International Journal of Electrical, Electronics and Computer Engineering **1**(2): 28- 33(2012)

**ISSN No. (Online) : 2277-2626**

# Statistical Signal Processing for Latent Variable

***Vasim Khan* and Prof. Gaurav Gupta*****

*Department Electronics and Communication Engineering,*

**Vikrant Institute of Technology and Management, Indore, (MP)*

*** Mahakal Institute of Technology, Ujjain, (MP)*

*(Received 05 November, 2012, Accepted 02 December, 2013)*

**ABSTRACT: Statistical Signal processing may broadly be considered 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 behavior of the system, the development of appropriate techniques for estimation of model parameters, and the assessment of model performances. This paper evaluates and compares the performance of Ways needed in fast ICA algorithm for decorrelation of the separating matrix can be deflationary or symmetric orthogonalization. Simulation studies reveal that symmetric approach has a better performance as compared to deflation approach, in terms of CPU time. The performance of the real-time applications such as speech signal enhancement and EEG/MEG essential features extraction for brain computer interface (BCI) based on MATLAB (R 2011a) process models.**

**KEYWORDS**: Independent Component Analysis (ICA), Non-Gaussianity, Principal Component Analysis, Orthonormalization Kurtosis, Centering.

## I. INTRODUCTION

Statistical Signal Processing basically refers to the analysis of random signals using appropriate statistical techniques. The main purpose of this Paper is to introduce different signal processing models and different statistical and computational issues involved in solving them for Latent variable. statistical signal processing technique having emerging new practical application areas, such as latent signal separation such as mixed voices, identify aircrafts and interference from their mixtures such as Electroencephalogram (EEG), Magneto encephalography (MEG), and Electrocardiogram (ECG) or images, analysis of several types of data or feature extraction. Independent Component Analysis (ICA) is a statistical signal processing technique separates the independent sources from their mixtures by measuring non-gaussian. A random or stochastic process is a mathematical model for a phenomenon that evolves in time in an unpredictable manner from the viewpoint of the observer. It may be unpredictable because of such effects as interference or noise in a communication link or storage medium, or it may be an information-bearing signal.

Deterministic from the viewpoint of an observer at the transmitter but random to an observer at the receiver [1-6]. The theory of random processes quantifies the above notions so that one can construct mathematical models of real phenomena that are both tractable and meaningful in the sense of yielding useful predictions of future behavior.

## II. LITERATURE REVIEW

Two array signal processing techniques are combined with independent component analysis to enhance the performance of blind separation of acoustic signals in a reflective environment such as rooms. The first technique is the subspace method which reduces the effect of room reflection. The second technique is a method of solving permutation, in which the coherency of the mixing matrix in adjacent frequencies is utilized [1].

Address the imminent problem which arises when researchers injudiciously use a linear and instantaneous (memory less) model for the source mixing structures of independent component analysis (ICA), also known as blind source separation (BSS), in pursuit of separating noisy and frequently no stationary combined mother and fatal electrocardiogram (ECG) signals from cutaneous measurements under the following false assumptions. (1) Sensors (electrodes) are instantaneous linear mixtures of mother and fatal source signals. (2) Noise is an additive Gaussian perturbation. (3) Mother and fetal ECG signals\ are assumed to be stationary and linear, mutually statistically independent and statistically independent from noise. (4) Most of the second-order (SO) and fourth-order (FO) blind source separation (BSS) methods developed this last decade assume that third-order cumulants vanish hence the need to use FO. All these assumptions are not valid and will be challenged. We will expose these vices without providing any significant contributions for overcoming them.

Rather, we provide a framework for investigations which are based on conformal mapping of nonlinear mixtures and novel dynamic nonlinear structures with time-variant memory to cater for quadratic coupling between mothers and fatal which is quasi-periodical and the concomitant (quasi) cyclostationarity [2]. Results given here show linear ICA shortfalls in non stationary environment which is precipitated by quadratic coupling between mother and fatal ECGs during events of synchronized QRS complexes and P-waves and account for more than 20% of the 100,000 maternal cardiac cycles obtained from several clinical trials.

