Spatial Filter and Feature Selection Optimization based on EA for multi-channel EEG

Yubo Wang[1], Krithikaa Mohanarangam[1], Rammohan Mallipeddi[1*] and K. C. Veluvolu[1]

Abstract— **The EEG signals employed for BCI systems are generally band-limited. The band-limited multiple Fourier linear combiner (BMFLC) with Kalman filter was developed to obtain amplitude estimates of the EEG signal in a pre-fixed frequency band in real-time. However, the high-dimensionality of the feature vector caused by the application of BMFLC to multi-channel EEG based BCI deteriorates the performance of the classifier. In this work, we apply evolutionary algorithm (EA) to tackle this problem. The real-valued EA encodes both the spatial filter and the feature selection into its solution and optimizes it with respect to the classification error. Three BMFLC based BCI configurations are proposed. Our results show that the BMFLC-KF with covariance matrix adaptation evolution strategy (CMAES) has the best overall performance.**

I. Introduction

Brain-computer interface (BCI) is defined as an alternative communication pathway that translates the measured brain activity into control commands. In EEG based BCI systems, different control commands are associated with the EEG signals that differs in either mental strategies or in intensity level of a mental task [1].

It has been shown that the class information can be decoded from frequency domain of corresponding EEG signal, many methods have been employed for this purpose [2]. The EEG signal employed in BCI applications is generally band-limited. To decompose such kind of signal into frequency domain and keep the balance between temporal and frequency resolution, the band-limited multiple Fourier linear combiner (BMFLC) has been developed [4]. As it can provide optimal time and frequency resolution, its application to BCI has been successful [3] .

In a nutshell, the BMFLC is employed to adaptively estimate amplitude of the EEG signal in a pre-defined frequency band. Then, the obtained amplitude estimation from all the channels of interest are cascaded to form a feature vector. To improve the performance of BMFLC based BCI system, a visual inspection was conducted to find a narrower sub-band. Due to the involvement of the visual inspection, only EEG data from C3 and C4 are considered [2]–[4].

The performance of the BMFLC based BCI system can be improved by further incorporating data from multiple EEG channels. However, the involvement of multiple channels EEG causes two issues in the implementation: 1) Due to the volume conduction of skull, scalp and various noise sources, the EEG signal generated by the cortex of interest may not be directly related to the targeted neural activity. Hence, EEG signal purification is required before its any further usage [5]; 2) Application of BMFLC to handle multiple channels EEG data can increase the dimensionality of the feature vector, feature subset selection is required to stabilize the performance of the classifier.

The popular method for handling the first problem is to employ a spatial filter algorithm. As labeled EEG data is available in BCI experiment, to explore the task-related information via labeled training data, common spatial filter (CSP) has been developed [6]. With the assumption that the variance of EEG conveys the class information, CSP tries to find a projection that maximizes signal variance in one class while minimizing the variance in the other class [6].

In a typical implementation of CSP [6], EEG signal is band-pass filtered in the band of interest. Applying CSP to the time-frequency features is problematic, as it requires to estimate spatial filter for each frequency components separately. This implementation may lead to amplitude boosting of irrelevant frequency components, thus jeopardizing the performance of BCI system.

In order to apply BMFLC for multiple EEG channel BCI systems, we propose to solve the spatial filter and feature selection problem with evolutionary algorithm. By incorporating a real-valued EA algorithm, we are able to directly optimize the spatial filter with respect to the classification accuracy. The second problem can be handled by encoding the feature selection procedure into the solution vector of the applied EA. Then, the solution vector, which is provided by the EA at the end of its execution, should provide us the estimated spatial filter and the selected feature subspace simultaneously. Application of the proposed method to movement imagery data highlight the performance.

II. Methods

This section discusses in brief the methods employed in this paper. Three different configurations based on BMFLC are provided for feature extraction. Then the necessary detail of various techniques are provided in the following subsection.

A. BMFLC based BCI System

According to the placement of BMFLC-KF in the signal processing chain, three different configurations have been proposed in this work. The structural diagram is given in Fig.1. The signal path in solid line is used in both training and testing phase whereas the dotted line is only used in training phase.

