UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES
Département d'informatique

Estimating emotions and tracking interest during movie watching, based on multimedia content and physiological responses

Joep Kierkels, Mohammad Soleymani, Guillaume Chanel, Eric Bruno, Stéphane Marchand-Maillet, Thierry Pun

CVML – CUI – University of Geneva

Outline

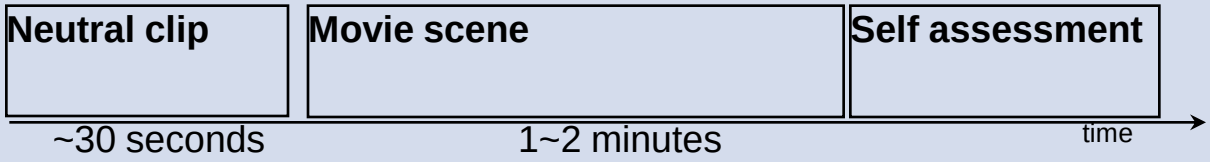
- Introduction
- Dataset
- Affective model
- Multimedia and physiological signals correlation
- Affective characterization
- Classification for interest detection
- Conclusion and future work

Introduction

- Affect and interest detection from movie scenes is useful for indexing, highlighting and user profiling.
- Find correlates features
- Assessment of user behaviour and emotions
  - User's emotion prediction by arousal/valence levels from audio-video content analysis
  - User's emotional characterization by arousal/valence levels from physiological signals
- Assess interest levels

Video dataset

- In the experimental data set we have 8 movies from 4 different genres:
  - Horror: The Ring, 28 days later
  - Action: Kill Bill VOL I, Saving private Ryan
  - Drama: Hotel Rwanda, The Pianist
  - Comedy: Mr. Bean's Holiday, Love actually
- 8 short video clips were extracted from these movies.
- Neutral clip between each two video clips to record baseline and let the participant to return to neutral state.
- Total duration ~2 hours.



Multimedia content features

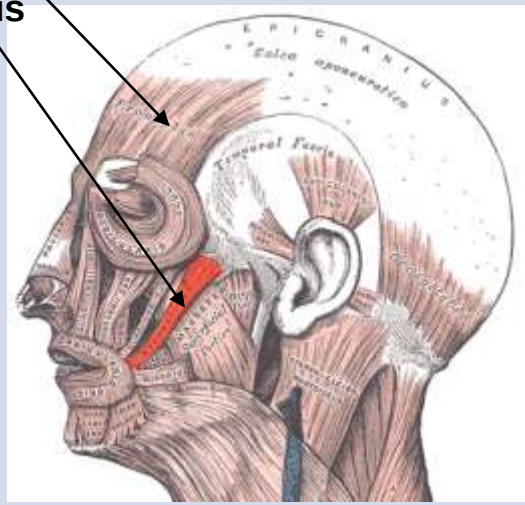
- All video content features are extracted with the help of OVAL and OMT software packages (*Viper*)
- Features extracted from multimedia content, namely:
  - Average shot duration
  - Shot change rate and variation
  - Color variance
  - Key lighting
  - Zero crossing rate (ZCR)
  - Audio energy
  - Audio type vector (music, speech, environment sound, and silence ratio over time)

Physiological feature extraction

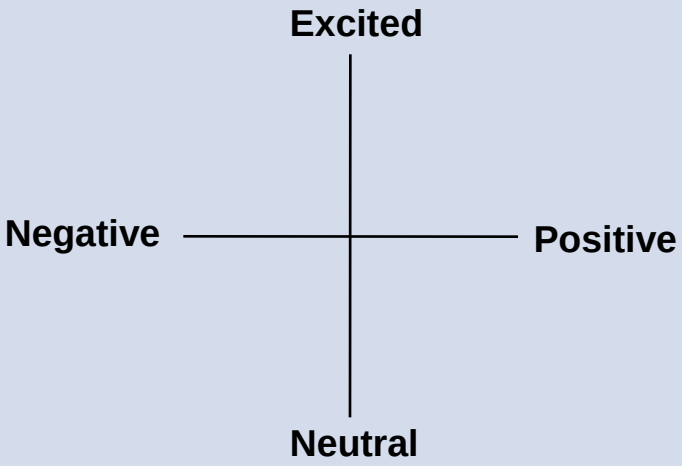
EMG	**Contraction Power**
ECG	**Heart Rate**
	Heart Rate Variability
	Inter-Beat- Interval
Blinks	**Blink Amplitude**
	Blink Duration
	Blink Frequency
Plethysmo-graph	**Heart Rate**
	Heart Rate Variability
	Inter-Beat- Interval
Respiration	**Respiration Depth**
	Respiration Rate
GSR	**GSR**
	Number of peaks
	Peak amplitude
	dGSR/dt
Temperature	**Temperature**
	dTemp/dt

