



SpudTV

Search Computing & Social Media Workshop
NEM Summit, Turino, September 28, 2011

Sander Koelstra
Queen Mary University of London

Introduction

Our work on affective state estimation

Developing a real-time affective music video
recommendation system

Who makes up SpudTV?



Sander Koelstra



Mohammad Soleymani
Guillaume Chanel



Ashkan Yazdani
Jong-Seok Lee
Eleni Kroupi

UNIVERSITY OF TWENTE.

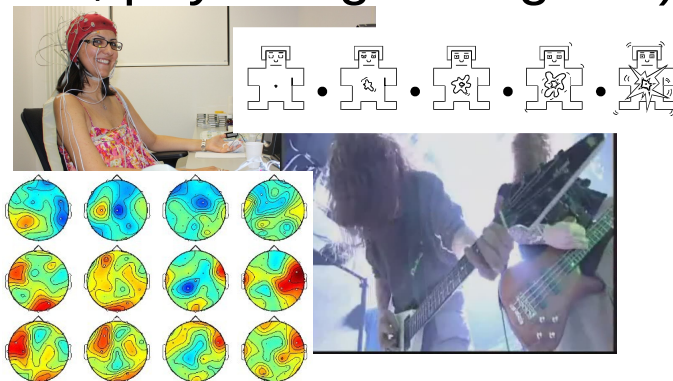
Christian Mühl



Engin Kurutepe

- A system that estimates the user's affective state and recommends music videos to match

Affect estimation
(Multimedia content analysis,
EEG, physiological signals)

