Foundations of Human Computing: Facial Expression and Emotion[1]

Jeffrey F. Cohn

Department of Psychology, University of Pittsburgh, 3137 SQ, 210 S. Bouquet Street, Pittsburgh, PA 15260 USA
jeffcohn@cs.cmu.edu
www.pitt.edu/~jeffcohn

Abstract. Many people believe that emotions and subjective feelings are one and the same and that a goal of human-centered computing is emotion recognition. The first belief is outdated; the second mistaken. For human-centered computing to succeed, a different way of thinking is needed. Emotions are species-typical patterns that evolved because of their value in addressing fundamental life tasks. Emotions consist of multiple components, of which subjective feelings may be one. They are not directly observable, but inferred from expressive behavior, self-report, physiological indicators, and context. I focus on expressive facial behavior because of its coherence with other indicators and research. Among the topics included are measurement, timing, individual differences, dyadic interaction, and inference. I propose that design and implementation of perceptual user interfaces may be better informed by considering the complexity of emotion, its various indicators, measurement, individual differences, dyadic interaction, and problems of inference.

Key words: Emotion, measurement, facial expression, automatic facial image analysis, human-computer interaction, temporal dynamics

1. INTRODUCTION

How can computers recognize human emotions? Is this even the correct question? By emotion, people often think of subjective feelings, but emotions are more than that and subjective feeling is in no sense essential. There is no *sin qua non* for emotion. Emotions are species-typical patterns consisting of multiple components that may include intentions, action tendencies, appraisals, other cognitions, neuromuscular and physiological changes, expressive behavior, and subjective feelings. None of these alone is necessary or sufficient for any given situation. In human-human interaction, intentions and action tendencies often are more important than what an individual may be feeling. People may or may not be aware of what they're feeling, and feelings often come about some time late in the temporal unfolding of an emotion.

A goal of human-centered computing is computer systems that can unobtrusively perceive and understand human behavior in unstructured environments and respond appropriately. Much work has strived to recognize human emotions. This effort is informed by the importance of emotion to people's goals, strivings, adaptation, and quality of life [1, 2] at multiple levels of organization, from intra-personal to societal [3]. Efforts at emotion recognition, however, are inherently flawed unless one recognizes that emotion – intentions, action tendencies, appraisals and other cognitions, physiological and neuromuscular changes, and feelings – is not readily observable. Emotion can only be

[1] A previous version of this paper was originally published in the *Proceedings of the ACM International Conference on Multimodal Interfaces,* Banff, Canada, 2006 (Copyright © ACM Press).

inferred from context, self-report, physiological indicators, and expressive behavior (see Figure 1). The focus of the current paper is on expressive behavior, in particular facial expression and approaches to its measurement, feature selection, individual differences, interpersonal regulation, and inference.

Facial expression has been a subject of keen study in behavioral science for more than a hundred years[4, 5], and within the past 10 years considerable progress has been made in automatic analysis of facial expression from digital video input [6-8].

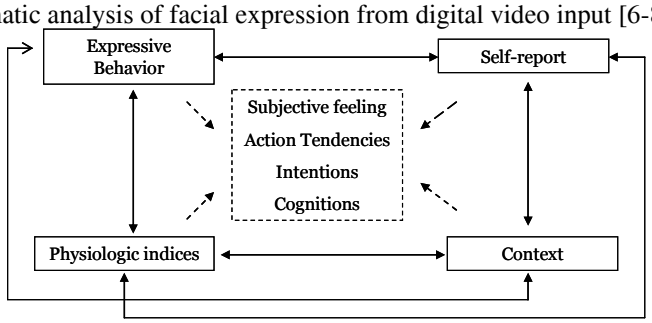

Fig. 1. Components and indicators of emotion. Solid boxes represent observables, dashed boxes latent variables. Solid arrows indicate observable correlations among indicators. Large correlations among multiple indicators indicate greater coherence among indicators. Dashed arrows represent inferential paths. Paths between emotion components are omitted. (©2006 ACM)

Facial expression correlates moderately with self-reported emotion [5] and emotion-related central and peripheral physiology [9, 10]. Facial expression has similar underlying dimensions (e.g., positive and negative affect) with self-reported emotion [11] and serves interpersonal functions by conveying communicative intent, signaling affective information in social referencing, and contributing to the regulation of social interaction [12, 13]. Expressive changes in the face are a rich source of cues about intra- and interpersonal indicators and functions of emotion [3, 14]. As a measure of trait affect and socialization, stability in facial expression emerges early in life [15]. By adulthood, stability is moderately strong, comparable to that for self-reported emotion [16].

Early work in automatic analysis and recognition of facial actions from input video focused on the relatively tractable problem of posed facial actions acquired under well-controlled conditions (e.g., frontal full-face view with minimal head motion and uniform lighting). Recent work has progressed to analysis and recognition of spontaneous facial actions with non-frontal views, small to moderate out-of-plane head motion, subtle facial actions, and variation in illumination [17-19]. Moreover, methods of analysis and synthesis of facial expression are beginning to merge. It is becoming possible to animate an avatar from shape and appearance measures of human facial actions in real time [20], which is likely to significantly impact human-centered computing. By separating identity from facial behavior, for instance, user confidentiality could be better protected.

