Incremental Versus Non-incremental Learning in Adaptive Common Spatial Patterns

Xiaomu Song
Department of Electrical Engineering
School of Engineering
Widener University
Chester, PA 19013

Suk-Chung Yoon
Department of Computer Science
College of Arts and Sciences
Widener University
Chester, PA 19013

***Abstract*—Extracting reliable and discriminative features remains a critical challenge in the development of brain computer interface (BCI) techniques. Common spatial patterns (CSP) is frequently used for spatial filtering and feature extraction in electroencephalography (EEG)-based BCI. It performs a supervised and subject-specific learning of EEG data acquired in two different task conditions. Incremental learning has been used in CSP to adapt to a target subject's data by including classified data in training data and re-estimating spatial filters. In practical circumstances where no user feedback is instantly available to provide true class labels of target trials, mis-classified EEG trials will be added to the training data of a wrong class, and potentially influence the training of spatial filters and feature extraction. In this study, incremental and non-incremental learning were investigated based on a recently developed adaptive CSP (ACSP) method using multi-subject EEG data. Their performances were compared in terms of intra- and inter-subject classification performances. Experimental results indicate that the non-incremental learning is a better option when true class labels of target data are not provided.**

I. Introduction

The analysis of single-trial multi-channel electroencephalography (EEG) data in EEG-based brain computer interface (BCI) is typically interfered with a variety of noise and artifacts. To attenuate noise disturbances and extract representative features, spatial filtering has been often used in EEG-based BCI systems. Common spatial patterns (CSP) is one of the most often used spatial filtering techniques [1]. It implements a supervised learning that needs training data from two different classes, and is subject/session-specific. EEG exhibits nonstationary behavior because of a variety of factors, such as changes of electrodes impedance/positions, subject attention/fatigue levels, and eye/facial muscle movements. As a result, spatial filters trained from a subject's data acquired at a certain time may not be suitable for new EEG data collected from the same subject at a different time or from different subjects performing the same mental task.

To improve the inter-session and inter-subject performance of CSP, variants of CSP have been proposed. Some of them using training data from different subjects and sessions [2], [3], others try to adapt to intra- and/or inter-subject variations [4], [5], [6]. Efforts have also been made to explore common stationary patterns across sessions and/or subjects [7], [8], [9]. In some of these approaches [4], [6], an incremental learning is usually considered to include classified EEG trials to training data based on which the spatial filters can be retrained and updated. The basic assumption that the incremental CSP learning can improve the BCI performance is that the classified data are assigned to the correct class. In real-time BCI applications, an instant feedback about the correctness of the past classification is not always available, and classified trials are usually assigned to the estimated class to update the training data. In such a case, mis-classified EEG trials will be added to a wrong class and influence the CSP learning.

In this work, this issue was investigated based on a recently developed adaptive CSP (ACSP) technique [10]. Two different incremental ACSP implementations were compared to a non-incremental implementation of the same method. Multi-subject EEG data from BCI competitions were used. Experimental results show that the non-incremental ACSP would be a preferred option in the aforementioned situation.

II. Common Spatial Patterns

The classic CSP method performs a supervised learning to maximize feature variation in one class while minimizing the variation in the other. If an M-channel EEG trial acquired in a task condition is denoted by an $M \times N$ matrix $\mathbf{E}_i(y)$, where i is the trial index, N is the number of time points, and $y \in \{1, 2\}$ is the class label, the normalized spatial covariance matrix $\mathbf{C}_y$ of class y is calculated as:

$$\mathbf{C}_y = \frac{1}{n_y} \sum_{i=1}^{n_y} \frac{\mathbf{E}_i(y)\mathbf{E}_i^T(y)}{trace(\mathbf{E}_i(y)\mathbf{E}_i^T(y))}, \tag{1}$$

where n_y is the number of EEG trials in class y. The CSP learning aims to solve the following generalized eigenvalue problem [1]:

$$\mathbf{C}_1\mathbf{W}^T = \mathbf{C}_2\mathbf{W}^T\mathbf{\Lambda}, \tag{2}$$

where the rows of $\mathbf{W}$ are spatial filters trained from $\mathbf{E}_i(y)$, $i = 1, \cdots, n_y$. The spatial filtering on the i^{th} EEG trial is performed as:

$$\mathbf{Z}_i = \mathbf{W}\mathbf{E}_i(y). \tag{3}$$

The columns of $\mathbf{W}^{-1}$ are termed as common spatial patterns that are treated as time-invariant EEG source distribution vectors. The first and last m rows of $\mathbf{Z}_i$ are the feature projections on $\mathbf{W}$ with maximal variations, and considered to be most discriminative. They can be used to construct a feature vector with the r^{th} feature calculated as:

$$x_r = \log[\frac{Var(z_r)}{\sum_{j=1}^{2m} Var(z_j)}], \tag{4}$$

where $Var()$ denotes the computation of variance, and z_r is the r^{th} row of $\mathbf{Z}_i$.

