Classification by Weighting for Spatio-Frequency Components of EEG Signal During Motor Imagery

Hiroshi Higashi   Toshihisa Tanaka

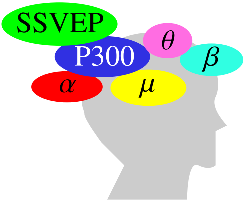
Tokyo University of Agriculture and Technology, Japan
higashi@sip.tuat.ac.jp, tanakat@cc.tuat.ac.jp
http://www.sip.tuat.ac.jp/

ICASSP 2011, May 27, 2011

General Goal

Extracting features associated brain condition from brain signals to exploit brain computer interfaces (BCI).

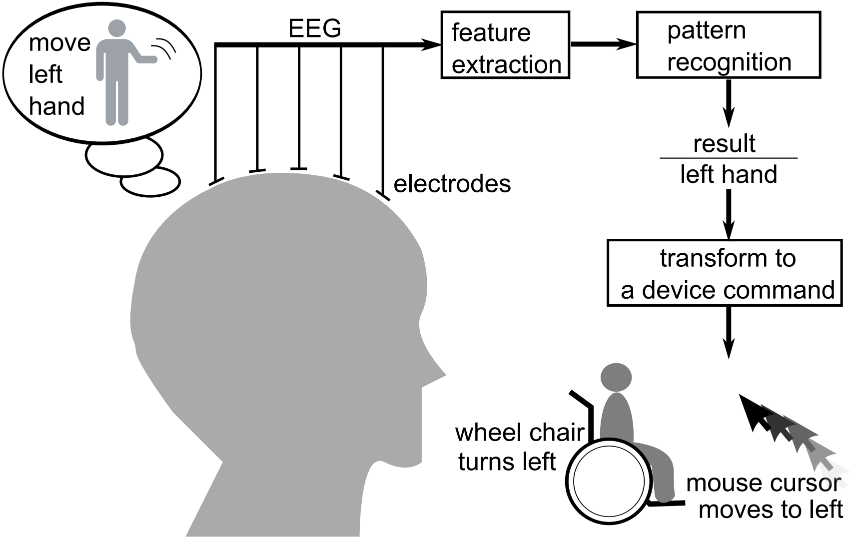
Several features corresponding to brain condition can be observed by the observation system such as EEG and fMRI.



BCIs capture such features and assign the device commands corresponding to each brain condition.

In this study, we focus on **EEG** as measurement method and the feature induced by **imaging movement of body**.

Motor Imagery Based BCI (MI-BCI)



EEG Features of MI-BCI

- An energy in frequency band called **mu rhythm** desynchronizes by imaging movement.
- The desynchronized location in the brain depends on imagined part.

If we know

- a change of the energy in a certain frequency band
- spatial location of its energy change

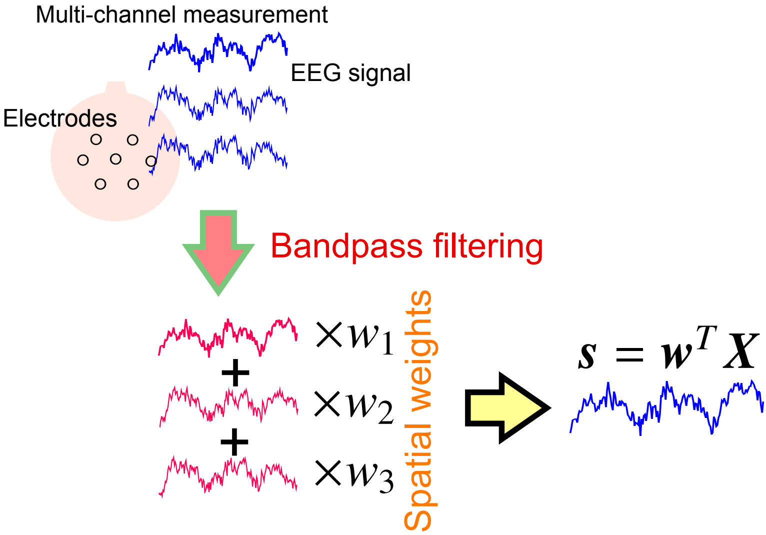
we can recognize imagined part of the body movement from EEG signals.

Effective method in 2-class MI-BCI

Common spatial pattern (CSP) [Ramoser, et al.]

