

# Multimodal Interfaces and Affective Computing

IS4300

GUEST LECTURER: LAZLO RING

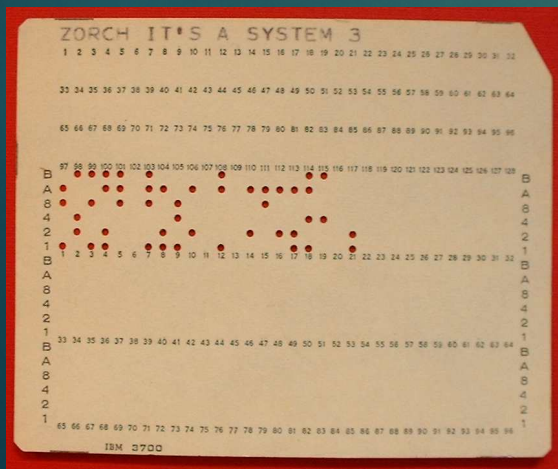
# What is a Multimodal Interface?

- ▶ Interfaces that allows for users to interact through multiple modalities
- ▶ Examples:
  - ▶ Augmented Reality
  - ▶ Virtual Reality
  - ▶ Speech Recognition
  - ▶ Tangible Interfaces
  - ▶ Gestural interfaces



# Why should we care?

- ▶ Different interfaces excel in different scenarios
- ▶ Interfaces can accommodate users with different abilities/disabilities



# Multimodal Interface Example:

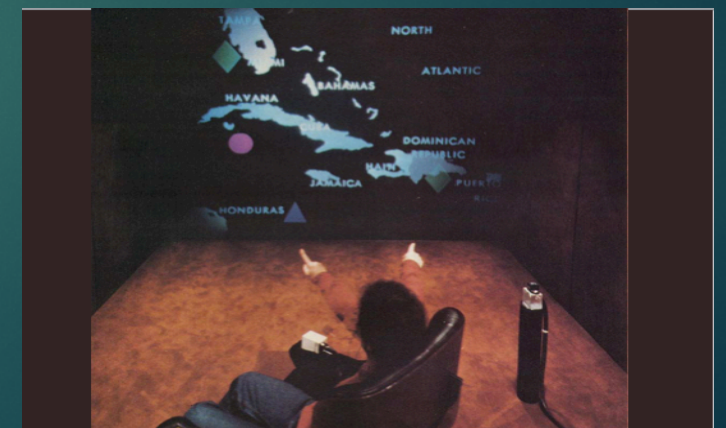
- ▶ Sixth Sense – Pranav Mistry

The image shows the logo for 'Sixth Sense' on a black background. The word 'sixthsense' is written in a bold, white, lowercase sans-serif font. Below it, the tagline 'a wearable gestural interface' is written in a smaller, lighter, lowercase sans-serif font.

**sixthsense**  
a wearable gestural interface

# Natural vs. Artificial Interfaces

- ▶ Natural user interfaces
  - ▶ Emulates naturally occurring interactions (I.E. Face to Face conversations)
  - ▶ Recognizes and Produces messages across a variety of channels
- ▶ Artificial interfaces
  - ▶ Extends standard computer interactions with new interface modalities



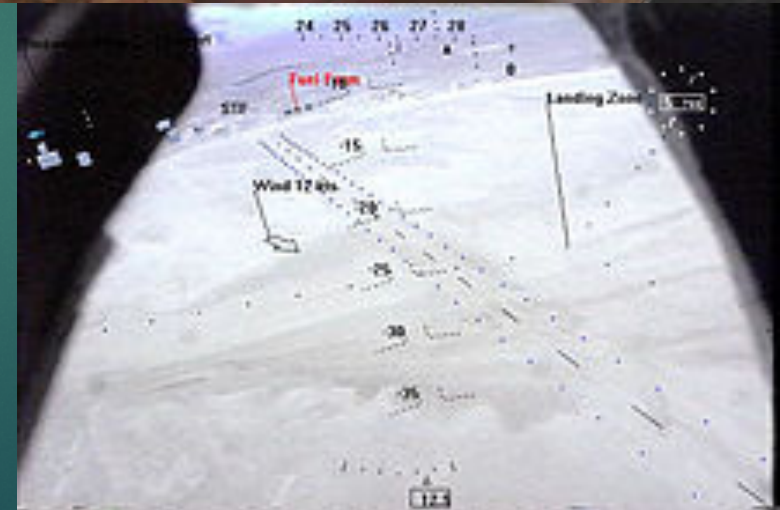
# Interaction Terms

- ▶ Message:
  - ▶ What are you trying to get across to your user?
- ▶ Medium
  - ▶ How are you delivering the message to the user?
- ▶ Modality
  - ▶ What senses are being used by the user to receive the message?



