Histogram of Gradient Orientations of EEG Signal Plots for Brain Computer Interfaces

Dissertation Defense

Supervisor
Dr. Juan Miguel Santos

Candidate
Rodrigo Ramele

Cosupervisor
Dra. Ana Julia Villar

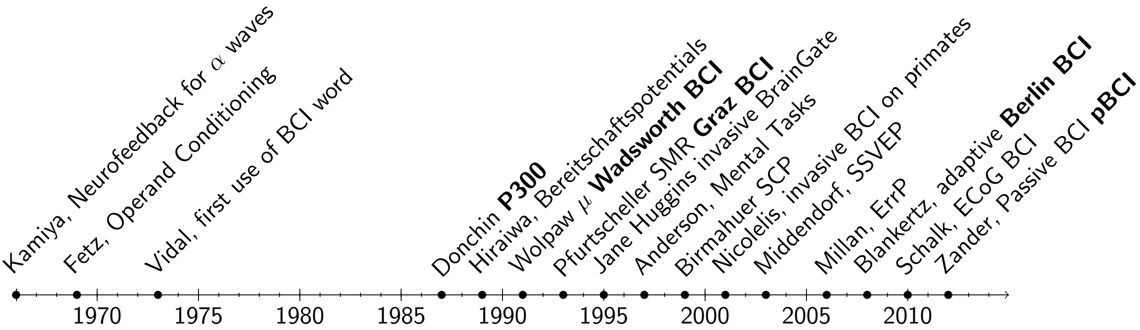
Doctorado en Ingeniería en Informática
Instituto Tecnológico de Buenos Aires

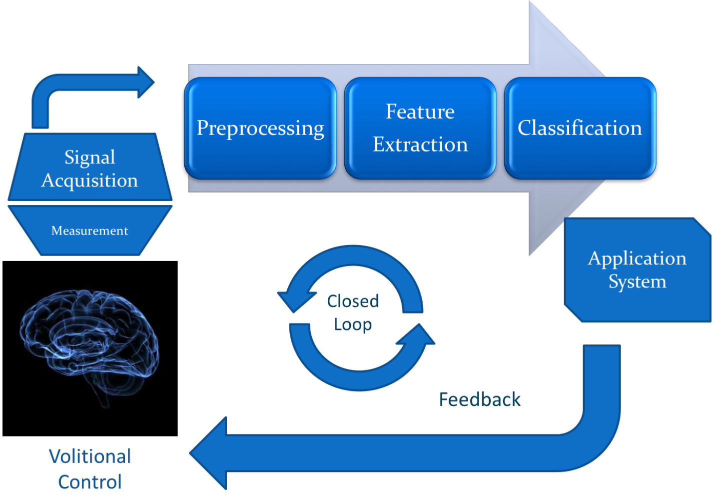
Noviembre 29 2018

Outline

Brain Computer Interfaces

Brain Computer Interfaces





General components of a BCI system.

Mapa de BCI con los diferentes paradigmas y las diferentes soluciones. También podría agregar una tabla con las velocidades alcanzadas.

Introduction

- BCI challenging technology.
- Outstanding advances but yet its push into mainstream technology has not fully materialized.
- More clinical and physician Involvement: devise mechanisms to help them stay in the loop.
- (+) speed, SNR, reliability, portability, usability.
- (-) biocompatibilities, setup, training time, calibration time, subject inter/intra variability.
- The search for practical, relevant, and invariant features that convey good-enough information [**Perdikis2014**].
- Ethical aspects should be revised [**Yuste2017**].

Evoked Potential
Steady State Visual Evoked Potential

SSVEP
SSAEP
SSSEP

Bereitschaftspotentials

Motor Imagery:
ERD/ERS: Event Related Desynchronization-Synchronization

Wadsworth BCI
Graz BCI

Selective Attention
P300,N400
Visual Spatial Evoked Attention

Mental Tasks

Tübingen BCI

Operant Conditioning:
Slow Cortical Potentials
ErrP

Berlin BCI

Digital and wearable electroencephalographs.

Motivation

- This thesis tries to unravel the following question: is it possible to analyze and discriminate electroencephalographic signals by automatic processing the shape of the waveforms using the Histogram of Gradient Orientations ?

What we aim to do

- A procedure to construct analyzable 2D-images based on one-dimensional signals.
- An enhancement over the Histogram of Gradient Orientation technique to allow non-squared patches and to adapt it to signal plots.
- A mapping procedure to link EEG time-series characteristics to features of 2D-images.
- A feature extraction method for EEG signals that can be used objectively to encode a representation of the waveform.
- A classification algorithm that use the encoded representation with the purpose of comparing and identifying waveforms for BCI applications.

Electroencephalography

- Permutation Entropy

Electroencephalography

- Permutation Entropy
- SHCC

Electroencephalography

- Permutation Entropy
- SHCC
- Matching Pursuit

Electroencephalography

- Permutation Entropy
- SHCC
- Matching Pursuit
- MIDS

Electroencephalography

- Permutation Entropy
- SHCC
- Matching Pursuit
- MIDS
- Local Binary Patterns

Electroencephalography

- Permutation Entropy
- SHCC
- Matching Pursuit
- MIDS
- Local Binary Patterns
- Peak Picking/aEEG/PAA

The Histogram of Gradient Orientations

- Signal Preprocessing

The Histogram of Gradient Orientations

- Signal Preprocessing
- Signal Segmentation

The Histogram of Gradient Orientations

- Signal Preprocessing
- Signal Segmentation
- Signal Plotting

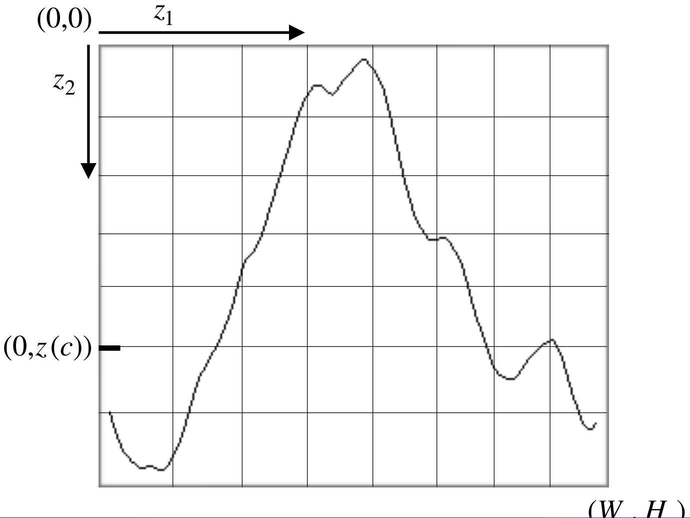
The Histogram of Gradient Orientations

- Signal Preprocessing
- Signal Segmentation
- Signal Plotting
- Keypoint Localization

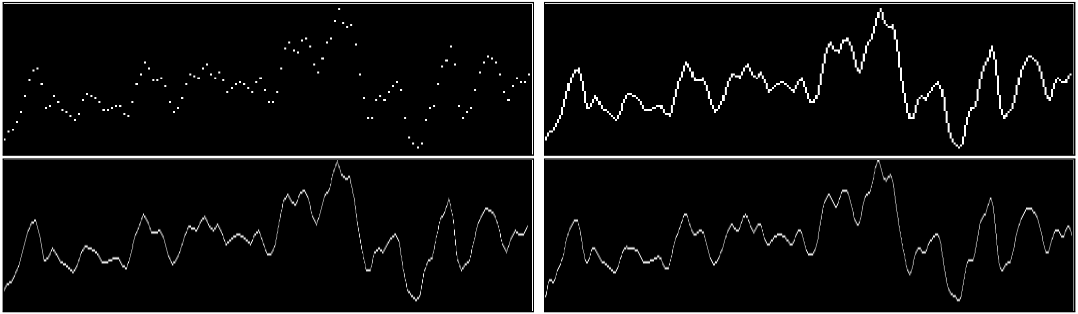
The Histogram of Gradient Orientations

- Signal Preprocessing
- Signal Segmentation
- Signal Plotting
- Keypoint Localization
- Calculation of the Histogram of Gradient Orientation

