Proceedings of the 7th IASTED International Conference
Biomedical Engineering (BioMED 2010)
February 17 - 19, 2010 Innsbruck, Austria

OPTIMIZING EEG CHANNEL SELECTION BY REGULARIZED SPATIAL FILTERING AND MULTI BAND SIGNAL DECOMPOSITION

Mahnaz Arvaneh[1, 2], Cuntai Guan[2], Kai Keng Ang[2] and Chai Quek[1]

[1]School of Computer Engineering, Nanyang Technological University, Singapore
[2]Institute for Infocomm Research, Agency for Science,Technology and Research (A*STAR), Singapore
[1]Email: mahn0001@ntu.edu.sg

ABSTRACT
Appropriate choice of number of electrodes and their positions are essential in Brain-Computer Interface applications since using less electrodes collects insufficient information for classification purposes whereas using more collects redundant information that could degrade BCI performance. This paper proposes a novel method of optimizing EEG channel selection by using the regularized Common Spatial Pattern (CSP) algorithm to discard redundant channels and multi band signal decomposition to select subject-specific frequency range. The performance of the proposed method is compared with EEG channel selection using Fisher criterion, mutual information, support vector and CSP on 9 subjects for two motor imagery tasks. Experiment results show the proposed method yields the highest accuracy in selecting 4 to 10 channels compared with the methods studied as well as using all the channels. The results also illustrate the proposed method significantly improves by multi band filtering and can achieve an average of 47% reduction of channels with only an averaged drop of 1.04% in classification accuracy.

KEY WORDS
Brain-Computer Interface, EEG channel selection, Regularized spatial filter

1. Introduction

A brain-computer interface (BCI) is a system for measuring, decoding and analyzing neural signals from the brain to help people with severe motor disabilities to use their brain signals for communication and control of objects in their environment [1, 2]. Among the various noninvasive BCIs, Electroencephalogram (EEG) is commonly used because it involves less expensive equipments. However, the use of EEG-based BCI is challenging due to the poor resolution of EEG and its multi-channel nature in the acquisition of brain signals[3]. Selecting too few channels could result in inoperability of the BCI due to insufficient information, and selecting too any channels could include noisy and redundant channels that degrade BCI performance.

One method to improve the performance of EEG-based BCI is to use the appropriate number of channels and appropriate positions on the scalp to acquire required brain signals. However good recording positions differs from patient to patient. In practical applications using a large number of electrodes suffers patients and needs intensive computations. Moreover reducing noisy and redundant electrodes may increase the accuracy [4]. Hence finding an intelligent method for optimal channel selection to achieve the highest accuracy with any given number of channels is critical for practical purposes.

Applying some recursive algorithms based on classifier such as SVM [4] or optimizing some criterion such as mutual information between channels and class labels EEG based Brain-Computer [5] are some popular methods for channel selection finding in literature. The performance of proposed method in [4] depends on accuracy of the applied classifier and properties of the features coming from channels. On the other hand the proposed method in [5] is a type of feature ranking methods working independent of classifiers.

The Common Spatial Pattern (CSP) algorithm [6] is shown to be effective in discriminating two classes of EEG measurements in BCI applications. Hence the spatial pattern coefficients is proposed in [7] for EEG channel selection. Since EEG are generally noisy from the contamination of various artifacts, channel selection using the CSP algorithm could result in poor overall accuracy of the BCI if the EEG are unfiltered or have been filtered with an inappropriately selected frequency range [8].

There are some researches on sparse representation and factorization of CSP [9, 10]. The authors in [9] proposed a regularized form of CSP to reduce EEG channels. They applied one of the sparse spatial filters on EEG signals which were filtered into 8 to 35 Hz and showed producing suitable sparse spatial filter can reduce EEG channels with a small effect on the accuracy of classification.

