Modeling and recognition of waveforms by multiresolution methods with application to hdEEG

Mauro Zucchelli

October 24, 2012

Contents

1. Introduction
2. Methods
3. Results and Discussion
4. Conclusions

Contents

1 Introduction

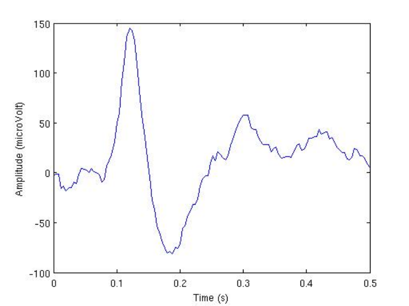
2 Methods

3 Results and Discussion

4 Conclusions

Objective

The purpose of this work was to focus on a particular pathology, namely temporal lobe epilepsy, in order to detect, analyze and model the so-called interictal spikes.



Background

- Temporal lobe epilepsy is a type of focal epilepsy, arising from one or both temporal lobes of the brain.
- Interictal spikes (also called spike and wave oscillations) are a particular type of waveform characterizing the EEG tracing of patients suffering particular types of epilepsy.
- hdEEG is a recording method that relies on numerous electrodes to increase spatial sampling (128-256 against "normal" 20-32) to enhance the spatial resolution of the signal.

Background

Interictal Spikes

Background

EEG vs hdEEG

EEG

hdEEG

State Of Art

SOA

1. Michel and Murray [2012] summarize the advantages of using EEG as a brain imaging tool;
2. Wilson and Emerson [2002] review of spike detection algorithm;
3. Casson A.J. et al., [2009] important metrics to asses the performance of a spike detection algorithm;
4. S. Mallat and Z. Zhang [1993] introduce Matching Pursuit;
5. C.G. Bénar et al., [2009] invent Consensus Matching Pursuit;

State Of Art

SOA

1. Michel and Murray [2012] summarize the advantages of using EEG as a brain imaging tool;
2. Wilson and Emerson [2002] review of spike detection algorithm;
3. Casson A.J. et al., [2009] important metrics to asses the performance of a spike detection algorithm;
4. S. Mallat and Z. Zhang [1993] introduce Matching Pursuit;
5. C.G. Bénar et al., [2009] invent Consensus Matching Pursuit;

State Of Art

SOA

1. Michel and Murray [2012] summarize the advantages of using EEG as a brain imaging tool;
2. Wilson and Emerson [2002] review of spike detection algorithm;
3. Casson A.J. et al., [2009] important metrics to asses the performance of a spike detection algorithm;
4. S. Mallat and Z. Zhang [1993] introduce Matching Pursuit;
5. C.G. Bénar et al., [2009] invent Consensus Matching Pursuit;

State Of Art

SOA

1. Michel and Murray [2012] summarize the advantages of using EEG as a brain imaging tool;
2. Wilson and Emerson [2002] review of spike detection algorithm;
3. Casson A.J. et al., [2009] important metrics to asses the performance of a spike detection algorithm;
4. S. Mallat and Z. Zhang [1993] introduce Matching Pursuit;
5. C.G. Bénar et al., [2009] invent Consensus Matching Pursuit;

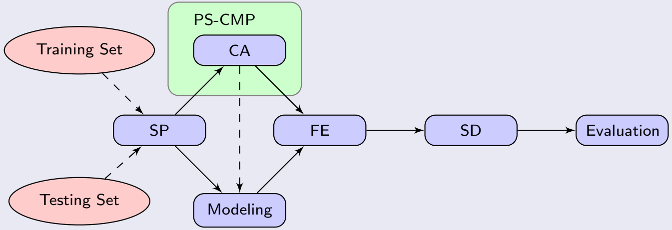
State Of Art

SOA

1. Michel and Murray [2012] summarize the advantages of using EEG as a brain imaging tool;
2. Wilson and Emerson [2002] review of spike detection algorithm;
3. Casson A.J. et al., [2009] important metrics to asses the performance of a spike detection algorithm;
4. S. Mallat and Z. Zhang [1993] introduce Matching Pursuit;
5. C.G. Bénar et al., [2009] invent Consensus Matching Pursuit;

Pipeline

The approach



A fully automatic approach is proposed, based on multiresolution sparse overcomplete representations that are used to both model the interictal waveforms and to derive features to be exploited for signal detection.

