ISSN:0975-9646

Jaya Kulchandani et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (5) , 2014, 6739-6743

Blind Source Separation via Independent Component Analysis : Algorithms and Applications

Jaya Kulchandani, Kruti J. Dangarwala

Department of Information Technology
Shri S'ad Vidya Mandal Institute of Technology
Bharuch 392-001, Gujarat, India

***Abstract*— Blind Source Separation (BSS) refers to the process of recovering source signals from a given mixture of unknown source signals were in no prior information about source and mixing methodology is known. Independent Component Analysis (ICA) is widely used BSS technique which allows separation of source components from complex mixture of signals based on certain statistical assumptions. This paper covers up the fundamental concepts of ICA and reviews different algorithms of Independent Component Analysis. In addition, the merits and demerits of each algorithm are outlined. Finally brief description of recent application in ICA is presented.**

Keywords—Blind Source Separation; Independent Component Analysis ; Mixing

I. INTRODUCTION

In today's signal processing applications a signal may contain unwanted information or mixture of signals which makes it very hard to get source signals that are of interest. Blind source separation is most common way of recovering source signal of interest from given mixtures of the unknown source signals. The word 'Blind' emphasizes the fact that no prior information regarding source signal or mixing process is known. The strength of the Blind Source Separation model is that only mutual statistical independence between the source signals is assumed and no prior knowledge about the characteristics of the source signals, source from which the signals have originated or the arrangement of the sensors is needed. Over last few years, Blind source separation has become a central topic of discussion in number of fields due to its wide range of applications such as analyzes of EEG or MEG data [1], analyzes of astronomical images[2], telecommunications [3], speech identification [4], image de-noising [5], image and speech encryption [6].

Some of recent research Trends in Blind Source Separation are:

- Blind separation of noisy mixed speech based on Independent Component Analysis and Neural Network[7]
- Underdetermined Blind Source Separation based on Fuzzy C-means and Semi-nonnegative Matrix Factorization[8]
- Blind Separation of Multiple Binary Sources from one Nonlinear Mixture[9]
- Convolutive Blind Source Separation Algorithm based on Higher Order Statistics[10]
- Algorithm for Nonlinear Blind Source Separation Based on Feature Vector Selection[11]
- Acoustic Vector Sensor based Speech Source Separation with Mixed Gaussian-Laplacian Distributions[12]

In this paper, Independent component analysis method of BSS is described in detail. ICA is broadly used Blind source separation method for underdetermined BSS which decomposes a complex data/signal into independent parts. Generative model of ICA makes an assumption that the components within given mixtures are statistically independent as well as non-Gaussian. These components are also known as sources that are extracted by ICA technique.

The reminder of the paper is organized as follows. Section II explains motivation factor behind BSS. Section III provides the introduction to ICA and brief overview of its fundamental concepts. Section IV focuses on different ICA algorithms. Section V presents different applications of ICA and Section IV concludes the paper.

II. MOTIVATION

"Cocktail Party Problem" is the key motivating factor for Blind Source Separation approach. Cocktail problem demonstrates a scenario of loud party full of people. During the conversation between two people the signal received by human ear is mixture of sound from different sources like music instruments, people talking around, some external noise etc…Though humans hear mixed signal, have power to un-mix signals and concentrate on a sole signal/sound. On the other hand machines like microphone and hearing aids gets confused. Such problems signify the need to develop a system that enables to obtain only required source signal. Blind Source Separation deals with problems that are closely related to "Cocktail Party Problem". General scenario of cocktail party problem is illustrated in Figure 1.

