ISSN No. 0976-5697

Volume 2, No. 1, Jan-Feb 2011

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Blind Signal Separation Algorithm with Independent Component Analysis (ICA) by Means of Neural Training: Design and Development with Newer Approaches

Ajoy Kumar Dey*
Department of Signal Processing,
Blekinge Tekniska Hogskola (BTH)
Karlskrona, Sweden
dey.ajoykumar@yahoo.com

Susmita Saha
Department of Signal Processing,
Blekinge Tekniska Hogskola (BTH)
Karlskrona, Sweden
dey.ajoykumar@yahoo.com

Abstract**—**Independent Component Analysis (ICA) and its mathematical ideas are presented for the problem of Blind Signal Separation (BSS) and multichannel blind deconvolution of independent source signals. BSS and ICA are emerging techniques that aspire to recover unobserved signals or sources from the observed mixtures. The aims of this paper are to review some new approaches and implement some new and unique idea regarding the problem of blind signal separation with ICA. Computer based simulations illustrate the performances of the developed algorithms.

Keywords- Blind source separation; Independent component analysis; Neural network; Learning algorithm.

I. INTRODUCTION

There are so many innovation and potential application of Neural networks to the higher level research of signal processing and its application and implementation as well. This paper mainly consists of the learning algorithms for blind signal separation with independent component analysis and deconvolution of signals. Here we consider a case in which a number of simultaneous observation of signals are available that are linear or nonlinear superposition's of separate independent signals from different sources [1]. In this kind of problem, we actually separating the waveforms of the original sources from any array of sensors without knowing transmission characteristics of the transmission channel and there is no assurance to get back the signals with the same waveforms like the source signals [3], [4].

In this paper, we focus on so many terms like independent component analysis which are commonly known as ICA, blind source separation, blind source deconvolution or blind estimation, Where Independent Component Analysis is a technique to extract the independent components from the observed multidimensional mixture of data. ICA is a signal processing technique whose goal is to express a set of random variables as linear combinations of statistically independent component variables [2], [5], [6]. By ICA, we can separate the original sources by their mixtures.

From the so many observed mixtures, BSS consists of recovering unobserved signals or sources. Typically the observations are obtained at the output of a set of sensors, where each sensor receives a different combination of the source signals. When the source signals are not observed and no information is available about the mixtures, then we called it blind signal. The lack of prior knowledge about the mixture is compensated by a statistically strong but often physically plausible assumption of independence between the source signals [7], [8]. The so-called blindness should not be understood negatively; the weakness of the prior information is precisely the strength of the BSS model, making it a versatile tool for exploiting the spatial diversity provided by an array of sensors. Promising applications can already be found in the processing of communications signals, in biomedical signals like ECG and EEG, geophysical exploration and image enhancement and recognition [1].

In this paper, we propose a general formulation of blind signal separation frame work for acoustic signals based on the Independent component analysis algorithm using the neural network platform. In particular, we emphasis some important issues: Algorithmically analysis and extraction efficiency regarding the BSS.

The paper organization is as follows: Section II describes the ICA algorithm and its implementation. Section III discusses the mixing process. Section IV describes about the Blind signal extraction and Section V shows a general case study. Finally, Section VI presents a conclusion and an indication towards the future scope of this work.

II. AN ALGORITHM OF ICA

Independent Component Analysis is a way to find the independent components of a multivariate random variables. These components are directions in which the elements of the random variables have no dependency. One ideal application of ICA is blind signal separation [2].
The Independent Component Analysis model of linear instantaneous mixture can be formulated as,

$$x(t) = A\,s(t) \qquad (1)$$

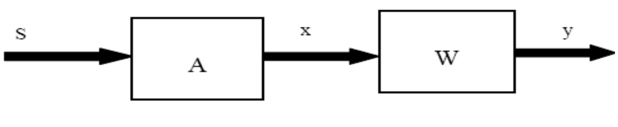


Figure 1. ICA Architecture.