The Histogram of Gradient Orientations

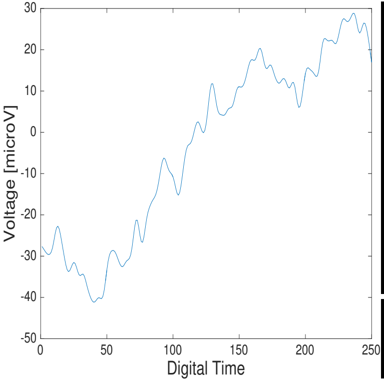


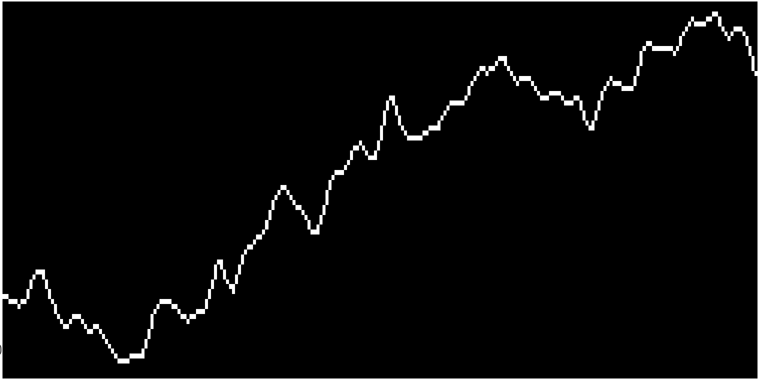
The Histogram of Gradient Orientations



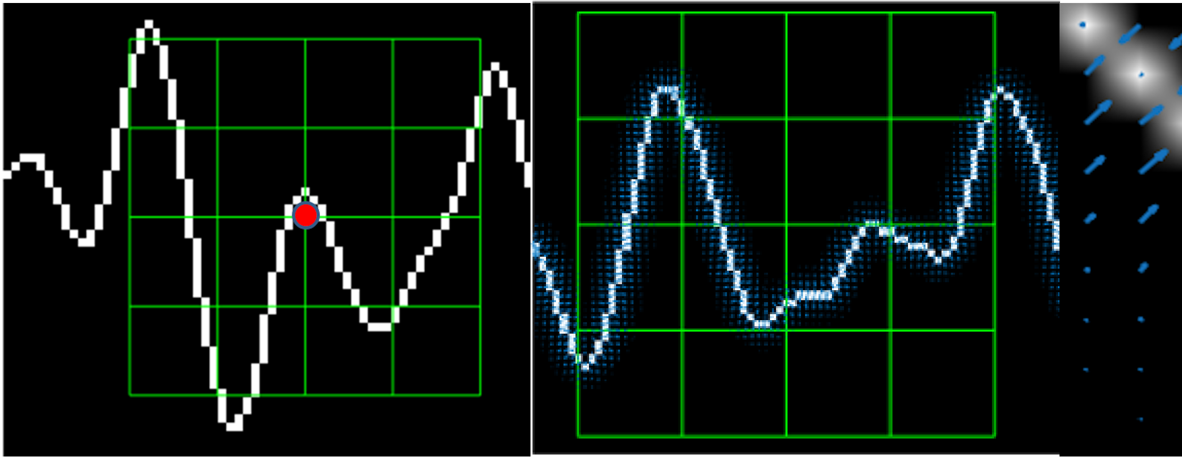
Generated images based on different interpolation schemes.

The Histogram of Gradient Orientations





The Histogram of Gradient Orientations



Patch and vector field of oriented gradients.

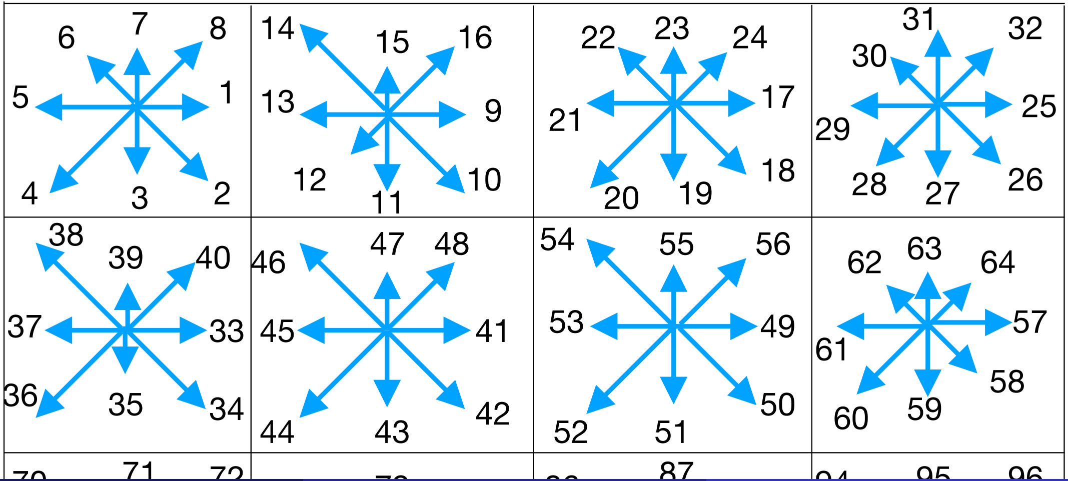
The Histogram of Gradient Orientations

Hence, for each spatial bin $i, j = \{0, 1, 2, 3\}$, corresponding to the indexes of each block $B_{i,j}$, the orientations are accumulated in a 3-dimensional histogram h through the following equation:

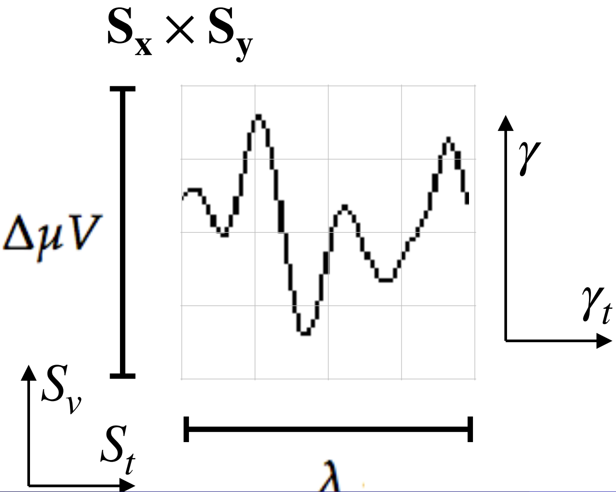
$$h(\theta, i, j) = \sum_{\mathbf{p}} \omega_{\text{ang}}(\angle J(\mathbf{p}) - \theta)\, \omega_{ij}\, (\mathbf{p} - \mathbf{kp})\ \|J(\mathbf{p})\| \qquad (1)$$

where $\mathbf{p}$ is a pixel from within the patch, θ is the angle bin with $\theta \in \{0, 45, 90, 135, 180, 225, 270, 315\}$, $\|J(\mathbf{p})\|$ is the norm of the gradient vector in the pixel $\mathbf{p}$, computed using finite differences, and $\angle J(\mathbf{p})$ is the angle of the gradient vector. The scalar $\omega_{\text{ang}}(\cdot)$ and vector $\omega_{ij}(\cdot)$ functions are linear interpolations used by [**Lowe2004**] and [**Vedaldi2010**] to provide a weighting contribution to eight adjacent bins.