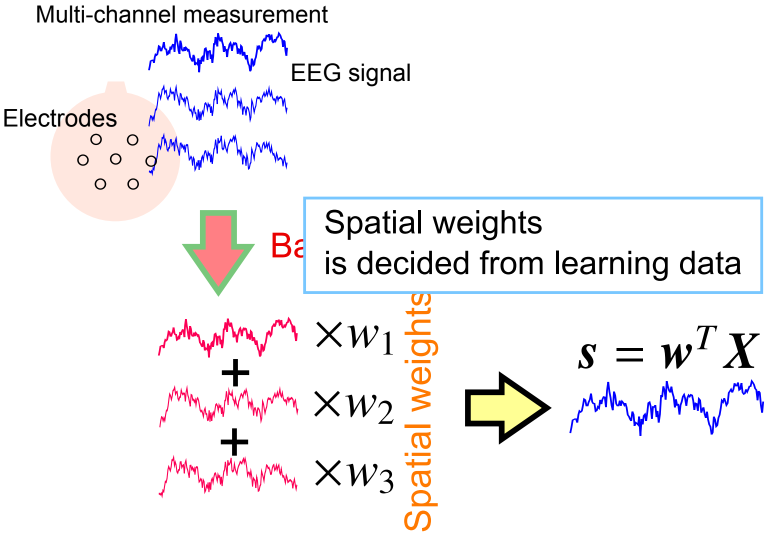
Common Spatial Pattern (CSP)

[H. Ramoser, et al., 2000.]



Common Spatial Pattern (CSP)

[H. Ramoser, et al., 2000.]



CSP: How to Find w (2-Class Problem)

The expectation of the variance of spatial-weighted signal belonging to class c is minimized.

- c: a class label, $c \in \{1, 2\}$

Optimization problem

$$\min_{\boldsymbol{w}} \quad E_{\boldsymbol{X} \in \mathcal{C}_c}[\mathrm{var}(\boldsymbol{w}^T \boldsymbol{X})]$$

$$\text{subject to} \quad \sum_{d=1,2} E_{\boldsymbol{X} \in \mathcal{C}_d}[\mathrm{var}(\boldsymbol{w}^T \boldsymbol{X})] = 1$$

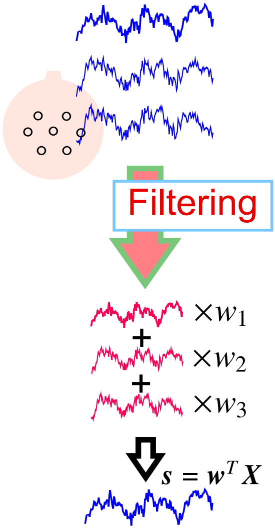
- C_c: a set of data belonging to class c
- $E_{\boldsymbol{X} \in C_d}$: expectation over C_d

For classification we can use a variance of the extracted signal $\boldsymbol{w}^T \boldsymbol{X}$ as feature value.

Problem: Filtering for CSP

CSP needs bandpass filtering as pre-processing.

The optimal frequency band depends on a measurement environment and/or a subject [J. Müller-Gerling, et al., 1999].



We want the best filter to extract components that mostly contain features related to the motor-imagery.

Design Methods for the Filter

Searching band with learning data by cross-validation

Search the most classifiable band out of candidates.

$\Rightarrow$ The number of the candidates is finite.

Huge computational cost is needed, when a lot of candidates.

Common Spatio-Spectral Pattern (CSSP) [S. Lemm, et al.]

Cannot provide complicated frequency selectivity.

Common Sparse Spectral Spatial Pattern (CSSSP)

[G. Dornhege, et al.]

Cost function includes the sparsity criteria.

Optimization is very complex and time-consuming.

Spectrally weighted CSP (SPEC-CSP) [R. Tomioka, et al.]

No guarantee that the optimization converges

Outline of the Proposed Method

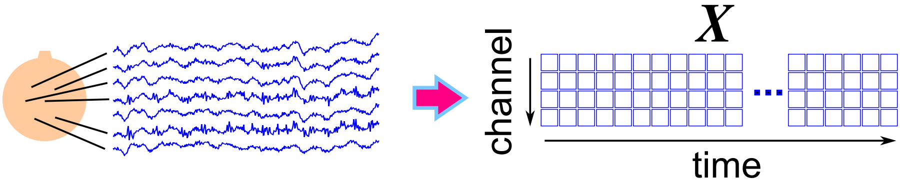
We propose a new simple method to design the spatio-frequency filters.

