

Facial Emotion Recognition

Maher Ben Moussa

Katarzyna Wac

Prof. Nadia Magnenat-Thalmann

Prof. Dimitri Konstantas

What are emotions?

- Defining emotions is an everlasting debate
 - Categorical [1]?
 - Dimensional [2]?
 - Appraisal [3]?
- Facial Emotion Recognition concentrates mainly on categorical approach

[1] Ekman, P. (1982). *Emotion in the human face*. Cambridge, UK: Cambridge University Press.

[2] Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39,1161–1178.

[3] Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.).(2001). *Appraisal processes in emotion: Theory, methods, research*. New York: Oxford University Press.

Facial Expressions

- A face does not only express emotions, but also other social signals [1], such as:

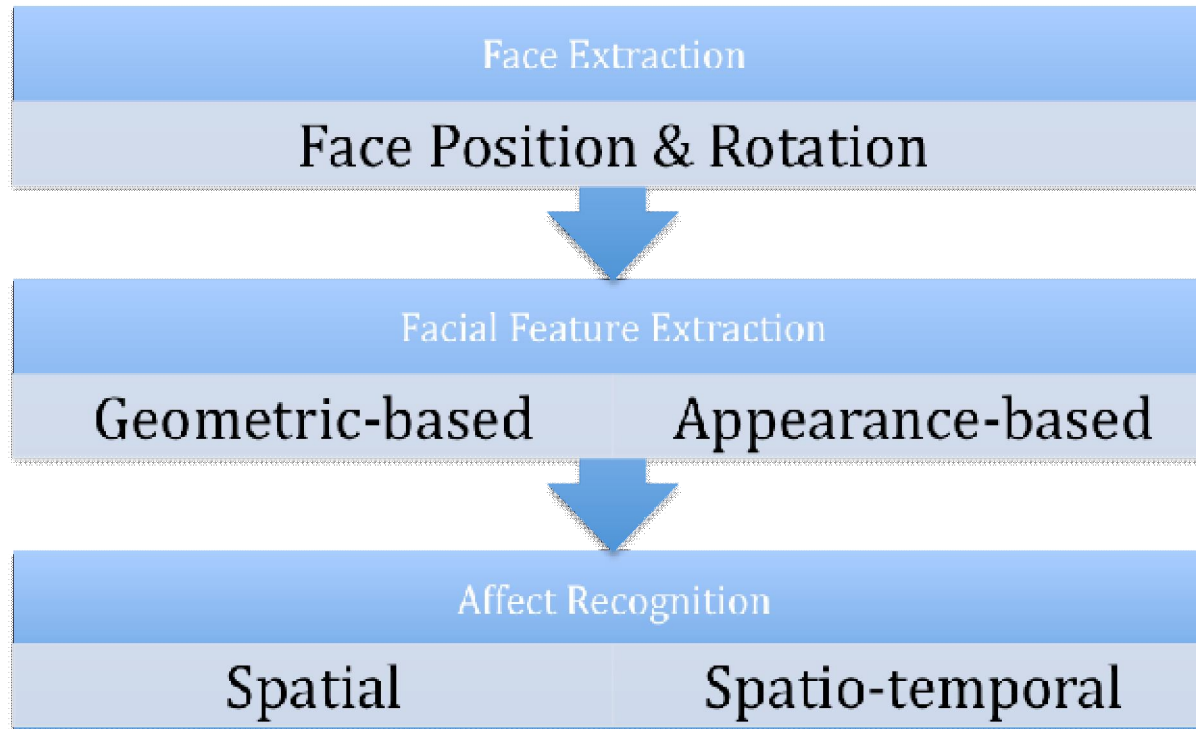
Personality	Status
Dominance	Persuasion
Regulation	Rapport
- Further, humans are in control of showing and hiding emotions

[1] Vinciarelli, A., Pantic, M., Bourlard, H. and Pentland, A. Social signals, their function, and automatic analysis: a survey, IEEE International Conference on Multimodal Interfaces (ICMI'08), 61-68, (2008)

Facial Emotion Recognition Systems

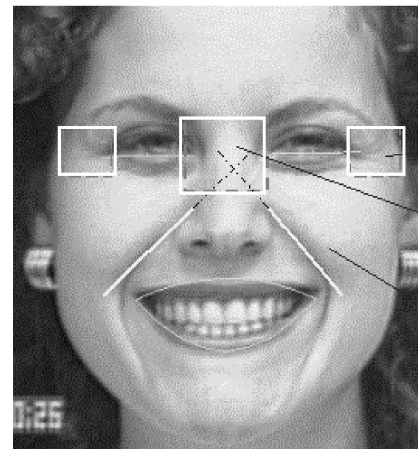
1. Analyse streaming video and track the facial feature points over time
2. Calculate feature vectors
3. Detect emotions according to trained models
4. Recognition models are trained using annotated database

