion is met

IV. Experiments and Simulation Results

In this section, the performance and effectiveness of the proposed approach will be discussed by comparing results of experiments and stimulations. Experiments and simulations were performed on synthetically generated signals using the proposed approach and some other approaches. In the simulations, sparse, super- and sub-Gaussian signals were separated from the underdetermined noisy mixtures in the challenging case where the true number of source signals is unknown. The types of noise that is considered in this paper is the white noise and pink noise.

The parameter inputs of the modified Modified gradient descent local Hierarchical alternating least squares algorithm are the observable mixtures matrix Y, and the mixing matrix A obtained by the method stated above. We choose the maximum number of iterations to be only 50 iterations. We investigate the performance of the proposed UBSS approach in the above mentioned cases by comparing its results with the results of approaches in Snoussi and Idier (2006) [18], Peng and Xiang (2010) [19], and S. Sun et al. (2012) [20]. Here, the simulation of the separation of a variety of sparse, non-sparse, and super- and sub-Gaussian signals are stated.

A. Separation of synthetic signals with additive noise

Here, the simulation of the separation of a variety of sparse, non-sparse, and super- and sub-Gaussian signals are stated. All these cases are in the presence of above mentioned kinds of noise.

1) Sparse, non-sparse, and super- and sub-Gaussian signals: The effectiveness of the proposed UBSS approach is investigated by comparing the results of the proposed approach with the methods mentioned above. We chose the number of mixtures to be only 2 and the number of sources to be 5 to create a more challenging case and to prove that the proposed approach can separate $I + 3$ source signals from I mixtures. The five source signals, two observable mixtures that contains additive white noise, and pink noise, and The estimated source signals are plotted in Figs. 1, 2, 3, and 4. The number of sampling time points is 10,000. The simulation results of the proposed approach in addition to those of the five different UBSS methods are shown in Figs. 5, 6, 7. The performance of the source recovery method can be evaluated by Eqs. (16) and (17).

$$SIR = -10\log\left(\frac{\left\|\hat{X}_i - X_i\right\|_{Fro}^2}{\|X_i\|_{Fro}^2}\right), i = 1, 2, ..., J \quad (16)$$

$$SNR = \frac{1}{J}\sum_i 10\log\left(\frac{\|X_i\|_{Fro}^2}{\left\|\hat{X}_i - X_i\right\|_{Fro}^2}\right), i = 1, 2, ..., J \quad (17)$$

where, J is the number of source signals. The efficiency of the separation results is good when SNR > 25 [21]

