Short-term emotion assessment in a recall paradigm

G. Chanel[1], J. Kierkels[1], M. Soleymani[1], D. Grandjean[2], T. Pun[1]

[1] CVML – CUI – University of Geneva

[2] Swiss Center for Affective Sciences - University of Geneva

Outline

Introduction

Data acquisition protocol

Features extraction

- EEG
- Peripheral

Classification

- Classifiers
- Fusion and rejection

Results

Conclusion and future work

Introduction - Objectives

Emotion assessment from signals of:

- the peripheral nervous system (GSR, blood pressure, respiration);
- the central nervous system (EEG).

Why ?

Advantages of EEG:

- part of emotional processes are cognitive;
- fast response and temporal resolution;
- the dynamic of the process can be better studied.

Advantages of using both modalities:

- physiological signals cannot be easily faked;
- fusion of modalities should improve results.

Applications: behavior prediction, monitoring of critical states, games …

Introduction - Brief state of the art

Reference Year	# participants	Modalities	Stimuli & time aspects	# classes & results
Healey 2000	1	periph.	Self-induction 3 to 5 min	2 classes 84%
Lisetti 2004	29	periph.	Clips 1 to 4 min.	6 classes 84%
Wagner 2005	1	periph.	Songs 2 min.	2 classes ≈90%
Leon 2007	9	periph.		3 classes 71%
Choppin 2000	20	EEG	Pictures, sounds 6-10 sec.	3 classes ≈60%
Takahashi 2004	12	EEG + periph.	Clips ?	5 classes 42%

Data acquisition protocol

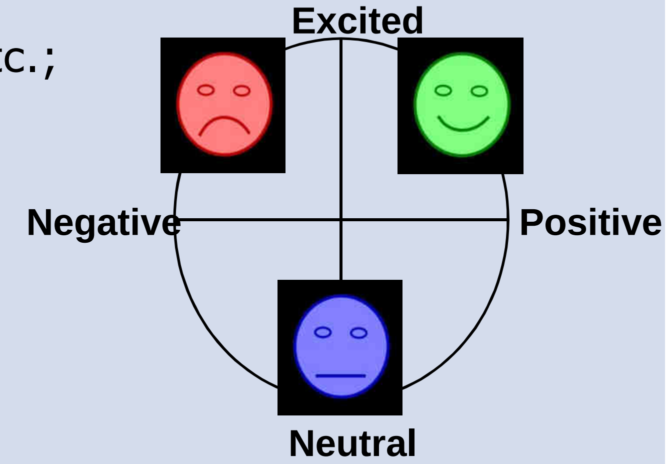
Model of emotions: valence-arousal space

Stimuli: imagination or recall of 3 different emotional events.

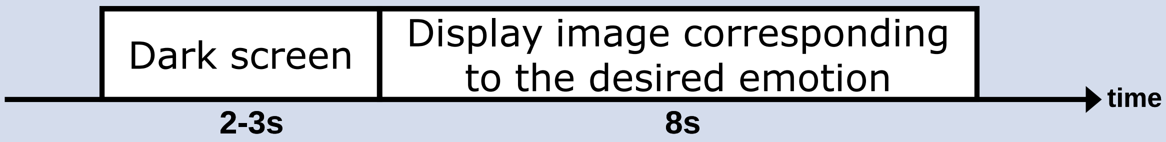
exciting positive: joy, hope, pride etc.;

exciting negative: disgust, anger, pain, hate, fear, etc.;

calm neutral: calm and neutral event.



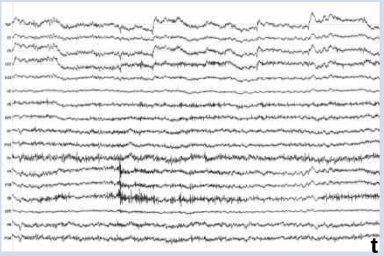
Protocol:



Trials are directly labelled into the 3 classes above.

