A Method for Enhancement of the Noisy Hybrid data using Independent Vector Analysis

Inam Ullah
Department of Electrical Engineering
University of Engineering & Technology
Peshawar, Pakistan

Taimur Ahmed Khan
Department of Electrical Engineering
University of Engineering & Technology
Peshawar, Pakistan

***Abstract* — This paper presents a method and technique of segregating noisy hybrid data using independent vector analysis (IVA) and minimum mean square error (MMSE) filter technique. Previously, method [1] is presented in which heterogeneous mixture is separated into components using independent vector analysis technique without including any type of noise in it. Similar work is continued with addition of additive white Gaussian noise (AWGN) in heterogeneous mixture (image and sound sources mixture). In proposed approach, IVA is applied to isolate the sources and for de-noising, minimum mean squared error (MMSE) filter is used. IVA is an effective method for segregating mixture of information that is originated from various hybrid sources into its components and it also evades permutation disorder, which usually occurs in independent component analysis (ICA). Previously homogenous sources were mostly mixed and segregated back into components using IVA & ICA techniques while in this paper, heterogeneous sources with the addition of AWGN noise are introduced. The improvement in the results is determined with signal-to-noise ratio (SNR).**

Keywords— Independent vector analysis (IVA), Heterogeneous sources, Blind Source Separation (BSS), Additive white Gaussian noise, minimum mean square error (MMSE).

I. INTRODUCTION

A. Blind Source Separation

Signal processing for estimating the actual signals from a mixture of observed signals is a challenging task where there is no adequate information about the sources and mixing environment. Blind source separation (BSS) technique is extensively used in this regard for extracting source signals from mixed speech. BSS leads to applications in many areas for estimation of source signals and images. Brain imaging is a superlative application for BSS. Biomedical engineering applications such as electro encephalogram (EEG) and Magneto encephalography (MEG) are intensively used nowadays to show the signals from muscle activity as well as the desired signal from brain activity. The two signals however can be separated using BSS to achieve the accurate estimation of Brain activity [2]. Other applications covering the communication side include speech improvement for the recognition of noise-robust speech, crosstalk segregation hearing-aid apparatus and Seismic waves Monitoring [3]. BSS generally applied in solving cocktail party scenario [4]. This technique is defined as converging on one auditory speech signal while filtering out the rest of the speech mixture.

Two generally used Blind source algorithms are IVA and ICA. For accurate implementation; both techniques are applied under few assumptions due to the fact that we are unaware regarding mixing atmosphere and source signals. These assumptions include norms that number of mixture signals, must be same or higher than the source signals.

B. Independent Vector Analysis (IVA)

Signal segregation may be performed using ICA algorithm but it faces the limitations of permutation disorder and scaling. Therefore, IVA method is used in the research as it does not face the problem of permutation and IVA technique remains dominant in estimation as it operates in frequency domain. Hence using short time Fourier transform, the signals are initially shifted to frequency domain and after separation of these signals from mixture the estimated signals using Inverse of short time Fourier transform are converted back to time domain. This technique has another advantage of using less memory. IVA model is described in the figure given below.
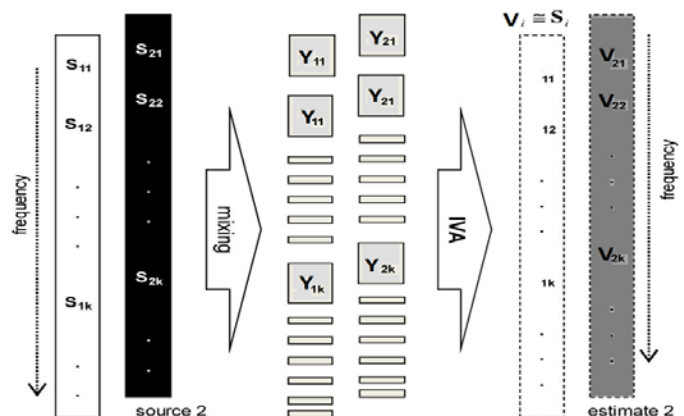

Fig. 1.   IVA functioning demonstrated in block diagram.

