NoE Peer-to-Peer tagged Media

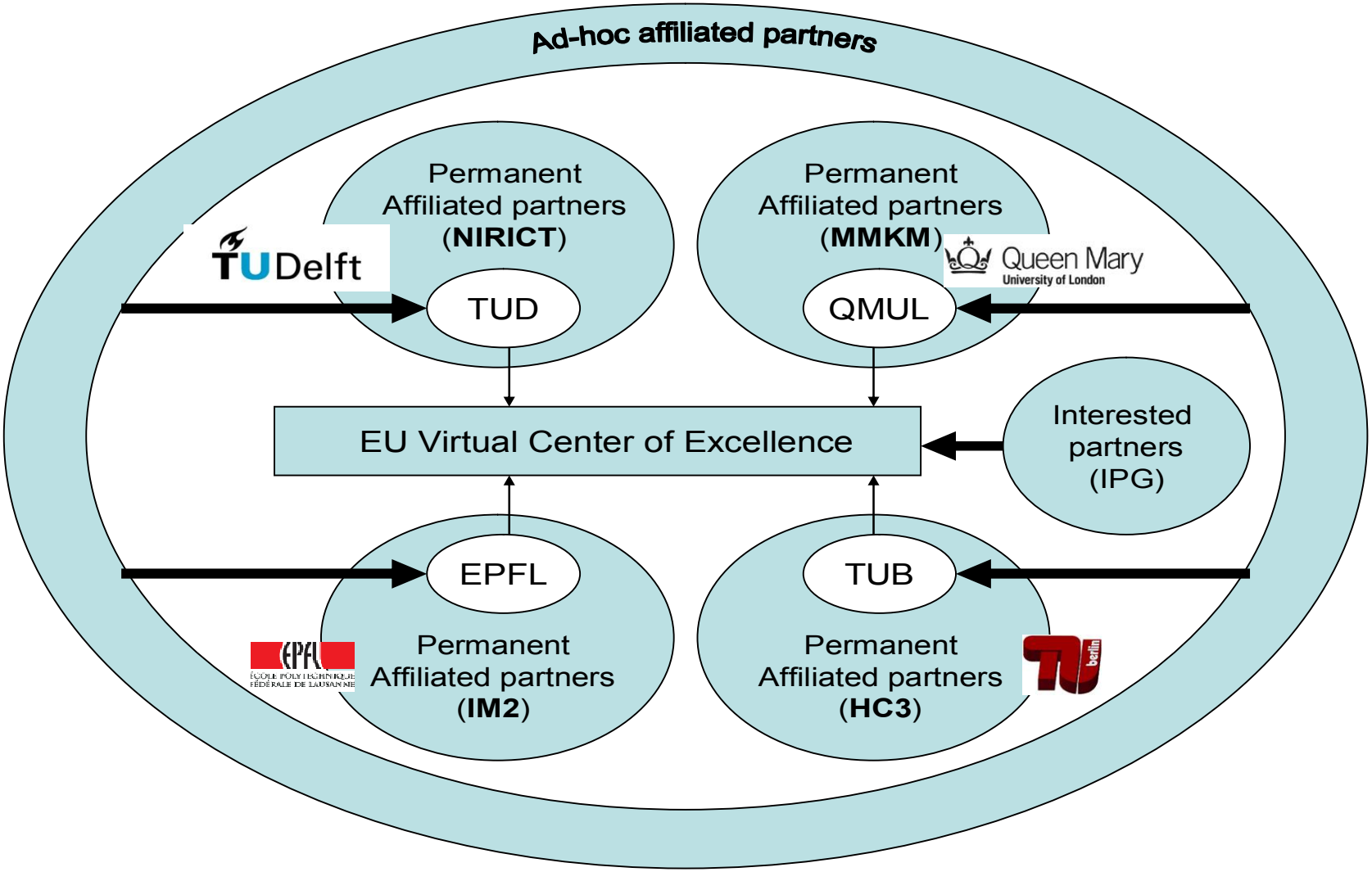
Human Centered Implicit tagging

CIVR, July 2010

Dr Ioannis (Yiannis) Patras

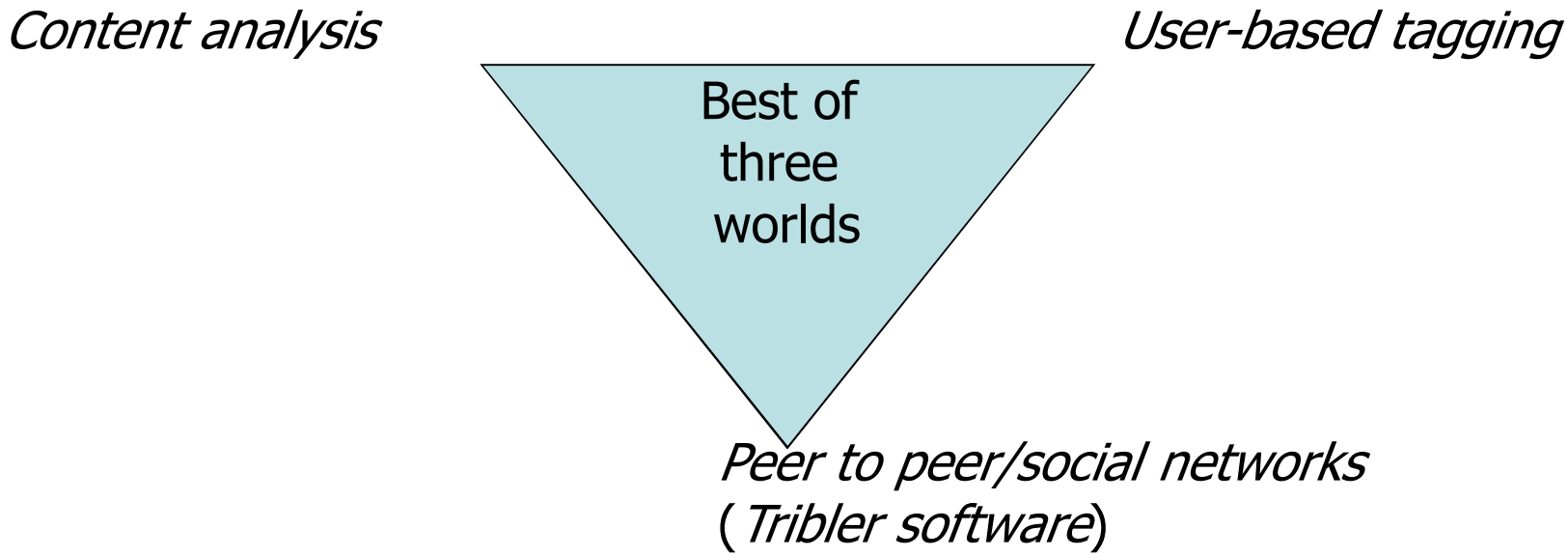
Queen Mary University of London
School of Electronic Engineering and Computer Science

