We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Recognizing Facial Expressions
Using Model-based Image Interpretation

Matthias Wimmer, Zahid Riaz, Christoph Mayer and Bernd Radig
Department of Informatics, Technische Universität München
Germany

1. Abstract

Even if electronic devices widely occupy our daily lives, human-machine interaction still lacks intuition. Researchers intend to resolve these shortcomings by augmenting traditional systems with aspects of human-human interaction and consider human emotion, behavior, and intention. This publication focuses on one aspect of the challenge: recognizing facial expressions. Our approach achieves real-time performance and provides robustness for real-world applicability. This computer vision task comprises of various phases for which it exploits model-based techniques that accurately localize facial features, seamlessly track them through image sequences, and finally infer facial expressions visible. We specifically adapt state-of-the-art techniques to each of these challenging phases. Our system has been successfully presented to industrial, political, and scientific audience in various events.

2. Introduction

Nowadays, computers are capable of quickly solving mathematical problems and memorizing an enormous amount of information, but machine interfaces still miss intuition. This aspect is even more important to interaction with humanoids, because people expect them to behave similar to real humans. A non-humanoid manner of the robot could generate confusions for the interacting person. Therefore, researchers augment traditional systems with human-like interaction capabilities. Widespread applicability and the comprehensive benefit motivate research on this topic. Natural language recognition has already been successfully deployed in commercial systems since a few years. However, to construct convenient interaction mechanisms robots must integrate further knowledge, e.g. about human behavior, intention, and emotion interpretation. For example in computer-assisted learning, a computer acts as the teacher by explaining the content of the lesson and questioning the user afterwards.

Awareness of human emotions will significantly rise the quality and success of these lessons. For a comprehensive overview on applications in emotion recognition, we refer to [16]. Today, dedicated hardware often facilitates this challenge [14, 26, 24]. These systems derive the emotional state from blood pressure, perspiration, brain waves, heart rate, skin temperature, etc.

In contrast, humans interpret emotion via visual and auditory information only. As an advantage, this information is acquired without interfering with people. Furthermore,

computers are able to easily acquire this information with general purpose hardware. However, precisely deriving human emotion from this vast amount of sensor data poses great challenges.
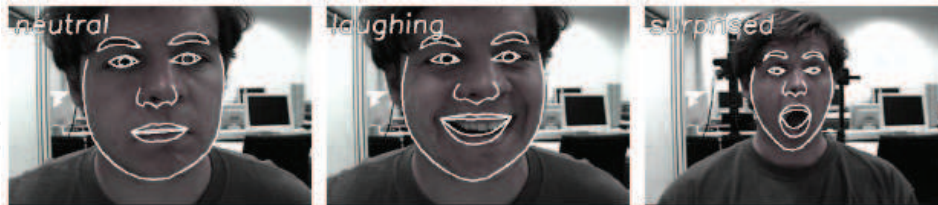

Figure 1. Interpreting facial expressions with a deformable face model

To tackle this challenge our goal ist to create a system that estimates facial expressions in real-time and that could robustly run in real-world environments. We develop it using model-based image interpretation techniques, which have proven its great potential to fulfill current and future requests on real-world image understanding. We take a three-step approach that robustly localizes facial features, tracks them through image sequences, and finally infers the facial expression. As explained in Section 4, some components require to be improved over the state-of-the-art, others are specifically adapted to the face interpretation scenario. Our experimental evaluation is conducted on a publicly available image database and is therefore well comparable to related approaches. Our demonstrator has been successfully applied to various real-world scenarios and it has been presented to the public and to industrial and political audience on various trade fairs.

This paper continues as follows. Section 3 explains the state-of-the-art of facial expression recognition covering both psychological theory and state-of-the-art algorithms. Section 4 describes the various components of our model-based approach. Section 5 experimentally evaluates our approach on a publicly available image database and draws a comparison to related approaches. Section 6 summarizes our achievement and points out future work.

3. Facial Action Recognition: State-of-the-art

This section elaborates on psychological aspects of facial expression recognition and indicates the state-of-the-art of current techniques.

3.1 Universal Facial Expressions and the Facial Action Coding System (FACS)

Ekman and Friesen [7] find six universal facial expressions that are expressed and interpreted in the same way by humans of any origin all over the world. They do not depend on the cultural background or the country of origin. As an example, Figure 1 shows how our technique distinguishes between the different facial expressions using a model-based approach.

The Facial Action Coding System (FACS) [8] precisely describes the muscle activities within a human face when facial expressions are displayed. Action Units (AUs) denote the motion of particular facial parts and state the facial muscles involved. Facial expressions are generated by the combinations of AUs. Extended systems like the Emotional FACS [12] specify the relation between facial expressions and emotions.