Blind Source Separation of acoustic mixtures aims at providing a solution to the classical cocktail-party problem. The inherent delays and convolutions in microphone recordings, entails a modification in the Independent Component Analysis (ICA), which achieves separation by making the assumption of statistical *independence* of source signals that are linearly combined. The Proposed Algorithm provides a solution for the blind source separation problem by shifting he domain of the problem to Time-Frequency domain and applying ICA to each of the Frequency components individually[3]. Satisfactory results were achieved for Speech-Music as well as Speech-Speech Separation by adopting the Time- Frequency domain ICA.

The first technique is the subspace method which reduces the effect of room reflection when the system is used in a room. Room reflection is one of the biggest problems in blind source separation (BSS) in acoustic environments. The second technique is a method of solving permutation. For employing the subspace method, ICA must be used in the frequency domain, and precise permutation is necessary for all frequencies [1].

In this method, a physical property of the mixing matrix, i.e., the coherency in adjacent frequencies, is utilized to solve the permutation. The experiments in a meeting room showed that the subspace method improved the rate of automatic speech recognition from 50% to 68% and that the method of solving permutation achieves performance that closely approaches that of the correct permutation, differing by only 4% in recognition rate [4].

Independent component analysis (ICA) is usually used for blind source separation (BSS), and the FastICA algorithm separates the independent sources from their mixtures by measuring nongaussianity using Kurtosis. In this paper, the field programmable gate array (FPGA) implementation of FastICA for real-time signal process is proposed and the sample rate of 192 kHz is reached under the presented architecture [5]. The floating-point arithmetic design provides better accuracy and higher dynamic performance than fixed-point design for implementation of digital signal processing algorithm. The FPGA design is based on a hierarchical concept, and the experimental results of the design are presented.

## III. METHODOLOGY

Observe N linear mixtures x1,…,xn of n independent components xj = $a_{j1}s_1 + a_{j2}s_2 + \ldots + a_{jn}s_n$, for all j, $a_j$ is the column of the mixing matrix A. Assume each mixture $x_j$ and each $s_k$ is a random variable Time difference between mixes dropped Independent components are latent variables. Cannot be directly observed [5-7].

Independent component analysis (ICA) is a well-known method of finding latent structure in data. 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 [8].

The ICA model is $x = f(\theta, s)$ Where $x = (x1, \ldots, xm)$ is an observed vector and f is a general unknown function with parameters θ that operates on statistically independent latent variables listed in the vector $s = (s1, \ldots, sn)$.
A special case is obtained when the function is linear, And we can write $x = As$, Where A is an unknown m × n mixing matrix. consider x and s as random vectors. When a sample of observations $X = (x1, \ldots, xN)$ becomes available, 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. 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 [6,7].

ICA Mixture model: x = As,A is mixing matrix; s is matrix of source signals. Goal: Find some matrix W, so that, s = Wx
W = inverse of A. Non gaussianity estimates independent
Estimation of $y = w^T x$, let $z = A^T w$, so $y = w^T As = z^T s$ y is a linear combination of $s_i$, therefore $z^T s$ is more gaussian than any of $s_i$,$z^T s$ becomes least gaussian when it is equal to one of the $s_i$ .$w^T x = z^T s$ equals an independent component [11] Kurtosis Fourth order cumulant. Classical measure of nongaussianity. $\mathrm{kurt}(y) = E\{y^4\} - 3(E\{y^2\})^2$. For gaussian y, fourth moment = $3(E\{y^2\})^2$.Kurtosis for gaussian random variables is 0[14].Entropy (H) degree of information that an observation gives. A Gaussian variable has the largest entropy among all random variables of equal variance.
Negentropy J: Based on the information theoretic quantity of differential entropy [5].

### (a) Data Preprocessing For ICA

It is often beneficial to reduce the dimensionality of the data before performing ICA. It might be well that there are only a few latent components in the high-dimensional observed data, and the structure of the data can be presented in a compressed format. Estimating ICA in the original, high-dimensional space may lead to poor results.

