PORT SAID ENGINEERING RESEARCH JOURNAL
Faculty of Engineering - Port Said University
Volume 17 No 2 September 2013 pp: 90 - 95

An Adaptive Neuro-Fuzzy Interface System for Classifying Sleep EEG

Noha E. El-Kafrawy [1], Fatma E.Z. Abou-Chadi [2] and Sameh E. Rehan [3]

ABSTRACT

In the present paper, classification of sleep stages of EEG by using Adaptive Neuro-Fuzzy. Six sleep EEG records for each of ten patients were selected from Cairo Canter of Sleep Disorder. Three methodologies of analysis were utilized for feature extraction. These include: autoregressive modelling (AR), bispectral analysis, and discrete wavelet transform (DWT), where principle component analysis (PCA) was used to reduce feature dimensionality. The features derived from the three methodologies of feature extraction were used as input feature vectors to the classifier. The classification rates reached are 89.5%, 92% and 90.8% for the AR modelling, the bispectral analysis, and DWT, respectively. To improve the classification accuracy a data fusion at the matching score was utilized. The total classification accuracy reached 94.3%.

KEYWORDS: Autoregressive Modelling, Bispectral Analysis, Discrete Wavelet Transform, Principal Component Analysis, Adaptive Neuro-Fuzzy Interface System.

1. INTRODUCTION

The use of Electroencephalogram (EEG) analysis in the study of sleep and sleep disorder is an ongoing research topic of vital importance to clinicians from wide medical interests [1]. Modern signal processing and pattern recognition techniques were combined in order to characterize important features of sleep EEG signals.

Researchers have developed many techniques to analyse and classify the sleep stages using EEG records. Mendez et al. [2] found that the time-invariant and time-variant parametric models resulted fine tools for diagnosing the obstructive sleep apnea (OSA) . Satoshi Hagihiraet et al. [3] confirmed the methodology of bispectral analysis of EEG, which analyses EEG waveform and calculates bispectrum, bicoherence, and power spectrum of EEG. T.Ning et al [4] investigated the EEG of the rat during various vigilance states to study the degree of Gaussianity of their amplitude distribution and to detect the presence of quadratic phase coupling. Moreover, several important characteristic waves in (EEGs) were reported using Daubechies wavelet by MinSoo Kim et al. [5].

Edgar Oropesa1 et al. [6] used a wavelet packet transformation to provide localized time-frequency information, and they used an artificial neural network for doing optimal classification. The classification results compared to those of a human expert reached 70 to 80% of agreement. Sheng Liu et al. [7] proposed a -robust system that combines multiple signal-processing methods in a multistage scheme: integrated adaptive filtering, wavelet transform and artificial neural network. The detection rate of epileptic events was 90.0%. BrijilChambayil [8] used an artificial Neural Network (ANN) classifier to detect the eye blinks artifacts. The performance of the Cascade-Forward Back-Propagation (CFBP) network is better compared to the Feed-Forward Back-Propagation (FFBP) networks in classifying a signal to an eye blink or not. Inanguler et al. [9] described the application of adaptive Neuro fuzzy interface system model for classification of EEG signals. The total classification accuracy of their model was 98.68%.

The present work aims to compare the performance of three feature extraction techniques for classifying the different stages of sleep EEG using an Adaptive Neuro-Fuzzy Interface system (ANFIS) in order to obtain the highest classification rate. These include: autoregressive modelling (AR), bispectral analysis, and discrete wavelet transform (DWT) in this case, principal component analysis (PCA) was used to reduce feature dimensionality of DWT. The features derived from the three methodologies of signal analysis were used as input feature vectors to the classifier. An Adaptive Neuro-fuzzy Interface system network was used for classification.

2. DATA SET

Ten patients aged between (35-50) years were selected from the Cairo center for sleep disorder to record their sleep EEG signals.EEG signals were recorded bilaterally using SAHC (sleep analyzers hybrid microcomputer system) from 6 p.m to about 3 a.m.

