Optimizing a Weighted Moderate Deviation for Motor Imagery Brain Computer Interfaces

J. Fumanal Idocin, J. Fernández, M. Gómez, C. Vidaurre, M. Papco, H. Bustince

Public University of Navarre

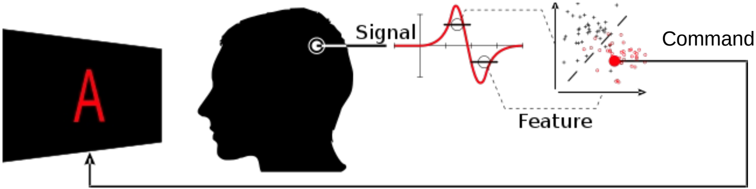
September 27, 2021

Introduction

1. Brain Computer Interfaces (BCIs)
2. Motor Imagery (MI)
3. Enhanced Multimodal Fusion Framework MI-BCI framework
4. Multichannel Weighted Moderate Deviations
5. Application and Results

Brain Computer Interfaces

- They are means to communicate **brain signals** and a **computer**.
- The target: to **understand** commands to execute in a external device.



Motor Imagery

- An individual rehearses or **simulates** a given **action**.
- We measure the **ElectroEncephalogram signal** while the simulation happens.
- We try to **identify** the imagined movement.

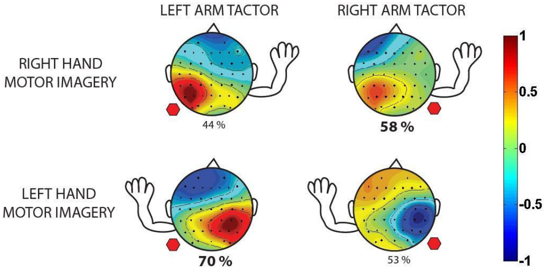


Image taken from: `http://www.traumatrucos.es/?p=395`

Enhanced Multimodal Fusion Framework

- A MI-BCI framework based on EEG signal[1].
- Tested on **left** hand, **right** hand, **tongue** and **foot** movement identification.
- Special focus on **signal derivation** and **decission making** based on the Multimodal Fuzzy Fusion framework[2].

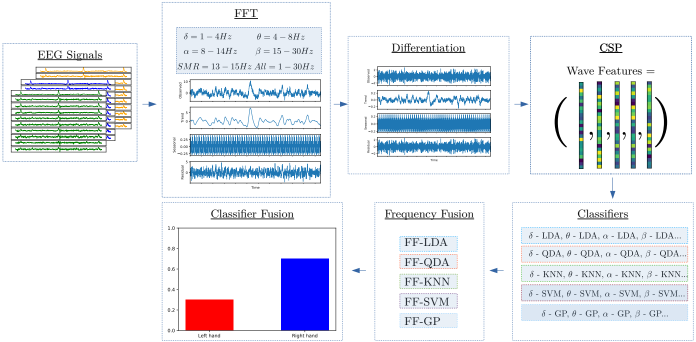
[1] Javier Fumanal-Idocin et al. "Motor-Imagery-Based Brain-Computer Interface Using Signal Derivation and Aggregation Functions". In: *IEEE Transactions on Cybernetics* (2021).

[2] Li-Wei Ko et al. "Multimodal Fuzzy Fusion for Enhancing the Motor-Imagery-Based Brain Computer Interface". In: *IEEE Computational Intelligence Magazine* 14.1 (2019), pp. 96–106.

Enhanced Multimodal Fusion Framework

1. EEG signal collection.
2. Fast Fourier Transform.
3. Differentiation.
4. Common Spatial Patterns.
5. Classifier training.
6. Extended Multimodal decision making phase.

EMF



Moderate Deviation-based aggregation

- Aggregation based on finding the point that minimizes sums of distances[3].
- Based on Darózcy means:
  A Darózcy mean is a mapping $M_D(\mathrm{x}) = y$, such that $\sum_{i=1}^{n} D(x_i, y) = 0, x_i, y \in R$.
- As D, we use a Moderate Deviation function. A mapping $I^2 \to R$ is said to be a Moderate Deviation function if and only if:
  (D1) for every $x \in I, D(x, \cdot) \to R$ is non-decreasing.
  (D2) for every $x \in I, D(\cdot, y) \to R$ is non-increasing.
  (D3) $D(x, y) = 0$ if and only if $x = y, x, y \in I$.
  Example: $D(x, y) = y - x$

[3] Abdulrahman H Altalhi et al. "Moderate deviation and restricted equivalence functions for measuring similarity between data". In: *Information Sciences* 501 (2019), pp. 19–29.

