

Affective Computing 2025

L1. Introduction to AC

Welcome words



Before we start..

Teaching Team



Main Lecturer
Assis. Prof. Haoyu Chen
chen.haoyu@oulu.fi
TS321

Haoyu Chen

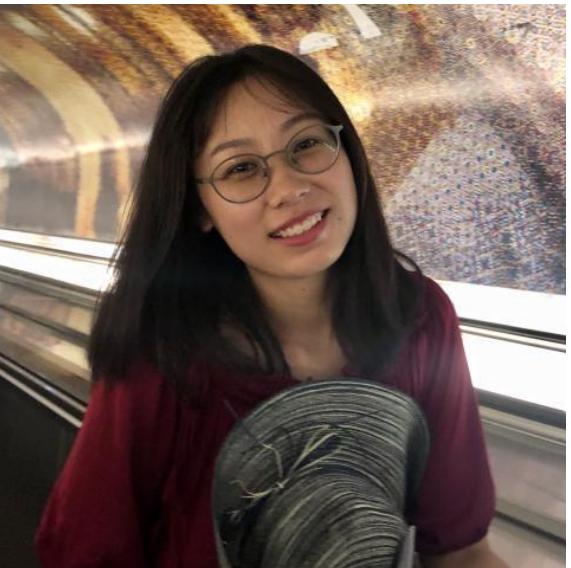


Assistant Lecturer
Dr. Yante Li
yante.li@oulu.fi,
TS311



Assistant Lecturer
Dr. Puneet Kumar
puneet.kumar@oulu.fi,
TS330

Teaching Team



Teaching Assistant

Fang Kang

fang.kang@oulu.fi;

TS327;

(in English)



Teaching Assistant

Dr. Matti Matilainen

matt.matilainen@oulu.fi

, TS324

(in Finnish)



Teaching Assistant

Mr. Yushun Xiao

yushun.xiao@oulu.fi,

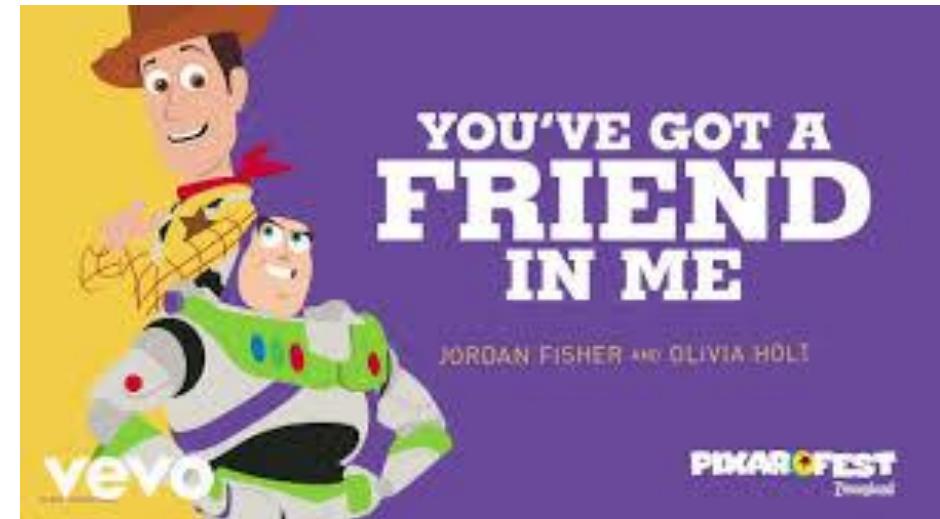
TS336

(in English)



We hate traditional lecturing!

Chat and Make friends with students around you now



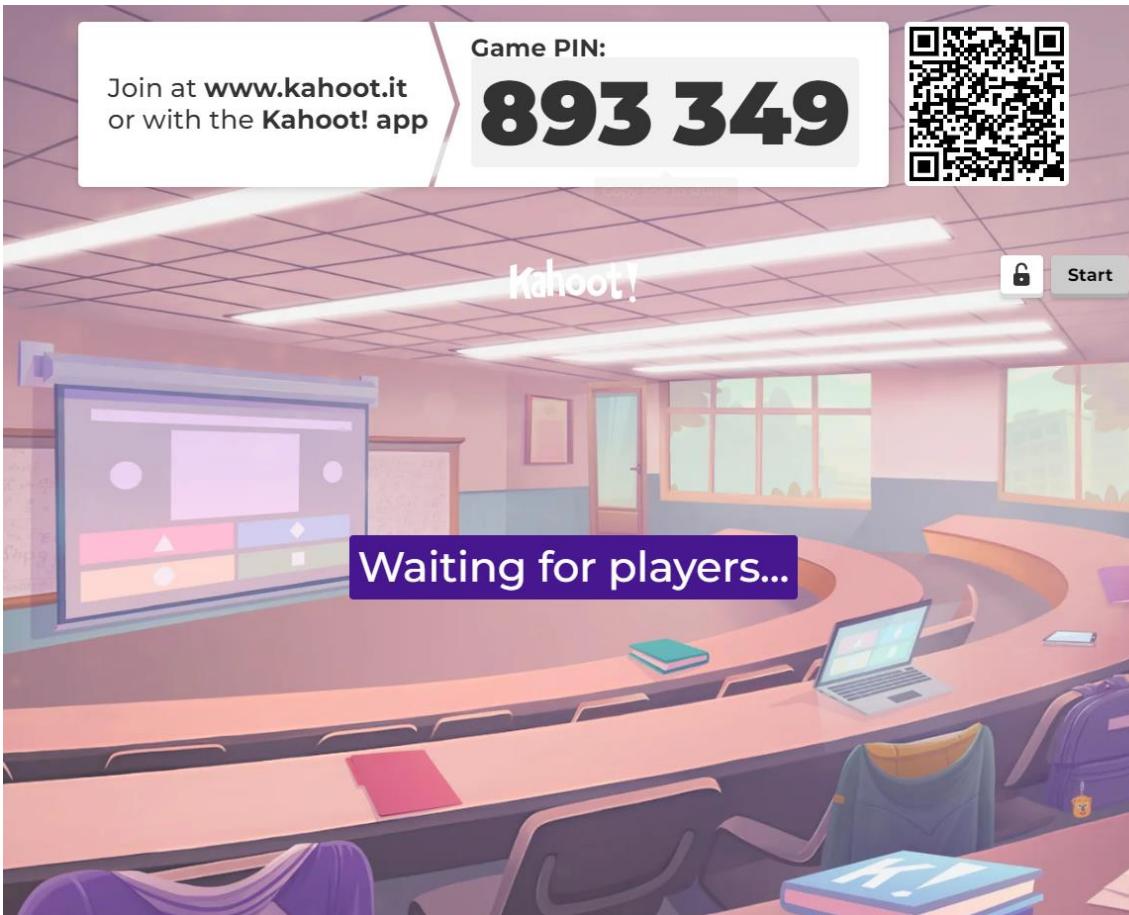
<https://theobservantmom.com/youre-right-traditional-education-sucks/>

Prerequisites

- Algorithms
- Introduction to AI
- Basic Data Structures
- Computer vision, Machine learning (Advanced courses)
- Programming skills using Python (Python 3.6)
 - If you don't know Python, you CAN take this class (but get ready to invest some time early on)



Kahoot time!



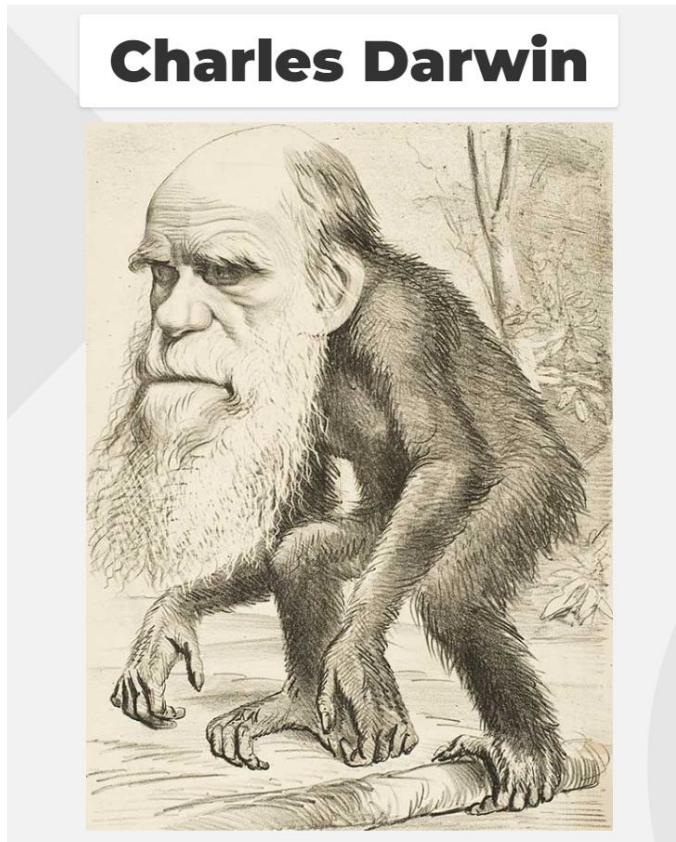
First principle in this course:

“If you are a real
seeker of truth, it is
necessary that at least
once in your life you
doubt, as far as
possible, all things.”

Renes Descartes



FYI, the fact is that:



cally researched emotions was **Charles Darwin**.
Man and Animals", from the biological aspects. He

Charles Darwin (1809-1882) is **NOT** the first scientist who studied emotions!

FYI, the fact is that:



Descartes starts his *Passions of the Soul* (1649)

“I shall be obliged to write just as if I were considering a topic that no one had dealt with before me”

The screenshot shows the Stanford Encyclopedia of Philosophy website. The header includes the logo, 'Browse', 'About', 'Support SEP', and a search bar. A sidebar on the left provides links for the current entry: 'Back to Entry', 'Entry Contents', 'Entry Bibliography', 'Academic Tools', 'Friends PDF Preview', and 'Author and Citation Info'. Below the sidebar, the main content area is titled 'Supplement to 17th and 18th Century Theories of Emotions' and 'Descartes on the Emotions'. It lists ten entries: 1. Introduction, 2. The Project of Theodicy and the Function of the Passions, 3. Passions, Medicine and Freedom, 4. The Passions and the Nature of Mind-Body Union, 5. The Definition of the Passions of the Soul, 6. The Classification of the Passions, 7. Wonder and Generosity, 8. The Remedy for the Passions, 9. The Influence of the Imagination, and 10. Influences on Later Authors.

More than 100 years earlier

Overview of the Course

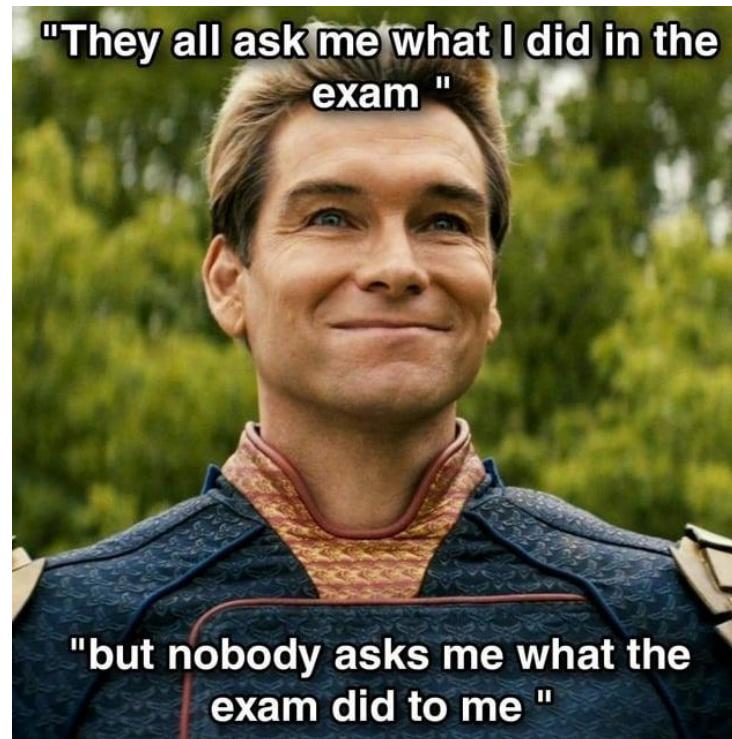
#	Time	Location	Topic	Lecturer
0.	Wed 03.09.2025 10:15-12:00		Skipped	
1.	Wed 10.09.2025 10:15-12:00	L6 Pohjola-sali+Online	Introduction to Affective Computing	Haoyu Chen
2.	Wed 17.09.2025 10:15-12:00	L6 Pohjola-sali+Online	Emotion Theories	Haoyu Chen
3.	Wed 24.09.2025 10:15-12:00	L6 Pohjola-sali+Online	Facial & Body Expression Analysis	Yante Li
4.	Wed 01.10.2025 10:15-12:00	L6 Pohjola-sali+Online	Emotion Recognition from Speech	Haoyu Chen
5.	Wed 08.10.2025 10:15-12:00	L6 Pohjola-sali+Online	Multimodal Learning for Emotion Analysis	Puneet Kumar
6.	Wed 15.10.2025 10:15-12:00	L6 Pohjola-sali+Online	Emotional Behavior Synthesis	Haoyu Chen
7.	Fri 24.10.2025 14:15-16:00	L6 Pohjola-sali+Online	Student final project presentation	

Attending the lecture is NOT mandatory!

How to pass the course?



No exam!

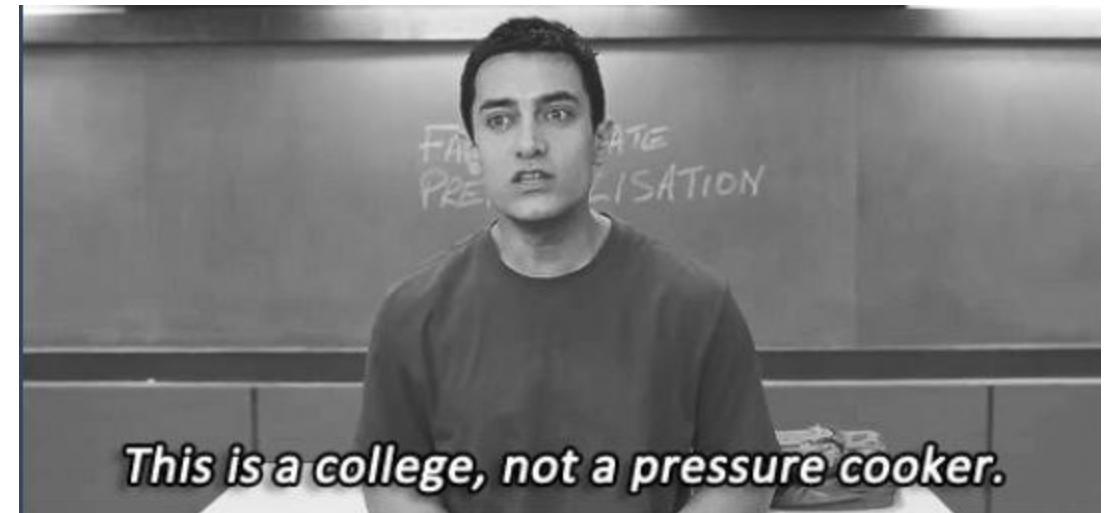


No exam!

DID YOU GO TO COLLEGE?