Architecture



Facial Feature Tracking

- Geometric features [1] (Our approach)
 - The shapes of the facial components (eyes, mouth, etc.)
- Appearance features [2]
 - The texture of the facial skin (wrinkles, bulges, and furrows)



Area lateral to
eye outer corner

Nasal root

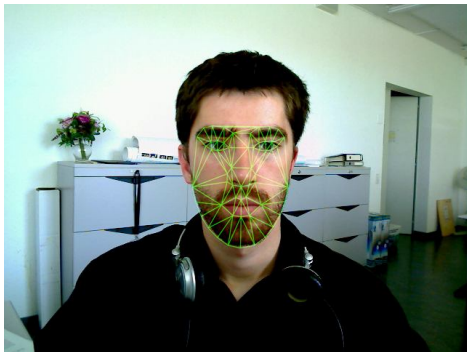
Nasolabial region

[1] M. Ben Moussa, N. Magnenat-Thalmann, Applying Affect Recognition in Serious Games: The PlayMancer Project. The Second International Workshop on Motion in Games (MIG 2009), Lecture Notes in Computer Science, Springer, 5884, pp. 53-62, 2009

[2] Tian, Y.L., Kanade, T. & Cohn, J.F. (2001). Recognizing action units for facial expression analysis. IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 23, No. 2, pp. 97-115.

Our tracking

1. Fitting the skeleton Active Shape Models
2. Refining lips by lips color segmentation
3. Refining tracking by template matching



Emotion Classification

- Using Spatial Information (Our approach)
 - Based on static facial expressions [1]
- Using Spatial-Temporal Information
 - Based on transformation of facial expressions over time [2]
 - Makes it possible to distinguish spontaneous emotional expressions from fake ones [3]
 - Difficulty in implementation results in lower accuracy

[1] L.S. Chen, R. Travis Rose, F. Parrill, X. Han, J. Tu, Z. Huang, M. Harper, F. Quek, D. McNeill, R. Tuttle, and T.S. Huang, "VACE multimodal meeting corpus," MLMI 2005

[2] M. Pantic and M.S. Bartlett, 'Machine Analysis of Facial Expressions (pdf file)', in Face Recognition, K. Delac & M. Grgic, Eds., Vienna, Austria: I-Tech Education and Publishing, 2007, pp. 377-416

[3] M.F. Valstar, H. Gunes and M. Pantic, 'How to Distinguish Posed from Spontaneous Smiles using Geometric Features ', in Proc. ACM Int'l Conf. Multimodal Interfaces (ICMI'07), pp. 38-45, Nagoya, Japan, November 2007

Databases

- The most well-known databases:
 - Cohn-Kanade AU-Coded Facial Expression Database [1]
 - FABO database [2]
 - MMI Facial Expression Database [3]
 - Belfast Naturalistic Emotional Database [4]
- Controlled Environment

[1] T. Kanade, J. Cohn, and Y. Tian, "Comprehensive Database for Facial Expression Analysis," Proc. IEEE Int'l Conf. Face and Gesture Recognition (AFGR '00), pp. 46-53, 2000.

[2] H. Gunes and M. Piccardi, "A Bimodal Face and Body Gesture Database for Automatic Analysis of Human Nonverbal Affective Behavior," Proc. 18th Int'l Conf. Pattern Recognition (ICPR '06), vol. 1, pp. 1148-1153, 2006.

[3] M. Pantic, M.F. Valstar, R. Rademaker, and L. Maat, "Web-Based Database for Facial Expression Analysis," Proc. 13th ACM Int'l Conf. Multimedia (Multimedia '05), pp. 317-321, 2005.

[4] E. Douglas-Cowie, N. Campbell, R. Cowie, and P. Roach, "Emotional Speech: Towards a New Generation of Database," Speech Comm., vol. 40, nos. 1/2, pp. 33-60, 2003.

