Automatic Analysis and Recognition of Brow Actions and Head Motion in Spontaneous Facial Behavior*

Jeffrey F. Cohn
University of Pittsburgh
jeffcohn@cs.cmu.edu

Lawrence Ian Reed
University of Pittsburgh
lirst6@pitt.edu

Zara Ambadar
University of Pittsburgh
ambadar@pitt.edu

Jing Xiao
Carnegie Mellon University
jxiao@cs.cmu.edu

Tsuyoshi Moriyama
Keio University
moriyama@ozawa.ics.keio.ac.jp

Abstract - *Previous efforts in automatic facial expression recognition have been limited to posed facial behavior under well-controlled conditions (e.g., frontal pose and minimal out-of-plane head motion). The CMU/Pitt automated facial image analysis system (AFA) accommodates varied pose, moderate out-of-plane head motion, and occlusion. AFA was tested in video of two-person interviews originally collected to answer substantive questions in psychology, and represent a substantial challenge to automatic recognition of facial expression. This report focuses on two action units, brow raising and brow lowering because of their importance to emotion expression and paralinguistic communication. For two-state recognition, AFA achieved 89% accuracy. For three-state recognition (brow raising, brow lowering, and no brow action), accuracy was 76%. Brow and head motion were temporally coordinated. These findings demonstrate the feasibility of action unit recognition* in spontaneous facial behavior.

Keywords: Facial expression recognition, multimodal coordination, timing, spontaneous facial behavior.

1. Introduction

Facial expression is one of the most powerful, natural, and immediate means for human beings to communicate their emotions, appraisals, and intentions [27]. Affective computing strives to make use of this information for human-computer and robot interaction. Within the past few years, significant effort using computer vision has occurred to make possible the goal of automatic recognition of facial expression for human- and robot-human interaction. Several systems [13] [31] have recognized under controlled conditions a small set of emotion-specified expressions, such as joy and anger. Others [1] [20] [21] [29] have achieved some success in the more difficult task of recognizing facial action units (AU) of the Facial Action Coding System (FACS) [12]. Action units (AU) are the smallest visibly discriminable changes in facial expression. The CMU/Pittsburgh group [9, 10, 20, 27, 28] has developed a system that recognizes 20 of approximately 30 action units that have a known anatomic basis and occur frequently in emotion expression and paralinguistic communication [17, 24].

A limitation of almost all research to date in automatic facial expression recognition is that it is limited to deliberate facial expression recorded under controlled conditions that omit significant head motion and factors such as facial occlusion from glasses or facial jewellery that complicate analysis. To make use of facial expression for human-computer and human-robot interaction, facial expression must be recognized in spontaneous facial behavior and in relation to non-verbal behaviour, especially head gesture, which can modify the message value of facial expression. Smiling in the context of downward head pitch, for instance, communicates embarrassment rather than joy [8]. Efforts to recognize facial expression in spontaneous facial behavior are only now getting under way [2, 6, 9]

In the present study, we investigate action unit recognition in spontaneous facial behaviour and the temporal relation between facial action and head motion. We focus on two of the most frequent and important facial actions in the brow region, brow raising and brow lowering because of their critical role in the expression of emotion, cognition, communicative intent [4]. Brow raising (AU 1+2) can communicate surprise, greetings, and interest, provide emphasis for speech acts, and contribute to the regulation of turn-taking in social interaction. Brow lowering (AU 4) is common in facial expression of negative emotions, such as anger and fear, and can communicate cognitive states of concentration and puzzlement. To date, these brow actions have been recognized only in posed facial behaviour. With the exception of blinks (AU 45) [7], previous work in automatic facial expression recognition has been limited to posed facial behaviour measured with frontal pose and at most small out-of-plane motion. The CMU/Pitt automated facial image analysis system (AFA) is the first we know of to recognize facial actions measured in spontaneous facial behaviour with non-frontal pose, moderate out-of-plane head motion, and occlusion.