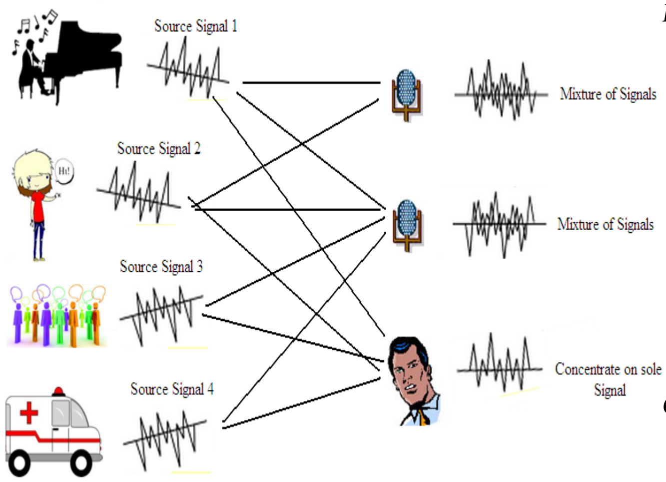

Fig. 1: Cocktail party Problem

III. INDEPENDENT COMPONENT ANALYSIS

Independent component analysis is a signal processing technique whose aim is to state a set of random variables as linear combinations of statistically independent component variables. Interesting information on sensor signals can be revealed by ICA by providing access to its independent components. ICA allows to blindly separation of original source signals only by their mixtures.

A. Problem Formalization

The ICA based BSS problem retrieves unknown source signals by making a simple assumption of **n** independent signals denoted as $\mathbf{s(t) = s_1(t),\ldots,s_n(t)}$ and observed mixture of signals denoted as $\mathbf{x(t)= x_1(t),\ldots,x_n(t)}$, assuming the mixture of signals as linear and instantaneous, mixing equation can be given as :

$$\mathbf{x(t) = As(t)} \tag{1}$$

where **A** is unknown mixing matrix and **t** denotes time instance. Here source signal **s(t)** and the procedure responsible for transformation of source signal into mixed signal is unknown which reflects the "Blindness" property of the problem. If the mixing matrix **A** is invertible i.e. no. of sources N is less than or equal to no. of sensors P (N<=P) then the source can be separated directly and output **y(t)** can be expressed as :

$$\mathbf{y(t) = Bx(t) = BAs(t)} \tag{2}$$

where **B** is a separating matrix or inverse of mixing matrix **A**. The general processing model for BSS can be represented as

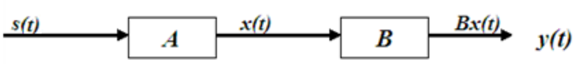

Fig. 2: Processing model

B. Independence

Components independence is the key assumption of ICA model. It states that for a given two variables or source signals s_1 and s_2, are said to be independent if knowing the value of s_1 does not provide any information of s_2.In mathematical term independence can be defined as :

$$\mathbf{Ps_1,s_2(s_1,s_2) = Ps_1(s_1)\, Ps_2(s_2)} \tag{3}$$

Where $\mathbf{Ps_1,s_2(s_1,s_2)}$ is the joint density of s_1 and s_2 , $\mathbf{Ps_1(s_1)}$ and $\mathbf{Ps_2(s_2)}$ are marginal probability densities of s_1 and s_2 respectively.

C. Objective Function

Objective function to ICA is an indicator of independence principle of ICA which allows estimating ICA data model by formulation. Objective function is sometimes also known as contrast function or cost function that can be minimized or maximized to obtain optimum results. The statistical properties and algorithmic properties highly depend upon the type of objective function being used. The different measures of independence used as ICA objective function are: Measure of Non-Gaussianity and Mutual Information (Information about random variables possessed by other member variables of set) [13].

D. ICA Ambiguities

Ambiguities in ICA model can be given as:

1) Varaiances of the independent cannot be deterimined;
Since both **S** and **A** are unknown, any scalar multiplier in one of the source $\mathbf{s_i}$ will get cancelled by diving it with corresponding $\mathbf{a_i}$ of **A**.

2) Order of the independent component cannot be determined :
Since both **S** and **A** are unknown, order of the terms can be liberally changed in the sum and the components can be called in any order.

E. Preprocessing for ICA

Preprocessing is performed in ICA to reduce noise from multidimensional dataset in order to avoid degradation of ICA performance. Moreover pre-processing step before ICA allows simplifying the algorithm as well as reduces the number of parameters to be estimated. Pre-processing steps that need to be performed before application of ICA algorithm can be given as:

1) Centering : It is the process to "center" the observation x by subtracting its mean vector i.e.,

$$\mathbf{x_c = x\text{-}m} \tag{3}$$

where x_c is the centered observation vector and m is the mean vector. Centering allows to simplify ICA algorithm and un-mixing matrix can be estimated by centered data.

2) Whitening: It is the process that allows to remove correlation in the data.Whitening process transforms the observed vector linearly so that a new vector obtained as a result of the process is white ,i.e. components are uncorrelated and varaiances of the component is equal to unity.Whitening process reduces the number of parameters to be estimated; in a way reduces complexity of the problem.

