다채널 뇌파 분류를 위한 주성분 분석 기반 선형동적시스템

이혜경[0] 최승진
포항공과대학교 컴퓨터공학과
{leehk[0], seungjin}@postech.ac.kr

PCA-based Linear Dynamical Systems for Multichannel EEG Classification

Hyekyoung Lee[0] 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. The classification of EEG data is an important task in EEG-based BCI. In this paper we present methods which jointly employ principal component analysis (PCA) and linear dynamical system (LDS) modeling for the task of EEG classification. Experimental study for the classification of EEG data during imagination of a left or right hand movement confirms the validity of our proposed methods.

1. Introduction

An important part in EEG-based BCI is the classification of circumscribed and transient EEG changes which are recorded during different types of motor imagery such as imagination of left-hand or right hand movement. Features such as band power, Hjorth parameters, or adaptive autoregressive parameters are extracted in EEG recordings of overlaying sensorimotor areas located over central and neighboring areas. For the classification of the features, linear discrimination analysis, neural networks, and hidden Markov models (HMMs) are used [1].

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.

HMM and LDS are widely-used probabilistic models for time series data and belong to a class of linear Gaussian models[2]. Both HMM and LDS can also be represented as graphical probabilistic models and describe probability distributions over sequences of observations. State-space models represent the past information through a real-valued hidden state vector, whereas discrete-valued states are employed in HMM. In fact, LDS can be viewed as a continuous-state analogue of HMM. The dependency between the present state vector and the previous state vector is specified through the dynamic equations of the system and the noise model. When these equations are linear and the noise model is Gaussian, the state-space model is also known as a LDS or Kalman filter model. Almost BCI research group have researched model using HMM, not confirming state dynamicity of EEG signal [3]. In this paper, we employ the LDS as an alternative to HMM for the task of EEG classification.

We use the PCA to preprocess the observation sequence before the data is fed into either HMM or LDS. Our experimental study shows that PCA-based preprocessing accelerates the convergence of learning LDS and improves the classification performance. Detailed description of our proposed methods is illustrated in Section 2.

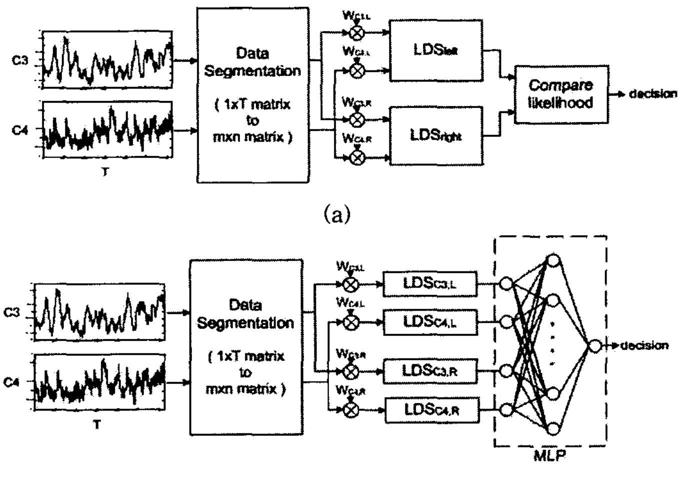
2. Proposed Methods

We consider C3 and C4 channels located in sensorimotor cortex related with (left or right) movement as well as imagination of movement. Figure 1 shows our proposed methods, PCA-LDS1 and PCA-LDS2, respectively. Both methods employ data segmentation and feature extraction using PCA. In the PCA-LDS1, only two LDS models are learned, corresponding either left or right movement. Binary classification is carried out by likelihood scoring. In PCA-LDS2, 4 different LDS models are trained. Unlike PCA-LDS1, PCA-LDS2 does not consider coupling between C3 and C4 channels. Thus, two LDS models for each channel results in 4 LDS models. The likelihood scores are fed into the MLP for final decision.

2.1 Feature Extraction - PCA

In order to apply the PCA to the EEG data, we decompose the data into N overlapping blocks to construct $M \times N$ data matrix, (see Figure 2).