The recordings were made when patient was lying down and relaxed in a quiet room with dim light, where the patient head was fixed in a stereotaxic frame

[1] Department of Electronics and Communication Engineering, Faculty of Engineering, Mansoura University, Mansura, Egypt.
*Email:*nelkafrawy@yahoo.com

[2] Department of Electronics and Communication Engineering, Faculty of Engineering, Mansoura University, Mansura, Egypt.
Email: fzmr4@hotmail.com

[3] Department of Electronics and Communication Engineering, Faculty of Engineering, Mansoura University, Mansura, Egypt.
Email: sameh_rehan@ieee.org

connected with ball-shaped silver electrodes with shielded cable for recording. The standard arrangement of the 10-20 system of electrodes was positioned on the human skull over the central and occipital lobes for the right and left sides. The ear or mastoid were used as recommended references for the electrodes measuring EEG to maximizes inter electrode distance and to avoid mixing activity from two different scalp areas. The sampling rate was 100 Hz. In the present works, six sleep stages are considered. These are stage wake, stage1, stage 2, stage 3, stage 4and stage REM.

3. FEATURE EXTRACTION

Three methodologies of analysis were utilized for feature extraction. These include: autoregressive modelling (AR), bispectral analysis, and discrete wavelet transform (DWT). The principal component analysis (PCA) was used to reduce feature dimensionality

3.1. Autoregressive Modelling (Ar)

AR modelling proceeds by a series of well-defined steps. The first step is to examine the stationarity of sleep EEG signals. The stationarity of the signal was tested using "Run test"[10]. The test results showed that signals which are non-stationary. Best-fit curve of orders ranging from 2 to 4 were applied to remove non-stationary of the sleep EEGs.

The second step is to estimate the coefficients of the AR using the Autocorrelation method [11].

The third step is to select the model order. The model order selection was performed using Yule Walker method for the six sleep stages. It has been found that the optimal model order which gives a minimum ratio between residual signal power and the original signal power is 6.The residual signal is defined as the difference between original and modelled signal.

The fourth step is to test the validity of the model. A Kolmogrov-sminrovone sample test (K-S) test was used to test the randomness of the residual signal [11].

3.2. Bispectral Analysis

For a zero-mean, stationary process *{X (k)}*, the third-order cumulant is defined as the expected value of the triple product [4].

$$R(m, n) = E\{x(k)\, x(k+m)\, x(k+n)\} \quad (1)$$

and the bispectrum is defined as the Fourier transform of the third-order cumulant [TOC] Sequence, i.e.

$$B(w_1, w_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} R(m,n) e^{-j(w_1 m + w_2 n)} \quad (2)$$

If the process *{ X (k) }* is truly Gaussian, then its third cumulants *R(m. n) = 0* for all *(m. n)* and its bispectrum *B(w1, w2)* = 0 for all (*w1, w2*). Consequently, a random process exhibiting a highly nonzero bispectrum can be considered as highly non-Gaussian. Another useful property of the bispectrum we have utilized here is its ability to detect the quadratic phase coupling.

$$D = \sum_{(w_1, w_2)} |B(w_1, w_2)| \quad (3)$$

To quantify the degree of a quadratic phase coupling, the bicoherence index is computed. This index is a function of the bispectrum *B (w1, w2)* and the power spectrum *P (w)* is defined as:

$$bic(w_1, w_2) = \frac{|B(w_1, w_2)|^2}{p(w_1)\, p(w_2)\, p(w_1 + w_2)} \quad (4)$$

Equation (4) denotes that the bicoherence index measures proportion of a specified of energy in each pair of frequencies *w1*and *w2* components of a specified sample of the time series and their sum *w3* that satisfies the definition of quadratic phase coupling , i.e. the phase of *w3* must equal the phase of *w1* and *w2*.

Since bicoherence is the proportion of the coupled to the uncoupled energy, it is normalized (by the product of the *w1* power **w2* power**w3* power) and it is independent of the total energy [4].

The bicoherence index should have significance level exceeds 0.15 to indicate the existence of phase coupling between the harmonic components of the process [4].

In general, it is apparent that the bispectrum involves the phase information that the conventional power spectrum does not involve.

