뇌파 분류에 유용한 주성분 특징

박성철[O] 이혜경 최승진
포항공과대학교 컴퓨터공학과
{schpark[O], leehk, seungjin}@postech.ac.kr

On Useful Principal Component Features for EEG Classification

Sungcheol Park[O] Hyekyoung Lee Seungjin Choi
Dept. of Computer Science & Engineering, POSTECH

Abstract

EEG-based brain computer interface(BCI) provides a new communication channel between human brain and computer. EEG data is a multivariate time series so that hidden Markov model (HMM) might be a good choice for classification. However EEG is very noisy data and contains artifacts, so useful features are expected to improve the performance of HMM. In this paper we addresses the usefulness of principal component features with hidden Markov model (HMM). We show that some selected principal component features can suppress small noises and artifacts, hence improves classfication performance. Experimental study for the classification of EEG data during imagination of a left, right, up or down hand movement confirms the validity of our proposed method.

1. Introduction

A brain computer interface (BCI) is a system which translates a subject's intentions into a control signal for a device, e.g. a computer application, a wheelchair or a neuroprosthesis[1]. Mainly BCI is carried out through EEG. Hence the classification of EEG plays an important role in BCI.

Several attempts have been made for EEG pattern recognition which consists of two procedures: (1) feature extraction; (2) classification. EEG data is a multivariate time series which contains noise and artifacts, linear dynamical systems (a.k.a Kalman filter) or HMM might be a useful method to model the EEG data for classification [2]. For feature extraction from EEG data, several different methods have already been tried. These include adaptive autoregressive (AAR) model, Hjorth parameters, and principal component features [2].

PCA is a well-known linear transformation for effective lower-dimensional representation for the data. Principal component directions are merely sought by the eigenvectors of the data covariance matrix having the largest eigenvalues. It is known that PCA minimizes the reconstruction error. Because of its simpleness and good performance, PCA has been used in many areas such as image processing, speech processing, etc for dimensionality reduction or feature extraction.

In this paper we employ the PCA for feature extraction from multivariate time series. The principal component features are used to train HMM for EEG pattern classification. In addition we investigate which principal component features are useful as features for EEG pattern classification. EEG data is very noisy and contains many artifacts (for example ocular artifiact which occurs during eye-blinking). Minor components are expected to correspond to small noise (with high frequency), so they are discarded. We expect that first several principal components contain contributions mainly from ocular artifact which has relatively large variation. We also throw away these several first principal components and show that this indeed improve the performance of HMM-based classification.

2. Data acquisition and analysis methods

2.1 Experiments

In this paper we analyze EEG data from experiments with a subjects called YSH. Two bipolar EEG-channels were recorded over left and right sensorimotor areas, close to electrode positions C3 and C4. The EEG are sampled at 200 Hz and bandpass filtered between 0.1 and 35Hz. During the experiment a cue was shown every trial on the computer screen. We used eight kinds of stimuli, stick, alphabet, rope, wall, eeg, button, puzzle and mouse (see Figure 1). All of the tasks except egg have four classes, up, down, left and right movement, and egg is two classes, left and right movement. Each task consists of 20 trials, which are presented to the subject in random order. We will call these 20 trials '1 run'. For each task, the subject experienced 2 runs, so there were 40 trials per task during one session. The subject experienced 3 session and he/she imagined dominant hand, both hands and both hands & language movement respectively. The protocol of the experiment is illustrated in Figure 2.

We consider C3 and C4 channels located in sensorimotor cortex related with (left, right, up or down) movement as well as imagination of movement. Figure 3 shows our proposed BCI system structure.

2.2 Feature Extraction

We segmented the imagination section from the EEG data at first. In order to apply the PCA to the segmented EEG data, we decompose the data into N overlapping blocks to construct $M \times N$ data matrix, (see Figure 4).

