EMBEDDED IMPLEMENTATION OF EEG ANALYSIS USING INDEPENDENT COMPONENT APPROACH

[1]FASEELA.K.P & [2]SUPRIYA.P

[1]Dept of ECE, AWH Engineering College, Calicut, India
[2]Dept of EEE, Amrita Vishwa Vidyapeetham, Coimbatore, India
E-mail: faseelakp@gmail.com, p_supri@yahoo.com

Abstract—Brain signals are important in diagnosing various disorders and abnormalities in the human body. These signals are recorded by scalp electrodes and are called as EEG signals. EEG signals are a mixture of signals from different brain regions which contain artefacts along with original information. These contaminated mixtures are analysed such that diagnosis of various diseases is possible. One of the effective methods available is Independent Component Analysis (ICA) for removing artefacts and for separation and analysis of the desired sources from within the EEGs. This paper focuses on the analysis of EEG signals using ICA approach. Two ICA algorithms- Pearson ICA and JADE ICA are analysed in this paper. Comparison of these ICA algorithms in removing artefacts from EEG has been carried out by simulation using MATLAB. Then the Pearson ICA algorithm simulation is done using Visual C#. The algorithm has been implemented in an Embedded Development Kit (EDK) using .NET Micro Framework and the results are presented.

Keywords-*Blind source separation, Electroencephalogram, Independent Component Analysis (ICA), Pearson system, Joint Approximate Diagonalization of Eigen Matrices, Embedded Development Kit.*

I. INTRODUCTION

Blind Source Separation (BSS) is a major problem that comes into existence in various fields such as signal and image processing and telecommunications.BSS problem involves decomposing a mixture of non Gaussian signals into its independent component signals with no prior information about the signals that constitutes the mixture. Independent Component Analysis (ICA) is a method used for achieving blind separation of sources. It is used to extract more meaningful individual signals from mixtures of signals.

The analysis of brain signals for diagnosing diseases is an important application in biomedical field. Electroencephalography (EEG) refers to the measurement of electrical activity of the brain which is used to give important information about the functioning of the brain. EEG is measured by connecting electrodes on the scalp surface. EEG signals are subject to noise and artefacts like eye movements and blinks which make difficulties in diagnosing various brain diseases like epilepsy, brain death, and coma and so on. ICA provides an efficient method for removing these artefacts which have high amplitude in comparison with the amplitudes of the brain signals.

ICA is a statistical decomposition technique which is being used as a method to find hidden and underlying features in any data. Various statistical decomposition techniques have been found in literature like Projection Pursuit [1], Blind Deconvolution [2], Principal Component Analysis (PCA) [3] and Linear Discriminant Analysis (LDA) [4].

Different types of ICA algorithms are found in literature [5] for EEG analysis. Papers [6, 7] use fast ICA for the analysis of EEG signals. But fast ICA cannot separate the signals if they are close to Gaussian distribution. This paper focuses on Pearson ICA approach which has the advantage of extracting the signals that are close to Gaussian distribution also. The performance of the Pearson ICA is compared with another ICA called JADE ICA which is based on the diagonalization of cumulant matrices. It is seen that if the signals are very near to Gaussian distribution, Pearson ICA separates them accurately than any other ICA algorithms.

Comparison of Pearson ICA with JADE (Joint Approximate Diagonalization of Eigen-matrices) ICA has been executed in this work. Computation time and Signal to Interference Ratio (SIR) of the two algorithms are compared. For EEG analysis, Pearson ICA gives better performance compared with JADE ICA. The algorithms for Pearson ICA and JADE ICA have been developed for the analysis of EEG signals and Matlab simulation is carried out for these two algorithms. Pearson ICA is simulated in visual C# and the waveforms are analysed.

For Hardware implementation of the Pearson ICA algorithm, Embedded Development Kit (EDK 2.0) is used. This kit uses .NET Micro Framework as the Development Platform with visual Studio 2005 as the IDE. EEG data is processed by the kit and the results are received by the PC. The obtained results are plotted in visual C# through a third party tool called ZedGraph.

