Facial Action Detection from Dual-View Static Face Images

Maja Pantic and Leon Rothkrantz
Delft University of Technology
Electrical Engineering, Mathematics and Computer Science
Mekelweg 4, 2628 CD Delft, the Netherlands
E-mail: M.Pantic@ewi.tudelft.nl, L.J.M.Rothkrantz@ewi.tudelft.nl

Abstract – **This paper presents an automatic system that we developed for automatic recognition of facial gestures (facial muscle activity) from static images of combined frontal- and profile-view of the face. For the frontal view, the face region is subjected to multi-detector processing which per facial component (eyes, eyebrows, mouth), generates a spatial sample of its contour. A set of 19 frontal-face feature points is then extracted from the spatially sampled contours of the facial features. For the profile view, 10 feature points are extracted from the contour of the face-profile region. Based on these 29 points, 29 individual facial muscle action units (AUs) occurring alone or in combinations in an input dual-view image are recognized using a rule-based reasoning. With each scored AU, the utilized algorithm associates a factor denoting the certainty with which the pertinent AU has been scored. A recognition rate of 86% is achieved.**

I. INTRODUCTION

The major impulse to investigate automatic facial expression analysis comes from the significant role of facial expressions in our social and emotional lives. They are conversational and interactive signals that clarify our current focus of attention and regulate our interactions with the environment and other persons in our vicinity [16]. They are our direct and naturally preeminent means of communicating emotions [16], [7]. Hence, automatic analyzers of facial expressions seem to have a natural place in various vision-based man-machine systems including automated tools for lip reading, bimodal speech analysis, videoconferencing, face / visual speech synthesis, affective computing, and next generation human-behavior-aware man-machine interfaces.

Approaches to automatic facial expression analysis attempt usually to recognize a small set of prototypic emotional facial expressions, i.e., fear, sadness, disgust, anger, surprise and happiness [9], [11]. This practice may follow from the work of Darwin and more recently Ekman [7], who suggested that basic emotions have corresponding prototypic expressions. In everyday life, however, such prototypic facial expressions occur relatively infrequently; emotions are displayed more often by subtle changes in one or few discrete facial features, such as raising the eyebrows in surprise [16]. To detect such subtlety of human emotions and, in general, to make the information conveyed by facial expressions available for the usage in various applications listed above, automatic recognition of facial gestures (i.e., atomic facial signals) is needed.

From several methods for recognition of facial gestures, the FACS system [5] is the most commonly used in psychological research. It is a system designed for human observers to describe changes in the facial expression in terms of visually observable activations of facial muscles. The changes in the facial expression are described with FACS in terms of 44 different Action Units (AUs), each of which is anatomically related to the contraction of a specific (set of) facial muscle(s). Using the FACS' rules for encoding AUs in a face image, a FACS coder (i.e., a human expert in using FACS) decomposes a shown facial expression into the AUs that produce the expression.

Though FACS provides a good foundation for AU-coding of face images by human observers, achieving this task by a computer is by no means a trivial task. A problematic issue is that AUs can occur in complex combinations, causing bulges (e.g., by the tongue pushed under one of the lips) and various in- and out-plane movements of permanent facial features (e.g., jetted jaw), that are difficult to differentiate from 2D face images. Approaches that have been reported for automatic recognition of AUs in images of faces are few [9]. Some researchers described patterns of facial motion that correspond to a few specific AUs, but did not report on actual recognition of these AUs [8], [2], [6]. Bartlett et al. [1] used optical flow and principal component analysis (PCA) to detect 6 individual AUs in portraits. Cohn et al. [3], [4] used optical-flow, PCA and Hidden Markov Models to recognize 8 individual AUs and 7 combinations of AUs in portraits. Tian et al. [19] used lip tracking, template matching and Neural Networks to recognize 16 AUs occurring alone or in a combination in nearly frontal-view face images. In brief, no automated system capable of recognizing all 44 AUs defined in FACS has been reported up to date. Systems presented in [12] and [14] perform the best in this aspect: they code 22 and, respectively, 20 AUs occurring alone or in a combination in frontal-view and, respectively, profile-view face images.