# Augmented Reality

- ▶ Systems that augment the user's view of the world by providing additional information
- ▶ Most commonly done through visual stimuli
- ▶ Heavily dependent on the fields of computer vision and object recognition



# Virtual Reality

- ▶ Systems that fully immerse a user in a virtual environment
- ▶ Commonly done through Head Mounted Displays (HMD) and Data Gloves
- ▶ Very hardware dependent





# Virtual Reality Example



# Challenges of Augmented/Virtual Reality

- ▶ System Reaction Time
  - ▶ Respond too slow and it will be non-immersive/sickening to users
- ▶ Creating/integrating elements into the environment
  - ▶ Scale of the world has to be accounted for
- ▶ Manipulation Techniques
  - ▶ Without the use of a data glove or gestural interface, the user may have trouble interacting with the interface.

# Sound Based Interfaces

- ▶ If done correctly it can reduce:
  - ▶ Cognitive load
  - ▶ The amount of information that needs to be displayed on the screen
  - ▶ Visual attention grabbers (I.E. flashing pop-ups)
- ▶ If done incorrectly...



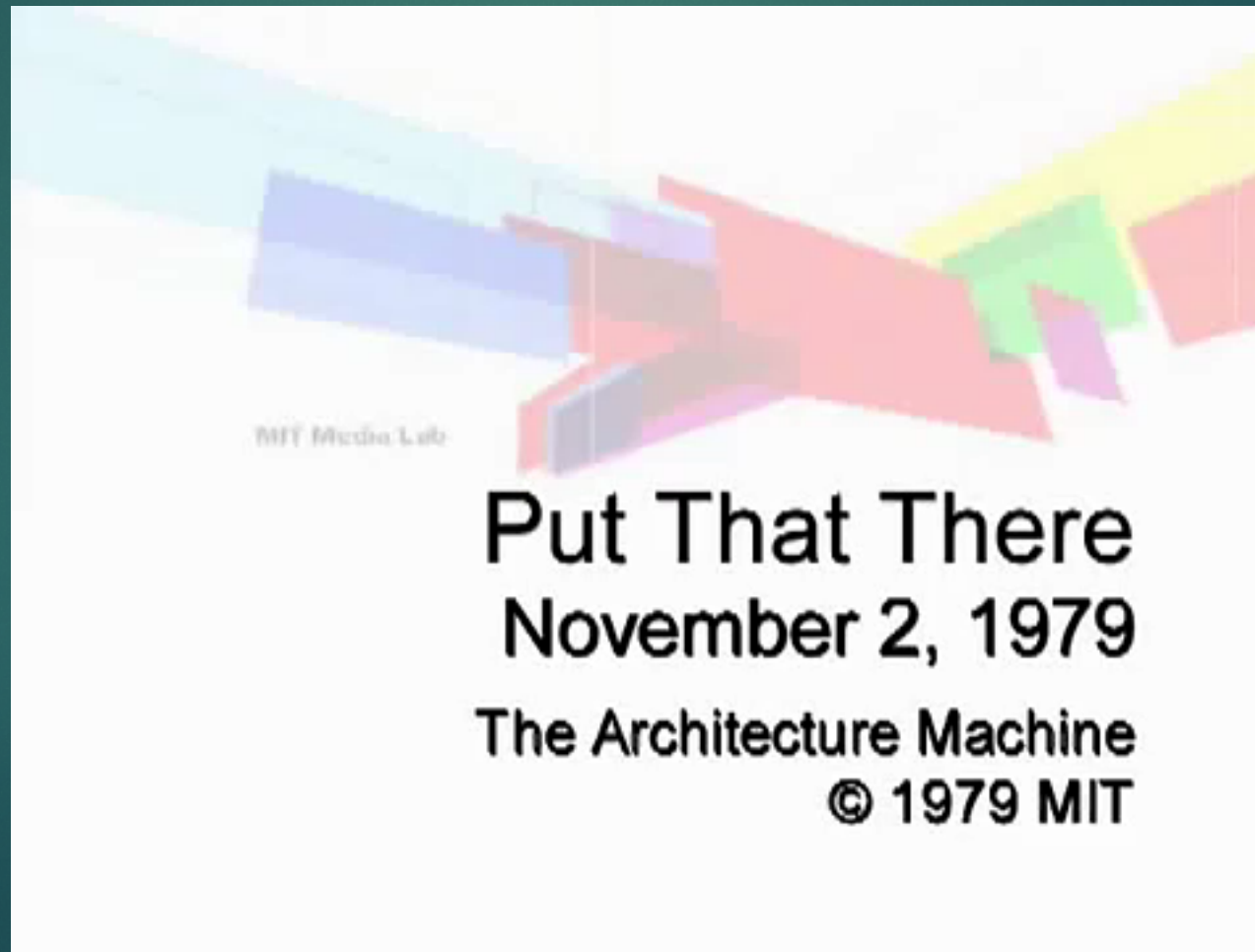
# Speech Recognition Systems

- ▶ Pros:
  - ▶ Allow for natural, voice based interactions
  - ▶ Can require little training on the user's part to use
- ▶ Cons:
  - ▶ Inaccurate
  - ▶ Can be computationally heavy, resulting in delayed interactions



# Speech Based Interfaces

- ▶ Put That There (Chris Schmandt – 1979)



# Navigational interfaces

MATCH: an architecture for multimodal dialogue systems  
- Johnston et al - 2001

Multimodal  
Access  
To  
City  
Help

# Tangible Interfaces

- ▶ Interfaces with physical components
- ▶ Allows for haptic feedback
- ▶ Allows users to understand physical relationships between objects



# Gestural Interfaces

- ▶ Interfaces that use multi-touch or physical movement to control
- ▶ Highly dependent on precise touch interfaces or computer vision
- ▶ Requires additional hardware

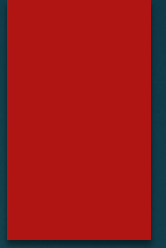




# Multimodal Discussion

- ▶ How would designing a multimodal system be different than designing for unimodal interfaces?
- ▶ What are the benefits of multimodal?
- ▶ What are the drawbacks?

