Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Classification of the motor imagery EEG signal using vector quantization and K-nearest neighbors' algorithm

Tae-Ung Jang, Wansu Lim, Yeon-Mo Yang, Byoeng Man Kim *

Kumoh National Institute of Technology, Gumi, South Korea

A R T I C L E I N F O

Article history:
Received 27 December 2015
Received in revised form
23 January 2016
Accepted 23 January 2016

Keywords:
Classification
EEG signal
K-nearest neighbors' algorithm

A B S T R A C T

Brain computer interface (BCI) is the mutual communication between the human brain and computers, and it is one of the most popular technologies in the human computer interface (HCI) research. In our propose algorithm, we first analyze time-frequency spectrum of EEG signals using short-time Fourier transform (STFT), and then we apply the LBG algorithm to extract the features of EEG signals via the vector quantization. Next, we calculate the degree of the similarity on the time series pattern of EEG signals. Finally, the motor imagery EEG signal is determined by using the method of the k-nearest neighbors (KNN). In our simulation, BCI competition II data is utilized, and as a result, the maximum performance of 88.57% is obtained.

1. Introduction

Brain computer interface (BCI) provides a novel and promising alternative communication method for interacting between the human brain and computers (Leeb et al., 2015; Higashi and Tanaka, 2013). The early BCI research was studied for the disabled person, and then BCI is extended to general person to improve the quality of life. To obtain the specific brain information, BCI systems measure the electroencephalograph (EEG) signal which is generated in the brain nervous system. After that, BCI extracts the specific features from the measured EEG signal. For example, ERP (event-related potentials), SCP (slow cortical potentials), sensorimotor activity, and VEP (visual evoked potentials) are used as EEG features.

Human brain signals appear in many different types such as SMR, Mu, and Beta waves depending on the frequency band (Leeb et al., 2015). Each brain area has different information. Specially, Brain waves about the kinesthesis like Motor Imagery is known for using the most in SMR wave area. Motor Imagery is typically called by the image training and it is one of psychological training techniques. The motor imagery is to conceive the specific movement such as lifting hand and shaking head. The brain signal related to the motor imagery has a similar signal pattern with the actual movement (Park et al., 2014; Bonnet et al., 2013).

In particular, we propose the novel classification algorithm to discriminate the left and right Motor Imagery EEG signal. Section 2 introduces the related works regarding the classification of the motor imagery EEG signals. The details of the classification algorithm are given in Sections 3 with respect to the feature extraction and the vector quantization. Section 4 describes the performance evaluations via the classification accuracy. Finally, conclusions are drawn in Section 5.

2. Relevant research

Schlögl (2003) described the competition result regarding the Motor Imagery part of BCI competition II's Third Session conducted in 2002~2003. The total seven teams participated, and they competed for their discriminant algorithm using Motor Imagery EEG provided by Graz University of Technology. The way of the competition is that each team submits the calculated answer for the experiment data using the learning data.

In three high rank teams, Christin Schafer gathered the feature of EEG signals using two channels (C3 and C4) and the Morlet-Wavelets algorithm in the frequency range of 10 to 22Hz. The motor imagery was discriminated by Bayesian error which is calculated based on the weight of the previous value obtained by a multivariate normal distribution of each class. As a result, the minimum error was measured as 10.71%.

The feature of Akash Narayana's team was defined as the AR-Spectral power obtained by the energy ratio between C3 and C4, and the linear discriminant analysis was used for the discrimination. As a result, the minimum error was measured as 15.71%. The last team, Amir Saffari, employed the AAR parameters as a feature, and they

* Corresponding Author.
Email Address: bmkim@kumoh.ac.kr (B.M. Kim)

discriminated a motor imagery EEG using the neural networks. As a result, the minimum error was measured as 17.14%. Ofner and Muller-Putz (2015) proposed the algorithm for the selection of optimal channels in the environment of multiple channels. Especially, (Ofner and Muller-Putz, 2015) drew the CFE map using the ERSP (Event-Related Spectral Perturbation) method and subsequently, the SVM algorim was applied to the discrimination of the EEG signals, which came from the selected channels. McCreadie et al. (2014) optimized the non-homogeneous spatial filter considering the non-stationary characteristics. CSP and LDA algorithms were operated to calculate the time-frequency segments and classification, respectively. Arvaneh et al. (2013) proposed the iCSSP filter and the performance results of the iCSSP were compared with the CSP, iCSP, and CSSP algorithm.

