IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 4, April 2017

A Survey on EEG Feature Extraction and Feature Classification methods in Brain Computer Interface

Mangala Gowri S G[1], Cyril Prasanna Raj P [2]

Assistant Professor, ECE Department, M S Engineering College, VTU, Bangalore, India[1]

Dean of R & D, ECE Department, M S Engineering College, VTU, Bangalore, India[2]

Abstract: Brain Computer Interfacing (BCI) is a methodology which provides a way for communication from outside world using brain signals. It detects the specific patterns in a person's ongoing brain activity which relates to the person's intention to initiate control. The BCI system translates these patterns into meaningful control command. To develop BCI system, various signal processing algorithms are proposed. Electroencephalogram (EEG) signals are used to extract the features and further it is classified. A survey of different Classification algorithms is used in EEG-based BCI research and to identify their critical properties. This paper is organized with a recent methodology of feature extraction and feature Classification algorithms. It also aims at addressing the methods and technology adapted in each phase of the EEG signal processing. It also highlights the pros and cons by reviewing literatures, books and other related documents. This survey helps in designing a suitable algorithm for the development and implementation of further classification of signals.

Keywords: Brain Computer Interface (BCI), Electroencephalogram (EEG), Feature Extraction, Wavelet Transform, Feature Classification.

1. INTRODUCTION

Brain-Computer Interface (BCI) is a method of communication based on the neuronal activity generated by the brain and its normal output pathways between nerves and muscles. It also emphasizes on Human-Machine Interaction between humans, computers and machines [12].This paper summarizes on the recent comprehensive survey of feature extraction methods and feature classification techniques. Electroencephalogram (EEG) signals are electrical signals collected from the scalp. They are frequently used in Brain-computer Interaction. However, EEG signals change over time and are highly non-stationary. The major challenges in BCI research are how to extract the features in time varying EEG signals. The EEG signals record as a weak potential by placing the electrodes on the scalp and analyse to establish a BCI.The acquired raw EEG signal is pre-processed. Figure 1, represents the raw EEG signal.

Pre-processing is a state of processing the EEG signal to remove the baseline and performing its average of the signal from the original signals. The noise free EEG signal is analysed by using wavelet transform to extract all the fundamental frequency components of EEG signal i.e. alpha, beta, gamma, delta and theta.

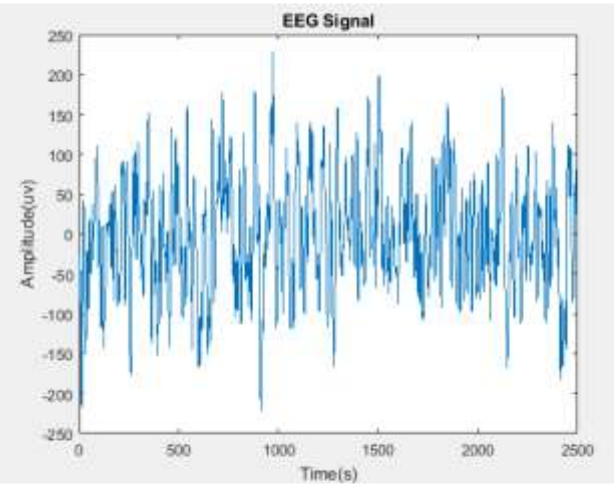

Figure 1: Raw EEG Signal

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 4, April 2017

From the BCI block diagram, an EEG signal has been acquired from the scalp of the brain using EEG acquisition set-up and the EEG Bands are extracted. The frequency sub bands are further classified and the signals are detected.
The separation of EEG for feature classification is based on the decomposition of the signal to certain levels. Figure 2, shows the Basic Block diagram of Brain Computer Interface System

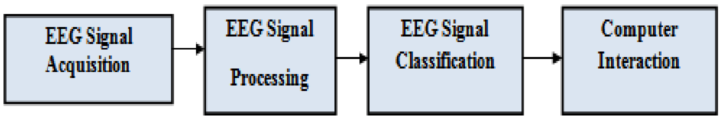

Figure 2: Brain Computer Interface System

From the EEG signal it is possible to differentiate gamma, Beta alpha, delta and theta bands. The gamma band is modulated by sensory input and internal processes which relates to memory and attention, which lies at a frequency of 30-100 Hz. The beta is located at frontal, central region which are associated with normal waking consciousness, which lies at a frequency of 13-30 Hz. Alpha waves have a frequency spectrum ranging from 8-13 Hz which lie in the occipital region. It refers to an awake person when the eyes are closed. Beta waves lie at a frequency ranging from 13-30 Hz and are detected in the parietal and the frontal lobes. The delta waves have the frequency range of 0.5-4 Hz and are detectable in infants and sleeping adults. The theta waves have the frequency range of 4-8 Hz and are obtained from children and sleeping adults. The following, **Table 1** illustrates the different EEG Band levels.