- **Add** into the cost function of CSP **parameters that are weights for frequency components**.
- **Alternating optimize** both a **spatial filter** (pattern) and a **frequency filter**.

  $\Rightarrow$ The cost function converges because the single cost.
- The optimization sub-problems are reduced to **generalized eigenvalue problems**.

Preliminary: Block Illustrates a Signal Element

For easily understanding, we illustrate an element of a signal matrix as a block.



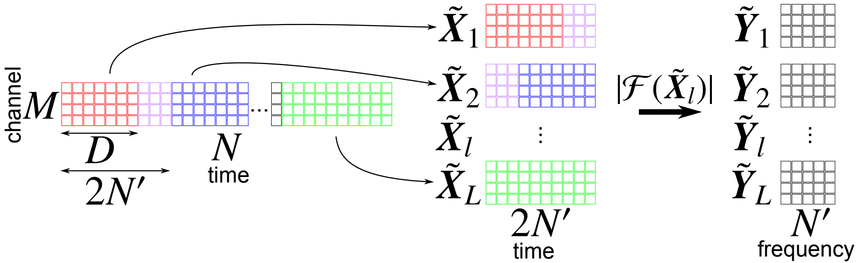
Preliminary: Time-Shifted DFT

Windowing a signal matrix $\boldsymbol{X}$ with length $2N'$ and shifting by D samples.

$$[\tilde{\boldsymbol{X}}_l]_{m,n} = [\boldsymbol{X}]_{m,(l-1)D+1+n}$$

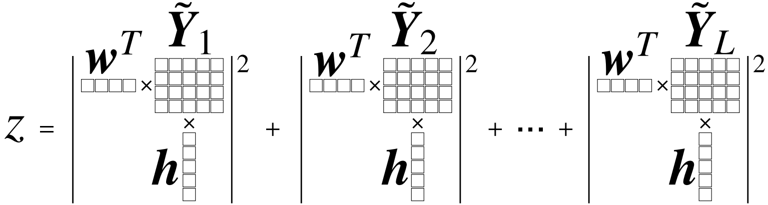
Transform $\tilde{\boldsymbol{X}}_l$ to spectrum $\tilde{\boldsymbol{Y}}_l \in \mathbb{R}^{M \times N'}$

$$[\tilde{\boldsymbol{Y}}_l]_{m,n} = |[\mathcal{F}(\tilde{\boldsymbol{X}}_l)]_{m,n}|$$



Definition the Feature Using Spectra

$$z(\boldsymbol{X}, \boldsymbol{w}, \boldsymbol{h}) = \sum_{l=1}^{L} |\boldsymbol{w}^T \tilde{\boldsymbol{Y}}_l \boldsymbol{h}|^2.$$



- $\boldsymbol{w}$: vector of spatial weights for channels.
- $\boldsymbol{h}$: vector of spectral weights for frequency components.

⇒ NEXT: How to find $\boldsymbol{w}$ and $\boldsymbol{h}$.

Optimization Criterion for Weights

Design the optimization problem for both weight vectors by extending CSP cost.

$\boldsymbol{w}_c$ and $\boldsymbol{h}_c$ are decided such that expectation of feature z belonging to class c is minimized in learning data.

Optimization problem

$$\min_{\boldsymbol{w},\boldsymbol{h}} \quad f(\boldsymbol{w}, \boldsymbol{h}) = E_{\boldsymbol{x} \in \mathcal{C}_c}[z]$$

$$\text{subject to} \quad \sum_{d=1,2} E_{\boldsymbol{x} \in \mathcal{C}_d}[z] = 1$$

- c: an optional class label, $c \in \{1, 2\}$

Alternating Optimization

Difficult to simultaneously optimize both $\boldsymbol{w}$ and $\boldsymbol{h}$.

$\Rightarrow$ Divide the problem into two sub-problems.

1. Optimization $\boldsymbol{w}$ while fixing $\boldsymbol{h}$.
2. Optimization $\boldsymbol{h}$ while fixing $\boldsymbol{w}$.

