Research Journal of Applied Sciences 9 (5): 262-273, 2014
ISSN: 1815-932X

Blind Techniques for Improving Speech Mixtures Using an Adaptive Method

[1]Jasmine J.C. Sheeja and [2]B. Sankara Gomathi
[1]JCT College of Engineering and Technology, Pichanur ECE, Anna University, Chennai, India
[2]National Engineering College, K.R. Nager, Kovailpatti, Tamil Nadu, India

Abstract: Speech signals are recorded by a distant microphone in an enclosed space, frequently tainted by interferences of many speech signals and reverberations of room. Accordingly, the separation of speech signals is important and allows the process of Blind Source Separation (BSS) and Blind De-reverberation (BD). The signified frequency-domain BSS and Independent Component Analysis (ICA) are utilized for separation process. Subsequently, the permutation ambiguities of the ICA solutions are prearranged and the separated signals are shaped accurately in the time domain. While most speech enhancement algorithms improve the excellence of speech but not increase speech transparency in reverberation. This motivates the improvement of an algorithm that can be modified for an acoustic environment and recover speech transparency. This study proposes the Adaptive Equalization Method for dereverberation and Savitzky-Golay filtering is used to reduce the noise.

Key words: Blind source separation, independent component analysis, blind dereverberation, adaptive equalization, Constant Modulus algorithm

INTRODUCTION

The sources emitted in a room as acoustic signals are recorded with microphones, signals are often coupled between source signals present in that room. The source signals are reverberated and so the signals are deformed and less intelligible. The speakers are in a little distance from microphone in a room. Hence, the mixtures of speech signals are observed at the microphones called convolutive mixtures. The convolutive mixtures have intelligible speech signals, detrimental effect on the apparent quality due to reverberation in closed room, which damages the effectiveness of automatic speech recognition (Nakatani *et al*., 2007; Furuya and Kataoka, 2007). Due to the interference and the occurrence of reverberations among the speech signals, the separation of speech signals from the distant microphone turns into a more complex and challenging process. Therefore, to recover a speech signals from reverberate and interfering signals and also to reduce the adverse acoustic effect a new technique is estimated. Recovering and separating source signals from the captured microphone signals, namely blind source separation and blind dereverberation of the convolutive speech mixtures will be useful for many wide ranges of audio applications like biomedical signal analysis, multiple-antenna wireless communications, acoustic, speech processing, etc.

Blind source separation: The task of Blind Source Separation (BSS) is to separate the sources using the known signals. Numerous BSS algorithms have been developed for speech separation. Many of the probable methods rely on the sparsity of signals in some domain such as the Time Frequency (TF) domain (Bao *et al*., 2013). The separations of source speech signals are executed on the principle of the convolutive speech mixtures with the supposition of mutual statistical independence between the source signals and so, it is named as a Blind Method. Blind source separation is a complex computational practical math problem in past few year research interests. The source separation technique has a prior knowledge of signal processing theories and math model to begin the study of current BSS Theory and methods. The signal-processing theories like signal decomposition, sampling, amplitude, frequency domain, time domain and filter banks and math models from statistics, linear algebra and probability are used in BSS Methods. BSS Method is a mixed process that is commonly defined with a mathematical model named Multiple-Input Multiple-Output (Huang *et al*., 2005). In the Collaborative Blind Source Separation (CBSS) technique, based on the location of the mixture components are converted to their corresponding sources using a pair of coincident microphone arrays. This assumes that more than two speech source must not

Corresponding Author: Jasmine J.C. Sheeja, JCT College of Engineering and Technology, Pichanur ECE, Anna University, Chennai, India

contribute the one time-frequency instant in the mixture. Based on the probable source pairs for first simultaneous microphone array, their corresponding estimated Direction of Arrival (DOA) for the second coincident microphone array is predictable (Zheng *et al*., 2013). Consider a two input and output BSS problem of convolutive mixtures of speech in the frequency domain. The performance of frequency-domain BSS in long reverberant environments is poor (Araki *et al*., 2003). There are assortments of potential multi input multi output applications of BSS in various fields including acoustic signal processing, biomedical signal processing, telecommunications, satellite image astronomical analysis, etc. In a reverberant environment, the acoustic signals are recorded and BSS is focused here to recover and separate the source signals. Generally, in each microphone, delayed version of signals and multiple attenuate are detained for speech signals in the blind separation process. This problem is a key challenge to separate source speech signals in various applications like mobile telephony or conferencing; automatic speech recognition and human intelligibility. The source speech signal and separation criteria capitalization have some certain properties to be required in BSS technique estimation. The BSS technique exploits an important attribute, the inherent non-stationarity of source speech signals. There are various, BSS based algorithms to explain the non-stationarity. Speech signals have been considered to the convolutive speech signal mixtures, the algorithms based on Time-Domain and Frequency-Domain Methods (Nion *et al*., 2010). However, the permutation ambiguity of the ICA solution becomes an important problem. There must be aligning the permutation of each frequency bin so that a separated signal in the time domain includes frequency components from the same source signal (Sawada *et al*., 2004).

Blind dereverberation: The speech signals are reverberated due to the distance between the microphones and the speech signals. The multiple diffracted and deflected signals obtained in enclosed acoustic environment also cause reverberation. The effect of this deformation in source signal is reduced by blind acoustic dereverberation, this concentrates on acoustic signal's spatial diversity (Habets *et al*., 2008). Speech intelligibility turns risky when the signal is reverberated. The speech signal frequency get masked by increases low-frequency energy, change in amplitude modulations correlated with frequency of speech, spectral cues, flattens formant transitions and blurs temporal. To recover and reconstruct the deformed source speech signal without any prior knowledge about room reverberation and its acoustical properties, well-known technique named blind dereverberation or blind reverberation cancelation is used (Kokkinakis and Loizou, 2009). In some cases, the dereverberation of speech signals for difficult acoustic scenarios. They are as follows (Haque *et al*., 2011):

- The microphone receiver receives unknown source speech signals with convolutive mixtures, so signals cannot be trained to calculate the impulse responses of long acoustic and consist of thousands of coefficients. Hence, dereverberation is called a blind problem
- The dereverberation problem is increased by degree of complexity in enclosed noisy environment as the source signals doesn't hang about collinear with the least mean-squared error solution
- Null values present in the frequency response of acoustic impulse responses which produce considerable noise amplification during the process of channel equalization
- The source speech speaker's small movement can produce the considerable changes in acoustic impulse responses in order of the tenth of the acoustic wavelength

Even though, there are a number of rich dereverberation techniques estimated; only some of them are considered as time varying and noise channels concurrently. Some of them are multi-microphone dereverberation, spectral subtraction, Single Input Multi Output Inverse Filtering Method, reduction of spectral coloration, pre-trained harmonic filters, etc. In digital signal processing the optimal filtering of the various unforeseen signal, noise or time-varying signals is not acquired by a two FIRs and IIR filter of fixed coefficients. Under such conditions adaptive filters are used to trace the changes of signal and noise (Borigsar and Kulkarni, 2010).

