VRIJE
UNIVERSITEIT
BRUSSEL

Next Generation User Interfaces
Implicit Human-Computer Interaction

Prof. Beat Signer

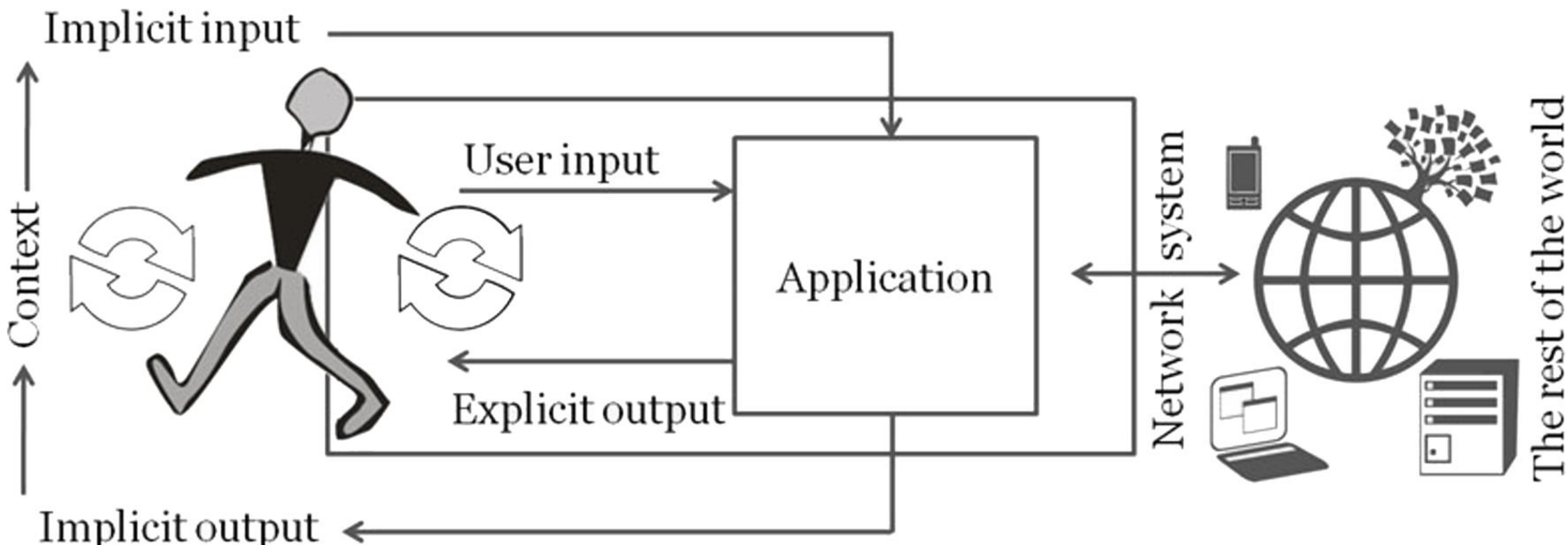
Department of Computer Science
Vrije Universiteit Brussel

http://www.beatsigner.com

Implicit Human-Computer Interaction

- Over the last decade, we have seen a clear trend towards *smart environments and living spaces* where *sensors and information processing* is *embedded into everyday objects* as foreseen in *Mark Weiser*'s vision of *ubiquitous computing* with the *goal to simplify the use of technology*
- In Implicit Human-Computer Interaction (IHCI), we try to use contextual factors (e.g. various sensor input) to build *human-centered anticipatory user interfaces* based on naturally occurring human interactive behaviour
- Context-aware computing can be used to design implicit human-computer interaction

Implicit Human-Computer Interaction …



[https://www.interaction-design.org/encyclopedia/context-aware_computing.html]

- Implicit Human-Computer Interaction (IHCI) is *orthogonal* to (traditional) explicit HCI
  - implicit communication channels (*incidental interaction*) can help in building more natural human-computer interaction

Context

- Context-aware systems often *focus on location* as the only contextual factor
- However, even if location is an important factor, it is only one context dimension

> *Context is any information that can be used to character-ize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves.*

A.K. Dey, 2000

VRIJE UNIVERSITEIT BRUSSEL

Example: Navigation

- Various contextual factors can be taken into account when designing the interface of a car navigation system
  - current location (GPS)
  - traffic information
  - daylight
    - automatically adapt screen brightness
  - weather
  - current user task
    - e.g. touch is disabled while driving and only voice input can be used
  - …

Everyday Examples

- Systems that take user actions as input and try to output an action that is a *proactive anticipation* of what the users need
  - simple motion detector at doors that open the door automatically to allow humans with shopping carts to pass through
  - escalators that move slowly when not in use but speed up when they sense a person pass the beginning of the escalator
  - smartphones and tablets automatically changing between landscape and portrait mode based on their orientation
  - smart meeting rooms that keep track of the number of people in a meeting room and alter the temperature and light appropriately
  - …

