Dictionary Update based Adaptive EEG Classification for Real Time Brain-Computer Interface Applications

Younghak Shin, Seungchan Lee and Heung-No Lee*
School of Information and Communications
Gwangju Institute of Science and Technology
Gwangju, Republic of Korea
shinyh@gist.ac.kr

***Abstract*— Due to the non-stationarity of EEG signals, classification performance is deteriorated during experimental sessions. Therefore, adaptive classification techniques are required for real-time BCI applications. In this paper, we propose simple adaptive sparse representation based classification (SRC) methods. We study supervised and unsupervised dictionary update schemes for new test data. The proposed methods are very simple and additional computation for the re-training of the classifier is not needed. We evaluate the proposed methods using an online BCI experimental dataset. The proposed methods are assessed by comparing classification results with the conventional SRC and other adaptive classification methods. We find that the proposed adaptive schemes show improved classification accuracy as compared to conventional methods without additional computation.**

***Keywords*— electroencephalogram (EEG); brain-computer interface (BCI); sparse representation based classification (SRC); L1 minimization; non-stationarity**

I. Introduction

Brain-computer interface (BCI) systems provide an alternative communication and control channel between human brain and external devices without any normal muscle movements. Scalp recorded electroencephalogram (EEG) signal is most widely used for non-invasive BCI systems [1].

Recently many wearable devices such as smart watch and EEG headsets are released. Development of portable EEG acquisition system is one of promising research area for the health care and medical applications. Much research effort have been focused on development of BCI applications for general public and dry electrodes which not need conductive gel for preparation of EEG recording [2]. However, for the portable EEG device going beyond laboratory researches, the most important issue is stable classification performance.

EEG signals have inherent non-stationary characteristics. Thus, there exist significant day-to-day and even session-to-session variability [3]. Due to this, classification performance is unavoidably deteriorated in BCI systems with time. To overcome the performance decrease caused by the non-stationarity of EEG signals, many adaptive classification methods are proposed. The study [3] proposes a bias adaptation scheme of linear discriminant analysis (LDA) classification using class labels of several test trials. They have shown that simple bias adaptation is effective for online test data. Similarly, [4] suggest unsupervised bias adaptation of LDA without using class label information. Previous studies for adaptive classification method need classifier re-adjustment (training) such as parameters and bias adaptation for new test trials. However, for this re-training, additional computation is needed in each update (adjustment) step.

Recently, sparse representation based classification (SRC) has shown an increased interest [5]. In the SRC framework, a test data from one class is predominantly represented by the same class training data from dictionary. The dictionary is composed by all class training data and usually underdetermined. Sparse representation of the test data using the dictionary can effectively be solved by the L1 minimization tool, and the classification is performed by comparing the representation error for each class. SRC scheme is first introduced for motor imagery based BCI application in [6]. They have shown that the SRC exhibits better classification performance than the conventional LDA method. Another study [7] also revealed that the SRC shows better classification accuracy and noise robustness than the well-known support vector machine (SVM) method.

In this study, with the unique classification mechanism of the SRC method we propose a simple dictionary update based adaptive SRC method for real-time BCI systems. We consider supervised and unsupervised dictionary update methods. Proposed dictionary update methods are very simple and additional computation for adaptation is not needed. Using online motor imagery based BCI experimental datasets, we evaluate classification performance of the proposed adaptive method by comparing with the conventional SRC and other adaptive classification methods.

II. Experiment

We performed online motor imagery based BCI experiment. Ten subjects participated in our online experiment. The experiment was performed on multiple days (two or three days). In each day, just one session experiment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (NRF-2015R1A2A1A05001826).

was executed. The number of sessions for each subject was determined by classification results and condition of each subject. Right hand (R) and left hand (L) motor imagery were performed for each subject. For this experiment, we used Active Two EEG measurement system made by Biosemi, Inc. The sampling rate of these datasets was 512 samples per sec and the number of EEG channels was 64. The channel positions were selected from international 10/20 standard.

The detailed experimental paradigm was illustrated in Fig. 1. The same paradigm was used for both training (calibration) and online testing (feedback) sessions. In each trial, the target bar was presented on 0 s at the right or left side of the screen corresponding to the right or left hand motor imagery. Two seconds after cue onset, the subject was instructed to perform the motor imagery task. During the training session, no feedback was provided. However, in the online testing session, the online feedback was provided in each trial. We collected 60 training trials and 75 online test trials for each class.

