

Eliminating Electroencephalogram Artefacts Using Independent Component Analysis[#]

Akram Rashid^{*1}, I.M. Qureshi², Sana Mukhtar³

Abstract—The elimination of artefacts from Electroencephalogram(EEG) has an important role many signal and image processing applications. The artefacts are the noises that appears during the acquisition of signals from the patient body. With the presence of these artefacts it become difficult for doctors and technicians to analyse the Electroencephalogram signals efficiently. The aim of this research work is to remove these artefacts using Independent Components Analysis(ICA). The scalp EEG is intensively used as an important clinical tool for diagnosis and treatment of diseases. The probabilistic modified ICA algorithm is used to separate EEG signals from artifacts for efficiently brain tumor detection. This research work aims to detect epileptic activity for an electroencephalogram having sixteen-channels. The research consists of three important stages First one is data collection from patients, second is feature extraction and third one is EEG signal analysis. In feature extraction the stress is to detect epileptic form of activity from the patient collected signals. In signal analysis stage the stress is to get information about the type of the brain tumor.

Keywords: Fuzzy Elimination, Electroencephalogram, Image processing, Diagnose, Tumor.

1. Introduction

The EEG signals carries significant information. The EEG signals obtained from the electrodes placed on the scalp are composed of electrical potentials. Brain cells(neurons) and artifacts are the main source of these potentials. The potentials obtained from different independent brain cells(neurons) carry great deal of information for researchers and physicians. Placing the EEG electrodes inside the human head comparatively give more useful information.

However this technique is not recommended as it may cause severe pain and serious damage to the neurons. So the recommended is to always get the EEG signal of interest from the scalp. Signal processing techniques such as data detection, data classification, noise reduction, data compression and feature extraction are being used to tackle serious EEG issues.

The analysis and study of the signals is of great importance not only for medical diagnosis and medical treatment but also for research purposes. Figure 1 below shows sample of EEG data.

The cerebral cortex nerve cells of brain produce electrical activity. This electrical activity is widely used for clinical brain tumor diagnoses. It is also widely employed for the diagnoses of epileptic discharges appears in electroencephalogram. This electrical activity pattern can also be used to diagnose many other factors which may can affect other brain abnormalities. These abnormalities include brain hemorrhage, head injuries or brain tumors. An efficient early brain tumor detection technique have been proposed in this research. EEG artefacts have been removed using Independent Component Analysis Technique. The basic EEG system include a filter circuit to remove the noises.

These noises are power line noises, muscle noises, lung noises, component noises.

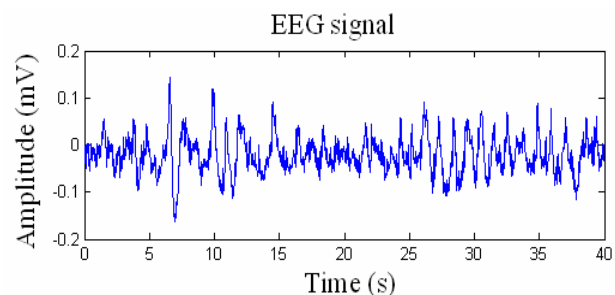


Figure 1 Real EEG signal mixed with EOG, ECG and Line interference

This research work has got its inspiration from previous research works available as removing noises from EEG and ECG signals for efficient detection of tumor. To eliminate the baseline noise from ECG signals number of techniques have been employed. ECG baseline artifacts are eliminated using Independent component analysis(ICA) technique[12].

For elimination of ocular artifacts many regression based techniques are devised. Simple time-domain regression is the base of these regression techniques[23]. The EEG and EOG are generally uncorrelated in regression techniques. However practically these regression methods are mostly correlated.

In these regression techniques the initial transfer function is established which actually necessitate the calibration trials between each ECG and EEG channels[23].

A general subtraction method has been recommended by Haas[24] This method is named as ARMAX. In this method measured EEG signal is expressed as a linear combination of a background Ocular artifacts and EEG signals. The estimation of background EEG is easily obtained by ARMAX technique.

