To appear in *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FG'04)*, Seoul, Korea, xxx-xxx.

Multimodal Coordination of Facial Action, Head Rotation, and Eye Motion during Spontaneous Smiles

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

Lawrence Ian Reed
University of Pittsburgh
lirst6@pitt.edu

Tsuyoshi Moriyama
Carnegie Mellon University
tmoriyam@cs.cmu.edu

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

Karen Schmidt
University of Pittsburgh
kschmidt@ucsur.pitt.edu

Zara Ambadar
University of Pittsburgh
ambadar@pitt.edu

Abstract

Both the configuration of facial features and the timing of facial actions are important to emotion and communication. Previous literature has focused on the former. We developed an automatic facial expression analysis system that quantifies the timing of facial actions as well as head and eye motion during spontaneous facial expression. To assess coherence among these modalities, we recorded and analyzed spontaneous smiles in 62 young women of varied ethnicity ranging in age from 18 to 35 years. Spontaneous smiles occurred following directed facial action tasks, a situation likely to elicit spontaneous smiles of embarrassment. Smiles (AU 12) were manually FACS coded by certified FACS coders. 3D head motion was recovered using a cylindrical head model; motion vectors for lip-corner displacement were measured using feature-point tracking; eye closure and horizontal and vertical eye motion (from which to infer direction of gaze or visual regard) were measured by a generative model fitting approach. The mean correlation within subjects between lip-corner displacement, head motion, and eye motion ranged from +/- 0.36 to 0.50, which suggests moderate coherence among these features. Lip-corner displacement and head pitch were negatively correlated, as predicted for smiles of embarrassment. These findings are consistent with recent research in psychology suggesting that facial actions are embedded within coordinated motor structures. They suggest that the direction of correlation among features may discriminate between facial actions with similar morphology but different communicative meaning, inform automatic facial expression recognition, and provide normative data for animating computer avatars.

1. Introduction

Both the configuration of facial features and the timing of facial actions are important to emotion and communication [8, 9]. Previous literature in automatic facial expression recognition has focused on the former. Little attention has been given to the timing of facial actions and its relation to head and eye motion. Head motion is typically considered more as a nuisance variable than as a meaningful component of nonverbal communication. Investigators have either selected image data in which head motion was absent [1, 21, 22] or relatively small and parallel to the image plane of the camera [25, 27]. Several investigators have developed approaches to recover 3D head motion [2, 4, 26]. Such models have been used to separate rigid (head) from non-rigid (expression) motion prior to automatic facial expression recognition [e.g., 7]

With few exceptions, the timing of facial actions and their relation to head and eye motion have received little attention. In a preliminary study, Cohn & Schmidt [6] found that the timing of lip-corner displacement varied between spontaneous and deliberate smiles, with only spontaneous smiles showing a highly consistent relation between maximum velocity and amplitude of lip-corner displacement. Recent work by the authors [23] suggests that human observers are highly sensitive to differences in timing of facial actions; human judgments of a smile's genuineness are strongly influenced by smile dynamics independently of configuration. In biometrics, Hill and Johnson [12] found that individual differences in rigid head motion inform person recognition. Fox and colleagues [10] more specifically found that the timing of mouth movement during speech conveys unique information about person identity. Person recognition was maximized by combining mouth movement with acoustic and facial appearance parameters.

In neuroscience, the coordination of head and eye movement is well established, and neuroanatomic mechanisms have been a research focus [16, 17]. Literature on the coordination of head or eyes with facial action, apart from case studies of facial paralysis, remains unexamined to our knowledge. The common path that coordinates head and eye movement, the medial longitudinal fasciculus, is located in the brainstem proximal to the facial nerve nucleus. Connections between them could enable coordination.

