

Blind Source Separation for Speech Music and Speech Speech Mixtures

K Prakash

M.Tech Student,

Gudlavalleru Engineering College, Gudlavalleru,
Krishna Dist, A.P., India, Pincode-521356

Hepzibha Rani D

Assistant Professor

Gudlavalleru Engineering College, Gudlavalleru
Krishna Dist, A.P., India, Pincode-521356

ABSTRACT

Separating one source from a mixture of sources is a problem, normally observed with parties. Here the sources may be all speech signals or one is speech and the other is music. To have a better understanding of speech, need to separate actual signal. This can be done by using blind source separation technique. It is hard to extract an interesting conversation from the background noisy crowd. Speech mixture is despoiled by the surrounding noise, interferences and additional speakers. Here an attempt for solving this separation problem, i.e. extracting one or more speech signals from a speech mixture. To eliminate or reduce the noise in speech signal in speech mixture is done by using wavelets. The wavelet output speech mixture processes for source separation by using, two techniques ICA and binary T-F masking. This separation technique is likewise applicable to segregate speech signal under reverberant conditions.

Keywords

Discrete time wavelets transform (DWT), Independent component analysis (ICA), and Time frequency masking (T-F Masking).

1. INTRODUCTION

Blind source separation is the segregation of a set of signals from set of mixed signals. The separation of signals can be done without knowing the information of source signals or mixing operation. The problem of extracting a single speaker from a mix of many speakers is known as “cocktail party problem”. Where the number of speakers is talking one is attempting to perceive sound from the one of the speakers. It is really difficult to discover the one actually speaking [1]. One signal processing system can be built up to evoke a single speech signal from a mix of many signs.

The mixing of sources can be obtained by using two mikes. Noise in the mixed signal is detached by using wavelet denoising technique. The denoised signal can be distinguished by using blind source segregation algorithm.

In this report, the implementation of blind source separation is done by using independent component analysis and binary time frequency masking. Here recovering of two independent source signals from a piece of music of the unknown speech mixture [2]. Through this an attempt is made to separate speech signals and recovered successfully.

Speech processing finds a vast range of applications in signal analysis, communications fields and this application use areas like noise cancellation, reverberation removal, echo filtering etc.

2. BLIND SOURCE SEPARATION

In trying to resolve the cocktail party problem here may be existing as a minimum of two difficulties, first one is the effect of echo makes the mixtures gathered by the microphones are convolutional and non linear, and the other is , these mixtures is not fully independent because the signal propagation is entirely nonlinear environment and there are some parasitic effects.

Blind source separation exhibit the problem of identifying N signals which are unidentified from a mixture of M recorded signals. Here the N unknown signals which are statically independent are $s(t) = [s_1(t), s_2(t), s_3(t), \dots, s_n(t)]$, and the recorded mixture signals $X(t) = [x_1(t), x_2(t), x_3(t), \dots, x_n(t)]$. The evaluation of source signals is performed on the basis of the output signals $Y(t) = [y_1(t), y_2(t), y_3(t), \dots, y_n(t)]$. The main objective of this system is that the original source signals are extracted by the adaptive system. The source signals extrication and the creative signal extracting without idea of the transmission channel characteristics, therefore the source signals can be extracted by a number of correlated problems: Blind source separation (BSS), Independent component analysis (ICA), Blind channel extraction (BSE) and Multichannel Blind Deconvolution (MBD).

The common blind source segregation form. The simplest case, p mixed signals linearly combine with q unknown signals. These unknown signals are mutually statistically independent, the source signals have zero mean and noise infected.

This can be in matrix notation as

$$X(k) = H s(k) + v(k) \quad (1)$$

Here $x(k)$ is vector of sensory signals,

$S(k)$ is source vector and $V(k)$ is an additive noise vector and H is an unidentified $m \times n$ mixing matrix

The branching system is presented with the unknown mixing process.

