2003 International Symposium on Advanced Intelligent Systems
September 25-28, 2003, Jeju, Korea

Feature extraction and Classification of EEG for BCI system

EungSoo Kim[1], HanBum Cho[2], EunJoo Yang[2], TaeWan Eum[2]
[1] Department of Computer & Communication Engineering, Daejeon University, Daejeon, Korea
[2] Department of Electronics Engineering, Graduate School, Daejeon University, Daejeon, Korea
email:eskim@dju.ac.kr

***Abstract* – EEG is an electrical signal, which occurs during information processing in the brain. These EEG signals has been used clinically, but nowadays we are mainly studying Brain-Computer Interface(BCI) such as interfacing with a computer through the EEG, controlling the machine through the EEG. The ultimate purpose of BCI study is specifying the EEG at various mental states so as to control the computer and machine. A BCI has to perform two tasks, the parameter estimation task, which attemps to describe the properties of the EEG signal and the classification task, which separates the different EEG patterns based on the estimated parameters. First, we have to do parameter estimation of EEG to embody BCI system. It is important to improve performance of classifier. But, It is not easy to do parameter estimation by reason of EEG is sensitivity and undergo various influences. Therefore, this research should do parameter estimation and classification of the EEG to use various analysis algorithm.**

I. INTRODUCTION

The EEG records measuring electrical change that is accompanied in activity of brain cell from outside and it is the most important indicator pointer that measure activity state of brain. The EEG includes useful information about brain activity and can be measured by non-invasive method. Therefore these EEG has been used research about function of brain and clinical etc., and recently that is used in research of an EEG-based BCI.

General architecture of the EEG based on BCI is as follow. First, Fix electrode on user's specification scalp surface. There, the signals – which simply are very small voltage potentials –are amplified and sent to a computer via an analog-digital (A/D) converter. And then, acquired data feature extraction and classifies applying various algorithm. After classified data is translated into appropriate commands by the computer, it is applied variously in simple TV On/Off, control of computer cursor, word processor etc. That is, BCI consists of recursive architecture of data acquisition module, signal processing module, target application module. The most important part in BCI's architecture is signal processing module that extract feature of the EEG and classifies. It is very difficult to classify and find feature because EEG does not generate signal clearly that is distinguished by mental state.

The Graz BCI has been based on the detection of the ERD and the ERS patterns during the motor imagery [5][6]. And the method of Jonathan Wolpaw and his colleagues is based on the self-regulation of the μ (8-12Hz) rhythms or the β (13-28Hz) rhythms[7]-[11].

In this study, we used EEG included facial muscle that can generate easily by short training time. And we did feature extraction using power spectrum and PCA, AR model to design signal processing module that can confide. Also, it compared each recognition rate to use linear discriminant analysis (LDA) and Multilayer perceptrons (MLP), Radial basis function networks (RBN).

II. METHODS

A. EEG Feature Extraction

The purpose of the feature extraction is to transform the information from an EEG device to more meaningful form for the classifier. For the analysis of oscillatory EEG components, we investigated the following feature extraction methods:

1) Power Spectrum

When the DFT is used to analyze a signal, one is often concerned not so much with the amplitude and phase of the signal's spectrum as with either its power or its magnitude. The power spectrum is given by:

$$Power[n] = (\text{Re}\ X[n])^2 + (\text{Im}\ X[n])^2 \qquad (1)$$

The power spectrum estimate was calculated using two different methods: The Welch method and Blackman-Tukey method. The welch method is a modified periodogram method where the inconsistency of the periodogram method, that the increasing the number of samples does not reduce variance, is downgraded by averaging overlapping consecutive samples of the signal. In Blackman-Tukey method the power spectrum of the signal is calculated from the discrete Fourier transform of the autocorrelation function of the data. Here the smoothing effect achieved from the autocorrelation function rather than from the averaged periodograms, therefore having better spatial resolution. In this method windowing has extra importance because at larger lags fewer data points are available for computation so those estimates are less accurate. Windowing emphasizes shorter lags, thus giving them greater weight when calculating spectrum estimate.

