SpudTV

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

Sander Koelstra
Queen Mary University of London
Talk outline

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
- Christian Mühl
- Engin Kurutepe
SpudTV Scenario

• 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:

- Pleasant/Unpleasant
- Valence
- Arousal

- Elated
- Happy
- Calm
- Content
- Neutral
- Excited
- Sleepy
- Sad
- Frustrated
- Afraid
- Angry

Inactive → Active
Stimuli selection (1)

• 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

Subjective annotation

Manual filtering
- is the tag valid?
- is there a video?
- is the song appropriate?
Experiment (1)

- 32 EEG electrodes
- 4 EOG electrodes
- 4 EMG electrodes
- Respiration
- GSR, BVP, temperature
- Face video
Experiment (2)

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

- **Arousal**
  - 4-7 Hz
  - 8-13 Hz
  - 14-29 Hz
  - 30-47 Hz

- **Valence**
  - 4-7 Hz
  - 8-13 Hz
  - 14-29 Hz
  - 30-47 Hz
Analysis (2)

- Binary single-trial classification of thresholded subjective ratings for extracted MCA, EEG, and physiological features.

- Results (F1-scores, **=p<0.01, *=p<0.05):

<table>
<thead>
<tr>
<th></th>
<th>Arousal</th>
<th>Valence</th>
</tr>
</thead>
<tbody>
<tr>
<td>EEG</td>
<td>0.583**</td>
<td>0.563**</td>
</tr>
<tr>
<td>Physiological</td>
<td>0.533*</td>
<td>0.608**</td>
</tr>
<tr>
<td>MCA</td>
<td>0.618**</td>
<td>0.605**</td>
</tr>
<tr>
<td>Fusion</td>
<td>0.629**</td>
<td>0.658**</td>
</tr>
<tr>
<td>Random class.</td>
<td>0.481</td>
<td>0.490</td>
</tr>
</tbody>
</table>
DEAP Dataset

- 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
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.
Recommendation system (1)

• 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.
• 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
Implementation Challenges

- 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
Feel free to come and see our live demo of the recommendation system.
Questions ?
Field Trials – Software design

Legend:
- Field trial software modules
- Research algorithms
- Libraries/APIs
- Hardware
- Used technology/software

Data stream:
- 

Components run in:
- Same thread
- Separate threads

Music videos:
- Rating screens
- Instructions

 Participant

TCP Socket stream

A/D Conversion

Biosemi hardware

Actiview software

SpudTV

28 September 2011
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