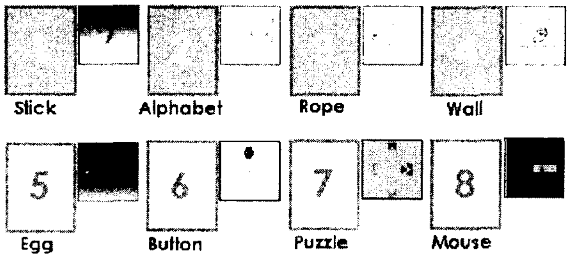

Figure 1 : Stimuli

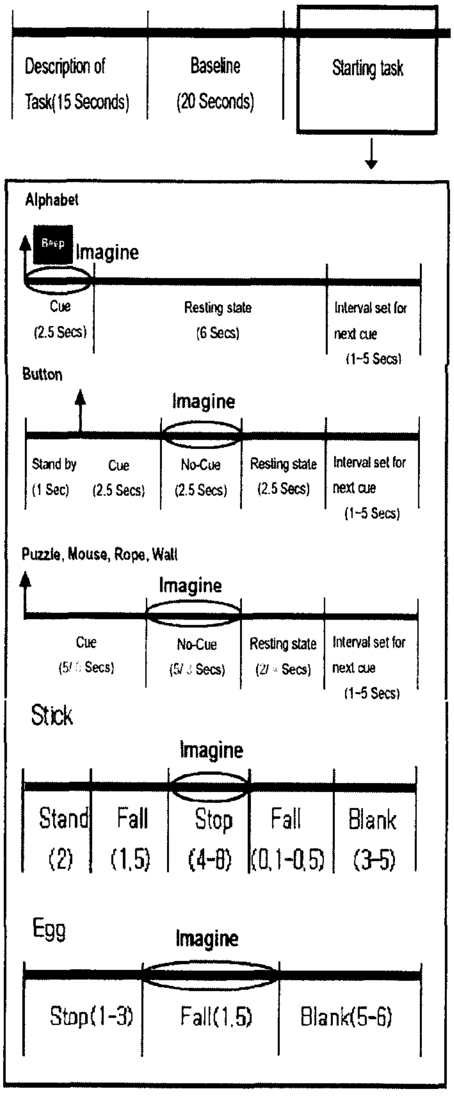

Figure 2 : experiment protocol

The PCA finds a linear transformation $v = Wu$, where W is a p by M matrix and its row vectors correspond to the normalized orthogonal eigenvectors of the data covariance matrix, $R_u = E[UU^T]$. Then the SVD of R_u gives $R_u = U_u D_u U_u^T$ where U_u is the eigenvector matrix (i.e.,modal matrix) and D_u is the diagonal matrix whose diagonal elements correspond to the eigenvalues of R_u. Then the linear transformation W for PCA is given by $W = U_u^T$. For dimensionality reduction, one can choose p dominant column vectors in U_u to construct a linear transform W. In our case, we should calculate $W_{C3,L}$, $W_{C4,L}$, $W_{C3,R}$, $W_{C4,R}$, $W_{C3,U}$, $W_{C4,U}$, $W_{C3,D}$ and $W_{C4,D}$ (where subscripts $C3$ and $C4$ denote channels, L, R, U and D correspond to left, right, up and down movement, respectively) in training phase. Then feature vectors are computed by $v_n = Wu_n$ where n is the integer $1,\ldots,N$[3].

EEG signals have many artifact : eye blinking, eye movement, muscle activity, interference of other channels, etc. Dimensionality reduction can reduce these artifact by eliminating redundant components and also can reduce computational complexity in HMM. Until now most researchers removed only eigenvectors having small eigenvalues. We deleted not only the eigenvectors of small eigenvalues but also the eigenvectors of large eigenvalues, because some artifact which affect the classification much have the largest variations(see Figure 5).

2.3 Classification

We classified a given set of feature vectors, $Y = (y_1, y_2, \cdots, y_N)$, $y_n = \{(v_{1,n}, \cdots, v_{p,n})_{C3}, (v_{1,n}, \cdots, v_{p,n})_{C4}\}$, with HMM for each movement. We calculate the likelihood, $P(Y|HMM_{left})$, $P(Y|HMM_{right})$, $P(Y|HMM_{up})$ and $P(Y|HMM_{down})$, and assign an appropriate class depending on which likelihood is larger.

To assess classification performance, the performance was estimated by 5-fold cross-validation.

3. Results

In order to show that proposed feature extracting method is good, we compared the cases that the n components of the largest eigenvalues are eliminated where n is the integer $0,\cdots,4$. Table 1 shows the results for various features and tasks. Generally, the accuracy is better when we removed one or more components. We can also catch that each task has the best number of eliminated components which have large eigenvalues. In the button and puzzle case, the accuracy gradually increase until 3 PCs, and it decrease when one more PC eliminated. Alphabet has similar pattern, and total mean confirm this pattern.