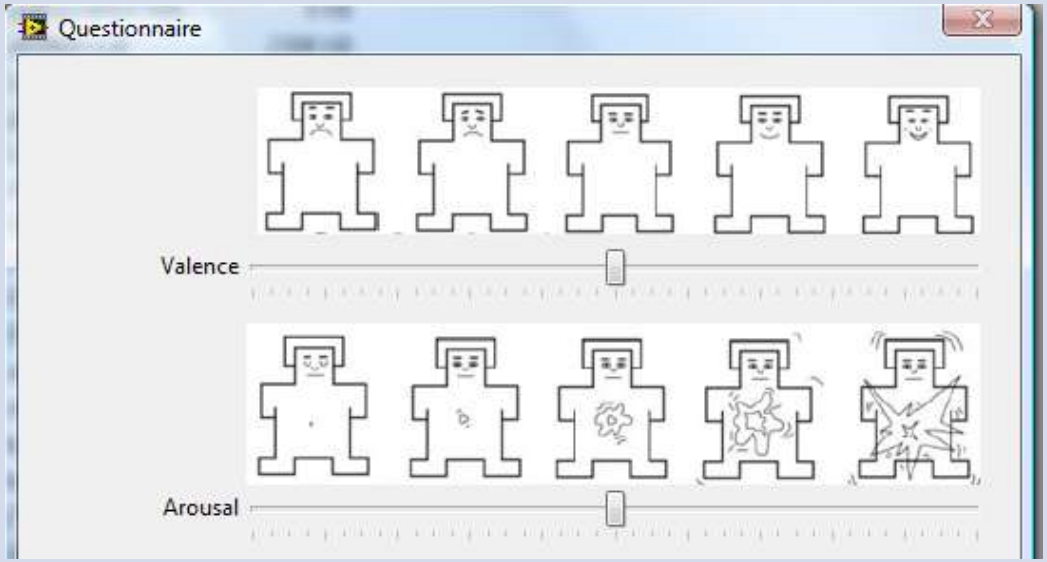
Frontalis

Zygomaticus



- Introducing the arousal valence space





Feature correlation

Correlation of selected multimedia features with physiological features of 8 participants				
	EMG Zygomatic. Energy/Key lighting	Skin temp.standard deviation /5th autocorrelation of MFCC coefficient	Skin temperature range/Shot length variation	EMG Zygomatic. energy/ 15th/20 bin of hue histogram
1	0.24	-	-	-0.41
2	0.62	0.44	0.42	-0.41
3	0.46	0.40	0.56	-0.34
4	0.40	0.32	0.43	-0.30
5	0.36	0.39	0.58	-
6	0.44	0.31	0.51	-0.32
7	0.47	0.34	0.27	-0.43
8	0.54	0.34	0.42	-0.45

Affect characterization

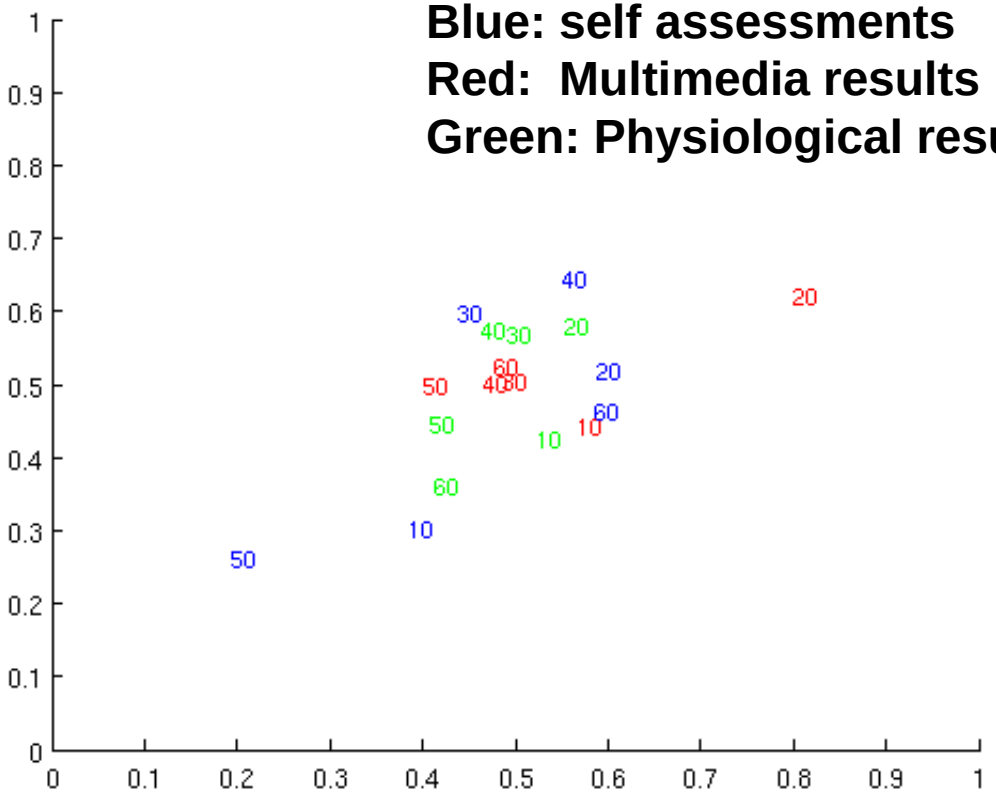
- Relevance vector machine is used to compute linear regression weights (w_i) to estimate arousal/valence scores. (Tipping toolbox)
- y_j is the estimation from arousal or valence value for j-th observation (video clip), x_i is the i-th feature

