



## **Separation of Independent Sources in Composite Signal**

*Irshad Ahamad, Ravi Pandit and Parag Parandkar  
Department of Electronics and Communication Engineering,  
Oriental University, Indore, (MP) India*

*(Received 1 October, 2013 Accepted 05 November, 2013)*

**Abstract:** This Paper addresses the problem to solve Separation of independent sources in composite signal, using fast independent component analysis (FICA). The term blind refers to the fact that no explicit knowledge of source signals or mixing system is available. In a statistical signal processing, signals from multiple sources arrive simultaneously at the receiver array, so that each receiver array output contains a mixture of source signals that is random signals. Sets of receiver outputs are processed to recover the source signals or to identify the mixing system. FICA approach uses statistical independence of the source signals to solve the (BSS) problems [13].

Estimating the original source signals without knowing the parameters of mixing and/or filtering processes. It is difficult to imagine that one can estimate this at all. In fact, without some a priori knowledge, it is not possible to uniquely estimate the original source signals. However, one can usually estimate them up to certain indeterminacies. ICA is one of the most widely used BSS techniques for revealing hidden factors that underlie sets of random variables, measurements, or signals [14].

**Keywords:** BSS (Blind source separation), Fast ICA (Independent sources Analysis).

### **I. INTRODUCTION**

Separation of independent sources in composite signal a new technique to statistically extract independent components from the observed multidimensional mixture of data. Statistical Signal Processing basically refers to the analysis of random signals using appropriate statistical techniques. To involve the recovery of information from physical observations, due to the random nature of the signal, statistical techniques play an important role in signal processing. Statistics is used in the formulation of appropriate models to describe the behaviour of the system, the development of appropriate techniques for estimation of model parameters, and the assessment of model performances. Statistical signal processing technique having practical application areas, such as latent variable separation, mixed voices or images, analysis of several types of data or feature extraction [1].

Statistical Signal Processing technique developed to deal with problems that are closely related to cocktail-party problems. The cocktail party problem is basically the scenario that a party with several conversations going on around the room. Even though hear these conversations as mixed signals in ears, humans are unable to unfix these signals into different components, so that concentrate on a sole conversation. The model of the problem to recover the source signals given only the mixtures. This problem is also commonly known as Blind Source Separation or Blind Signal Separation [2, 3]. The term blind refers to the fact that no explicit knowledge of source signals or mixing system is available. Independent component analysis approach uses statistical independence of the source signals to solve the blind signal separation problems [4]. The problem of source separation is an inductive inference problem. There is not enough information to deduce the solution, so one must use any available information to inter

the most probable solution. The aim is to process these observations in such a way that the original source signals are extracted by the adaptive system [5].

Application domains include communications, biomedical, audio, image, and sensor array signal processing. Fast ICA algorithm improves the efficiency of independent component analysis. However, most of the publication focused on offline signal processing using Fast ICA algorithm. It cannot be applied to real-time applications such as speech signal enhancement and EEG/MEG essential features extraction for brain computer interface (BCI). In order to realize the real-time signal processing, the Fast ICA algorithm can be implemented on a field-programmable gate array (FPGA) to speed up the computations involved.

### **II. LATENT VARIABLE**

A family of statistical models. It explains the correlations among observed variables by making assumptions about the hidden ('latent') causes of those variables. Consider situations in which a number of sources emitting signal which are interfering with one another. Familiar situations in which this occurs are a crowded room with many people speaking at the same time, interfering electromagnetic waves from mobile phones or crosstalk from brain waves originating from different areas of the brain. In each of these situations the mixed signals are often incomprehensible and to separate the individual signals. This is the goal of Blind Source Separation [6]. A classic problem in BSS is the cocktail party problem. The objective is to sample a mixture of spoken voices, with a given number of microphones - the observations, and then separate each voice into a separate speaker. In this encounter many problems, e.g. time delay between microphones, echo, amplitude difference, voice order in speaker and underdetermined mixture signal.