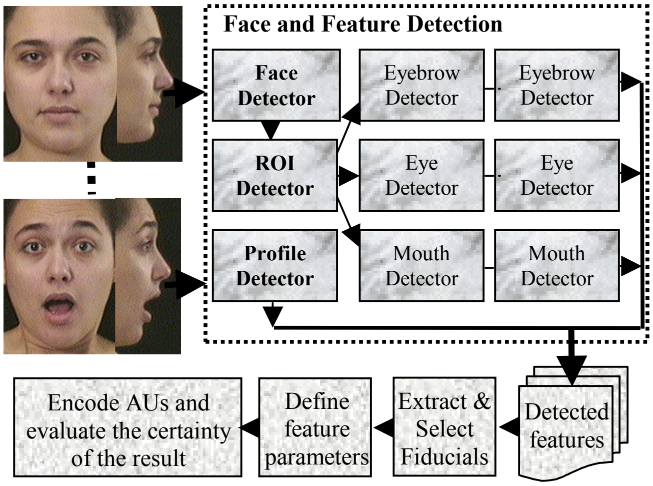

Fig. 1. Outline of the AU-recognition method

The research reported here addresses the problem of automatic AU coding from combined frontal- and profile-view face images. It was undertaken with two motivations:

1) In a portrait, facial gestures such as showing the tongue (AU19) or pushing the jaw forwards (AU29) represent out-plane non-rigid facial movements which are difficult to detect. Such facial gestures are clearly observable in a profile-view of the face. On the other hand, changes in the appearance of the eyes and eyebrows cannot be detected from the non-rigid changes in the profile contour, but are clearly observable from a frontal-view of the face. The usage of both frontal- and profile facial view promises, therefore, a quantitative increase of facial actions that can be handled.

2) A basic understanding of how to achieve automatic facial gesture analysis from multiple views of the human face is necessary if facial expression analyzers capable of handling partial and inaccurate data are to be developed [15]. Based on such knowledge, procedures of greater flexibility and improved quality can evolve.

Fig. 1 outlines the proposed method. First, a dual-view image of an expressionless face of the observed subject is processed. Then, each subsequent image of the observed subject is processed in the following manner. First, the face region and the face-profile region are extracted from the frontal-view image and, respectively, profile-view image. To do so, watershed segmentation with markers is applied on the morphological gradient of the color image. The contour of the segmented face-profile region is extracted as the face profile contour while the segmented frontal-view face region is subjected to a multi-detector processing. For each Region of Interest (eyebrows, eyes, mouth), one or more spatial samples of the contour of the relevant facial component are generated. Under the assumption that input images are non-occluded, scale- and orientation-invariant dual-views of the face (Fig. 1), we proceed with feature-points extraction. A set of 19 frontal face feature points is extracted from the spatially sampled contours of the facial components. For the profile view, 10 feature points are extracted from the contour of the segmented face-profile region. By performing an intra-solution consistency check, a certainty factor CF is assigned to each extracted point. A comparison of CFs assigned to frontal-face feature points leads to a selection of the most accurate of the redundantly extracted data. Subtle changes in the analyzed facial expression are measured next. Motivated by AUs of the FACS system, these changes are represented as a set of mid-level parameters describing the state and motion of the feature points and the shapes formed by certain feature points. Based on these parameters, a rule-based algorithm interprets the extracted facial information in terms of 29 AUs occurring alone or in a combination. With each scored AU, the utilized algorithm associates a factor denoting the certainty with which the pertinent AU has been scored. Face and face-profile detection, feature extraction, parametric representation, AU coding and experimental results are explained in sections II, III, IV, V and VI respectively.

II. FACE AND FACE-PROFILE DETECTION

The first step in automatic facial gesture analysis is to locate the face in the scene. This is addressed as a segmentation problem in two objects: the Face and the Background. For its low computational complexity and its good localization properties we choose the watershed segmentation with markers as the segmentation means.