Note z can be transformed the following forms.

$$
\begin{aligned}
z =& \boldsymbol{w}^T \left(\sum_{l=1}^{L} \tilde{\boldsymbol{Y}}_l \boldsymbol{h} \boldsymbol{h}^T (\tilde{\boldsymbol{Y}}_l)^T \right) \boldsymbol{w} \\
=& \boldsymbol{h}^T \left(\sum_{l=1}^{L} \tilde{\boldsymbol{Y}}_l \boldsymbol{w} \boldsymbol{w}^T (\tilde{\boldsymbol{Y}}_l)^T \right) \boldsymbol{h}
\end{aligned}
$$

Sub-Problems

Sub-problem 1: Optimization *w* while fixing *h*

$$\min_{\boldsymbol{w}} \ \boldsymbol{w}^T \boldsymbol{R}_c \boldsymbol{w}, \quad \text{subject to } \boldsymbol{w}^T(\boldsymbol{R}_1 + \boldsymbol{R}_2)\boldsymbol{w} = 1$$

- $\boldsymbol{R}_d = E_{\tilde{\boldsymbol{Y}}_l \in \mathcal{C}_d}\left[\sum_{l=1}^{L} \tilde{\boldsymbol{Y}}_l \boldsymbol{h}\boldsymbol{h}^T(\tilde{\boldsymbol{Y}}_l)^T\right]$

Sub-problem 2: Optimization *h* while fixing *w*

$$\min_{\boldsymbol{h}} \ \boldsymbol{h}^T \boldsymbol{Q}_c \boldsymbol{h}, \quad \text{subject to } \boldsymbol{h}^T(\boldsymbol{Q}_1 + \boldsymbol{Q}_2)\boldsymbol{h} = 1$$

- $\boldsymbol{Q}_d = E_{\tilde{\boldsymbol{Y}}_l \in \mathcal{C}_d}\left[\sum_{l=1}^{L} \tilde{\boldsymbol{Y}}_l \boldsymbol{w}\boldsymbol{w}^T(\tilde{\boldsymbol{Y}}_l)^T\right]$

Sub-problems can be reduced to generalized eigenvalue problems.

Iteration Steps for Optimization

Step 1 Initialize $\boldsymbol{h}$.

Step 2 Optimize $\boldsymbol{w}$ by solving Sub-problem 1.

Step 3 Optimize $\boldsymbol{h}$ by solving Sub-problem 2.

Step 4 Repeat Step 2 and Step 3 until cost function $f(\boldsymbol{w}, \boldsymbol{h})$ is converged.

Classification Rule

The extracted feature value are classified by following rules.

Method 1: use different filter *h* between classes

$$u = \operatorname*{argmin}_{c} z(\boldsymbol{X}, \boldsymbol{w}_c, \boldsymbol{h}_c) \implies \boldsymbol{X} \in \mathcal{C}_u.$$

Method 2: use common filter *h* in all classes

- $\boldsymbol{w}_g$ and $\boldsymbol{h}_g$ are found in a class g where g is an optional class label.
- Solve Sub-problem 1 with $\boldsymbol{R}(\boldsymbol{h}_g)$ in other class $\bar{g}$, and get $\boldsymbol{w}_{\bar{g}}$.

$$u = \operatorname*{argmin}_{c} z(\boldsymbol{X}, \boldsymbol{w}_c, \boldsymbol{h}_g) \implies \boldsymbol{X} \in \mathcal{C}_u.$$

Experiment

Dataset: BCI Competition III dataset IVa

2 classes EEG (right hand and right foot motor imagery)
5 subjects (*aa, al, av, aw, ay*)
118 electrodes (extended 10-20 method)
Sampling frequency: 100 Hz.
The number of trials in each class: 140
Duration of each trial: 3.5 seconds

Compared Feature extraction methods

CSP1 (with the bandpass filter with the passband of 7–30 Hz)
CSP2 (with the bandpass filter manually-optimized by CV)
CSSP
SPEC-CSP

⇒ In the methods, we adopted simple classification rules which is to compare the feature values corresponding to each class.

$$u = \operatorname*{argmin}_{c} \operatorname{var}(\boldsymbol{w}_c^T \boldsymbol{X})$$