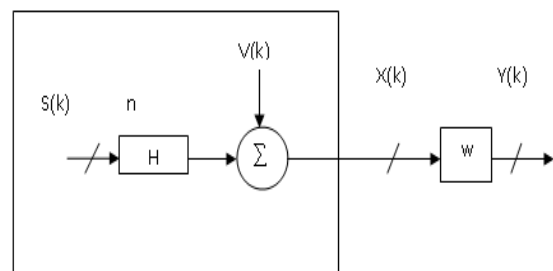


Fig.1 source separation system

Generally assumed that the number of sources, here simply the sensory vector signal $x(k)$ is presented. To estimate source signals, feed-forward and adaptive algorithm is needed to enable to estimate source signals, to recognize the mixing and separating matrices (H and w).

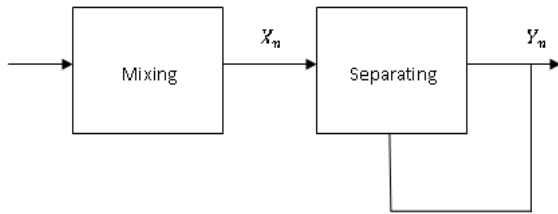


Fig.2 matrices of mixing and separating

3. PREPROCESSING

For easy and fast denoising discrete wavelet transform is used for noise reduction in the source signal. Denoising is done by considering the threshold, by considering the highest coefficients of discrete wavelet transform and its inverse to obtain a noise free signal.

Here choosing the proper wavelet scaling is important to denoised the signal. Daubechies 16 and 32 wavelets are used and both results are compared, in which both transforms gives satisfying result. The denoising or noise reduction is through wavelet shrinkage. Denoising process consisting of three steps: linear forwarding, shrinkage and inverse wavelet transform for linear signal. In parametric models the parameters must estimated to a particular model in advance or in priori.

Let the captured data

$$X(t) = S(t) + N(t) \quad (2)$$

$S(t)$ is the original signal, $N(t)$ the noise signal added to the original signal. Assume $W(X)$ and $W^{-1}(X)$ are the wavelet transformations in forward and inverse operations. Assume $D(Th)$ denotes the noise reduction operator, noise reduced here by using soft threshold operator 'Th'[3]. The denoising is done to $X(t)$ by using wavelet shrinkage to estimate the $S'(t)$, here $S'(t)$ is the estimated signal for original signal $S(t)$.

$$Y=W(X) \quad (3)$$

$$Z=D(Y,Th) \quad (4)$$

$$S'=W^{-1}(Z) \quad (5)$$

This implementation does not make known details of the operators D, or choice of threshold.

4. INDEPENDENT COMPONENT ANALYSIS

The independent component analysis is a method a method for extrication source signal from it is mixed by some additional signals. This independent component analysis works on the principle of independence and non gaussianity. The ICA assigns that every source at specified time instant is a random variable and independent from other signals i.e. statically independent, means that one source cannot correlated with other sources.

This algorithm divided into three parts: core procedure, separation stage and merging stage. In common, ICA cannot make out the definite number of source signals, a exclusively

right order of the source signals, nor the appropriate scaling of the source signals. This algorithm requires need of sensors equal to number of source signals, consider mixtures are recorded and they are modeled as,

$$X_1=e s_1(t)+f s_2(t) \quad (6)$$

$$X_2= g s_1(t)+h s_2(t) \quad (7)$$

In matrix notation this can be written as

$$x=As \quad (8)$$

The main goal is to find the original source signals so to get source signals the inverse should applied.

$$S=A^{-1}x \quad (9)$$

Consider in a room two persons are speaking, the speech was gathered using two microphones, these recorded signals are assigned as $x_1(t)$ and $x_2(t)$. The recorded signals are the subjective summation of the speech signals emitted by the two speakers, and denoted as $s_1(t)$ and $s_2(t)$. This can be written as in linear equation form shown as follows,

$$x_1(t) = p_{11}s_1 + p_{12}s_2 \quad (10)$$

$$x_2(t) = p_{21}s_1 + p_{22}s_2 \quad (11)$$

Where p_{11} , p_{21} , p_{12} and p_{22} are the parameters that are depend upon the characteristics of the microphone and space of the microphone. To estimate the original signals here discarding the time delay or a number of factors from the basic mixing model. Some statistical properties is used to recover the original signals estimation.