Partner Concept of PetaMedia: EU Network of National Networks



Technical Scope of PetaMedia

- What: Developing new paradigms and technologies for efficient and effective access to multimedia content
  How: Multimedia content analysis meets P2P/social networks



Use Scenarios and Field Trials

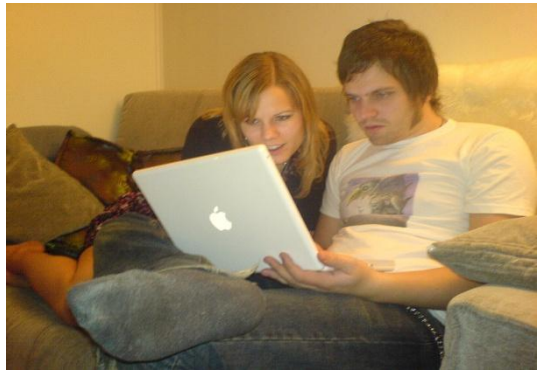
- IRPs and second phase IRPs feed into field trials. Field trials are based on the three use scenarios
- Field trials and use scenarios:
  - Means of communication to outside world
  - Testing of our technologies
  - People perspective on PetaMedia: "what the technology is good for"

User goal

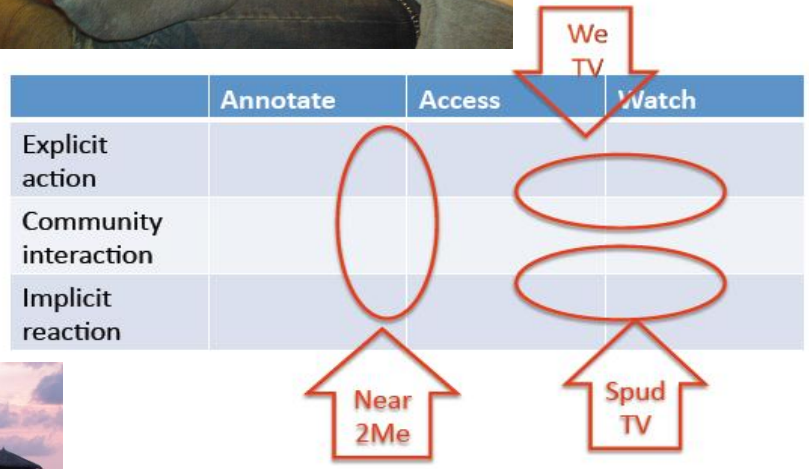
Social signals exploited

	Annotate	Access	Watch
Explicit action	Create tags using content analysis	Exploit user tags for access	Tag-driven scalability
Community interaction	Refine tags using social information	Improve access using social information	Content delivery in P2P
Implicit reaction	Create tags by observing user reactions	Leverage affect to improve access	Perceptual quality of content

Three Use Scenarios



WeTV: Compiling personal narratives about an event (e.g., concert) using user-contributed multimedia



SpudTV: Watching recommended content without supplying explicit preference information

Near2Me: Accessing social multimedia that brings users near to far-away places

Photos from Flickr users larskflem, novecentino, & stillmemory


MUSIC TELEVISION®
2.0

Project Teams

- EPFL: Touradj Ebrahimi, Ashkan Yazdani, Jong-Seok Lee
- QMUL: Ioannis (Yiannis) Patras, Sander Koelstra
- TU-Berlin: Kai Clüver, Engin Kurutepe
- University of Geneva: Thierry Pun, Mohammad Soleymani
- University of Twente: Anton Nijholt, Ronald Poppe, Christian Mühl

Couch Potato?

MTV 2.0

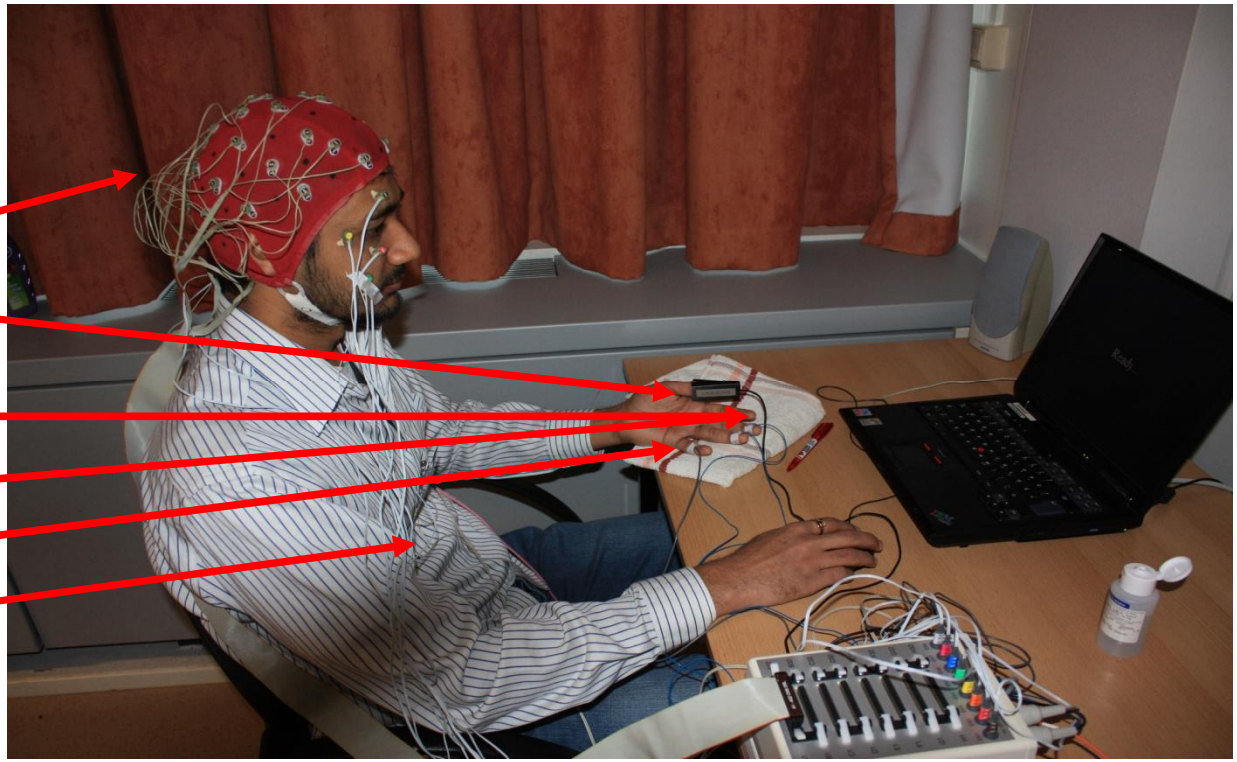
- **Playback controlled by:**
  - User profile
  - Song metadata
  - Physiological sensors