The PCA finds a linear transformation $\boldsymbol{v} = \boldsymbol{Wu}$, where $\boldsymbol{W}$ is a p by M matrix and its row vectors correspond to the normalized orthogonal eigenvectors of the data covariance



(a)

(b)

Figure 1: Schematic diagram for (a) PCA-LDS1 and (b) PCA-LDS2

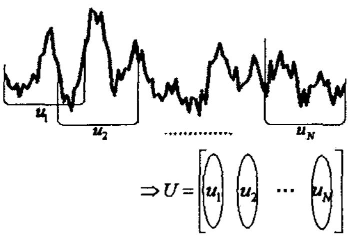


Figure 2: Data segmentation

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 $W_{C3,L}$, $W_{C4,L}$, $W_{C3,R}$ and $W_{C4,R}$ (where subscripts $C3$ and $C4$ denote channels, L and R correspond to left and right movement, respectively) in training phase. Then feature vectors are computed by $v_n = Wu_n$ where n is the integer $1,\ldots,N$

EEG signal has 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 LDS.

2.2 Classification - LDS

Linear time-invariant dynamical systems (also known as linear Gaussian state space models[4]) are described by

$$\begin{aligned} x_{t+1} &= Ax_t + w_t, & w_t &\sim N(0, Q), \\ y_t &= Cx_t + e_t, & e_t &\sim N(0, R), \end{aligned} \tag{1}$$

where $A \in R^{k\times k}$ is the state transition matrix and $C \in R^{p\times k}$ is the output matrix. The output y_t is a linear function of the state x_t which evolves through first-order Markov chain. Both state and output noise, w_t and e_t are zero-mean normally distributed random variables with covariance matrices Q and R, respectively. Only the output of the system is observed, the state and all the noise variables are hidden.

In the case of PCA-LDS1, feature vector in each LDS is $y_n = \{(v_{1,n},\cdots,v_{p,n})_{C3}, (v_{1,n},\cdots,v_{p,n})_{C4}\}$ Let us denote by LDS_{left} and LDS_{right}. Two LDS models learned from a training set of data recorded during left-hand and right-hand movement imagination, respectively. Given a set of feature vectors obtained from test data, $Y = \{y_1, y_2, \cdots, y_N\}$ we compute the likelihood, $P(Y \mid LDS_{left})$ and $P(Y \mid LDS_{right})$, to assign an appropriate class depending on which likelihood is larger.

In the case of PCA-LDS2, feature vector for each LDS is given by $y_n = \{(v_{1,n}, v_{2,n},\cdots,v_{p,n})\}$ For each LDS model, we compute likelihoods, $P(Y \mid LDS_{C3,left})$ $P(Y \mid LDS_{C3,right})$ $P(Y \mid LDS_{C4,left})$ and $P(Y \mid LDS_{C4,right})$ These likelihood scores are fed into the MLP to make a decision. MLP is trained in such a way that if the data is left-imagination, then the output is -1, otherwise the output is +1.

Although the PCA-LDS1 considers the interaction between channels, the dimension of its feature vector is twice larger than PCA-LDS2, which cause more complexity.

3. Experimental Results

Two bipolar EEG-channels were recorded over left and right sensorimotor areas, close to electrode positions C3 and C4. The EEG are sampled at 128 Hz and bandpass filtered between 0.5 and 30 Hz. Course of the experimental trial is followed this way. From 0 to 2 s a fixation cross was presented, followed by the cue at 2 s. At 3 s an arrow was displayed at the centre of the monitor for 1.25 s. Depending

	HMM1	HMM2	LDS1	LDS2
PCA	77.50	77.50	75.25	76.50
RAW	60.63	64.38	64.44	71.25
HJORTH	56.88	62.50	58.75	59.50

Figure 3 : Classification accuracy(%)

on the direction of the arrow presented left or right the subject was instructed to imagine a movement of either the left or the right hand. And then, feedback session continues from 4.25 to 8.0 s. One session constitutes 40 times repeating the course of the trial (20-left and 20-right). The total session is 4, so the number of trial is 160: 80-left and 80-right. We did not use feedback session. So the data from 3 to 4.25 s are used [5].