1.1. Artefacts Present in EEG

In practice the artifacts are always present in EEG raw data. Artifacts are divided in to following three main categories

[#] This paper has been presented at the International Conference on Advanced Technology & Sciences (ICAT'14) held in Antalya (Turkey), August 12-15, 2014.

^{1,2,3} Faculty Department of Electrical Engineering Air University Islamabad, Pakistan,

^{*}akram@mail.au.edu.pk

external, internal and physiological artifacts. Internal artifacts are caused by the actions of the subject itself[13]. External artifacts are caused by the actions of the outer sources.

1.2. External Artefacts:

The main supply line frequency may cause an artifact of 50 Hz frequency component to appear in EEG signal. This line noise or harmonics make its interference with Raw EEG signals because both occur in the same frequency band. Then notch filters are used to filter out these artifacts. During this filtering process many useful information of the EEG signals may be eliminated. The data of either one electrode or many many electrodes may be corrupted by line noises[11]. Constant Electric field is being produced by the surrounding walls and cables which also disturb the main EEG signal[20].

The electric and the magnetic fields external electric devices may also disturb the EEG signal [24].The artifacts have much stronger amplitude as compare to the original EEG signal. The verse thing is that the frequency range also overlaps to some extent. Being a significant component obtained with the EEG signal, it is therefore significant to make special considerations to eliminate these artifacts from the original signal for efficient detection of tumor[23]. Sometimes the Electrodes are not properly attached to the skin area because of miniaturizing or electrodes are attached to the hair area. There is a possibility that Electrodes are pushed or placed improperly.

1.3. Internal Artefacts:

The main source of movement artifact is the imperfect contacts of electrodes with the patient's body. The division of accelerometer signal is made in x and y directions. The observed patient is awake and moving. The accelerometer is attached to the forehead of the patient. It is receiving the signals of different potentials with the movement of the patient[22]. The accelerometers used are the prototypes. The signal is collected with the pressure receptors

Muscle activity is the result of activity usually neck and facial muscles. Signals from these muscles comes in large range of frequencies. Moreover these obtained signals are distributed over over all scalp electrodes. Location of the source muscles plays a vital role in introducing noises in the respective electrodes[19]. Muscle potential on the scalp occur in different ways. Muscle relaxants are used to minimize the effect of the artifacts. This can also be reduced by posture changing of the subject. The facial movements or the Jaw movements are the main source of muscular potentials.

1.4. Heart-Beat or ECG Artifacts:

Varying voltages are induced on to the recordings with the expansion and contraction of blood vessel[21]. The frequency of this artifact is about 1.2Hz. the heart beat artifact appear either as spike or as a smooth wave. An electric field is being produced by A pulsating heart. This electrical activity of the heart is conducted to the scalp and from there onward electrodes are used to sense this electrical activity of the heart. Potential changes in the heart are the main sources of the occurrence of the electrical activity which is a measured potential of EEG signals. Similar to Muscular Potentials or EMG(Electromyogram) the ECG records potential differences in the cardiac muscle cell which appear during depolarization of the heart and repolarization of the heart within one cycle of the beat of the heart. In other words we can say that the recorded electrical activity of the heart also significantly effect the EEG.s

1.5. Physiological Artefacts

The muscular activity also Introduce physiological artifacts owing to heart and Ocular. Electrooculogram(EOG) aer very vital artifacts. The external artifacts can be eliminated by enhancing technologies. The physiological artifacts are only eliminated after the recording process. EOG is produced when an electric field is being varied around eyes. Hence an electric signal is generated because of this electric field. This electric signal is named as EOG. When this signal spread on the scalp it appears as a noise on the recorded EEG. To stop its interferences with the background EEG signal it need to be eliminated. [20]. Blinks and eye-movement not only appear in the EOG recording, but also influence the EEG signal.

2. Problem Statement

With the following assumptions the independent components of ICA model can be estimated meaningfully.