This paper seeks to find an algorithm for optimally selecting a few desired number of channels. Hence first an optimum multi band signal decomposition filter is proposed to reduce noise by identifying the subject-specific frequency range, then reducing redundant and useless channels is performed by introducing a regularized spatial filtering encouraged sparsity in both spatial filters corresponding to two classes. Next a pair-wise channel selection is proposed based on the CSP estimated from reduced channels.

The remainder of this paper is organized as follows: Section 0 describes the proposed method of optimizing EEG channel selection by regularized spatial filtering and multi band signal decomposition. Section 0 and 4 presents a comprehensive comparison between different channel selection methods in the literature and the proposed method. Section 5 concludes the paper.

2. Method

General structure of proposed method for optimizing EEG channel selection is shown in Fig.1. In this method, first multi band signal decomposition filtering is applied to full channel EEG signals, thereafter proposed regularized CSP reduces redundant channels and finally pair-wise channel selection is performed by CSP on remained channels.

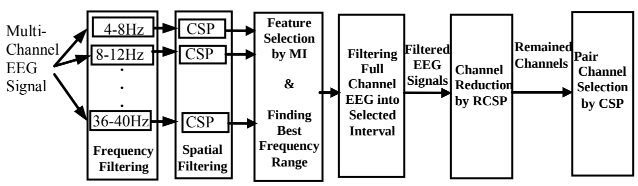

Fig.1. Proposed method of optimizing EEG channel selection

2.1 Filtering by multi band signal decomposition

The subject specific multi band filter used in our method [11] comprises four progressive stages: As it can be seen in Fig.1 the first stage employs a filter bank for bandpass filtering of EEG measurements into multiple bands using zero-phase Chebyshev Type II. The second stage performs spatial filtering on each of these bands using the CSP algorithm. Thus, each pair of bandpass and spatial filters yields CSP features which are specific to the frequency range of the bandpass filter. Third stage employs a feature selection algorithm based on Mutual Information based Best Individual Features (MIBIF) [11] to select the best discriminative CSP features from the filter bank. During fourth stage an Elliptic bandpass filter, filters the raw EEG signals into the frequency range specified by the best discriminative CSP features. This multiband decomposition reduces noise by filtering EEG signals in appropriate subject-specific frequency range, therefore we hope for an improvement in channel selection results.

2.2 CSP and regularized CSP for channel selection

CSP [6] is an effective technique to discriminate between two classes of multichannel data. The aim of CSP is to project raw EEG signal to a filtered signal Z as given in (1) which maximizes the variance of one class while the variance of the other class is minimized.

$$Z = WX \tag{1}$$

$X \in R^{N \times T}$ is a matrix representing the EEG signal of a single trial; N and T are the number of channels and the number of measurement samples per channel respectively. The rows of projection matrix, W, are the stationary spatial filters and the columns of W^{-1} are the common spatial patterns. Spatial patterns derived from CSP method can be seen as EEG source distribution vectors. In [7] authors assumed that the first and last spatial pattern vectors specify which channels have been correlated with the performed tasks, whereas channels corresponding to the maximum coefficients of them would be the most correlated channels with the corresponding task. As our experiment results will show due to use of two CSP weight vectors (first and last), channels from both related brain areas are selected. Although this feature helps the method achieving better results, selected neighbour channels have much redundant information and reduce the performance.

In this paper a regularised CSP (RCSP) algorithm is proposed to improve the results. Let $X=\{X_1, X_2, ..., X_T\}$ where $X_i \in R^{N\times T}$ has been centered, scaled and denotes the i^{th} trial of EEG signal. The CSP problem can be expressed as

$$\begin{aligned} &\underset{w}{Minimize} \quad \sum_{i\in C_1} \mathrm{var}(wX_i) \\ &Subject \quad to \quad \sum_{i=1}^{T} Var(wX_i) = 1 \end{aligned} \tag{2}$$

where C_1 represents all EEG trials of first class and $w \in R^N$ is the unknown weight vector of the spatial filter. With only the first constraint in (2), the optimization problem reduces to original CSP expressed by Reyleigh coefficient [6]. The cost function in equation (2) can be represented by using variance definition as (3)

$$\begin{aligned} \mathrm{var}(wX_i) &= wE\left\{\left(X_i - E\{X_i\}\right)\left(X_i - E\{X_i\}\right)^T\right\} w^T \\ &= w\Sigma_i w^T \end{aligned} \tag{3}$$

