Imperial College London

Affective multimedia analysis: introduction, background and perspectives

Mohammad Soleymani
Intelligent Behaviour Understanding Group (iBUG)
Department of Computing
Imperial College London, UK
m.soleymani@imperial.ac.uk

Outline

Initial ideas

Affect and multimedia

Definitions

State of the art

Open issues

Best practices

Perspectives

Initial idea

Picard, R. (1995). Affective computing. Technical Report 321, MIT Media Laboratory, MIT Media Laboratory: Perceptual Computing; 20 Ames St., Cambridge, MA 02139.

There were similar ideas prior to this paper, e.g., Kensai Eng.

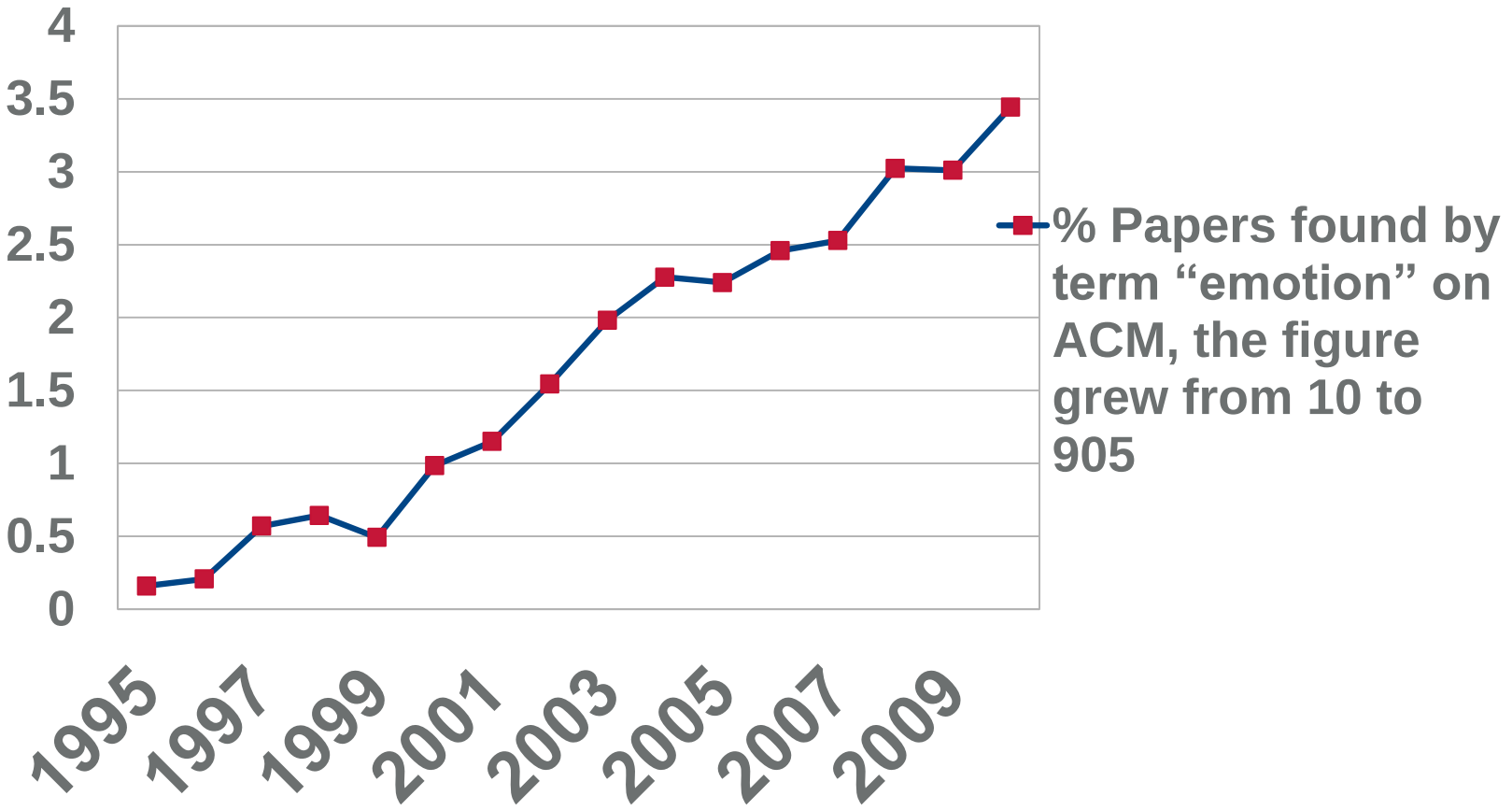
Rosalind Picard coined the term “affective computing”