3. Proposed algorithm

The purpose of BCI system is to recognize the human intent from human brain signal and control various machines. Fig. 1 represents the BCI modeling process to build the practical BCI devices. Numerous BCI systems follow the format of the above model. Especially, the field of feature generator and feature classification is being the most actively researched. Therefore, our proposed BCI system is performed in this research area.

The data used in our research is a motor imagery EEG signal provided from BCI competition II consisting two types of signals such as learning signal and test signal. We first establish a classification system using learning signals, and then distinguish the motor imagery information along with the test signal as shown in Fig. 2.

In the learning session of Fig. 2, we eliminate the noise signal in the pre-treatment process and extract the specific features by using the spectrum analysis of brain signals. Next, we classify the extracted features to K groups in the clustering section and create the K codebooks based on the group representative values. In the test session, after removing the noise and obtaining the feature as the same procedure with the learning session, we carry out the vector quantization with the extracted feature and the codebook coming from the learning session. Finally, the part of the feature classification determines the human intent with the quantized features.

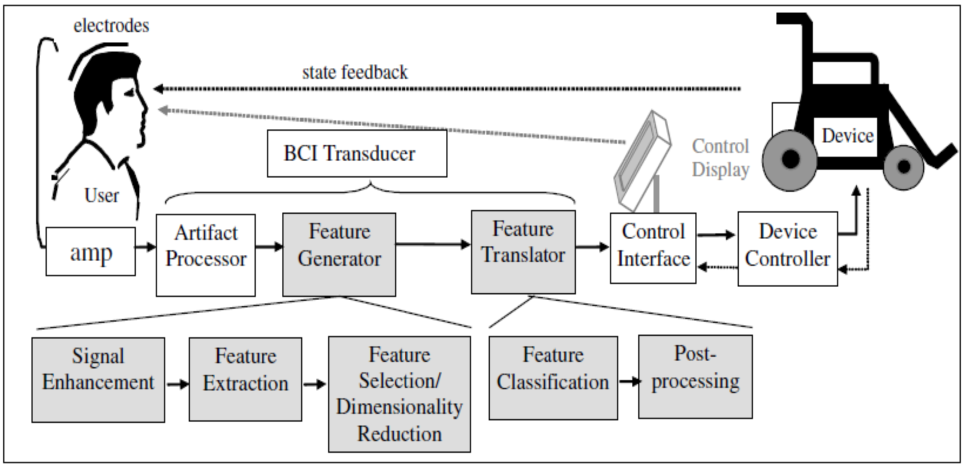

Fig. 1: BCI modeling process

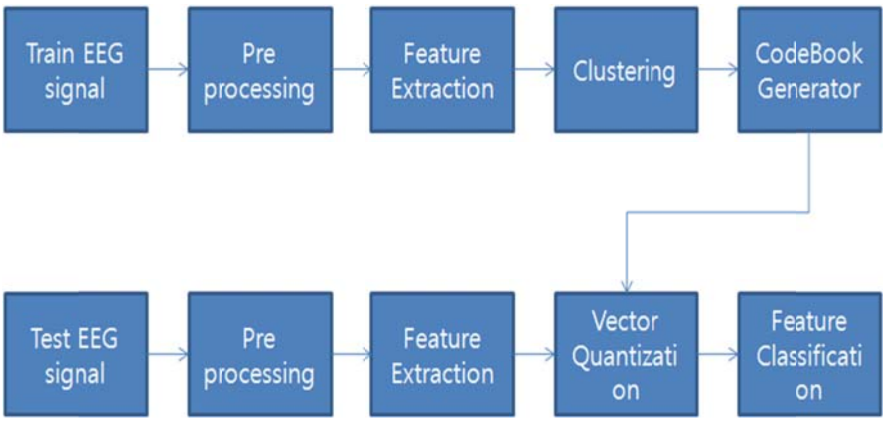

Fig. 2: Configuration of the motor imagery classification system

3.1. Pre-treatment procedure

The main purpose of the pre-treatment process is to eliminate the noise, which means unnecessary signal information in the experiment. In the proposed algorithm, we only use the Alpha (8-30Hz), SMR and Beta signal information among the experiment data of 1-64Hz. Other frequency band signals are removed because they are considered as a noise signal. Particularly, we apply the FIR filter in the elimination of noise in company with the Hamming window as a window function. Moreover, we multiply a constant value of 1.37 to the filtered signal in order to compensate the signal power degradation.