In behavioral science, several investigators using manual human-observer based methods have found reliable correlation between head and eye motion and facial actions; and the pattern of correlations varies with emotion and context. Smiles of enjoyment occur while turning toward another person [18], while smiles of embarrassment occur as gaze and head orientation are directed away from another person [15]. Brow-raising often occurs with visual search [3] and upward gaze and head motion [20]. In infants, surprise expressions in the absence of the actual emotion can be induced by changes in visual attention and orientation alone [5]. Surprise expressions and head tilt occur as infants visually track objects that are moving vertically. Findings such as these, while qualitative and based on human observation rather than quantitative instrumentation, are consistent with the hypothesis that head and eye motion are highly correlated with facial action. And further, that the meaning of morphologically similar facial actions (e.g., smiling) can be disambiguated by attending to differential patterns of head motion.

We developed an automatic facial expression analysis system that quantifies the multimodal timing of head and eye motion with facial actions. We tested the hypotheses that (1) these modalities are highly correlated during spontaneous behavior and (2) that smiles in a context likely to elicit embarrassment show a negative correlation between head motion and smile intensity. To assess coherence among modalities, we recorded and analyzed spontaneous smiles in 62 young women ranging in age from 18 to 35 years. We found moderate correlations within subjects over the course of spontaneous smiles between lip-corner displacement, a measure of smile intensity, head motion (rotation and translation) and eye motion (horizontal and vertical translation of the iris). Lip-corner displacement in the smiles we studied was negatively correlated with head pitch and, in a trend, positively correlated with vertical iris translation.

These findings are consistent with recent research in psychology, which suggests that facial actions are embedded within coordinated motor structures. They have implications for automatic facial expression recognition and interpretation and can provide normative data for more realistic animation of computer avatars.

2. Automatic facial expression analysis

2.1. System overview

Figure 1 shows an overview of the system for analysis of facial feature trajectories, head translation and rotation, and eye state. Eye state includes both eye opening and closure and eye position within the socket (i.e., horizontal and vertical translation of the eye). From a digitized image sequence, the face is automatically detected and facial features are localized. 3D head tracking using a cylindrical head model recovers 6*df* of head motion. These parameters are used to stabilize the face image for tracking facial features and to correlate head rotation with facial feature trajectories and eye motion across an image sequence. Rate of eye closure (i.e., blinking) is a measure of autonomic arousal and is believed to increase during deceptive behavior; eye position indicates direction of visual regard or gaze. Our primary interest is in the relation between head rotation, lip-corner displacement, and horizontal and vertical eye motion (i.e., translation) within the eye socket during spontaneous (i.e., naturally occurring) smiles.

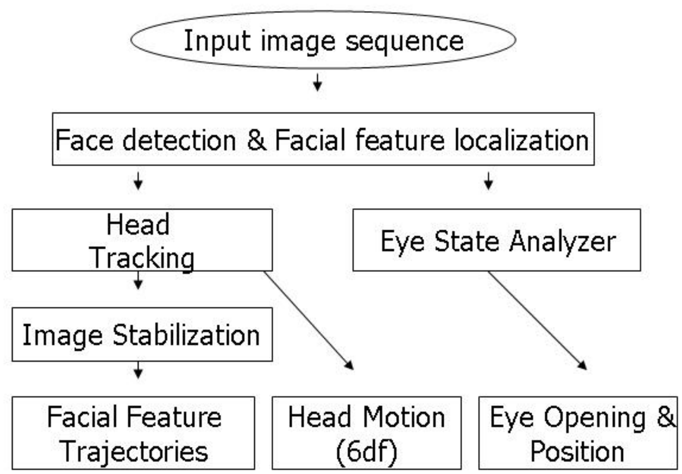

Figure 1. Automated Facial Feature, Head Motion, and Eye State Analysis System.

2.2. Face detection and facial feature localization

To detect the face and permanent facial features, we use the approach of Zhou, Gu, & Zhang [28]. 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 and facial feature tracking

Head and facial feature tracking are accomplished concurrently. For head tracking, the initial head pose is given either automatically, as described in the preceding paragraph, or by manually initializing a single reference image. Following [26], a generic cylindrical model then is utilized to estimate the 6 degrees of freedom of head motion in real-time. For any given frame, the template is the head image in the previous frame that is projected onto the cylindrical model. Then the template is registered to the head appearance in the given frame to recover the full motion of the head. The iteratively re-weighted least squares technique is used to account for non-rigid motion and occlusion. The template is dynamically updated to accommodate gradual changes in lighting and self-occlusion. This step enables the system to work well even when most of the face is occluded. 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, please see [26].

