Face expression recognition: A brief overview of the last decade

Conference Paper · May 2013

DOI: 10.1109/SACI.2013.6608958

Cătălin Căleanu
Polytechnic University of Timisoara

Face Expression Recognition: a Brief Overview of the Last Decade

Cătălin-Daniel Căleanu
Faculty of Electronics and Telecommunications
"POLITEHNICA" University from Timişoara
Timişoara, România
catalin.caleanu@etc.upt.ro

Abstract—**The huge research effort in the field of face expression recognition (FER) technology is justified by the potential applications in multiple domains: computer science, engineering, psychology, neuroscience, to name just a few. Obviously, this generates an impressive number of scientific publications. The aim of this paper is to identify key representative approaches for facial expression recognition research in the past ten years (2003-2012).**

I. INTRODUCTION

The scientific study of facial expression showed that it represents one of the most powerful and immediate means for emotions and intentions communication [1]. Due to its potential applications e.g., natural human-machine interface, behavioral science, clinic practice, automatic facial expression recognition has attracted much attention which has been manifested through a large number of scientific publications. As a consequence, numerous methods were proposed. For comprehensive surveys of the past efforts in the field, readers are referred to [2], [3] or [4]. Despite these efforts, recognizing facial expression with a high accuracy remains to be difficult and still represents an active research topic [5].

The present article is intended to be a high-level survey over the facial expression recognition research that has been carried out in the past ten years, between year 2003 and 2012. Due to length restrictions, only a small sample of recognition techniques is explicitly referred to. More exactly, we have selected 10 papers that we felt were important and different from each other. Also the novelty of the proposed technique, the number of the citations and the number of downloads were taken into account in the selection process.

The interest in creating such an overview is multifarious. By selecting the most interesting approaches, we want to focus the attention to new techniques and methodologies that may be of high interest to the researchers in the field of facial imagery. Moreover, this selection can be a useful indicator of the areas that will constitute the future research trends.

The organization of the remainder of this paper is as follows. We begin, in Section II, with a brief description of approaches which enable us to elaborate the taxonomy of face expression recognition methods. We then review approaches with respect to image acquisition and pre-processing operations (Section III). In Section IV, feature extraction and selection techniques are presented. The classification solutions are discussed in Section V along with the experimental results (Section VI). Conclusions are stated in Section VII.

II. TAXONOMY AND DESCRIPTION OF APPROACHES

In the last decade, the field of expression analysis has grown immensely. The proposed approaches can be classified from many perspectives, which will be analyzed below.

A. Taxonomy

From the feature extraction point of view, the techniques aiming the recognition of face expression can be categorized into methods that use appearance features and methods that use geometric features, although some hybrid-based approaches are also possible. The use of a face model/template constitutes a distinct approach, and it is typically referred as a template-based method.

In the first category, the face images are processed by an image filter or filter banks on the whole face or some specific regions of the face to extract changes in facial appearance. Typically, the entire facial images or some specific facial regions are convolved with some filters, e.g. Gabor wavelets, and the extracted filter responses at the fiducial points (manually selected, most of the times) form vectors that are further used for facial expression classification. Also, principal component analysis (PCA), variants of independent component analysis (ICA) or local binary pattern (LBP) are most often used in a holistic manner.

In the geometric feature extraction system, the shape, distances, angles or the coordinates of the fiducial points form a feature vector that represents facial geometry.

It seems that the highest recognition rates have been obtained when both the responses methods were combined [6].

From the temporal perspective, facial expression recognition techniques are represented by static (typically using still images) and dynamic (image sequences) approaches.

Other categories are the global techniques - which analyze the texture of the whole face without having explicit knowledge about the location of single facial features - versus local approaches - which try to extract local features of the face or to fit any holistic face model containing a set of feature points to the face.

Finally, we could classify the methods according to the image data type in 2D and 3D approaches.

B. Description of Approaches

The facial action coding system developed by Ekman [7] is a source of inspiration for many research papers. In the first paper that has been analyzed, Cohen et al. [8]

978-1-4673-6400-3/13/$31.00 ©2013 IEEE

present a method for recognizing the emotions through facial expressions presented in a video sequence is discussed. The novelty of this work consists in introducing the Tree-Augmented-Naive Bayes (TAN) classifier which incorporates the dependencies between features. This is in opposition to the Naïve Bayes (NB) assumption in which features are considered to be independent as in work of Sebe et al. [9].