- **Analyze and cluster music and present videos by navigating music clusters according to user's satisfaction (User- Perceived Enthusiasm).**

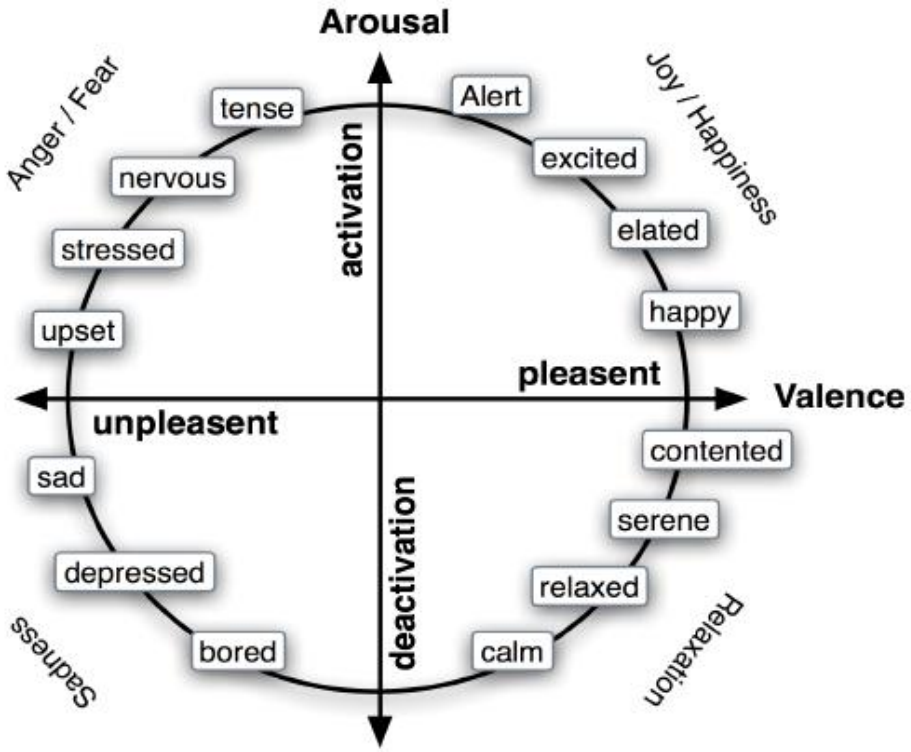
- **Personalizing recommendations based on affective reactions.**
- **Exploiting brain signals to create a user taste profile.**
- **How can self assessment from multiple participants with different backgrounds (profiles) be utilized for the evaluation of implicit tagging?**

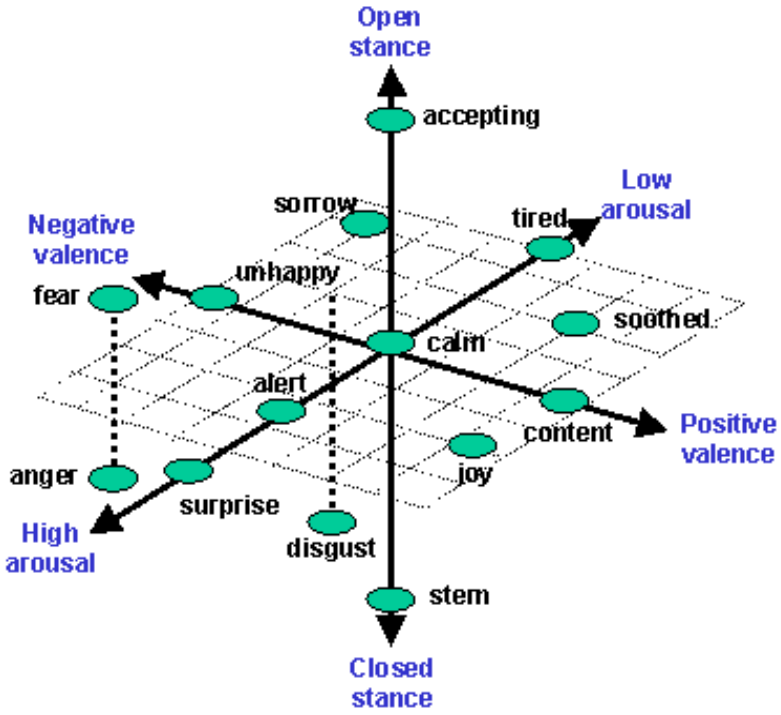
From Multimodal Human Sensing...

EEG Sensor cap
Plethysmograph (bloodflow)
Galvanic skin response
Heart rate
Temperature sensor
Respiration sensor



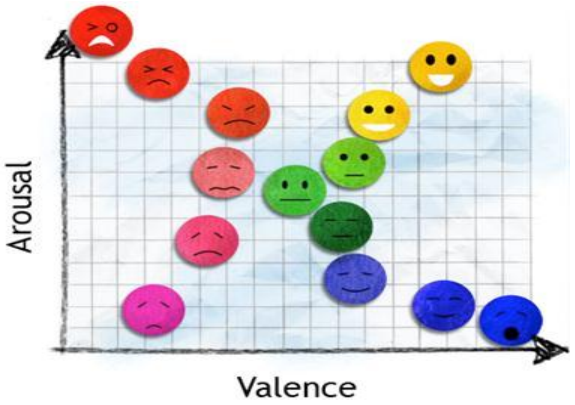
…to Emotion Models





Open Research Questions

- Prediction of valence/arousal from physiological.
- Satisfaction from valence/arousal data.