For example, several of the original dimensions may contain only noise. Also, over learning is likely to take place in ICA if the number of the model parameters (i.e., the size of the mixing matrix) is large compared to the number of observed data points. Care must be taken, though, so that only the redundant dimensions are removed and the structure of the data is not flattened as the data are projected to a lower dimensional space. In this section two methods of dimensionality reduction are discussed: principal component analysis and random projection [10].

In addition to dimensionality reduction, another often used preprocessing step in ICA is to make the observed signals zero mean and decor relate them. The decor relation removes the second-order dependencies between the observed signals. It is often accomplished by principal [10-11].

## IV. SIMULATION AND RESULTS

Simulation is based on MATLAB (R2011a) process models.

**(a) Simulation 1**.

In the Simulation1Generate Information Signal adaptive cancellation using the MATLAB functions ANFIS and GENFIS1.shown in fig.1

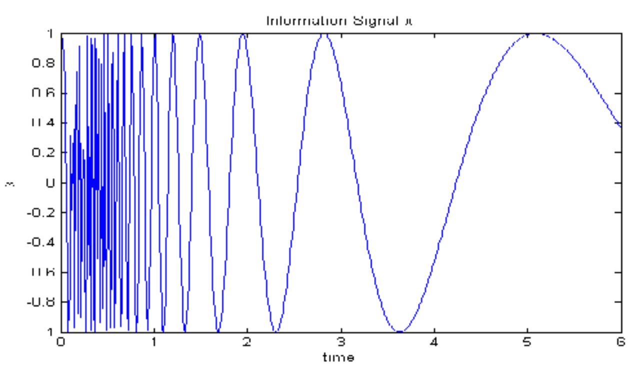


**Fig. 1.** Information Signal.

Unfortunately, the information signal x cannot be measured without an interference signal n2, which is generated from another noise source n1 shown in fig.2 via a certain unknown nonlinear process.

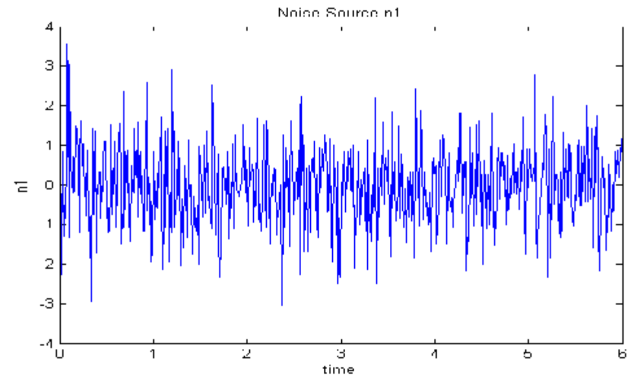


**Fig. 2.** Noise Source (n1)

The interference signal n2 that appears in the measured signal is assumed to be generated via an unknown nonlinear equation: n2(k) = 4*sin(n1(k))*n1(k-1)/(1+n1(k-1)^2),shown in fig.3.

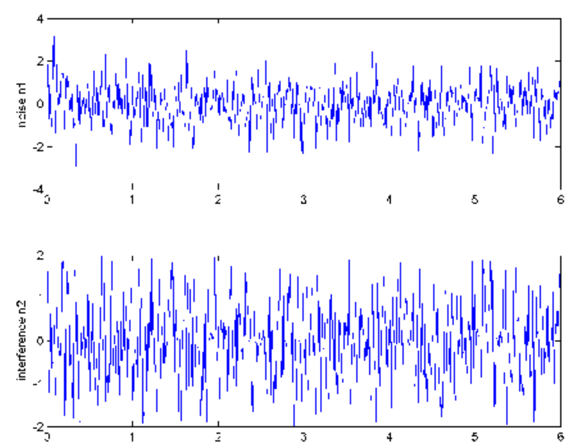


**Fig. 3.** n2(k) = 4*sin(n1(k))*n1(k-1)/(1+n1(k-1)^2).

The measured signal m is the sum of the original information signal x and the interference n2. However, we do not know n2. The only signals available to us are the noise signal n1 and the measured signal m, and our task is to recover the original information signal x. In the measured signal m that combines x and n2.shown in fig.4.

