Appeared in the 2002 IEEE International Conference on Multimodal Interfaces (ICMI 2002)

Individual Differences in Facial Expression: Stability over Time, Relation to Self-Reported Emotion, and Ability to Inform Person Identification

Jeffrey F. Cohn[1,2], Karen Schmidt[1], Ralph Gross[2], and Paul Ekman[3]

[1]*Department of Psychology, University of Pittsburgh*
[2]*Robotics Institute, Carnegie Mellon University, Pittsburgh*
[3]*Department of Psychiatry, University of California at San Francisco*

jeffcohn@pitt.edu, kschmidt@pitt.edu, rgross@cs.cmu.edu, ekman@compuserve.com

Abstract

The face can communicate varied personal information including subjective emotion, communicative intent, and cognitive appraisal. Accurate interpretation by observer or computer interface depends on attention to dynamic properties of the expression, context, and knowledge of what is normative for a given individual. In two separate studies, we investigated individual differences in the base rate of positive facial expression and in specific facial action units over intervals from 4- to 12 months. Facial expression was measured using convergent measures, including facial EMG, automatic feature-point tracking, and manual FACS coding. Individual differences in facial expression were stable over time, comparable in magnitude to stability of self-reported emotion, and sufficiently strong that individuals were recognized on the basis of their facial behavior alone at rates comparable to that for a commercial face recognition system (FaceIt from Identix). Facial action units convey unique information about person identity that can inform interpretation of psychological states, person recognition, and design of individuated avatars.

Keywords: Facial Expression, Individual Differences, Face Recognition

1 Introduction

Within the past decade, significant effort has occurred in automatic recognition of facial expression [8,20,21]. Motivating this effort is the goal of inferring emotion and communicative intent from an individual's facial expression. Because the face may express emotion [5], action tendencies and intentions [9,10], and individual differences in reactivity [13], accurate inference requires many sources of information such as context and dynamic properties of facial expression [18], and individual differences among people [19]. In some cultures, for instance, explicit expressions of anger are discouraged; those instances of anger that do occur may carry far more consequence than in those for which anger displays are less regulated. Differences in rates of anger expressions in response to an age-appropriate stressor, for instance, have been found between Chinese, Japanese, and Euro-American children as young as 10 months of age [3]. In general, accurate interpretation of facial expression benefits from knowledge of what is normative for a specific individual [6, 16]. The goal of the research reported below was to assess individual differences in rates of positive expression and in specific configurations of facial action units. We show that individual differences in facial expression are moderately stable over time and are comparable in stability to what has been found for self-reported emotion. These individual differences in facial expression are sufficiently strong that individuals may be accurately recognized on the basis of their facial behavior, which suggests that facial expression may be a valuable biometric.

We report results from two studies. In the first, rates of positive affect in response to film clips on two occasions at a 12-month interval were observed in 65 young adults. They were observed alone while watching short films intended to elicit emotion. Positive affect was quantified using convergent measures: facial electromyographic (EMG) recording from the *zygomatic major* muscle and feature-point tracking using optical flow. The *zygomatic major* is the principal muscle involved in smiling, which is the most common of facial expressive movements [19] and is a reliable index of emotion valence and intensity [2].

In study 2, 85 middle-aged to older adults were observed at a 4-month interval in a clinical interview. Facial expression was manually coded using the Facial Action Coding System [7]. We report Pearson correlation coefficients for specific facial actions and using pattern recognition demonstrate the utility of facial behavior signatures for person recognition.

The two studies afford evaluation of individual differences in multiple contexts (alone versus 2-person interview), using multiple measures (facial EMG, manual FACS coding, and lip-corner tracking in Study 1, manual FACS coding alone in Study 2), time delays of 4- to 12 months, comparison of facial behavior with self-reported emotion (Study 1), and biometric analysis of facial behavior (Study 2) in person recognition.