An adaptive equalizer is an equalization filter that repeatedly adjusts to time-varying properties of the channel. It is a filter that self-adjusts its transfer function according to an Optimizing algorithm (Malik and Sappal, 2011) There are certain methods such as Least Mean Square (LMS), Normalised Least Mean Square (NLMS) and Recursive Least Square (RLS) algorithm for noise cancellation using error correction coding. The experiments show acceptable outcomes in severely faded Nakagami-m channels. This creates a method for developing assured insight into the use of error control coding and adaptive filtering to scrap fading in the wireless channel (Das and Sarma, 2012). In this research, the enhanced speech reverberation by adaptive method can be analyze and tested by the MATLAB functions. For noise reduction technique the pre-defined MATLAB noise filter which is Savitzky-Golay filter is used. Both the

Blind Source Separation (BSS) and Blind Dereverberation (BD) are jointly used to separate the source signal and de-reverberate that captured by distant microphone in enclosed room.

LITERATURE REVIEW

A number of methods are being addressed by various researchers in the past in this study, some of the methods related to recovering and separations of acoustic methods are provided in this study.

Nion *et al*. (2010) presented a frequency-domain technique based on Parallel Factor (PARAFAC) analysis that achieved a multichannel Blind Source Separation (BSS) of speech mixtures. PARAFAC algorithms were shared with a dimensionality reduction step to significantly abridged computational complexity. The identifying ability potential of PARAFAC was exploited to obtain a BSS algorithm for under-determined case (more speakers than microphones), combining PARAFAC analysis with time-varying Capon beam forming. Finally, a low-complexity adaptive version of the BSS algorithm was projected which can track alterations in the mixing environment. Signal to interference ratio improvements of up to 6 dB were shown compared to state of the art BSS algorithms, at a class of magnitude lesser computational complexity. The disadvantage is that there is slow convergence due to ill conditioned data. And, perfect separation of signal is not theoretically possible.

Habets *et al*. (2008) presented the novel spectral variance estimator which was the resultant for the late reverberation of the near-end speech. A spectral variance estimator for the so-called late residual echoes that resulted from the deficient length of adaptive filter was derived. The model parameters depended on the reverberation time of the room which could be obtained using the estimated acoustic echo path. A novel post filter was developed which suppressed late reverberation of the near-end speech, residual echo and background noise and sustained a constant residual background noise level. Experimental consequences demonstrated the beneficial use of the developed system for reducing reverberation, residual echo and background noise. The demerit of this research is if the distance of the speaker and the microphone increases there is the spectral distortion and it is not sensitive for these cases.

Kokkinakis and Loizou (2009) considers a fresh approach for two-microphone enhancement of speech ruined by reverberation. Their approach steered computational resources to filter coefficients having the largest brunt on the error surface and therefore, only updated a subset of coefficients in all iterations. Experimental results carried out in a practical reverberant setup indicated the performance of the projected algorithm was equivalent to the performance of its full-update counterpart. In this method, the value of the speech originality must be improved by some other techniques. Comparatively, adaptive equalization technique gives better results.

Haque *et al*. (2011) addressed the problem of speech dereverberation considering a noisy and gradually time-varying atmosphere. The expected Multi-Microphone Speech Dereverberation Model utilized the estimated Acoustic Impulse Responses (AIRs) to de-reverberate the speech as well as enhanced the signal to noise ratio. The Eigen filter was efficiently computed to avoid the tedious cholesky decomposition, solely from the approximation of the AIRs. The design of the Eigen filter also incorporated a frequency-domain constraint that improves the quality of the speech signal, resisted spectral nulls in addition to develop the Signal to Noise Ratio (SNR). A Zero Forcing Equalizer (ZFE) was used to de-reverberation of speech signal by eliminating the distortion caused by the AIRs as well as the Eigen filter. The ZFE was employed in block-adaptive form which made the projected technique appropriate for speech dereverberation in a time-varying situation. The simulation results established the superior performance of the expected method as compared to the state of the art dereverberation techniques in provisions of Log-Likelihood Ratio (LLR), segSNR, Weighted Spectral Slope (WSS) and Perceptual Evaluation of Speech Quality (PESQ).

Nesta and Omologo (2012) estimated each mixing matrix along with the physical meaning of the Frequency-Domain Blind Source Separation (FD-BSS) by Independent Component Analysis (ICA) contained information on the physical acoustic propagation associated to every source and then can be used for localization principles. In this study, they analyzed the generalized state coherence transform that is a non-linear transform of the space represented by the whole de-mixing matrices. The transform permitted an accurate estimation of the propagation time-delay of multiple sources in multiple dimensions. Furthermore, it was shown that with suitable nonlinearities and a statistical model for the reverberation, GSCT can be considered an approximated kernel density estimator of the acoustic propagation time-delay. Experimental results corroborated the good properties of the transform and its effectiveness in addressing multiple source TDOA detection. This is not useful for multiple sources the microphone detection rate is fewer. After the separation of the signals, high dereverberation occurs.