Σ_i is mean covariance matrix for the signals belonging to C_i sets. In this study to minimize number of hired electrodes, which is minimizing the number of non zero entries of w vector, a regularized term is added as bellow

$$\begin{aligned} &\underset{w}{Minimize} \quad w\sum_i w^T + p\frac{\|w\|_1}{\|w\|_2} \\ &Subject \quad to \quad \sum_{i=1}^{m} w\sum_i w^T = 1 \end{aligned} \tag{4}$$

where m is the number of classes and $p>0$ is a regularization parameter. The minimum value of the regularized term as shown in (5) is equal to zero when all elements of w vector (w_i) except one of them become zero. Our proposed regularization term can be applied to parsity problems instead of $l_0 - norm$ which is an NP-hard problem [9].

$$\frac{\|w\|_1}{\|w\|_2} = \left(1 + \frac{2\sum_{i=1}^{N}\sum_{j>i}^{N} w_i w_j}{\sum_{i=1}^{N} w_i^2}\right)^{0.5} \tag{5}$$

The proposed optimization problem clearly depends on p value which specifies a tradeoff between number of reduced channels and decreased accuracy of classification. So p is optimally defined by a parallel optimization problem with (4). This optimization problem maximizes the number of removed channels subject to keeping the classification accuracy more than a predefined threshold (in our study 0.95 of full channel accuracy in training data has been selected).

To solve (4) we used iterative nonlinear optimization toolbox of Matlab which computes a quasi-Newton approximation. Furthermore the optimized subject specific value of p is selected by varying from 0.00 to 0.50. For the body of your document, use Times New Roman font, 10-point type size, single-spaced. The whole document should be fully justified (not only left-justified). Headings should be 12-point, upper- and lower-case, bold or 10pt upper case, bold. Subheadings should be 10-point upper- and lower-case.

3. Experiments

3.1 Data description

The EEG data used in this study consisted of two classes: right and left hand motor imageries. They were provided by BCI Competition IV, datasets 2a [13]. The EEG signals were recorded from nine subjects using 22 electrodes per subject. During each experiment, the subject was given visual cues that indicated four motor imageries should be performed: left hand, right hand, feet and tongue. Only EEG trials for right and left hand from 0.5 to 2.5 seconds after cue were provided. Each class of EEG signals consists of 140 trials.

3.2 Data pre-processing and channel selection methods

In our study first the raw EEG signals were filtered into the best subject-specific frequency range extracted by the multi band signal decomposition algorithm. Then each of right and left hand regularized spatial filters was obtained by minimizing (4). Next, channels whose coefficients were zero in both filters corresponding to two classes were removed. Finally remained channels were ranked by calculating CSP on their signals as follows:

Optimal channels for every motor imagery task are determined through the maximums of the absolute value of the concerned spatial pattern. Let SP_{Ri} and SP_{Hi} denote i[th] optimal channels of spatial pattern for right and left hand motor imagery respectively, therefore (6) is calculated to obtain overall ranking, where i varies from 1 to the number of remained channels. Finally as every channel has been iterated twice in CH, the lower rank is discarded. The proposed pair-wise channel selection results in selecting channels from both activated brain areas.

$$\begin{aligned} CH_{2i-1} &= Find(Max(\mathrm{SP}_{\mathrm{Ri}}, SP_{Li})) \\ CH_{2i} &= Find(Min(\mathrm{SP}_{\mathrm{Ri}}, SP_{Li})) \end{aligned} \tag{6}$$

To consider performance of the proposed method, some previous EEG channel selection methods using Fisher Criterion [4], Mutual Information [5], Support vector channel selection [4] and CSP channel selection without channel reduction [7] were applied to EEG multi band filtered data. In first three methods covariance of each channel was introduced as the feature coming from a channel.