Where A is an unknown mixing Vector, $x(t)$ and $s(t)$ are n dimension observed random vector and unknown source vector. The ICA model is a generative model, and it describes how the observed data are generated by a process of mixing the components. The independent components are

Latent variables that they cannot be directly observed, and the mixing matrix A is assumed to be unknown [17]. The aim of ICA is to recover the unobserved sources $s(t)$ from the observations $x(t)$.

ICA is very closely related to the method called blind source separation (BSS) or blind signal separation. Here Blind means that there is no specific idea about the mixing matrix A and few assumption about the source signals. ICA is the most popular and widely used for the performing the BSS. The blind source separation is to find a separating matrix W and to recover the source $s(t)$ by,

$$y(t) = Wx(t) \qquad (2)$$

Where, $y(t)$ is the estimation of source vector. We do not know the sources and the mixing process. In general, we assume that: the sources are independent from each other, the mixing matrix is nonsingular and at most one of the sources is Gaussian.

III. MIXING PROCESS

According to the problem statement and specific research application, there are many mathematical models and physical models that we can mixtures the input sources $s_j\,(t)\;(j = 1, 2, \dots, n)$. Now here we will focus on the simplest model where m observed signals $x_i(t)$ are instantaneous linear combination of the $n\;(m \geq n)$ unknown source signals, which are assumed in this paper to be zero mean and statistically independent. So now we have,

$$x_i\,(t) = \sum_{j=1}^{n} h_{ij} s_j(t) + n_i(t), \qquad (i = 1,2,\dots,m) \quad (3)$$

Or in the matrix notation,

$$x(t) = Hs(t) + n(t) \qquad (4)$$

Where, $x(t) = [x_1(t) \dots x_m(t)]^T$ is the sensor vector at discrete time t, $s(t) = [s_1(t) \dots s_n(t)]^T$ is the source signal vector, $n(t) = [n_1(t) \dots n_m(t)]^T$ is the noise vector, and H is an unkown full rand $m \times n$ mixing matrix [12], [13]. Now we use the simple Linear separating system to recover the original source signal,

$$y(t) = W(t)x(t) \qquad (5)$$

Where $y(t) = [y_1(t) \dots y_n(t)]^T$ is an estimate of $s(t)$ and $W(t)$ is a $n \times m$ separating matrix, which is often called the synaptic weight matrix. A number of adaptive learning algorithms have been proposed by which the matrix $W(t)$ is expected to converge to the separating matrix. However, the problem is ill-conditioned, and we cannot obtain H^{-1} even when $m = n$. Hence, $W(t)$ should converge at best to $W*$ that satisfies,

$$T = W * H = PD$$

Where, D is a scaling diagonal nonsingular matrix and P is any $n \times m$ performance matrix [6], [10]. The performance of the source separation is evaluated by the composite matrix $T(t) = W(t)H$ which describes the total or global mixing separating model such that $y(t) = T(t)s(t)$. The separation is perfect when it tends to a generalized permutation matrix PD which has exactly one nonzero element in each row and each column. This corresponds to the indeterminacies of the scaling and order of the estimated signals $y_i(t)$. This indeterminacy is not usually a serious problem since the most relevant information about the source signals is contained in the waveforms of the signals rather than their magnitudes and orders in which they are arranged. The performance matrix, P, should equal H, possibly rescaled and permutated, identity matrix. This means that the condition of the performance matrix should go towards one during learning [10], [11].