Moderate Deviation-based aggregation

Different **versions** of **Moderate Deviation**-based aggregation:

- For **numerical** data:
  Abdulrahman H Altalhi et al. "Moderate deviation and restricted equivalence functions for measuring similarity between data". In: *Information Sciences* 501 (2019), pp. 19–29.
- For **interval-valued** data:
  Javier Fumanal-Idocin et al. "Interval-valued aggregation functions based on moderate deviations applied to Motor-Imagery-Based Brain Computer Interface". In: *IEEE Transactions on Fuzzy Systems* (2021).
- For **multi-valued** data using **weights**:
  Martin Papčo et al. "A fusion method for multi-valued data". In: *Information Fusion* 71 (2021), pp. 1–10. ISSN: 1566-2535.

(We opted for the multi-valued data fusion)

Weighted Moderate Deviation-based aggregation

Defined as:

$$MD_{D,W} = \frac{1}{2}\left(\sup y \in I | \sum_{i=1}^{s}\sum_{j=1}^{t} w_{ij} D(x_{ij}, y) < 0 + \inf y \in I | \sum_{i=1}^{j}\sum_{j=1}^{t} w_{ij} D(x_{ij}, y) > 0\right) \tag{1}$$

Using as deviation function D:

$$D_\epsilon = (x+\epsilon)(y-x) \tag{2}$$

Then, we know the resulting MD:

$$MD_\epsilon(u,v) = \frac{u(u+\epsilon)+v(v+\epsilon)}{u+v+2\epsilon} \tag{3}$$

For example, for $I^4 \rightarrow R$ (2 x 2 patch):

$$MD_{\epsilon,w}(X) = MD_\epsilon(wX) = \frac{\sum_{i=1}^{2}\sum_{j=1}^{2} wx_{ij}(wxij+\epsilon)}{4w\epsilon + \sum_{i=1}^{2}\sum_{j=1}^{2} wx_{ij}} \tag{4}$$

Weighted Moderate Deviation-based aggregation

- Multimodal fusion scheme:

α Band - LDA

β Band - LDA

$\vdots$

All Band - LDA

Sample

(for each class c)

$$X_{LDA} = (x_{alpha}, x_{beta}, \ldots, x_{all})$$
$$X_{QDA} = (x_{alpha}, x_{beta}, \ldots, x_{all})$$
$$X_{KNN} = (x_{alpha}, x_{beta}, \ldots, x_{all})$$
$$X_{SVM} = (x_{alpha}, x_{beta}, \ldots, x_{all})$$
$$X_{GP} = (x_{alpha}, x_{beta}, \ldots, x_{all})$$

$$W_c = (w_1, w_2, w_3, w_4, w_5) \qquad w_{cf}$$

$$X = \begin{pmatrix} X_{LDA} \\ X_{QDA} \\ X_{KNN} \\ X_{SVM} \\ X_{GP} \end{pmatrix}$$

Weighted Moderate Deviation-based aggregation

- Multimodal fusion scheme:

α Band - LDA

β Band - LDA

⋮

All Band - LDA

Sample

(for each class c)

$X_{LDA} = (x_{alpha}, x_{beta}, \ldots, x_{all})$

$X_{QDA} = (x_{alpha}, x_{beta}, \ldots, x_{all})$

$X_{KNN} = (x_{alpha}, x_{beta}, \ldots, x_{all})$

$X_{SVM} = (x_{alpha}, x_{beta}, \ldots, x_{all})$

$X_{GP} = (x_{alpha}, x_{beta}, \ldots, x_{all})$

$W_c = (w_1, w_2, w_3, w_4, w_5)$ $\quad w_{cf}$

$$X = \begin{pmatrix} X_{LDA} \\ X_{QDA} \\ X_{KNN} \\ X_{SVM} \\ X_{GP} \end{pmatrix}$$

$$MD_{\epsilon}(X_{LDA}) = \frac{\sum_{i=1}^{N_{bands}} w_1 x_i (x_i \epsilon)}{N_{bands} w_1 \epsilon + \sum_{i=1}^{N_{bands}} w_1 x_i}$$

Weighted Moderate Deviation-based aggregation

- Multimodal fusion scheme:

(for each class c)