II. PROPOSED METHOD

IVA algorithm and minimum mean square error technique is implemented for segregation of noisy heterogeneous sources i.e. mixture of image and sound data. It is prominent that the data of images will be 2-dimensional and that of sound will be 1 dimensional. From literature survey it can be concluded that IVA and ICA had been applied to various homogenous mixtures. However, in the proposed method the implementation of IVA will be done on heterogeneous mixtures assuming that each source in mixture is degraded by

additive white Gaussian noise. Mainly two sorts of mixtures i.e. Instantaneous and convolutive mixtures could be used for segregation by BSS algorithms. Instantaneous mixtures does not take into account the effect of reflections while convolutive mixtures are those mixtures that are formed by including the effects of signal reverberations. In our proposed method instantaneous heterogeneous mixtures having source signals deteriorated by AWGN would be used for segregation and MMSE estimation technique would be applied for de-noising. Quality of segregated output will be measured by SNR.
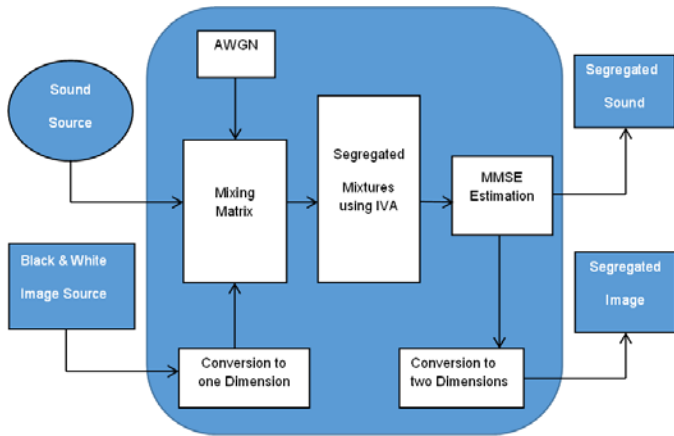
A. Methodology

Fig. 2.   IVA functioning demonstrated in block diagram.

Above block diagram explains the methodology of the proposed work. Initially, sound and image sources are taken and noisy mixture is formed. These signals are different in nature as, image is two dimensional and sound is one dimensional. Initially image source is converted to one dimension which is then mixed with sound source. We have two source signals a_1 and a_2 so we should have at least two mixtures x_1 and x_2. Mathematically as:

$$x_1 = a_1 h_{11} + a_2 h_{12} \qquad (1)$$

$$x_2 = a_1 h_{21} + a_2 h_{22} \qquad (2)$$

We can rewrite equation as

$$x = H.a \qquad (3)$$

Where $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $H = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix}$ & $a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ Matrix h is basically mixing matrix

$$x_i = \sum_j^L h_{ij} \circ a_j \qquad (4)$$

$$a_i \approx y_i = \sum_j^M w_{ij} \circ x_j) \qquad (5)$$

In expression (5), ○ is element-wise product, while L denotes number of sources and M represents number of mixtures.

$$x^{(d)} = H^{(d)}.a^{(d)}$$

AWGN is added to the mixture, where SNR specifies the signal to noise ratio per sample in dB

$$x\_noisy = awgn\,(x, snr) \qquad (6)$$

Once this noisy instantaneous mixture is generated out of heterogeneous sources, IVA technique is applied which will segregate our sources but these segregated sources are still noisy because of the addition of AWGN. Therefore, minimum mean square error filtering is used to remove noise and an output of segregated sound and 1-dimensional segregated image is achieved.

For the evaluation, we have opted for SNR, overall ∆SNR as:

$$\Delta SNR = SNR_{out} - SNR_{in}$$

Where SNR_{in} is measured between original sources signal and noisy mixture and SNR_{out} is measured between segregated output signals and noisy mixture. Improvement in SNR is depicted by ∆SNR.

III.   EXPERIMENTAL RESULTS AND DISCUSSION

Results in the form of ∆SNR are evaluated in this section, after applying IVA technique and MMSE filtering on noisy heterogeneous mixtures. Initially 04 image sources and 04 sound sources are randomly selected.
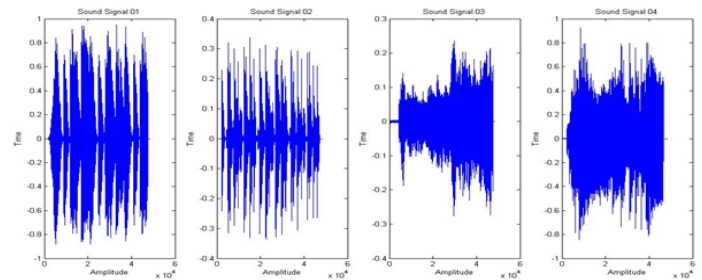

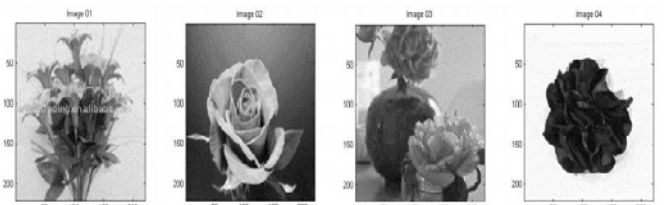Fig. 2.   Four randomly selected sound signals