2. Automated facial image analysis

2.1. System overview

Figure 1 depicts the overall structure of the CMU/Pitt automated facial image analysis system (AFA) system for recognition of facial action units and analysis of their dynamics. A digitized image sequence is input to the system. The region of the face and location of individual facial features are delineated in the initial frame. This step is performed automatically for frontal images and manually for non-frontal images. Head motion then is recovered automatically and used to warp (or stabilize) the face image to a standard (i.e., canonical) view. Changes in both permanent (e.g., brows) and transient (e.g., furrows) facial features are automatically detected and tracked offline throughout the image sequence. Informed by FACS, facial features are grouped into separate collections of feature parameters. Parameters include feature displacement, velocity, and appearance. The extracted facial feature and head motion trajectories are fed to a classifier for action unit recognition. In addition to action unit recognition, the system quantifies the timing of facial actions and head gesture for studies of timing of facial actions.

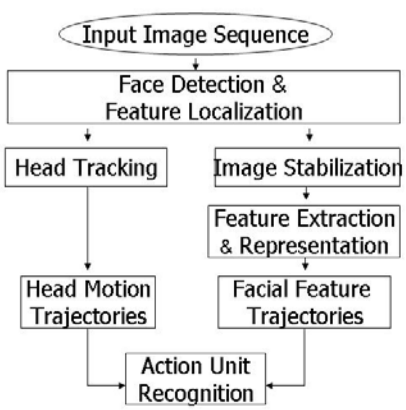

Figure 1. Overview of CMU/Pitt Automated Facial Image Analysis (AFA) System.

2.2. Face detection and feature localization

To detect the face and permanent facial features, we use the approach of Zhou, Gu, & Zhang [32]. This procedure consists of two processes, that is, face detection and facial component localization. The former is accomplished using a classifier trained on the local features, which are selected by the AdaBoost algorithm on the subspace representation of local non-negative matrix factorization (LNMF). In the latter, geometric registration and tangent shape estimation are simultaneously achieved by the estimation maximization (EM) algorithm supported by a continuous shape regularization function and a confidence measure to ensure that shape parameters are stable and accurate. The procedure performs well for frontal face images. For non-frontal face images, manual adjustment is easily performed in the initial face image.

2.3. Head tracking and image stabilization

Following [30], AFA uses a cylindrical head model to estimate the 6 degrees of freedom of head motion, whose parameters are horizontal and vertical position, distance to the camera (i.e., scale), pitch, yaw, and roll. A cylindrical model is fit to the initial face region, and the face image is cropped and "painted" onto the cylinder as the template of head appearance. For any given subsequent frame, the template is projected onto the image plane assuming the pose has remained unchanged from the previous frame. We then compute the difference between the projected image and the current frame, and the difference provides the correction on the estimate of pose.

Because head poses are recovered using templates that are continually updated and the pose estimated for the current frame is used in estimating the pose in the next frame, errors would accumulate unless otherwise prevented. To solve this problem, the system automatically selects and stores one or more reference frames and associated head poses from the tracked images. Whenever the difference between the estimated head pose and that of a reference frame is less than a preset threshold, the system rectifies the current pose estimate by re-registering that frame with the reference. The re-registration prevents errors from accumulating and enables the system to recover head pose when the head reappears after occlusion, such as when the head moves momentarily out of the camera's view. This procedure recovers head translation, scaling, and rotation for each frame. These parameters are of interest in understanding the behavior of the head and are used to stabilize the face image to a frontal view for facial and eye feature tracking. For details, see [30].

An example of automatic recovery of 3D head motion and image stabilization is shown in Figure 2. From the input image sequence (Figure 2A), the head is tracked and its pose recovered (Figure 2B). The system stabilizes the face region by transforming the image to a common orientation (Figure 2C). Note that even though the head pitches forward in the image sequence, size and orientation of the stabilized face region remain the same.

