A Comparative Study of Features and Classifiers in Single-channel EEG-based Motor Imagery BCI

Suguru Kanoga, Atsunori Kanemura, and Hideki Asoh
AIST Japan
Motor Imagery Brain—Computer Interfaces

◆ Use rhythmic EEG features by motor imagery

\( \mu \) and \( \beta \) rhythms in motor areas are discriminative features

◆ Integrate spatial filter to emphasize the features

But, … the filter needs multichannel information

Single(/Few)-channel measurement devices

- Recently developed
  - Portable
  - Cheap
  - Limited channel (Ultimately single)

→ We can not use spatial filter for feature extraction

- Only allow specialized feature extraction technique to emphasize rhythmic features

→ Comparative study by using the same dataset is required

Motivation

Investigating how to build the best single-channel motor imagery BCI

◆ Research objective
  To find the best combination of channel, feature, and classifier

◆ Materials
  EEG dataset: BCI competition IV dataset 2a [3]

◆ Analysis
  - Epoch segmentation
  - Artifact rejection
  - Feature extraction: 3 types (PS, SCCSP, GLCM)
  - Classifiers: 6 types (LDA, k-NN, GMM, RF, MLP, SVM)
  - Evaluation: 2-class classification with 10-fold CV

Materials (1/2)

◆ Open-access EEG dataset

BCI competition IV dataset 2a

- # of channels: 22
- # of subjects: 9
- Task: Image a movement
- # of classes: 2 (left or right hand)
- Duration: 4 s
- # of trials: 288
- Sampling rate: 250 Hz
- Bandpass filter: 4 and 40 Hz
  (Fourth-order Butterworth filter)

Materials (2/2)

◆ Epoch segmentation

- 0.5-2.5 s after the cue\cite{Ang2012}

→ 288 2-s epochs

◆ Artifact rejection

- Some epochs were contaminated artifacts (e.g. muscular and ocular)
- The labels are given by the BCI competition (maybe visual inspection)

→ 221 to 279 epochs were used for classification

Feature Extraction (1/2)

◆ Power spectra (PS)
  Fourier transform
  - 100 sample points (and 28 zeros)
  - 50% overlap Hamming window
  - Log-scaled variance ranged 4 - 40 Hz

◆ Single-channel common spatial pattern (SCCSP)\(^5\)
  Calculates filter \( W \) which maximize the variance of
  two-class frequency bins
  - Similar to classical CSP
    (frequency bins as channels)
    → Common ‘spectral’ pattern…?
  - Log-scaled variance

\[^5\] S. Ge et al., “Classification of four-class motor imagery employing single-channel electroencephalography,” *PloS ONE*, vol. 9, no. 6, e98019, 2014.
Feature Extraction (2/2)

◆ Gray-level co-occurrence matrix (GLCM) \(^6\)
  - Concatenates PS to be a spectrogram
  - Translates spectrogram into 8-level values
  - Sums co-occurrence counts from the 8-level values
    4 directions \((0, 45, 90, \text{and } 135^\circ)\)
  - Used for 4 text descriptors
    1. Contrast
    2. Correlation
    3. Energy
    4. Homogeneity

→ To capture

co-occurrence power fluctuation (by event-related (de)synchronization)

Classification

◆ Six classifiers:
  1. Linear discriminant analysis (LDA)
  2. \(k\)-nearest neighbor (\(k\)-NN)
  3. Gaussian mixture model (GMM)
  4. Random forest (RF)
  5. Multi-layer perceptron (MLP)
  6. Support vector machine (SVM)

◆ Assessments
  - Independent validation for each subject and channel
  - 10-fold cross validation (10 times to avoid selection bias)

The best combination of channel, feature, and classifier
Results and Discussion (1/2)