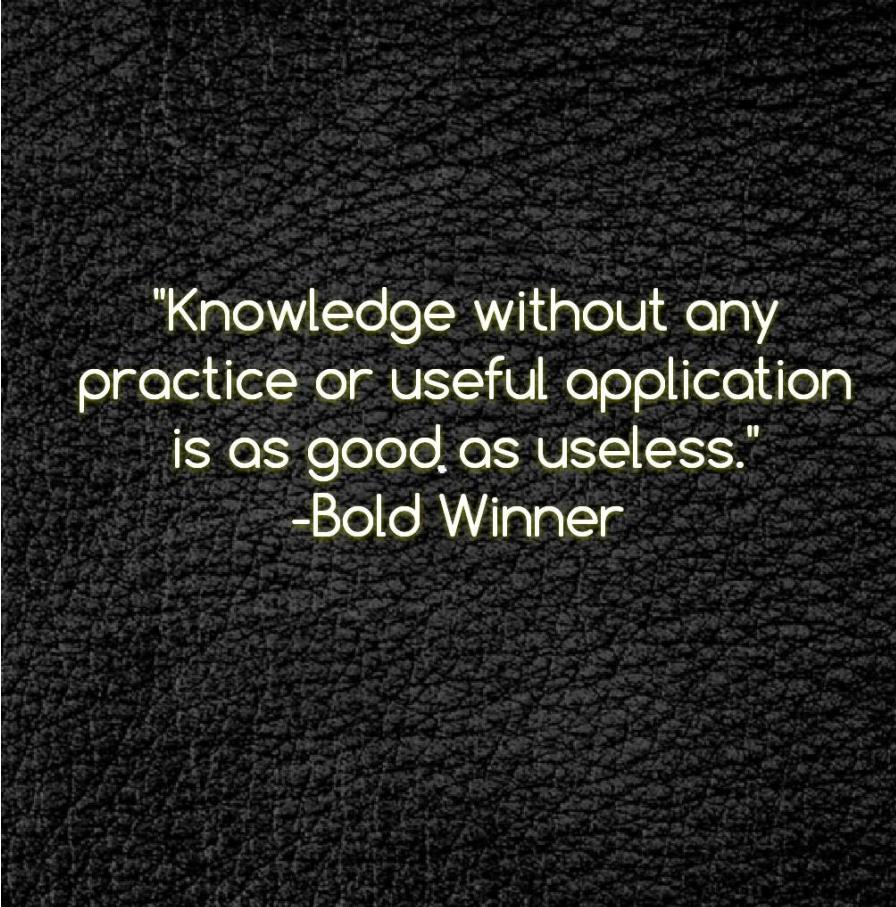


ChatGPT is changing the game



Free to think

The world is transforming



"Knowledge without any
practice or useful application
is as good as useless."
-Bold Winner

Grading

Assessment:

- 1. Three exercises: 20 points (min 12 points)
- 2. Final project: 50 points (min 30 points)
- 3. Final report: 30 points (min 18 points)
- 100 points in total

You need to finish all these three parts to pass the course

1. Programming exercises

We will introduce all the exercises in a brief mini-lecture on-site on Friday 12.09.2025 12:15-14:00 L8.

Release of the exercises

#	Release Time	Topic
1.	12.09.2025	Exercise 1 - Data annotation
2.	19.09.2025	Exercise 2 - Facial Expression Analysis
3.	26.09.2025	Exercise 3 - Multimodal Learning for Emotion Analysis

Deadlines

#	Deadline
Exercise 1	19.09.2025
Exercise 2	26.09.2025
Exercise 3	10.10.2025

2. Final project

1. Make a team of 6 students.

2. Select a project/research topic.

3. Give presentation on the final lecture.

4. Students directly mark your score, we get things serious!!!

Deadlines

#

Exercise 1

Deadline

19.09.2025

Exercise 2

26.09.2025

Exercise 3

10.10.2025

Project/research plan

23.09.2025

Project/research presentation

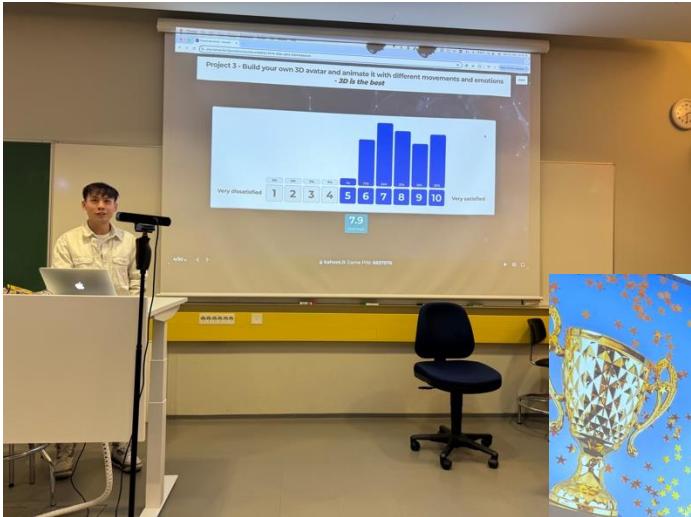
24.10.2025

Project/research report (final)

31.10.2025

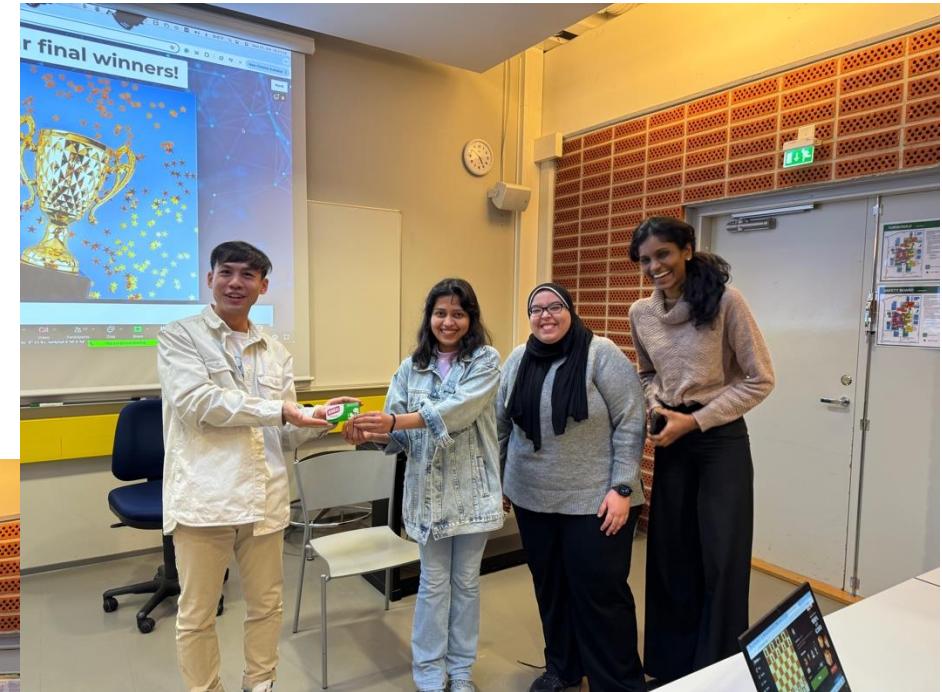
More instructions on Moodle

2. Final project



More instructions on Moodle

Haoyu Chen



Imagine you
are working on
a startup!

3. Final project report

1. One team one report.
2. No plagiarism, no abuse of ChatGPT.
3. Research report plan (one page, free form) by 24 September 2025.
4. Research report by 31 October 2025.

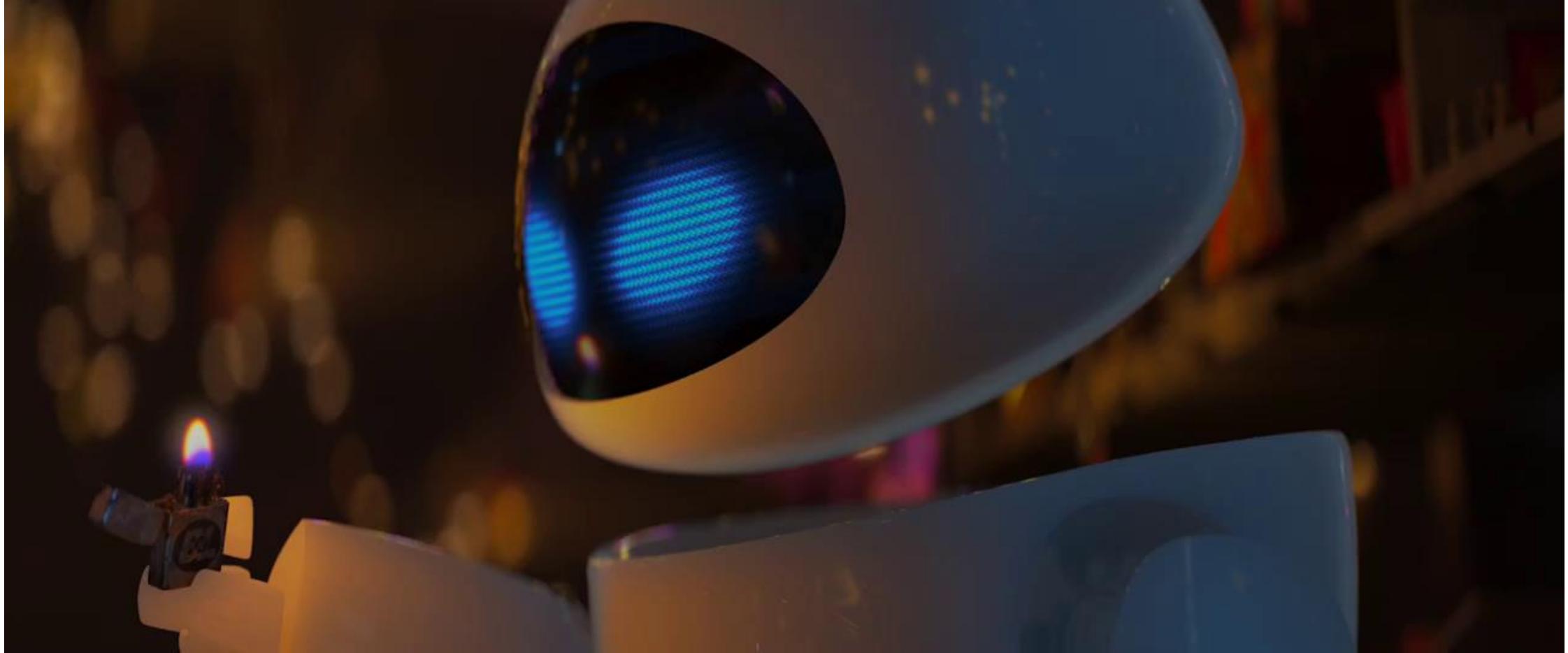
Important deadline

Deadlines

#	Deadline
Exercise 1	19.09.2025
Exercise 2	26.09.2025
Exercise 3	10.10.2025
Project/research plan	23.09.2025
Project/research presentation	24.10.2025
Project/research report (final)	31.10.2025

Now let's start the ~~lecture~~
discussion!

No golden answer, we will just show how complicated the topic is.



Affective Computing ultimate goal: Machines with Emotional Intelligence



Where the idea of
affective computing
originated?

Pioneers

- Rosalind Picard introduced the term “**Affective Computing**” in 1995
 - Her book with that title published in 1997

Originally defined as

- “.....computing that relates to, arises from, and deliberately influences emotion”



Long and scattered..

Communities and associations

- **A community of practice:** HUMAINE Association, 2007
Represented by a professional society.
- Association for the Advancement of Affective Computing
A professional, world-wide association for researchers in Affective Computing, Emotions and Human-Machine Interaction
- International Conference: Affective Computing & Intelligent Interaction (ACII)
Bi-annual International Conference, started in 2005
- International Journal (started in 2010)



Affective computing is about:

- **An Interdisciplinary field of research:** computer science, psychology, cognitive science, physiology, neuroscience, etc.

It is the study and development of systems and devices that can ***recognize, interpret, process, and simulate*** human affects.

- – How affective sensing can inform machine-understanding of people
- – How affect influences human-computer and human-robot interaction and machine usability
- – How to make computers more human-like
- – The ethics of “giving” machines emotional capabilities

Covers but is not limited to the topics involving:

- – Psychology and behavior as they relate to affective computing
- – Behavior generation and user interaction

The first work proposes to
study emotions
in a computational way

First principle in this course:

“If you are a real
seeker of truth, it is
necessary that at least
once in your life you
doubt, as far as
possible, all things.”

Renes Descartes



“Doubt is
the origin of
wisdom”

Rene Descartes

Do emotions really exist?/what it is?

Definition (maybe no)

Can emotion be computed?

Implementation (maybe no)

Why emotions are important for CS?

Motivation (maybe yes)

Researchers are doing



Motivations (why emotions are important)

“Emotions are fine for art, entertainment and certain social interactions, but keep them out of science and computing” --Affective Computing By Rosalind W. Picard



Motivations (why emotions are important)

“Emotion are fine for art, entertainment and certain social interactions, but keep them out of science and computing” --Affective Computing By Rosalind W. Picard

- Research shows that human intelligence is not independent of emotion. **Emotion and cognitive functions are un-separately integrated** into the human brain.

Norman, D.A. (1981). ‘Twelve issues for cognitive science’ Picard, R., & Klein, J. (2002). Computers that recognize and respond to user emotion: Theoretical and practical implications.

Taleb, T.; Bottazzi, D.; Nasser, N.; , "A Novel Middleware Solution to Improve Ubiquitous Healthcare Systems Aided by Affective Information,"

Motivations (why emotions are important)

Emotions affect decision making!

It's in our every day lifes.

Imagine without emotions,
where you will be now?



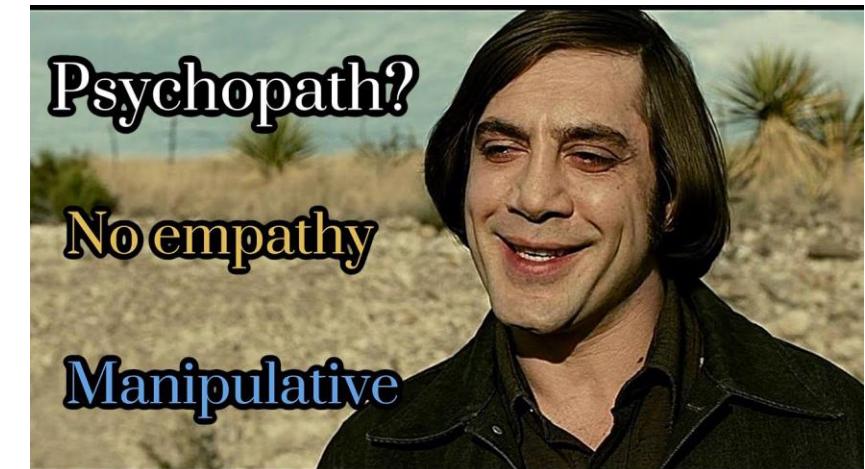
Goals

Automatic assessment of human emotional/affective state.

Emotion is an **internal psychological experience**,

which is not visible!

While facial expression is an **external physical manifestation** of that emotion.

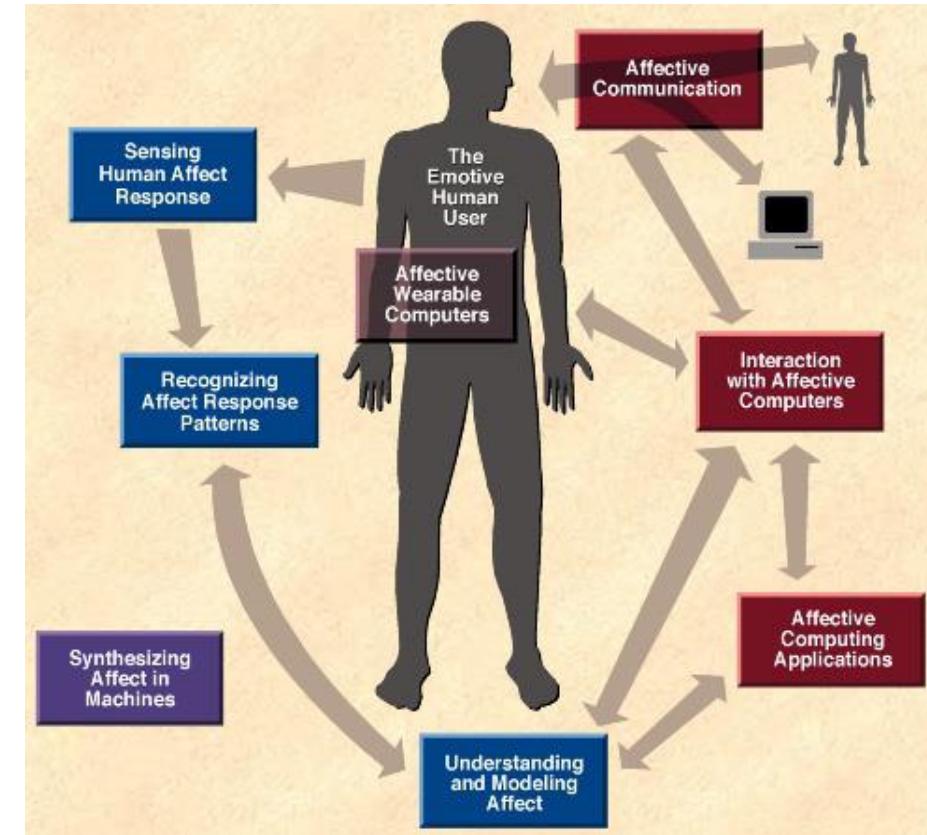


Norman, D.A. (1981). 'Twelve issues for cognitive science' Picard, R., & Klein, J. (2002). Computers that recognize and respond to user emotion: Theoretical and practical implications.