Exercise: Context-aware Digital Signage

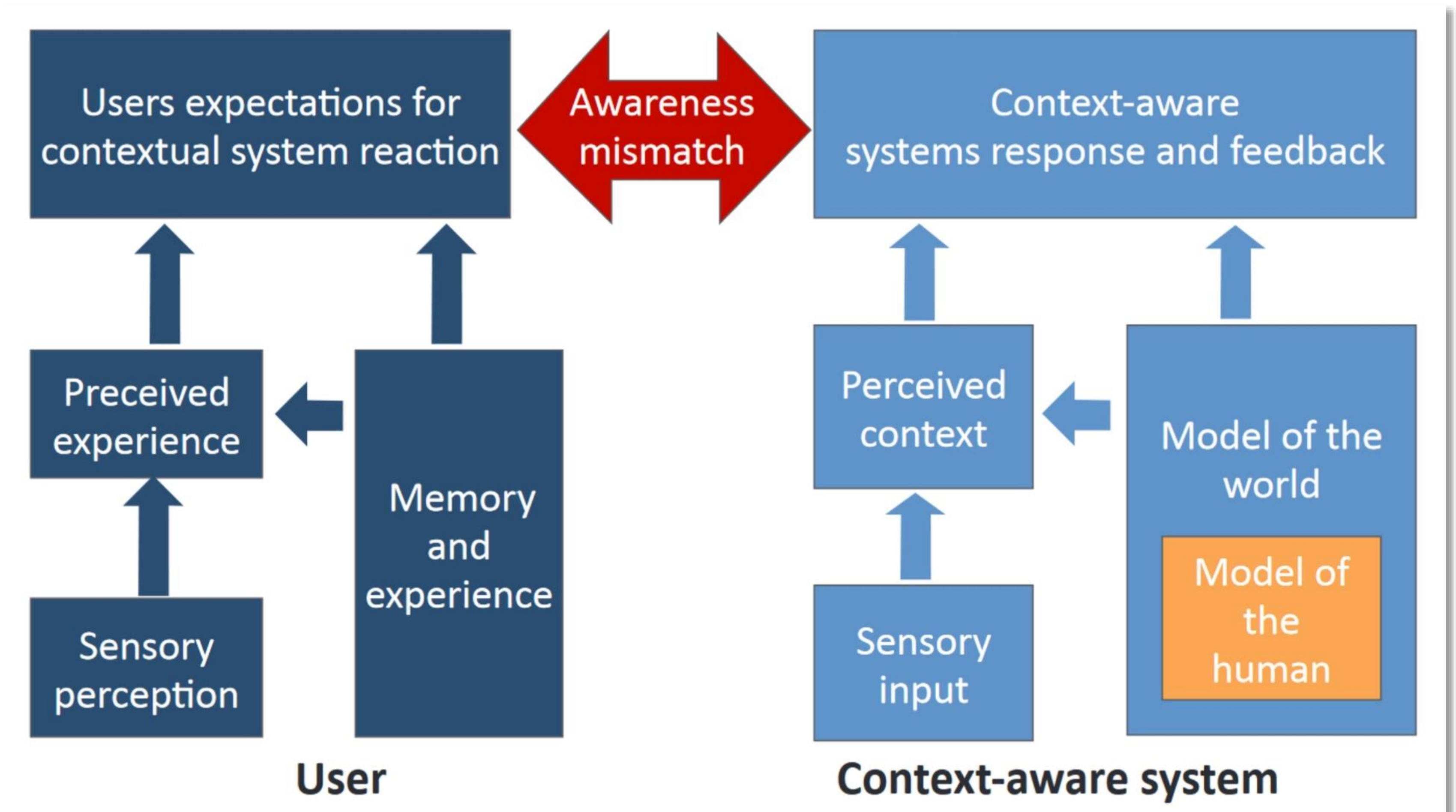
Contextual Factors

- Human factors
  - user
  - social environment
  - task
  - ..
- Physical environment
  - location
  - infrastructure
  - conditions
  - …

From Sensor Input to Context

- *How do we compute the perceived context* from a single or multiple sensor inputs?
  - machine learning techniques?
  - rule-based solutions?
  - …
- How should we *model context*?
  - e.g. generic context models without application-specific notion of context
- How to *trigger implicit interactions* based on context?
- How to *author* new context elements?
  - relationships with sensor input, existing context elements as well as application logic

User-Context Perception Model (UCPM)



Musumba and Nyongesa, 2013

Things Going Wrong

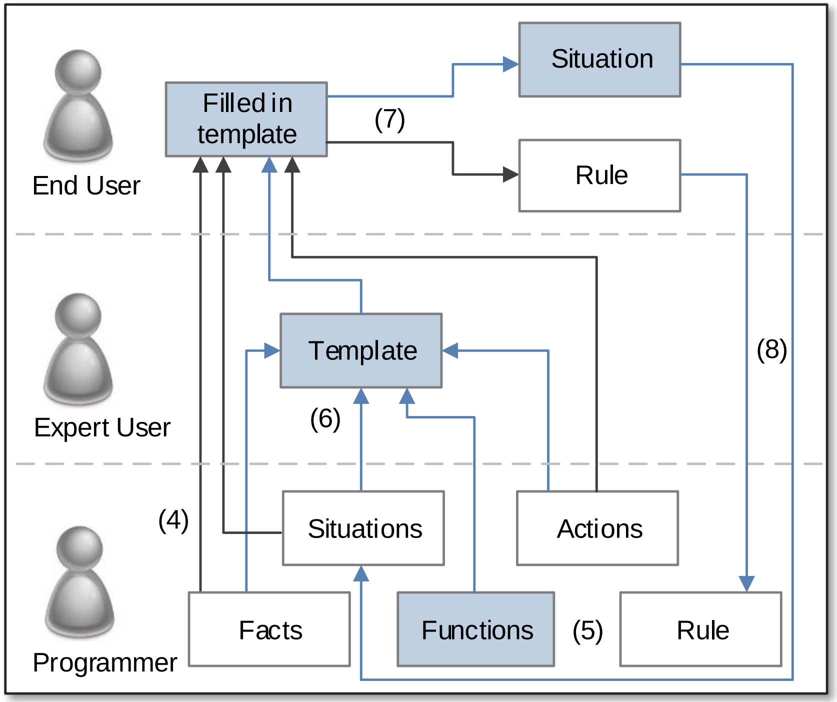
- What if the implicit interaction with a system goes wrong?
  - is it really the wrong system behaviour or is the user just not aware of all factors taken into account (awareness mismatch)?
- The *quality* of implicit human-computer interaction as perceived by the user is directly *related to the awareness mismatch*
- Fully-automated vs. semi-automated systems
  - sometimes is might be better to not fully automate the interaction since wrong implicit interactions might result in a very bad user experience
  - *take the user into the loop*

Intelligibility

- Improved system intelligibility might increase a user's trust, satisfaction and acceptance of implicit interactions
- Users may ask the following questions (Lim et al., 2009)
  - *What*: What did the system do?
  - *Why*: Why did the system do W?
  - *Why Not*: Why did the system not do X?
  - *What If*: What would the system do if Y happens?
  - *How To*: How can I get the system to do Z, given the current context?
- *Explanations* should be provided *on demand* only in order to avoid information overload
  - feedback easier for rule-based solutions than for machine learning-based approaches