This paper is organized as follows. Section II gives a brief discussion on ICA. Section III describes Pearson ICA and JADE ICA algorithms. Performance of these algorithms for the analysis of EEG signals is presented in section IV. It also shows the simulation results and hardware implementation details along with its results. The last section brings out the future scope and conclusion of the work.

II. BASICS OF ICA

ICA approach consists of recovering unknown independent signals from the linear combinations of their mixtures, with no prior information about the sources or the mixing mechanism. The ICA model assumes that there are N independent source signals given by

$$s = \{s_1(t), s_2(t), \ldots\ldots, s_N(t)\}$$

(1)

These signals are mixed with a random mixing matrix A. There are N different recorded mixtures which are independent and linear mixtures of original source signals. These mixtures are represented as

$$x = \{x_1(t), x_2(t), \ldots\ldots, x_N(t)\}$$

(2)

If there are two sources and two mixtures, then ICA problem can be represented as

$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t)$$

(2)

$$x_2(t) = a_{21}s_1(t) + a_{22}s_2(t)$$

(3)

It can be represented in matrix form as

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix}$$

(4)

So for N sources and N mixtures, ICA problem can be represented as

$$\begin{pmatrix} x_1(t) \\ \vdots \\ x_N(t) \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{pmatrix} \begin{pmatrix} s_1(t) \\ \vdots \\ s_N(t) \end{pmatrix} \quad (5)$$

The general representation is given as

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s} \quad (6)$$

The aim of ICA is to extract the independent sources **s** and mixing matrix **A** from its mixture **x**. For this a demixing matrix **W** is used which is able to extract independent sources from their mixtures. It is given as

$$\hat{\boldsymbol{s}} = \boldsymbol{W}\boldsymbol{x}$$

(7)

The obtained $\hat{\boldsymbol{s}}$ is the scaled version of **s**.

The ICA architecture is given in Fig.1

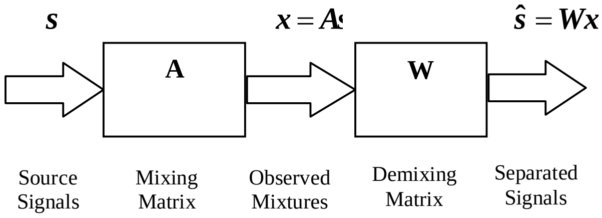

Figure 1. ICA architecture

Before applying ICA algorithm, it is necessary to have some pre-processing techniques like centering and whitening to make ICA estimation simpler. Centering brings the samples to the zero mean and whitening reduces the correlation among the components. Centered data is given by

$$\mathbf{x} = \mathbf{x} - E\{\mathbf{x}\}$$

(9)

If **D** and **E** are the Eigen value and Eigen matrices of the covariance matrix of **x**, whitening is given by the Eigen Value Decomposition of the covariance matrix such that **x** is transformed to a $\tilde{\mathbf{x}}$

$$\tilde{\mathbf{x}} = \mathbf{E}\mathbf{D}^{-1/2}\mathbf{E}^{T}\mathbf{A}\mathbf{s} = \tilde{\mathbf{A}}\mathbf{s}$$

(10)

III. METHODOLOGY

The proposed Pearson ICA and JADE ICA algorithms are described in this section.

A. Pearson ICA

Pearson system consists of a number of distributions of various types. It is used to model a broad class of source distributions. Pearson ICA uses Pearson system and is based on Maximum likelihood approach [8]. Pearson ICA combines two techniques called Maximum Likelihood Approach and fixed contrast functions for Independent Component Analysis. Pearson system is used in the maximum likelihood approach to model source distributions. Fixed contrast functions are employed to improve the speed and stability of the Pearson system. Pearson system is defined by the differential equation given in [10] as:

$$f'(y) = \frac{(y-a)f(y)}{b_0 + b_1 y + b_2 y^2} \quad (11)$$

where a, b_0, b_1 and b_2 are the parameters of the distribution.
The score function of the source distribution is used as a contrast in the maximum likelihood approach. The score function is given as:

$$\varphi(y) = -\frac{f'(y)}{f(y)} = -\frac{(y-a)}{b_0 + b_1 y + b_2 y^2} \qquad (12)$$