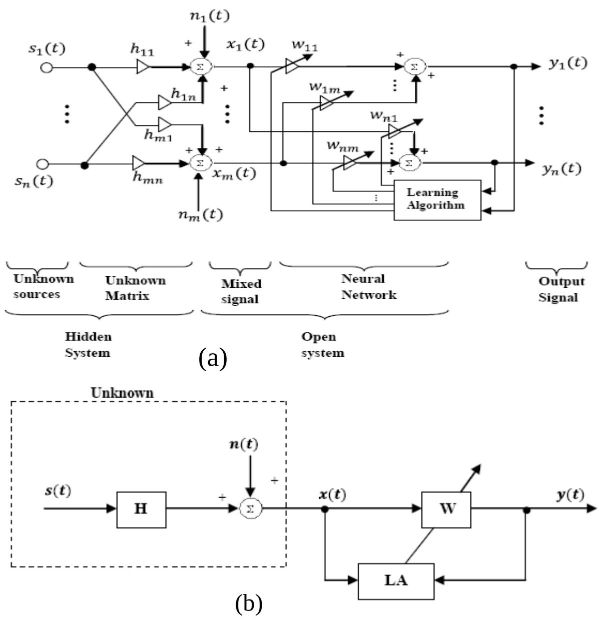


Figure 2. (a) Architecture of mixing and separating models and (b) general block diagram of learning algorithm.

For every individual mixing matrix of H we get the different condition of performance matrix and from that we can get a comparative study between the matrix of H and the characteristics of condition of performance matrix.

TABLE I. TABLE FOR H MATRIX

Serial No	H Matrix
1.	$H_1 = \begin{matrix} -2.2851 & -0.3922 \\ -2.5672 & 0.9902 \end{matrix}$
2.	$H_2 = \begin{matrix} 0.2899 & -1.3522 \\ 0.7182 & -0.5507 \end{matrix}$

For different combination of matrix H, we get the different condition of performance matrix which is shown in the figure 3.

IV. BLIND SIGNAL EXTRACTION

There are several approaches have been recently developed for blind signal extraction and blind deconvolution [14], [15]. But we use the process is called the blind signal extraction in contrast to BSS. A single processing unit is used in the first step to extract one source signal with specified stochastic properties. In the next step, a deflation technique is used in order to eliminate the already extracted signals from the mixtures. Another technique employs a hierarchical neural network which ensures that the signals are extracted in a specified order without extracting the same source signals twice [18], [19]. This is achieved by using inhibitory synapses between the units. Let us consider a single processing unit described by

$$y_1(k) = w_1^T(k)\, x_1(k) = \sum_{j=1}^{m} w_{1j}(k) x_{1j}(k) \qquad (6)$$

Where $x_1 = x$ or $x_1 = Qx$ (where Q is prewhitening matrix). The unit successfully extracts a source signal, say the jth signal, if $w_1(\infty) = w_{1*}$ that satisfies the relation $w_{1*}^T H = e_j^T$, where e_j denotes the $j$thcolumn of a $n \times n$ nonsingular diagonal matrix. A possible loss function is

$$l(w_1) = -\frac{1}{4}\, |k_4(y_1)| \qquad (7)$$

Where $k_4(y_1)$ is the normalized kurtosis defined by

$$k_4(y_1) = \frac{E[|y_1|^4]}{\{E|y_1|^2\}^2} - 3 \qquad (8)$$

It is easy to show that for the prewhitened sensor signals x_1 the normalized kurtos = is satisfies the following relations:

$$k_4(y_1) = k_4\,(w_1^T\, H\, s) = k_4(e^T s) \qquad (9)$$

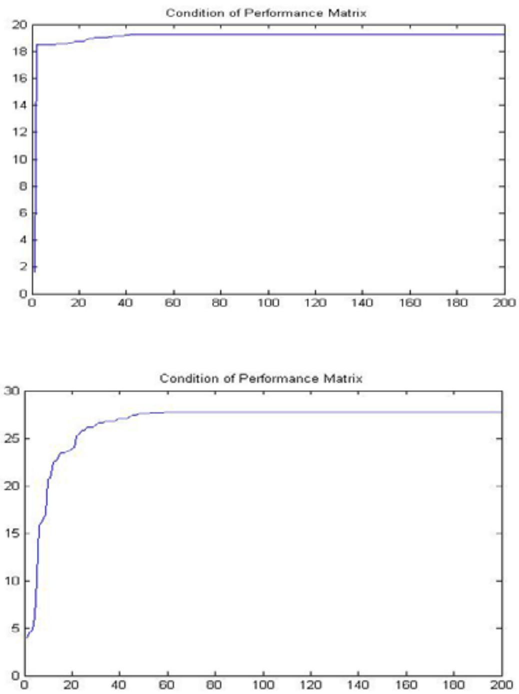


Figure 3. (a) Condition of Performace matrix for matirx H_1 (b) Condition of Performace matrix for matirx H_2.