5. DISCRETE WAVELET TRANSFORMS

Discrete wavelets are used to analyze the signals. Discrete wavelet transforms formed as a choice for shot time fourier transform to beat the problems related to its time and frequency resolution characteristics. The Short time Fourier transform produces uniform resolutions for all frequencies, the discrete wavelet transform provides low time resolution for high frequencies and high time resolution for low frequencies[5]. The DWT is efficient in time frequency representation of signal, and DWT is a particular case of wavelet transform. The DWT is represented as follows:

$$W(j,k) = \sum_j \sum_k X(k)Q^{-j/2} \quad (12)$$

Consider finite energy and fast decaying signal $y(t)$, it is also called as the mother wavelet. The DWT algorithm is performed Multirate filter bank, consider the filter bank with the octave spacing between the centers of filters. Every sub band has half the samples of the adjacent higher frequency bands with unlike resolution by reconstructing the signal interested in a common estimation.

The signal is again processed by the alike wavelet in disintegration step. This breakdown is done by using highpass and low pass filtering of the signal at that particular time instant.

$$Y_{high}[k]=\sum_n x[n] a[2k-n] \quad (13)$$

$$Y_{low}[k]=\sum_n x[n] b[2k-n] \quad (14)$$

The above shows the outputs of low pass and highpass filters, after associate sampling of 2. Output of these filters is produces new coefficients, and applied threshold to get the original signal when reconstruction.

Dwt is well-organized algorithm for extracting original information and spectral properties of nonstationary signals

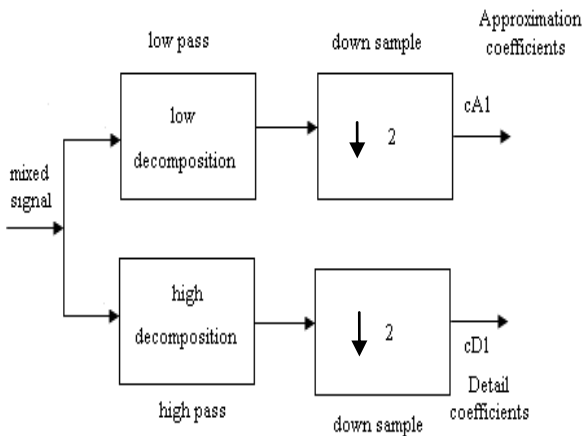


Fig. 3 block diagram of Discrete wavelet transform

Where $\downarrow 2$ denotes down sampling and keeps even indexed terms.

6. INVERSE DISCRETE WAVELET TRANSFORMS

It is a process of extracting the original signal from combining the dwt outputs. Here the coefficients are interpolated by adding zeros among the coefficients, by this we can doubling the length of the signal. These are convolved with rebuilding scaling filter for estimate and detailed coefficients. These results are added mutually to get the creative signal.

We have to make the coefficients before doing DWT and convolving to get estimation of actual signal, this is through

by using $N/2 - 1$ coefficients from the DWT coefficients, and adding these coefficients at the end, here N is the scaling feature.

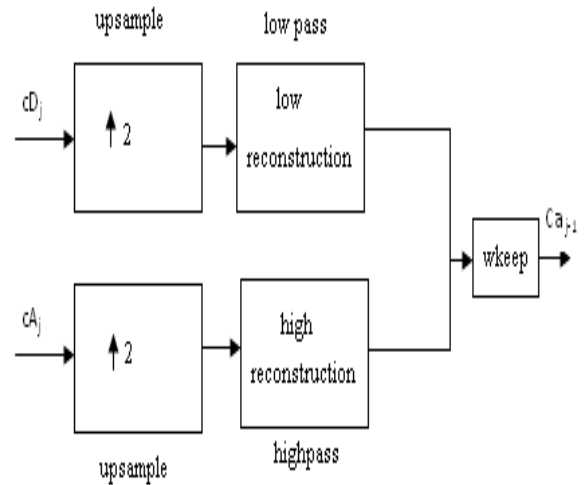


Fig.4 Block diagram of inverse discrete wavelet transform

Where, $\uparrow 2$ up sampling and keeps even indexed terms.

7. PROPOSED ALGORITHM

The block diagram for source segregation process is shown below. First the signal gathered from the microphone is considered as mixed signal and it taken as input to the algorithm. The denoising is done by using wavelet denoising and decomposed by discrete wavelet transform. The output of DWT is given to ICA for segregation process using statistical properties. Further the signal is reconstructed by using inverse discrete wavelet transform.

