Analysis of Electroencephalogram using Radial Basis Function

P.PADMAPRIYA

Dept of Biomedical Engineering,

Rajalakshmi Engineering College,

Chennai, India.

ppriyasmile@gmail.com

A.N.NITHYAA

Dept of Biomedical Engineering,

Rajalakshmi Engineering College,

Chennai, India.

nithyaa.an@rajalakshmi.edu.in

Abstract- **The project proposes an automatic support system for classification using artificial neural network for brain tumor detection for medical application. The detection of the brain tumor is a challenging problem, due to the structure of the tumor cells. Brain is the most incredibly complex organ in our body. It is made up of several neurons; it's the functional unit of brain. All over the world brain is still in research to analysis the EEG signals and to diagnose the several diseases. The Electroencephalogram (EEG) is the process of monitoring the activity of brain activities for diagnosing and treating physiological disorders such as brain tumors, multiple personality and depression. The EEG Machine is used to record the electrical potential generated by the nerve cells in the cerebral cortex. The scalp EEG is to be attractive tool in medical diagnosis since it is non invasive and real time depication of brain function. But at the same time EEG is in additive with various artifacts and noise. The artificial neural network will be used to classify the brain EEG signal that is tumor case or normal. The manual analysis of the signal is time consuming, inaccurate and requires intensive trained person to avoid diagnostic errors. For the EEG analysis neural network plays an important role to estimate the functions that do not have explicit mathematical models. Since the neural network can find the hidden information in the data and detect the abnormalities which lead the system to be robust.**

Keywords- **Electroencephalogram (EEG), Brain signals, artificial neural network (ANN), Principal component extraction (PCA), Radial Basis Function (RBF)**

I.INTRODUCTION

Automated classification and detection of Tumors in different medical signals is motivated by the necessity of high accuracy when dealing with a human life. Also, the computer assistance is demanded in medical institutions due to the fact that it could improve the results of humans in such a domain where the false negative cases must be at a very low rate. It has been proven that double reading of medical images could lead to better Tumor detection. But the cost implied in double reading is very high, that's why good software to assist humans in medical institutions is of great interest nowadays. Due to large number of patients in intensive care units and the need for continuous observation of such conditions, several techniques for automated diagnostic systems have been developed in recent years to attempt to solve this problem. In this paper the automated classification of Brain signals by using some prior knowledge like intensity and some anatomical features is proposed. Currently there are no methods widely accepted therefore automatic and reliable methods for Tumor detection are of great need and interest.

A.EEG

There are several biomedical signals extracted from our body. One such kind of the important signal is the EEG. The recording of electrical activity of the brain along the scalp is known as electroencephalogram. Its characteristics are frequency range is 0.5-100HZ, voltage is 2-100uV. They are different kinds of the brain signals alpha, beta, theta, delta. There are various advantages of EEG signals some of them can be stated as follows:
• Temporal resolution of EEG signal is high.
• EEG is a non-invasive procedure.
• EEG has ability to analyze brain activity.
The various uses of EEG signals are as follows:
• Diagnose epilepsy and see what types of seizures are occurring. EEG is the most useful and important test in confirming a diagnosis of epilepsy.
• Check for problems with loss of consciousness or dementia.
• Help find out a person's chance of recovery after a change in consciousness.
• Find out if a person who is in a coma is brain-dead.
• Study sleeps disorders, such as narcolepsy.
• Help find out if a person has a physical problem (problems in the brain, spinal cord, or Nervous system) or a mental health problem.

P.Padmapriya,   A.N.Nithyaa

473

II.TUMOR CELLS

All organs and tissues of the body are made up of building blocks known as cells. Nowadays tumor is the major disease along many people's which lead to death. Brain tumor is an abnormal growth of cells within the brain which may be cancerous or non cancerous. Although cells in different parts of the body may look and work differently, most repair themselves in the same way, by dividing to make more cells. Normally, this turnover takes place in an orderly and controlled manner. For some reason, the process gets out of control; the cells will continue to divide, developing into a lump, which is called a tumor.