Contents

1 Introduction

2 Methods

3 Results and Discussion

4 Conclusions

Matching Pursuit

MP

Given a dictionary of waveforms $D = \Psi(\vec{p})$ of size P which at least contains N linearly independent functions (with $P > N$), the corresponding sparse regression problem aims at finding signal expansions of the form:

$$s(t) = \sum_{i=1}^{I} a_i \Psi_{\vec{p}_i}(t) + N(t) \tag{1}$$

Matching Pursuit

Dictionary

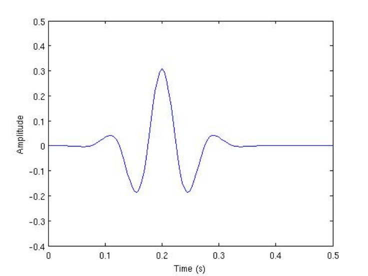
An over-complete dictionary D of Gabor wavelet is chosen, with changing modulation frequency f_0, spread σ and temporal jitter u and extending the Bénar proposing dictionary the phase ρ:

$$\Psi_{f_0,\sigma,\rho,u}(t) = \Psi_{f_0,\sigma,\rho}(t-u) \quad (2)$$

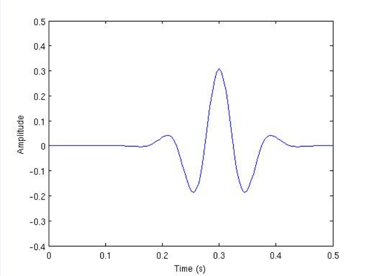
$$\Psi_{f_0,\sigma,\rho}(t) = \frac{1}{(4\pi\sigma)^{1/4}} \exp\{2j\pi f_0 t + j\rho\} \exp\left\{\frac{-t^2}{2\sigma^2}\right\} \quad (3)$$