Table 1: Different EEG Bands with their locations

EEG BAND	FREQUENCY RANGE	LOCATION
DELTA	0-4 HZ	Frontal Lobe
THETA	4-7 HZ	Midline Temp
ALPHA	8-13 HZ	Frontal Occipital
BETA	13-30 HZ	Frontal Central
GAMMA	30-100HZ	Parietal Lobe

2. RELATED WORK

Emotions play a significant role in affecting computing. The emotions are happy, sad, surprise, angry etc, which are used to find the mental stress and mental disorders. In human brain each and every cell performs a specific function. Each and every function is used to analysis the decision making for a particular problem. There are many works which are proposed to understand the EEG analysis which relates to feature extraction and a classification techniques, a survey of many works has been discussed. In [1] different kinds of emotions and their EEG features and different feature extraction methods like Discrete wavelet transform (DWT) and SVM are implemented [2] Dimensionality reduction methods like PCA, LDA, and CFS are adopted and compared on the feature set. The best average classification accuracy of 91.77% was achieved. In [3], adaptive methods for EEG signal segmentation in time-frequency domain was implemented. C.Petrantonais used higher order crossings(HOC) for feature extraction and a robust classification method by name HOC-Emotion classifier (HOC-EC) implemented using four different classifiers namely quadratic discriminant analysis (QDA), k-nearest neighbour, Mahalanobis distance, and support vector machines (SVMs), in order to accomplish efficient emotion recognition. In [5] SVM's are used to classify those emotions into different group like negative and positive emotions The emotions are identified using short term assessment and particular time related emotions only extract using STFT and MI and classify those emotions using SVM, RVM. All these techniques and methods are implemented using MATLAB and different subjects were used to analysis the emotions. Mandeep singh et.al in [6] considers EEG signals for 4 different subjects. The extracted EEG signals are decomposed into different sub bands using discrete wavelet transform. EEG bands namely gamma, beta,

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 4, April 2017

alpha, theta and delta are extracted which is classified into 2 classes of emotions namely High arousal (HA) and Low arousal (LA). In [7] EEG signals are classified using two emotions (i.e., positive and negative) by giving an external stimulus. The power spectrum features, are analysed with an accuracy rate of about 85.41% by using SVM Classifier. All these methods proposed by these authors will give a brief summary of all the techniques used for both feature extraction and feature classification of EEG signals.

3. 10-20 ELECTRODE PLACEMENT SYSTEM

The placement of electrodes over the scalp for data capturing is done based on the Standard International System for placement of electrodes by the name 10-20 Electrode Placement System, 10-10 System and 10-5 System**[23][24]**.The 10-20 system of placement of electrodes is shown below in Fig.3. Many of the researchers have used 10-20 system to capture data. Each electrode placed on the scalp is defined using alphabets. The frontal lobe is designated as F, parietal lobe as P, Temporal lobe as T and Occipital lobe is designated as O.All the above researchers used 10-20 system for placement of electrodes. It is an internationally recognized method for the placement of electrodes for the EEG signals to capture EEG data. From Figure 3, the Vertical imaginary line is drawn from nasion to the inion and a horizontal line from left ear lobe to right ear lobe. From 10% above the nasion and inion, along vertical line a circle is drawn around the head, other electrodes are positioned maintaining a 20% inter electrode as indicated by the 20.20% up from the circle from the nasion is Fz, and another 20% further along is the top of head labelled Cz. Pz is positioned on the vertical line in a similar manner.
C3, T3, C4 and T4 are positioned in the same way along the horizontal mark. The electrodes on the imaginary circle are also at a 20% distance from each other, while keeping T3 and T4 on the horizontal line. The remaining electrodes are placed equidistant between the vertical line and the circle, filling the horizontal lines of the frontal and parietal

B. Principal Component Analysis (PCA): It is a type of dimensional reduction analysis technique used for feature extraction. In PCA, dimensional reduction is achieved by projection to lower dimensional space using linear transformation. Although PCA is a simple and classical method, it can often effectively reduce redundant information. PCA assumes that the data is linear.

