Emotion from facial expression recognition

Manuel Graña,
Andoni Beristain

Computational Intelligence group
University of the Basque Country

JCIS 2007, Salt Lake City

Contents

- Motivation
- Facial expressions
- Automatic Facial Expression Analysis
- Emotional databases
- Representative Facial Expression Recognition Systems
- Conclusions
- References

Contents

- **Motivation**
- Facial expressions
- Automatic Facial Expression Analysis
- Emotional databases
- Representative Facial Expression Recognition Systems
- Conclusions
- References

Motivation

- Non verbal information prevails over words themselves in human communication (M. Pantic, L. J.M. Rothkrantz ,B. Fasel, J. Luettin,…)
- Ubiquitous and universal use of computational systems, requires improved human-computer interaction.
- Humanize computers

Motivation (II)

- Affective Computing: *Affective computing is computing that relates to, arises from, or deliberately influences emotions (R. W. Picard).*

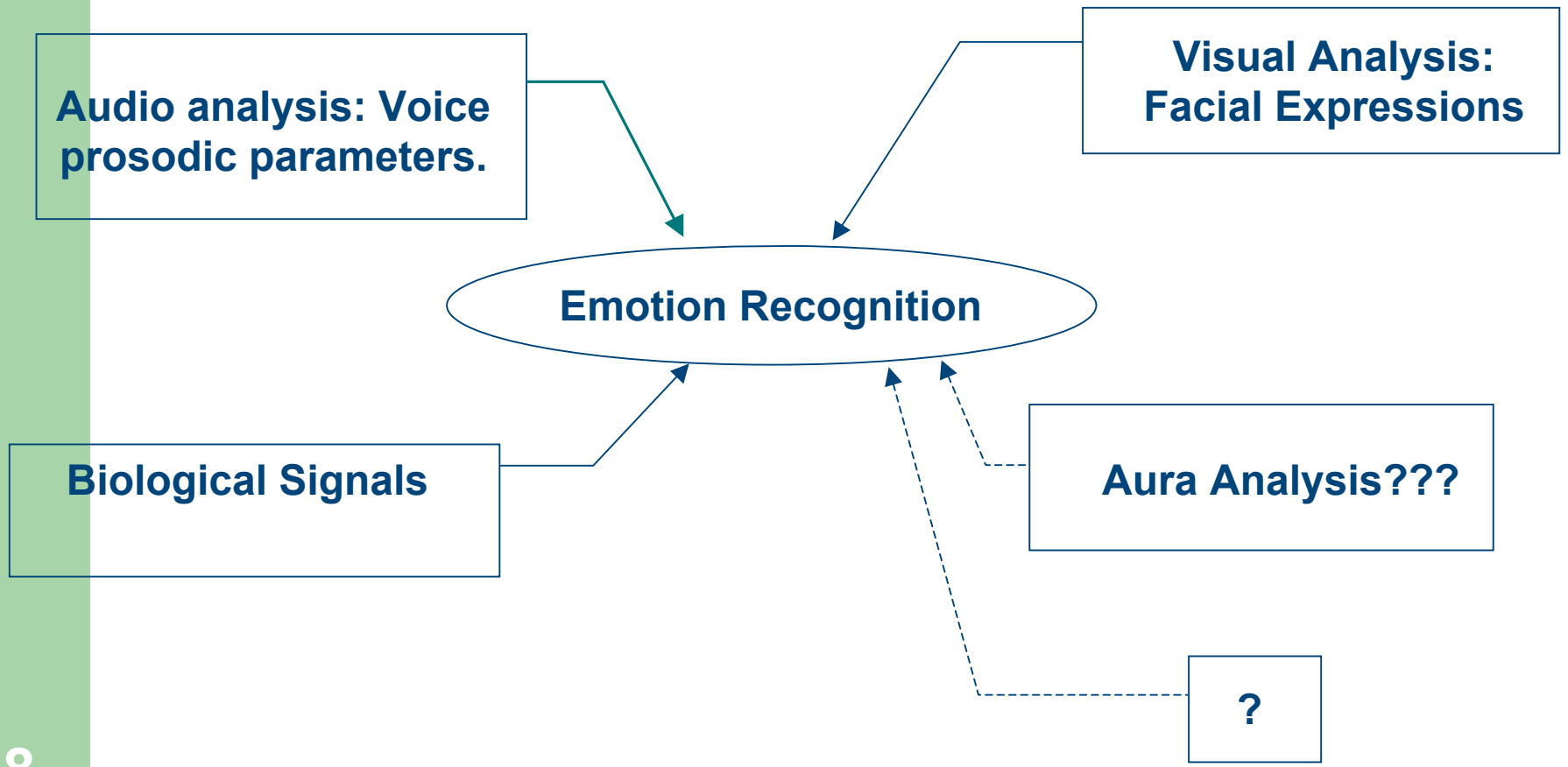
Motivation (III)

- Automatic emotion recognition doesn't begin until 1990:
  - Affordable computer power
    - Signal processing.
    - Classifier system construction
    - Face detection
  - Foundations from
    - Face detection and analysis
    - Machine learning
  - Reduced noise sensors.
  - Voice recognition.

Motivation (IV)

- Application :
  - Predictive environments (Ambient Intelligence).
  - More human-like human-computer, and human-robot interaction (e.g: emotional avatar).
  - Emotional Mirror (Affective Computing).
  - Treatment for people with psycho-affective illnesses (e.g.: autism).
  - Distance learning

Motivation (V)



Contents

- Motivation
- **Facial expressions**
- Automatic Facial Expression Analysis
- Emotional databases
- Representative Facial Expression Recognition Systems
- Conclusions
- References

Facial expressions

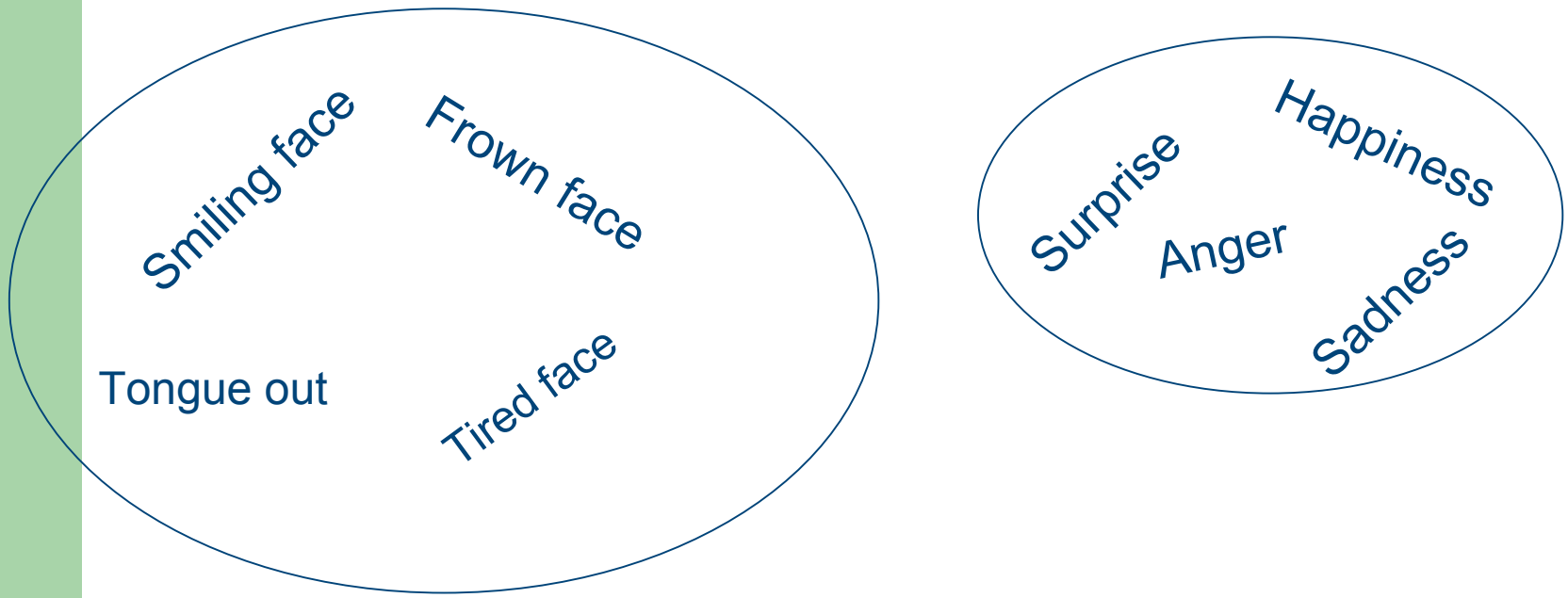
- Facial muscle movements.
- Wrinkles.
- Temporary deformation of facial features.
- Short in time, a few seconds.
- 3 stages: initiation, intensification, transition
- Strength of facial expressions.