There are numerous methods to diagnose a brain tumor which include a magnetic resonance, computerized axial tomography (CT) scan, an angiogram. However MRI and CT have the disadvantages that patient is exposed to large dose of radiation which affects the human health. Some patients get fear of seeing the apparatus setup. But the scalp EEG signal has the advantage that it is non invasive and it used to detect those tumors cells effectively.
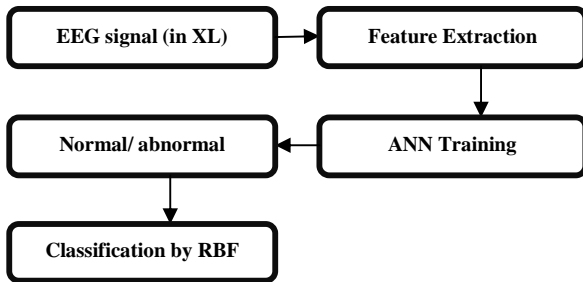
III.BLOCK DIAGRAM

Fig.1 Block diagram of tumor detection using ANN

A. Feature Extraction

Feature extraction is the collection of relevant information from the signal. In this project the feature is been extracted by using Principal Component Analysis.

B. Principal Component Analysis

PCA is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to the preceding components. Depending on the field of application, it is also named the discrete Karhunen–Loève transform (KLT), the Hotelling transform or proper orthogonal decomposition (POD).

C. Algorithm Flow of PCA Steps

1. Input.
2. Subtract the mean
   a) m = mean(T,2);
   b) temp = double(T(:,i)) - m;
3. Calculate the covariance matrix
   a) L = A'*A;
4. Calculate the eigenvectors and Eigen values of the covariance matrix
   a. [V D] = eig(L);
   b. L_eig_vec = [L_eig_vec V(:,i)];
   c. Eigenfaces = A * L_eig_vec;
5. Choosing components and forming a feature vector
6. Deriving the new data set.

Technically, a principal component can be defined as a linear combination of optimally-weighted observed variables. In the course of performing a principal component analysis, it is possible to calculate a score for each subject on a given principal component. For example, in the preceding study, each subject would have scores on two components: one score on the satisfaction with supervision component, and one score on the satisfaction with pay component. The subject's actual scores on the seven questionnaire items would be optimally weighted and then summed to compute their scores on a given component. Now it is mostly used as a tool in exploratory data analysis and for making predictive models. PCA can be done by Eigen value decomposition of a data covariance (or correlation) matrix or singular value decomposition of a data matrix, usually after mean centering (and normalizing or using Z-scores) the data matrix for each attribute. The results of a PCA are usually discussed in terms of component scores, sometimes called factor scores (the transformed variable values corresponding to a particular data point), and loadings (the weight by which each standardized original variable should be multiplied to get the component score).

P.Padmapriya, A.N.Nithyaa

474
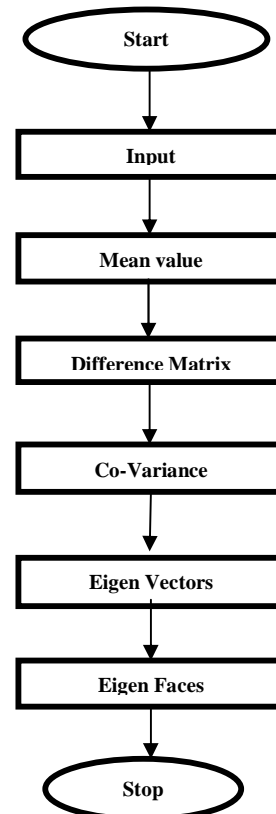
Fig.2 Algorithm flow of the PCA

Define a data matrix X^T, with zero empirical mean (the empirical (sample) mean of the distribution has been subtracted from the data set), where each of the *n* rows represents a different repetition of the experiment, and each of the *m* columns gives a particular kind of datum (say, the results from a particular probe). (Note that X^T is defined here and not X itself, and what we are calling X^T is often alternatively denoted as X itself.)

The singular value decomposition of X is $X = W\Sigma V^T$, where the $m \times m$ matrix W is the matrix of eigenvectors of the covariance matrix XX^T, the matrix Σ is an $m \times n$ rectangular diagonal matrix with nonnegative real numbers on the diagonal, and the $n \times n$ matrix V is the matrix of eigenvectors of X^TX. The PCA transformation that preserves dimensionality (that is, gives the same number of principal components as original variables) is then given by:

$$Y^T = X^T W$$
$$= V\Sigma^T W^T W$$
$$= V\Sigma^T$$

V is not uniquely defined in the usual case when $m < n - 1$, but Y will usually still be uniquely defined. Since W (by definition of the SVD of a real matrix) is an orthogonal matrix, each row of Y^T is simply a rotation of the corresponding row of X^T. The first column of Y^T is made up of the "scores" of the cases with respect to the "principal" component; the next column has the scores with respect to the "second principal" component.

