TENSOR CLASSIFICATION FOR P300-BASED BRAIN COMPUTER INTERFACE

Akinari Onishi[⋆†] *Anh Huy Phan*[†] *Kiyotoshi Matsuoka*[⋆] *Andrzej Cichocki*[⋆†]

[⋆] Kyushu Institute of Technology, Graduate School of Life Science and System Engineering,
Department of Brain Science and Engineering,
2-4, Hibikino, Wakamatsu, Kitakyushu, Fukuoka, Japan
[†] RIKEN, Brain Science Institute, Laboratory for Advanced Brain Signal Processing,
2-1, Hirosawa, Wako, Saitama, Japan

ABSTRACT

Classification methods have been widely applied in most brain computer interfaces (BCIs) that control devices for better quality of life. Most existing classification methods for P300-based BCIs extract features based on temporal structure related to P300 components of event-related potentials (ERPs). Some others exploit the spatial distribution of ERPs optimally selected by recursive channel elimination. However, none of them employed multilinear structures which exploit hidden features in P300-based BCI data. In this paper, we propose a new feature extraction method based on tensor decomposition for ERP-based BCIs. The method seeks an optimal feature subspace simultaneously spanned by temporal and spatial bases, and additional bases which indicate a variant of ERPs obtained by different degrees of polynomial fittings. The proposed method has been evaluated by both the BCI competition III data set II and the affective face driven paradigm data set, and achieved 92% and 95% classification accuracies respectively, which were better than those of most existing P300-based BCI algorithms.

Index Terms— Brain-Computer Interface (BCI), P300-based BCI, facial image, tensor, higher order discriminant analysis (HODA), electroencephalography (EEG), event-related potentials (ERPs).

1. INTRODUCTION

Brain-computer interfaces (BCIs) control devices such as wheelchair by means of brain signals [1]. Feature extraction and classification of electroencephalography (EEG) signals are prerequisite for most BCIs such as P300-based BCIs [2] and event-related desynchronization/synchronization-based (ERD/ERS-based) BCIs [3]. Since BCIs do not require any muscle movements, many researchers investigate BCIs for medical usage such as motor function substitution.

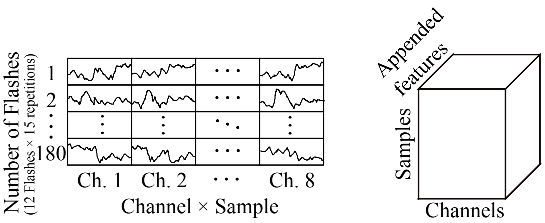
Two major existing approaches have been taken to improve P300-based BCI classification performance. One is to optimize stimulator by changing its stimuli or the paradigm. Takano et al. used green/blue flicker [4]. Hoffmann et al. applied images such as radio, lamp, TV and so on as a flash by changing its brightness [5]. Another approach is to apply appropriate feature extraction methods and classification techniques. Pires et al. applied Max-SNR beamformer, Fisher's criterion beamformer, and the combination of the two beamformers respectively before the classification by means of Bayesian classifier [6]. Hoffman et al. showed that Bayesian LDA can achieve better performance than Fisher's LDA [5].

Nevertheless, the existing approaches have not exploited multilinear structure which expresses a relation between P300 components over channels, or connection between variants of features. We note that features for classification of P300 signals are characterized by several different modes (dimensions) including samples, channels, flashes, repetitions and feature types such as low-pass filter and nonlinear regression. Recently multiway feature extraction has been applied to ERD/ERS-based BCIs, which is a promising tool for BCIs [7].

In this research, a novel method to extract multiway feature is proposed for ERP-based BCIs including P300-based BCIs. Instead of vectorizing all data, our approach expands the observed data into high dimensional tensors, and employs multilinear discriminant analysis methods [7] to find the optimal feature subspace from training P300 signals. Experimental data set of the BCI competition III data set II [8] and affective face driven paradigm (AFDP) [9] have been analyzed and we confirmed that our method was valid and achieved high classification performance for both data sets.