The parameters of the score function namely a, b_0, b_1, and b_2 are estimated by the method of moments. The third order and fourth order moments are defined in (13) and (14) as:

$$\alpha_3 = \sum_{i=1}^{n} (y_i - \bar{y})^3 / (n\sigma^3) \qquad (13)$$

$$\alpha_4 = \sum_{i=1}^{n} (y_i - \bar{y})^2 / (n\sigma^4) \qquad (14)$$

where $\sigma = \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2 / n}$

Using (13) and (14), parameters of the Pearson system can be estimated using the following equations.

$$b_1 = a = -\frac{\alpha_3(\alpha_4 + 3)}{C} \qquad (15)$$

$$b_0 = -\frac{(4\alpha_4 - 3\alpha_3^2)}{C} \qquad (16)$$

$$b_2 = -\frac{(2\alpha_4 - 3\alpha_3^2 - 6)}{C} \qquad (17)$$

where $C = 10\alpha_4 - 12\alpha_3^2 - 18$

The score function is optimized by using natural gradient [9] or relative gradient algorithm with a weight updating formula as given in (18)

$$W_{k+1} = W_k + \eta\left(I - \varphi(y) y^T\right) W_k \qquad (18)$$

where η is the learning rate
Or a fixed point weight updating formula [10] as given in (19)

$$W_{k+1} = W_k + D\left(E\{\varphi(y)y^T\} - diag\left(E\{\varphi(y_i)y_i\}\right)\right)W_k \qquad (19)$$

where

$D = diag\left(1/\left(E\{\varphi(y_i)y_i\} - E\{\varphi'(y_i)\}\right)\right)$ is employed.

The details of the algorithm are given below for completeness:
1. Calculate the third and fourth sample moments α_3 and α_4 for the current data.
2. Based on the value of α_4, select the appropriate contrast functions.
   a. If α_4 is less than 2.6, tanh contrast is used.
   b. If α_4 is greater than or equal to $\alpha_3^2 + 4$, tanh contrast is used.
   c. If α_4 is in between these two values, Pearson system is used.
3. Estimate the parameters of the distribution by moment method.
4. Calculate scores $\varphi(y_i)$ for the contrast functions
5. Calculate the weight updating matrix W_{k+1} using equation (18) or (19)
6. Repeat until convergence of the weight matrix occurs or when maximum iteration reaches.

B. JADE ICA

JADE ICA approach is based on the joint diagonalization of cumulant matrices [11].This algorithm uses second and fourth order cumulants to separate the independent sources. The second order cumulants is used to make the data as decorrelated as possible. This produces a whitening matrix P and the obtained sources are called whitened sources. There is no parameter tuning or weight upadation to obtain better performance. JADE ICA performs well when the number of sources to be extracted is less. When the number of sources increases, its performance become poor compared to Pearson ICA. The details of the JADE ICA algorithm are given below:
1. Calculate the Whitening matrix **P**.
2. For the whitened sources, calculate the cumulant matrices and find the maximum set among these matrices.
3. Find orthogonal matrix **R** which minimizes the contrast function.
4. Estimate the demixing matrix as $\boldsymbol{A} = \boldsymbol{R}^{-1}\boldsymbol{P}$.
5. Estimate the independent sources $\boldsymbol{s} = \boldsymbol{A}^{-1}\boldsymbol{x}$ where **x** is the mixed signal.

IV. RESULTS

This paper concentrates on the analysis of EEG signals rather than its measurement. The various signals recorded from different electrodes (FP1, FP2, F7 etc) are taken and the two ICA algorithms- Pearson ICA and JADE ICA are applied on these signals to get the exact signals. While comparing the performances of Pearson ICA and JADE ICA for EEG signals, it is seen that Pearson ICA separates more accurately than JADE ICA.

