Motor Imagery based BCI Classification via Sparse Representation of EEG Signals

Brain Computer Interface

Younghak Shin, S. Lee, M. Ahn, S. C. Jun and Heung-No Lee

Department of Information and Communications,
Gwangju Institute of Science and Technology (GIST), South Korea

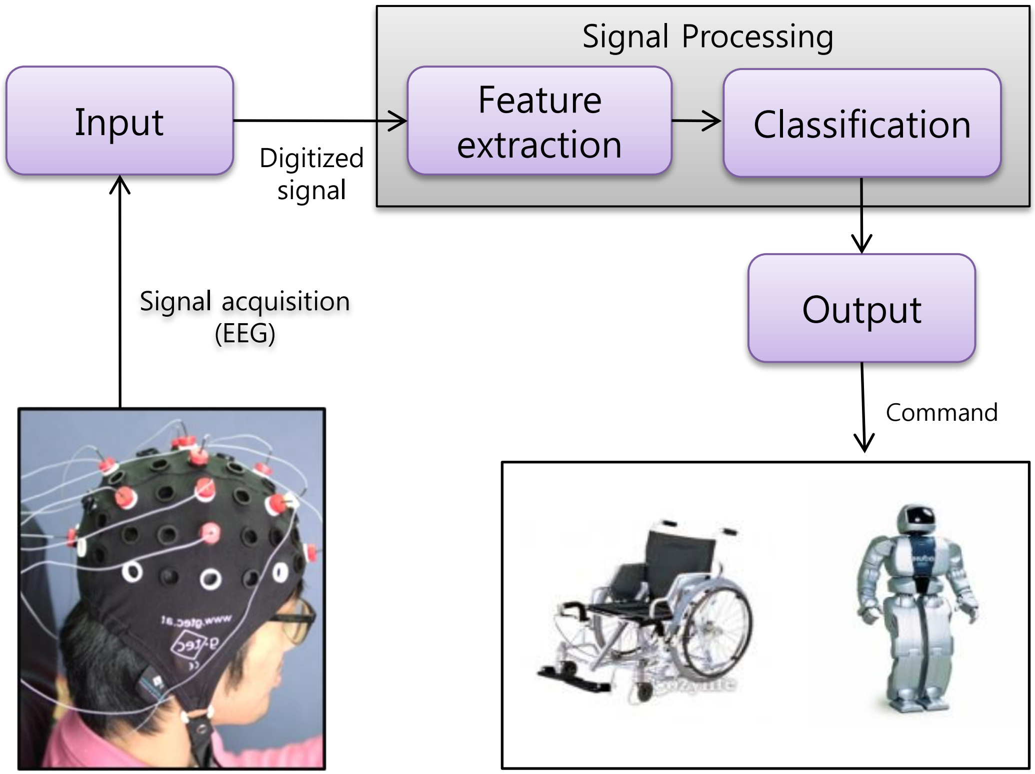
May 13, 2011

Outline

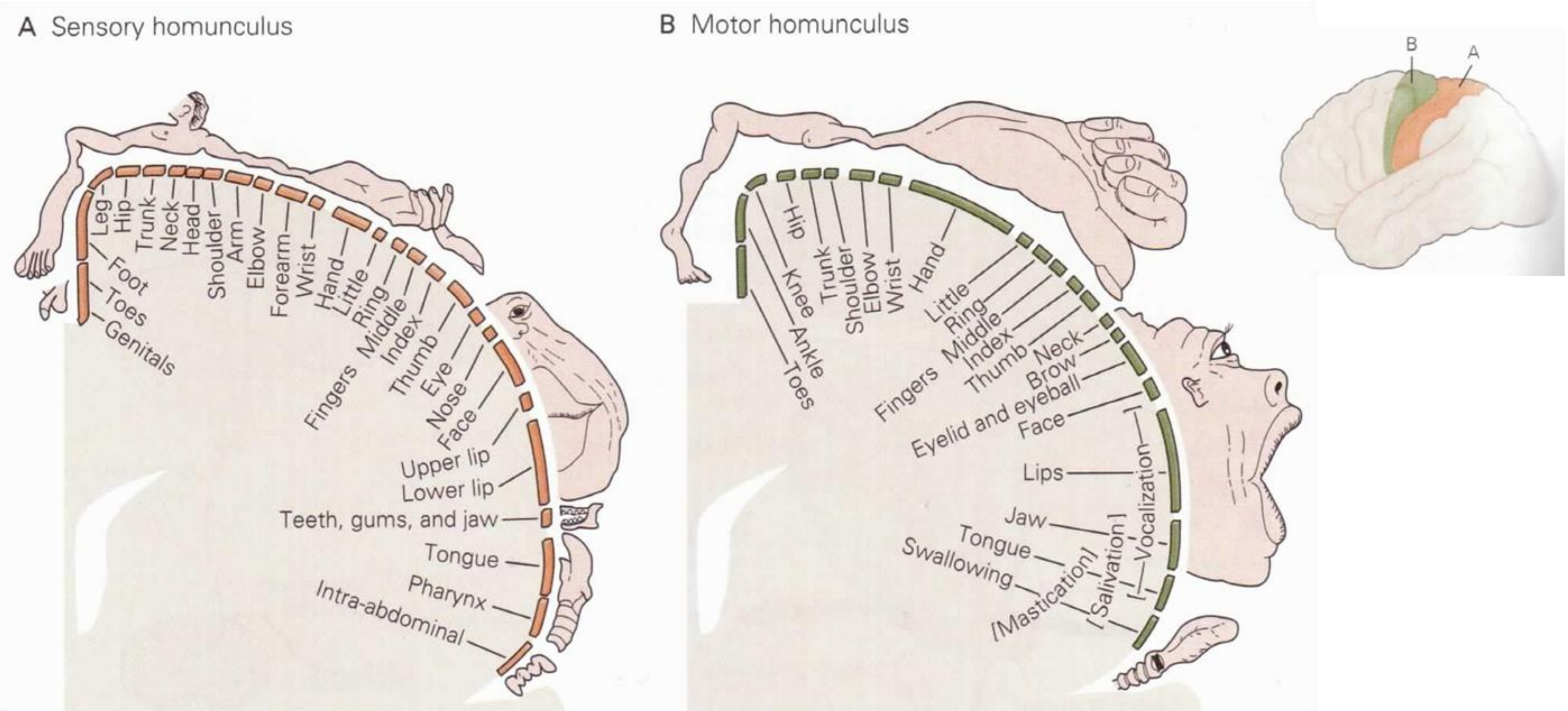
- Introduction
- Experiment
- Preprocessing
- Proposed Scheme
- Sparse Representation
- Results and Discussions
- Conclusions