Results

Nr	Type	Classes	Accuracy
[1]	Spatial	6	73%
[2]	Spatial	8	80%-100%
[3]	Spatial	4	93%-95%
[4]	Spatial-Temporal	6	77%
[5]	Spatial-Temporal	2	94%

- [1] L. Cohen, N. Sebe, A. Garg, L. Chen, and T. Huang, "Facial Expression Recognition from Video Sequences: Temporal and Static Modeling," *Computer Vision and Image Understanding*, vol. 91, nos. 1-2, pp. 160-187, 2003.
- [2] H. Gunes and M. Piccardi, "Affect Recognition from Face and Body: Early Fusion versus Late Fusion," *Proc. IEEE Int'l Conf. Systems, Man, and Cybernetics (SMC '05)*, pp. 3437-3443, 2005.
- [3] N. Sebe, M.S. Lew, I. Cohen, Y. Sun, T. Gevers, and T.S. Huang, "Authentic Facial Expression Analysis," *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition (AFGR)*, 2004.
- [4] R. El Kaliouby and P. Robinson, "Real-Time Inference of Complex Mental States from Facial Expression and Head Gestures," *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition (CVPR '04)*, vol. 3, p. 154, 2004.
- [5] M.F. Valstar, H. Gunes, and M. Pantic, "How to Distinguish Posed from Spontaneous Smiles Using Geometric Features," *Proc. Ninth ACM Int'l Conf. Multimodal Interfaces (ICMI '07)*, pp. 38-45, 2007.

