Research Article
A Modified Infomax ICA Algorithm for fMRI Data Source Separation

[1]Amir A. Khaliq, [2]I.M. Qureshi, [1]Suheel A Malik and [1]Jawad A. Shah
[1]Department of Electronic Engineering, International Islamic University Islamabad, Pakistan
[2]Department of Electronic Engineering Air University, ISSS Islamabad, Pakistan

Abstract: This study presents a modified infomax model of Independent Component Analysis (ICA) for the source separation problem of fMRI data. Functional MRI data is processed by different blind source separation techniques including Independent Component Analysis (ICA). ICA is a statistical decomposition method used for multivariate data source separation. ICA algorithm is based on independence of extracted sources for which different techniques are used like kurtosis, negentropy, information maximization etc. The infomax method of ICA extracts unknown sources from a number of mixtures by maximizing the negentropy thus ensuring independence. In this proposed modified infomax model a higher order contrast function is used which results in fast convergence and accuracy. The Proposed algorithm is applied to general simulated signals and simulated fMRI data. Comparison of correlation results of the proposed algorithm with the conventional infomax algorithm shows better performance.

Keywords: Blind source separation, functional Magnetic Resonance Imaging (fMRI), independent component analysis, medical image processing

INTRODUCTION

The study of the human brain can be made by PET, SPECT, fMRI etc. fMRI is the latest technique used for measuring the functionality of the brain using blood oxygen level dependent signal (BOLD) (Ogawa *et al.*, 1990). In fMRI the magnetic properties of the brain tissues associated with an activity are recorded in the form of 3-D images of the brain (Ogawa *et al.*, 1990). For a specific activity a set of neurons becomes active in performing some task, thus it needs more oxygenated blood as compared to nearby idle neurons. The ratio of the oxygenated blood between active and idle neuron changes and the change in magnetic properties of the oxygenated blood is recorded by the fMRI scanner in the form of Blood-Oxygen-Level-Dependent (BOLD) signal (Ogawa *et al.*, 1990). A number of images are taken rapidly by the fMRI scanner during the experiment. This multi image data is then statistically analyzed for active and non-active voxels and the results are presented on a statistical map. However, due to small variations of the BOLD signal, SNR values of functional MR images are too low (Friman *et al.*, 2001) and thus it requires a de-noising step before classification.

During fMRI experiment subjects are normally performing activity and non-activity during predefined cycles of time. The BOLD signal from each brain voxel is recorded along with time thus creating time series signals which are mixtures of activity non-activity and noise and structural information etc. Location of activated voxels in this mixed data is a challenging task (McKeown and Sejnowski, 1998).

Analysis of fMRI data has been done by different approaches including Statistical Parametric Mapping (SPM), Principal Component Analysis (PCA), ICA correlation analysis, Non-negative matrix factorization etc.

fMRI data analysis approaches can be broadly divided in two categories, one is model dependent and other one is data dependent. In model dependent approaches knowledge of experimental model is required a priori. These approaches consist of SPM, time frequency analysis, correlation analysis etc.

Statistical Parametric Mapping (SPM) is a univariate fMRI data analysis approach in which the signal at each voxel is tested using t-test, f-test (Friston, 1996). In time frequency analysis approach assumption is made that task related noise and physiological activities spectra's are different and thus can be extracted from the mixed observed data (Mitra *et al.*, 1997). In Correlation analysis approach, the voxels which are responsible for activity/non-activity are extracted using a priori knowledge of the experimental paradigm. A correlation of the experimental model and each time series is performed and the voxels are termed as active voxels in case of strong correlation results (Bandettini *et al.*, 1993). Data driven approaches of fMRI analysis do not require any experimental model a priori like Principal Component Analysis (PCA) (Backfrieder,

1996) and Independent Component Analysis (ICA) (McKeown *et al.*, 2003), NMF (Amir *et al.*, 2012) etc.