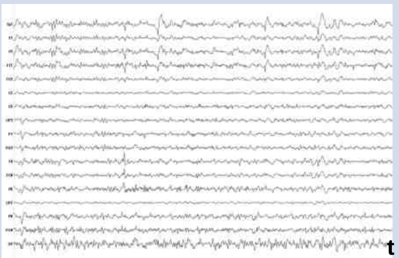
10 participant took part in the study.

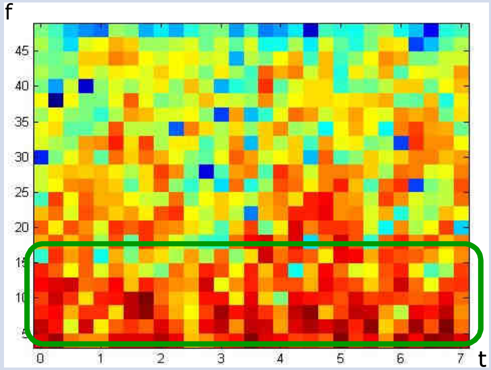
Features extraction - EEG signals



Fs = 1024 Hz
64 electrodes
8 seconds

[4-45Hz] filtering
Laplacian computation





29 time frames
257 frequency bands
(0Hz - 512Hz)
Δf = 2Hz

STFT for each signal
(electrode)
Window = 512 spls
Half overlap

9 Frequency
band selection:
4Hz – 20Hz

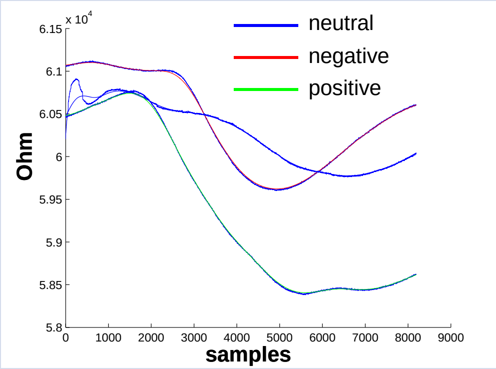
16704
STFT features

Mutual information computation
between each pair of electrodes

2016
MI features

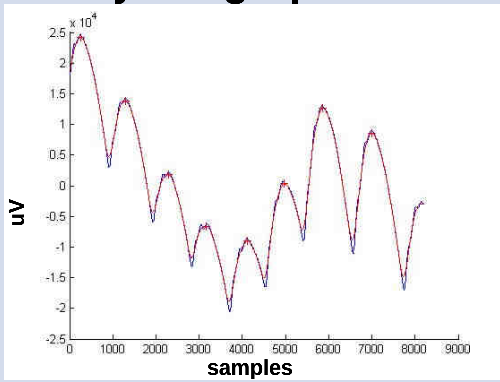
Features extraction - Peripheral signals

GSR



average decrease rates:
- during decay time
- whole trial

Plethysmograph



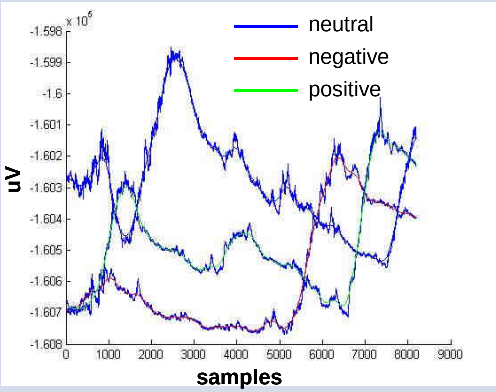
Mean

Mean of derivative

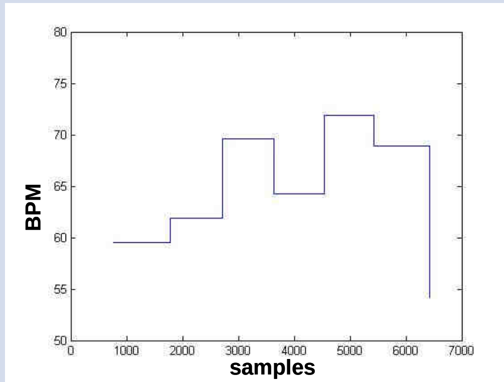
Standard deviation

Identification of maxima

Respiration



Heart rate



Fast fourier transform
10 freq. bands [0.25Hz – 2.5Hz]
$\Delta f = 0.25$Hz
max - min