1. All components are statistically independent and must be mixed linearly.
2. Among the latent variables there must not be more than one Gaussian signal and the cumulative density function of latent variables should not be quite opposite than that of *logistic sigmoid*.
3. The total number of the latent variables n are equal or less than the total number of the observed variables $m(i.e.m \geq n)$. Many efforts have been made to solve this problem but no concrete solution yet. There appears a redundancy in the mixed signals when $m > n$. when $m = n$ the ICA works ideally.
4. The rows of the mixing matrix must be linearly independent because The mixing matrix is always full column rank. The mixed signals will be linear multiples of one another, when the mixing matrix are not of full rank then
5. Mixing matrix has negligible propagation delay.

The first assumption is of much significance while the other assumptions ensure that the estimated independent components are unique. Consider a mixture of N sources as

$$s_1(t), s_2(t), \dots, s_N(t) \dots \dots (1)$$

which are unknown but are assumed to be statistically independent from the observation of the M signals as

$$X_1(t), X_2(t), \dots, X_N(t) \dots \dots (2)$$

These signals result from a mixture of underlying sources. ICA needs at least as many mixtures as there are independent sources ($M \geq N$). In our research we considered $M = N$. We shall try to extract the desired EEG signal from the interference. The mixture is supposed to be linear and instantaneous, so that observations at time t instant result from a linear combination of the sources at that instant.

$$x_i(t) = \sum_{j=1}^N h_{ij} s_j(t) \quad i = 1, \dots, M \dots \dots (3)$$

Where h_{ij} are the unknown or the blind mixing coefficients that produce $x_i(t)$. For our research work we have modelled the system as $N \times M$ system and $N = M$. It may be a difficult for us to assume in the beginning that EEG data received for M electrodes is synthesized by exactly N statistically independent components as we ultimately can not know the exact number of independent

components embedded in the EEG data.

3. Proposed Solution

To eliminate the artefacts from EEG signals we have employed Analogue filtering method. In this method the reference signal EOG is used as input and EOG artefacts of EEG and ECG base line noise are subtracted. The analogue filtering is being performed by ICA technique. Removing these artefacts in EEG signals using ICA or FASTICA algorithm will significantly increase the chances of Early tumor detection in brain. In our research work we present an efficient method of removing artefacts from EEG signals which in turn detect brain tumor efficiently. EEG records carry information about abnormalities and responses of human brain and heart respectively. The EEG signals are very vulnerable and highly effected by many artefacts(noises). These artifacts are caused by various factors, like line interference, subject movements, loose electrode contacts with human body, muscle activity, heart beat and field around the eyes(EOG). Because of the presence of these artefacts it becomes difficult to analyse EEG signals and thus to obtain effective useful clinical information. In order to get the useful information we must have to remove these artefacts.

First to eliminate the artefacts from EEG signals we have employed Analogue filtering method. In this method EOG signal is always used as reference input signal and subtracts EOG artefacts of EEG and base line noise of ECG. The function of analogue filtering is being performed by Independent component analysis(ICA)Technique.

4. Results

source matrix having two statistically independent uniform random variables as:

$$S = [s_1, s_2]^T \dots (4)$$

The best representation of ICA latent variables is in the form of Matrix. According to equation 2 where matrix H producing X can represent the original multivariate. The algorithm FASTICA contains the following four steps

- 1) Zero mean the available all mixed signals.
- 2) Make the available data whiten.
- 3) Unit Norm the initial weight vector.
- 4) For fulfilment of the above three stated steps Consider

$$w_{new} = aE\{mg(w^T m)\} - aE\{mg'(w^T m)\}w \dots (5)$$

Here contrast function is represented by g .

- 5) Suppose

$$w_{new} = w_{new} / \|w_{new}\| \dots (6)$$

Where w_{new} is representing weight vector. At every iteration it is updated.

