A Wavelet based ICA Algorithm for Separation of Linearly Mixed Speech Signals

Mohit Kumar
Department of Electronics & Communication Engineering
Galaxy Global Group of Institution

Asst. Prof. Saranjeet
Department of Electronics & Communication Engineering
Galaxy Global Group of Institution

Abstract

In this paper, an algorithm is purposed which can be used for separation of linearly mixed speech signals. Initially, this task was done by a popular algorithm called Independent Component Analysis (ICA). ICA is an effective algorithm to decompose mixed in to independent components. But it is observed that, when number of independent speaker increases, ICA is not able to extract the independent components completely. Hence the performance of ICA is satisfactory only when two individual speakers are present. The problem of extracting three speech signals more accurately is solved by using a Discrete Wavelet Transform based ICA. As a result of wavelet decomposition, wide band signals are converted into narrow band which decreases the probability of finding two independent signals in same frequency band. Recorded signals by the microphones are fed to an analysis filter bank which decomposes the signals in to approximation and detail frequency bands. These two groups of signals are further processed by ICA separately and then combined by synthesis filter bank to get the independent source signals. Further statistical parameters of the DWT based ICA are compared with conventional ICA on the basis of their performance.
Keywords: ICA, Discrete Wavelet Transform, Signal Separation, Non-Gaussian mixture, Mutual Information

---

I. Introduction

The fundamental purpose of speech is communication, i.e., the transfer of messages. Human speech is a complex signal to understand and synthesize due to its dual nature. Speech signal processing faces a lot of atmospheric interference. It is quite difficult and sometimes impossible to record speech signals from only a single source when more than one sound source is present in vicinity. The first and most difficult task in speech signal processing is to separate the desired speech signal from other speech signals which are mixed with it. When a group of n peopleare assumed to talk simultaneously with each other, the mixture of signals is recorded by n microphones. The goal is to extract the speech signals of the individual speakers or the sources from the mixture of speech signals without the knowledge of sources and the mixing process. When the number of original source signals and mixing matrix is unknown, the problem becomes more difficult. We only have the received signal and we have to extract the unknown source signals. Common techniques used in speech processing are short time energy, average zero crossing rate, pitch period estimation, STFT, wavelet transform, principle component analysis and independent component analysis.

ICA is an effective algorithm to decompose mixed signals in to independent components. But it is observed that, when number of independent speaker increases, ICA is not able to extract the independent components completely. The performance of ICA can be increased by using a Discrete Wavelet Transform based ICA. As a result of wavelet decomposition, wide band signals are converted into narrow band which decreases the probability of finding two independent signals in same frequency band. Recorded signals by the microphones are fed to an analysis filter bank which decomposes the signals in to approximation and detail frequency bands. These two groups of signals are further processed by ICA separately and then combined by synthesis filter bank to get the independent source signals.

Jutten and Herault in 1991 first proposed the framework of blind separation of sources by decomposing signals into independent signals. The term "ICA" was coined by Comon in 1994 in his paper on the theory of linear ICA. Linsker proposed the Infomax-principle in1992 which was later explained by Bell and Sejnowski in 1995. Later Amari introduced the concept of a natural gradient simplifying the Infomax learning rule in 1998. The most efficient ICA algorithm developed was the Fast ICA-algorithm proposed by Hyvarinen in 1999 using negentropy and kurtosis as a contrast function. The BSS problem with more sources than measured signals ($n > m$) is referred to as overcomplete ICA, presented by Lewicki and Sejnowski in 1998. ICA can be used along with other signal processing techniques to solve the problem of extracting required speech signal from mixture of different speech signals.

ICA is similar to PCA but it works on non-Gaussian mixture of signals. PCA tries to find the orthogonal components from a given set of data which are least correlated with each other by using second order statistics of data like mean and variance. The concept of Independent variables is broader than being uncorrelated. ICA works on extracting the independent components from mixed signals. The two broad definitions of independence for ICA are

- Maximization of Non-Gaussianity (based on Kurtosis)
- Minimization of Mutual Information (based on Negentropy)

All rights reserved by www.ijirst.org

II. INDEPENDENT COMPONENT ANALYSIS

Due to the complexity in analysis of mixed speech signals as discussed above in section 1.5, some special algorithms are used to decompose the mixed speech signals in to different components so that they can be further used in other applications like speech coding, text to speech conversion and pattern recognition [29]. A lot of techniques have been proposed for separation of linearly mixed speech signals. Independent Component Analysis (ICA) is one of the most efficient and popular technique to solve this problem.The idea of decomposing mixed signals into independent signals was first proposed by Jutten and Herault in 1991 in the framework of blind separation of sources, while the term "ICA" was later coined by Comon in 1994 in his fundamental paper on the theory of linear ICA.

