

A Fast Adaptive Speech Extraction Method using Blind Source Separation for Audio Signal Processing



Mandli Rami Reddy, M L Ravi Chandra, Alam Siva sankar

Abstract: The adaptive signal processing methods are used in several applications like channel estimation, Noise removal and extraction of signals also. The methods vary on time, frequency and statistical approach. In this paper, the source speech signals are separated using different methods like FastICA, PCA and kICA. Comparison of original signal and estimated signals are evaluated for different methods. The implementation was done in MATLAB. The spectrogram, Negentropy and Kurtosis waveforms are plotted for different methods.

KeyWords: BSS, ICA, noise, speech, spectrogram, Negentropy, Kurtosis, statistical.

I. INTRODUCTION

The issue in Blind Source Separation (BSS) is increasing faster by day to day usage. This issue is found similar in other application such as, multi-path channel identification, equalization and direction of arrival (DOA), speech enhancement estimation and crosstalk removal in multichannel in sensor arrays. By using this application the higher-order statistics is improved by generating new technique for identifying statistically independent signals in signal modeling. The separation of source issue that is harmful at the heart which is developed by the signal processing and also by machine learning which is driven mainly as a density estimation task. The BSS is one of the main uses of separating the signals. The process of separation of voice signals of people at same time called BSS. The main problem in voice signal is the cocktail party problem. This problem is rectified by algorithm called the independent component analysis (ICA) technique. The main process technique used in this method is to detect the sound with single object in various sound environments. Figure. 1 shows the cocktail party problem which is the best example of two vocal signals. The voices have two types of source that are recorded from two independent source signals. Hence the problems are carried out and solved by extracting the original signals using independent component analysis (ICA) technique.

The ICA is an important techniques used for extraction. This technique works on an extension of the principal component analysis (PCA). The PCA is a technique which optimizes the covariance matrix of the data in statistics of second-order. Hence higher order statistics can optimize ICA as kurtosis. This can be processed to find uncorrelated components with independent components. Thus PCA performs at the higher-order correlations in which it is extracted independently when the sources of mixture data are insignificant.

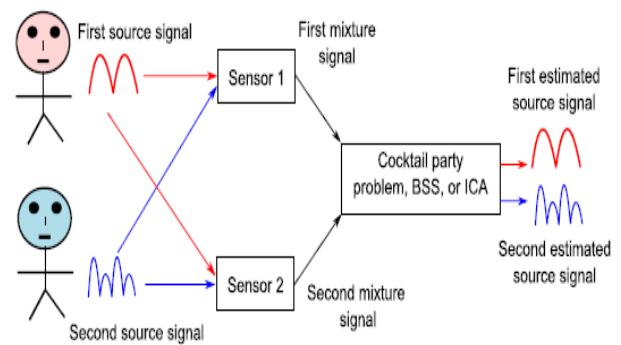


Fig. 1 The cocktail party problem.

From the figure 1 found that two source signals are generated from separate individuals. The two sensors are then recorded by microphones and later on the two source signals is mixed. Thus by using this process the original signals is recovered from the mixed signals.

II. LITERATURE SURVEY

Parra et al, (2000) performs optimization technique using algorithm utility for automatic speech recognition. This methodology simulates the acoustic signal that is recorded in a reverberant environment. The recorded signals are the sums of differently convolved sources. Thus the process identifies the unknown channel and optimizes the channel. Y. Yang et al, (2011) introduce a temporal predictability based Blind Source Separation (BSS). This method used to separate the signal from the mixed one. Thus the simulation shows that the signals are separated to get individual noise source signals. Zhinong Li et al, (2008) compares the machine faults in linear BSS method. The source separation method is presented which is based on linear BSS. If the machine is in nonlinear mixing source then it is effective in BSS method.

Revised Manuscript Received on February 28, 2020.