3.3. Discrete Wavelet Transform

Discrete wavelet decomposition was performed for signals of length 4096 from sleep EEG signal for the six sleep stages to the tenth level of resolution using “haar” mother function [13]. For each of the 11 resolution and approximation coefficients, the average energy content of the coefficients at each resolution was computed. There were a total of 9 sub bands from which features were extracted. The i^{th}element of a feature vector was given [13]

$$vi^{dwt} = \frac{1}{ni} \sum_{j=1}^{ni} w_{i,j}^2 \qquad i=1.2\ldots12 \quad (5)$$

Where $n_1 = 2^{11}, n2 = 2^8$, $n3 = 2^7$,…, $n_{12} = 2^0$, v_i^{dwt} is the *ith* feature element in a DWT feature vector; n_i is the number of samples individual sub band; and $w_{i,j}^2$ is the *j th* coefficient of the *ith* sub band. As a result, a DWT feature vector is formed as given by [14]

$$v^{dwt} = \{v1^{dwt}, v2^{dwt}, \ldots..v10^{dwt}\} \quad (6)$$

Principal component analysis (PCA) is one of the most important recent methods of dimensionally reduction, where the representation of the input vector is in terms of eigenvectors is of their covariance matrix [16].

PCA has been applied to DWT coefficient matrices. This reduces the number of features to the first ten principal component of DWT.

4 ADAPTIVE NEURO-FUZZY INTERFACE SYSTEM (ANFIS)

4.1 Architecture Of ANFIS

The ANFIS is a fuzzy Sugeno model put in the framework of adaptive systems to facilitate learning and adaptation [17]. Such framework makes the ANFIS modelling more systematic and less reliant on expert knowledge. To present the ANFIS architecture, two fuzzy if-then rules based on a first order Sugeno model are considered:

Rule 1: If *(x is A1)* and (*y is B1*) then *(f1 = p1x + q1y + r1)*

Rule 2: If (*x is A2*) and *(y is B2*) then *(f2 = p2x + q2y + r2)*

Where x and y are the inputs, A_i and B_i are the fuzzy sets, f_i are the outputs within the fuzzy region specified by the fuzzy rule, the p_i, q_i and r_i are the design parameters that are determined during the training process. The ANFIS architecture to implement these two rules is shown in Fig. 1, in which a circle indicates a fixed node, whereas a square indicates an adaptive node. In the first layer, all the nodes are adaptive nodes. The outputs of layer 1 are the fuzzy membership grade of the inputs, which are given by:

$$o_i^1 = \mu A_i(x) \quad i=1,2 \tag{7}$$

$$o_i^1 = \mu B_{i-2}(y) \quad i=3,4 \tag{8}$$

Where $\mu A_i(x), \mu B_{i-2}(x)$ can adopt any fuzzy membership function. For example, if the bell shaped membership

.

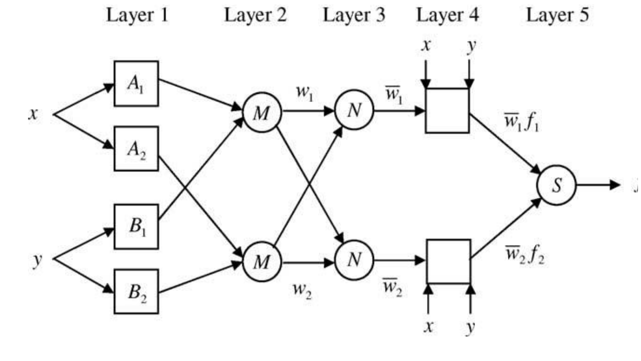


Fig.1.ANFIS architecture

Function is employed, $\mu A_i(x)$ is given by:

$$\mu A_i(x) = \frac{1}{1+\left\{\left(\frac{x-c_i}{a_i}\right)^2\right\}^{b_i}} \tag{9}$$

Where a_i, b_i and c_i are the parameters of the membership function, governing the bell shaped functions accordingly.

In the second layer, the nodes are fixed nodes. They are labeled with *M*, indicating that they perform as a simple multiplier.

The outputs of this layer can be represented as:

$$o_i^2 = \omega_i = \mu A_i(x) \quad i=1,\ 2 \tag{10}$$

Which are the so-called firing strengths of the rules

In the third layer, the nodes are also fixed nodes. They are labeled with *N*, indicating that they play a normalization role to the firing strengths from the previous layer.