ICA is similar to PCA but it works on non-Gaussian mixture of signals. PCA tries to find the orthogonal components from a given set of data which are least correlated with each other by using second order statistics of data like mean and variance [22]. The concept of Independent variables is broader than being uncorrelated. ICA works on extracting the independent components from mixed signals. The problem described in this dissertation is similar to famous cocktail party problem,which is a blind source separation (BSS) problem. A group of n peopleare assumed to talk simultaneously with each other. This mixture of signals is recordedby n microphones. The goal is to extract thespeech signals of the individual speakers or the sources from the mixture of speech signals without the knowledge of the sources and the mixtureprocess. ICA can be applied on given recorded set of data to recover the voice ofeach speaker just by using the mixed signals.

In practical situations like the Cocktail Party Problem where a lot of speakers are speaking continuously and independently, the microphones placed in the room will not be able to pick sound signals from individual speakers separately. Instead, what they will record is the linear combination of signals from various speakers produced as a result of interference and mixing of signals from different speakers. These speech signals are needed to be separated in to independent original signals from various individual sources. ICA proves to be a very useful tool in separation of Speech Signals which have been mixed linearly. [2] Independent component analysis (ICA) is an algorithm for revealing hidden factors from the mixture of data. Data for ICA may be sets of random variables, measurements, or signals. Observed Data in case of ICA is assumed to be linear mixture of unknown data variables [25]. Cocktail party problem can be visualised as convolutive mixture of N statistically independent speech signals can be represented in matrix form by a N dimensional independent sources as-

$$S(t) = [s_1(t), \dots , s_N(t)]$$

M microphones receive linearly mixed sources. This process in matrix form is X=AS, where X is a row matrix containing the observed signal, S is a matrix containing unknown original signals as row vectors. A is unknown weight matrix which when multiplied by original signal matrix gives the observed signal matrix. The ICA algorithm is trying to find as independent signals as possible from the mixture of signals x(t) by using an unmixing matrix W, and that the independent components can be written as $y(t) = W. x(t)$. For simplicity, only the case where the number of sources equals the number of observation ($m = n$) is considered.Now the aim is to estimate a matrix W which when multiplied by received mixed signals gives back original independent signals such that $W=A^{-1}$. Independence is the key factor in ICA. The extracted components by ICA should be as independent as possible.

III. RESULT AND DISCUSSION

It was observed through experimentation on different sets of recorded speech data that ICA performs well for two independent sources/speakers but the performance decreases abruptly when number of speakers increase. Also the algorithm is very sensitive towards selection of data sets.The separation of speech signals by only using ICA algorithm alone can be outperformed by clubbing it with some other algorithm. This idea coined in to mind by literature survey which shows it is always more accurate to study and process speech signals in parts instead of the whole signal at one go [20]. One such technique is Wavelet transform which is used to decompose the received signals in to sub frequency bands. Discrete wavelet transform uses filter banks to achieve this and the signals given as output of wavelet transform can be analyzed separately by applying ICA on them separately.

As a result of wavelet decomposition, wide band signals are converted into narrow band which decreases the probability of finding two independent signals in same frequency band. X(n) is divided into two sub bands and after a down-sampling, ICA1and ICA2, which represent ICA algorithm extract independent signals from the decomposed signals of wavelet filters resulting in X1'(n)and X2'(n) separately. These signals can be then fed to reconstruction filter bank to get the results. The complete process can be visualized by following block diagram.