2. FEATURE EXTRACTION FOR P300-BASED BCI

During typical P300-based BCI experiments, a matrix of letters are presented and EEG is recorded when subjects silently count how many times the indicated target letter is intensified [10]. A waveform of the target class contains a P300 component which has a positive peak that appears approximately 300 ms after a target stimulus (see Fig. 2). Normally, the observed waveforms of the EEG signals are processed 700 ms time window after a flash, removing base correlation using 100 ms pre-stimulus waveform over multi-channels. Assuming that the interface requires $\#flashes$ times intensification to identify individual letters, and that all the flashes repeated



(a) Feature matrix (b) Feature tensor

Fig. 1. Feature matrix and feature tensor for P300 detection. (a) Each row vector of the feature matrix consists of EEG window recorded from multi-channels, and is labeled as a target or non-target signal. (b) The observed signals are tensorized into 3-D feature tensors whose frontal slices are compatible with row vectors in (a). Additional dimensions can be augmented such as regressions, frequency bins various time-frequency transformations, dictionaries. The tensor features enable us to apply tensor classification algorithms.

for $\#repetitions$ times to average them, there are in total $\mathrm{M} = \#flashes \times \#repetitions$ waveforms of size T samples $\times$ S channels in an epoch. It is obvious that the number of samples is large when EEG is recorded with a low sampling frequency (e.g., 240 Hz) and a few channels (e.g., 11 channels). The most common technique to reduce the dimensionality of P300 waveform is to downsample the signals after applying a moving average filter [5, 11]. The downsampled data then can be used as features for classification.

Most feature extraction methods for P300 waveform are based on the temporal structure of P300 component [2, 10], and they have not yet fully exploited the hidden structure in ERPs. That is because EEG segments of T samples on S channels are often concatenated into a long vector T $\times$ S (shown in Fig. 1(a)). We note that spatial information also affects the P300 detection. For example, we can select dominant channels by a recursive elimination [11].

In this paper, we propose a tensorization and a dimensionality reduction method for P300-based BCIs which extracts simultaneously dominant temporal and spatial information from the training data. In the sequence, we will show that seeking an optimal feature subspace for multiway samples is converted to a supervised feature extraction based on tensor decomposition with additional constraints.

2.1. Tensorization by Data Expansion

Unlike vectorizing waveforms as shown in Fig. 1(a), we expand the data into high dimensional tensors. The tensorization makes the features abundant, and enables us to take advantage of simultaneous multidimensional decompositions along all dimensions (modes) for feature extraction, which showed significant improvement in BCI classification performance [7]. In this research, we expanded data into 3-D tensor by appending a new dimension *degrees of polynomial fittings*

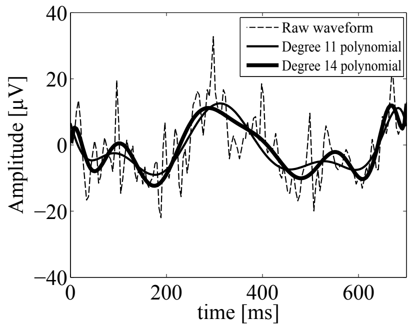


Fig. 2. Illustration of EEG waveform that contains P300 and its approximations to polynomial models. The fitting to the degree 11 polynomial model extracts slow waveform while fitting to the degree 14 polynomial yields more oscillations.

(see Fig. 1(b)). Fig. 2 illustrates a raw target EEG waveform, and its two approximates with degree 11 and 14 polynomials. The positive or negative peaks of ERPs reflect the cognitive procedure in human's brain such as perception of visual stimulus and emotion processing. Since averaged ERP waveform shows a smooth curve like a polynomial function in a limited time range, we can approximate data to a polynomial model by a least square method. The degree of a polynomial decides how many peaks or how well the system should extract the signal for the classification.

