Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

FrP2A1.2

Time-Variant Spatial Filtering for Motor Imagery Classification

Haihong Zhang, Chuanchu Wang, Cuntai Guan
Institute for Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613
Email: {hhzhang, ccwang, ctguan}@i2r.a-star.edu.sg

Abstract—**Effective spatial filtering plays a key role in motor imagery classification. This paper presents a novel approach to spatial filtering of EEG signal by modelling time-variant spatial patterns. This is in contrast to conventional Common Spatial Pattern which assumes static spatial patterns in a motor imagery trial. We define the model such that it accounts for relatively higher order dynamics in EEG. Furthermore, we formulate the training of the model as a dual optimization problem, and we derive an iterative optimization algorithm using quadratically constrained quadratic programming. Our experimental results on healthy subjects indicates that the proposed method is able to produce higher classification accuracy.**

Index Terms—**Common Spatial Pattern, Time-Variant Filtering, Motor Imagery, EEG**

I. INTRODUCTION

Motor imagery classification provides an important basis for brain-computer interfaces (BCIs) by directly conveying people's imagination of movements to the outside world [1]. It is especially appealing to severely paralyzed patients, since motor ability is no longer a prerequisite for this communication. It also offers a promising tool for normal people to enhance their interaction with computers.

Motor imagery, or imagination of movement, is actually the mental rehearsal of a motor act without any real motor output. Evidence from a psychophysiology study with positron emission tomography (PET) shows that, different from in real performance, in motor imagery the execution would be blocked at cortical-spinal level [2]. It can change the functional connectivity within the cortex and leads to an amplitude suppression (event-related desynchronization, or ERD) or an amplitude enhancement (event-related synchronization, or ERS) of μ and β rhythms [3]. The ERD and ERS can be apparent in grand trial averages, but they may not be clearly present in individual motor imagery trials [4].

In 1997, Pfurtscheller et al. demonstrated the feasibility of using EEG to differentiate between imagination of right and left hand movements [5]. Since then, the topic of motor imagery classification has grown into the focus of EEG-based BCI research. The importance of the topic is also indicated in, e.g. the latest BCI Competition III (2005) where 6 out of 8 data sets are motor imagery signals [6].

This paper focuses on the feature extraction for motor imagery classification. In this regard, spatial filtering methods are of particular interest: Combining data from multiple channels (locations on scalp) is expected to provide a means to focus on activities with particular spatial distributions. In particular,

the common spatial pattern (CSP) algorithm has become a popular method for distinguishing motor imagery EEG signals (e.g. [7], [8]). The algorithm aims to find linear projections that maximize variance for one class while at the same time minimizing variance for the other class. Thus, it may bring out the difference in spatial distributions of EEG between e.g. left/right hand movement imaginations. Basic CSP deals with binary classification. To address multi-class classification of EEG signals, a few possible extensions to CSP were studied in [9], with the results suggesting the use of a stack of pair-wise CSP projections. Conventional CSP algorithms process band-pass filtered EEG signals while the frequency band is fixed. In [10], Lemm et al. presented an extension to CSP that can emphasize specific frequency bands, by simply concatenating input signals and their delayed counterparts. More recently, they improved the method by integrating an FIR filter into the CSP framework, and the new algorithm is able to optimize simultaneously a spatial filter and a spectral filter [11].

However, priori arts explored static spatial patterns only. In other words, they employ a spatial pattern filter that is fixed throughout a motor imagery trial. On one hand, this methodology simplies the filtering problem; On the other hand, it neglects time-variant spatial information which may be critical for motor imagery classification. Actually the temporal structure of spatial pattern can be inconsistent. For example, ERD and ERS usually occur at different time in a trial.

This paper presents an approach to address the problem by introducing time-variant spatial filtering. We propose a linear model to describe time-variant spatial patterns. And we derive an optimization algorithm for searching for optimal filters, using quadratically constrained quadratic programming. Our preliminary experiment indicates that the proposed method is able to produce higher classification accuracy.

The rest of the paper is organized as follows. Section 2 revisits the conventional CSP method. Section 3 proposes a new time-variant filtering model. In Section 4 we derive a mathematical solution to training the model. And Section 5 presents our experimental results and discussions, followed by conclusion in Section 6.

