Emotion assessment from physiological signals for adaptation of games difficulty

Guillaume Chanel[1], Cyril Rebetez[2], Mireille Betrancourt[2], Thierry Pun[1]

[1] Multimodal Interaction Group,
Computer Vision and Multimedia Laboratory, CS Department
[2] TECFA, Faculty of Psychology
University of Geneva, Switzerland

Content

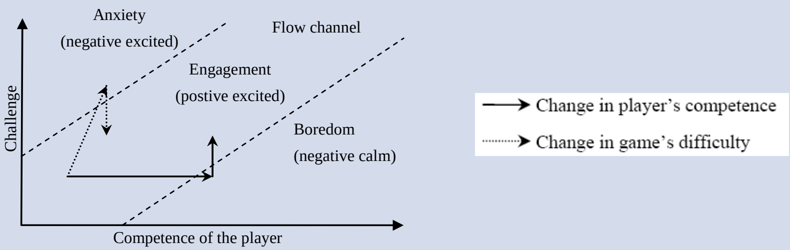
Objectives and hypotheses
Acquisition protocol
Feature extraction
- EEG's
- Peripheral signals

Analysis of questionnaires
Analysis of physiological features
Classification
- Features selection
- Peripheral signals, EEG's
- Fusion
- Effect of trial duration

Game over
Conclusions and future work

Objectives and hypotheses

“Flow of experience” theory [Csikszentmihályi]
- complete engagement in task, positive feelings, loss of sense of time;
- appears when the challenge of a task meets the “user’s” skill.



Objectives and hypotheses

Maintain the level of involvement and pleasure by:
- assessing the emotional state of the user through monitoring of physiological signals;
- controlling the difficulty of the task to influence challenge;
- here: task is a game, Tetris.

Why Tetris?
- known to elicit strong emotional responses;
- possibility to control the difficulty of the task (25 speed levels);
- well known so different gamer competences available;
- can be played with one hand.

Objectives and hypotheses

Hypotheses

- H1 : playing at different levels of difficulty induces one of 3 emotional states (boredom, engagement, anxiety);
- H2 : as the skill increases, the player will switch from the engagement state to the boredom state;
- H3 : these emotional states can be assessed using central and peripheral signaling.

Validation:

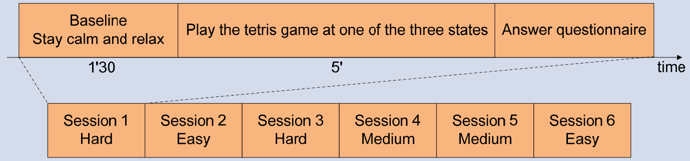
- from questionnaires and physiological data analysis;
- 20 participants (incl. 14 with EEG recordings).

Acquisition protocol

Determination of 3 gaming conditions (threshold method):

- medium (engagement) : levels 11 to 20;
- hard (anxiety) : medium level + 8, max 25;
- easy (boredom) : medium level – 8, min 5.

Schedule of the protocol :



Random permutation of difficulty for each participant

Acquisition protocol

Physiological signals from:

- peripheral nervous system: GSR, blood pressure, respiration, temperature;
- central nervous system: EEG, 19 electrodes (Bioesmi Active II).

Why ?

- physiological signals cannot be easily faked;
- part of emotional processes are cognitive;
- fusion of modalities improves results.

Feature extraction – EEG's

EEG features:

- energy in 3 bands alpha, beta, theta, related to emotional processes (e.g. alpha lateralization for approach-withdrawal);
- *EEG_W*, related to workload, engagement, attention, fatigue.

Feature for electrode i	Frequency band
θ_i	4-8 Hz
α_i	8-12 Hz
β_i	12-30 Hz

$$EEG\_W = \log\left(\frac{\sum_{i=1}^{N_e} \beta_i}{\sum_{i=1}^{N_e} \theta_i + \alpha_i}\right)$$

Feature extraction - Peripheral

GSR:
- mean value;
- mean of derivative;
- sum of negative derivatives;
- % of negative samples in derivative.

Respiration:
- main frequency;
- max – min (range);
- standard deviation.

Temperature:
- mean value;
- mean of derivative.

Blood pressure:
- mean value;
- standard deviation.

Heart rate:
- mean value;
- variance;
- mean of derivative.