Given a set of points in Euclidean space, the first principal component corresponds to a line that passes through the multidimensional mean and minimizes the sum of squares of the distances of the points from the line. The second principal component corresponds to the same concept after all correlation with the first principal component has been subtracted from the points. The singular values (in Σ) are the square roots of the Eigen values of the matrix XX^T. Each Eigen value is proportional to the portion of the "variance" (more correctly of the sum of the squared distances of the points from their multidimensional mean) that is correlated with each eigenvector. The sum of all the Eigen values is equal to the sum of the squared distances of the points from their multidimensional mean. PCA essentially rotates the set of points around their mean in order to align with the principal components. PCA is often used in this manner for dimensionality reduction. This advantage, however, comes at the price of greater computational requirements if compared, for example and when applicable, to the discrete cosine transform. Nonlinear dimensionality reduction techniques tend to be more computationally demanding than PCA. PCA is sensitive to the scaling of the variables. If we have just two variables and they have the same sample variance and are positively correlated, then the PCA will entail a rotation by 45° and the "loadings" for the two variables with respect to the principal component will be equal. But if we multiply all values of the first variable by 100, then the principal component will be almost the same as that variable, with a small contribution from the other variable, whereas the second component will be almost aligned with the second original variable. This means that whenever the different variables have different units (like temperature and mass); PCA is a somewhat arbitrary method of analysis. (Different results would be obtained if one used Fahrenheit rather than Celsius for example.). In mathematical terms we want to find Eigen vectors and Eigen values of a covariance matrix of images. Eigenvectors with highest Eigen values are the principle component of the Image set. We may lose some information if we ignore the components of lesser significance. But if the Eigen values are small then we won't lose much. Using those set of Eigen vectors we can construct Eigen faces.

IV. NEURAL NETWORK

Neural networks are predictive models loosely based on the action of biological neurons. The selection of the name "neural network" was one of the great PR successes of the Twentieth Century. It certainly sounds more exciting than a technical description such as "A network of weighted, additive values with nonlinear transfer functions". However, despite the name, neural networks are far from "thinking machines" or "artificial Brains". A typical artificial neural network might have a hundred neurons. In comparison, the human nervous system is believed to have about $3x10^{10}$ neurons. We are still light years from "Data". The original "Perceptron" model was developed by Frank Rosenblatt in 1958. Rosenblatt's model consisted of three layers, (1) a "retina" that distributed inputs to the second layer, (2) "association units" that combine the inputs with weights and trigger a threshold step function which feeds to the output layer, (3) the output layer which combines the values. Interest in neural networks was revived in 1986 when David Rumelhart, Geoffrey Hinton and Ronald Williams published "Learning Internal Representations by Error Propagation". They proposed a multilayer neural network with nonlinear but differentiable transfer functions that avoided the pitfalls of the original perceptron's step functions. They also provided a reasonably effective training algorithm for neural networks.

A. *Back Propagation Algorithm*

Consider a network with a single real input x and network function F. The derivative F'(x) is computed in two phases:
Feed-forward: the input x is fed into the network. The primitive functions at the nodes and their derivatives are evaluated at each node. The derivatives are stored.

Back propagation: The constant 1 is fed into the output unit and the network is run backwards. Incoming information to a node is added and the result is multiplied by the value stored in the left part of the unit. The result is transmitted to the left of the unit. The result collected at the input unit is the derivative of the network function with respect to x.