2) AR model

The autoregressive model is one of a group of linear prediction formulas that attempt to predict an output y[n] of a system based on the previous outputs (y[n-1],y[n-2]...) and inputs (x[n], x[n-1], x[n-2]...). Deriving the linear prediction model involves determining the coeffiecients a1,a2,.. and b0,b1,b2,... in the equation:

$$ye[n]\ (\text{estimated}) = a1*y[n-1] + a2*y[n-2]... + b0*x[n] + b1*x[n-1] + ... \quad (2)$$

Note the remarkable similarity between the prediction formula and the difference equation used to describe discrete linear time invariant systems. Calculating a set of coefficients that give a good prediciton ye[n] is tantamount to determining what the system is, within the constraints of the order chosen. A model which depends only on the previous outputs of the system is called an autoregressive model (AR).

3) Principal Components Analysis (PCA)

Principal component analysis is a linear procedure to find the direction in input space where most of the energy of the input lies. In other words, PCA performs feature extraction. The projections of these components correspond to the eigenvalues of the input covariance matrix. Principal component analysis(PCA) is a statistical technique falling under the general title of factor analysis. The purpose of PCA is to identify the dependence structure behind a multivariate stochastic observation in order to obtain a compact description of it. When applied to m-dimensional data set X, it performs forward and backward mapping with linear transforms,

$$V = W^T X \quad (3)$$

$$X^* = WV \quad (4)$$

where $W = [w_1, w_2, \Lambda, w_P]$,is the linear transform, V is a p-dimensional feature vector representation of X, X^* is the reconstructed X . If vectors of W are chosen to be the p eigenvectors corresponding to p largest eigenvalues of $X^T X$,then the approximation error $\|X - WW^T X\|$ will be minimized.

B. Classification

1) Multilayer perceptrons (MLP)

Multilayer perceptrons (MLPs) are feedforward neural networks trained with the standard backpropagation algorithm. They are supervised networks so they require a desired response to be trained. They learn how to transform input data into a desired response, so they are widely used for pattern classification. With one or two hidden layers, they can approximate virtually any input-output map. They have been shown to approximate the performance of optimal statistical classifiers in difficult problems. Most neural network applications involve MLPs.

2) Linear discriminant analysis (LDA)

Discriminant analysis is a technique for classifying a set of observations into predefined classes. The purpose is to determine the class of an observation based on a set of variables known as predictors or input variables. The model is built based on a set of observations for which the classes are known. This set of observations is sometimes referred to as the training set. Based on the training set , the technique constructs a set of linear functions of the predictors, known as discriminant functions, such that L = b1x1 + b2x2 + … + bnxn + c , where the b's are discriminant coefficients, the x's are the input variables or predictors and c is a constant.

These discriminant functions are used to predict the class of a new observation with unknown class. For a k class problem k discriminant functions are constructed. Given a new observation, all the k discriminant functions are evaluated and the observation is assigned to class i if the ith discriminant function has the highest value.

3) Radial basis function networks (RBN)

Radial basis function (RBF) networks have a static Gaussian function as the nonlinearity for the hidden layer processing elements. The Gaussian function responds only to a small region of the input space where the Gaussian is centered. The key to a successful implementation of these networks is to find suitable centers for the Gaussian functions. This can be done with supervised learning, but an unsupervised approach usually produces better results.