- 6) Make comparison of old with new one, if there is convergence start another from new row, otherwise perform step 4.

$$\Delta \bar{w} = \bar{w}(n) - \bar{w}(n-1) \quad \text{if} \quad \|\Delta \bar{w}\| < \delta \dots (7)$$

Fixed point iteration scheme is the main base of Fast ICA in order to find out the maxima of the non-gaussianity for $W^T X$ as measured in previous equations. This can also be obtained as an

approximate Newton iteration denoted by g , which is the derivative of the function g which is a non-quadratic function is used in equation (6). The derivative of equations(7) is

$$g_1(u) = \tanh(a_1 u), \quad g_2(u) = u \exp(-u^2 / 2) \dots (8)$$

Where $1 \leq a_1 \leq 2$. As suitable constant it is $a_1 = 1$. The basic steps for the FASTICA algorithm are following.

1. An initial(random) weight vector is chosen
2. Let

$$W^+ a = aE\{Xg(W^T X)\} - aE\{g'(W^T X)\}W \dots (9)$$

$$S = [s_1, s_2, s_3 \dots s_n]^T \dots (10)$$

According to equation(2) ICA tries to determine unmixing matrix W as per following equation

$$W \approx H^{-1} \dots (11)$$

This gives the resulting matrix as:

$$Y = WX = W(HS) = S' \approx S \quad (W \approx H^{-1}) \dots (12)$$

This template was provided by courtesy of Causal

$$3. \text{ Consider} \quad W = \frac{W^+}{\|W^+\|} \dots (13)$$

4. When no convergence then go back to step

It is important to note that by convergence we understand that the new and the old values of matrix W are moving in the one and in same direction this shows that both have dot product 1. But this does not mean that the vector converge to a single value. Since the components are independent so W and $-W$ have Independent components and can be defined only up to a multiplicative sign. It is also assumed that the data is pre-whitened.

Table 1. Results Comparison of Fast ICA and PPGA

Algorithms	Mean Error	Standard Error
ICA	0.00039	0.0407
PPGA	1.0087e-004	0.0210
ICA	0.00038	0.0408
PPGA	1.0087e-004	0.0211

On the same grounds we again mix the noise signal and fast ICA algorithm is applied. We obtain the signal as shown in equation 12 above. The results for both PPGA and ICA are summarized in the table1. For EOG artifacts the data set I shown in figure 2 contains 36 seconds of data with sampling frequency $F_s=239\text{Hz}$. There are 26 channels of data(19 channels shown). The data was collected from electrodes placed at the scalp on slandered locations using international system. As shown in the figure the data contains a seizure onset around $t=1-3$ which is evident on T8-T9 channels with the wavy waveforms. Muscles artifacts are appearing on all channels from $t=304-315$ shown on T4 & F8 channels. Occipital or rear head movement artifacts are appearing around $t=0-5$. Eye blink artifacts at Fp1,Fp2 are appearing around $t=290-295$. Compare this data with figure 2. Showing onset rhythmic wave on electrodes F4, C4 with dark blue area in figure 2.