[1]College of IT Engineering, Kyungpook National University, 1370 Sanyuk-dong, Daegu, South Korea 702-701
[*]mallipeddi.ram@gmail.com

978-1-4244-9270-1/15/$31.00 ©2015 IEEE

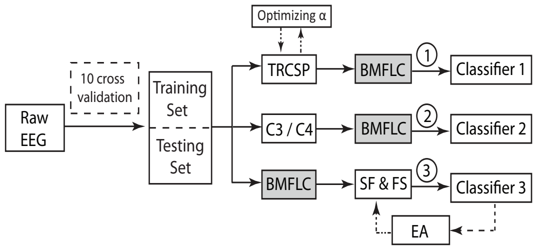

Fig. 1. Architecture of signal processing chain

Common to all three methods, the raw EEG data is divided into training and testing sets according to the selected cross validation scheme. The classifier and spatial filter are trained solely based on the training data. In the first configuration, the Tikhonov regularized CSP (TRCSP) is applied to the band-pass filtered training EEG signal. BMFLC-KF is then applied to the spatial filtered data to form the feature vectors for the classifier. In the second configuration, the electrodes which are considered relevant to the class are selected manually, *i.e.* C3 and C4 corresponds to the left and right hand movement imagery. The third configuration employs BMFLC-KF to first decompose all available EEG signal. Then the EA is employed to find spatial filter (SF) and frequency selection (FS) on the obtained BMFLC weights simultaneously. We can observe that in the configurations 1 and 2, the channel reduction is carried prior to the application of BMFLC. Whereas in configuration 3, the BMFLC is applied to all EEG channels which preserves the maximum amount of information in frequency domain.

B. BMFLC-KF based EEG pre-processing

The previously developed BMFLC divides a pre-defined frequency band $[\omega_1, \cdots, \omega_n]$ into n equally distributed divisions with frequency spacing Δ_f, and estimates the amplitude of each frequency component by using Kalman filter [3]. A brief description of BMFLC-KF is given below.

We first re-write the truncated Fourier series in state-space form as:

$$y_k = \mathbf{x}_k^T \mathbf{w}_k + v_k \tag{1}$$

$$\mathbf{w}_{k+1} = \mathbf{w}_k + \eta_k \tag{2}$$

where $\mathbf{x}_k = [\sin(\omega_1 k), \cdots, \sin(\omega_n k), \cos(\omega_1 k), \cdots, \cdots]^T$ and $\mathbf{w}_k = [a_{1k}, \cdots, a_{nk}, b_{1k}, \cdots, b_{nk}]^T$ are frequency of interests and its corresponding amplitude receptively. T denotes the matrix transpose.

Assume that v_k and η_k are independent Gaussian process with 0 mean and covariance of R and $\mathbf{Q}$ respectively. The Kalman filter iteration is used to find the weights estimation $\hat{\mathbf{w}}_k = \mathbf{E}[\mathbf{w}_k|\mathbf{y}_{k-1}]$ [3]. The weight vectors of BMFLC represents the Fourier coefficients of the band-limited signal. The estimated weights are further combined as:

$$\mathbf{w}_k^f = \left[\begin{array}{ccc} \sqrt{a_{1k}^2 + b_{1k}^2} & \cdots & \sqrt{a_{nk}^2 + b_{nk}^2} \end{array} \right]^T \tag{3}$$

where $\mathbf{w}_k^f$ is the absolute weight vector of the frequency components at time instant k. The time-frequency mapping of a given signal is obtained by cascading the weights in $\mathbf{D} = [\mathbf{w}_1^f, \ldots, \mathbf{w}_n^f]$.

C. Covariance matrix adaptation evolution strategy (CMAES)

CMAES, is a real-valued evolutionary algorithm, explores the second order statistics estimated from the solution space in order to find the optimal value of the target function. It is well suited for optimizing the function with complicated structure. It has been applied to both artificial functions and real practical problems [7].

The (μ, λ) CMAES requires λ number of solutions in each generation and selects the best μ number of solutions for updating the value in next step. The CMAES algorithm generates the solution $s_k^{(g)}$ by sampling from a Gaussian distribution with mean m, step-size δ and covariance estimation $\mathbf{C}$, where g and k are indexes for generation and solution respectively. To find the solution for next generation, CMAES estimates the mean, step-size and covariance with the solutions in the current generation, by considering the history of evolution path. After each generation, the solutions are ranked based on their fitness value. The update of mean is carried out by a weighted summation of the μ best solutions in the current generation. The value of $\mathbf{C}$ is determined by the covariance of the estimated evolution path denoted by p_c and the covariance estimated from the μ number of solutions in the current generation.

