An Optimization of Spatio-Spectral Filter Bank Design for EEG Classification*

Masanao Obayashi[1], **Takuya Geshi**[2], **Takashi Kuremoto**[3] **and Shingo Mabu**[4]
Information Science and Design Engineering, Yamaguchi University,Tokiwadai 2-16-1
E-mail: {[1]m.obayas,[2]t015vk, [3]wu,[4]mabu}@yamaguchi-u.ac.jp

Abstract

How to select the appropriate frequency band to classify EEG signal by motor imagery is discussed in this paper. Our proposal is an improvement of the conventional Bayesian Spatio-Spectral Filter Optimization (BSSFO). Defect of BSSFO is on the way to generate the renewal particle of the filter bank, such a random number generation. To avoid a local optimum, an evolutional update method of particles is introduced. It is shown that performance of the EEG classification ability is improved.

Keywords: spatio-spectral filter, EEG, classification, .optimization, mutual information, common spatial filter

1. Introduction

Recently, researches using brain computer interface (BCI) have been actively studied. To precisely identify EEG signal, it is necessary to remove the artifact and noise by using appropriate spatial and spectral filter, Refs. 1-2. Furthermore, the best frequency bands identifying EEG signals depend on individuals and measurement environment, Ref. 3. Bayesian Spatio-Spectral Filter Optimization (BSSFO) is known as a powerful method to solve these problems, Ref. 4. However, BSSFO has also drawbacks that the obtained solution by it falls into sub-optimum. To overcome this drawback, we propose improvements of the preprocessing and update method of the filter bank in Ref. 4, to result in confirming effectiveness of our proposal.

2. Classification System

The classification system of EEG signals proposed by K. Suk, et al. Ref. 1 that is improved in this paper by us is shown in Fig.2. According to the flow of Fig.1, the contents are described in the following subsections.

2.1 Preprocessing

Laplacian smoothing is applied to all the EEG signals to reduce artifacts and noise as follows. The weight of the data in attention electrode surrounded by a green circle is 4 and that of each of four surrounding electrodes surrounded by red circles is -1. (See Fig.1) These weights are changed to optimal values in Section 3 by us (Improvement 1).

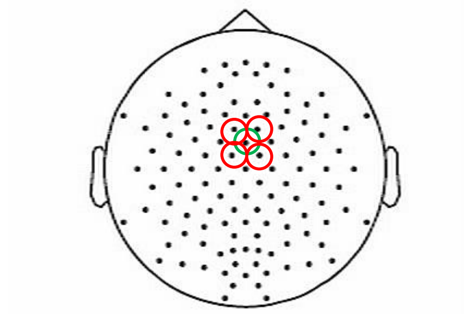

Fig. 1 Explanation of the weights of the EEG electrodes in smoothing of the EEG signal in the extended international 10/20 system used in this study

*For the title, try not to use more than 3 lines. Typeset the title in 12 pt Times Roman, uppercase and boldface.

2.2 Spectral Filtering

The filter used in this paper is a fifth-order Butterworth bandpass-filter. The EEG signals of all the data sets are bandpass-filtered between 4 Hz and 40 Hz covering both the $\mu-\mathrm{rhythm}(8-14\mathrm{Hz})$ and $\beta-\mathrm{rhythm}(14-30\mathrm{Hz})$.

2.3 Common Spatial Pattern

Common Spatial Pattern (CSP) proposed by H. Ramoser, et al. in Ref. 2 is used in this paper. CSP is applied to the signals after bandpass filtering. CSP algorithm is for searching for the spatial weight to multiply to the EEG signals. The spatial weight W is gotten by solving the optimization problem of following a generalized eigenvalue problem:

$$\arg\max_{W} \frac{\mathrm{W}^{\mathrm{T}}\Sigma_1 W}{\mathrm{W}^{\mathrm{T}}\Sigma_2 W}, \tag{1}$$

where Σ_1 and Σ_2 are covariance matrices of each class. Σ_1 and Σ_2 are calculated as follows:

$$\Sigma_{1,2} = \mathbf{X}\mathbf{X}^T / \mathrm{trace}(\mathbf{X}\mathbf{X}^T), \tag{2}$$

where $\mathbf{X}$ means EEG signal matrix, and following:

$$\mathbf{X} = \begin{pmatrix} x_{1,1} & \cdots & x_{1N} \\ \vdots & \ddots & \vdots \\ x_{C,1} & \cdots & x_{C,N} \end{pmatrix}, \tag{3}$$

where N is the number of signal samples, C means the number of electrodes. The first and last row vectors are taken as spatial patterns for class 1 and class 2, respectively.