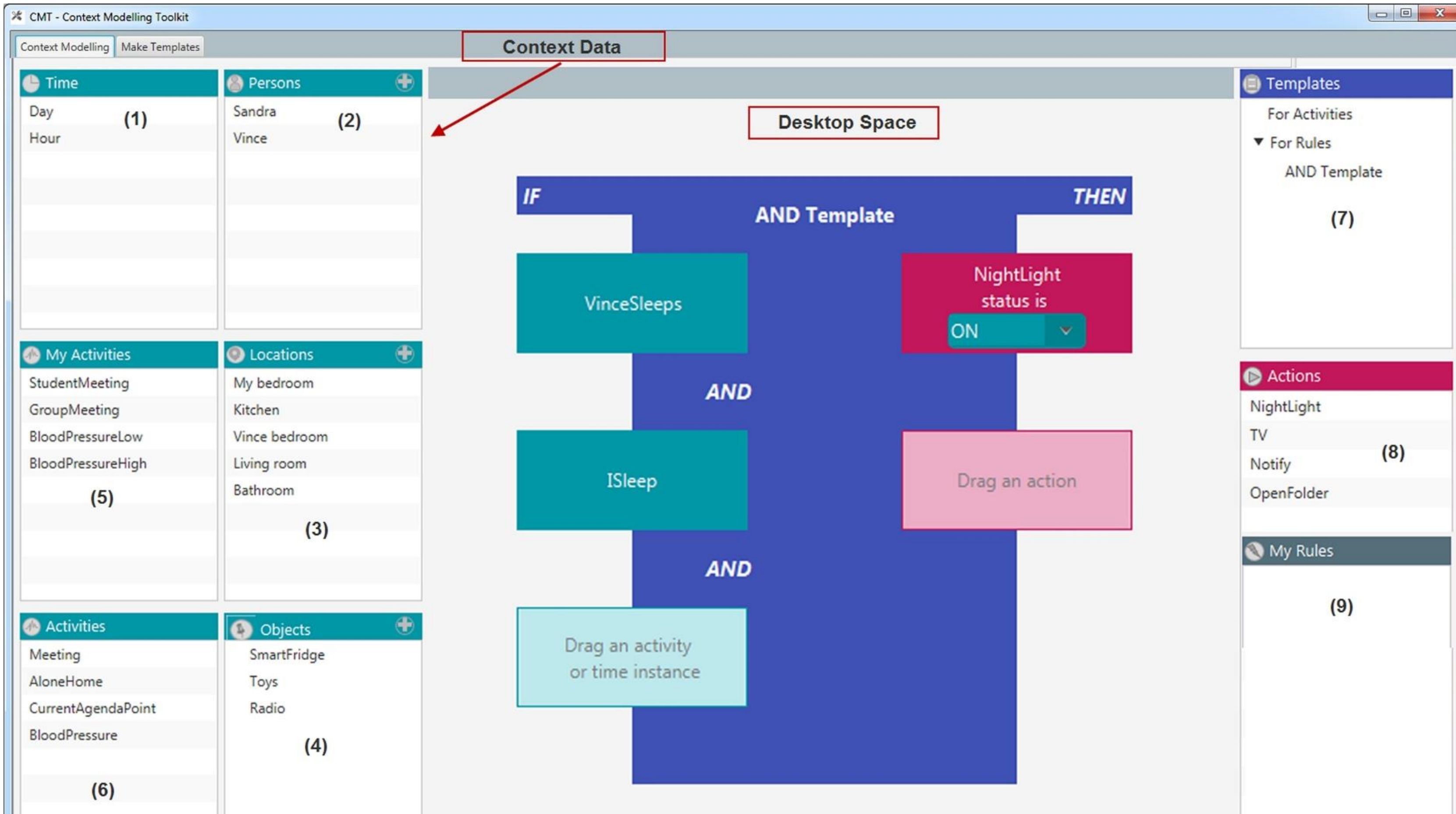
Context Modelling Toolkit (CMT)

- Multi-layered context modelling approach
  - seamless transition between end users, expert users and programmers
- Beyond simple *"if this then that"* rules
  - reusable situations
- Client-server architecture
  - server: context reasoning based on Drools
  - client: sensor input as well as applications



Trullemans and Signer, 2016

Context Modelling Toolkit (CMT) …



Trullemans and Signer, 2016

HCI and HCII in Smart Environments

Smart meeting room in the WISE lab

Some Guidelines for Implicit HCI

- Always first investigate what users want/have to do
  - as a second step see what might be automated
  - use context-awareness as a source to *make things easier*
- The definition of a *feature space* with factors that will influence the system helps in realising context-aware implicit interactions
  - find parameters which are characteristic for a context to be detected and find means to measure those parameters
- Always try to *minimise* the *awareness mismatch*
  - increase *intelligibility* by providing information about the used sensory information (context) in the user interface

Some Guidelines for Implicit HCI …

- Designing proactive applications and implicit HCI is a very difficult task because the system has to anticipate what the user wants
  - always investigate whether a fully-automated solution is best or whether the user should be given some choice (control)

Affective Computing

- Computing that takes into account the *recognition, interpretation, modelling,* processing and synthesis of human affects (*emotions*)
- Implicit human-computer interaction can be based on recognised human emotions

Rosalind W. Picard

Emotions

Emotions are episodes of coordinated changes in several components (neurophysiological activation, motor expression, subjective feelings, action tendencies and cognitive processes) in response to external or internal events of major significance to the organism.

Klaus R. Scherer, Psychological Models of Emotion, 2000

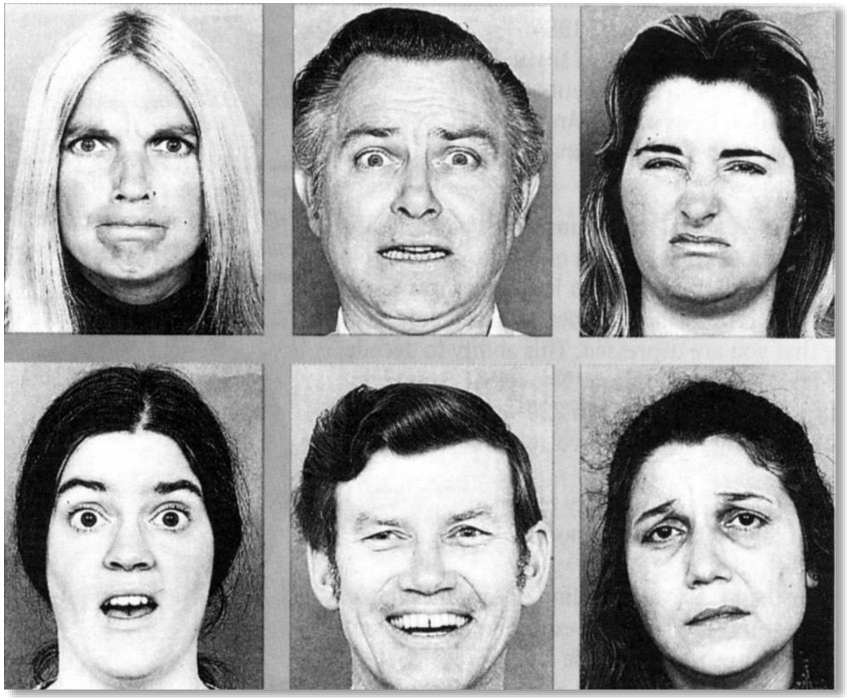
- External events
  - behaviour of others, change in a current situation, …
- Internal events
  - thoughts, memories, sensations ...