# Affective Computing



# Why should we care about Emotions?



# Why should computers care about Emotions?

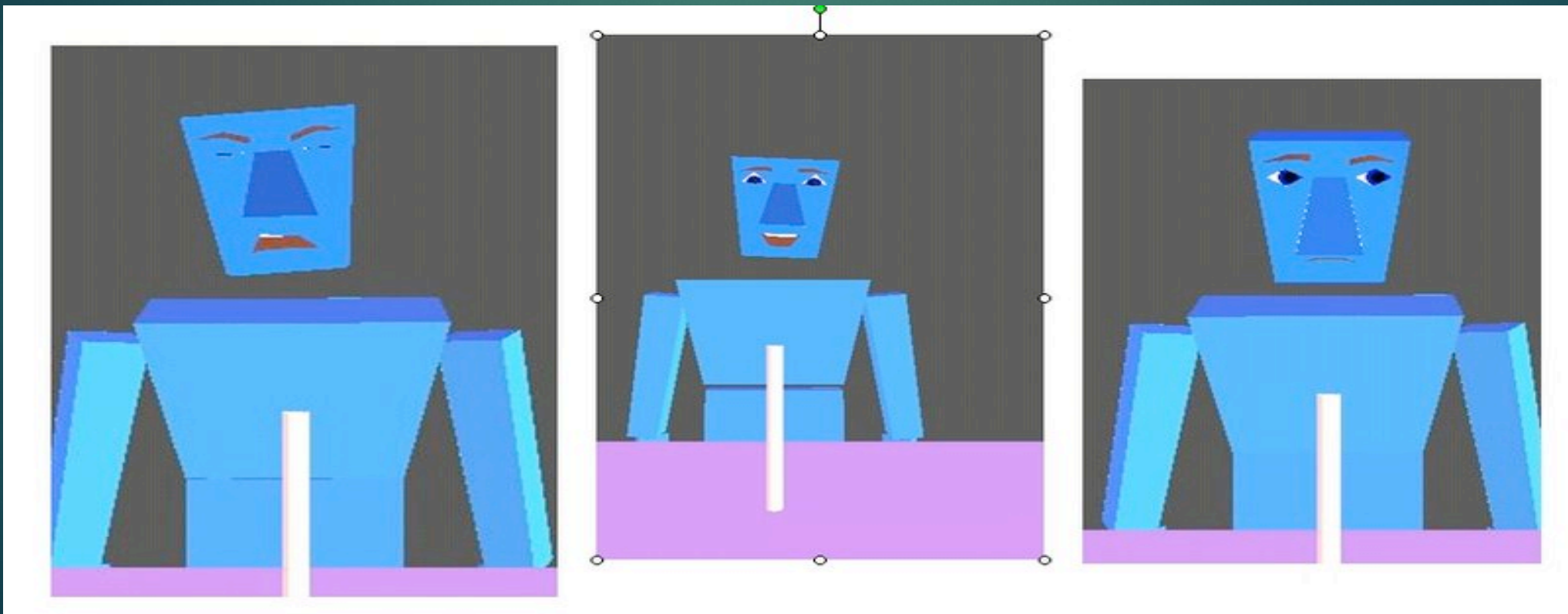