3.2 Benchmark DB

The Cohn-Kanade-Facial-Expression database (CKFE-DB) contains 488 short image sequences of 97 different persons performing the six universal facial expressions [15]. It provides researchers with a large dataset for experimenting and benchmarking purpose. Each sequence shows a neutral face at the beginning and then develops into the peak expression. Furthermore, a set of AUs has been manually specified by licensed FACS-experts for each sequence. Note that this database does not contain natural facial expressions, but volunteers were asked to act. Furthermore, the image sequences are taken in a laboratory environment with predefined illumination conditions, solid background and frontal face views. Algorithms that perform well with these image sequences are not immediately appropriate for real-world scenes.
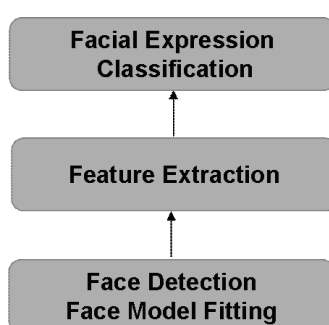

Figure 2. The common three-phase procedure for recognizing facial expressions, see [19]

3.3 The Common Three-phase Procedure

According to the survey of Pantic et al. [19], the computational task of facial expression recognition is usually subdivided into three subordinate challenges: face detection, feature extraction, and facial expression classification as shown in Figure 2. Chibelushi et al. [1] added a pre- and a post-processing step. This section presents several state-of-the-art approaches, which accomplish the involved steps in different ways. For a more detailed overview we refer to Chibelushi et al.

Phase 1, the human face and the facial components have to be accurately located within the image. This is often accomplished by computing a bounding box that roughly specifies the location and the extent of the entire face. More elaborate approaches make use of a fine grain face model, which has to be fitted precisely to the contours of the visible face. As an advantage, the model-based approach provides information about the relative location of the different facial components and their deformation, which turns out to be useful for the subsequent phases.

On the one hand, automatic algorithms compute the location of the visible face as in [18, 10, 4]. On the other hand, humans specify this information by hand, because the researchers rather focus on the subsequent interpretation task itself, as in [23, 2, 22, 25].

Phase 2, knowing the exact position of the face, features that are descriptive for facial expressions are extracted from the image data. Facial expressions consist of two important aspects: the muscle activity while the expression is developping and the shape of the peak expression. Algorithms focus on extracting features that represent these aspects.

Michel et al. [18] extract the location of 22 feature points within the face and determine their motion between an image of the neutral face and an image of the peak expression. They use feature points that are mostly located around the eyes and around the mouth. The very

similar approach of Cohn et al. [3] uses hierarchical optical flow in order to determine the motion of 30 feature points. They term their approach feature point tracking. Littlewort et al. [17] utilize a bank of 40 Gabor wavelet filters at different scales and orientations to extract features directly from the image. They perform convolution and obtain a vector of magnitudes of complex valued responses.

Phase 3, the facial expression is derived from the previously extracted features. Mostly a classifier is learned from a comprehensive training set of annotated examples. Some approaches first compute the visible AUs and then infer the facial expression by rules stated by Ekman and Friesen [9]. Michel et al. [18] train a Support Vector Machine (SVM) that determines the visible facial expression within the video sequences of the CKFE-DB by comparing the first frame with the neutral expression to the last frame with the peak expression. Schweiger and Bayerl [22] compute the optical flow within 6 predefined regions of a human face in order to extract the facial features. Their classification is based on supervised neural network learning.
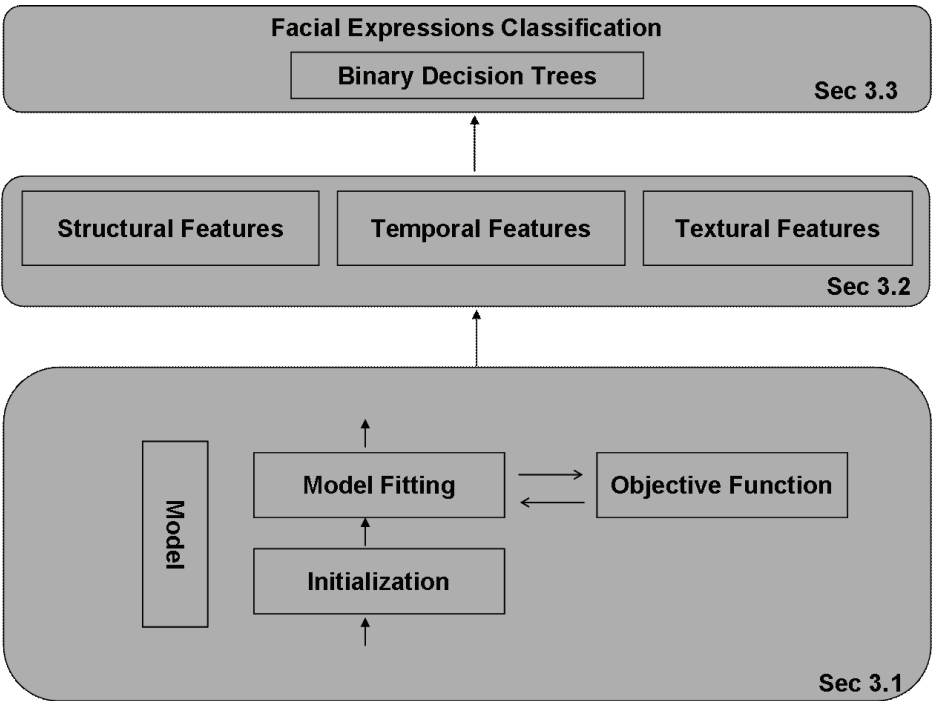