The first manuscript was rejected two times

Her first paper includes several ideas regarding using affective computing for multimedia content management:

- Summarization
- Profiling/preferences
- Annotation, recommendation and retrieval
- Aesthetics, interactive emotional content

Affective computing



Affect and multimedia

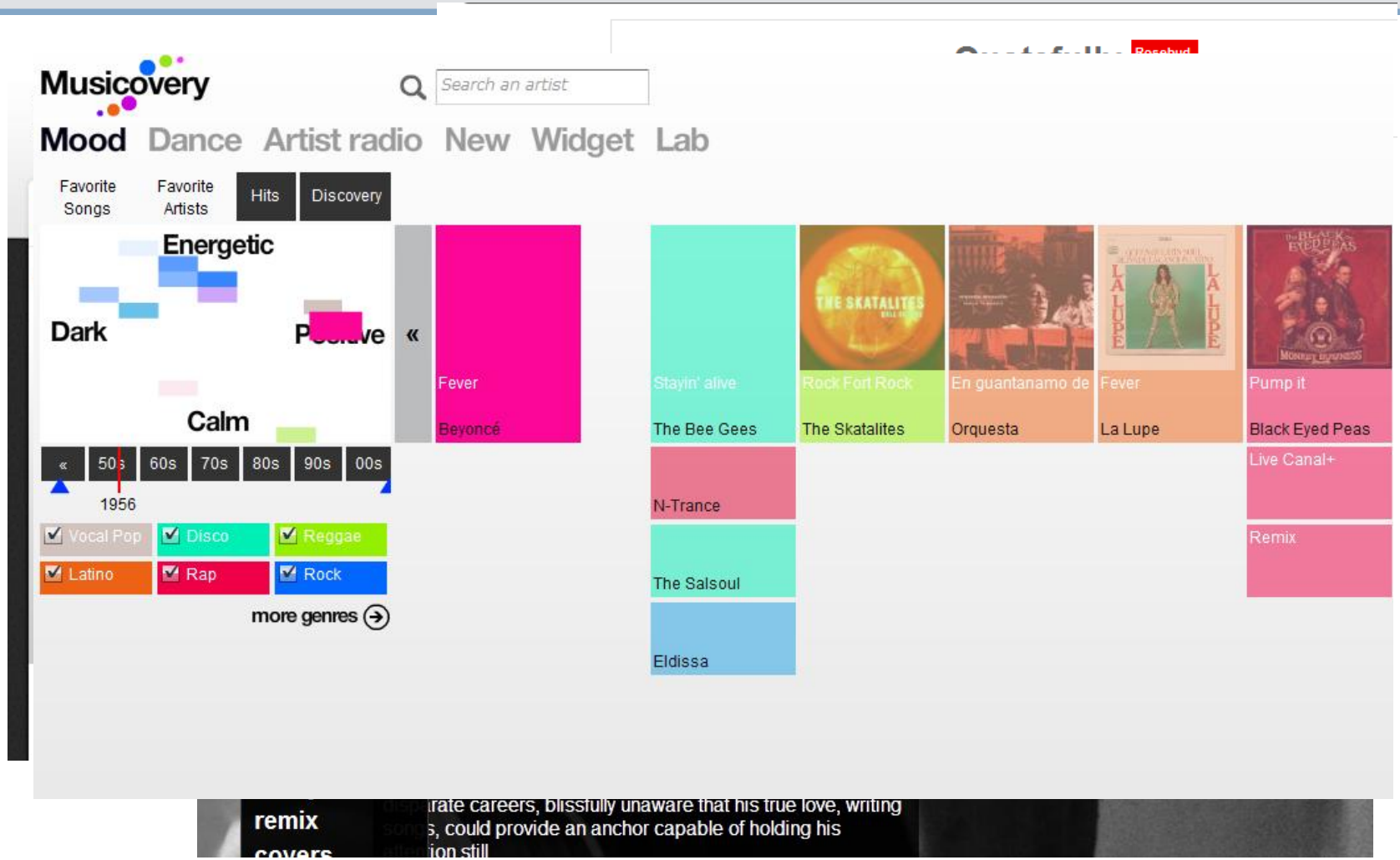
Affective responses are personal and user-dependent

There is a more common response to content, i.e. genre ,e.g., most of you are bored.