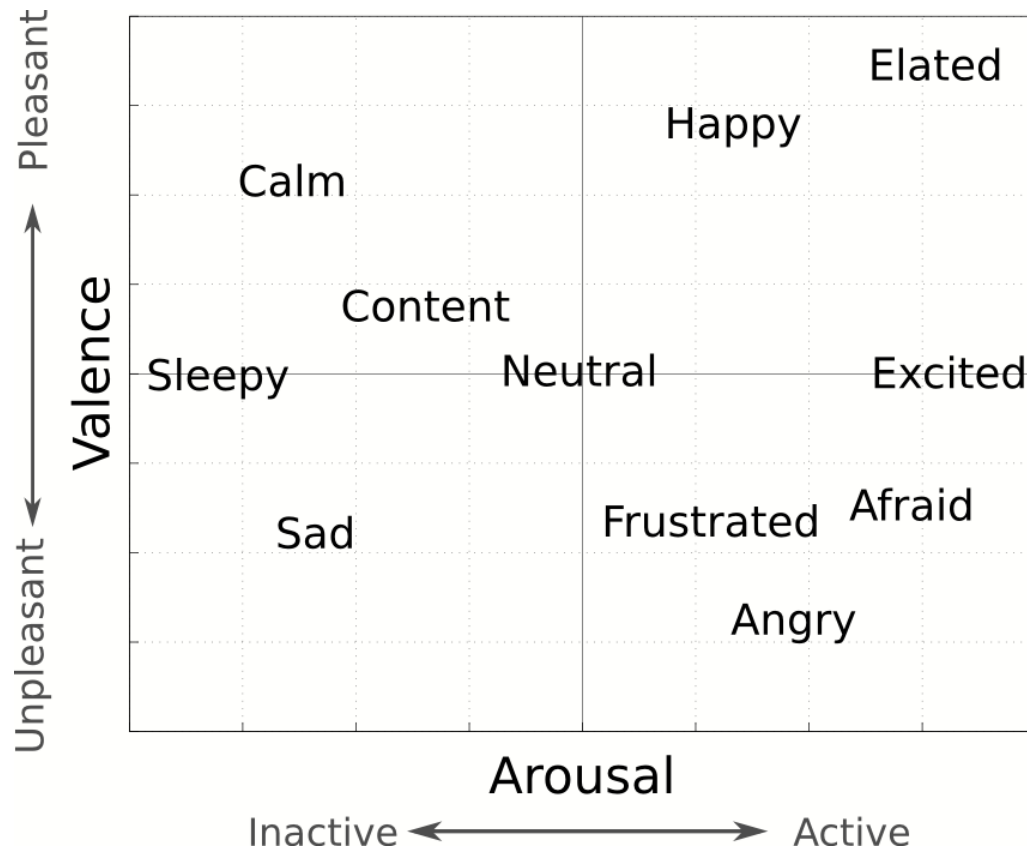


Content
recommendation

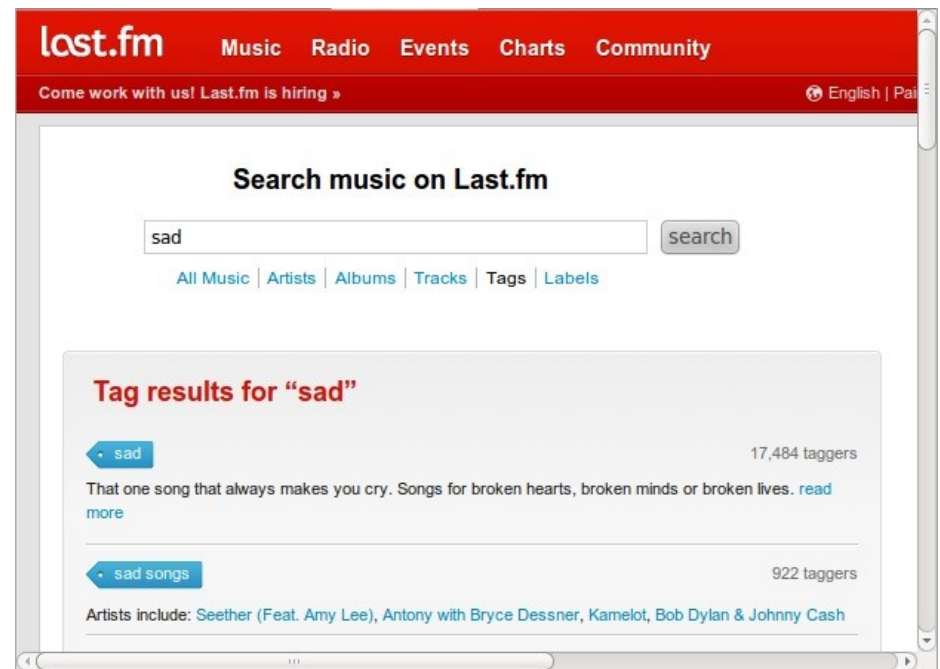


Affect representation

Russell's valence-arousal space:



- How to select music videos that elicit strong emotions?
- Use Last.fm to find videos with affective tags.



Stimuli selection (2)

gleeful, jolly, jolliness, jovial, joviality, love, lovely, loving, affect, affective, affection, adoration, adorable, adore, adoring, fondness, fondly, liking, likable, like, attraction, attractive, attract, caring, careful, tenderness, tender, tenderly, compassion, compassionate, compassionately, sentimentality, sentimental, sentimentally, sentiment, lust, arousal, arousing, desire, desirable, passion, infatuation, longing, joy, joyful, joyous, cheer, cheerful, cheerfulness, amusement, amusing, bliss, blissful, gay, gaiety, glee, gleeful, jolly, jolliness, jovial, joviality, delight, delightful, enjoy, enjoyment, enjoyable, glad, gladness,