The method presented by Ma and Khorasani [10] is based on a combination of a two-dimensional discrete cosine transform (2D-DCT) and constructive one-hidden-layer feedforward neural network. The novelty of the approach comes from the proposed pruning technique which substantially reduces the size of the neural network while improving the generalization capability and the recognition rate.

A novel low-computation discriminative feature space is introduced in [11] by Shan et al. It is based on Local Binary Patterns (LBP) features which could be extracted faster. Moreover, it is shown that LBP features are robust to low resolution. As possible solutions for the classification stage, template matching and Support Vector Machine (SVM) were considered and compared. Also, a comparison with the previous mentioned work of Cohen et al. [8] (geometric features +TAN) proved the superiority of the proposed technique.

Since the year 2006, little work has been done in 3D based face expression recognition. One of the first attempts is represented by the work of Zeng et al. [12] who employed a 3D face tracker for feature extraction. In the same year, in the approach proposed by Wang et al. [13], the classification of the prototypic facial expressions is performed by extracting primitive 3D facial expression features, and by calculating the feature distribution. Other innovative 3D solutions are the two approaches of Dornaika et al. [14] and Kotsia & Pitas [15] which use AUs together with a 3D face model named Candide, initially proposed by Ahlberg [16]. Some of the most important advantages of these techniques are the texture independence, the view independent (since the used tracker simultaneously provides the 3D head pose and the facial actions) and a simple learning phase which only need to fit second-order auto-regressive models to sequences of facial actions. For an excellent 3D facial expression survey please consult the work of Fang et al. [17].

One of the most interesting approaches is presented by Panning in [18]. The authors propose a novel approach from multiple perspectives e.g., it use facial feature detection in color image sequences and the feature extraction is initialized without manual input. For expression classification a three layer feed forward artificial neural network is employed. In the next work, Buciu et al. [19] present a comparison of ICA approaches whereas the classification stage is implemented using either a Cosine Similarity Measure (CSM) or a SVM classifier.

In the paper proposed by Jabid and Chae [20], a new appearance based feature descriptor, called the local directional pattern (LDP), has been proposed to represent facial geometry. Template matching and support vector machine are used for classification. A novel facial expression recognition technique based on a sparse representation is proposed by Jia in [21]. It is also shown that a multi-layer sparse representation provides better experimental results than the conventional sparse representation.

The last selected paper, proposed by Valstar and Pantic [22], enables the detection of much larger range of facial expressions.

III. Image Acquisition and Pre-Processing

Facial data can be acquired from a database, a live video stream or other sources, in 2D or 3D, both in a static or dynamic mode. The most popular type of pictures is the 2D grey scale facial images. Typically, this step is followed by some pre-processing (noise removal, light compensation, detection, normalization, tracking, etc.) operations. In the context of the selected papers, these two steps are detailed in the following.

In the work of [8] the images are acquired from a video stream using a face tracking algorithm and a 3D wireframe model (16 surface patches embedded in Bézier volumes) proposed by Tao and Huang [23].

A normalization process was employed by Ma et al. [10] in which the centers of the eyes and mouth are taken as the reference points. A fixed distance d between the centers of the eyes represents a first normalization criterion in [11]. An interesting observation is further formulated and applied as a second pre-processing step. It refers to the face dimensions: the width of selected face is roughly 2d and the height is roughly 3d. No illumination compensation is required, due to the LBP's gray scale invariance.

In [13], the surface feature analysis is based on the triangle meshes of faces, which are created by a 3dMD static digitizer [24], which uses the principle of light pattern projection. . In another 3D approach [14], the local facial actions and deformations are calculated using an appearance-based tracker which simultaneously computes the 3D head pose and the facial actions.

The work of [18] is innovative from the perspective of pre-processing step due to the particular modality in implementing the face detection stage. More exactly, in order to minimize the false positives detections in complex backgrounds, a combination of Haar-like-Feature detection and skin color detection has been proposed.

The following pre-processing steps were proposed in [19]: a registration operation based on the manually identified position of eyes followed by a rotation of the image to horizontally align the face according to eyes; the final 60 x 45 pixels facial image was obtained by cropping and downsampling operations.

The images used in [20] were cropped from the original C-K database one using the positions of two eyes and resized into 150×110 pixels. The height of the image is 2.7d with level of eye located 2d apart from bottom boundary, where d represents the distance between the eyes. No attempt was made for illumination compensation, since LDP is robust in illumination change.