Yoshioka and Nakatani (2012) generalized the existing Dereverberation Methods using sub band domain multi channel linear prediction filters so that a Multiple Input Multiple Output (MIMO) room impulse response can be blindly shortened by the Resultant Generalized algorithm between a set of unknown amount of sources and a microphone array. Unlike existing Dereverberation Methods, the offered algorithm was developed without assuming specific acoustic conditions and offered a firm theoretical underpinning for the Sub-Band Domain Multi Channel Linear Prediction Methods applicability. The generalization was achieved by using a new cost function to estimate the prediction filter and an Efficient Optimization algorithm. The Projected Generalized algorithm made it easier to realize the common background underlying different dereverberation methods and future practical development. Experimental consequences were reported, presents that the Generalized algorithm successfully achieved blind MIMO impulse response shortening particularly in a mid to high frequency range. The Analyzed algorithm alone is not a good performer, so it will be combined with some other robustness technique to achieve reverberation.

Chien and Hsieh (2012) presented the Independent Component Analysis (ICA) which is essential for un-supervised learning and Blind Source Separation (BSS). The ICA unsupervised learning procedure endeavors to demix the observation vectors and made out the salient features or mixture sources. A convex divergence quantity was developed by relating the protuberant functions to the Jensen's inequality. Modifiable with a convexity parameter, this inequality-based divergence measure had a wide range of the steepest descents to reach its minimum value. A convex divergence ICA(C-ICA) was composed and a non-parametric C-ICA algorithm was resultant with diverse convexity parameters where the non-Gaussianity of source signals was illustrated by the Parzen window-based distribution. Experimental outcomes indicated that the dedicated C-ICA significantly reduced the number of learning epochs during estimation of the demixing matrix. The convergence speed was enhanced by using the Scaled Natural Gradient algorithm. Experiments on the BSS of direct, noisy and convolutive mixtures of speech and music signals additionally established the dominance of the anticipated C-ICA to JADE, Fast-ICA and the non-parametric ICA based on common information.

Gunther (2012) exploited non-stationarity of signals as a meant to separate groups of independent signals from observed mixtures. The semi-blind separation setting was beneficial for blind echo cancellation because the unknown near-end signal appeared in a group by itself. Therefore, successful semi-blind separation recovered the near-end signal from the measured microphone signal which, during double-talk was a mixture of the near-end signal as well as echoes of the far-end signal. Thus, semi-blind separation could adopt an echo canceling filter to change in the echo path response during a double-talk event. This study displays the disadvantages are slow convergence and after implementing the adaptive techniques this method just tracking the changes in the echo path of the channel.

Anderson *et al.* (2012) intended to employ the multivariate Gaussian source prior to attain joint blind source separation of sources that were linearly dependent across datasets. Analysis within the study yielded the local steadiness conditions, non-identifiability conditions and induced Cramer-Rao lower bound on the attainable interference to source ratio for independent vector analysis with multivariate Gaussian source priors. Furthermore, by exploiting a novel non orthogonal decoupling of the independent vector analysis cost function, he introduced both Newton and Quasi-Newton's Optimization algorithms for the general autonomous vector analysis framework.

Lin *et al.* (2012) introduced the perception of obligatory spectral diversity to mitigate the negative effects of near-common zeros in blind system identification when used for subsequent speech dereverberation. It had addressed the common zeros, problem for Blind System Identification Algorithms based on cross-relation and established how near-common zeros affect channel diversity as well as the performance of a frequently Employed Adaptive Blind System Identification algorithm and Normalized Multichannel Frequency do Main Least Mean Squares (NMCFLMS) algorithm. In the suggested FSD approach, the MINT algorithm is applied to the modified system which can be identified with less error. This is not robust for dereverberation.

OVERVIEW OF THE CONCEPT

Speech signals are detained by distant microphone in an enclosed atmosphere (for example-auditorium) are often contaminated by interferences of many speech signals and reverberations of the room. The technology includes automatic speech recognition and teleconferencing applications are greatly restricted its effectiveness range in the speech processing system. Separation of the source speech signals from the convolutive speech mixtures of microphone signals is being a risky and challenging task because of interference and room reverberation makes more unconformity. Nevertheless, it is necessary to alleviate critical acoustic disturbance. Throughout last decade, Blind Source Separation (BSS) and Blind Dereverberation (BD) were the two conventional approaches used to alleviate the speech processing.

Blind source separation divides the source sound from the convolutive mixture of acoustic signals and blind dereverberation removes the deflections and diffractions, i.e., presence reverberation effect from the speech signals. This study proposes Adaptive Equalization algorithm for the dereverberation of signals. The dereverberation for the speech signal is estimated by considering the one signal and the echoes or signal deflections can be removed. The sources are estimated by the Constant Modulus algorithm. This recommends joint method using adaptive equalization with the optimized parameter process in dereverberation and separation for convolutive multi-sound signal mixtures.

PROPOSED METHOD

The mixed signals are implied as Y_i recorded through the microphones. The process of signal separation and dereverberation is explained in following steps.

Time to frequency conversion: Signals in time instants are converted to frequency domain signals. The conversion is offer by using the FFT. The conversion of signals from time to a frequency domain makes the process in frequency-domain BSS approach. After, this signals are whiten for preprocess of ICA.

Source separation: Signals are separated by using the ICA analysis. Source signals from the different source mix together and have to perform separation of signals without knowledge of source signal. This can be known as the Blind separation. After separation of the signals again it is converted to the time signals by using IFT.

Alignment of signals: The separated outputs should be aligned by scaling and permutation. Otherwise, signals will be in the form of the rearranged manner.

Dereverberation: To remove the reverberant effect of the separated signals by adaptive equalization technique the CMA algorithm is employed. After source separation of signals, this forms to remove the effect of echoes and reverberation. This is called as the blind dereverberation. The General Adaptive Filtering Method in which the digital filter carries filtering on the input signal, produce an output signal. Adaptive algorithm adjusts the filter coefficient included in the vector, in order to let the error signal to be the smallest. Error signal is the difference of desired signal and the filter output. Consequently, adaptive filter involuntarily goes on a proposal based the attribute of the input signal and the desired signal.

By means of this technique, adaptive filter can be modified to the environment place by these signals. The CMA is mainly used algorithm in adaptive filtering. It is a Gradient Descent algorithm; it regulates the adaptive filter taps moderating them by an amount comparative to the instantaneous approximate to the gradient from the error surface. Its iterative process includes computing the output of a Finite Impulse Response (FIR) filter formed by a set of filter coefficients.