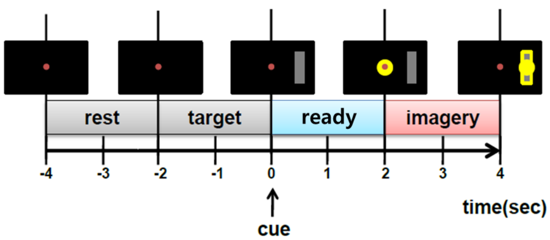

Fig. 1. One trial time sequence for motor imagery experiment

III. Methods

A. Preprocessing and Feature Extraction

Using the obtained motor imagery dataset of each subject, we perform the data preprocessing. After an instruction (left or right hand) appears at the monitor screen, the time samples from 1 to 2 second are collected for all trial data. We then band pass filter the trial data to eliminate the frequencies which are not related to motor imagery signals. In this study, sensorimotor rhythm, 8 to 15 Hz, is used for band pass filtering. We then reduce the dimension of EEG signal using the common spatial pattern (CSP) filtering which is a widely used feature selection method for motor imagery based BCIs [6].

The CSP filtering algorithm finds the filters $\mathbf{W} \in \mathbb{R}^{C \times C} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_C]$ which transforms the EEG data $\mathbf{X} \in \mathbb{R}^{C \times S}$ (C and S denote the number of EEG channels and time samples) into a spatially filtered space: $\mathbf{X}_{CSP} = \mathbf{W}^T \cdot \mathbf{X}$. Generally, $\mathbf{W}$ is computed by simultaneous diagonalization of the covariance matrices, Σ_1 and Σ_2, of the two classes data. This is equivalent to solving the generalized eigenvalue problem, i.e., $\Sigma_1 \mathbf{w} = \lambda \Sigma_2 \mathbf{w}$, where λ is eigenvalue. In practice, first and last k columns of the $\mathbf{W}$ corresponding to the k largest and k smallest eigenvalues are used for CSP filtering. For fair comparison, we set the k equal to five for all our datasets in this study. The obtained CSP filters maximize the variance of the spatially filtered signal for one class data while minimizing it for the other class data. Detailed information about the CSP filtering algorithm can be found in [6]. After CSP filtering, for each CSP filter, we compute the frequency band power of sensorimotor rhythm which is the final feature vector for classification.

B. Sparse Representation based Classification

The SRC method can be categorized as sparse coding step and identification step. The sparse coding step is formulated as $\mathbf{y} = \mathbf{A}\mathbf{x}$. Where, $\mathbf{y}$ and $\mathbf{A}$ indicate a test feature vector and a collection of training feature vectors. Also, $\mathbf{x}$ is an unknown coefficient vector. $\mathbf{A}$ is called a dictionary formed by class-dictionary $\mathbf{A}_i = [\mathbf{a}_{i,1}, \mathbf{a}_{i,2}, \ldots, \mathbf{a}_{i,N_i}]$, where $i = 1, 2, \ldots, C$ represents class information and N_i denotes the number of training trials for class i. In this study, C is equal to 2. $\mathbf{a}_{ij} \in \mathbb{R}^{m \times 1}$ is the j-th training feature vector of dimension m=2k from the class i. In this study, each element of $\mathbf{a}$ is the band power feature of the CSP filtered data. The dictionary $\mathbf{A}$ is formed by $\mathbf{A} := [\mathbf{A}_1; \mathbf{A}_2] \in \mathbb{R}^{m \times N}$, where N denotes the total number of training trials and. Thus, in this study, $N = 2N_i$ for two class problems.

In the SRC algorithm, first, the columns of dictionary $\mathbf{A}$ are normalized to have a unit L2 norm. Then, in the sparse coding step, unknown coefficient vector $\mathbf{x}$ can be recovered by solving following optimization problem via L1 norm minimization tool:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x}, \quad (1)$$

Note that (1) is an under-determined system. The literature of compressive sensing (CS) shows that the L1 norm minimization algorithm can solve this optimization problem effectively in polynomial time [8]. Using the recovered coefficient vector $\mathbf{x}$ by L1 minimization, class identification is performed as follows:

$$\text{class}(\mathbf{y}) = \min_i r_i(\mathbf{y}), \quad (2)$$

C. Dictionary Update based SRC method

To overcome non-stationarity of EEG signals, many adaptive classification schemes are proposed. The main concept of the adaptive classification is re-adjustment (re-training) of the classifier for the new test data. On the other hand, in the SRC scheme, one important characteristic is that training (or parameter decision) of a classifier is not needed unlike in other decision hyper-plane based classification methods such as LDA and SVM [7]. Thus, in the SRC scheme, a dictionary is simply formed by collecting the training feature vectors as columns of the dictionary. Then, using the dictionary sparse coding step is performed for each test data. Due to this unique classification mechanism, a simple intuitive method for adaptive SRC is dictionary update.