C. Wavelet Transform (WT): The EEG signal is non-stationary and is best suited for time-frequency methods. The transforms are a family of functions derived from a generating function called mother wavelet using translation and dilation operations. Wavelet transform has the advantage of having a varying window size which is broad at low frequency and narrow at high frequency. Short time windows are accustomed to get high frequency data. Wavelet transform gives precise data at high frequencies [13].This makes the wavelet transform suitable to get the frequency data at low frequencies and precise time for the analysis of irregular knowledge patterns, such as impulses occurring at various time instances and also uses multivariate signals.

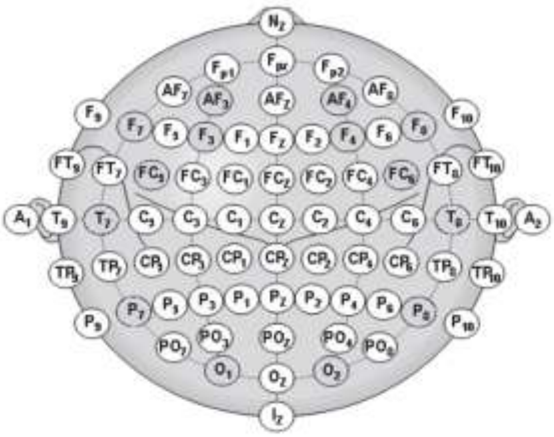

Figure 3:The 10-20 system of placement of electrodes [22]

D. Continuous Wavelet Transform (CWT):
A continuous wavelet transform (CWT) is employed to divide a continuous time function into wavelets. Unlike Fourier transform, the continuous wavelet transform acquires the power to construct a time frequency representation of a signal that gives excellent time frequency localization. However, its major weakness is that scaling parameter and translation parameter of CWT change continuously. Thus, the coefficients of the wavelet for all available scales after calculation will consume a lot of effort and yield a lot of unused information [14]

4. FEATURE EXTRACTION METHODS

Feature Extraction is the process of identifying a particular information form EEG which is been measured by the neuronal activity from the brain. Features are characteristics of a signal that are able to differentiate emotions .The main task of feature extraction is to derive the salient features which can map the EEG data into consequent emotion states. There are different feature extraction methods used for analyzing the EEG signals. After obtaining the noise-free signals from the signal enhancement phase, essential features from the brain signals were extracted. Some of the proposed methods include discrete wavelet transform (DWT), Fast Fourier Transforms (FFT), Adaptive Auto Regressive parameters (AAR), Bilinear AAR, multivariate AAR, PCA, ICA, Genetic Algorithms. Among these FFT, DWT, ICA, PCA are the most commonly used feature Extraction methods. The different Extraction techniques are discussed below, to analyse and choose the best Feature Extraction method in our Research.

A. Independent Component Analysis (ICA): ICA is one of the most commonly used feature extraction techniques.ICA is a blind source separation technique (of knowing the source and channel characteristics), and used for the extraction of multiplying source signal with unknown mixing matrix. The aim of ICA will be to obtain un-mixing matrix that original source signals can be reconstructed.

B. Discrete Wavelet Transform (DWT): DWT are wavelets which are discretely sampled. By considering a signal x(n), as shown in Fig.4 , the signal is decomposed into high pass filter and low pass filter , where the outputs are giving the detailed co-efficient from high pass filter and approximation co-efficient from the low pass filter. The two filters are related to each other and are known as "**Quadrature Mirror Filter"**. DWT works on the principle of multi scale feature representation.

C. Fast Fourier Transform (FFT): The FFT is an important and efficient tool for the feature extraction. FFT algorithm is involved a wide range of mathematical operation from simple real and complex numbers arithmetic to group theory. The calculation is very complex and time consuming to reduce the operation time and increasing the speed by using FFT. They are used in a wide variety of applications, from digital signal processing and solving partial differential equations to algorithms for quick multiplication of large integers.

5. FEATURE CLASSIFICATION TECHNIQUES

Different extraction methods are understood to select a better method, the extracted features are fed as input for further classification. It is done by a suitable classifier for recognition and detection of emotions. A classifier is a system that divides some data into different classes, and gives the relationship between the features and the emotion that belongs to that part of the EEG signal. There are several methods of classification like Artificial Neural Network, Support Vector Machine, and Linear Discriminant Classifier-NN classifiers and PNN Classifier are used for classification which can be understood and can be chosen from the different classifiers. After extracting, the feature classification is used to group the related emotions. The following methods are discussed below.