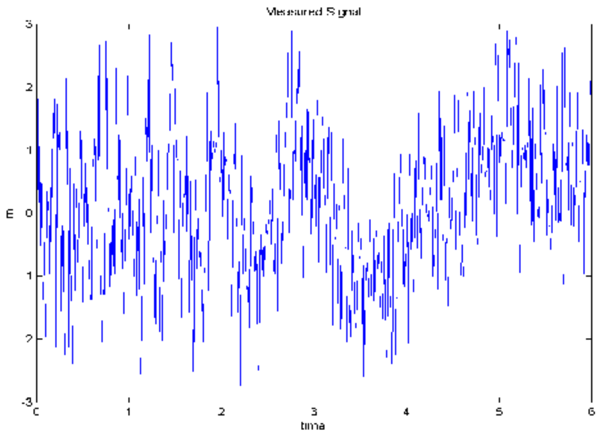


**Fig. 4.** Measured Signal.

**(b) Simulation 2.**

In the simulation 2, the problem of signal recovery from latent noisy data. The general de-noising procedure involves three steps. The basic version of the procedure follows the steps Decompose: Choose a level N. Compute decomposition of the signal at level N.

Threshold detail coefficients: For each level from 1 to N, select a threshold and apply soft thresholding to the detail coefficients.

Reconstruct: Compute reconstruction using the original approximation coefficients of level N and the modified detail coefficients of levels from 1 to N.

Thresholding can be done using the function WTHRESH which returns soft or hard thresholding of the input signal. Hard thresholding is the simplest method but soft thresholding has nice mathematical properties shown in fig.5.

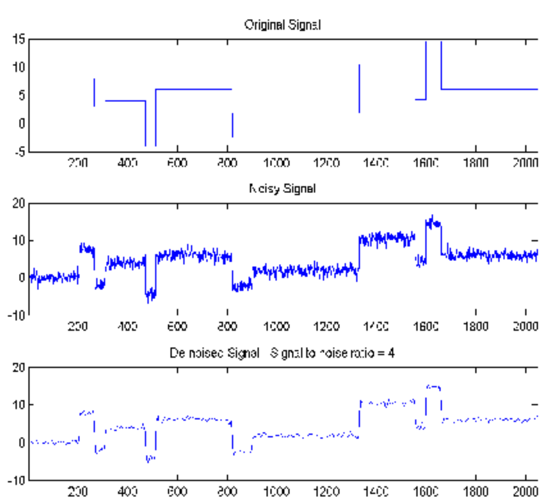


**Fig. 5**. Original Signal, Noisy Signal, De-noised Signal.

**(c) Simulation 3. (Real Time Signal Processing)**

Biomedical signals such as electroencephalogram (EEG), magneto encephalography (MEG), and electrocardiogram (ECG) are generally measured from clinical sensors or instruments; however measured signals are polluted by the aircrafts and other unknown noise signals, such as eye movements, muscle noise, and power noise from instruments. This problem can be solved by independent component analysis(ICA) algorithm, which identifies aircrafts from the measured signals, shown in fig.6.