The fixed distance of 60 pixels between the eyes represents a normalization criterion for the work presented in [21]. The final cropped face had the width of two times this distance and the height roughly three times.

The first step of the system proposed in [22] is represented by the well-known Viola-Jones face detector [25]. For the next step, characteristic facial point detection, the solution proposed by Vukadinovic and Pantic [26] was chosen. Registration and smoothing operations represents other steps in the processing scheme proposed by Valstar in [22].

IV. Feature Extraction

One of the most critical aspects for any successful facial expression recognition system is extracting the best features to describe the physical phenomena. The efficiency (minimizing within-class variations of expressions while maximizing between-class variations, low-dimensional feature space, etc.) and effectiveness (can be easily extracted from the raw face image) representation of the facial images would provide robustness during recognition process.

From this perspective, the work of Cohen et al. [8] uses the features proposed by Sebe et al. [9] which consist of 12 facial motion measurements relative to a 3D wireframe model (16 surface patches embedded in Bézier volumes) proposed by Tao and Huang [23].

The 2-D discrete cosine transform (DCT) compression technique was employed in [10] to the difference image obtained by subtracting a neutral image from a given expression image. In the frequency domain, only the coefficients of the lower frequencies (having large amplitudes) were considered as input vectors for the following classifier.

An extended LBP (neighborhood of different sizes and uniform patterns) is proposed in [11] as feature extraction method. In this way a face is divided into small regions from which LBP histograms are extracted and concatenated.

Research on face expression recognition has mainly relied on 2D images and, obviously, has certain limitations. See Liu and Ward [27] for arguments in the favor of 3D facial imagery processing. The next two approaches employed 3D facial data. Wang et al. [13] use a geometric feature based facial expression descriptor in the 3D Euclidian space. Dornaika et al. [14] inferred the facial expression from the temporal representation of a 3D model which includes head pose (three rotations and three translations) and a facial action vector. Obviously, the system performance will be dictated by the ability to accurately track the local facial actions/deformations.

Panning [18] consider, in an innovative manner, both static (10 distances) and transient (3 regions for texture analysis) feature types. The set of distances are measured between static feature points and the transient features are detected by the appearance of edges in predefined areas (forehead, the nose bridge, and the nasolabial fold).

The work of Buciu [19] take ICA approach as baseline for feature extraction and compare it with five additional ICA flavors.

Jabid et al. [20] proposes the LBP operator which encodes the micro-level information of edges, spots, and other local features in an image. It. computes the edge response values in different directions and uses these to encode the image texture. After computing all the LDP code for each pixel, the input image is represented by an LDP histogram which represents a descriptor of that image.

Jia et al. [21] divide the image into several 20x20 pixel patches, then an enhanced LBP operator is used as feature descriptor. Because some face regions provide more important information than others, different face regions should be given different weights.

The features computed in [22] from the 20 tracked fiducial points are: the positions of these points, the distances between pairs of points and the angle that the lines make with the horizontal axis. Finally, some temporal information is added.

V. Classification

Expression categorization is a process of assigning observed data to one of predefined facial expression categories by a classifier. A wide range of classifiers, covering parametric as well as non-parametric techniques, has been applied to the automatic expression recognition problem: nearest neighbor classifiers, Fisher's linear discriminant, ANNs, HMMs, SVMs or random forests are typical examples .

The classification scheme of [8], called Tree-Augmented-Naive Bayes (TAN) classifier, is based on an acyclic graphical model with the class and features as the nodes, and the dependencies represented by the directed edges in the graph between the nodes, forming in this way a Bayesian network. The simple template matching with weighted Chi square statistic and SVM are adopted as classification solutions in [11] and further compared. The generalization performance for SVM using RBF kernel provided a substantial increase in the recognition rate when compared with template matching or the previous presented classification scheme, TAN [8].

In [10], for implementing the classification stage, the application of an adaptive constructive one-hidden layer feedforward neural network to face expression recognition was considered. In this approach new hidden units/layers are added incrementally. Another ANN based approach is presents in [18] where, for classification purpose, a full connected feed-forward artificial neural network is trained and used. It is implemented using the open source FANN library [28].

The work presented in [13] used in the experiments four popular classifiers: Quadratic Discriminant Classifier (QDC), Linear Discriminant Analysis (LDA), Naive Bayesian Classifier (NBC), and Support Vector Classifier (SVC) with RBF kernel.