5.1. ARTIFICIAL NEURAL NETWORK: An Artificial neural Network (ANN) is a computational model based on the structure and functions of biological neurons. The information is passed through the network which affects the structure of the ANN as the network changes and learns the different paradigms based on the input and output. [15] Information flows from one node to another node and neurons are called as elements. NN has the input layer, hidden layer and output layer. It has following features, Computer Based Learning, Ability to arrange nodes itself, Real Time Processing, Ability to with stand in failure. Neural network finds the pattern to analyse the particular emotion. There are three types of learning from the neural network they are Supervised Training, Unsupervised Training, and Reinforcement Training. For Supervised Learning there is a target class which is used to find the output exactly we want. Unsupervised Training there is no target class we have to find the output, which is related to the problem. Reinforcement is the combination of both. Based on these NN we have to find a particular pattern for emotion.

5.2. K-NEAREST NEIGHBOR (K-NN): KNN is a non-parametric method that classifies the data by comparing the training data and testing data based on estimating the feature values. Nearest-neighbour classifiers are based on learning by comparing test tuple with training tuples that are similar to it. When tuple is not familiar then **k**-nearest-neighbour classifier searches the pattern space for the **k** training tuples that are closest to the unknown tuple. These **k** training tuples are the **k** "nearest neighbours" of the unknown tuple. The closeness is defined by using Euclidean distance formula given by:

$$\mathbf{Dist(X1, X2)} = \sqrt{\sum (\mathbf{x1i} - \mathbf{x2i})^2} \ \ldots\ldots\ldots\ldots \ \mathbf{(1)}$$

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 4, April 2017

The Euclidean space between two points or tuples, say, **X1 = (x11, x12… x1n)** and **X2 = (x21, x22 ...x2n)**, is where, x1i and x2i represents the training and testing data respectively. Different attributes are measured on different scales, so if the Euclidean distance formula is used directly, the effect of some attributes might be completely dwarfed by others that have larger scales of measurement. After feature extraction process the EEG training data and test data is passed to the classification process. Then Euclidean distance is calculated between each EEG training sample and testing sample. The class for first K neighbors is considered and the majority vote is the classified class. The accuracy for the KNN is high as compared to the other classifiers. [15]

5.3. SUPPORT VECTOR MACHINE (SVM): The SVM is extensively used for classification, regression and density. It maps the input patterns into a higher dimensional feature space through some non-linear mapping. A linear decision surface is then constructed in this high-dimensional-feature space. SVM is considered as a linear classifier in the parameter space, but it becomes a nonlinear classifier as a result of the nonlinear mapping of the space of the input patterns into the high-dimensional feature space. Training the SVM is a quadratic-optimization problem. The construction of a hyper plane given by **wT x + b = 0** where (**w** is the vector of hyper plane coefficients, b is a bias term) so that the margin between the hyper plane and the nearest point is maximized and can be posed as the quadratic-optimization problem.SVM has been shown to provide high-generalization ability. A proper kernel function for a certain problem is dependent on the specific data and till now there is no good method on how to choose a kernel function.

5.4. LINEAR DISCRIMINANT ANALYSIS (LDA): Linear discriminant analysis (LDA) is one of the most popular classification algorithms for Brain Computer Interface applications, and has been used successfully in a large number of systems. LDA linearly transforms data from high dimensional space to low dimensional space. Finally the decision is made in the low dimensional space. Thus the definition of the decision boundary plays an important role in classification process.

5.5. PROBABILISTIC NEURAL NETWORK (PNN): A single PNN is capable of handling multiclass problems. This is opposite to SVM, which decompose a multiclass classification problem into dichotomies and each chotomizer has to separate a single class from all others. The PNN architecture is composed of many interconnected processing units or neurons organized in successive layers. The input layer unit does not perform any computation and simply distributes the input to the neurons in the pattern layer. On receiving a pattern x from the input layer, the neuron $\mathbf{x_{ij}}$ of the pattern layer computes its output given by the equation:

$$\mathbf{\Phi ij(x) = \frac{1}{(2\pi)^{d/2}\sigma^{d}} exp\left[-\frac{(x-x_{ij})^{T}(x-x_{ij})}{2\sigma^{2}}\right]} \text{ ---------- } \mathbf{(2)}$$

where **d** denotes the dimension of the pattern vector **x**, **σ** is the smoothing parameter, and $\mathbf{x_{ij}}$ is the neuron vector.