MultiModal Affect Fusion

Maher Ben Moussa

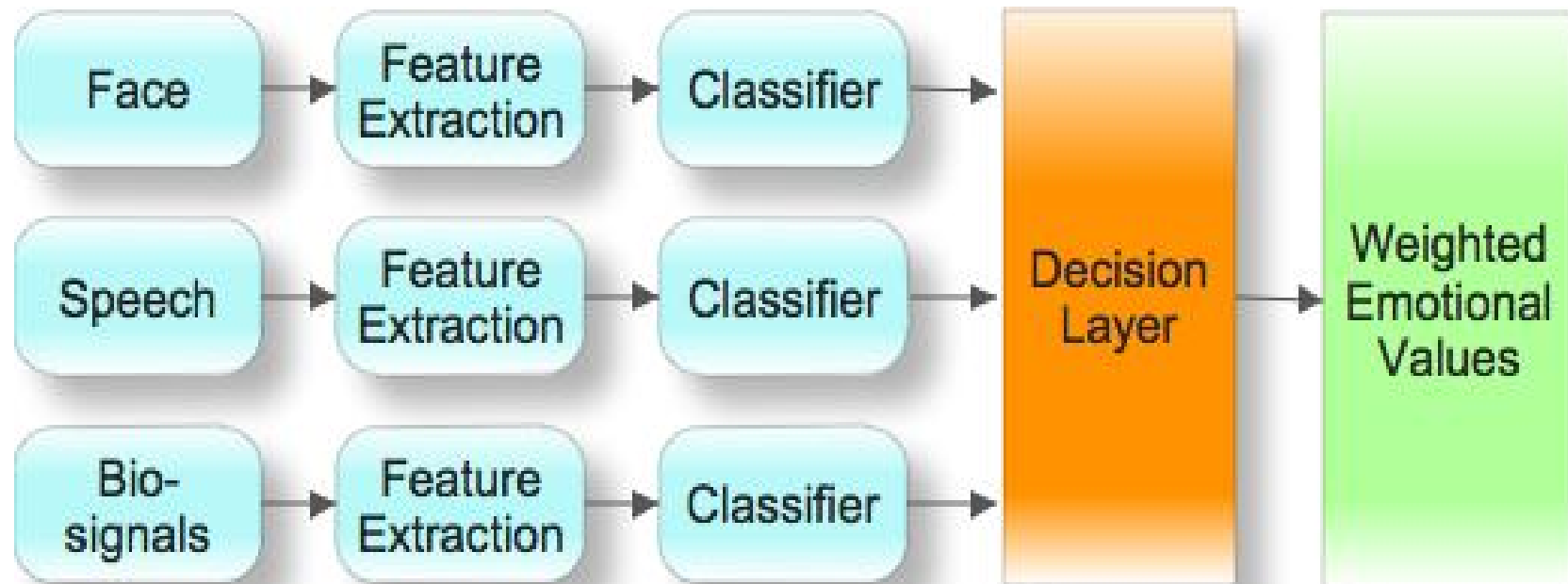
Katarzyna Wac

Prof. Nadia Magnenat-Thalmann

Prof. Dimitri Konstantas

Fusion of Multi-Modalities

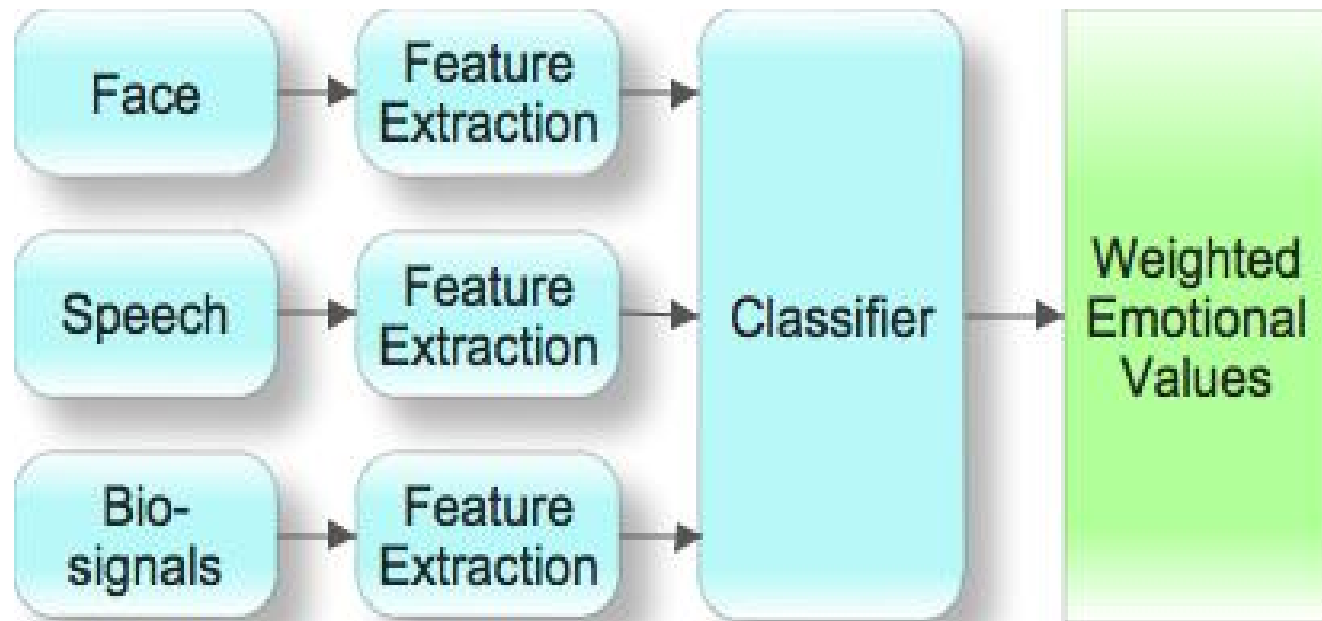
- Decision Level Based



[1] M. Ben Moussa, N. Magnenat-Thalmann, Applying Affect Recognition in Serious Games: The PlayMancer Project. The Second International Workshop on Motion in Games (MIG 2009), Lecture Notes in Computer Science, Springer, 5884, pp. 53-62, 2009

Fusion of Multi-Modalities

- Feature Level Based



[1] J. Kim, E. Andre, M. Rehm, T. Vogt and J. Wagner, Integrating information from speech and physiological signals to achieve emotional sensitivity, in Proc. 9th Eur. Conf. Speech Communication and Technology, Lisbon, Portugal (2005)

Static Fusion

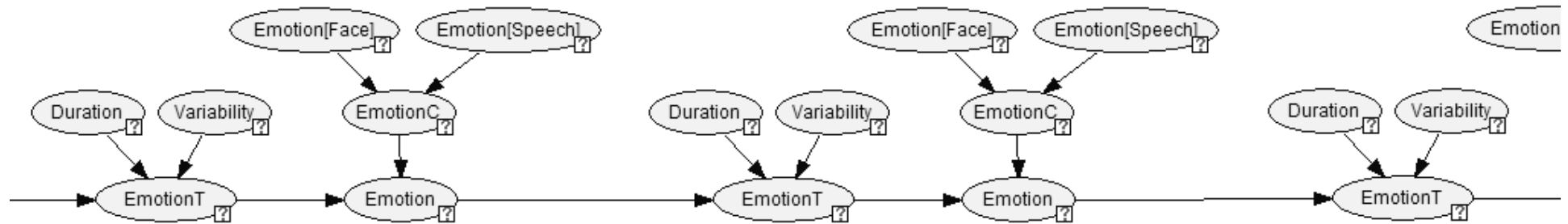
- Feature Level Based [1]
 - Features from different modalities are classified by one classifier
 - Dimensionality reduction is applied
- Decision Level Based [2]
 - Classified emotional values are fused according to the confidence
 - Trained on ground truth values

[1] J. Kim, E. Andre, M. Rehm, T. Vogt and J. Wagner, Integrating information from speech and physiological signals achieve emotional sensitivity, in Proc. 9th Eur. Conf. Speech Communication and Technology, Lisabon, Portugal (2005)

[2] M. Ben Moussa, N. Magnenat-Thalmann, Applying Affect Recognition in Serious Games: The PlayMancer Project. The Second International Workshop on Motion in Games (MIG 2009), Lecture Notes in Computer Science, Springer, 5884, pp. 53-62, 2009

Temporal Fusion

- Next to confidence levels, historical information is also taken into consideration to determine the next emotion
- Feature Level Based [1]
 - Changes in the features over time
- Decision Level Based [2] (our approach)
 - Duration of previous detected emotion
 - Variability of previous detected emotion



[1] M. Song, J. Bu, C. Chen, and N. Li, "Audio-Visual-Based Emotion Recognition: A New Approach," Proc. Int'l Conf. Computer Vision and Pattern Recognition (CVPR '04), pp. 1020-1025, 2004.