B. *Steps of the Algorithm*

The back propagation algorithm is used to compute the necessary corrections, after choosing the weights of the network randomly. The algorithm can be decomposed in the following four steps:
i) Feed-forward computation
ii) Back propagation to the output layer

P.Padmapriya, A.N.Nithyaa

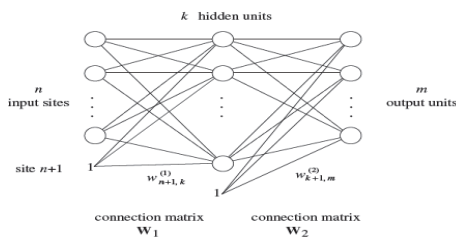iii) Back propagation to the hidden layer
iv)Weight updates

Fig.3 BPN architecture

C. How BPN Network work and Disadvantages

Although the implementation is very different, back propagation networks are conceptually similar to K-Nearest Neighbor (k-NN) models. The basic idea is that a predicted target value of an item is likely to be about the same as other items that have close values of the predictor variables. One of the disadvantages BPN models compared to multilayer perceptron networks is that BPN models are large due to the fact that there is one neuron for each training row. This causes the model to run slower than multilayer perceptron networks when using scoring to predict values for new rows. It is slow and inefficient. It can get stuck in local minima resulting in sub-optimal solutions.

D. Radial Basis Function

This is becoming an increasingly popular neural network with diverse applications and is probably the main rival to the multi-layered perceptron. Much of the inspiration for RBF networks has come from traditional statistical pattern classification techniques. A radial basis function network is an artificial neural network that uses radial basis functions as activation functions. The output of the network is a linear combination of radial basis functions of the inputs and neuron parameter.

Radial basis function networks are used for function approximation, time series prediction, and system control. Radial basis function (RBF) networks are feed-forward networks trained using a supervised training algorithm. They are typically configured with a single hidden layer of units whose activation function is selected from a class of functions called basis functions. While similar to back propagation in many respects, radial basis function networks have several advantages. They usually train much faster than back propagation networks. They are less susceptible to problems with non-stationary inputs because of the behavior of the radial basis function hidden units.

The major difference between RBF networks and back propagation networks (that is, multi layer perceptron trained by Back Propagation algorithm) is the behavior of the single hidden layer. Rather than using the sigmoidal or S-shaped activation function as in back propagation, the hidden units in RBF networks use a Gaussian or some other basis kernel function. Each hidden unit acts as a locally tuned processor that computes a score for the match between the input vector and its connection weights or centers. In effect, the basis units are highly specialized pattern detectors. The weights connecting the basis units to the outputs are

used to take linear combinations of the hidden units to product the final classification or output.
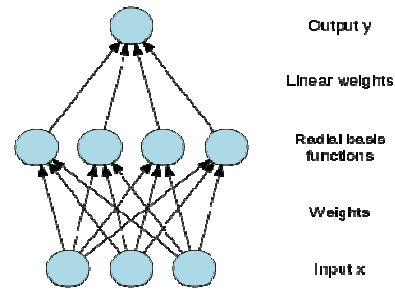
E. Network Architecture

Fig.4 Architecture of a radial basis function network.

The basic architecture for a RBF is a 3-layer network, as shown in Fig. The input layer is simply a fan-out layer and does no processing. The second or hidden layer performs a non-linear mapping from the input space into a (usually) higher dimensional space in which the patterns become linearly separable. Output layer is the final layer performs a simple weighted sum with a linear output. If the RBF network is used for function approximation (matching a real number) then this output is fine. However, if pattern classification is required, then a hard-limiter or sigmoid function could be placed on the output neurons to give 0/1 output values. The input can be modeled as a vector of real numbers X ε R^n. The output of the network is then a scalar function of the input vector, $\varphi: R^n \to R$, and is given by $\varphi(X) = \sum_{i=1}^{N} a_i \rho(\|X - C_i\|)$ where N is the number of neurons in the hidden layer, C^i is the center vector for neuron i, and a^i is the weight of neuron i in the linear output neuron. In the basic form all inputs are connected to each hidden neuron.

F. Clustering

The unique feature of the RBF network is the process performed in the hidden layer. The idea is that the patterns in the input space form clusters. If the centres of these clusters are known, then the distance from the cluster centre can be measured. Furthermore, this distance measure is made non-linear, so that if a pattern is in an area that is close to a cluster centre it gives a value close to 1. Beyond this area, the value drops dramatically. The notion is that this area is radically symmetrical around the cluster centre, so that the non-linear function becomes known as the radial-basis function.

G. Gaussian Function

The most commonly used radial-basis function is a Gaussian function. In a RBF network, r is the distance from the cluster centre.