2 Study 1: Facial expression during solitary viewing of films

2.1 Data collection

Original data were collected from a sample of subjects in a psychophysiological study of emotion [4, 15]. Manual facial action coding (FACS) [7], automated facial analysis using feature tracking (AFA) [20], and facial electromyography (EMG), were used to investigate the dynamic properties of smiling [18]. Video and EMG data were collected on two occasions separated in time by approximately 12 months. At each session, subjects were video-recorded with their consent during baseline (seated with eyes open) and viewing of film clips. The camera was camouflaged behind dark glass. Film clips were selected based on previous literature and our own pilot testing. Facial expression was recorded during the viewing of a comedy routine, and visual data (FACS and AFA) were collected in the period following each of the first three jokes. EMG data were collected for the duration of the session, including a baseline portion and the entire comedy routine (5 minutes).

2.2 Facial expression measurement

2.2.1 Facial Action Coding System.

Smiles (n=195 from 94 subjects) were manually coded using the Facial Action Coding System (FACS) [7], and the presence of other facial actions overlapping with AU 12 (*zygomatic major* ”smile”) were recorded. Agreement between two certified FACS coders for the presence or absence of action units 6, 12, 14, 15, 17, and 23 during smiles was 0.92 (n=27 smiles). Agreement for the order of action unit appearance was 0.86. Appearance of rigid head movements (AUs 51-58) and their relationship to AU 12 were also coded.

2.2.2 Facial electromyography.

Facial electromyographic (EMG) data were collected using standard placement of electrodes over the *zygomatic major* muscle, digitized at 512 Hz, bandpass filtered between 15 and 90 Hz, and down-sampled at 1/30 second intervals to correspond with the length of a video frame. While down-sampling at 1/30 second may introduce aliasing error in the higher frequency portion of the resulting signal, the original EMG signal contains little energy at high frequencies and our interest is in the mean value of the time series (EMG), which is primarily determined by the low frequency component. Accordingly, aliasing error may be assumed to have little or no influence on the scores we analyzed. EMG values reported here are z scores for *zygomatic major* activity obtained during the comedy film clip, as compared with the mean values for that muscle during an eyes open baseline period. Values for a second identical session one year later are also reported. This group included subjects with video data (n=31) and additional subjects from the same study (n=35).

2.2.3 Automated Facial Analysis.

Videotaped smiles were digitized at 30 frames per second. Automated Facial Analysis [20] using feature point tracking was performed on 50 smiles at session 1. The position of lip corners in a video series was recorded for the duration of AU 12 in the smile, and the longest continuous sequence of rising values in r (polar coordinate of lip corner position) was recorded. This sequence of rising values was designated as the smile onset. The right panel of Figure 1 shows an example. The left panel of Figure 1 shows a frame from the corresponding image sequence from [4].

2.3 Relation between *zygomatic major* EMG and polar coordinate of lip-corner motion.

Zygomatic major EMG and polar coordinate of lip corner motion were in agreement for AU 12 in 72% of cases with distinct EMG onset. In the 72% of smiles for which both EMG and visually tracked onset were detectable, onsets were highly correlated ($r = 0.95$, $p < .01$), with visible onset occurring an average of 0.23 seconds after the EMG onset (See Figure 1 for an example). Lip-corner motion indicated a significant effect of gender on smile latency ($p < .01$), with women smiling (visible lip width onset) more quickly following the joke ($mean = 0.72s$) than men ($mean = 1.08s$).

2.4 Stability of *zygomatic major* muscle activity over time

Mean level of *zygomatic major* (z score relative to the mean raw *zygomatic major* activity during baseline) decreased between sessions over the 12-month interval. At visit 1, mean *zygomatic major* EMG was 18.9 (SD 18.9) in comparison with mean of 14.3 (SD 20.1) one year later ($p < .05$). While the mean intensity decreased between visits, the rank ordering of individuals at each visit remained similar. The correlation for zygomatic major EMG between these two sessions was 0.578 (n=65 subjects) (See Table 1). There were no sex differences among subjects for these variables.

To our knowledge, only one other study has shown stability in facial behavior over 12 months or more. That study was of mothers' expression of emotion with their first and second infants when they were each 8 weeks old [16]. That study used manual coding rather than objective measures of facial activity. Mother-infant interaction is a special context in which facial behavior and speech are exaggerated. Ours is the first study to show such stability over a 12-month or longer interval in typical adult behavior.