Taleb, T.; Bottazzi, D.; Nasser, N.; , "A Novel Middleware Solution to Improve Ubiquitous Healthcare Systems Aided by Affective Information,"

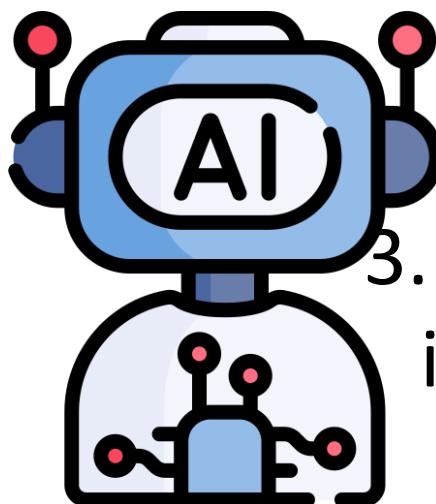
Research regions

Traditional division of research regions of affective computing



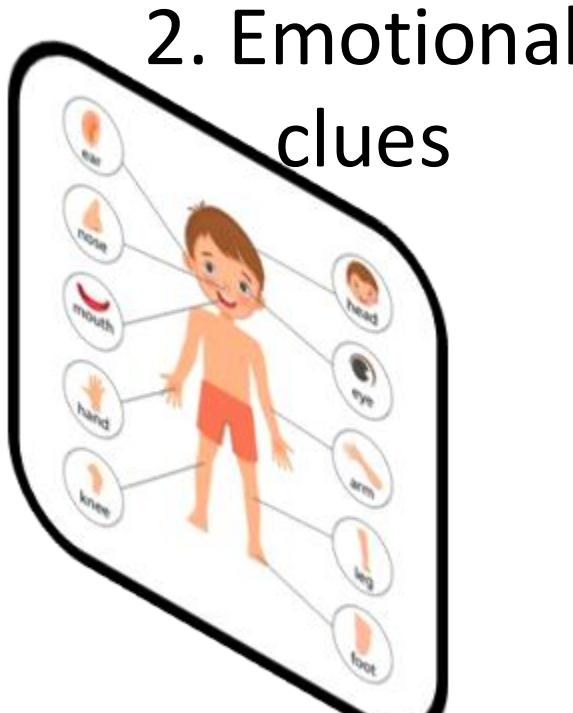
The research areas of affective computing visualized by MIT (2001)

A simplified model



Machines

3. Emotional
inference



2. Emotional
clues

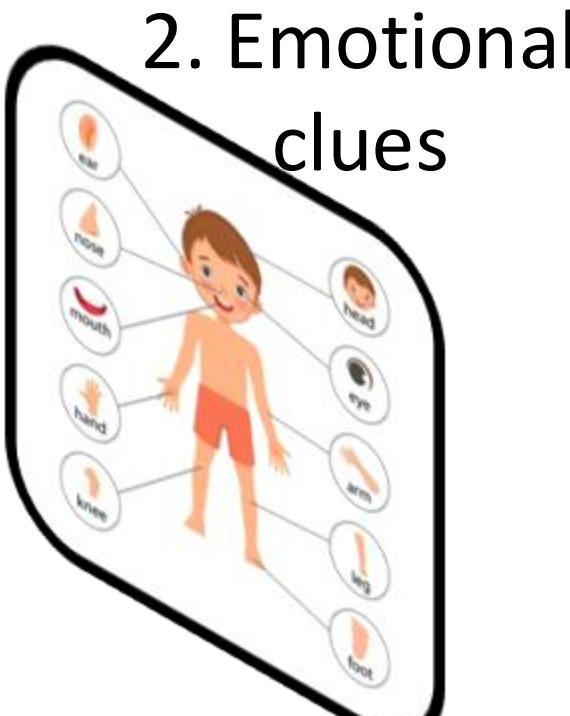
An unknow
blackbox room

1. Emotion
&cognition



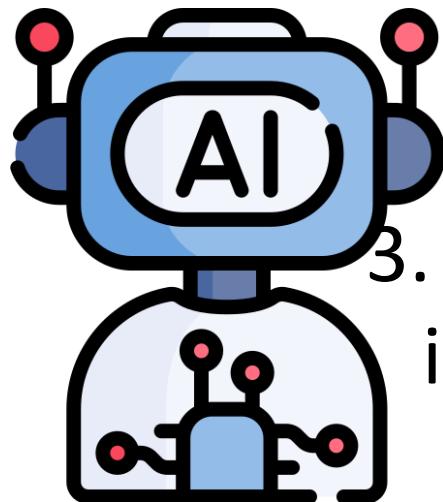
A simplified model

How to model inside activity



2. Emotional
clues

Machines



3. Emotional
inference

An unknow
blackbox room

1. Emotion
&cognition



How Emotion/Affection is Modeled?

Emotions are experiences

Emotions are ‘just’ hormones

Emotions are inherently socially constructed

...

<https://online.uwa.edu/news/emotional-psychology/>

“How Emotions Are Made” by Lisa Feldman-Barrett

Haoyu Chen

The mind takes a combination of
internal experiences (hormone levels, heart rate, etc)
and
external experiences (meaning drawn from sensory
information) at any given moment,
and uses
a word (mad, sad, happy, etc) **constructed** by our
brains to understand what's going on

Emotions are conceptualized
outcomes, violated by our
language/culture



How Emotion/Affection is Modeled?

Sisu in Finnish: be resilient/ambitious



热闹 (re nao) in Chinese: Fun and lively

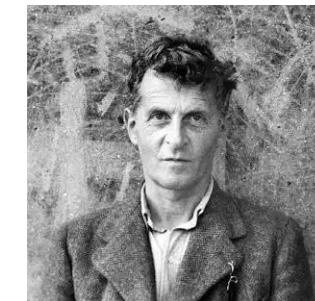


тоска in Russian: homesick or gloomy



Emotions are conceptualized outcomes, violated by our language/culture

정 in Korean: a deep love or affection that builds up as you get to know someone



<https://online.uwa.edu/news/emotional-psychology/>
How Emotions Are Made" by Lisa Feldman-Barrett

Philosophers like Ludwig Wittgenstein and Jacques Derrida have explored the limitations and power of language

How Emotion/Affection is Modeled?



In affective computing, affect is often seen as another kind of information:

- discrete or continuous units or states internal to an individual that can be transmitted in a loss-free manner from people to computational systems and back.

Affection description perspectives: how can we quantify emotion – from **Psychology**

- **Discrete Emotion Description**

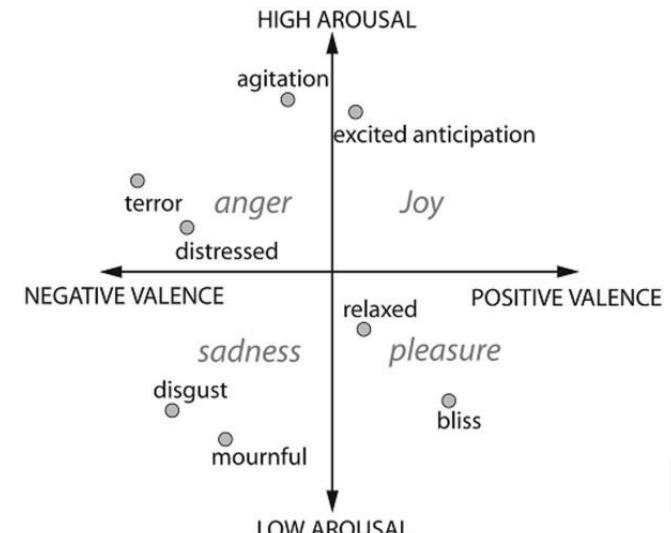
- Happiness, fear, sadness, surprise, disgust, anger



- **Dimensional Description**

- Valence: the pleasantness
- Arousal: the intensity of emotion that was evoked/INTRIGUED
- Dominance: the degree of control exerted

Very limited!
More in L2



How Emotion/Affection is Modeled?

Our Emotion Instruction Data Sample

Context type 1: Caption:
Two young women dancing in a kitchen.

Context type 2: Attribute:
Emotion: contentment Human action: stretching arm
Brightness: 0.7 Object: ['Cupboard', 'Food', 'Countertop']
Colorfulness: 0.4 Facial expression: happy

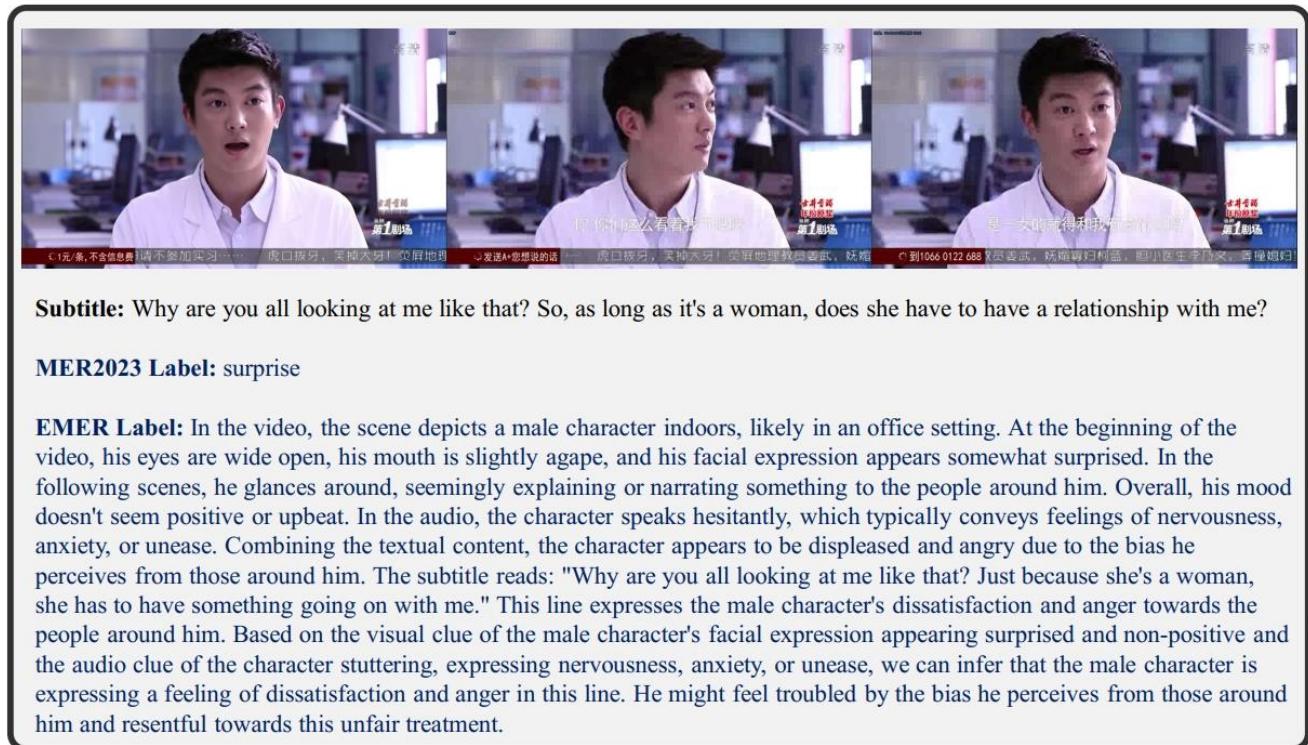
Categorical
Question: Please select the emotion closest to the image from the following options: amusement, anger, awe, contentment, disgust, excitement, fear and sadness.
Answer: Predict emotion: contentment.

Conversation Basic Interaction
Question: Can you describe the activity happening in the image?
Answer: The image depicts two young women enjoying themselves by dancing in a kitchen. One of the women is seen stretching her arm, most likely as part of their dance routine.

Conversation Advanced Interaction
Question: Based on the given emotion of contentment and the activities in the image, can you guess about the possible context or story of this scenario?
Answer: These women are good friends or roommates, spending quality time together. The act of dancing, especially in an unconventional place like a kitchen, suggests they are comfortable with each other and are enjoying a break from routine.

Reasoning
Question: Based on the setting and activities of the two young women, what might have prompted this impromptu dance session in the middle of a kitchen?
Answer: The impromptu dance could have been prompted by a favorite song playing, or simple spontaneous fun amidst preparing a meal or organizing the kitchen.

<https://github.com/aimmemotion/EmoVIT>



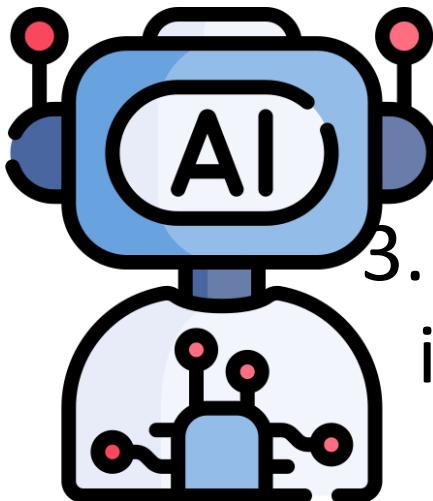
<https://github.com/zeroQiaoba/AffectGPT>

Large language Model (e.g., ChatGPT) to reason the emotions

A simplified model

What if the little brain give fake instructions to body?

Machines



3. Emotional inference



2. Emotional clues

An unknow room

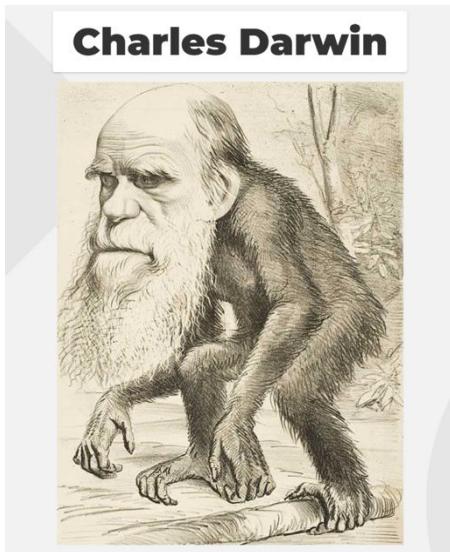
1. Emotion &cognition



A simplified model

What if the little brain give fake instructions to body?

What is reliable emotional clues?



Early research on emotions was Charles Darwin.
in his book "The Expression of the Emotions in Man and Animals", from the biological aspects. He

Emotions and facial expressions are hard-wired

Haoyu Chen

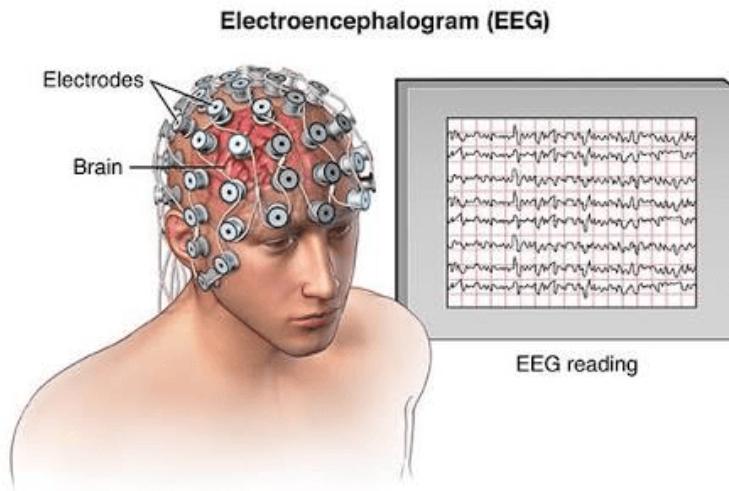


Facial expressions can be made up
but there are some true expressions.

A simplified model

What if the little brain give fake instructions to body?

What is reliable emotional clues?



EEG



MRI

A simplified model

What if the little brain give fake instructions to body?