III. Adaptive Common Spatial Patterns

To adapt to the intra- and inter-subject variations in new EEG data, various ACSP methods have been developed [4], [6], [5]. A common process in these approaches is to perform an initial classification of the new data, and to assign the classified data to the estimated class to update the spatial filters in this class. The improvement brought from this procedure is based upon an assumption that the new data is correctly classified, which cannot be guaranteed. In this study, a different ACSP method was used [10]. This method first estimates the similarity between training and target data, and updates spatial filters for both classes simultaneously.

Given a target EEG trial with an unknown class label, the updated class covariance matrices are calculated as:

$$\begin{aligned} \bar{\mathbf{C}}_1 &= \frac{\phi_1}{n_1 + sgn(\phi_1)}\mathbf{C}_t + \frac{n_1}{n_1 + sgn(\phi_1)}\mathbf{C}_1, \\ \bar{\mathbf{C}}_2 &= \frac{\phi_2}{n_2 + sgn(\phi_2)}\mathbf{C}_t + \frac{n_2}{n_2 + sgn(\phi_2)}\mathbf{C}_2, \end{aligned} \tag{5}$$

where $\mathbf{C}_t$ is the normalized spatial covariance matrix of the target trial, n_1 and n_2 are numbers of training trials from the two classes. $\phi_1 \geq 0$ and $\phi_2 \geq 0$ are the similarity measures between the target and training data in the two classes. The Kullback Leibler divergence (KLD) is used to estimate ϕ_y, $y \in \{1, 2\}$. KLD measures the distance between two probability distributions. If each EEG trial is normalized to zero mean and standard deviation, then the EEG data can be characterized by a zero mean M-dimensional multivariate Gaussian distribution. Let $f_t = \mathcal{N}(0, \mathbf{C}_t)$ and $f_y = \mathcal{N}(0, \mathbf{C}_y)$ denote the distributions of the target and training data, the KLD between them is:

$$\begin{aligned} &KL(f_t, f_y) = \\ &\frac{1}{2}\{trace(\mathbf{C}_y^{-1}\mathbf{C}_t) - \log[\frac{det(\mathbf{C}_t)}{det(\mathbf{C}_y)}] - M\}. \end{aligned} \tag{6}$$

KLD is not a symmetric distance measure and typically symmetrized by averaging $KL(f_t, f_y)$ and $KL(f_y, f_t)$:

$$KLD(f_t, f_y) = \frac{1}{2}[KL(f_t, f_y) + KL(f_y, f_t)] \tag{7}$$

The similarity measure ϕ_y is determined as:

$$\phi_y = 1 - \frac{KLD(f_t, f_y)}{KLD(f_t, f_1) + KLD(f_t, f_2)}. \tag{8}$$

After calculating $\bar{\mathbf{C}}_y$ using (5), the remaining steps are the same as the CSP method to obtain updated spatial filters for feature extraction. From (5) it can be found that the weights for $\mathbf{C}_y$ and $\mathbf{C}_t$ are also influenced by the number of training trials. More training trials result in less weights for the target trial, which implies a greater chance to include more information of target subjects and a less possibility that the new trial might have very different characteristics.

IV. Incremental and Non-incremental ACSP

Two implementations of incremental ACSP was considered in this study. The first one is a class-specific approach where each classified trial is added to the training data of the estimated class. In the second incremental ACSP implementation, the classified trials are added to the training data of both classes using the weights shown in (5). For the non-incremental ACSP, classified trials won't be included in the existing training data. In the experimental study, the first class-specific incremental ACSP implementation was termed as "IncACSP-I", and the second incremental implementation was denoted as "IncACSP-II". The non-incremental one was called "NincACSP". The following is the complete procedure of the incremental ACSP:

- Step 1: Calculate $\mathbf{W}$ using the training data and CSP.
- Step 2: Input a target trial.
- Step 3: Compute the covariance matrix $\mathbf{C}_t$, and the KLD between $\mathbf{C}_t$ and $\mathbf{C}_y$, $y \in \{1, 2\}$ using (7).
- Step 4: Estimate ϕ_1 and ϕ_2 using (8).
- Step 5: Compute $\bar{\mathbf{C}}_1$ and $\bar{\mathbf{C}}_2$ using (5), and update $\mathbf{W}$.
- Step 6: Extract features using (3) and (4).
- Step 7: Train/retrain a data classifier using the updated features to classify features extracted from the target trial.
- Step 8: Add the classified target trial to the training data.
- Step 9: Go to Step 2 for the next target trial.