As we mentioned in Section III-*B*, the dictionary $\mathbf{A}$ is formed by class-dictionary $\mathbf{A}_i = [\mathbf{a}_{i,1}, \mathbf{a}_{i,2}, ..., \mathbf{a}_{i,N_i}]$ in the SRC method. Each column vector $\mathbf{a}_{ij}$ is a *j*-th training feature vector of class *i*. Therefore, for each test trial in the online testing phase, a feature vector of a new test trial $\mathbf{y}$ can be easily updated as a new column of the dictionary. Then, characteristics of the test feature can be applied into the dictionary while the online testing experiment is performed. And therefore, we can expect the classification performance of the online testing phase is not deteriorated.

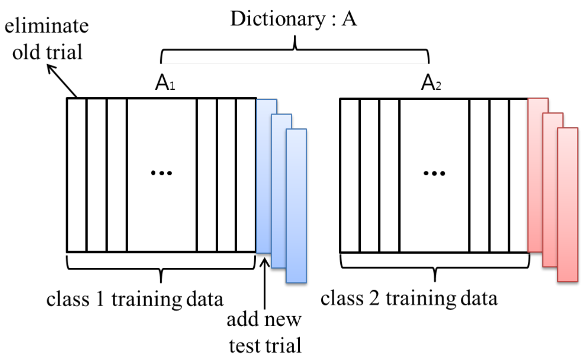

Fig. 2. Proposed dictionary update concept

Fig. 2 shows the proposed dictionary update rule. In this study, we consider two types of dictionary update rule, supervised and unsupervised update. In the supervised update rule, the target class label of test trials is used for updating the online test trials. Thus, a new test trial which has same class label of training trials in the class-dictionary is updated into the corresponding class-dictionary at the same time the oldest training trial, i.e., the first training trial of the class-dictionary is eliminated. On the other hand, in the unsupervised update rule, class label information of the test trial is not used. Thus, each test trial is updated into the corresponding class-dictionary based on the estimated result of the current classifier, which is represented by the direction of the yellow ball movement shown in Fig. 1.

IV. Results

To evaluate classification performance of the proposed adaptive SRC methods, we compare classification accuracy (%) of proposed methods with that of conventional SRC method using the online experimental dataset from ten subjects. From the multi session datasets of ten subjects, twelve session datasets are selected for evaluation of proposed methods. In this study, for the two class classification problems of the conventional SRC method, the dimension of the dictionary $\mathbf{A}$ is i.e., m = 10 CSP features and N = 120 training trials. For each subject, 150 test trials where each has the same 10 dimension features are evaluated with dictionary $\mathbf{A}$. Table 1 shows the classification accuracy (%) of the SRC and the proposed dictionary update based SRC methods for each session dataset.

From the results of the Table 1, proposed supervised and unsupervised dictionary update methods show improved mean classification accuracy than the conventional SRC method. Therefore, proposed simple dictionary update methods are efficient for online classification problem. Note that the supervised SRC method shows better mean classification accuracy than the unsupervised SRC method. In the unsupervised dictionary update method, the class labels of the test trials are determined by the results of the current classifier. Unfortunately, the classifier usually does not provide perfect classification results for all test trials because of the non-stationarity of EEG. Few incorrectly classified test trials are also updated in a different class-dictionary with the original target class. These trials affect the sparse coding step in the SRC algorithm. Therefore, this might be the reason that the unsupervised methods exhibit lower mean classification accuracy than the supervised methods.

TABLE I. Classification accuracy of conventional SRC and proposed adaptive SRC schemes, Supervised SRC (S_SRC) and Unsupervised SRC (U_SRC) for 12 session datasets.