Facial expressions (III)

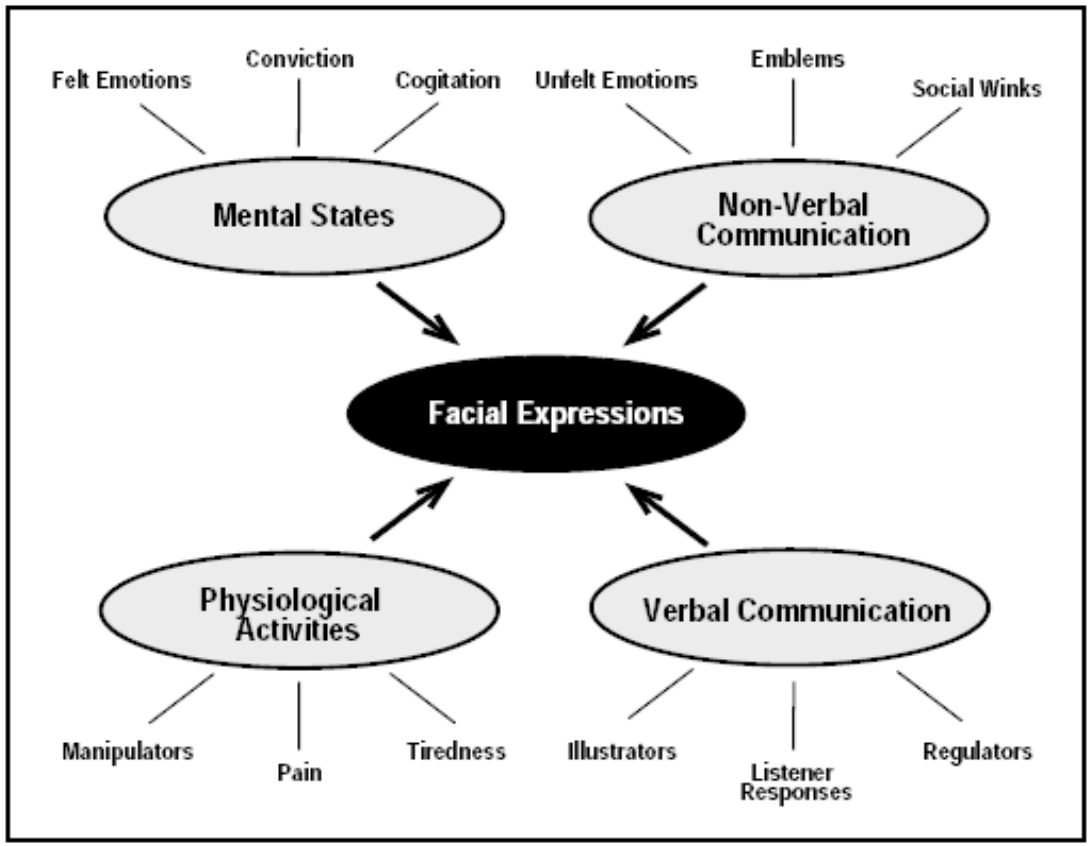
- Paul Ekman's 6 universal emotions:
  - Same facial expressions for everybody.
  - Surprise, Fear, Anger, Disgust, Happiness, Sadness.
- Neutral facial expression and neutral emotion.

Facial expressions (IV)

Facial expressions ≠ Emotion



Facial expression (V)



Fassel 2003

Contents

- Motivation
- Facial expressions
- **Automatic Facial Expression Analysis**
- Emotional databases
- Representative Facial Expression Recognition Systems
- Conclusions
- References

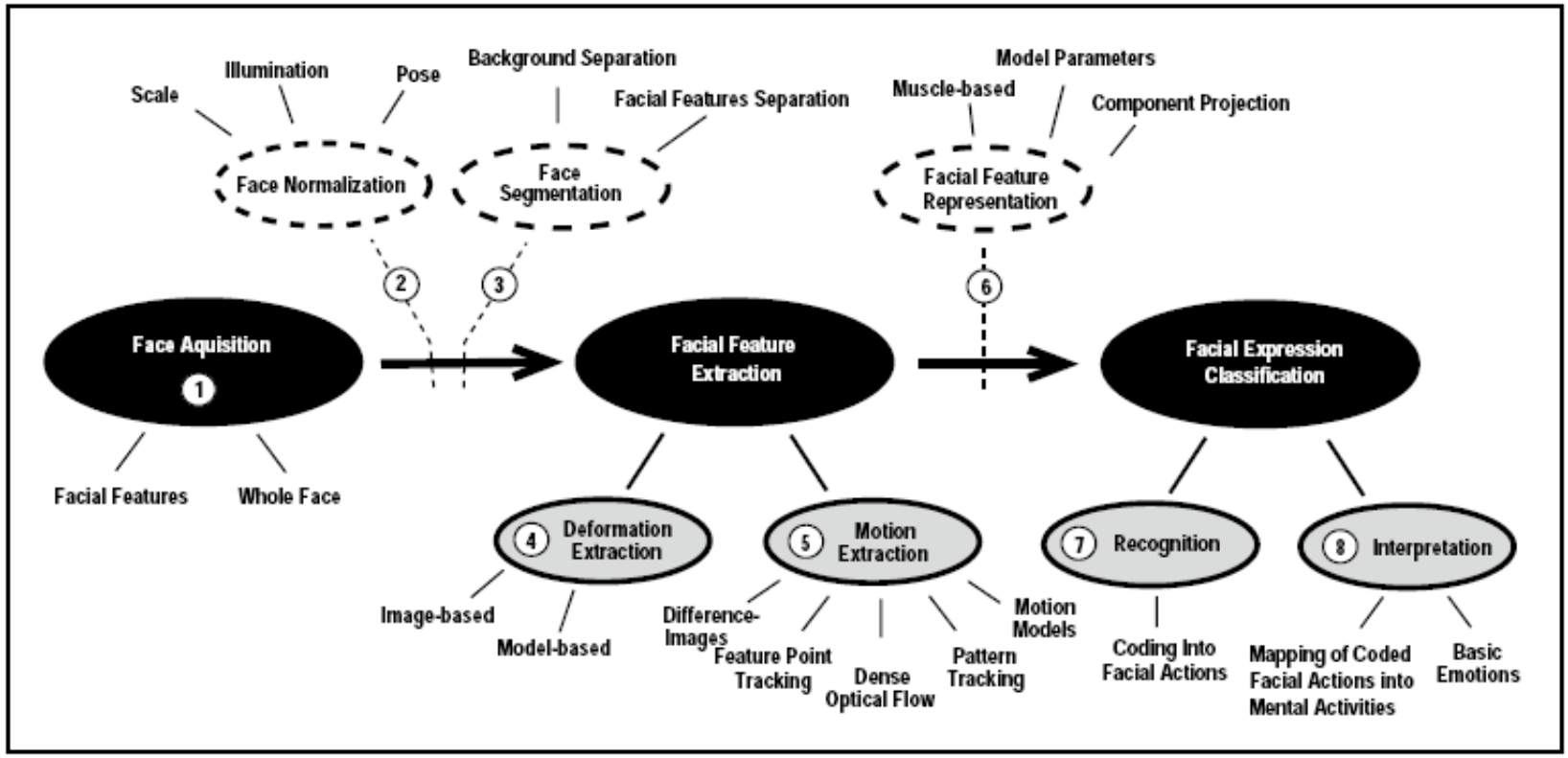
Automatic Facial Expression Analysis

- Ideal System:
  - Automatic facial image acquisition.
  - Subjects of any age, ethnicity and appearance.
  - Robust to variation in lightning.
  - Robust to partially occluded faces.
  - No special markers/make-up required.
  - Deals with rigid head motions.
  - Automatic face detection.
  - Automatic facial expression feature extraction.
  - Deals with inaccurate facial expression data.
  - Automatic facial expression classification.
  - Discriminates all possible expressions.
  - Deals with unilateral facial changes.
  - Obeys anatomical rules.