The Histogram of Gradient Orientations



The Histogram of Gradient Orientations



BCI Paradigms

- Alpha Waves Wiggles

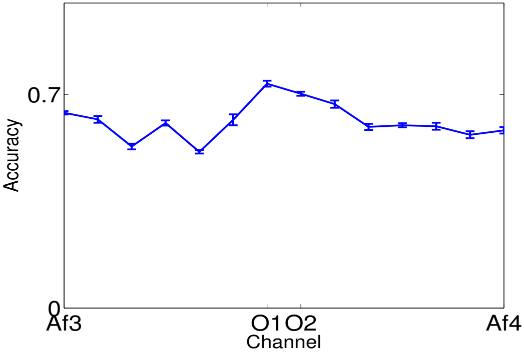
BCI Paradigms

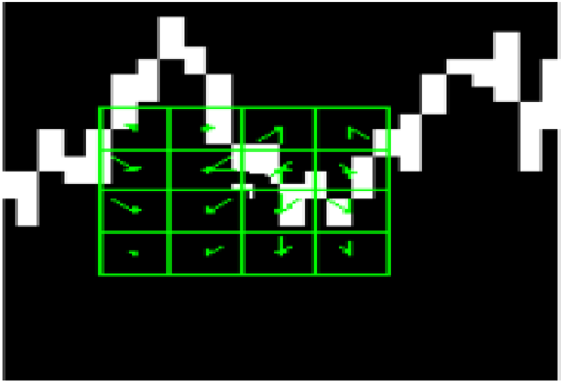
- Alpha Waves Wiggles
- Mu Greek Letter

BCI Paradigms

- Alpha Waves Wiggles
- Mu Greek Letter
- The P300 Wave

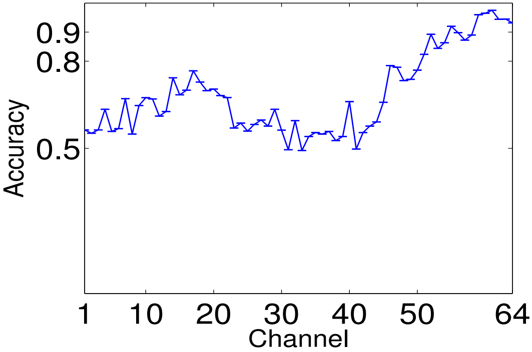
Alpha Waves wiggles

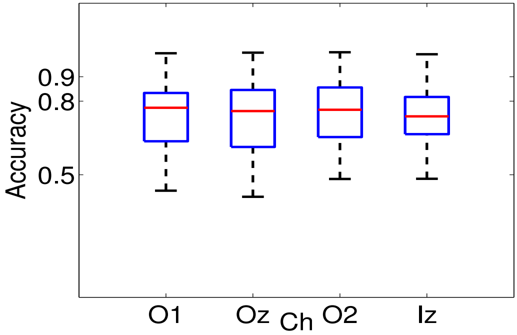




Dataset I: The classification accuracy is maximum on occipital channels O1 and O2. The horizontal patch scale S_t and the vertical patch scale S_v are set to 1, whereas γ and γ_t are set to 2, which corresponds to a variation of $\Delta \mu V = 10$ microvolts in the signal amplitude during $\lambda = 0.08$ seconds.

Alpha Waves wiggles



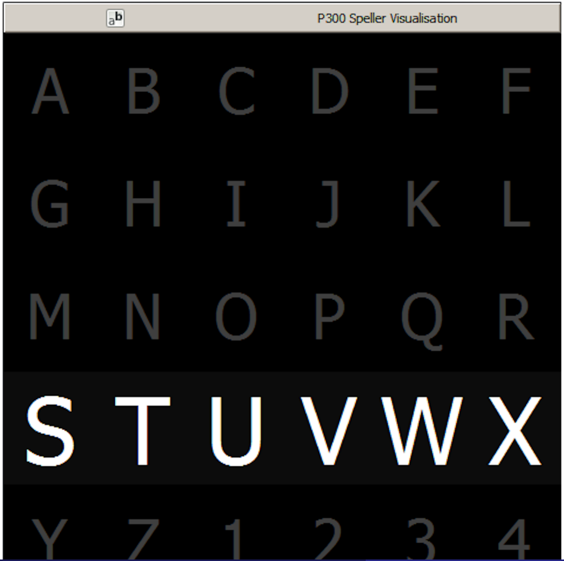


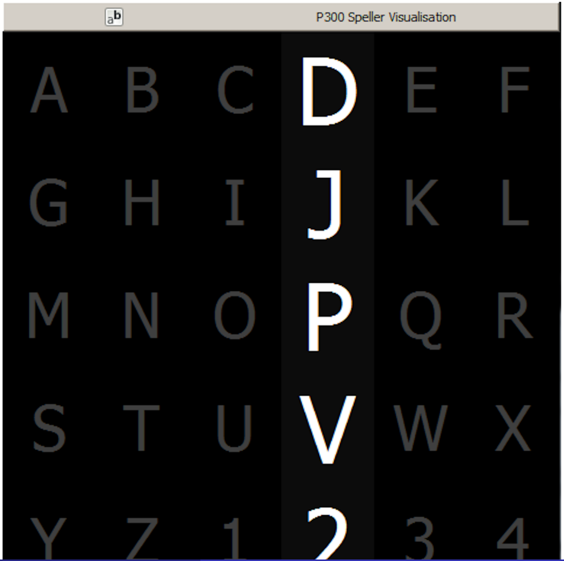
Dataset II: Classification Accuracy for segments of 1s ($N = 160$) of EEG, between class 1 and class 2. In this case as the sampling frequency F_s is lower, the signal span is $\lambda = 0.06$ s

.

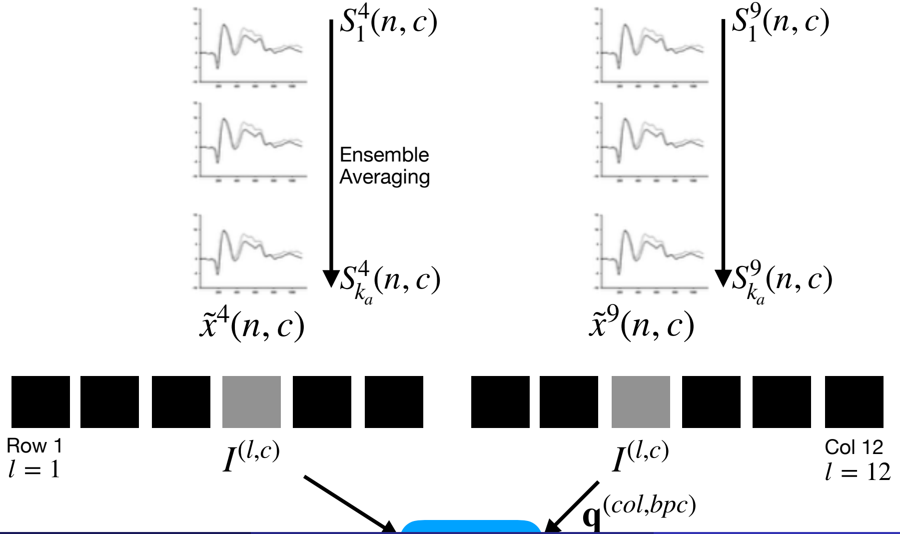
Mu a Greek letter

The P300 Wave





The P300 Wave



Step D: Match to the calibration template T^{bpc} by computing

$$\hat{row} = \arg \min_{l \in \{1,\ldots,6\}} \sum_{h=1}^{k} \left\| \mathbf{q}^{(l,bpc)} - \mathbf{d}_h^{(bpc)} \right\|^2 \quad (2)$$

and

$$\hat{col} = \arg \min_{l \in \{7,\ldots,12\}} \sum_{h=1}^{k} \left\| \mathbf{q}^{(l,bpc)} - \mathbf{d}_h^{(bpc)} \right\|^2 \quad (3)$$

with $\mathbf{d}_h^{(bpc)}$ belonging to the set $N_T(\mathbf{q}^{(l,bpc)})$, which is defined, for the best performing channel, as $N_T(\mathbf{q}^{(l,bpc)}) = \{\mathbf{d}_h^{(bpc)} \in T^{bpc} / \mathbf{d}^{(bpc)}$ is the k-nearest neighbor of $\mathbf{q}^{(l,bpc)}\}$. This procedure is a unary classification scheme, an adapted version of the algorithm described in Section **??** to the letter identification required in the BCI-Based P300 Speller implementation.