Fig. 3.   Four randomly selected images

These sound signals and 1-D image signals are mixed together in various combinations. Appended below are series of results achieved from mixing of sound signal 01 and 1-D image signal.
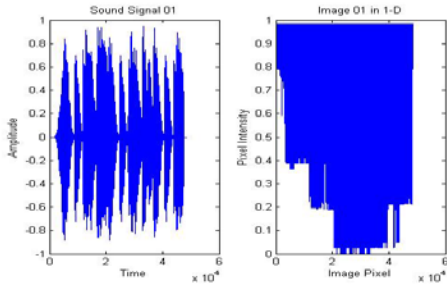
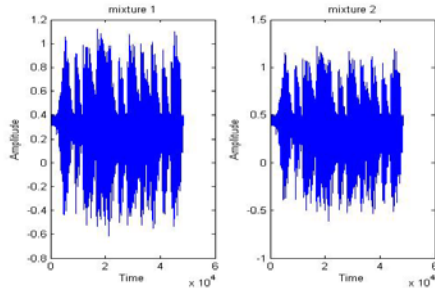Fig. 3. Sound signal 01 and Image 01 in 1-D form

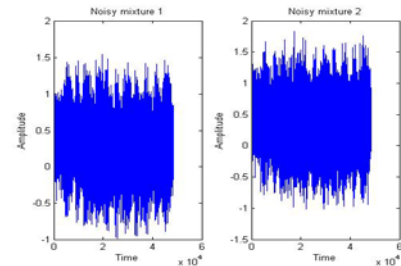Fig. 4. Mixture 1 & 2 formed from sound signal 01 and 1-D image 01

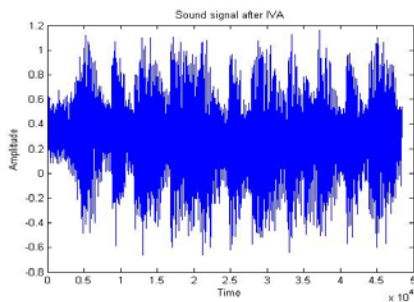Fig. 5. Noisy mixture 1 & 2 after addition of AWGN
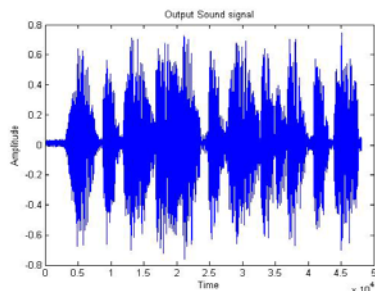

Fig. 6. Noisy sound signal after applying IVA

Fig. 7. Output sound signal after using MMSE filtering on segregated noisy sound signal

Results are evaluated in the form of SNR

$$\Delta \text{SNR} = SNR_{out} - SNR_{in} \qquad (7)$$

Where SNR_{in} and SNR_{out} (Jan et al., 2011) in time domain are:

$$\text{SNR}_{in} = 10\log_{10} \frac{\sum_t (a_i[t])^2}{\sum_t (a_i[t] - x_i[t])^2} \qquad (8)$$

$$\text{SNR}_{out} = 10\log_{10} \frac{\sum_t (x_i[t])^2}{\sum_t (x_i[t] - y_i[t])^2} \qquad (9)$$

In expression (8) & (9), $a_i[t]$ is original source component, $x_i[t]$ is noisy mixture and, $y_i[t]$ is segregated output in time domain. Appended below are segmental and overall SNR of various tests.

TABLE 1. Overall and segmental performance evaluation in SNR form between original and segregated sound sources

Test	Img	Wav	SNR Segmental (dB)			SNR Overall (dB)		
			SNR_{ou}	SNR_{in}	ΔSNR	SNR_{ou}	SNR_{in}	ΔSNR
1	01	01	-1.3	-9.9	8.5	1.20	-16.3	17.5
2	02	02	-0.6	-8.6	7.9	2.06	-9.42	11.4
3	03	03	-0.5	-8.7	8.2	0.94	-11.3	12.3
4	04	04	-3.9	-9.7	5.7	-2.61	-14.3	11.7
5	01	04	-0.8	-9.6	8.8	0.61	-13.5	14.1
6	02	03	-0.4	-8.5	8.0	1.02	-9.88	10.9
7	03	02	-0.6	-9.0	8.4	1.62	-10.8	12.4
8	04	01	-5.7	-9.8	4.1	-4.5	-17.1	12.6

In table 1, segment wise and overall SNR values are shown. ΔSNR is the performance evaluation of separated sound source compared to original source and noisy heterogeneous mixture. From data it can be observed that ΔSNR value varies for different mixtures and tests and on average 14dB improvement in signal quality is achieved.