Here, I present key issues to consider in designing interfaces that approach the naturalness of face-to-face interaction. These include approaches to measurement, types of features, individual differences, interpersonal regulation, and inference.

2. APPROACHES TO MEASUREMENT

Two major approaches are sign- and message judgment [21]. In message judgment, the observer's task is to make *inferences* about something underlying the facial behavior, such as emotion or personality. In measuring sign vehicles, the task is to *describe* the surface of behavior, such as when the face moves a certain way. As an example, upon seeing a smiling face, an observer with a judgment-based approach would make judgments such as "happy," whereas an observer with a sign-based approach would code the face as having an upward, oblique movement of the lip corners. Message judgment implicitly assumes that the face is an emotion "read out." Sign-based measurement is agnostic and leaves inference to higher-order decision making.

2.1 Message Judgment

Message judgment approaches define facial expressions in terms of inferred emotion. Of the various descriptors, those of Ekman have been especially influential. Ekman [22] proposed six "basic emotions." They are joy, surprise, sadness, disgust, fear, and anger. Each was hypothesized to have universally displayed and recognized signals, universal elicitors, specific patterns of physiology, rapid, unbidden onset, and brief duration, among other attributes. Of the universal signals, prototypic expressions were described for each emotion (Figure 2). Most research in automatic recognition of facial expression [26, 27] and much emotion research in psychology [28] has concentrated on one or more of these six emotions. This list, however, was never intended as exhaustive of human emotion. It is not. Rather, it was proposed in terms of conformity with specific criteria noted, such as having a universal display (i.e., prototypic expression). Other emotions that may be inferred from facial expression include embarrassment and contempt among others. Indicators of cognitive states, such as interest and confusion, have been described as well [29].

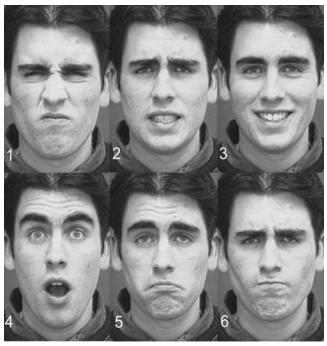

Fig. 2. Emotion-specified expressions: disgust, fear, joy, surprise, sadness, and anger. From [13]. Individual images are from the Cohn-Kanade FACS-Coded Image Database [25]. (© Jeffrey Cohn).

While much literature has emphasized expressions of one or another emotion, expressions may include blends or combinations of two or more [30]. For purposes such as detecting deception, expressions that include traces of contradictory emotions are of particular interest. Masking smiles[31], in which smiling is used to cover up or hide an underlying emotion are the best known. Figure 3 shows examples of two masking smiles (middle and on the right) and a "felt" or Duchenne smile (on the left) for comparison. Duchenne smiles are believed to express felt positive emotion. In the two masking smiles there is indication of sadness in one, suggested by the downward pull of the lip corners and slight pouching of the skin below them (due to AU 15 in FACS; see next section), and disgust in the other, suggested by the upward pull of the lip medial to the lip corner and philtrum and the altered shape of the nasolabial furrow (shape and appearance changes due to AU 10 in FACS) These examples come from a study in which nurses watched a movie clip intended to elicit joy and other movie clips intended to elicit disgust, revulsion, or sadness [31]. The nurses were instructed to hide their negative emotions so that an observer would be unable to tell which movie they were watching.

Fig. 3. From left to right, example of a "felt" smile, a smile "masking" disgust and one "masking" sadness. Corresponding FACS AUs are AU 6+12+26, AU 10+12+26, and AU 12+15, respectively. Reprinted with permission from [31] (© Paul Ekman).

2.2 Sign Measurement

Cohn & Ekman [32] review manual methods for labeling facial actions. Of the various methods, the Facial Action Coding System (FACS) [24, 33] is the most comprehensive, psychometrically rigorous, and widely used [32, 34]. Using FACS and viewing video-recorded facial behavior at frame rate and slow motion, coders can manually code nearly all possible facial expressions, which are decomposed into action units (AUs). Action units, with some qualifications, are the smallest visually discriminable facial movements. By comparison, other systems for measuring facial actions are less thorough and fail to differentiate between some anatomically distinct movements or consider as separable movements that are not anatomically distinct [35].

The most recent version of FACS specifies 9 action units in the upper face, 18 in the lower face, 11 for head position and movement, nine for eye position and movement, and additional descriptors for miscellaneous actions, gross body movement, and

supplementary codes. (For a complete list together with anatomic basis for each AU, see [21, 36]).