[2] D. Datcu and L.J.M. Rothkrantz, "Semantic audio-visual data fusion for automatic emotion recognition," in Euromedia' 2008, Porto, 2008.

Results

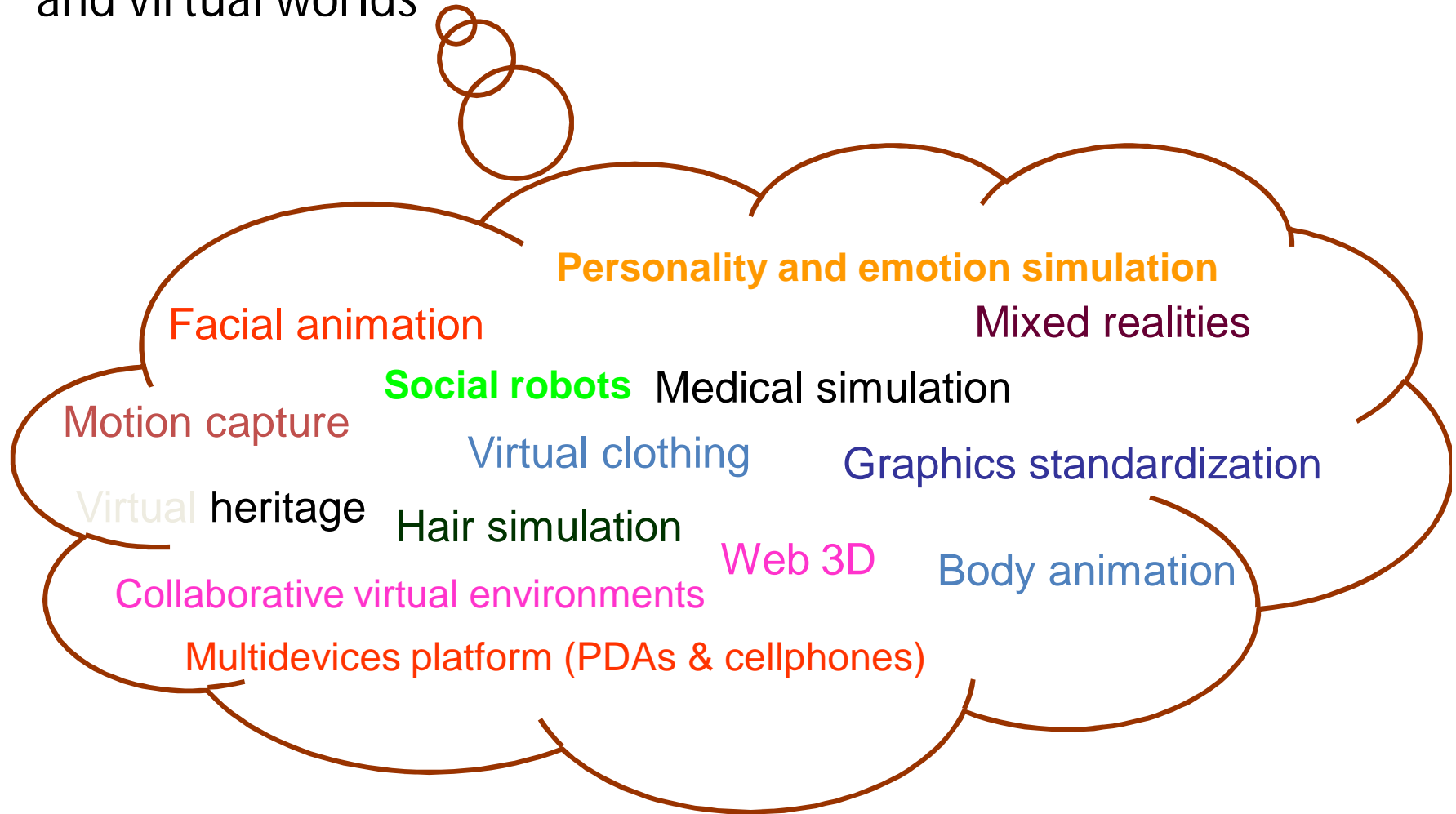
Nr	Face	Speech	Bio	Level	Type	Acc.
[1]	X	X		Decision	Static	95%-98%
[2]	X	X		Feature	Static	84%-87%
[3]	X	X		Decision	Temporal	89%
[4]	X	X		Feature	Temporal	84%
[5]		X	X	Decision	Static	54%
[5]		X	X	Feature	Static	51%