Examples 1

Latency u



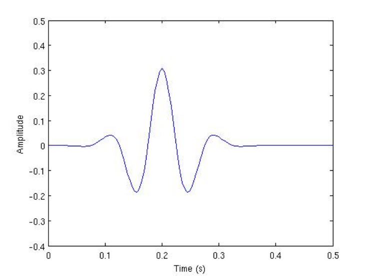
u=200 ms



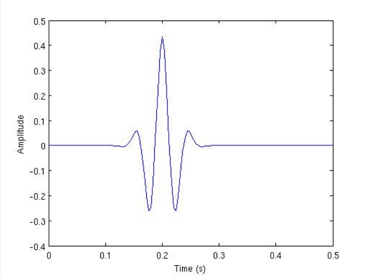
u=300 ms

Examples 2

Central Frequency f_0



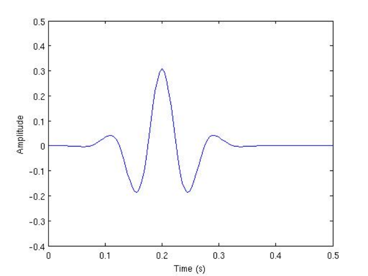
f_0=10 Hz



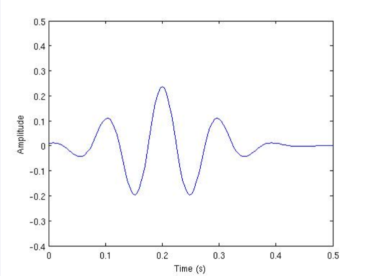
f_0=20 Hz

Examples 3

Oscillation ξ



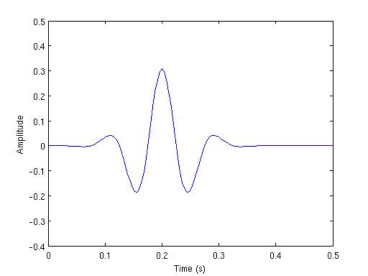
ξ=3



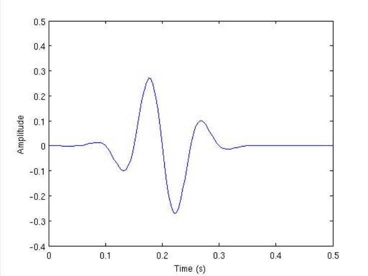
ξ=5

Examples 4

Phase ρ



$\rho=0$



$\rho=\pi/2$

Matching Pursuit

MP

The decomposition of a signal s at step $i+1$ is given by the residue of the signal at step i minus the atom found at iteration i multiplied by its amplitude coefficient $\langle s^i, \Psi_{\vec{p}_i} \rangle$:

$$s^{i+1} = s^i - \langle s^i, \Psi_{\vec{p}_i} \rangle \Psi_{\vec{p}_i}; \tag{4}$$

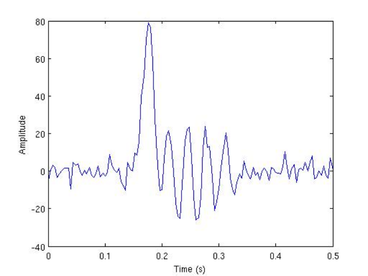
where the parameters $\vec{p}_i$ of the atoms for each iteration i is given by:

$$\vec{p}_i = \underset{\vec{p}_i}{argmax} \, |\langle s^i, \Psi_{\vec{p}_i} \rangle|; \tag{5}$$

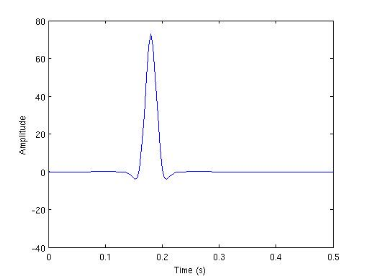
n.b. Of course at step 0, s_0 is equal to the original signal.

Matching Pursuit: an example

s_0



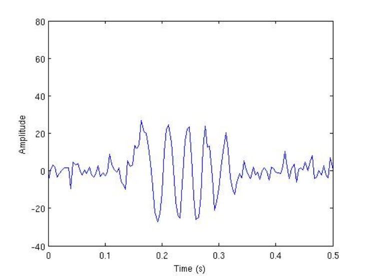
s_0



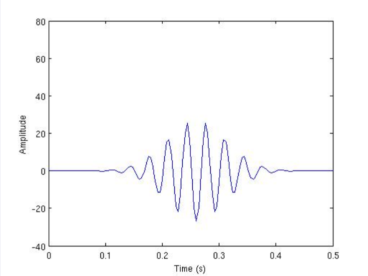
Ψ_1

Matching Pursuit: an example

s_1



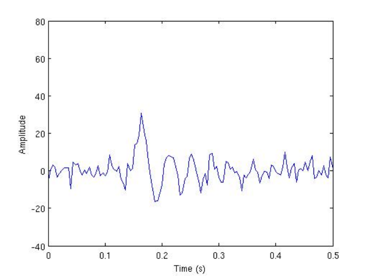
s_1



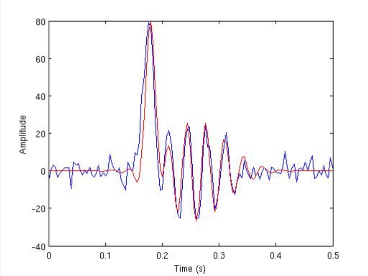
Ψ_2

Matching Pursuit: an example

s_2



s_2



Reconstruction

Consensus Matching Pursuit

CMP

Consensus Matching Pursuit is an evolution of classical MP to work on multitrial signals, in CMP is replaced by:

$$s_k^{i+1} = s_k^i - M_k^i \Psi_{\vec{\hat{p}}_i} \tag{6}$$

where the amplitudes M_i^k are given by

$$M_i^k = \langle s^i, \Psi_{\vec{\hat{p}}_i} \rangle \tag{7}$$

The goal is to find atoms identified by the parameters set $\vec{p}_i^C(k)$ that form a set of prototypes and to derive the set of its neighbors best matching each realization.