We have tested the head tracker in image sequences that include maximum pitch and yaw as large as 40^{o} and 75^{o}, respectively, and time duration of up to 20 minutes [30]. We compared the recovered motion with ground truth obtained by a position and orientation measurement device that used markers attached to the head (Optotrak® 3020 Position Sensor). The AFA head tracker was highly consistent with ground truth measurements; for example, for 75^{o} yaw, absolute error was 3.86^{o} [30]].

2.4. Feature extraction and representation

Contraction of the facial muscles produces changes in the appearance and shape of facial landmarks (e.g., brows) and in the direction and magnitude of motion on the surface of the skin and in the appearance of transient facial features (e.g., wrinkles). In the present study, we focus on two types of muscle action. Contraction of the *frontalis* muscle raises the brows in an arch-like shape and produces horizontal furrows in the forehead (AU 1+2 in FACS). Contraction of the *corrugator supercilii* and *depressor supercilii* muscles draws the inner (i.e., medial) portion of the brows together and downward and causes vertical wrinkles to form or deepen between the brows (AU 4 in FACS). To extract changes in the displacement and appearance of facial features, we use feature tracking [27] and Gabor wavelet representation [27, 28]. Head motion parameters are extracted and represented as described in section 2.3.

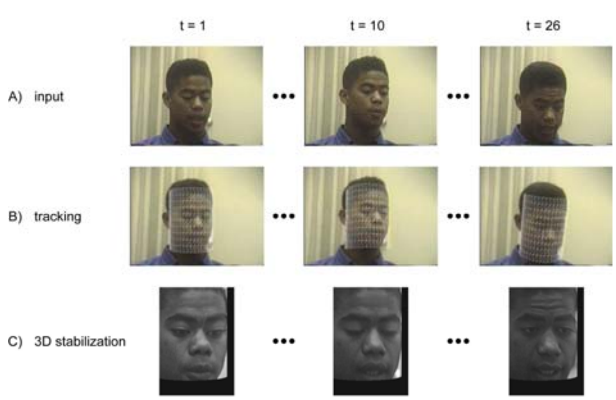

Figure 2. Automatic recovery of 3D head motion and image stabilization. A) Frames 1, 10, and 26 from the original image sequence. B) Automatic face tracking in the corresponding frames. C). Stabilized face images. From [9].

2.4.1. Facial feature tracking

To track facial features, we use the Lucas-Kanade feature tracking algorithm. Because the 3D face shape can be represented by a linear combination of shape bases or structures, the 2D image location of facial features must lie in a low-dimensional linear space. The dimension of the space is dependent on the number of shape bases. Therefore, we can apply subspace constraints on the locations of facial features to attenuate noise [16].

Pixel coordinates of the feature points in the stabilized brow region are computed relative to the most frontal and upright image in the sequence. In the current study, we focus on feature points in the brow region. Upward brow motion, referred to as *brow raise,* is represented as the mean vertical displacement of feature points in the brows. *Medial contraction* (AU 4) is computed as described in [5, 25] as

$$\Delta d = \sqrt{\Delta x^2 + \Delta y^2} \quad (1)$$

The system computes maximum velocity of these features as well.

2.4.2. Gabor wavelet representation

To extract appearance features in the stabilized image sequence, we use multi-scale, multi-resolution Gabor wavelets. For the experiment reported below, Gabor filters were applied at 6 specific locations, which were the nasal root region, the region above the center of each brow, and the medial, left and right sides of the forehead. The response image of the Gabor filter is the correlation of the input image $I(x)$ with the Gabor kernel as described in [28]. In the current study, we used 8 orientations ranging from 0 to π differing by $\pi/8$ and one spatial frequency, $k_i = \pi/8$ to avoid redundancy [26].