Figure 3. The three-phase procedure with our concrete implementation. Our work especially focuses on Phase 1, where we use model-based techniques. This approach splits the challenge of image interpretation into four computationally independent modules: the model, initialization, model fitting, and the objective function. Our work contributes to designing robust objective functions

4. Facial Expressions Recognition via Model-Based Techniques

Our approach makes use of model-based techniques, which exploit a priori knowledge about objects, such as their shape or texture, see Figure 3. Reducing the large amount of image data to a small set of model parameters facilitates and accelerates the entire process of facial expression interpretation. Our approach sticks to the three phases stated by Pantic et al. [19]. In Section 4.1. we consider fitting a face model to the camera image. Section 4.2

describes the extraction of meaningful features and Section 4.3 shows how we derive the facial expressions visible from these features.

4.1 Model-based Image Interpretation

Model-based techniques consist of four components: the model, the initialization algorithm, the objective function, and the fitting algorithm [28], see Figure 4.

Figure 4. Deformation by change of just one parameter in each row. Topmost row: b1 rotates the head. Middle row: b3 opens the mouth. Lower-most row: b10 moves pupils in parallel

Our approach makes use of a statistics-based deformable **model**, as introduced by Cootes et al. [5]. The model contains a parameter vector **p** that represents its possible configurations, such as position, orientation, scaling, and deformation. Models are mapped onto the surface of an image via a set of feature points, a contour, a textured region, etc. Referring to [6], deformable models are highly suitable for analyzing human faces with all their individual variations.

Its parameters $p = (t_x, t_y, s, \theta, b)^T$ comprise the translation, scaling factor, rotation, and a vector of deformation parameters $b = (b_1, \ldots, b_m)^T$. The latter component describes the configuration of the face, such as the opening of the mouth, roundness of the eyes, raising of the eye brows, see Figure 4.

The **initialization algorithm** automatically starts the interpretation process by roughly localizing the object to interpret, see Figure 5. It computes an initial estimate of the model parameters that needs to be further refined by the subsequent fitting algorithm. Our system integrates the approach of Viola and Jones [27], which is able to detect the affine transformation parameters $(t_x, t_y, s, \text{and } \theta)$ of frontal faces.
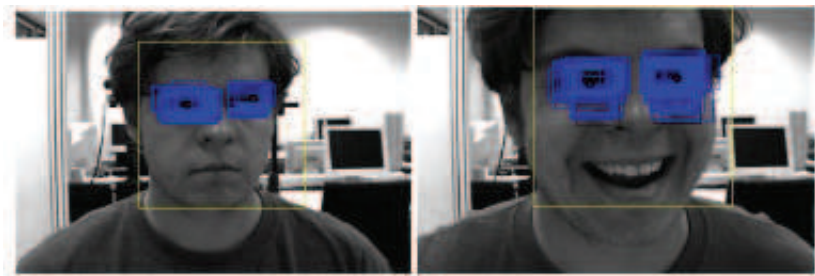

Figure 5. Localizing face and eyes using Viola and Jones' boosting algorithm

In order to obtain higher accuracy, we apply a second iteration of the Viola and Jones object detector to the previously determined image region of the face. In this iteration the algorithm is utilized to localize facial components, such as eyes and mouth. This extension allows to roughly estimate the deformation parameters b as well. The algorithm has been trained on positiv and negativ training examples. In the case of the eyes, our positive training examples contain the images of eyes, whereas the negative examples consist of image patches in the vicinity of the eyes, such as the cheek, the nose, or the brows. Note that the resulting eye detector is not able to robustly localize the eyes in a complex image, because it usually contains a lot of information that was not part of the training data. However, it is highly appropriate to determine the location of the eyes within a pure face image or within the face region of a complex image.