II. CSP REVISITED

The CSP method explores topographic patterns of brain rhythm modulations. For example in the case of discimination between left hand and right hand motor imagery, CSP tries to maximize variance for the class of right hand and at the same

time minimizing variance for left hand trials. Let the brain signal be $\mathbf{x}$. The CSP aims to search for an optimal projection vector $\mathbf{w}$ by

$$\underset{\mathbf{w}}{\operatorname{argmax}} \quad E[(\mathbf{w}^T\mathbf{x}_i - E[\mathbf{w}^T\mathbf{x}])^2 | \mathbf{x} \in O_1] \text{ s.t.}$$
$$E[(\mathbf{w}^T\mathbf{x}_i - E[\mathbf{w}^T\mathbf{x}])^2] = 1 \tag{1}$$

This can be reformulated as

$$\underset{\mathbf{w}}{\operatorname{argmax}} \quad \mathbf{w}^T\Sigma_{O_1}\mathbf{w} \text{ s.t. } \mathbf{w}^T(\Sigma_{O_1} + \Sigma_{O_2})\mathbf{w} = 1 \tag{2}$$

where Σ_{O_i} denotes the covariance matrix for class O_i. This optimization problem can be solved using joint diagonalization, which consists of calculating a matrix W and a diagonal matrix D:

$$W\Sigma^{(1)}W^T = D, \quad W\Sigma^{(2)}W^T = I - D \tag{3}$$

The matrices W and D can be obtained using the following procedure.

1) Calculate the matrix P which whitens the matrix $\Sigma_{O_1} + \Sigma_{O_2}$: $P(\Sigma_{O_1} + \Sigma_{O_2})P^T = I$; this is can be carried out using SVD and normalization;
2) Calculated the whitened matrix $\Sigma'_{O_1} = P\Sigma_{O_1}P^T$;
3) Calculated the eigen vector matrix Q for Σ'_{O_1}: $\Sigma'_{O_1} = QDQ^T$;
4) Calculate the projection matrix W: $W = P^TQ$

It can be seen that the W satisfies Eq. 3. The columns of W are the optimal CSP projection vectors. Typically one uses only the vectors with the highest eigenvalues for the positive class (i.e. O_1) and those with the lowest eigenvalues for the negative class (i.e. O_2).

In the CSP method, EEG samples are treated irrespective of their time. Thus, it does not account for inconsistent spatial patterns in an instance (also referred to a trial) of motor imagery EEG. In view of this issue, we propose a time-variant Spatial Filtering in the following section.

III. TIME-VARIANT SPATIAL FILTERING

Let the multi-channel signal from a bandpass filter be $\mathbf{x}(t)$ where t denotes the time. The idea of time-variant spatial filtering is to find a dynamic spatial filter $\mathbf{w}(t)$ which enhances the variance of one class while suppressing the variance of the other class. And we define the integrated variance measure as follows.

$$y(\mathbf{w}) = \frac{1}{L}\int E[\mathbf{w}(t)^T\mathbf{x}(t) - E[\mathbf{w}(t)^T\mathbf{x}(t)]]^2 dt \tag{4}$$

where L is the length of a motor imagery trial.

Let's say that an optimal $\mathbf{w}(t)$ functional should maximize the variance for class 1 whereas the variance for both class together should remain constant. Hence, finding the optimal functional $\mathbf{w}(t)$ for class O_1 against class O_2 can be expressed as a constraint optimization problem:

$$\max_{\mathbf{w}} y^{\{O_1\}}(\mathbf{w}) \text{ s.t. } y^{\{O_1\}}(\mathbf{w}) + y^{\{O_2\}}(\mathbf{w}) = 1 \tag{5}$$

Though numerous forms of functional $\mathbf{w}(t)$ are possible, in this work we consider a linear function because of its simplicity.

$$\mathbf{w}(t) = \mathbf{w}_0 + t \cdot \mathbf{w}_1 \tag{6}$$

The EEG signal is given as a sequence of discrete-time samples: $\mathbf{x}_t$. And the variance measure can be reformulated as

$$
\begin{aligned}
y(\mathbf{w}) &= \frac{1}{L}\sum_t E[(\mathbf{w}_0 + t\cdot\mathbf{w}_1)^T\mathbf{x}(t) - E[\cdot]]^2 \\
&= \frac{1}{L}\sum_t (\mathbf{w}_0 + t\cdot\mathbf{w}_1)^T\Sigma_t(\mathbf{w}_0 + t\cdot\mathbf{w}_1) \\
&= \frac{1}{L}\Big[\mathbf{w}_0^T\sum_t \Sigma_t\mathbf{w}_0 + 2\mathbf{w}_0^T\sum_t t\Sigma_t\mathbf{w}_1 + \\
&\quad \mathbf{w}_1\sum_t t^2\Sigma_t\mathbf{w}_1\Big] \tag{7}
\end{aligned}
$$

where Σ_t is the covariance of $\mathbf{x}_t$ at time t.