What is reliable emotional clues?



Recent research topics (from TAC)

Emotion Recognition in:

- Speech:
 - Emotion in natural speech
 - Depression detection
- Face
 - Recognizing expressions
 - Intensity analysis
- Physiology
 - Inferring response to music via EEG
 - Detecting stress from skin conductance
- Text
 - Opinions in twitter; blogs
 - Emoticons

Synthesis

- Emotional speech
- Emotional facial expressions

Games/Entertainment computing

- Responses to victory and defeat
- Affective music player
- Boredom detection

Modeling

- Modeling emotional influences on decision-making
- Modeling factors that elicit emotions

Applications

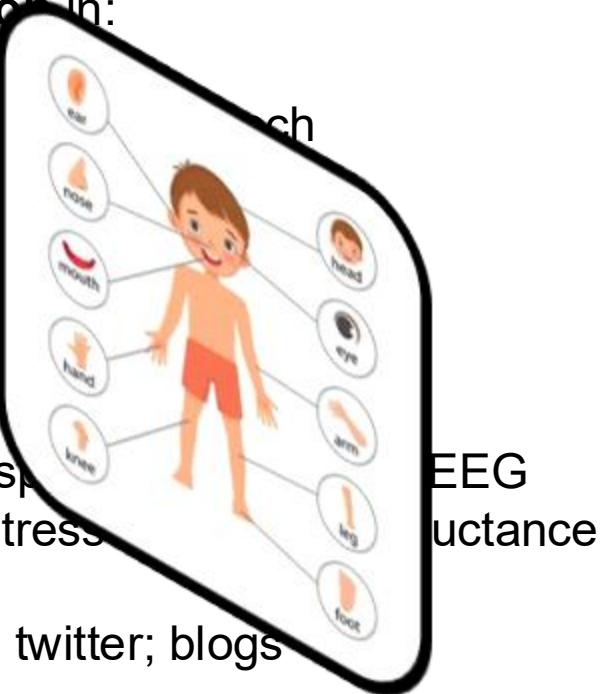
- Health: detection and shaping
- Games/entertainment: detection and shaping;
synthesis/realism
- Education: detections; shaping

Behavioral Science

Recent research topics (from TAC)

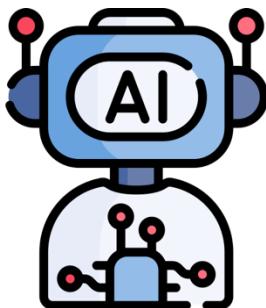
Emotion Recognition in:

- Speech:
Emotion in speech
Depression
- Face:
Recognizing Intensity and
Intensity and
- Physiology
Inferring responses
Detecting stress
- Text
Opinions in twitter; blogs
Emoticons



Synthesis

- Emotional speech
- Emotional facial expressions



Games/Entertainment computing

- Responses to victory and defeat
- Affective music player
- Boredom detection



Modeling

- Modeling emotions
- Modeling factors

Applications

- Health: detection and shaping
- Games/entertainment: detection and shaping; synthesis/realism
- Education: detections; shaping

Behavioral Science

Emotional computer

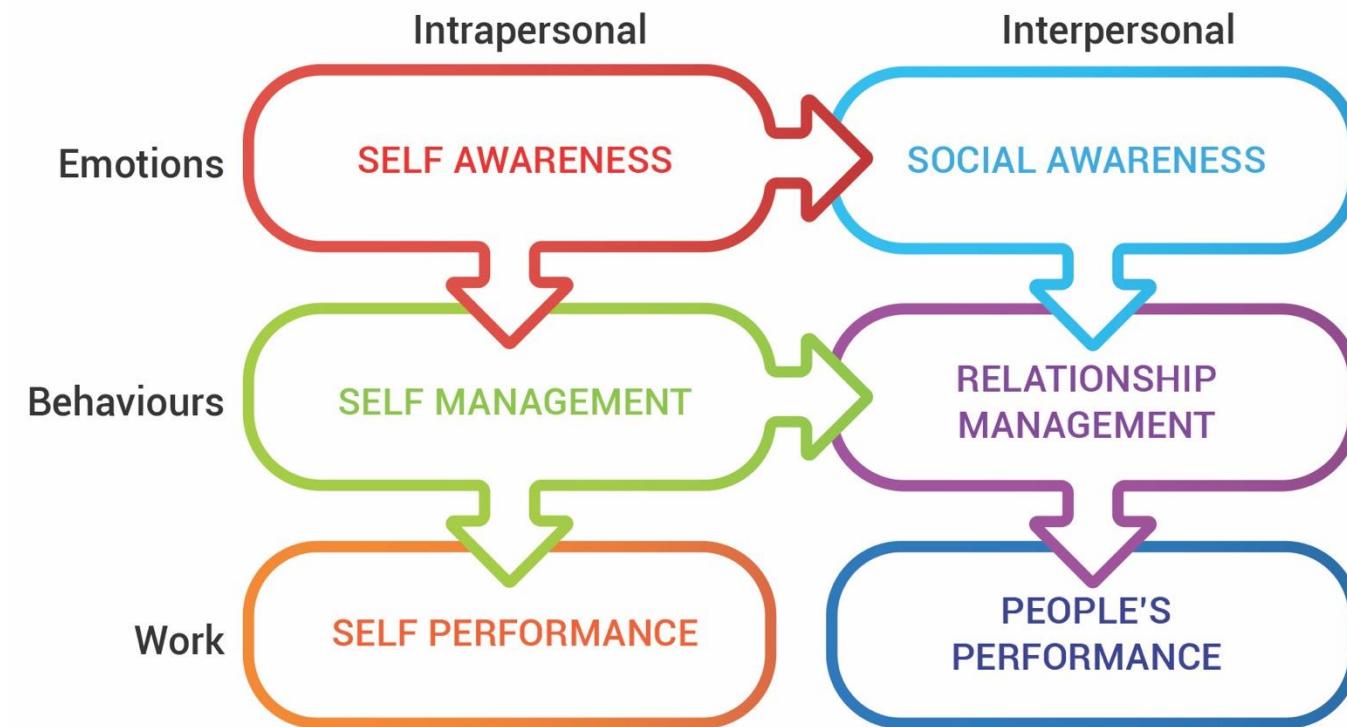


Including facial expression, speech and gesture analysis and facial expression and speech synthesis.

Computer recognizes emotions

Skills of Emotional Intelligence

- Aware of emotions
- Expressing emotions
- Regulating emotions
- Recognizing emotions
- Utilizing emotions



Copyright © EBW Global

Affection Detection and Recognition

Data sources:

- Speech/Vocal expression
- Facial Expression
- Bio-signals (Psychological sensors, Wearable sensors)
 - Brain Signal, skin temperature, blood pressure, heart rate, respiration rate
- Gesture
 - Limbic movements
- Text

Rafael A. Calvo, Sidney D'Mello, "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications"
Leon, E.; Clarke, G.; Sepulveda, F.; Callaghan, V., "Optimised attribute selection for emotion classification using physiological signals"

Affection Detection and Recognition

Modalities:

- Unimodal
- Multimodal
 - provide a more natural style for communication

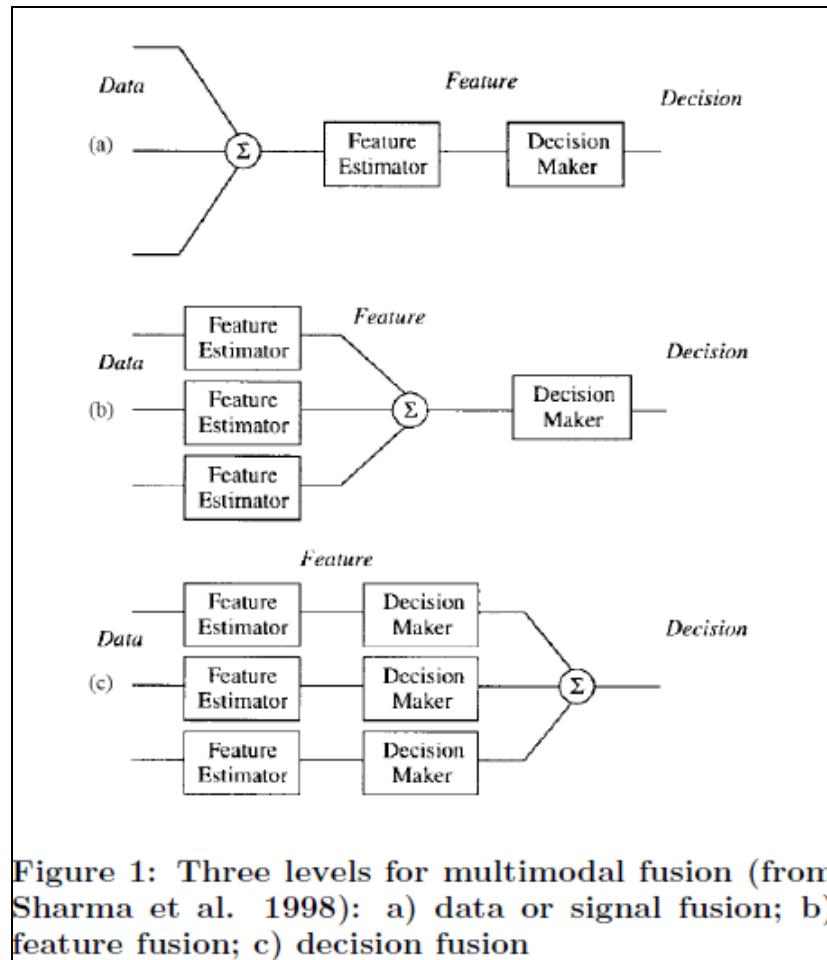
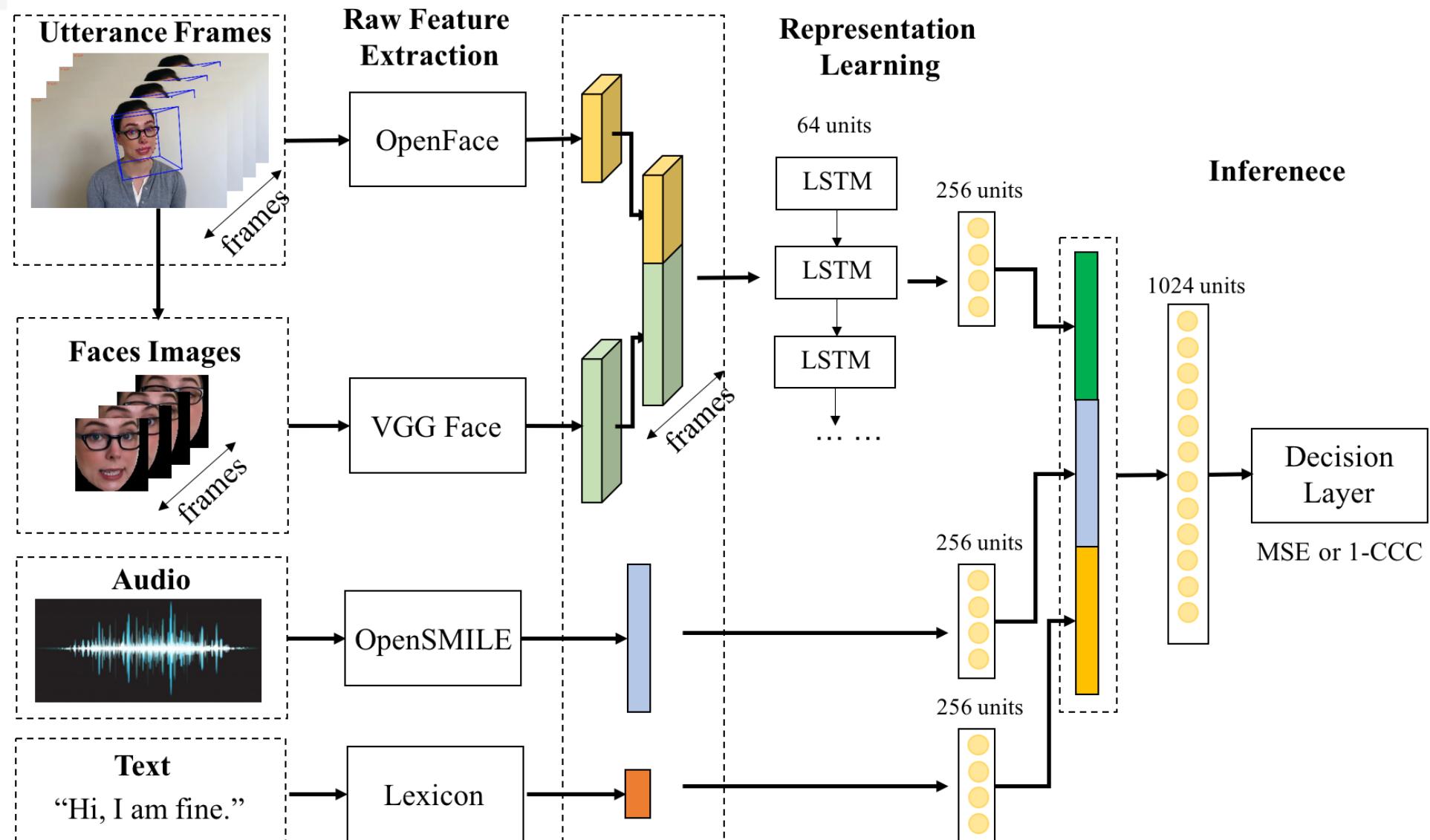


Figure 1: Three levels for multimodal fusion (from Sharma et al. 1998): a) data or signal fusion; b) feature fusion; c) decision fusion



Speech in affective computing

Speech in affective computing

Speech in one major way of expressing Affective messages: **verbal + vocal**

- - Lexical (explicit)
 - Affective words, stress, etc.
- - Paralinguistic (implicit)
vocal signals beyond the basic verbal message, convey meaning beyond the words and grammar used.
 - Speech prosody
 - voice quality
 - Spectral characteristics

Paralinguistic Features of Speech – *how is it said?*

- Prosodic features (e.g., **pitch-related feature, energy-related features, and speech rate**)
- Spectral features (e.g., MFCC - Mel-frequency cepstral coefficient and cepstral features)
- Spectral tilt, LFPC (Log Frequency Power Coefficients)
- F0 (fundamental frequency of speech), Long-term spectrum
- Speech disfluencies (e.g., filler and silence pauses)
- Context information (e.g., subject, gender, and turn-level features representing local and global aspects of the dialogue)
- Nonlinguistic vocalizations (e.g., laughs and cries, decode other affective signals such as stress, depression, boredom, and excitement)

Acted Corpora

- Word/content vs. Emotions:

happy 

sad 

angry 

confident

frustrated

friendly

interested 

anxious

bored 

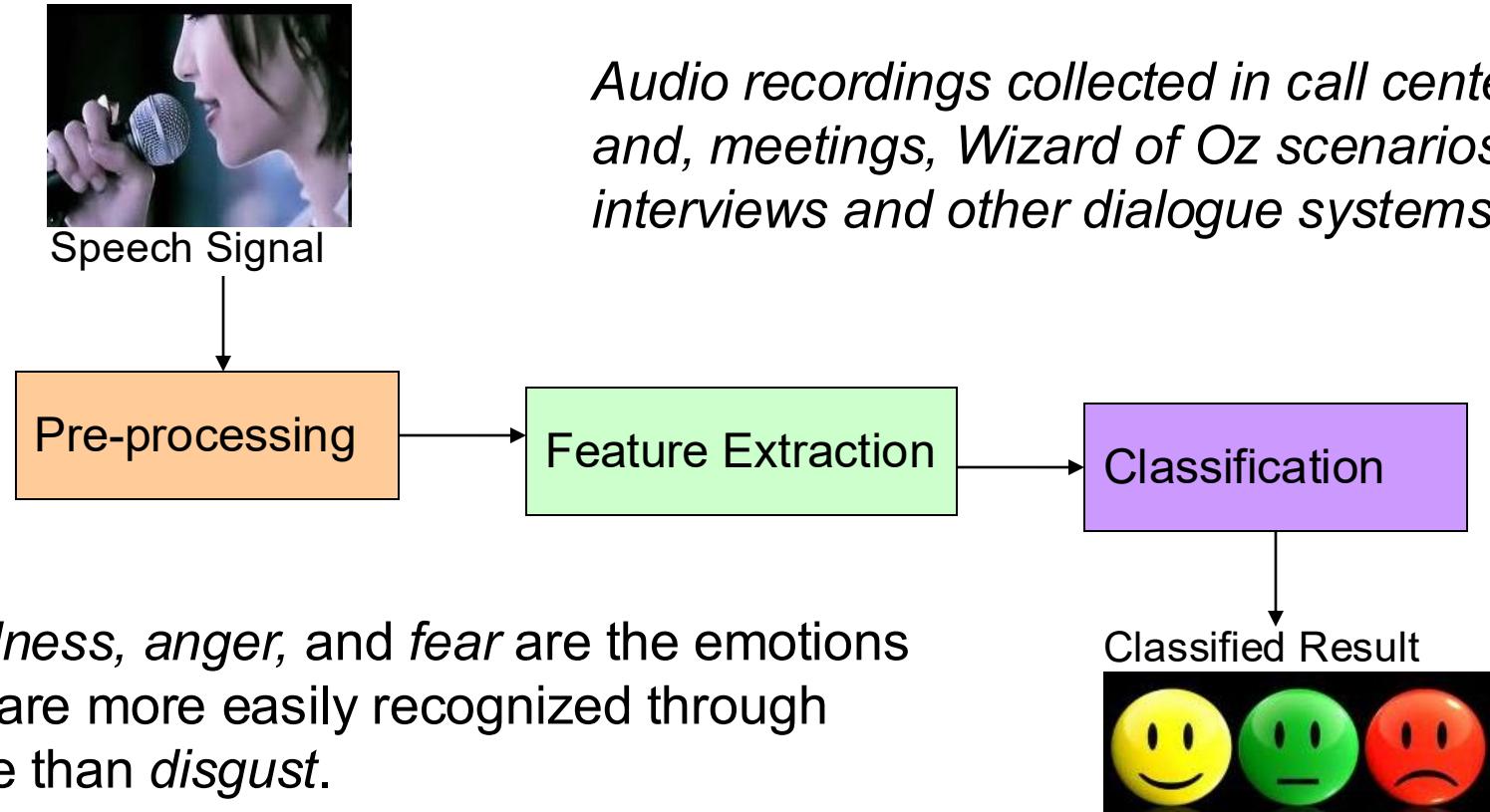
encouraging



Speech Recognition Architecture



Speech Signal



M. Pantic, N. Sebe, J. F. Cohn, and T. Huang. Affective multimodal human-computer interaction. In *ACM International Conference on Multimedia (MM)*, 2005.

