International Journal of Scientific Engineering and Technology
Volume No.4 Issue No.4, pp: 264-267

(ISSN : 2277-1581)
01 April. 2015

Blind Source Separation using ICA for Additive Mixing in Time and Frequency domain

Swapnil Mohan Mahajan[1], Suneet Kishore Betrabet[2]
Communication Engineering,Vellore Institute of Technology, Chennai
[1]Swapnil.mohan2013@vit.ac.in, [2]suneet.kishore2013@vit.ac.in

Abstract: This paper presents BSS for additive mixing where every recordings consist of differently weighted signal. Therefore, by using ICA for both time-domain and frequency-domain, we are going to separate source signals from mixed signal. The aim of our analysis is to perform undetermined convolutive BSS via frequency bin-wise clustering and permutation alignment where convolutive mixture are most delayed and weighted. So, ICA in time-domain is fails to separate signals. Hence, instead of this we use ICA in frequency-domain which playing vital role in separation of audio signals by using MATLAB which is our future work.

Keywords— Compressive Sensing; Sparsity; GPSR; K-means; L1-magic.

I. INTRODUCTION

A. Problem Description

The Cocktail Party Problem means several sources are mixed together and it is difficult to find out what is the original signal were. And the concept of blind source separation [1] additionally called blind signal separation is that the separation of a group of signals from a group of mixed signals, while not the help of knowledge (or with little information) concerning the source signals or the blending method. Blind signal separation depends on the belief that the supply signals don't correlate with one another. For instance, the signals are also reciprocally statistically independent or decorrelated. Blind signal separation so separates a group of signals into a group of different signals, specified the regularity of every resulting signal is maximized, and therefore the regularity between the signals is decreased.

B. Preliminary Work

Blind separation and blind deconvolution are related issues. From fig it is clear that a set of sources $S_1(t),\ldots,S_N(t)$ (different peoples speaking, music, etc) are mixed together by a matrix A. We are aware about the sources and mixing process. All we receive is the N mixed signals, $X_1(t),\ldots,X_N(t)$. So, our aim is to recover original signal from mixed signal by finding a square matrix. In blind deconvolution [1], an unknown signal s(t) is convolved with an unknown tapped delay-line filter, $a_1,\ldots,a_L$ giving a corrupted signal $x(t)=a(t)*s(t)$ where a(t) is the impulse response of the filter. So, the task is to recover s(t) by convolving x(t) with a learnt filter $w_1,\ldots,w_L$ which reverses the effect of filter a.

So, in simple way from fig 1, the two sources “Hello” and “Morning” are getting added and convolved with respective of impulse responses to the both sensors (microphone). At last we get the separated original sources with the help of unmixing matrix W.

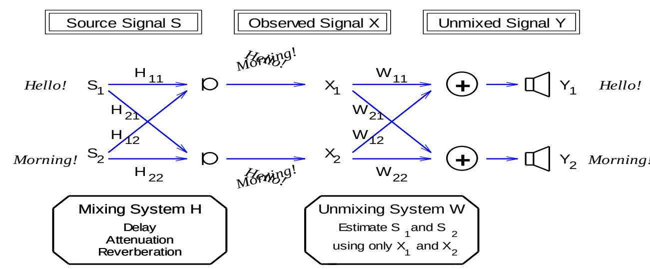

Fig 1 A diagrammatic representation of a mixing and an unmixing system

II. METHODOLOGY

The Blind source separation (BSS) [1] has received nice attention owing to its numerous application fields such as array processing, speech recognition, communications, feature extraction, image processing, and biomedical analysis. Independent Component Analysis (ICA) [2], [3], [4] is one of the ways used for blind source separation. The goal of ICA [2], [3], [4] is to get recover independent sources given sensor outputs within which the sources are linearly mixed. ICA not solely decorrelates the signal however reduces the higher-order statistical dependencies, trying to make the signals as independent as possible.

The general assumptions in BSS [1] regarding the sources are:

- The source signals are unit statistically independent and are stationary processes.
- At the most one of the source signals is Gaussian distributed.
- It is only possible to recover the source signals modulo scale, and polarity.