TABLE 2. Overall and segmental performance evaluation in SNR form between original and segregated sound sources

Test	Img	Wav	SNR Segmental (dB)			SNR Overall (dB)		
			SNR_{ou}	SNR_{in}	ΔSNR	SNR_{ou}	SNR_{in}	ΔSNR
1	01	01	4.62	0.00	4.62	4.80	0.00	4.80
2	02	02	4.39	0.00	4.39	4.70	0.00	4.70
3	03	03	4.38	0.00	4.38	4.72	0.00	4.72
4	04	04	4.26	0.00	4.26	4.77	0.00	4.77
5	01	04	4.57	0.00	4.57	4.75	0.00	4.75
6	02	03	4.37	0.00	4.37	4.69	0.00	4.69
7	03	02	4.39	0.00	4.39	4.72	0.00	4.72
8	04	01	4.31	0.00	4.31	4.80	0.00	4.80

In table 2, segment wise and overall SNR values are shown. ΔSNR is the performance evaluation of separated image source compared to original source and noisy heterogeneous mixture. From data it can be observed that ΔSNR value varies for different mixtures and tests and on average 4.7dB improvement in signal quality is achieved. Due to scaling issue of IVA and MMSE limitations, ΔSNR value for separated image sources is found less.

A. Comparison with Previous Work

Test	Img	Wav	SNR Segmental (dB)			SNR Overall (dB)		
			SNR_{ou}	SNR_{in}	ΔSNR	SNR_{ou}	SNR_{in}	ΔSNR
1	01	01	-15.0	-16.3	1.30	1.20	-16.3	17.5
2	02	02	-8.28	-9.4	1.14	2.06	-9.4	11.4
3	03	03	-10.1	-11.3	1.21	0.94	-11.3	12.3
4	04	04	-13.2	-14.3	1.10	-2.61	-14.3	11.7
5	01	04	-12.3	-13.5	1.20	0.61	-13.5	14.2
6	02	03	-8.77	-9.88	1.11	1.02	-9.88	10.9
7	03	02	-9.75	-10.8	1.12	1.62	-10.8	12.4
8	04	01	-16.1	-17.2	1.11	-4.5	-17.2	12.7

In heterogeneous Blind source segregation method [1], overall 15 dB SNR was achieved without incorporating any type of noise. Therefore in the proposed approach, removal of noise from noisy heterogeneous mixture with post-processing MMSE approach was a challenging task.

As compared to the results obtained by [1] without applying MMSE post-processing, our proposed method has shown considerably higher performance. Out of random 50 experiments and tests, 08 results are shown and it can be observed that with the addition of AWGN in the method [1] mixtures, the SNR gain is decreased from average 15 dB values to 1dB. However with incorporation of MMSE post-processing, the overall SNR gain of the segregated sources is improved.

IV.   REFERENCES

[1] M. Shah, T. Jan, A. Jehangir and A. Ali, "A novel method for blind segregation of speech and image data using independent vector analysis," Sindh Univ. Res. Jour. (Sci. Ser.) Vol.46 (3) 283-286 (2014).

[2] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short time spectral amplitude estimator," IEEE Trans. Acoust, Speech, and Signal Process., vol. 32, no. 6, pp. 1109–1121, 1984.

[3] S. Amari, A. Cichocki, and H. H. Yang., "A new learning algorithm for blind signal separation, Advances Neural information Processing Systems," vol. 8 pp 757-763, 1996.

[4] M. Wu and D. L. Wang, "A two-stage algorithm for one-microphone reverberant speech enhancement," IEEE Trans. Audio, Speech, Lang. Process., vol. 14, pp. 774–784, May 2006.

[5] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise assisted data anlysis method," Advances in Adaptive Data Analysis, vol. 1, pp. 1–41, Jul. 2008.

[6] A. Hyvarinen, and E. Oja, "Independent Component   Analysis", John Wiley and Sons, New York, USA 2002.

[7] M. Z. Ikram and D. R. Morgan, "A Beamforming Approach to Permutation Alignment for Multichannel Frequency Speech Separation", Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Orlando, USA, 2002, pp 881-84.

[8] T. Jan and W. Wang, "Empirical mode decomposition for joint denoising and dereverberation, 19th European Signal Processing Conference (EUSIPCO 2011)", Barcelona, Spain, 206-210.

[9] N. Murata, S. Ikeda, and A. Ziehe, "An approach to blind source separation based on temporal structure of speech signals", Neurocomputing, vol. 41, pp 1–24, 2001.