INTERNATIONAL SCIENTIFIC AND VOCATIONAL JOURNAL (ISVOS JOURNAL)

Vol.: 5 Issue: 1 Date: 30.06.2021 Received: 22.12.2020 Accepted: 01.02.2021 Final Version: 11.06.2021

ISVOS Journal, 2021, 5(1): 1-6

The Role of Feature Selection in Significant Information Extraction from EEG Signals

Eda Dağdevir [a], Mahmut Tokmakçı [b,1]

[a] Kayseri University, Vocational School of Technical Sciences, Department of Biomedical Device Technologies, Kayseri 38280, Turkey
ORCID ID: 0000-0001-7065-9829
[b] Erciyes University, Faculty of Engineering, Department of Biomedical Engineering, Kayseri 38039, Turkey
ORCID ID: 0000-0001-5786-7359

Abstract

Information extraction from EEG signals for use in Brain Machine Interface systems has been a highly effective research topic recently. Due to the complexity, high dimension, and subject specific behavior of the EEG signals make feature extraction and selection very important. For this reason, there are many studies in the direction of feature extraction and selection which affect the performance of the Brain Machine Interface system at a high level. In this study, different statistical characteristics were obtained from wavelet coefficients obtained by wavelet transform by using BCI Competition IV-2b data set. The selection of the efficient ones of these features is provided by Principal Component Analysis. The fitness of logistic regression model established with both feature groups was measured by Akaike Information Criteria. The results indicated that relatively better statistical performance can be obtained by using fewer features thanks to PCA. These results are important in terms of statistical comparison and demonstration of the success in extracting information from EEG signals.

Keywords: “EEG, Brain Machine Interface, Feature Extraction, Feature Selection, Akaike Information Criteria”

1. Introduction

Brain Machine Interface (BMI) is a system that enables human and machine interaction for establish a connection between the brain and an external mechanism [1]. BMI are systems that enable people to use a computer, an electromechanical arm or various neuroprosthesis without using their motor nervous system. The information created by the communication of a large number of neurons with each other can be recorded with electrodes that will be placed in different parts of the brain. BMI systems for chronic neuromuscular disorders, amyotrophic lateral sclerosis (ALS), stroke, high-level spinal cord injury, motor or tetraplegia patients to control a vehicle with Electroencephalogram (EEG) signals taken from the relevant part of their brain are becoming increasingly widespread [2]. These systems, called motor imagery-based BMI (MI-BMI) systems, should have in excellent performance in terms of real-time data processing capability. The EEG method is the most commonly used method as the signal collection method in MI-BMI systems due to its easy applicability and non-invasiveness. In studies using MI-BMI systems, mu (8-13 Hz) and beta (13-30 Hz) rhythms are widely researched due to their high temporal resolution and ability to identify mental tasks associated with different movement [3].

Signal processing algorithms developed for use in MI-BMI systems are a very common research area. The signal processing procedure to be used for MI-BMI systems consists of five main steps. These can be listed as signal acquisition, preprocessing, feature extraction, feature selection and classification [4]. In this study, public access MI-BMI data set was used. The records in the data set, whose details can be accessed from [5] were preprocessed with 0.5 Hz-100 Hz Butterworth filter and 50 Hz Notch filter. From the wavelet coefficients obtained by wavelet transform from these filtered records; Statistical features such as mean absolute value, mean square root, standard deviation and variance were obtained. Principal Component Analysis (PCA) was used to determine those that better represent the data among all the features obtained separately for each subject. The performance of the logistic regression model established with all features and the features reduced by PCA was examined. Also, the fit of the models was compared with the Akaike Information Criterion [6].

[1] Corresponding Author. Tel.: +0-352-207-6666 ; fax: +0-352-437-5784 .
E-posta adresi: tokmakci@erciyes.edu.tr

2. Material and Methods

2.1. Data Set

The data set used in this study was collected by “Graz University of Technology” within the scope of BCI Competition IV [5]. Of the records taken from 9 people, 2 sessions without feedback were used. 10 trial records were taken from each subject for 6 times in each session. Records were used for a total of 120 trials. 3-channel bipolar EEG signal was taken with the help of Ag-AgCl electrode. For these recordings, channels named C_3, C_z and C_4 were used according to the 10-20 electrode placement system, while F_z was used as the ground electrode. Since 3 EEG records are used, there are 360 trial records for each of the 18 subjects. From these records, in 180 trials, subjects were given a visual stimulus in the form of a right arrow to imagine raising their right hand, and a visual stimulus in the form of a left arrow to imagine raising their left hand for the remaining 180 trials. Recordings with a sampling frequency of 250 Hz were applied a 0.5 Hz-100 Hz Butterworth filter and a 50 Hz Notch filter.