BLIND SOURCE SEPARATION

The aim of blind source separation is to separate the mixture of signals and yield the separated signals without knowing the original signals. This compares to the cocktail party problem if that is the case, there will be some speakers and several microphones. The original signals are the pure signals which are in the form of discrete time signals. Now, the mixture of audio signals symbolized as S(t) have to be separated. In BSS Method, there are two methods can be usually used which are Time-Domain BSS and Frequency-Domain BSS. In Time-Domain BSS Method, the convergence rate is low and the computational cost can be high. So, at present the frequency-domain BSS is applied for separation (Fig.1).

The block diagram exploits the working of the proposed method, initially the signals are converted to frequency domain and the ICA analysis is processed. After the separation, alignment by permutation and scaling then convert to time domain signals. Finally, for removing noise and echo the Blind Reverberation Method is applied. Thus, separated signals are free of noise and echo is attained.

Frequency-domain BSS: In frequency-domain BSS, the mixture of signals in the time domain is changes to frequency-signals. This can yield by the Fast Fourier Transform (FFT). The number of signals must be equal in a number of microphones. ICA analysis is used to separate the mixture signals. The separation is achieved by using the separation matrix. The observed signals or mixture of signals can be given as in the form of column vector. The preprocessing methods whitening and center limiting theorem are applied before ICA.

At this point, it consists of N sources and M microphones, to perform ICA based analysis for the source separation consider N = M:

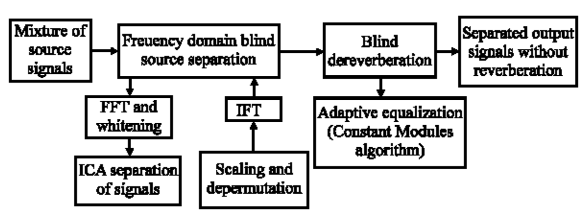

Fig. 1: Block diagram of proposed concept

$$I(t) = [I_1(t), ..., I_n(t)] \quad (1)$$

Equation 1 designates the input signal in the form of column vector. I_i the ith source signal 't' stands for the instant of time these inputs are independent to each other so it's represented as I and the observed signals at M sensors is:

$$S(t) = [S_1(t), ..., S_n(t)] \quad (2)$$

From the source signals, the observed signals S(t) vector is attained. In Eq. 2, S_i is ith mixture of signals. The Mixing Model of the signals in the time domain is described:

$$S_i(t) = \sum_{j=1}^{N} \sum_{i} A_{ij}(l) I_j(t\text{-}l) \quad (3)$$

Equation 3 expresses $A_{ij}(l)$ the impulse response from the source to the sensor. The separation system consists of MIMO filters, to get the output as the input given. The Demixing Model or separation to this signal is multiplication with the D matrix:

$$E_k(t) = \sum_{i=1}^{M} \sum_{l=1}^{L-1} D_{ki}(l) S_i(t\text{-}l) \quad (4)$$

In the Eq. 4, l is the length of the filter to generate the separated signal of length equal to the input length. The value of D is inverse of matrix A but the system only observes the value of S. D is named as the separation matrix. So, to achieve the separation matrix, the Independent Component Analysis (ICA) technique is performed. E signifies the estimated output after separation.

Before performing the ICA, to change the signals into the frequency-domain, convert the time domain signals into the frequency domain signals. For transferring time signals to frequency signal employ FFT. Using this Mixing Model can be expressed as:

$$S(\omega, t) = A(\omega) I(\omega, t) \quad (5)$$

Where:

ω = The frequency length
t = The frame index of the signal

A denotes the frequency response from source to sensor. The Demixing Model is estimated by the following equation:

$$E(\omega, t) = D(\omega) S(\omega, t) \quad (6)$$

The separation matrix is multiplied with the mixture of source to yield the separated signals. This is estimated by multiplication of separation matrix with mixed signals given in the Eq. 6. The separation matrix can be calculated by using the subsequent steps. The architecture of blind source separation is given (Fig. 2).

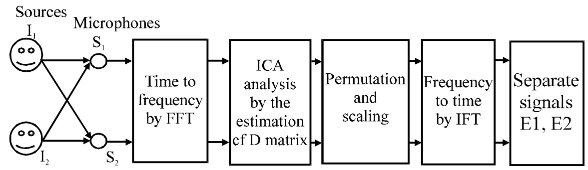

Fig. 2: Frequency domain BSS architecture

The source speech of two persons is I, mixed signals are S, the frequency domain conversion is by the FFT. The ICA process is done for estimation of separation of signals, permutation and scaling is for the alignment of the separated signals. After the separation signals are again changed to time domain by IFT. The separated signals are given as E.

Independent component analysis: Independent Component Analysis (ICA) is a method for finding fundamental factors or components from multivariate (multi-dimensional) arithmetical data. Independent component analysis is used to perform the separation by calculating the separation matrix values. This researh based on considering that the source signals are independent so only chosen ICA analysis. There are several methods to execute ICA analysis, the value of sources can be extracted from the mixture signals.

Preprocessing of ICA: After, converting signal into frequency domain for preprocess the ICA whitening of the signal is performed. This is also known as the half part in ICA where the correlation between the two signals is removed and the signals are independent. This is by calculating the eigenvectors of the covariance matrix of the mixed signal and the matrix with eigenvectors. After whitening only the signal is prepared to perform ICA.