# Why should computers care about Emotions?



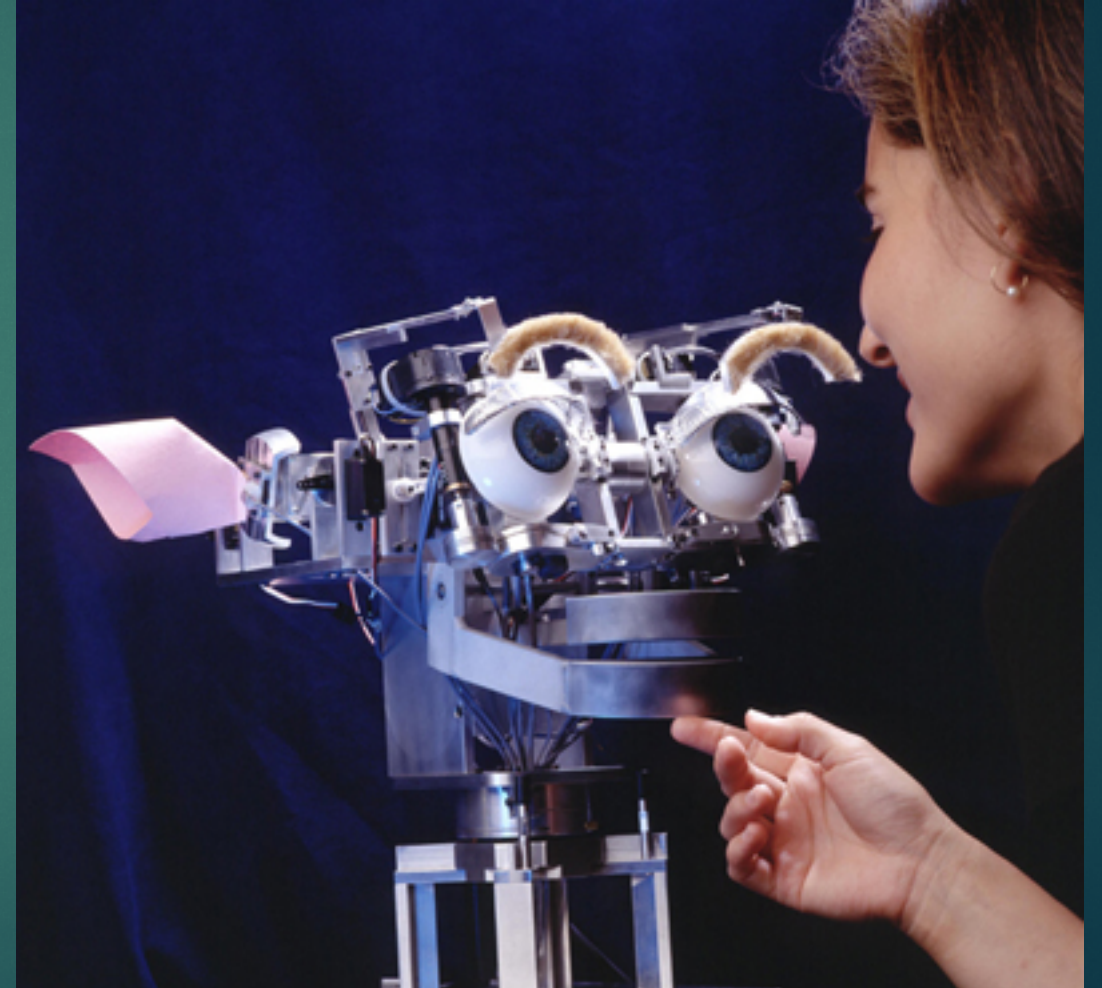
# What is Affective Computing?