The outputs of this layer can be represented as:

$$o_i^3 = \varpi_i = \frac{\omega_i}{\omega_1+\omega_2} \quad i=1,2 \tag{11}$$

Which are the so-called normalized firing strengths.

In the fourth layer, the nodes are adaptive nodes. The output of each node in this layer is simply the product of the normalized firing strength and a first order polynomial (for a first order Sugeno model). Thus, the outputs of this layer are given by:

$$o_i^4 = \varpi_i f_i = \varpi_i(p_i x + q_i y + r_i) \quad i=1,2 \tag{12}$$

In the fifth layer, there is only one single fixed node labeled with *S*. This node performs the summation of all incoming signals. Hence, the overall output of the model is given by:

$$o_i^5 = \sum_{i=1}^{2} \varpi_i f_i = \frac{\sum_{i=1}^{2} \omega_i f_i}{\omega_1+\omega_2} \tag{13}$$

It can be observed that there are two adaptive layers in this ANFIS architecture, namely the first layer and the fourth layer. In the first layer, there are three modifiable parameters *{ai, bi, ci}*, which are related to the input membership functions. These parameters are the so-called premise parameters. In the fourth layer, there are also three modifiable parameters *{pi, qi, ri}*, pertaining to the first order polynomial. These Parameters are so-called consequent parameters [17, 18].

4.2. Learning Algorithm

The task of the learning algorithm for this architecture is to tune all the modifiable parameters, namely $\{a_i, b_i, c_i\}$ and $\{p_i, q_i, r_i\}$, to make the ANFIS output match the training data. When the premise parameters a_i, b_i and c_i of the membership function are fixed, the output of the ANFIS model can be written as [17]:

$$f = \frac{\omega_1}{\omega_1+\omega_2} f_1 + \frac{\omega_2}{\omega_1+\omega_2} f_2 \tag{14}$$

Substituting Eq. (11) into Eq. (14) yields:

$$f = \varpi_1 f_1 + \varpi_2 f_2 \tag{15}$$

Substituting the fuzzy if-then rules into Eq. (15), it becomes:

$$f = \varpi_1(p_1x+q_1y+r_1)+\varpi_2(p_2x+q_2y+r_2) \qquad (16)$$

After rearrangement, the output can be expressed as:

$$f = (\varpi_1 x)p_1 + (\varpi_1 y)q_1 + (\varpi_1)r_1 + (\varpi_2 x)p_2 + (\varpi_2 y)q_2 + (\varpi_2)r_2 \qquad (17)$$

This is a linear combination of the modifiable consequent parameters p1, q1, r1, p2, q2 and r2. The least squares method can be used to identify the optimal values of these parameters easily. When the premise parameters are not fixed, the search space becomes larger and the convergence of the training becomes slower. A hybrid algorithm combining the least squares method and the gradient descent method is adopted to solve this problem. The hybrid algorithm is composed of a forward pass and a backward pass. The least squares method (forward pass) is used to optimize the consequent parameters with the premise parameters fixed. Once the optimal consequent parameters are found, the backward pass starts immediately. The gradient descent method (backward pass) is used to adjust optimally the premise parameters corresponding to the fuzzy sets in the input domain. The output of the ANFIS is calculated by employing the consequent parameters found in the forward pass. The output error is used to adapt the premise parameters by means of a standard back propagation algorithm.

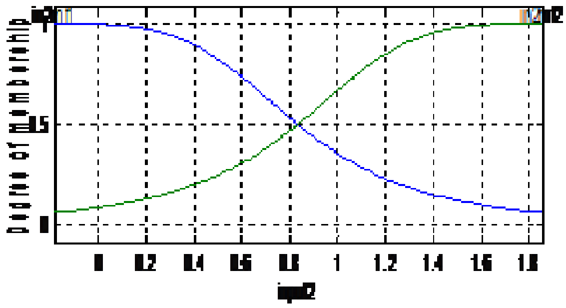
It has been proven that these hybrid algorithms are highly efficient in training the ANFIS [17, 18].