2.5. Action unit recognition

In previous work, we have experimented with various approaches of classifying feature parameters into action units. These include hidden Markov models (HMM) and discriminant analysis [20], rule-based recognition [9], and neural networks [27]. In the work presented here, discriminant analysis was used for action unit recognition. Measurements included vertical displacement and maximum velocity of *brow raise,* displacement and maximum velocity of the *medial contraction* of the brow, head pitch velocity, and difference in Gabor coefficients between initial and target image. Each measurement was represented by a *2p* dimensional vector by concatenating *p* brow displacements and velocities, head pitch velocity, and change in Gabor coefficients; that is

$$\boldsymbol{D} = (\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots\ldots, \boldsymbol{d}_p) = (d_{1x}, d_{1y}, d_{2x}, d_{2y}, \ldots d_{px}, d_{py}) \quad (2)$$

The discrimination between action units was done by computing and comparing the *a posteriori* probabilities of action units AUs; that is

$$\boldsymbol{D} \rightarrow AU_k \quad \text{if} \quad p(AU_k \mid \boldsymbol{D}) > p(AU_j \mid \boldsymbol{D}) \quad (3)$$
$$j \neq k$$

where

$$p(AU_i \mid \mathrm{D}) = \frac{p(\mathrm{D} \mid AU_i)\,p(AU_i)}{p(\mathrm{D})} = \frac{p(\mathrm{D} \mid AU_i)\,p(AU_i)}{\sum_{j=1}^{k} p(\mathrm{D} \mid AU_j)\,p(AU_j)} \quad (4)$$

The discriminant function between AU_i and AU_j is therefore the log-likelihood ratio:

$$f_{ij}(\mathrm{D}) = \log \frac{p(AU_i \mid \mathrm{D})}{p(AU_j \mid \mathrm{D})} = \log \frac{p(\mathrm{D} \mid AU_i)\,p(AU_i)}{p(\mathrm{D} \mid AU_j)\,p(AU_j)} \quad (5)$$

The $p(\boldsymbol{D} \mid AU_i)$ are assumed to be a multivariate Gaussian probability distribution $N(\boldsymbol{u}_i, \Sigma_i)$, where the mean $\boldsymbol{u}_i$, and the covariance matrix Σ_i are estimated by the sample means and sample covariance matrices of the training data. While the d_i are not independent, [11]has shown that Bayes classifiers are robust to violations of this assumption and perform surprisingly well. The discriminant function is a quadratic discriminant function in general; but if covariance matrices Σ_i and Σ_j are the same, it reduces to a linear discriminant function.

3. Experimental procedures

3.1. Participants and procedures

Video data were from two studies of spontaneous facial behavior by Ekman and colleagues. In the first [14] 20 young-adult men of varied ethnicity lied or told the truth to an interviewer about whether they had stolen a large sum of money. Data from 16 were available for analysis. Nine of the 16 were of European heritage, 4 of African heritage, and 3 Asian. Four wore glasses. In the other [23], 85 older men and women were interviewed about health-related behavior while undergoing a cardiac scan. Data from 5 (4 men and 1 woman of European ancestry) were available for analysis. None wore glasses.

All image sequences were recorded using an S-Video camera. Head orientation to the camera was about plus/minus 20-30 degrees from frontal, and out-of-plane head motion was common. Video sequences were digitized into 640x480 pixel arrays with 16-bit color resolution. In the original studies, certified FACS coders manually FACS-coded start and stop times for all action units in 1- to 2-minutes of facial behavior. Certified FACS coders from the University of Pittsburgh confirmed all coding. Sequences for which manual FACS coding was not confirmed were excluded.

3.2. Facial actions

We focus on automatic recognition of AU 1+2 (inner and outer brow are raised), AU 4 (inner brows are pulled together and lowered), and randomly selected comparison sequences (i.e., AU 0 or neutral) in which no brow motion occurred. AU 1+2 and AU 4 are the two most common action units in the brow region. AU 1+2 communicates surprise, greetings, and provides paralinguistic emphasis for speech[4]. AU 4 is common to negative emotions (e.g., sadness, anger, and fear), concentration and cognitive effort[4]. Other action units in the brow region occurred too infrequently to include for analysis. These action units included AU 1 (inner brow raise in absence of AU 2), AU 1+4, found typically only in sadness, and AU 1+2+4, typically only in fear. These expressions were relatively rare in the image data from 2-person interviews we analyzed.