- ▶ “Computing that relates to, arises from, or influences the users emotions” – Picard 1995



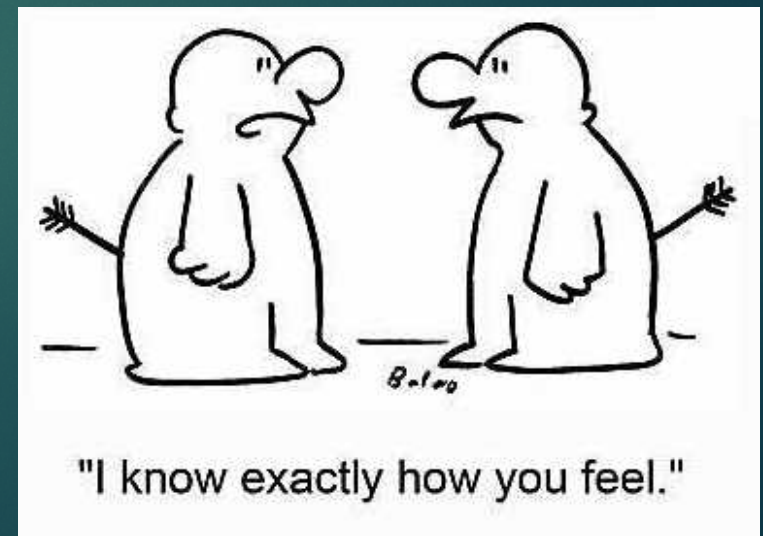
# Types of Affective Computing

- ▶ Systems that elicit affect
  - ▶ Systems that use cognitive models to understand the user's emotional state
- ▶ Systems that recognize affect
  - ▶ Commonly done through sensors
- ▶ Systems that react to affect
  - ▶ Systems that behavior differently based your emotional state



# So, Why do we care again?

- ▶ Klein, J., Moon, Y., & Picard, R. W. (2002). This computer responds to user frustration:: Theory, design, and results. *Interacting with computers*, 14(2), 119-140.
  - ▶ Game designed to elicit frustration
  - ▶ “Questionnaire” either:
    - ▶ Ignored them
    - ▶ Let them vent
    - ▶ Empathized with them