The stability in smiling that we observed was comparable to that of self-reported emotion over the same 12-month interval. Table 1 shows the correlation within and between sessions between *zygomatic major* EMG and self-reports of positive emotion. Note that while smiling and self-reported emotion are both moderately stable, the correlation between them while still significant is relatively low. This suggests that caution is required in inferring subjective emotion from upward lip corner motion alone.

In Study 2, we investigated stability in facial expression in finer detail. Rather than describing facial expression in terms of global emotion-specified categories (positive or negative emotion), we used the manual Facial Action Coding System (FACS) [7] to describe specific facial action units. Action units (AUs) are the smallest visibly discriminable change in facial movement. Using combinations of action units, all possible facial expressions can be described. Asymmetries in facial movement, such as occur when one but not the other brow is raised, may be described as well. We used FACS-coded interviews to assess stability of fine-grained facial expression at a time delay of 4 months. We tested the hypothesis that stability in facial action units is sufficiently high to afford person recognition from this measure alone.

3 Study 2: Facial expression during two-person interviews

3.1 Data collection

FACS-coded data were obtained from a study by Rosenberg, Ekman, & Blumenthal [17] of facial expression in patients with heart disease. Subjects were 85 men and women with a history of transient myocardial ischemia (TMI) who were interviewed on two occasions at a 4-month interval. They averaged 59 years of age (SD = 8.24) and were predominantly Euro-American. Spontaneous facial expressions were video-recorded during a clinical interview with the Type A Structured Interview (SI Interview: cited in [17]). This interview elicits action units related to disgust, contempt, and other negative facial emotions associated with Type A personality as well as smiles.

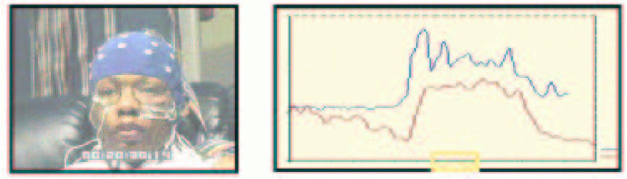

Figure 1. Temporal relation (right panel) between *zygomatic major* EMG (blue line, upper time series) and polar coordinator of lip-corner motion (red line, lower time series). r = 0.95. The left panel shows corresponding example from observational scenario.

Table 1. Concurrent & predictive correlations for self-reported emotion and facial expression over 12 months interval between visit 1 and visit 2.

	Visit 1		Visit 2	
	Self-reported emotion	*Zygomatic* EMG	Self-reported emotion	*Zygomatic* EMG
Visit 1 Self-reported emotion	—	.36*	**.56***	.32*
Visit 1 *Zyg.* EMG		—	.14	**.58***
Visit 2 Self-reported emotion			—	.31*

3.2 Manual FACS coding

Two-minute segments were manually FACS coded. For comparison purposes, one third of the interviews were comparison coded. Percent agreement for action units and action unit combinations was 80%. We focus on action units and action unit combinations common to negative emotion and social smiles that include action unit (AU) 12.

Table 2. Stability coefficients (Pearson r) > 0.40, indicating moderate to strong stability.

Action Units(AUs)	Description	Stability Coefficient
Symmetric AUs		
AU 1	Inner corners of eyebrows raised	.60
AU 4	Eyebrows drawn medially and down	.47
AU 5	Eyes widened	.43
AU 6	Cheeks raised, eyes narrowed	.92
AU 7	Lower eyelids raised and drawn medially	.52
AU 10	Upper lip raised	.42
AU 14	Lip corners tightened	.40
AU 15	Corners of the mouth pulled downward and inward	.45
AU 18	Lips pursed	.50
AU 20	Lip corners pulled laterally	.41
AU 24	Lips pressed together	.48
Asymmetric AU's		
Left AU 1		.77
Aggregate of upper face right AU's		.80
Right AU 10		.45
Left AU 10		.50
AU Combinations		
AU 1+4		.48
AU 1+2+4		.66
AU 12+15		.53
All $p < .05$		