Principal Component Analysis (PCA) is an important method used under data driven methods. Main assumption made when applying PCA is that the underlying sources variability. Normally the number of principal components representing the data are less than the actual dimension of the data thus the dimensionality of the data is also reduced by discarding the principal components with lowest variance. There are other multi-variate data analysis approaches, but the most promising approach is ICA, which is a data driven and powerful exploratory tool for fMRI data analysis (Qiu-Hua *et al.*, 2010; Dyrholm *et al.*, 2007). ICA converts multidimensional vectors into statistically independent components. Therefore this technique is broadly used in other blind source separation problems like image processing, text processing, communication and other blind source separation problems.

In this study ICA infomax algorithm is modified by making the contrast function to converge faster using its higher order cumulant. Simple fixed point and gradient descent based optimization is used for finding the un-mixing matrix W. Speed and accuracy of the proposed scheme are demonstrated using general simulated data and fMRI simulated data.

ICA (INFOMAX) ALGORITHM

Independent Component Analysis (ICA) is generally used for multivariate data analysis with applications in various fields such as medical image analysis, bioinformatics, pattern recognition, digital communications (Mohammad *et al.*, 2011), computational neuroscience and other areas where the blind source separation problem exists. ICA decomposes multivariate data into a linear sum of non-orthogonal basis vectors with coefficients being statistically independent (Dyrholm *et al.*, 2007). Since fMRI observed data consist of sources linearly mixed by the time courses of activity, noise and physiological activity. Thus the ICA model assumes that the observed fMRI data X is a linear mixture of the source vector S and the time courses matrix T :

$$X = TS \tag{1}$$

where, X is the observed data of dimension $n \times m$, T is a mixing matrix of dimension $n \times k$ and S consist of independent sources of dimension $k \times m$.

The goal of ICA is to find the un-mixing matrix A such that:

$$\hat{S} = Y = AX \tag{2}$$

where, $A = T^{-1}$ with the constraint that $A^T A = I$

In most of the ICA algorithms, the observed data is normally centered (zero mean) and whitened before implementing the ICA algorithm. Centering of the observed data is done by subtracting the mean from the observed data X :

$$X^{'} = X - E[X] \tag{3}$$

where, X = The observed data $X^{'}$ = The centered data

Whitening can be done by multiplying the observed data X with some whitening matrix U so that the correlation and covariance matrix of the sources becomes identity i.e., $E[YY^T] = I$. This can be done using the eigen-decomposition (Aapo *et al.*, 2001) and is given as under:

$$Z = UX^{'} \tag{4}$$

$$U = ED^{-\frac{1}{2}}E^T \text{ and } EDE^T = E\left[X^{'}X^{'T}\right]$$

E is the orthogonal matrix of eigenvectors of

$$E\left[X^{'}X^{'T}\right]$$

and D is the diagonal matrix of its eigenvalues.

Another preprocessing step for fMRI data is also required i.e. the dimension reduction due to its high dimensionality. In this case we have used Singular Value Decomposition (SVD) as a dimension reduction technique (Aapo *et al.*, 2001).

The data is now ready for ICA algorithm in which a contrast function is maximized/minimized so that maximum independence of extracted sources is achieved. There are different flavors of ICA which consists of kurtosis maximization, negentropy, information maximization (Infomax).

Now sources can be extracted sequentially i.e., one by one or symmetrically. Sequential approach is known as deflation approach and was first proposed by Delfosse and Loubaton (1995).

In the deflation based approach ICA algorithm looks for such an A which maximizes some contrast function which is basically the measure of non-Gaussianity. Different approaches are used in the literature (Aapo *et al.*, 2001) for measuring non-Gaussianity like kurtosis, nagentropy, information maximization etc. In this study we are considering only Infomax algorithm which is discussed as under.