For each input face image (either in frontal or in profile view), the markers of the two objects are extracted as follows. First, a color-based segmentation extracts the skin region as the largest connected image component with Hue, Saturation and Value within the range [5, 35], [0, 0.7] and [0.1, 0.9] respectively (Fig. 2) [13]. A binary erosion of the skin region with a small structuring element (3×3) yields the Face marker. In the absence of a similar model for the color of the Background, its marker is extracted as the bounding box of the skin region. Once the markers of the two objects are extracted, we apply watershed segmentation on the morphological gradient of the input color image. The gradient is estimated as the color difference between the morphological opening and closing operators, each of which is applied separately to each of the three components of the color image. We choose the Euclidian distance in the Lu*v* color space as a metric of the color difference, since

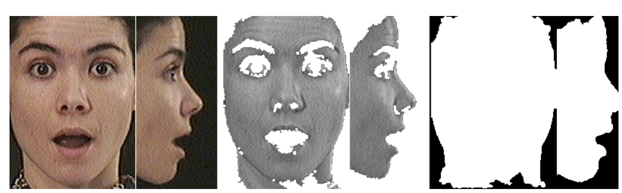

Fig. 2. Dual-view face image, HSV color-based segmentation for Face markers and segmented face and face-profile regions

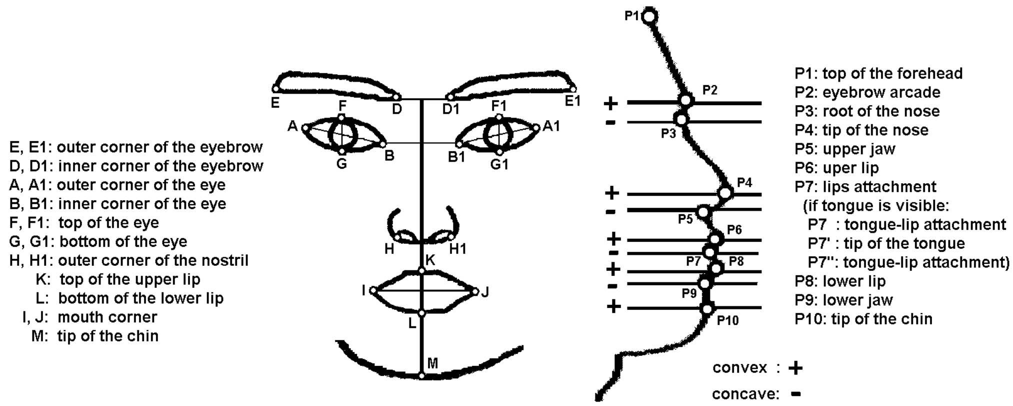

Fig. 3. Feature points (fiducial points of the face components' contours and of the profile contour)

the Lu*v* space is perceptually uniform under this metric [18].

A typical result of the application of the utilized algorithm is shown in Fig. 2. The algorithm yields a good localization of the face given that the most prominent color edge between the two markers is indeed the face contour.

III. Feature extraction

The face region and the face-profile region, extracted from an input dual-view face image as described above, are used for further analysis of shown facial gestures. We proceed with feature extraction under the assumption that input face images are non-occluded, scale and orientation invariant, and that profile-images are in right profile view (e.g., Fig 1).

A. Profile Face Feature Extraction

The contour of the segmented face-profile region is treated as the face profile contour in further processing.

To extract the feature points from the face profile contour, we move from image to function analysis and treat the right-hand side of the face profile contour (up to the point P1, Fig. 3) as a profile contour function. We extract the extremities of this function (the zero-crossings of the function's 1st order derivative) as the feature points (Fig. 3). To ascertain correct extraction of the feature points when the tongue is visible (P7' and P7'' exist), we extract the feature points in the particular order (i.e., P1, P4, P2, P3, P10, P5, P9, P7 or P7' and P7", P6, P8).

To handle inaccuracies in feature points' detection (i.e., to handle false positives), we exploit both the knowledge about facial anatomy and the information extracted from the image of a neutral facial expression of the observed subject. A standard "search" window W_P has been defined for each fiducial P with respect to anatomically possible directions and magnitudes of the motion on the skin surface affecting the temporal location of P. Fiducial P_I is determined further for face profile image I such that it represents a specific zero crossing (Fig. 3) of the 1st order derivative of the profile contour function defined for image I and belongs to the W_P set around the location of P_N discerned for the face-profile image N of a neutral expression of the observed subject.

B. Frontal Face Feature Extraction

Multi-detector processing of the face region segmented from an input frontal-view face image is used to spatially sample the contours of the facial components.