All rights reserved by *www.ijirst.org*

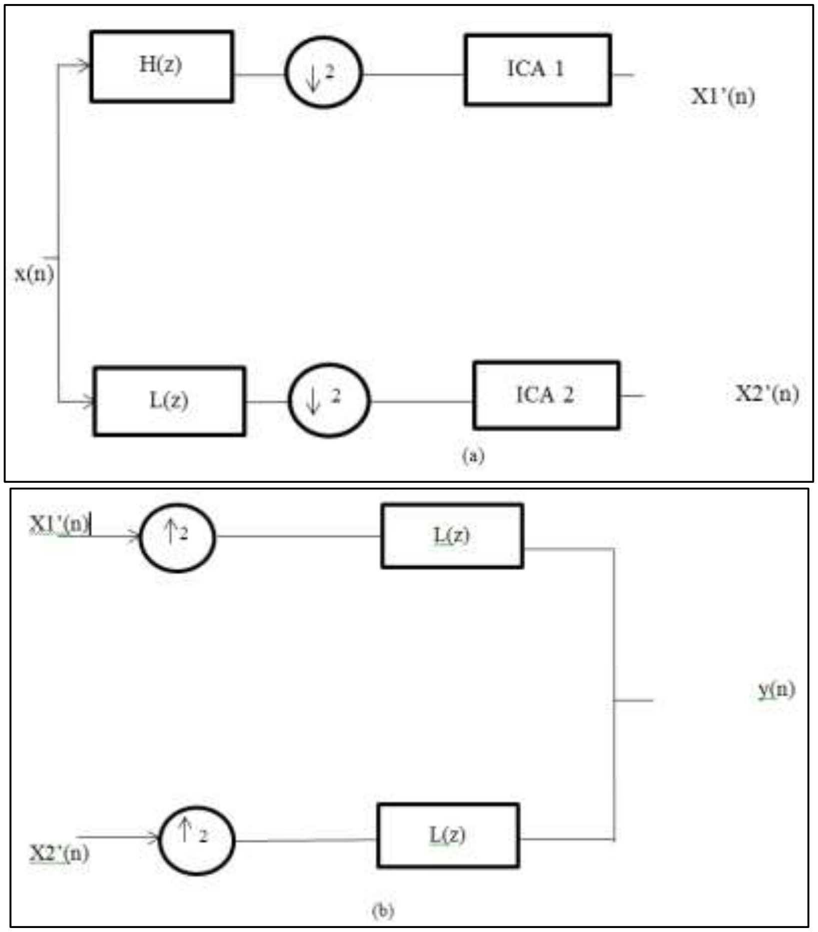
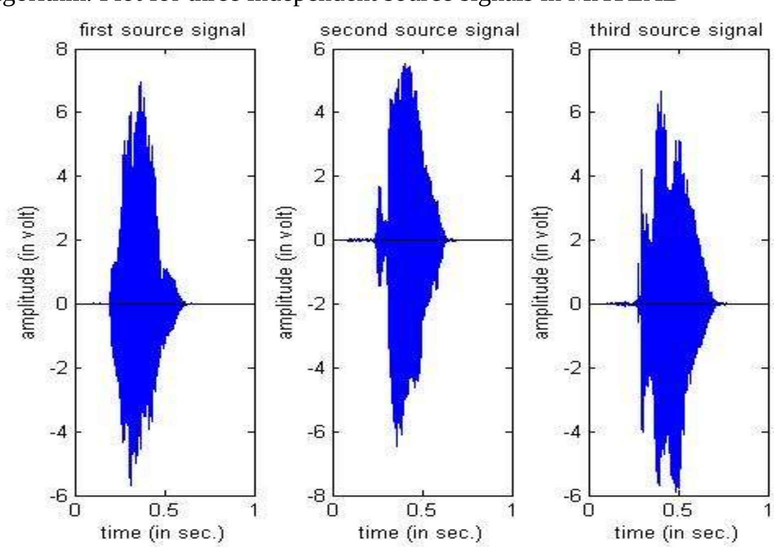


Fig. 4.2: Block diagram of DWT based ICA algorithm. Fig. (a) shows the Synthesis filter bank and two separate ICA algorithms on the decomposed signals giving X1'(n) and X2'(n) as independent components and Fig. (b) shows the Analysis filter bank to reconstruct the original source signals.

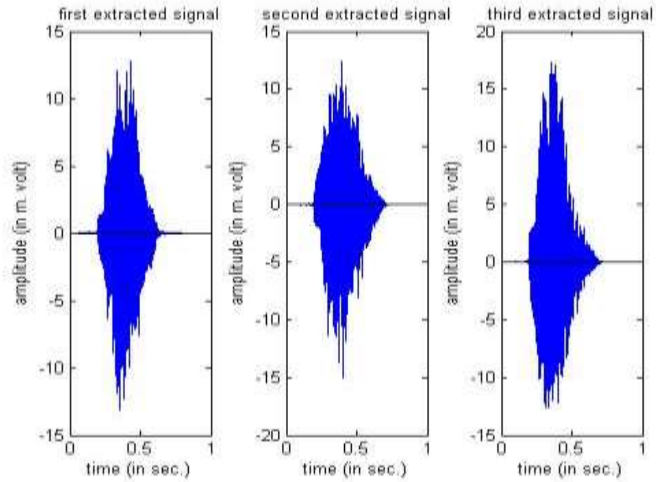
The following results show how the DWT based ICA algorithm outperforms the basic ICA algorithm. The correlation coefficient matrix between the original independent source signals and recovered/extracted signals is the factor to compare the two algorithms. It shows how much the source and recovered signals are correlated with each other. Due to the normalization of signals in both the algorithms, correlation only takes in to account the shape of signals and not the amplitude. Hence it can be used as a mathematical tool to compare conventional ICA with the DWT based ICA. Following are the plots of different signals taken under consideration or extracted as a result of algorithm. Plot for three independent source signals in MATLAB