To track facial features in the stabilized image, we use the Lucas-Kanade [19] feature tracking algorithm. Because the 3D face shape can be represented by a linear combination of certain 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 [13].

2.4. Tracking of eye opening and eye position

We track eye opening and closure and the position of the iris within the eye region. From the position of the iris, we can infer direction of gaze or visual regard. When the iris is centered within the eye region, we infer that gaze is directed ahead. When the iris is to the side, we infer that gaze is to the side.

To track eye opening and closure and position of the iris, we use a generative model fitting approach. Figure 2 shows the generative eye model that generates the appearance of an eye. The model has 3 parameters in this implementation: iris location (x_{iris}, y_{iris}) and height of upper eyelid relative to the lower eyelid, h. Let $T(\mathbf{x};\mathbf{p})$ represent the template, where $\mathbf{x} = (i, j)^T$ is a column vector containing the pixel coordinates on the template coordinate system, $\mathbf{p} = (h, x_{iris}, y_{iris})^T$ is the 3 model parameters. To fit the eye model to the input eye region image, Lucas-Kanade algorithm [19] is extended, which basically minimizes the mean square error between the input image and the template, fine-tuning the template location to the input eye region. In other words, the problem of finding the optimal $\mathbf{p}$ is translated into that of minimizing the following objective function D:

$$D = \sum [T(\mathbf{x};\mathbf{p} + \delta\mathbf{p}) - I(W(\mathbf{x};\mathbf{t} + \delta\mathbf{t}))]^2 \qquad (1)$$

W denotes the parameterized set of allowed warps and $\mathbf{t}$ is a parameter vector of the warp. In this implementation, $\mathbf{t}$ contains only translation $\mathbf{t} = (t_1, t_2)^T$.

$$\frac{\partial D}{\partial \mathbf{p}} = 0, \qquad \frac{\partial D}{\partial \mathbf{t}} = 0 \qquad (2)$$

$$\Leftrightarrow \quad \sum [\nabla T]^T \left[T(\mathbf{x};\mathbf{p}) + \nabla T \delta\mathbf{p} - I(W(\mathbf{x};\mathbf{t})) - \nabla I \frac{\partial W}{\partial \mathbf{t}} \delta\mathbf{t} \right] = 0 \qquad (3)$$

$$\sum \left[\nabla I \frac{\partial W}{\partial \mathbf{t}} \right]^T \left[T(\mathbf{x};\mathbf{p}) + \nabla T \delta\mathbf{p} - I(W(\mathbf{x};\mathbf{t})) - \nabla I \frac{\partial W}{\partial \mathbf{t}} \delta\mathbf{t} \right] = 0. \qquad (4)$$

Simultaneous equations (3) and (4) can be rewritten as follows:

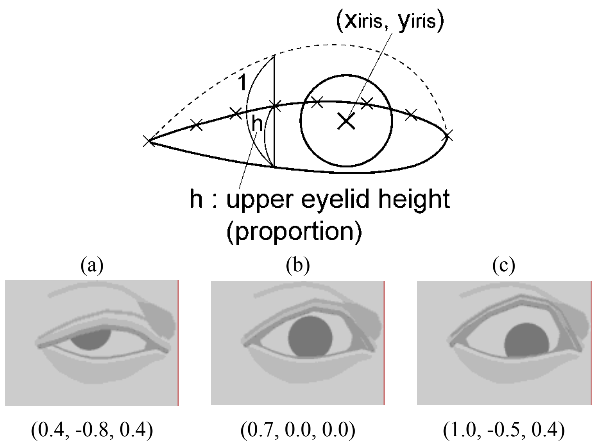

Figure 2. Generative parameterized eye template with example parameter combinations.

