Cheering you up in empathic rooms

Intelligent Systems Laboratory — Winter Term 2013/2014

Christian Limberg & Benjamin Wolff

Supervisors: Hana Boukricha, Felix Hülsmann

Bielefeld University, Faculty of Technology

Universität Bielefeld

Abstract

We present a system that automatically detects faces in a webcam image and perfoms real-time emotion-detection, with respect to four emotions (neutral, happy, sad, surprise, disgust). Depending on the emotion, the system plays a suitable melody. Emotion-detection is done on each frame and divided into two steps. First, we perform a Support vector machine-based recognition of Action Units, according to the Facial Action Coding System (FACS)[3]. Second, we map these Action Units to emotions. We measured performance of our system compared to an average human observer.

Introduction

The human face is a rich and powerful source. In face-to-face communication it reflects our feelings and emotions, thus it is often called ”the window to the soul”. Therefore, the use of facial expressions in Human Computer Interaction gains more and more interest. How can an intelligent systems recognize and use these informations? How can an intelligent system adapt to the users emotion? Which features can be extracted for processing? Our focus in Winter-Term 2013/2014 is on how supervised learning techniques can be applied to detect emotions in video-streams, rather than actuate components in the Intelligent Systems Laboratory.

Main Objectives

1. Perform face detection and feature extraction from webcam-image
2. Learn FACS-labelled data from CK+ database [5]
3. Create ruleset for mapping Action Units to emotions
4. Develop real-time face-detection with online emotion classification
5. Actuate music depending on recognized emotion
6. Evaluate generalization performance