3.2. Features extraction procedure

There are several techniques for the features extraction from brain signals for instance spectral parameters, parametric modeling, and time-frequency representation (Arvaneh et al., 2013). In our algorithm, we use the power spectrum analysis in both time and frequency domains by utilizing short-time Fourier transform (STFT), which is one of the time-frequency representation approaches.

Fourier transform is the method that converts the brain signal from time domain to frequency domain and it permits the analysis of the power variation in each frequency range of the EEG. However, typical Fourier transform loses all time information in case of converting the entire signal. Thus, to overcome this drawback our proposed algorithm employs the STFT method in the analysis of the power variation of frequency domain according to the time. STFT divides the signal in a short duration signal using a window function. After that, we conduct Fourier transform with the derived short duration signal to analyze both time and frequency. STFT has a trade-off between the resolution of the frequency and the division duration of the time, as shown in Eq. 1. As a result, when we reduce the division duration of the time, the resolution quality of the frequency decreases, and vice versa.

$$STFT\mathrm{x}[\mathrm{n}](\mathrm{m},\omega) \equiv \mathrm{X}(\mathrm{m},\omega) = \sum_{n=-\infty}^{\infty} x[n]\omega[n-m]e^{-j\omega n} \quad (1)$$

In our research, we set the division duration by 128 samples for keeping the frequency resolution of 1Hz and a window function in the division of the time is performed as the Hamming window. Moreover, the half duration of the divided signal is overlapped with the next divided signal to reduce information loss of the signal end. This overlapped duration forms the additional sequence data, resulting in 17 sequence data as the same with 9 seconds. The experimental data is measured in three channels at the same time and they are collected from C3, Cz, C4 by 10-20 system rule. In our research, the observed data from C3 and C4 are examined, and each of C3 and C4 combines 17 additional sequences.

3.3. Clustering

In clustering process, the random data set is built into specific K groups by using K-mean or LBG algorithm. In this research, the feature vector, which represents an energy value in the range of 8Hz to 30Hz, is employed. However, due to various types of feature vectors, it can cause a performance degradation of the classification method. Thus, we operate a quantization of entire signal via the codebook and LBG algorithm. The following section explains the detail quantization process.

Step 1: Initialization: We calculate the initial center vector as Eq. 2 and then get the value of the distortion measure as Eq. 3.

$$c_1^* = \frac{1}{M}\sum_{m=1}^{M} X_m \quad (2)$$

$$D_{ave}^* = \frac{1}{Mk}\sum_{m=1}^{M} |X_m - c_1^*|^2 \quad (3)$$

Step 2: Splitting: Eq. 4 perform the binary separation of the codebook obtained, resulting in the double expansion of the codebook.

$$c_i^{(0)} = (1+\epsilon)c_i^*, c_{N+i}^{(0)} = (1-\epsilon)c_i^* \quad (4)$$

Step 3: Iteration: Eq. 5 calculates the similarity between the newly obtained codebook in the step 2 and all feature vectors, and then the most similar feature vector with the codebook is selected in Eq. 6. Finally, the pair of codebook and feature vector is generated.

$$\left|X_m - c_n^{(i)}\right|^2 \quad (5)$$

$$\mathrm{Q}(X_m) = c_n^{(i)} \quad (6)$$

Once finishing the pair process in Eq. 6, Eq. 7 recalculates the representative value

$$c_n^{i+1} = \frac{\sum_{Q(X_m)=c_n^{(i)}} X_m}{\sum_{Q(X_m)=c_n^{(i)}} 1} \quad (7)$$