$$\mathbf{A}\cdot\mathbf{v}=\mathbf{b} \quad \Leftrightarrow \quad \mathbf{v}=\mathbf{A}^{-1}\cdot\mathbf{b}, \qquad (5)$$

where $\mathbf{v}=[\delta\mathbf{p}\ \delta\mathbf{t}]^T$. Given initial parameter set $\mathbf{t}_0=(t_{01},t_{02})^T$ and $\mathbf{p}_0=(h_0,x_0,y_0)^T$, $\mathbf{p}$ is updated frame by frame until $\delta\mathbf{p}$ and $\delta\mathbf{t}$ get less than predefined thresholds:

$$\mathbf{p}\leftarrow\mathbf{p}+\delta\mathbf{p}, \quad \mathbf{t}\leftarrow\mathbf{t}+\delta\mathbf{t} \qquad (7)$$

The initial frame in this process is ideally one in which the face image is the most frontal and upright, which can be determined either automatically or manually. Because this module is integrated with a face and facial feature detector, as described above, it can directly use a frame selected by face detector or by manual selection.

2.5. Coherence among lip-corner displacement, head motion, and eye (iris) motion

With minor modifications, we followed the approach of [24] to calculate lip-corner displacement. Pixel coordinates of the left lip corner in the stabilized image sequence were computed as described in Section 2.3 relative to the most frontal and upright frame in the sequence. The initial (x,y) coordinate of left lip corner in each participant's segment was designated as the origin point, (0,0), of which the displacement of the lip corner from this origin point was recalculated for each frame as:

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

Displacement values were collected for each frame of each participant's segment, forming a time series of lip corner positions. Although this calculation will only give a value for displacement and not direction, it is already known that the participants smiled during the segment, thus direction may be inferred. Results from the time series data were smoothed using the SPSS T4253H algorithm. The same smoothing function also was applied to time series of head translation and rotation and eye position.

For each participant, we computed Pearson correlation coefficients between the smoothed time series for lip-corner displacement, head translation and rotation, and eye (iris) translation from the beginning to the end of the smile (AU 12).

3. Participants and observational procedures

3.1. Participants

Participants were 62 self-selected, female undergraduate students from a larger study [14] who participated to earn course credit. They ranged in age from 18 to 35 years and were of both Euro- and African American heritage.

3.2. Observational Procedures

Subjects were observed while performing a series of directed facial action tasks [14]. During the directed facial action task, spontaneous smiles occurred during the course of the recording session. Spontaneous smiles were defined as instances of FACS (Facial Action Coding System) [9] action unit 12 as determined by certified FACS coders that occurred in the absence of specific instructions to smile. Any smiles that occurred just prior to, during, or immediately after the instructions to produce a deliberate smile were excluded. Spontaneous smiles were excluded if the onset was non-neutral, if the participant was talking, or there was occlusion (i.e. hands to the face). A neutral display was required in the initial frame of the image sequence, particularly in the lips, because the automatic measurements were based on the assumption of a neutral starting point. Smiles occurring during speech were not included because speech production may alter the temporal pattern of the smile. For each subject, we analyzed the first spontaneous smile.

4. Results

We first present descriptive statistics and then present data on the relation between facial action, head motion, and eye motion during spontaneous smiles.

4.1. Descriptive statistics

Mean smile duration was 2.15 seconds (*sd* = 1.24), which is within the range reported previously for spontaneous smiles [11]. Head pitch, roll, and yaw ranged from small to moderate (An example is shown in Figure 3).

To appear in *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FG'04)*, Seoul, Korea, xxx-xxx.

4.2. Temporal coordination of lip-corner displacement, head rotation, and eye motion

Figure 3 shows an example of lip-corner displacement, head motion, and eye motion in a spontaneous smile. The smile begins at around frame 170. As smile intensity increases, the head pitches down and to the left (frame 185); smile intensity then decreases as the head rotates upwards (frame 198) and then turns toward the camera and experimenter. Figure 4 shows the corresponding time series for selected parameters. For economy of presentation, not all parameters are shown in the example. Note the close inverse association between lip-corner displacement and head pitch, which is characteristic of embarrassment smiles.