3.3 Stability of FACS action units

With the possible exception of AU 6, the action units were all ones associated with negative emotion, which was expected given the nature of the clinical interview. Two common action unit combinations also are associated with negative emotion (AU 1+2+4 and AU 12+15). They had stability coefficients of 0.50 or higher. Action units may occur asymmetrically, as when inner brow raise (AU 1) occurs on one side of the face but not the other. Strong stability was found for asymmetric AU 1 on the left side and an aggregate of asymmetric upper face action units occurring on the right side. Table 2 shows action units and selected action unit combinations having stability coefficients in the moderate to strong range ($r > .40$).

3.4 Human ID from facial expression

In a recent study, Liu et al. [14] found that facial asymmetry, as quantified by right-left pixel differences across the entire face, contributed unique variance for human identification. We tested the hypothesis that individual differences in use of specific action units could inform person identification. Using nearest neighbor classification, action unit frequencies at the first and second interview were used to recognize individuals at the second interview based on their previous facial behavior.

Figure 2 shows the cumulative response curve for person recognition at interview 2 based on action units. For exact match, the probability is approximately 50%, which is markedly higher than the 1% accuracy that would be expected by chance (1/85 = 1%).

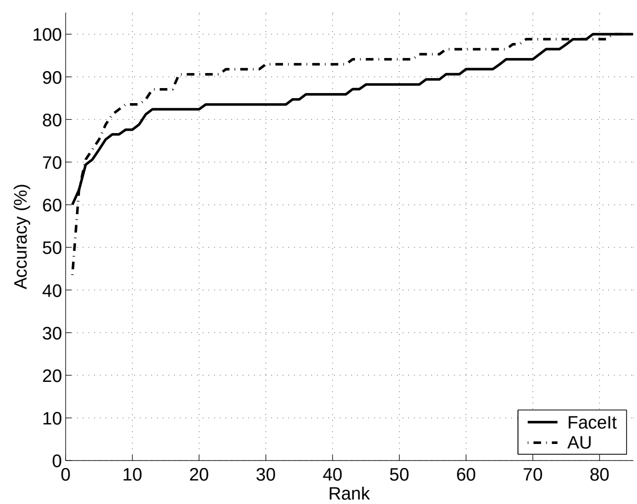

Figure 2. Cumulative response curves for AU-based classifier and for FaceIt.

For comparison, we analyzed the face image data using FaceIt [12]. FaceIt was the top performing face recognition algorithm in the recent Facial Recognition Vendor Test [1].

FaceIt's face detection module was run on every 30th frame in the 2 minutes of image data per subject (i.e., 120 samples at Visit 1 and at Visit 2). The 3 frames with the highest face detection confidence scores were selected for training (Visit 1) and testing (Visit 2). Out of the three testing frames per subject, we selected the one with the highest face recognition confidence score to determine the recognition result for each subject.

AU-based person recognition and FaceIt achieved comparable accuracy, and their errors appeared to be relatively independent. While the AU-based classifier did slightly better on those subjects who were correctly identified by FaceIt

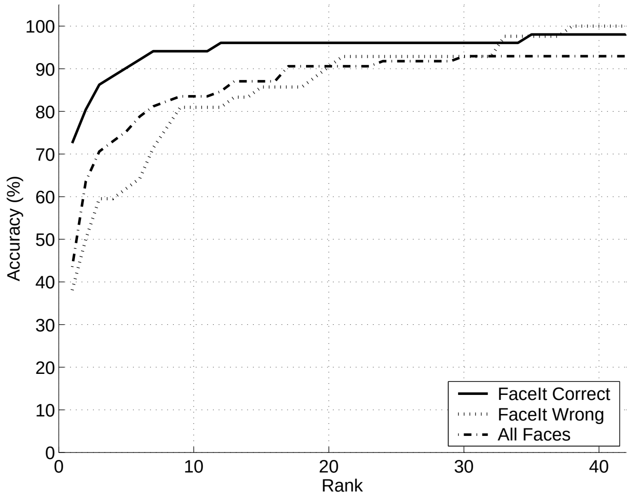

Figure 3. Cumulative response curves for AU-based classifier for subjects correctly and incorrectly recognized by FaceIt and for original cumulative response curve.