Consensus Matching Pursuit

Step 1

For each trial k, the projection of the data $s_k(t)$ on all atoms of the dictionary is computed and the magnitude is extracted:

$$M_k^i(\vec{p}) = | < s_k^i(t), \Psi(\vec{p}) > | \tag{8}$$

the local maxima are then selected and retained:

$$L^i(k) = \{\vec{p} | M_k^i(\vec{p}) \ locally \ maximum \ at \ \vec{p}\} \tag{9}$$

Consensus Matching Pursuit

Step 2

Each local maximum $\vec{p}_l$ votes in an accumulator map $V^i(\vec{p})$ by adding to the accumulator a kernel C, centered on this local maximum.

$$V^i(\vec{p}) = \sum_k \sum_{\vec{p}_l \in L^i(k)} M_k^i(\vec{p}_l) C(\vec{p}, \vec{p}_l). \tag{10}$$

The kernel $C(\vec{p}, \vec{p}_l)$ defines a similarity measure, that indicates the spread of each local maximum. The consensus atom $\vec{A}_i$ at iteration i is defined as:

$$\vec{A}_i = \underset{\vec{p}}{argmax} \left\{ V^i(\vec{p}) \right\} \tag{11}$$

Consensus Matching Pursuit

Step 3

For each individual trial k, the parameter most similar to $\vec{A}_i$ is selected among the local maxima $L^i(k)$:

$$\vec{\hat{p}}_i(k) = \underset{\vec{p} \in L^i(k)}{argmax} \left\{ M_i^k(\vec{p}) C(\vec{p}, \vec{A}_i) \right\} \tag{12}$$

With the parameter vector $\vec{\hat{p}}_i(k)$ the atom $\Psi_{\vec{\hat{p}}_i(k)}$ can be calculated and used to find the residue s_k^{i+1}.

Consensus Matching Pursuit

CMP reconstruction

$\Psi_{\vec{\hat{p}}_i}$ is used to calculate the reconstruction of the signal:

$$r_k(t) = \sum_{i=1}^{I} \Psi_{\vec{\hat{p}}_i} M_k(\vec{\hat{p}}_i); \tag{13}$$

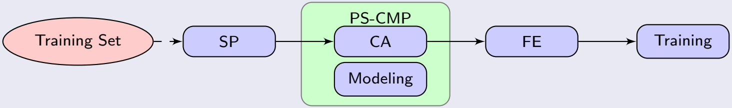
With which is possible to calculate the goodness of fit:

$$GOF_k = 1 - \frac{\|s_k(t) - r_k(t)\|^2}{\|s_k(t)\|^2}; \tag{14}$$

For each trial k.

PS-CMP

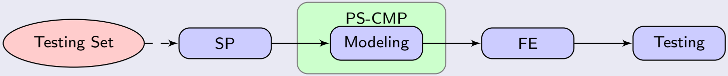
Training set



In the first phase CMP is performed on a group of signals of the same type called training set. Important measures like GOF and the amplitude of the atoms are $M_k(\vec{\hat{p}}_i)$ are extracted from CMP and will be used a features of the training set of the classifier.

PS-CMP

Testing set



Waveforms that are similar in shape to those in the training set will be called *w*, while all the others will be called *b*. PS-CMP is run on this set, but using the Consensus Atoms found with the training phase. This means that only *Step 1* and *Step 3* are executed.

PS-CMP

Features

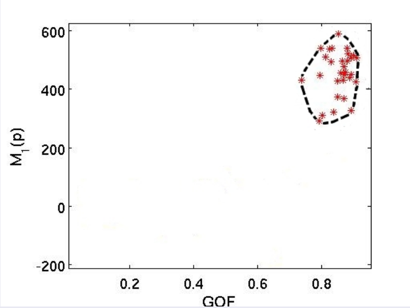
It has been observed that:

1. The cardinality of atoms (I) used for the decomposition must be the lowest possible;
2. GOF_{target} is generally lower of $GOF_{non-target}$
3. $M_k(\vec{\hat{p}}_i)$ must be considered an important feature.