- [1] H.J. Go, K.C. Kwak, D.J. Lee, and M.G. Chun, "Emotion Recognition from Facial Image and Speech Signal," Proc. Int'l Conf. Soc. of Instrument and Control Engineers, pp. 2890-2895, 2003.
- [2] Z. Zeng, Z. Zhang, B. Pianfetti, J. Tu, and T.S. Huang, "Audio-Visual Affect Recognition in Activation-Evaluation Space," Proc. 13th ACM Int'l Conf. Multimedia (Multimedia '05), pp. 828-831, 2005.
- [3] Z. Zeng, J. Tu, M. Liu, T.S. Huang, B. Pianfetti, D. Roth, and S. Levinson, "Audio-Visual Affect Recognition," IEEE Trans. Multimedia, vol. 9, no. 2, pp. 424-428, Feb. 2007.
- [4] M. Song, J. Bu, C. Chen, and N. Li, "Audio-Visual-Based Emotion Recognition: A New Approach," Proc. Int'l Conf. Computer Vision and Pattern Recognition (CVPR '04), pp. 1020-1025, 2004.
- [5] Kim, J., André, E., Rehm, M., Vogt, T., & Wagner, J. (2005). Integrating information from speech and physiological signals to achieve emotional sensitivity. Proc. of the 9th European Conference on Speech Communication and Technology.