The **objective function** f(I, p) yields a comparable value that specifies how accurately a parameterized model p describes the content of an image I. It is also known as the likelihood, similarity, energy, cost, goodness, or quality function. Without losing generality, we consider lower values to denote a better model fit. Traditionally, objective functions are manually specified by first selecting a small number of simple image features, such as edges or corners, and then formulating mathematical calculation rules. Afterwards, the appropriateness is subjectively determined by inspecting the result on example images and example model parameterizations. If the result is not satisfactory the function is tuned or redesigned from scratch. This heuristic approach relies on the designer's intuition about a good measure of fitness. Our earlier works [29, 30] show that this methodology is erroneous and tedious. This traditional approach is depicted to the left in Figure 6.
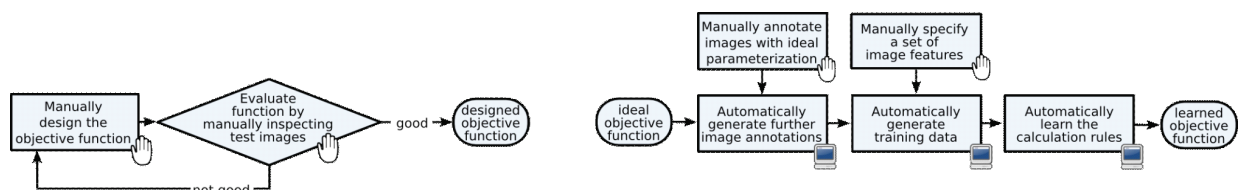

Figure 6. The traditional procedure for designing objective functions (left), and the proposed method for learning objective functions (right)

To avoid these drawbacks, we recently proposed an approach that learns the objective function from annotated example images [29]. It splits up the generation of the objective function into several partly automated tasks. This provides several benefits: firstly, automated steps replace the labor-intensive design of the objective function. Secondly, this approach is less error prone, because giving examples of good fit is much easier than explicitly specifying rules that need to cover all examples. Thirdly, this approach does not rely on expert knowledge and therefore it is generally applicable and not domain-dependent. The bottom line is that this approach yields more robust and accurate objective functions, which greatly facilitate the task of the fitting algorithm. For a detailed description of our approach, we refer to [29].

The **fitting algorithm** searches for the model that best describes the face visible in the image. Therefore, it aims at finding the model parameters that minimize the objective function. Fitting algorithms have been the subject of intensive research and evaluation, e.g. Simulated Annealing, Genetic Algorithms, Particle Filtering, RANSAC, CONDENSATION, and CCD, see [13] for a recent overview and categorization. We propose to adapt the objective function rather than the fitting algorithm to the specifics of our application. Therefore, we are able to

use any of these standard fitting algorithms, the characteristics of which are well-known, such as termination criteria, runtime, and accuracy. Due to real-time requirements, our experiments in Section 5 have been conducted with a quick hill climbing algorithm. Note that the reasonable specification of the objective function makes this local optimization strategy nearly as accurate as a global optimization strategy, such as Genetic Algorithms.

4.2 Extraction of Structural, Temporal and Textural Features

Two aspects generally characterize facial expressions i.e. their structural contribution and temporal behavior: they turn the face into a distinctive state [17] and the involved muscles show a distinctive motion [22, 18]. Our approach considers both aspects by extracting structural and temporal features. Furthermore, Textural features represent context knowledge about the person, such as the general face shape, age or gender. This large amount of feature information provides a profound basis for the subsequent classification step, which therefore achieves great accuracy.