$$= \sum_{i=1}^{m} \frac{e_i^4}{\|e\|^4}\, k_4\,(s_i) = \sum_{i=1}^{m} e_i^{-4}\, k_4\,(s_i) \qquad (10)$$

Where $e = w_1^T\, H$ and $\|\bar{e}\| = 1$. In order to derive a global convergent (a) learning rule, we apply the standard stochastic gradient descent technique. Minimization of the loss function leads to a simple learning rule [14].

$$w_1(k+1) = w_1(k) - \eta_1(k)\, f\big(y_1(k)\big) x_1(k) \qquad (11)$$

Where x is prewhitened and the nonlinear activation function is evaluated adaptively as

The higher order moments m_2 and m_4 and the sign of the kurtosis $k_4(y_1)$ can be estimated on-line by using the

$$f(y_1) = sign\,[k_4(y_1)] \left[y_1 - \frac{m_2(y_1)}{m_4(y_1)}\, y_1^3\right] \frac{m_4(y_1)}{m_2^3(y_1)} \qquad (12)$$

The higher order moments m_2 and m_4 and the sign of the kurtosis $k_4(y_1)$ can be estimated on-line by using the following averaging formula,

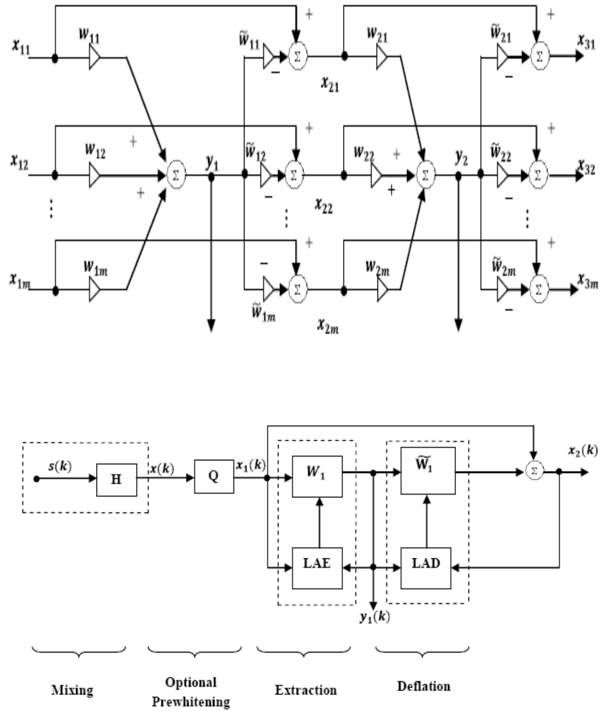


Figure 4. (a) Neural network for blind extraction of signals and independent component analysis (b) block diagram illustration extraction and deflation process of the first extracted signal.

$$m_p(k+1) = (1-\eta) m_p\,(k) + \eta |y_1(k)|^p \qquad (13)$$

With $\eta > 0$ and $m_p\,(k) \cong E[|y_1(k)|^p], p = 2, 4$. This procedure may be recursively applied to extract sequentially the rest of the estimated source signals. This means that we require an on line linear transformation, $x_{1+j}(k) = x_j(k) - \tilde{w}_j(k) y_j(k), \quad j = 1, 2, \dots \dots$

(14)