* Correspondence Author

Mandli Rami Reddy*, ECE Department, Srinivasa Ramanujan Institute of Technology, Ananthapuram, Andhra Pradesh, India. Email: mandliramireddy@gmail.com

Dr M L Ravi Chandra, ECE Department, Srinivasa Ramanujan Institute of Technology, Ananthapuram, Andhra Pradesh, India. Email: mlravigates@gmail.com

Dr Alam Siva sankar, ECE Department, Srinivasa Ramanujan Institute of Technology, Ananthapuram, Andhra Pradesh, India. Email: alamsivasankar1971@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Thus the result is based on the separation of source signals. S. Van Vaerenbergh and I. Santamaria (2006) describe the different nonlinearities system. This system inverts the linear BSS which is based on clustering approach to solve problems in underdetermined post nonlinear blind source separation (PNL BSS). Thus the method transforms the nonlinear mixture component to solve undetermined BSS problem.

H. Sawada et al, (2007) introduced a blind source separation (BSS) for optimizing the group frequency components. This method analysis ICA results for all estimation sensors for each source. Thus the process shows the effective separation in several sources that are configured in low moderate. Bin Zhao et al, (2005) discussed a novel blind separation method to determine the separated signals. At the receiver end the signals separation numbered to separate when direction-of-arrival (DOA) is obtained. Thus the signal is separated by the individual communication signals on source separation. O. Shifeng et al, (2009) presented a novel variable step size algorithm to restructure. The performance index has been restructured nonlinear based algorithm for updating rule of step-size. This algorithm is used by adopting an auxiliary separation system. Thus the algorithm performs the steady state in both stationary and non-stationary system. Qi Lv and Xian-Da Zhang (2006) simulate the speech signal using Blind source separation (BSS) to validate the higher applications. In this simulation BSS is implemented without prior assumption on the number of sources. Hence C prototypes algorithm is used as the new type of BSS method to estimate the mixing matrix.

K. J. Faller et al, (2017) investigate source separation algorithms to improve intelligibility of speech. This method can enhance the spatial hearing of hearing aids. Thus the BSS algorithm will modify the speech source that is simulated as spatial audio. Y. Zhang and S. A. Kassam (2010) discuss complex blind source separation. This algorithm performs via EASI system. The process is carried out with the QAM signal which is separated. This technique is based on magnitude-phase which represents the complex signals and circularly symmetric source. Z. Li et al, (2010) presented the whitening and non-linear de-correlation based Blind source separation algorithm. The ICA is also be used in this to take results in neural network and signal processing. Thus the method simulated and analyzed using BBS algorithm and the corresponding result gives better convergence speed and steady-state error. Yoshihiro Sakai et al, (2007) discuss the BSS algorithm to improve the convergence rate. This methodology uses the blind signal separation circuit to reduce failure during double-talk in the echo canceller. Thus the circuit describes the constitution method to get improve characteristic of blind signal separation.

J. Ma and X. Zhang (2008) presented blind separation algorithm to get instantaneous linear mixture signals for low computational complexity. This method will characterize based on Signal Noise Ratio (SNR) at maximal condition. The source signals and noises have same eigenvalue (GE) problem. Thus the method is compared with the algorithm to have effective low complexity in computation. B. Xia and H. Xie (2007) discuss two main problems that affect blind source separation in temporal

correlated signals. Instantaneous BSS and Convolutional BSS are the two problems that can simplify residual signals by cost function. Thus the technique compares the simulation and undergoes the separation process to extract temporal structure in residual part of source signals. Tao Xu and Wenwu Wang (2009) presented K-means clustering algorithm to estimate unknown mixing matrix from audio mixture. The separation of audio signals occur some problem while address the sparse signal representation. The algorithm is processed under two stages K-means clustering algorithm and conventional approaches. Thus the method gives better performance by comparing recent sparse representation approach.

III. BACKGROUND METHODOLOGY

3.1. PRINCIPAL COMPONENT ANALYSIS

The principal component analysis (PCA) is used to estimate the average value in sample. This technique consist of observed vector x to remove its mean. After the process of removal the vector will transform into a new vector. Some of the possibilities of vector are lower dimension whose elements are uncorrelated with each other. Hence the process is carried out by evaluating the covariance matrix by Eigen value decomposition which is found by the linear transformation. Thus the covariance matrix C_x with a zero-mean vector x is shown in equation (1).

$$C_x = E\{XX^T\} = EDE^T \quad (1)$$

Where,

$$E = (e_1, e_2, \dots, e_n) = C_x$$

C_x represents the eigenvectors in orthogonal matrix.

$D = \text{Diag}(\lambda^1, \lambda^2, \dots, \lambda^n) = C_x$ in Diagonal matrix of Eigen value.