First, we apply a simple analysis of image histograms to locate 6 regions of interest (ROI): two eyebrows, two eyes, nose, and mouth. Then, to spatially sample the contour of a certain facial component, we apply one or more facial-feature detectors to the pertinent ROI. For example, the contours of the eyes are localized in the ROIs of the eyes by using a single detector representing an adapted version of the method for hierarchical-perceptron feature localization (Fig. 4) [20]. On the other hand, the contour of the mouth is localized in the mouth ROI by applying both a 4-parameters deformable template and a method that fits three 2nd degree parabolas (Fig. 4) [13]. For further details about these and other detectors employed to spatially sample the contours of the facial components, readers are referred to [10].

We proceed with feature points' extraction. For the cases where multiple detectors were used to localize the contour of a certain facial component, a relevant set of fiducial points is extracted from each spatially sampled contour of the pertinent facial component. For instance, from each localized mouth contour, we extract 4 feature points (Fig. 3). In total, we extract 19 different feature points corresponding to the vertices and/or the apices of the contours of the facial components (Fig. 3).

C. Data Certainty Evaluation and Feature Selection

We utilize an "intra-solution consistency check" to assign a certainty factor to each of the extracted feature points. For example, to assign a certainty factor $CF_A = CF_F$

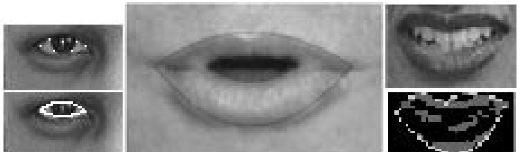

Fig. 4. Curve fitting on eye micro-features, mouth template matching, fitting three 2nd degree parabolas to the mouth

$= CF_G = CF_B \in [0, 1]$ to the fiducials of the right eye, we measure first the distance between the currently detected inner corner $B_{current}$ and point $B_{neutral}$ detected in the neutral expression image of the observed subject. Then we calculate the pertinent CF_B by using the following functional form:

$$CF_B = \boldsymbol{sigm}(\boldsymbol{d}(B_{current}, B_{neutral}); 7; 3.5)$$

where d(p1, p2) is the distance between points p1 and p2 and $sigm(x; \mu; \sigma)$ is a Sigmoid function whose parameters are determined under the assumption that there are 60 to 80 pixels across the width of the subject's eye. The major impulse for the usage of the inner corners of the eyes as the referential points for calculating CFs of the fiducial points of the eyes comes from the stability of these points with respect to non-rigid facial movements: facial muscles' contractions do not cause physical displacements of these points. For the same reason, the referential features used for calculating CFs of the fiducial points of the profile contour, eyebrows, nose/ chin and mouth are the tip of the nose (point P4, Fig. 3), the size of the relevant eyebrow area, the inner corners of the nostrils, and the medial point of the mouth, respectively.

Eventually, in order to select the best of sometimes redundantly available solutions (e.g., for the fiducial points belonging to the mouth), we perform an inter-solution check. We compare, namely, the CFs of the feature points extracted from the contours spatially sampled by different detectors of the same facial component. The feature points having the highest CF are used for further analysis of shown AUs.

IV. Parametric Representation

Each AU of the FACS system is anatomically related to the contraction of a specific facial muscle [5]. Contractions of facial muscles induce motion in the skin surface and deform the shape and location of the facial components. Some of these changes in facial expression are observable from the changes in the position of the feature points. To classify detected changes in the position of the feature points in terms of facial muscle activity, the pertinent changes should be represented first as a set of suitable mid-level parameters.

We defined 6 mid-level parameters in total: 2 describing the motion of the feature points, 2 describing their state, and 2 describing shapes formed by certain feature points. The definitions of the parameters are given in Fig. 5. They are calculated for various feature points by comparing the currently extracted points with the related points extracted from the dual-view image of a neutral expression.

We assign a certainty factor $CF \in [0, 1]$ to each calculated mid-level parameter. We do so based on the CFs associated with the selected feature points (see section 3.3), whose state or motion are described by the pertinent mid-level parameter. For example:

$CF_{\boldsymbol{up/down}(P6)} = CF_{\boldsymbol{in/out}(P6)} = CF_{P6}$ ($= CF_{P4}$),
$CF_{\boldsymbol{increase/decrease}(BD)} = \boldsymbol{min}(CF_B, CF_D)$,
$CF_{angular(P6P8)} = CF_{increased\_curvature(P5P6)} = CF_{P6}$ ($= CF_{P4}$).