This ensures minimization of the generalized energy (loss) function

$$l_j(\tilde{w}_j) = \frac{1}{p}\, \|x_j + 1\|^p \qquad (15)$$

Where, $y_j = w_j^T\, x_j$ and

$$w_j(k+1) = w_j(k) - \eta_j(k) f\left[y_j(k)\right] x_j(k) \qquad (16)$$

$$f(y_j) = sign\left[k_4(y_j)\right]\left[y_j - \frac{m_2(y_j)}{m_4(y_j)}\, y_j^3\right]\frac{m_4(y_j)}{m_2^3(y_j)} \qquad (17)$$

Minimization of the loss function leads to the simple local learning rule,

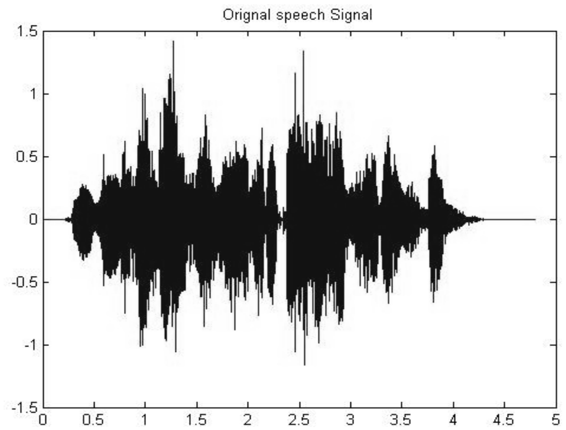
$$\tilde{w}_j(k+1) = \tilde{w}_j(k) + \tilde{\eta}_j(k)\left[y_j(k)\right] g\left[x_{j+1}(k)\right] \qquad (18)$$

$j = 1, 2, \ldots ..$

It is easy to show that vector $\tilde{w}_j$ converges to one the column vector of the global mixing matrix $A = QH$. The procedure can be continued until the amplitude of each signal x_{j+1} is below some threshold [16], [20]. This procedure means that it is not required to know the number of source signals in advance, but it is assumed that the number is constant and the mixing system is stationary.

V. CASE STUDY

Here we consider a model for speech signals where we get an original signal which is originally mixtures of the three speech signals. The duration of the speech signal is 4 second and the sample rate is 12 KHz. Now we have to do perform of separation of this mixture and create three separated and individual speech signals. We use the same



non-linearity as in the first part and if the separation is not satisfactory at all after the one iteration, iterate it several times. The original speech signal and the three separated and individual speech signals are shown in the fig. 5.

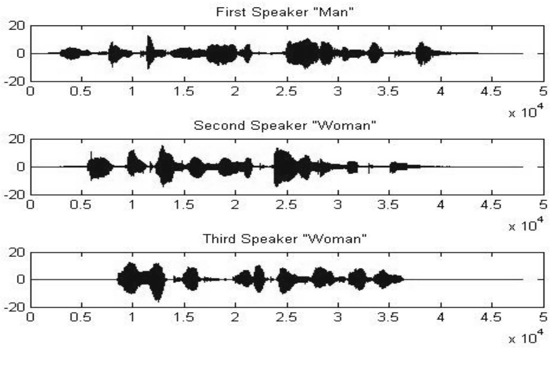


Figure: 5 (a) Original Speech signal (b) Three seperated speech signal extractign from the orginal one.

VI. CONCLUSION

Blind Signal Separation and Independent Component Analysis is a powerful and useful statistical tool for extracting independent source given only observed data that are mixtures of the unknown sources. The efficiency of the proposed method of Independent Component Analysis, has been tested with simulated signals as well as with clinical EEG. Theoretical Investigation and experimental results indicate that ICA constitutes a comprehensive and promising tool for the signal processing and control.

VII. REFERENCES

[1] S Amari and A Cichocki, “Adaptive blind signal processing-neural network approaches,” Proceedings of the IEEE, Issue. 10, pp. 2026-2048, Oct 1998.