In ICA the general idea is to separate the signals, assuming that the original underlying source signals are mutually independently distributed. When regarding ICA, the basic framework for most researchers has been to assume that the mixing is instantaneous and linear, as in informal. ICA is often described as an extension to PCA that uncorrelated the signals for higher order moments and produces a non orthogonal basis [7].

### III. INDEPENDENT COMPONENT ANALYSIS (ICA)

Independent Component Analysis (ICA) is a statistical technique, means that knowing the value of one of the components does not give any information about the other components perhaps the most widely used, for solving the blind source separation problem. This section, presents the basic Independent Component Analysis model and show under which conditions its parameters can be estimated.

#### A. ICA model

The general model for ICA is that the sources are generated through a linear basis transformation, where additive noise can be present. Suppose that N statistically independent signals

$$s_i(t) \text{ where } i = 1, 2, 3, 4, \dots, n.$$

Assume that the sources themselves cannot be directly observed and that each signal,  $s_i(t)$ , is a realization of some fixed probability distribution at each time point t. Also, suppose observe these signals using sensors, and then obtain a set of n observation signals

$$x_i(t), \text{ where } i = 1, 2, 3, 4, \dots, n$$

That is mixtures of the sources. A fundamental aspect of the mixing process is that the sensors must be spatially separated (e.g. microphones that are spatially distributed around a room) so that each sensor records a different mixture of the sources. With this spatial separation assumption in mind, model the mixing process with matrix multiplication as follows.

$$x(t) = As(t) \quad \dots(1)$$

where A is an unknown matrix called the mixing matrix and  $x(t)$ ,  $s(t)$  are the two vectors representing the observed signals and source signals respectively as shown in Figure 1. Incidentally, the justification for the description of this signal processing technique as blind is that no information on the mixing matrix, or even on the sources themselves. The objective is to recover the original signals,  $s_i(t)$ , from only the observed vector  $x_i(t)$ . Obtain estimates for the sources by first obtaining the “unmixing matrix” W, where  $W = A^{-1}$ . This enables an estimate,  $s(t)$ , of the independent sources to be obtained:

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE and SI do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

$$s(t) = Wx(t) \quad \dots(2)$$

A key concept that constitutes the foundation of independent component analysis is statistical independence. To simplify the above discussion consider the case of two different random variables  $s_1$  and  $s_2$  as shown in fig. 1.

The random variable  $s_1$  is independent of  $s_2$ , if the information about the value of  $s_1$  does not provide any information about the value of  $s_2$ , and vice versa. Here  $s_1$  and  $s_2$  be a random signals originating from two different physical processes that are not related to each other.

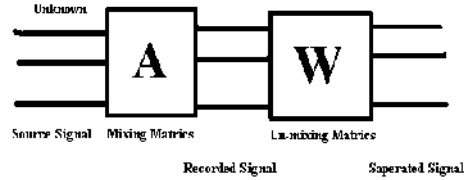


Fig. 1. Blind source separation (BSS) block diagram.

#### B. Independence

Mathematically, statistical independence is defined in terms of probability density of the signals. Consider the joint probability density function (pdf) of  $s_1$  and  $s_2$  be  $p(s_1, s_2)$ . Let the marginal pdf of  $s_1$  and  $s_2$  be denoted by  $p_1(s_1)$  and  $p_2(s_2)$  respectively.  $s_1$  and  $s_2$  are said to be independent if and only if the joint pdf can be expressed as:

$$p_{s_1, s_2}(s_1, s_2) = p_1(s_1)p_2(s_2) \quad \dots(3)$$

Similarly, independence could be defined by replacing the pdf by the respective cumulative distributive functions as;

$$E\{p_1(s_1)p_2(s_2)\} = E\{g_1(s_1)\}E\{g_2(s_2)\} \quad \dots(4)$$

Where  $E\{.\}$  is the expectation operator. In the following section we use the above properties to explain the relationship between uncorrelated and independence.