Table 2 shows the classification accuracy for various features and session, dominant hand, both hands and both hands & language. In the dominant hand case, all of the accuracy of abnormal, 1~4 PCs, are better than normal, 0

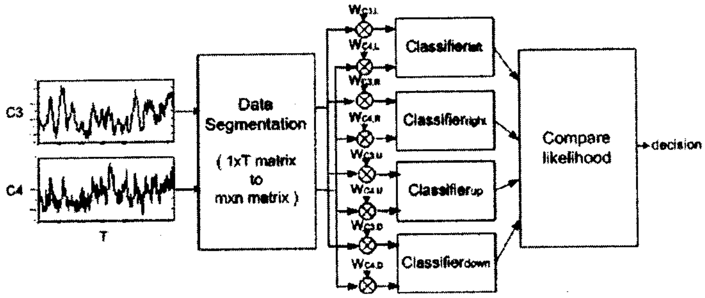

Figure 3 : overall structure

PC. In both hands and both hands & language case, the performance gradually progressed until 3 PCs, and it became worse off when one more PC removed.

In our results, we can notice that removing principal components having large eigenvalues make performance better for EEG classification, and that there are best number of the eliminated principal components at each task or each session.

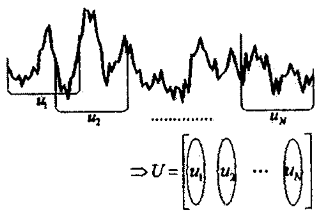

Figure 4 : Data segmentation

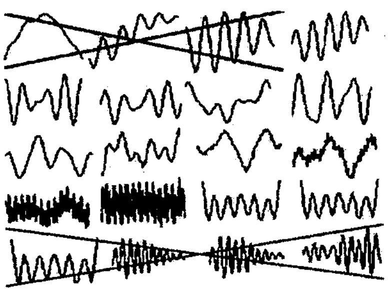

Figure 5 : principal component

$\mathbf{u}_{u,1}^T, \cdots, \mathbf{u}_{u,p}^T$

4. Conclusion

In this paper we presented the method of making useful principal component features for EEG classification. We noticed that some artifacts have larger variations than imagination EEG signals, so we eliminated some principal components having largest eigenvalues. Fortunately, we got much better accuracy with this approach and the computation complexity was much less expensive.

Table 1 : classification accuracy for each task

classification result (test/whole data) [%]	0 PCs	1 PCs	2 PCs	3 PCs	4 PCs
stick	91.7/94.5	83.5/97.2	98.3/99.2	98.3/99.3	98.3/99.3
alphabet	82.5/89.8	85.8/93.3	78.3/91.2	79.2/90.5	81.7/90.8
rope	88.3/91.7	87.5/96.5	88.2/92.5	94.2/97.3	88.3/98.7
wall	90/99	95.8/98.7	94.2/98.3	96.7/98.3	95.8/98.2
button	92.5/98.8	95/99	97.5/99	91.3/98.5	90/98
puzzle	89.6/90	90.4/96.3	93.3/97.7	93.3/98.5	90.4/96.7
egg	100/100	100/100	100/100	100/100	100/100
mouse	95/99.7	99.2/100	100/100	100/99.8	100/99.8
total mean	91.2/95.4	92.0/97.6	93.7/97.2	94.1/97.8	93.1/97.7

Table 2 : classification accuracy for each session

classification result (test/whole data) [%]	0 PCs	1 PCs	2 PCs	3 PCs	4 PCs
Dominant Hand	85.9/95.6	91.8/96.7	92.4/97.5	90.6/97.4	92.5/97.4
Both Hands	89.3/91.3	86.8/96.4	89.8/94.3	97.6/96.2	90.4/96
Both Hands & Language	97.9/99.6	98.3/99.9	99.3/100	100/100	96.2/99.9

5. Acknowledgments

This work was supported by ETRI.

6. Reference

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control", *Clin. Neurophysiol.*, vol. 113, pp. 767-791, 2002.
[2] H. Lee and S. Choi, "PCA-based linear dynamical systems for multichannel EEG classification", in Proc. ICONIP, pp. 745-749, Singapore, Nov. 2002.
[3] Jolliffe, I.T., "Principal Component Analysis", 2nd Edition, Springer, 2002.