Significance:

- ACM MM grand challenges
  - » HP Challenge: Understanding the Emotional Impact of Images and Videos.
- Growing interest from industry from cosmetics to broadcasting.
- Real world examples:

Real world examples -online radios



Affect definition and terminology

We are emotions? No consensus among psychologists

According to the cognitive theory of emotions:

> " An episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems **in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism**."
>
> Klaus Scherer

Stimulus→ perception of an event/action → evaluation according to norms → subjective feeling

e.g. you are jogging and hear a dog barking

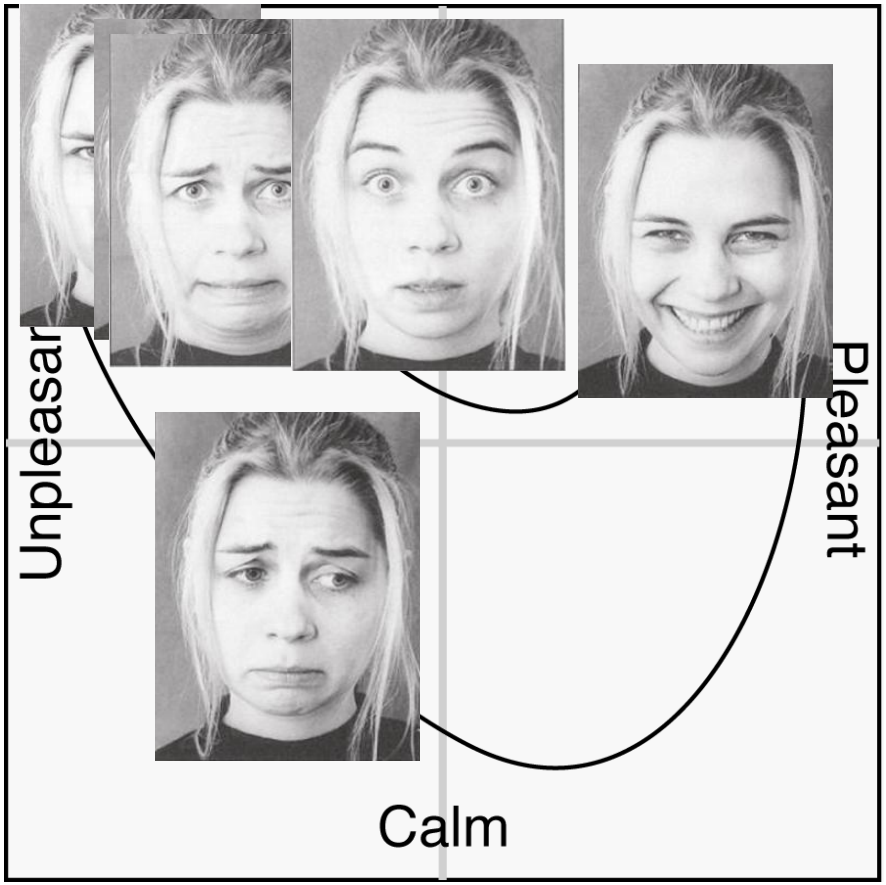
→ fear/anger depending on the size of the dog

Definition of affective terms

- **Emotion or affect:**
  - »Utilitarian: supported by Darwinian perspective. Examples: fear, anger, etc
  - »Aesthetic: present in art, e.g., appreciation

- **Mood**: diffused affective state, long term, often <u>without a certain stimulus</u>

Emotional representation

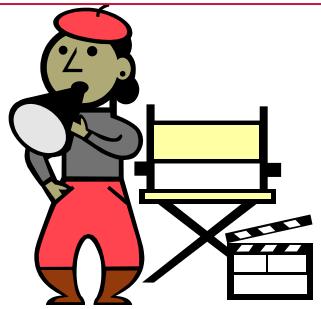


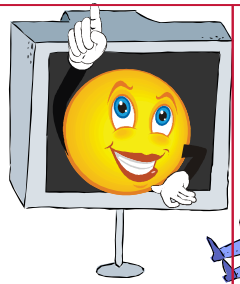
Dietz & Lang, 1999

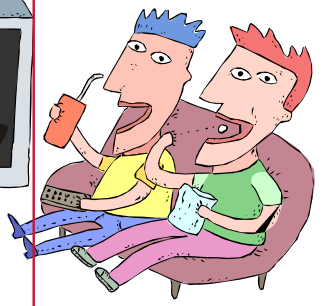
- Arousal and valance are correlated
- There are overlapping emotions
- Some set of basic emotions are biased toward the negative emotions