Rafael A. Calvo, Sidney D'Mello, "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications

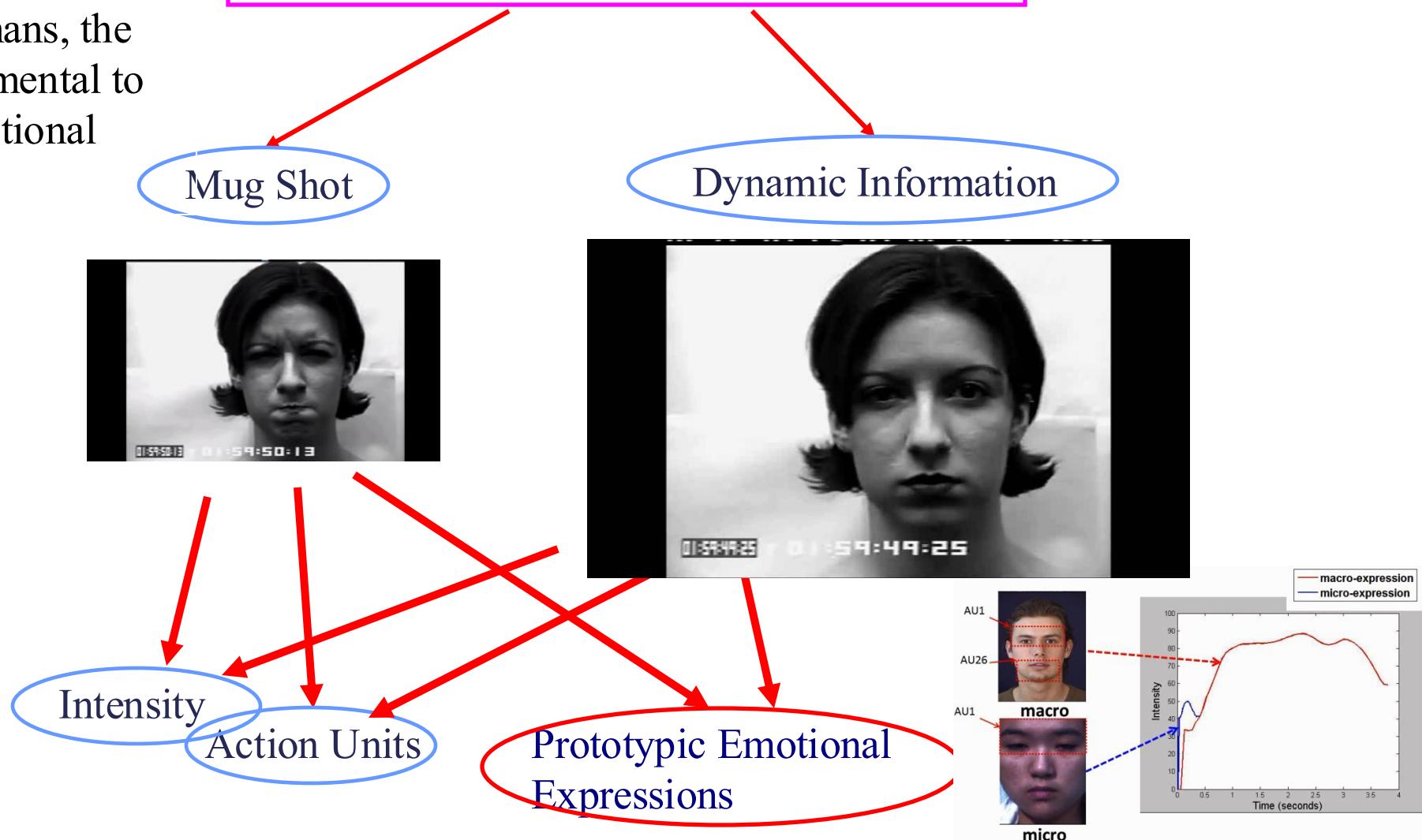
Facial expressions



Haoyu Chen

Psychological studies have suggested that for humans, the facial motion is fundamental to the recognition of emotional states.

Facial Expression Recognition

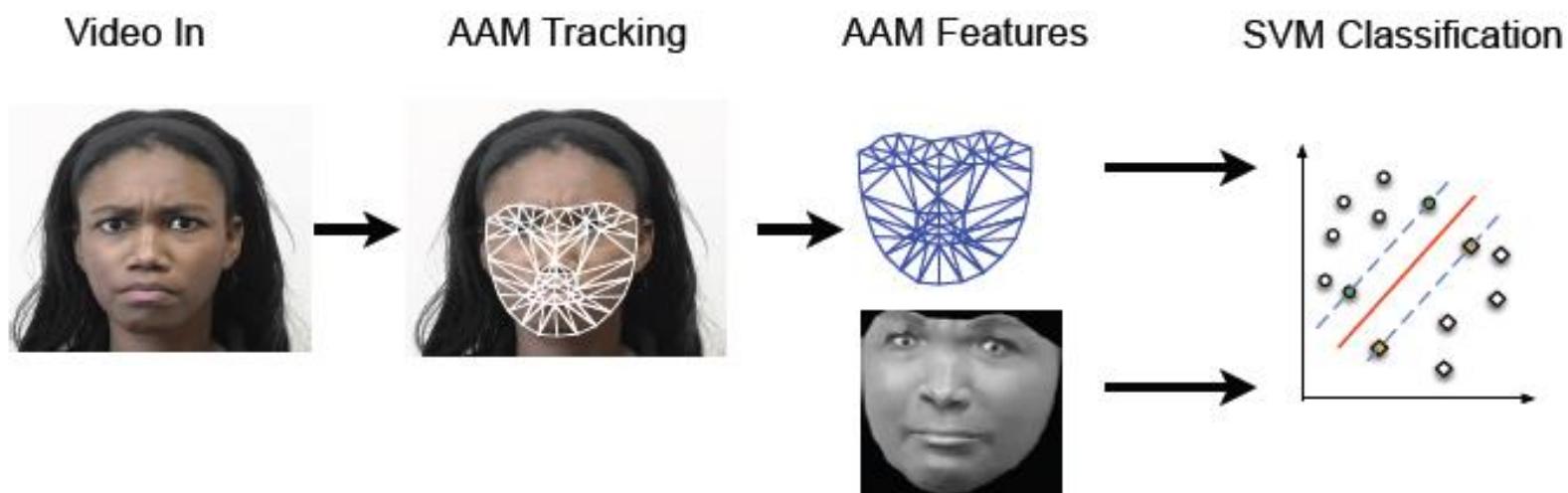


Facial features:

- **Appearance** features
 - Local binary pattern
 - Local binary pattern from three orthogonal planes
 - Gabor, HOG, SIFT(*Scale-invariant feature transform*), etc.
- **Shape** features
 - Distance between facial landmarks, angle, coordinates of facial landmarks, etc.
 - To locate facial landmarks using ASM, AAM, CLM, DRMF(*Discriminative Response Map Fitting*), etc.
- Deep learning features
CNN, RNN.....

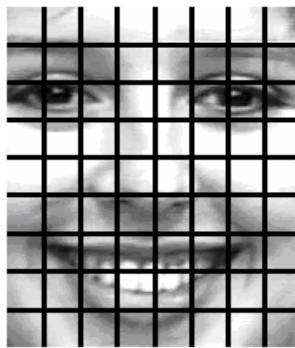
Example: Active Appearance Model (AAM)

(AAM) based system which uses AAMs to detect and track the face and extract visual features.

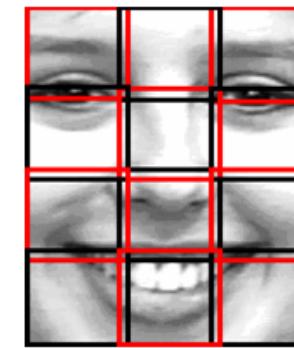


Appearance feature? Shape feature?

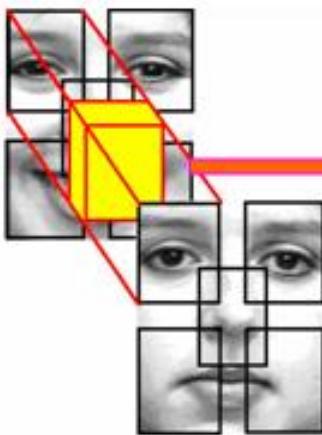
Example: LBP-TOP



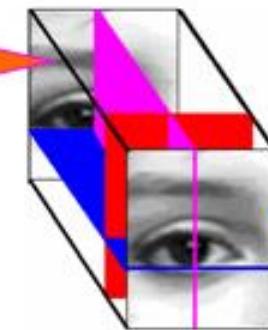
(a) Non-overlapping blocks(9 x 8)



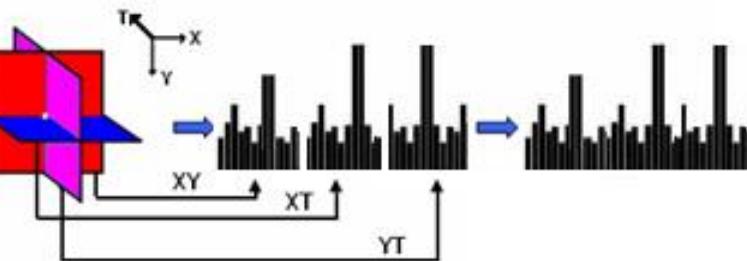
(b) Overlapping blocks (4 x 3, overlap size = 10)



(a) Block volumes



(b) LBP features
from three orthogonal planes



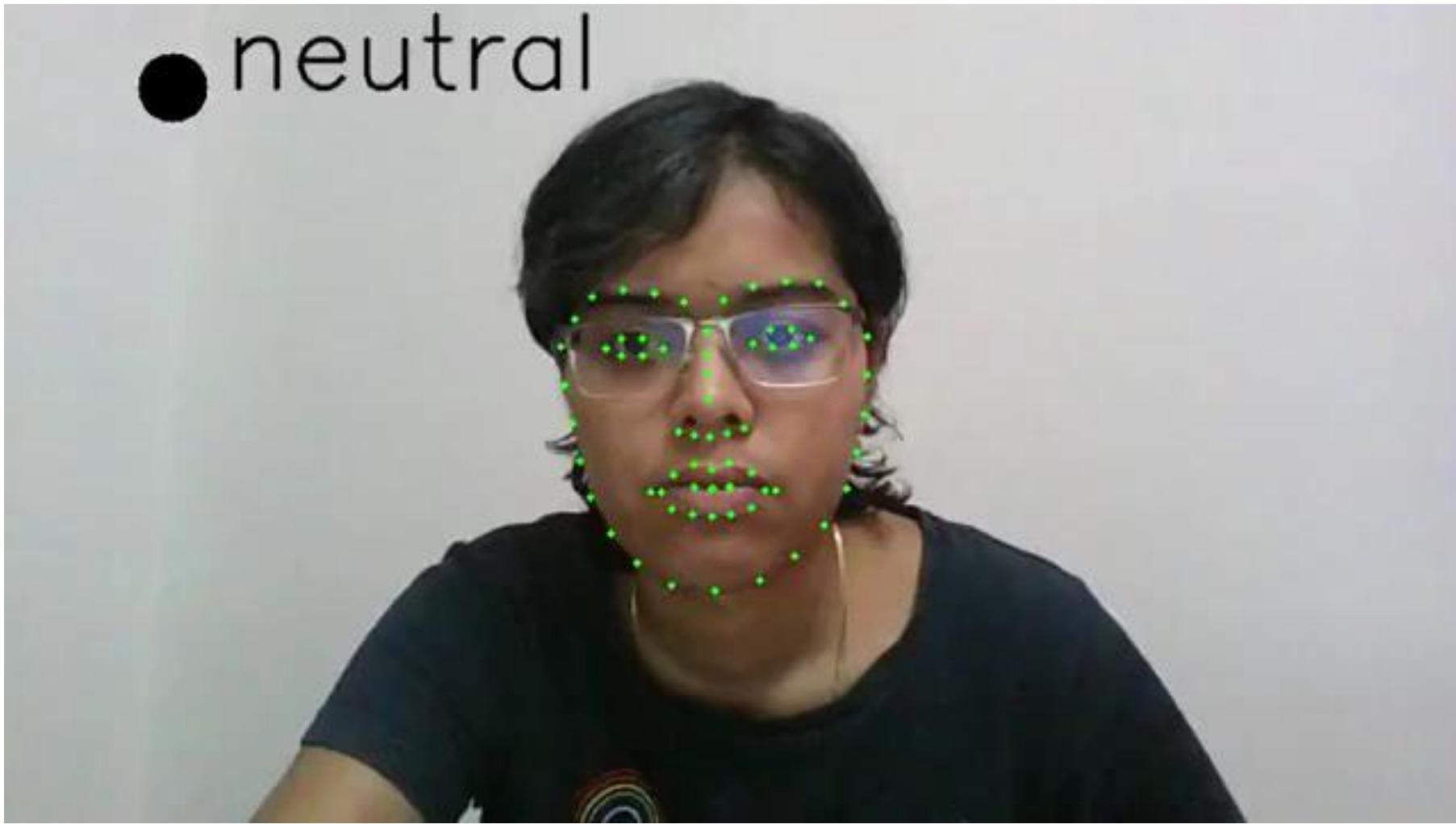
(c) Concatenated features for one block volume
with the appearance and motion