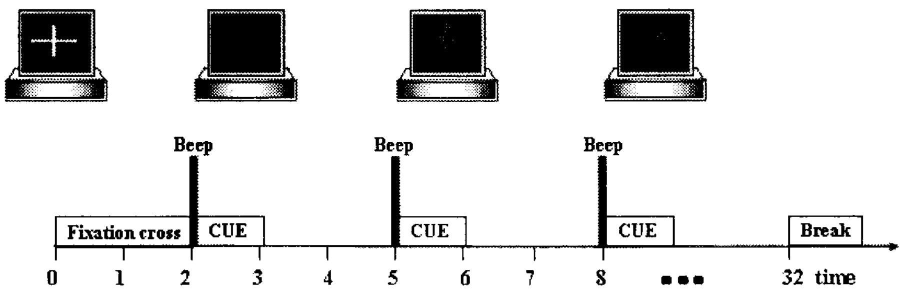

Figure 1. Scheme of the acquisition paradigm.

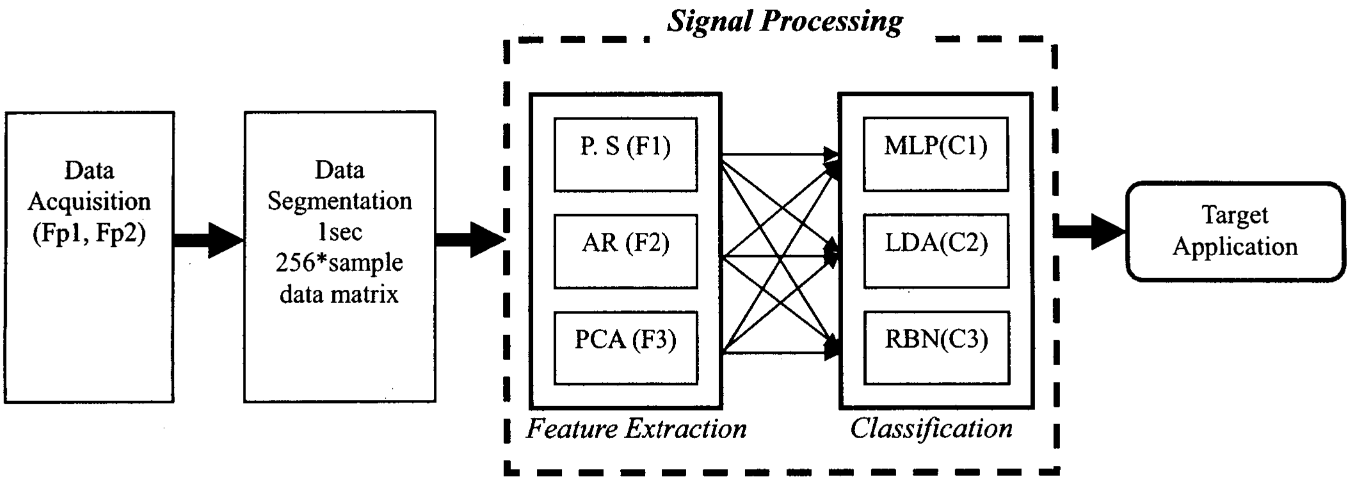

Figure 2. Signal Processing module of system Architecture

III. EXPERIMENTS

Twenty two subjects participated in this research. Two EEG channels were recorded using electrode positions Fp1 and Fp2 according to the international 10-20 system (Figure 3). We set up sampling frequency 256Hz, sensitivity $7\mu V/mm$, high frequency filter 60Hz, low frequency filter 1Hz.

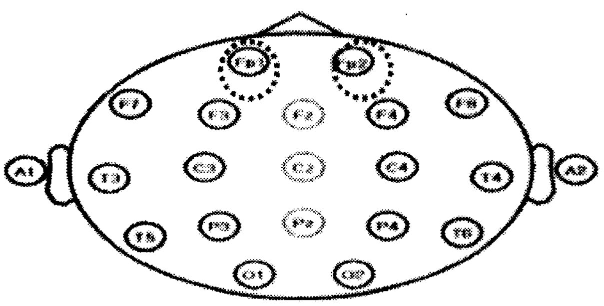

Figure 3. Electrode position

The subjects were sitting in a comfortable armchair looking at the center of a monitor placed approximately 1.5 m in front of them. We did to see the monitor that explain detailed contents about an experiment to subject during 3 minutes before measure. Each trial started with the presentation of a fixation cross at the center of the monitor(Fig. 2). After 2 second, arrow was displayed at the center of the monitor with short warning tone. And then subject generate facial muscle. Subjects generate facial muscle during each time to 3 seconds interval. 10 sample data are measured in one trial.

