Motor Imagery Classification Using Multiresolution Analysis and Sparse Representation of EEG Signals

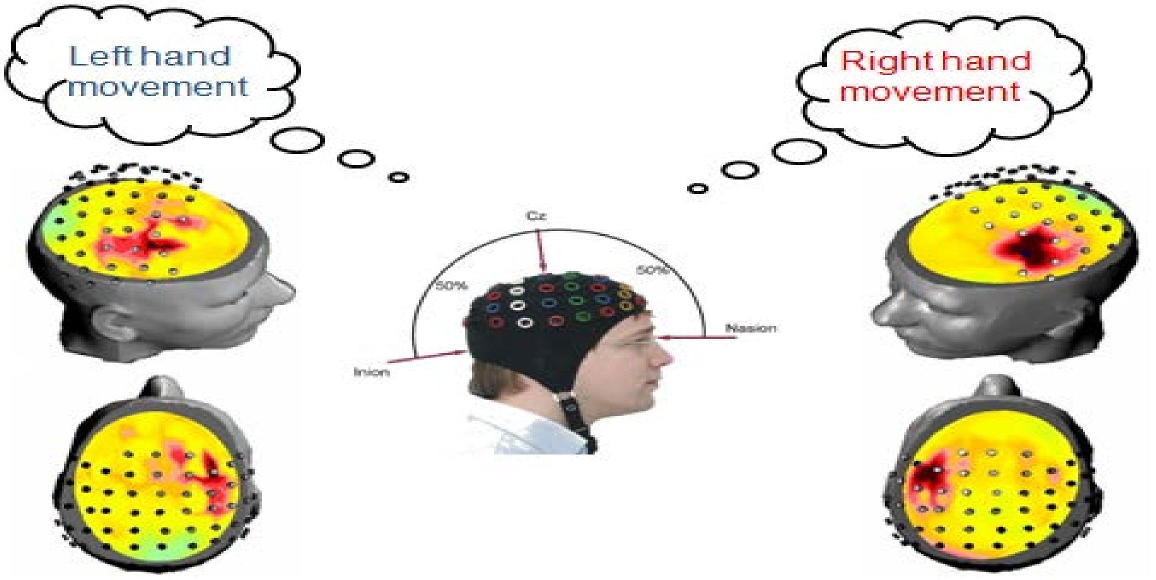
Pouria Saidi, George Atia, Alan Paris, Azadeh Vosoughi

Department of Electrical and Computer Engineering
University of Central Florida

Global SIP 2015

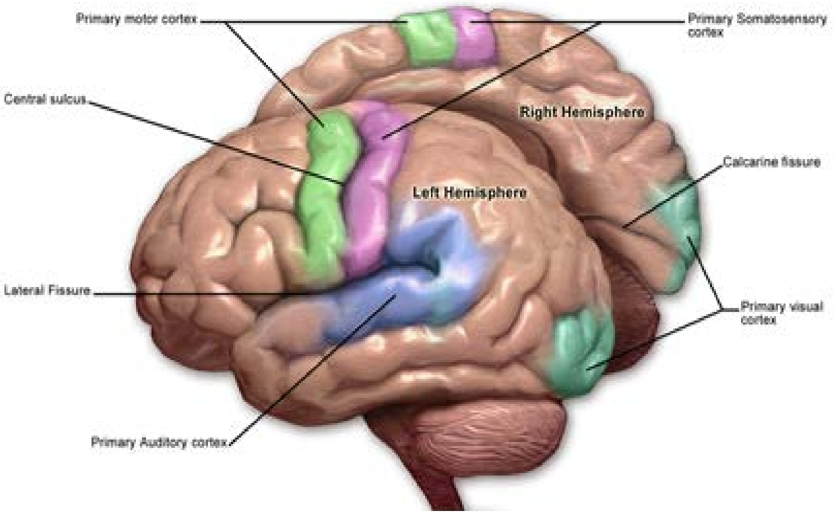
Motor Imagery Brain Signals

- Problem: Classifying motor imagery brain signals (imagined movement of limbs)



Motor Imagery Brain Signals

- Goal: Use less data and efficient algorithms to support real-time BCI.
- Approach:
  - Exploit sparse characteristics of EEGs.
  - Energies in different frequency sub-bands of the Wavelet Packet decomposition of EEG trials from few electrodes near the sensorimotor cortex.