Criteria for analysis were confirmation of the original FACS coding by a certified FACS coder in our laboratory, action unit intensity of 'b' or higher on a scale from 'a' (trace) to 'e' (maximal intensity), and no co-occurrence of action unit 9 with AU 4 or of action unit 4 with AU 1+2. AU 4 in the presence of AU 9 was omitted because AU 9 also lowers the brows. Recognition of AU 4 in the presence of AU 9 could serendipitously result from the action of either AU 4 or AU 9. Thus, omitting this combination provided a more conservative rigorous test of AU 4 recognition. We considered including AU 4+9 as a separate AU, but too few were available. Similarly, too few occurrences of AU 1+2+4 were available for analysis. Most action units were of small to moderate intensity.

4. Results

4.1. Coordination of head motion and facial actions

Head motion was highly coordinated with brow motion, with Pearson correlation coefficients in the moderate range. Brow raising was more likely to occur with forward head pitch. Otherwise, the number of positive and negative correlation coefficients was comparable for brow raising and lowering. Table 1 reports the average Pearson correlation among parameters ignoring the sign of correlation.

Figure 3 shows an example of system output for a sequence in which brow lowering occurs as the head pitches slightly upward; this is followed by sustained brow lowering and head orientation followed by a pattern of alternating brow contraction and relaxation of coordinated changes in pitch and yaw beginning at about frame 133.

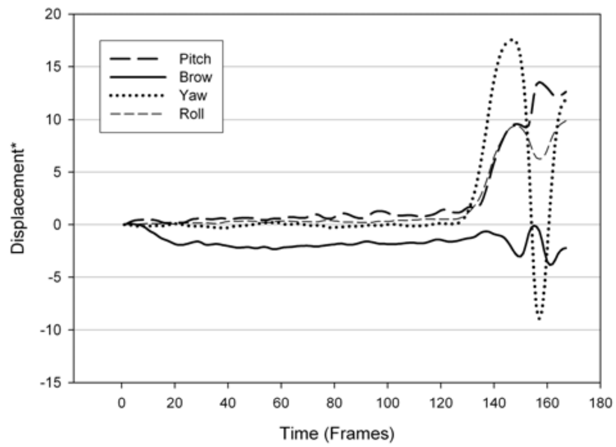

Figure 3. An example of system output for AU 4. *Note: Displacements of brow and head motion are in pixels and degrees, respectively.

The mean differences between start of brow motion and pitch, roll, and yaw were 0.12, 2.16, and 1.41 frames, respectively, with a mode of 0 frames. Standard deviations were relatively large, however, ranging from 12.94 to 17.36, respectively. Mean differences in start time did not vary between AU 1+2 and AU 4.

Table 1. Average Pearson correlation between head motion and brow motion

	Pitch	Roll	Yaw	Brow
Pitch	--	.59	.58	.43
Roll		--	.65	.44
Yaw			--	.56

Note. Sign of the correlation omitted in computing average Pearson correlation. Direction and magnitude of correlation coefficients among brow and head motion parameters were similar in AU 1+2 and AU 4.

4.2. Action unit recognition

The analyses used 99 samples of the three action units (AU 1+2, AU 4, and AU 0) that occurred 25 or more times in the 21 participants. The *a priori* probabilities $p(AU_i)$s were based on the relative frequency of each action unit in the sample data. Cross-validation was by leave-one-out procedure.

Table 2. Automatic recognition of action units in spontaneous facial behavior.

	Automated Facial Image Analysis		
Manual FACS coding	*AU 1+2*	*AU 4*	*AU 0*
AU 1+2 (brow raiser)	82%	7%	12%
AU 4 (brow lowerer)	13%	60%	27%
AU 0 (no brow action)	4%	12%	84%

Note. Overall recognition rate for three-state recognition = 76%. Recognition rate for two-state classification (AU 1+2 versus AU 4) = 89%.