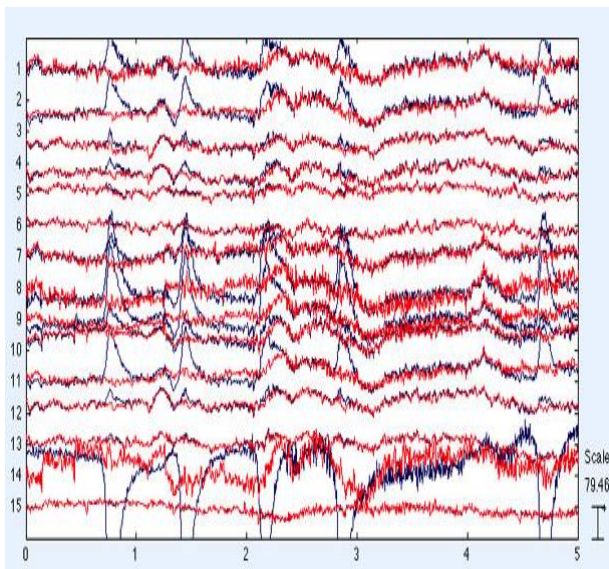


Figure 2. The corrected and the original EEG superimposed

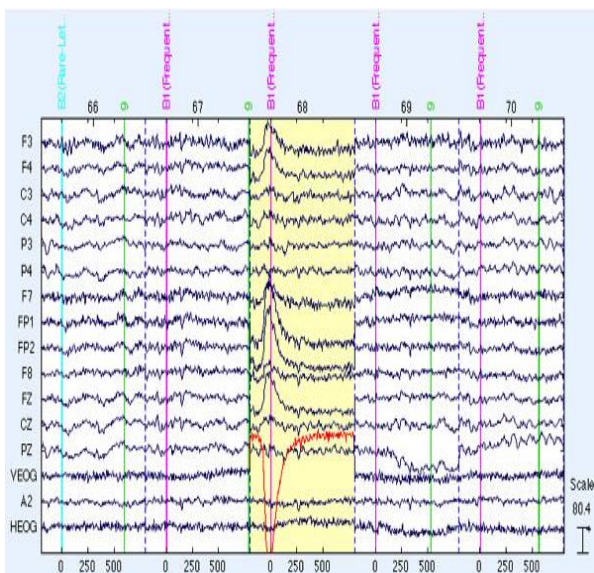


Figure 3. The time of stimulus presentation at the occurrence of eye blink,

We wish to exclude in uncorrected EEG data from the detected trials. For this purpose we must save the detected list of events by a text file so that whenever we want it back we import it easily back for ICA dataset which is already corrected. Before epoching the event was number 132. And after epoching the event was number 62. Because epochs have not been generated to bins as response events were have not been assigned yet. We can assure that these were the detected events. In our data both the 2nd and 4th flags have made their presence we must keep it in mind that whenever an artefact is detected for the bin the 2nd flag is flagged.

Also we label the detected artefacts. Finally the event list is imported for the corrected dataset of ICA. Only by this way we surely be able to the blinking occurred trials at the time of stimulus during creation of ERPs from the corrected dataset. We may not be able to remove all artefacts during this process and some other artefacts may show their presence in the data for example artefacts due to patient movements. Hence the artefact detection procedure must be rerun before making averaging. The existence of epileptiform activity is difficult to detect in a most of patients. Rather it is impossible to determine with surety. The breach rhythm presence may not indicate an existence of seizures

in the patient[17]. The presence of breach rhythm in patients before operation is a universal phenomenon.

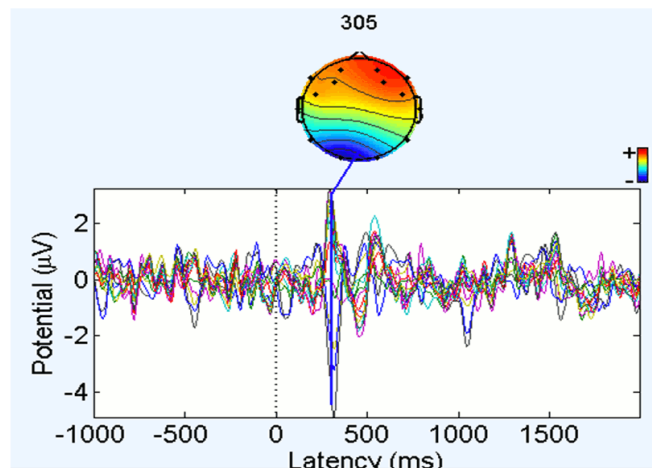


Figure 4. ERP waveforms of all channels for target stimulus

Figure above shows the recordings of a patient which are Preoperative and postoperative of a patient. The figure shows the changes in oligodendroglioma of patient. The signal showing such type of rhythmic activity may indicate the presence of a seizure. Such type of signal does not carry the characteristics such as frequency, amplitude or ictal discharge distributions.