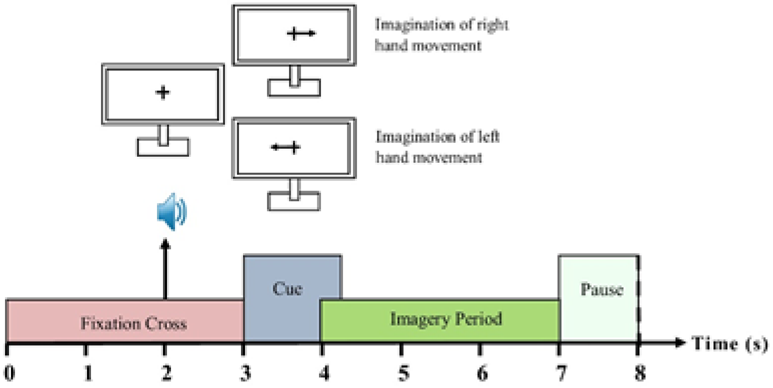
Imagination of right hand movement
Imagination of left hand movement
Fixation Cross
Cue
Imagery Period
Pause
Time (s)
0
1
2
3
4
5
6
7
8

Figure 1. Experimental paradigm

In the experiment procedure, as can be seen in Figure 1, each trial starts with the presentation of the auditory stimulus (1 kHz, 70ms) simultaneously with the appearance of the plus signal on the screen. Then, the right or left arrow mark seen in Figure.1 is randomly displayed for 1.25s. The subject is expected to imagine raising his hand in the direction of the arrow indicated for 4 s. After each trial, there is a break of at least 1.5 s. This break is changed up to 1s to prevent focus. In the trial records taken from each subject, randomly 60 of them are shown right arrow and 60 of them left arrow. Figure 2 shows the 40th trial record of the subject numbered B01, which is known to show the left arrow signal. Here, the 1st second represents the moment when the auditory stimulus is given. Between 2 and 3.5 seconds, it is the time period in which the arrow is displayed on the screen. The subject is asked to imagine raising his hand in the direction indicated by the arrow from the second to the end of the experiment.

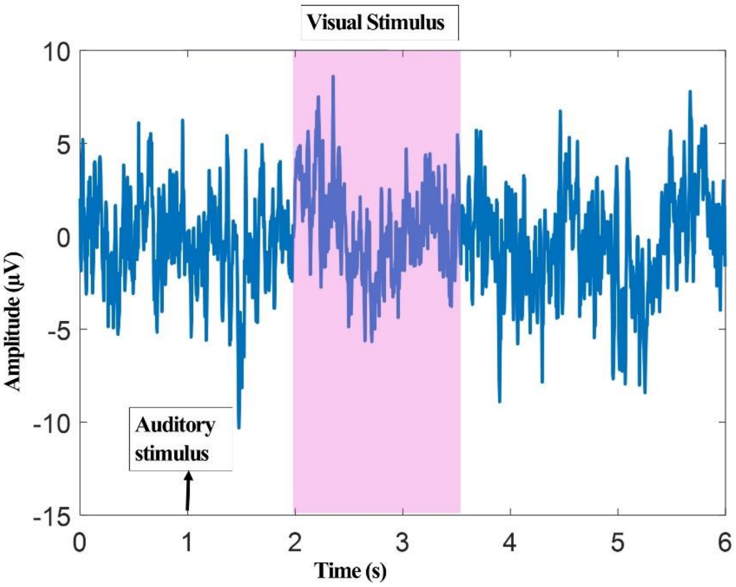

Figure 2. A sample of EEG signal during the experiment

2.2. Preprocessing

The recordings filtered with 0.5 Hz-100 Hz Butterworth filter and 50 Hz Notch filter were analyzed in wavelet transform. It is very important to use the appropriate wavelet function and decomposition level in order to obtain efficient features. Although wavelet families such as Daubechies 2-10, Biorthogonal 2.2, 3.3, 4.4, Coiflets 1-5, Morlet and Mexican Hat are used for EEG signal decomposition to classify left and right hand movement, Daubechies 4 wavelet providing the best kappa value was preferred in this study. The level of decomposition was selected as 7, taking into account the maximum decomposition level given in Equation (1) [7].

$$S = int(log_2 N) \tag{1}$$

where, S is the maximum decomposition level, N is the time series length of signals, int (.) is rounding an integer. Thus, each trial of the EEG signal of each subject is divided into the detailed sub-band D1-D7 and the approximate sub-band A7.