The difference between the two incremental ACSP implementations lies in Step 8. The non-incremental ACSP follows the same procedure except that Step 8 is excluded. The incremental ACSP would outperform the non-incremental one as previous studies have shown improvements brought by including target subjects' data in the CSP learning [2], [3]. This is typically true if there is a user feedback loop in the BCI system to know true class labels of target trials. When the true class label of a target trial is not provided right after the classification, however, the trial will be assigned to the wrong class if a mis-classification occurs. As a result, the updated spatial filters could influence the feature extraction and classification of following trials. The non-incremental ACSP does not include classified target trials in the training data, and would be an alternative option to the incremental ones in such a situation. A comparison study was performed between the incremental and non-incremental ACSP implementations using multi-subject EEG data and the results are elaborated in section VI.

V. Experimental Study

A. EEG Data

Three datasets from BCI Competitions III (dataset IVa) and IV (datasets IIa and IIb) were used in this work [11], [12].

Dataset IVa was collected from five subjects using 118 EEG channels at a sampling rate of 1000 Hz. One of three motor imagery tasks, including left hand, right hand, and right foot, was performed during the data collection. All trials in this dataset were band-pass filtered from 0.05 Hz to 200 Hz, and down-sampled to 100 Hz. The numbers of training and testing trials vary across subjects for each task and are listed in Table I, where the first row is the subjects' IDs. Dataset IIa was acquired from 9 subjects using 22 EEG channels. The subjects were performing one of four motor imagery tasks: left hand, right hand, foot, and tongue. For each subject, two data sessions were collected and there are 288 trials (72 trials for each task) in each session. The first session was used for training, and the second for the evaluation. Dataset IIb was collected from the same subjects in IIa using 3 EEG channels. The subjects were performing either left or right hand task. There are five sessions for each subject. Three sessions were used for training, and two sessions for the evaluation. There are 60-80 trials for each task condition in each session. All data in datasets IIa and IIb were sampled at 250 Hz and band-pass filtered between 0.5 and 100 Hz. Eye movement in IIa and IIb was regressed out using the simultaneous electrooculogram recordings. In the experimental study, only the trials from the left and right hand tasks were used. These trials were band-pass filtered from 8 to 32 Hz. The filtered trials have a zero mean and were normalized to standard deviation.

TABLE I
THE NUMBER OF TRAINING AND TESTING TRIALS IN DATASET IVA.

	aa	**al**	**av**	**aw**	**ay**
Training	168	224	84	56	28
Testing	112	56	196	224	252

B. Evaluation

The intra- and inter-subject classification performances were investigated for the incremental and non-incremental ACSP. In the intra-subject classification, training and testing data are from the same subject. Two types of inter-subject classifications were performed. In the first one, training and testing data are from all subjects in a dataset. The second is more challenging: training data are from only one subject in a dataset and the testing data are from all other subjects in the same dataset. A cross validation was performed so that each subject's training data were used once. A linear support vector machine was used as the classifier. Classification accuracy (denoted as P_a) and Cohen's kappa (κ) coefficient were computed to evaluate the classification performances.

VI. RESULTS

The P_a (%) and κ values computed from the intra- and inter-subject classifications are listed in Table II. For the two incremental ACSP implementations, it was observed that IncACSP-II outperforms IncACSP-I for all datasets in the intra- and inter-subject classifications. This indicates that in the cases where no true class label is provided after the classification, the weighted update of training data for both classes is a better option than the class-specific update of the training data when the incremental ACSP learning is considered.