In summary:

- ✓ **Completely automatic**
- ✓ **Person independent**
- ✓ **Robust to any environmental condition**

Automatic Facial Expression Analysis (II)



Fassel 2003

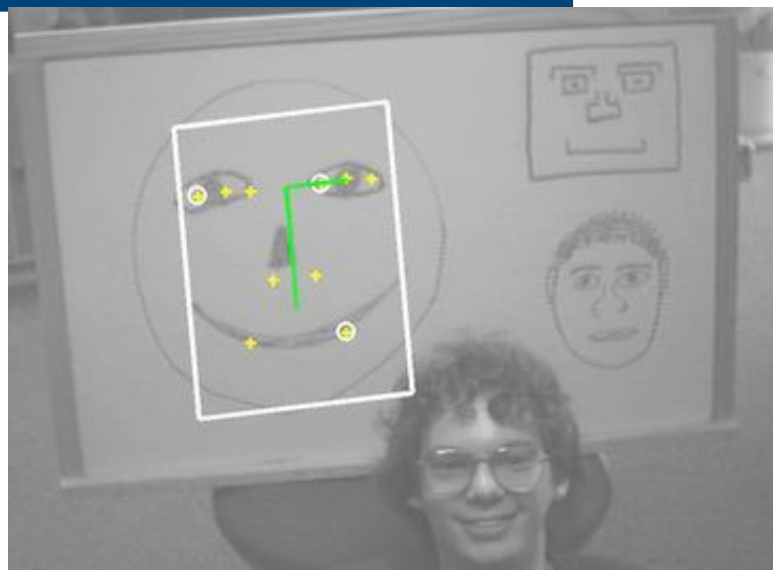
Automatic Facial Expression Analysis: Face acquisition

- Segment face from scene.
- Bounding rectangle or blob.
- 2D and 3D detection.
- Real time 2D solutions: Haar features, SVM, Adaboost,…

Automatic Facial Expression Analysis: Face acquisition (II)

Automatic Facial Expression Analysis: Face acquisition (III)

Automatic Facial Expression Analysis: Face acquisition (IV)



Automatic Facial Expression Analysis: Face acquisition (V)

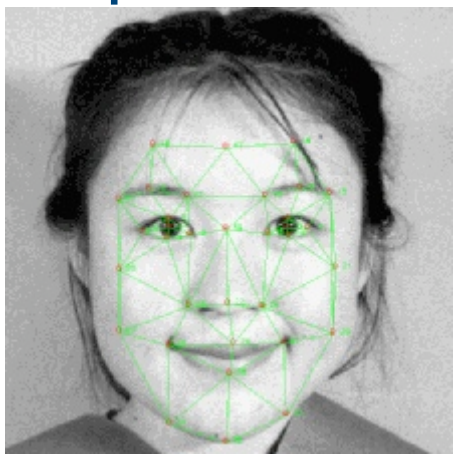
- Face detection is still an ongoing research area.
- Same problems as other artificial vision applications.
- Interpersonal appearance variability.

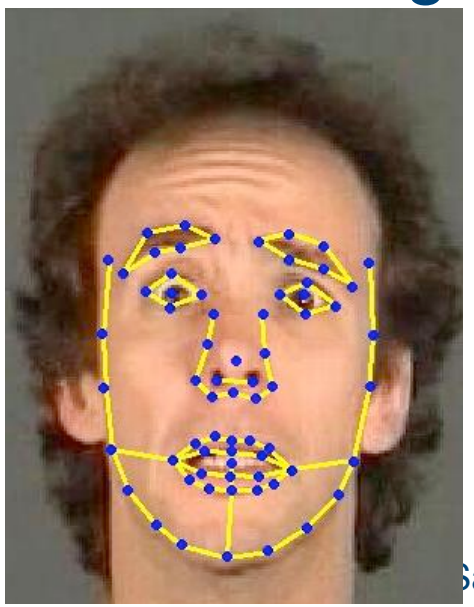
Automatic Facial Expression Analysis: Facial Feature Extraction

- Still Image based methods
  - For both images and videos.
  - Video frames considered independently.
- Video based methods
  - Only for video.
  - Motion information considered.

Still Image based methods

- Facial feature as graph deformation.
- Furrow presence detection.
- Comparison with reference face image.





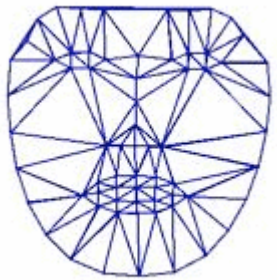
Faculty of Technology Bielefeld University

Still Image based methods

- Recognize facial features:
  - Colour information.
  - Edge information.
  - Shape information.
- Recognize furrows:
  - Edge information.
  - Texture information.

Video based methods

- Motion analysis: Optical flow, tracking algorithms (Kalman, Condensation,…).
- Only for video.
- Require more computer power

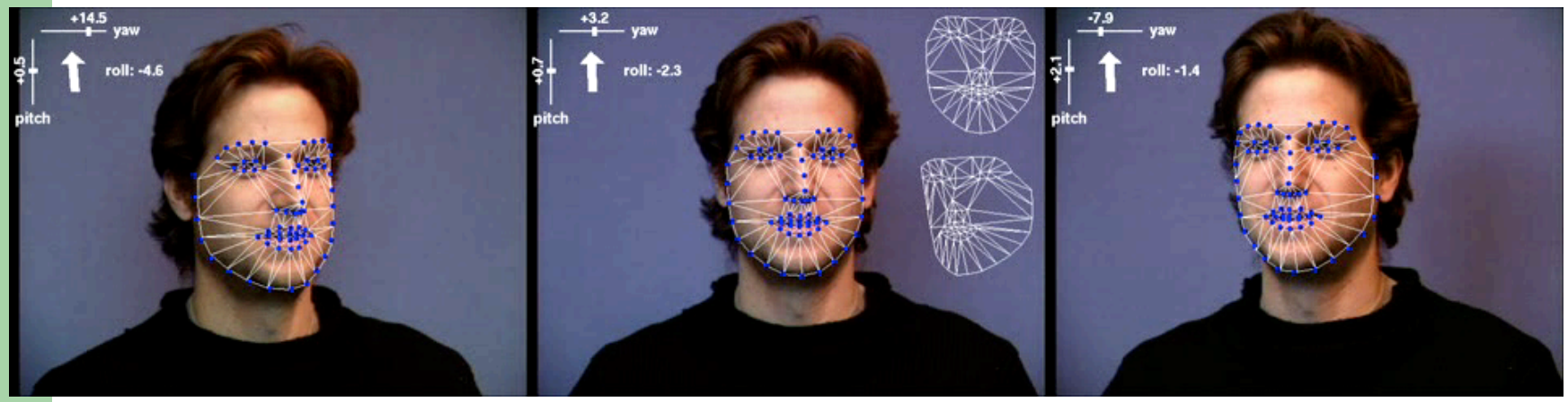


Carnegie Mellon University web

Video based methods

- Active Appearance Models (AAM).
- Carnegie Mellon University.
- Training required.
- Person specific training offer good results.
- Interpersonal training offers poor results.

Video based methods



Carnegie Mellon University web

Automatic Facial Expression Analysis: Facial Feature Extraction

	Holistic	Local
Still image	-PCA -Edges -Colour -Gabor wavelet	-Active Contours -Blobs -Colour -Edges -Gabor wavelet -Local PCA -Template
Video based	-PCA -2D Discrete Cosine Transform (DCT) -Optical Flow -Image difference	-Local PCA -Local Optical Flow -Active Contours

Automatic Facial Expression Analysis: Classification

- Classes
  - Ekman's 6 universal emotions + neutral expression.
  - Every face configuration, when using a coding approach.
- Categories:
  - Based on spatial features.
  - Based on spatiotemporal features.

Classification based on spatial features

- Usually applied after reducing the data dimensionality (PCA, ICA, Gabor filters).
- Artificial Neural Networks (ANN).
- Support Vector Machines (SVM) _ Relevance Vector Machines (RVM).

Classification based on spatiotemporal features

- Facial expressions are something dynamic.
- There is also a pre-processing for noise filtering.
- Hidden Markov Models (HMM).
- Recurrent Neural Networks.
- Motion-energy templates.

Classifiers in Facial expression recognition

- Face expression is used as benchmark to test new classifiers.
- Sometimes non feasible approaches are proposed naively.
- Under laboratory conditions.