2.3. Feature Extraction

Utilizing statistical features during information extraction from EEG signals is a frequently used approach [8],[9]. At this stage, instead of using the wide frequency band, three sub-bands containing frequencies belonging to the mu and beta band which are frequently used to extract information from MI-BMI systems were used. Here, detailed sub-bands covering mu and beta frequencies are D4, D5, D6. Statistical features are extracted for each of these sub-bands. These are four different statistical properties: mean absolute value, mean square root, standard deviation and variance [3].

2.4. Feature Selection

Principal Component Analysis (PCA) was used for feature selection. PCA provides reduced size of linear feature spaces using a statistical approach [8],[9]. PCA is the arrangement of the axes of the data to keep the high dimensional inter data variance at the highest level. The component with the highest variance value is selected as the main component and the other components are ranked in descending order according to the variance [10]. In this study, PCA was applied to the feature vector obtained by feature extraction and the features were reduced by selecting the features that explain the data with 95% variance.

2.5. Classification

Feature vectors applied and not applied PCA were classified by Logistic Regression method separately for each subject, and the classification performance of the methods was recorded [11].

3. Results and Discussion

Since the number of statistical features extracted from the Wavelet coefficients obtained using the data set is 4 and the number of EEG channels is 3, the sizes of the feature vector obtained at the end of the feature extraction process are (720x13) for the B01, the B02, the B03, the B06, the B07and the B09 subjects, (780x13) for the B04 and the B05 subjects and, (840x13) for the B08 subject. Here, 720,780 and 840 are the number of trials that the subject imagined raising his right or left hand, giving the row size of feature matrix. The last column of the feature vector contains the right or left directional behavior information in the relevant experiment. It is indicated by 1 if the subject imagined of raising his right hand in the relevant trial, and 0 if he imagined of raising his left hand. Among these features, the number of columns for the feature vector formed by selecting the features that explain the EEG data with 95% variance with the PCA method is 6. Table 1 shows the accuracy values of the logistic regression model. Table 2 shows the logistic regression model kappa values. The accuracy values given in Table 1 are calculated according to Equation (2) [12].

$$B = 1 - \frac{\sum_{d=1}^{D} |t[d] - y[d]|}{D} \tag{2}$$

where, B is classification accuracy criterion, D is the total number of trials for each subject, t [d] is the predicted response of the logistic regression model, and y [d] shows the actual behavioral information, respectively. y [d] is denoted by 1 if the subject imagined of raising his right hand in the relevant trial, and 0 if he imagined of raising his left hand. The kappa values given in Table 2 are calculated according to Equation (3).

$$K = \frac{B - B_0}{1 - B_0} \tag{3}$$

where, K shows the kappa value criterion, B shows the classification accuracy calculated for the established classification model and B_0 shows the expected accuracy value, respectively. For a two-class problem, the expected accuracy value is 0.5. These results show that when the real-time data processing procedure is considered, it is possible to extract information from EEG signals with relatively close performance by reducing the number of features. In addition, the Akaike Information Criterion [6] values of logistic regression models in which all features are used together and features reduced by PCA are given in Table 3.

Accordingly, in terms of Akaike Information Criterion [6], there is no significant difference between the model in which all features are used and the model using reduced features with PCA. However, the Akaike Information Criterion values of the features using PCA are relatively low. This situation shows that the model in which the features reduced by PCA analysis is used is a relatively more suitable model. The lower the AIC value, the better the logistic regression parameters established represent the EEG data. Akaike Information Criterion values are calculated as given Equation (4).

$$AIC = 2\log\big(likelihood(left, right)\big) + 2k + \frac{2k(k+1)}{D-k-1} \tag{4}$$

where, $likelihood(left, rigth)$ is likelihood function of regression parameters, k is the number of predicted logistic regression parameters, D is total number of trials.