Related works

- Using Wavelet transforms to extract features. (G. Garcia et al. 2003)
- Using Autoregressive coefficients (R. Boostani, et al. 2007)
- Most related work

Sparse representation-based classification scheme for motor imagery-based brain–computer interface systems(Y Shin, et al. 2012)

Outline

- EEG characteristics
- Feature extraction technique
- Proposed method based on sparse characteristics of EEG signals
- Results
- Conclusion

EEG Characteristics

- Two types of rolandic mu rhythm can be distinguished in the alpha band.

1. The lower-frequency mu rhythm between 8-10 Hz.
2. The higher frequency mu rhythm between 10-13 Hz.

EEG Characteristics

- Event-driven changes in the power of the EEG signals in particular frequency sub-bands are shown to improve the performance of BCI. (Pfurtschler 2003)

- In this paper we use energies, related to different frequency sub-bands motivated by the existence of different levels within the alpha band.

Pre-processing

- One of the most promising techniques in EEG signal processing is Common Spatial Patterns (CSP)[Ramoser-2000].

- CSP aims to project the data along a direction for which the trials from one class have maximum variance and the trials from the other class have minimum variance.

Wavelet Packet Decomposition

- Using time-frequency methods for non-stationary signals such as EEG can improve the performance of the classification techniques.
- Wavelet Packet Decomposition can be described using the filter-bank approach.