$$\hat{y}(j) = \sum_i w_i x_i + w_0$$

- Leave one out cross validation

Results of affect characterization

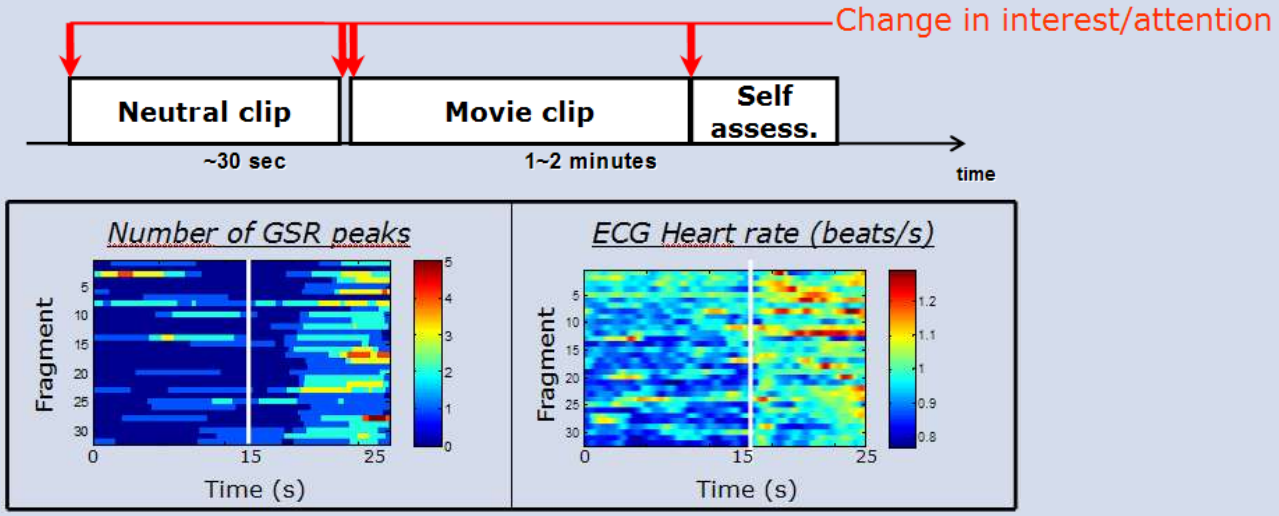
Average ts and
self-a or



Participar		ces with eatures
1		
2		
3		
4		
5		
6		
7		
8	0.17	0.19

Interest detection

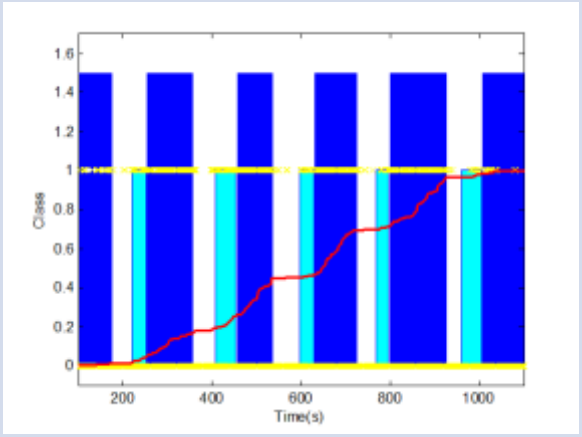
Need for training data which contains relevant features for at least two levels of interest.

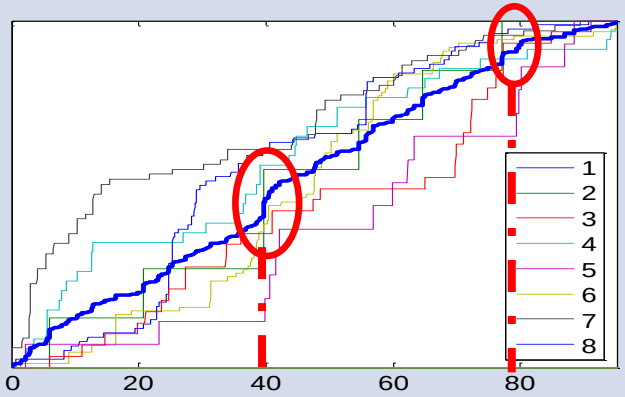


LDA Classifier output: 72% accuracy

Interest detection during movies

- Analyze full experiment.
- Compare participants.
- Evaluate.





00 :00

Conclusion

- The correlation between stimuli (multimedia) and response (physiological signals) was shown.
- Results show the ability of the physiological signals for affective characterization
- Temporal detection of interesting movie fragments is feasible.

Future work

- Implementation of head-movement and facial expressions as features.
- Obtaining objective measure for detecting interesting episodes instead of visual inspection.
- Using prior information for affect estimation (genre, users rating, etc)
- Modeling and estimating dynamics of affect in movie watching by multimedia features.