Emotion Classification

- Different models to classify emotions
- *Discrete models* treat emotions as discrete and different constructs
  - Ekman's model
  - …
- *Dimensional models* characterise emotions via dimensional values
  - Russell's model
  - Plutchik's model
  - PAD emotional state model
  - …

Ekman's Emotions Model

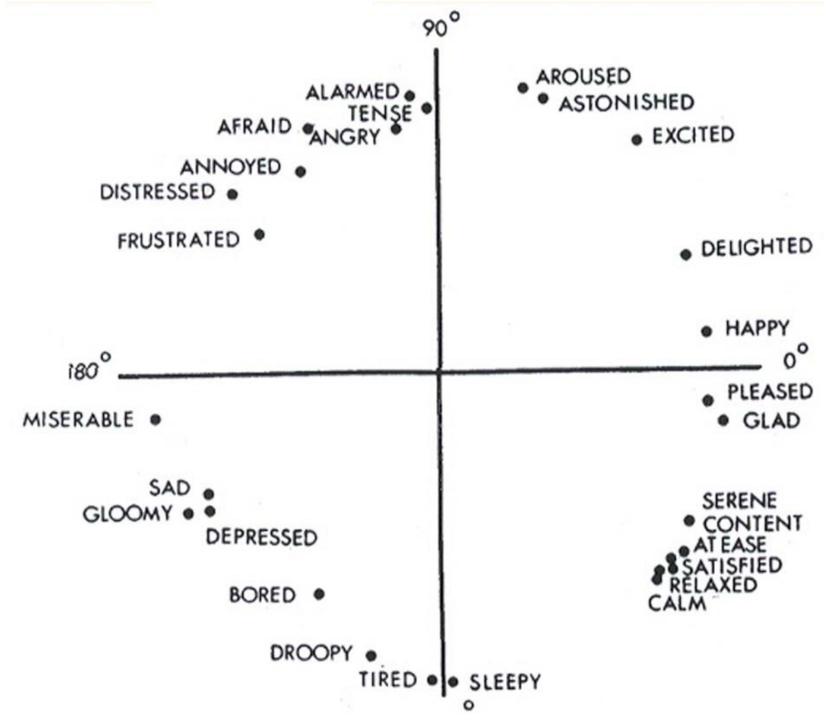
- Theory of the universality of six basic facial emotions
  - anger
  - fear
  - disgust
  - surprise
  - happiness
  - sadness



- Discrete categories can be used as labels for emotion recognition algorithms
  - multiple existing databases rely on Ekman's model

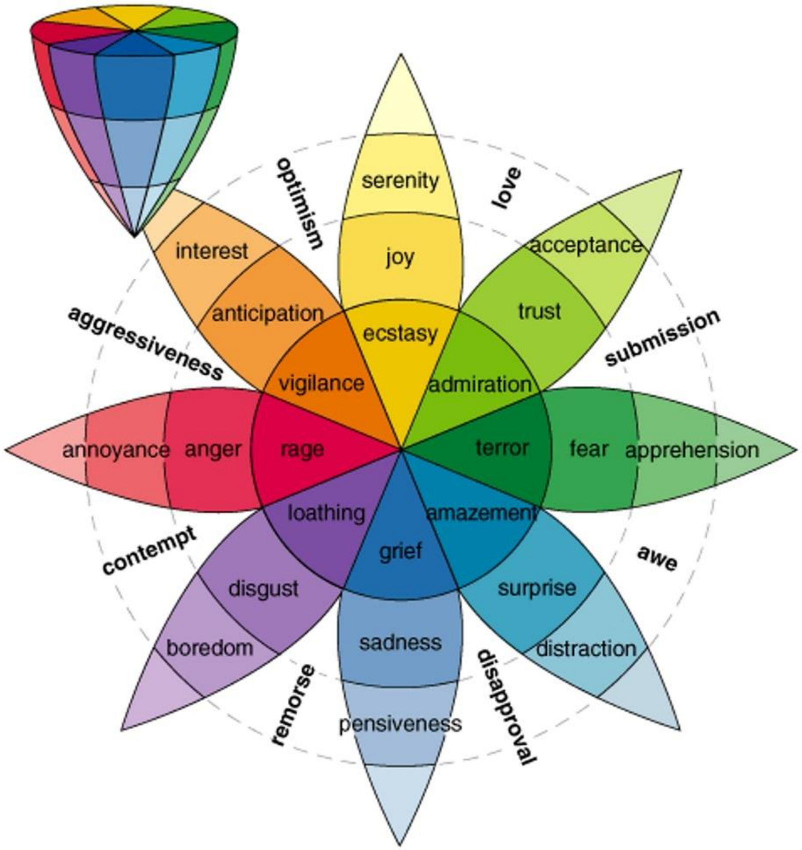
Russell's Circumplex Model of Affect

- Emotions are mapped to two dimensions
  - *valence* (x-axis)
    - intrinsic attractiveness or aversiveness
  - *arousal* (y-axis)
    - reactiveness to a stimuli