Classification

Databases of 300 examples (100 ex. per class):

- 22 Peripheral features with associated labels;

$$\mathbf{x}_{Periph} = [x_{Plet}^{1} \cdots x_{Plet}^{4} \;\; x_{GSR}^{1} \cdots x_{GSR}^{4} \;\; x_{Resp}^{1} \cdots x_{Resp}^{14}]$$

- 16704 EEG STFT features with associated labels.
- 2016 EEG MI features with associated labels.

Classifiers:

- Linear and Quadratic Discriminant analysis (LDA / QDA);
- Support Vector Machine (SVM) with linear kernel;
- Linear Relvance Vector Machine (RVM).

Few examples per class ⇨ *Leave one out* cross-validation strategy:

- ✓ maximize the number of samples for learning;
- ✗ only mean of classifier accuracy is computable.

Classification

Data or feature level:

- Electrode combination using MI as a criteria for hierarchical clustering
  ⇨ slight loss in accuracy;
- Concatenation of the feature sets: $\mathbf{x}_{conc} = [\mathbf{x}_{EEG\_FFT}\mathbf{x}_{EEG\_MI}\mathbf{x}_{perpiph}]$
  ⇨ results are near to those of the bigger feature set.

Decision level / Opinion fusion with summation rule:

Each classifier c_i provide a confidence measure $p_{i,j}$ on each possible decision ω_j for the features vector $\mathbf{x}$, these measures are than combined to determine a score s_j :

$$s_j = \sum_{i=1}^{n} w_i p_{i,j} \quad with \quad w_i = \frac{1}{n}$$

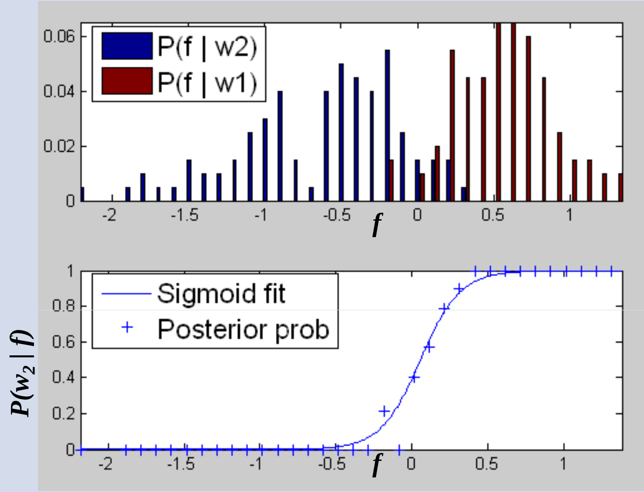
$p_{i,j}$ needs to be standardized, generally $p_{i,j} = P(\omega_j \mid \mathbf{x}, c_i)$.
The class ω_j with the maximum score s_j is then chosen.

Use of the posterior probalities of classes for fusion:

- LDA, QDA and RVM posterior probabilities are directly available;
- SVM produce only uncalibrated output f;