Classification

One Class - SVM

One Class Support Vector Machine, are a variation of the normal SVM that adapts the classical SVM idea to the description a single class of data: in this case the sphere with minimal volume (or minimal radius) containing all objects has to be found.



Contents

1 Introduction

2 Methods

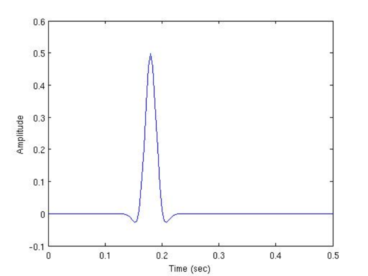
3 Results and Discussion

4 Conclusions

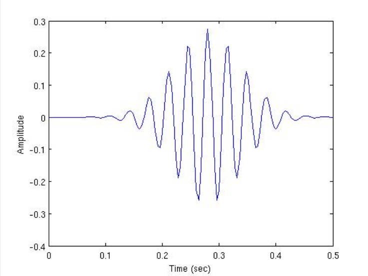
Simulated dataset

Training set: Gabor bases

In training set τ each signals was obtained by the combination of two Gabor atoms respectively representing a transient signal (spike) and a slow oscillation.



$\langle\ f_0 = 10\text{Hz},\ \xi=1,\ \rho=0\ \rangle$

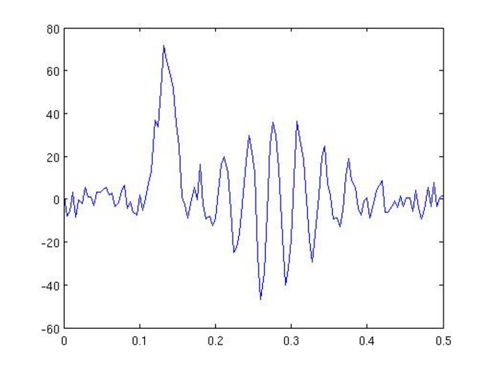


$\langle\ f_0 = 30\text{Hz},\ \xi=11,\ \rho=0\ \rangle$

Simulated dataset

Training set

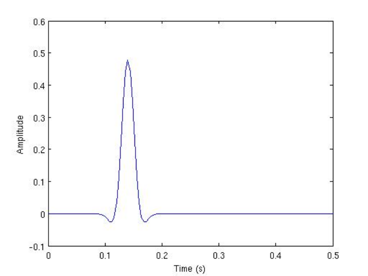
With this parameters but varying latency, amplitude and noise, a 30 trials training set is created.

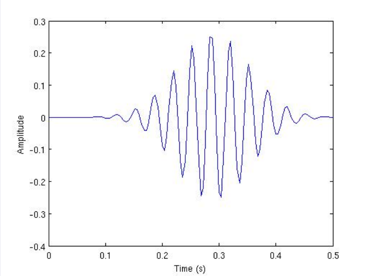


Simulated dataset

Training set: CA

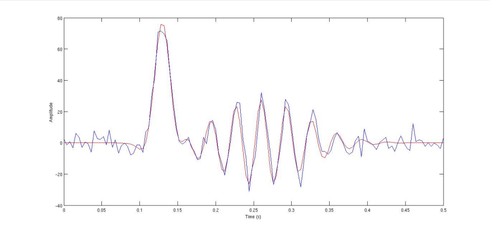
PS-CMP was then performed on τ. Only two iterations were performed for decomposing the signals, leading to only two atoms.





