International Journal of Image Processing and Vision Science

Volume 2 | Issue 1 Article 10

January 2013

Source Separation using ICA

P. Shivani Sahoo
ECE Dept.,ITER,Bhubaneswar, pshivanisahoo@gmail.com

A. Barik
ECE Dept.,ITER,Bhubaneswar, abarik@gmail.com

Padmini Mishra
Dept. Of ECE., GIET, Gunupur, Odisha, India, padmini@gmail.com

Source Separation using ICA

P.Shivani Sahoo[1], A.Barik[2], Padmini Mishra[3]

[1,2]*ECE Dept.,ITER,Bhubaneswar*

pshivanisahoo@gmail.com

[3] *Assistant Professor, Dept. Of ECE., GIET, Gunupur, Odisha, India*

Abstract — Independent Component Analysis (ICA) is a statistical signal processing technique having emerging new practical application areas, such as blind signal separation such as mixed voices or images, analysis of several types of data or feature extraction. Fast independent component analysis (Fast ICA) is one of the most efficient ICA technique. Fast ICA algorithm separates the independent sources from their mixtures by measuring non-gaussianity.In this paper we present a method that can separate the signals as individual channels from other channels and also remove the noise using fast ica algorithm. The method is to decompose a multi channel signal into statistically independent components.

.

Keywords — **ICA, Fast ICA,BSS**

I. INTRODUCTION

Independent component analysis (ICA) is a new technique to statistically extract independent components from the observed multidimensional mixture of data. Many successful examples of ICA application in the field of signal processing are reported recently. Independent component analysis (ICA) was originally developed to deal with problems that are closely related to cocktail- party problems.ICA is a powerful and useful statistical tool for extracting independent source given only observed data that are mixtures of the unknown sources. ICA has been studied by many researchers in neural networks and statistical approaches during the past 10 years. Independent component analysis is a signal processing technique whose goal is to express a set of random variables as linear combinations of statistically independent component variables.

By ICA we can separate the original sources blindly only by their mixtures.

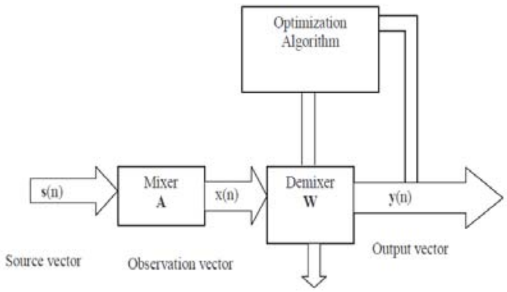
.

II. MATHEMATICAL MODEL FOR SOURCE SEPARATION

ICA can be applied to a collection of stastically independent sources which are recorded from different users. Suppose we have N statistically independent signals, $s_i(t)$, where i=1,2,……,N. where N is the number of source signal. This is represented compactly by the mixing equation

$x(t)=As(t)$ …..(1) where s(t) is a column vector collecting the source signals, vector x(t) similarly collects the N observed signals, and the square n×n "mixing matrix" contains the mixture coefficients. The ICA problem exists in recovering the source vector using only the observed data the assumption of independence between the entries of the input vector, and possibly some prior information about the probability distribution of the inputs. The objective is to recover the original signals, $s_i(t)$, from the observed vector x(t). we obtain estimates for the sources by first obtaining the ' unmixing matrix' W, where $W=A^{-1}$………(2)

This enables an estimate, y(t), of the independent sources to be obtained. $Y=Wx$….(3) where the time index t has been omitted for notational simplicity



BASIC BLOCK DIAGRAM OF THE BSS

III. INDEPENDENT COMPONENT ANALYSIS

ICA involves estimating the linear transformation that maximizes the independence of the signals [8]. This linear transform is reffered to as the unmixing matrix, W. since the original sources , $s_i(t)$, were assumed to be independent, we know that maximizing the independence of the components of y from equation(3) we will obtain estimates of the original sources[9].

At this point it is important to place this intuitive notation of ICA on a more rigorous foundation. Suppose the observation vector x is formed according to equation(1), that is, x is a linear combination of independent components. To estimate the one of the independent components we consider a linear combination of the x_i terms. Let y_i be one of the estimates, then we have:

$y_i=W^T x$…………(4)

A. DATA PREPROCESSING FOR ICA

The preprocessing consists further of two steps known as centering and whitening. The centering step is done by sub-tracting the mean of the observed data x, so as to make x a zero-mean variable. A whitening step is done to remove the correlation between the components of the observed data. One popular method for whitening is to use the eigenvalue de-composition (EVD) of the covariance matrix of the mixed sig-nal. The final step is the FastICA algorithm which is brief. 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.