The equation represents a Gaussian bell-shaped curve, as shown in Fig.
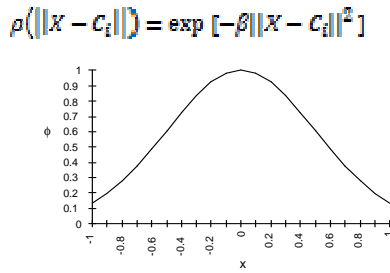
P.Padmapriya,   A.N.Nithyaa

476

$$\rho\left(\|X - C_i\|\right) = \exp\left[-\beta\|X - C_i\|^2\right]$$

Fig.5 Gaussian Function

H. Training
1) Training Hidden Layer

The hidden layer in a RBF network has units which have weights that correspond to the vector representation of the centre of a cluster. These weights are found either using a traditional clustering algorithm such as the *k*-means algorithm, or adaptively using essentially the Kohonen algorithm. In either case, the training is unsupervised but the number of clusters that you expect, *k*, is set in advance. The algorithms then find the best fit to these clusters. The *k*-means algorithm will be briefly outlined. Initially *k* points in the pattern space are randomly set. Then for each item of data in the training set, the distances are found from all of the *k* centres. The closest centre is chosen for each item of data - this is the initial classification, so all items of data will be assigned a class from 1 to *k*. Then, for all data which has been found to be class 1, the average or mean values are found for each of co-ordinates.

2) Training Output Layer

Having trained the hidden layer with some unsupervised learning, the final step is to train the output layer using a standard gradient descent technique such as the Least Mean Squares algorithm.

3) Advantages/Disadvantages

RBF trains faster than a MLP.Another advantage that is claimed is that the hidden layer is easier to interpret than the hidden layer in an MLP. Although the RBF is quick to train, when training is finished and it is being used it is slower than a MLP, so where speed is a factor a MLP may be more appropriate. It finds the input to output map using local approximators. Usually the supervised segment is simply a linear combination of the approximators. Since linear combiners have few weights, these networks train extremely fast and require fewer training samples.

I. Mathematical Model

In summary, the mathematical model of the RBF network can be expressed as:

$\mathbf{x} = \mathbf{f}(\mathbf{u})$, $\mathbf{f}: \mathbb{R}^N \to \mathbb{R}^M$

$X_j = f_j(u) = w_{0j} + \Sigma\, w_{ij}G(\|u - c_j\|)$, j=1,2,.., *M*

Where is the Euclidean distance between $\mathbf{u}$ and c_i

V. RESULT

The EEG signals obtained from different patients. The features of those are extracted by using PCA.

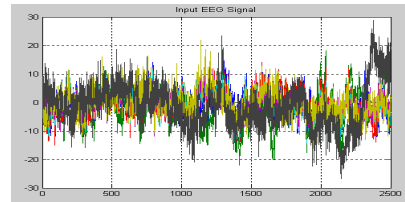### A. EEG of Brain Tumor Person Measured at Every Channel Locations

Fig. 6 Output from normal person with artifacts

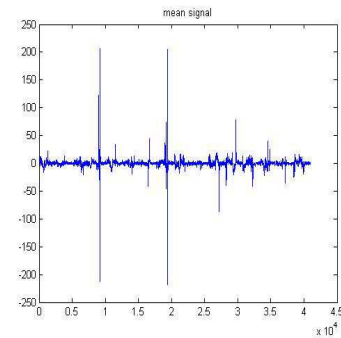### B. Feature Extracted Output

Fig.7 Output obtained after calculating from the statistical parameters

B. Training and Testing using RBF

Thus the extracted features are trained using the neural network tool. The input is fed to test whether the corresponding signal is normal or abnormal using RBF.
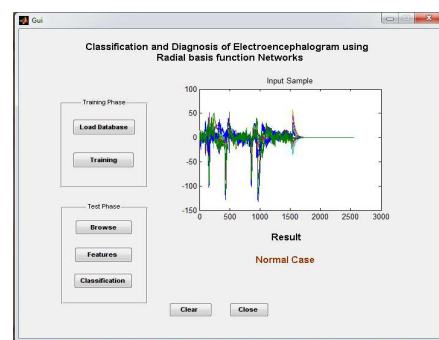