Simulated dataset

Training set: reconstruction



GOF_k was calculated for every trial k:

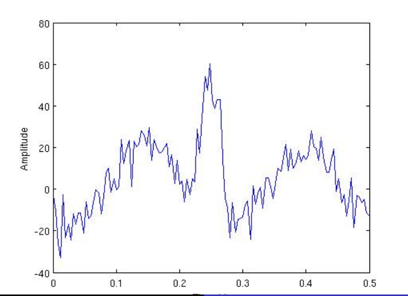
	GOF
mean	0.8246
std	0.0600

Simulated dataset

Testing set

A **testing set** T of 100 signal was built:

- 50 were composed in the same way as the training set (w)
- 50 were obtained by the sum of two Gabor functions with random parameters (b).



Simulated dataset

Testing set: GOF

The GOF mean and standard deviation for w and b:

	GOF_w	GOF_b
mean	0.7820	0.2200
std	0.0767	0.2007

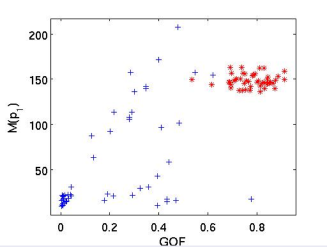
Projection coefficients of the waveforms onto the atoms, in particular the one corresponding to the first atom $M_k(\vec{\hat{p}}_1)$:

	$M(\hat{p}_1^w)$	$M(\hat{p}_1^b)$
mean	148.4558	56.5587
std	6.8270	55.2301

Simulated dataset

Testing set: classification

OCSVM was trained on τ and the classification was performed on the testing set T.

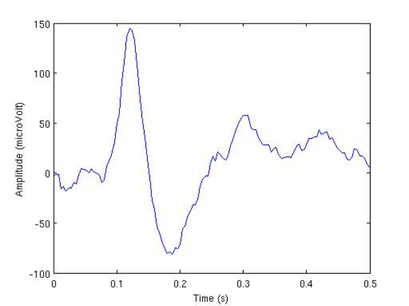


This classifier equipped with a proper parameters setting performed well on 97 trials out of 100

Real dataset

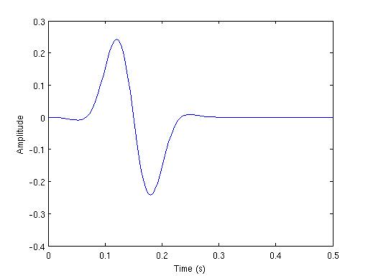
Training set

A training set of 28 spikes was manually selected from a channel of the temporal lobe of an epileptic patient.

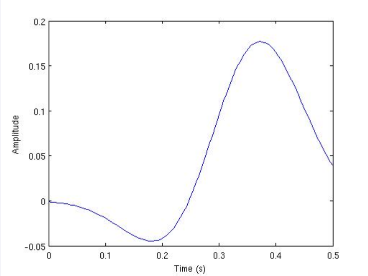


Real dataset

Training set: Consusus Atoms



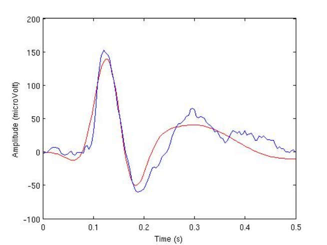
$\vec{A}_1$



$\vec{A}_2$

Real dataset

A signal reconstruction:



Mean GOF for the reconstructed signals:

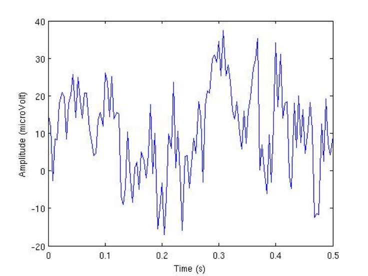
	GOF
mean	0.8431
std	0.0820

Real dataset

Testing set

A first testing set T_1 was created selecting different signals from a single channel:

- 30 spikes (w)
- 30 non-spike (b).