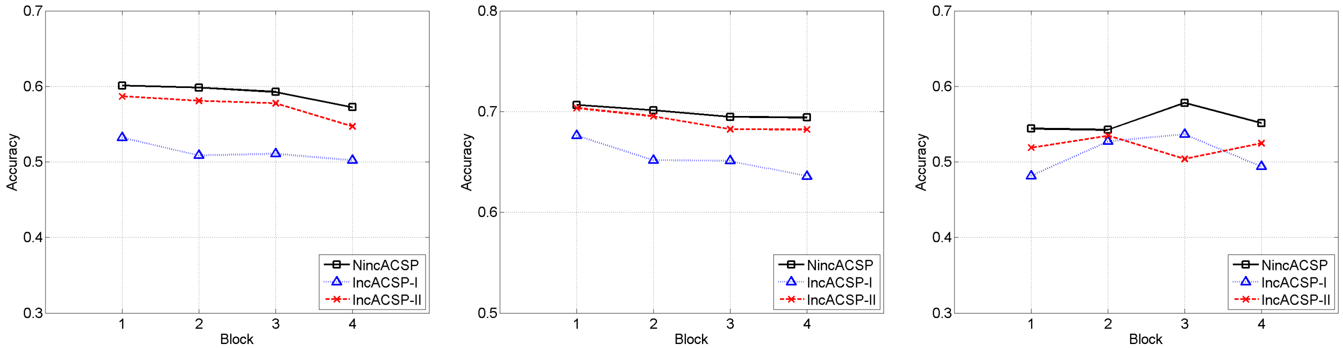
TABLE II
A COMPARISON OF THE INCREMENTAL AND NON-INCREMENTAL ACSP IN THE INTRA-SUBJECT STUDY AND THE INTER-SUBJECT STUDY USING SINGLE- AND MULTI-SUBJECT TRAINING DATA.

Intra-subject	IIa		IIb		IVa	
	P_a	κ	P_a	κ	P_a	κ
NincACSP	71.84	0.44	75.53	0.51	76.55	0.53
IncACSP-I	62.11	0.24	74.96	0.49	58.21	0.16
IncACSP-II	72.38	0.45	75.42	0.51	68.69	0.37
Inter-subject (Multi-subject)	IIa		IIb		IVa	
	P_a	κ	P_a	κ	P_a	κ
NincACSP	72.76	0.46	75.04	0.5	58.69	0.18
IncACSP-I	70.83	0.42	74.19	0.48	56.67	0.13
IncACSP-II	72.53	0.45	74.26	0.49	63.1	0.26
Inter-subject (Single-subject)	IIa		IIb		IVa	
	P_a	κ	P_a	κ	P_a	κ
NincACSP	59.11	0.18	70.02	0.4	53.84	0.08
IncACSP-I	51.34	0.03	65.48	0.31	50.0	0.0
IncACSP-II	57.31	0.15	69.19	0.38	53.15	0.06

The non-incremental ACSP outperforms the two incremental ones in most cases except for dataset IIa in the intra-subject classification and dataset IVa in the inter-subject classification using multi-subject training data, where IncACSP-II provided the highest accuracies. NincACSP and IncACSP-II are quite similar to each other because both of them perform the same weighted update of spatial covariance matrices for both classes, and the only difference is that the classified data won't be added to the training data in NincACSP. It was also found that in the inter-subject classification using single subject's training data, the non-incremental method outperforms the incremental ones for all three datasets. Considering that the training data size keeps increasing in the incremental ACSP implementations, the non-incremental ACSP is computationally more efficient and would be preferred in multi-subject BCI applications when the true label of classified trials are not instantly available.

To study how the incremental ACSP approaches influence the classification performance, the EEG data from each evaluation session in datasets IIa, IIb, and IVa were divided into four blocks, and all blocks in each session have the same number of EEG trials. The inter-subject classification using single-subject training data was performed. When datasets IIa and IIb were used, the number of EEG trials in each block was the same for all subjects, and each subject's training data were used once. When dataset IVa was used, the number of EEG trials in each block is different across subjects. In addition, only one of three subjects' ("av", "aw", and "ay") training data were used in each run, and the testing data were from the other four subjects. Table I shows that these three subjects have less number of EEG trials in the training data than the other two subjects ("aa" and "al"). The classification accuracy of each block was computed as the accuracy of the inter-subject classification averaged over all subjects' evaluation sessions in each dataset.