Infomax is a method of ICA based on information theory which aims to find independent source signals by maximizing entropy. The general strategy of Infomax algorithm begins with Eq. (1), where, the extracted sources *Y* are obtained from signal mixtures *X* by optimizing a un-mixing matrix *A* such that the extracted signals are source signals if they are mutually independent.
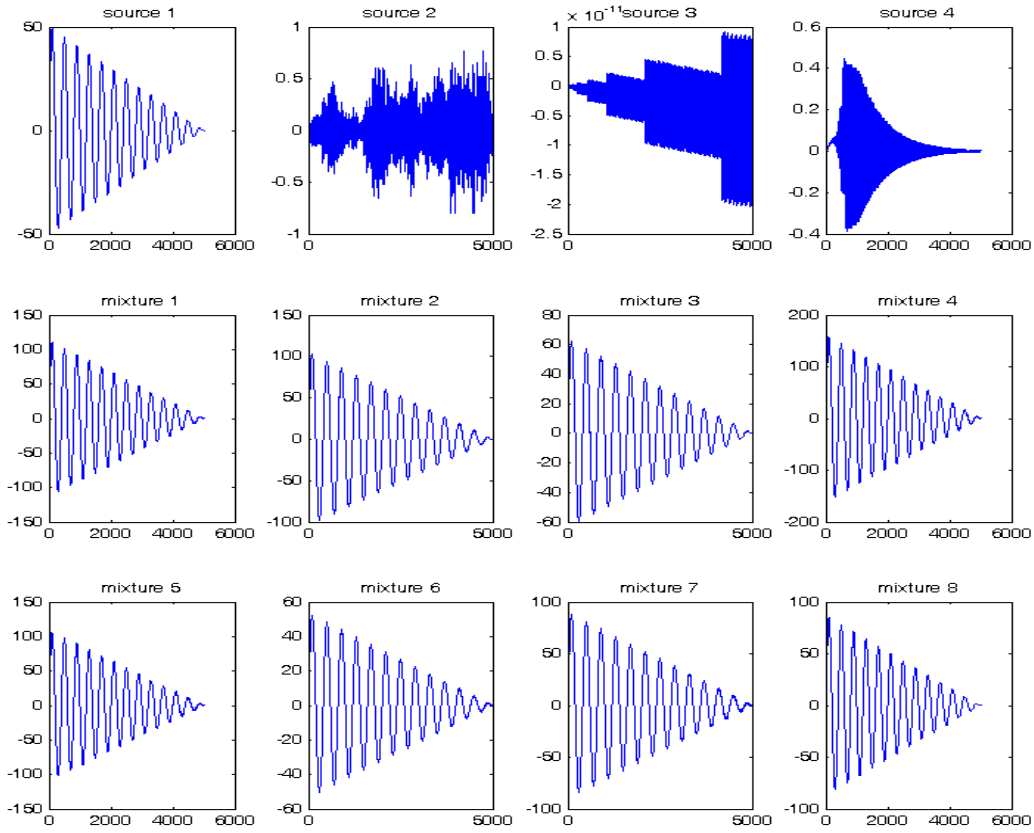
Fig. 1: Four simulated sources and eight mixtures

Entropy is related to independence in such a way that maximum entropy implies independent signals. Therefore, the objective of ICA is to find the un-mixing matrix A that maximizes the entropy in the extracted signals Y. Infomax (Aapo *et al.*, 2001) involves the maximization of the output entropy:

$$Z = g(Y) \text{ where, } Y = AX$$

and g($\cdot$) is a squashing function $g(Y) = \tanh(Y)$. It was shown that infomax contrast maximization is equivalent to the minimization of KL divergence between the distribution of $Y = AX$ (Bell and Sejnowski, 1995).

Different approximations have been proposed in literature for measuring the entropy of the signals for ICA framework. Here we will consider the approximation as narrated by (Dyrholm *et al.*, 2007):

$$h(Y) = E\left[\sum_{i=1}^{m} \ln p_s(y_i)\right] + \ln|A| \tag{5}$$

where,

$h(Y)$	=	The entropy of the source matrix
$p_s(y_i)$	=	The pdf of the ith source
$|A|$	=	The determinant of un-mixing matrix $|A|$

Now using the gradient algorithm A can be updated iteratively by maximizing the cost function $h(Y)$ such that:

$$A(n+1) = A(n) + \nabla(h) \tag{6}$$

where, $\nabla(h) = \dfrac{\partial h(Y)}{\partial A} = \dfrac{\partial h(AX)}{\partial A}$