2.2. Model and Method for Dimensionality Reduction

Consider a set of K multiway samples $\boldsymbol{\mathcal{X}}^{(k)} \in \mathbb{R}^{\mathrm{T}\times\mathrm{S}\times\mathrm{O}}, k = 1, ..., \mathrm{K}$ which consist of EEG segments of T samples or their approximations with O degrees in S channels. Each sample belongs to a non-target class ($c = 1$) or target class ($c = 2$), which have K_1 and K_2 samples respectively. The model for dimensionality reduction of multiway samples is expressed as

$$\boldsymbol{\mathcal{X}}^{(k)} \approx \boldsymbol{\mathcal{G}}^{(k)} \times_1 \mathbf{T} \times_2 \mathbf{S} \times_3 \mathbf{V}, \quad k = 1, ..., \mathrm{K} \qquad (1)$$

where $\times_n$ denotes the tensor-matrix multiplication along mode-n [12], basis matrices $\mathbf{T} \in \mathbb{R}^{\mathrm{T}\times\mathrm{R_t}}$, $\mathbf{S} \in \mathbb{R}^{\mathrm{S}\times\mathrm{R_s}}$ and $\mathbf{V} \in \mathbb{R}^{\mathrm{O}\times\mathrm{R_o}}$ are projected filters for $\boldsymbol{\mathcal{X}}^{(k)}$ along time (samples), channel and fitting degree. The core tensor $\boldsymbol{\mathcal{G}}^{(k)} \in \mathbb{R}^{\mathrm{R_t}\times\mathrm{R_s}\times\mathrm{R_o}}$ consists of compressed features of $\boldsymbol{\mathcal{X}}^{(k)}$ projected onto the feature subspace spanned by $\mathbf{T}$, $\mathbf{S}$ and $\mathbf{V}$. Our purpose is to estimate projection matrices $\mathbf{T}$, $\mathbf{S}$ and $\mathbf{V}$ such that $\boldsymbol{\mathcal{G}}^{(k)}$ maximize difference between two classes. In general, we can maximize the Fisher ratio between the core tensors $\boldsymbol{\mathcal{G}}^{(k)}$ to find the basis factors $\mathbf{T}$, $\mathbf{S}$ and $\mathbf{V}$:

$$\varphi = \arg\max_{\mathbf{T},\mathbf{S},\mathbf{V}} \frac{\left\|\bar{\boldsymbol{\mathcal{G}}}_1 - \bar{\boldsymbol{\mathcal{G}}}_2\right\|_{\mathrm{F}}^2}{\displaystyle\sum_{k=1}^{\mathrm{K}} \left\|\boldsymbol{\mathcal{G}}^{(k)} - \bar{\boldsymbol{\mathcal{G}}}_{c_k}\right\|_{\mathrm{F}}^2}, \qquad (2)$$

where $c_k \in \{1, 2\}$ denotes the the class indices to which the k-th training sample $\boldsymbol{\mathcal{X}}^{(k)}$ belongs, and assuming that $\mathcal{I}_c$ is a set of sample indices k in each classes, $\bar{\mathcal{G}}_c$ is the mean tensors of each classes defined as

$$\bar{\mathcal{G}}_c = \frac{1}{\mathrm{K}_c} \sum_{k \in \mathcal{I}_c} \mathcal{G}^{(k)}, \quad c = 1, 2. \tag{3}$$

From eq. (1), the core tensors $\mathcal{G}$ can be approximately expressed by projections:

$$\begin{aligned} \mathcal{G}^{(k)} &= \mathcal{X}^{(k)} \times_1 \mathbf{T}^{\mathrm{T}} \times_2 \mathbf{S}^{\mathrm{T}} \times_3 \mathbf{V}^{\mathrm{T}} \\ &= \mathcal{Z}_1^{(k)} \times_1 \mathbf{T}^{\mathrm{T}} = \mathcal{Z}_2^{(k)} \times_2 \mathbf{S}^{\mathrm{T}} = \mathcal{Z}_3^{(k)} \times_3 \mathbf{V}^{\mathrm{T}}. \end{aligned} \tag{4}$$