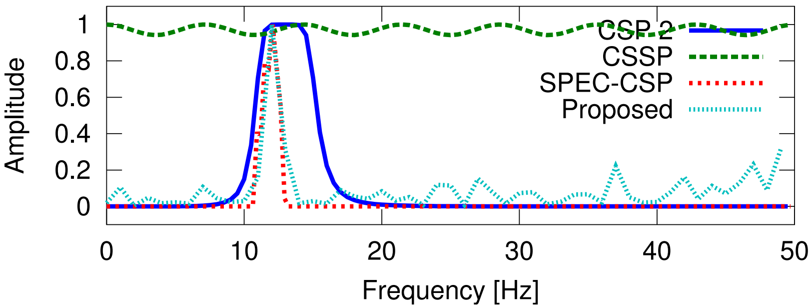
Classification Accuracy

Table: Classification accuracy [%] by 5×5 CV

Method	Subject *aa*	*al*	*av*	*aw*	*ay*	Ave.
CSP1	69.3	89.9	49.1	89.0	80.1	75.5
CSP2	82.6	96.9	**52.7**	**96.9**	81.5	82.1
CSSP	76.2	93.6	51.3	96.5	84.4	80.4
SPEC-CSP	79.6	94.8	49.4	96.4	84.3	80.9
Method1	82.4	95.3	50.8	91.6	90.1	82.0
Method2	**83.2**	**97.6**	49.1	94.5	**90.6**	**83.0**

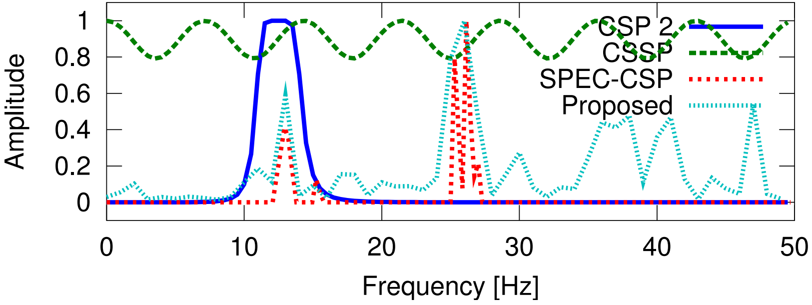
Parameters in the proposed method: Frame length 50, shift sample 5.

Amplitude Characteristic (subject *al*)



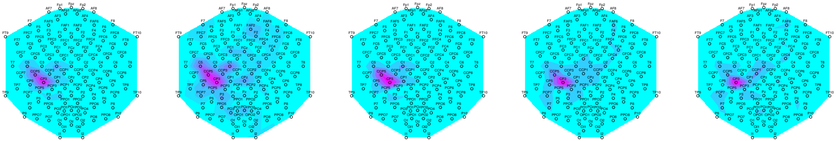
- The proposed method gave the same passband as that manually optimized.

Amplitude Characteristic (subject *aa*)



- The proposed method also have large weights in band higher than the manually-optimized passband.
- The higher bands look the harmonics of 10–15 Hz.

Spatial Weight Pattern (subject *aa*)

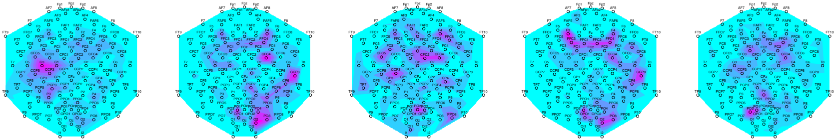


(a) CSP1 (b) CSP2 (c) CSSP (d) SPEC-CSP (e) Proposed

Figure: Subject *aa*

- All methods gave almost same spatial weights.

Spatial Weight Pattern (subject *ay*)



(a) CSP1 (b) CSP2 (c) CSSP (d) SPEC-CSP (e) Proposed

Figure: Subject *ay*

- Different frequency components gave different spatial patterns.

Conclusions

- Propose a method to **design the weights for channels and frequency components** using learning data.
- The solution of the weight optimization can be obtained by **alternating solving sub-problems**, **generalized eigenvalue problems**.

  $\Rightarrow$ The cost function is non-increasing in each iteration.
- The proposed method achieves **higher classification accuracy** in motor-imagery based BCI.