Features in each block volume

CMVS research center is famous in the field



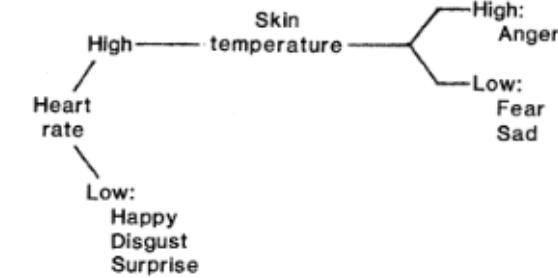
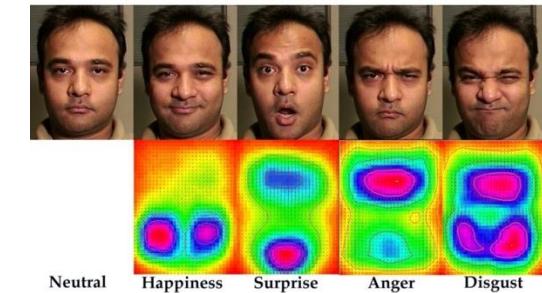
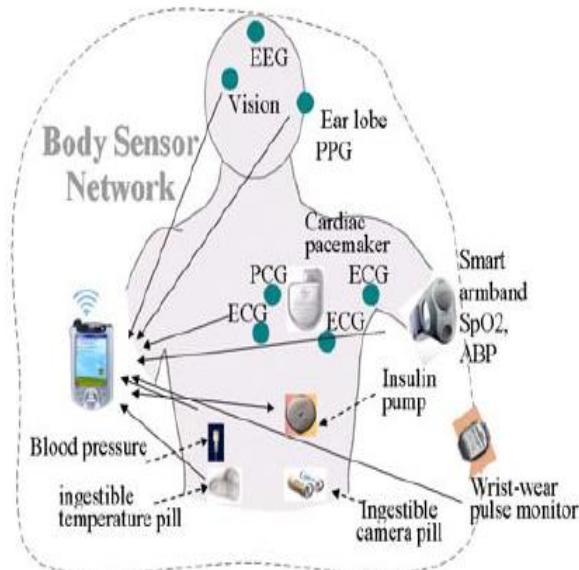
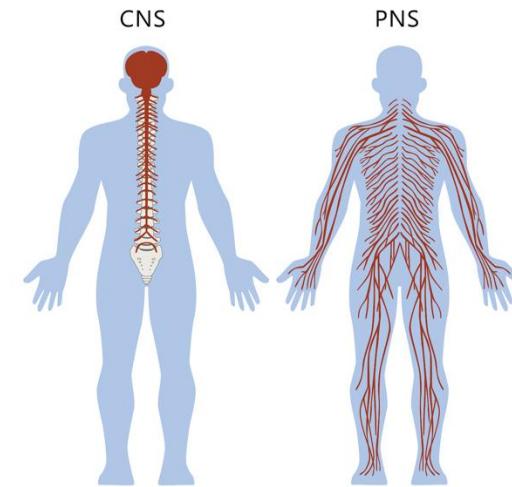
- neutral



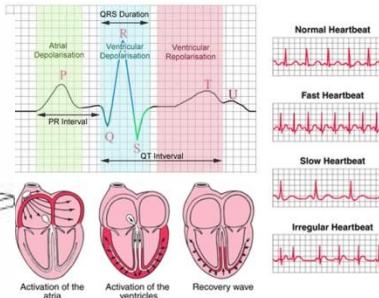
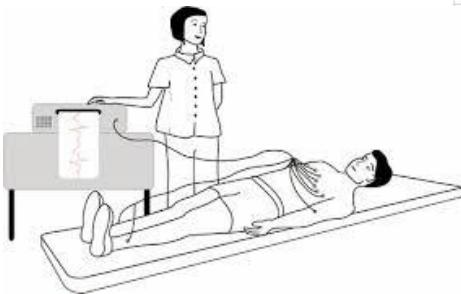
Physiological Signals / Bio-Signals

Physiological Signals / Bio-Signals

- Physiological signals derived from e.g., *Central Nervous System* (CNS) and *Peripheral Nervous System* (PNS) of human body.
 - Fear for example increases heartbeat and respiration rates, causes palm sweating, etc.
- Physiological information used:
 - GSR - Galvanic Skin Resistance
 - RESSP - Respiration
 - BVP - Blood volume pulse
 - Skin Temperature
 - Electrocardiography (ECG)
 - Electrodermal activity (EDA)
 - Electromyogram (EMG)
 - Electroencephalogram (EEG)



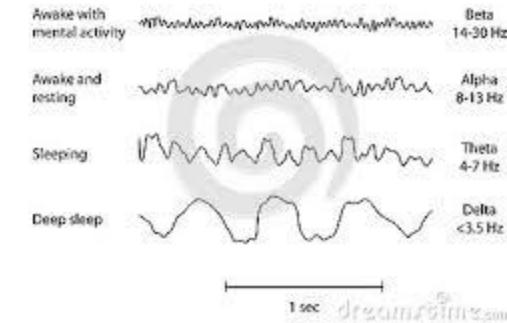
Electroencephalogram (EEG), Electrocardiography (ECG), Electrodermal activity (EDA), Electromyogram (EMG)



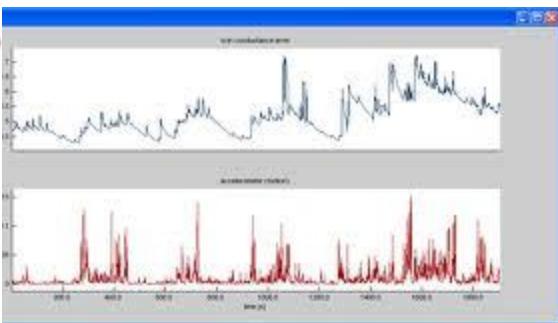
ECG: recording of the **electrical activity of the heart**



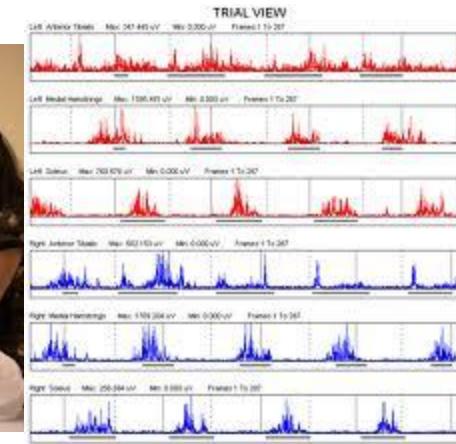
Normal Adult Brain Waves



EEG: recording of **electrical activity along the scalp**



EDA: to measure the electrical **conductance of skin**, which varies depending on the amount of sweat-induced moisture on the skin.

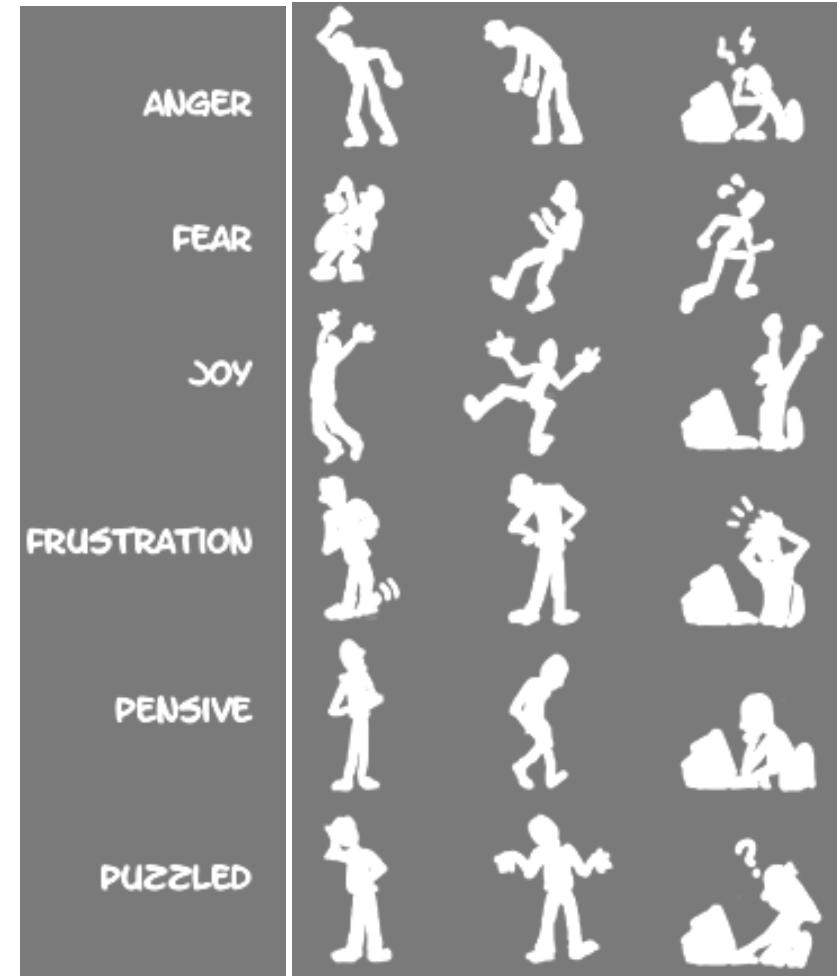


EMG: for evaluating and recording the **electrical activity produced by skeletal muscles**.

Gesture / Body Motion

- Pantic et al.'s survey shows that gesture and body motion information is an important modality for human affect recognition. Combination of face and gesture is more accurate than facial expression alone.
- Two categories of Body Motion based affect recognition
 - ▣ Stylized
 - The entirety of the movement encodes a particular emotion.
 - ▣ Non-stylized
 - More natural - knocking door, lifting hand, walking etc.

Body language: human vs. animals



Frequently used Modeling Techniques

- Fuzzy logic
- Neural Networks (NN)
- Hybrid: Fuzzy + NN
- Tree augmented Naïve Bayes
- Hidden Markov models (HMM)
- K-Nearest neighbors (KNN)
- Linear discriminant analysis (LDA)
- Support vector machines (SVM)
- Gaussian mixture models (GMM)
- Discriminant function analysis (DFA)
- Principal component analysis (PCA)
- Non-linear manifold modeling
- Deep learning
- Cross-validation, Structural risk minimization (SRM)

Computer expresses emotions (synthesis)

- Facial expression
- Speech

Facial expression synthesis

Is needed for:

- **Interact** with virtual agents: help desks, museum guides, online shopping, and distance learning
- We would like such virtual characters to behave **realistically**. This involves understanding the signals we use in non-verbal communication, and enabling the virtual character to display such signals.
- Such understanding can be achieved through the **analysis of data** collected from humans interacting with each other or with computers. This data can then be used to **teach** the virtual characters to express themselves in an appropriate manner.

FE synthesis demos (2012)



Query



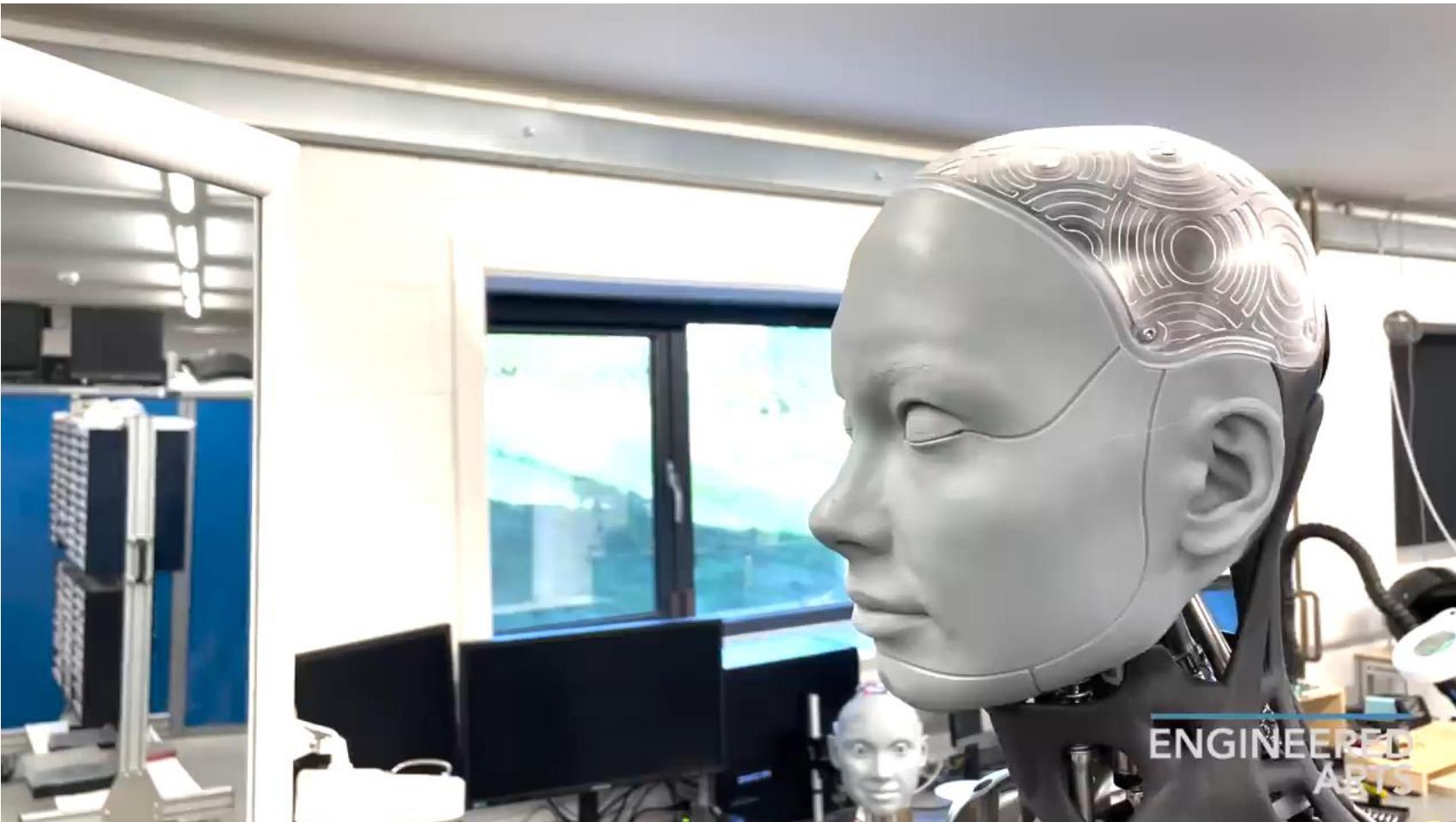
Final Result

FE synthesis demos (2022)



FE synthesis demos (2022)

Ameca Robot



Emotional speech synthesis

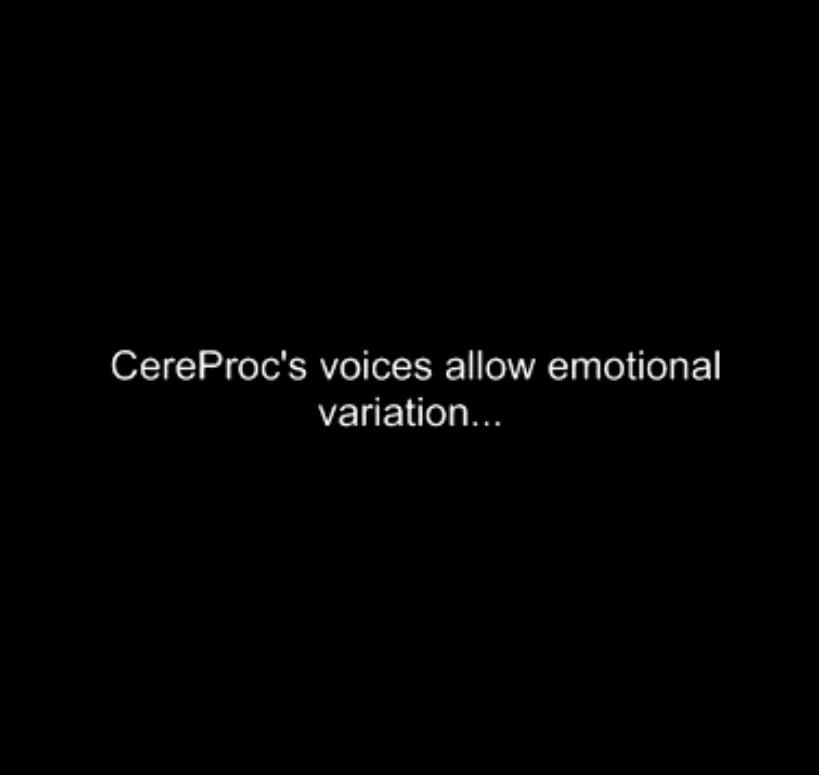
- The overall goal of the speech synthesis research community is to **create natural sounding synthetic speech**.