# How do we classify Emotion?

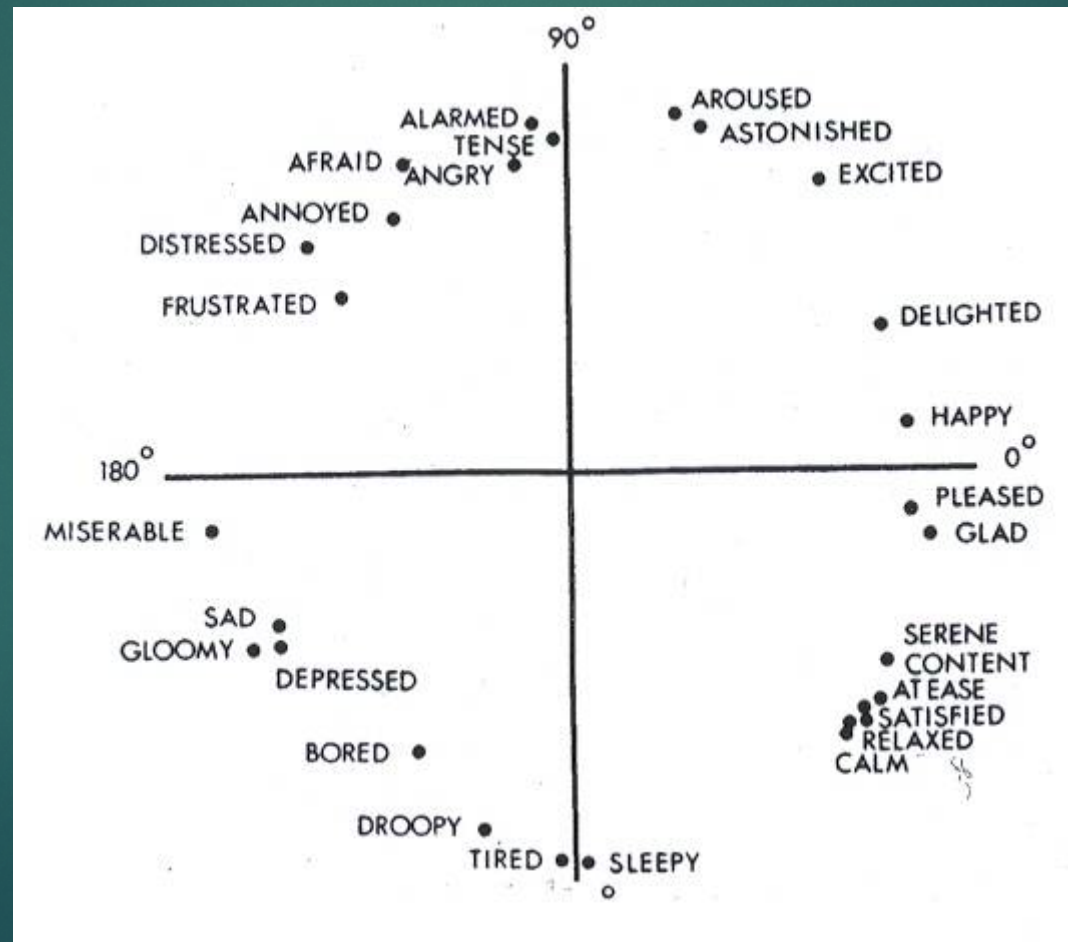
- ▶ **Ekman's list of basic emotions:**

- ▶ *Anger*
- ▶ *Disgust*
- ▶ *Fear*
- ▶ *Happiness*
- ▶ *Sadness*
- ▶ *Surprise*


















# How do we classify Emotion?

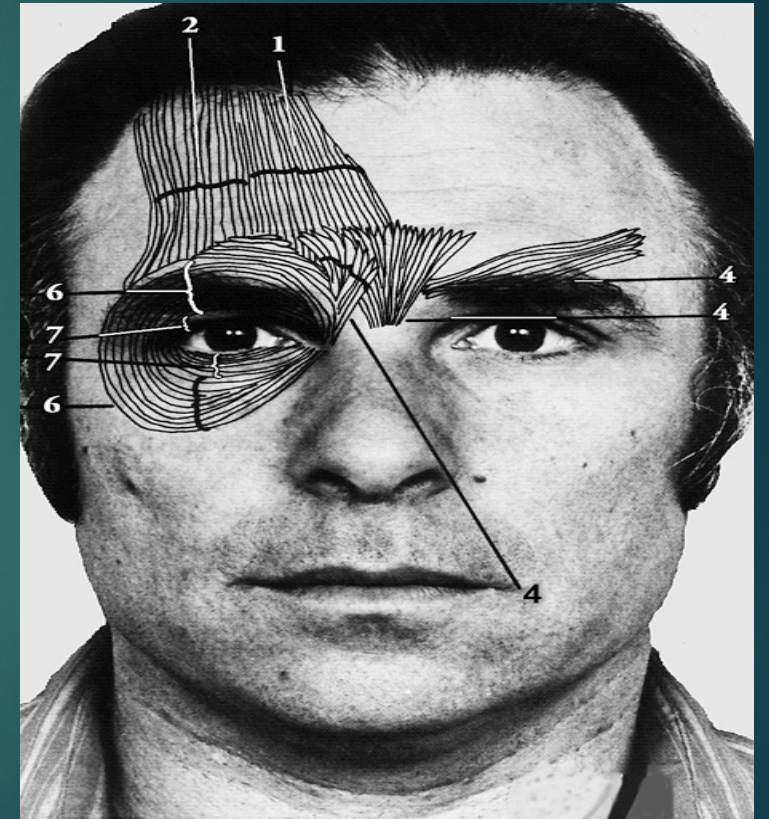
- ▶ Classified through **Valence** and **Arousal** via the **Circumplex Model**



# Detecting Emotion: Facial Expressions

## ► Ekman's Facial Action Coding System

<b>AU1</b>  Inner brow raiser	<b>AU2</b>  Outer brow raiser	<b>AU4</b>  Brow Lowerer	<b>AU5</b>  Upper lid raiser	<b>AU6</b>  Cheek raiser
<b>AU7</b>  Lid tighten	<b>AU9</b>  Nose wrinkle	<b>AU12</b>  Lip corner puller	<b>AU15</b>  Lip corner depressor	<b>AU17</b>  Chin raiser
<b>AU23</b>  Lip tighten	<b>AU24</b>  Lip presser	<b>AU25</b>  Lips part	<b>AU26</b>  Jaw drop	<b>AU27</b>  Mouth stretch