Comparison of Pearson and JADE ICA with different number of random signals and also with different number of samples is shown in Table. I. Both algorithms differ in performance based on the number of sources and also on the signal length. When the numbers of sources are less in number, both algorithms separate the signals well.

The Signal to Interference ratio (SIR) and computation time for both Pearson ICA and JADE ICA has been calculated and the comparison with different number of EEG sources and signal length is shown in Table II. SIR gives the performance value in dB. It gives the separation capability of the algorithm. A high value of SIR means good separation. SIR is given by

$$SIR = -10\log\left(\frac{\|y_i - s_i\|^2}{\|s_i\|^2}\right) \qquad (20)$$

Both Pearson ICA and JADE ICA take less time when the number of sources is less in number. But as number of sources is very large in number, JADE takes more time to converge and have poor separation of sources. As signal length increases, SIR also increases which represents better separation performance of the algorithms. But sometimes, as signal length increases, SIR also decreases which is clearly indicated in Table II.

TABLE I
COMPARISON IN TERMS OF SIR AND CONVERSION SPEED

No. Of Signals	Signal Length	SIR(dB)		Conversion speed(seconds)	
		Pearson ICA	JADE ICA	Pearson ICA	JADE ICA
4	200	15	16	0.06	0.06
4	5000	33	29	0.10	0.063
10	200	10	<5	0.32	0.62
10	5000	25	26	0.27	0.16
20	5000	21	21	0.53	1.08
40	2000	14	<5	3.5	6347.2
40	5000	17	17	2.4	120
75	5000	15	<5	33	>6000
100	10000	19	<5	55	>6000
100	20000	22	<5	64	>6000

TABLE II
COMPARISON IN TERMS OF SIR AND CONVERSION SPEED FOR EEG DATA

No. Of Signals	Signal Length	SIR(dB)		Conversion speed(seconds)	
		Pearson ICA	JADE ICA	Pearson ICA	JADE ICA
2	100	32.41	17.09	0.00942	0.05953
2	200	38.32	25.16	0.00986	0.05235
2	500	37.65	13.63	0.00983	0.05131
2	1000	29.33	17.28	0.01045	0.05574
2	2000	27.81	17.43	0.01129	0.05605
2	5000	27.72	16.70	0.01411	0.05692
3	100	21.80	9.37	0.05274	0.06254
3	1000	24.34	9.94	0.05525	0.06056
3	10000	25.72	12.96	0.05692	0.06562
4	10000	12.12	7.96	0.09822	0.06080
4	20000	12.36	8.12	0.09505	0.08456

The implementation of the algorithms has been done in (i) MATLAB (ii) Visual C# (iii) EDK. Simulation outputs show the validity of the algorithms for EEG analysis. Simulation has been done with different signal length such as 100, 200, 1000 and 5000 and so on.

(i) MATLAB Simulation

Both Pearson ICA and JADE ICA have been developed in MATLAB and results are analysed. Fig. 2 and Fig.3 represent the separation output for two sources for a signal length of 200 and shows 100 samples for clear visibility of samples. From Fig.2 and Fig.3, it is clear that Pearson output tracks the input source more accurately than JADE ICA.A detailed analysis using MATLAB is mentioned in [13].

(ii)Simulation in VISUAL C#

The Pearson ICA has been developed in Visual C# which uses Microsoft Visual Studio 2005 as the IDE (Integrated Development Environment), which is a successor of Visual Studio 6. The input data is prepared on an excel sheet. This data has been read from excel for loading to the program. Then the entire processing is done and the output data is available and is plotted using a third party tool called Zed Graph. Fig. 4 shows simulation results for 1000 samples and zoomed to plot for 50 samples.