One of the gains of frequency-domain BSS is it makes use of any ICA algorithm for instant mixtures, for instance the information maximization approach united with the natural gradient, Fast ICA, JADE or an algorithm based on the non-stationarity of signals. ICA is to peak the source signal only using observation (mixed) signals to restore the each independent component of source signal is processed by the iteration step. Presently, the measures of separation are depending on either kurtosis or negentropy. Now maximizing the negentropy based analysis is derived. The final estimate of the separation matrix is given by Newton's Iterative Method which is given below. The value for the separation matrix can be calculated for each frequency bins:

$$D_{new} = D\Big(E\big\{g(|S^H Y_W|^2)+(|S^H Y_W|^2)g'(|S^H Y_W|^2)\big\} - E\big\{g(|S^H Y_W|^2)(S_{W^H}D)\big\}\Big) \quad (7)$$

Where:
w = The previous value of separation matrix
w^H = The conjugate transpose of w matrix

The iterative equation can be initialized by selecting the w value by means of the random variables. S_w is the whitened signal of the mixture. E specifies the expectation and the g value the nonlinear function and g' the derivative of the function g:

$$g(u) = \tanh(au) \quad (8)$$

In the Eq. 8, 'a' takes the value of 0.01, where u is the function variable. The separation matrix can be calculated for the frequency bins with the stopping criterion of the D. Stopping criterion is:

$$\delta = (|D_{old} - D_{new}|)^2 \quad (9)$$

The stopping criterion is given in Eq. 9. This is very less near to 0.01. After every calculation of separation matrix, the w values are normalized as:

$$D_{new} = \frac{D_{new}}{|D_{new}|} \quad (10)$$

In the Eq. 10, $|D_{new}|$ indicates the norm function. After iteration the value of the D for the each frequency bins, the w should be converging to same point. So that the de-correlation of the previous value to new value can be calculated by using the equation:

$$D_{new} = D_{new} - \sum_{j=0}^{new-1} (D_{new}^T D_j)D_j \quad (11)$$

Equation 11 declares D_j, the previous frequency bins D values. Then, the calculation of w stops, the output of the signal is evaluated. The ICA output persist permutation and scaling ambiguities. Without correcting the order of the signal the output of the signal is impossible. The permutation and scaling in each frequency bin is given as:

$$E(\omega,t) = D(\omega)S(\omega,t) = W(\omega)P(\omega)I(\omega,t) \quad (12)$$

Where:
W = The diagonal matrix for scaling
P = The permutation matrix
I = The original source signal

To align the order of the signal de-permutation and scaling are performed.

Scaling: After performing ICA, the scaling and permutation problems should be solved. Scaling the columns of the separation matrix can be stated as the scaling problem. This can be used to prevent the separated signals from the distortion. This can align the signals in order to get the aligned signals as the output. The scaling ambiguity instigates a filtering effect during the time domain. The output signals turn into natural, so $D(\omega)$ is generated on specified criteria. Scaling is performed to correct the alignment of the separated signal. The aligned filtered signals are estimated by rescaling the separation matrix. There is a minimal and practical solution for the scaling problem. The scaling problem may be solved based on the minimal distortion principle:

$$D(\omega)' = \mathrm{diag}(D(\omega)^{-1}) \times D(\omega) \quad (13)$$

Through Eq. 13, the solution will be output signal E_i becomes an evaluation of the reverberant description of source I_i received at sensor i. Alternatively, the permutation problem is convoluted, particularly while the number of source signals is large then the number of probable permutations increases to the factorial of N. The correction for permutation is essential as otherwise different signals will be restored at different frequencies and the whole process will fail.

Permutation alignment: As regards the permutation problem, many methods are available although it still appears to be an open issue. The permutation problem refers to the necessity of permuting the columns of the separation matrices to align the orders of the separated spectral components over all frequency bins. Commonly there are different types of algorithms for the permutations direction of arrival, correlation method and dyadic sorting approach. Before creating output signals in the time domain, align the permutation so that all frequency bins contain frequency components from one source signal.

In this method, the comparison between two signals are taken. Now the two frequencies of the separated signals are compares and the permutation measure is calculated. When the permutation occurs then again rescaling is maintains for the signals. By scaling the columns of the separation matrix then the output the signal is de-permuted. The general equation for calculating the permutation ratio is given:

$$P_{kl} = \frac{\rho_{pp}(\omega_k,\omega_l)+\rho_{qq}(\omega_k,\omega_l)}{\rho_{pq}(\omega_k,\omega_l)+\rho_{qp}(\omega_k,\omega_l)} \quad (14)$$

From Eq. 14, the value of the P_{kl} is calculated. The ratio between the comparison of signal frequency bins is in the range of $P_{kl}>1$. If the value of p is <1 then there will be a possible for permutation occurs. The permutation is removed by rescaling the signal. For each frequency bins the permutation can be checked. ρ is the value of comparison between two bins k, l. Different signals are denotes the indices p and q. The value of for the signal can be calculated using the following equation. Currently the $V(\omega, t) = |E(\omega, t)|$ for each frequency bins this function is calculated and compared with the signal:

$$\rho_{qp}(\omega_k \omega_l) = \frac{\sum_{t=0}^{t-1} V_q(\omega_k, t) V_p(\omega_l, t)}{\sqrt{\sum_{t=o}^{t-1} V_q^2(\omega_k, t)}\sqrt{\sum_{t=0}^{t-1} V_p^2(\omega_l, t)}} \tag{15}$$

In Eq. 15, the values of the frequency bins are summed and divided by the same frequency bins square roots. The frequency range of each frequency bin from starting to end point is represented as ‘t’. This can be calculated for every frequency bins and compared both for same signals and the different signals. This scheme is done step by step process at the first step only pairs of bins are depermuted. In the second step, these pairs are aligned. This scheme is continued until all bins are processed. This process can be first check the number of possible permutations using the above formulas. If there is a permutation take place attempt to correct it. If the frequency is unnecessary then that can be dropped. Then, the remaining frequencies are rescaled. And the signals are reconstructed.

BLIND DEREVERBERATION

Reverberation means the echo or any some other disturbances which affects the clear speech. The recurrence of speech creates reflection of sound waves and some other noises also included in this especially for room acoustics. Reflection of sound waves or reverberation must be probable inside the room. So, the disturbances in speech dereverberation and the Inter Symbol Interference (ISI) have to be conquered. For most of the cases in dereverberation considered as there is only one sound in the room. And try to recover the original signals. Adaptive Equalization Method offers to approximate the recovery from the reverberation signals. Adaptive equalization is the finest solution for deduction of reverberation and inter symbol interference. Adaptive equalization techniques can be worked under the adaptive algorithms. This method can automatically changes the values for the filter coefficients. An adaptive filter is a self-modifying digital filter which alters the coefficients value to minimize the error function. This can be referred as the cost function which measures the difference between the reference or desired signal and the output of the adaptive filter.