The P300 Wave

- Dataset I - P300 ALS Public Dataset

The P300 Wave

- Dataset I - P300 ALS Public Dataset
- Dataset II - P300 for healthy subjects

The P300 Wave

- Dataset I - P300 ALS Public Dataset
- Dataset II - P300 for healthy subjects
- Dataset III - P300 Pseudo-Real Dataset Generation

The P300 Wave

- Dataset I - P300 ALS Public Dataset
- Dataset II - P300 for healthy subjects
- Dataset III - P300 Pseudo-Real Dataset Generation
- Dataset IV - P300 Dataset IIb BCI Competition II (2003)

The P300 Wave

Table: Dataset I: Character recognition rates for the public dataset of ALS patients using HIST calculated from single-channel plots. Performance rates using single-channel signals with the SVM classifier are shown for comparison. The best performing channel *bpc* for each method is visualized

Participant	*bpc*	HIST	*bpc*	Single Channel SVM
1	Cz	35%	Cz	15%
2	Fz	85%	PO8	25%
3	Cz	25%	Fz	5%
4	PO8	55%	Oz	5%
5	PO7	40%	P3	25%
6	PO7	60%	PO8	20%
7	PO8	80%	Fz	30%
8	PO7	95%	PO7	85%

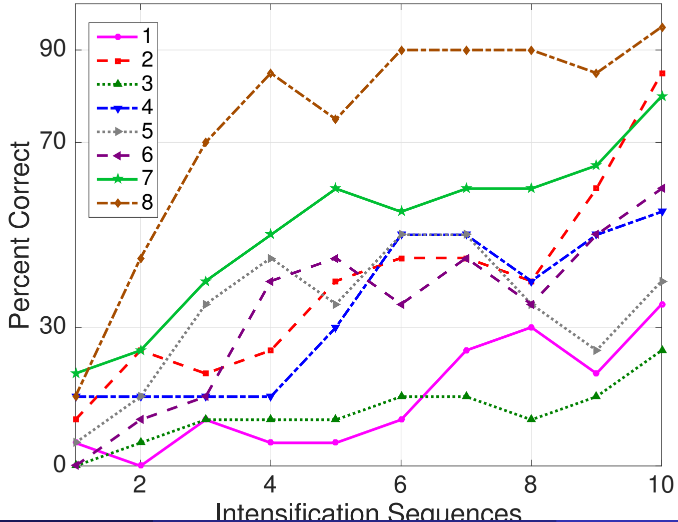

1
2
3
4
5
6
7
8
Percent Correct
90
70
30
0
2
4
6
8
10
Intensification Sequences

Table: Dataset II: Character recognition rates and *bpc* using HIST calculated from single-channel signals. Performance rates using single-channel signals with the SVM classifier are shown for comparison.

Participant	*bpc*	HIST	*bpc*	Single Channel SVM
1	Oz	40%	Cz	10%
2	PO7	30%	Cz	5%
3	P4	40%	P3	10%
4	P4	45%	P4	35%
5	P4	60%	P3	10%
6	Pz	50%	P4	25%
7	PO7	70%	P3	30%
8	P4	50%	PO7	10%

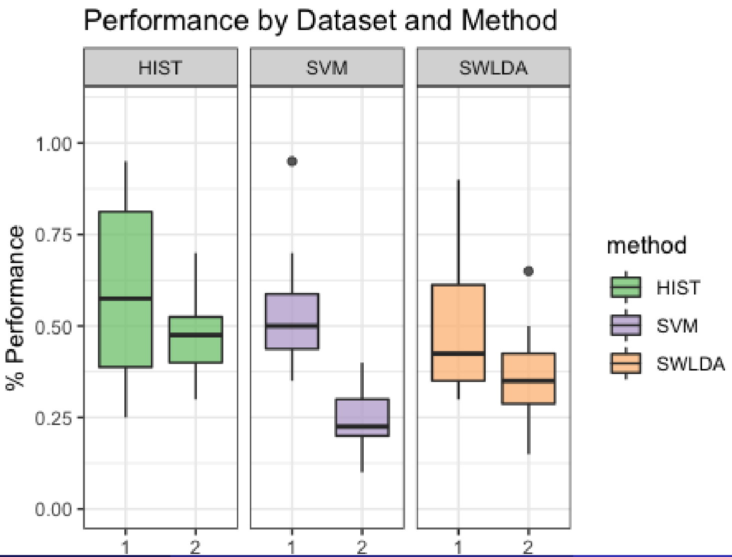
HIST method has an improved performance at letter identification than SVM that process the signals on a channel by channel strategy (Wilcoxon signed-rank test, $p = 0.004$ for both

The P300 Wave

Table: Character recognition rates and the best performing channel *bpc* for the public dataset I using the HIST versus performance rates obtained by SWLDA and SVM classification algorithms with a multichannel concatenated feature.

Participant	*bpc* for HIST	HIST	Multichannel SWLDA	Multichannel SVM
1	Cz	35%	45%	40%
2	Fz	85%	30%	50%
3	Cz	25%	65%	55%
4	PO8	55%	40%	50%
5	PO7	40%	35%	45%
6	PO7	60%	35%	70%
7	PO8	80%	60%	35%
8	PO7	95%	90%	95%

The P300 Wave



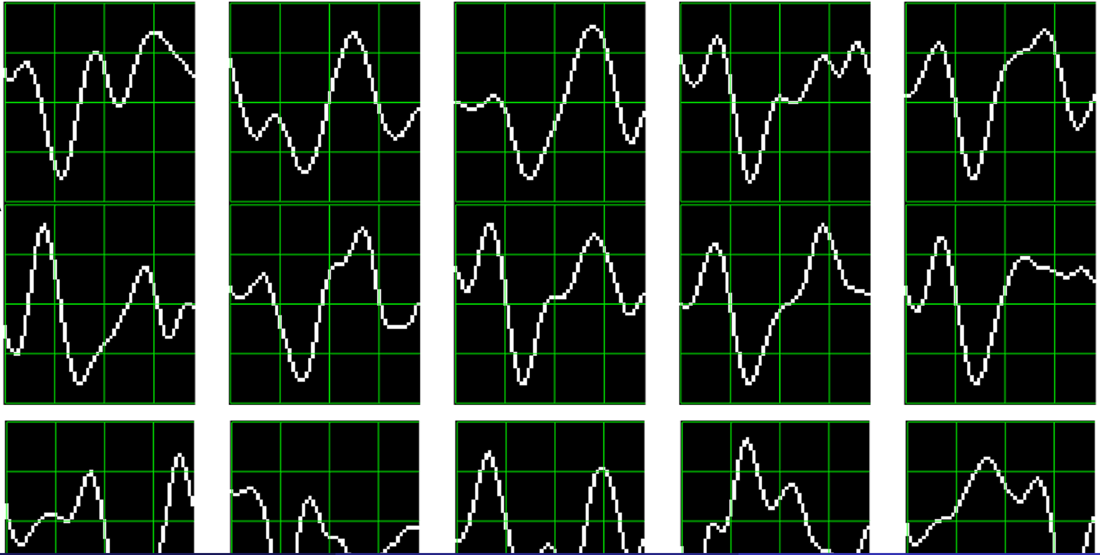
The P300 Wave

Table: Character recognition rates and the best performing channel *bpc* for the dataset II using HIST versus performance rates obtained by SWLDA and SVM classification algorithms with a multichannel concatenated feature.