IV. ICA ALGORITHMS

A. MS-ICA

MS-ICA [14] is an algorithm proposed by Molgedey and Schuster, makes use of time-delayed correlation function for source separation process. The task of separation of linearly superimposed uncorrelated signals and its mixing matrix determination process is addressed by eigen value problem. Eigen value problem simultaneously diagonalize (convert square matrix into a diagonal matrix by multiplying it with second matrix and its inverse) two symmetric matrix whose elements are measurable time delayed correlation function. The diagonalzed matrix is determined from objective function which has its number of minima equivalent to number of degenerated solutions. The algorithm allows determining the nonlinearities present in the mixture of signals. Moreover, linear and nonlinear coefficients can be detected from correlation function considering the time delays .The approach can be implemented with minimal user interaction for tuning parameters. The main advantage of this approach is that the solution is simple as well as constructive and use of time-delayed function improves source separation process.

B. Kernel-ICA

Kernel-ICA[15] is an algorithm based on the entire function space which contains candidate nonlinearities and makes use of objective functions based on canonical correlations analysis to reproduce Hibert Kernel space. Once the Hibert Kernel space is reproduced using objective function, "kernel trick" is used to search over the space efficiently. Different sources can be adapted by making use of function and more robust solution is provided to varying source distributions. Kernel ICA methods are significantly more robust to outliers than the other ICA algorithms, including FastICA. Main drawback of Kernal-ICA is that it is relatively slower to other algorithms.

C. RADICAL

The RADICAL (Robust, Accurate, Direct Independent Component Analysis Algorithm) [16] is based on efficient entropy estimates. It estimates the independent sources using differential entropy estimator based on 'm'-spacing estimator. The evolution of the approach was based upon several principles like direct estimation of entropy obviates the need for density estimation as an intermediate step and over the space of smooth densities there are unavoidable local minima in the commonly used K-L divergence based optimization landscape. Initially, algorithm directly minimizes the measure of departure from independence according to the estimated Kullback-Leibler divergence between the joint distribution and the product of the marginal distributions. Then this approach is paired with efficient entropy estimators from the statistics literature. The algorithm based on this estimator is simple, computationally efficient and intuitively appealing. Moreover algorithm can be easily used in higher dimensions. Algorithm takes no advantage of local changes in sorting order due to local changes in rotation. Consequently, the application of standard sorting algorithms for such scenarios would be expected to greatly reduce the computational complexity of the analysis.

D. EGLD

The EGLD [17] is imposed on modeling source distributions. Skewness of the distributions is taken in practice by EGLD which can be considered as main excellence. EGLD covering the whole space of the third and the fourth moments fits for the broader distributions. A marginal distribution is used to estimate source distribution. The score function just like the contrast function looks as a standardization to judge the convergence. The main advantage of this approach is that wide class of source signals like sub- and super-Gaussian can be separated by EGLD-ICA.

E. Fast-ICA

The Fast-ICA [18] belongs to the family of fix-point algorithms for ICA, which is based on the iteration to search for the maximum of the non-Gaussianity of variables. The main advantage of the fixed point algorithms is that their convergence can be shown to be very fast (cubic or at least quadratic). Combining the good statistical properties (e.g. robustness) of the new contrast functions, and the good algorithmic properties of the fixed point algorithm, Fast-ICA provides a very appealing method for ICA. Simulations as well as applications on real-life data have validated the novel contrast functions and algorithms introduced.

F. SICA

[19] Paper introduces a new ICA algorithm for speckle signal. Initially, ICA methods was developed to deal with noise free signals but in speckle environments signals noise is multiplicative; so general noise free model of ICA does not provide effective results in such scenario. SICA was developed to deal presence of speckle noise and obtain the mixing matrix from mixture of signals. To do so, SICA instead of finding non-correlation and non-order correlation between the outputs, finds a specific structure in the second and third order statistic of the output that considers multiplicative model. SICA provides the improved version of ICA for speckle signals.

V. ICA APPLICATIONS

With the emerging trends in ICA, the technique finds its application in number of areas. Some of applications of Blind Source separation via Independent Component Analysis can be given as:

A. Ultrasonic NDE

Ultrasonic NDE (non destructive evaluation) identifies the location of the defect such as cracks, lags and porosity using reflection and transmission properties of ultrasonic waves in material. Moreover, evaluates the size of the defects on the basis of echoed energy [20]. ICA in ultrasonic NDE allows classifying defects, modeling of the system and reduction of noise.