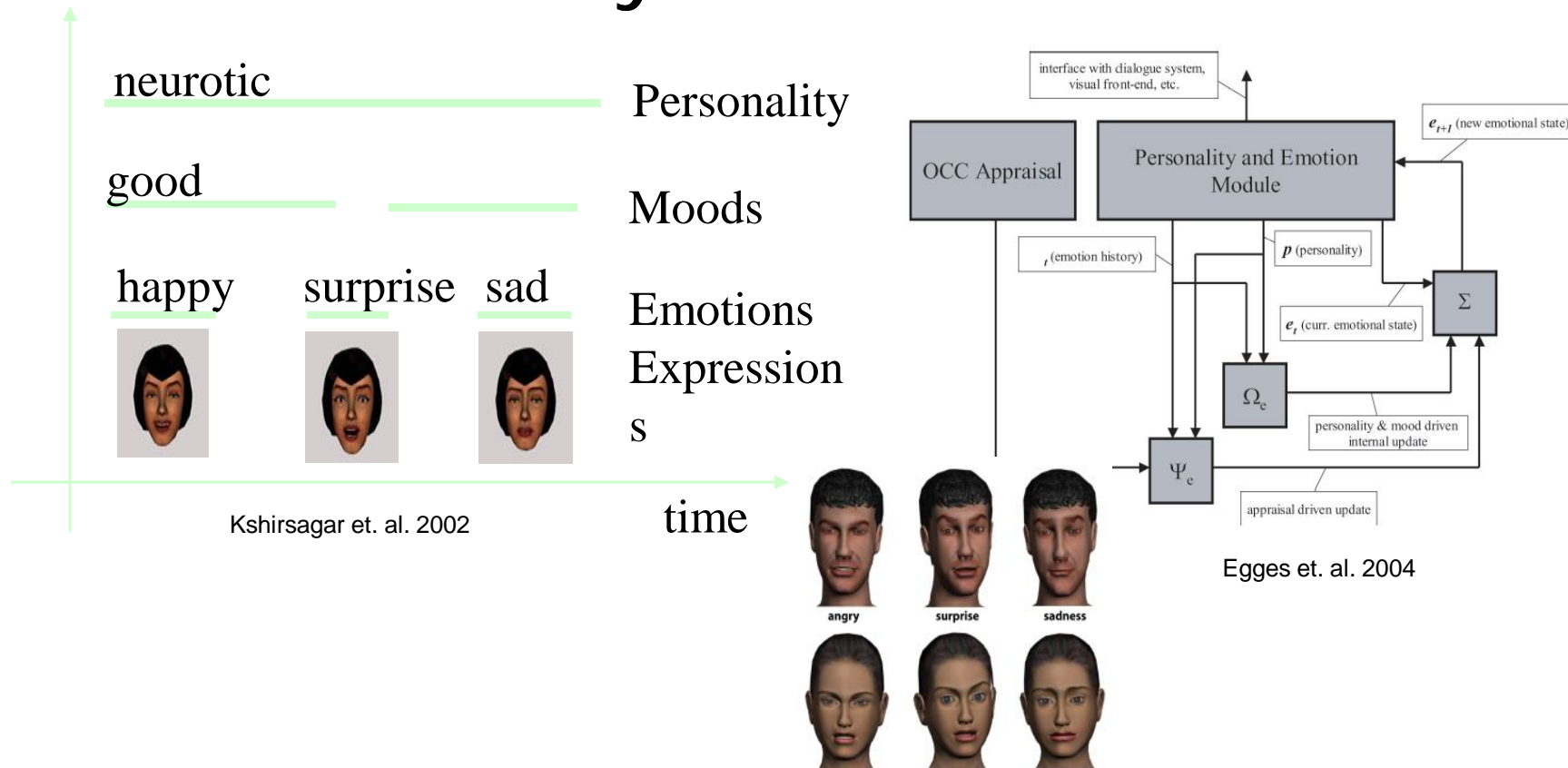
Other Research About Emotions

MIRALab

- A pluridisciplinary lab working on virtual human simulation and virtual worlds



Personality/Emotion Simulation



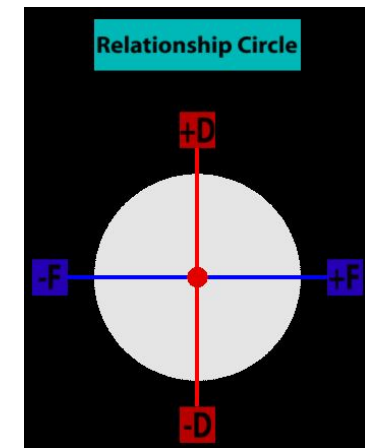
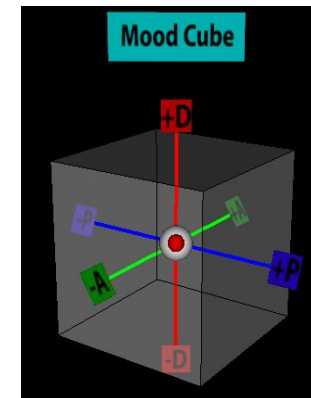
-A. Egges, S. Kshirsagar, N. Magnenat-Thalmann. **Generic Personality and Emotion Simulation for Conversational Agents**. *Computer Animation and Virtual Worlds, John Wiley and Sons, Vol. 15, No. 1, pp. 1-13, January 2004.*

-S. Kshirsagar, N. Magnenat-Thalmann. **A Multilayer Personality Model**. *Proceedings of 2nd International Symposium on Smart Graphics, pp. 107-115, June 2002.*

-Z. Kasap, M. Ben Moussa, P. Chaudhuri, N. Magnenat-Thalmann. **Making Them Remember—Emotional Virtual Characters with Memory**. *IEEE Computer Graphics and Applications, 29, 2, pp. 20-29, March 2009.*

Emotion Engine

- OCEAN personality*
 - Openness, Conscientiousness, Extroversion, Agreeableness, Neuroticism
- OCC emotions**
 - Positive: Joy, hope, relief, pride, gratitude, love
 - Negative: Distress, fear, disappointment, remorse, anger, hate
- Mehrabian Mood***
 - 3 dimensional: pleasure, arousal, dominance
 - Positive: Exuberant, dependant, relaxed, docile,
 - Negative: Bored, disdainful, anxious, hostile
- Relationship****
 - Friendliness and dominance