Figure 3. Selected frames from image sequence showing the relation between lip-corner displacement, head rotation, and eye motion.

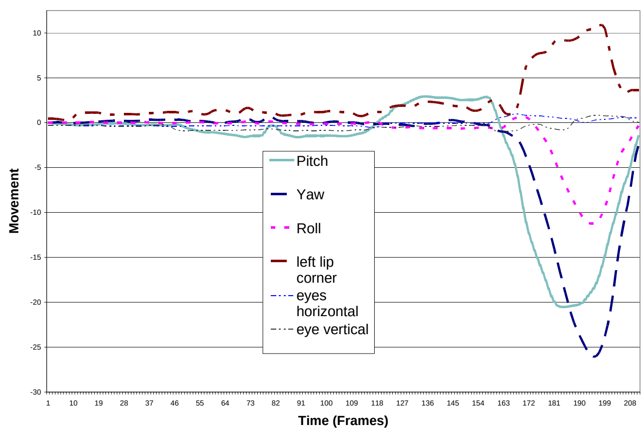

Figure 4. Time series of lip-corner displacement, head rotation, and eye motion for the example shown in the preceding figure.

To quantify the association between modalities, for each subject we computed Pearson correlation coefficients between lip-corner displacement and head and eye motion. We found moderate coherence between all pairs of variables. Table 1 shows mean correlation coefficients between lip-corner displacement and head rotation and eye motion. Mean correlation coefficients for lip-corner displacement and head translation are not shown but were of comparable magnitude (absolute r = .37 to .40). In each case, we found strong evidence for coherence between facial action and head and eye motion (6 *df*). Head pitch, roll, and yaw were moderately correlated with each other as well (correlations not shown in the Table but are available by request). A high level of coordination occurred across all modalities.

Table 1. Correlations between lip-corner displacement, head rotation, and eye motion (iris translation).

	Mean Positive Correlation	Mean Negative Correlation	Percent with Negative Correlation
Head			
Pitch	.36	-.50	73%*
Roll	.39	-.42	48%
Yaw	.47	-.43	53%
Iris Translation			
Horizontal	.37	-.42	55%
Vertical	.41	-.35	65%*

Note. All correlations, $p < .05$.
*Percentage of smiles with negative correlation for head pitch, $p = .01$; percentage with negative correlation for vertical iris translation, $p = .11$.

The standard deviations for the means presented in Tables 1 and 2 fell with a narrow range of between 0.21 and 0.32. Variation in correlation coefficients may relate to individual differences in style of expression or to differences in the meaning of expression, which is a focus of ongoing research.

We found two consistent patterns between the direction of head and eye motion and smile intensity (i.e., lip-corner displacement). One, increase in smile intensity was about 3 times as likely to occur when the head was rotating downward (i.e., pitch decreasing) than when the head was rotating upward (i.e., pitch increasing) (z = 2.51, p = .01), as in the example in Figure 4. As participants began to smile, their head pitched downward as smile intensity increased and then pitched upward (i.e., pitch increasing) as smile intensity decreased. Vertical movement of the eyes (i.e., vertical translation of the iris) showed a trend toward the opposite pattern (z = 1.62, p = .11). Subjects were twice as likely to move their eyes vertically when smile intensity was increasing. This latter finding is likely related to concurrent changes in head pitch. That is, as the head rotated downwards, participants brought their eyes up in order to maintain eye contact with the experimenter.

5. Conclusions

We developed an automatic face analysis system that quantifies the multimodal coordination of facial action and head and eye motion during spontaneous smiles. In the spontaneous facial behavior observed here, head rotation was common and ranged from small to moderate. This observation suggests that efforts to automatically recognize facial actions that assume planar head motion are unlikely to generalize to real-world applications in which moderate out-of-plane head motion is common, such as the smiles studied here.

We found strong support for the hypothesis that facial action is coordinated with head and eye motion. Facial action, as indicated by lip-corner displacement during spontaneous smiles, was moderately correlated with all 6 *df* of head motion and with eye motion. These findings are consistent with qualitative data from observer-based research suggesting that facial action is embedded within coordinated motor structures.