With three action units, there are two possible discriminant functions. Wilks' lambda for both discriminant functions was .288, $p < .001$. Approximately 76% of action units were correctly recognized (Table 2). Most disagreements were between AU 4 and comparison sequences in which no brow action occurred (AU 0). For the two-state classification between AU 1+2 and AU 4, recognition accuracy was 89%.

5. Conclusions

Brow raising and lowering are important to paralinguistic communication and emotion expression. We developed and implemented an automatic facial image analysis system that recognizes these actions in spontaneous facial behavior. The actions were of low to moderate intensity and embedded in the natural flow of behavior during a two-person interview. For the 3-state classification (AU 1+2, AU 4, and AU 0 or neutral) AU 1+2 was recognized with accuracy approaching what has been reported for posed expressions measured under more controlled conditions. Accuracy for AU 4 was lower. Most errors for AU 4 were with AU 0 (neutral). For the 2-state classification of AU 1+2 versus AU 4, 89% accuracy was achieved.

Several factors may account for decreased accuracy relative to what has been reported previously for action unit recognition in posed facial behaviour. In contrast to previous research in automatic facial expression recognition, the image data were originally collected for research in psychology and not for purposes of training and testing computer vision algorithms. The image data were from interview situations in which orientation to the camera varied as much as about plus/minus 30 degrees, approximately 20% of subjects wore glasses, which partially occluded the brows and nasal root, moderate out-of-plane head motion was common, and action unit intensity typically was low to moderate. In contrast, in studies of posed facial behaviour, pose is uniform across subjects, orientation is frontal, out-of-plane head motion is absent or small, and intensity ranges from moderate to large. Real world conditions provide a more challenging context for action unit recognition. Given these differences between spontaneous, naturally occurring facial behaviour and posed facial behaviour, decreased accuracy is not surprising. Further research is needed to understand the specific challenges presented by each source of natural variation (i.e., pose, out-of-plane head motion, occlusion, and action unit intensity). The present findings support the feasibility of automatic recognition of naturally occurring facial actions and their potential use in human- and robot-human interaction.

Novel to the current work was explicit attention to head and facial feature velocities in action unit recognition. These factors previously have been either ignored or considered as sources of error. We found that velocity measures added to the predictive power of traditional approaches for action unit recognition. Brow raising, in particular, was more likely to occur as the head pitched forward (e.g., Figure 2). This pattern may result from attempts to maintain focus on the partner as the head pitches forward and down, and is opposite to what would be anticipated in expressions of surprise, in which brow raising occurs in the context of upward head pitch. The findings suggest that the meaning of facial actions may best be disambiguated by attending to co-occurring head motion. More generally, they support the hypothesis that facial actions occur as part of coordinated motor routines [7]. In future work, it will be important to investigate the heterogeneity of head motion and facial feature coordination and its relation to communicative intention and meaning.

Neuro-anatomic studies have found close coordination between head motion and gaze [18, 19]. The present study suggests that a high degree of coordination is possible between facial expression and head motion as well; yet it also suggests significant variability. While the modal difference between the start of brow motion and pitch, roll, and yaw was zero frames, the range of difference scores was relatively large. From a psychological perspective, it is likely that increasing lag times between the start of facial expression and head motion alter their communicative meaning and influence how observers parse the stream of behavior.

In the computer graphics literature, there has been considerable interest in creating life-like computer avatars. While many of the challenges are technical, such as those of realistic image rendering, a major limitation of work to date is the lack of fusion among different modalities. Head motion in avatars typically is absent or appears to occur randomly, which undercuts verisimilitude and can corrupt the intended message value. To create more realistic and higher fidelity avatars, normative data about the fusion of facial expression and head as well as body motion is needed. Because both the configuration and timing of facial expression and head and body motion are important, we would urge that future work consider both aspects. The research described in this report and in related work [5, 7] is a start in this direction. Our research program is an initial attempt to address the timing as well as configuration of behavior by providing quantitative data about the relation between brow motion, smiling and head and eye motion [7]. Further work is needed that samples additional interactive contexts and includes additional facial features.