Expression recognition approaches

- Direct approach:
  - Feature vector -> emotion
- Coding approach:
  - Feature vector -> facial feature configuration -> facial expression -> emotion

Direct approach

- Feature vector -> Emotion
- Advantages:
  - Lower complexity.
  - Less computer demanding.
- Disadvantages:
  - Difficult to extend with more emotions.
  - Less precise.
  - Difficult to generalize to new data

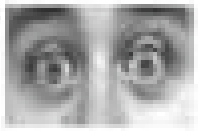
Coding approach

- Feature vector -> facial configuration -> facial expression -> emotion
- Advantages:
  - Precise.
  - Versatile.
  - Extensible.
- Disadvantages:
  - More computer processing required.
  - More complexity.

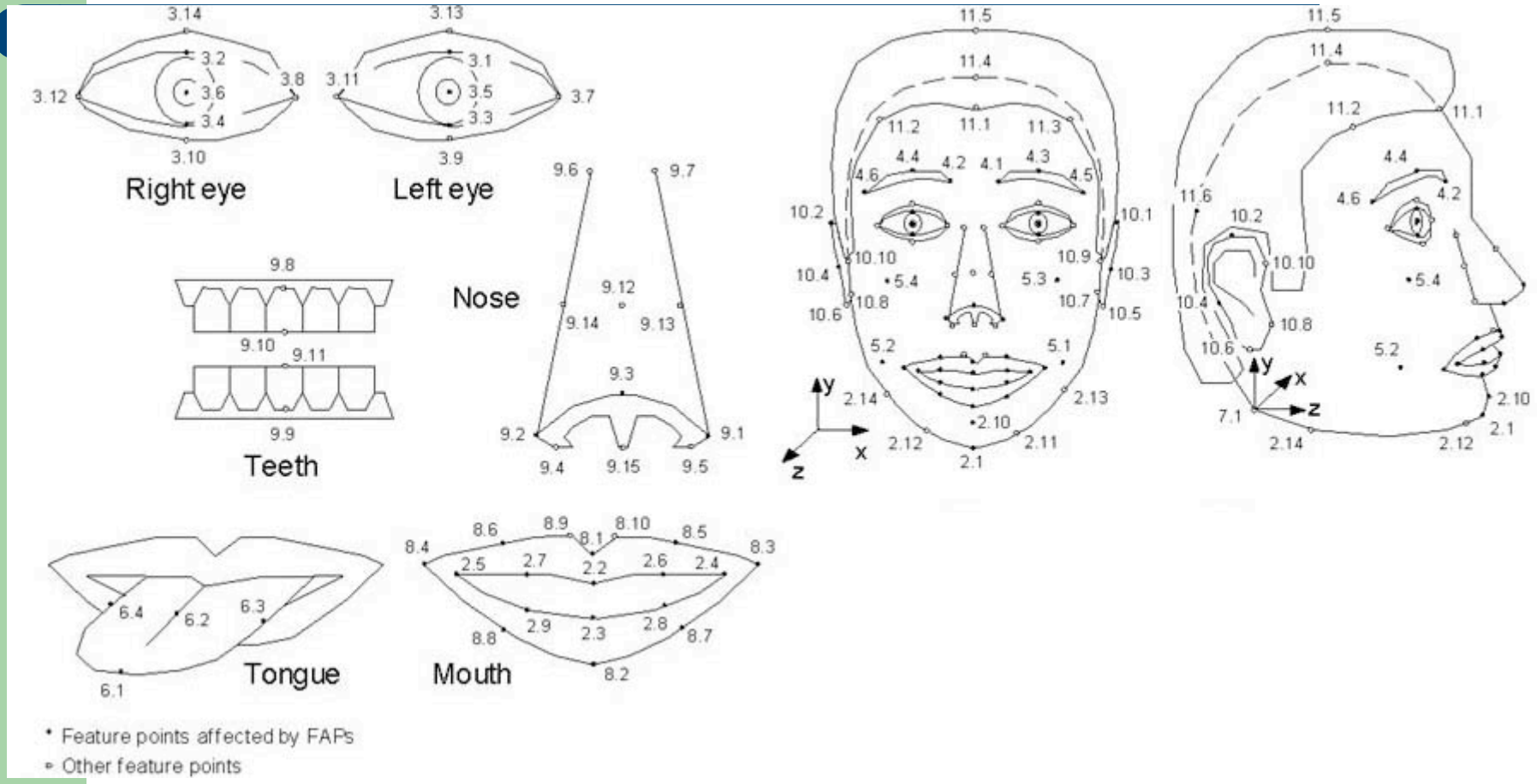
Coding approach (II)

- Facial expression coding systems:
  - Facial Action Coding System (FACS):
    - Origin in psychology, to objectively label video sessions.
    - Partitions facial expressions in terms of specific facial muscle and muscle group movements.
    - Developed by P. Ekman and W. Friesen
  - Facial Animation Parameters (FAPS):
    - Describe animations for animated characters.
    - Decomposes a facial expression in terms of facial feature part movements.
    - Element of the MPEG-4 standard.

Facial Action Coding System (FACS) Example

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

Facial Animation Parameters (FAPS): Example



Contents

- Motivation
- Facial expressions
- Automatic Facial Expression Analysis
- **Emotional databases**
- Representative Facial Expression Recognition Systems
- Conclusions
- References

Emotional databases

- It is essential to have test data to check new approaches and to compare them with previous systems.
- Spontaneous behaviour recordings are required.
- Ethical problems to record some of the universal emotions.

Emotional databases

- Problems labelling the media.
- Different human coders means different labelling.
- Reduce subjectivity, using coding systems (FACS).

Emotional database examples

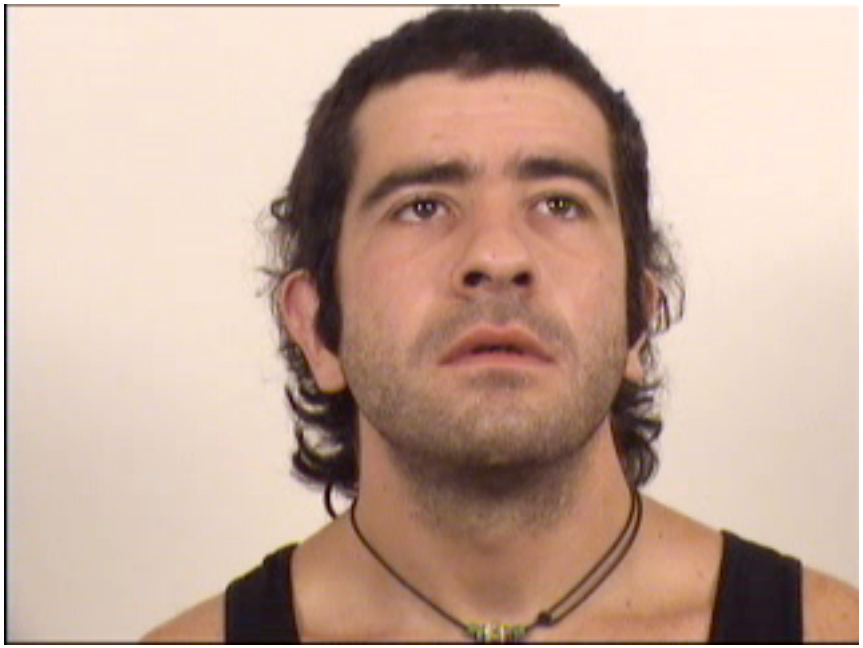
- Cohn-Kanade AU-Coded Facial Expression Database:
  - FACS coded by certified Facial Action Coding System (FACS) coders for either the entire sequence or target Action Unions (Aus
- The PIE (Pose, Illumination and Expression) Database. Human ID Group (Carnegie Mellon University).
- The Vision and Autonomous Systems Center's Image Database
  - Set of Databases
  - The PIE database is also included in this database.
- The FERET Database.
- The AR Face Database from the Computer Vision Center (CVC) at the U.A.B
- FEEDTUM database, JAFFE database, ...
- Our multimedia emotional database.