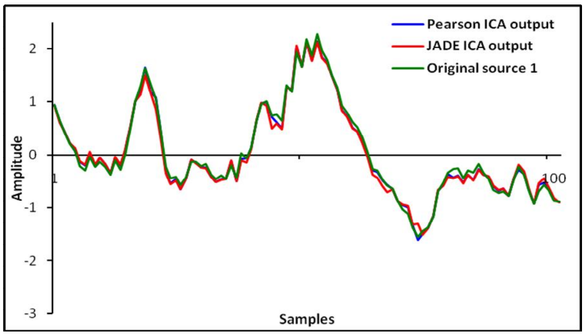

Figure 2. Separation output for source 1 when number of sources=2

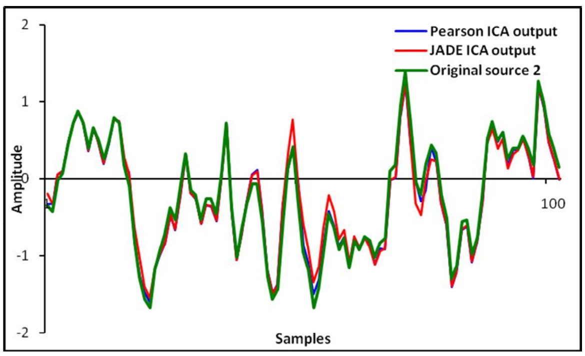

Figure 3. Separation output for source 2 when number of sources=2

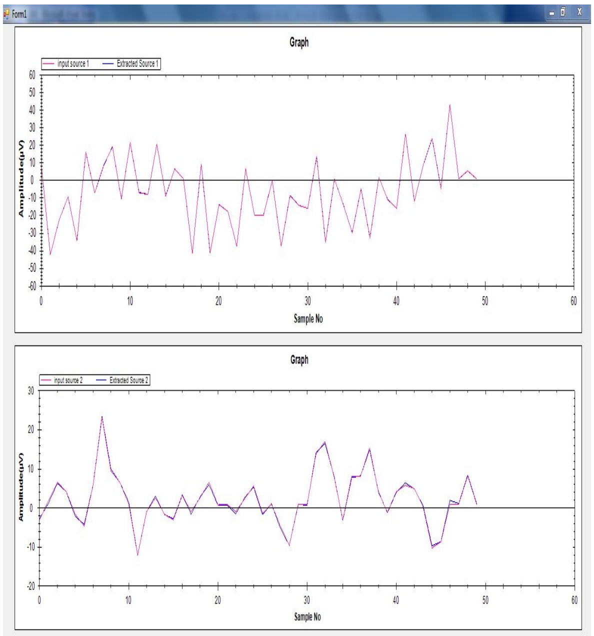

Figure 4. Simulation Results for 1000 samples

(iii). Hardware implementation in EDK

The entire processing of the Pearson ICA has been carried out on the EDK kit. The kit works with the available input data. It processes this data. All the resulting samples are passed to the serial port of the PC. The EDK has two COM ports namely COM1, which is used to download and debug the application and COM2 which is used for standard RS-232 communication [12]. Once downloading of the program is completed, it is required to add a start bit, stop bit, and byte array size with the string that is going to be sent to the COM port. [12].

The entire processing of the kit is takes place with the RS 232 port and once deployment to the kit succeeds, the RS-422, RS-423, RS-449 and RS-485 port is used to send the data to the serial port of the PC.

SJJ_COMM Lite or the serial terminal software can be used to display the dataset that has reached to the serial port of the PC. SJJ_COMM was created since Windows Vista and Windows 7 do not support the HyperTerminal. SJJ_COMM is employed to capture the generated debugged output. SJJ_COMM Lite is set for 115200. SJJ_COMM Lite or the serial terminal software can be used to display the dataset that has reached the serial port of PC. The implementation of the algorithm in the hardware is portrayed in Fig. 5.

Matlab simulation outputs for 50 samples are shown in Fig.6 for both the sources which show the performance efficiency of the algorithm. This shows that both give the same results which indicate the good separation characteristics of the algorithm. Small variations in the output are due to less number of samples.