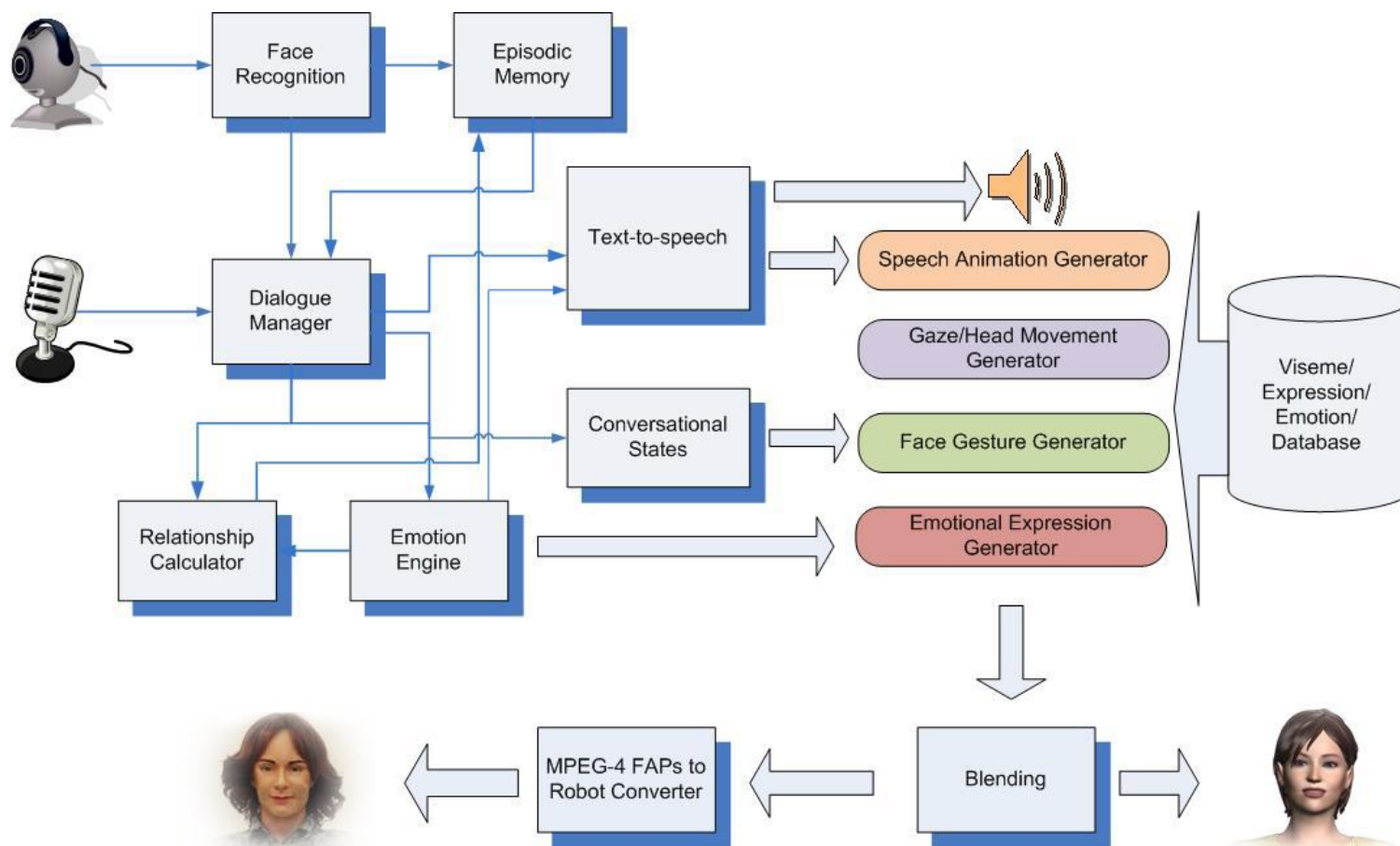
* S. Hampson, **State of the art: Personality**. In *The Psychologist*, 12(6):284--290, 1999.

** A. Collins, A. Ortony, and G. L. Clore., **The Cognitive Structure of Emotions**. Cambridge University Press, New York, 1988.

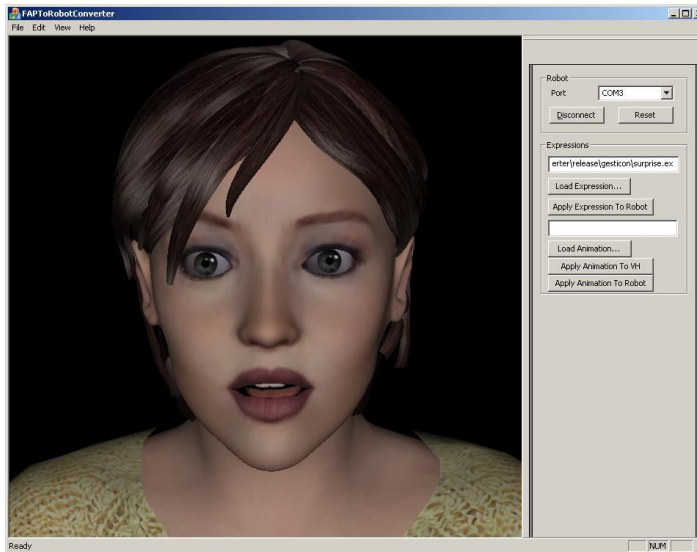
*** A. Mehrabian, **Analysis of the big-five personality factors in terms of the PAD temperament model**, *Australian Journal of Psychology*, 48(2):86--92, 1996.

**** M. Argyle, **Bodily Communication**. Second Edition, Methuen and Co Ltd., London

Ongoing Research



Virtual Humans & Robotics



- Thank you for your attention