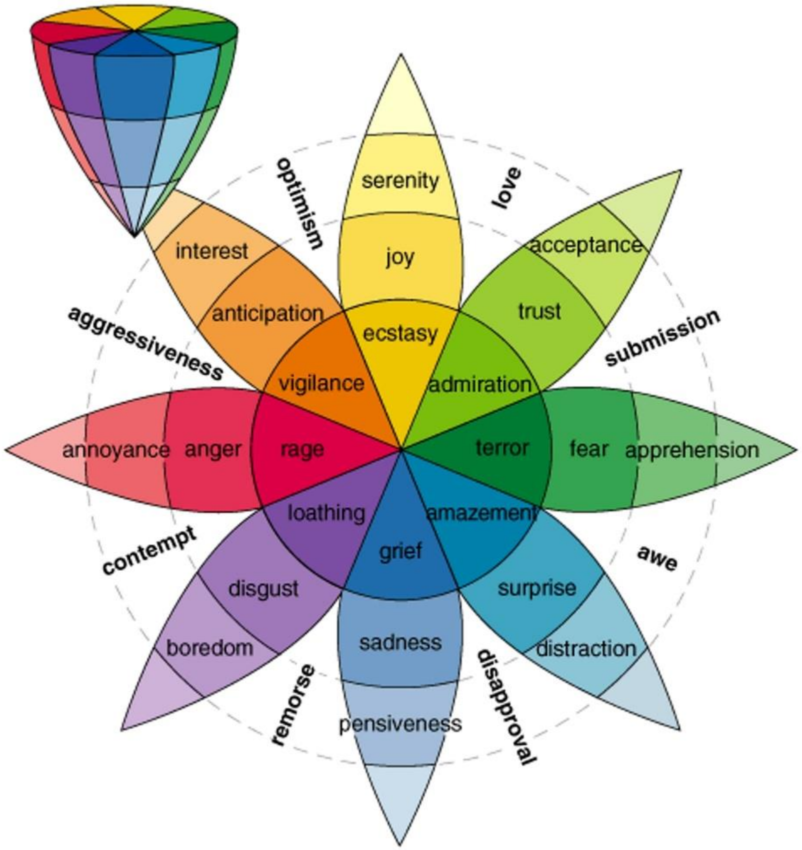
Pluchik's Wheel of Emotions

- Three-dimensional "extension" of Russell's circumplex model
- *8 basic emotions*
  - joy vs. sadness
  - trust vs. disgust
  - fear vs. anger
  - surprise vs. anticipation
- *8 advanced emotions*
  - optimism (anticipation + joy)
  - love (joy + trust)
  - submission (trust + fear)



Pluchik's Wheel of Emotions …

- *8 advanced emotions*
  - awe (fear + surprise)
  - disapproval (surprise + sadness)
  - remorse (sadness + disgust)
  - contempt (disgust + anger)
  - aggressiveness (anger + anticipation)

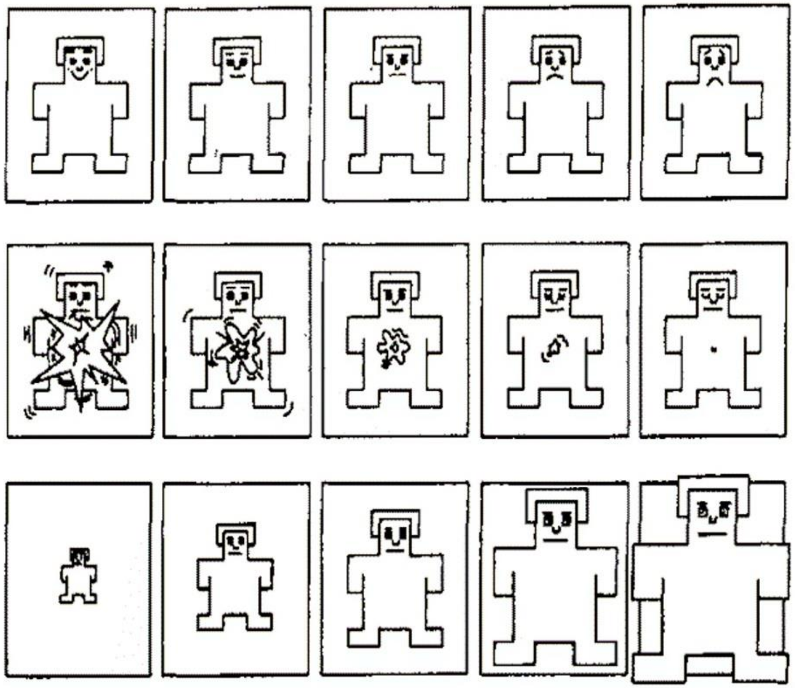


PAD Emotional State Model

- Representation of emotional states via three numerical dimensions
  - *pleasure-displeasure*
  - *arousal-nonarousal*
  - *dominance-submissiveness*
- Example
  - anger is a quite unpleasant, quite aroused and moderately dominant emotion

Self-Assessment of PAD Values

- *Self-Assessment Manikin (SAM)* is a language neutral form that can be used to assess the PAD values
  - each row represents five values for one of the dimensions
    - pleasure
    - arousal
    - dominance

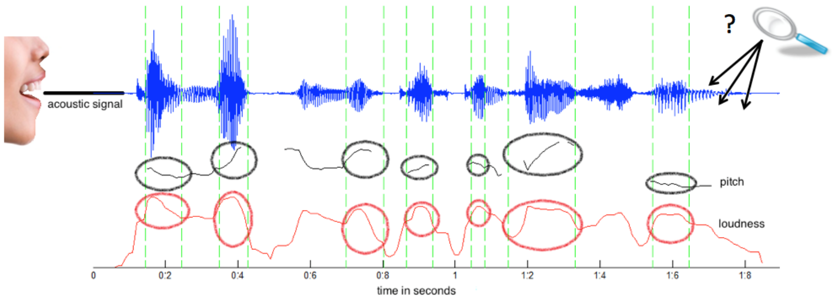


Emotion Recognition

- Emotions can be manifested via different *modalities*
  - acoustic features (voice pitch, intonation, etc.)
  - verbal content (speech)
  - visual facial features
  - body pose and gestures
  - biosignals (physiological monitoring)
    - pulse, heart rate, …
- In general, artificial intelligence algorithms are used for an accurate recognition of emotions
- Potential *multimodal fusion* of multiple modalities
  - improve emotion recognition accuracy by observing multiple modalities

Acoustic Feature Recognition

- Behaviour and evolution of acoustic features over time is meaningful for emotion detection
- Typical features
  - intonation
  - intensity
  - pitch
  - duration