RekEmozio

- Voice and video
- Mixed population
  - Actors and amateurs
  - Men and women
  - Spanish and Basque
  - Frontal and lateral views
- Six basic emotions + neutral expression
- Diverse sentences
  - Related and unrelated to the emotion

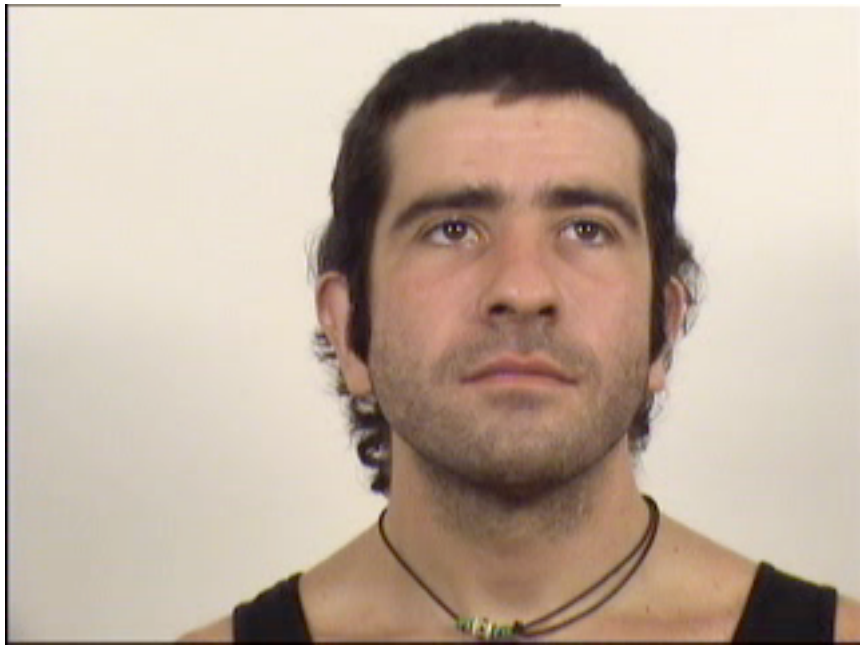
Database instances

- Fear

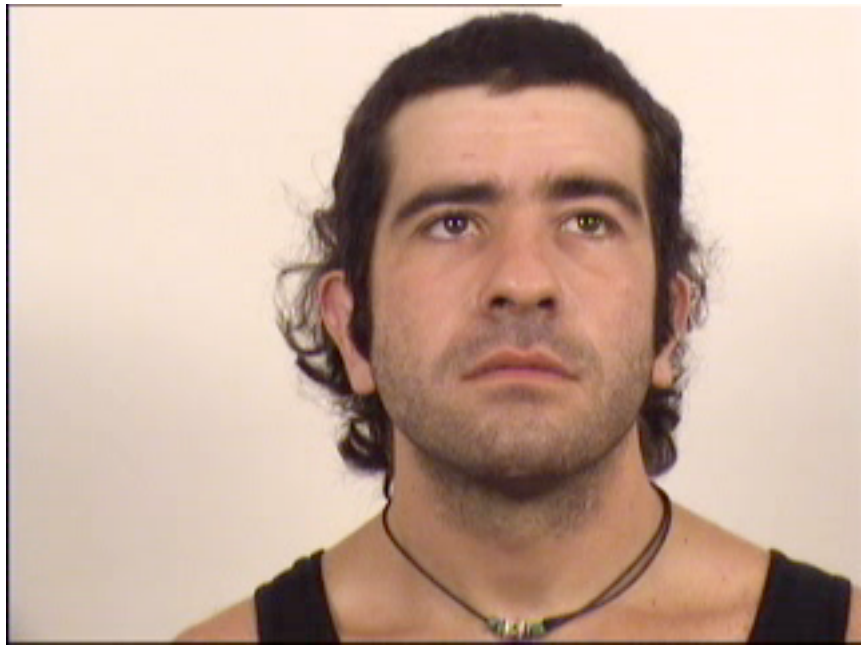


- Disgust

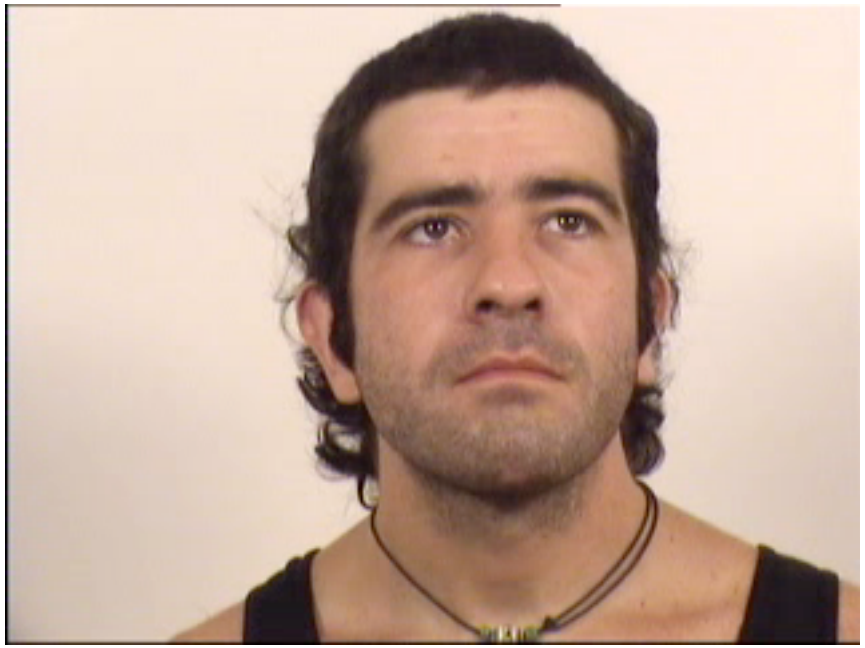
- Happiness



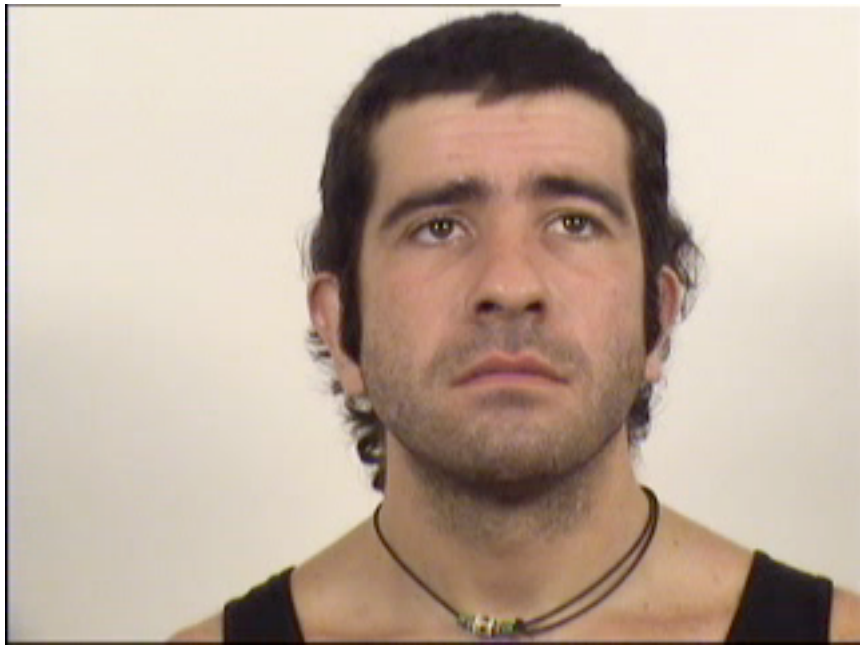
- Surprise



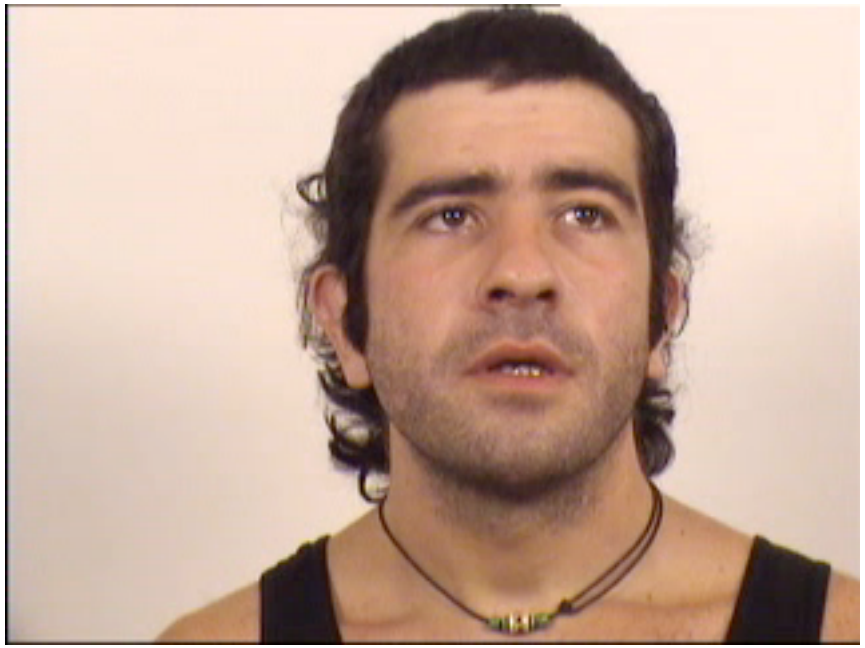
- Anger



- Sadness



- neutral



Languages	*BSQ*	*SPA*	*Women*	*Men*	*Total*
♯ *Actors*	7	10	8	9	17
♯ *Amateurs*	2	10	5	7	12
Total	9	20	13	16	29