- Clustering music using (last.fm, Jango, etc.) metadata and social network.
- How to best navigate music clusters?
- How to connect sensor measurements with navigation?
- How to validate the final system?

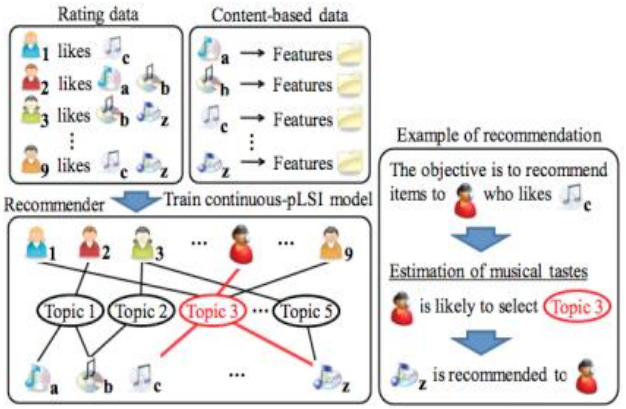
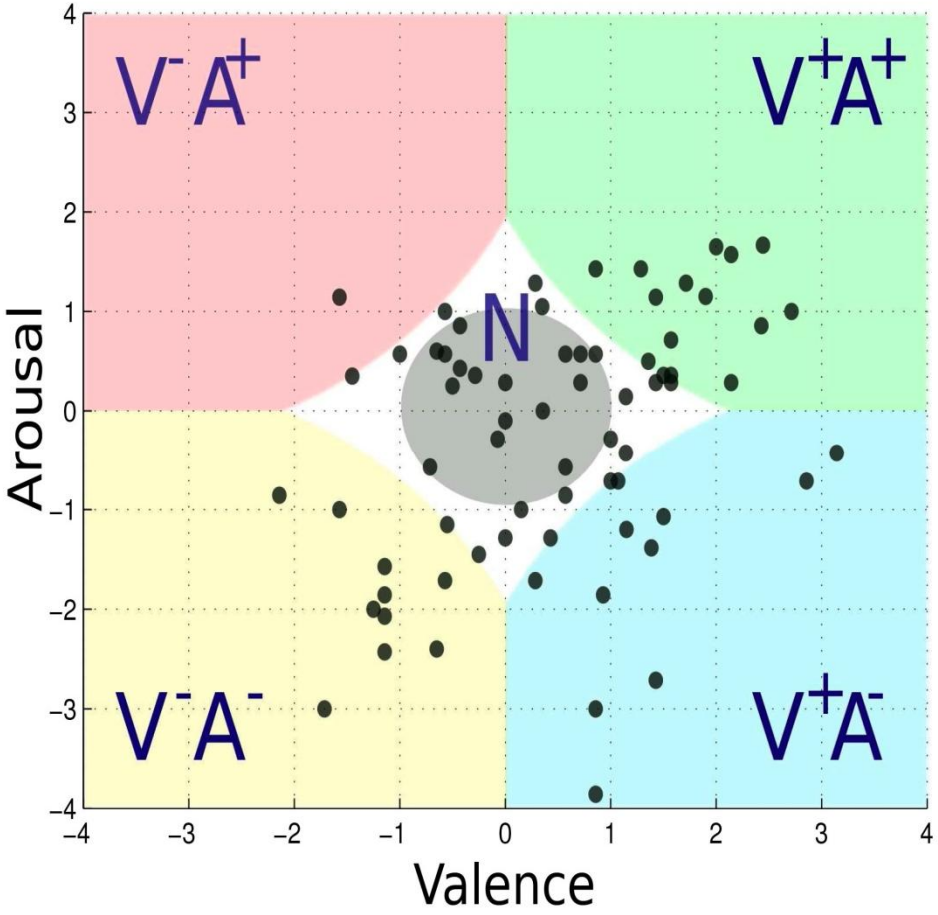


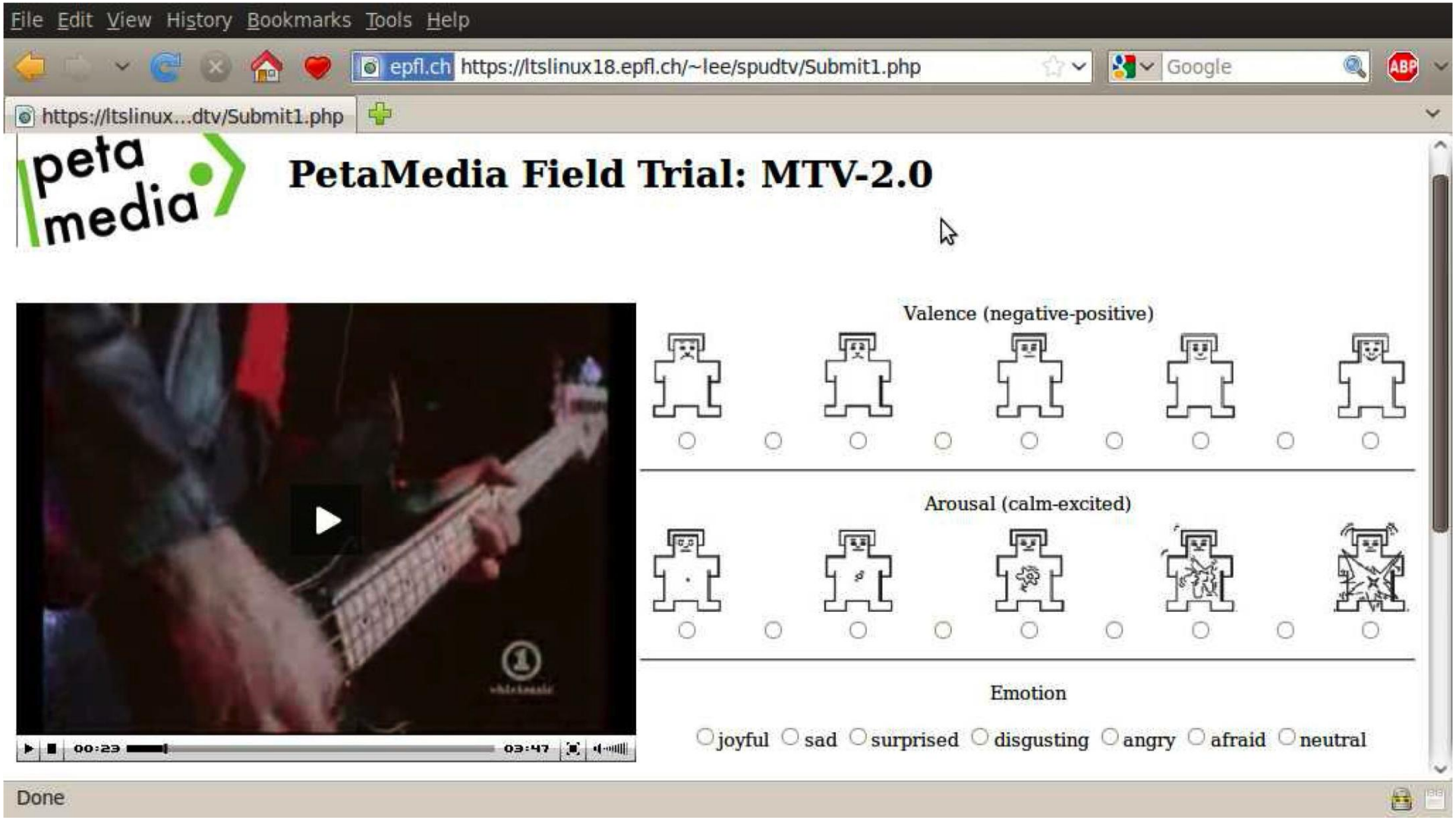
Figure 1. Hybrid recommender based on continuous pLSI.