In signal process, independent component analysis (ICA) [2], [3], [4] could be a procedure technique for separating a variable signal into additive subcomponents by assumptive that the subcomponents are non-Gaussian signals which they're all statistically independent from one another. ICA could be a special case of blind source separation. The BSS [1] based on independent component analysis (ICA) [2] are often classified into two groups in terms of process domain such as,

A. *ICA in time-domain*

B. *ICA in frequency-domain*

ICA in time-domain [3] and ICA in frequency-domain [4] are combined can do a superior separation of the sources under reverberant condition i.e., in acoustic space. But, generally

International Journal of Scientific Engineering and Technology (ISSN : 2277-1581)
Volume No.4 Issue No.4, pp: 264-267 01 April. 2015

speaking, conventional time-domain ICA is fails to separate sources under heavily reverberant conditions as a result of the low convergence within the low iterative learning of the inverse of mixing process. On the opposite hand, frequency-domain ICA also degrades significantly because the independence assumptions of narrow-band signals collapses once the number of sub-bands will increases. So, our main aim is to separate the signals from mixed signal in reverberant condition i.e., by using convolutive mixing [5], but before that we see the algorithms for additive mixture using time-domain ICA [3] and frequency-domain ICA [4].

III. ADDITIVE MIXING

A. ICA in time-domain

Let us consider m independent signals $s_{1,\ldots,}s_M$ summarized in a vector $s=(s_{1,\ldots,}s_M)^T$, where T denotes the transposition. The m independent sources generate signals $s(t)$ at discrete times $t = 1, 2, \ldots$. Let us assume that we can observe only their n linear mixtures, $x=(x_{1,\ldots,}x_N)^T$

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

Or, in component form,

$$x_i(t) = \sum_k (A_{ik} * s_k(t)) \tag{2}$$

From the given observed signal, x(1),…..x(t),we would like to recover s(1),….,s(t) without knowing the matrix A. When n = m, the matter reduces to on-line estimation of A or its inverse, W.

We know that, x= As

and y= wx

Therefore, y= wAs

To minimize cost function [3], let us first consider a candidate A of the mixing matrix within the overcomplete case and put

$$y = A^T x \tag{3}$$

So, $$x = Ay \tag{4}$$

Here, y is an estimate of original signal s. However there are number of y satisfying equation which does not give original s even when A is the true mixing matrix.

Let us consider, the probability density function p(y,A) of y determined by A ϵ $S_{m,n}$. So, our target is to make the parts of y as independent as possible. Let us, choose an independent distribution of y

$$q(y) = \prod_{a=1}^{m} q_a(y_a) \tag{5}$$

By using Kullback divergence [3] between two distributions p(y,A) and q(y), we can reduce cost function

$$CA = KL[p(y,a):q(y)]$$
$$= \int p(y,a).\log \frac{p(y,a)}{q(y)} dy \tag{6}$$

This shows that how far the p(y, A) is from the independent distribution q(y) and is reduced by y = A^Tx are independent under a certain condition. The entropy term is,

$$-H = \int p(y,a).\log p(y,a) dy \tag{7}$$

does not depend on A. Hence, this is equivalent to following cost function term,

$$C(A) = -E[\sum_{a=1}^{m} \log q_a(y_a)] - c \tag{8}$$

Where, c is the entropy of A.

The gradient [3] of

$$l(y,a) = -\sum \log q_a(y_a) \tag{9}$$

is calculated by

$$dl = \varphi(y)^T dy = \varphi(y)^T dA^T x \tag{10}$$

Where, $\varphi(y)$ is called activation function and as speech is super-Gaussian distribution,

$$\varphi(y) = [\varphi_1(y_1), \ldots, \varphi_n(y_n)]$$
$$\varphi_i(y_i) = -\frac{d}{dy_i} \log q_i(y_i)$$

and $dy = dA^T x$

is used. We then have ordinary gradient

$$\nabla l = x\varphi(y)^T = Ay\varphi(y)^T \tag{11}$$

But, the steepest descent direction of the cost function C [3] is given by natural gradient [3] $\tilde{\nabla} l$, which is given as

$$\tilde{\nabla} l = \nabla l - A(\nabla l)^T A$$
$$= A(y\varphi(y)^T - \varphi(y)y^T A^T A \tag{12}$$

Hence, the gradient [3] is

$$\nabla l = \varphi(y) x^T \tag{13}$$

Therefore, by using this natural gradient [3] is

$$\tilde{\nabla} l = \nabla l - W\{\nabla l\}^T W$$
$$= \varphi(y)x^T - y\varphi(y)^T W \tag{14}$$

The learning rule [3] is

$$\nabla W_t = -n_t \tilde{\nabla} l = -n_t \{\phi(y_t) x_t^T - y_t \phi(y_t)^T W_t\} \tag{15}$$

In this way, when n=m and no prewhitening process takes place, we get the natural gradient [3] formula which is given by

$$\tilde{\nabla} l = (I - \phi(y) y^T) W \tag{16}$$

This is the algorithm which we followed.

Result

At last, we are able to separate original signals from mixed signal. To implement this we used two source signals i.e., male 1 and male 2.