List of emotions



corresponding
affective tags and
→ most tagged songs →

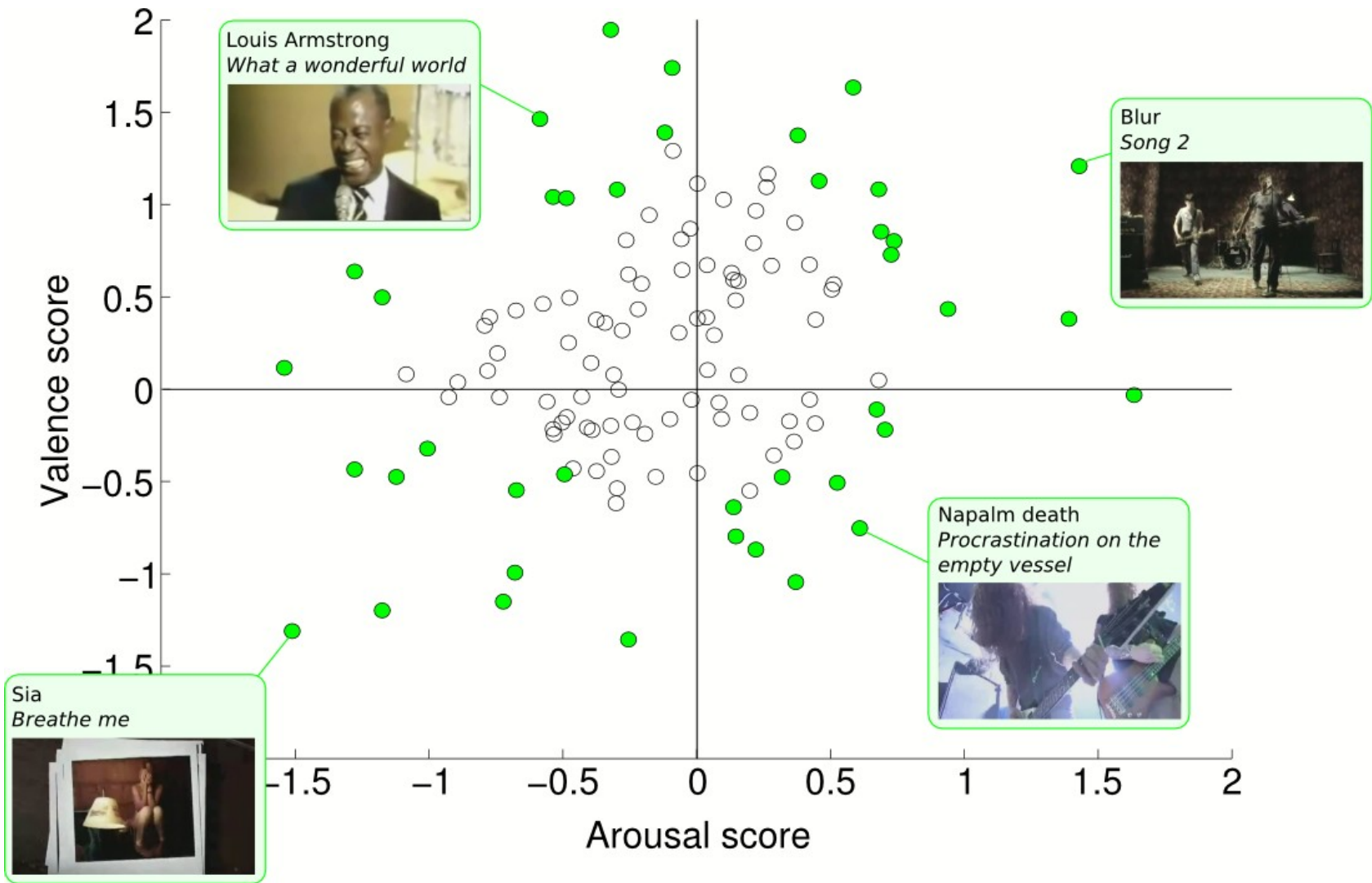
Manual filtering

- is the tag valid?
- is there a video?
- is the song appropriate?