$$f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b$$

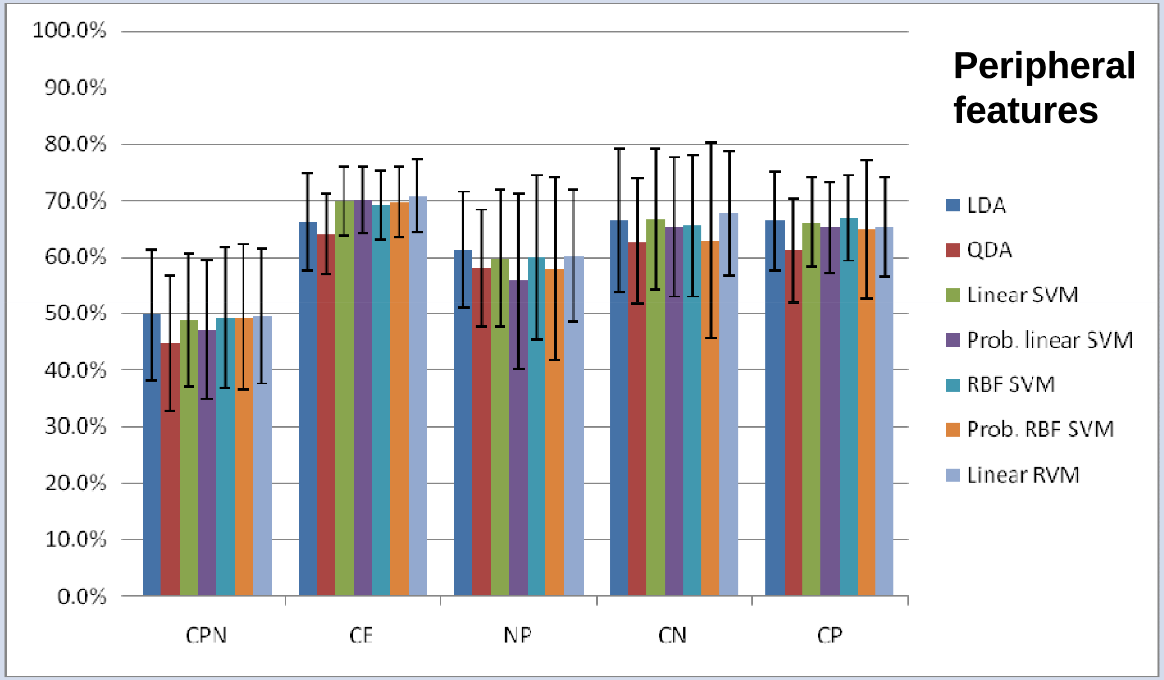
- posterior probability estimate based on distribution of f values [Platt 1999, Wu 2004].



Rejection of samples with low confidence value:

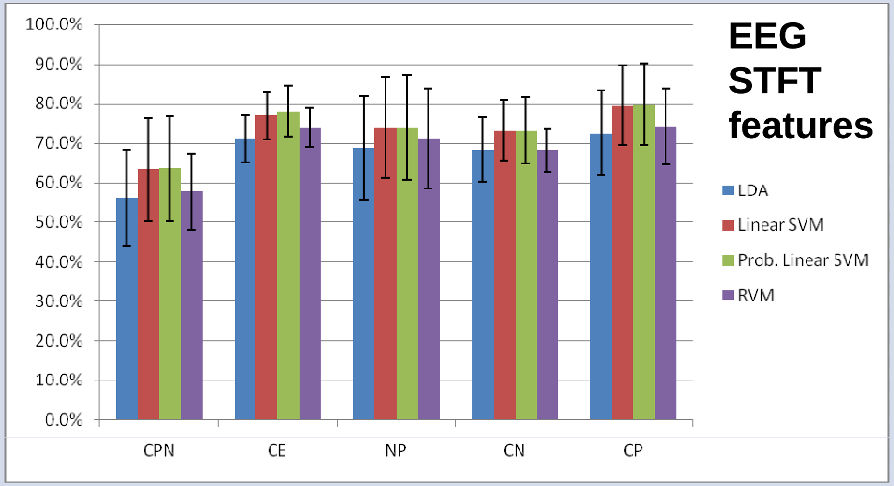
- reject samples where the fusion score s_j is inferior to a threshold δ;
- the new accuracy is then computed only on the non-rejected samples.