EEG based BCI

- Brain Computer Interface system



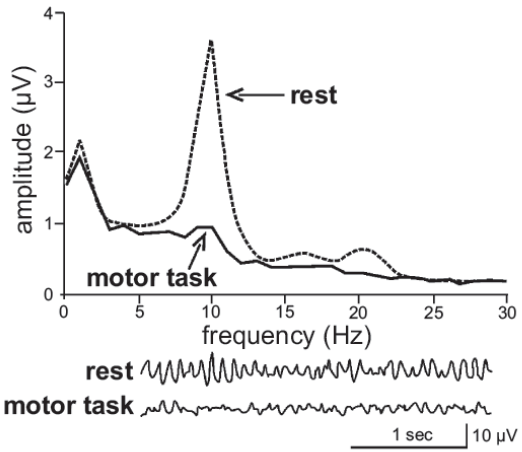
Sensory and Motor mapping



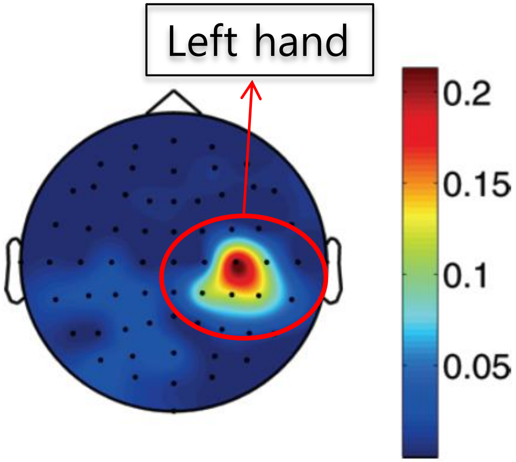
[Neuroprosthetics: Theory and Practice, K.W. Horch & G.S. Dhillon]

Motor imagery based BCI

- Mu and Beta rhythm(SMR) analysis [Wolpaw 1991]
  - The mu(8-12Hz) and/or Beta(15~30Hz) rhythms originate above the motor cortex.
  - When we imagine the left hand movement, amplitude attenuation of mu rhythm appears at the contralateral area of cortex.



< Spectral focus >



< Position focus >

Sparse Representation (SR)

- Sparse Representation has received a lot of attentions in recent years.
- The problem of the SR is to find for the most compact representation of a signal in terms of linear combination of atoms in an overcomplete dictionary. [Huang 2006]
- The problem of SR is to find the coefficient $\mathbf{x} \in \mathbb{R}^{N \times 1}$:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{x}$$

where, $\mathbf{A} \in \mathbb{R}^{M \times N}$ is given over-complete dictionary$(M \ll N)$

$\mathbf{y} \in \mathbb{R}^{M}$ is given signal

- Where $\|\mathbf{x}\|_0$ is the ell-0 norm. Solving this minimization is NP hard
- Recently developed Compressive Sensing theory [Donoho 2006] reveals that if solution is sparse enough, ell-1 norm solution is equivalent to the ell-0 norm solution.

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{x}$$

Motivation and Purpose

- The sparse representation can be used for a number of applications including noise reduction, compression, and pattern recognition.
- Classification based on Sparse Representation was studied in Face Recognition area [Wright 2009].
- Feature extraction and classification are also key parts of BCIs.
- In this study, we apply the sparse representation to the motor imagery based BCI classification.
- Using Mu and Beta rhythms as a BCI feature, we aim to develop a new Sparse Representation based Classification (SRC) method.

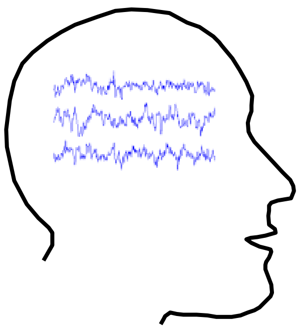
Data acquisition

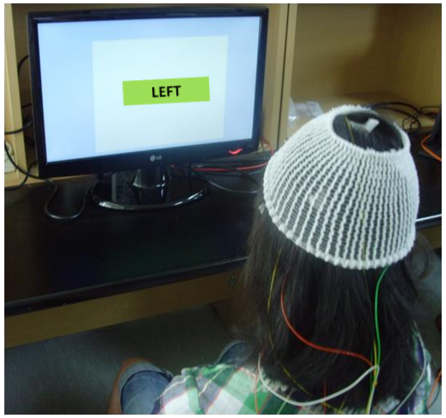
PZ3 Amplifier

RZ5 Workstation

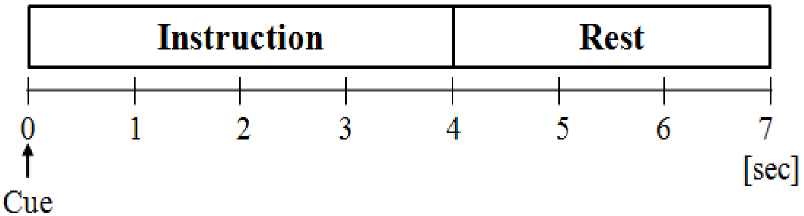
Desktop

EEG signal





Training



- We have two classes

  : Left and Right hand imaginary movement

CSP(Common Spatial Pattern)