Fig. 1 illustrates the obtained accuracies of all blocks



(a) Dataset IIa (b) Dataset IIb (c) Dataset IVa

Fig. 1. The block-wise classification accuracies obtained from all subjects' testing data in datasets (a) IIa, (b) IIb, and (c) IVa using the non-incremental and two incremental ACSP implementations.

from the three datasets using the NincACSP, IncACSP-I, and IncACSP-II methods. It was observed that the block accuracies from IncACSP-I and IncACSP-II are lower than those obtained by NincACSP. This indicates that mis-classified trials can deteriorate the learning of spatial filters. On the other hand, NincACSP has no risk to assign the data to a wrong class. The block accuracies obtained from IncACSP-I are lower than those from IncACSP-II for datasets IIa and IIb. This is consistent to the results shown in Table II. When datasets IIa and IIb were used, the obtained block accuracies show a trend of decrease from block 1 to 4, while the block accuracies obtained from dataset IVa do not exhibit a consistent decrease. A possible reason of this observation is the varying number of training and testing trials from different subjects in dataset IVa. Apparently, the block accuracies from the two incremental ACSP implementations are lower than those from the non-incremental one, indicating that the inclusion of classified data into the training data leads to lower classification accuracy when true class label is not provided. In IncACSP-II, the weighted update of training data for both classes results in the update of spatial covariance matrices for both classes instead of one class. Although the accumulation of partial spatial covariance from the opposite class could influence the performance of updated spatial filters, compared to the inclusion of an entire spatial covariance from the opposite class in IncACSP-I, mis-classified trials have less effects on the training of spatial filters in IncACSP-II.

VII. Conclusion

It was investigated whether the inclusion of classified EEG trials in training data could improve the learning of spatial filters when true class labels of the classified data are not instantly available in BCI applications. The study was based upon an adaptive CSP (ACSP) method that integrates spatial covariance of target data into the learning of spatial filters. Two incremental and one non-incremental ACSP implementations were compared using motor imagery data from BCI competitions. The experimental results show that (1) the non-incremental ACSP (NincACSP) outperforms the incremental ones in most cases, and (2) the incremental ACSP that performs a weighted update of the training data (IncACSP-II) outperforms the other that performs the class-specific update of the training data (IncACSP-I). It was also noted that information of both classes is involved to the updates of spatial filters in NincACSP and IncACSP-II. Since the non-incremental ACSP does not involve additional training data, it is computationally more efficient and would be a better option for motor imagery BCI tasks where true class labels of target trials are not promptly provided.

References

[1] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Müller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 41–56, 2008.

[2] F. Lotte and C. Guan, "Learning from other subject helps reducing brain-computer interface calibration time," in *Proc. 35th IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2010, pp. 614–617.

[3] D. Devlaminck, B. Wyns, M. Grosse-Wentrup, G. Otte, and P. Santens, "Multisubject learning for common spatial patterns in motor-imagery BCI," *Computational Intelligence and Neuroscience*, vol. 2011, no. 217987, pp. 1–9, 2011.

[4] S. Sun and C. Zhang, "Adaptive feature extraction for EEG signal classification," *Med. Bio. Eng. Comput.*, vol. 44, pp. 931–935, 2006.

[5] Y. Li and C. Guan, "An extended EM algorithm for joint feature extraction and classification in brain-computer interfaces," *Neural Computation*, vol. 18, pp. 2730–2761, 2006.

[6] A. Bamdadian, C. Guan, K. Ang, and J. Xu, "Online semi-supervised learning with KL distance weighting for motor imagery-based BCI," in *Proc. 34th International Conference of IEEE EMBS*, 2012, pp. 2732–2735.

[7] B. Blankertz, M. Kawanabe, R. Tomioka, F. Hohlefeld, V. Nikulin, and K.-R. Müller, "Invariant common spatial patterns: alleviating nonstationarities in brain-computer interfacing," *Adv. Neural Inf. Process. Syst.*, vol. 20, pp. 1–8, 2008.

[8] W. Samek, C. Vidaurre, K. R. Müller, and M. Kawanabe, "Stationary common spatial patterns for brain computer interfacing," *J. Neural Eng.*, vol. 9, pp. 026 013:1–14, 2012.

[9] M. Arvaneh, C. Guan, K. Ang, and C. Quek, "Optimizing spatial filters by minimizing within-class dissimilarities in electroencephalogram-based brain-computer interface," *IEEE Trans. Neural Networks and Learning Systems*, vol. 24, no. 4, pp. 610–618, 2013.

[10] X. Song and S.-C. Yoon, "Improving brain-computer interface classification using adaptive common spatial patterns," *Computers in Biology and Medicine*, vol. 61, pp. 150–160, 2015.

[11] B. Blankertz, K.-R. Müller, and et al., "The BCI competition III: Validating alternative approaches to actual BCI problems," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 14, pp. 153–159, 2006.

[12] M. Tangermann, K.-R. Müller, and et al., "Review of the BCI competition IV," *Front. Neurosci.*, vol. 6, no. 55, pp. 1–31, 2012.