Whitening can be shown as

$$Z = P^* X \quad (2)$$

Where, P denotes the whitening matrix and Z denotes white new matrix.

P is represents as;

$$P = D^{-0.5} \times E^T \quad (3)$$

3.2. ICA

The signals are varies from time and it is represented as, $s_i = \{s_{i1}; s_{i2}; \dots; s_{iN}\}$, the number of time steps is denoted by N and s_{ij} is the amplitude of the signal, s_i at the j th time. Given two independent source signals $s_1 = \{s_{11}; s_{12}; \dots; s_{1N}\}$ and $s_2 = \{s_{21}; s_{22}; \dots; s_{2N}\}$ (see Fig. 1). Both signals are given by;

$$S = \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} = \begin{pmatrix} (S_{11}, S_{12}, \dots, S_{1N}) \\ (S_{21}, S_{22}, \dots, S_{2N}) \end{pmatrix} \quad (4)$$

Where, $S \in R^{P \times N}$ denotes the space and also defines the source signal.

The source signal indicates p.

Both S1 and S2 are the mixed source signals as $X_1 = a * s_1 + b * s_2$. Here a and b are the mixing coefficients in x_1 . Hence the mixture x_1 is the weighted as sum of the two source signals. X_2 mixture is repeated as the same process and the distance between the source signals and the sensing device is changed for measuring. The mathematical representation is shown as $X_2 = c * S_1 + d * S_2$. Where, c and d are mixed coefficients. The mixed coefficients are different from coefficients c and d due to sensing devices that has both signals in different locations. Thus the source signals are measured with each sensor in a different mixture. Hence the corresponding output in source signal has different impact which is represented as follows:

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} aS_1 + bS_2 \\ cS_1 + dS_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} = AS \quad (5)$$

Here $X \in R^{n \times N}$ denotes the mixture signals. Where n is the number of mixtures. Thus (figure 1) the mixing coefficients such as a; b; c, and d are utilized for transforming linearly source signals. This source signals are mixed to space signals S in X space, $S \rightarrow X : X = AS$, where $A \in R^{n \times P}$ is the mixing coefficients matrix:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (6)$$

Properties of Mixed Signals:

- 1. Independence:** If signals are shared between the mixtures then it is independent when the source signals are independent to mixture signals.
- 2. Gaussianity:** The gaussianity is a process of mixing signals in histogram that are bell shaped. This can be used for searching for non-Gaussian signals within mixture signals. The signals are extracted independently when they must be non-Gaussian. Hence the signals are estimated independently when they have fundamental restriction in ICA.
- 3. Complexity:** This is more complex than source signals which is shown from the previous example of mixed signals. The extracted signals are independent and they are non-Gaussian histograms with these signals which represent source signals.

3.2.1. MEASURING NON GAUSSIANTY

Kurtosis:

The Kurtosis is used to measure non-Gaussianity by the absolute value of kurtosis. This theorem has central limit which is strong measure signal with traditional higher order statistics that uses kurtosis independent. The kurtosis is defined by zero-mean random variable v given as;

$$Kurt(v) = E\{v^4\} - 3(E\{v^2\})^2 \quad (7)$$

Where $E\{v^4\}$ = Fourth moment of v, and $E\{v^2\}$ = Second moment of v.

Here, the $E\{v^4\}$ equals $3(E\{v^2\})^2$ where v is the Gaussian random variable. From the equation (7) if v is a non-Gaussian random variable then the kurtosis is zero. Here a particular kurtosis value can be either positive or negative. The positive value is called Super-gaussian and negative value is called Sub-gaussian. Super-gaussian random variables and sub-gaussian random variables both have a spiky probability density function and flat probability density function.

3.2.2 Negentropy

Negentropy is another main technique used to measure the non-gaussianity. The Negentropy works in different entropy based on the information theoretic quantity. The entropy means that it can be interpreted on random variable which is the basic concept of information theory as the degree of information. Entropy can observe the random variable that is unpredictable and unstructured in the larger entropy. This technique is closely related to the coding length of the random variable.