V. Action Unit Recognition

The last step in automatic facial expression analysis is to translate the extracted facial information (i.e., the calculated mid-level parameters) into a description of displayed facial changes such as an AU-coded description of shown facial expression. To achieve this, we apply the fast direct chaining inference process [17] to two separate sets of rules.

1) A set of 21 rules for encoding 21 AUs (AU1, AU4, AU8, AU9, AU10, AU12, AU13, AU15-AU20, AU23-AU29, AU36) occurring alone or in a combination in an input face-profile image. A full list of the utilized rules can be found in [10].

2) A set of 22 rules for encoding 22 AUs (AU1, AU2, AU4-AU8, AU12, AU13, AU15, AU18, AU20, AU23-AU28, AU35, AU38, AU39, AU41) occurring alone or in a combination in an input frontal-face image. For a full list of the used rules, see [12].

Motivated by the FACS system, each rule is defined in terms of the predicates of the mid-level representation and each encodes a single AU in a unique way according to the

Fiducial points motion	
up/down (P) = $y_{P_{neutral}} - y_{P_{current}}$ if ***up/down(P)*** < 0, ***P*** moves up	***in/out(P)*** = $x_{P_{neutral}} - x_{P_{current}}$ if ***in/out(P)*** > 0, P moves outward
Fiducial points state	
If P9 equals P7, ***absent(***P9***)***. If there is no maximum of f'' between P5 and P7, ***absent(***P6***)***. Similarly for P7', P7'' and P8 (see Fig. 3).	***increase/decrease(AB)*** = $\boldsymbol{AB}_{neutral} - \boldsymbol{AB}_{current}$, where $\boldsymbol{AB} = \sqrt{\{(x_A - x_B)^2 + (y_A - y_B)^2\}}$ If ***increase/decrease(AB)*** < 0, distance ***AB*** increases.
Shapes formed by fiducial points	
The physic meanings of ***angular(P6P8)*** = *true* and ***increased_curvature(P5P6)*** are shown below.	

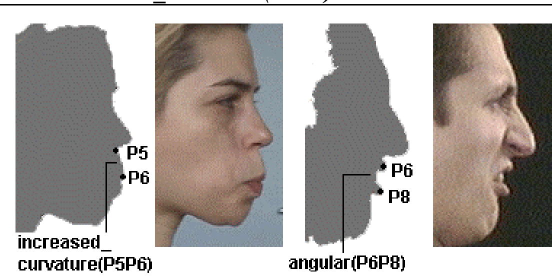

Fig. 5. Mid-level feature parameters for AU recognition

relevant FACS rule. For example, the rule used for coding AU12 in a face-profile image, which is described in the FACS system as an oblique upward pull of the lip corners (i.e., smile), is the following:

IF ***in/out***(P6) < 0 AND ***in/out***(P8) < 0 AND ***increase/decrease***(P5P6) > 0 THEN AU12.

Similarly, the rule utilized for coding AU12 in a frontal-view face image is the following:

IF (***increase/decrease***(IB) > 0 AND ***increase/decrease***(C[1]I) < 0) OR(***increase/decrease***(JB1) > 0 AND ***increase/decrease***(CJ) < 0) THEN AU12.

With each scored AU, the utilized algorithm associates a factor CF $\in$ [0, 1] denoting the certainty with which the pertinent AU has been scored. Its value equals the overall certainty factor CF_p of the premise ***p*** of the rule whose firing caused the AU in question to be scored. The certainty factor CF_p of the premise ***p*** of a fired rule is calculated as follows.

1) If ***p*** contains ***c1 AND c2***, then CF_p = ***min*** (CF_{c1}, CF_{c2}).
2) If ***p*** contains ***c1 OR c2***, then CF_p = ***max*** (CF_{c1}, CF_{c2}).
3) If ***p*** contains just clause ***c***, then $CF_p = CF_c$.
4) ($\forall$***c***) $CF_c = CF_{fp}$, where ***fp*** is the feature parameter to which clause ***c*** is related.