6. ADVANTAGES AND DISADVANTAGES OF FEATURE EXTRACTION TECHNIQUES

Table 2: Different Feature Extraction Methods with their advantages and disadvantages

Method Name	Advantages	Disadvantages	Analysis Method	Suitability
Discrete Wavelet Transform	i)DWT represents multi-resolution time-scale analysis ability ii)It is best suited for analysis of transient signal changes iii)It has a varying window size which is distributed at low frequencies and narrow at low frequencies and narrow at high frequencies.	i)Selection of an Appropriate Mother Wavelet	i) Both time, frequency domain and linear methods.	i)Transient and stationary signal
Independent Component Analysis	i) Compared to Principal component Analysis, ICA removes the constraint of orthogonality and forces components to be approxamately independent rather than simply uncorrelated.	i) ICA components lack the important variance maximization property possessed by PCA components.	i)Multivariat e Signals	i)Non Gaussian signals
Fast Fourier Transform	i)Good tool for stationary signal processing ii) Uses complex computations. It is suitable for narrowband signal iii)It has an enhanced speed over virtually all other available methods in real-time applications	i) Fourier transform is not suitable for location information, but suitable for frequency information and it is difficult in representing transients. ii) FFT suffers from large noise sensitivity and does not have shorter duration data record.	i)Frequency domain	i)Narrowband and Stationary signals
Principal Component Analysis	i) PCA is a statistical data analysis for feature extraction for linear combination of principal components.	i) Reduces the dimensionality of data and the data is complicated.	i)Multivariat e Dataset	i)Gaussian Distribution

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 4, April 2017

The above Table 2, illustrates the different feature methods with its advantages and disadvantages, which helps us to understand the EEG signal and analyse the emotion related features. From this we can identify the particular characteristic of the methods and also identify the methods.

7. ADVANTAGES AND DISADVANTAGES OF CLASSIFICATION METHODS

The following Table 3, explains that different feature classification method and its advantages, disadvantages, which helps us to understand the different Classification Algorithms and later can be used for emotion classification, where the emotions are detected and the accuracy is performed after classification

Table 3: Different Feature Classification Methods with their advantages and disadvantages

Method Name	Advantages	Disadvantages
Linear Discriminant Analysis	i)Extremely Fast ii) Low iii)LDA linearly transforms data from high dimensional space to low dimensional space	i)Fail to discriminate functions and variety of features ii)The distributions are non-Gaussian and the LDA projections preserve complex structure in data iii)LDA will also fail if discriminatory information is not in the mean but in the variance of the data
Support Vector Machine	i)Good Generalization ii)More Performance	i)Computational Complexity high
KNN	i)Easy to understand ii)Easy to Implement	i)Poor runtime performance ii)Sensitive to irrelevant and redundant features
Neural Networks	i)Easy to train ii)Accurate Pattern Classification	i)Needs training to operate ii) Requires high processing time for large network.

8. CONCLUSION

This Survey paper gives a clear review of different Feature Extraction and Classification methods in Brain Computer Interface, which can be implemented for the detection of human emotions, for identifying the best algorithm. Researchers will improvise the efficiency and accuracy in determining the appropriate method which can be used in Emotion Classification. The emotions are identified by EEG signal with different feature extraction techniques and classification methods. The emotions accuracy may be varied from one extract technique to another. Different extraction techniques that combine with the classification method provide better results. This accuracy level is used to analyse which method is suitable to classify the emotions in different genres of people.