2.4 Bayesian Spatio Spectral Filter Optimization (BSSFO)

<u>**Algorithm** : BSSFO filter optimization algorithm</u>

Input Data: $\{\mathbf{X}, \mathbf{\Omega}\}$, K, m,

$\mathbf{X} = \{\mathrm{x}_i\}_{i=1}^{D}$: Set of EEG signal, D: amount of trial,

$\mathbf{\Omega} = \{\omega_i\}_{i=1}^{D}$: Set of class labels, where $\omega_i \in \{+1, -1\}$,

K : The number of particles,

m: The half number of spatial patterns to be determined in a spatial pattern learning algorithm

Output Data :

$\hat{B} = \{\hat{\mathbf{b}}_j, \hat{\pi}_j\}_{j=1}^{\eta}$: Set of optimal particles,

$\hat{W} = \{\hat{W}_j\}_{j=1}$: Set of optimal spatio filters,

η : the number of particles

Optimization:

Initialization :

- $\hat{\mathbf{B}}^{old} = \{\mathbf{b}_k^{old}, \pi_k^{old}\}_{k=1}^{K}$

  - $\mathbf{b}_k^{o\;l} \stackrel{d}{=} \{b_k^s, b_k^e\}$

  - $\pi_k^{o\;l} \stackrel{d}{=} \frac{1}{K}$: Weight of kth particle

while stopping criterion not satisfied **do**

if the first iteration **then**

$B^{new} = B^{old}$

else

$B^{old} = B^{new}$

$\psi(k) = 0, \forall k \in \{1, \cdots, K\}$

for k = 1 to K **do**

Generate a random number $r \in \{0, 1\}$, uniformly distributed.

Find the largest j for which $r \geq \sum_{n=1}^{j} \pi_n$,

$\psi(j) = \psi(j) + 1$,

Improve ment 2 ⇨ $\mathbf{b}_k^{new} = \begin{cases} \mathbf{b}_j^{old} + N(\mathbf{0}, \mathbf{R}) & if, \psi(j) > 1 \\ \mathbf{b}_j^{old} & \text{otherwise} \end{cases}$

end

end

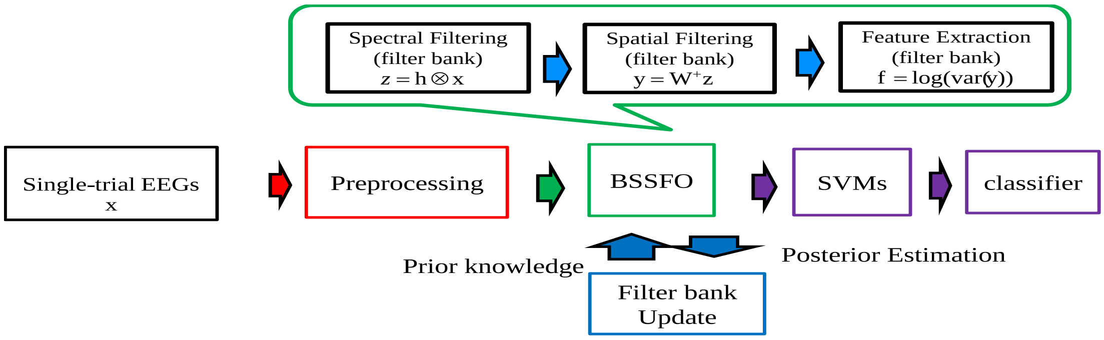

Fig. 2 Whole structure of the EEG classification system proposed by K. Suk, et al. Ref. 1

for k=1 to K **do**

$\mathbf{Z}_k = \mathbf{h}_k^{new} \otimes \mathbf{X}$ /* Perform a bandpass filter */

Solve $\mathbf{W}_k(\sum^{(+)} + \sum^{(+)})\mathbf{W}_k = \mathbf{I}$

/* Perform a CSP algorithm */

$\hat{\mathbf{W}}_k =$ the first m and the last m column vectors in $\mathbf{W}_k$

for i = 1 to D **do**

$\mathbf{f}_k^i = \log[\mathrm{var}(\hat{\mathbf{W}}_k^+ \mathbf{z}_k^i)]$ /* $\mathbf{F}_k = \{\mathbf{f}_k^i\}_{i=1}^D$ */

end

$\boldsymbol{I}(\mathbf{F}_k;\boldsymbol{\Omega}) = \boldsymbol{H}(\mathbf{F}_k) - \{\boldsymbol{H}(\mathbf{F}_k \mid \omega=+1)\} + \boldsymbol{H}(\mathbf{F}_k \mid \omega=-1)\}$