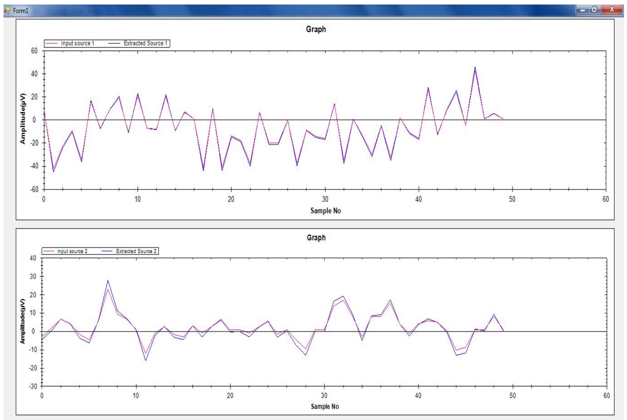

Figure 5. Separation output for 50 samples using EDK

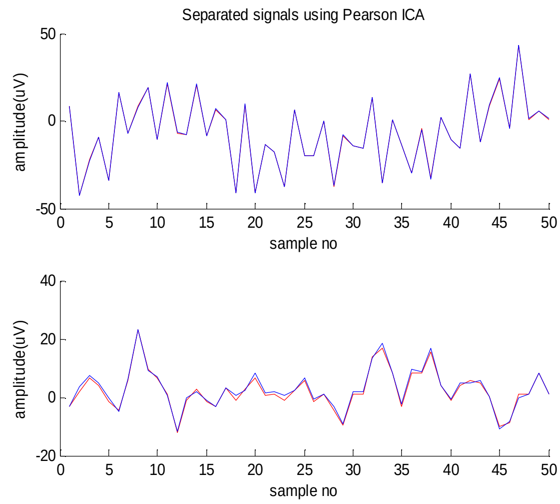

Figure 6. MATLAB Simulation for 50 samples

For Hardware implementation, two sources are considered. So for each source, one port pin is used for indication of the input. Two LEDs connected to the port pins check the threshold value of the input signals. If a specific signal has five continuous zeros, the LED turns ON indicating abnormality in signal detection.

Fig.7 represents the hardware implementation set up used in the paper. Initially both LEDs are in OFF position.

After successful BUILD of the program, deployment of the program is done which will download the application to the iPac-9302. When the downloading of the application is completed, the current running application is stopped and the new application is burned to flash memory by erasing the current application flash memory area. Next, unplug the power supply from the board, unplug the null modem cable from the debug serial port, connect it to the COM2 port and wait for one second and then plug the power back in.

Whenever EEG values are coming as five continuous zeros for the input source, the corresponding LED is turned ON.Fig.7 represents the initial condition before applying the input signals. When the signal detection is accurate, the two LEDs are turned OFF. Fig.8 indicates that first signal detection has some error and needs to be verified.

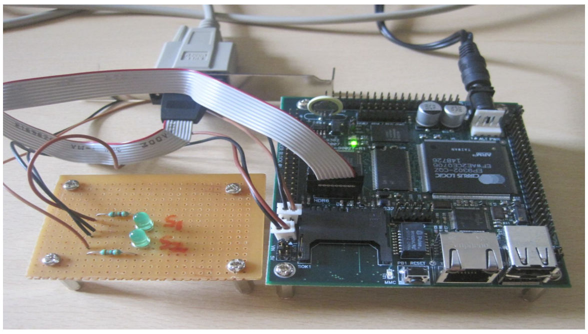

Figure 7. Initial LED indication for sources

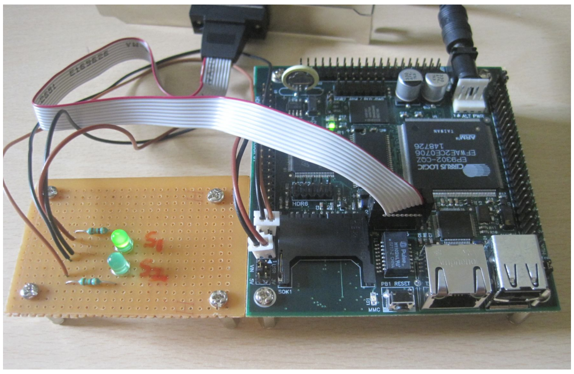

Figure 8. LED indication for sources when source 1 is providing incorrect measurement