α Band - LDA

β Band - LDA

⋮

All Band - LDA

Sample

$X_{LDA} = (x_{alpha}, x_{beta}, \ldots, x_{all})$

$X_{QDA} = (x_{alpha}, x_{beta}, \ldots, x_{all})$

$X_{KNN} = (x_{alpha}, x_{beta}, \ldots, x_{all})$

$X_{SVM} = (x_{alpha}, x_{beta}, \ldots, x_{all})$

$X_{GP} = (x_{alpha}, x_{beta}, \ldots, x_{all})$

$W_c = (w_1, w_2, w_3, w_4, w_5)$ $\quad w_{cf}$

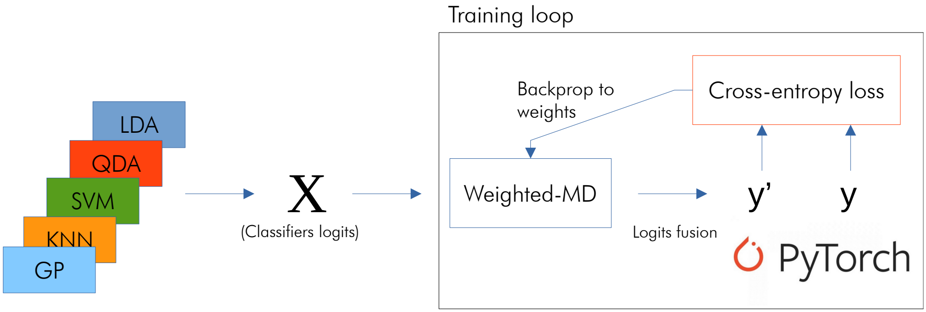
$$X = \begin{pmatrix} X_{LDA} \\ X_{QDA} \\ X_{KNN} \\ X_{SVM} \\ X_{GP} \end{pmatrix}$$

$$MD_\epsilon(X_{LDA}) = \frac{\sum_{i=1}^{N_{bands}} w_1 x_i (x_i \epsilon)}{N_{bands} w_1 \epsilon + \sum_{i=1}^{N_{bands}} w_1 x_i} \qquad MD_\epsilon(X_F) = \frac{\sum_{i=1}^{N_{classifiers}} w_{cf} x_i (x_i \epsilon)}{N_{classifiers} w_{cf} \epsilon + \sum_{i=1}^{N_{classifiers}} w_{cf} x_i}$$

Weighted Moderate Deviation-based aggregation

- Once we apply the weighted-MD:
  (1) **Fixing** the weights for all the channels to the same value.
  (2) **Learning** the weights using an optimization algorithm.

We tried both, using backpropagation with Pytorch in the second case:



Test performed

- We use the BCI IV 2a dataset to test the new aggregation proposal[4].
- We randomly generate 90 50%/50% train/test splits for evaluation.

Two classes: left/right hand.

Four classes: left hand, right hand, tongue, foot.

[4] Michael Tangermann et al. "Review of the BCI competition IV". In: *Front. in neuroscience* 6 (2012), p. 55.

Results for binary classification

Table 1: Results for the binary classification using the EMF with different configurations.

Aggregation	Accuracy	F1 Score
Weighted-MD learnt	$86.97\% \pm 3.98$	$86.95\% \pm 4.00$
Weighted-MD fixed	$85.80\% \pm 4.08$	$85.77\% \pm 4.09$
Arithmetic Mean	$85.80\% \pm 4.04$	$85.77\% \pm 4.04$
Choquet Integral	$86.39\% \pm 4.20$	$86.37\% \pm 4.21$
Sugeno Integral	$81.39\% \pm 4.39$	$81.37\% \pm 4.39$

Results for 4 classes classification

Table 2: Results for the four classification using the EMF with different configurations.

Aggregation	Accuracy	F1 Score
Weighted-MD learnt	$65.91\% \pm 13.15$	$65.91\% \pm 13.15$
Weighted-MD fixed	$72.93\% \pm 2.29$	$72.93\% \pm 2.29$
Arithmetic Mean	$72.22\% \pm 2.31$	$72.22\% \pm 2.31$
Choquet Integral	$72.93\% \pm 1.85$	$72.93\% \pm 1.85$
Sugeno Integral	$64.45\% \pm 2.66$	$64.45\% \pm 2.66$

Conclusions

- Multivalued MD-aggregations seem to be effective in this problem in any case.
- **Learning** the weights seem to be **effective** in the **binary** case, but not in the four classes. Probably **more data** is needed.
- Further learning algorithms for the weights should be studied.

References & Questions

Aggregations implementation library:
`https://github.com/Fuminides/Fancy_aggregations`

BCI frameworks implementation (WIP):
`https://github.com/Fuminides/athena`

Main references:

1. Javier Fumanal-Idocin et al. "Motor-Imagery-Based Brain-Computer Interface Using Signal Derivation and Aggregation Functions". In: *IEEE Transactions on Cybernetics* (2021)
2. Martin Papčo et al. "A fusion method for multi-valued data". In: *Information Fusion* 71 (2021), pp. 1–10. ISSN: 1566-2535