D. Optimization of spatial filter and frequency band with CMAES

CMAES permits us to encode the spatial filter and feature selection into a single solution vector. We then optimize the solution vector with respect to the performance measure of a BCI system, *i.e.* classification accuracy.

To optimize the spatial filter and frequency selection simultaneously, the solution vector for CMAES is constructed as:

$$s_k = [\overbrace{sf_{11}, \ldots, sf_{1M}, \quad sf_{k1}, \ldots, sf_{kM},}^{\text{Spatial Filter}} \quad \overbrace{FS, \quad BW}^{\text{Frequency Selection}}]$$

where sf_{ij} denotes the jth component in ith spatial filter, FS is the starting frequency and BW is the bandwidth. M is the number of total EEG channels. To combine the above procedure with the time-frequency mapping estimated from BMFLC, FS has to lie in the range of $[f_1, f_n - BW]$ where f_1 and f_n are the corner frequencies from BMFLC respectively. As shown in [4], majority of the subjects have a reactive frequency band less than 3Hz in bandwidth. Thus BW is constrained to have a maximum 3Hz band.

The classification error obtained from the linear discriminant analysis (LDA) is employed as the cost function for CMAES. The frequency weights obtained from each channel are combined to the following form:

$$\mathbf{W}_k = \begin{bmatrix} \mathbf{w}_{k,1}^{fs} & \cdots & \mathbf{w}_{k,M}^{fs} \\ \vdots & \vdots & \vdots \\ \mathbf{w}_{k,1}^{fe} & \cdots & \mathbf{w}_{k,M}^{fe} \end{bmatrix}^T \tag{4}$$

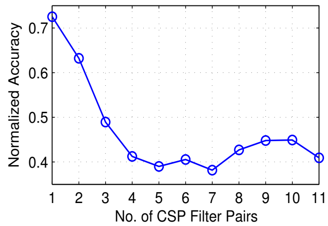

Fig. 2. Normalized Accuracy for 11 pairs of CSP filter averaged over all subjects

where $\mathbf{w}_{k,M}^{fs}$ indicates the weights of frequency fs at time instant k in channel M. $fs = FS$ and $fe = FS + BW$ are the selected frequency range by the employed EA. The spatial filter in the solution vector is extracted into the following form:

$$\mathbf{SF} = \begin{bmatrix} sf_{11} & \dots & sf_{1M} \\ sf_{21} & \dots & sf_{2M} \end{bmatrix} \tag{5}$$

Thus, the resultant feature vector after applying frequency selection and spatial filter is just the matrix multiplication between $\mathbf{SF}$ and $\mathbf{W}_k$. To avoid over-fitting, the classification error obtained from validation data is used as the cost function. The solution vector, which contains the spatial filter and frequency selection, is therefore optimized simultaneously.

III. Results

A. Parameter Selection

All three configuration that are developed in this paper are tested on BCI Competition IV dataset IIa [8]. The dataset contains 22 channel EEG data for 9 subjects performing motor imagery tasks. Only the hand motion classes are used in this work.

As the EEG α band is the region of interest for brain motor imagery, the frequency range for BMFLC-KF is set to $f_1 = 6$Hz and $f_n = 14$Hz. Δ_f is set to 0.5Hz as it has been shown to offer better results in EEG signal decomposition [3]. The CMAES implementation employed in this work is similar to [7]. The population size λ is set to 5 times the solution dimension. Other parameters for CMAES are left unchanged as it has shown to offer robust performance.

To obtain a reliable performance measure, the EEG data is divided into training and testing data based on 10-times cross validation. For optimizing the common spatial filter and applying CMAES, only the training set is used. The final classifier is also trained upon the same training data. The final classification accuracy is obtained by using testing data for the overall performance estimation.

B. Selecting the Optimal Number of Spatial Filters

As the dataset used in this work consists of 22 EEG recordings during the motor imagery tasks, theoretically we can obtain up to 11 pairs of common spatial filters. However, the spatial filters that correspond to large eigenvalue contain meaningful class information.