Present Works

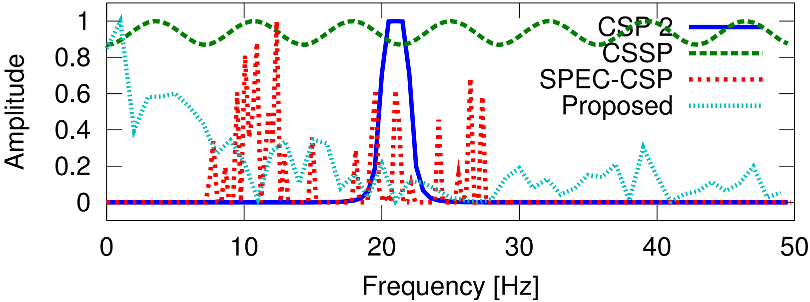
We already developed

- a temporal filter design method to easily apply realtime application,
- a method to design a bank of spatio-temporal filters to extract plural EEG frequency features.

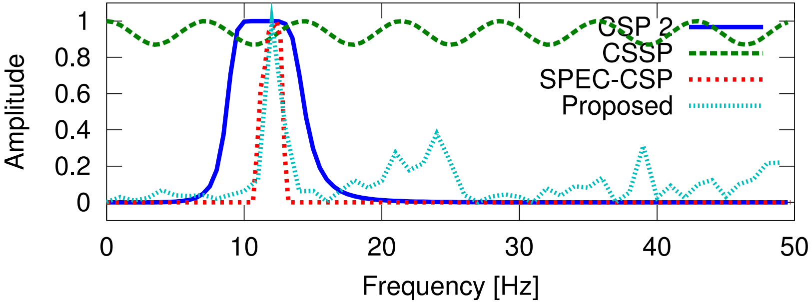
The works are to be presented.

Thank you for your attentions!
Any questions?

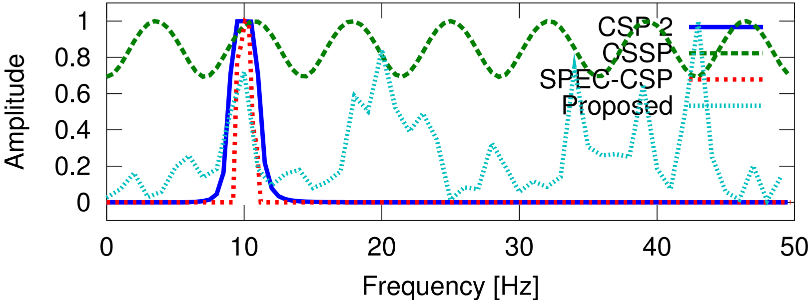
Amplitude Characteristic (subject *av*)



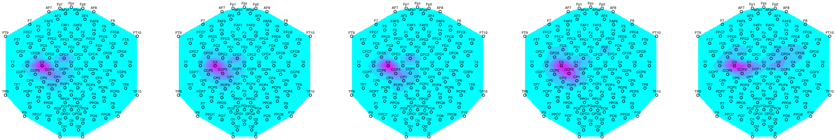
Amplitude Characteristic (subject *aw*)



Amplitude Characteristic (subject *ay*)



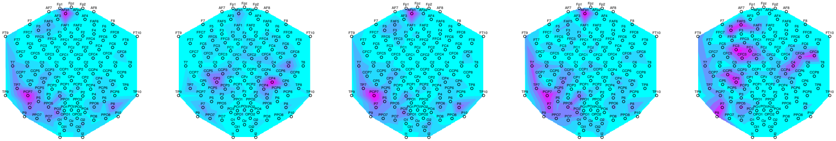
Spatial Weight Pattern (subject *al*)



(a) CSP1 (b) CSP2 (c) CSSP (d) SPEC-CSP (e) Proposed

Figure: Subject *al*

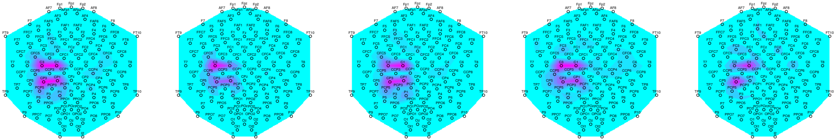
Spatial Weight Pattern (subject *av*)



(a) CSP1 (b) CSP2 (c) CSSP (d) SPEC-CSP (e) Proposed

Figure: Subject *av*

Spatial Weight Pattern (subject *aw*)



(a) CSP1 (b) CSP2 (c) CSSP (d) SPEC-CSP (e) Proposed

Figure: Subject *aw*

Relation of Parameters and Classification Accuracy (subject *aa*)

N': signal lengths, D: shift samples