1. *Original signals:-*

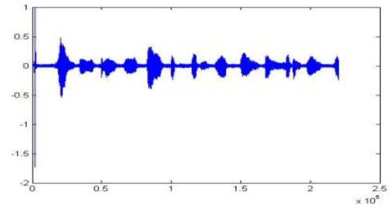

Fig.A.1 Original male 1

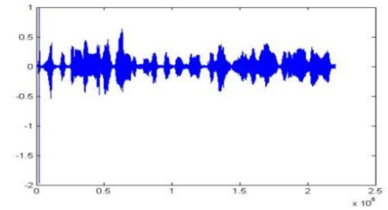

Fig.A.2 Original male 2

2. *Separated signals:-*

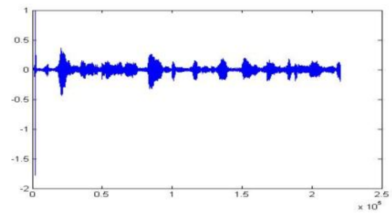

Fig.A.3 Separated male 1

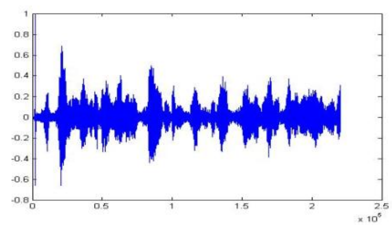

Fig.A.4 Separated male 2

B. ICA in frequency-domain

To overcome the problems which we are facing in ICA in time-domain, we used ICA in frequency-domain [4]. As we know that the convolution in time-domain is the multiplication in frequency-domain.

Therefore, $$X_i(\omega) = \sum_k A_{ik}(\omega).S_k(\omega) \tag{17}$$

Where, $X_i(\omega)$, $A_{ik}(\omega)$ and $S_k(\omega)$ are the Fourier transform of x(t), h(t) and s(t) respectively.

Here we divide the signals into number of frames of say 40 ms. The difficult problem during this is the ambiguities [4] of permutation and amplitude. So to resolve this drawback we tend to used envelope of detector. Therefore, we propose a technique based on windowed-Fourier transform [4] i.e., known as spectrogram [4].

Spectrogram [4] is a visual illustration of the spectrum of frequencies as they vary with the time or another variable. Graph shows: the horizontal axis is time, the vertical axis is frequency; a third dimension indicating the amplitude of a particular frequency at a particular time.

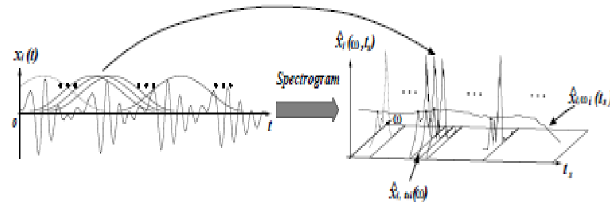

Fig B.1 The windowed-Fourier transform

If the delay and reflections aren't too long, we are able to ignore the convolutions. However, the new drawback will occurs i.e., ambiguities [4] of amplitude and permutations. Therefore, we have to remove these ambiguities so as to reconstruct the signal. Our plan is to use the inverse of the decorrelating matrices and also the envelope of the speech signal. This can be possible because of the temporal structure of the acoustic signal that it is stationary for a short period but not stationary for a long term.

Therefore by following windowed-Fourier transform [4],

$$X(\omega, t_s) = A(\omega).S(\omega, t_s) \tag{18}$$

Where, $X(\omega,t_s)$ and $S(\omega,t_s)$ are the windowed-Fourier transform [4] of x(t) and s(t). The windowed-Fourier transform is defined as,

$$X(\omega, t_s) = \sum e^{-j\omega t} x(t)\omega(t - t_s) \tag{19}$$

Where, ω is the frequency, N is the number of points of discrete Fourier transform, t_s is the window position.

Let us define $X(\omega,t_s)$ and $S(\omega,t_s)$ for a fixed frequency ω as $X_\omega(t_s)$ and ω as $S_\omega(t_s)$. Equation (18) becomes

$$X_\omega(t_s) = A(\omega)S_\omega(t_s) \tag{20}$$

It shows that convolutive mixtures are an instantaneous mixture for fixed frequency ω. Therefore, we can apply any ICA algorithm for each frequency and easily separate the signals. So, we have a separated time sequence for each frequency.

$$u_\omega(t_s) = B(\omega)X_\omega(t_s) \tag{21}$$

By aligning this sequence for every frequency and by applying inverse Fourier-transform, we are able to reconstruct separated signal. But, ICA [4] cannot solve the ambiguity of amplitude and permutation. So, we tend to see the way to solve those two problems.

- The drawback of irregular amplitude are often solved by putting back the separated independent components to the sensor device input with the inverse matrices $B(\omega)^{-1}$.
- Due to the similarities of independent components in many frequencies, these components are classified and also the permutation is solved.