To maximize the objective function Eq. 5 with the variance measure expressed in Eq. 7, we derive an iterative optimization algorithm in below.

IV. OPTIMIZATION ALGORITHM

The optimization task is to search for optimal $\mathbf{w}_0$ and $\mathbf{w}_1$ that satisfy the Eq. 5. To this end, we first rewrite the variance function Eq. 7 in two forms: the function of $\mathbf{w}_0$ and the function of $\mathbf{w}_1$.

$$y(\mathbf{w}_0) = \mathbf{w}_0^T A_0\mathbf{w}_0 + \mathbf{b}_0^T\mathbf{w}_0 + c_0 \tag{8}$$

where

$$
\begin{aligned}
A_0 &= \frac{1}{L}\sum_t \Sigma_t \\
\mathbf{b}_0 &= 2\sum_t t\Sigma_t\mathbf{w}_1 \\
c_0 &= \mathbf{w}_1^T\sum_t t^2\Sigma_t\mathbf{w}_1
\end{aligned}
\tag{9}
$$

And

$$y(\mathbf{w}_1) = \mathbf{w}_1^T A_1\mathbf{w}_1 + \mathbf{b}_1^T\mathbf{w}_1 + c_1 \tag{10}$$

where

$$
\begin{aligned}
A_1 &= \frac{1}{L}\sum_t t^2\Sigma_t \\
\mathbf{b}_1 &= 2\sum_t t\Sigma_t\mathbf{w}_0 \\
c_1 &= \mathbf{w}_0^T\sum_t \Sigma_t\mathbf{w}_0
\end{aligned}
\tag{11}
$$

For a given $\mathbf{w}_1$(or $\mathbf{w}_0$), $y(\mathbf{w}_0)$(or $y(\mathbf{w}_0)$) is a quadratic function of $\mathbf{w}_0$(or $\mathbf{w}_1$). From Eq. 5, searching for the optimal spatial filter becomes a dual quadratically constrained quadratic programming (QCQP) problem. For example, if we want to look for optimal filter for class O_1, there are two objectives with respect to $\mathbf{w}_0$ and $\mathbf{w}_1$ respectively.

$$\max_{\mathbf{w}_0} \left\{ \mathbf{w}_0^T(A_0^{\{O_1\}})^T\mathbf{w}_0 + (\mathbf{b}_0^{\{O_1\}})^T\mathbf{w}_0 + c_0^{\{O_1\}} \right\} \text{ subject to}$$
$$\mathbf{w}_0^T(A_0^{\{O\}})^T\mathbf{w}_0 + (\mathbf{b}_0^{\{O\}})^T\mathbf{w}_0 + c_0^{\{O\}} = 1 \tag{12}$$

where $A_0^{\{O\}} = A_0^{\{O_1\}} + A_0^{\{O_2\}}$, $\mathbf{b}_0^{\{O\}} = \mathbf{b}_0^{\{O_1\}} + \mathbf{b}_0^{\{O_2\}}$, and $c_0^{\{O\}} = c_0^{\{O_1\}} + c_0^{\{O_2\}}$.

The other objective for $\mathbf{w}_1$ takes similar form and is omitted here to save space.

Because the 2nd-order coefficient matrices here are all positive definite, the problem is convex and can be readily solved using semidefinite programming (see Section IV.B).

To obtain both optimal $\mathbf{w}_0$ and $\mathbf{w}_1$, we devise the following iterative algorithm.

1) Initilize $\mathbf{w}_0$. This can be done using the same dual diagonalization method used in conventional CSP, except the coefficient matrix is slightly different.
2) Calculate the quadratic coefficients in Eq 12
3) Calculate $\mathbf{w}_1$ using QCQP.
4) Compute the change δ of the variance function Eq. 7. If $\delta < \zeta$ where ζ is a preset small value, stop.
5) Calculate the quadratic coefficients in Eq 10.
6) Calculate $\mathbf{w}_0$ using QCQP.
7) Compute the change δ of the variance function Eq. 7. If $\delta < \zeta$ stop.
8) Go to step 2.

V. Perliminary experiment and Results

Three healthy subjects participated in our preliminary experiments. The subjects were seated in a comfortable armchair, and during the course of the data collection sessions, remained still – without large body movements. We employed a 40-channel EEG amplifier (termed Neuroscan-40) from Compumedics. The sampling rate was 250Hz.

In each session, the subjects were asked to do 20 groups of left hand presses and 20 groups of right hand presses. In each group the subjects were intructed to press 3 times a computer keyboard.