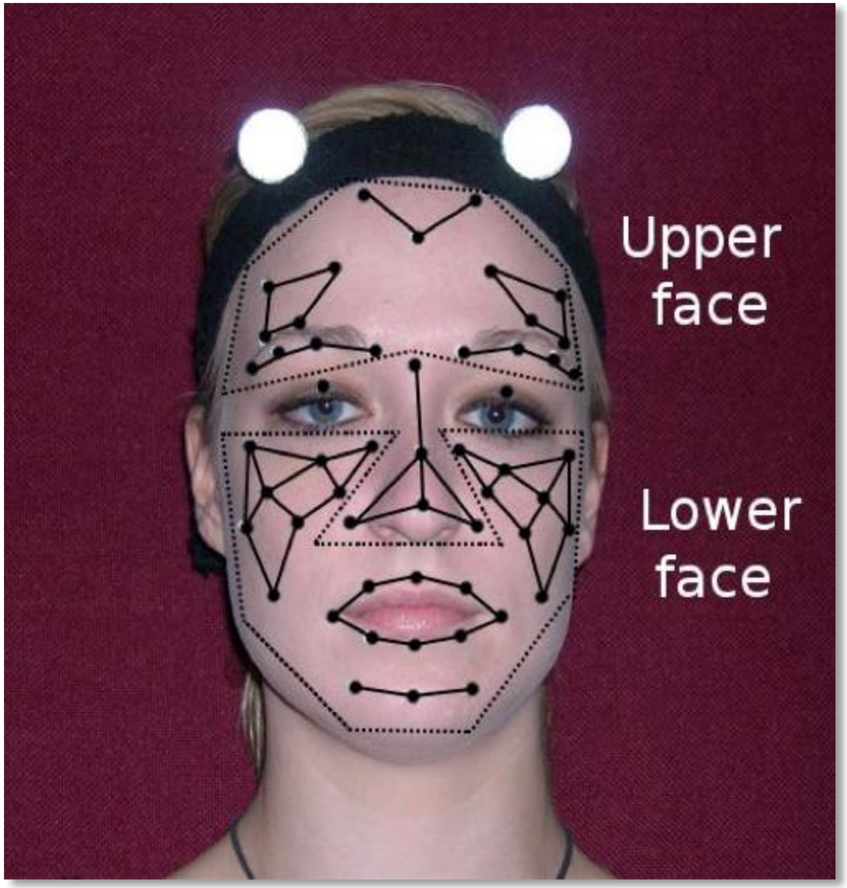


Speech-based Emotion Recognition

- Recognition of emotions from speech content (e.g. via speech recogniser) is based on typical methods such as
  - bag of words (unigrams)
  - n-gram language models
- Typical features
  - emotion dictionaries
  - lattices
  - orthography (punctuation, capitalisation, emoticons)
  - wordnet
  - syntax
  - semantic roles
  - world knowledge

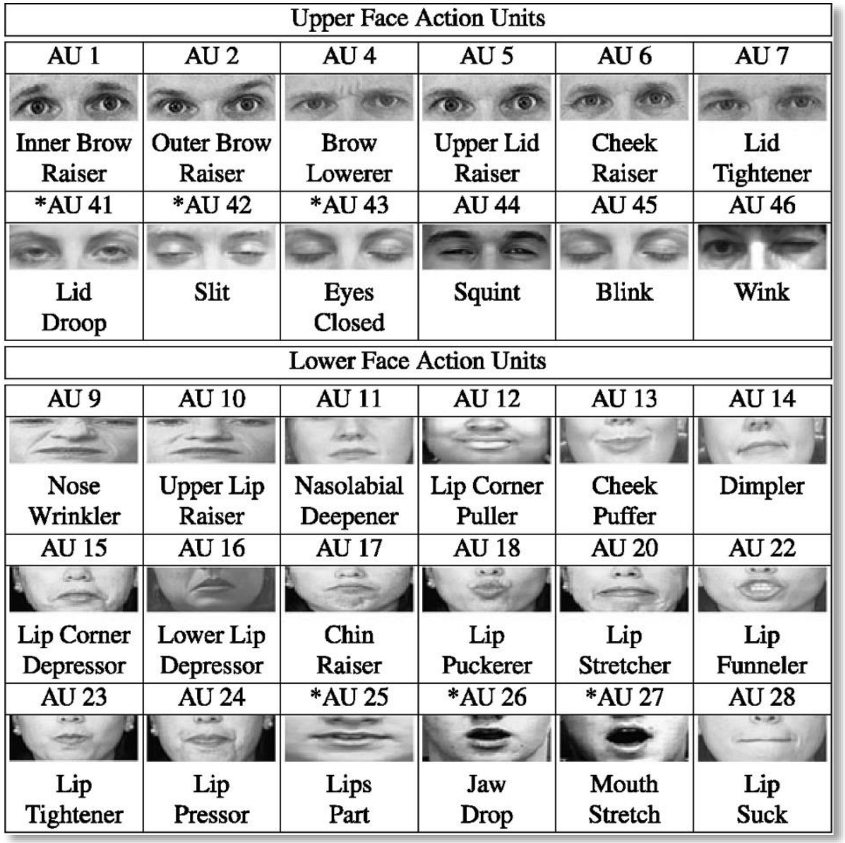
Facial Emotion Recognition

- Find face parts
  - use orientation or prominent features such as the eyes and the nose
- Extract facial features
  - geometry based
  - appearance based (textures)
- Classification through
  - support vector machines
  - neural networks
  - fuzzy logic systems
  - active appearance models



Facial Action Coding System (FACS)

- Used to describe changes, contraction or relaxations of muscles of the face
- Based on so-called *Action Units (AUs)*
  - description for component movement or facial actions
  - combination of AUs leads to facial expressions
    - e.g. sadness = AU 1+4+15
  - `http://www.cs.cmu.edu/~face/facs.htm`



Body Pose and Gestures

- Body language carries rich emotional information
  - body movement, gestures and posture
  - relative behaviour (e.g. approach/depart, looking/turning away)
- Detailed features extracted from motion capture