(a) (b) (c)

Fig. 4.8: Independent source signals.

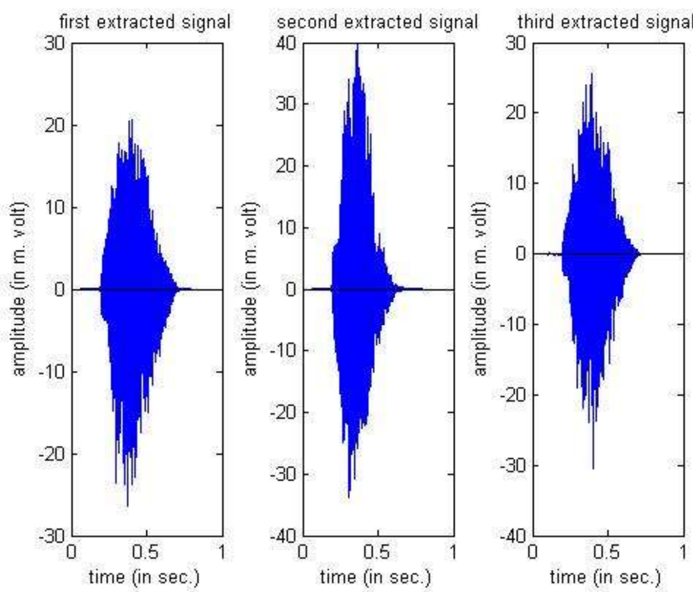
All rights reserved by www.ijirst.org



(d) (e) (f)

Fig. 4.9: Mixed signals as a result of linear mixing of Sources with unknown matrix.

ICA is applied on above shown mixed signals and independent source signals recovered by using 'tanh', 'exp' and 'Y^3' non-linearities (Iterations = 100, Step size = 1) are shown in fig 4.9, 4.10 and 4.11 respectively. Subplot (a), (b) and (c) are first, second and third extracted signals using 'tanh' nonlinearity respectively. Subplot (d), (e) and (f) are first, second and third extracted signals using 'exp' nonlinearity respectively. Subplot (g), (h) and (i) are first, second and third extracted signals using 'Y^3' nonlinearity respectively.



(a) (b) (c)

Fig. 4.10: Recovered signals by using 'tanh' non-linearity (Iterations = 100, Step size = 1)

Following table shows results of Correlation matrix by basic ICA and DWT based ICA.

Table - 4.1

Correlation Coefficients for a set of data containing three independent speech signals with different non-linearity, step size and number of iterations.

Non linearity	*Max iteration*	*Step size*	*Correlation matrix obtained by conventional ICA*	*Correlation matrix obtained by DWT based ICA*
Tanh	*100*	*1*	*0.7575 0.9903 0.6417* *0.1173 0.1484 0.6182* *0.6485 0.0673 0.4815*	*0.7860 0.2979 0.3172* *0.3974 0.7562 0.1988* *0.4700 0.5667 0.9150*
Tanh	*100*	*-1*	*0.8825 0.9434 0.8866* *0.2224 0.2138 0.2468* *0.4357 0.2815 0.4143*	*0.4776 0.4893 0.5032* *0.6651 0.6722 0.6893* *0.5762 0.5583 0.5251*
Tanh	*100*	*.5*	*0.9297 0.9295 0.2233* *0.0822 0.0406 0.1032*	*0.0508 0.9804 0.0946* *0.9663 0.0767 0.1715*