Structural features: The deformation parameters b describe the constitution of the visible face. The examples in Figure 4 illustrate the relation between the facial expression and the value of b. Therefore, we consider b to provide high-level information to the interpretation process. These features are assembled in a feature vector. In contrast, the transformation parameters tx, ty, s, and θ are not related to the facial expression and therefore, we do not consider them as features.

$$t_0 = (b_1, \ldots, b_m)^T$$

Figure 7. Fitting a deformable face model to images and inferring different facial expressions by taking structural and temporal image features into account

Temporal features: Since facial expressions emerge from muscle activity, the motion of particular feature points within the face gives evidence about the facial expression. Real-time capability is important, and therefore, a small number of feature points is considered

only. The relative location of these points is connected to the structure of the face model. Note that we do not specify these locations manually, because this assumes a good experience of the designer in analyzing facial expressions. In contrast, we automatically generate G feature points that are uniformly distributed. We expect these points to move descriptively and predictably in the case of a particular facial expression. We sum up the motion $g_{x,i}$ and $g_{y,i}$ of each point $1 \leq i \leq G$ during a short time period. We set this period to 2 sec to cover slowly expressed emotions as well. The motion of the feature points is normalized by the affine transformation of the entire face (tx, ty, s, and θ) in order to separate the facial motion from the rigid head motion.

In order to determine robust descriptors, Principal Component Analysis (PCA) determines the *H* most relevant motion patterns (principal components) visible within the set of training sequences. A linear combination of these motion patterns describes each observation approximately correct. This reduces the number of descriptors (H ≤ 2G) by enforcing robustness towards outliers as well. As a compromise between accuracy and runtime performance, we set the number of feature points to G = 140 and the number of motion patterns to H = 14. Figure 7 visualizes the obtained motion of the feature points for some example facial expressions.

The feature vector **t** is assembled from the *m* structural and the *H* temporal features as mentioned in equation below. This vector represents the basis for facial expression classification.

$$t_1 = (b_1,......,b_m,h_1,......h_H)^T$$

Textural features are obtained by applying PCA on extracted texture from face images. These texture examples cover the full face boundary excluding the forehead, hair and the ears. In contrast to shape deformations, that are caused by facial expression changes and varying head poses, texture variations arise mainly due to illumination changes. Therefore, various training images of the same person consider different types of texture variations. However, shape parameters have to be estimated first before recording the texture parameters. Figure 8 demonstrates the warping of the texture extracted from an example image onto the reference shape.
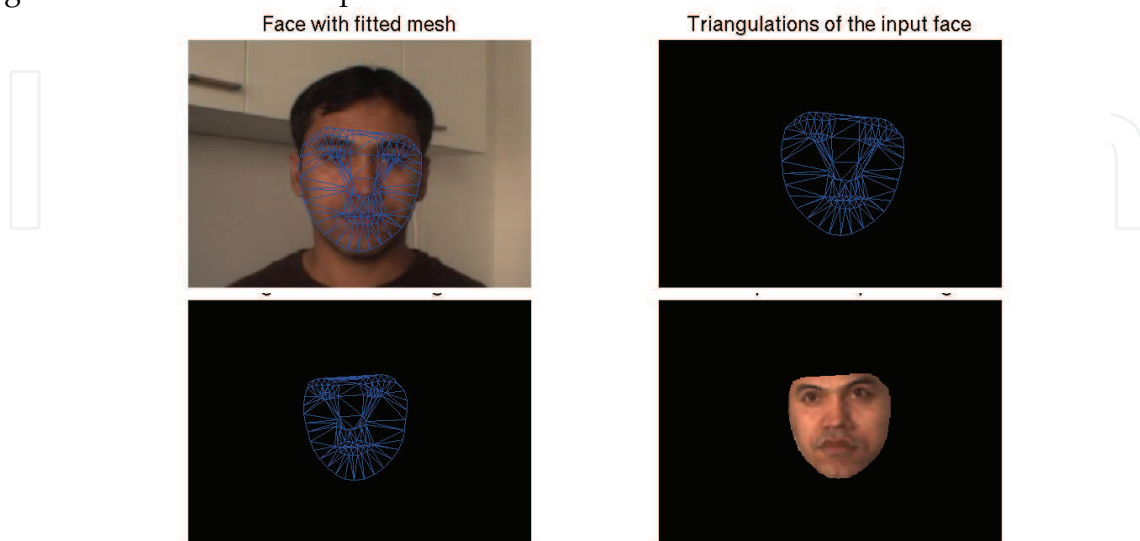

Figure 8. Face with fitted Model (Top Left) , Model of the input image (Top Right) , Reference shape (Bottom Left) ,Texture warping results on the reference shape (Bottom Right)