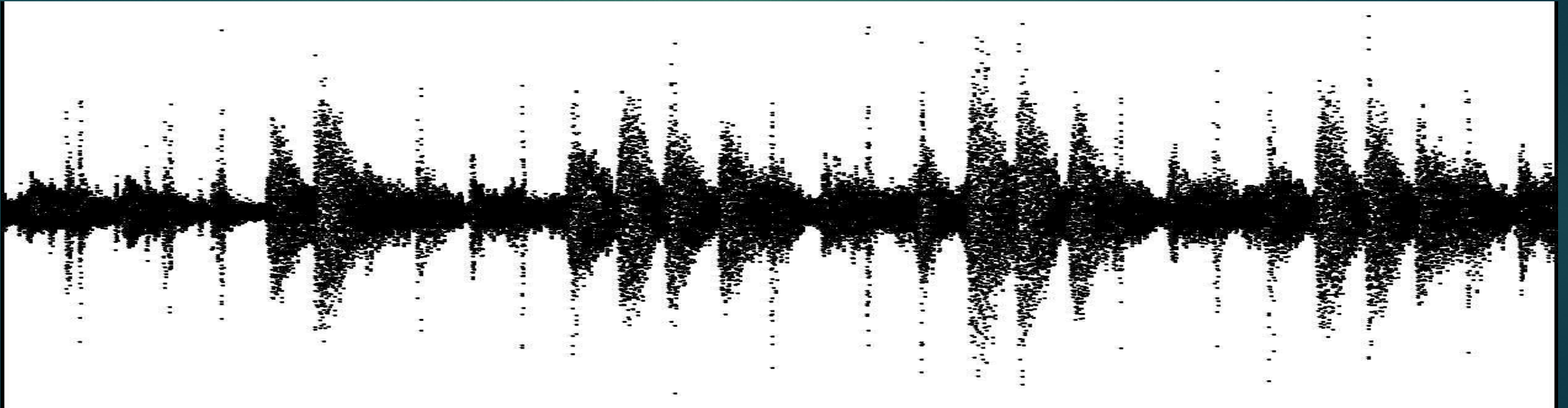
# Detecting Emotion: Facial Expressions



# Detecting Emotion: Audio Processing

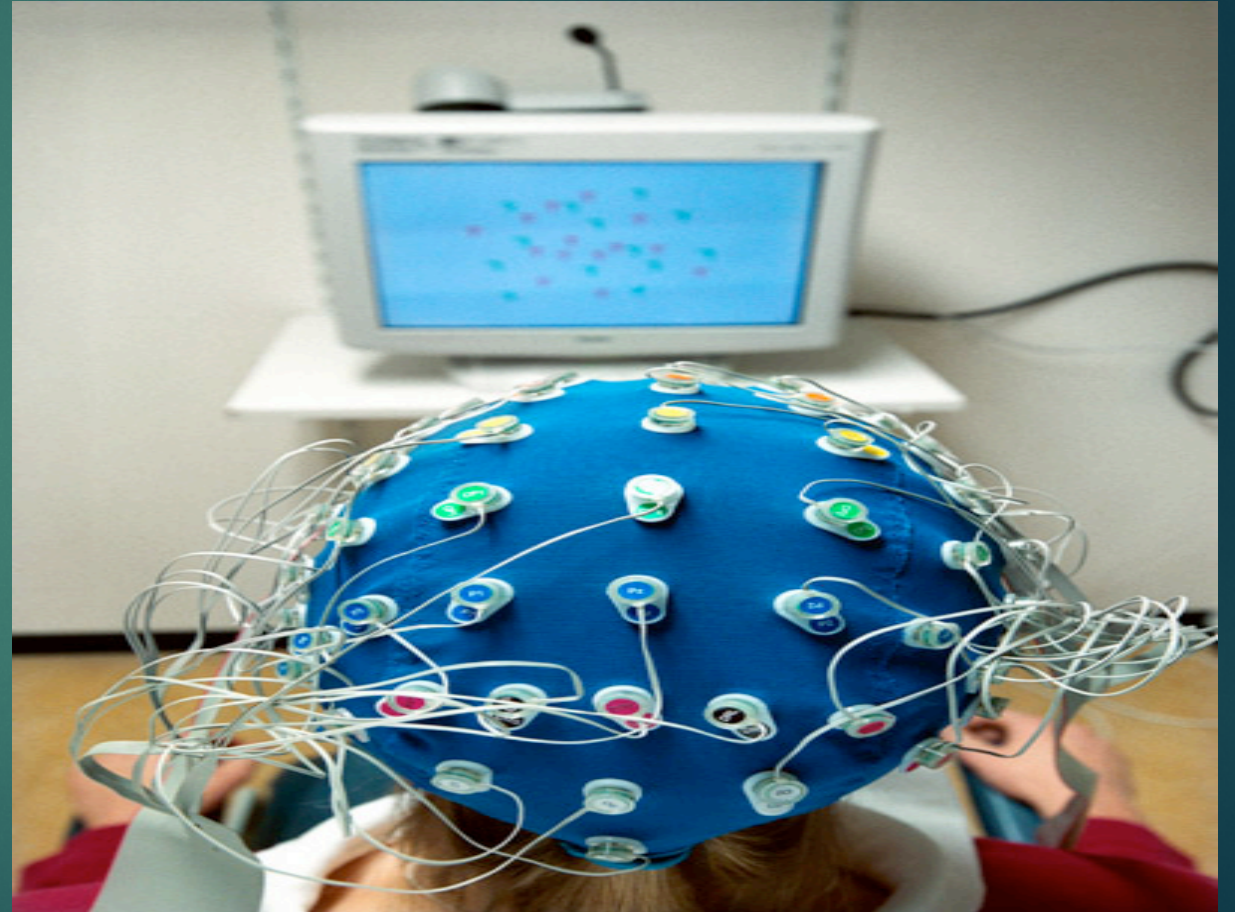
**Audio data can be analyzed to detect emotions:**