B. Biomedical Applications of ICA

Magneto encephalography (MEG) is a neuro-imaging technique that allows mapping activities of brain by recording magnetic fields produced by electrical current that occurs naturally in brain. In case when MEG record is used for medical purpose, it might become difficult to obtain necessary features from neuromagnetic signals because of presence of artifacts. [21] Paper introduces new method that enables to separate artifacts and brain signals from the mixture of signals using ICA. More successful applications based on ICA can be given as:

- Fetal Electrocardiogram extraction, i.e. removing/filtering maternal electrocardiogram signals and noise from fetal electrocardiogram signals [22, 23].
- Separation of transplanted heart signals from residual original heart signals [24]
- Separation of low level myoelectric muscle activities to identify various gestures [25]

C. Telecommunication

In telecommunication, BSS based ICA technique can be used in CDMA in order to separate user's signal from another user's signal.

D. Audio signal processing

Most audio signals are mixtures of several audio sources (speech, music, noises). Blind Source Separation (BSS) consists in recovering one or several source signals from a given mixture signal. Direct applications include real-time speaker separation for simultaneous translation, sampling of musical sounds for electronic music composition. Many derived applications aim to modify the mixture signal, for example speech enhancement within hearing aids, voice cancellation for karaoke, rendering of stereo CDs on multichannel devices.

E. Image Processing

Recently, number of image processing application implemented using Independent Component Analysis (ICA) have been proposed [26]. Its main is aim is to capture the statistical structure in images from higher-order statistics that cannot be obtained by second order information.[27] Introduces an algorithm based on Independent Component Analysis to detect faces. Approach allows to efficiently discriminate between face and non face images. As the size of the training set is increased along with the size of image better performance is obtained. The goal of ICA in face recognition is to train a system that can recognize and classify familiar faces, given a different image of the trained face.

VI. CONCLUSION

This paper provides basic idea about Blind source separation based on Independent Component analysis. Fundamental concepts of ICA are discussed briefly which includes problem formulation of ICA, objective function, ambiguities in ICA and Pre-processing steps required for ICA implementation. Paper focuses mainly on review of different ICA algorithms that will help new researchers to get basic understanding about algorithms and allow them to make efficient use of algorithms based on their applications. Further, applications of BSS based ICA are presented that show the use of ICA concepts in real world situations.

REFERENCES

[1] Scott Makeig, Tzyy-Ping Jung, Anthony J. Bell, Dara Ghahremani, and Terrence J. Sejnowski. Blind separation of auditory event-related brain responses into independent components. Proc. Natl. Acad. Sci. USA, (94):10979{10984, 1997.

[2] D. Nuzillard and A. Bijaoui, “Blind source separation and analysis of multispectral astronomical images,” Astron. Astrophys. Suppl. Ser., vol. 147, pp. 129–138, Nov. 2000.

[3] L. D. Lathauwer, J . Castaing, Tensor-based techniques for the blind separation of DS-CDMA signals, Signal Process. , 87(2): 322–336, 2007.

[4] M. Z. Ikram and D. R. Morgan, “Permutation inconsistency in blind speech separation: Investigation and solutions,” IEEE Trans. Speech Audio Process., vol. 13, no. 1, pp. 1–13, Jan. 2005.

[5] M. Naceur, M . Loghmari, M. Boussema, The contribution of the sources separation method in the decomposition of mixed pixels, IEEE Trans.Geosci.Remote Sens. , 42(2): 2642–2653, 2004.

[6] Qiu-Hua Lin,Fu-Liang Yin,Tie-Min Mei,Hualou Liang,A blind source separation based method for speech encryption,IEEE Trans. Circuits and Systems .,Volume:53 Issue:6 ,pp .1320 -1328 ,June 2006.

[7] Hongyan Li, Xueying Zhang,“Blind separation of noisy mixed speech based on independent component analysis and neural network”,IEEE Trans. Computing, Measurement, Control and Sensor Network,pp. 105-108 ,2012

[8] Ossama. S. Alshabrawy, W. A. Awad, Aboul Ella Hassanien, “Underdetermined Blind Source Separation based on FuzzyC-Means and Semi-Nonnegative Matrix Factorization”, Proceedings of the Federated Conference on Computer Science and Information Systems, pp. 695–700,2012.