(Figure 3), the difference was relatively small and the shape of the cumulative recognition curves was similar. Across all subjects, the AU-based classifier exceeded FaceIt for all but exact match.

4 Conclusions

We found strong evidence for stable individual differences in facial expression. In both solitary (film viewing) and interpersonal (interview) contexts, facial expression was stable between time periods from 4- to 12 months. The stability we observed was comparable to what has been reported previously for self-reported emotion and related personality dimensions (e.g., extraversion and neuroticism). Stability was sufficiently robust that individuals could be recognized at far above chance levels based solely on their use of facial action units (which may be considered facial behavior signatures).

Small but significant correlations were found between facial expression and self-reported emotion. It is likely that higher correlation would have been found had we used FACS action units for describing facial expression in relation to self-reported emotion instead of less differentiated measurement as represented by *zygomatic major* EMG. Previous literature, for instance, suggests that self-reports of emotion are more affectively positive when cheek tightening and crows feet wrinkles (AU 6, also known as ”Duchenne marker”) occur in the presence of oblique lip-corner motion (AU 12). Detailed description of facial expression is needed to provide strong basis for inferring emotion.

Consistent with prior literature, we found evidence that women were more expressive than men. While there were no differences in overall levels of *zygomatic major* EMG in Study 1, feature point tracking revealed that women were faster than men in responding to the comedy routines. Individual differences in response latencies have been reported previously in clinical depression [22]; with few exceptions [6], individual differences in latency are relatively unexplored in interpretation of facial behavior. These differences suggest that it is important in design of multi-modal interfaces to quantify these factors and use them in interpreting departures from normal set points. In addition, by including unique variance in timing and base rate of specific action units in animating computer avatars, personal identify may be embodied.

Identification of individuals via FaceIt was lower than the 80% accuracy reported by Gross et al. [11] for time delay face image data. Several factors may have contributed to this difference. Because the image data were not collected with the intention of face recognition in mind, pose, and illumination were not standardized between interviews. Gross et al. [11] found that face recognition lacks robustness to variations in pose in excess of 30- to 40 degrees. Some subjects wore glasses at one but not the other interview, and the face outline often was partially occluded. Because we selected for analysis those images for which FaceIt had high confidence in face detection, however, this latter factor was likely mitigated. We anticipate that under more controlled conditions face recognition would have been higher. At the same time, the conditions extant here are ones common in field settings, such as variation in pose, illumination, and appearance over time. In field settings, too, face size relative to the image and image clarity are likely to be more challenging than the relatively full-face video we analyzed. Thus, the face recognition results for FaceIt may not have been atypical of what would be found outside of controlled laboratory conditions. Importantly, the present findings suggest that individual differences in facial expression are an informative biometric that appear to be relatively independent of features used by commercial face recognition algorithms. By combining facial action units with face recognition algorithms, more robust person identification in natural settings may be achieved.

Critical challenges in the development of multi-modal interfaces are automatic recognition and interpretation of facial and other non-verbal behavior. The emphasis to date has been on detecting emotion-specified expressions (e.g., joy or anger) without regard to how characteristic they may be of a user within a particular context. By knowing what is normative for a specific person, both in terms of specific action units and their temporal organization, we can improve the accuracy with which underlying psychological states and intentions are inferred. Ekman [6], for instance, has shown that in detecting deception it is essential to evaluate nonverbal behavior against an individual's own base rate. Similarly, unique personalities may be conveyed more effectively in avatars by endowing them with unique behavioral signatures.