#### C. Uncorrelated

Two random variables  $s_1$  and  $s_2$  are said to be uncorrelated if their covariance  $C(s_1, s_2)$  is zero.

$$\begin{aligned} C(s_1, s_2) &= E\{(s_1 - ms_1)(s_2 - ms_2)\} \\ &= E\{s_1s_2 - s_1ms_2 - s_2ms_1 + ms_1ms_2\} \\ &= E\{s_1s_2\} - E\{s_1\}E\{s_2\} \\ &= 0 \end{aligned} \quad \dots(5)$$

Where  $ms_1$  is the mean of the signal. Equation 4 and 5 are identical for independent variables taking  $g_1(s_1) = s_1$ .

Hence independent variables are always uncorrelated. However the opposite is not always true. The above discussion proves that independence is stronger than uncorrelatedness and hence independence is used as the basic principle for ICA source estimation process. However uncorrelatedness is also important for computing the mixing matrix in ICA

#### D. Central limit theorem

A simple pre-processing step that is commonly performed is to “centre” the observation vector  $x$  by subtracting its mean vector  $m = E\{x\}$ . That is obtaining the cantered observation vector,  $x_c$ , as follows:

$$x_c = x - m$$

This step simplifies ICA algorithms by allowing us to assume a zero mean. Once the unmixing matrix has been estimated using the cantered data, obtain the actual estimates of the independent components as follows:

$$s(t) = A^{-1}(x_c+m)$$

**E. Whitening**

Another step which is very useful in practice is to prewhiten the observation vector  $x$ . Whitening involves linearly transforming the observation vector such that its components are uncorrelated and have unit variance. Let  $x_w$  denote the whitened vector, then it satisfies the following equation:

$$E\{x_w x_w^T\} = I$$

where  $\{x_w x_w^T\}$  is the covariance matrix of  $x_w$ . Also, since the ICA framework is insensitive to the variances of the independent components, assume without loss of generality that the source vector,  $s$ , is white, i.e

$$E\{ss^T\} = I$$

A simple method to perform the whitening transformations to use the Eigen value decomposition (EVD) of  $x$ . That is, we decompose the covariance matrix of  $x$  as follows

$$E\{xx^T\} = VDV^T$$

Where  $V$  is the matrix of eigenvectors of  $E\{xx^T\}$ ,

And  $D$  is the diagonal matrix of Eigen values, i.e.

$$D = \text{diag}\{ \lambda_1, \lambda_2, \dots, \lambda_n \}.$$

The observation vector can be whitened by the following transformation:

$$x_w = VD^{-1/2}V^T x$$

Where the matrix  $D^{-1/2}$  is obtained by a simple component wise operation as

$$D^{-1/2} = \text{diag}\{ \lambda_1^{-1/2}, \lambda_2^{-1/2}, \dots, \lambda_n^{-1/2} \}.$$

Whitening transforms the mixing matrix into a new one, which is orthogonal

$$x_w = VD^{-1/2}V^T A s = A_w s$$

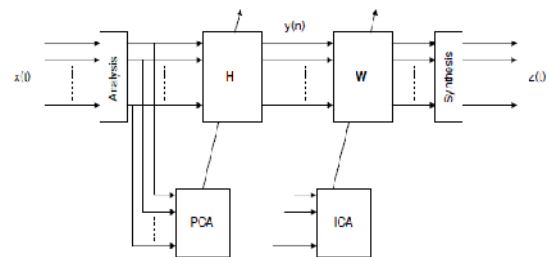
hence,

$$\begin{aligned} E\{x_w x_w^T\} &= A_w E\{ss^T\} A_w^T \\ &= A_w A_w^T \\ &= I \end{aligned}$$

Whitening thus reduces the number of parameters to be estimated. Instead of having to estimate the  $n^2$  elements of the original matrix  $A$ , only need to estimate the new orthogonal mixing matrix, where orthogonal matrix has  $n(n-1)/2$  degrees of freedom. One can say that whitening solves half of the ICA problem. This is a very useful step whitening is a simple and efficient process that significantly reduces the computational complexity of ICA. An illustration of the whitening process with simple ICA source separation process is explained in the later section.