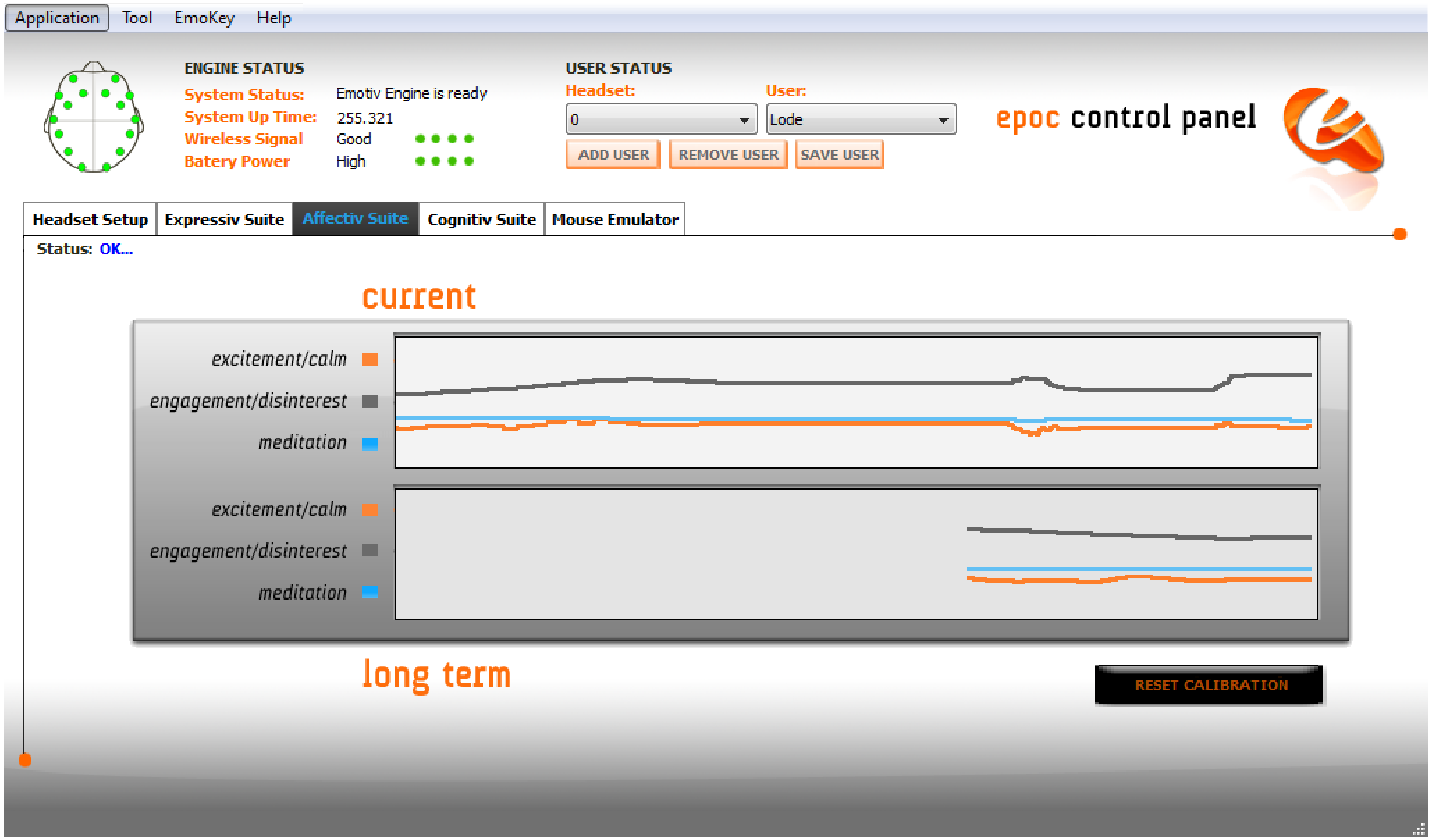
Biosignals

- Different emotions lead to different biosignal activities
  - anger: increased heart rate and skin temperature
  - fear: increased heart rate but decreased skin temperature
  - happiness: decreased heart rate and no change in skin temperature
- Advantages
  - hard to control deliberately (fake)
  - can be continuously processed
- Disadvantages
  - user has to be equipped with sensors
- Challenge
  - wearable biosensors

Emotiv EPOC Neuroheadset

- Non-invasive EEG device
- 14 sensors
- Integrated gyroscope
- Wireless
- Low cost
- Average sensor sensibility
  - mainly due to sensor non-invasiveness

Emotiv EPOC Neuroheadset …

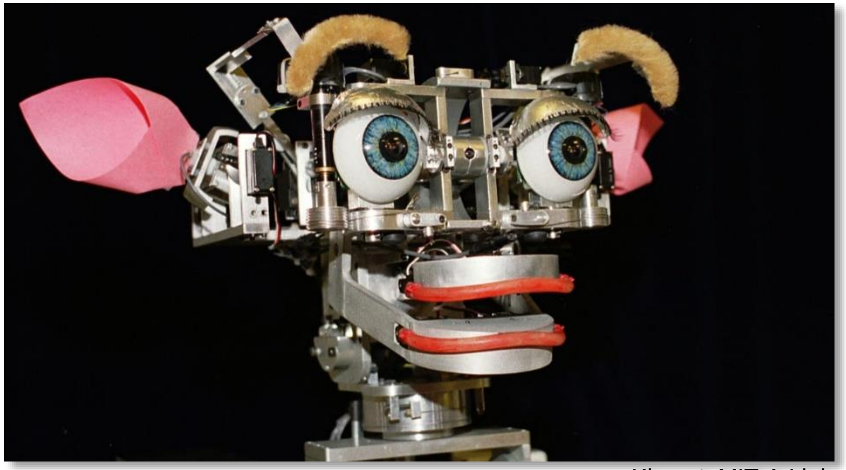


From Signals to Labelled Emotions

- Five potential channels
  - visual: face
  - visual: body movement
  - acoustic: speech content
  - acoustic: acoustic features
  - physiological: heart rate, blood pressure, temperature, GSR, EMG
- Associating *emotion descriptors*
  - machine learning problem
  - SVMs, HMMs, NNs?
  - rely on only single modality or fusion of multiple modalities?
  - associate emotion descriptors before or after fusing the modalities?
    - i.e. feature- or decision-level fusion?