Participant	*bpc* for HIST	HIST	Multichannel SWLDA	Multichannel SVM
1	Oz	40%	65%	40%
2	PO7	30%	15%	10%
3	P4	40%	50%	25%
4	P4	45%	40%	20%
5	P4	60%	30%	20%
6	Pz	50%	35%	30%
7	PO7	70%	25%	30%
8	P4	50%	35%	20%

The P300 Wave

A



Additionally, the stability of the P300 component waveform has been extensively studied in patients with ALS [**SellersandEmanuelDonchin2006**, **TomohiroMadarame2008**, **Nijboer2009**, **Mak2012**, **McCane2015**] where it was found that these patients have a stable P300 component, which were also sustained across different sessions. In line with these results we do not find evidence of a difference in terms of the performance obtained by analyzing the waveforms, by using the HIST method, for the group of patients with ALS and the healthy group of volunteers (Mann-Whitney U Test, $p = 0.46$). Particularly, the best performance is obtained for a subject from the ALS dataset for which, based on visual observation, the shape of they P300 component is consistently identified.

Dissertation Contributions

- A procedure to construct analyzable 2D-images based on one-dimensional signals.
- The introduction of the HIST method as an enhancement on the SIFT technique to allow non-squared patches and to adapt it to signal plots.
- A mapping procedure to link EEG time-series characteristics to features of 2D-images.
- A feature extraction method for EEG signals that can be used objectively to encode a representation of the waveform.
- A classification algorithm that use the encoded representation with the purpose of comparing and identifying waveforms for BCI applications.

Conclusion

- EEG Waveforms can be analyzed by this method.
- Oscillatory processes can be studied by this methodology.
- The stability of ERP transient components can be studied objetively with the proposed methology.

References

Questions

Questions?

Sample frame title

This is a text in first frame. This is a text in first frame. This is a text in first frame.

Sample frame title

This is a text in second frame. For the sake of showing an example.

- Text visible on slide 1

Sample frame title

This is a text in second frame. For the sake of showing an example.

- Text visible on slide 1
- Text visible on slide 2

Sample frame title

This is a text in second frame. For the sake of showing an example.

- Text visible on slide 1
- Text visible on slide 2
- Text visible on slide 3

Sample frame title

This is a text in second frame. For the sake of showing an example.

- Text visible on slide 1
- Text visible on slide 2

- Text visible on slide 4

Sample frame title

$$\omega_{\mathrm{ang}}(\alpha) = \sum_{r=-1}^{1} \omega\left(\frac{8\alpha}{2\pi} + 8r\right) \tag{4}$$

Sample frame title

In this slide, some important text will be highlighted beause it's important. Please, don't abuse it.

Remark

Sample text

Important theorem

Sample text in red box

Examples

Sample text in green box. "Examples" is fixed as block title.

Method

Single Channel transformation of the EEG multichannel time series matrix into an image

$$\mathscr{T}\{x(t, c, V) = 0\} \mapsto I(z_1, z_2, \mathcal{C}, \mathcal{I}) = 0 \qquad (5)$$

where t is time, c is the specified channel, V is the voltage value for the signal, $\mathcal{C}$ is the color channel for an image and $\mathcal{I}$ is the pixel value intensity.

Plot Generation: The EEG matrix is transformed to a binary bidimensional image $(t, c, V) \mapsto (t, V, Grey, \mathcal{I})$ with $\mathcal{I} = 0$ or $\mathcal{I} = 255$ for each c.

Signal Transformation: Visually centering the signal over the image.

First the non-zero media is removed from the signal.

$$\tilde{x}(t,c) = \lfloor \delta.(x(t,c) - \bar{x}(c)) \rfloor \tag{6}$$

And the signal is centered on the image

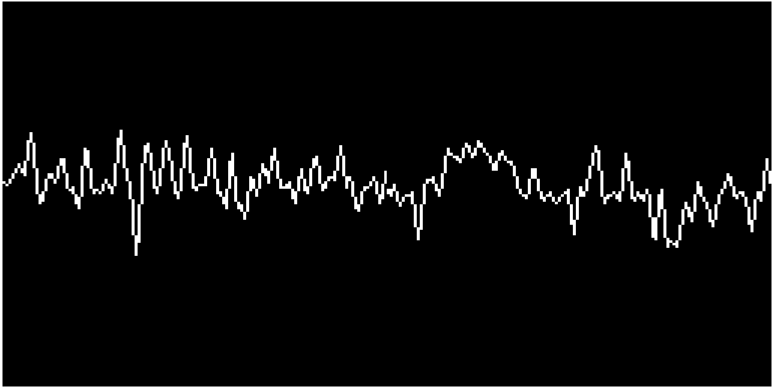
$$h(c) = (\max \tilde{x}(t,c) - \min \tilde{x}(t,c)) + \sigma \tag{7}$$

$$Z(c) = \lfloor \frac{h(c)}{2} \rfloor - \lfloor \frac{\max \tilde{x}(t,c) + \min \tilde{x}(t,c)}{2} \rfloor \tag{8}$$

where t is time, δ is scale factor, c is the channel parameter, $x(t,c)$ is the EEG matrix whereas $\bar{x}(c)$ is the mean value for each channel, $h(c)$ is the height of the image in pixels, σ is the descriptor size and $Z(c)$ is the horizontal pixel at which the zero value of the signal will be located.

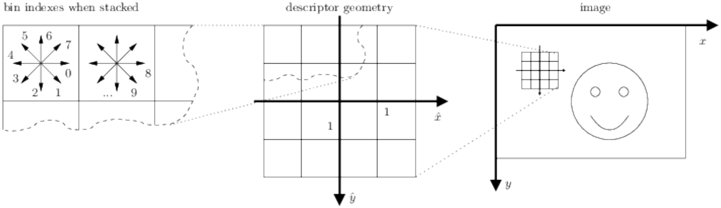
Signal Transformation: Binary Image generation.

$$I(z_1, z_2) = \begin{cases} 255 & z_1 = \delta \cdot t; z_2 = \tilde{x}(t,c) + Z(c) \\ 0 & \text{otherwise} \end{cases} \quad (9)$$



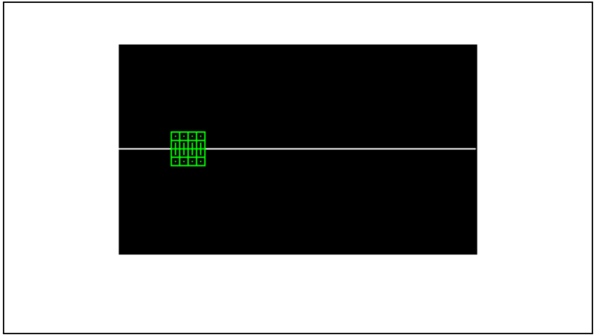
Features: SIFT[1] Descriptors

Scale Invariant Feature Transform Descriptors are local features of an image that represents gradient changes in intensities. They are 128-dimensional vectors that contains the histograms of relative gradient directions on each of the blocks that each patch is divided (4x4=16 blocks, 8 rotational directions on each). A single scale ($\sigma = 1$) is composed of 4 blocks of 3 pixels on each side (12x12).