Synthesized speech benefits from emotions by *delivering certain content in the right emotion* (e.g., good news are delivered in a happy voice), therefore making the speech and the content more believable.

Emotions can make the interaction with the computer more natural because the system reacts in ways that **the user expects**.

E.g., audio books, narratives (audio track) for videos...

Emotional speech synthesis demos



CereProc's voices allow emotional variation...

Online/App: <https://app.typecast.ai/en/editor/character-casting?type=add>

Emotional speech synthesis demos



Veo3. Kling Al. Nano banana...

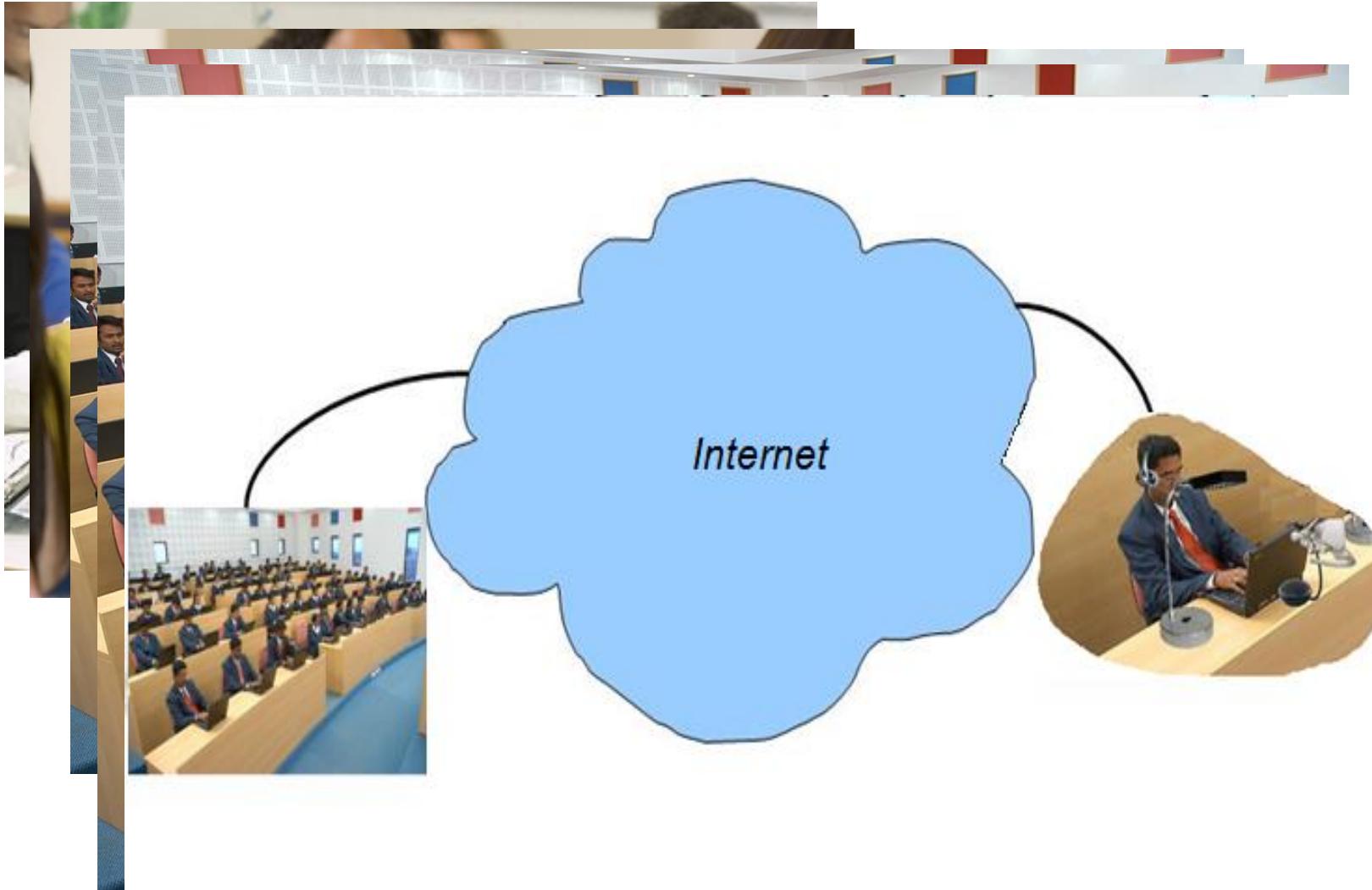


Advertisement for my AI restaurant

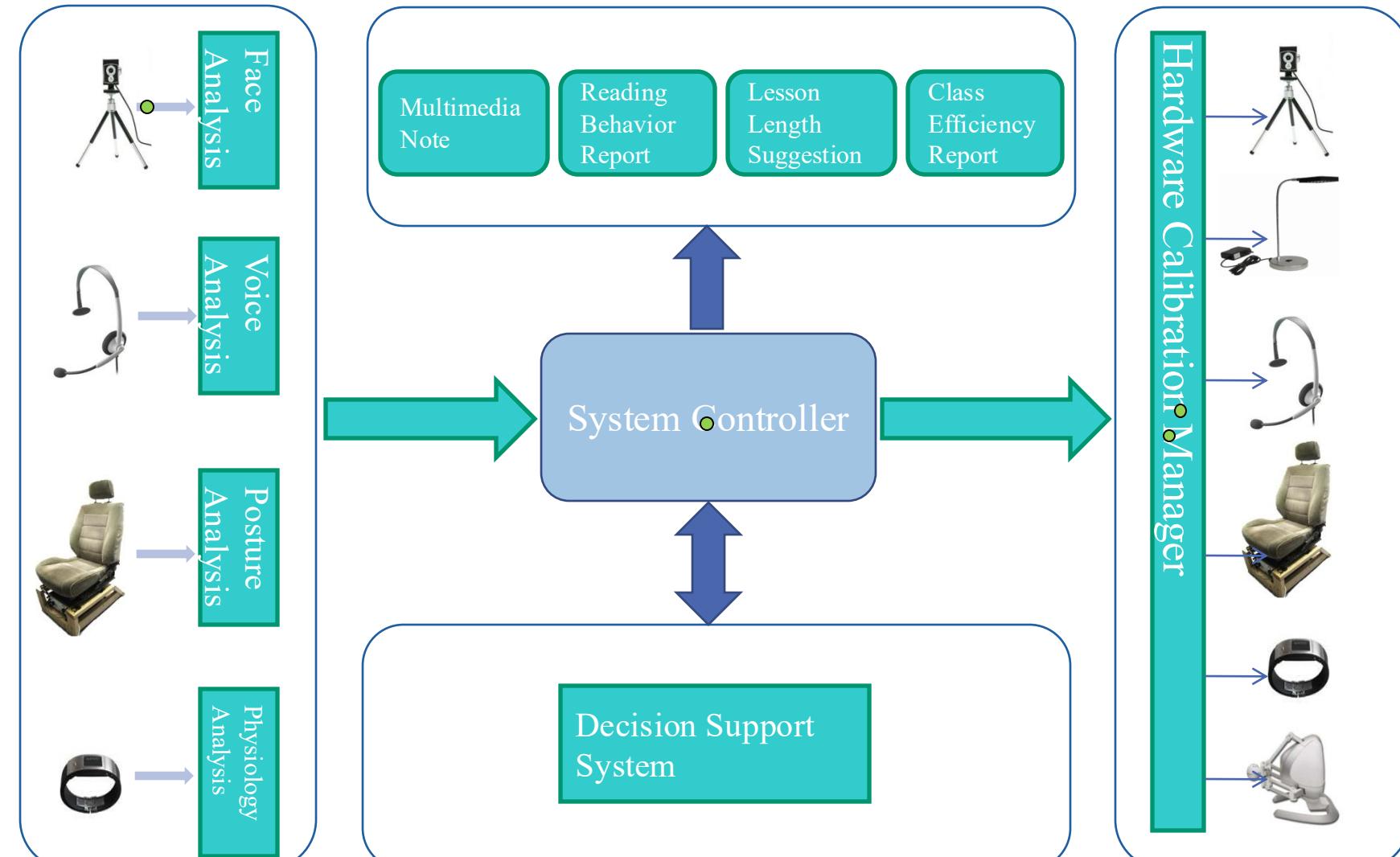
AF Applications:

- **Security sector:** affective behavioural cues play a crucial role in establishing or detracting from credibility
- **Medical sector:** affective behavioural cues are a direct means to identify patients' status.
- **Neurology/psychiatry:** in studies of mental dysfunctions or disorders.
- **Dialog/Automatic call center Environment:** to reduce user/customer frustration
- **Education** and learning.
- **Human Computer Interaction (HCI) / Human Robot Interaction (HRI)**
- **Game:** animated characters / improve user experiences

Case1: Context Aware Multimodal Affection Analysis Based Smart Learning Environment



Affective Computing



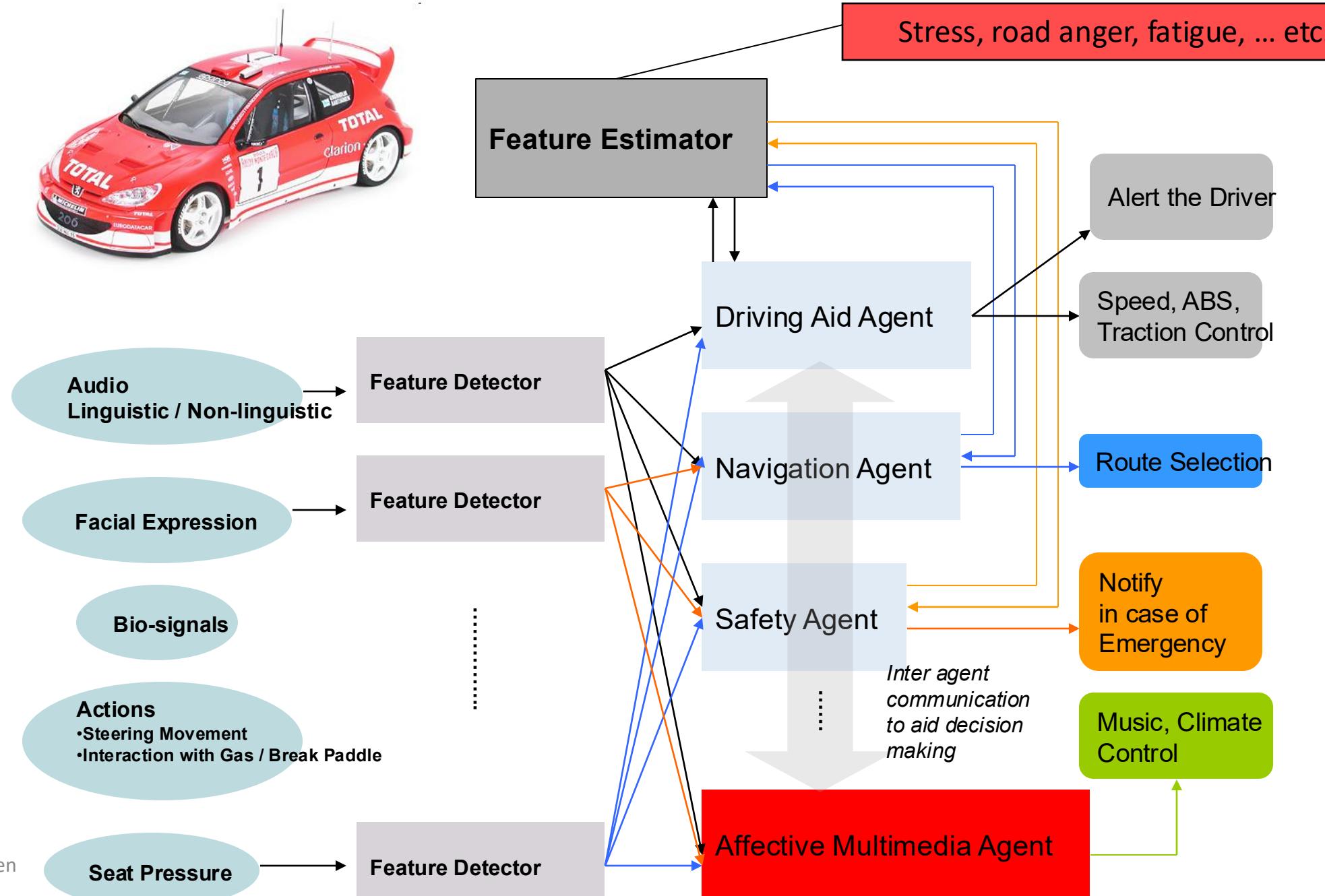
Multimodal Affect Input

Parameter Adjustment

Application System Architecture

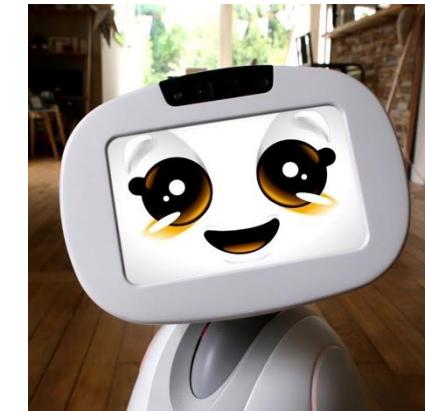
Case 2: Driver Emotion Aware Multiple Agent Controlled Automatic Vehicle

Affective Computing



Concerned Issues

- Privacy concerns
 - Data sharing vs. personal privacy
 - Measurement with/out noticing.
- Ethical concerns
 - Self-aware Robot/computer, interfering human emotions/decisions, erroneous consequences.
- Risk of misuse of the technology
 - In the hand of impostors for bad-illegal purposes
 - Misusage caused by mistakes
- Complex technology
 - Effectiveness is still questionable, risk of false interpretation



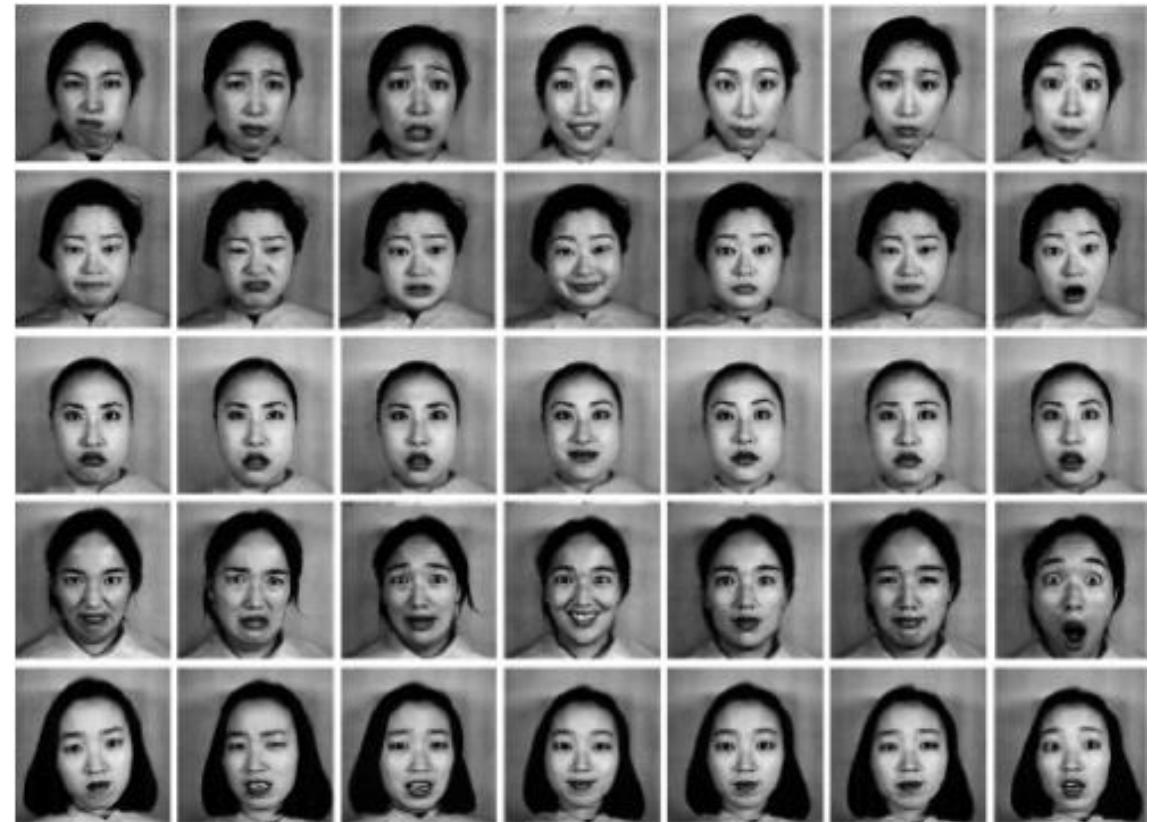
Future Research Directions

- So far **Context** has been overlooked in most Affection Computing researches
- Collaboration among Affection researchers from **different disciplines**
- Fast **real-time** processing
- **Multimodal** detection and recognition to achieve higher accuracy
- ...