Real dataset

Testing set: features

GOFs mean and standard deviation were calculated for w and b

	GOF_w	GOF_b
mean	0.8536	0.3884
std	0.0404	0.2157

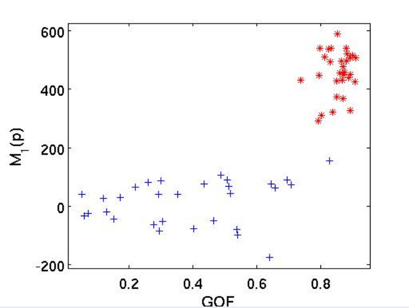
In the same way mean and standard deviation were calculated on the first atoms amplitudes (spike) $M(\hat{p}_1^w)$ and $M(\hat{p}_1^b)$.

	$M(\hat{p}_1^w)$	$M(\hat{p}_1^b)$
mean	456.1289	68.2913
std	76.3499	35.5233

Real dataset

Testing set: classification

Using the *GOF* and $M(\hat{p}_1)$ it was possible to generate a first dataset for classification.



As it is possible to see the two classes are perfectly separable.

Real dataset

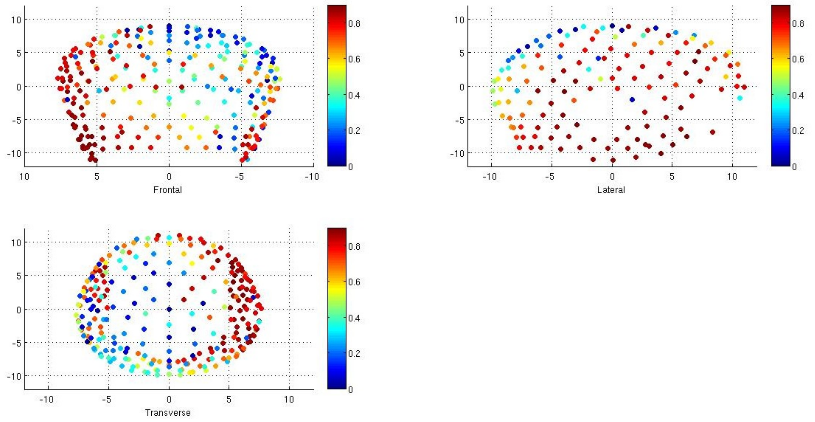
All channels classification

The validation of the method required to test the performance on all 256 channels, at least over a given time window containing a spike in the temporal region.

The pipeline of the algorithm was applied to this testing set in order to extract the features needed for the classification. The same consensus atoms found with τ and previously used in T_1 were used. *GOF*s were calculated and projected on the electrodes map.

Real dataset

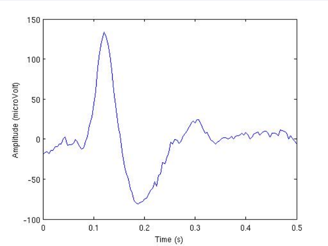
All channels classification: GOF



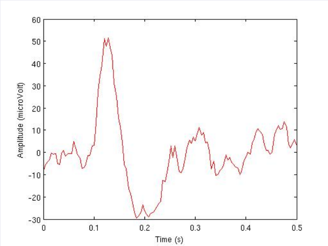
Real dataset

All channels classification

Goodness of fit has its highest values (red) in the right temporal lobe, but it largely spreads in the nearby regions. An interesting fact is that the contralateral lobe presented values of GOF particularly high.



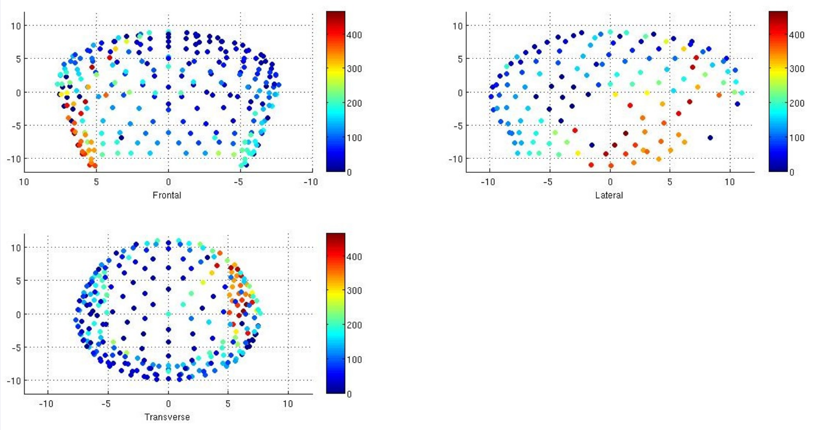
(a)