$\nabla(h)$ can be calculated as:

$$\nabla(h) = A^{-T} - 2\tanh(AX)X^T \tag{7}$$

PROPOSED MODIFIED INFOMAX BASED HIGHER ORDER GRADIENT FUNCTION

By looking into gradient based optimization approach, it is clear that when the gradient is stable, then it must point in that direction of A which maximizes the cost function, in this case it is the entropy of the transformed sources. Also the gradient is basically the inverse transpose of A along with the $\tanh(Y)$ sources multiplied by the transpose of the observed data which again gives us A. Thus if we take some higher order cumulant of the gradient it will
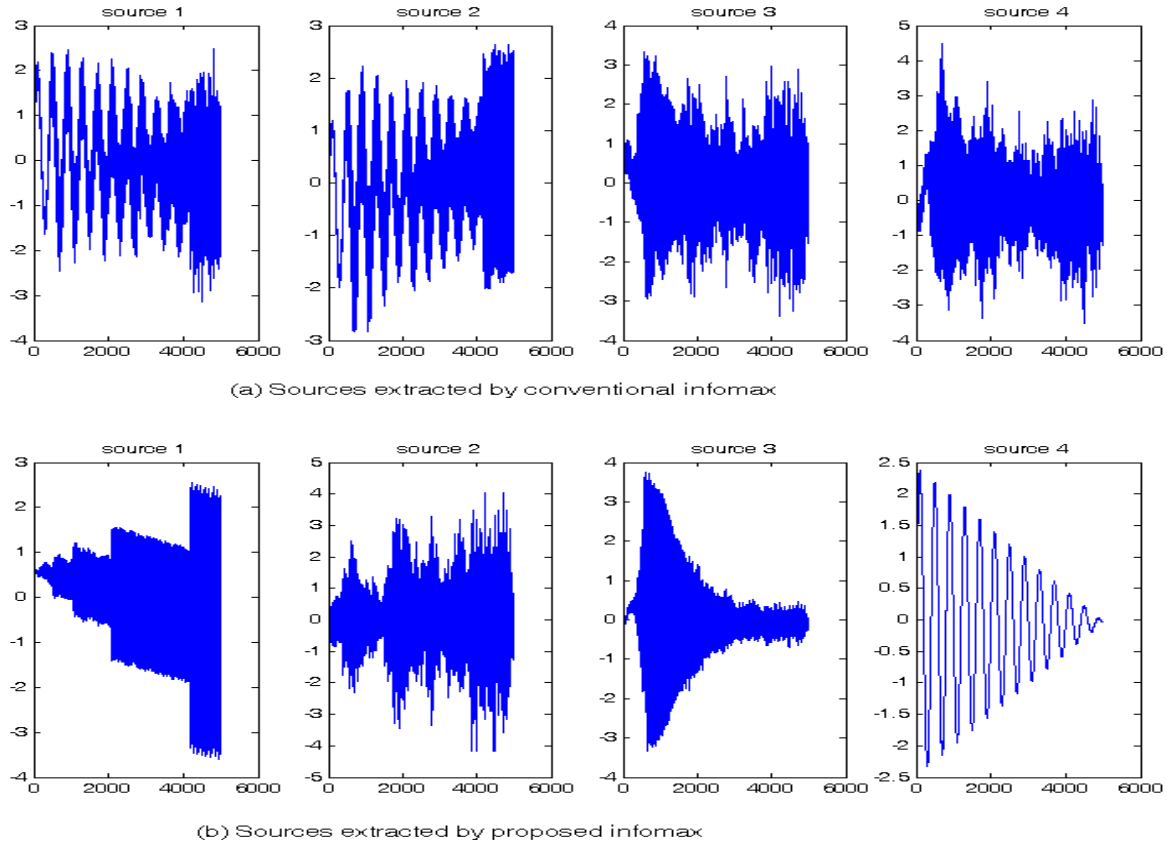
Fig. 2: Four sources extracted by the conventional infomax (b) Four sources extracted by the proposed modified infomax

update A with a faster rate. Also all vectors of A will not change their direction as A will be normalized in each iteration and a convergence is achieved fast (Aapo *et al.*, 2001) and ultimately independent sources will be extracted with greater accuracy in less time. Keeping these points in mind, equation (8) fulfills the criteria of convergence:

$$A(n+1) = A(n) + (\nabla(h))^n = A(n) + A^{-3T} - 2\tanh(AX)^3 X^{3T} \quad (8)$$

It should be kept in mind that the value of n should be odd so that sign of elements in the matrix X and A are not lost.