Totally 15 electrodes channels around sensorimotor areas were selected for data analysis, including 'F3', 'Fz', 'F4', 'FC3', 'FCz', 'FC4', 'C3', 'Cz', 'C4', 'CP3', 'CPz', 'CP4', 'P3', 'Pz', 'P4'. The 4-second data corresponding to left/right hand press were segmented after the filtering process mentioned above.

We compared the conventional CSP and the proposed time-variant spatial filtering using 10 fold cross validation. We employed the SeDuMi toolbox (developed by MacMaster University,Canada. http://sedumi.mcmaster.ca/) to perform QCQP. As the features from CSP or our spatial filtering are sensitive to filter bands, we just tested the mu (8-12Hz), low-beta (12-15Hz), middle beta (15–18Hz), and the full frequence range (i.e. without band pass filtering). And we picked up the best band which gave rise to the highest classification accuracy, for each subject and each method respectively.

The energy values of EEG after filtering were calculated and served as the inputs to a classifier. In this study we used a naive Bayesian parzen window method for the classification between left and right hand press.

The final results in terms of statistics of classification error are given in Table I.

A. Discussions

The preliminary results suggest that the proposed method can produce more effective spatial filtering, where it yielded

	CSP		Our Method	
	mean	std	mean	std
Subject 1	18.5	7.7	17.6	6.6
Subject 2	21.7	7.7	19.55	7.3
Subject 3	24.6	7.3	22.9	8.7

TABLE I

COMPARATIVE CLASSIFICATION ERROR RATE (%).

up to approximately 10% relative reduction in classification error rate. In this study we did not explicitly optimize the frequency band selection for individual subjects is crucial for highly accurate motor imagery classification, though it's widely accepted that the system performance is sensitive to the filter band. It's interesting to note that a novel method was presented in[11] which automatically and simultaneously optimized a CSP and a spatial filter. Currently we are studying on how to effectively optimize our time-variant spatial filtering and spatial filters simultaneously.

VI. Conclusion

In this paper we proposed a novel approach to spatial filtering for motor imagery classification. By allowing time-variant spatial filters, the approach accounts for relatively higher order temporal dynamics in EEG. Furthermore, we derived an optimization algorithm for searching for optimal 1st order time-variant spatial filters using quadratically constrained quadratic programming. Our preliminary experiment indicates that, compared with conventional CSP, the proposed method is able to produce higher classification accuracy.

References

[1] J.R. Wolpaw, N. Birbaumer, D.J. MacFarland, G. Pfurtscheller, and T.M. Vaughan. Brain-computer interface for communication and control. *Clinical Neurophyiology*, 113:767–791, 2002.

[2] J. Decetry, D. Perani, M. Jeannerod, V. Bettinardi, B. Tadardy, R. Woods, J. Mazziotta, and F. Fazio. Mapping motor representations with positron emission tomography. *Nature*, 371:600–602, 1994.

[3] G. Pfurtscheller and A. Aranibar. Event-related cortical desynchronization detected by power measurements of scalp eeg. *Electroencephalography Clinical Neurophysiology*, 42:817–826, 1977.

[4] S. Lemm, C. Schafer, and G. Curio. BCI competition 2003 – data set III: Probabilistic modeling of sensorimotor μ rhythms for classification of iamginary hand movements. *IEEE Trans. on Biomedical Engineering*, 51(6), 2004.

[5] G. Pfurtscheller, C. Neuper, D. Flotzinger, and M. Pregenzer. Eeg-based discrimnation between imagination of right and left hand movement. *Clinical Neurophysiology*, 103:642–651, 1997.

[6] B. Blankertz, K.-R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlogl, G. Pfurtscheller, J. del R. Millan, M. Schroder, and N. Birbaumer. The bci competition iii: Validating alternative approaches to actual bci problems. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2):153–159, 2006.

[7] J. Muller-Gerking, G. Pfurtscheller, and H. Flyvbjerg. Designing optimal spatial filtering of single trial EEG classification in a movement task. *Clinical Neurophysiology*, 110:787–798, 1999.

[8] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Transactions on Rehabilitation Engineering*, 8(4):441–446, 2000.

[9] G. Dornhege, B. Blankertz, G. Curio, and K.-R. Müller. Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms. *IEEE Transactions on Biomedical Engineering*, 51(6):993–1002, 2004.

[10] S. Lemm, B. Blankertz, G. Curio, and K.-R. Müller. Spatio-spectral filters for robust classification of single trial EEG. *IEEE Transactions on Biomedical Engineering*, 52(9):1541, 2005.

[11] G. Dornhege, B. Blankertz, M. Krauledat, F. Losch, G. Curio, and K.-R. Müller. Optimizing spatio-temporal filters for improving brain-computer interfacing. In *Advances in Neural Information Processing Systems*, 2006.