Next, Eq.8 recalculates the value of the distortion measure.

$$D_{ave}^{(i)} = \frac{1}{Mk}\sum_{m=1}^{M} |X_m - Q(X_m)|^2 \quad (8)$$

Eq. 9 computes the degree of the optimization by the comparison between the newly obtained distortion and the previous distortion until the further optimization is not made. If the repetition of Eq. 9 stops, the final codebook is stored in Eq. 10.

$$\frac{\left(D_{ave}^{(i-1)} - D_{ave}^{(i)}\right)}{D_{ave}^{(i-1)}} > \epsilon \quad (9)$$

$$c_n^* = c_n^{(i)} \quad (10)$$

Step 4: Repeat: Until attaining the K codebooks, the steps 2-3 are repeated.

Using the methods described above, we conduct a quantization process by generating the K codebooks and the learning data with 17 sequence information. In our research, the length of the codebook is set in the range of 8 to 128 to confirm the optimum parameter. The number of the repetition is limited to the maximum 100 times according the value of the distortion measure.

3.4. Vector quantization.

In the vector quantization process, the value of the similarity of each codebook in the clustering process is calculated by using the Euclidean distance approach, and then each codebook replaces the value of the nearest code. Table 1 shows the example of the vector quantization. Especially, this example demonstrates the quantization result of the first test data, resulting in the first and second sequence quantized by 98th and 102th codebook, respectively.

Table 1: Example of the vector quantization with the first test data

Sequence	1	2	3	4	5	6	7
Index	98	102	11	57	51	16	101

3.5. Classification algorithm

The method of the feature classification operates k-nearest neighbors (KNN) algorithm with the learning data. The KNN algorithm calculates the degree of the similarity with each time-frequency data and accordingly the aggregated value of all similarity is regarded as a final similarity. Next, KNN determines the degree of the similarity with respect to the test data. Table 2 shows the example of the similarity between codes in the codebook.

Table 2: Similarity between codes

	Code 1	Code 2	Code 3	Code 4	Code 5
Code 1	1	0.5764	0.6605	0.1596	0.4652
Code 2	0.5764	1	0.7591	0.1765	0.7982
Code 3	0.6605	0.7591	1	0.3247	0.6946
Code 4	0.1596	0.1765	0.3247	1	0.2054
Code 5	0.4652	0.7982	0.6946	0.2054	1
Code 6	0.1994	0.4677	0.4282	0.2735	0.4505
Code 7	0.4677	0.7579	0.6778	0.2143	0.7083
Code 8	0.3205	0.4262	0.5	0.5601	0.4348
Code 9	0.4821	0.7922	0.6897	0.271	0.805
Code 10	0.446	0.4484	0.5117	0.2884	0.3423

In Table 2, the value of one represents the maximum degree of similarity. In the case of the first code, the similarity with the second and third code is 0.576 and 0.6605, respectively. Thus, the first code is more similar with the third code than that of the third code. After building the similarity table for the train data, we operate the quantization process for the test data, and consequently we calculate the similarity between the train data and the test data. For example, Fig. 3 illustrates the calculation of the similarity between the 19th train data and the first test data.

0.57 * … * 1

1st test data

Sequence	1	2	3	4	5	6
Index	98	102	11	57	51	16

19th test data

Sequence	1	2	3	4	5	6
Index	23	20	39	114	11	16

Fig. 3: Calculation between learning signal and test signal

The x-axis of Fig. 3 represents the time. The value of the first and second sequence of the test data is quantized into 98th code and 102nd code, respectively. In the calculation of the similarity, we bring 98th code for the test data and 19th code for the train code from the each similarity table, and as a result, the similarity value between two data becomes 0.57. The final similarity is obtained by multiplication of all similarity. The calculation of the degree of similarity is completed for the entire learning signal to a single test signal. By sorting the obtained similarity values in order of their size, we select the K number of high rank values, and based on the selected high K values, we decide the label of the test signal, as shown in Table 3. Table 3 displays the similarity result of the first test data with K of 5. The 81st training data provides the highest similarity with the value of 0.9533. The second high similarity is from the 86th training data as 0.938. The last row in the table is the labels of each train data, for instance, the 81st and 86th training data represent the left and right motor imagery, respectively. As a result, the first test data is labeled as the left motor imagery because the similarity with the 86th training data is highest