Pseudo code:

step1	=	Load data and put it equal to X
step2	=	Make the data zero mean and white
Step3	=	Initiate the un-mixing matrix A randomly
Step4	=	Find $\nabla(h)$ using Eq. (7)
Step5	=	Update weight matrix A using Eq. (8)
Step6	=	If the change in un-mixing matrix A is less than some minimum value Stop, else go to step 5.

SIMULATION RESULTS AND DISCUSSION

In this study a modified version of infomax is presented. To verify the validity of the proposed algorithm, it is applied to general simulated and fMRI simulated source signals. First row of Fig. 1 shows four simulated source signals. These signals are mixed by linearly combining them with random signals thus making eight mixed signals as shown in row 2 and row 3 of Fig. 1. The conventional infomax algorithm is applied to mixed signals for extracting the underlying source signals as shown in Fig. 2a. It is quite clear that there is no one to one correspondence of the extracted sources with the actual sources even visually. Figure 2b shows the results of the extracted sources using the proposed modified infomax algorithm which are quite similar to the original sources even visually.

To check the validity of the algorithm on simulated fMRI data, proposed and conventional algorithms are also applied to extract the hidden sources in fMRI simulated data. This data is downloaded from http://mlsp.umbc.edu/simulated_fmri_data.html. This data was basically developed by Correa *et al.* (2005) for their algorithm testing. Figure 3 shows simulated fMRI sources and corresponding time courses while Fig. 4 shows extracted sources and time courses of simulated fMRI data using conventional infomax. The results are
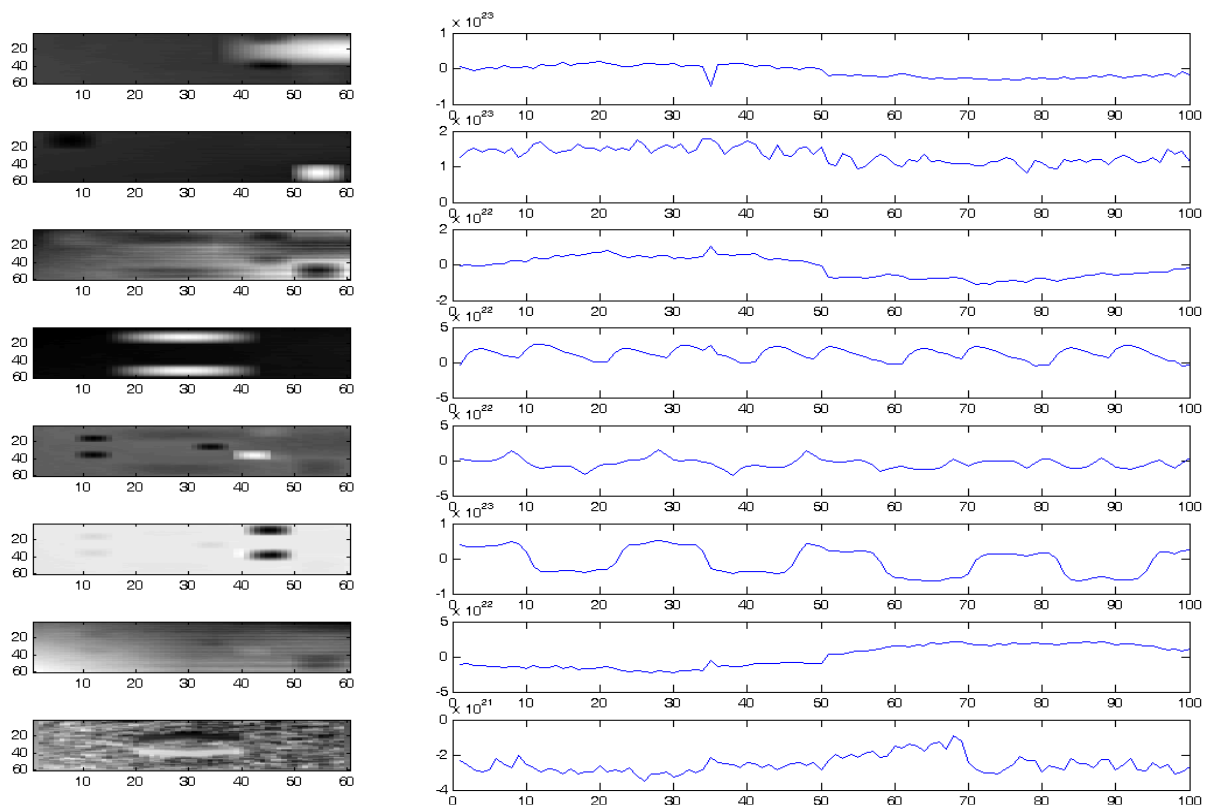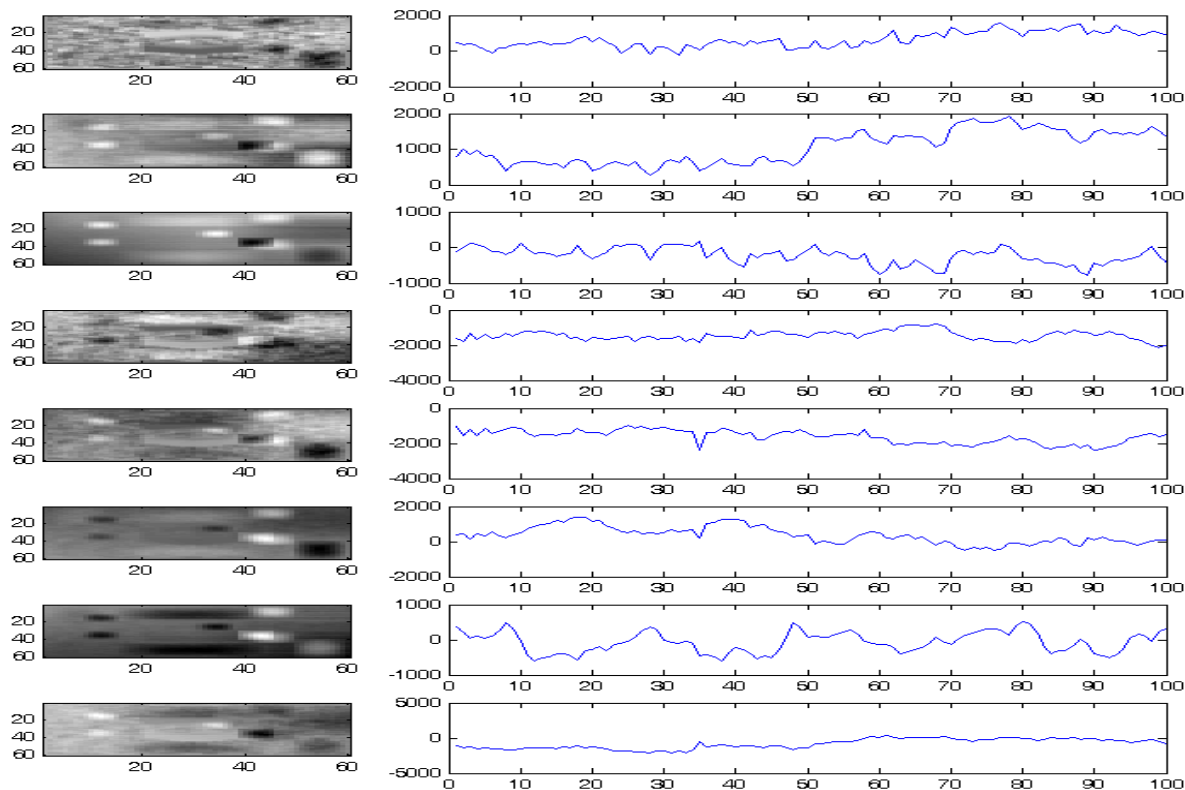
Fig. 3: Simulated fMRI sources and time courses