Noting that

$$\|\mathcal{G}^{(k)} - \bar{\mathcal{G}}_{c_k}\|_{\mathrm{F}}^2 = \mathrm{tr}\left[\mathbf{T}^{\mathrm{T}} \langle \mathcal{Z}_1^{(k)} - \bar{\mathcal{Z}}_{1_{c_k}}, \mathcal{Z}_1^{(k)} - \bar{\mathcal{Z}}_{1_{c_k}} \rangle_{-1} \mathbf{T}\right],$$
$$\|\bar{\mathcal{G}}_1 - \bar{\mathcal{G}}_2\|_{\mathrm{F}}^2 = \mathrm{tr}\left[\mathbf{T}^{\mathrm{T}} \langle \bar{\mathcal{Z}}_{1_1} - \bar{\mathcal{Z}}_{1_2}, \bar{\mathcal{Z}}_{1_1} - \bar{\mathcal{Z}}_{1_2} \rangle_{-1} \mathbf{T}\right],$$

where $\bar{\mathcal{Z}}_{1_{c_k}}, c_k \in \{1, 2\}$ is the mean tensor defined for $\mathcal{Z}_1^{(k)}$ as in eq. (3), and $\langle \mathcal{A}, \mathcal{B} \rangle_{-n}$ denotes the contracted product between $\mathcal{A}$ and $\mathcal{B}$ in all modes except for mode-n. Hence, in order to update $\mathbf{T}$ while $\mathbf{S}$ and $\mathbf{V}$ are fixed, we optimize

$$\varphi = \arg\max_{\mathbf{T}} \frac{\mathrm{tr}\left[\mathbf{T}^{\mathrm{T}} \mathbf{S}_{\mathrm{b}} \mathbf{T}\right]}{\mathrm{tr}\left[\mathbf{T}^{\mathrm{T}} \mathbf{S}_{\mathrm{w}} \mathbf{T}\right]}, \tag{5}$$

where

$$\mathbf{S}_{\mathrm{b}} = \langle \bar{\mathcal{Z}}_{1_1} - \bar{\mathcal{Z}}_{1_2}, \bar{\mathcal{Z}}_{1_1} - \bar{\mathcal{Z}}_{1_2} \rangle_{-1}, \tag{6}$$

$$\mathbf{S}_{\mathrm{w}} = \sum_{k=1}^{\mathrm{K}} \langle \mathcal{Z}_1^{(k)} - \bar{\mathcal{Z}}_{1_{c_k}}, \mathcal{Z}_1^{(k)} - \bar{\mathcal{Z}}_{1_{c_k}} \rangle_{-1}. \tag{7}$$

As a consequence, $\mathbf{T}$ is $\mathrm{R_t}$ leading left generalized eigenvectors of the generalized eigenvalue decomposition $\mathbf{S}_{\mathrm{w}}\mathbf{T} = \lambda \mathbf{S}_{\mathrm{b}}\,\mathbf{T}$ or $\mathrm{R_n}$ leading eigenvector of matrix $(\mathbf{S}_{\mathrm{b}} - \varphi\, \mathbf{S}_{\mathrm{w}})$ [13].

Similarly, we can alternatively update $\mathbf{S}$ and $\mathbf{V}$. This optimization problem is related to multilinear discriminant analysis or high order discriminant analysis (HODA) [7]. Recently, the NFEA toolbox [14] has been developed for supervised and unsupervised feature extraction for multiway data based on tensor decompositions with various constraints including HODA. By applying HODA to the given training data $\mathcal{X}_{\mathrm{train}}^{(k)}, k = 1, ..., \mathrm{K}$, we can retrieve the basis factors $\mathbf{T}$, $\mathbf{S}$ and $\mathbf{V}$. The feature tensor $\mathcal{G}^{(k)}$ of a test or training data $\mathcal{X}^{(k)}$ is obtained via a simple projection written in eq. (4). The number of features in $\mathcal{X}^{(k)}$ is reduced to $\mathrm{R_t} \times \mathrm{R_s} \times \mathrm{R_o}$ entries. In fact, we don't use all the features for training a classifier. Only a few significant features are chosen from among the entire features based on their Fisher scores. The features extracted from the training waveforms are ready to train a classifier such as LDA or SVM.

The whole procedures for both training and test stages are illustrated in Fig. 3. Output values from the classifier are passed through a P300 decoder to predict an intended letter in a P300-based BCI.