Pilot Study Preparation

- 80 music videos were initially manually chosen by 8 different people
- Each video was rated for valence and arousal by 10-20 volunteers
- 4 videos selected from each quadrant in the arousal-valence space as well as 4 neutral videos
- 6 subjects viewed 2 minutes of each video, rated valence/arousal and gave the video a general rating on a 9-point scale

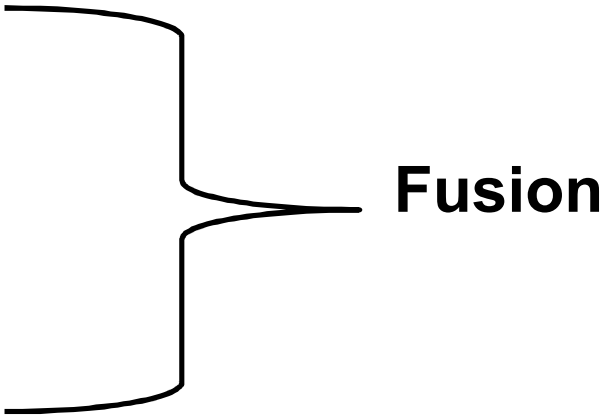


Pilot Study Preparation



Work Division

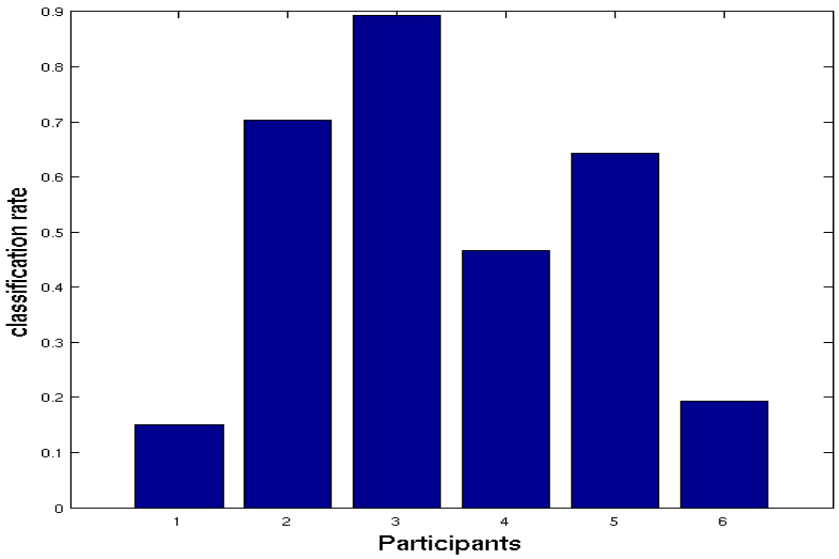
- TUB – Clustering and navigation.
- UT – EEG – neurophysiological data analysis.
- UniGe – Physiological signals
- EPFL & QMUL – EEG



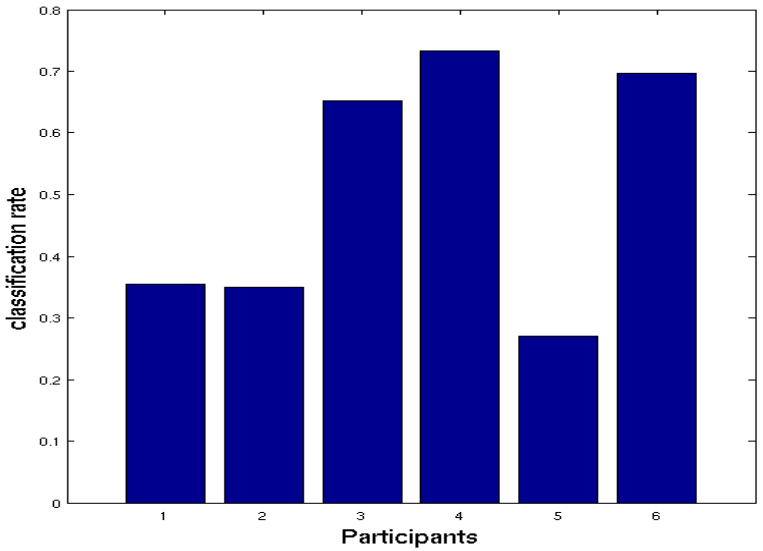
Peripheral Single Trial Classification

- Features extracted:
- EMG and EOG
  - Energy of the signal in the 20-400 Hz
- GSR
  - Statistical moments, Number of peaks, percentage of decreasing samples in the signal, Spectral power features
- Plethysmograph
  - Heart rate, HRV and its spectral features, blood volume pressure, Spectral power features
- Respiration
  - Statistical moments, Respiration rate, Spectral power features
- Temperature
  - Statistical moments, Spectral power features