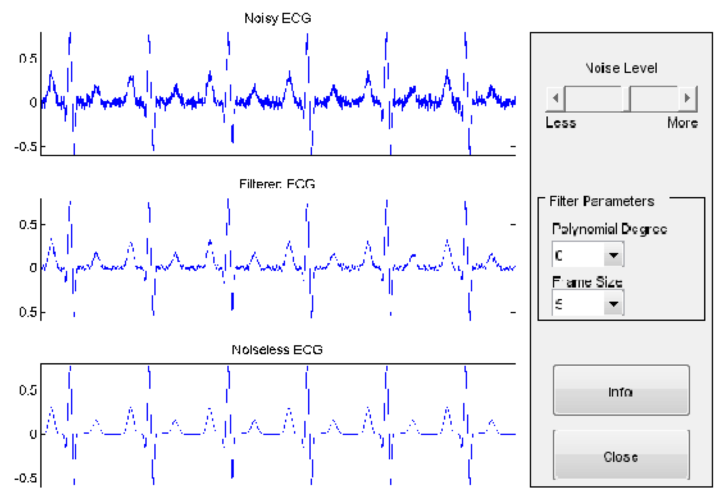


**Fig. 6.** Noiseless ECG.

**(d) Simulation 4. (Acoustic echo cancellation)**

In the simulation 4 application of adaptive filters to acoustic echo cancellation (AEC).Acoustic echo cancellation is important for audio teleconferencing when simultaneous communication (or full-duplex transmission) of speech is necessary. In acoustic echo cancellation, a measured microphone signal contains two signals:

- the near-end speech signal v(n)
- the far-end echoed speech signal dhat (n)

The goal is to remove the far-end echoed speech signal from the microphone signal so that only the near-end speech signal is transmitted. First describe the acoustics of the loudspeaker-to-microphone signal path where the speakerphone is located .The following sequence of commands generates a random impulse response that is not unlike what a conference room would exhibit a system sampling rate of fs = 8000 Hz shown in fig.7

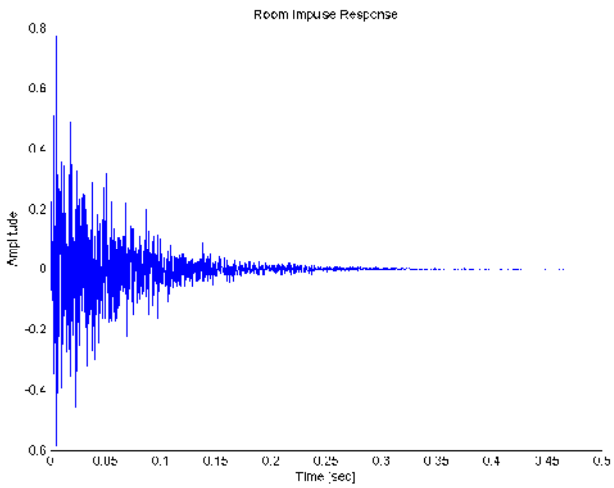


**Fig. 7**. Room Impulse Response.

Now we describe the path of the far-end speech signal. A male voice travels out the loudspeaker, bounces around in the room, and then is picked up by the system's microphone. Let's listen to what his speech sounds like if it is picked up at the microphone without the near-end speech present.(shown in fig.8)

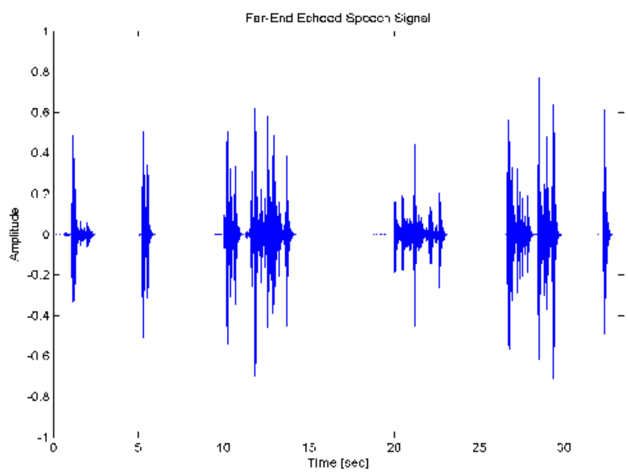


**Fig. 8.** Far-End Echoed Speech Signal.

The Near-End Speech Signal: The teleconferencing system's user is typically located near the system's microphone. Here is what a male speech sounds like at the microphone (shown in fig.9).