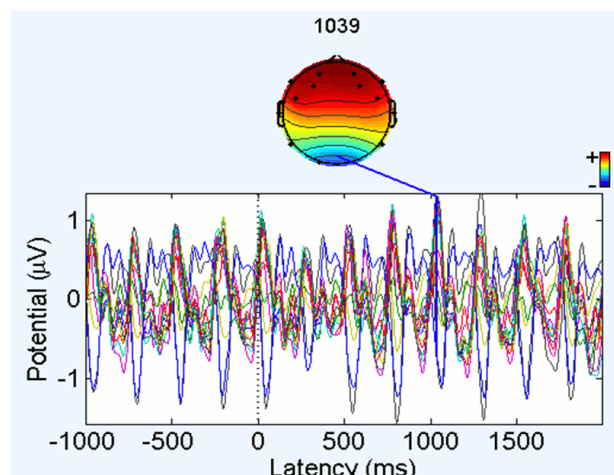


Figure .5 Waveforms of all channels for non-target events

The figure above shows the preoperative EEG in of a patient who have a oligodendroglioma with a right temporal. With some centroparietal spread Polymorphic delta activity is localized mainly over the right anterior-mid temporal region. (Right temporal chain is bottom 4 channels longitudinal bipolar montage;. Calibration: 1 second, 2 µV.) There is a temporal oligodendroglioma after tumor resection. Bilateral alpha is not present, the patient is drowsy. but right posterior alpha like activity With intermixed sharp theta and right posterior activity creates a breach rhythm. It is very much clear that polymorphic delta activity is not much prominent and also Beta also is largely increased on that side (this shows that there is an irregular eye movement which produce artefact in the recorded EEG data)

5. Conclusion

This is a basic tool used to measure brain activity. EEG is widely used for brain tumor detection and other Biomedical researches. This paper presents a system of efficient Malignant tumor detection involving a two-steps process. In the first step

All artifacts effecting EEG signal have been eliminated successfully as a pre-diagnosis stage of malignant tumor. EEG signals are highly contaminated from subject, equipments and from many other electromagnetic sources. We have employed EEG simulation data set I,II.

The noise corrupted signals are cleaned Employing FASTICA. . When the cells are originated from the brain itself Malignant tumor is called Primary tumor. When cell spread to brain from other locations of human body then tumors are termed as secondary. Primary Malignant brain tumor appear as hypo intense dark areas in fMRI T1 weighted images(figures 2 and 3) and hyper intense bright areas T2 weighted images figure 2. Blind source separation technique applied for removal of this noise is quite simple and converges in a few steps as compare to other adaptive filtering techniques and hence FASTICA employed EEG signals.

Acknowledgements

This study is a research work in Biomedical Engineering using Image processing and signal processing. I hereby thanks Dr. Aamer Sleemchaudhury Professor at University of Lahore Islamabad Campus and Dr. Aqdas Naveed Malik Deen Faculty of Engineering IIUI for their kind cooperation in research.

References

[1] Akram Rashid, Zahooruddin and Dr. I.M. Qureshi “Electrocardiogram Signal Processing for baseline Noise Removal using Blind Source Separation Techniques: A Comparative Analysis” IEEE Trans on Machine Learning and Cybernetics Guilin, Guangxi, China July 2011. Volume 4, pp1756

[2] A.E.H. Emery, “Population frequencies of inherited neuromuscular disease- A world survey”, *Neuromuscular Disorders*, Vol 1, No 1, pp. 19-29, 1991

[3] Prabhakar K. Nayakand Niranjana U . Cholayya“Independent component analysis of Electroencephalogram”, *IEE Japan Papers of Technical Meeting on Medical and Biological Engineering* Vol.MBE-06, No.95-115,pp.25-28,2006.

[4] “Brain Tumor” from http://en.wikipedia.org/wiki/Brain_tumor

[5] H.G. Hosseini, H. Nazeran and K.J Reynolds, “ECG Noise Cancellation Using Digital Filters,” 2nd International Conference on Bioelectromagnetism, Australia, pp. 151-152, 1998.

[6] MA Mneimneh, Eeyaz, “An adaptive kalman filter for removing baseline wandering in ecg signal.” *Computers in Cardiology*, 2006;33:253–256.

[7] Steven Pacia, MD, “Brain Tumors and Epilepsy”, *Fight against childhood epilepsy & seizures*.