[2] Weidong Zhou and Lei Jia, “Blind signal separation using ICA,” International Conference on Info-tech and Info-net, 2001. Proceedings. ICII 2001 - Beijing. 2001 , vol. 4, pp. 237-241, Nov. 2001.

[3] S. Amari, “A theory of adaptive pattern classifiers,” IEEE Trans.Electro. Comput., vol. EC-16, pp. 299–307, Mar. 1967.

[4] S. Amari, “Neural theory of association and conceptformation,” Biological Cybern., vol. 26, pp. 175–185, 1977.

[5] C. Jutten and J. Herault, “Blind separation of sources—PartI: An adaptive algorithm based on neuromimetic architecture,” Signal Processing, vol. 24, pp. 1–20, 1991.

[6] L. Tong, R.-W. Liu, V. C. Soon, and J. F. Huang, “Indeterminacy and identifiability of blind identification,” IEEE Trans.Circuits Syst. vol. 38, pp. 409–509, 1991.

[7] D. B. Chan, P. J. W. Rayner, and S. J. Godsill, “Multi-channel signal separation,” in Proc. ICASSP, May 1996, pp. 649–652.

[8] E. Weinstein, M. Feder, and A. Oppenheim “Multichannel signal separation by decorrelation,” IEEE Trans. Speech Audio Processing, vol. 1, no. 4, pp. 405–413, 1993.

[9] L. Tong, Y. Inouye, and R.-W. Liu, “Waveform preserving blind estimation of multiple independent sources,” IEEE Trans. Signal Processing, vol. 41, pp. 2461–2470, 1993.

[10] S. Amari, A. Cichocki, and H. H. Yang, “Recurrent neural networks for blind separation of sources,” in Proc. Int. Symp. Nonlinear Theory and its Applications, NOLTA-95, Las Vegas, Dec.1995,pp.37–42.

[11] A.Cichocki and R. Unbehauen, “Robust neural networks with on-line learning for blind identification and blind separation of sources,” IEEE Trans. Circuits Syst. I, vol. 43, pp. 894–906, Nov. 1996.

[12] A. Cichocki, S. Amari, M. Adachi, and W. Kasprzak, “Self adaptive neural networks for blind separation of sources,” in Proc. 1996 IEEE Int. Symp. Circuits and Systems, May 1996, vol. 2, pp. 157–160.

[13] A. Cichocki, S. Douglas, S. Amari, and P. Mierzejewski,“Independent component analysis for noisy data,” in Proc. Int. Workshop Independence and Artificial Neural Networks, Tenerife, Feb. 9/10, 1998, pp. 52–58.

[14] A. Cichocki, R. Thawonmas, and S. Amari, “Sequential blind signal extraction in order specified by stochastic properties,” Electron. Lett., vol. 33, no. 1, pp. 64–65, Jan. 1997.

[15] J. J. Shynk and R. P. Gooch, “The constant modulus array for co-channel signal copy and direction finding,” IEEE Trans Signal Processing, vol. 44, pp. 652–660, Mar. 1996.

[16] A. Cichocki, S. Amari, and J. Cao, “Neural network models for blind separation of time delayed and

convolved signals," Japanese IEICE Trans. Fundamentals, vol. E-82-A, no. 9, pp. 1057–1062, Sept. 1997.

[17] P. Comon , "Independent component analysis, A new concept", Signal Processing, vol. 36, pp.287 -314,1994.

[18] V. Capdevielle, C. Serviere, and J. Lacoume, "Blind separation of wide-band sources in the frequency domain," in Proc. ICASSP, May 1995, pp. 2080–2083.

[19] A. Cichocki, S. Douglas, S. Amari, and P. Mierzejewski, "Independent component analysis for noisy data," in Proc. Int. Workshop Independence and Artificial Neural Networks, Tenerife, Feb. 9/10, 1998, pp. 52–58.

[20] D. T. Pham, P. Garrat, and C. Jutten, "Separation of a mixture of independent sources through a maximum likelihood approach," in Proc. EUSIPCO'92, 1992, pp. 771–774.