In summary, individual differences in facial expression were stable over time and mode of measurement, which included facial EMG and feature point-tracking in Study 1 and manual FACS coding in Study 2. Individual differences in facial expression were moderately strong in both individual and interpersonal contexts and comparable in stability to what has been reported previously for self-reported emotion. Indeed, individuals could be recognized on the basis of their facial behavior alone. Facial expression conveyed unique information about person identity that was relatively independent of that obtained by commercial face recognition algorithms, as represented by FaceIt. Individual differences in facial expression can inform interpretation of psychological states and design of individuated avatars.

References

[1] D. M. Blackburn, M. Bone, and J. Phillips. Facial recognition vendor test 2000: Evaluation report.

[2] J. T. Cacioppo, R. E. Petty, M. E. Losch, and H. S. Kim. Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, pages 260–268, 1986.

[3] L. A. Camras, H. Oster, J. Campos, R. Campos, T. Ujiee, K. Miyake, L. Wang, and Z. Meng. Production of emotional facial expressions in American, Japanese and Chinese infants. *Developmental Psychology*, 34:616–628, 1998.

[4] J. F. Cohn, N. A. Fox, and M. Kovacs. [EEG, EMG, and ECG during baseline and emotion challenge conditions]. Unpublished raw data.

[5] P. Ekman. Facial expression and emotion. *American Psychologist*, 48:384–392, 1993.

[6] P. Ekman. *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*. W.W. Norton, New York, 2001.

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

[8] I. Essa and A. Pentland. Coding, analysis, interpretation and recognition of facial expressions. *IEEE PAMI*, 7:757–763, 1997.

[9] N. Fridja and A. Tcherkassof. Facial expressions as modes of action readiness. In J. A. Russel and J. M. Fernandez-Dols, editors, *The psychology of facial expression*. Cambridge University Press, Cambridge, 1997.

[10] A. J. Fridlund. *Human facial expression: An evolutionary view*. Academic, NY, NY, 1994.

[11] R. Gross, J. Shi, and J. F. Cohn. Quo vadis face recognition? In *Third Workshop on Empirical Evaluation Methods in Computer Vision*, pages 119–132, Kauai, Hawaii, 2001.

[12] Identix. FaceIt version 2.0.0.16, from Identix (formerly Visionics), http://www.identix.com.

[13] R. W. Levenson. Expressive, physiological, and subjective changes in emotiojn across adulthood. In S. H. Qualls and N. Abeles, editors, *Psychology and the aging revolution: How we adapt to longer life*, pages 123–140. American Psychological Association, Washington, D.C., 2000.

[14] Y. Liu, K. Schmidt, J. F. Cohn, and R. L. Weaver. Human facial asymmetry for expression-invariant facial identification. In *Proc. of the Fifth IEEE Internationcal Conference on Automatic Face and Gesture Recognition*, pages 208–214, Washington, D.C., 2002.

[15] A. Miller, N. A. Fox, J. F. Cohn, E. E. Forbes, J. T. Sherrill, and M. Kovacs. Regional brain activity in adults with a history of childhood-onset depression: Sex differences and clinical variability. *American Journal of Psychiatry*, 159:934–940, 2002.

[16] G. A. Moore, J. F. Cohn, and S. B. Campbell. Mothers' affective behavior with infant siblings: Stability and change. *Developmental Psychology*, 33:856–860, 1997.

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

[18] K. Schmidt, J. F. Cohn, and Y. L. Tian. Signal characteristics of spontaneous facial expressions: Automatic movement in solitary and social smiles. submitted.

[19] K. L. Schmidt and J. F. Cohn. Human facial expressions as adaptations: Evolutionary questions in facial expression. *Yearbook of Pysical Anthropology*, 44:3–24, 2001.

[20] Y. Tian, T. Kanade, and J. F. Cohn. Recognizing action units for facial expression analysis. *IEEE PAMI*, 23:97–115, 2001.

[21] Y. Yacoob and L. Davis. Recognizing human facial expression from long image sequence using optical flow. *IEEE PAMI*, 18:636–642, 1997.

[22] A. Zlochower and J. F. Cohn. Vocal timing in face-to-face interactions of clinically depressed and nondepressed mothers and their 4-month-old infants. *Infant Behavior and Development*, 19:373–376, 1996.