- **Classification accuracy**

<table>
<thead>
<tr>
<th>Subject</th>
<th>Feature</th>
<th>LDA</th>
<th>k-NN</th>
<th>GMM</th>
<th>RF</th>
<th>MLP</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3</td>
<td>PS</td>
<td>70.6 ± 0.8 (C3)</td>
<td>67.9 ± 0.8 (C3)</td>
<td>60.7 ± 0.7 (C3)</td>
<td>69.2 ± 0.8 (C3)</td>
<td><strong>71.4 ± 0.7 (C3)</strong></td>
<td>70.8 ± 0.7 (C3)</td>
</tr>
<tr>
<td></td>
<td>SCCSP</td>
<td>67.0 ± 0.7 (C3)</td>
<td>64.4 ± 0.8 (C3)</td>
<td>62.1 ± 0.9 (C6)</td>
<td>65.6 ± 0.9 (C3)</td>
<td>71.1 ± 0.8 (C3)</td>
<td>66.6 ± 0.6 (C3)</td>
</tr>
<tr>
<td></td>
<td>GLCM</td>
<td>64.4 ± 0.7 (C4)</td>
<td>64.3 ± 1.2 (C3)</td>
<td>60.5 ± 0.5 (C3)</td>
<td>63.5 ± 0.8 (C3)</td>
<td>64.0 ± 0.7 (C3)</td>
<td>65.9 ± 0.7 (CP4)</td>
</tr>
<tr>
<td>S9</td>
<td>PS</td>
<td>85.0 ± 0.5 (C3)</td>
<td>82.2 ± 1.4 (C4)</td>
<td>80.6 ± 0.8 (C4)</td>
<td>83.0 ± 0.7 (C4)</td>
<td>85.4 ± 0.7 (C4)</td>
<td>85.0 ± 0.7 (C4)</td>
</tr>
<tr>
<td></td>
<td>SCCSP</td>
<td>83.8 ± 0.5 (C3)</td>
<td>84.2 ± 0.5 (C4)</td>
<td><strong>86.6 ± 0.6 (C4)</strong></td>
<td>85.0 ± 0.6 (C4)</td>
<td>86.2 ± 0.5 (C4)</td>
<td><strong>86.6 ± 0.4 (C4)</strong></td>
</tr>
<tr>
<td></td>
<td>GLCM</td>
<td>68.0 ± 0.8 (CP3)</td>
<td>72.3 ± 1.0 (CP4)</td>
<td>69.1 ± 0.9 (CP4)</td>
<td>71.6 ± 0.9 (C4)</td>
<td>71.7 ± 0.9 (CP4)</td>
<td>73.7 ± 0.8 (C4)</td>
</tr>
<tr>
<td>Mean</td>
<td>PS</td>
<td>61.8 ± 0.5</td>
<td>62.3 ± 0.4</td>
<td>60.0 ± 0.4</td>
<td>61.8 ± 0.4</td>
<td>63.1 ± 0.4</td>
<td><strong>63.5 ± 0.4</strong></td>
</tr>
<tr>
<td></td>
<td>SCCSP</td>
<td>61.8 ± 0.4</td>
<td>62.2 ± 0.4</td>
<td>61.6 ± 0.4</td>
<td>61.2 ± 0.42</td>
<td><strong>63.5 ± 0.4</strong></td>
<td>63.3 ± 0.4</td>
</tr>
<tr>
<td></td>
<td>GLCM</td>
<td>61.4 ± 0.4</td>
<td>60.7 ± 0.3</td>
<td>58.5 ± 0.3</td>
<td>58.7 ± 0.3</td>
<td>60.3 ± 0.4</td>
<td>61.4 ± 0.3</td>
</tr>
</tbody>
</table>

- **PS** and **SCCSP** showed superior performance than GLCM \((p<0.001)\)
- **MLP** with SCCSP / **SVM** with PS effectively classified EEG data
  → Average: 63.5 ± 0.4%
  Best : **86.6 ± 0.4%** (C3 or C4 position)
Results and Discussion (2/2)

- Absolute values of the coefficients in LDA

μ and β rhythms would have large coefficients …?

→ The relationship between features and classifiers should be further investigated.
Conclusions

Investigating how to build the best single-channel motor imagery BCI

◆ Results
- PS and SCCSP showed superior performance than GLCM ($p<0.001$)
- MLP with SCCSP / SVM with PS effectively classified EEG data
  
  Average: $63.5 \pm 0.4\%$
  
  Best: $86.6 \pm 0.4\%$ (C3 or C4 position)

◆ Limitation
Original dataset used multichannel environment
(Not used single(/few)-channel measurement device)
Appendices