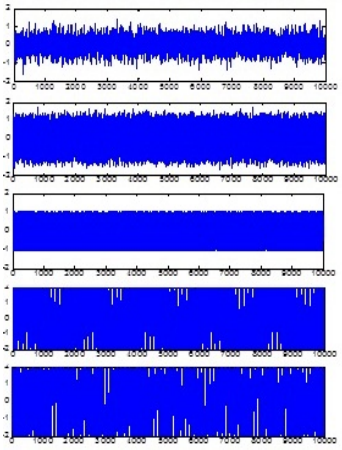

Figure 1. The source signals

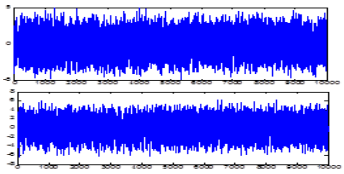

Figure 2. The mixed signals with additive white noise

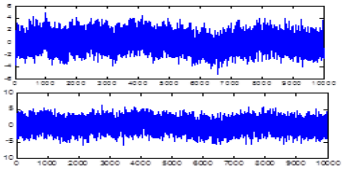

Figure 3. The mixed signals with additive pink noise

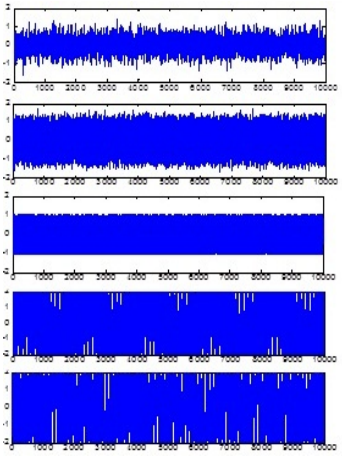

Figure 4. The estimated source signals

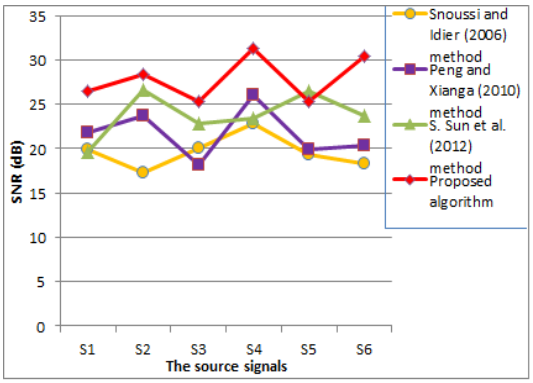

Figure 5. Performance estimation of the source signals from 3 observable mixtures with additive white noise

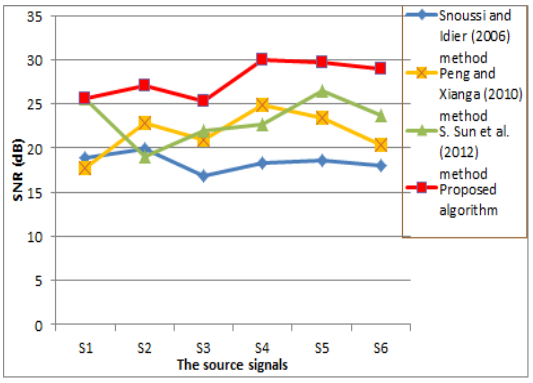

Figure 6. Performance estimation of the source signals from 3 observable mixtures with additive pink noise

We note from Fig. 5 that the proposed approach achieves about 4 dB higher SNR for J=7 sources with only two mixtures than the highest performance algorithm among the other five approaches. Likewise, the proposed approach achieves higher performance in case of Pink noise. Another comparison of the proposed approach with the other three approaches is presented using the SNR index for each kind of noise and is demonestrated in the following Fig. 7

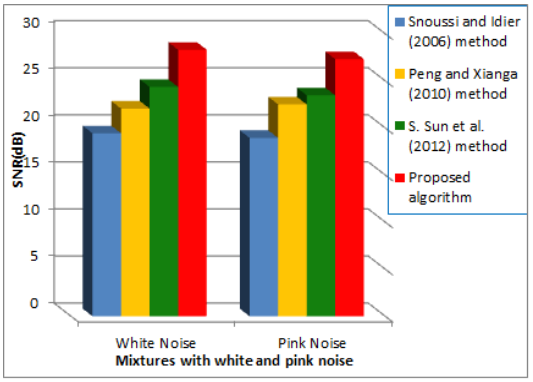

Figure 7. SNR index for the source signals from 3 observable mixtures with white and noises

From the results in Figs. 1, 2, 3, 4, 5, 6, 7, we can conclude that the separation performance of the proposed approach is very high, has faster convergence, and can separate sparse, non-sparse, and super- and sub-Gaussian signals in addition to sparse and non-sparse signals when compared with the other approaches.

B. Separation of real-world signals

To further measure the estimation performance of the source recovery approach in the previous Subsection, other comparisons are performed with the three other approaches using a dataset of real-world signals that are available to download from http://www.bsp.brain.riken.jp/ICALAB/ICALABSignalProc/benchmarks. This benchmark is EEG19, which contains 19 electroencephalogram (EEG) signals with clear heart, eye movement, and eye blinking artifacts. Only five signals are chosen as shown in Fig. 8. Note that the first four signals X_1, X_2, and X_3 are super-Gaussian while the last two signals X_4 and X_5 are sub-Gaussian. Fig. 9, Fig. 10 shows the observable mixtures signals and Fig. 11 shows the estimated source signals. The mixing matrix A is the same as in the previous experiments. Figs. 12, 13 shows the SIR results for all methods tested.