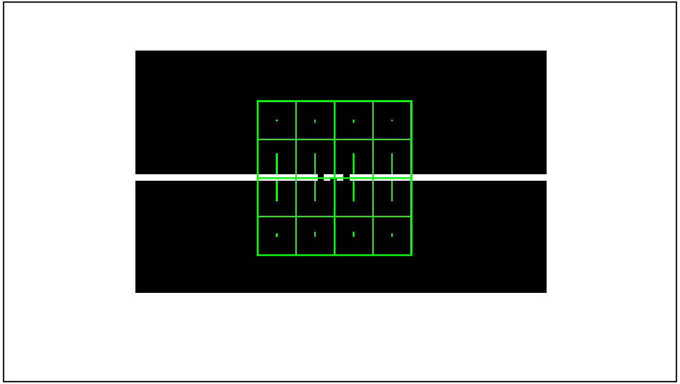
[1] **Lowe 2004**.

SIFT Descriptors



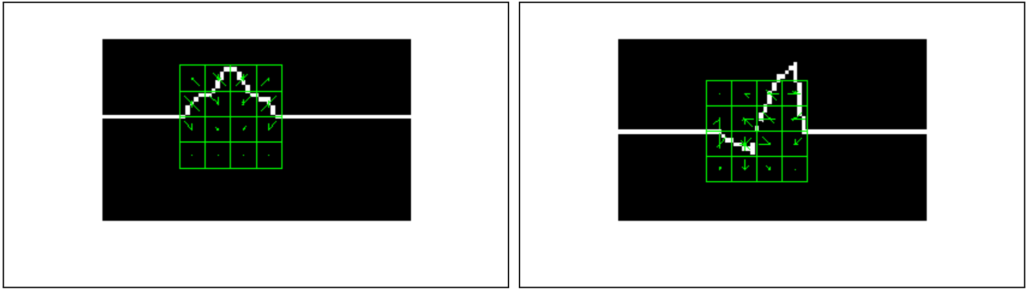
SIFT Descriptor $[z_1, z_2, \theta, \sigma]$ where (z_1, z_2) are the 2D coordinates where the *Keypoint* is located, θ is the descriptor general orientation and σ is the descriptor size.

SIFT Descriptors



SIFT Descriptor with its corresponding gradient tendencies.

SIFT Descriptors



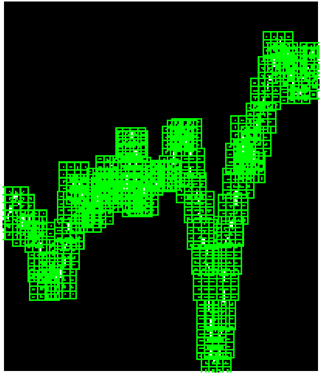
Sample Descriptors from artificial signals.

SIFT Descriptors

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	22	22	17	173	173	173	173	40	51	51	40	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	40	51	51	40	173	173	173	173	17	22	22	17
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Sample descriptor values of the given patch.

Keypoint Localization



Classification

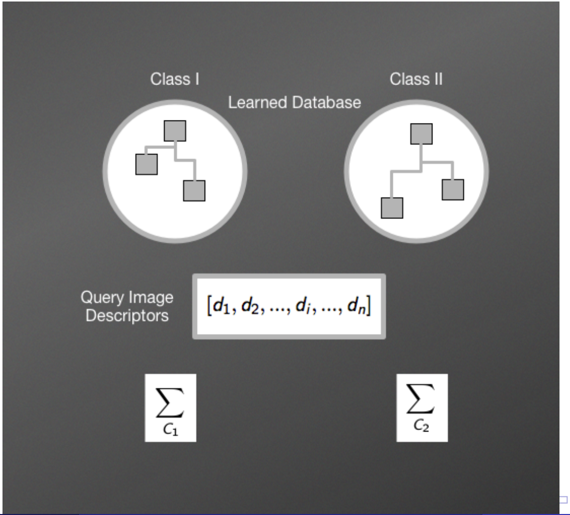
- Discriminative Semi-supervised classification method was used: Naive Bayes Nearest Neighbor, NBNN[2] algorithm:

$$\hat{C} = \arg\min_{C} \sum \|d_i - NN_C(d_i)\|^2 \tag{3}$$

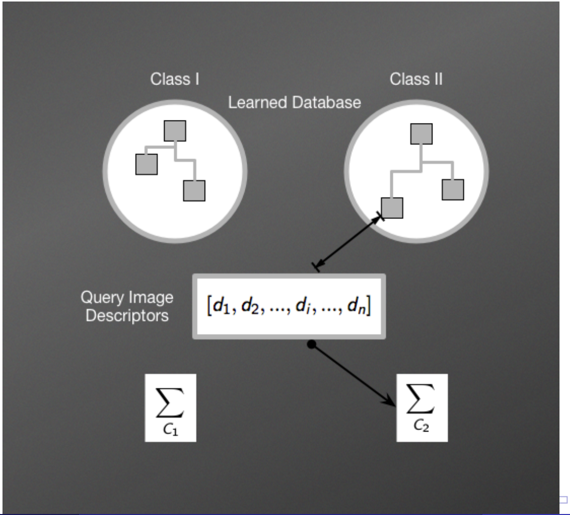
where $\hat{C}$ is estimated Class to which this image (and underlying EEG signal windows) should belong whereas d_i is the i-th descriptor obtained from the query image and $NN_C(d_i)$ is the near neighbor descriptor for each class.

[2] **Boiman, Shechtman, Irani 2008**.