B. ICA ALGORITHMS

In this section, two of the most important ICA algorithms are presented in some detail. Oja and Hyv¨arinen'sFastICA algorithm[11] and Bell and Sejnowski's information maximization algorithm. These two algorithms are probably the most widely used and they each illustrates the important principle of ICA.

C. FastICA algorithm

The algorithm is based on using non-Gaussianity as a metric for independence as discussed. FastICA is based on using entropy as a measure of non-Gaussianity. A fundamental result of information theory is that a Gaussian random variable has the greatest entropy of all random variables of equal variance.

As a result, entropy can be used as measure of non-Gaussianity. To be precise FastICA is not based on entropy, but on negentropy. Negentropy defined by

$J(y)= H(y_{gauss})-H(y)$…………………(6)

Where H(.) denotes the entropy of a random variable, J(.) denotes negentropy and $y_{gauss}$is a Gaussian random vectorwith the same covariance matrix as y. negentropy is always non-negative since

$H(y_{gauss}) \geq H(y)$.

fastICA is a fixed point algorithm that maximizes negentropyusing newton's iterative method, the FastICA algorithm is given below

1. Choose an initial (e.g. random) weight vector W.
2. Let W+=E{Wg(WTX)- E{Wg'(WTX)}W
3. Let W=W+/‖W‖
4. If not converged, go back to 2.

The function g can be almost any non quadratic function but hyperbolic tangent function s have been shown to behave well in practice.

IV. RESULTS

Two speech signals are recorded and considered as the source signals. The order in which the outputs are recovered could not be predicted. The variances of the source signals are assumed to be equal to one.

Two signals, mixed and then separated using the FastICA algorithm.

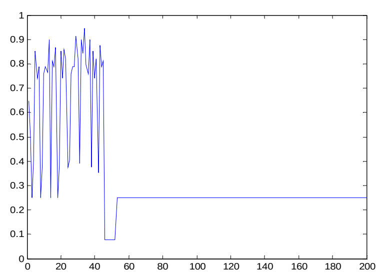


Fig. 3 Source signal 1

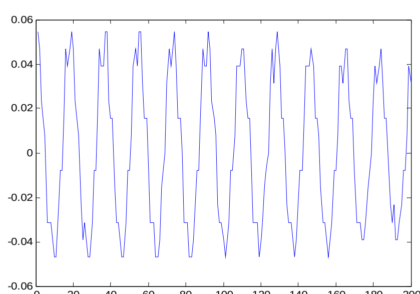


Fig .4 Source signal 2

The recovered signals are obtained by applying ICA algorithm in the demixer side.The recovered source signal for the source inputs are given in fig 5 and fig 6 respectively.

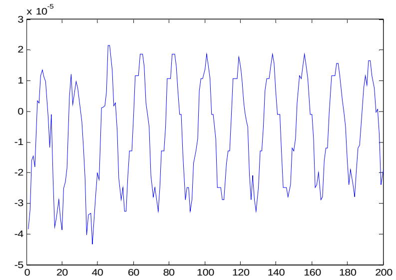


Fig.5 Recovered signal for source signal 2

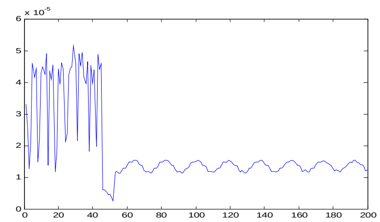


Fig. 6 Recovered signal for source signal 1

V. CONCLUSIONS

In this paper, two souce signals taken the preprocessed and mixed together and then the mixed signal separation using the FastICA algorithm is shown .The result shows that the Fast ICA algorithm is an effective tool for denosing of mixed signal. The error signal between the original and recovered signal decreases by increasing the number of iterations.

REFERENCES

[1] A.Hyv¨arinen, J. Karhunen, and E. Oja, Independent component analysis . John wiley and sons, Inc., 2001.

[2] R. Lambert, "multichannel blind deconvolution: Fir matrix algebra and separation of multipath mixtures," Ph.D.dissertation, University of Southern California,1996.

[3] A.Hyv¨arinen and E. Oja, "A fast fixed point algorithm for independent component component analysis", Neural computation, vol.9, no.6, pp.1483-92, 1997.

[4] Hyvärinen,A. (1999a). Fast and robust fixed-point algorithms for independent component analysis. IEEE Transactions on Neural Networks, 10(3):626–634

[5] Cardoso, J.-F. (1997). Infomax and maximum likelihood for source separation. IEEE Letters on Signal Processing,4:112–114..