Adaptive equalization: Mainly Adaptive Equalization Method can be consists of two modes. The first one is training mode and the other one is called tracking mode or decision directed mode. In the training mode, the trained signals are passed through the filter and the error value can be updated. To achieve the dereverberation for the unknown signal this estimated error value should be removed from the signal by using this only Adaptive Equalization algorithms. The error signal is employed by the Adaptation algorithm to update the adaptive filter coefficient vector in accordance with various functioning principle. In Adaptive Method, there are several algorithms for noise cancellation and channel equalization. The basic main thing is to get the mean square error for this optimization.

Blind adaptive equalization: In digital communication, equalizer was premeditated to balance the channel distortions, through a process known as equalization. Nearly two categories of equalization which are: trained equalization and blind (self-recovering) equalization. Blind equalization finds significant application in data communication systems. In data communications, linear channel distortions and effect of restricted channel bandwidth, multipath and fading is often the most considerable distortion in communication system.

Blind equalization improves system bandwidth proficient by avoiding the use of training sequence. The linear channel distortion, known as the Inter-Symbol Interference (ISI), can relentlessly damage the transmitted signal and make it complicated for the receiver to directly recover the transmitted data. Channel equalization and classification has proven to be an effective means to recompense the linear distortion by removing much of the ISI.

Constant modulus algorithm: Blind channel equalization is also known as a self-recovering equalization. The purpose of blind equalization is to improve the unknown input sequence to the unknown error values and weight tap vectors on the statistical properties of the input sequence. The receiver can harmonize to the received signal and to regulate the equalizer not including the training sequence. The term blind is used in this equalizer because it function the equalization on the data devoid of a reference signal. Instead, the blind equalizer relies on information of the signal structure and its statistic to complete the equalization. The Constant Modulus algorithm is one such algorithm working for the blind adaptation problem.

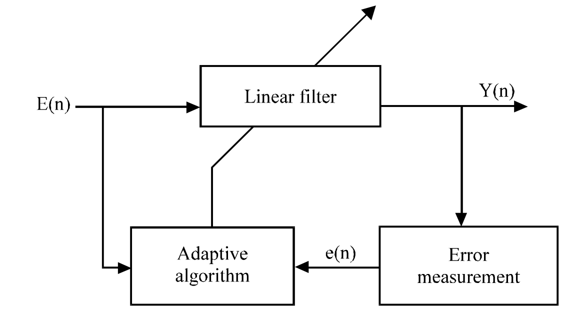

Fig. 3: Constant Modulus Algorithm Filter Model

The CMA projected by Godard, the most popular algorithm for blind equalization of signals. For equalization of the signal in the Adaptive Method, Constant Modulus algorithm can be used. LMS, RLS and CMA are the main adaptive techniques for equalization. Consider the model of a digital communication channel characterized by Finite Impulse Response (FIR) filter. The received signal of equalizer is x(n). In order to remove a effect of channel distortion, the equalizer is defined to eliminate its effect. The filter output is y(n).

The working model of the constant modulus algorithm, E(n) denotes the signal from the ICA estimation, Adaptive algorithm means the CMA update function. The linear filter is calculates the output of the filter using CMA function, error measurement calculates the error of the filter (Fig. 3).

CMA is performed without the information of the reference signal. This is similar to that of the LMS algorithm. The calculation of Constant Modulus algorithm is defined.

Error calculation:

$$e(n) = (|\,y(n)\,|^2 - 1)^2 \tag{16}$$

Where:
y(n) = The output from the filter
e(n) = The error value in Eq. 16

For blind condition the constant is considered as one.

Output signal:

$$y(n) = F(n)^H E(n) \tag{17}$$

In Eq. 17, F(n) the function to yield the output, x(n) the filter input (the output from the separation technique). The f(n) value is updated by the Eq. 18 using the error value.

Update equation:

$$F(n+1) = F(n) - \mu \times E(n) y(n)^* \times e(n) \tag{18}$$

Where:
y(n)* = The transpose of matrix y(n)
E(n) = The input to the filter
μ = The step size parameter
F(n) = The form of the matrix of length n

Initialization of f(n) is assumed by the center tap values of the matrix is unit:

$$E(n) = [E(n), E(n-1), E(n-2), \ldots, E(n-M+1)]^T \tag{19}$$

Equation 19 illustrates the input signal to the equalizer filter:

$$F(n) = [F_0(n), F_1(n), F_2(n), \ldots, F_{M-1}(n)] \tag{20}$$

Equation 20 is the matrix of update function. An initial value of F(n) is given as follows:

$$F(n) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \multicolumn{5}{c}{\text{- - - - - - - - - - - - -}} \\ 1 & 1 & 1 & 1 & 1 \\ \multicolumn{5}{c}{\text{- - - - - - - - - - - - -}} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{21}$$

Equation 21 demonstrates the center tap values of the F(n) are assumed to be one. After getting the result of CMA due to some noise variations employ simple de-noising technique for the removal of noises. Savitzky-Golay filter is employed to reduce the noise present in the echo removed signal. The filter which is known as least squares smoothing filter used in the large data streams commonly.

Least squares are the main process used to remove the noises and reverberations in the adaptive equalization techniques. This filter is also works as well as the adaptive techniques, now it removes the noise by minimizing the mean squared approximation error. This also the model performs as matrix functions. The resultant graphical notation is explained in the experimental results given.

EXPERIMENTAL RESULTS

The experimental performance of the ICA-CMA is illustrated in this study. The input mixture signal of the ICA is convert to frequency format and for the preprocessing of ICA whitening of the mixture signal is carried out. Here, the speech signal is of 3 sec of the length nearly 160000 data points is taken for experimentation. For separation of the signals, the signals convert to frequency domain by applying FFT, the signals are separated to the length of 1024, 2048, 4096 of the equal parts for performing ICA. The separation matrix is accomplished by initiating the assumed value of unknown measures. The separation matrix is calculated in the loop extension till it converges to some point. The separation matrix is calculated for all frequency bins. The steps of ICA completed, the signal is converted to time format by the application of IFT. After the conversion to time varying signal, the separated signals from the ICA are taken separately and process CMA functions. At this point the reverberation removing matrix is function of output estimate from the previous state. The input signals of the mixed sources are shown in the graphical representation.