The patterns of correlation we found may be specific to contexts involving embarrassment. Participants typically smiled after performing directed facial action tasks. Keltner [15] used such tasks to elicit embarrassment, and our findings were consistent with his. We too found that smiles occurred after directed facial action tasks and that they occurred as part of a coordinated motor routine involving head rotation directed away from the camera and experimenter. Keltner [15] also reported “suppressor movements” around the mouth as subjects attempted to control or minimize their smiles. While we have not yet analyzed suppressor movements (e.g., AU 23/24, AU 17), they appeared common. Indeed, in the example shown above (Figure 3), suppressor movements can be seen in frame 198. We cannot say with certainty that the smiles we observed were related to feelings of embarrassment and relief at the task’s completion, as we did not collect self-report measures.

We would anticipate that smiles associated with the experience of joy and especially surprise would show an opposite pattern; that is, head pitch and smile intensity increasing and decreasing together. Further work is needed to determine the relation between specific patterns of cross-modal covariation in relation to self-reported emotion, communicative intention, and context.

Previous efforts in automatic facial expression recognition have focused on image sequences in which head motion is small or parallel to the image plane of the camera. We found that in spontaneous facial behavior moderate head motion is common. While head motion presents a challenge for facial expression recognition, our findings suggest that head motion may be of potential benefit as well. To the extent that there are systematic relationships between particular facial actions and head and eye motion, providing classifiers with this additional information may result in more robust facial expression recognition. It will be important in this regard to include attention to context, as we propose above, because specific patterns of cross-modal covariation are likely to vary with context. A negative correlation between head pitch and lip-corner displacement may be typical of smiles of embarrassment while a positive correlation may be typical of smiles of enjoyment.

An important goal of automatic facial expression recognition is to enable accurate inferences about emotion and communicative intention. Our findings point to the hypothesis that the communicative meaning of morphologically similar facial actions may be disambiguated by attending to specific patterns of coordination with head rotation and eye motion.

In computer graphics, there has been great 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 coordination among different modalities. Head motion typically is absent or appears to occur randomly in avatars. Until now, normative data about cross-modal covariation have been lacking. Our study is a first attempt to address that problem by providing quantitative normative data about the relation between smiling and head and eye motion. Further work is needed that samples other interactive contexts and includes additional facial features.

In summary, we developed an automatic facial expression analysis system that quantifies the multimodal timing of facial action, head rotation, and eye motion. Using this system, we found strong support for the hypothesis that these modalities are highly coordinated during the course of spontaneous smiles. For the smiles studied here, we found a pattern of correlation consistent with smiles of embarrassment. We propose that specific patterns of cross-modal covariation may be specific to particular interpersonal contexts. By including covariation between multimodal features in classifiers, more robust automatic facial expression recognition might be obtained and more accurate inferences about communicative meaning realized. Normative data on cross-modal coordination can contribute to the development of more realistic computer avatars.

6. References

[1] Bartlett, M.S. et al. (1999). Measuring facial expressions by computer image analysis. *Psychophysiology, 36*, 253-263.

[2] Basu, S., Essa, I., & Pentland, A. (1996). Motion regularization for model-based head tracking. *International Conference on Pattern Recognition, ICPR96.*

[3] Blurton Jones, N., & Konner, M. (1971). An experiment on eyebrow raising and visual searching in children. *Journal of Child Psychology & Psychiatry, 11,* 233-240.

[4] Bregler, C. & Malik, J. (1998). Tracking People with Twists and Exponential Maps*, IEEE Conference on Computer Vision and Pattern Recognition, CVPR98*, pp. 8-15.

[5] Camras, L.A., Lambrecht, L., & Michel, G.F. (1996). Infant "surprise" expressions as coordinative motor structures. *Journal of Nonverbal Behavior, 20*, 183-195.

[6] Cohn, J. F. & Schmidt, K. S. (2003). The timing of facial motion in posed and spontaneous smiles. *Proceedings of the 2nd International Conference on Active Media Technology (ICMAT 2003),* Chongqing, China. *2,* 57-72.