5. CLASSIFICATION RESULTS

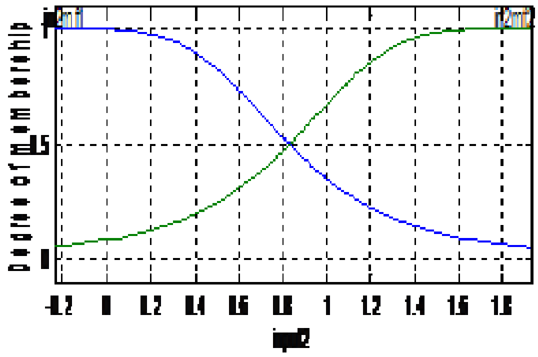
The training and testing patterns were derived from 15 records for each of the 10 patients and using three feature vectors: spectral analysis using AR modelling, the bispectral analysis, and DWT using PCA, separately.

An ANFIS classifier was designed for each set of extracted features and trained using the back-propagation gradient descent method in combination with the least squares method. Samples with target outputs sets for the six sleep EEG stages to be classified were given the binary target values of (0, 0, 0, 0,0, 1),(0, 0, 0,0, 1, 0), (0, 0, 0, 1, 0, 0), (0, 0, 1, 0, 0, 0),(0, 1, 0, 0 ,0, 0, 0) and (1, 0, 0, 0,0, 0) , respectively.

The fuzzy logic process consists of converting each input data to degrees of membership, aggregating all outputs into a single fuzzy set, and then converting fuzzy [17].Membership information into a single output. The data sets were divided into two separate data sets—the training dataset and the testing data set. The hold-out method [18] was used in performance evaluation where 50% of the records were used in the training phase and 50% were used in the testing phase .Each ANFIS used 75 training data for each stage and the step size for parameter adaptation had an initial value of 0.01.

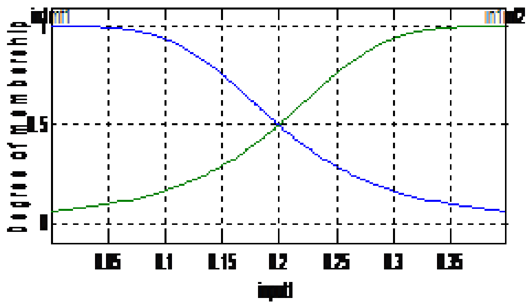


2.(a)

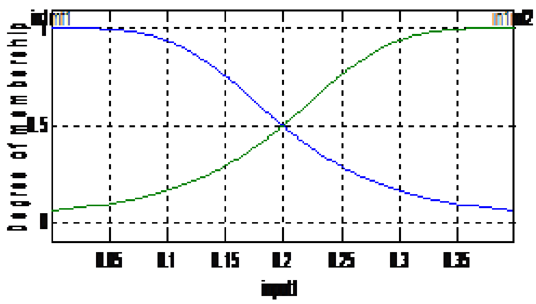


2.(b)

Fig.2.Initial and final generalized bell shaped membership function of second coefficient of AR

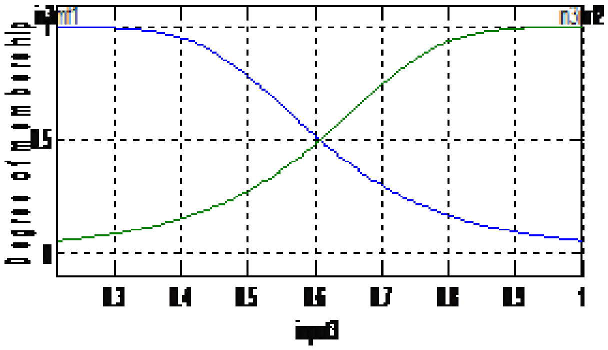


3.(a)

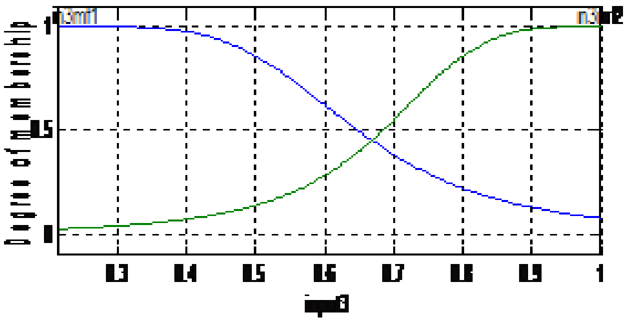


3.(b)