In order to evaluate the affect state, the authors of [14] calculate the cosine of the angle between two vectors which represents the actual expression and the synthesized universal expression trajectories. Based on this parameter, a normalized similarity measure is determined and used in the recognition process. In [19], two different classifiers are employed. The first is represented by the CSM classifier which is based on the nearest neighbor rule; the second classifier is the Support Vector Machine. SVM makes binary decisions, thus, in the work of Jabid [20], the one-against rest technique is adopted.

The key concept of the sparse representation expression recognition presented in [21] is the evaluation of the classes from the training samples for the minimum reconstruction error based on the sparse coding coefficients. Furthermore, in order to deal with subtle facial expressions, a multi-layer sparse representation method is proposed to improve recognition performance of multi- intensity expressions.

The approach presented in [22] perform a SVM based AU recognition. The choice is motivated by the high dimensionality of the feature space which, in this case, will not affect the training time. For the problem of AU temporal model detection, two approaches were compared: multiclass SVM (mc-SVM) and Hybrid SVM-HMM.

VI. Experimental Results

The development of robust facial expression recognition algorithms requires well labeled databases of sufficient size that include carefully controlled variations of pose, illumination and resolution. These databases can be categorized into 2-D image, e.g., JAFFE database [29] or MMI database [30], 2-D video (arguable the most used database of this type is Cohn-Kanade AU-Coded database [31]), 3-D static (BU-3DFE [32]) and 3-D dynamic (BU-4DFE [33]).

Some researchers have employed custom/in-house database. For example, the experiments described in [8] are based on data collected from video sources in which a set of five people are displaying six emotions. Two types of evaluations were performed: person dependent and independent experiments. The average expression recognition accuracies reported were 83.31% and 65.11% respectively. Another custom database was created in [10]. It contains images of size 128 x 128 with 256 gray levels expressing just four affects: smile, anger, sadness, and surprise are the four specific facial expressions of interest. The reported testing recognition rate was 93.75%.

In [18] for training purpose the authors interactively placed landmarks on images of 15 different people of the "Feedtum Facial Expression Database" from the Technical University of Munich [34]. The reliability of the feature detection and tracking system was tested on image sequences of the Smart-Kom database [35]. On recall of the training data as well as other additional 30 samples (which not had been included for learning) it achieved good generalization with almost 100% correct classification.

The next paragraph refers mainly to the result reported against Cohn-Kanade (C-K) Facial Expression Database. The comparisons provided in the work of [11] showed that the LBP-based SVM (87.6%) outperform both Gabor-based SVM (86.9%) and Geometric Feature + TAN (73.2%) on C-K database. The principal advantage of this approach lies in the simplicity of LBP histogram which requires much less computational resource. The recognition rates reported by [14] were obtained using two different training sets, the C-K database and a custom data set consists of five 30 second videos. From 101 played expressions there were 14 misclassified expressions leading to a recognition rate of 86.14%. The experiments reported in [19] have been performed using two databases of facial expression images: C-K AU-coded facial expression database in which thirteen persons expressed six basic emotions and Japanese female facial expressions (JAFFE). The best results obtained with leave-one-out are 87.6% for C-K database using fastICA + SVM polynomial and 81% for JAFFE using extended infomax and SVM-RBF. Jabid et al. [20] carried out a 7-fold cross-validation scheme where each dataset is randomly partitioned into seven groups separately. Six groups were used as a training dataset to train the classifiers or model their templates, while the remaining groups were used as testing datasets. The results of [21] were reported against C-K+ (Extended Cohn-Kanade Dataset) using a 10-fold cross validation testing scheme. The best reported result is the recognition rate of 87% using the expression recognition based on sparse representation. In the work of [22], four databases were used (the C-K database, the MMI facial expression database, the DS118 data set of spontaneous facial displays, and the triad data set of spontaneous human behavior) for various sets of experiments When evaluated for the C-K database (six-basic-emotion detection system on 153 videos), a 91.7% recognition rate was reported and a 95.3% accuracy for MMI database (244 videos). These results are among the best ever reported and show that appearance-based approaches do not necessarily outperform geometric-feature-based approaches.

To date, there are few publicly available 3D databases designed specifically for expression analysis. This is probably the reason for having fewer reported results. From our survey, the test of [13] is on the six prototypic expressions using the data from BU-3DFE database captured from 60 subjects with two high intensity models for each expression. The highest correct recognition rate of 83.6% was obtained using LDA classifier.