These signals are mixed together and recorded through the 2 microphones. These signals are mixed and consist of noise and echo effects. This signal is separated from the mixed source signals using ICA then the echo removal is estimated by CMA algorithm for the Blind Dereverberation Method. The graph itself confirms the difference between the input and the output of the speech signals (Fig. 4).

Figure 5 details the signals 1 and 2 are free from echoes and reverberation effects of room acoustics. Here, adaptive technique removes the effect of the ISI, Adaptive Methods perform better consequent than other removal methods. After, finishing equalization, the noise present in the signals is removed by the noise removing filers. The figure below presents the signals after removal of noise.

Figure 6 exhibits the output of the speech signals after separation, dereverberation and denoising. This is the output signal after separation and dereverberation by ICA-CMA Methods. The noise removal is achieved by Savitzky-Golay filter. The process is experimented by using MATLAB and the experiment results are displayed. The designed functions for the ICA and CMA algorithms and the filter for noise removal the graphs are plotted against the signal length vs. frequency of the signal. The results are obtained without knowing original input, location of the microphones and sources, properties of speech and added noise and the echoes. Now the signals are free from noise and other interferences. The SDR, SIR, segSNR, values are compared to the other methods. Table 1 and 2 illustrate the respective comparison values.

The SDR, SIR asset value of the proposed method is compared with the DCT, K-SVD, PCA, ideal DUET

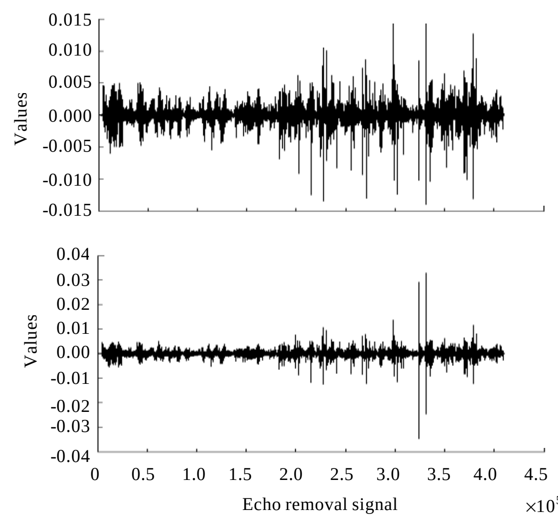

Fig. 5: a, b) Echo removed signals

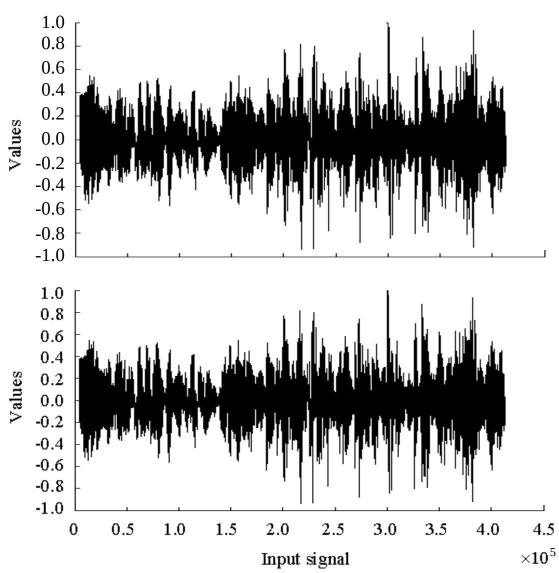

Fig. 4: a, b) Input signals 1 and 2

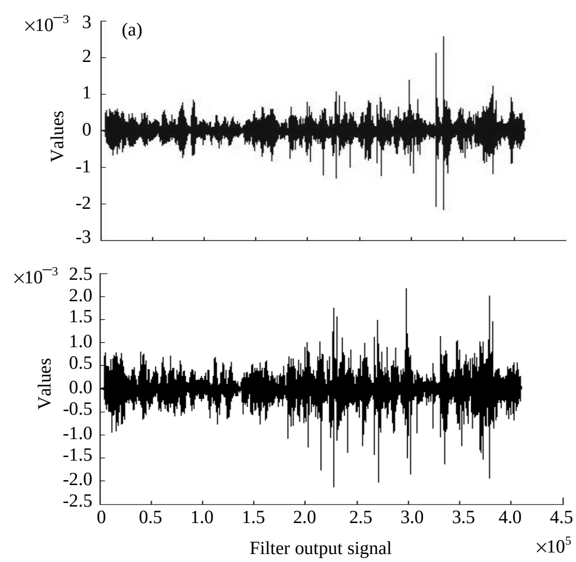

Fig. 6: a, b) Noise removed signals 1 and 2

Table 1: Average SDR, SIR values

	Measures			
	SDR		SIR	
Speech	Source 1	Source 2	Source 1	Source 2
DCT	5.7843	16.9233	3.1930	14.3660
Ksvdonly	8.6525	18.2549	4.9738	15.8018
Duet	-7.8259	9.9026	-12.6961	8.1667
Duetiter	-2.0453	15.2677	-8.0876	8.8332
Ovlpduet	-9.8183	10.7482	-12.7126	7.4321
duetideal	8.2099	21.9263	2.6693	10.9246
Two Layer Sparsity Model	13.2140	19.5834	8.0808	17.7373
Proposed	15.1186	23.7980	11.7851	16.6172

Table 2: Average segSNR values

	Measures (segSNR)	
Speech	Source 1	Source 2
rev	-3.44	-4.51
∞-norm	-4.93	-4.97
ISS	-1.31	-1.29
MLP	-2.03	-3.69
BAE	2.41	-0.55
Proposed	3.01	0.67