Ground truth	Classified as by t_0						Recognition rate
	Surprise	Happiness	Anger	Disgust	Sadness	Fear	
Surprise	61	0	0	1	3	17	85.9%
Happiness	1	57	2	3	0	**11**	77.0%
Anger	0	0	11	5	12	7	31.4%
Disgust	2	7	8	9	3	7	25.0%
Sadness	8	0	10	5	27	9	45.8%
Fear	5	**17**	2	9	8	17	29.8%
Average recognition rate							**49.1%**
Ground truth	Classified as by t_1						Recognition rate
	Surprise	Happiness	Anger	Disgust	Sadness	Fear	
Surprise	28	1	1	0	0	0	93.3%
Happiness	1	26	2	2	3	**4**	70.3%
Anger	1	1	14	2	2	1	66.7%
Disgust	0	2	1	10	3	1	58.8%
Sadness	1	2	2	2	22	1	73.3%
Fear	1	**5**	1	0	2	13	59.1%
Average recognition rate							**70.3%**
Ground truth	Classified as by t_2						Recognition rate
	Surprise	Happiness	Anger	Disgust	Sadness	Fear	
Surprise	64	0	1	2	3	1	90.1%
Happiness	1	59	0	1	0	**13**	79.7%
Anger	0	0	14	5	10	6	40.0%
Disgust	3	3	5	13	6	6	36.1%
Sadness	2	0	9	4	35	9	59.3%
Fear	4	**15**	0	8	1	30	51.7%
Average recognition rate							**59.5%**

Table 1. Confusion matrix and recognition rate of our approach

Given a set of shape points x of the input example image and x_{ref} of the reference image, we compute the relative position in the example image to every pixel position in the reference shape and sample the texture values to gain the texture vector l_{text}. The texture vector is normalized to remove global lighting effects. Piecewise affine transform is used to warp the texture of the example image on the reference shape [35]. From the warped texture and the PCA data the textural features are estimated. Texture's parameters variations cause changes in the appearance of the faces similar to the eigenfaces approach [36]. The combination of shape and texture parameters is well-known Active Appearance Model (AAM) introduced by Cootes et al [37]. Since this approach utilizes both, shape and texture information its feature vector contains shape parameters b as well as the textural features l.

$$t_2 = (b_1, \ldots\ldots, b_m, l_1, \ldots\ldots l_L)^T$$

4.3 Classification of Facial Expressions Using Decision Trees

With the knowledge of the feature vector **t**, a classifier infers the correct facial expression. We learn a Binary Decision Tree [20], which is a robust and quick classifier. However, any other multi-class classifier that is able to derive the class membership from real valued features can be integrated as well, such as a k-Nearest-Neighbor classifier. We take 67% of the image sequences of the CKFE-DB as the training set and the remainder as test set, the evaluation on which is shown in the next section.

5. Experimental Evaluation

In order to evaluate the accuracy of our approach, we apply it to the previously unseen fraction of the CKFE-DB. Table 1 shows the recognition rate and confusion matrix of each facial expression. The facial expressions happiness and fear are confused most often. The reason for this confusion is the similar muscle activity around the mouth. This coincidence is also reflected by FACS.

Facial expression	Our results t_0	Our results t_1	Our results t_2	Approach of Michel et al. [18]	Approach of Schweiger et al. [22]
Anger	31.4%	66.7%	40.0%	66.7%	75.6%
Disgust	25.0%	58.8%	36.1%	58.8%	30.0%
Fear	29.3%	59.1%	51.7%	66.7%	0.0%
Happiness	77.0%	70.3%	79.7%	91.7%	79.2%
Sadness	45.8%	73.3%	59.3%	62.5%	60.5%
Surprise	85.9%	93.3%	90.1%	83.3%	89.8%
Average	49.1%	70.3%	59.5%	71.8%	55.9%

Table 2. Recognition rate of our approach compared to the results of different algorithms

The accuracy of our approach is comparable to the one of Schweiger et al. [22] who also conduct their evaluation on the CKFE-DB, see Table 2. For classification, they also favor motion from different facial parts and determine principal components from these features.

However, Schweiger et al. manually specify the region of the visible face whereas our approach performs an automatic localization via model-based image interpretation. Michel et al. [18] also focus on facial motion by manually specifying 22 feature points that are predominantly located around the mouth and around the eyes. They utilize a support vector machine (SVM) for determining one of the six facial expressions.

In order to account for local variation and to fulfill the lack of ground truths, an appearance model based approach is used to train a classifier. This improves the results further.

6. Summary and Outlook

Automatic recognition of human facial gestures has recently attained a significant place in multimodal human-machine interfaces and further applications. This paper presents our proof-of concept for Facial Expression Interpretation, which is real-time capable and robust enough to be deployed to real-world scenarios.

We exploit model-based techniques and adapt them specifically to facial expression recognition. First, we extract informative features about faces from the image data that describes skin color and lip color regions. Furthermore, we learn the objective function for model fitting from example images as opposed to constructing it manually.

The obtained system shows promising referring to its recognition results and we have successfully conducted live demonstrations that were attended by industrial and political audience, such as in [11].