- ▶ Variance in voice energy
- ▶ Amount of breathing



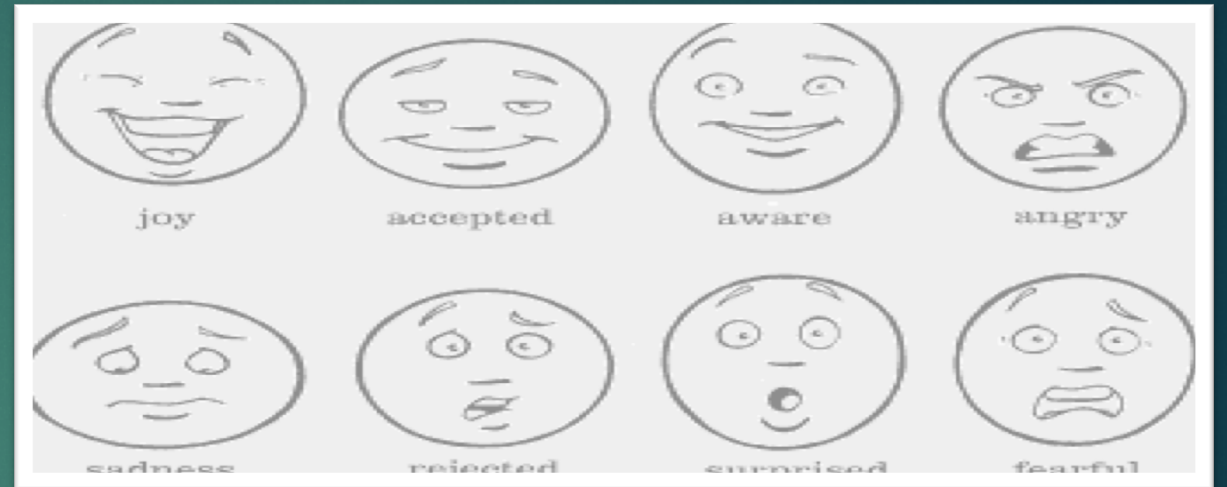
# Detecting emotion: Physiological Sensors

- ▶ Heart Rate
- ▶ Skin Conductance
- ▶ EKG



# Detecting Emotion: Other techniques

- ▶ Posture
- ▶ Just ask!
  - ▶ Via Likert Scale Measures
  - ▶ Via Circumplex Model



# What do we do after we know their emotional state?

- ▶ Burleson, W. (2006). *Affective learning companions: strategies for empathetic agents with real-time multimodal affective sensing to foster meta-cognitive and meta-affective approaches to learning, motivation, and perseverance* (Doctoral dissertation, Massachusetts Institute of Technology).
  - ▶ Mirroring Emotions





# What else can we do with their emotional state?

- ▶ Longitudinal Affect Computing – Ring, Bickmore, Schulman, IVA 2012
  - ▶ Interacted with an embodied conversational agent for a month
  - ▶ One of two dialogue variants used to ask participants to take a walk



# What about displaying emotions?

- ▶ ***Length of Smile Apex as Indicator of Faked Expression, McDaniel & Si, IVA 2014***
- ▶ Systemically exploring the display of fake and natural smiles in virtual agents
- ▶ 6 smile variations
  - ▶ Adjusted Duration of the smile at different points



# What would you do?

- ▶ Gather into team groups
  - ▶ Talk about how you could use multimodal/affective interfaces in your project
  - ▶ 10 minutes

# Your groups ideas

▶ Discuss!

Questions?