V. CONCLUSION

In this paper, analysis of independent components of EEG using Pearson ICA and JADE ICA approach has been described. One of the assumptions of ICA is that it works well with non Gaussian signals only. Pearson ICA can however separate signals that are close to Gaussian distribution also. EEG signals are analysed using these algorithms which are developed in MATLAB and it found that the algorithms have good separation quality. It is observed that when EEG signals are analyzed Pearson ICA provides better results compared with JADE ICA. It is found that if the numbers of source signals are less, JADE and Pearson ICA algorithms are similar in performance with respect to separation. But if the sources are more in number, Pearson ICA provides more accurate results than JADE ICA.

The algorithm for Pearson ICA has been developed in Visual C# and the results are plotted using ZedGraph.

The EDK board (Version 2.0) does not support file operations (such as reading from excel, text), the input data needs to be hardcoded into the EDK. However with advanced EDK boards, the data can be read using excel sheet. The analysis of EEG signals has been done in the present work. A measurement module can be developed and interfaced to this kit using its ADC module.

ACKNOWLEDGMENT

The authors thank Dr.T.N Padmanabhan Nambiar, Chairperson, Department of Electrical and Electronics Engineering, Amrita Vishwa Vidyapeetham, Coimbatore for his constant inspiration throughout the work. The authors also thank Microsoft for supporting the experimental set up available in the embedded laboratory of the University.

REFERENCES

[1] P.Diaconis, and J.Salzman, " Projection pursuit for discrete data,"1987

[2] Alexander M. Bronstein, Michael M. Bronstein, and Michael Zibulevsky, "Relative optimization for blind deconvolution," IEEE Trans. Signal Processing, vol. 44, May 2003

[3] Martin Kermit, "Independent Component Analysis: Theory and Applications," April 2000.

[4] K. Fukunaga, "Introduction to Statistical Pattern Recognition", academic Press, Boston, 2nd edition, 1990.

[5] V, Krishnaveni, S.Jayaraman and P.M.Manoj Kumar, "Comparison of Independent Component analysis for removal of artefacts from EEG recordings," Measurement Science Review, vol 5, 2005

[6] Ricardo Vigario, Jaakko Sarela, Veikko Jousmaki, Matti Hamalainen, and Erkki, "Independent Component Approach to the analysis of EEG and MEG recordings,"IEEE Trans. Biomedical Engineering,vol. 47,May2000.

[7] M. Ungureanu, C. Bigan, R. Strungaru, V. Lazarescu, " Independent Component Analysis Applied in Biomedical Signal Processing," Measurement Science Review, vol 4, 2004

[8] Karvanen, J., Eriksson, J., and Koivunen, V, "Pearson System Based Method for Blind Separation," Proc. Second International Workshop on Independent Component Analysis and Blind Signal Separation, Helsinki 2000, pp. 585-590

[9] Jean-Francois Cardoso, "Blind Signal Separation: Statistical Principles," Processing of IEEE, Vol.86, No.10, Oct., 1998: pp. 2009-2025.

[10] A.Hyvarinen, "The fixed algorithm and maximum Likelihood estimation for independent component analysis," Neural Processing Letters, vol.10,issue 1, August 1999, pp.1-5.

[11] Jean-Francois Cardoso, "High-order contrasts for independent component analysis,"Neural Computation, vol. 11, no 1, Jan. 1999, pp. 157—192. Sean Liming and John.R.Malin, "Embedded Development Kit for .NET Microframework," 4 th ed,2008, pp 22-130.

[12] K.P.Faseela, and P.Supriya, "EEG Signal analysis using Blind Sign Processing Techniques", Proc. First International conference on Modern Trends in Instrumentation and Control,PSGCT,Sep 2-3,2011,pp 320-326