[7] Cohn, J. F., Xiao, J., Moriyama, T., Ambadar, Z., & Kanade, T. (2003). Automatic recognition of eye blinking in spontaneously occurring behavior. *Behavior Research Methods, Instruments, and Computers, 35,* 420-428.

[8] Ekman, P. (1992). *Telling lies: Clues to deceit in the marketplace, politics, & marriages.* New York, NY: Norton.

[9] Ekman, P. Friesen, W., & Hager, J. (2002). *Facial Action Coding System*. Salt Lake City, Utah: Research Nexus.

[10] Fox, N., Gross, R., de Chazal, P., Cohn, J.F., & Reilly, R. (2003). Person identification using multi-modal features: Speech, lip, and face. *ACM Multimedia Workshop in Biometrics Methods and Applications (WBMA 2003), Berkeley, California, xxx-xxx.*

[11] Frank, M., Ekman, P., & Friesen, W. (1993). Behavioral markers and recognizability of the smile of enjoyment. *Journal of Personality & Social Psychology, 64*, 83-93.

[12] Hill, H. & Johnson, A. (2001). Categorizing sex and identity from the biological motion of faces. *Current Biology, 11,* 880-885.

[13] Irani (1999). Multi-frame optical flow estimation using subspace constraints. *IEEE International Conference on Computer Vision, ICCV99.*

[14] Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. *IEEE International Conference on Automatic Face and Gesture Recognition, FG'00*, *Grenoble, France,* 46-53.

[15] Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality & Social Psychology, 68,* 441-454.

[16] King, W.M., Lisberger, S.G., & Fuchs, A.F. (1976). 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, 39,* 1135-49.

[17] Klier, E.M., Hongying, W., & Crawford, J.D. (2003). Three-dimensional eye-head coordination is implemented downstream from the superior colliculus. *Journal of Neurophysiology, 89,* 2839-2853.

[18] Kraut, R.E. & Johnson, R. (1979). Social and emotional messages of smiling: An ethological approach. *Journal of Personality and Social Psychology, 37,* 1539-1553.

[19] Lucas, B. & Kanade, T. (1981). An interactive image registration technique with an application in stereo vision. *International Joint Conference on Artificial Intelligence, pp. 674-679.*

[20] Michel, G., Camras, L., & Sullivan, J. (1992). Infant interest expressions as coordinative motor structures. *Infant Behavior and Development, 15,* 347-358.

[21] Pantic, M. (August 2001). *Facial expression analysis by computational intelligence techniques*. Doctoral dissertation, Delft University of Technology.

[22] Pantic, M. & Rothkrantz, M. (2000). Expert system for automatic analysis of facial expression. *Image and Vision Computing, 18,* 881-905.

[23] Schmidt, K., Ambadar, Z., & Cohn, J.F. (In preparation). Human perception of genuineness of smiles in relation to their temporal organization. Affect Analysis Group, University of Pittsburgh.

[24] Schmidt, K., Cohn, J. F., & Tian, Y. L. (2003). Signal characteristics of spontaneous facial expressions: Automatic movement in solitary and social smiles. *Biological Psychology, 65, 49-66.*

[25] Tian, Y.L., Kanade, T., & Cohn, J.F. (2001). Recognizing action units for facial expression analysis. *IEEE PAMI, 23,* 97-116.

[26] Xiao, J., Moriyama, T., Kanade, T., & Cohn, J.F. (2003). Robust full-motion recovery of head by dynamic templates and re-registration techniques. *International Journal of Imaging Systems and Technology, 13*, 85-94.

[27] Yacoob, Y. & Davis, L. (1994). Computing spatio-temporal representations of human faces. *IEEE Conference on Computer Vision & Pattern Recognition, CVPR94*, 70-75.

[28] Zhou, Y., Gu, L., & Zhang, H.J. (2003). Bayesian tangent shape model: Estimating shape and pose parameters via Bayesian inference. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR2003*, pp. xxx-xxx.