**F. PCA-ICA algorithm**

The reason for combining both PCA and ICA is to reduce the dimension of the problem before implementing the ICA algorithm, thus making it easier for the ICA algorithm to solve the problem. The proposed structure combined for the PCA and ICA is shown in Figure 2, seen as a two step algorithm, the first part of the algorithm is of second order statistics and decorrelates the data. The second part of the algorithm is of higher order statistics and separates the data. As such, the two algorithms complement one other, since if only PCA is used separation cannot be accomplished, due to uncorrelated does not mean independent. If just ICA is applied alone the problem becomes more difficult, since the dimensions of the problem are too large and ICA will be slow to converge. In the derivations of ICA an assumption has been made that the input data is uncorrelated.



**Fig. 2.** Structure for combining PCA and ICA.

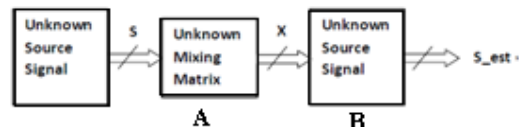
ICA is a statistical method that expresses a set of multidimensional observations as a combination of unknown latent variables. These underlying latent variables are called sources or independent components and they are assumed to be statistically independent of each other. The ICA model is

$$x = f(\theta, s) \dots \dots \dots (1)$$

Where  $x = (x_1, \dots, x_m)$  is an observed vector and  $f$  is a general unknown function with parameters  $\theta$  that operates on statistically independent latent variables listed in the vector  $s = (s_1, \dots, s_n)$ . A special case of (1) is obtained when the function is linear, and can be write

$$x = As \dots \dots \dots (2)$$

Where  $A$  is an unknown  $m \times n$  mixing matrix. In Formulae and consider  $x$  and  $s$  as random vectors. When a sample of observations  $X = (x_1, \dots, x_N)$  becomes available, write  $X = AS$  where the matrix  $X$  has observations  $x$  as its columns and similarly the matrix  $S$  has latent variable vectors  $s$  as its columns shown in Figure 3.



**Fig. 3.** BSS using ICA algorithm.

The mixing matrix  $A$  is constant for all observations. If both the original sources  $S$  and the way the sources were mixed are all unknown, and only mixed signals or mixtures  $X$  can be measured and observed, then the estimation of  $A$  and  $S$  is known as blind source separation (BSS) problem.

The normalized value  $w_{new}$  is compared with the old value  $w_{old}$  and if the values do not match then  $w_{new}$  fed back to the input of the block and also stored as  $w_{old}$  in a register for the purpose of comparison. When  $w_{new} = w_{old}$  then this value is given to the output as the converged vector  $w$  which gives one independent component. For finding the other independent component a new random vector  $w$  is assumed and it is decorrelated with the earlier  $w$  and is again put to the iteration process for getting an optimized converged value For  $I = 1$  and 2 (for two mixed signals).

The process of centering is to subtract the mixed signal means  $\mu_1$  and  $\mu_2$  from  $x_1$  and  $x_2$ , respectively. First the mixed signal elements are accumulated one by one. After getting the summation of  $x$ ,  $\mu$  is obtained by dividing the summation by the sample length. In order to speed up the processing, multiplication operation (multiply by  $1/\text{sample length}$ ) is used instead of the division. Second, the mean is subtracted from the mixed signal data for achieving centering shown in Figure 5.6. The operation is formulated as:

$$x(i) = x(i) - \left( \sum_{i=1}^{\text{samplelength}} x_i \right) \times (1/\text{samplelength})$$

Where  $i = 1, 2, \dots, \text{sample length}$

The first step of whitening is to find the whitening matrix

**P**. **P** is given by :

$$P = D^{-1/2} X E^T$$

Where  $D = \text{diag}(\lambda_1, \lambda_2) =$  diagonal matrix of the covariance matrix  $C_X$ 's eigenvalues.