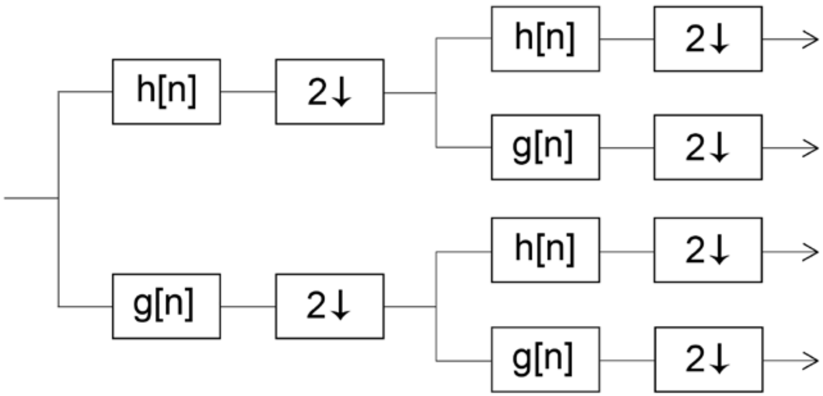


Fig-1 Wavelet Packet Decomposition

Feature Extraction

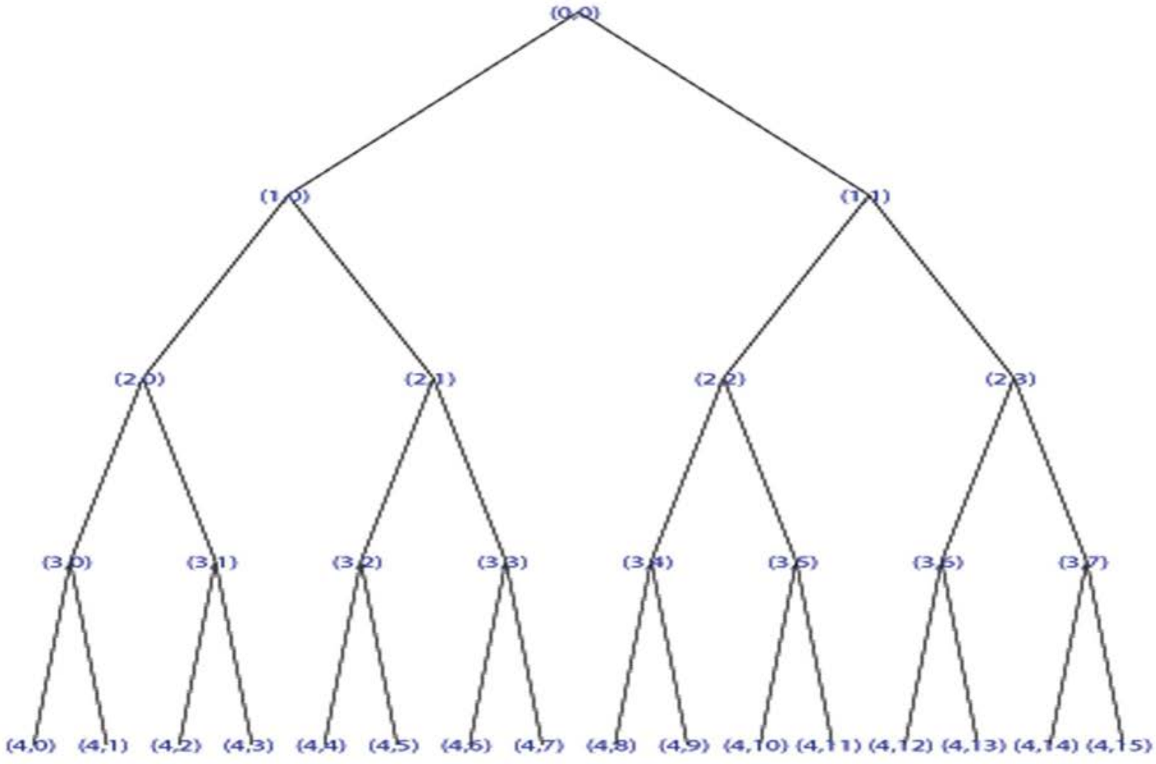


Fig 2 Energies are computed In 16 frequency sub-bands

University of Central Florida

Feature Extraction

- The entropy of a signal z is calculated from the wavelet coefficients, using

$$Entropy(z) = -\sum_i s_i^2 \log s_i^2$$

where s_i is the i-th wavelet coefficient of z obtained from WPT.

$$\begin{bmatrix} y_1^1 & y_2^1 & \dots & y_Q^1 \\ \vdots & & & \\ y_1^N & y_2^N & \dots & y_Q^N \end{bmatrix} \xrightarrow{CSP} [z_1 \quad \dots \quad z_Q] \xrightarrow{wavelet\ coef\ energy} [x_1^{apx} \quad x_1^d \quad \dots \quad x_L^d] \xrightarrow{Entropy\ concatenated} [x_1^{apx} \quad x_1^d \quad \dots \quad x_L^d \quad ent]$$

sparse representation of EEG signals

- In this work, we approximate the measurement vectors by linear combinations of a small number of atoms from a dictionary.

$$\begin{bmatrix} y_1^1 & y_2^1 & \dots & y_Q^1 \\ \vdots & & & \\ y_1^N & y_2^N & \dots & y_Q^N \end{bmatrix} \xrightarrow{CSP} [z_1 \quad \dots \quad z_Q] \xrightarrow{wavelet\ coef\ energy} [x_1^{apx} \quad x_1^d \quad \dots \quad x_L^d]$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_B \end{bmatrix} \cong \begin{bmatrix} a_{11}^m & a_{21}^m & \dots & a_{N_m 1}^m \\ a_{12}^m & a_{22}^m & \dots & a_{N_m 2}^m \\ \vdots & \vdots & \vdots & \vdots \\ a_{1B}^m & a_{2B}^m & \dots & a_{N_m B}^m \end{bmatrix} \begin{bmatrix} \alpha_1^m \\ \alpha_2^m \\ . \\ . \\ . \\ \alpha_{N_m}^m \end{bmatrix}$$

sparse representation of EEG signals

Therefore, the test signal is approximated using K atoms from the dictionary as

$$x = \alpha_{\lambda_1} a_{\lambda_1} + \alpha_{\lambda_2} a_{\lambda_2} + \cdots + \alpha_{\lambda_k} a_{\lambda_k}$$

where $\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_k\}, k = 1, \dots, K$ is the support of the sparse vector.

sparse representation of EEG signals

- Training samples from M classes generate M sub-dictionaries of a $B \times N$ dictionary A,

where $N = \sum_{m=1}^{M} N_m$.