Some AUs could be scored twice due to the existence of the related rules in each of the two employed sets of rules (e.g., AU12). Hence, the last processing step of the utilized algorithm deals with those redundantly available scores. For each such pair of the redundantly inferred conclusions, it discards the one with which a lower CF has been associated.

VI. Experimental Evaluation

Most of the existing approaches to facial expression analysis assume that the presence of the face in the input image is ensured [9], [11]. However, in most of the real-life situations where such automated systems are to be employed (e.g., videoconferencing) the location of the face in the scene is not known a priori. The presence of a face can be ensured either by employing an existing method for automatic face detection in arbitrary scenes or by using a camera setting that will ascertain the assumption at issue. The method proposed here does not perform face detection in an arbitrary scene; it operates on dual-view face images acquired by two head-mounted CCD digital PAL cameras (Fig. 6). The camera set in front of the face acquires frontal-view images while the second camera, placed on the right side of the face, acquires face-profile images. The utilized camera setting ascertains the assumption that the examined images are orientation and scale invariant and that the face-profile-images are in right profile view (e.g., Fig 1, Fig. 2).

The test data set has been created in office environments with the help of 8 certified FACS coders drawn from college personnel. The subjects of both sexes (60% female) differed in age (20 to 35 years) and ethnicity (European, Asian and South American). The subjects were asked to display series of expressions that included single AUs and combinations of those. A total of 560 dual-view images of subjects' faces were recorded during sessions which began with displaying a neutral expression. Metadata were associated with the acquired test images given in terms of AUs scored by two FACS coders. As the actual test data set, we used 454 images for which the coders agreed about the displayed AUs. The human judgments of these 454 test images were compared further to those generated by our method. The result of the comparison is given in Table 1. It is interesting to note that, if we consider only the images in which the AUs were encoded with a CF > 0.3 (there are in total 423 such images), agreement between the generated conclusions and the pertinent human judgments is even 91%.

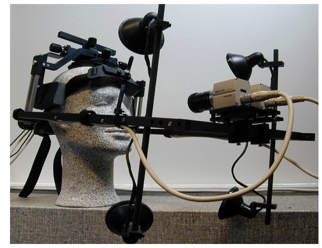

Fig. 6. Head-mounted two-cameras device

Table I

The results of facial action coding of 454 test images measured for the upper face AUs (AU1, AU2, AU4-AU7, AU41), the AUs affecting the nose (AU9, AU38, AU39), the AUs affecting the jaw (AU17, AU26, AU27, AU29), the AUs affecting the mouth (AU8, AU10, AU12, AU13, AU15, AU16, AU18-AU20, AU23-AU25, AU28, AU35, AU36), and overall:

C denotes the number of images for which the generated conclusions where identical to those scored by human coders,

PC denotes the number of images coded partially correct (some AU-codes were missing or were recognized additionally),

IC denotes the number of incorrectly coded images.

	C	PC	IC	Rate
upper face	422	32	0	93.0%
nose	443	10	1	97.6%
mouth	423	28	3	93.2%
jaw	436	17	1	96.0%
overall (all AUs)	392	58	4	**86.3%**

VII. Conclusion

In this paper, we proposed a novel, automatic method for analyzing subtle changes in facial expression based upon changes in contours of facial components and face profile contour detected in a dual-view face image. The significance of this contribution is in the following:

1) The presented approach to automatic AU recognition extends the state of the art in automatic facial gesture analysis in several directions, including the number of AUs (29 in total), the difference in AUs, and the data certainty propagation handled. Namely, the previously reported automatic AU analyzers do not assign certainty

[1] C is the middle point between the feature points H and H1.

measures to the inferred conclusions (let alone varying them in accordance with the certainty of the input data), cannot detect out-plane non-rigid movements such as the jetted jaw (AU29) and, at the best, can detect 22 AUs.

2) This paper provides a basic understanding of how to achieve automatic AU coding in both frontal-face and face-profile images. It exemplifies how, based on such knowledge, procedures of greater flexibility and improved quality can evolve (e.g., inaccurate/partial data from one facial view can be substituted by data from the other view). Hereupon further research on facial gesture analysis from multiple facial views can be based.