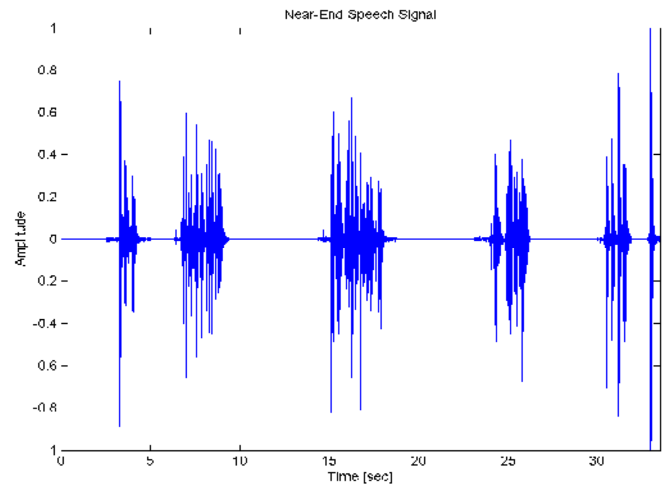


**Fig. 9.** Near-End Speech Signal.

The signal at the microphone contains both the near-end speech and the far-end speech that has been echoed throughout the room. The goal of the acoustic echo canceler is to cancel out the far-end speech, such that only the near-end speech is transmitted back to the far-end listener (shown in fig.10).

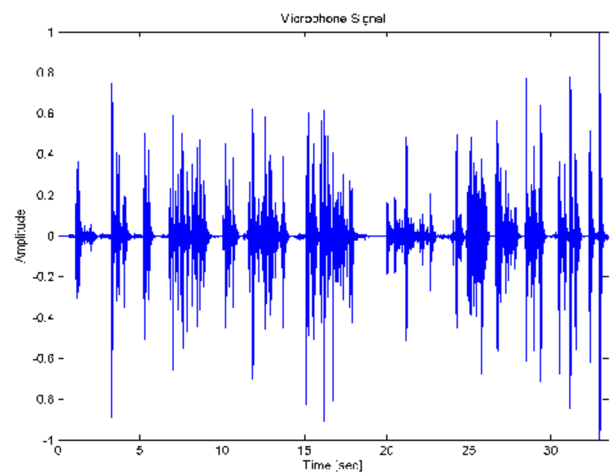


**Fig. 10.** Microphone Signal.

The FDAF uses a fast convolution technique to compute the output signal and filter updates. This computation executes quickly in MATLAB (R2011a). It also has improved convergence performance through frequency-bin step size normalization. Pick some initial parameters for the filter and see how well the far-end speech is cancelled in the error signal (shown in fig.11).

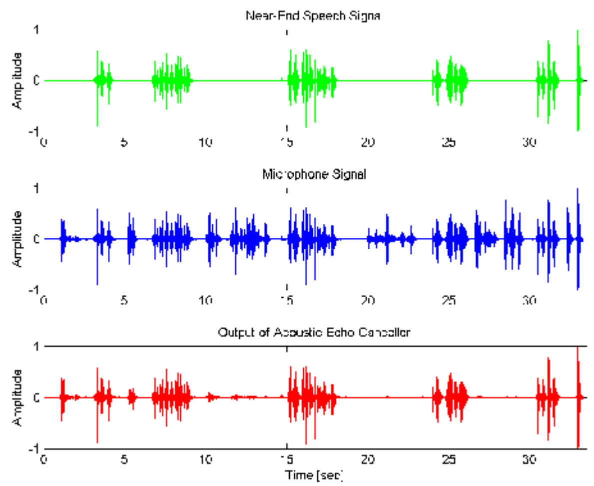


**Fig. 11**. Output of Acoustic Echo Canceller.

## V. CONCLUSION

In this research, Statistical signal Processing for latent variable decorrelates the separating matrix can be deflationary or symmetric orthogonalization. In some applications, it may be preferable to use the fast ICA algorithm with symmetric orthonormalization, in which every vector is impartially treated and the parallel computation of independent components is enabled. Fast ICA algorithm improves the efficiency of independent component analysis. Extensive simulation studies reveal that symmetric approach has a better performance as compared to deflation approach. Source signals are statistically independent, Knowing the value of one of the components does not give any information about the others. ICA algorithm offers many features such as high processing speed, which is extremely desired in many applications. In order to reduce the complexity.