- To reduce the dimension of feature vector and make distinguishable features, we use the CSP method.
- CSP is a powerful signal processing technique suitable for EEG-based BCIs [Blankertz 2008].
- CSP filters maximize the variance of the spatially filtered signal under one condition while minimizing it for the other condition.
  - Compute covariance matrix

$$C_R = X_R X_R^T$$

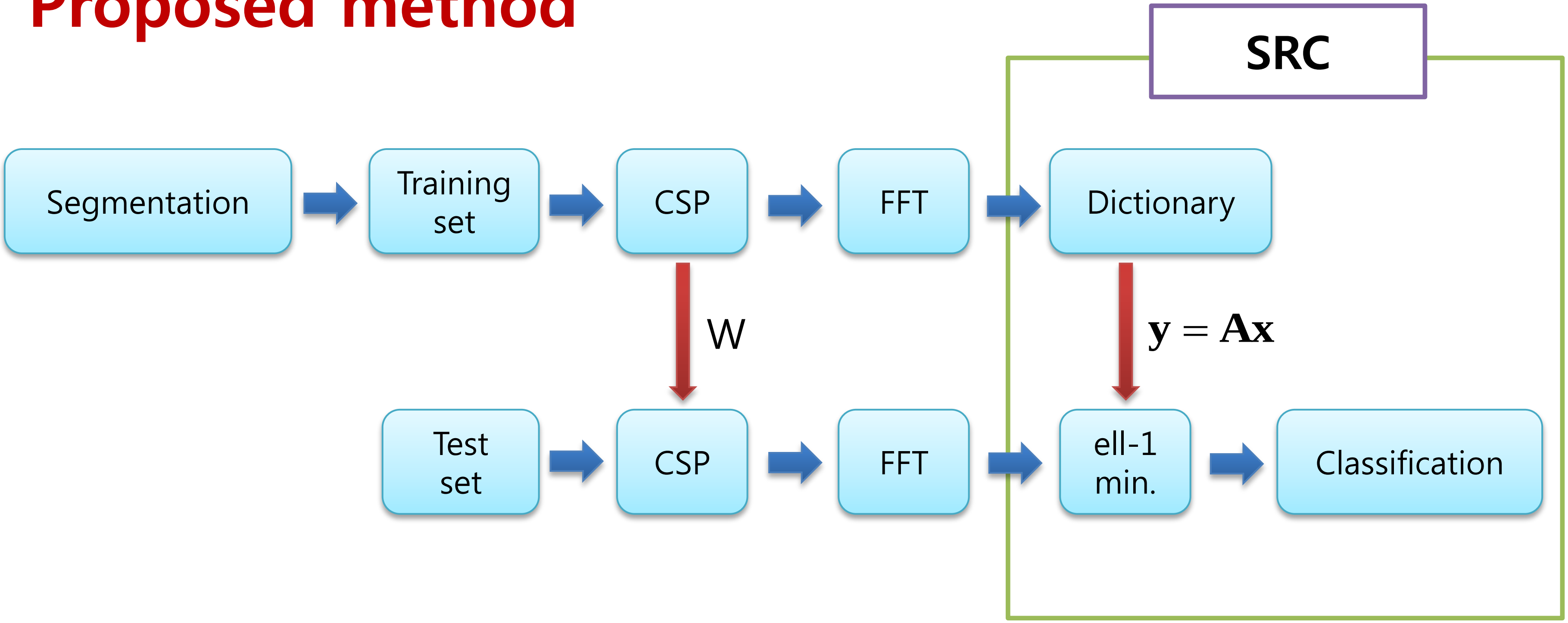
$$C_L = X_L X_L^T$$

  - Find vectors w satisfying the following optimization problems

$$\max_w \left(\frac{w_i^T C_R w_i}{w_i^T C_L w_i} \right)$$

$$\Rightarrow \min_w \left(-w_i^T C_R w_i \right) \quad \text{subject to } w_i^T C_L w_i = 1$$