	Actors		*Amateurs*		*Total*	
	BSQ	*SPA*	*BSQ*	*SPA*	*BSQ*	*SPA*
♯ *Sentences*	1067	1511	41	207	1108	1718
Total	2578		248		2826	

	Actors		*Amateurs*		*Total*	
	Women	*Men*	*Women*	*Men*	*Women*	*Men*
♯ *Sentences*	1205	1373	103	145	1308	1518
Total	2578		248		2826	

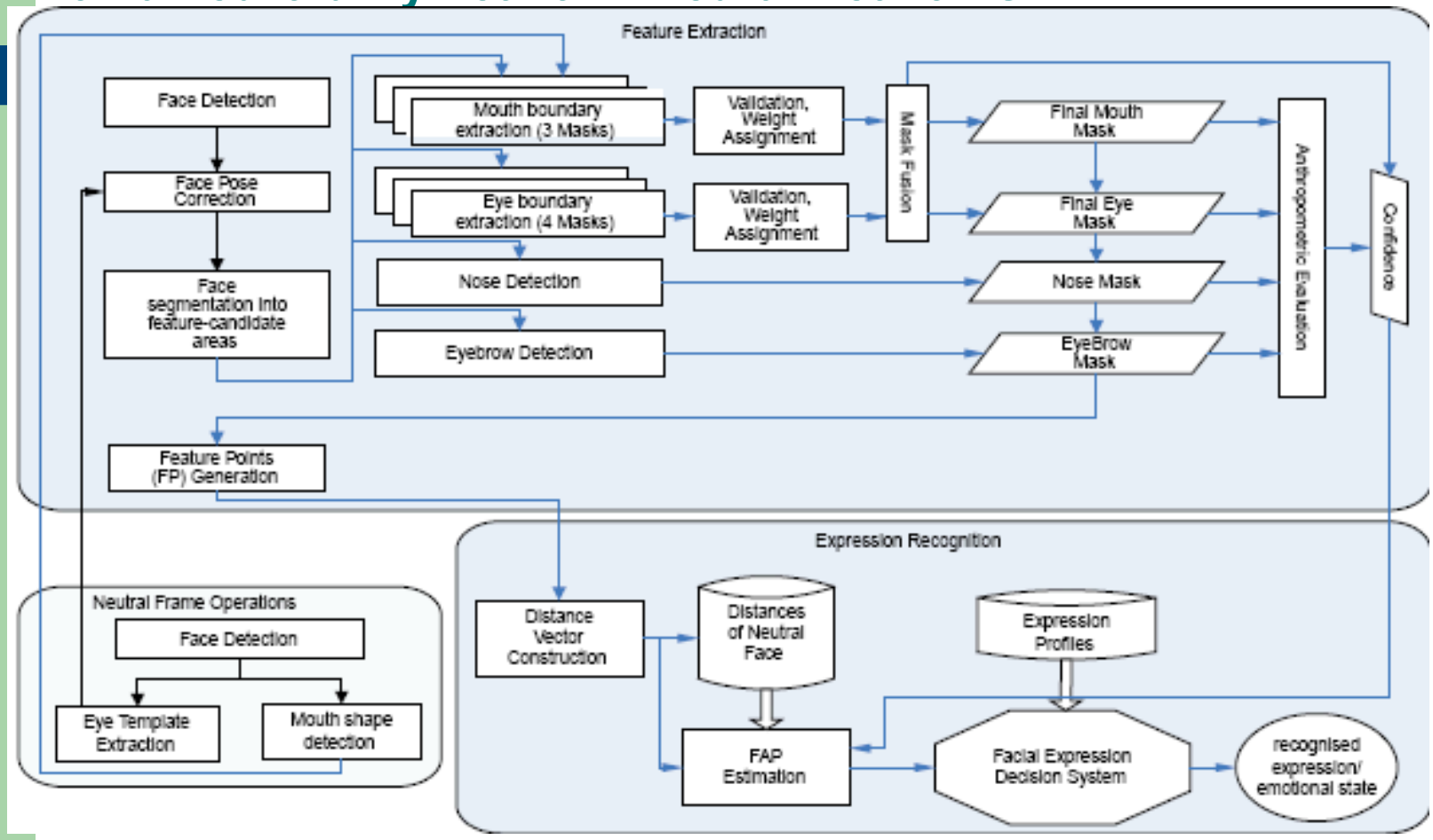
Contents

- Motivation
- Facial expressions
- Automatic Facial Expression Analysis
- Emotional databases
- **Representative Facial Expression Recognition Systems**
- Conclusions
- References

Representative Facial Expression Recognition Systems

- Still image based System:
  - Ioannou, S., et al., *Emotion recognition through facial expression analysis based on a neurofuzzy network.* Neural Networks, 2005. **18**(2005 Special Issue): p. 423-435.
  - 78% of emotion recognition rate in Humane Network of Excellence database.

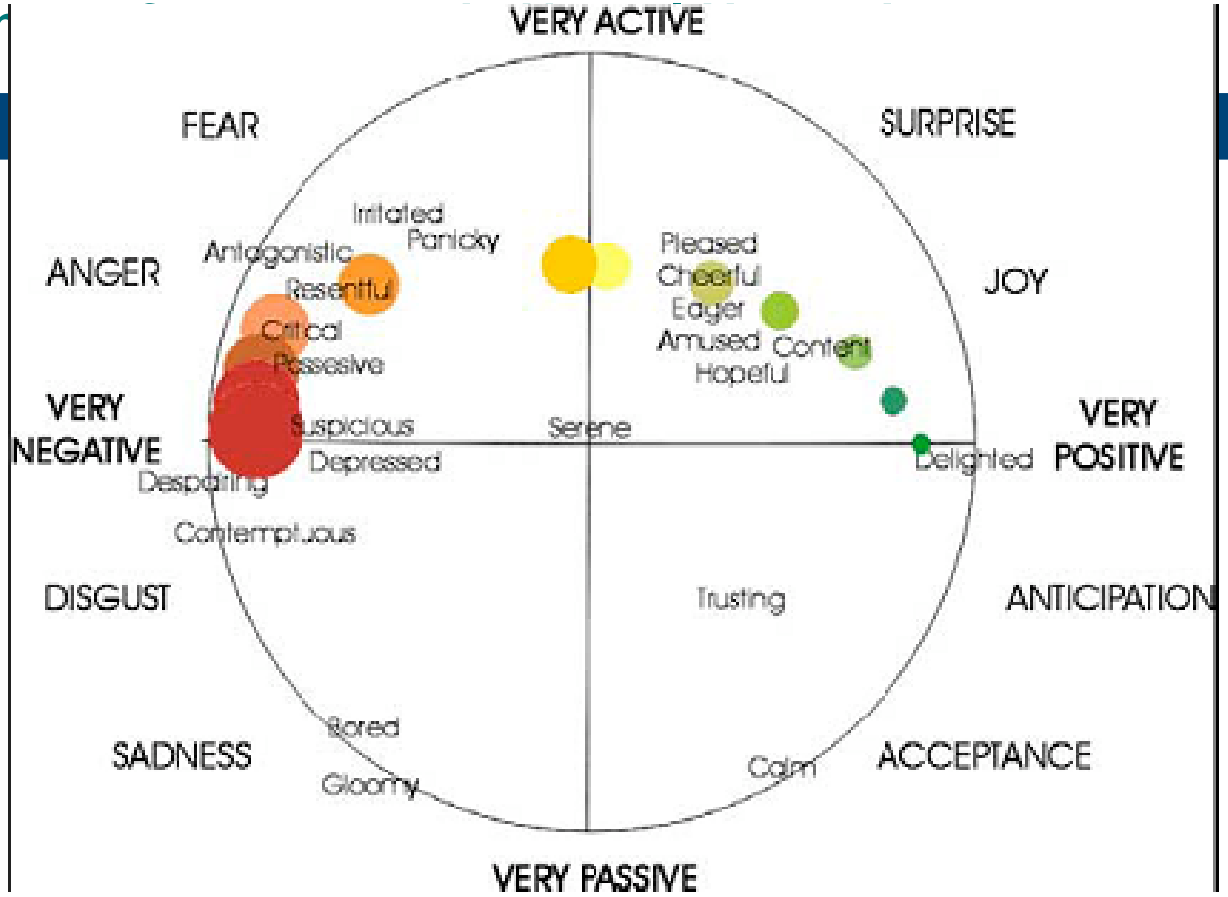
Emotion recognition through facial expression analysis based on a neurofuzzy network. Neural Networks