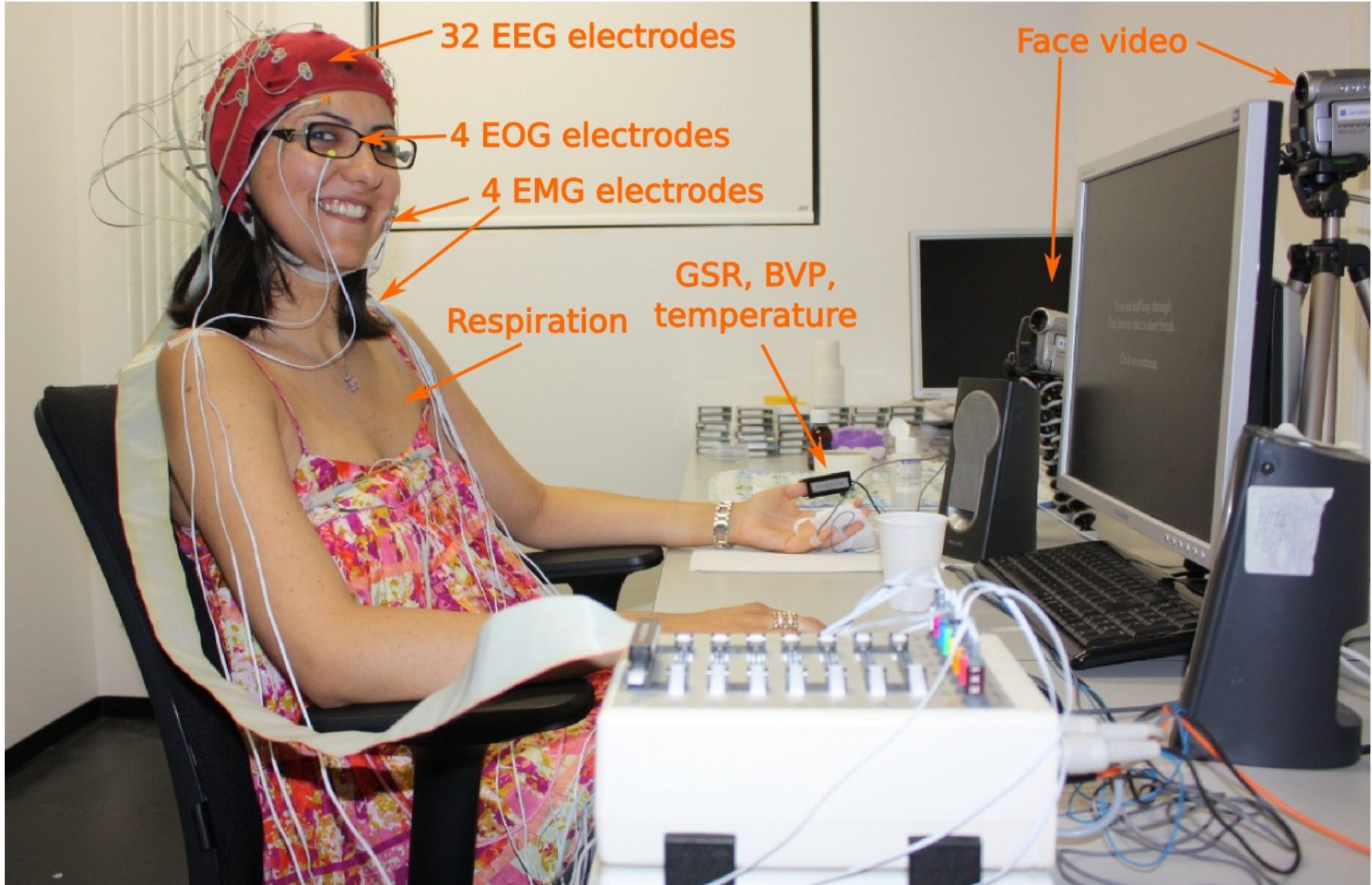
Subjective annotation

The screenshot shows a video player interface. The video frame displays a man in a white hat and a woman in a red jacket. Below the video frame is a valence scale titled "Valence (negative-positive)". The scale consists of five stick figures with different facial expressions, ranging from a sad face on the left to a happy face on the right. Below each figure is a radio button. The fourth radio button from the left is selected, indicated by a red dot.

Stimuli selection (3)



Experiment (1)