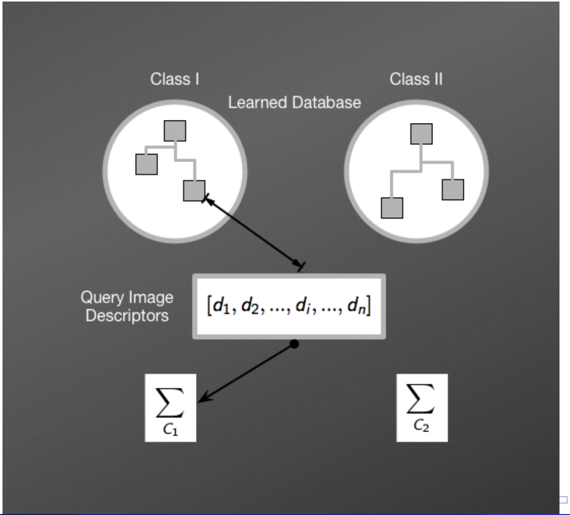
Classification



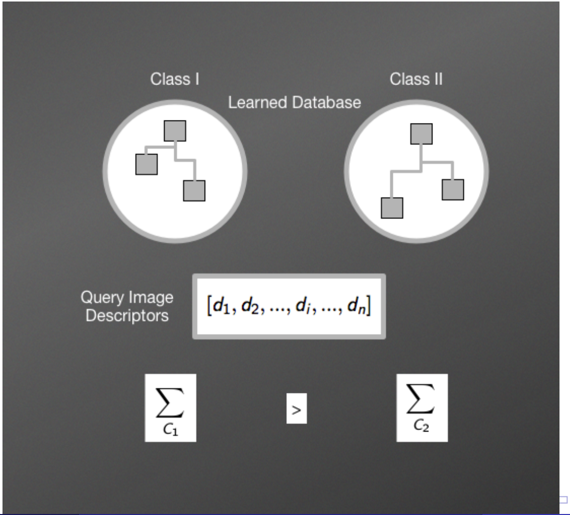
Classification

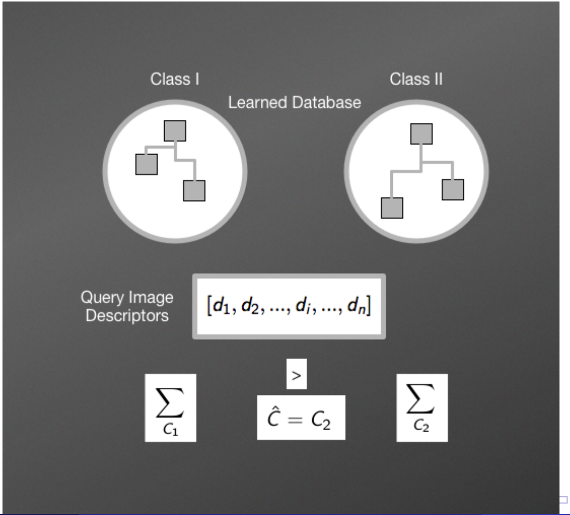


Classification



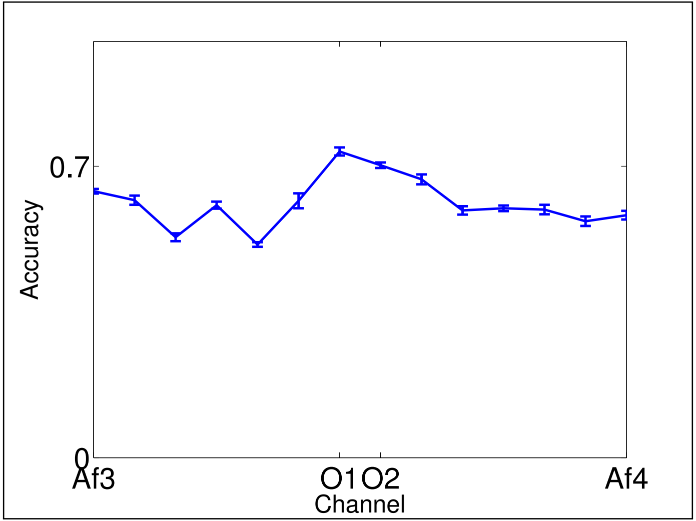
Classification




Class I
Learned Database
Class II
Query Image Descriptors
$[d_1, d_2, ..., d_i, ..., d_n]$
>
$\sum_{C_1}$
$\hat{C} = C_2$
$\sum_{C_2}$

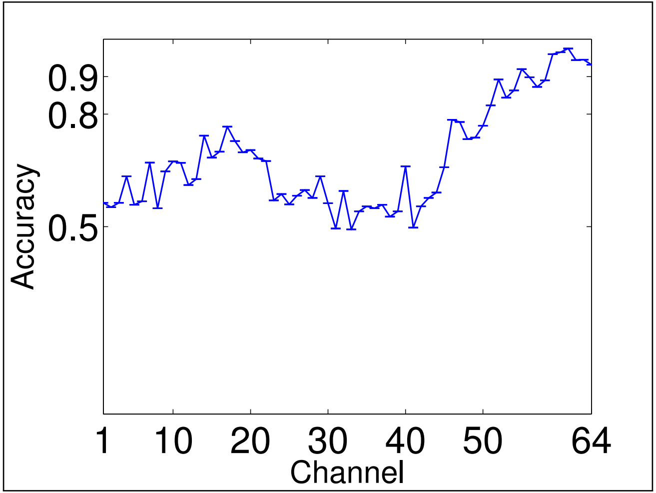
Offline Results

Dataset I: Subject Independent α Waves[3]



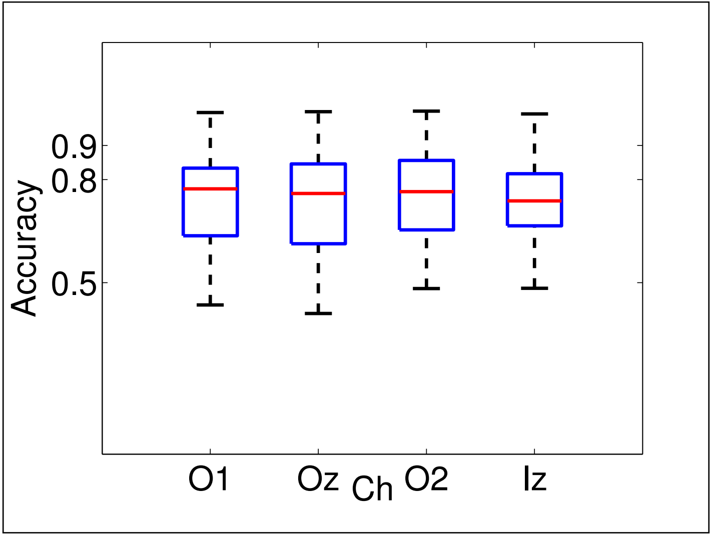
10-Fold Cross validated accuracy values for 10 subjects.

Dataset II: EEG Dataset, Runs 1 and 2[4]

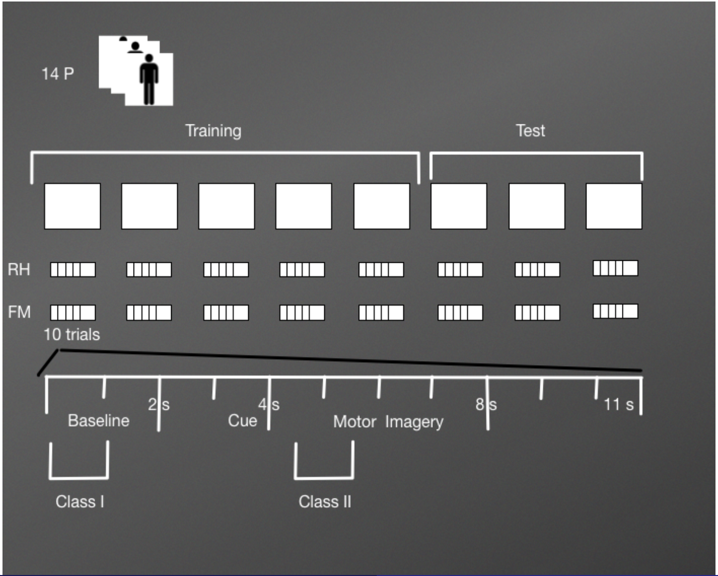


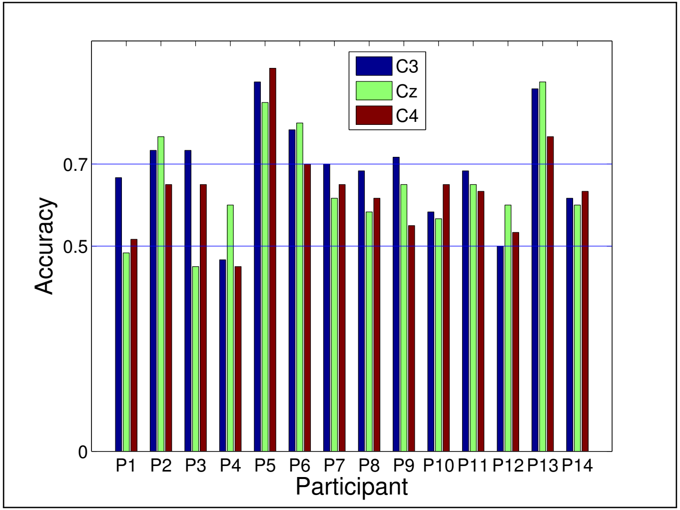
10-Fold Cross validated accuracies values for one random subject.