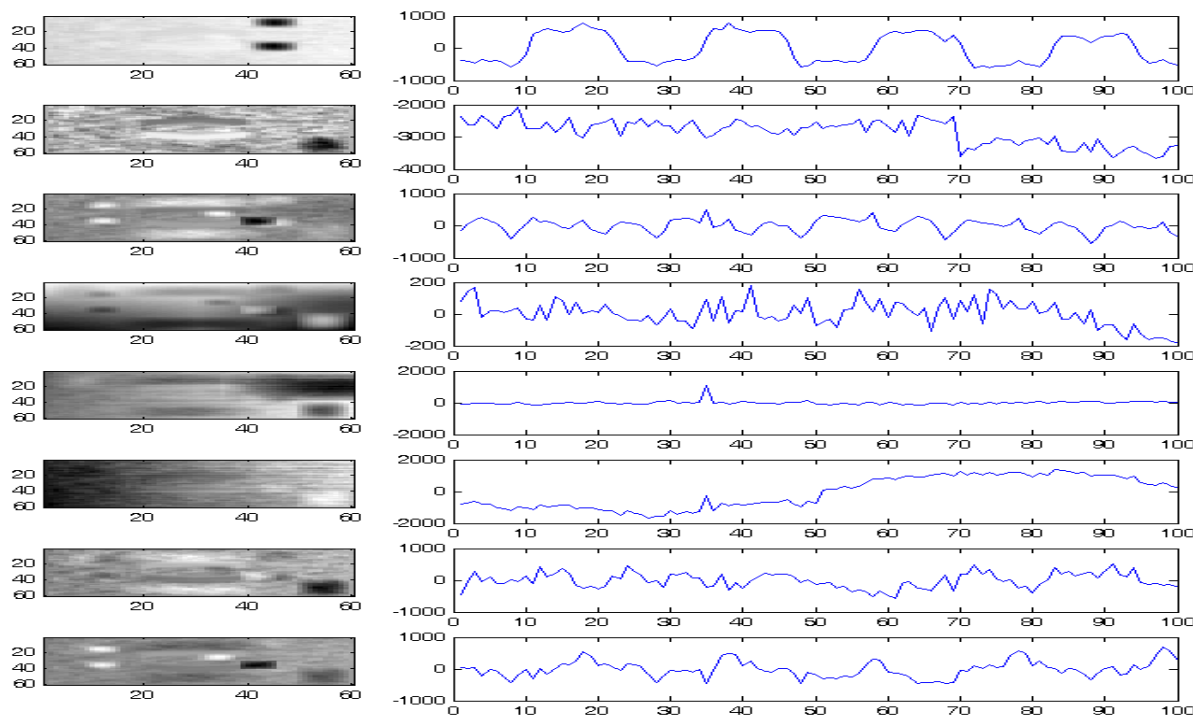Fig. 4: fMRI extracted sources and time courses by conventional informax

Fig. 5: fMRI extracted sources and time courses by the proposed infomax

Table 1: Correlation results of simulated fMRI extracted 8 sources (S1-S4) and time courses (TC1-TC8) using infomax and proposed infomax

Sources/Algo:	Time (sec)	S1	S2	S3	S4	S5	S6	S7	S8	Avg:
InfoMax	0.77	0.50	0.45	0.61	0.64	0.42	0.45	0.52	0.45	50
Proposed	0.77	0.60	0.60	0.96	0.76	0.59	0.68	0.87	0.63	71
Sources/Algo:	Tim (sec)	TC1	TC2	TC3	TC4	TC5	TC6	TC7	TC9	Avg:
InfoMax	0.77	0.40	0.34	0.89	0.71	0.24	0.28	0.78	0.30	49
Proposed	0.77	0.97	0.64	0.98	0.87	0.36	0.62	0.78	0.90	76

Table 2: Correlation results of extracted sources and actual simulated sources using infomax and proposed infomax

Algorithm	Time (sec)	Source 1	Source 2	Source 3	Source 4	Avg:
Infomax	0.88	0.64	0.65	0.62	0.66	64
Proposed modified infomax	0.52	0.98	100	0.99	0.99	99

not good even visually. Figure 5 shows extracted sources and time courses using the proposed infomax algorithm. By looking to the extracted sources and time courses it is quite clear that there exist one to one correspondence of sources and time courses with the actual sources of Fig. 3. Table 1 shows the execution time and correlation results of the conventional and proposed infomax algorithm for general signals of Fig. 1 and 2. Table 2 shows the execution time and correlation results of the conventional and proposed algorithm in case of fMRI data for the conventional and proposed algorithm.