To determine the optimal number of spatial filter pairs that maximizes the overall performance, we apply TRCSP to each subject for all possible pairs. We first rank each common spatial filter based on their corresponding eigenvalue, then we add one pair of spatial filter at a time in the ranked sequency. For each subject, we then obtained classification accuracy for each possible spatial filter pairs. As the comparison has to be performed across subjects, the obtained 11 classification accuracies for each subject is normalized into the range [0 to 1]. The final result is the averaged of normalized classification accuracy across subjects and is shown in Fig.2. It is clear that the first pair of the spatial filter offers the best performance. The later added common spatial filter pairs deteriorate the classification accuracy significantly. Thus, we selected 1 pair of spatial filter for all methods in the following comparison.

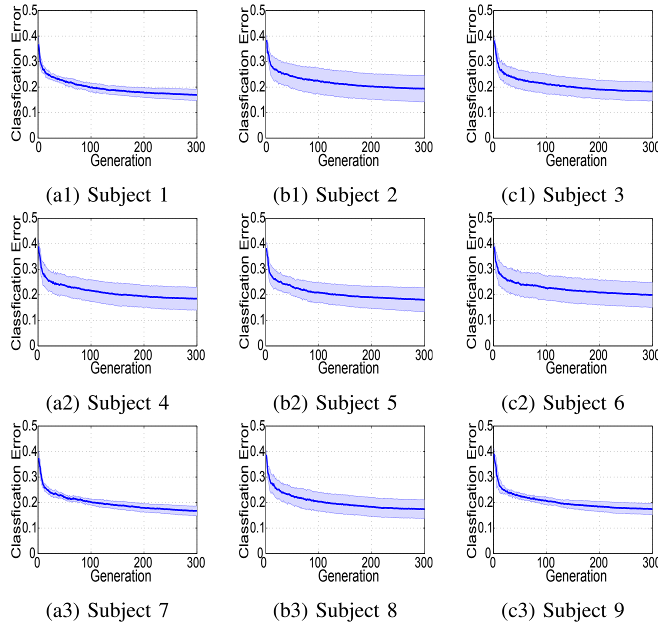

Fig. 3. Evolution of training of CMAES for 300 generations. Shaded area indicates the standard deviation obtained from 10 cross validation runs.

C. Performance CMAES based Spatial Filter and Frequency Selection Optimization

The value of the cost function of CMAES at each generation for each subject is shown in Fig.3. The shaded area indicates the standard deviation of the obtained classification error obtained from 10-times cross validation scheme. For all the subjects, we observe that the classification error converges at the end of the evolution process. It can be also noted that the CMA-ES improves the classification error rapidly in the first around 20 generations. It is then followed by a slow improvement up to 100 generations. After 100 generations, the classification error becomes stable. The simulations for 500 generations also yielded similar results.

Another observation from Fig.3 is that the standard deviation varies significantly across subject. A small standard deviation can be found in Fig.3(a1), (a3) and (c3) corresponding to subject 1, subject 7 and subject 9 respectively. Whereas a substantially large standard deviation can be observed in Fig.3(b1), (a2), (b2) and (c2) which correspond to

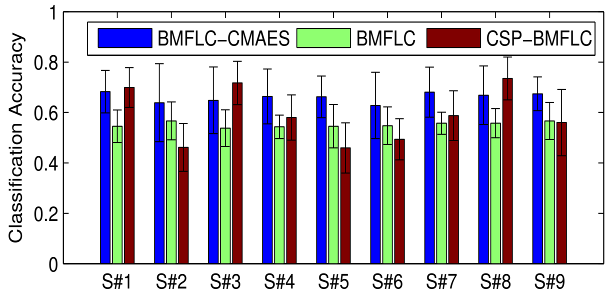

Fig. 4. Classification accuracy on testing set of all configurations

subject 2, subject 4, subject 5 and subject 6 respectively. An enlarged standard deviation indicates that CMAES converges to a different value. It also demonstrates the complexity of the cost function being employed. However, we also notice that this divergence in the evolution process does not affect the performance of the final classifier.