In order to show that PCA is a good feature extractor, we compare the PCA-based features with Hjorth parameters [3] and raw data. We also compare LDS to continuous HMM. LDS is replaced by HMM, then the resulting methods are referred to as PCA-HMM1 and PCA-HMM2. Methods based on the Hjorth parameter or raw data, are called as RAW-HMM1,HJORTH-LDS1.

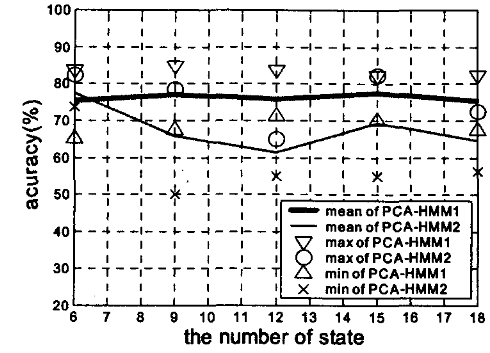
Figure 3 shows classification accuracy for various features and classifiers. Figure 4 shows classification accuracy when the number and dimension of state are changed using HMM and LDS, respectively. In the case of PCA, the window size is 0.5 s with overlapped portion being 0.875, and the dimension are reduced by half of window size. Hjorth parameters are obtained in the same way. In all case, we don't use the feedback session, but use the cue session between 3 and 4.25 s.

In the case using Hjorth parameter, the results are worse than the result using raw data. But in the case using PCA, we observed that the performance was improved by almost 10%, that the convergence speed also was faster than others. So we confirm PCA is a suitable feature extractor for EEG signal. Both HMM and LDS as classifier showed similar performance, which might imply that the state dynamicity of EEG signal is not either purely continuous or purely discrete. However, LDS required less complexity than HMM in the context of learning and LDS had more stable result than it to test data.

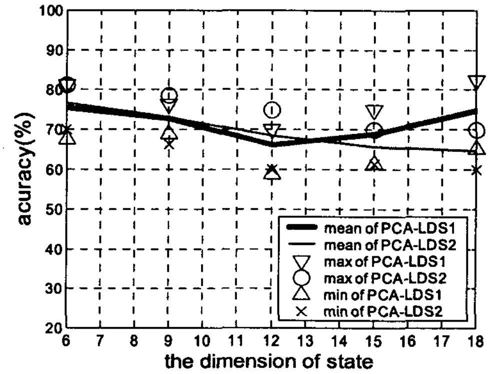
4. Conclusion

In this paper we presented LDS/HMM-based methods for multichannel EEG classification. We also employed PCA-based preprocessing so that LDS/HMM were trained from PCA-based features. We observed that PCA-based LDS/HMM have better performance than the others, whereas the LDS is less expensive than HMM in complexity. Currently we are investigating switching state space models for EEG classification.

5. Acknowledgment



(a)



(b)

Figure 4: Result of (a) PCA-HMM and (b) PCA-LDS

We thank Graz BCI research group for sharing their EEG data. This work was supported by ETRI and by Ministry of Education of Korea for its financial support toward the Electrical and Computer Engineering Division at POSTECH through its BK21 program.

6. References

[1] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of IEEE*, vol.89, no.7, pp. 1123-1134, July 2001.

[2] S. Roweis and Z. Ghahramani, "A unifying review of linear Gaussian models," *Neural Computation*, vol. 11, no. 2, pp. 305-345, 1999.

[3] B. Obermaier, C.Guger, C. Neuper, and G. Pfurtscheller, "Hidden Markov models for online classification of single trial EEG data," *Pattern Recognition Letters*, vol. 22, pp. 1299-1309, 2001.

[4] Z. Ghahramani, and G.E. Hinton, "Parameter estimation for linear dynamical systems," *University of Toronto Technical Report*, CRG-TR-96-2, 1996.

[5] C. Guger, A. Schlogl, C. Neuper, D. Walterspacher, T. Strein, and G. Pfurtscheller, "Rapid prototyping of an EEG-based Brain-Computer-Interface (BCI)," *IEEE Trans. Rehab. Engng.*, vol.9, no.1, pp. 49-58, 2001.