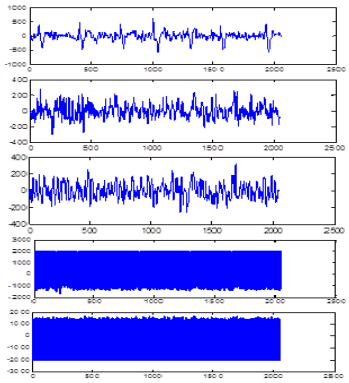

Figure 8. The source signals (EEG)

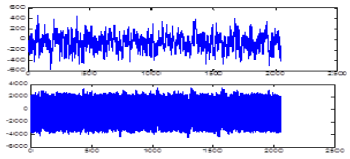

Figure 9. The mixed signals with additive white noise (EEG)

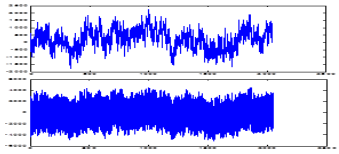

Figure 10. The mixed signals with additive pink noise (EEG)

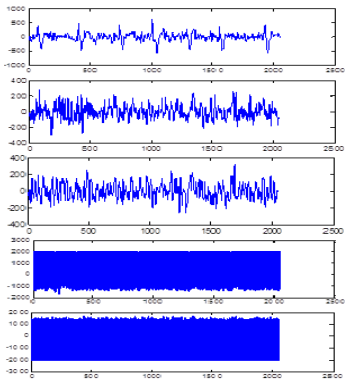

Figure 11. The estimated source signals

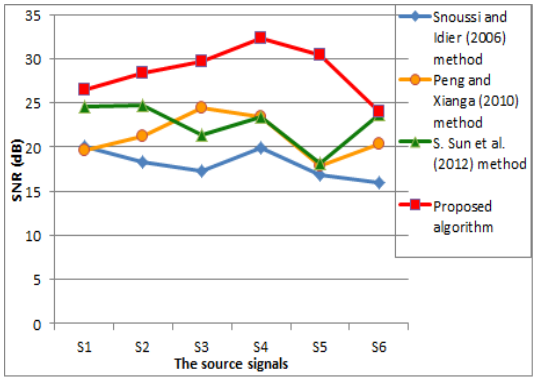

Figure 12. Performance estimation of the seven EEG19 source signals from 3 observable mixtures with additive white noise

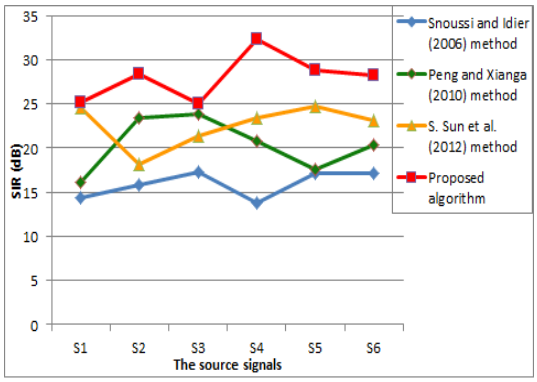

Figure 13. Performance estimation of the seven EEG19 source signals from 3 observable mixtures with additive pink noise

V. Conclusion

In this paper, we addressed the problem of underdetermined blind source separation with the challenging case that to separate I+3 source signals from I mixtures with addtive white and pink noises. A new two-step approach for optimum estimation of the source signals. In this approach, STFT is combined with rough fuzzy c-means clustering to estimate the mixing matrix. Then the source signals are estimated by a modified gradient descent local Hierarchical alternating least squares based general matrix factorization. Simulation experiments demonstrated the validity and superior performance of the proposed approach.

Acknowledgment

The authors would like to thank all SRGE members who contribute in this paper. Special thanks are expressed to Dr. Nashwa El-Bendary for her untold contributions, helping me through this work and her continuous cooperation.

References

[1] Ossama S. Alshabrawy, M. E. Ghoniem, W. A. Awad and Aboul ella hassanien, Underdetermined Blind Source Separation based on Fuzzy C-Means and Semi-Nonnegative Matrix Factorization. IEEE Federated Conference on Computer Science and Iinformation Systems (FedCSIS), Wroclaw, Poland, pp. 695-700, 9-12 September, 2012

[2] Ossama S. Alshabrawy, M. E. Ghoniem, A. A. Salama, Aboul ella hassanien, Underdetermined Blind Separation of an Unknown Number of Sources Based on Fourier Transform and Matrix Factorization, IEEE Federated Conference on Computer Science and Iinformation Systems (FedCSIS), Krakw, Poland, pp. 19-25, 8 - 11 September, 2013

[3] A. Hyvarinen, J. Karhunen, E. Oja, Independent Component Analysis, Wiley, New York, 2001.