All rights reserved by *www.ijirst.org*

			0.3752 0.3802 0.9662	*0.0094 0.0912 0.9483*
Y^3	*100*	*1*	*0.9028 0.9431 0.8807* *0.1619 0.1978 0.3118* *0.4177 0.2932 0.3839*	*0.4942 0.5253 0.3827* *0.5945 0.6327 0.6957* *0.5713 0.3625 0.4873*
Y^3	*100*	*-1*	*0.9123 0.9142 0.9180* *0.2272 0.2115 0.1932* *0.3649 0.3690 0.3685*	*0.4757 0.5026 0.4683* *0.6842 0.6913 0.6745* *0.5548 0.5222 0.5722*
Y^3	*100*	*0.5*	*0.8543 0.7503 0.9841* *0.0093 0.5407 0.1279* *0.5307 0.4126 0.1542*	*0.4941 0.2437 0.3463* *0.4416 0.7247 0.8668* *0.3561 0.0378 0.1072*
exp(y)	*100*	*1*	*0.5996 0.7607 0.9255* *0.8112 0.3912 0.2169* *0.0532 0.5405 0.3077*	*0.2000 0.6830 0.1378* *0.7245 0.7087 0.1667* *0.6540 0.2116 0.9738*
exp(y)	*100*	*-1*	*0.9266 0.6645 0.4117* *0.3683 0.5647 0.7786* *0.1523 0.5152 0.4784*	*0.1999 0.0566 0.6095* *0.5002 0.2609 0.1196* *0.8381 0.9361 0.7809*
exp(y)	*100*	*0.5*	*0.6011 0.7612 0.9247* *0.8101 0.3917 0.2203* *0.0530 0.5394 0.3073*	*0.2001 0.6829 0.1409* *0.7255 0.7082 0.1685* *0.6529 0.2137 0.9730*

IV. CONCLUSION

In this dissertation, algorithm for separation of more than two linearly mixed speech signals by ICA is improved by using discrete wavelet transform based ICA. It contains two parallel processes, one process takes the high-frequency wavelet part of observations as its inputs, and the other process takes the low frequency part. From the results, it becomes evident that DWT based ICA outperforms the basic ICA algorithm in terms of correlation of source signals with the extracted signals. In future work, this algorithm can be improved by more detail study of deciding the nonlinearity and step size. The algorithm is still not fully capable of extracting more than three individual speakers properly. Hence some more improvement can be done in the future work on this part. Hardware implementation of algorithm by FPGA or other means is also the future scope for this work.

REFERENCES

[1] A. Belouchrani, K. A. Meraim, J.-F. Cardoso, and E. Moulines, (1997) "A blind source separation technique based on second order statistics," IEEE Trans. on Signal Processing, 45(2):434-444, 1997.
[2] A. Hyvarinen, (1999) "Fast and robust fixed-point algorithms for independent component analysis," IEEE Trans. on Neural Networks, 10(3):626-634, 1999.
[3] A. Hyvarinen, J. Karhunen, and E. Oja, Text book, (2001) "Independent Component Analysis," John Wiley &Sons, 2001.
[4] Benny Sällberg, NedelkoGrbic, and Ingvar Claesson (2008) "Complex-Valued Independent Component Analysis for Online Blind Speech Extraction," IEEE transections on Audio, Speech, and Language Processing, VOL. 16, NO. 8, 2008.
[5] D.P. Acharya and G. Panda, (2008) "A Review of Independent Component Analysis Techniques and their Applications," IETE TECHNICAL REVIEW, Vol. 25, ISSUE 6, 2008.
[6] Lawrence R. Rabiner and Ronald W. Schafer, (2007) "Introduction to Digital Speech Processing," Foundations and Trends in Signal Processing, Vol. 1, Nos. 1–2, 2007.
[7] Michele Gubian, Lou Boves, Francesco Cangemi, (2011) "Joint Analysis of F_0 and Speech Rate with Functional Data Analysis," IEEE (ICASSP 2011)
[8] Radoslaw Mazur and Alfred Mertins, (2011) "A Sparsity Based Criterion for solving the Permutation Ambiguity in Convolutive Blind Source Seperation," (ICASSP 2011) IEEE.
[9] SoroushJavidi, Clive Cheong Took, and Danilo P. Mandic, (2011) "Fast Independent Component Analysis Algorithm for Quaternion Valued Signals," IEEE TRANSACTIONS ON NEURAL NETWORKS, Vol. 22, No. 12, 2011.
[10] Zhengmao Ye, HabibMohamadian, and Yongmao Ye, (2008) "Independent Component Analysis for Spatial Object Recognition with Application of Information Theory Synthesis," International Joint Conference on Neural Networks (IJCNN 2008)
[11] Xing Fan and John H.L. Hansen, (2010), "Acoustic Analysis for Speaker Identification of Whispered Speech," IEEE (ICASSP 2010)