Proposed method



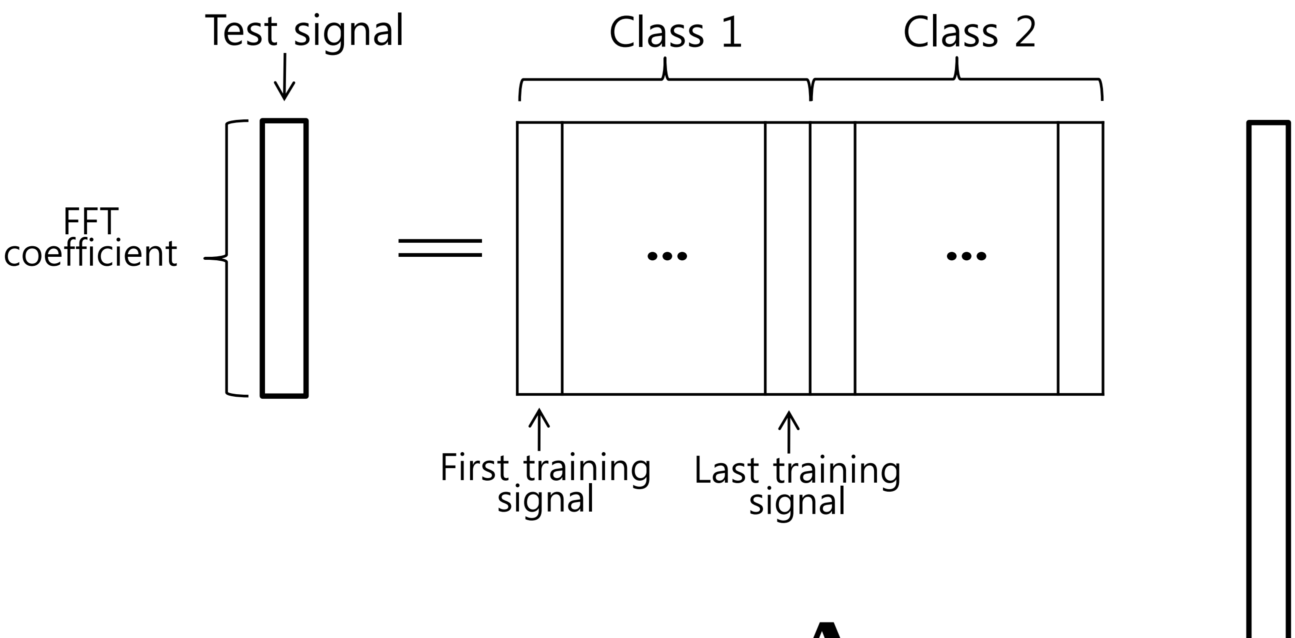
- Using training data set we can find CSP filter W .
- To use mu and beta rhythms as a BCI feature, we have taken the FFT
- We propose a new sparse representation based classification (SRC) method using dictionary and ell-1 minimization