CONCLUSION

In this study infomax algorithm is studied and applied on general signals and fMRI simulated signals. A higher order Infomax contrast function is also presented with simple fixed point iteration algorithm. Since the contrast function is higher order, therefore, it converges rapidly and extracted sources having a good correlation with actual sources. The proposed algorithm is tested on general signals and fMRI simulated signals and its application can be extended to other BSS problems.

ACKNOWLEDGMENT

REFERENCES

Aapo, H., J. Karhunen and E. Oja, 2001. Independent Component Analysis. Wiley, New York.

Amir, A.K., I.M. Qureshi and J.A. Shah, 2012. Un-mixing Functional Magnetic Resonance Imaging (FMRI) data using matrix factorization. Int. J. Imag. Syst. Technol., 22(4): 195-199.

Backfrieder, W., 1996. Quantification of intensity variations in functional MR images using rotated principal component. Phys. Med. Biol., 4: 1425-1438.

Bandettini, P.A., A. Jesmanowicz, E.C. Wong and J.S. Hyde, 1993. Processing strategies for time-course data sets in functional MRI of the human brain. Magn. Reson. Med., 30: 161-173.

Bell, A. and T. Sejnowski, 1995. An information maximization approach to blind separation and blind de-convolution. Neural Comput., 7: 1129-1159.

Correa, N., Y.O. Li, T. Adali and V. Calhoun, 2005. Comparison of blind source separation algorithms for fMRI using a new Matlab toolbox: GIFT. Proceeding of IEEE International Conference on Acoust., Speech, Signal Processing (ICASSP). Philadelphia, 5: 401-403.

Delfosse, N. and P. Loubaton, 1995. Adaptive blind separation of independent sources: A deflation approach. Signal Proces., 45: 59-83.

Dyrholm, M., S. Makeig and L.K. Hansen, 2007. Convolutive ICA for spatio-temporal analysis of EEG. Neural Comput., 19: 934-955.

Friman, O., J. Cadefamn, P. Lundberg, M. Borga and H. Knutsson, 2001. Detection of neural activity in functional MRI using canonical correlation analysis. Magnet. Reson. Med., 45(2): 323-330.

Friston, K.J., 1996. Statistical Parametric Mapping and Other Analyses of Functional Imaging Data. In: Toga, A.W. and J.C. Mazziotta (Eds.), Brain Mapping: The Methods. Academic Press, San Diego, pp: 363-396.

McKeown, M.J. and T.J. Sejnowski, 1998. Independent component analysis of fmri data: Examining the assumptions. Hum. Brain Mapp., 6: 368-372.

McKeown, M.J., L.K. Hansen and T.J. Sejnowski, 2003. Independent component analysis of functional MRI: What is signal and what is noise? Curr. Opin. Neurbiol., 13: 620-629.

Mitra, P.P., S. Ogawa, X. Hu and K. Ugurbil, 1997. The nature of spatio-temporal changes in cerebral hemodynamics as manifested in functional magnetic resonance imaging. Magnet. Reson. Med., 37: 511-518.

Mohammad, E., J.A. Jahanshahi and S.A. Ghorashi, 2011. Joint ICA-based blind detection and parameter assessment in DS-CDMA systems. Res. J. Appl. Sci. Eng. Technol., 3(9): 857-865.

Ogawa, S., T.M. Lee, A.R. Kay and D.W. Tank, 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proceedings of the National Academy of Sciences of the United States of America, 87: 9868-9872.

Qiu-Hua, L., J. Liu, Y.R. Zheng, H. Liang, V.D. Calhoun, 2010. Semi-blind spatial ICA of fMRI using spatial constraints. Hum. Brain Mapp., 36: 1076-1088.