Dataset II: EEG Dataset, Runs 1 and 2[5]

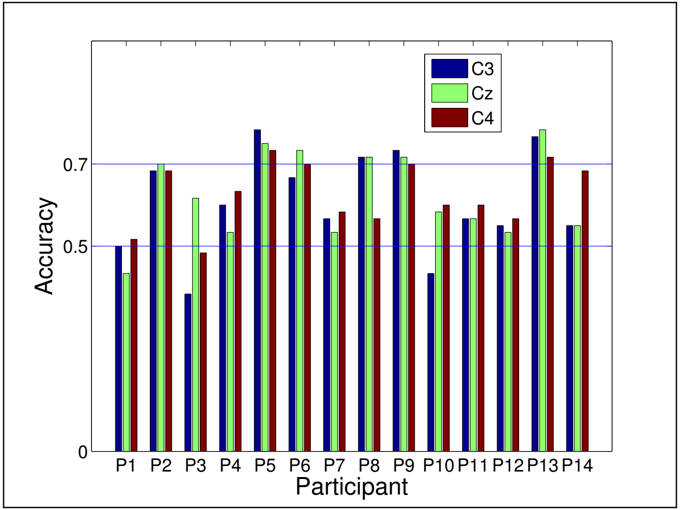


Dataset III: Motor Imagery[6]

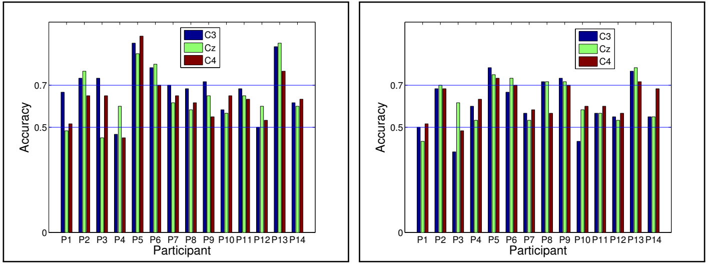




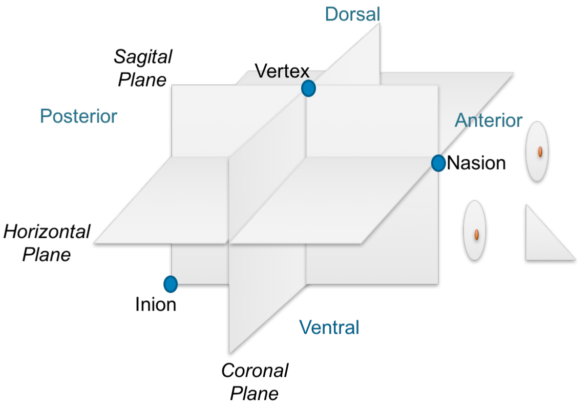
Accuracy for the BCI Simulation classifying Baseline vs. RH (Right Hand) motor imagery.

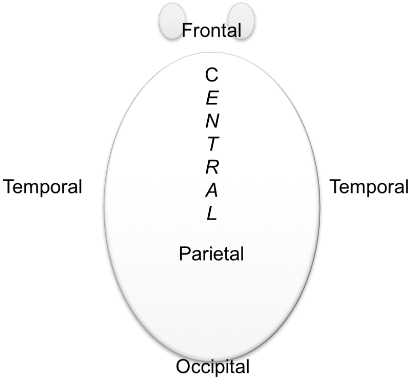


Accuracy for the BCI Simulation classifying Baseline vs. FM (Feet Movement) motor imagery.

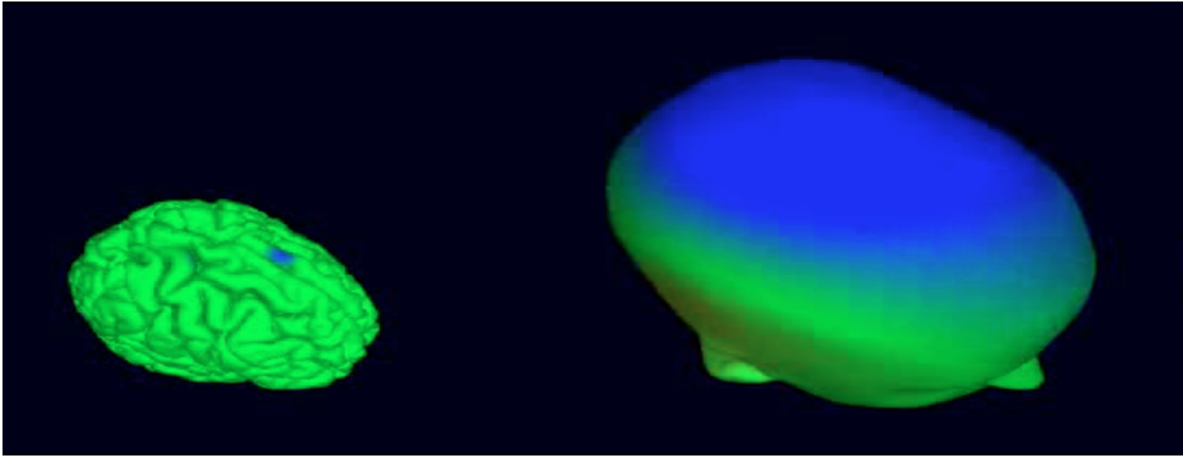


Comparative results obtained for the Offline BCI Simulation using MI RH (left) and MI FM (right)



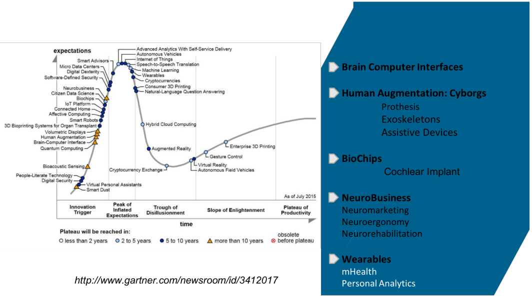


Neuronal Planes regularly used in neuroscience research. In BCI they are used to understand electrode location and spatial filters.



A source signal with positive/negative polarity is generated in a very specific region of the brain but due to volume conduction their influence affects a widespread area of the scalp where sensors are located (Image of the brain from Swartz Center for Computational Neuroscience).

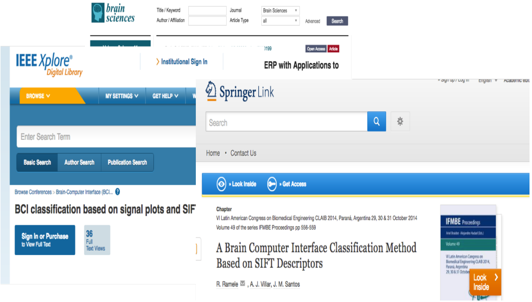
Digital and wearable electroencephalographs.



Digital and wearable electroencephalographs.

Digital and wearable electroencephalographs.

Digital and wearable electroencephalographs.



Digital and wearable electroencephalographs.

Table: EEG waveforms descriptions found in the surveyed literature.

Method	Phenomena
Positive Rounded Component	α-Waves, Epilepsy
Rising and Falling Phase	Epilepsy
Terminal plateau	Epilepsy
Ripples and Wiggles	Epilepsy, ERP
Sinusoidal Shape	Epilepsy
Sawtooth	Motor Imagery, Sleep
Sharpness or Spike-like	Epilepsy
Rectangular	Epilepsy
Line length	Anomaly Detection
Root Mean Square	Anomaly Detection
Wicket Shape	Epilepsy
Peak and Trough Sharpness Ratio	Epilepsy
Symmetry between rise and decay phase	Epilepsy
Slope Ratio	Sleep

Table: EEG waveforms descriptions found in the surveyed literature.

Method	Phenomena
Positive/Negative Peak Amplitude	ERP
Positive vs Negative Ratio	Sleep K-Complex
Base-to-Peak Amplitude	ERP
Peak-to-Peak Amplitude	ERP
Positive/Negative Peak Latency	ERP
Integrated Activity	ERP, Epilepsy,ICU
Cross-Correlation	ERP, Epilepsy, Sleep
Coupling	
Cross Frequency, Phase-Amplitude, Phase-Phase	Sleep
Period Amplitude Analysis	ERP, Epilepsy