VII. Conclusion

This paper's objective is to survey and discuss recent advances in face expression recognition. More exactly, face expression recognition systems have improved a lot over the past decade. It is almost impossible to cover all of the published work so we have selected 10 papers, published since 2003 till date, that we think were relevant for this topic. The main emphasis has been on a review of the recent developments in: data acquisition and pre-processing, feature extraction and selection, and subsequent classification. Experimental results reported against different facial databases were also presented.

This analysis enables us to:

- Identify the weak points and define the desirable characteristics of such systems. The lack of a basis for benchmarks (comprehensive, accessible reference set of expression displays) and associated evaluation procedures for all different efforts in the research on machine analysis of human face expression recognition represents arguable the most important current issue. In our opinion, the next big problem is that many of the systems still require manual intervention. The challenge is to make the system fully automatic. Some unrealistic assumptions are still present in many papers e.g., frontal view image of faces without hair, glasses and all facial expression displayed using six basic emotion categories.

- Predict the future trends in the domain of facial imagery. In our opinion, the most important research direction concerns the development of the 3D dynamic datasets using new 3D sensors such as structure light cameras or time-of-flight cameras. This direction is quite promising for real-time segmentation. Also the inclusion of the temporal information will improve accuracy and robustness. In the future, the focus will shift from considering posed expression recognition to the development of methods for spontaneous expression recognition along with the deployment in a system with real-time capability. More work has to be done with regard to the integration of other communication channels such as voice and gesture.

We conclude by saying that the technology of facial expression recognition has enormous market potential and, in the near future, will enhance most human-computer interfaces.

REFERENCES

[1] Y. Tian, L. Brown, A. Hampapur, S. Pankanti, A. Senior, and R. Bolle, "Real world real-time automatic recognition of facial expression," in IEEE PETS, Australia, March 2003.

[2] M. Pantic and L. J. M. Rothkrantz, "Automatic analysis of facial expressions: The state of the art," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, pp. 1424–1445, December 2000.

[3] B. Fasel and J. Luettin, "Automatic facial expression analysis: a survey," Pattern Recognition, vol. 36, no. 1, pp. 259–275, 2003.

[4] Y.-L. Tian, T. Kanade, and J. F. Cohn, "Facial expression analysis," in Handbook of Face Recognition, S. Z. Li and A. K. Jain, Eds. Springer New York, 2005, pp. 247–275.

[5] B. J. Matuszewski, W. Quan, L-K Shark, "Facial Expression Recognition" in "Biometrics – Unique and Diverse Applications in Nature, Science, and Technology", M. Albert (Ed.). InTech, 2011.

[6] Y.-Li Tian, T. Kanade, and J. Cohn, "Evaluation of Gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity," in Proc. Fifth IEEE Int. Conf. Automatic Face and Gesture Recognition, May, pp. 229–234, 2002.

[7] P. Ekman and W. Friesen, Facial Action Coding System, Palo Alto, CA: Consulting Psychologists Press, 1978.

[8] I. Cohen, N. Sebe, A. Garg, S. C. Lawrence, S. H. Thomas, "Facial expression recognition from video sequences: temporal and static modeling", Computer Vision and Image Understanding, Special Issue on Face Recognition, vol. 91, pp. 160-187, Issues 1-2, July-August 2003.

[9] N. Sebe, I. Cohen, A. Garg, M.S. Lew, T.S. Huang, "Emotion Recognition Using a Cauchy Naive Bayes Classifier", International Conference on Pattern Recognition (ICPR'02), vol I, pp. 17-20, Quebec City, Canada, August 2002.

[10] L. Ma, K. Khorasani, "Facial expression recognition using constructive feedforward neural networks," Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on , vol. 34, no. 3, pp.1588-1595, June 2004.

[11] C. Shan, S. Gong, and P. W. McOwan, "Robust Facial Expression Recognition Using Local Binary Patterns," Proc. IEEE International Conference on Image Processing (ICIP'05), Genova, Italy, September 2005.

[12] Z. Zeng, Y. Fu, G. I. Roisman, Z. Wen, Y. Hu, and T. S. Huang, "Spontaneous Emotional Facial Expression Detection," Journal of Multimedia, vol. 1, no. 5, pp. 1-8, 2006.

[13] J. Wang, L. Yin, X. Wei, and Y. Sun, "3D facial expression recognition based on primitive surface feature distribution," in Proc. Conf. Computer Vision and Pattern Recognition, 2006, pp. 1399–1406.