Table 3: Example of the labeling process for the first test data with five high similarity rank values (K=5)

Sequence	1	2	3	4	5
Index	23	20	39	114	11
Similarity	0.9533	0.938	0.9336	0.9331	0.9307
Label	1	2	2	2	2

4. Performance evaluation

In order to evaluate the performance of the proposed algorithm, BCI competition II is utilized as the motor imagery data. This data set consists of data for EEG measurements in the left and right movement of the 20 women in the imaginary data provided by the Graz University of Technology. Experimental data has the same sequence structure and consist, as shown in Fig. 4.

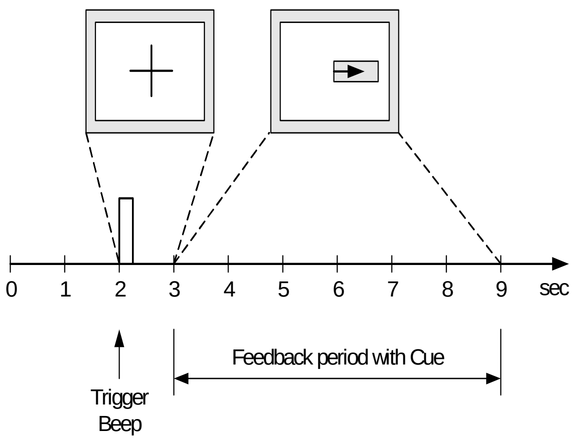

Fig. 4: Time flow of EEG experiment

The experimental time is 9 seconds, consisting of the initial two seconds as preparing time, subsequent one second as the waiting time with the cross symbol on the monitor, and the remaining 6 seconds as the motor imagery experiment with the arrow symbol. When the left or right arrow is displayed on the screen, the subjects think the movement in the direction of the arrow. The sampling rate is 128 Hz (128 samples per second) and the number of total samples is 1152 obtained during 9 seconds. In addition, the multi-channel system is operated according to the 10-20 international standard system, measuring the brain signal on C3, Cz, C4 at the same time. Each channel has 1152 samples followed by the time sampling rule. BCI competition II has 140 number of train and test data, individually, and the discriminant model is created by the train data. Finally, we decide the label of the test data by using our proposed classification model.

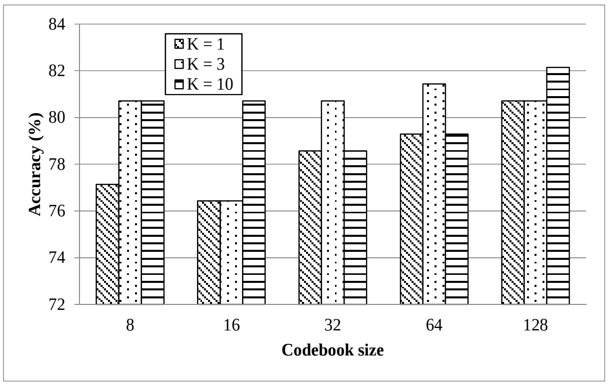

Fig. 5: Accuracy of classification according to the codebook size during 9 seconds

We evaluate our proposed algorithm as two experiments. In the first experiment, two parameters such as the size of the codebook and the K value of KNN method are changed, as shown in Fig. 5. In the second experiment, we conduct the further simulation with the data obtained in the range of 4 to 6 second.

Fig. 5 displays the classification accuracy as a function of the codebook size. Three experimental scenarios have been simulated to demonstrate the performance trend following the K value of 1, 3, and 10 in the KNN algorithm. The performance of K=1 is observed that the accuracy figures keep increasing as the codebook becomes high value size, justifying the more precise representation of the detected signal among various labels. In the cases of K=3 and K=10, the accuracy graph does not gradually increase. Especially, in the case of K=10 the accuracy decreases at the codebook size of 32. Thus, in order to optimize the classification, we need to find the appropriate value of K. The highest accuracy is 82.14% as a result of the code book size of 128 and K=10.