REFERENCES

1) Murugappan, N. Ramachandaran, and Y. Sazali, Mohd “Classification of Human Emotion from EEG discrete wavelet” J. Biomed. SCI. Eng., vol. 3, pp-390-396, April 2010.
2) Xiao-Wei Wang, Dan Nile, “Emotional State Classification from EEG data using machine Learning Approach” Elsevier, 2013.
3) C.Petrantonais, “Adaptive Emotional Information Retrieval from EEG signals in the Time Frequency Domain” IEEE, 2012
4) C. Petrantonais, “Emotion recognition from EEG using higher order” IEEE, vol. 14, pp. 390-396, 2010.
5) Mohammed Soleymani, “Short term Emotion assessment in a recall paradigm” Elsevier, 2009.
6) Mandeep Singh, Mooninder Singh and Surabhi Gangwar, “Feature Extraction from EEG for Emotion Classification”, International Journal of IT and Knowledge Management (IJITKM), vol. 7, No.1, pp. 6-10,2013.
7) Noppadon Jatupaiboon, Setha Pan-ngum, Pasin Israsen,“Emotion Classification using Minimal EEG Channels and Frequency Bands”, 10th International Joint Conference on Computer Science and Software Engineering (JCSSE),2013
8) Varun Bajaj, Ram Bilas Pachori “Human Emotion Classification from EEG Signals using Multiwavelet Transform” International Conference on Medical Biometrics, 2014.
9) XW Wang, D Nie, and BL Lu, ‘‘EEG-Based Emotion Recognition Using Frequency Domain Features and Support Vector Machines,’’Neural Information Processing, Lecture Notes in Computer Science. Springer vol. 7062, pp. 734-743, 2011.
10) Nasehi and H. Pourghassem, “An optimal EEG-based Emotion recognition Algorithm using Gabor Features,” WSEAS Transactions on Signal Processing, vol. 8, pp. 87-99, July 2012.

DOI10.17148/IJARCCE.2017.64133

11) Amjed S. Al-Fahoum , Ausilah A. Al-Fraihat , "Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains", Biomedical Systems and Informatics Engineering Department,Jordan Hindawi Publishing Corporation, ISRN Neuroscience, Volume 2014
12) Mohamed Rizon , "Discrete Wavelet Transform Based Classification of Human Emotions Using Electroencephalogram Signals American Journal of Applied Sciences ,878-885, ISSN 1546-9239 , Science Publications, 2010
13) M. R. N. Kousarrizi, A. A. Ghanbari, M. Teshnehlab, "Feature extraction and classification of EEG signals using wavelet transform, SVM and artificial neural networks for brain computer interfaces," in Proceedings of the International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing (IJCBS '09), pp. 352–355, August 2009.
14) D.Cvetkovic, E.D. ¨ Ubeyli, I.Cosic, "Wavelet transform feature extraction from human PPG, ECG, and EEG signal responses to ELF PEMF exposures: a pilot study," Digital Signal Processing,vol. 18, no. 5, pp. 861–874, 2008.
15) J. Preethi, M.Sreeshakthy, A.Dhilipan, "A Survey on EEG Based Emotion Analysis using various Feature Extraction Techniques", International Journal of Science, Engineering and Technology Research (IJSETR), Department of Computer Science, Anna University Regional Centre ,Volume 3, Issue 11, November 2014
16) M. Ungureanu, C. Bigan, R. Strungaru, V. Lazarescu,"Independent Component Analysis Applied in Biomedical Signal Processing," Measurement Science Review, vol. 4, section 2, 2004.
17) Yuan-Pin Lin, Chi-Hong Wang, Tien-Lin Wu, Shyh-Kang Jeng,"Support Vector Machine for EEG signal classification during listening to emotional music," Multimedia Signal Processing, IEEE 10th Workshop, 2008.
18) Ali S. AlMejrad. "Human Emotions Detection using Brain Wave Signals: A Challenging", 2010.
19) Garrett, D., Peterson, D. A., Anderson, C. W., & Thaut, M. H."Comparison of linear, nonlinear, and feature selection methods for EEG signal classification". Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 11(2), 141-144, 2003.
20) Yuen, C. T., San, W. S., Seong, T. C., & Rizon, M. "Classification of Human Emotions from EEG Signals using Statistical Features and Neural Network". International Journal of Integrated Engineering, 2011
21) N. Hazarika, J. Z. Chen, A. C. Tsoi, and A. Sergejew, "Classification of EEG signals using the wavelet transform," Signal Processing, vol. 59, no. 1, pp. 61–72, 1997.
22) Sean Jenkins, Raymond Brown and Neil Rutterford, "Comparing Thermographic, EEG, and Subjective Measures if Affective Experience during Simulated Product Interactions", vol. 3, no.2, 2009.
23) D.O.Bos, "EEG-based Emotion Recognition," The Influence of Visual and Auditory Stimuli, pp.1-17, 2006.
24) G.E Chatrian, E. Lettich, P.L Nelson,, "Ten percent electrode system for topographic studies of spontaneous and evoked EEG activity", Am. J. EEG Technol. 25, pp. 83–92,1985.