$p(\mathbf{F}_k, \boldsymbol{\Omega} \mid \mathbf{B}_k) \equiv \exp\{I(\mathbf{F}_k;\boldsymbol{\Omega})\}$

/* Update the weight of particles */

end

$\mathbf{B}^{new} = \left\{\mathbf{b}_k^{new}, \hat{\pi}_k^{new}\right\}_{k=1}^K$

end

$S = \bigcup_k (\pi_k > \tau), k \in \{1,2,\cdots,K\}$

$\hat{\mathbf{B}} = \left\{\mathbf{b}_j^{new}, \hat{\pi}_j^{new}\right\}_{j\in S}, \hat{W} = \{\hat{\mathbf{W}}_j\}_{j\in S}$

2.5 Feature Extraction

After spectral and spatial filtering, we get the feature vector $\mathbf{F}_k$ of *k*th particle as follows:

$$\mathbf{F}_k = \{\mathbf{f}_k^i\}_{i=1}^D, \quad \mathbf{f}_k^i = \log[\mathrm{var}(\hat{\mathbf{W}}_k^+ \mathbf{z}_k^i)]. \tag{4}$$

Using $\mathbf{F}_k$, mutual information $I(\mathbf{F}_k;\boldsymbol{\Omega})$ between $\mathbf{F}_k$ and class label $\boldsymbol{\Omega}$ is calculated as follows:

$$I(\mathbf{F}_k;\boldsymbol{\Omega}) = \boldsymbol{H}(\mathbf{F}_k) - \{\boldsymbol{H}(\mathbf{F}_k \mid \omega=+1)\} + \boldsymbol{H}(\mathbf{F}_k \mid \omega=-1)\}, \tag{5}$$

where $\boldsymbol{H}(F_k)$ and $\boldsymbol{H}(F_k \mid \omega=c)$ are defined as following equations.

$$\boldsymbol{H}(F_k) \cong -\frac{1}{D}\sum_{i=1}^{D} log\left[\frac{1}{D}\sum_{j=1} \hat{p}(\mathbf{f}_k)\right], \tag{6}$$

$$\boldsymbol{H}(F_k \mid \omega=c) \cong -\frac{1}{D_c}\sum_{i\ s.s.t\omega_i=c} log\left[\frac{1}{D_c}\sum_{j\ s.t.\ \omega_j=c1} \hat{p}\,(\mathbf{f}_k^i - \mathbf{f}_k^j, \nu)\right], \tag{7}$$

$$\hat{p}(\mathbf{f}_k) = \frac{1}{D}\sum_{i=1}^{D}\varphi(\mathbf{f}_k - \mathbf{f}_k^i, \nu), \tag{8}$$

$$\varphi(\mathbf{a},\nu) = \frac{1}{(2\pi)^{d/2}\nu^d\,|\Sigma|^{1/2}} exp\left[-\frac{\mathbf{a}^+\Sigma^{-1}\mathbf{a}}{2\nu^2}\right], \tag{9}$$

where D and D_c are the total number and class c of trials and Σ is the covariance matrix. The weight π_k of the classification result weight of particle k is calculated as follows:

$$\pi_k = \frac{\exp[I(\mathbf{F}_k;\Omega)]}{\sum_j \exp[I(\mathbf{F}_j;\Omega)]}. \tag{10}$$

2.6 Classifier

A Gaussian kernel-based SVM is used in this paper. An optimal filter bank S with the set of class-discrimination frequency bands selected by the following rule:

$$S = \bigcup_k (\pi_k > \tau), \tag{11}$$

where $k \in \{1,2,\cdots,K\}$ and τ denotes a threshold parameter that is determined empirically. The class label is determined by the following rule:

$$\hat{c} = \underset{c\in\{+,-\}}{\arg\max}\left\{\sum_{k=1}^{|S|}\pi_k \cdot \Phi_k^c(\mathbf{f}_k^*)\right\}, \tag{12}$$

where |S| denotes the size of the optimal filter bank S, $\mathbf{f}_k^*$ denotes the feature vector from the input signal-trial EEG x_i^*, and $\Phi_k^c(\mathbf{f}_k^*)$ is the result of a SVM which classifies the EEG into the class c, in the *k*th frequency band.

3. Improvement of the method

We propose two improvements of the method mentioned above.

3.1 Improvement 1

The first is that in the Laplacian smoothing, the weight value 4 of the data in attention electrode surrounded by a green circle in Fig. 1 is changed to 5 to enhance the signal of the attention electrode. .