Result

Here, we successfully separated the signals form mixed signal.

1. *Mixed Signal:-*

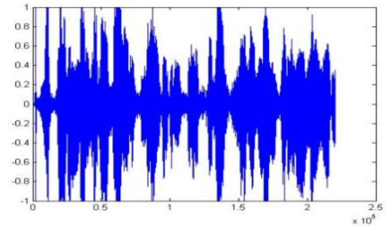

Fig. B.2 mixed signal 1

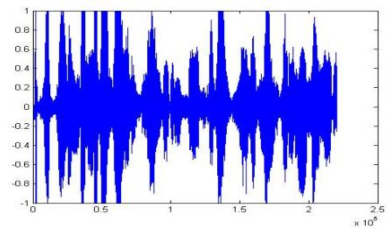

Fig. B.3 Mixed signal 2

2. *Separated Signals:-*

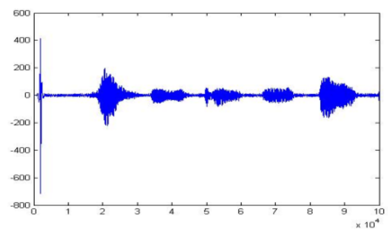

Fig. B.4 Separated male 1 signal

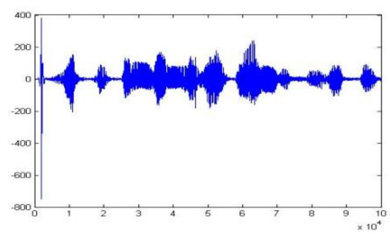

Fig. B.5 Separated male 2 signal

IV. RESULT ANALYSIS

As we know that ICA in time-domain is fails to separate the signals under reverberant conditions because of delayed and weighted signals, we are going to use ICA in frequency-domain [5].

The BSS technique is applied to an undetermined [5] case wherever the number of microphones are less than sources. The separation technique is performed within the frequency-domain that comprises two stages.

- In the very first stage, the mixture of frequency-domain samples is clustered into each source by an expectation-maximization (EM) algorithm [5].
- Since the cluster is completed in a very frequency bin-wise manner, the permutation ambiguities [5] of these bin-wise clustered samples ought to be aligned. This is going to solved in second stage by using the probability on how likely each sample belongs to the assigned class.

So, to separate such an undetermined mixtures, we uses T-F masking [5]. Also the EM algorithm [5] is an iterative methodology for locating most chance or most posterior estimates of parameters in statistical models.

The advantage of this two stage structure of clustering part improves the separation of signals from mixtures compared with the methods based on TDOA. Also, this permutation alignment technique performs better than a traditional technique based on amplitude envelopes.

CONCLUSION

Thus, we got the separated signals from mixed signal by applying ICA method in time-domain and frequency-domain for the case of additive mixing. We also solved the ambiguities of amplitude and permutation in frequency-domain. Our future study will include the undetermined convolutive blind source separation using ICA in frequency-domain contains frequency bin-wise clustering and permutation ambiguities which improves the separation performance.

ACKNOWLEDGMENT

This project is result of the hard work put in by us. This project has been a great learning experience for us and its timely completion would not have been possible without the help our school faculties and facilities provided by VIT management.

REFERENCES

i. *Michael Syskind Pedersen, Jan Larsen, Ulrik Kjems, and Lucas C. Parra, "A survey of convolutive blind source separation methods.", Springer Handbook on Speech Processing and Speech Communication, pp. 1-34.*

ii. *Tsuyoki NISHIKAWA, Hiroshi SARUWATARI, Kiyohiro SHIKANO, "Comparision of Time-Domain ICA, Frequency-Domain ICA and Multistage ICA for Blind Source Separation" Graduate School og Information Science, Nara Institute of Science and Technology, 8916-5,Takayama-cho, Ikoma-shi, Nara, 630-0101, JAPAN.*

iii. *Shun-ichi Amari - RIKEN Brain Science Institute, Wako-shi, Hirosawa, Saitama 351-01, Japan:-" Natural Gradient Learning for Over- and Under-Complete Bases in ICA" Neural Computation 11, 1875–1883 (1999) .*

iv. *Shiro Ikeda and Noboru Murata, "A method of ICA in time-frequency domain," In Proceedings of International Workshop on Independent Component Analysis and Blind Signal Separation (ICA'99) , pp.365-371, Aussions, France, January 1991.*

v. *Hiroshi Sawada,Shoko Araki and Shoji Makino, Fellow, IEEE:-"Underdetermined Convolutive Blind Source Separation via Frequency Bin-Wise Clustering and Permutation Alignment":-IEEE Transaction paper on audio,speech and language processing (2010)***.**