3.3. FastICA:

This method is highly efficient for computing the signal by using FastICA algorithm. The FastICA is used to estimate ICA performance which uses a fixed-point iteration scheme that are found. For ICA the method could be 10-100 times faster than conventional gradient descent method. The advantage of FastICA algorithm is that it can be used to perform projection pursuit as well as providing a general-purpose data analysis method.

The Steps involves Fast ICA algorithm:

1. It makes the mixed data available at zero mean.
2. Whiten the data.
3. The initial weight vector w of unit norm is taken.

$$w_{norm} = \frac{w}{\|w\|} \quad (8)$$

Let

$$w_{new} = E\{m_i g(w^T m_i)\} - E\{m_i g'(w^T m_i)\} w \quad (9)$$

The equation shows the basic weight update, where g is the contrast function.

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

Thus the normalization step makes the new w as unit norm and it will update in each iteration. Compare w_{new} with the old vector, if converged than move ahead, if not go to step 4.

IV. PROPOSED METHODOLOGY

4.1. KERNEL INDEPENDENT COMPONENT ANALYSIS

The Kernel ICA (KICA) is an algorithm that process non-linear transformation by combining KPCA with ICA. The ICA deal with the sample data using basic idea of the KICA which is mapped to high dimension feature space by using nonlinear characteristics. This method confirms to Mercer condition. In this technique the KPCA uses the linear principle component analysis to deal with the sample data for mapping high dimension feature using a



nonlinear mapping transformation.

Common use of kernel functions;

1. The Radial Basics Gaussian Function is given in equation (11) which defines infinite dimension in feature space.

$$k(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad (11)$$

2. Equation (12) represents the Polynomial Kernel Function defines finite dimensions

$$k(x, y) = ((x.y) + c)^2 \quad (12)$$

V. RESULT AND DISCUSSION

Thus the experiment shows the uses of source speech signals. Different methodology of separations are mixed and then used for evaluation. The experiment is performed in MATLAB platform and the corresponding results taken using BSS algorithms that are FastICA, PCA and KICA. The waveform and spectrogram of source signals is shown in figure (2).

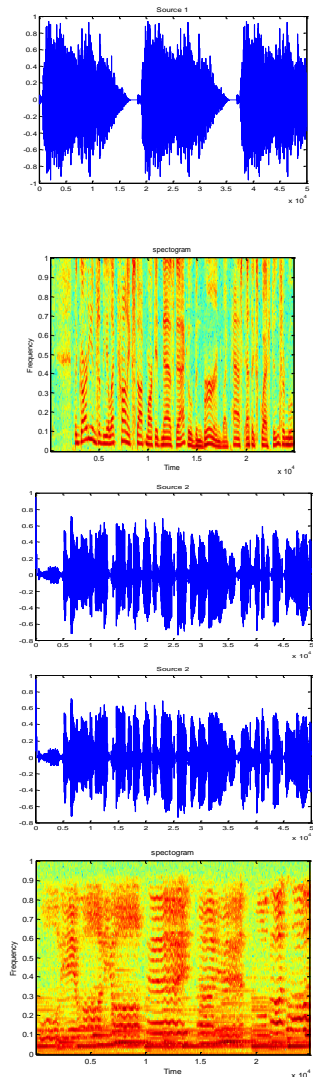


Fig 2. Source signal and its spectrogram

The corresponding output waveform of principle component analysis is shown in figure (3). Hence, fig 3(a) represents principle components of signal and comparison of original signal and estimated PCA signal is shown in fig 3(b).