Peripheral Single Trial Classification

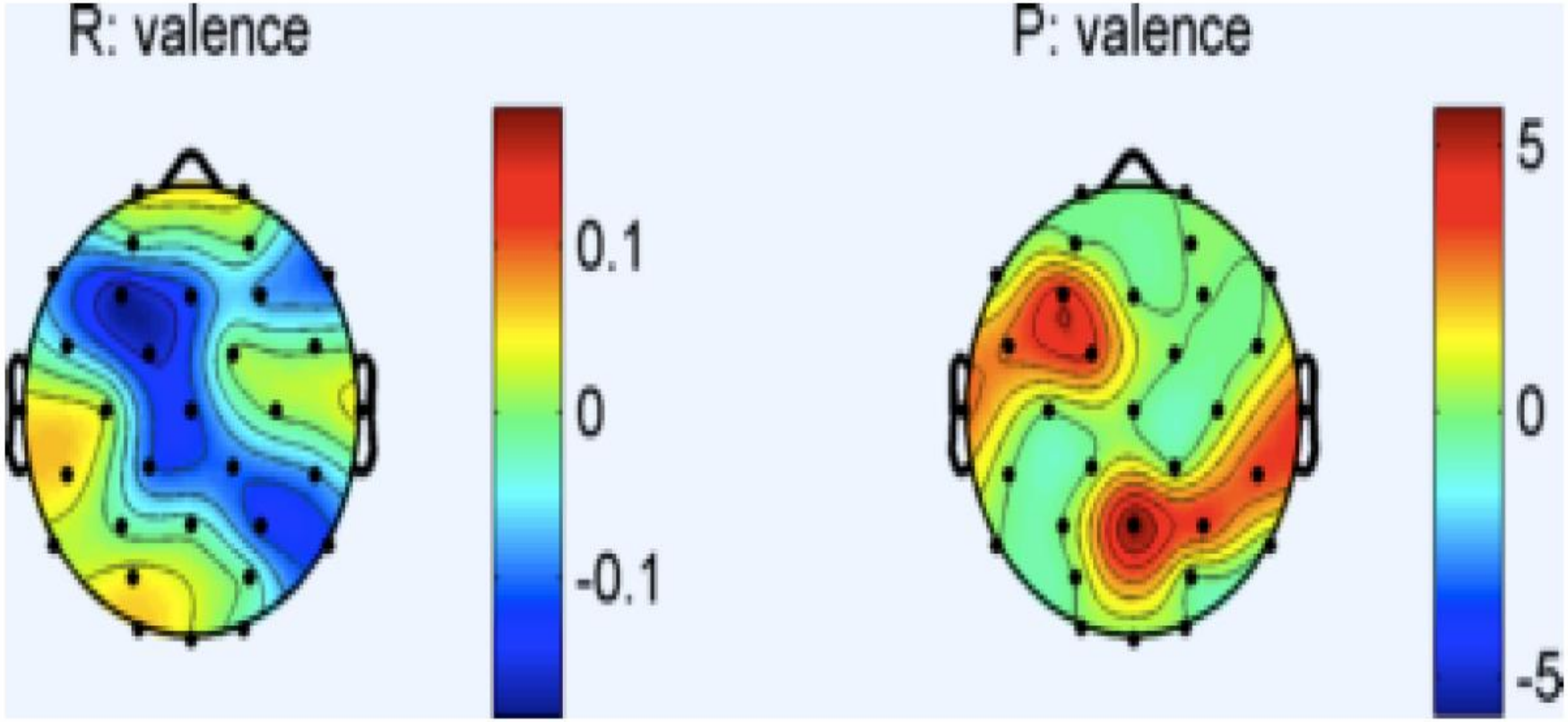


Valence (Pleasant/unpleasant) classification rate, peripheral, naïve Bayesian classifier



Arousal(Calm Excited) classification rate, peripheral, naïve Bayesian classifier

Correlations of alpha power and valence



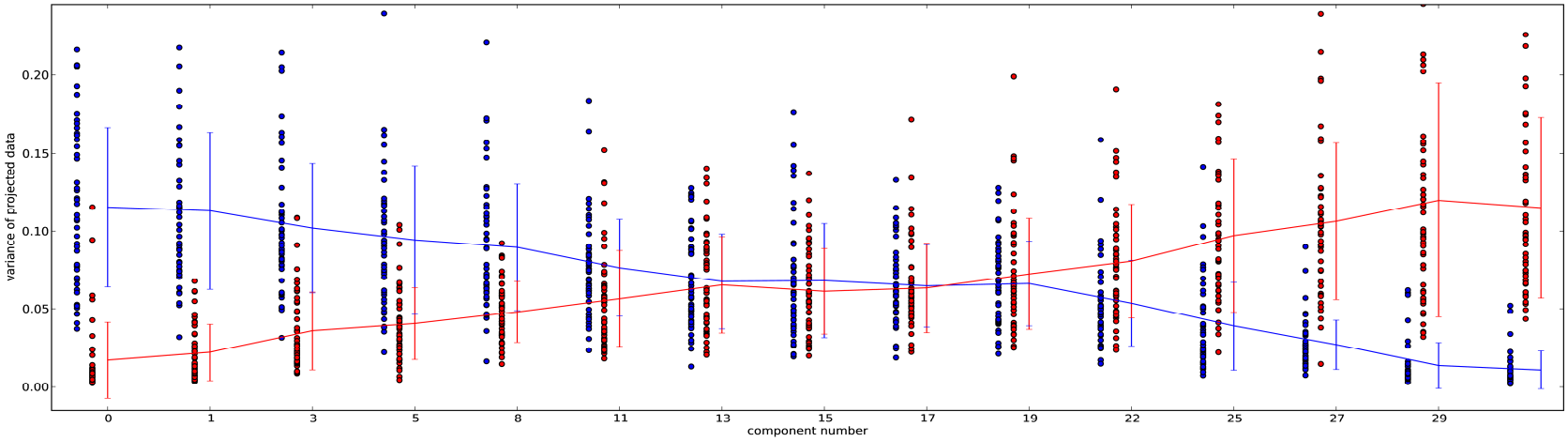
EEG Single Trial Classification (1)

- Goal: Predict valence, arousal and general rating for each video.
- Threshold subjects' arousal/valence/general ratings into two classes (i.e. positive or negative arousal)
- Extract features using common spatial patterns algorithm.
- Use linear SVM classifier for classification.
- Segment each video into 10 samples and test using leave-one-video-out cross-validation.