Feature extraction - Peripheral

Peripheral signal	Feature name	Extracted feature	Comment
GSR	μ_{GSR}	Mean skin resistance	Estimate of general arousal level
	δ_{GSR}	Mean of derivative	Average GSR variation
	$f_{GSR}^{DecRate}$	Mean of derivative for negative values only	Average decrease rate during decay time
	$f_{GSR}^{DecTime}$	Proportion of negative samples in the derivative vs. all samples	Importance and duration of the resistance fall
	$f_{GSR}^{NbPeaks}$	Number of resistance falls in the signal	-
Blood pressure	μ_{BVP}	Mean value	Estimate of general pressure
	σ_{BVP}	Standard deviation	Blood pressure variation

Feature extraction - Peripheral

Heart rate	μ_{HR}	Mean of heart rate	-
	δ_{HR} σ_{HR}	Mean of heart rate derivative Standard deviation of heart rate	Estimations of heart rate variability
	f_{HR}^{LF}	Energy in 0.05Hz-0.15Hz band	Parasympathetic and sympathetic activity
	f_{HR}^{HF}	Energy in 0.15Hz-1Hz band	Parasympathetic activity
	$f_{HR}^{LF/HF}$	Ration of energy in the LF and HF bands	Ratio of parasympathetic and sympathetic activity
Respiration	f_{Resp}^{Rate}	Frequency with the highest energy	Respiration rate
	σ_{Resp}	Standard deviation	Variation of the respiration signal
	f_{Resp}^{DR}	Maximum value minus minimum value	Dynamic range or greatest breath
Skin Temperature	μ_{Temp}	Mean value	-
	δ_{Temp}	Mean of derivative	Estimation of temperature variability

Analysis of questionnaires

Description:
- 30 questions, with Likert scale ranging from 1 to 7;
- related to emotions : “I was stressed”, “I had pleasure”,…
- related to involvement : “I was focused on the game”, “I was motivated”, …

Factor analysis on the 30 dimensions to obtain axes with maximum variance:
- 56% of variance with first 2 components;
- 1st component >0 correlation with pleasure, interest, motivation, focus: valence?
- 2nd component >0 correlation with excitation, pressure, <0 correlation with calm, control: arousal?

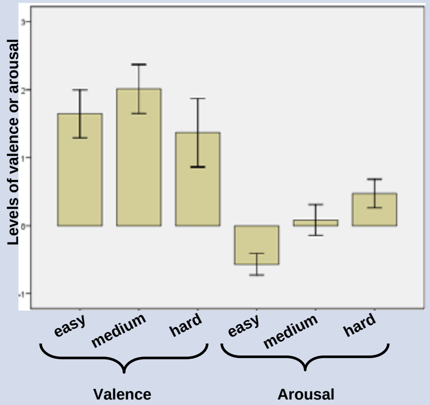
Analysis of questionnaires

Results:

- participants felt lower valence (pleasure, focus) in the easy and hard difficulties than in the medium one;
- the more difficult the game is, the higher is the arousal.

Three conditions do exist:

⇨ **H1 validated by self-assements**



Analysis of physiological features

Peripheral physiological features:

- ANOVA on physiological features to search for differences in activations for the different conditions and analyze the relevance of those features for emotion assessment;
- check for differences between the easy and medium conditions, and between the medium and hard condition.

Results:

- increase of arousal btw. medium and hard conditions but less than btw. the easy and medium conditions;
- increased arousal for increasing game difficulty;
- peripheral physiological data also **supports H1** to some extent.

Analysis of physiological features

Trend on the most relevant physiological features, easy-to-medium, and medium-to-hard conditions:

Feature	F-value	p-value	Trend of the mean
μ_{GSR}	4.4	0.01	↘→
δ_{GSR}	2.7	0.07	↘→
$f_{GSR}^{DecRate}$	3.1	0.05	↘→
$f_{GSR}^{DecTime}$	6.7	< 0.01	→↗
$f_{GSR}^{NbPeaks}$	18.3	< 0.01	↗→
μ_{HR}	3.4	0.04	→↗
f_{HR}^{LF}	2.4	0.09	↘↗
σ_{Resp}	5.8	< 0.01	→↗
μ_{Temp}	9.4	< 0.01	↘↘
δ_{Temp}	10	< 0.01	↘↘

- increase for easy-to-medium,
- stable for medium-to-hard

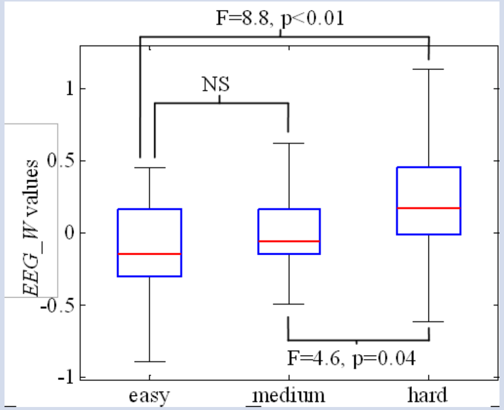
Analysis of physiological features

EEG features:

- alpha-band: no difference btw. the 3 conditions;
- beta and theta bands: several features show differences;
- EEG_W:
  - median increases as difficulty increases;
  - significant differences btw. easy and hard conditions, medium and hard conditions;
  - median not higher for medium condition: EEG_W more related to workload than engagement;
- EEG data also **supports H1** to some extent.