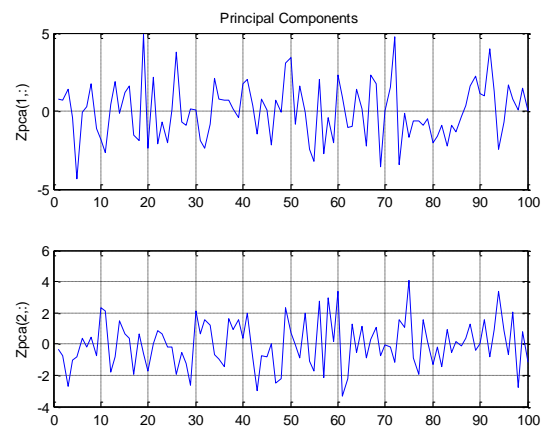


Fig 3(a): Principle components of source signal

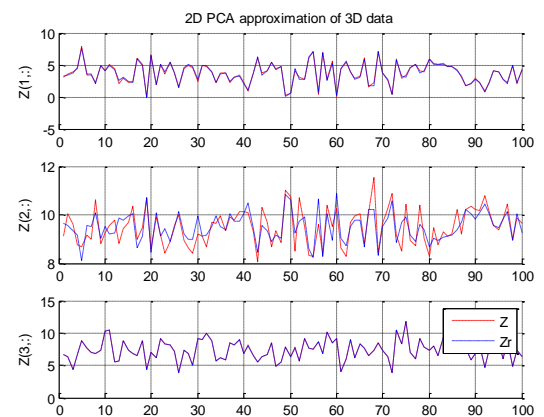
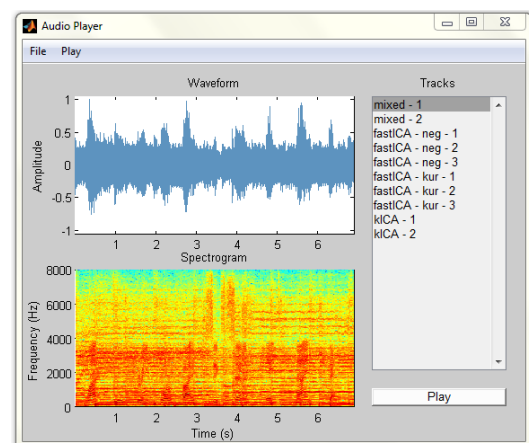


Fig 3 (b): original and estimated signal



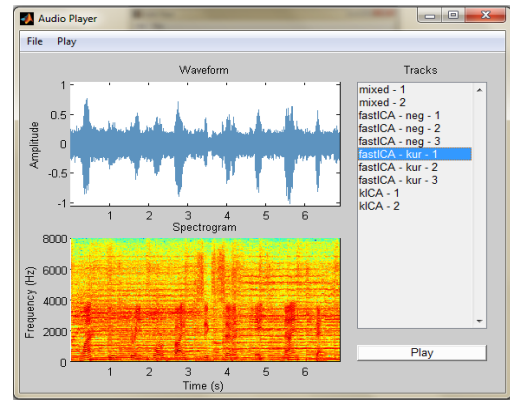
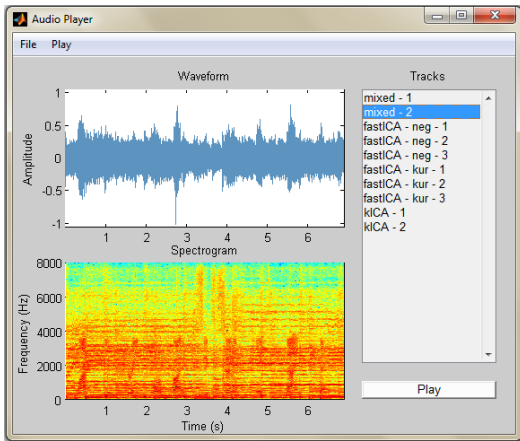


Fig 4: waveform and its Spectrogram of mixed signals
 Here the waveform and spectrogram of mixed signals is shown above in figure (4). Figure 5 represents the signals that are processed using FastICA based on Negentropy and Kurtosis components.