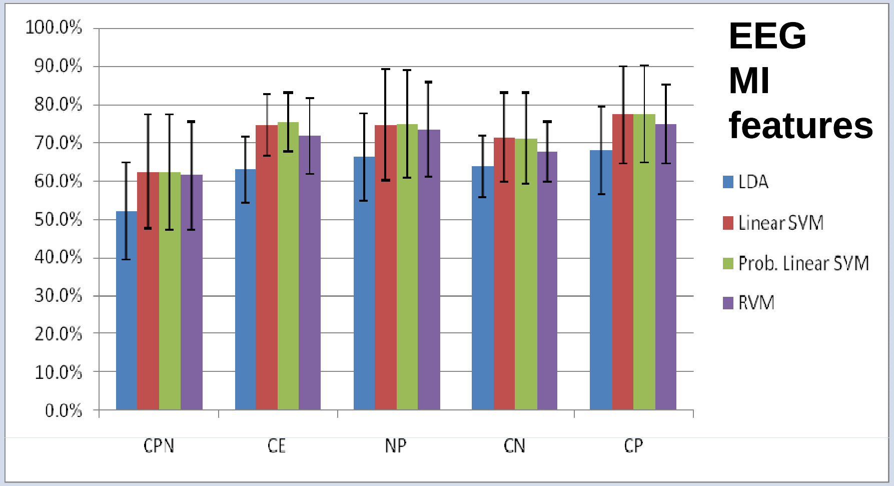
Results – Peripheral



- Average results are higher than the random level, but for NP classification task 2 participants are at the random level.
- Arousal classes are detected with higher accuracy than valence.
- LDA is choosen for fusion.

Results – EEG

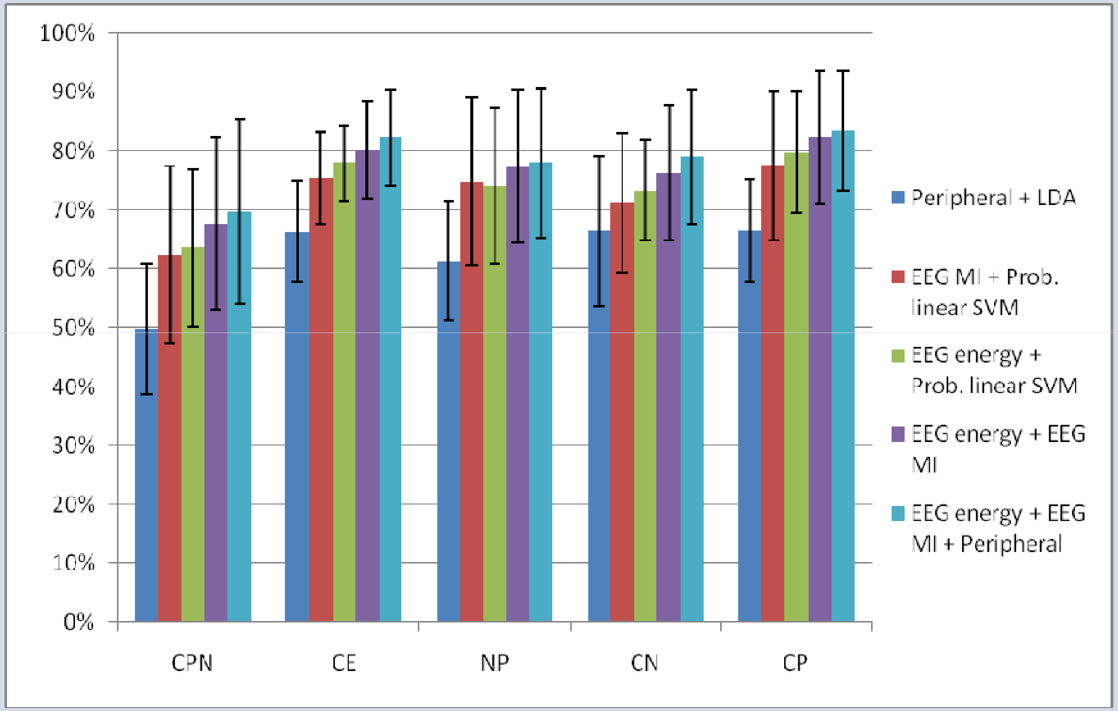




	Prob. Linear SVM	
	Best	**Worst**
Pos / Calm /Neg	**89%**	**41%**
Neg / Pos	**94%**	**54%**
Calm/ Excited	**93%**	**66%**