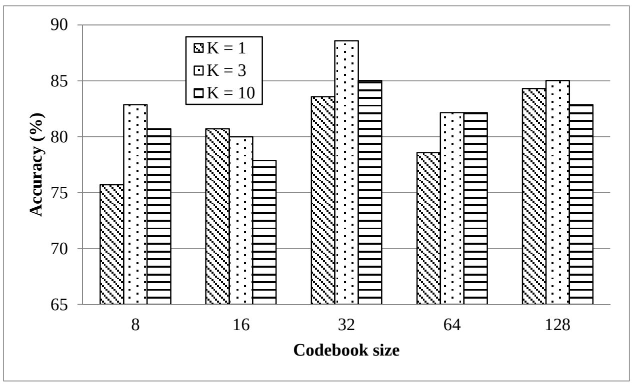

Fig. 6: Accuracy of classification according to the codebook size in the range of 4 to 6 second

In Fig. 6, we extract the data in the range of 4 to 6 sec from the time schema structure. This approach is experimented by Akash Narayana. The reason to select 4-6 sec is that this time duration is the most effective time to imagine the left or right arrow which is displayed on the screen, as shown in Fig. 6. Thus, if we experiment the classification test with 4-6 sec duration, the reliability and correctness of the performance result becomes high. Compared with the first experiment at Figure xx, the general performance of the second experiment is improved. In particular, at the codebook size of xx with KNN xx the accuracy performance with the first experiment at Fig. 6 is xx % compared to xx % for the second experiment at Fig. 6. This is because that the time duration of 4-6 sec clearly generates the motor imagery. In addition, it justifies that as the experiment time increases, the more redundancy date is collected, and as a result, the performance deteriorates. In other words, the optimized experimental time significantly could improve the classification accuracy.

5. Conclusion

In this paper, we proposed the classification algorithm of the motor imagery EEG signals by operating the spectrum analysis based on the STFT technique and the vector quantization for the feature extraction. In addition, we calculated the degree of the similarity between EEG signals in time domain by using the KNN scheme. Our proposed algorithm was evaluated with the BCI competition II dataset in terms of the classification accuracy according to the variation of the codebook size and the value of K in KNN. Especially, to obtain the improvement result, the proposed algorithm executes with the data, which is obtained during 4-6 sec. As a result, the maximum value of the accuracy is observed as 88.57% at the codebook size of 32 and KNN with K of 3.

Acknowledgment

This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the “ICT Consilience Creative Program” (IITP-2015-R0346-15-1007) supervised by the IITP (Institute for Information and communications Technology Promotion).

References

Arvaneh M, Guan C, Ang KK and Quek C (2013). EEG data space adaptation to reduce intersession nonstationarity in brain-computer interface. Neural Computation, 25(8): 2146-2171.

Bonnet L, Lotte F and Lécuyer A (2013). Two brains, one game: design and evaluation of a multiuser BCI video game based on motor imagery. Computational Intelligence and AI in Games, IEEE Transactions on, 5(2): 185-198.

Higashi H and Tanaka T (2013). Simultaneous design of FIR filter banks and spatial patterns for EEG signal classification. Biomedical Engineering, IEEE Transactions on, 60(4): 1100-1110.

Leeb R, Tonin L, Rohm M, Desideri L, Carlson T and Millan JDR (2015). Towards independence: a BCI telepresence robot for people with severe motor disabilities. Proceedings of the IEEE, 103(6): 969-982.

McCreadie KA, Coyle DH and Prasad G (2014). Is sensorimotor BCI performance influenced differently by mono, stereo, or 3-D auditory feedback? Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 22(3): 431-440.

Ofner P and Muller-Putz GR (2015). Using a noninvasive decoding method to classify rhythmic movement imaginations of the arm in two planes. Biomedical Engineering, IEEE Transactions on, 62(3): 972-981.

Park C, Took CC and Mandic DP (2014). Augmented complex common spatial patterns for classification of noncircular EEG from motor imagery tasks. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 22(1): 1-10.

Schlögl A (2003). Outcome of the BCI-competition 2003 on the Graz data set. Berlin, Germany: Graz University of Technology.