[14] F. Dornaika and B. Raducanu, "Inferring facial expressions from videos: Tool and application," Image Commun., vol. 22, pp. 769–784, October 2007.

[15] I. Kotsia, I. Pitas, "Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines," Image Processing, IEEE Transactions on, vol. 16, Issue 1, pp. 172 – 187, Jan. 2007.

[16] J. Ahlberg, "CANDIDE-3 -- an updated parameterized face," Report No. LiTH-ISY-R-2326, Dept. of Electrical Engineering, Linköping University, Sweden, 2001.

[17] T. Fang, X. Zhao, O. Ocegueda, S. K.. Shah and I. A. Kakadiaris, "3D Facial Expression Recognition: A Perspective on Promises and Challenges", in Proc. 9th IEEE International Conference on Automatic Face and Gesture Recognition (FG), Santa Barbara, CA, Mar. 21-25, 2011.

[18] A. Panning, A. Al-Hamadi, R. Niese, and B. Michaelis, "Facial expression recognition based on haar-like feature detection," Pattern Recognition and Image Analysis, vol. 18, pp. 447–452, 2008.

[19] I. Buciu, C. Kotropoulos, and I. Pitas, "Comparison of ICA approaches for facial expression recognition," Signal, Image and Video Processing, vol. 3, pp. 345–361, 2009.

[20] M. H. K. T. Jabid and O. Chae, "Robust facial expression recognition based on local directional pattern," ETRI Journal, vol. 32, no. 5, pp. 784–794, Oct 2010.

[21] Q. Jia, Y. Liu, H. Guo, Z. Luo, Y. Wang, "A sparse representation approach for local feature based expression recognition," Multimedia Technology (ICMT), 2011 International Conference on, pp.4788-4792, 26-28 July 2011.

[22] M.F. Valstar, M. Pantic, "Fully Automatic Recognition of the Temporal Phases of Facial Actions," Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on , vol.42, no.1, pp.28-43, Feb. 2012.

[23] H. Tao and T. S. Huang, "Connected vibrations: A modal analysis approach to non-rigid motion tracking," CVPR, pp. 735–750, 1998.

[24] 3dMD Ltd., http://¬www.3dmd.com.

[25] P. Viola and M. Jones, "Robust real-time object detection," Int. J. Comput.Vis.., vol. 57, no. 2, pp. 137–154, May 2004.

[26] D. Vukadinovic and M. Pantic, "Fully automatic facial feature point detection using Gabor feature based boosted features," in Proc. IEEE Int. Conf. Syst., Man, Cybern., 2005, pp. 1692–1698.

[27] C. Liu and J. Ward, "Advantages of 3d methods for face recognition research in humans," in Proceedings of the Second international conference on Analysis and Modelling of Faces and Gestures (AMFG'05), W. Zhao, S. Gong, and X. Tang (Eds.), Springer-Verlag, Berlin, Heidelberg, 2005, pp. 244-254.

[28] S. Nissen and E. Nemerson, "Fast Artifical Neural Network," http://leenissen.dk/fann/.

[29] M. J. Lyons, J. Budynek, S. Akamatsu, "Automatic classification of single facial images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.12, pp. 1357-1362, 1999.

[30] M. Pantic, M. F. Valstar, R. Rademaker, L. Maat, "Web-based database for facial expression analysis," in Proc. IEEE International Conference on Multimedia and Expo (ICME'05), Amsterdam, Netherlands, 2005.

[31] T. Kanade, J. F. Cohn, Y. Tian, "Comprehensive database for facial expression analysis," IEEE International Conference on Automatic Face and Gesture Recognition, 2000.

[32] L. Yin, X. Wei, Y. Sun, J. Wang, M. Rosato, "A 3D facial expression database for facial behavior research," in 7th International Conference on Automatic Face and Gesture Recognition (FG2006), IEEE Computer Society TC PAMI, 2006, pp. 211-216.

[33] L. Yin, X. Chen, Y. Sun, T. Worm, M. Reale, "A High-Resolution 3D Dynamic Facial Expression Database," The 8th International Conference on Automatic Face and Gesture Recognition (FGR08), 17-19 September 2008.

[34] F. Wallhoff, "Facial expressions and emotion database," Technische Universität München, Tech. Rep., 2006.

[35] SmartKom, "Smartkom database," 2003. [Online]. Available: www.smartkom.org