Design a dictionary

- CSP filtering
- FFT (8~15Hz)
- vectorization

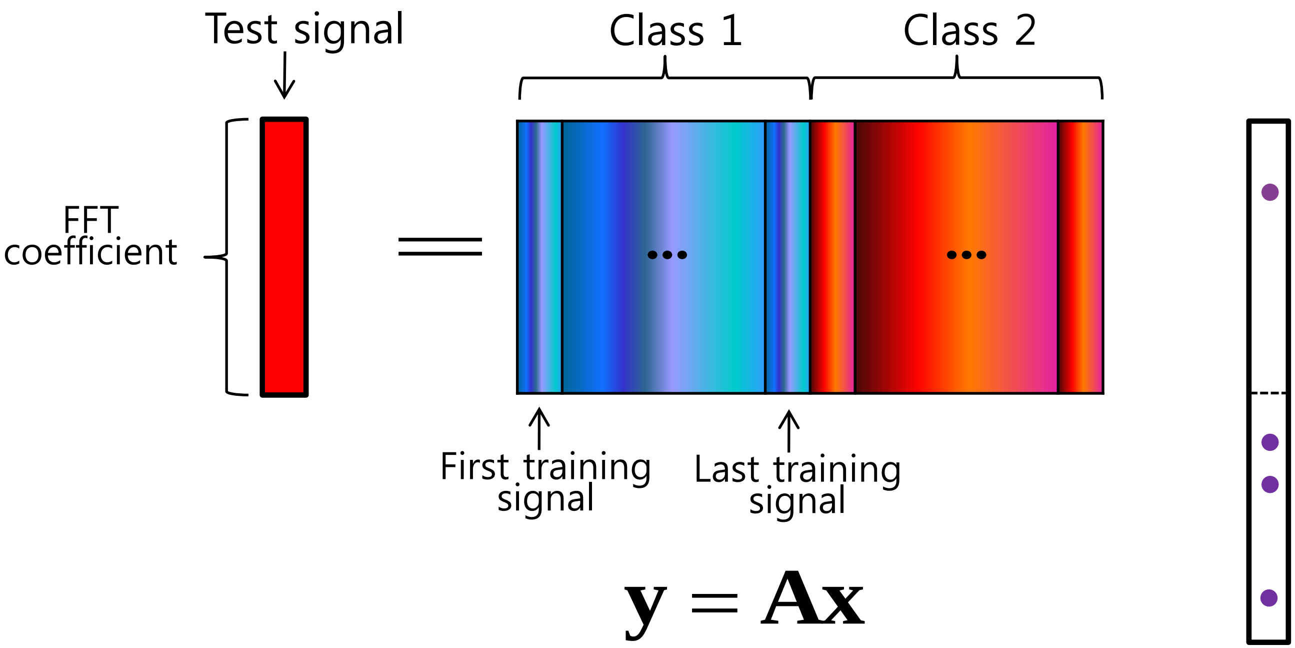
$$\mathbf{A}_i = [\mathbf{a}_{i,1}, \mathbf{a}_{i,2}, \ldots, \mathbf{a}_{i,N_t}]$$

$$\mathbf{A} := [\mathbf{A}_L ; \mathbf{A}_R]$$



$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

Sparse representation



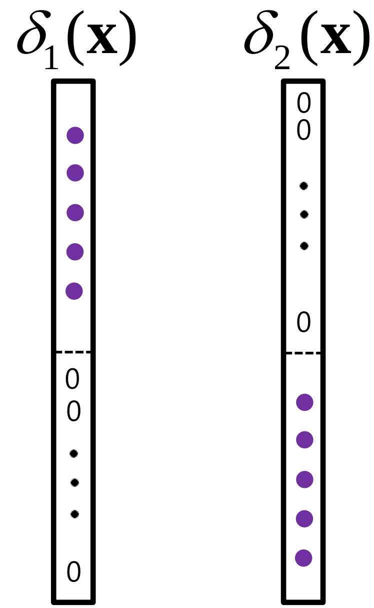
$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

- This sparse representation can be solved by ell-1 minimization [Candès 2006].
- For example, a test signal **y** of class 2 can be sparsely represented as the training signals of class 2.
- However, EEG signals are very noisy, nonzero coefficients may appear in the indices corresponding to the class 1.

Classification based on Sparse Representation

- To make use of the sparse representation result in a classification problem we use characteristic function δ.