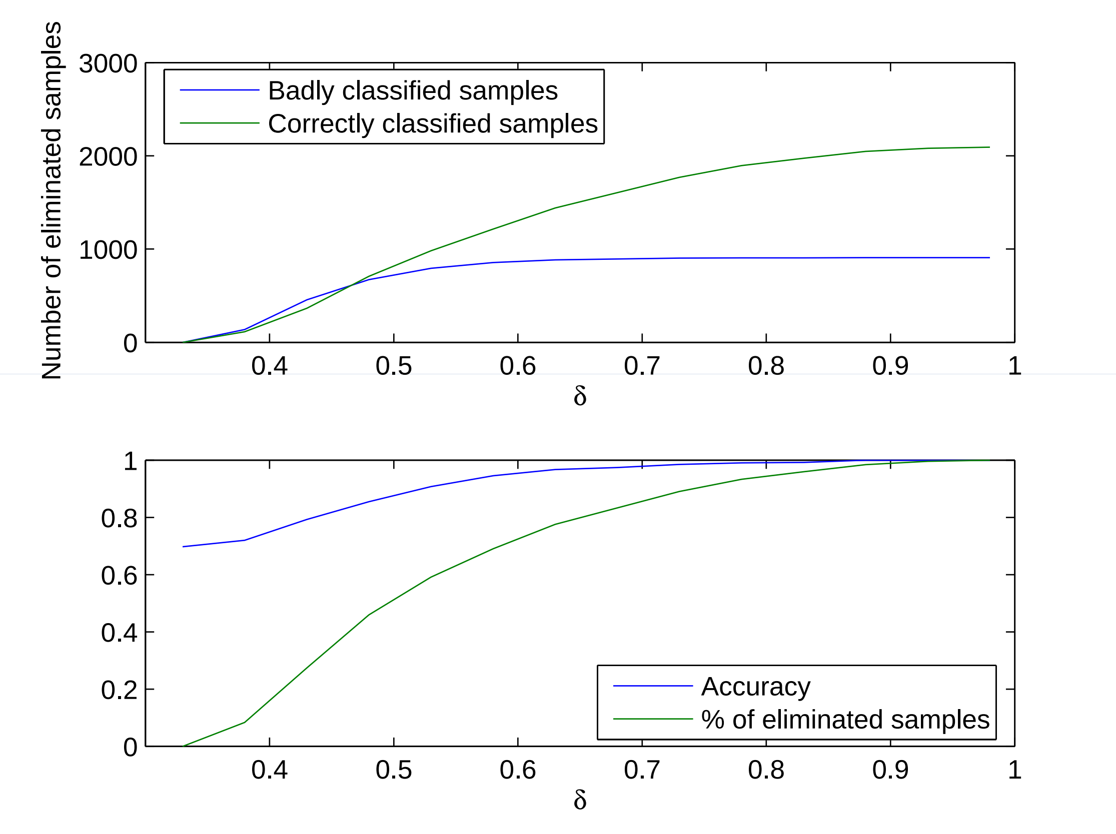
- Better accuracy with EEG than peripheral features.
- Valence classes and arousal classes are detected with similar accuracies.
- Probabilistic SVM is choosen for fusion because of its similar performance to SVM.

Results - Fusion of the 3 modalities



- Fusion of modalities always improves accuracy of the best single modality.
- Interest of fusion of different EEG features as well as fusion of central and peripheral phyisological signals
- Still high variability across participants.

Results - Rejection



- Choosen limit for δ : 0.47 because most of the badly classified samples are rejected at this point.
- At this point the accuracy is of nearly 80% but 40% of the samples are rejected.

Conclusion and future work

Conclusions:

- Results shows the usefulness of EEG signals, compared to peripheral signals, in emotion assessment and short time, highly cognitive conditions;
- fusion of the modalities improves mean results for all classes formulations;
- 10% accuracy improvement by rejecting 40% of samples.

Future works:

The question of time in physiological features:

- Performance analysis in different time resolution;
- Synchronization analysis of the different modalities.

Fusion and rejection:

- Different weights for the fusion;
- Strategy to find the best threshold for rejection.