[4] Yadong, Liu, Zongtan, Zhou, Dewen, Hu. A novel method for spatio temporal pattern analysis of brain fMRI data, Science in China Series F: Information Sciences, vol. 48, no. 2, pp. 151160, 2005.

[5] Araki, S., Makino, S., Blin, A., Underdetermined blind separation for speech in real environment with sparseness and ICA, In Proceedings of the ICASSP04, Montreal, Canada, pp.881884, 2004.

[6] Ohnishi, Naoya, Imiya, Atsushi, Independent component analysis of optical flow for robot navigation, Neurocomputing, vol.7, nos. 1012, pp. 21402163, 2008.

[7] Tonazzini, Anna, Bedini, Luigi, Salerno, Emanuele, A Markov model for blind image separation by a mean-field EM algorithm, IEEE Transactions on Image Processing, vol. 15, no.2, pp.473482, 2005.

[8] Er-Wei, Bai, QingYu, Li, Zhiyong ,Zhang, Blind source separation channel equalization of nonlinear channels with binary inputs, IEEE Transactions on Signal Processing, vol. 53, no. 7, pp.23152323, 2005.

[9] Dezhong Peng, Yong Xiang, Underdetermined blind separation of nonsparse sources using spatial time-frequency distributions, Digital Signal Processing, vol. 20, pp. 581596, 2010.

[10] Fengbo Lu, Zhitao Huang,Wenli Jiang, Underdetermined blind separation of non-disjoint signals in timefrequency domain based on matrix diagonalization, Signal Processing, vol. 91, pp. 15681577, January 2011.

[11] Chaozhu Zhang, Cui Zheng, Underdetermined Blind Source Separation Based on Fuzzy C-Means Clustering and Sparse Representation, International Conference on Graphic and Image Processing (ICGIP), Proc. of SPIE vol. 8285, 2011.

[12] SangGyun Kim, Chang D. Yoo, Underdetermined Blind Source Separation Based on Subspace Representation, IEEE Transaction on Signal Processing, vol. 57, no. 7, July 2009.

[13] SangGyun Kim, Chang D. Yoo, Underdetermined Blind Source Separation Based on Subspace Representation, IEEE Transaction on Signal Processing, vol. 57, no. 7, July 2009.

[14] A. Cichocki, A.H. Phan, C. Caiafa, FlexibleHALSalgorithms for sparse non-negative matrix/tensor factorization, In Proc. of 18th IEEE workshops on Machine Learning for Signal Processing, Cancun, Mexico, 1619, 2008.

[15] A. Cichocki, R. Zdunek, and S.-I. Amari. Hierarchical ALS algorithms for nonnegative matrix and 3-D tensor factorization. Springer, Lecture Notes on Computer Science, LNCS-4666, pages 169176, 2007.

[16] Andrzej Cichocki, Rafal Zdunek, Anh Huy Phan, Shunichi Amari, Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation, John Wiley, 2009

[17] Mohammed E. Fathy, Ashraf Saad Hussein, M. F. Tolba, Fundamental matrix estimation: A study of error criteria, Pattern Recognition Letters 32(2): 383-391, 2011

[18] Hichem Snoussi, Jerome Idier. Bayesian blind separation of generalized hyperbolic processes in noisy and underdeterminate mixtures, IEEE Transactions on Signal Processing, vol. 54 no. 9, pp. 32573269, 2006.

[19] Dezhong Peng, Yong Xiang. Underdetermined blind separation of non-sparse sources using spatial time-frequency distributions, Digital Signal Processing, vol. 20, pp. 581596, 2010.

[20] Shijun Sun, Chenglin Peng, Wensheng Hou, Jun Zheng, Yingtao Jiang, Xiaolin Zheng. Blind source separation with time series variational Bayes expectation maximization algorithm, Digital Signal Processing, vol. 22, pp. 1733, 2012.

[21] Chaozhu Zhang, Cui Zheng, Underdetermined Blind Source Separation Based on Fuzzy C-Means Clustering and Sparse Representation, International Conference on Graphic and Image Processing (ICGIP), Proc. of SPIE vol. 8285, 2011.