[9] Konstantinos Diamantaras, Theophilos Papadimitriou,Gabriela Vranou, “BLIND SEPARATION OF MULTIPLE BINARY SOURCES FROM ONE NONLINEAR MIXTURE”,IEEE Trans.,pp. 2108-2111,2011.

[10] Hongzhi Wang, Aiqi Bi, Peixin Xu, Can Gao, “Convolutive Blind Source Separation Algorithm based on Higher Order Statistics”, IEEE Trans. Intelligent System Design and Engineering Applications,pp. 487-490 ,2012.

[11] ZHENG Mao, ZHANG Wenxi, ZHENG Linhua, “Algorithm for Nonlinear Blind Source Separation Based on Feature Vector Selection”,IEEE Trans.,pp. 575-578 ,2010.

[12] Xiaoyi Chen, Atiyeh Alinaghi, Xionghu Zhong and Wenwu Wang,“Acoustic Vector Sensor based Speech Source Separation with Mixed Gaussian-Laplacian Distributions”,IEEE Trans.,2013

[13] D. P. Acharya and G. Panda, “A Review of Independent Component Analysis Techniques and their Applications”, IETE Technical Review,vol. 25,issue 6, pp. 320-332, Nov-Dec 2008

[14] L. Molgedey and H. Schuster, “Separation of independent signals using time-delayed correlations,” Physical Review Letters, vol. 72, no. 23, pp. 3634–3637, 1994.

[15] F. R. Bach and M. I. Jordan, “Kernel independent component analysis”, Journal of Machine Learning Research, 3:1–48, 2002.

[16] Erik G. Miller and John W. Fisher III, “Independent components analysis by direct entropy minimization”, Tech. Rep. UCB/CSD-03-1221, University of California at Berkeley, January 2003.

[17] J.Eriksson,J.Karvanen,V.Koivunen,“ Source Distribution Adaptive Maximum Likelihood Estimation of ICA Model”, Proceedings of Second International Workshop on Independent Component Analysis and Blind Signal Separation,Helsinki 2000, pp. 227—232,2000.
[18] A. Hyva¨rinen, “ A Family of Fixed-point Algorithms for Independent Component Analysis”, IEEE Trans.,pp.3917-3920,1997.
[19] D. Blanco, B. Mulgrew, S. McLaughlin,“ICA METHOD FOR SPECKLE SIGNALS”,IEEE Trans.,pp. 821-824 ,2004.
[20] Ying WU, Tian-tai GUO, Bo ZHANG, Lin GUO, Jun-jie WU, Liang ZHAO, “Application of Blind Source Separation in Ultrasonic NDE,Second International Workshop on Education Technology and Computer Science,pp.115-118,2010.
[21] R. Vigário,, J. Särelä,, and Oja, E., “Independent component analysis in wave decomposition of auditory evoked fields”, Proc. Int. Conf. on Artificial Neural Networks, pp. 287-292, Skövde, Sweden, 1998.
[22] E. Niedermeyer and F. L. Da Silva, Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Lippincott Williams and Wilkins; 4th edition , January 1999.
[23] J. C. Rajapakse, A. Cichocki, and Sanchez, “Independent component analysis and beyond in brain imaging: Eeg, meg, fmri, and pet,” in Neural Information Processing, 2002. ICONIP ’02. Proceedings of the 9th International Conference on, vol. 1, 2002, pp. 404–412 vol.1.
[24] J. Wisbeck, A. Barros, and R. Ojeda, “Application of ica in the separation of breathing artifacts in ecg signals,” 1998.
[25] S. Calinon and A. Billard, “Recognition and reproduction of gestures using a probabilistic framework combining pca, ica and hmm,” in ICML ’05: Proceedings of the 22nd international conference on Machine learning. ACM, 2005, pp. 105–112.
[26] A. J. Bell and T. J. Sejnowski, “The "independent components" of natural scenes are edge filters.” Vision Res , vol. 37, no. 23, pp. 3327–3338, December 1997.
[27] Vaidehi, Annis Fathima, Teena Mary Treesa, Rajasekar , Balamurali and Girish Chandra, “An Efficient Face Detection and Recognition System”,IMECS,Proceedings of the International MultiConference of Engineers and Computer Scientists,vol I,,March 2011