Figure 4 is the EEG of normal state and facial muscle state. Measured data cut to 256 datas for 1 second and then classification recognition rate is compared MLP, LDA, RBN after get feature input vector using power spectrum, AR Model, PCA as figure 2. F1, F2, F3 speak as each Power Spectrum, AR model, Principal component analysis. C1, C2, C3 are each multilayer perceptrons, linear discriminant analysis, radial basis function networks. Signal processing module composed F1 - C1, F1 - C2, F1 - C3, F2 - C1, F2 - C2, F2 - C3, F3 - C1, F3 - C2, F3-C3. In the power spectrum, the input vector were extracted distribution(0~128Hz) of each frequency component that compose EEG. Degree of AR model chose by $p = 20$ in $N/20 < p < N/5$ range(N are data number). The autoregressive coefficient is used by input vector. In the principal component analysis(PCA), input vector were principal component of minimum number to have information of EEG data. Input vector that occupy 10% of whole among PCA pattern was used. Each feature vector used in input of MLP, LDA, RBN. MLP used Two Hidden Layer. Training function is thing to decide connection weight's update direction. That used resilient backpropagation (Trainrp).

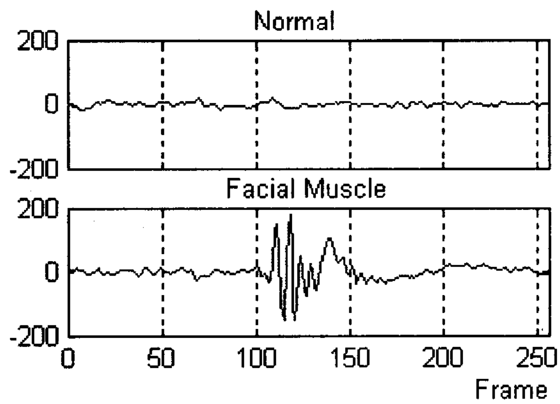

Figure 4. EEG in general state and EEG with facial muscle

IV. RESULTS

The different methods of EEG preprocessing and classification is compared. First, we display the recognition rate of MLP in Table I. The recognition rate is average of 22 subjects in Table I. We changed numbers of hidden layer and were trained. Variation by number of hidden layer did not appear. The recognition rate is highest after do power spectrum.

In the case of F1-C1, F2-C1, F3-C1, we got the high recognition rate more than 98%. In the case apply LDA and RBN, recognition rate of F2 (feature extraction method 2, AR) is high more than power spectrum and recognition rate of F3 (feature extraction method 3, PCA) is low. AR model among feature extraction method was better than other. MLP of classification method was the best. F1 -C1 among each signal processing module displayed the highest recognition rate.

Table I. Classification accuracy of MLP

Recognition Rate	Two hidden layer				
	15-10	15-5	15-3	10-5	10-3
AR	99	98.8	98.6	98.5	98.6
PCA	99.8	99.6	99.5	99.8	99.2
P.S	100	100	100	99.8	99.5

Table II. Result of classification accuracy (performance in %) that use various signal processing

Recognition Rate		Feature Extraction		
		P.S(F1)	AR(F2)	PCA(F3)
Classifier	MLP(C1)	99.87	98.72	99.61
	LDA(C2)	93.44	95.56	60.45
	RBN(C3)	79.89	93.56	70.44

V. CONCLUSION

In this research, when use the EEG that included facial muscle in BCI system, we compared the recognition rate to find the most suitable signal processing module using various algorithm. We construct nine kinds of signal processing module. And then recognition rate of each module is compared. The result of comparison, we got high recognition rate of 99.87% that classified by MLP after feature extraction using power spectrum. High recognition rate is displayed 92% except C2 - F3, C3 - F1, C3 - F3 generally. Because EEG with facial muscle signal is divided enough visually and only distinguish with the EEG of general state. We could acquire appropriate feature extraction method and combination of algorithm in same condition from result of this research.