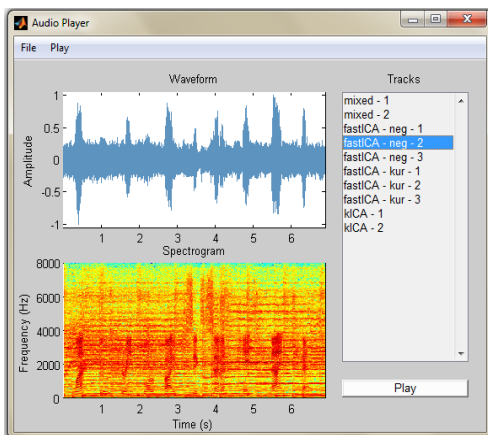
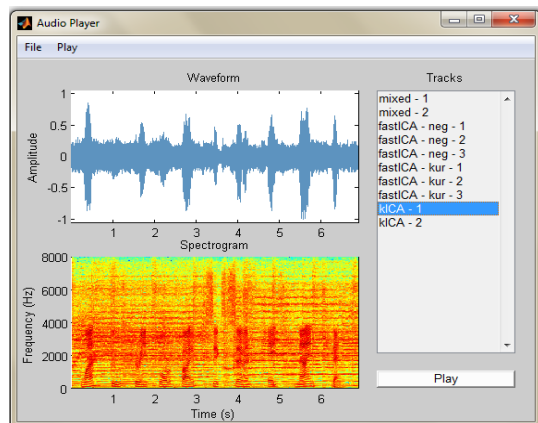
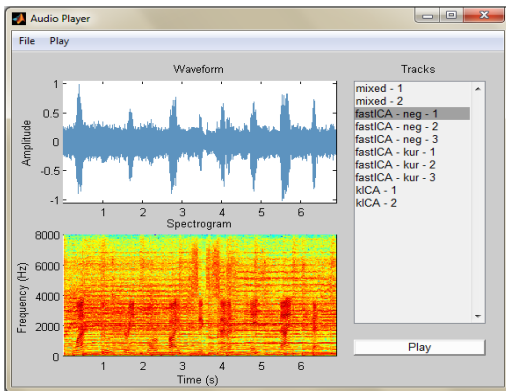
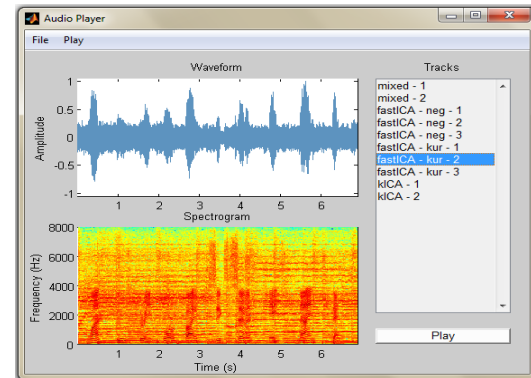


Fig 5(b): Kurtosis waveforms using FastICA
 The KICA output waveform and its spectrogram is shown in figure (6).

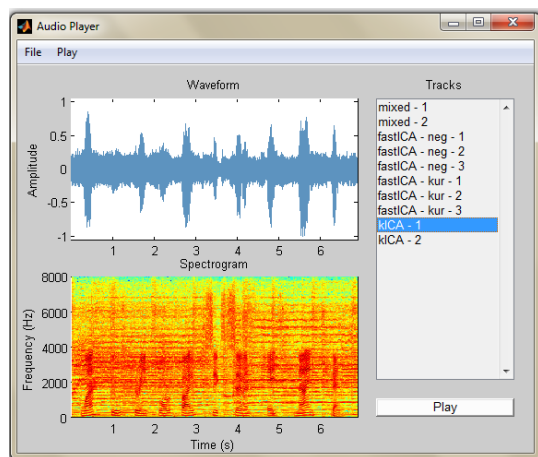
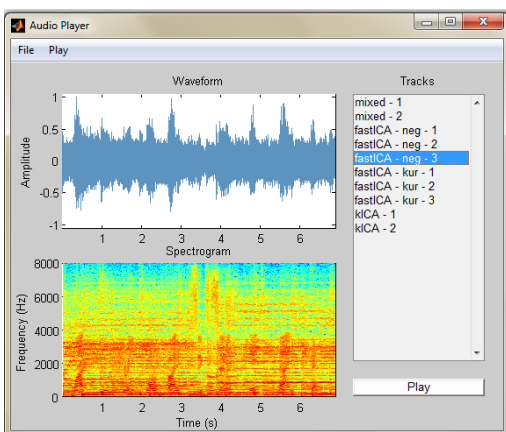