In future work, we aim at integrating multi-modal feature sets, for which we have already preliminary results [21]. Furthermore, we will present structural, motion and textural features to the classifier. We will apply Early Sensor Fusion in order to keep all knowledge for the final decision process and to exploit the ability of a combined feature-space optimization.

7. References

C. C. Chibelushi and F. Bourel. Facial expression recognition: A brief tutorial overview. In *CVonline: On-Line Compendium of Computer Vision*, editor Robert Fisher, January 2003. [1]

Isaac Cohen, Nicu Sebe, L. Chen, A. Garg, and T. Huang. Facial expression recognition from video sequences: Temporal and static modeling. *Computer Vision and Image Understanding (CVIU) special issue on face recognition*, 91(1-2):160–187, 2003.[2]

Jeffrey Cohn, Adena Zlochower, Jenn-Jier James Lien, and Takeo Kanade. Feature-point tracking by optical flow discriminates subtle differences in facial expression. In *Proceedings of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pages 396 – 401, April 1998. [3]

Jeffrey Cohn, Adena Zlochower, Jenn-Jier James Lien, and Takeo Kanade. Automated face analysis by feature point tracking has high concurrent validity with manual facs coding. *Psychophysiology*, 36:35 – 43, 1999 [4]

Tim F. Cootes and Chris J. Taylor. Active shape models – smart snakes. In *Proceedings of the 3rd British Machine VisionConference*, pages 266 – 275. Springer Verlag, 1992. [5]

G. J. Edwards, Tim F. Cootes, and Chris J. Taylor. Face recognition using active appearance models. In H. Burkhardt and Bernd Neumann, editors, *5th European Conference on Computer Vision*, volume LNCS-Series 1406–1607, pages 581–595,Freiburg, Germany, 1998. Springer-Verlag. [6]

Paul Ekman. Universals and cultural differences in facial expressions of emotion. In J. Cole, editor, *Nebraska Symposium on Motivation* 1971, volume 19, pages 207–283, Lincoln, NE, 1972. University of Nebraska Press. [7]

Paul Ekman. Facial expressions. In T. Dalgleish and M. Power, editors, *Handbook of Cognition and Emotion*, New York,1999. John Wiley & Sons Ltd. [8]

Paul Ekman and Wallace Friesen. The Facial Action Coding System: A Technique for the Measurement of Facial Movement. *Consulting Psychologists Press*, San Francisco, 1978. [9]

Irfan A. Essa and Alex P. Pentland. Coding, analysis, interpretation, and recognition of facial expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):757–763, 1997. [10]

Stefan Fischer, Sven Döring, Matthias Wimmer, and Antonia Krummheuer. Experiences with an emotional sales agent. In Elisabeth Andr´e, Laila Dybkjær, Wolfgang Minker, and Paul Heisterkamp, editors, *Affective Dialogue Systems*, volume3068 of Lecture Notes in Computer Science, pages 309–312, Kloster Irsee, Germany, June 2004. Springer. [11]

Wallace V. Friesen and Paul Ekman. *Emotional Facial Action Coding System*. Unpublished manuscript, University of California at San Francisco, 1983. [12]

Robert Hanek. Fitting Parametric Curve Models to Images Using Local Self-adapting Seperation Criteria. *PhD thesis*, Department of Informatics, Technische Universität München, 2004. [13]

Curtis S. Ikehara, David N. Chin, and Martha E. Crosby. A model for integrating an adaptive information filter utilizing biosensor data to assess cognitive load. In *User Modeling*, volume 2702/2003, pages 208–212. Springer Berlin / Heidelberg, 2003. [14]

Takeo Kanade, John F. Cohn, and Yingli Tian. Comprehensive database for facial expression analysis. In *International Conference on Automatic Face and Gesture Recognition*, pages 46–53, France, March 2000. [15]

Christine L. Lisetti and Diane J. Schiano. Automatic facial expression interpretation: Where human interaction, articial intelligence and cognitive science intersect. Pragmatics and Cognition, *Special Issue on Facial Information Processing and Multidisciplinary Perspective*, 1999. [16]

Gwen Littlewort, Ian Fasel, Marian Stewart Bartlett, and Javier R. Movellan. Fully automatic coding of basic expressions from video. *Technical report*, March 2002. [17]

P. Michel and R. El Kaliouby. Real time facial expression recognition in video using support vector machines. In *Fifth International Conference on Multimodal Interfaces*, pages 258–264, Vancouver, 2003. [18]