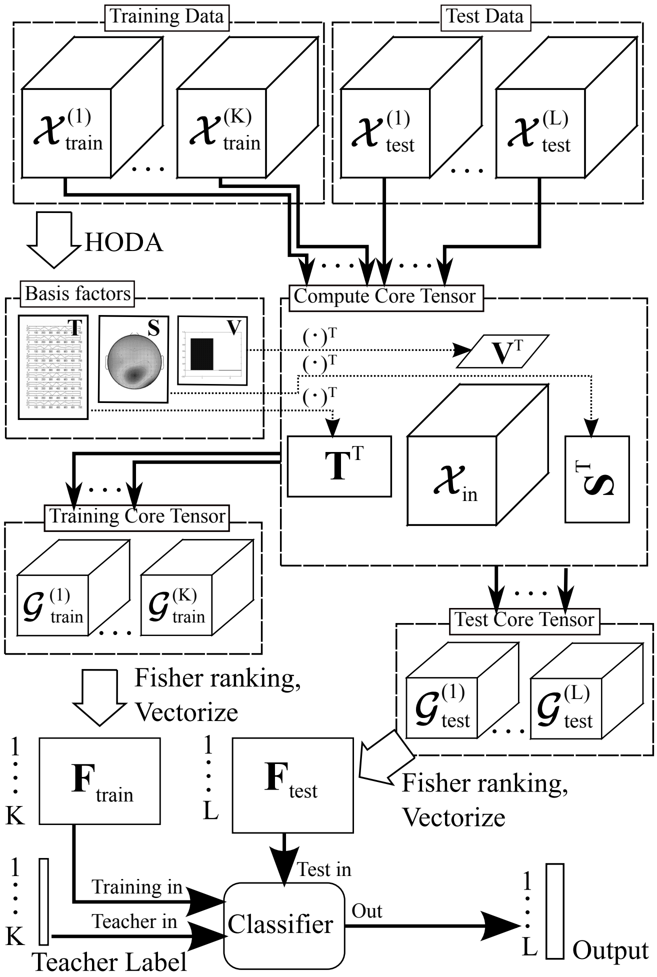


Fig. 3. Diagram for multiway feature extraction applied to a P300 speller system. (i) *Finding basis factors*: Decompose a set of K training tensors with category information (labels) to find 3 basis factors $\mathbf{T}$, $\mathbf{S}$ and $\mathbf{V}$ such that eq. (2) is maximized. (ii) *Projections*: Compute projections by basis factors $\mathbf{T}$, $\mathbf{S}$ and $\mathbf{V}$ to project data tensor onto the feature subspace (test core tensors). Dominant features are selected based on Fisher ranking. (iii) *Classification*: Learn a classifier by training features. Outputs are posterior probabilities corresponding to waveforms in one trial which can be passed through a P300 decoder.

3. SIMULATIONS

The proposed method was applied to the following two ERP-based BCI data sets. We firstly used *BCI competition III, data set II* [8], where the clear P300 components of ERPs were recorded. We used EEG data trimmed 700 ms after each stimulus onset for 15 repetitions, which contain 11 channel data (Fz, Fz, P5, Pz, P6, PO7, PO8, FC3, FC4, C3 and C4). We secondly employed *affective face driven paradigm (AFDP) data set* [9], which had ERPs elicited by a facial image presented in random order [9]. The ERPs of AFDP have multiple components that contribute to achieve high classification per-

Table 1. Classification accuracies of ERP-based BCI

Data sets	BCI competition III [8]	AFDP [9]
Smoothing	89%	92.5%
HODA + regression	92%	95%

formance. The same classification method can be applied for both data sets. The number of training data $\mathrm{K_{train}}$ is 85 and 80 respectively.

We evaluated the following two methods. As for the first method (smoothing), EEG signals were downsampled to a sampling rate of 20 Hz. That is, there were only 14 samples per waveform per channel. Linear discriminant analysis (LDA) was applied to extract features and classify the data. Regarding the second method (HODA + regression), EEG signals were first tensorized by degree 11 and 14 polynomials for BCI competition III data set II [8], and degree 7, 15 and 21 polynomials for AFDP data set. The same downsampling rate was applied to reduce the data size. Hence, training data were 3-D tensors $\boldsymbol{\mathcal{X}}^{(k)} \in \mathbb{R}^{14\times 11\times 2}$, $k = 1, \ldots, 85$ for BCI competition III data set II and $\boldsymbol{\mathcal{X}}^{(k)} \in \mathbb{R}^{15\times 8\times 3}$, $k = 1, \ldots, 80$ for AFDP data set. The core tensor size were $9 \times 10 \times 1$ and $15 \times 8 \times 1$ respectively. In fact, we selected a few dominant features features selected by the Fisher ranking.