Fig.3: Initial and final generalized bell shaped membership function of Bispectrum

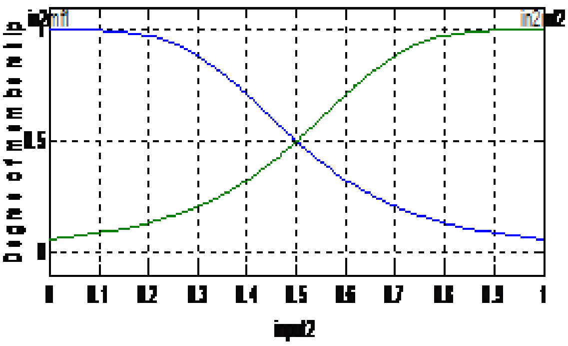


4.(a)

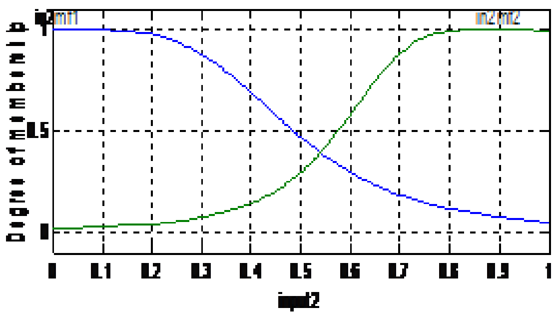


4.(b)

Fig.4: Initial and final generalized bell shaped member ship function of second coefficient of dwt



5.(a)



5.(b)

Fig.5: Initial and final generalized bell shaped member ship function of eighth coefficient number of dwt

Table 1. Classification rates of ANFIS using AR

Sleep stage	Classification accuracy %
Stage wake	90.6%
Stage 1	89.3%
Stage 2	88%
Stage 3	88%
Stage 4	92%
Stage REM	89.3%

Table 2. Classification rates of ANFIS using bispectral

Sleep stage	Classification accuracy %
Stage wake	89.3%
Stage 1	92%
Stage 2	90.6%
Stage 3	93.3%
Stage 4	93.3%
Stage REM	92%

Table 3. Classification rates of ANFIS using DWT

Sleep stage	Classification accuracy %
Stage wake	92%
Stage 1	90.6%
Stage 2	89.3%
Stage 3	92%
Stage 4	90.6%
Stage REM	90.6%

To improve the classification results, a data fusion at the matching score level was adopted. A normalization step is generally necessary before the raw scores originating from different matchers can be combined in the fusion stage [20].

A raw matcher score will be denoted as *s* from the set of *S* of all scores for that matcher and the corresponding normalized score as *n*. The Min-Max (MM) method maps the new scores to the [0, 1] range. Max(S) and Min (S) specify the end points of the score range as follows [20]:

$$n = \frac{s - \min(S)}{\max(S) - \min(S)} \tag{18}$$

In general, there are different fusion methods such as simple sum (ss), min max (MM), max score (MS) and matcher Weighting (MW) [20]. It has been found that the simple sum (ss) gives the highest classification rate. Let n_i^m represents the normalized score for the matcher *m* *(m=1,2, ...,M)* where *M* is the number of different matchers) and for the *user i (i=1,2, ...)* and *I* is the number of segments in the database. The fused score is denoted as f_1 Simple Sum (SS) Scores for an individual are summed [20]:

$$f_i = \sum_{m=1}^{M} n_i^m . \forall i \tag{19}$$

Table 4 depicts the results of using the data fusion approach for the score of the three ANFIS models.

Table 4. Percentage of correct classification using fusion

Sleep Stages	SS
stage wake	93.89%
stage1	93.6%
stage 2	92.9%
stage3	97.25%
stage4	94.6%
stage REM	93.89%

6. CONCLUSION

In this work a new ANFIS is introduced for the classification of six different stages of sleep EEG signals; stage wake, stage 1, stage 2, stage 3, stage 4 and stage REM. Three feature extraction techniques were used: autoregressive modelling (AR), bispectral analysis, and discrete wavelet transform DWT. The principal component analysis (PCA) was used to reduce feature dimensionality. An Adaptive Neuro-Fuzzy Interface System (ANFIS) classifier was used to classify the six stages of sleep EEG.