[Wright2009]



$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{A}\mathbf{x}$$

$$r_i(\mathbf{y}) := \|\mathbf{y} - \mathbf{A}\delta_i(\mathbf{x})\|_2$$

- For each class *i*, $\delta_i(\mathbf{x})$ is obtained by nulling all the elements with corresponding to the other class.
- We can obtain residuals for each class.
- We determine the class *i* that has minimum residuals.

$$\text{class}(\mathbf{y}) = \arg \min_i r_i(\mathbf{y})$$

Accuracy evaluation

- We have analyzed four data sets.

Subjects	# of EEG channels	# of total trials
A	32	200
B	32	100
C	12	60
D	12	60

- All subjects are novice in the BCI experiment.
- We compare performance of the proposed SRC with the LDA(linear discriminant analysis) method.
- To make fair comparison, we also use the CSP filtering, the FFT, and the Mu(8~14Hz) and/or the Beta(15~30Hz) rhythms in the LDA classification method

Classification accuracy of subject A and B

Subject	# of trainin g signals	# of test sig nals	Accuracy (%)	
			LDA	*SRC*
A (total 200 signals)	150	50	63.50	71.25
	160	40	67.75	75.50
	180	20	68.50	77.75
	190	10	69.25	79.50
	199	1	68.75	79.00
B (total 100 signals)	50	50	67.50	71.50
	80	20	68.00	80.00
	90	10	69.50	82.50
	95	5	72.00	82.00
	99	1	72.50	82.00

- For both subject A and B, the accuracy of the SRC is better than that of the LDA method

Classification accuracy of subject C and D

Subject	# of trainin g signals	# of test sig nals	Accuracy (%)	
			LDA	*SRC*
C (total 60 signals)	40	20	89.17	91.67
	50	10	89.17	91.67
	55	5	89.17	90.84
	58	2	88.33	91.67
	59	1	88.33	91.67
D (total 60 signals)	40	20	80.83	77.50
	50	10	85.83	81.67
	55	5	86.67	78.33
	58	2	84.17	85.00
	59	1	85.00	86.67

- For subject C, SRC is better than LDA.
- For subject D, when the number of training signals increases, SRC is better than LDA.
- In this case, the number of training trials is not large enough.

Conclusions

- In this study, we propose a new sparse representation based classification (SRC) method for the motor imagery based BCI system.
- This method needs a well constructed dictionary matrix consisted of training signals.
- We use the CSP filtering and the FFT to produce the columns of the dictionary matrix.
- Our proposed SRC method shows better classification accuracy than LDA method which is well known for BCI classification.

Thank you
Any questions?

Reference

[Wolpaw1991] J. R. Wolpaw, D. J. McFarland, G. W. Neat, and C. A. Forneris, "An EEG-based brain-computer interface for cursor control," Electroencephalogr. Clin. Neurophysiol., vol. 78, pp. 252-259, 1991.

[Huang 2006] K. Huang and S. Aviyente, "Sparse Representation for Signal Classification," Neural Information Processing Systems, 2006.

[Donoho2006] D.Donoho, "Compressed sensing," IEEE Trans. Information Theory, vol. 52, pp. 1289-1306, 2006

[Wright2009] John Wright, Allen Y. Yang, Arvind Ganesh, S. Shankar Sastry, Yi Ma, "Robust Face Recognition via Sparse Representation" IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 31, no. 2, pp. 210~227, February 2009

[Blankertz2008] Benjamin Blankertz, Ryota Tomioka, Steven Lemm, Motoaki Kawanabe, Klaus-Robert Müller. "Optimizing Spatial Filters for Robust EEG Single-Trial Analysis," IEEE Signal Proc. Magazine, 25(1):41-56, January 2008.

[Candès 2006] E. Candès, J. Romberg, and T. Tao, "Stable Signal Recovery from Incomplete and Inaccurate Measurements," Comm. Pure and Applied Math., vol. 59, no. 8, pp. 1207-1223, 2006