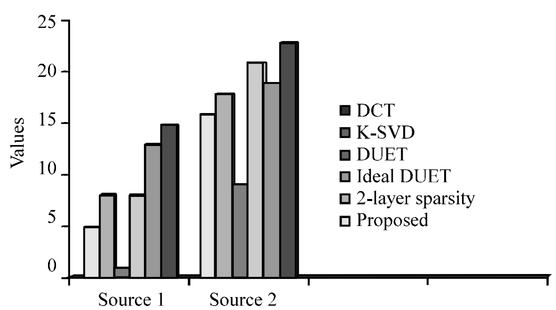

Fig. 7: SDR graph

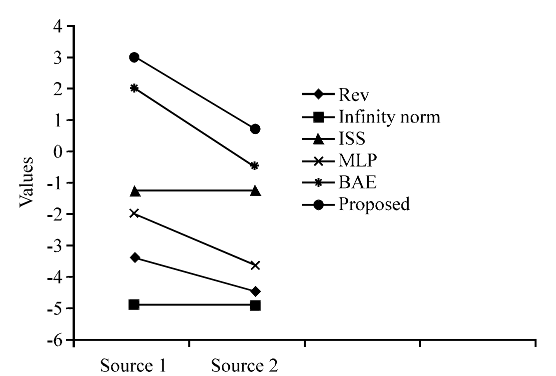

Fig. 8: segSNR graphical representation

approaches. The segSNR measure is to achieve frequency domain blind source separation evaluated with the Inversion Spectral Subtraction (ISS), Multichannel Linear Prediction (MLP), Infinity-Norm Minimization algorithms and Blind Adaptive Estimation Methods. The SNR ratio of the proposed method varies from -1.12 to 28, at different db (Fig. 7 and 8).

CONCLUSION

The study justifies the separation of independent signals from the observed mixtures of speech. The Developed algorithm presents the source separation and reverberation in the joint manner. Initially, the preprocessing of ICA is completed through a whitening process followed by, the explanations of ICA steps and permutation alignment in detail. ICA results are recovered from distortion because of applying permutation and scaling process. The dereverberation of signal with the aid of adaptive equalization technique is established in the Constant Modulus algorithm with the assumed initial values. This method is easier and achieve that it works without knowing any previous knowledge of original signals. Experimental analysis is carried out with MATLAB Software and results are displayed in a graphical format. This method of equalization for the reverberated signal provides superior results. Noise is removed using Savitzky-Golay filter the outcome of the algorithm is de-reverberated separated signal.

REFERENCES

Anderson, M., T. Adali and X.L. Li, 2012. Joint blind source separation with multivariate gaussian model: Algorithms and performance analysis. IEEE Trans. Signal Process., 60: 1672-1683.

Araki, S., R. Mukai, S. Makino, T. Nishikawa and H. Saruwatari, 2003. The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech. IEEE Trans. Speech Audio Process., 11: 109-116.

Bao, G., Z. Ye, X. Xu and Y. Zhou, 2013. A compressed sensing approach to blind separation of speech mixture based on a two-layer sparsity model. IEEE Trans. Audio Speech Lang. Process., 21: 899-906.

Borigsar, K.R. and G.R. Kulkarni, 2010. Simulation and comparative analysis of LMS and RLS algorithms using real time speech input signal. Global J. Res. Eng., 10: 44-47.

Chien, J.T. and H.L. Hsieh, 2012. Convex divergence ICA for blind source separation. IEEE Trans. Audio Speech Lang. Process., 20: 302-313.

Das, S. and K.K. Sarma, 2012. Noise cancellation in stochastic wireless channels using coding and adaptive filtering. Int. J. Comput. Appl., 46: 21-25.

Furuya, K. and A. Kataoka, 2007. Robust speech dereverberation using multichannel blind deconvolution with spectral subtraction. IEEE Trans. Audio Speech Lang. Process., 15: 1579-1591.

Gunther, J., 2012. Learning echo paths during continuous double-talk using semi-blind source separation. IEEE Trans. Audio Speech Lang. Process., 20: 646-660.

Habets, E.A.P., S. Gannot, I. Cohen and P.C.W. Sommen, 2008. Joint dereverberation and residual echo suppression of speech signals in noisy environments. IEEE Trans. Audio Speech Lang. Process., 16: 1433-1451.

Haque, M.A., T. Islam and K. Hasan, 2011. Robust speech dereverberation based on blind adaptive estimation of acoustic channels. IEEE Trans. Audio Speech Lang. Process., 19: 775-787.

Huang, Y.A., J. Benesty and J. Chen, 2005. A blind channel identification-based two-stage approach to separation and dereverberation of speech signals in a reverberant environment. IEEE Trans. Speech Audio Process., 13: 882-895.

Kokkinakis, K. and P.C. Loizou, 2009. Selective-tap blind dereverberation for two-microphone enhancement of reverberant speech. IEEE Signal Process. Lett., 16: 961-964.

Lin, X.S., A.W.H. Khong and P.A. Naylor, 2012. A forced spectral diversity algorithm for speech dereverberation in the presence of near-common zeros. IEEE Trans. Audio Speech Lang. Process., 20: 888-899.

Malik, G. and A.S. Sappal, 2011. Adaptive equalization algorithms: An overview. Int. J. Adv. Comput. Sci. Appl., 2: 62-67.

Nakatani, T., K. Kinoshita and M. Miyoshi, 2007. Harmonicity-based blind dereverberation for single-channel speech signals. IEEE Trans. Audio Speech Lang. Process., 15: 80-95.

Nesta, F. and M. Omologo, 2012. Generalized state coherence transform for multidimensional TDOA estimation of multiple sources. IEEE Trans. Audio Speech Lang. Process., 20: 246-260.

Nion, D., K.N. Mokios, N.D. Sidiropoulos and A. Potamianos, 2010. Batch and adaptive PARAFAC-based blind separation of convolutive speech mixtures. IEEE Trans. Audio Speech Lang. Process., 18: 1193-1207.

Sawada, H., R. Mukai, S. Araki and S. Makino, 2004. A robust and precise method for solving the permutation problem of frequency-domain blind source separation. IEEE Trans. Speech Audio Process, 12: 530-538.

Yoshioka, T. and T. Nakatani, 2012. Generalization of multi-channel linear prediction methods for blind MIMO impulse response shortening. IEEE Trans. Audio Speech Lang. Process., 20: 2707-2720.

Zheng, X., C. Ritz and J. Xi, 2013. Collaborative blind source separation using location informed spatial microphones. IEEE Signal Process. Lett., 20: 83-86.