A limitation of the present work for human-computer interaction is that it involved person-person rather than human-computer interaction. When interacting with a computer, facial expression may differ relative to social interaction. In the absence of social context, facial expression may be more muted. Fridlund [15] found that facial expression occurred with less intensity when people when viewing video clips alone rather than with other people. Designers of human-computer interfaces may want to consider how to increase the social salience of the experience. While [22] has argued that people tend to imbue computers and other devices with persona and treat them as if they were animate, thus implying social behaviour, little is known about whether people are equally expressive whether interacting with computers or people under varying conditions. Fridlund's findings would suggest not. Efforts to increase the social salience and animus of computers may prove essential to successful development of more natural interfaces. In the case of robot-computer interaction, we suspect that social context would be highly salient. Initial work in this area suggests this is the case [3].

In summary, we found high coordination between brow and head motion and encouraging results in regard to automatic recognition of facial actions in spontaneous facial behavior with variable pose, moderate out-of-plane head motion, and partial occlusion. Brow raising, lowering, and non-brow action comparison sequences were discriminated with 76% accuracy. The latter result may be a conservative estimate of what can be achieved. Inclusion of additional features and alternative pattern recognition approaches may increase accuracy.

6. Acknowledgements

This research was supported by Grant R01 MH51435 from the National Institute of Mental Health and CIA Contract 2000-A128400-000 to Jeffrey F. Cohn and Takeo Kanade. Paul Ekman, Mark Frank, Erika Rosenberg, and James Blumenthal generously provided image- and metadata and consulted on various aspects of the work.

7. References

[1] M. S. Bartlett, P. Ekman, J. Hager, C., and T. J. Sejnowski, "Measuring Facial Expressions by Computer Image Analysis," *Psychphysiology*, vol. 36, 1999.

[2] M. S. Bartlett, G. Littlewort, and J. R. Movellan, "The next generation of automatic facial expression analysis for human-comptuer interaction and behavioral science.," presented at IEEE Conference on Systems, Man, and Cybernetics, The Hague, The Netherlands, 2004.

[3] C. Breazeal, *Designing sociable robots.* Cambridge, MA: MIT, 2002.

[4] J. F. Cohn and P. Ekman, "Methods for measuring facial actions," in *Handbook of nonverbal behavior research methods in the affective sciences*, J. A. Harrigan, R. Rosenthal, and K. Scherer, Eds. New York: Oxford, In press.

[5] J. F. Cohn and T. Kanade, "Use of automated facial image analysis for measurement of emotion expression," in *The handbook of emotion elicitation and assessment*, *Oxford University Press Series in Affective Science*, J. A. Coan and J. B. Allen, Eds. New York, NY: Oxford, In press.

[6] J. F. Cohn, L. I. Reed, Z. Ambadar, J. Xiao, and T. Moriyama, "Automatic analysis and recognition of facial actions and out-of-plane head motion in spontaneous facial behavior," presented at IEEE Conference on Systems, Man, and Cybernetics, the Hague, the Netherlands, 2004.

[7] J. F. Cohn, L. I. Reed, T. Moriyama, J. Xiao, K. Schmidt, and Z. Ambadar, "Multimodal coordination of facial action, head rotation, and eye motion," presented at Sixth IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, Korea, 2004.

[8] J. F. Cohn and K. Schmidt, "The timing of facial motion in posed and spontaneous smiles," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 2, pp. 1-12, 2004.

[9] J. F. Cohn, J. Xiao, T. Moriyama, Z. Ambadar, and T. Kanade, "Automatic recognition of eye blinking in spontaneous facial behavior," *Behavior Research Methods, Instruments, and Computers*, vol. 35, pp. 420-428, 2003.

[10] J. F. Cohn, A. Zlochower, J. Lien, and T. Kanade, "Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding," *Psychophysiology*, vol. 36, pp. 35-43, 1999.

[11] P. Domingos and M. Pazzani, "On the optimality of the simple Baysean classifier under zero-one loss," *Machine learning*, vol. 29, pp. 103-130, 1997.