Maja Pantic and Leon J. M. Rothkrantz. Automatic analysis of facial expressions: The state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1424–1445, 2000. [19]

Ross Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Mateo, California, 1993. [20]

Björn Schuller, Matthias Wimmer, Dejan Arsic, Gerhard Rigoll, and Bernd Radig. Audiovisual behavior modeling by combined feature spaces. In *IEEE International Conference on Acoustics, Speech, and Signal Processing* (ICASSP), volume 2, pages 733–736, Honolulu, Hawaii, USA, April 2007. [21]

R. Schweiger, P. Bayerl, and Heiko Neumann. Neural architecture for temporal emotion classification. In *Affective Dialogue Systems* 2004, LNAI 3068, pages 49–52, Kloster Irsee, June 2004. Elisabeth Andre et al (Hrsg.). [22]

Nicu Sebe, Michael S. Lew, Ira Cohen, Ashutosh Garg, and Thomas S. Huang. Emotion recognition using a cauchy naïve bayes classifier. *In Proceedings of the 16th International Conference on Pattern Recognition*, volume 1, pages 17–20, Washington, DC, USA, 2002. IEEE Computer Society. [23]

Elizabeth Marie Sheldon. Virtual agent interactions. *PhD thesis*, 2001. Major Professor-Linda Malone. [24]

Ying-Li Tian, Takeo Kanade, and Jeffrey F. Cohn. Recognizing action units for facial expression analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):97–115, February 2001. [25]

Rita M. Vick and Curtis S. Ikehara. Methodological issues of real time data acquisition from multiple sources of physiological data. In *Proceedings of the 36th Annual Hawaii International Conference on System Sciences*, page 129.1, Washington, DC, USA, 2003. IEEE Computer Society. [26]

Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition*, volume 1, pages 511–518, Kauai, Hawaii, 2001. [27]

Matthias Wimmer. Model-based Image Interpretation with Application to Facial Expression Recognition. *PhD thesis*, Technische Universitat M¨unchen, Institute for Informatics, December 2007. [28]

Matthias Wimmer, Freek Stulp, Sylvia Pietzsch, and Bernd Radig. Learning local objective functions for robust face model fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (PAMI), 30(8), 2008. to appear. [29]

Matthias Wimmer, Freek Stulp, Stephan Tschechne, and Bernd Radig. Learning robust objective functions for model fitting in image understanding applications. In Michael J. Chantler, Emanuel Trucco, and Robert B. Fisher, editors, *Proceedings of the 17th British Machine Vision Conference* (BMVC), volume 3, pages 1159–1168, Edinburgh, UK, September 2006. BMVA. [30]

Okabe, A.; Boots, B. ; and Sugihara, K. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, New York: Wiley, 1992 [31]

Gil Zigelman , Ron Kimmel, and Nahum Kiryati, Texture mapping using Surface Flattening via Multidimensional Scaling, *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, No. 2, April-June 2002 [32]

Volker Blanz , Thomas Vetter, A morphable model for the synthesis of 3D faces, *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, p.187-194, July 1999 [33]

Stan Z. L, Anil K. J. *Handbook of Face Recognition*, Springer, 2004. [34]

Stegmann M.B., Active Appearance Models: Theory Extensions and Cases, *Master Thesis*, Technical University of Denmark, 2000. [35]

Turk M. A., Pentland A.P., Eigenfaces for Recognition, *Journal of Cognitive Neuroscience* **3** (1): 1991, pp 71–86. [36]

Cootes T.F., Edwards G.J., Taylor C.J., Active Appearance Models, *Proc. European Conference on Computer Vision* Vol. 2, pp. 484-498, Springer 1998. [37]

Advances in Human Computer Interaction

Edited by Shane Pinder

ISBN 978-953-7619-15-2

Hard cover, 600 pages

Publisher InTech

Published online 01, October, 2008

Published in print edition October, 2008

In these 34 chapters, we survey the broad disciplines that loosely inhabit the study and practice of human-computer interaction. Our authors are passionate advocates of innovative applications, novel approaches, and modern advances in this exciting and developing field. It is our wish that the reader consider not only what our authors have written and the experimentation they have described, but also the examples they have set.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Matthias Wimmer, Zahid Riaz, Christoph Mayer and Bernd Radig (2008). Recognizing Facial Expressions Using Model-based Image Interpretation, Advances in Human Computer Interaction, Shane Pinder (Ed.), ISBN: 978-953-7619-15-2, InTech, Available from:
http://www.intechopen.com/books/advances_in_human_computer_interaction/recognizing_facial_expressions_using_model-based_image_interpretation

INTECH
open science | open minds