$E = (e_1, e_2) =$  Orthogonal matrix of  $C_X$ 's eigenvectors.

$C_X = E\{XX^T\}$  is a  $2 \times 2$  matrix.

#### IV. SIMULATION AND RESULTS

In the simulation, source signals sine and square waves generated from MATLAB. Then, the mixed signals are produced by multiplying random mixing matrix and source signals. The sample length of mixed signal  $X$  and estimated independent components  $S$  are both 1000 in the simulation. Fig. 4(a) are the source signals sine and square waves. Fig. 4(b) shows the mixed signals. The demixing matrix  $B$  is found by the Fast ICA algorithm and the estimated independent component signals, shown in Fig. 4(c), are derived by multiplying the demixing matrix  $B$  and mixed signals.

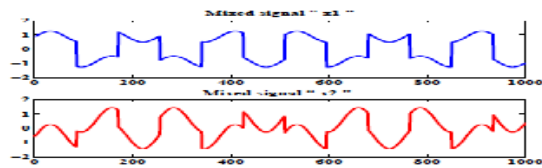


Fig. 4. (a) Independent sources  $s_1$  and  $s_2$ .

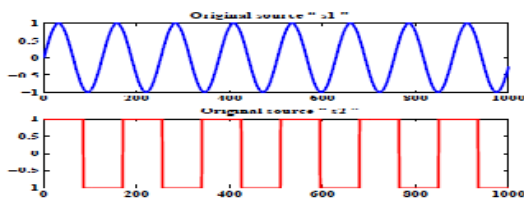


Fig. 4. (b) Observed signals,  $x_1$  and  $x_2$ .

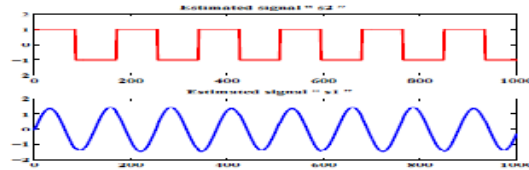


Fig. 4. (c) Estimates of independent components.

(From an unknown linear mixture of unknown independent components).

The independent components are then mixed, an arbitrarily chosen mixing matrix  $A$ , where

$$A = \begin{pmatrix} 0.3816 & 0.8678 \\ 0.8534 & -0.5853 \end{pmatrix}$$

The resulting signals from this mixing are shown in Fig. 4 (b). Finally, the mixtures  $x_1$  and  $x_2$  are separated using Fast ICA to obtain  $s_1$  and  $s_2$ , shown in Fig. 4 (c). By comparing Fig. 4 (c) to Fig. 4 (a) it is clear that the independent components have been estimated accurately and that the independent components have been estimated without any knowledge of the components themselves or the mixing process.

#### V. TWO SPEECH SIGNAL

In this simulation two speech signal are taken as the source signals shown in Fig.5(a). The source signal recovered by Tanh with symmetric orthonormalization based fixed point algorithm and deflation orthonormalizations shown in Fig.5(b). Algorithms using Tanh with symmetric orthonormalization have pretty much the same statistical performance shown in Fig 5 (b) recovered by fixed point algorithm

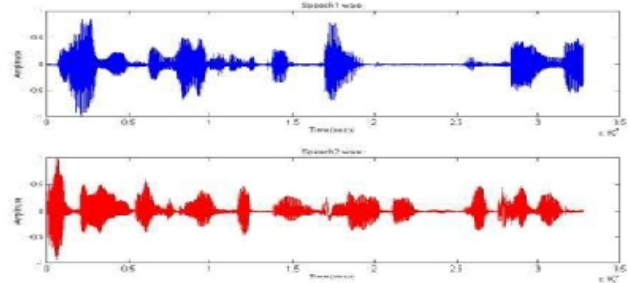


Fig. 5. (a) Original speech signals.