3.3 Feature extraction and classification

In order to consider performance of the applied methods, accuracy of classification was estimated with different number of optimal channels. First the spatial filter was hired to project the signals, then variances of first and last rows of the projected signal [12] were determined as inputs of a SVM classifier. Finally, a 10×10-fold cross validation was used to estimate the accuracy of classification.

4. Results

In this study the sparsest spatial filters were selected by varying regularized parameter p given in (4) subject to keeping acceptable accuracy. In order to illustrate the positive effect of multi band filtering on the proposed method, RCSP channel reduction was applied on two groups of signals 1- multi band filtered EEG signals and 2- 8 to 35 Hz filtered EEG signals.

In table 1, first row presents the averaged 10×10 fold classification accuracy for full channel EEG. Averaged number of removed channels by RCSP for multi band filtered signals and achieved accuracy after removing those redundant channels are indicated on Second and third rows respectively. Finally averaged number of removed channels and achieved accuracy after removing them for 8 to 35 Hz filtered signals are presented in two last rows.

EEG signals except for subject 2, 4 and 6 were successfully classified with more than 75% accuracy. Results in table 1 show the regularized CSP on multi band filtered signals was significantly successful. In comparison with full channel results, it decreased the number of electrodes on average to 47% (of the 22 electrodes) while the average reduction in accuracy was only 1.04%. Also indicated results in table 1 proof significant effect of multi band filtering on improving the performance.

Table 1
Performance comparison of EEG channel reduction by RCSP after multi band and 8 to 35 Hz filtering (Ch: Channels, Acc: Accuracy)

Subject	1	2	3	4	5	6	7	8	9
Full Ch Acc (%)	87.3	56.8	93.1	63.6	87.6	62.6	77.1	94.2	93.8
#Removed Ch - Multi band	10	14	13	16	10	15	8	13	14
Acc of remained Ch (%)- Multi band	84.7	56.3	92.3	67	83.4	63.1	75.6	92.8	92.2
#Removed Ch - 8 to35 Hz	10	15	13	16	7	15	8	15	15
Acc of remained Ch (%)- 8 to 35 Hz	80.8	52.2	90.4	68.6	59.8	55	65.1	92.1	92.2

After channel reduction, we applied the proposed CSP channel selection explained in section 3 to rank remained channels.

Fig. 2 depicts achieved accuracy versus different number of channels (from 2 to 22) chosen by 4 different channel selection methods based on Fisher Criterion (FC) [4], Mutual Information (MI) [5], Support Vector Machine (SVM) [4], Common Spatial Pattern (CSP) [7] and compares the results with the proposed channel selection method.

As Fig.2 presents, it seems subject 2 could no perform motor imageries correctly so that the classification accuracy is around 50% and the results of channel selection are significantly scattered. Considering the other subjects concludes our proposed method and CSP channel selection method are capable of selecting relevant channels, whereas FC, MI and SVM methods fail for some subjects. Especially in selecting 4 to 10 channels our proposed method is almost superior over the others.

In fact, methods FC and MI perform rather a channel ranking than a channel selection. They rank channels individually without considering the relevancy between channels, so they cannot select a few channels well. Because for example two best individual channels do not necessarily make the best subset. The ranking methods such as applied MI and FC select channels without considering correlation and complementary information between channels.

As the accuracy of applied SVM method strongly depends on performance of the classifier, it works unsuccessfully in selecting a few number of channels. As it is visible in Fig.2, a sharp decrease of accuracy around 6 to 2 channels is obvious for hired SVM method.

Visualization of the channel positions according to their ranks may support the analysis of our applied methods. As the experimental paradigm is well known we investigated whether the best selected channels are those situated over or close to motor areas.