Synthesis of Emotions

- Intelligent agents support social interactions with users (showing emotions)
  - real life (robots)
  - virtual reality (virtual agents)
- "Characters with a brain"
  - reason about environment
  - understand and express emotion
  - communicate via speech and gesture
  - applications
    - e-learning
    - robots and digital pets
    - …



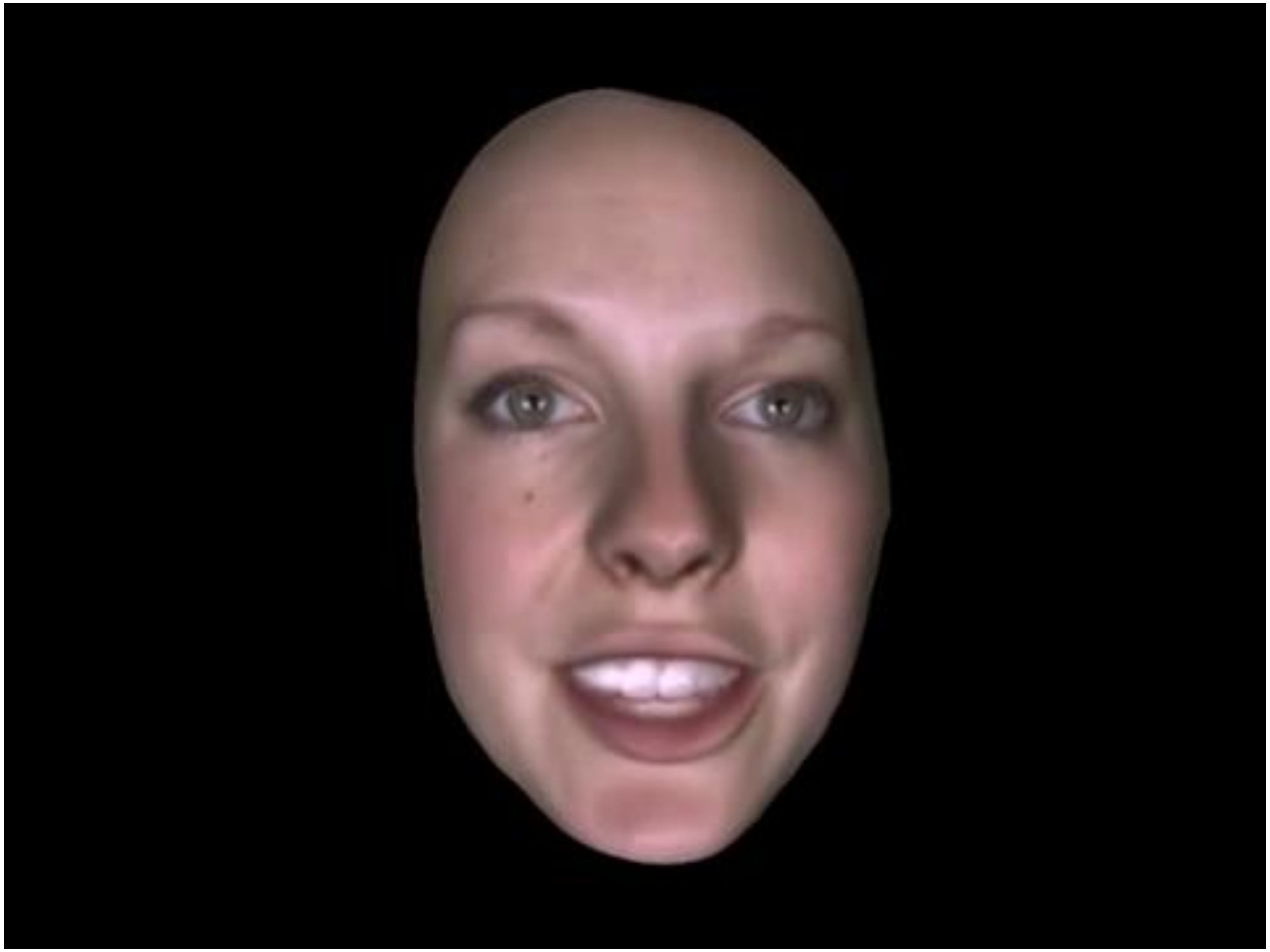
Kismet, MIT A.I lab

Virtual Characters

- Virtual character with human behaviour that supports face-to-face human-machine interaction
- Basic physical behaviour
  - walking, grasping
- Non-verbal expressive behaviour
  - gestures, facial expression (emotion), gaze
- Spontaneous and reactive behaviour
  - responsiveness to events

Max Headroom, 1987

Video: Text-driven 3D Talking Head



Effectors in Emotion Synthesis

- Facial expressions
  - emotion categories have associated facial action programs
  - Facial Action Coding System (FACS)
- Gestures
  - deictic, iconic, …
  - timing and structure are important
- Gaze
  - roles of gaze: attention, dialogue regulation, deictic reference
  - convey intentions, cognitive and emotional state
- Head movement
  - during conversation head is constantly in motion
  - nods for affirmation, shakes for negation, …

EmoVoice Framework

- Real-time recognition of emotions from acoustic speech properties
  - uses features from pitch, energy, duration, voice quality and spectral information
  - uses the Open Sound Control (OSC) protocol
  - mirroring of emotions to the user
    - http://www.informatik.uni-augsburg.de/lehrstuehle/hcm/projects/tools/emovoice/

Homework

- Read the following paper that is available on PointCarré (papers/Weiser 1991)
  - M. Weiser, *The Computer for the 21st Century*, ACM Mobile Computing and Communications Review, July 1991

References

- M. Weiser, *The Computer for the 21st Century*, ACM Mobile Computing and Communications Review, July 1991
  - http://dx.doi.org/10.1145/329124.329126
- A. Schmidt, *Context-Awareness, Context-Aware User Interfaces and Implicit Interactions*
  - https://www.interaction-design.org/encyclopedia/context-aware_computing.html
- G.W. Musumba and H.O. Nyongesa, *Context Awareness in Mobile Computing: A Review*, International Journal of Machine Learning and Applications, 2(1), 2013
  - http://dx.doi.org/10.4102/ijmla.v2i1.5

References ...

- B.Y. Lim, A.K. Dey and D. Avrahami, *Why and Why Not Explanations Improve the Intelligibility of Context-aware Intelligent Systems*, Proceedings of CHI 2009, Boston, USA, April 2009
  - https://doi.org/10.1145/1518701.1519023
- S. Trullemans and B. Signer, *A Multi-layered Context Modelling Approach for End Users, Expert Users and Programmers*, Proceedings of SERVE 2016, International Workshop on Smart Ecosystems cReation by Visual dEsign, Bari, Italy, June 2016
  - http://beatsigner.com/publications/trullemans_SERVE2016.pdf

References …

- J.A. Russel, *A Circumplex Model of Affect*, Journal of Personality and Social Psychology, 39(6), 1980
  - https://www2.bc.edu/james-russell/publications/Russell1980.pdf
- R.W. Picard, *Affective Computing*, MIT Technical Report No. 321, 1995
  - http://affect.media.mit.edu/pdfs/95.picard.pdf
- Expressive Text-driven 3D Talking Head
  - http://www.youtube.com/watch?v=TMxKcbQcnK4

Next Lecture

Course Review