D. Performance of the Final Classifier for all Configurations

After obtaining spatial filter from TRCSP and CMAES, a final classifier is built by using the same training set and then evaluated on the testing set. Three configurations are denoted as BMFLC-CMAES, BMFLC and CSP-BMFLC respectively. The performance for each subject averaged over 10-times cross validation is shown in Fig.4.

It can be observed that BMFLC-CMAES outperforms other algorithms in 6 out of 9 subjects. It also has the best performance averaged over all subjects, and passes the Friedman test with a confidence level of $\alpha = 0.05$. The performance of CSP-BMFLC varies dramatically across subjects. This result indicates that the performance of CSP is highly subject dependent. Despite the unsatisfactory performance of applying BMFLC alone, it is the most stable one among all three configurations.

IV. Discussion and Conclusion

In this work, three configurations of BMFLC based BCI system have been proposed. As most of BCI systems require processing EEG signal in quasi real-time, therefore BMFLC-KF is used to decompose the EEG signal. Without any optimization, the solely applied BMFLC has the most stable performance. There are two possible explanations for this result. First, it is a confirmation that the BMFLC-KF can decompose the EEG signal with high accuracy which provides the classifier a stable feature set. Secondly, the stability of results may indicate the amount of information that can be extracted from just two channels is limited. This result also highlight the necessity of using spatial filter.

It is evident that the performance of CSP-BMFLC is subject dependent. For 3 out of 9 subjects, the CSP-BMFLC has good testing accuracy. However, the CSP-BMFLC performs similar or even worse than BMFLC alone in the remaining subjects. As CSP algorithm relies on the training data to find the spatial filter, the degenerated test performance may due to the over-fitting of the training accuracy.

The superior performance of BMFLC-CMAES as compared to other configurations highlight the ability of CMAES in optimizing complex function. Our results also suggest that it is preferred to estimate spatial filter and feature selection simultaneously.

Among all methods employed in this work, BMFLC-CMAES has the highest computational needs. Prior to applying evolutionary algorithm, EEG data from the whole montage have to be decomposed by using BMFLC-KF. Although the BMFLC in this work only considers the μ band, it is clear that simultaneous processing of data from increased number of channels poses a challenge in real-time data processing. In a practical system, the computational needs can be alleviated by employing the parallel implementation of BMFLC algorithm.

In this work, the combination of CSP and CMAES with BMFLC-KF is applied to a two-class BCI application. The CMAES is employed to optimize the spatial filter and feature selection simultaneously. The performances of BMFLC-CMAES, CSP-BMFLC and BMFLC alone are compared. Our results indicate that BMFLC-CMAES has the best performance across the subjects and is also very stable as compared to the other algorithms.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology under the Grant NRF-2014R1A1A2A10056145.

References

[1] G. Pfurtscheller, G. R. Muller-Putz, R. Scherer, and C. Neuper, "Rehabilitation with Brain-Computer Interface Systems," *Computer*, vol. 41, pp. 1–8, 2008.

[2] Y. Wang, K. C. Veluvolu, J.-H. Cho, and M. Defoort, "Adaptive estimation of EEG for subject-specific reactive band identification and improved ERD detection," *Neuroscience letters*, vol. 528, no. 2, pp. 137–142, 2012.

[3] Y. Wang, K. C. Veluvolu, M. Lee, and L. Min-Ho, "Time-Frequency Analysis of Band-limited EEG with BMFLC and Kalman Filter for BCI Applications," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, no. 109, p. 109, 2013.

[4] K. C. Veluvolu, Y. Wang, and S. S. Kavuri, "Adaptive estimation of EEG-rhythms for optimal band identification in BCI," *Journal of Neuroscience Methods*, vol. 203, pp. 163–172, 2012.

[5] 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. 41, pp. 41–56, 2008.

[6] F. Lotte and C. Guan, "Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms." *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 355–362, 2011.

[7] N. Hansen, a.S.P. Niederberger, L. Guzzella, and P. Koumoutsakos, "A method for handling uncertainty in evolutionary optimization with an application to feedback control of combustion," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 1, pp. 180–197, 2009.

[8] C. Brunner, M. Naeem, R. Leeb, B. Graimann, and G. Pfurtscheller, "Spatial filtering and selection of optimized components in four class motor imagery data using independent components analysis," *Pattern Recognition Letters*, vol. 28, pp. 957–964, 2007.