Common Spatial Patterns

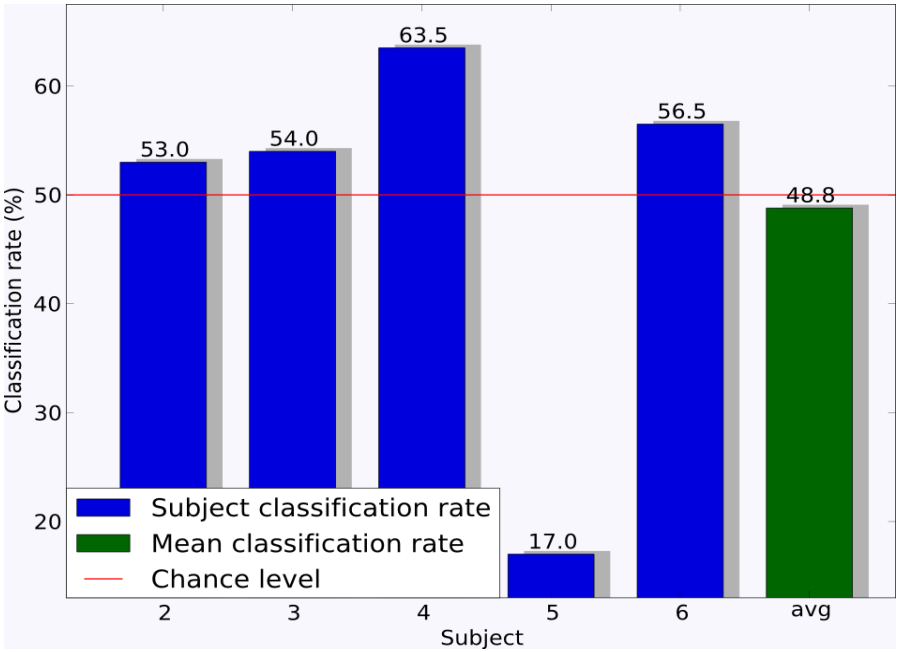
- **Method to decompose the signal into a number of components based on the variance of the signal that takes into account the class labels.**

- **Attempts to extract components for which the variance is maximal for one class and minimal for the other.**