Analysis of physiological features

EEG_W values for the 3 gaming conditions:



Analysis of physiological features

(H2: as skill increases, switch from engagement to boredom.)

Test on data from 2 consecutive medium-condition games:

- questionnaires: significant decrease for questions “I had pleasure to play” and “I had to adapt to the interface”;
- peripheral signals: decrease in GSR peaks, increase in average temperature and average derivative of temperature.

Conclusions:

- decrease of arousal and increase in skills btw. successive games;
- **tends to validate H2**, but is the game boring or is the competence increased ?

Classification

Classifiers:

- LDA – Linear Discriminant Analysis;
- QDA – Quadratic Discriminant Analysis;
- SVM – Support Vector Machines with RBF kernel.

Ground truth: 3 difficulties corresponding to 3 emotional states.

Classifiers:

- "small" number N of users (20): classifiers trained on N-1, tested on 1 (cross-validation);
- yields 1 classifier per user;
- allows to estimate what would be the accuracy of a single classifier designed for large N, e.g. 200: training set 100, test set 100.

Classification

Features selection:

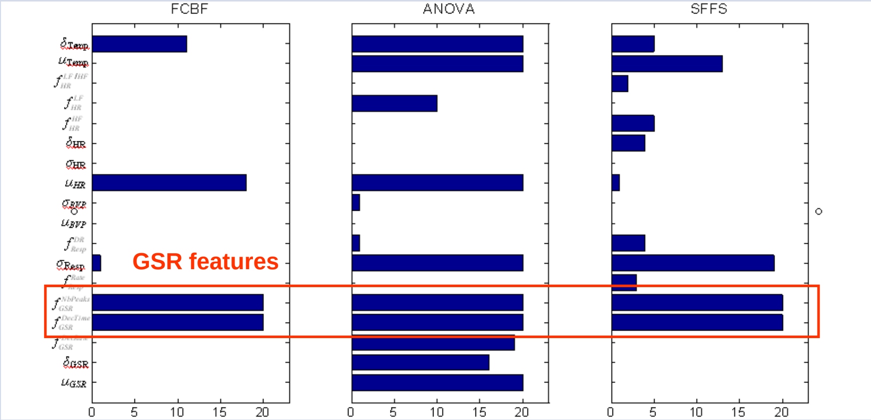
- FCBF - Fast Correlation Based Filter: removes features having low correlation with class concept;
- SFFS – Sequential Floating Forward Search: evaluates feature subsets.

Fusion of EEG and peripheral signals:

- at the decision level;
- Bayes belief integration: classifiers weighted by their average error.

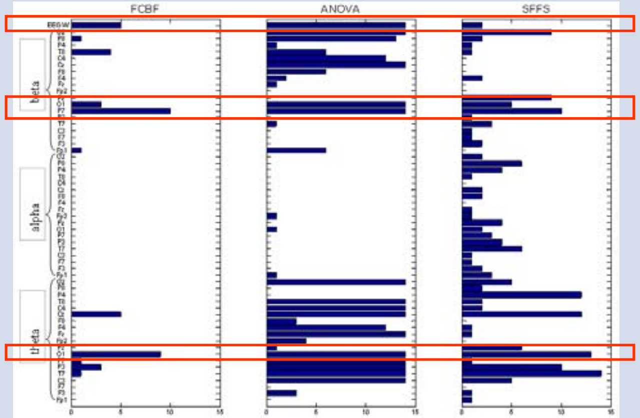
Classification – Features selection

Peripheral features selection (nr. of times a feature was selected):



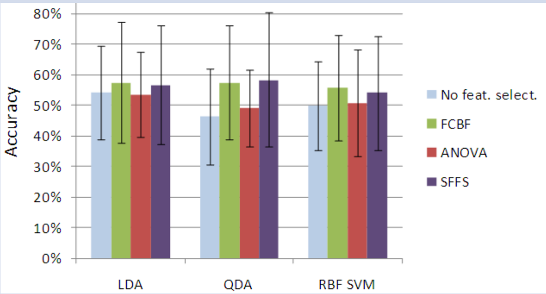
Classification – Features selection

EEG features selection (nr. of times a feature was selected):



Classification – Peripheral signals

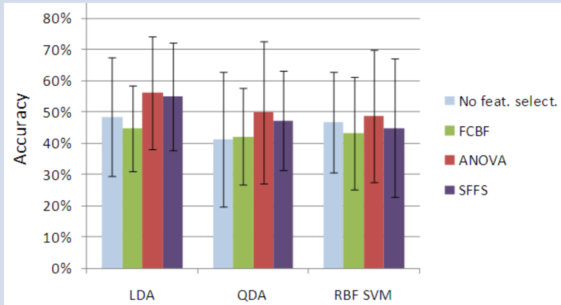
Peripheral features classification, confusion matrix for “FCBF + QDA”:



Estimated / True	Easy (Boredom)	Medium (Engagement)	Hard (Anxiety)
Easy (Boredom)	80%	10%	10%
Medium (Engag.)	37%	33%	30%
Hard (Anxiety)	21%	19%	60%

Classification – EEG's

EEG features classification, confusion matrix for “ANOVA + LDA”:



Estimated / True	Easy (Boredom)	Medium (Engagement)	Hard (Anxiety)
Easy (Boredom)	57%	43%	0%
Medium (Engag.)	21%	50%	29%
Hard (Anxiety)	19%	19%	62%

Classification – Fusion

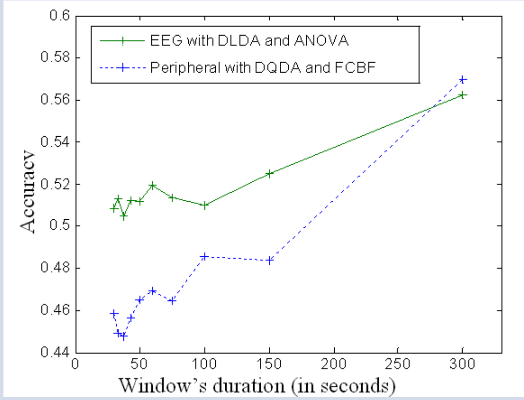
Fusion of EEG and peripheral features with Bayes Belief Integration:

- 5% increase of average accuracy;
- 2% (7%) increase for the easy (hard) conditions;
- 11% decrease for medium condition comp. with EEG's, but 6% increase comp. with peripherals features;
- only 4% error in classifying easy as hard, or hard as easy.

True \ Estimated	Easy (Boredom)	Medium (Engagement)	Hard (Anxiety)
Easy (Boredom)	82%	14%	4%
Medium (Engag.)	29%	39%	32%
Hard (Anxiety)	4%	27%	69%

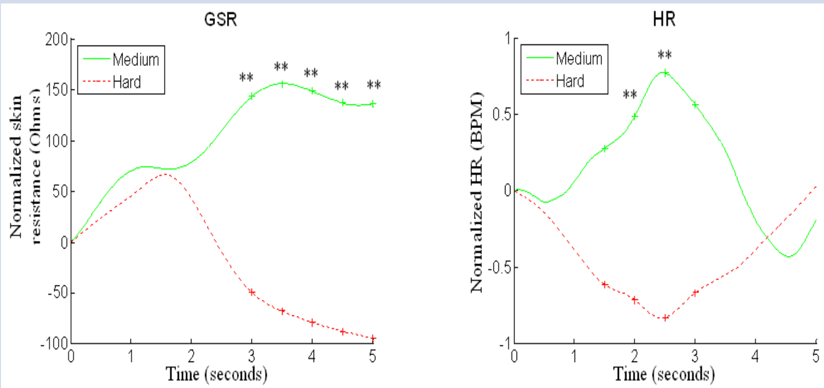
Classification – Effect of trial duration

Classification accuracy as a function of trial duration:
EEG features are more robust for short-term assessment.



Game over

GSR and HR in the 5s following game over:



Game over

Possible interpretations:

- GSR: low for high difficulty, which might indicate more arousal and stress to start a new game known to be difficult;
- HR: often higher for unpleasant stimuli, which might indicate here deception of losing a game at medium difficulty.

Thus different patterns of peripheral activity between sessions where users reported:

- higher motivation and pleasantness, or
- high pressure and less motivation.

Could be used to distinguish engaged from stressed states.

Conclusions and future work

Research hypotheses:

- H1 verified: playing Tetris at different levels of difficulty induces different emotional states;
- H2 somehow verified: is the change of emotional state due to boredom, or to an increase of competence?
- H3 somehow verified: accuracy >> chance levels, interest of fusion, of EEG's for short-term analysis.

Emotional states:

- easy condition: boredom (low pleasure, pressure, arousal, and motivation);
- medium condition : engagement;
- hard condition : anxiety (high arousal and pressure, low pleasure).

Conclusions and future work

Others:

- engagement can decrease if game difficulty does not change;
- analysis of game-over: distinct patterns of GSR and HR in 5sec following end-game (to distinguish engagement from stress?).

Future work:

- increase classification accuracy, e.g. feature selection;
- have a cross-participants model/classifier;
- fusion with other modalities, e.g. voice, body;
- use of context and temporal information, e.g. mood, previous emotional state, previous game events;
- develop affective Tetris (and other games).