3.2 Improvement 2

The second is that the improvement as to the update method of the particles which are for band start and end positions of bandpass filter. Use of the update method of the conventional method results in that particles with semi higher amount of information are remained, as a result, particles would be biased. For improvement, remove the half of all particles from the lower amount of information. Then, two particles with higher amount of information are selected stochastically, crossing them, new particles are generated. After that, particles between 1/4 and half from highest amount of information are re-initialized shown in Fig. 3.

Re-initializing of particles is to generate particles by use of following probability density function:

$$p(\mathbf{B}) = \frac{1}{2}\mathrm{N}(\boldsymbol{\mu}, \Sigma_{\boldsymbol{\mu}}) + \frac{1}{2}\mathrm{N}(\boldsymbol{\beta}, \Sigma_{\boldsymbol{\mu}}) \tag{13}$$

The crossing method of two particles is as follows:

$$(b_k^s, b_k^e)^{new} = \alpha_1(b_k^s, b_k^e)^{old1} + \alpha_2(b_k^s, b_k^e)^{old2}, \tag{14}$$

where the weights α_1, α_2 are set to 0.9 and 0.1, respectively, in next simulation to avoid generation of particles that both particles become far away.

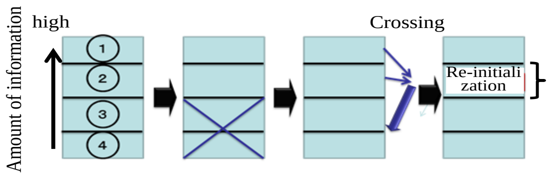

Fig. 3 Our proposed update method of particles

4. Computer simulation

4.1 Data set, simulation condition and results

EEG data set used in this simulation is Data set Iva in the BCI Competition III, in Ref. 5. The conditions to take these data is shown in Table 1. Table 2 , 3, and 4 show the simulation results that update method of particles of both conventional and improvement 2 are common, but without preprocessing, with conventional preprocessing, and with preprocessing improvement 1, respectively. In all cases, our proposed method is superior to conventional one.

Table 1 Simulation condition

EEG data	
the number of sampling	200
sampling rate	100 [Hz]
used time of EEG data	0.5[s]~2.5[s]
the number of electrodes	118
BSSFO	
the number of particles	15
width of frequency band	4 ~ 40 {Hz]
The number of loop algorithm	10

Table 2 Simulation results of both update methods of particles without preprocessing

update method		subjects					average
		aa	al	av	aw	ay	
conventional	average	70.54	99.82	60.56	76.25	56.67	72.77
Improvement 2	average	72.86	99.64	63.21	70.58	69.13	75.08

Table 3 Simulation results of both update methods of particles with conventional preprocessing

update method		subjects					average
		aa	al	av	aw	ay	
conventional	average	40.29	87.5	56.68	64.06	49.6	61.43
Improvement 2	average	53.57	91.61	53.67	65.54	50.36	63.95

Table 4 Simulation results of both update methods of particles with preprocessing with improvement 1

update method		subjects					average
		aa	al	av	aw	ay	
conventional	average	72.23	98.75	60.51	677.1	55.36	72.79
Improvement 2	average	73.3	99.64	61.99	74.91	69.88	75.9

5. Conclusion

In this paper, we intended to improve the update method of particles of the conventional method “BSFFO” that mean improvement of frequency bandpass filters and also to improve the preprocessing method. As a result, it is verified that our method is useful.

References

1. H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, “Optimal Spatial Filtering of Single Trial EEG during Imagined Hand Movement”, IEEE Trans. Rehabilitation Eng., vol. 8, no. 4, pp.441-446, Dec. 2000.
2. Fikri Goksu, Nuri F. Ince, Ahmed H. Tewfik, “Greedy solutions for the construction of sparse spatial and spatio-spectral filters in brain computer interface applications”, Neurocomputing, vol. 108, no. 2, pp.69-78, May. 2013.
3. H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, “Optimal Spatial Filtering of Single Trial EEG during Imagined Hand Movement”, IEEE Trans. Rehabilitation Eng., vol. 8, no. 4, pp.441-446, Dec. 2000.
4. Heung-Il Suk, Seong-Whan Lee, “A Novel Bayesian Framework for Discriminative Feature Extraction in Brain-Computer Interfaces”, IEEE Transactions. Pattern Analysis and Machine Intelligence, vol. 35, no. 2, pp.286-299, Feb. 2013.
5. http://bbci.de/competition/iii/desc_IVa.html