Nonetheless, the presented algorithm has some drawbacks. It assumes the usage of a head-mounted camera device, which reduces the freedom with which the subject can move around. It cannot analyze face images of subjects having facial hair or wearing glasses. Finally, it does not take into account the temporal nature of facial gestures. Yet, when discussing the later, it is interesting to note that the proposed method could greatly speed up the time-consuming (manual) process of acquiring AU-labeled data on which models that can capture the temporal nature of facial gestures could be trained (e.g., HMM for AU recognition). Devising such a generative probability model for temporal reasoning about AUs occurring in a face image sequence represents the main focus of our further research on this topic.

ACKNOWLEDGMENT

The current work of Maja Pantic is supported by the Netherlands Organization for Scientific Research (NWO) Grant EW-639.021.202.

REFERENCES

[1] M.S. Bartlett, et al., “Measuring facial expression by computer image analysis”, *Psychophysiology*, vol. 36, pp. 253-263, 1999.
[2] M. Black and Y. Yacoob, “Recognizing facial expressions in image sequences using local parameterized models of image motion”, *Computer Vision*, vol. 25, no. 1, pp. 23-48, 1997.
[3] J.F. Cohn, et al., “Automated face analysis by feature point tracking has high concurrent validity with manual faces coding”, *Psychophysiology*, vol. 36, pp. 35-43, 1999.
[4] J.F. Cohn, et al., “Detection, tracking and classification of Action Units in Facial Expressions”, *J. Robotics and Autonomous System*, vol. 31, pp. 131-146, 2000.
[5] P. Ekman and W. Friesen. *Facial Action Coding System*. Palo Alto: Consulting Psychologist Press, 1978.
[6] I. Essa and A Pentland, “Coding, analysis, interpretation and recognition of facial expressions”, *IEEE TPAMI*, vol. 19, no. 7, pp. 757-763, 1997.
[7] D. Keltner and P. Ekman, “Facial expression of emotion”, *Handbook of Emotions*. New York: Guilford Press, pp. 236-249, 2000.
[8] K. Mase, “Recognition of facial expression from optical flow”, *IEICE Transactions*, vol. E74, no. 10, pp. 3474-3483, 1991.
[9] M. Pantic and L.J.M. Rothkrantz, “Automatic analysis of facial expressions: The state of the art”, *IEEE TPAMI*, vol. 22, no. 12, pp. 1424-1445, 2000.
[10] M. Pantic. *Facial expression analysis by computational intelligence techniques*. PhD thesis, Delft University of Technology, 2001.
[11] M. Pantic and L.J.M. Rothkrantz, “Toward an Affect-Sensitive Multimodal HCI”, Proceedings of the IEEE, vol. 91, no. 9, pp. 1370-1390, 2003.
[12] M. Pantic and L.J.M. Rothkrantz, “Facial gesture recognition in face image sequences: A study on facial gestures typical for speech articulation”, *Proc. IEEE Int’l Conf. SMC*, vol. 6, p. 6, 2002.
[13] M. Pantic, M. Tomc and L.J.M. Rothkrantz, “A hybrid approach to mouth features detection”, *Proc. IEEE Int’l Conf. SMC*, pp. 1188-1193, 2001.
[14] M. Pantic, I. Patras and L.J.M. Rothkrantz, “Facial action recognition in face profile image sequences”, *Proc. IEEE Int’l Conf. ICME*, pp. 37-40, 2002.
[15] A. Pentland, “Looking at people”, *IEEE TPAMI*, vol. 22, no. 1, pp. 107-119, 2000.
[16] J. Russell and J. Fernandez-Dols. *The psychology of facial expression*. New York: Cambridge Univ. Press, 1997.
[17] M. Schneider et al. *Fuzzy expert system tools*. West Sussex: John Wiley & Sons, 1997.
[18] L. Shafarekno et al, “Automatic watershed segmentation of randomly textured color images”, *IEEE Trans. Image Processing*, vol. 6, no. 11, pp. 1530-1544, 1997.
[19] Y. Tian, et al., “Recognizing AUs for facial expression analysis”, *IEEE TPAMI*, vol. 23, no. 2, pp. 97-115, 2001.
[20] J.M. Vincent, et al., “Image feature location in multi-resolution images using multi-layer perceptrons”, *Neural Networks for Vision, Speech & Natural Language*. London: Chapman Hall, pp. 13-29, 1992.