Affect in response to Multimedia

Intrinsic emotion







Intended emotion/ genre

Multimedia content

Felt emotion / extrinsic

Identify:

Content Analysis

Self report or Sensors

Content

Human

State of the art

[Extrinsic] Affective tagging or summarization:

- User generated; users can assign tags e.g., funny, disgusting, etc

- **Content based, based on content, color, rhythm, etc.**

- Implicit (behavior based from sensors)
  - » Facial expressions
  - » Physiological signals
  - » Speech prosody and utterances

State of the art, cntd.

Study	Dataset	Emotional rep.	Annotators/Annotations
Hanjalic & Xu 2005	2 movies and 2 soccer matches	Continuous	None/None
Wang & Cheong, 2006	36 Full-length movies	Discrete	3/ Ekman basic emotions
Yang et al, 2008	Full-length tracks	Continuous	253/ pairwise, valence
Xu et al, 2008	720 minutes selected scenes	Discrete	?/discrete
Soleymani et al, 2009	21 full length movies	Both	1/continuous
Irie et al, 2010	206 scenes	Discrete	16/discrete
	Too small		**Too few/uncertain**

Open issues

Dataset development:

Psychological datasets for inducing are irrelevant, e.g. IAPS

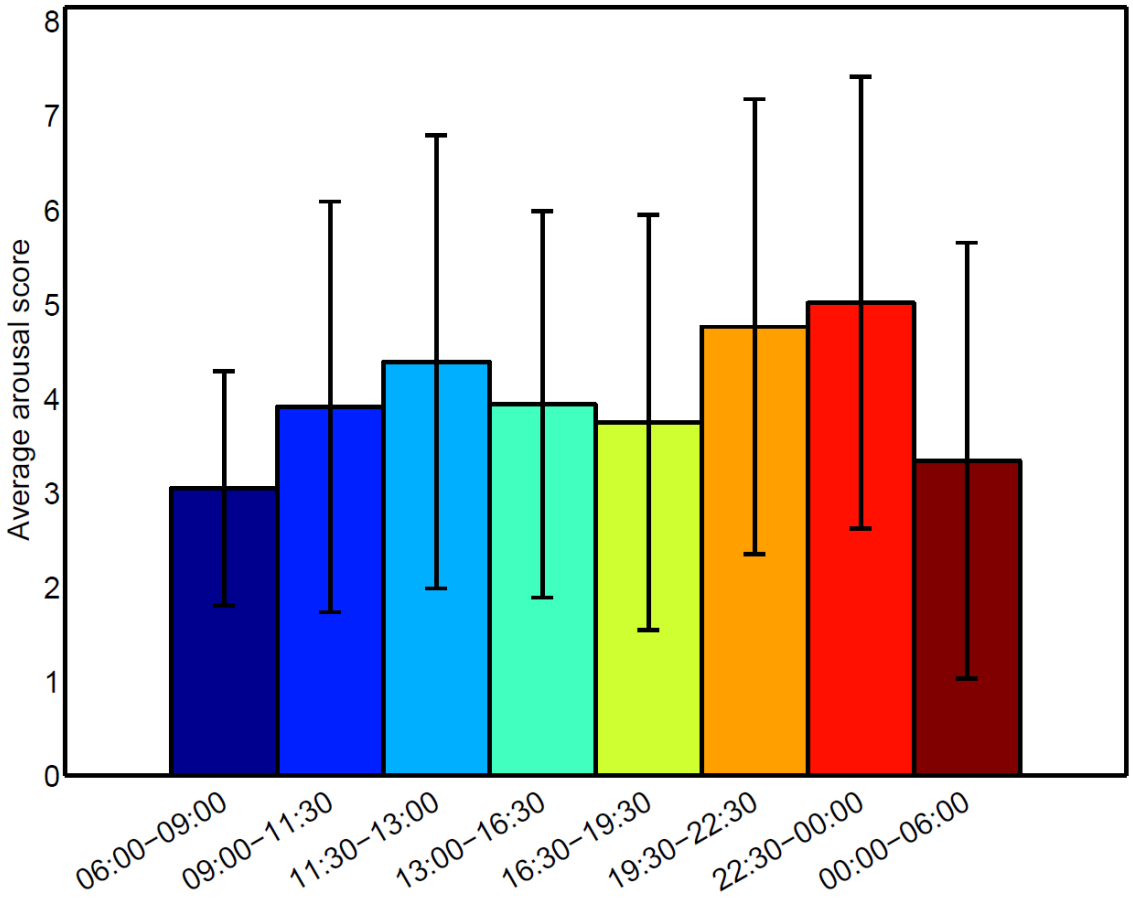
Restricted by copyright and small databases