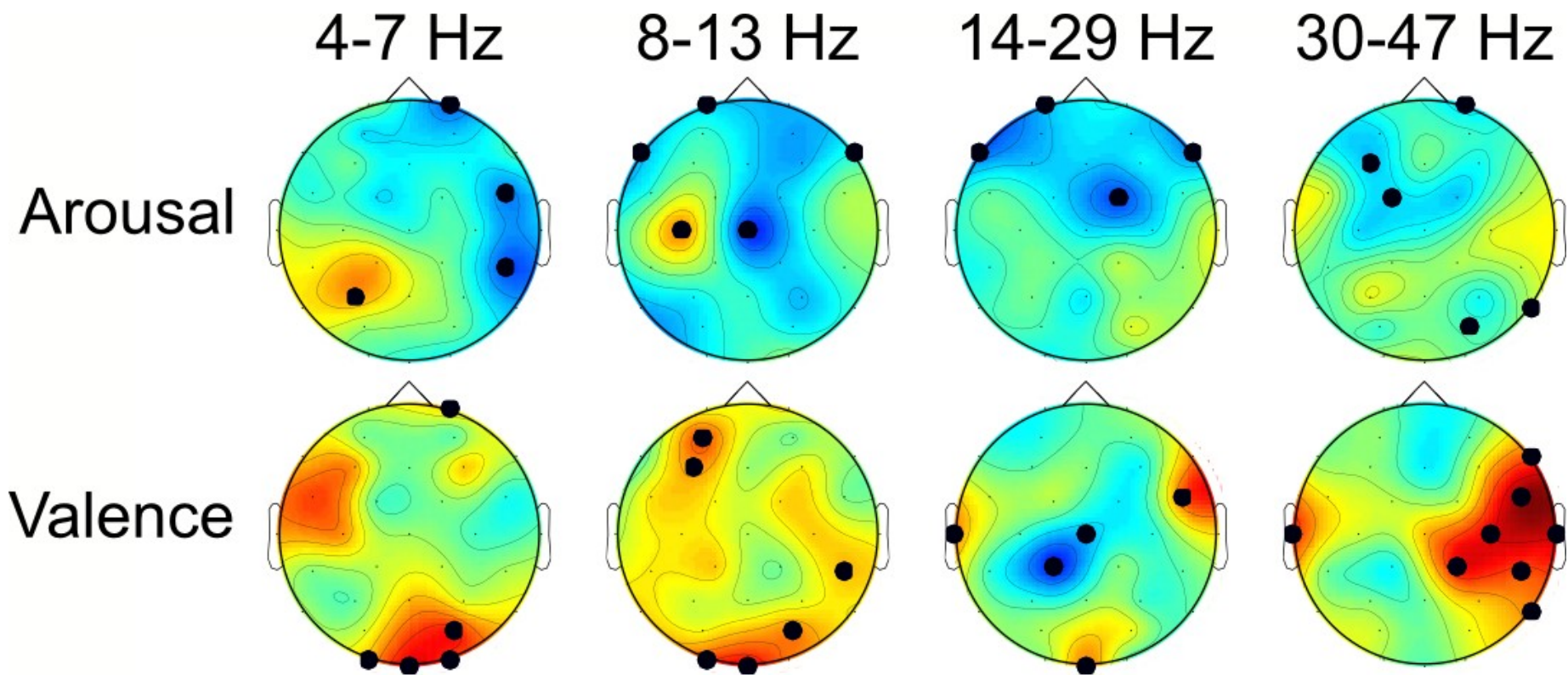


- Dataset contents
 - 32 subjects watching 40 one-minute videos
 - 32 channel EEG @ 512Hz
 - 12 channel physiological signals
 - DV-quality face video
 - Self-assessment per video:
 - arousal, valence, liking

This dataset has more recorded participants than any other publicly available EEG/affect dataset.

Analysis (1)

Spearman correlations of bandwidth power with subjective ratings ($p < 0.05$ highlighted):



- Binary single-trial classification of thresholded subjective ratings for extracted MCA, EEG, and physiological features.
- Results (F1-scores, **= $p < 0.01$, *= $p < 0.05$):

| | Arousal | Valence |
|---------------|---------|---------|
| EEG | 0.583** | 0.563** |
| Physiological | 0.533* | 0.608** |
| MCA | 0.618** | 0.605** |
| Fusion | 0.629** | 0.658** |
| Random class. | 0.481 | 0.490 |

- Dataset available online at:
<http://www.eecs.qmul.ac.uk/mmv/datasets/deap>
- So far:
 - 3 publications by SpudTV members
 - Requests for access from 36 different institutions

The screenshot shows the DEAP dataset website. The header includes the title "DEAPdataset" and a subtitle "a dataset for emotion analysis using eeg, physiological and video signals". Navigation links for "home", "dataset description", "download", and "contact" are visible. The main content area features a photograph of a participant wearing an EEG cap, a 3x4 grid of EEG topographic maps, and an abstract section. The abstract describes the dataset's purpose and methodology. A scatter plot of Valence score (y-axis, -1.5 to 2) versus Arousal score (x-axis, -1.5 to 2) is shown, with several data points highlighted and labeled with video clips: "Lionel Armstrong: What a wonderful world", "Star Trek 2", "Reggie death: Presentation on the Reggae music", and "Beethoven: No. 9". A citation for the dataset's publication is provided at the bottom.

DEAPdataset
a dataset for emotion analysis using eeg, physiological and video signals

home dataset description download contact

abstract

We present a multimodal dataset for the analysis of human affective states. The electroencephalogram (EEG) and peripheral physiological signals of 32 participants were recorded as each watched 40 one-minute long excerpts of music videos. Participants rated each video in terms of the levels of arousal, valence, like/dislike, dominance and familiarity. For 22 of the 32 participants, frontal face video was also recorded. A novel method for stimuli selection was used, utilising retrieval by affective tags from the last.fm website, video highlight detection and an online assessment tool.

The dataset is made publicly available and we encourage other researchers to use it for testing their own affective state estimation methods. The dataset was first presented in the following paper:

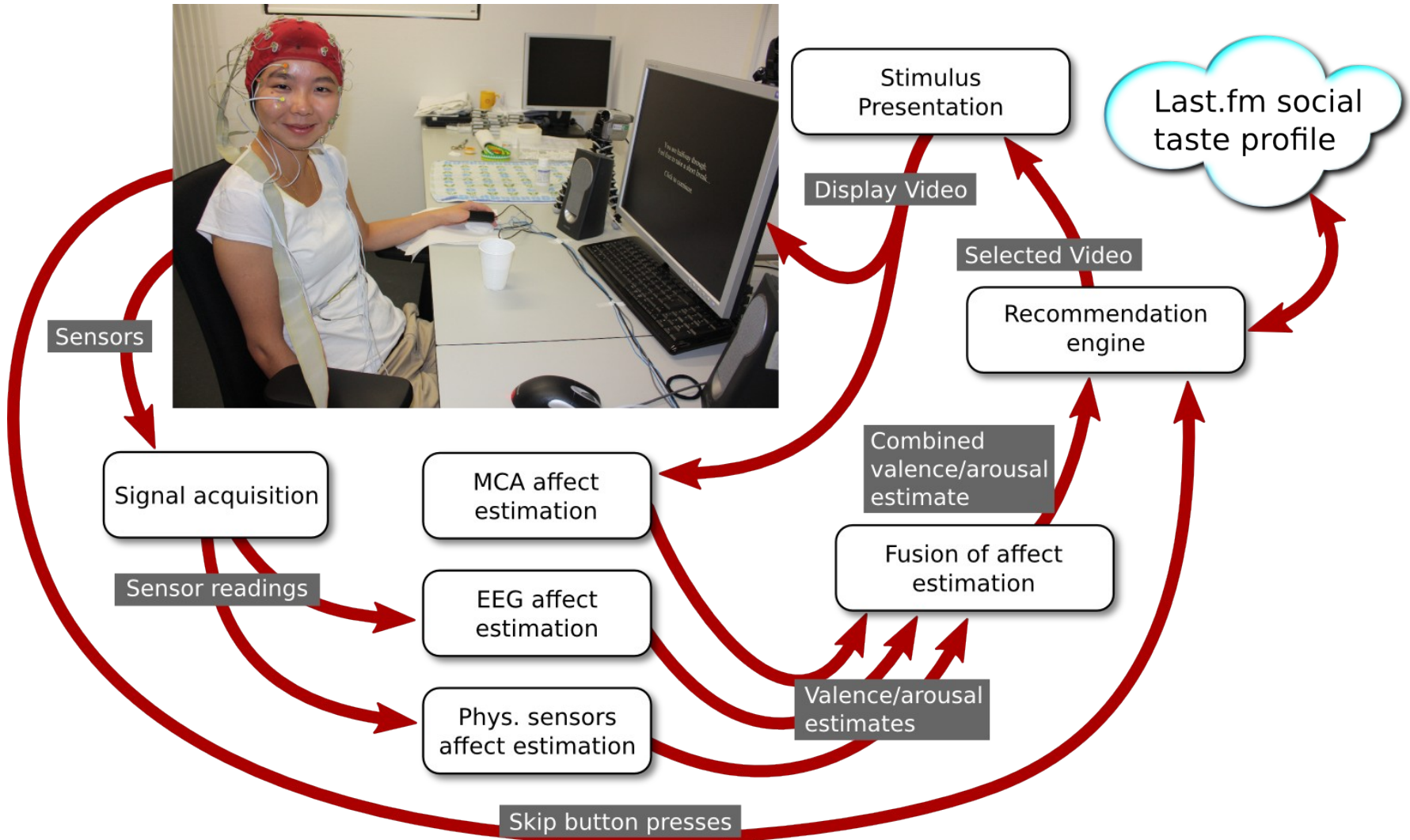
• "DEAP: A Database for Emotion Analysis using Physiological Signals (PDF)", S. Koelstra, C. Muehl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, J. Patras, IEEE Transaction on Affective Computing, Special Issue on Naturalistic Affect Resources for System Building and Evaluation, in press

Goal:

- Design and demonstrate a real-time system that uses affective feedback estimated from EEG, MCA, and physiological signals combined with taste profiles to recommend music videos.



System Overview



- Retrieve 10000 most recent songs from Last.fm profile.
- Clusters songs based on Last.fm tags.
- Hidden Markov Model used to navigate cluster space and recommend new songs.

Personal MTV Demo



U2
I'll Go Crazy If I Don't Go Crazy
Tonight
Kings of Convenience
Misread

Portishead
All Mine

Emiliana Torrini
Wednesday's Child

Jack Johnson
Bubble Toes

Feist
My Moon My Man



- How can we combine the recommender with affective state estimation?
 - Say the user is sad, should we play a happy song?
Or another sad song?
 - Defining a good strategy is **hard!**

- We choose to recommend songs that match the user's current affective state.
- Affective state extra dimension in clustering step
 - Use subjective ratings from online annotations
 - Estimated affective state influences HMM transition probabilities

- Large dataset preferable but difficult
Need manual annotations as recommendation ground truth
- Independently developed building blocks
Different programming languages, platforms, etc.
- Analysis algorithms are resource hungry and developed for offline usage