Materials, Methods, Domain Models, Architecture

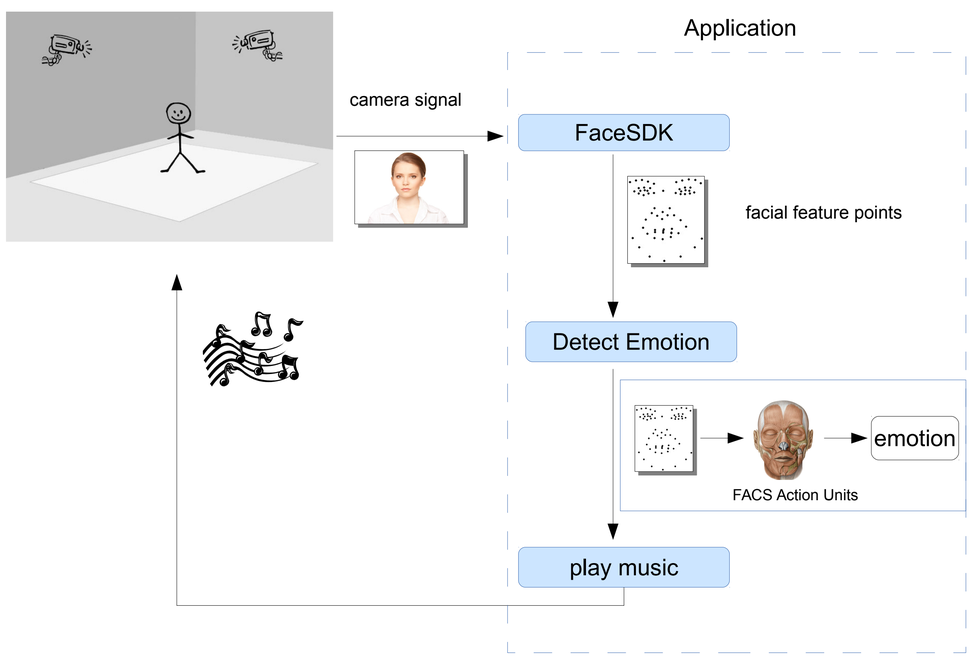

Figure 1: Online-classification system with camera input and audio feedback

In our work, we used Luxand FaceSDK[1], an application programming interface, which provides methods for face detection and face tracking. Furthermore FaceSDK easily allows to extract facial features. The facial features consist out of 66 2D-Points, placed on defined locations in the face. We used this method for feature extraction. For supervised learning we decided to use support vector machines with linear kernel. We used libsvm [2] for training and online-classification. The extended Cohn-Kanade-Database (CK+) was used to train our SVMs for Action Unit(AU) detection. The CK+-Database contains 593 FACS-labelled sequences from 123 subjects with a resolution of 640x490 pixels. We used a linear non-vs-all two-class SVM for each AU-Detection. The training-set of the linear SVM consisted of three peak-frames from sequences containing the AU as positive samples. All neutral frames and peak-frames not containing the AU were used as negative samples. For our System, we used 13 Action Units (AU1, AU2, AU4, AU5, AU6, AU7, AU9, AU10 ,AU12, AU15, AU20, AU24, AU25). We decided to resize all facial features to a interocular distance of 77px and scaled the features in the range [-1,1].

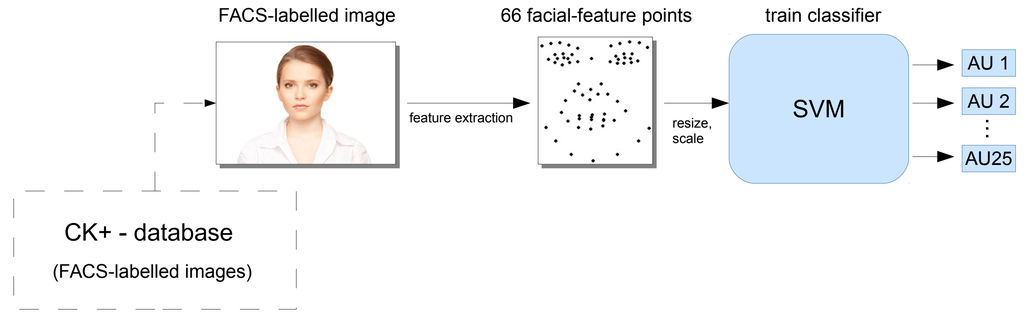

Figure 2: SVM training based on CK+-Database

We used the Karolinska Directed Emotion Faces (KDEF) [6] to measure generalization performance of our system. Furthermore we compared our results with an an average human observer, determined in a validation study on KDEF-database[4].

Results

Generalization of Action Units was tested using cross-validation on CK+-database. The system achieved a performance of 94,17%.
To estimate the performance of emotion detection we used the KDE-Database. The results are shown in table 1 and compared to an average human observer.

Emotion	Human Observer	CUEP-System	CUEP-Performance
Neutral	62.64%	90.00%	+27.36%
Happy	92.65%	92.90%	+0.25%
Sad	76,70%	72.70%	-4.0%
Surprised	96.00%	95.70%	-0.3%
Disgust	72.17%	53.60%	-18.57%

Table 1: Emotion detection performance on KDEF-Database compared to an average human observer

Conclusions and Future Work

- For happy, sad and surprised emotion our system achieved classification results that are comparable to the human and even a better performance in detecting neutral emotion.
- System is error prone to outside influences (e.g. lighting, camera quality, camera-to-face-distance, fast head movements,...)-
- For summer term we plan to actuate mulitple components in the intelligent room.
- We plan to integrate two more emotions: Anger and fear.
- We will continue classification improvements.

References

[1] Luxand facesdk. http://www.luxand.com/facesdk/.

[2] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011.

[3] P Ekman, WV Friesen, and JC Hager. Facs manual. *A Human Face*, 2002.

[4] Ellen Goeleven, Rudi De Raedt, Lemke Leyman, and Bruno Verschuere. The karolinska directed emotional faces: a validation study. *Cognition and Emotion*, 22(6):1094–1118, 2008.

[5] Patrick Lucey, Jeffrey F Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, and Iain Matthews. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 94–101. IEEE, 2010.

[6] Daniel Lundqvist, A Flykt, and A Öhman. The karolinska directed emotional faces. *Stockholm, Sweden: Karolinska Institute*, 1998.