Table 1. Comparison of the Accuracy Criteria

Subject	All of Features		Features of reduced with PCA	
	Accuracy	Number of Features	Accuracy	Number of Features
B01	56	12	55.72	4
B02	62.78	12	61.97	4
B03	56.53	12	57.67	3
B04	59.36	12	59.46	4
B05	55.64	12	56.13	4
B06	58.89	12	58.64	4
B07	58.75	12	56.42	4
B08	58.33	12	57.55	3
B09	59.86	12	55.31	3
Mean	58.46	12	57.65	~ 4

Table 2. Comparison of the Kappa Value Criteria

Subject	All of Features		Features of reduced with PCA	
	Kappa	Number of Features	Kappa	Number of Features
B01	0.12	12	0.11	4
B02	0.26	12	0.24	4
B03	0.13	12	0.15	3
B04	0.19	12	0.19	4
B05	0.11	12	0.12	4
B06	0.18	12	0.17	4
B07	0.18	12	0.13	4
B08	0.17	12	0.15	3
B09	0.20	12	0.11	3
Mean	0.17	12	0.15	~ 4

Table 3. Comparison of Akaike Information Criteria Values

Subject	All of Features	Features of reduced with PCA	Difference
	Akaike Information Criteria Values	Akaike Information Criteria Values	
B01	1000	993	7
B02	970	965	5
B03	1008	999	9
B04	1078	1070	8
B05	1093	1084	9
B06	985	981	4
B07	997	994	3
B08	1166	1157	9
B09	1000	998	2

4. Conclusions

In this study, it is aimed to extract information with similar performance by using fewer features in order to speed up BMI systems. For this purpose, statistical features were obtained by using Wavelet coefficients and using all trials separately for each subject. PCA has been used to reduce these attributes. Performance criteria were calculated by modeling all features and the feature reduced with PCA using logistic regression. In addition, Akaike Information Criteria values were calculated and compared for both cases. All these results show that efficient information can be extracted from EEG signals with similar performance by using fewer features. In addition, higher performance can be obtained by using other feature reduction methods in the literature instead of PCA. In future studies, the present study will be developed with different feature matrices and different feature selection methods. The current study is also a positive step towards accelerating BMI systems.

Acknowledgement

The authors acknowledge the financial, laboratory and infrastructure support provided by the Scientific Research Projects Coordination Unit, Erciyes University, Turkey (FDK-2020-9876).

References

[1] P. K. Pattnaik and J. Sarraf, “Brain Computer Interface issues on hand movement,” J. King Saud Univ. Inf. Sci., vol. 30, no. 1, pp. 18–24, 2018.

[2] M. Fatourechi, A. Bashashati, R. K. Ward, and G. E. Birch, “EMG and EOG artifacts in brain computer interface systems: A survey,” Clin. Neurophysiol., vol. 118, no. 3, pp. 480–494, 2007.

[3] N. S. Malan and S. Sharma, “Feature selection using regularized neighbourhood component analysis to enhance the classification performance of motor imagery signals,” Comput. Biol. Med., vol. 107, pp. 118–126, 2019.

[4] S. Aggarwal and N. Chugh, “Signal processing techniques for motor imagery brain computer interface: A review,” Array, vol. 1, p. 100003, 2019.

[5] R. Leeb, C. Brunner, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller, “BCI Competition 2008–Graz data set B,” Graz Univ. Technol. Austria, pp. 1–6, 2008.

[6] K. Burnham and D. Anderson, “Model Selection and Multimodel Inference,” Technometrics, vol. 45, pp. 181–181, 2003.

[7] M. Yang, Y.-F. Sang, C. Liu, and Z. Wang, “Discussion on the choice of decomposition level for wavelet based hydrological time series modeling,” Water, vol. 8, no. 5, p. 197, 2016.

[8] Raza, H., H. Cecotti, Y. Li and G. Prasad, “Adaptive learning with covariate shift-detection for motor imagery-based brain–computer interface,”. Soft Computing, vol. 20,no. 8, pp. 3085–3096, 2016.

[9] Sayed, K., M. Kamel, M. Alhaddad, H.M. Malibary and Y.M. Kadah, “Characterization of phase space trajectories for Brain-Computer Interface,” Biomedical Signal Processing and Control, vol.38, pp.55–66, 2017.

[10] R. K. Chaurasiya, N. D. Londhe, and S. Ghosh, “Statistical wavelet features, PCA, and SVM based approach for EEG signals classification,” Int. J. Electr. Comput. Electron. Commun. Eng., vol. 9, no. 2, pp. 182–186, 2015.

[11] E. Dagdevir and M. Tokmakci, “Determination of Effective Signal Processing Stages for Brain Computer Interface on BCI Competition IV Data Set 2b: A Review Study,” IETE Journal of Research,1914204, 2021.

[12] E. Dagdevir, M. Kocaturk, and M. Okatan, “Likelihood-Based Amplitude Thresholding in Extracellular Neural Recordings,” 27th Signal Processing and Communications Applications Conference, 2019, pp. 1–4.