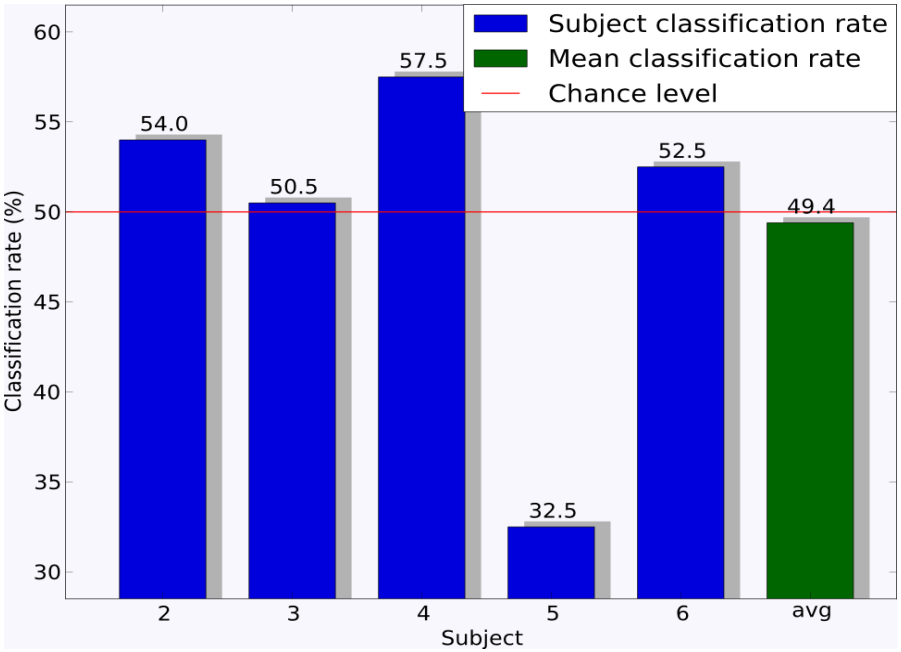


Results - General rating

- **Subject 1 excluded (19/20 rated high)**



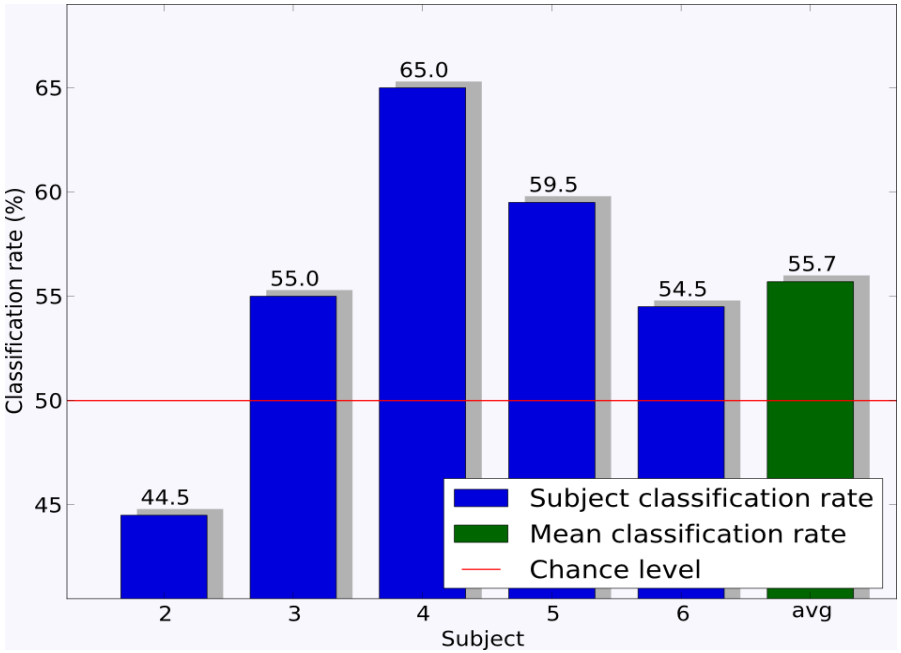
Common spatial patterns

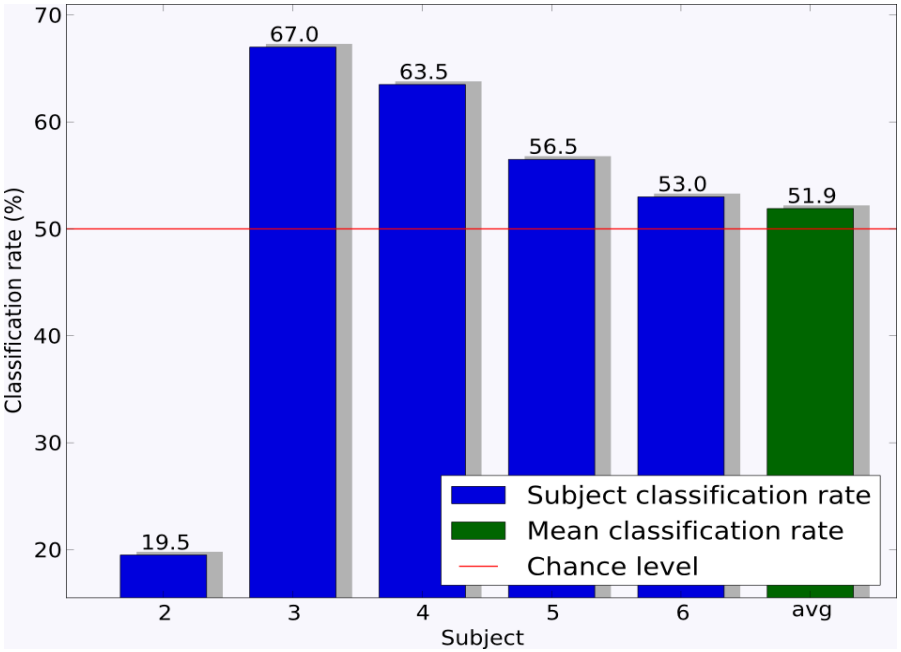


Power spectral density

Results - Arousal

- **Subject 1 excluded (17/20 rated high)**



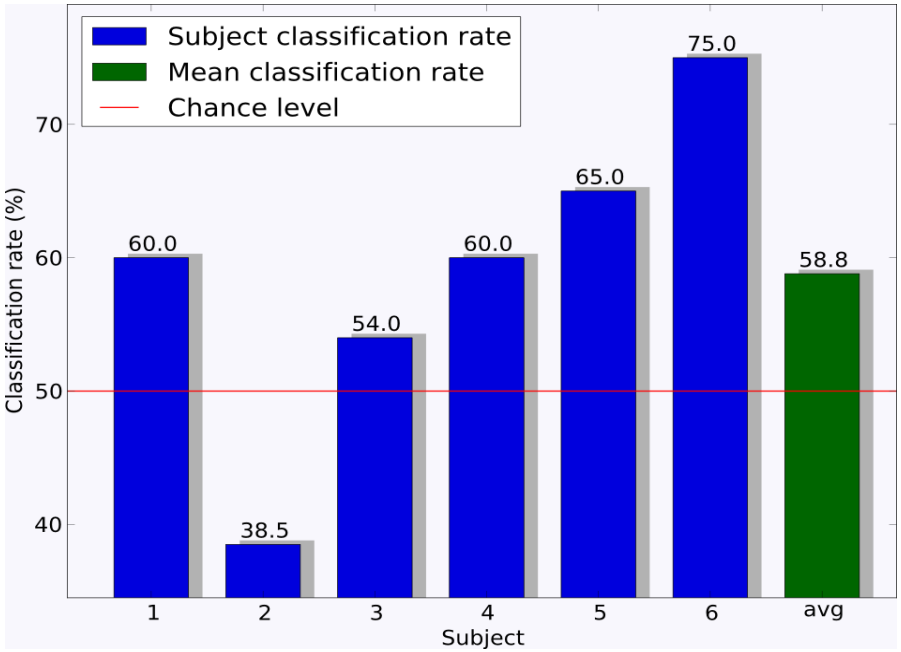


Common spatial patterns

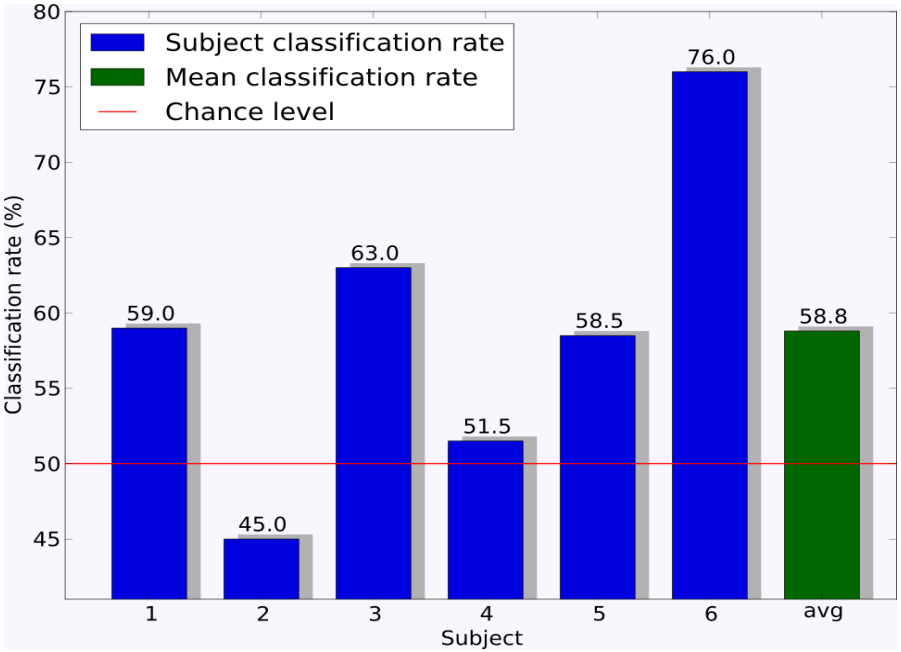
Power spectral density

Results - Valence

- **Subject 1 excluded (19/20 rated high)**



Common spatial patterns



Power spectral density

Conclusion

- **Early results show feasibility of arousal and valence prediction to some degree.**
- **But, current results not good enough yet for reliable single trial prediction.**
- **Future work:**
  - Regression instead of classification
  - Studying and implementing different windowing and feature extraction methods
  - Confidence measure of trial classification
  - Fusion with physiological signals (Geneva)