$$
\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_B \end{bmatrix} \cong
\begin{bmatrix}
\overbrace{\begin{matrix} a_{11}^{1} & \cdots & a_{N_m 1}^{1} \\ a_{12}^{1} & \cdots & a_{N_m 2}^{1} \\ \vdots & \vdots & \vdots \\ a_{1B}^{1} & \cdots & a_{N_m B}^{1} \end{matrix}}^{\text{Right Hand}} &
\overbrace{\begin{matrix} a_{11}^{2} & \cdots & a_{N_m 1}^{2} \\ a_{12}^{2} & \cdots & a_{N_m 2}^{2} \\ \vdots & \vdots & \vdots \\ a_{1B}^{2} & \cdots & a_{N_m B}^{2} \end{matrix}}^{\text{Foot}}
\end{bmatrix}
\begin{bmatrix} \alpha_1^2 \\ . \\ \alpha_{N_m}^2 \\ \alpha_1^2 \\ . \\ \alpha_{N_m}^2 \end{bmatrix}
$$

sparse representation of EEG signals

- After obtaining the sparse representation of a test signal, it can be classified by computing residuals as

$$r^m(x) = \left\|x - A^i \hat{a}^m\right\|_2, m = 1,2,\dots,M$$

where $\hat{a}^m$ denotes the entries of the sparse vector associated with the m-th-class sub-dictionary.

$$class(x) = arg_{m=1,2,\dots,M} min\, r^m(x)$$

Test Data

Train Data

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_B \end{bmatrix} \cong \begin{bmatrix} \overbrace{a^1_{11} & \dots & a^1_{N_m 1}}^{\text{Right Hand}} & \overbrace{a^2_{11} & \dots & a^2_{N_m 1}}^{\text{Foot}} \\ a^1_{12} & \dots & a^1_{N_m 2} & a^2_{12} & \dots & a^2_{N_m 2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a^1_{1B} & \dots & a^1_{N_m B} & a^2_{1B} & \dots & a^2_{N_m B} \end{bmatrix} \begin{bmatrix} \alpha^2_1 \\ . \\ \alpha^2_{N_m} \\ \alpha^2_1 \\ . \\ \alpha^2_{N_m} \end{bmatrix}$$

sparse representation of EEG signals

- To recover the sparse vector α, we need to solve the following optimization problem:

$$min\|\alpha\|_0$$

$$subject\ to\ A\alpha = x$$

- This problem is generally NP-hard. It can be written as

$$min\|A\alpha - x\|_2$$

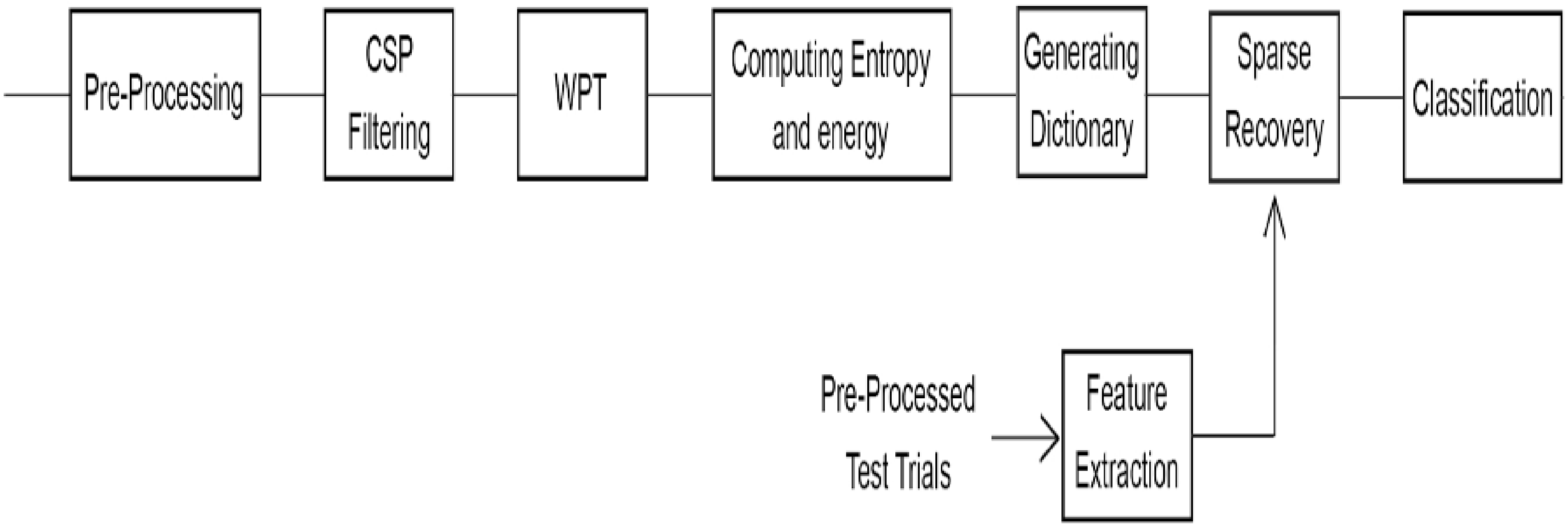
$$subject\ to\ \|\alpha\|_0 \leq K_0$$

where K_0 is an upper bound on the sparsity level.