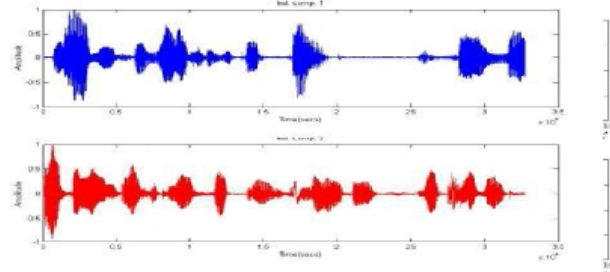


Fig. 5. (b) Tanh with symmetric orthonormalization based fixed point algorithm

## CONCLUSION

Ways needed in Fast ICA algorithm for decorrelation of the separating matrix can be deflationary or symmetric orthonormalization.

As for statistical performance the best results are obtained by using Tanh symmetric orthonormalization. All algorithms using Tanh with symmetric orthonormalization based fixed point algorithm have pretty much the same statistical performance and can also be used in complex applications like harmonic separation in power system and interference suppression in CDMA systems. Performance comparison in terms of CPU time shows that deflation orthonormalization approach of the Fast ICA algorithm is clearly inferior. Extensive simulation studies reveal that symmetric orthonormalization approach has a better performance as compared to deflation approach. Looking at the computational load, one sees clearly that Fast ICA requires the smallest amount of computation. In some applications, it may be preferable to use the fast ICA algorithm with deflation orthonormalization in which every vector is impartially treated and the parallel computation of independent components is enabled.

The proposed research can be extended in following dimensions. Pipeline architecture of the Fast ICA algorithm can be developed for real-time sequential mixed signal processing. An extended implementation of Fast ICA algorithm based on the proposed modules can be done for higher dimension (more than two sources and mixtures). VLSI implementation of Fast ICA algorithm offers many features such as high processing speed, which is extremely desired in many applications. In order to reduce the complexity, the Fast ICA block can be divided into several sub modules and each of the sub modules is developed by HDL coding. VLSI implementation of different ICA technique can be carried out.