Fig 5(a): Negentropy waveforms using FastICA

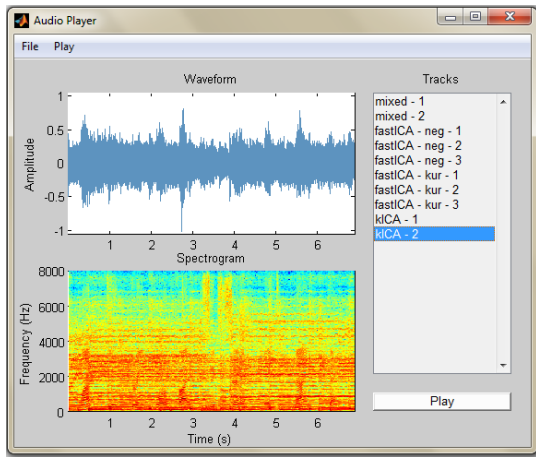


Fig 6: Estimated waveform and Spectrogram using kICA

Table 1: Mean Square Error(MSE) for PCA, FastICA and kICA

	MSE
PCA	0.0718
FastICA (negentropy)	0.0296
FastICA (kurtosis)	0.0292
kICA	0.0254

Table 1 shows the Mean square error (MSE) for different technique.

VI. CONCLUSION

The paper presented the efficiency of the blind source separation (BSS) method on signal separation process. The basic idea is the separation of sources which are statistically independent. The methods are complex in nature and execution but found efficient. Several methods like FastICA, PCA and kICA are implemented in MATLAB. The methods are tested using sources of speech signals. The spectrogram, Negentropy and Kurtosis waveforms are plotted for different methods.

REFERENCES

1. Parra, Lucas C. and Clay D. Spence. "Convolutive blind separation of non-stationary sources." IEEE Trans. Speech and Audio Processing, Vol.8, No.3, pp.320-327,2000.
2. Y. Yang, Z. Li, X. Wang and D. Zhang, "Noise source separation based on the blind source separation," 2011 Chinese Control and Decision Conference (CCDC), Mianyang, 2011, pp. 2236-2240.
3. Zhinong Li, Yudong Zeng, Tao Fan, Yaping Lv and Jing Ren, "Source Separation method of Machine Faults Based on Post-Nonlinear Blind Source Separation," 2008 7th World Congress on Intelligent Control and Automation, Chongqing, 2008, pp. 1786-1789.
4. S. Van Vaerenbergh and I. Santamaria, "A spectral clustering approach to underdetermined postnonlinear blind source separation of sparse sources," in IEEE Transactions on Neural Networks, vol. 17, no. 3, pp. 811-814, May 2006.
5. H. Sawada, S. Araki, R. Mukai and S. Makino, "Grouping Separated Frequency Components by Estimating Propagation Model Parameters in Frequency-Domain Blind Source Separation," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 15, no. 5, pp. 1592-1604, July 2007.
6. Bin Zhao, Jun-An Yang and Min Zhang, "Research on blind source separation and blind beamforming," 2005 International Conference on Machine Learning and Cybernetics, Guangzhou, China, 2005, pp. 4389-4393 Vol. 7.
7. O. Shifeng, G. Ying, J. Gang and Z. Xuehui, "Variable Step Size Algorithm for Blind Source Separation Using a Combination of Two Adaptive Separation Systems," 2009 Fifth International Conference on Natural Computation, Tianjin, 2009, pp. 649-652.

8. Qi Lv and Xian-Da Zhang, "A unified method for blind separation of sparse sources with unknown source number," in IEEE Signal Processing Letters, vol. 13, no. 1, pp. 49-51, Jan. 2006.
9. K. J. Faller, J. Riddle and E. Grubbs, "Automatic blind source separation of speech sources in an auditory scene," 2017 51st Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, 2017, pp. 248-250.
10. Y. Zhang and S. A. Kassam, "Complex blind source separation: optimal nonlinearity and approximation," 2010 44th Annual Conference on Information Sciences and Systems (CISS), Princeton, NJ, 2010, pp. 1-6.
11. Z. Li, J. An, L. Sun and M. Yang, "A Blind Source Separation Algorithm Based on Whitening and Non-linear Decorrelation," 2010 Second International Conference on Computer Modeling and Simulation, Sanya, Hainan, 2010, pp. 443-447.
12. Yoshihiro Sakai, Kota Takahashi and Wataru Mitsuhashi, "An algorithmic study on blind source separation for preprocessing of an acoustic echo canceller," SICE Annual Conference 2007, Takamatsu, 2007, pp. 1400-1405.
13. J. Ma and X. Zhang, "Blind Source Separation Algorithm Based on Maximum Signal Noise Ratio," 2008 First International Conference on Intelligent Networks and Intelligent Systems, Wuhan, 2008, pp. 625-628.
14. B. Xia and H. Xie, "Blind Source Separation of Temporal Correlated Signals," 2007 Third International IEEE Conference on Signal-Image Technologies and Internet-Based System, Shanghai, 2007, pp. 549-555.
15. Tao Xu and Wenwu Wang, "A compressed sensing approach for underdetermined blind audio source separation with sparse representation," 2009 IEEE/SP 15th Workshop on Statistical Signal Processing, Cardiff, 2009, pp. 493-496.
16. A. Siva Sankar, Dr.T. Jayachandra Prasad, Dr.M.N.Giriprasad, "LSB based Lossless Digital Image Watermarking using Polynomials in Spatial Domain for DRM", International Journal of Computer Association (IJCA), New York, USA, ISBN: 978-93-80747-76-3, No. 13, pp. 18-23, 2011. www.ijcaonline.org/proceedings/icwet/number13
17. Mandli Rami Reddy, T. Keerthi Priya, k. Prasanth, S.Ravindrakumar "Robust Adaptive Estimator using Evolutional algorithm for Noise Cancellation In Multichannel System " Chapter DOI: 10.1007/978-981-13-8942-9_55