The highest correct classification rate obtained from the Adaptive Neuro-Fuzzy Interface System is 92% using features derived from Bispectral analysis.

In order to improve the accuracy of the classification results obtained, a data fusion approach at the matching score was adopted. It has been found that the simple sum (ss) gives the highest classification rate. The total classification accuracy reached 94.3%.

REFERENCES

[1] M. Akay. Biomedicalsignal processing. Academic Press Inc., 1994.

[2] M.O. Mendez, O.P. Villantieri, A.M. Bianchi, and S. Cerutti, "Sleep Analysis forwearable Devices Applying Autoregressive Parametric Models", Politecnico di Milano, Italia Proceedings of the 2005.

[3] Satoshi Hagihira, Masaki Takashina, Takahiko Mori, Takashi Mashimo, and IkutoYoshiya, "Practical Issues in Bispectral Analysis of Electroencephalo-graphic Signals", Technical Communication, June 15 2001.

[4]T. Ning and J. Bronzino, "Bispectral analysis of the rat EEG during Various Vigilance States", IEEE Trans. Biomed. Eng., Vol. 36, No. 4, 1989.

[5]Min Soo Kim, Young Chang Cho, Abibullaev Berdakh and Hee Don Seo, "Analysis of Brain Function and Classification of Sleep Stage EEG using Daubechies Wavelet", Sensors and Materials, Vol. 20, No. 1, 2008.

[6]Edgar Oropesa , Hans L. Cycon , Marc Jobert, " Sleep Stage Classification usingWavelet Transform and Neural Network", International Computer Science Institute, March 30, 1999.

[7] He Sheng Liu, Tong Zhang, and Fu Sheng Yang, "A Multistage, Multimethod Approach for Automatic Detection and Classification of Epileptiform EEG", IEEE Transaction on Biomedical Engineering, Vol. 49, NO. 12, December 2002.

[8]BrijilChambayil, Rajesh Singla, R. Jha, "EEG Eye Blink Classification Using Neural Network", Proceedings of the World Congress on Engineering 2010, Vol. IWCE 2010, June 30 - July 2, 2010, London, U.K.

[9] InanGuler, ElifDeryaYbeyli." Adaptive neuro Fuzzy interface system for classification of EEG signals using wavelet coefficients", Journal of NeuroscinceMethods (2005).

[10] J.S.M. Bendant and piersol. Random data: analysis and measurement procedures.New York: Willy, 1985.

[11] S. Seigel. Non-parametric statistics for behavioral sciences. McGraw-Hill, 1956.

[12]D.R. Brillinger, "Some basic aspects and uses of higher order spectra", IEEE sig. Proc., Vol. 36, pp. 239-249, 1994.

[13] StephaneMallat. A Wavelet Tour of Signal Academic Press, dec. 2008.

[14] R.M. Rao and A.S. Bopardikar. Wavelet Transforms: Introduction to theory,Algorithms and applications. Wesley Longman, Inc., pp.1-30, 1998.

[15] Pablo Faundez Hoffman, Alvaro Fuentz, SigitasDagilis. K-complex detection usingcontinuous wavelet transforms. Department of control Engineering, AalborgUniversity,December 2000.

[16] R. Schukoff. "Pattern Recognition: Statistical, Structural, and Neural Approaches".John Wiley and Sons, November 1992.

[17] Jang J-SR. "Self-learning fuzzy controllers based on temporal back propagation" .IEEE Trans Neural New 1992;3(5):714–23.

[18] Jang J-SR, "ANFIS: Adaptive-network-based fuzzy inference system for classification of EEG signals using wavelet coefficients", IEEE Trans Syst Man Cybern 1993; 23(3):665–85.

[19] Arun Ross and Anil Jain. "Score normalization In Multimodal Biometrics System", Michigan State University, East Lansing. 21 october 2004.

[20] M. Indovina, U. Uladag, R. Sneliek, A. Mink, A. jain. "Multimodel Biometric Authentication methods: A COTS Approach". " Proc. MMUA 2003, Workshop on Multimodal User Authentication, pp. 99-106, Santa Barbara, CA, December 11-12, 2003.