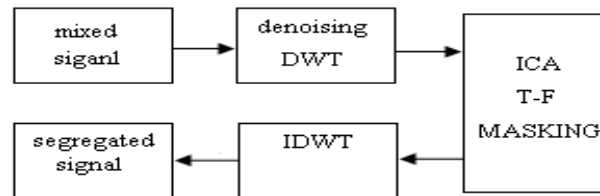
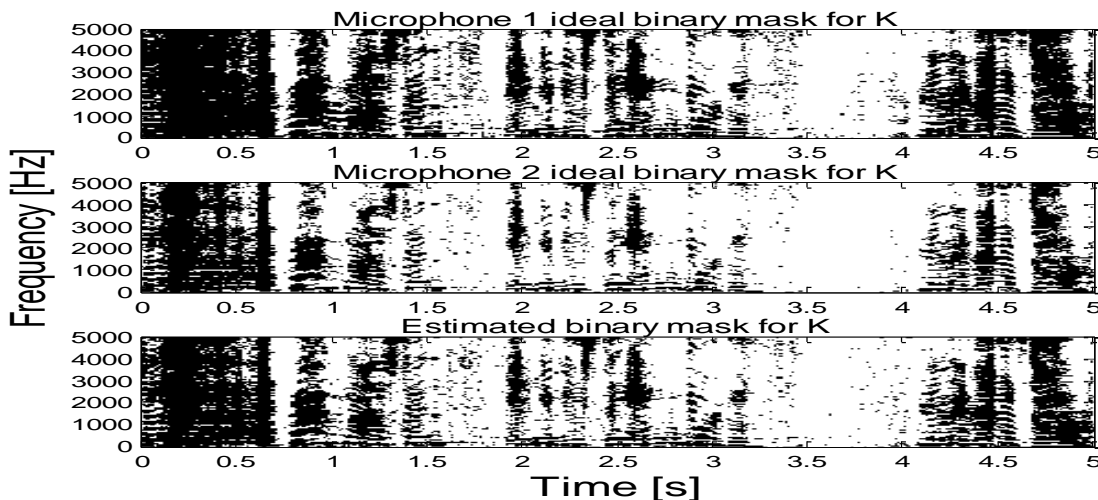
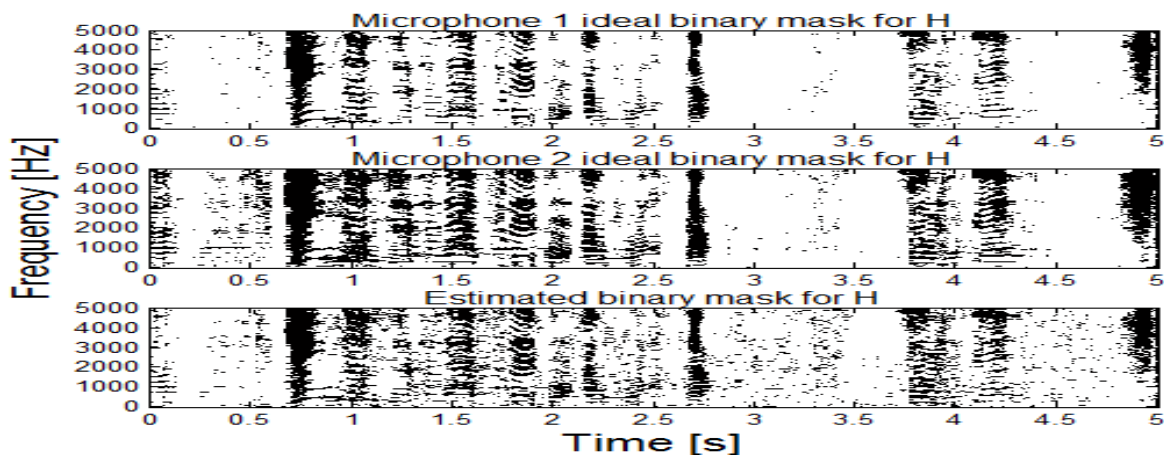
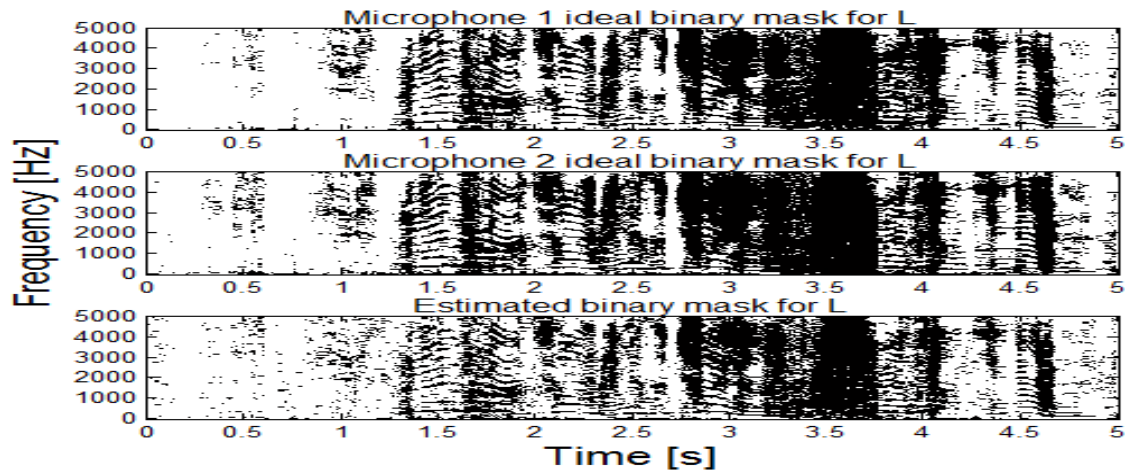