## REFERENCES

[1]. F'utoshi Asano', Shiro Ikeda, Michiaki Ogawa, Hideki Asoh', Nobuhiko Kitawaki A Combined Approach Of Array Processing And Independent Component Analysis For Blind Separation Of Acoustic Signals 1) Electrotechnical Laboratory, 2) PREST, JST, 3) Tsukuba University.

[2]. Virtues and Vices of Source Separation Using Linear Independent Component Analysis For Blind Source Separation Of Non-Linearly Coupled And Synchronised 2001 *Proceedings of the 23rd Annual EMBS International Conference*, October 25-28, Istanbul, Turkey.

[3]. Dr. S. Jayaraman (Asst. Professor), G. Sitaraman, R. Seshadri, Blind Source Separation Of Acoustic Mixtures Using Time frequency Domain Independent Component Analysis. Dept. of Electronics and Communication Engg., P.S.G. College of Technology, Coimbatore-641004.

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

[5]. 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 Ho*tel, Vancouver, BC, Canada July 16-21, 2006.

[6]. Blind Separation Of Noisy Mixed Speech Signals Based On Wavelettransform And Independent Component Analysis Hongyan Li, Huakui Wang, Baojin Xiao *College Of Information Engineering Of Taiyuan University Of Technology, Taiyuan 030024 Icsp2006 Proceeding.*

[7]. 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 2008 *International Conference on Computer Science and Software Engineering*.

[8]. G. Darmois, "generale des liaisons stochastiques," Rev Inst Int Stat. 21, 2-8, 1953. [2] P. Comon, "Independent Component Analysis, *A New Concept," Signal Processing*, vol. **36**, no. 3, pp. 287-314, 1994.

[9]. J. Herault and C. Jutten, "Space Or Time Adaptive Signal Processing By Neural Models," *Proc AIP Conf on Neural Networks for Computing, American Institute of Physics*, pp. 206-211, 1986.

[10]. R. Linsker, "An Application Of The Principal Of Maximum Information Preservation To Linear Systems," Advances in neural information processing systems, Morgan Kaufmann, San Francisco, CA, pp. 186-194, 1989.

[11]. A.J. Bell and T.J. Sejnowski, "An Information-Maximization Approach To Blind Separation And Blind Deconvolution," *Neural Comput* **7**, pp. 1129–1159, 1995.

[12]. S. Amari, A. Cichocki, and H.H. Yang, "A New Learning Algorithm For Blind Signal Separation," Advances in neural information processing systems 8 MIT Press, 1996.

[13]. A. Hyvarinen and E. Oja, "A Fast Fixed-Point Algorithm For Independent Component Analysis," *Neural Comput.* **9**, pp. 1483–1492, 1997.

[14]. M. Gaeta and J.L. Lacoume, "Source Separation Without Prior Knowledge: The Maximum Likelihood Solution," Proc Eusipco'90, pp. 621–624, 1990.

[15]. M. Lewicki and T. Sejnowski, "Learing Overcomplete Representations," Advances in neural information processing systems, vol. 10, MIT Press, Cambridge, MA, pp. 556–562, 1998.

[16]. M. Zibulevsky and B.A. Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural Comput*. **13,** pp. 863–882, 2001.

[17]. F.R. Bach, M.I. Jordan, "Kernel Independent Component Analysis," *Journal of Machine Learning Research*. **3**, pp. 1–48, 2002.

[18]. Michael M. Broristeiti, Alexander M. Bronstein, Michael Zibirlevshy and Yehoshzta Y. Zeevi, "Separation of Reflections Via Sparse Ica," *IEEE conference on Image Processing*, pp. 313–316, 2003.

[19]. Masaki. Yamazaki, Yen-Wei Chen and Gang Xu, "Separation of Reflection Components By Kernel Independent Component Analysis," IJCSNS. 6, 2006.