AUTHORS PROFILE



M.Rami Reddy, received the bachelor degree in Electronics and communication engineering from Sri Venkateswara College of engineering and technology (SVCET) Chittoor in 2005 and received the post graduate in communication and signal processing from G.Pulla Reddy engineering college (GPREG), Kurnool in 2010. He is currently working as Assistant professor in the srinivasa ramanujan institute of technology (SRIT). His research activities are in adaptive signal processing with a particular interest sub band least mean square algorithm. He handled various subjects for post graduate students. He is having 13 years of teaching experience. He has published 9 international papers and four national paper. He is life member of ISTE, IEI, IETE.



Dr.M.L.Ravi Chandra, obtained his B.Tech degree in Electronics and Communication Engineering, from KSRMCE, Affiliated to S.V. University and Master of Technology in Digital Systems and Computer Electronics from JNTUACE, Anantapur, A.P. India. He obtained his Ph.D. Degree from JNTUHCE, Hyderabad, Telanagana, India. He worked at various engineering colleges in Andhra Pradesh and Telangana as Assistant Professor, Associate Professor, Professor and Head of the Department. Presently he is working as a Professor and Head for ECE Department at Srinivasa Ramanujan Institute of Technology, Anantapur, A.P., India from December, 2015 to till date. He is having 20 years of teaching experience. He has 15 technical publications in International Journals and 7 publications in International Conferences. He is a life member of IE(I) and Member of IEEE.



Alam Siva Sankar, obtained his Diploma in Electronics and communication Engineering at S.V. Government Polytechnic, Tirupati, A.P. and obtained his AMIE in Electronics and Communication Engg., from The Institution of Engineers (India), Kolkata 700 020 , and Master of Technology degree in Digital Electronics and Advanced Communication from Manipal Institute of Technology(MIT), Manipal – 576104, Karnataka state, India. He obtained his Ph.D. Degree (Digital Rights Management/Digital Image Watermarking) in ECE from JNTUACE, Anantapur, A.P., India. He worked in Hindustan College of Engineering, Chennai from Jun, 1997 to April, 2002 as a lecturer. He worked as a Lecturer in ECE Dept. in Anand Institute of Higher Technology, Chennai, from Sep, 2003 to June, 2004. He worked in Gokula Krishna College of Engineering, sullurpet from July, 2004 to April, 2012 in various positions such as Assistant Professor and HOD, Associate Professor and HOD and Professor and HOD. He worked in Priyadarshini College of Engineering, Sullurpet from May, 2012 to April, 2015 as a Professor and Principal. Worked in Priyadarshini Institute of Technology, Tirupati from May, 2015 to March, 2016 as a Professor and Academic Dean. Worked in KMM institute of Technology and Science, Tirupati, from April, 2016 to June, 2018 as a Professor and Principal. Presently working as a Professor in ECE Dept. in Srinivasa Ramanujan Institute of Technology, Anantapur, A.P., India to till date. He is having more than 20 years of teaching experience and has more than 16 technical publications in International Journals and Conferences. He is a life member of ISTE (India), Associate Member of Institution of Engineers (India). He is a Member of IEEE.