Dataset	**Classification methods**		
	SRC	***S_SRC***	***U_SRC***
1	66	66	66
2	86	88	82.7
3	88.7	89.3	90.7
4	96.4	97.1	96.4
5	83.3	96.0	94.7
6	82.7	84.0	80.7
7	77.3	78.7	79.3
8	73.3	89.3	84.7
9	70.0	73.3	70.0
10	62.0	67.3	68.0
11	84.0	88.7	88.0
12	96.7	97.3	96.7
Mean	80.5	**84.6**	83.1
Std.	11.13	10.99	10.84

Fig. 3 shows the comparison results of the proposed methods with conventional adaptive LDA and SVM methods. The LDA and SVM are widely used classification methods in many EEG based BCI researches. For the adaptive LDA and SVM methods, first, linear decision hyper-plane is chosen from training data. Then in the testing session, the decision hyper-plane is re-trained for new test sample. We only consider supervised adaptation for the LDA and SVM methods.

From the results presented in Fig. 3, the proposed dictionary update based adaptive SRC methods show competitive classification accuracy than the other adaptive LDA and SVM methods. Note that even though the mean classification accuracy of the unsupervised adaptive SRC method is a little bit lower than the adaptive SVM method, in the conventional adaptive methods, re-training (re-adjustment) of the decision hyper-plane for new test data is time consuming process. However, in the proposed methods, dictionary update for adaptation of each test sample is very simple process and re-training of classifier is not needed.

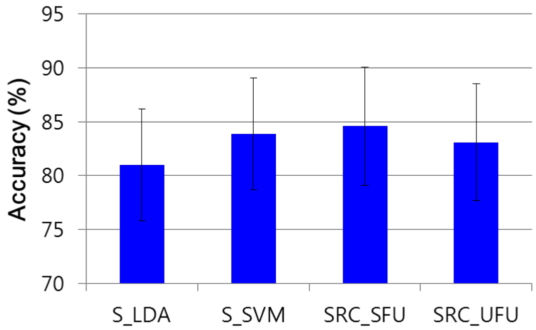

Fig. 3. Comparison of classification accuracy (%) between conventional adaptive methods and proposed methods.

V. CONCLUSION

In this paper, we propose simple dictionary update based adaptive SRC schemes to compensate for the non-stationary effects of EEG signals. We consider supervised and unsupervised dictionary update rules. In the proposed dictionary update methods, the test data are easily updated and utilized for the classification of other new test data without requiring any additional computation. By using an online motor imagery based BCI experimental dataset, we evaluate the classification performance of the proposed adaptive schemes. From the results, we show that the proposed adaptive SRC schemes show improved classification results compared to the conventional SRC. In addition, unsupervised adaptive SRC scheme that is more practically applicable in BCI exhibit competitive classification accuracy than other supervised adaptive LDA and SVM methods.

REFERENCES

[1] J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, “Brain-computer interfaces for communication and control,” Clin. Neurophysiol. vol.113 (6), pp. 767–791, 2002.

[2] L.D. Liao, C.Y. Chen, I.J. Wang, S.F. Chen, S.Y. Li, B.W. Chen, J.Y. Chang, and C.T. Lin, “Gaming control using a wearable and wireless eeg-based brain-computer interface device with novel dry foam-based sensors,” J. Neuroeng. Rehabil., vol. 9, p. 5, 2012.

[3] P. Shenoy, M. Krauledat, B. Blankertz, R.P.N. Rao, K.R. Müller, “Towards adaptive classification for BCI” J. Neural Eng., vol. 3, pp. R13–R23, 2006.

[4] C. Vidaurre, M. Kawanabe, P. von Bünau, B. Blankertz and K.R. Müller, “Toward unsupervised adaptation of LDA for brain–computer interfaces”, IEEE Trans. Biomed.Eng. vol.58 pp.587–597, 2011.

[5] K. Huang and S. Aviyente, “Sparse representation for signal classification”, Adv. Neural Inf. Process. Syst. vol.19, pp. 609–616, 2006.

[6] S. Younghak, L. Seungchan, L. Junho, L. Heung-No, “Sparse representation-based classification scheme for motor imagery-based brain-computer interface systems” J. Neural Eng. 9, 056002, 2012.

[7] S. Younghak, L. Seungchan, A. Minkyu, C. Hohyun, J. Sung Chan and L. Heung-No, “Noise robustness analysis of sparse representation based classification method for non-stationary EEG signal classification”, Biomed Signal Process Control vol.21, pp.8-18, 2015.

[8] D. Donoho, “Compressed sensing” IEEE Trans. Inf. Theory vol.52, pp.1289–306, 2006.