Implementation details

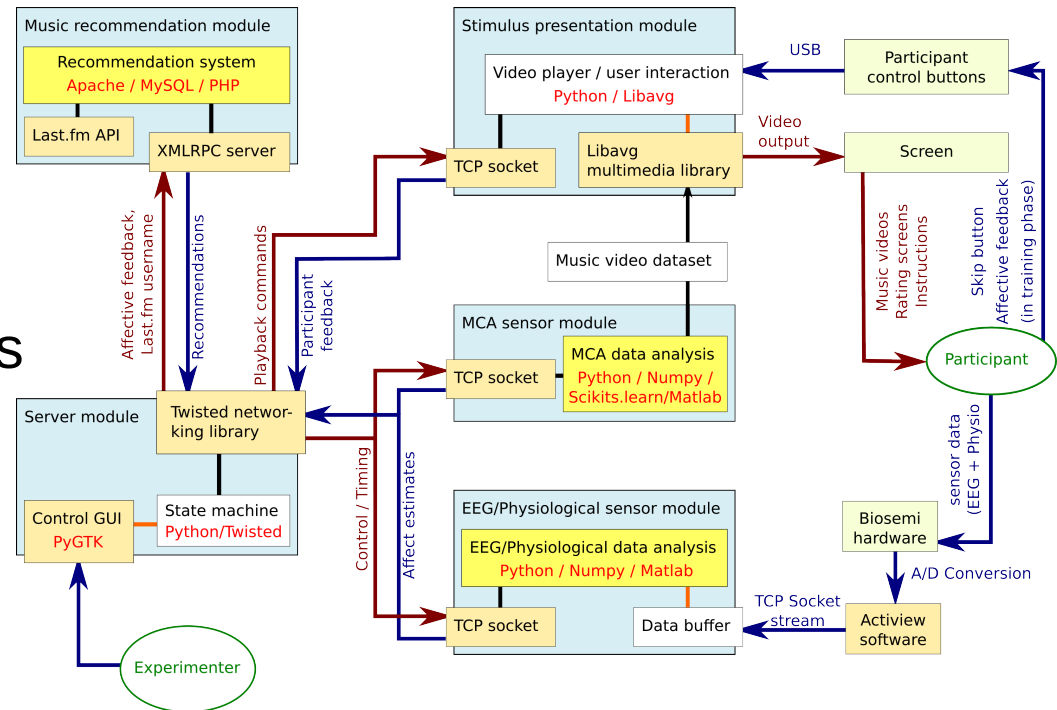
Dataset: 300 videos manually annotated for affect

Estimators adapted for real-time use

Modules run on different PC's

Communication over TCP sockets

Python as glue language to wrap existing implementations



Demo



run_server.py

Subject information
 Subject ID: sander Last.fm User Name: koelie Initialize

Training
 Number of training videos: 20 Start training run

Dummy/Sensor runs
 Run duration (mins): 30 Dummy run Sensor run

Controls
 Stop current action
 End Session

Status
 Current action: Sensor run: 02 / 30 minutes
 Connected modules: Player Dummy889_5 Dummy840_4 Dummy300_6

Valence estimates

| | | | | | | | |
|------------|------|------|------|------|------|------|------|
| Dummy300_6 | 0.44 | 0.41 | 0.27 | 0.10 | 0.13 | | |
| Dummy840_4 | 0.63 | 0.99 | 0.82 | 0.47 | 0.07 | 0.11 | 0.58 |
| Dummy889_5 | 0.21 | 0.38 | 0.04 | 0.26 | 0.64 | | |
| Final | 0.31 | | | | | | |

Arousal Estimates

| | | | | | | | |
|------------|------|------|------|------|------|------|------|
| Dummy300_6 | 0.44 | 0.41 | 0.27 | 0.10 | 0.13 | | |
| Dummy840_4 | 0.63 | 0.99 | 0.82 | 0.47 | 0.07 | 0.11 | 0.58 |
| Dummy889_5 | 0.21 | 0.38 | 0.04 | 0.26 | 0.64 | | |
| Final | 0.10 | | | | | | |