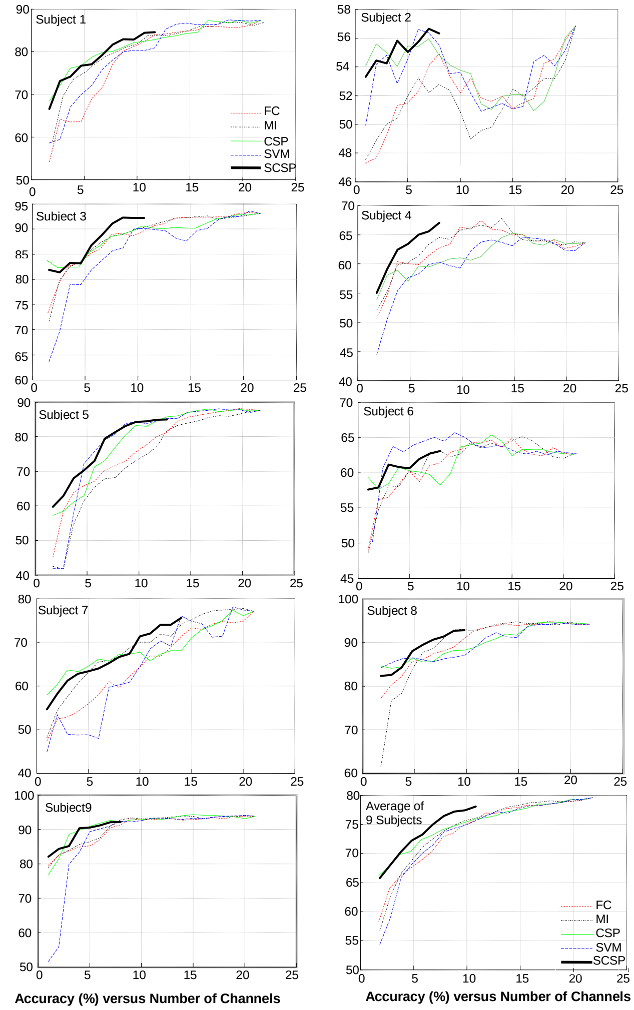

Fig.2. Comparison of 5 EEG channel selection methods

Fig. 3 visualized the selected channels of subject 1 for five considered methods where darker colors show more important channels which are selected earlier and lighter ones show less important channels. It should be noted in this step cross validation was not applied.

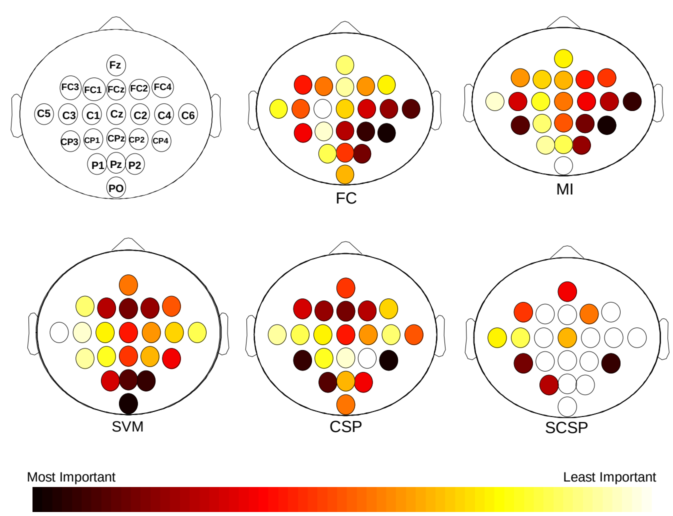

Fig.3. Visualization of channels importance for subject 1

In FC method the best channels are neighbor and just in one side of brain (down-right in subject 1), so selecting a few number of channels are full of redundant information without supporting both task activities. MI channel positions are a bit better than FC but still quiet near to each others. As it can be seen in Fig. 3 FC and MI methods may perform quietly well as a channel removal. In subject 1, SVM recognized top and down of brain channels as the most important ones. The preference of CSP based method is selecting channels pair wisely from both sides of brain. In subject 1, best channels are CP4 and CP3 but after a while some channels from top are selected. SCSP channel selection by maintaining the advantages of CSP method, selects just some of neighbor electrodes thus redundant information is reduced and performance increases.