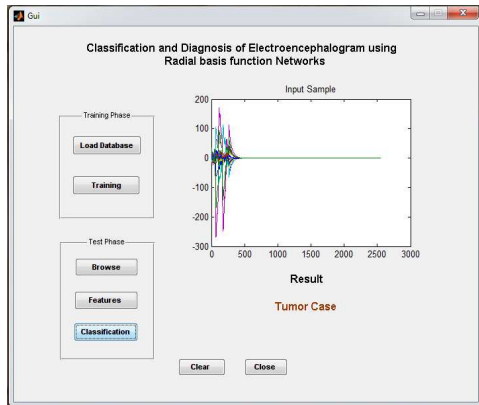*Fig.8 Output obtained after RBF training and classification for normal case*

P.Padmapriya,   A.N.Nithyaa

Fig.9

*Output obtained after RBF training
and classification for tumor case*

VI CONCLUSION

Thus the tumor detection is made simple and effective in this paper. The scalp EEG signal plays an important role in detecting the tumor. It helps in medical diagnosis and reduces the time consuming and error occurring due to the carelessness of the human. The RBF is used to train and test the signals of the patient. It gives the result in efficient manner which tells about the normal and abnormal cases.

REFERENCES

[1] J.R.Wolpaw,N.Birbaumer,D.J.McFarland, G.Pfurtscheller,T.M.Vaughan,*"Brain-computer interfaces for communication and control",* (invited review) *J.Clinical Neurophysiology Elsevier*, Vol. 113, pp. 767-791, 2002.

[2] E.Ben George,M.Karnan,*"Feature Extraction and Classification of Brain Tumor Using Bacteria Foraging Optimization Algorithm and Back Propagation Neural Networks"*, European Journal of Scientific Research, vol.88 No 3, Oct 2012, pp.327-333.

[3] Seenwasen Chetty, Ganesh K.Venayagamoorthy, *"A Neural Network Based Detection of Brain Tumor Using Electroencephalogram"*, International Conference Artificial Intelligence and Soft Computing, July 2002, pp.391-396.

[4] A.S.Miller, B.H.Blott and T.K.Hames, *"Review of Neural Network Applications in Medical Imaging and Signal Processing",* Proc. of The International Society of Computer Science, vol.2103, 1994, pp.89-92.

[5] E.Niedermeyer and F.Lopes da silva, *"Electroencephalogram, Basic Principles, Clinical Applications and Related Fields",* William and wilkins, 1993.

[6] Sharanreddy.M and Dr.P.K.Kulkarni, *"Review of Significant Research on EEG Based Automated Detection of Epilepsy Seizures and Brain Tumor",* International Journal of Scientific and Engineeering Research, vol2,Aug-2011,pp.1-9.

[7] Hojjat Adeli, Samanwoy Ghosh and Dastidat, *"Automated EEG Based Diagnosis of Neurological Disorders",* CRC Press, 1 edition, 2010.

[8] Indu sekhar samant, Guru Kalyan Kanungo, Santosh Kumar Mishra, *"Desired EEG Signals For Detection Brain Tumor Using LMS Algorithm and Feedforward Network",*International Journal of Engineering Trends and Technology, vol 3- 2012, pp.718-723.

[9] Fadi N.Karameh, Munther A.Dahleh, *"Automated Classification of EEG Signal In Brain Tumor Diagnotics",* IEEE Proceedings of Amercian Control Conference, Chiago, June 2000.

[10] Mercedes Cabrerizo, Melvin Ayala, Prasanna Jayakar and Malek Adjouadi, *"Classification and Medical Diagnosis of Scalp EEG Using Artificial Neural Networks",* International Journal of Innovative Computing, Information and Control, vol 7, No12, Dec-2011, pp.6905-6918

[11] A.E.H.Emery, *"Population frequencies of inherited neuromuscular diseases – A world survey",* Neuromuscular Disorders, Vol. 1, No. 1, pp. 19-29, 1991

[12] J.R.Wolpaw, N.Birbaumer, W.J.Heetderks, D.J.McFarland, P.H. Peckham, G.Schalk, E.Donchin, L.A.Quatrano, C.J.Robinson and T. M. Vaughan, *"Brain-computer interface technology: A review of the first international meeting", IEEE Trans. on Rehab. Eng*., Vol. 8, No. 2, pp. 164-173, June 2000.

P.Padmapriya,   A.N.Nithyaa