For two methods, compressed features were used to train linear discriminant analysis (LDA). In order to detect intended letters, the same P300 decoder was applied to two methods: posterior probabilities of the classification result were averaged over repetitions and the output was identified by finding the maximum averaged posterior probability.

Accuracies for two methods are given in Table 1. Classification using features extracted by HODA achieved an accuracy of 92% for BCI competition III data set [8], which is compatible with that of the second winner for the same data set. Also the classification with HODA achieved 95% for affective face driven paradigm.

4. DISCUSSION AND CONCLUSIONS

A multilinear discriminant method has been proposed to extract features for ERP-based BCIs. We confirmed by simulations that our method achieved a 92% and 95% classification accuracies respectively for BCI competition III data set II and face driven paradigm data set. Samples in the experiments are 3-D tensors with extra dimension expressing the fitting degree. The model can be straightforwardly applied to higher dimensional tensors when additional dimensions are appended. In order to choose suitable parameters for polynomial fitting and multilinear discriminant analysis, a set of validation samples can be used. The validation samples can be independent of or split out of the training samples. Moreover, elimination of channels for P300 can also be performed with the validation samples. This might improve the performance.

5. REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–91, 2002.

[2] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.

[3] Q. Zhao, L. Zhang, and A. Cichocki, "Multilinear generalization of common spatial pattern," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 525–528.

[4] K. Takano, T. Komatsu, N. Hata, Y. Nakajima, and K. Kansaku, "Visual stimuli for the P300 brain-computer interface : A comparison of white/gray and green/blue flicker matrices," *Clinical Neurophysiology*, vol. 120, no. 8, pp. 1562–1566, 2009.

[5] U. Hoffmann, J. Vesin, T. Ebrahimi, and K. Diserens, "An efficient P300-based brain-computer interface for disabled subjects," *Journal of Neuroscience Methods*, vol. 167, no. 1, pp. 115–25, 2008.

[6] G. Pires, U. Nunes, and M. Castelo-Branco, "Statistical spatial filtering for a P300-based BCI: tests in able-bodied, and patients with cerebral palsy and amyotrophic lateral sclerosis," *Journal of Neuroscience Methods*, vol. 195, no. 2, pp. 270–81, 2011.

[7] A. H. Phan and A. Cichocki, "Tensor decompositions for feature extraction and classification of high dimensional datasets," *Nonlinear Theory and Its Applications, IEICE*, vol. 1, no. 1, pp. 37–68, 2010.

[8] BCI competition website, "Data set II: P300 speller paradigm," . *http://www.bbci.de/competition/iii/*

[9] A. Onishi, Y. Zhang, Q. Zhao, and A. Cichocki, "Fast and reliable P300-based BCI with facial images," in *Proceedings of the 5th International Brain-Computer Interface Conference 2011*, 2011, pp. 192–195.

[10] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger, "How many people are able to control a P300-based brain-computer interface (BCI)?," *Neuroscience Letters*, vol. 462, no. 1, pp. 94–8, 2009.

[11] A. Rakotomamonjy and V. Guigue, "BCI competition III: dataset II- ensemble of SVMs for BCI P300 speller," *IEEE Transactions on Bio-Medical Engineering*, vol. 55, no. 3, pp. 1147–54, 2008.

[12] A. Cichocki, R. Zdunek, A. H. Phan, and S. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, Wiley, 1st edition, 2009.

[13] H. Wang, S. Yan, D. Xu, X. Tang, and T. Huang, "Trace ratio vs. ratio trace for dimensionality reduction," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

[14] A. H. Phan, "NFEA: Tensor toolbox for feature extraction and applications," *http://www.bsp.brain.riken.jp/˜phan/nfea/nfea.html*, 2011.