[12] P. Ekman, W. Friesen, and J. C. Hager, "Facial action coding system," Research Nexus, Network Research Information, Salt Lake City, UT, 2002.

[13] I. Essa and A. Pentland, "Coding, analysis, interpretation and recognition of facial expressions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, pp. 757-763, 1997.

[14] M. G. Frank and P. Ekman, "The ability to detect deceit generalizes across different types of high-stakes lies.," *Journal of Personality and Social Psychology*, vol. 72, pp. 1429-1439, 1997.

[15] A. J. Fridlund, "Sociality of solitary smiling: Potentiation by an implicit audience," *Journal of Personality and Social Psychology*, vol. 60, pp. 229-240, 1991.

[16] M. Irani, "Multi-frame optical flow estimation using subspace constraints," presented at IEEE International Conference on Computer Vision, 1999.

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

[18] W. M. King, S. G. Lisberger, and A. F. Fuchs, "Response of fibers in medial longitudinal fasciculus (MLF) of alert monkeys during horizontal and vertical conjugate eye movements evoked by vestibular or visual stimuli," *Neurophysiology*, vol. 39, pp. 1135-49, 1976.

[19] E. M. Klier, W. Hongying, and J. D. Crawford, "Three-dimensional eye-head coordination is implemented downstream from the superior colliculus," *Journal of Neurophysiology*, vol. 89, pp. 2839-2853, 2003.

[20] J. J. J. Lien, T. Kanade, J. F. Cohn, and C. C. Li, " Detection, tracking, and classification of subtle changes in facial expression," *Journal of Robotics & Autonomous Systems*, vol. 31, pp. 131-146, 2000.

[21] M. Pantic and M. Rothkrantz, "Facial action recognition for facial expression analysis from static face images," presented at IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2004.

[22] B. Reeves and C. Nash, *The media equation: How people treat computers, television and news media like real people and places*. Cambridge: Cambridge University, 1996.

[23] E. Rosenberg, P. Ekman, and J. A. Blumenthal, "Facial expression and the affective component of cynical hostility in male coronary heart disease patients." *Health Psychology*, vol. 17, pp. 376-380, 1998.

[24] M. A. Sayette, J. F. Cohn, J. M. Wertz, M. A. Perrott, and D. J. Parrott, "A psychometric evaluation of the Facial Action Coding System for assessing spontaneous expression.," *Journal of Nonverbal Behavior*, vol. 25, pp. 167-186, 2001.

[25] K. Schmidt, J. F. Cohn, and Y. Tian, "Signal characteristics of spontaneous facial expressions: Automatic movement in solitary and social smiles," *Biological Psychology*, vol. 65, pp. 49-66, 2003.

[26] Y. Tian, "Evaluation of face resolution for expression analysis," presented at IEEE Workshop on Face Processing in Video, Washington, DC, 2004.

[27] Y. Tian, T. Kanade, and J. F. Cohn, "Recognizing action units for facial expression analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 97-115, 2001.

[28] Y. L. Tian, T. Kanade, and J. F. Cohn, "Evaluation of Gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity," presented at Proceedings of the IEEE Conference on Automatic Face and Gesture Recognition, 2002, May.

[29] M. F. Valstar, M. Pantic, and I. Patras, "Motion history for facial action detection in video," presented at Proceedings of the IEEE Conference on Systems, Man, and Cybernetics, the Hague, the Netherlands, 2004.

[30] J. Xiao, T. Kanade, and J. F. Cohn, "Robust full motion recovery of head by dynamic templates and re-registration techniques," *International Journal of Imaging Systems and Technology*, vol. 13, pp. 85-94, 2003.

[31] Y. Yacoob and L. Davis, "Recognizing human facial expression from long image sequence using optical flow," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 636-642, 1997.

[32] Y. Zhou, L. Gu, and H. J. Zhang, "Bayesian tangent shape model: Estimating shape and pose parameters via Bayesian inference," presented at IEEE Conference on Computer Vision and Pattern Recognition, 2003.