REFERENCES

[1] Tommi Nykopp, Statistical Modelling Issues for The Adaptive Brain Interface, 2001

[2] Janne Lehtonen, EEG-based Brain Computer Interface, 2002

[3] Jessica D. Bayliss, A Flexible Brain-Computer Interface. 2001

[4] Cristoph Guger et al, "Rapid prototyping of an eeg-based brain-computer interface(bci)." IEEE Transactions on Rehabilitation Engineering, 9(1):49-58, 2001

[5] J.R Wolpaw. "An eeg-based brain-computer interface for cursor control." Electroencepephalogaphy and Clinical Neurophysiology, 1994

[6] J. R. Wolpaw, D.J.McFarland, and T.M. Vaughan "Brain-Computer Interface Research at the Wadsworth Center", IEEE Trans. on Rehabilitation Engineering, Vol. 8, No .2, June 2000.

[7] C. Guger, H. Ramoser, G. Pfurtscheller, "Real-Time EEG Analysis with Subject-Specific Spatial Patterns for a Brain-Computer Interface (BCI)" IEEE Trans. on Rehabilitation Engineering, Vol. 8, No. 4, 2000.

[8] Pfurtscheller G, Woertz M, Krausz G, Neuper C. : "Distinction of different fingers by the frequency of stimulus induced beta oscillations in the human EEG", Neurosci Lett. 2001; 307: 49-52

[9] G. Pfurtscheller and C. Neuper, "Motor Imagery and Direct Brain-Computer Communication" Proceeding of the IEEE, Vol. 89, No..7, July 2001.

[10]G.Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser. Schlögl, B. Obermaier, and M. Pregenzer, "Current Trends in Graz Brain–Computer Interface (BCI) Research", IEEE Vol. 8, No. 2, June 2000

[11] G. Pfurtscheller, J.Kalcher, Ch. Neuper, D.Flotzinger, M.Pregenzer, "On-line EEG classification during externally-paced hand movements using a neural network-based classifier", Electroencephalography and clinical Neurophysiology 99(1996) 416-425

[12] T. Felzer and B. Freisleben, "HaWCoS : The "hands-free" wheelchair control system", In ASSETS 2002 – Proceeding ACM SIGCAPH Conference on Assistive Technologies, Edinburgh, Scotland, 2002, ACM Press.

[13] T.Felzer and B.Freisleben, "BRAINLINK : A software tool upporting the development of and EEG-based brain-computer interface" , submitted for publication, 2002.

[14] T.Felzer and B.Frisleben, "An input device for human-computer interface based on muscle control" , submitted for publication, 2001

[15] C. W. Anderson, S.V. devulapalli, and E.A. Stolz, "Determining mental state from EEG signals using neural networks", Scientific Programming-Special Issue on Applications Analysis, vol. 4, no. 3. pp. 171-183, 1995.

[16] J. J. Tecce, J. Gips, C. P. Olivieri, L. J. Pok, and M. R. Consiglio, "Eye movement control of computer functions", International Journal of Psychophysiology, vol. 29, pp. 319-325, 1998.

[17] K. S. Park and K. T. Lee, "Eye-controlled human/computer interface using the line-of-sight and the intentional blink", Computers & Industrial Engineering, vol. 30, no. 3, pp. 463-473, 1996.

[18] Schlögl A., Neuper C. Pfurtscheller G. : "Estimating the mutual information of an EEG-based Brain-Computer-Interface", Biomedizinische Technik 47(1-2): 3-8, 2002.