Fig. 5 block diagram for source segregation

8. RESULT ANALYSIS





9. CONCLUSION

The segregation algorithm produces better estimations than the ICA alone is used for separation of mixed signal. The mixed signal is gathered by mixing the voice signals and music signals recorded by the sensors. The segregation is mainly done by using statistical properties independence and nonlinearity. The signal is first diagnosed by using wavelet transforms and by processing the resultant of dwt to ICA algorithm which uses time frequency masking for segregation. Estimated signal is reconstructed by using inverse discrete wavelet transforms. The resultant segregated sources are better estimations than the previous techniques. This technique lead to very efficient segregation of source signals from mixtures.

10. ACKNOWLEDGEMENTS

The authors place on record their thanks to the authorities of Gudlavalleru Engineering College, A.P for the facilities they provided

11. REFERENCES

- [1] M.Akay ,time frequency and wavelets in biomedical signal processing .Piiscataway, NJ: IEEE Press, 1998 ,pp. 1. Alexis favot and Markus Erne, " improved cocktail-party processing", proc. Of the 9th Int conference on digital audio effects(DAFx-06),Montreal, Canada September 18-20,2006
- [2] Kenneth E . Hild and David Pinto,"Convolute blind source separation by minimizing mutual information between segments of signals", IEEE transactions on circuits and systems ,regular papers , vol.52, No 10, October 2005
- [3] Robi polikar, "the engineer's ultimate guide to wavelet analysis," hosted by Rowan university , college of engineering web servers , last major updates January 2001.
- [4] Dr. Michael lewicki ,"Michael lewicki computational perception and scene analysis course," webserver.
- [5] David L. donoho ,"Denoising via soft thresholding .IEEE Transactions on information theory , 41:613-627, may 1995.
- [6] Subband representation " , IEEE Transactions on speech and audio processing, Vol.9.no.5 123-135.
- [7] David L. Donoho and lain M. Johnstone,"Idel spatial adaption via wavelet shrinkage. Biometrika", 81:425-455, September 1994.
- [8] E. Visser and T. W. Lee, "Speech enhancement using blind source separation and two channel energy based speaker detection,"IEEE Int.Conf. On Acoust. Speech and Signal Process. vol. 1, pp. 884–887, April 2003.
- [9] Tomasz Rutkowski and Andrzej Cichocki,"Speech extraction from interferences in real environment using bank of filters and blind source separation", IEEE, Neural networks, 2002.
- [10] William Addison and Stephen Roberts, "Blind Source Separation with Non Stationary Mixing Using Wavelets, "Pattern Analysis Research Group, the University of Liverpool, 2006.