Log

Legend: Recommender StateMachine GUI Player Estimator Error Sensors

Log level: 0 1 2

```

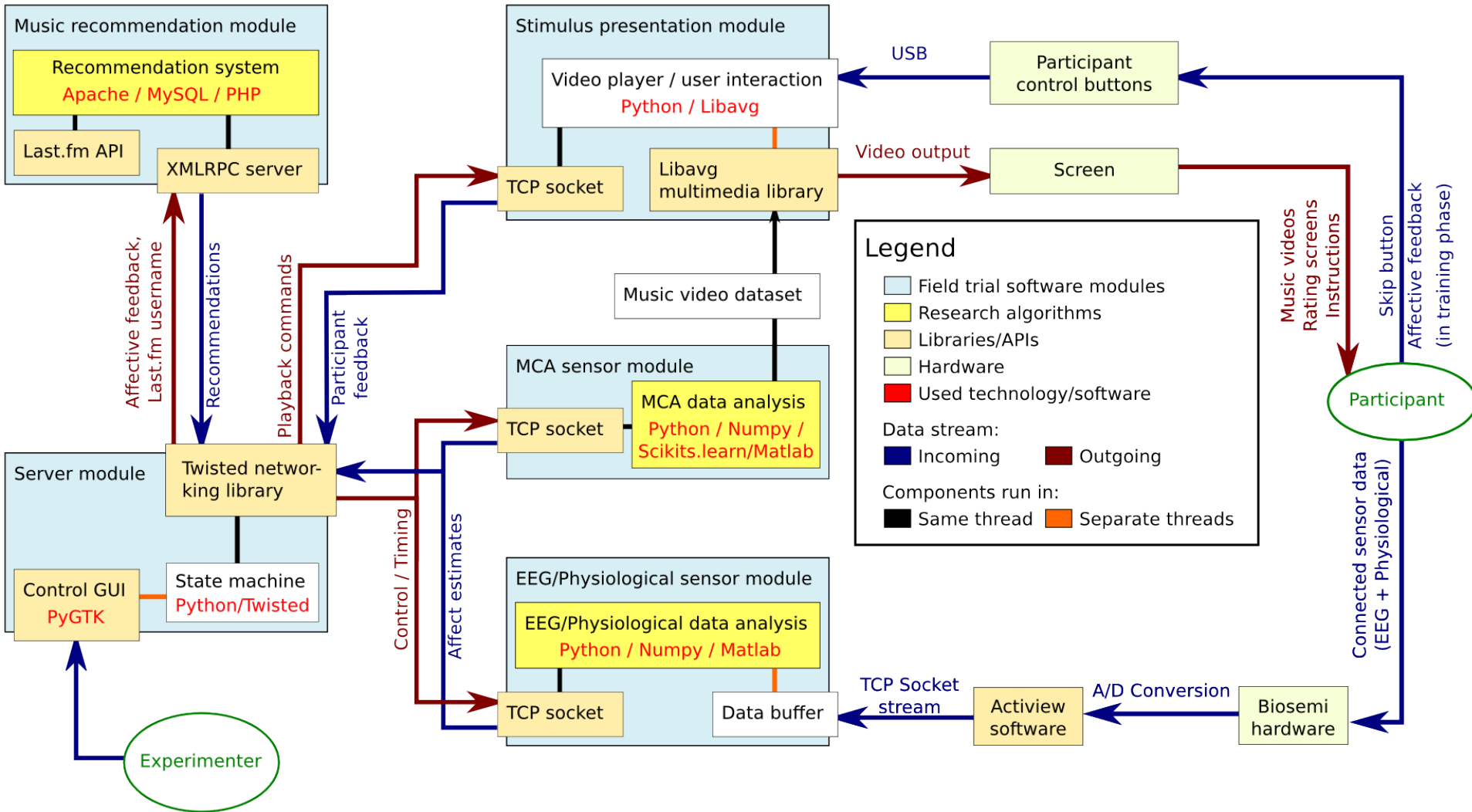
07-06-2011 13:56:30.968782 L02: Recommended songs: [187, 263]
07-06-2011 13:56:30.969248 L02: SEND 127.0.0.1:1 QUEUE 187 263
07-06-2011 13:56:34.778912 L02: RECV Dummy840_4 ESTIMATE 1307447790.778160 1307447794.7781
07-06-2011 13:56:34.779419 L02: Received estimate from Dummy840_4, period 1307447790.778160-1.
07-06-2011 13:56:35.702106 L02: State change: Sensor run: 02 / 30 minutes
07-06-2011 13:56:35.770464 L02: RECV Dummy889_5 ESTIMATE 1307447790.769564 1307447795.7695
07-06-2011 13:56:35.770971 L02: Received estimate from Dummy889_5, period 1307447790.769564-1.
07-06-2011 13:56:35.970563 L02: Calling getNextSong: koelie 154 -1 1
07-06-2011 13:56:35.980688 L02: Recommended songs: [275, 110]
07-06-2011 13:56:35.981145 L02: SEND 127.0.0.1:1 QUEUE 275 110
  
```



Feel free to come and see our live demo of the recommendation system.

Questions ?

Field Trials – Software design



peer to peer / Social

Music video recommendation using social user taste profiles from last.fm website

tagging

Implicit tagging by recognising users' affective states from EEG, MCA, Physiological signals



Multimedia content analysis

MCA analysis for estimating affective content and highlight detection