[8] Dr. YW Fan, Dr. Gilberto KK Leung, “ Management of seizure associated with brain tumor”, *Medical Bulletin*, Vol. 11, No. 4, April 2006.

[9] Aapo Hyvarinen, “A Family Of Fixed-Point Algorithms For Independent Component Analysis,” 0-8186-0/97 1997 IEEE.

[10] Mangano FT, McBride AE, and Schneider SJ., “Brain tumors and epilepsy”, In: Ettinger AB and Devinsky O, eds. *Managing epilepsy and co-existing disorders*. Boston: Butterworth-Heinemann; pp. 175-194, 2002.

[11] Aapo Hyvarinen “One-Unit Contrast Functions for

Independent component analysis: A Statistical Analysis,”Helsinki University of Technology, Laboratory of Computer and Information Sciences, Finland.

[12] Akram Rashid, Zahooruddin, Ijaz Mansoor Qureshi, Aamer Saleem “Electrocardiogram Signal Processing for Baseline Noise Removal USING Blind source separation techniques” *International Conference on Machine Learning and Cybernetics*, Vol. 4, pp 1756-1761, 2011

[13] Akram Rashid “Multiuser Detection in DS-CDMA using Evolutionary Techniques”*International Conference on Computer, Electrical, System Science and Engineering*, France Paris July 28-30. 2010.

[14] C. Majos, M. Julia-Sape, J. Alonso, M. Serrallonga, C. Aguilera, J. Acebes, C. Arus and J. Gili, “Brain tumorclassification by proton MR Spectroscopy: Comparison of Diagnostic Accuracy at Short and Long TE”, *American Journal Neuroradiology*, 25:1696-1704, November–December 2004.

[15] Jie Li, Guan Han, Jing Wen and Xinbo Gao, “Robust tensor subspace learning for anomaly detection”, *International Journal of Machine Learning and Cybernetics*, March 2011, DOI:10.1007/s13042-011-0017-0.

[16] Alfons Schuster and Yoko Yamaguchi, “From foundational issues in artificial intelligence to intelligent memristivenano-devices”, *International Journal of Machine Learning and Cybernetics*, March 2011, DOI: 10.1007/s13042-011-0016-1.

[17] Weiguo Yi, Mingyu Lu and Zhi Liu, “Multi-valued attribute and multi-labeled data decision tree algorithm”, *International Journal of Machine Learning and Cybernetics*, March 2011, DOI: 10.1007/s13042-011-0015-2.

[18] Jie Zhu, Xiaoping Li and Weiming Shen, “Effective genetic algorithm for resource-constrained project scheduling with limited preemptions”, *International Journal of Machine Learning and Cybernetics*, March 2011. DOI: 10.1007/s13042-011-0014-3.

[19] Shumei Zhang, Paul McCullagh, Chris Nugent, HuiruZheng and Matthias Baumgarten, “Optimal model selection for posture recognition in home-based healthcare”, *International Journal of Machine Learning and Cybernetics*, Vol. 2, No. 1, pp. 1-14, 2011.

[20] Yi Tang, Pingkun Yan, Yuan Yuan and Xuelong Li, “Single-image super-resolution via local learning”, *International Journal of Machine Learning and Cybernetics*, Vol. 2, No. 1, pp. 15-23, 2011.

[21] Fadi N Karamah, Munther A. Dahleh, “Automated classification of EEG/ECG signals in tumor diagnostic” , *Proceedings of American control conference*, Chicago, Illinois, June 2012.

[22] R. Verleger, T. Gasser, & J. Mocks, “Correlation of EOG artifacts in eventrelated potentials of EEG: Aspects of reliability and validity” , *psychophysiology*, Vol. 9, pp 472-480,2011.

[23] M. Murugesan, Mrs. R. Sukanesh “ Towards Detection of Brain Tumor in Electroencephalogram Signals using Support Vector Machines”, *International Journal of Computer Theory and Engineering*, Vol. 1 No.5, December 2011.