Emotion recognition through facial expression analysis based on a neurofuzzy network. Neural Networks

- Face location using SVM.
- Facial feature extraction:
  - Eyebrows (morphological edge detection).
  - Eyes (ANN + refinement with Canny and region growing)
  - Nostrils (localized dark areas)
  - Mouth (ANN + mophological gradient + thresholding)
- Coded approach (MPEG-4 FAPS).
- Classifier based on a neurofuzzy network.
- Use of quadrant of emotion's wheel.

Emotion recognition through facial expression analysis based on a r



Quadrant of emotion's wheel.

Representative Facial Expression Recognition Systems

- Video based System:
  - Yeasin, M., B. Bullot, and R. Sharma, *Recognition of facial expressions and measurement of levels of interest from video.* Multimedia, IEEE Transactions on, 2006. **8**(3): p. 500-508.
  - 90.9% of emotion recognition rate in Cohn-Kanade database.

Recognition of facial expressions and measurement of levels of interest from video

- Face location using ANN.
- Pre-processing to normalize size and lighting.
- Optical Flow for motion detection (PCA).
- HMM for classification.
- Direct Approach.

Recognition of facial expressions and measurement of levels of interest from video

Levels of interest

Representative Facial Expression Recognition Systems

- Multimodal system:
  - Sebe, N., et al. *Emotion Recognition Based on Joint Visual and Audio Cues*. in *18th International Conference on Pattern Recognition* 2006.
  - 90.9% of emotion recognition rate in Beckman Institute for Advanced Science and Technology database.

Emotion Recognition Based on Joint Visual and Audio Cues

- Voice and facial appearance input.
- 6 Ekman' universal emotions and some cognitive/motivational states.
- Voice:
  - Features: logarithm of energy, syllable rate, and pitch.
- Facial Appearance:
  - Face location: 3D model adapted manually.
  - 2D motion information.

Emotion Recognition Based on Joint Visual and Audio Cues

Emotion Recognition Based on Joint Visual and Audio Cues

- Combination of information from both inputs is done just after the feature vector extraction, not after emotion classification.
- Bayesian Network for classification.

Innovae Emotional Trainer

Developed by an spin off of the research group

Motivation for Innovae Emotional Trainer

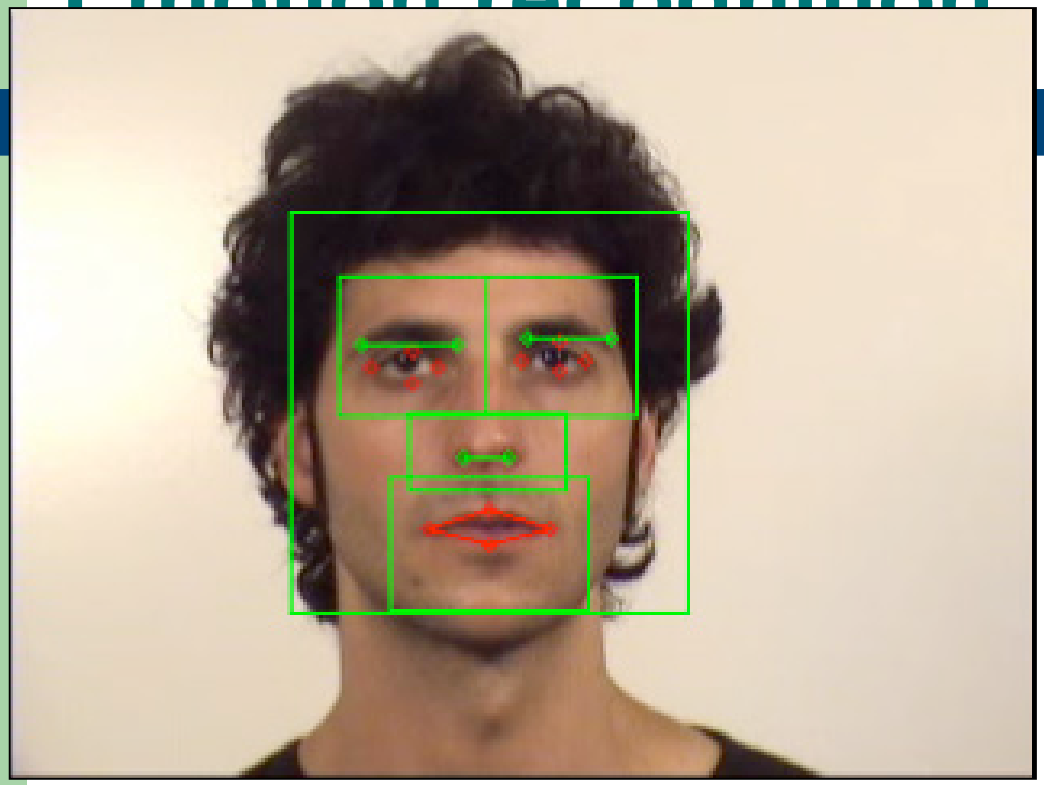
- Measure and improve people's acting skills and expressiveness using:
  - Support multimedia:
    - Descriptive text
    - Sample image
    - Sample video
  - Imitation and self observation
  - Online and offline application's feedback evaluating user's performance.

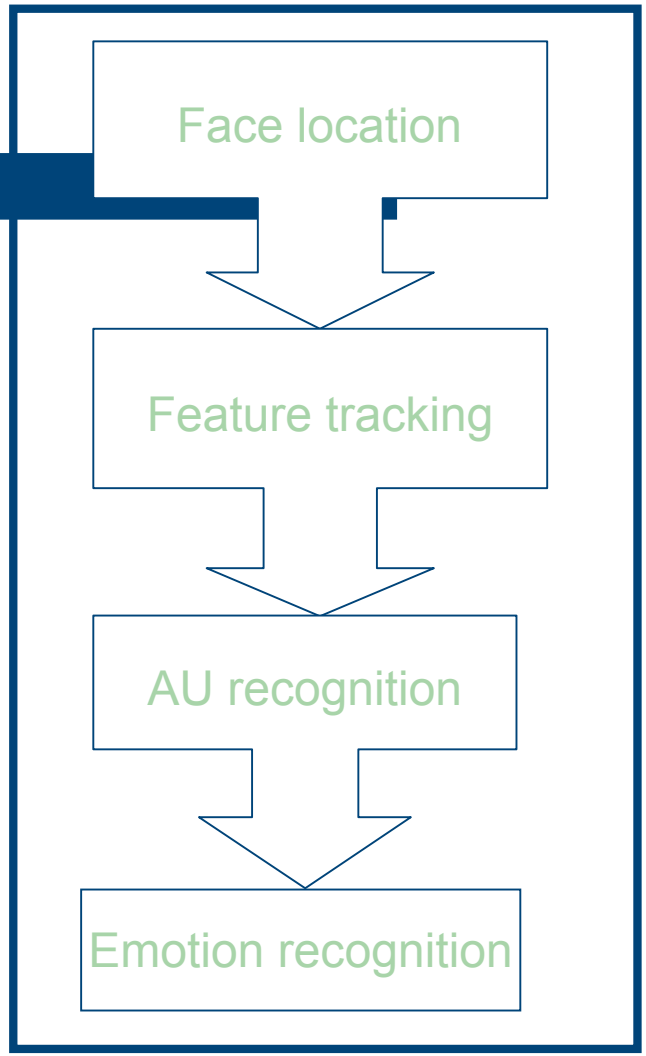
System description

System goals

- Goals:
  - Recognize Ekman's 6 emotions
  - Evaluation and improvement of people's acting skills
- Constraints
  - 1 frontal view of face.
  - Low rotation and translation resistance.
  - Real-time emotion recognition
  - The complete procedure should take less than 5 minutes.