Resources: databases & tools

Databases

- JAFFE (Easy)
 - Static facial expression images
 - 213 images of 7 facial expressions
 - 10 Japanese females
- Protocol
 - Leave-one-subject-out
 - Leave-one-sample-out



Databases

- Cohn-Kanade (Medium)
- Dynamic facial expression
 - 372 facial expression sequences
 - Six facial expressions (Anger, Disgust, Fear, Happiness, Sadness and Surprise)
- Protocol
 - Leave-one-subject-out



Databases

- MAHNOB (Hard)
- Spontaneous facial expression
 - **Multimodalities:** Video+audio+Bio
 - Nine emotion keywords
 - Valence and arousal mapping
 - Selected session: 267 samples from 14 subjects
- Protocol
 - Leave-one-subject-out
 - Valence & Arousal dimension



Arousal classes	Emotional keywords
Calm	sadness, disgust, neutral
Medium arousal	joy and happiness, amusement
Excited/Activated	surprise, fear, anger, anxiety
Valence classes	Emotional keywords
Unpleasant	fear, anger, disgust, sadness, anxiety
Neutral valence	surprise, neutral
Pleasant	joy and happiness, amusement

Databases

- AFEW (Hard)
- Dynamic temporal facial expression data
 - Close to the **real-world** condition
 - Seven facial expressions
 - Audio & Visual
 - More difficult
- Protocol
 - 10-fold-cross-validation



Databases

- Affectnet (Hard)
- Dynamic facial expression data
 - More than **one million** images from the Internet
 - 450,000 images have manually annotated labels for eight basic expressions
- Protocol
 - Test set: Annotated images
 - Validation set: Randomly selected 500 samples of each category
 - Training set: The rest of images

TABLE 11
Samples of Annotated Categories for Queried Emotion Terms

	Queried Expression						
Annotated Expression	Happy	Sad	Surprise	Fear	Disgust	Anger	Contempt
Neutral							
Happy							
Sad							
Surprise							
Fear							
Disgust							
Anger							
Contempt							
None							
Uncertain							

Databases

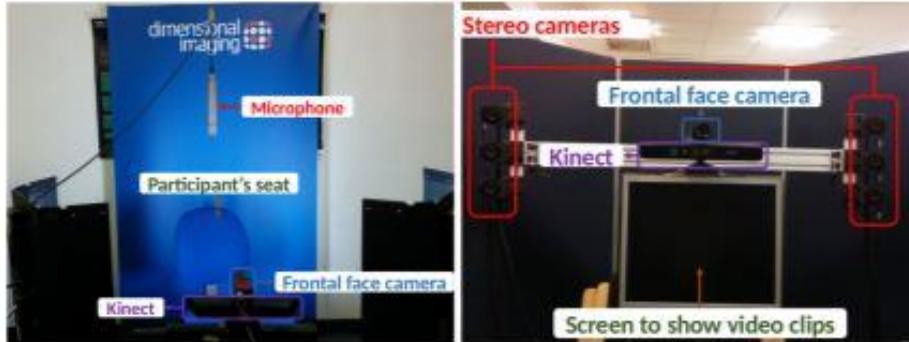
- RAF-DB (Hard)
- Dynamic facial expression data
 - 29672 images
 - Six basic expressions plus neutral and 12 compound expressions
- Protocol
 - 5-fold cross-validation
 - leave-one-subject-out



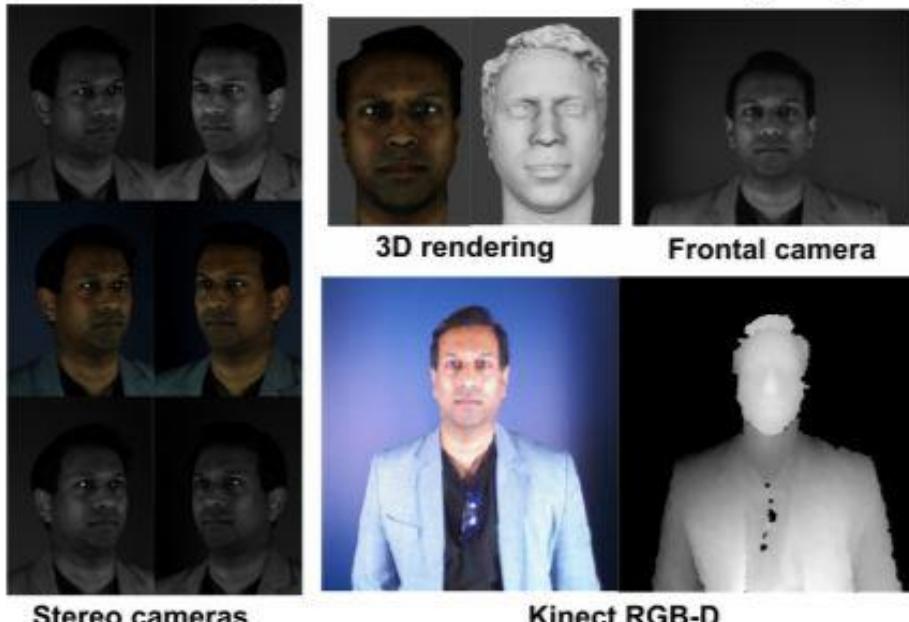
Fig. 5. Examples of six-class basic emotions and twelve-class compound emotions from RAF-DB. The detailed data proportion and class distribution of RAF-DB are attached to each expression class.

Databases

- 4DFAB (Hard)
- Dynamic facial expression data
 - 180 subjects
 - Seven basic expression
 - **1.8 million 3D faces**
- Protocol
 - 10-fold-cross-validation



(a) Frontal (right) and rear (left) view of recording setup.

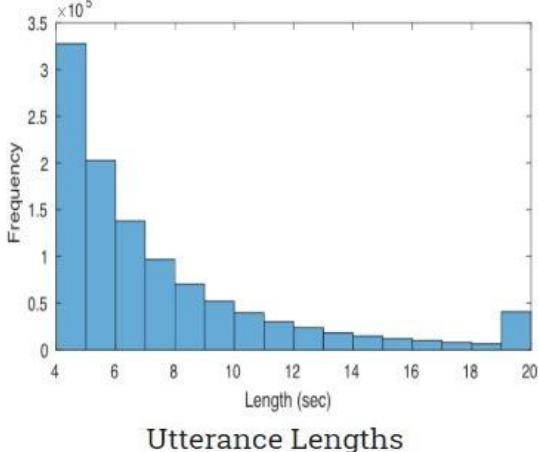


(b) Exemplar data from a single capture.

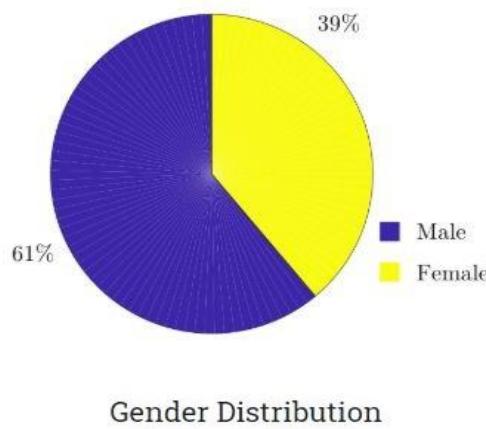
Figure 1: Overview of capturing system.

Databases

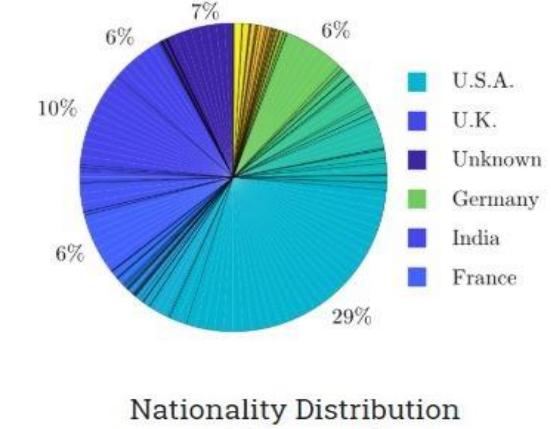
- VoxCeleb (Hard)
 - Dynamic temporal facial expression data
 - human speech extracted from interview videos
 - 7000+ speakers
 - 1 million + utterances
 - 2,000 + hours Each segment is at least 3 seconds long.
- Protocol
 - No emotion label
 - Unsupervised learning



Haoyu Chen



Gender Distribution



Nationality Distribution

Databases



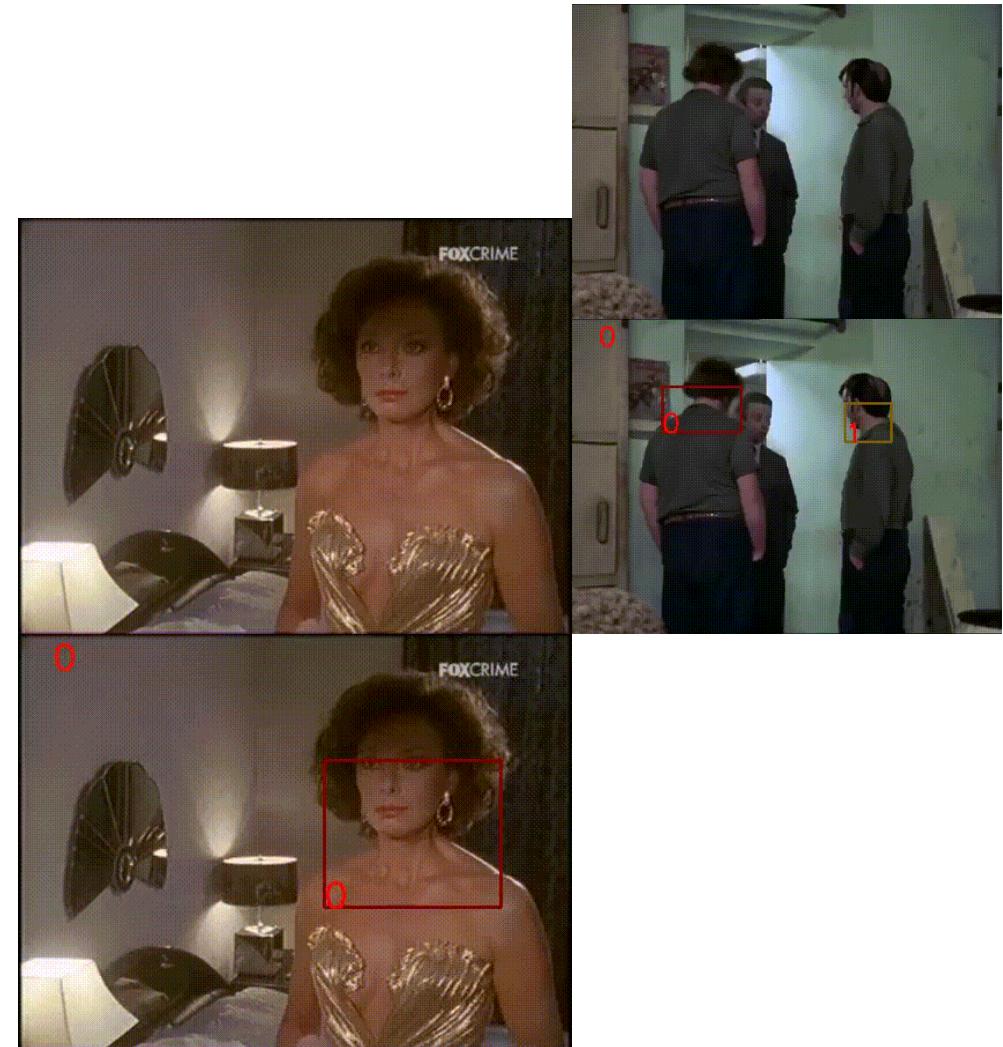
Large and **growing** dataset containing annotated short-video samples of **bodily expression** of emotions.

<https://cydar.ist.psu.edu/emotionchallenge/index.php>

9,827 video clips

13,239 instances, i.e., the person of interest in these clips.

Each instance is a data sample.



Available emotional speech databases

- MediaTeam emotional speech corpus (easy)
 - Finnish, acted (stereotypical), single modal, basic emotions (4-7 discrete classes), proprietary
- Hytke (challenging)
 - Finnish, spontaneous (emphasized), multimodal, SAM scale (dimensional annotation: valence, activation)
 - Speech, facial video, heart rate (RR), eye tracking, posture

Available biosignal databases

DEAP (<http://www.eecs.qmul.ac.uk/mmv/datasets/deap/>)

- Electroencephalogram (EEG) and peripheral physiological signals (also some facial videos)
- Dimensional annotation (self-report and online ratings)
 - Database descriptions
- <http://www.eecs.qmul.ac.uk/mmv/datasets/deap/readme.html>
- http://www.eecs.qmul.ac.uk/mmv/datasets/deap/doc/tac_special_issue_2011.pdf

Multimodal databases

- AFEW (challenging)
 - Spontaneous, multimodal, seven categorized emotions
 - **Facial video & Speech**
 - 325 video clips with seven emotions (surprise, sadness, neutral, sadness, neutral, happiness, fear, disgust and anger)
- Hytke (challenging)
 - 24 females
 - Finnish, spontaneous (emphasized), multimodal, SAM scale (dimensional annotation: valence and activation)
 - **Speech, facial video, heart rate**, eye tracking, posture
- MAHNOB (Challenging)
 - Spontaneous, multimodal, dimensional annotation: valence and arousal
 - **Facial video, EEG signal (32 channels), ECG, respiration amplitude and skin temperature**
 - The physiological signals are stored using BDF/MAT format, which is readable by EEGLAB, Matlab, EDFB browser, etc. (We provide BDF and mat format for physiological signals)
- MMSE (Challenging)
 - Multimodal
 - Emotion surprise, sadness, fear, embarrassment, anger, disgust, happiness, startle, sceptical, and physical pain
 - **Synchronized 3D models, 2D videos, thermal, physiological sequences, FACS codes, the heart rate, blood pressure, electrical conductivity of the skin (EDA), and respiration rate**

Tools: facial expression tools

- Face detection
 - http://dlib.net/face_detector.py.html
 - <https://www.pyimagesearch.com/2018/09/24/opencv-face-recognition/>
- Appearance features: LBP/LBP-TOP
 - <http://www.cse.oulu.fi/CMV/Downloads/LBPMatlab>
- Landmark Detection for geometric features
 - Discriminative Response Map Fitting (DRMF): <http://ibug.doc.ic.ac.uk/resources/drmf-matlab-code-cvpr-2013/>
 - ASM, AAM
 - HRNet-Facial-Landmark-Detection: <https://github.com/HRNet/HRNet-Facial-Landmark-Detection>

Tools: speech tools

- Praat
- Phonetics toolbox/software, for prosody and general speech analysis
- <http://www.fon.hum.uva.nl/praat/>

- Voicebox
- MATLAB toolbox for speech processing, feature extraction etc.
- <http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html>

Tools: Biosignal tools

- EEGLAB
 - Graphical EEG signal processing environment/toolbox for MATLAB
 - <http://sccn.ucsd.edu/eeglab/>
- BioSig
 - MATLAB Toolbox for biosignal processing, feature extraction etc.
 - <http://biosig.sourceforge.net/>

Fusion and Classification

- Fusion
 - Decision-level
 - SUM, MAX, PRODUCT, MIN, MEAN, etc.
 - Feature-level
 - Multiple kernel learning (MKL), Canonical correlation analysis (CCA), etc.

- Classifiers:

SVM

- Libsvm
 - <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

HMM

- <http://www.cs.ubc.ca/~murphyk/Software/HMM/hmm.html>

CRF

- <http://www.cs.ubc.ca/~murphyk/Software/CRF/crf.html>

KNN