- To solve the optimization problem, Orthogonal Matching Pursuit (OMP) greedy algorithm is used.

University of Central Florida

Methodology



Dataset

dataset 4a*:* provided by Fraunhofer FIRST, Intelligent Data Analysis Group and the Charite-University Medicine Berlin, Department of Neurology, Neurophysics Group.

- This data set consists of signals of five healthy subjects.

The visual indicator lasts for 3.5 seconds

a rest period begins with a random length of 1.75 to 2.25 seconds.

Dataset

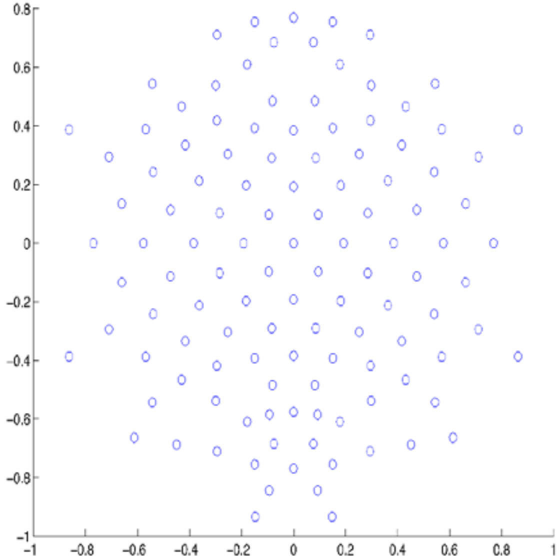


Fig 3-a Position of all the 118 electrode

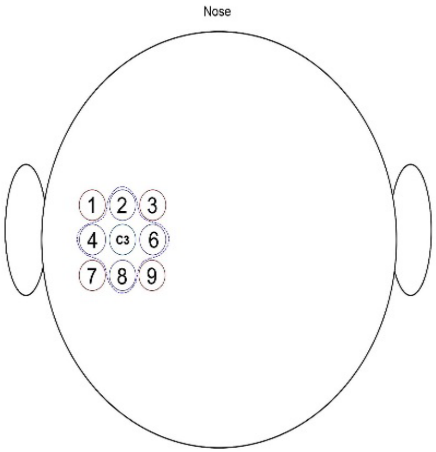


Fig 3-b Position of the five electrodes that are used.

Results

Table 1 Classification Accuracy rate (%)

Features	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
Wavelet Coefficients	64.46	73.89	54.11	75.71	64.96
Energy	64.79	85.50	61.51	73.11	59.36
Energy & Entropy	64.71	89.71	64.25	93.07	83.71
Method proposed by Y. Shin (2012)	57.29	87.25	60.14	75.07	83.43

Conclusion

- In this work, we proposed an algorithm to classify motor imagery EEG signals to support real time BCI.
- Dimensionality is reduced by selecting only five electrodes.
- We leverage the Sparse representation of the EEG trials in a multiclass dictionary learned from wavelet characteristics of the signals.
- Energy and Entropy related features enables efficient classification.

Conclusion

- This underscores the relevance of the energies and their distribution in different frequency sub-bands.

CSP

- Covariance matrices are transformed Using a whitening transformation derived from the eigenvector and eigenvalue factorization of the composite spatial covariance to S_1 and S_2.
- $S_1 = V\Sigma_1 V^T$ and $S_2 = V\Sigma_2 V^T$, then $\Sigma_1 + \Sigma_2 = 1$,
- where V Is the eigenvector matrix and Σ_1 and Σ_2 are the digonalized eigenvalue matrix.
- Hence:

The eigenvectors corresponds to the largest eigenvalue of one class, also corresponds to the smallest eigenvalue of the second group.