## REFERENCES

- [1]. Kuo-Kai Shyu, Ming-Huan Lee, Yu-Te Wu, Po-Lei Lee Implementation of pipelined FastICA on FPGA for Real-time Blind source separation. *IEEE Trans. Neural Netw.*, vol. **19**, no.6, June 2011.
- [2]. J.T. Chien, H.-L. Hsieh, and S. Furui, "A new mutual information measure for independent component analysis," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, 2010, pp. 1817–1820.
- [3]. Blind Separation Method Of Noised Image Based On Neural Network Nonlinear Filtering And Independent Component Analysis Tian Yanan, Wang Xu School Of Information Science & Engineering Northeastern University Shenyang, China E-Mail: Tianyanan@Ise.Neu.Edu.Cn 2010 *International Conference On Computer Science And Software Engineering*.
- [4]. H. Du, H. Qi And X. Wang, Comparative Study Of Vlsi Solutions To Independent Component Analysis, *IEEE Trans. Industrial Electronics*, Vol. **54**, No.1, 2009.
- [5]. Jen-Tzung Chien, Senior Member, IEEE, And Hsin-Lung Hsieh - Convex Divergence Ica For Blind Source Separation. *IEEE Transactions On Audio, Speech, And Language Processing*, Vol. **20**, No. 1, January 2012.
- [6]. Zhan-Li Sun And Kin-Man Lam, Member, IEEE - Depth Estimation Of Face Images Based On The Constrained Ica Model. *IEEE Transactions On Information Forensics And Security*, Vol. **6**, No. 2, June 2011.
- [7]. Javier Vía, Member, IEEE, Daniel P. Palomar, Senior Member, IEEE, Luis Vielva, And Ignacio Santamaría, Senior Member, *IEEE - Quaternion Ica From Second-Order Statistics. IEEE Transactions On Signal Processing*, Vol. **59**, No. 4, April 2011.
- [8]. Lin Wang, Heping Ding, Senior Member, IEEE, And Fuliang Yin - A Region-Growing Permutation Alignment Approach In Frequency-Domain Blind Source Separation Of Speech Mixtures. *IEEE Transactions On Audio, Speech, And Language Processing*, Vol. **19**, No. 3, March 2011.
- [9]. Hualiang Li And Tülay Adah, Fellow, IEEE - Algorithms For Complex MI Ica And Their Stability Analysis Using Wirtinger Calculus. *IEEE Transactions On Signal Processing*, Vol. **58**, No. 12, December 2010.
- [10]. Chengjun Liu And Jian Yang - Ica Color Space For Pattern Recognition. *IEEE Transactions On Neural Networks*, Vol. **20**, No. 2, February 2009
- [11]. J. Cook, V. Chandran, And S. Sridharan, "Multiscale Representation For 3-D Face Recognition," *IEEE Trans. Inf. Forensics Security*, Vol. **2**, No. 3, Pp. 529–536, Sep. 2007.
- [12]. Fpga Implementation Of Fastica Based On Floating-Point Arithmetic Design For Real-Time Blind Source Separation Kuo-Kai Shyu And Ming-Huan Li 2006 International Joint Conference On Neural Networks Sheraton Vancouver Wall Centre Hotel, Vancouver, Bc, Canada July 16-21, 2006.
- [13]. J.T. Chien And B.C. Chen, A New Independent Component Analysis For Speech Recognition And Separation, *IEEE Trans. Speech Audio Process.*, Vol. **14**, No.4, Pp.1245-1254, July 2006.
- [14]. Huy Nguyen And Rong Zheng, Senior Member, IEEE "Binary Independent Component Analysis With Or Mixtures". *IEEE Transactions on Signal Processing*, Vol. **59**, No. 7, July 2006.
- [15]. Y. Chen, "Blind Separation Using Convex Function," *IEEE Trans. Signalprocess.*, Vol. **53**, No. 6, Pp. 2027–2035, Jun. 2005
- [16]. C. M. Kim, H. M. Park, Y. K. Choi, And S. Y. Lee, Fpga Implementation Of Ica Algorithm For Blind Source Separation And Adaptive Noise Cancelling, *IEEE, Trans, Neural Netw.*, Vol. **14**, No.5, Pp.1038-1046, Sep.2003.
- [17]. Scott C. Douglas A Statistical Convergence Analysis Of The Fastica Algorithm For Two-Source Mixtures", *IEEE Trans. Neural Netw.*, Vol. **14**, Pp. 943-949, July 2003
- [18]. A Combined Approach Of Array Processing And Independent Component Analysis For Blind Separation Of Acoustic Signals Futoshi Asano', Shiro Ikeda, Michiaki Ogawa, Hideki Asoh, Nobuhiko Kitawaki 1) Electrotechnical Laboratory, 2) Prest, Jst, 3) Tsukuba University.

[19]. Virtues And Vices Of Source Separation Using Linear Independent Component Analysis For Blind Source Separatio Of Non-Linearly Coupled And Synchronised 2001 *Proceedings Of The 23rd Annual Embs International Conference*, October 25-28, Istanbul, Turkey.

[20]. Blind Source Separation Of Acoustic Mixtures Using Timefrequency Domain Independent Component Analysis Dr. S. Jayaraman (Asst. Professor),Gsitaraman, Rseshadri, Dept. Of Electronics And Communication Engg., P.S.G. College Of Technology, Coimbatore-641004.

[21]. Combined Approach Of Array Processing And Independent Component Analysis For Blind Separation Of Acoustic Signals Futoshi Asano, Member, IEEE, Shiro Ikeda, Member, IEEE, Michiaki Ogawa, Hideki Asoh, Member, IEEE, And Nobuhiko Kitawaki, Member, IEEE.