Models and representations:

Ekman basic emotions are utilitarian and based on facial expression and basic emotions are biased toward negative emotions

Context should be taken into account, time, weather, geo-location, social dimension, age, etc

Context - time



Online study on movies, Soleymani et al, under preparation

Best practices

Audio music mood classification (MIREX 2007-2012)

http://www.music-ir.org/mirex

Acoustic features

- Toolbox: MIRToolbox

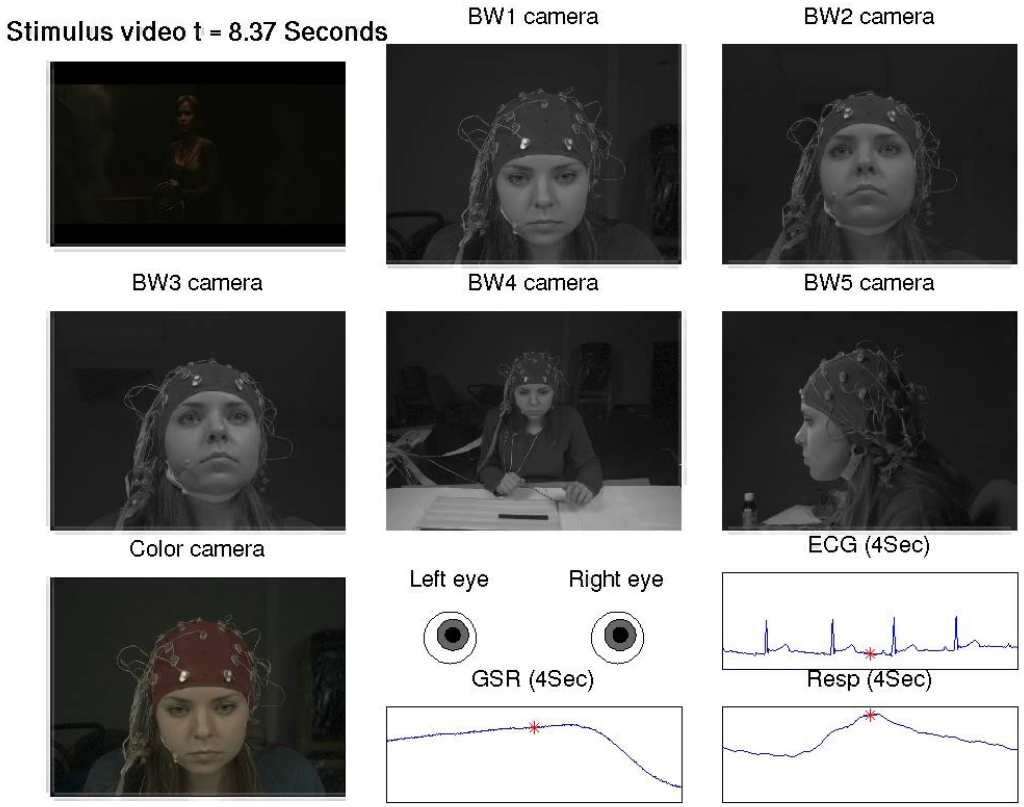
Mediaeval 2010-2012 (focused on one type)

- Boredom detection
- Violence detection
  - » Use case from Technicolor, movies

Public mood/sentiment detection via social media, e.g., http://wefeelfine.org/

Best Practices -- databases

MAHNOB HCI available at http://mahnob-db.eu/



Perspectives

Commercial products will emerge

- New technologies will facilitate the research, e.g., sensors will help tagging, Kinect face tracker, Emotiv & Neurosky helmets.

More collective work on standard benchmarking development.

- Large databases with annotations

Better understanding between communities:

- IEEE Trans. Affective Computing
- **Affective Computing and Intelligent interaction 2013, Geneva, Switzerland** (http://www.acii2013.org)

Imperial College London

Acknowledgement

Marie Curie program
Intra-European Fellowship

European Research Council

Established by the European Commission