(b)

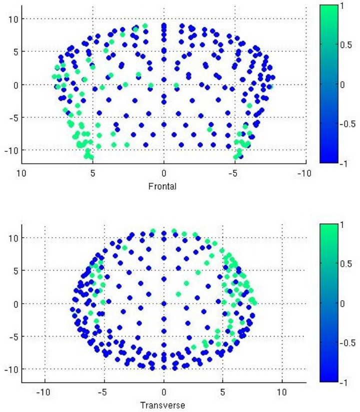
Real dataset

All channels classification: $M(\hat{p}_1)$



Real dataset

All channels classification: detection



Real dataset

All channels classification

This validation was only qualitative. In order to obtain a quantitive validation, 60 out of the 256 channels were selected from all the skull and then the extracted waveforms were manually labelled by an expert neurologist as either *spike* or *non-spike*.

Results of the golden test on the classification.

	Spike	**Non-Spike**
True	15	0
False	2	43

Real dataset

All channels classification: performance indexes

Sensitivity that indicates the fraction of TP find among the total number of spikes in the testing set.

$$Sensitivity = \frac{TP}{TP + FN} = \frac{15}{15 + 2} = 0.88; \tag{15}$$

Specificity as the fraction of TN among the total number of non-spike.

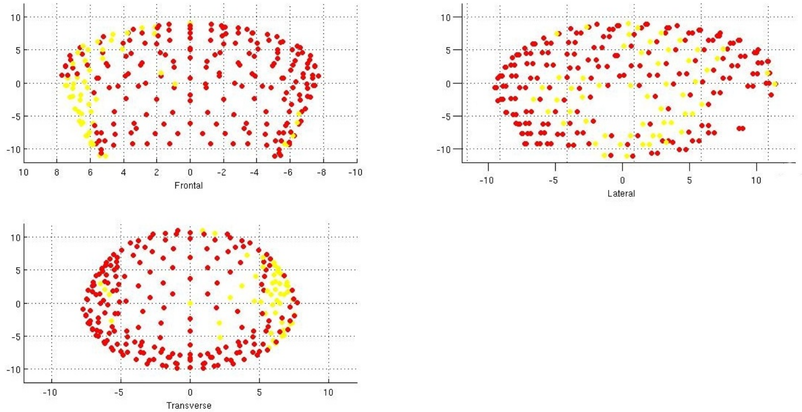
$$Specificity = \frac{TN}{TN + FP} = \frac{43}{43 + 0} = 1; \tag{16}$$

Selectivity is equal to the number of true spikes found by the classifier over the total number of signal classified as spike.

$$Selectivity = \frac{TP}{TP + FP} = \frac{15}{15 + 0} = 1; \tag{17}$$

Real dataset

Second patient



Real dataset

Results of the golden test on the classification.

	Spike	Non-Spike
True	11	4
False	14	31

	Sensitivity	Specificity	Selectivity
Value	0.44	0.69	0.73

Contents

1 Introduction

2 Methods

3 Results and Discussion

4 Conclusions

Conclusions

Conclusions

This work proposes an effective method for the detection and modeling of interictal prototypical signal patterns in temporal lobe epilepsy.

- Phase is added to CMP version proposed by Bénar;
- PS-CMP is adapted to become a powerfull feature extraction method;
- The classification pipeline was sucessfully tested on simulated and real dataset of different patient;
- This innovative technique captured the nature of the spikes and allowed to retrieve their position on the tracings and their spread across the cortex.

Conclusions

Future Works

Future steps will include:

- Implementation of a more efficient local maximum detection method;
- Investigation of different similarity functions;
- More features will be tested;
- The creation of a multi-patient training set.

Thanks for your attention!