Emotion recognition





JCIS 2007, Salt Lake City

System steps

1. Face location to activate the application
2. Initial context information
3. Calibration
4. For each emotion:
    a. Sample image and facial expression descriptive text.
    b. Sample video.
    c. Acting time.
5. Performance summary

Results on Innovae Emotional Trainer

- Different experiments for different goals:
- Experiment 1:
  - Estimate emotion recognition rate
- Experiment 2:
  - Prove didactic potential of the application.

Results: Experiment 1. Emotion recognition rate

- 20 subjects' video recordings showing the 6 emotions each.
- Image samples of Ekman's emotion image DB.
- 3 evaluators chosen to validate the recorded videos.
- Assume the Innovae emotional trainer as the 4th evaluator.

Results: Experiment 1. Emotion recognition rate (II)

	Human evaluators' average			Innovae Emotional Trainer		
	Right	Doubtful	Wrong	Right	Doubtful	Wrong
SADNESS	61%	12%	27%	18%	9%	73%
FEAR	36%	24%	40%	9%	18%	73%
HAPPINESS	100%	0%	0%	82%	18%	0%
SURPRISE	61%	27%	12%	73%	27%	0%
DISGUST	58%	30%	12%	36%	36%	28%
ANGER	79%	9%	12%	73%	18%	9%

Results: Experiment 2. didactic potential of the application

- 15 subjects' video recordings showing the 6 emotions each.
- 3 recording sessions:
  - Before using the application.
  - After using the application once.
  - After using the application twice.
- 4 evaluators marked the expressiveness in each session.

Results: Experiment 2. didactic potential of the application (III)

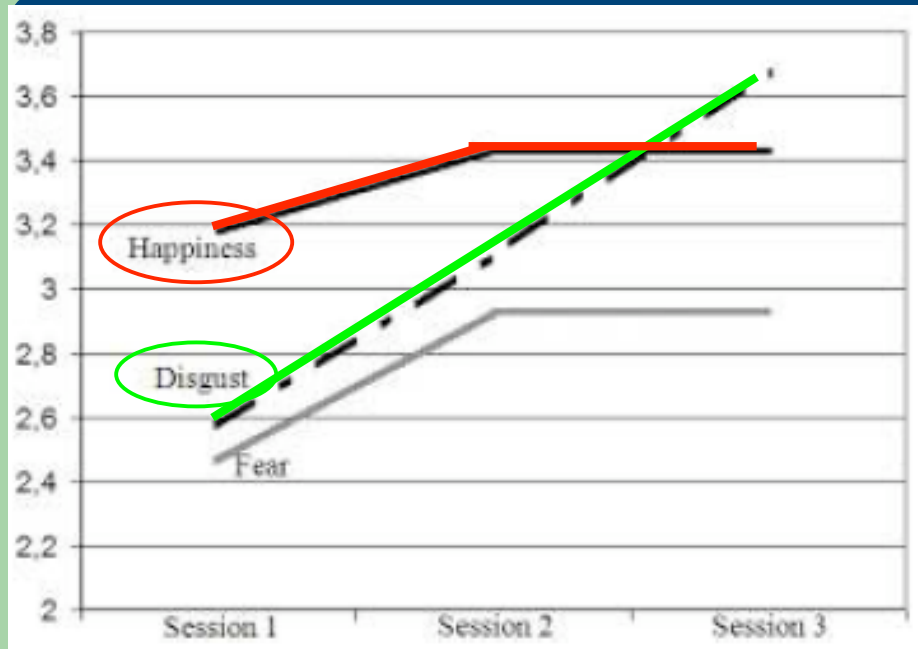


Figure 8: Mark average for Happiness, Disgust and Fear

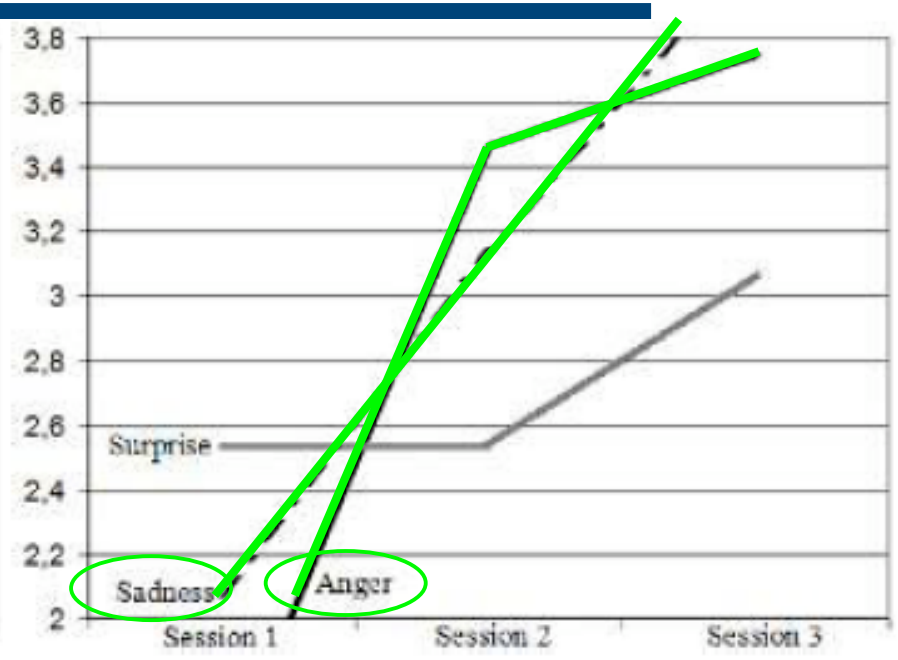


Figure 9: Mark average for Surprise, Sadness and Anger

71% of the videos showed improvement between the first session and the two others

98% of anger videos showed improvement between the first session and the two others

86% of sadness videos showed improvement between the first session and the two others

Conclusions for Innovae Emotional Trainer

- Designed as a “game” but with a teaching purpose:
- Mounted at the “Museo de la ciencia de Valladolid” (Valladolid, Spain)
- Future work:
  - Better recognition rates
  - More recognizable emotions

Contents

- Motivation
- Facial expressions
- Automatic Facial Expression Analysis
- Emotional databases
- Representative Facial Expression Recognition Systems
- **Conclusions**
- References

Conclusions

- Actual trend and desirable future:
  - Video Based.
  - 3D facial tracking.
  - HMM or similar for classification.
  - Coded Approach (FACS, FAPS).

Conclusions

- Ideal System:
  - Automatic facial image acquisition.
  - Subjects of any age, ethnicity and appearance.
  - Robust to variation in lightning.
  - Robust to partially occluded faces.
  - No special markers/make-up required.
  - Deals with rigid head motions.
  - Automatic face detection.
  - Automatic facial expression data extraction.
  - Deals with inaccurate facial expression data.
  - Automatic facial expression classification.
  - Distinguishes all possible expressions.
  - Deals with unilateral facial changes.
  - Obeys anatomical rules.

Contents

- Motivation
- Facial expressions
- Automatic Facial Expression Analysis
- Emotional databases
- Representative Facial Expression Recognition Systems
- Conclusions
- **References**

References

- Fasel2003: Fasel, B. and Luettin, J., *Automatic Facial Expression Analysis: A Survey.* Pattern Recognition, 2003. 36 (1). p:259-275
- Ioannou, S., et al., *Emotion recognition through facial expression analysis based on a neurofuzzy network.* Neural Networks, 2005. **18**(2005 Special Issue): p. 423-435.
- Yeasin, M., B. Bullot, and R. Sharma, *Recognition of facial expressions and measurement of levels of interest from video.* Multimedia, IEEE Transactions on, 2006. **8**(3): p. 500-508.
- Sebe, N., et al. *Emotion Recognition Based on Joint Visual and Audio Cues*. in *18th International Conference on Pattern Recognition* 2006.

Thank you for coming !