5. Conclusion

In this study, we focused on demonstrating an intelligent method to select the best given number of electrodes with an acceptable reduction in accuracy. Also our method specifies optimal number and positions of electrodes for a practical application in order to achieve the best trade of between number of channels used and BCI performance. These abilities are achieved by introducing a regularized spatial filter algorithm to remove redundant channels thereafter applying a CSP based algorithm to select given number of channels from remained ones. Meanwhile a subject specific multi-band filter, filters the raw EEG signal to reduce noise and increase the performance.

It was demonstrated that a suitable estimation of regularization parameter p can reduce the averaged number of electrodes from 22 to 12 whereas the classification accuracy decrease is only 1.04%. A comprehensive comparison between the proposed method and previous channel selection methods using Fisher criterion, mutual information, support vector and CSP on 9 subjects for two motor imagery tasks showed superior capability of proposed method in selecting a few given number of channels especially in selecting around 4 to 10 channels. Visualization of the electrode positions illustrated our method improves the results by removing some of neighbor channels and selecting those from both sides of brain.

References

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

[2] N. Birbaumer, "Brain-computer-interface research: Coming of age," *Clin Neurophysiol, vol. 117*, 2007, pp. 479- 483.

[3] A. Al-Ani and A. Al-Sukker, "Effect of Feature and Channel Selection on EEG Classification," in *Proc. IEEE/ EMBS'06*, 2006, pp. 2171-2174.

[4] T. N. Lal, M. Schroder, T. Hinterberger, J. Weston, M. Bogdan, N. Birbaumer, and B. Scholkopf, "Support vector channel selection in BCI," *IEEE Transactions on Biomedical Engineering, vol. 51*, 2004*,* pp. 1003-1010.

[5] L. Tian, D. Erdogmus, A. Adami, M. Pavel, and S. Mathan, "Salient EEG Channel Selection in Brain Computer Interfaces by Mutual Information Maximization," *Proc. IEEE/ EMBS '05*, 2005, pp. 7064-7067.

[6] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. R. Muller, "Optimizing Spatial filters for Robust EEG Single-Trial Analysis," *IEEE Signal Processing Magazine, vol. 25*, 2008, pp. 41-56.

[7] Y. Wang, S. Gao, and X. Gao, "Common Spatial Pattern Method for Channel Selelction in Motor Imagery Based Brain-computer Interface," *Proc. IEEE/ EMBC'05*, 2005, pp. 5392-5395.

[8] Q. Novi, C. Guan, D. Tran Huy, and X. Ping, "Sub-band Common Spatial Pattern (SBCSP) for Brain-Computer Interface", *3rd International IEEE/EMBS Conference on Neural Engineering,* 2007, pp. 204-207.

[9] X. Yong, R. K. Ward, and G. E. Birch, "Sparse spatial filter optimization for EEG channel reduction in brain-computer interface," in *IEEE/ ICASSP 2008*, pp. 417-420.

[10] J. Farquhar, N. J. Hill, T. N. Lal, and B. Schölkopf, "Regularised CSP for Sensor Selection in BCI," *Proceedings of the 3rd International Brain-Computer Interface Workshop and Training Course,* pp. 14- 15, 2006.

[11] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan, "Filter Bank Common Spatial Pattern (FBCSP) in Brain-Computer Interface," in *Proc. IEEE/ IJCNN '08*, 2008, pp. 2390-2397.

[12] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering, vol. 8*, 2000, pp. 441-446.

[13] “Data Sets 2a for the BCI Competition IV”, http://www.bbci.de/competition/iv/#datasets.