Action units may occur singly or in combinations, and combinations may be additive or non-additive. In additive combinations, the appearance of each action unit is independent; whereas in non-additive combinations they modify each other's appearance. Non-additive combinations are analogous to co-articulation effects in speech, in which one phoneme modifies the sound of ones with which it is contiguous. An example of an additive combination in FACS is AU 1+2, which often occurs in surprise (along with eye widening, AU 5) and in the brow-flash greeting [37]. The combination of these two action units raises the inner (AU 1) and outer (AU 2) corners of the eyebrows and causes horizontal wrinkles to appear across the forehead. The appearance changes associated with AU 1+2 are the product of their joint actions.

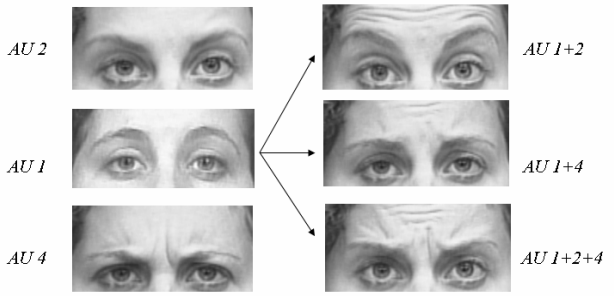

Fig. 4. Examples of individual action units and action unit combinations. AU 1+2 is an additive combination. AU 1+4 and AU 1+2+4 are non-additive, comparable to co-articulation effects in speech. (© Jeffrey Cohn).

An example of a non-additive combination is AU 1+4, which often occurs in sadness [4] (see Figure 4). When AU 1 occurs alone, the inner eyebrows are pulled upward. When AU 4 occurs alone, they are pulled together and downward (Figure 4). When AU 1 and AU 4 occur together, the downward action of AU 4 is modified. In AU 1+4, the inner eyebrows are raised and pulled together. This action typically gives an oblique shape to the brows and causes horizontal wrinkles to appear in the center of the forehead, as well as other changes in appearance that are characteristic of sadness. Automatic recognition of non-additive combinations such as this presents similar complexity to that of co-articulation effects in speech. Failure to account for non-additive combination in automatic recognition exploits the correlation among AUs and can lead to inflated estimates of algorithm performance. For further reading on FACS, see [21, 32].

2.3 Automatic measurement

Most work in automatic analysis and recognition of facial actions has followed the message-judgment approach, with the goal of recognizing the six prototypic expressions

proposed by Ekman. Relatively few investigators have pursued the sign-based approach in which specific facial action units are recognized [7, 27]. Of the two approaches, message judgment is better suited to recognition of prototypic expressions provided the number of classes is relatively small and between-class variation relatively large. These conditions typically obtain for prototypic expressions because they are few in number and vary from each other in multiple regions of the face. Using a sign-based approach to learn AU combinations, one would have to learn first the individual muscle actions and then learn the mapping of clusters of muscle actions onto prototypic expressions. While prototypic expressions can be learned either way, the former is more efficient.

In comparison with sign-based approaches to automatic facial image analysis, the generalizability of message based approaches is more limited. Prototypic facial actions, with the exception of joy, are relatively rare in naturally occurring behavior, and when they occur, their intensity is often reduced relative to when posed. Low-intensity, or subtle, expressions are more difficult to detect for both machine learning and human observers [38]. Another issue is that emotion is more often indicated by a smaller number of facial actions than is assumed by the message judgment approach. Anger, for instance, may be communicated by slight tightening of the lower eyelids, which a message-judgment approach would be ill-prepared to detect. Masking expressions, as well, are not readily learned using message-judgment. While their component facial actions may occur with adequate frequency for learning, the specific combinations are too varied and infrequent for many machine learning approaches. With a sign-based approach, when training and testing samples for masking or other complex expressions are small, rule-based classifiers informed by human experts may be used for expression recognition. For these and other reasons, the sign-based measurement approach may prove more productive for human-centered computing.

2.4 Reliability of meta-data

The reliability of manually labeled images (i.e., behavioral descriptors or meta-data) is a critical concern for machine learning algorithms. If ground truth is contaminated by 20-30% error, which is not uncommon, that is a significant drag on algorithm performance. For both message judgment and sign-based approaches, similar concerns arise. Using AUs as an example, at least four types of reliability (i.e., agreement between observers) are relevant to the interpretation of substantive findings. These are reliability for occurrence/non-occurrence of individual AUs, temporal precision, intensity, and aggregates. Most research in automatic facial expression analysis has focused on occurrence/non-occurrence [7, 8].

Temporal precision refers to how closely observers agree on the timing of action units, such as when they begin or end. This level of reliability becomes important when examining features such as response latency and turn taking (see Section 5). Action unit intensity becomes important for questions such as whether facial expression is influenced

by audience effects [39]. Several groups have found, for instance, that people tend to smile more intensely in social contexts than when they are alone [39, 40].

Aggregates refer to combinations of action units, which as noted may be additive or non-additive. By assessing the reliability of aggregates directly, one can more accurately estimate both their reliability and the reliability of component action units that occur in isolation.

For each of these types of reliability, a number of metrics appear in the literature. Percentage of times that two observers agree (i.e., percent agreement) is the most-often used but least informative because it fails to correct for agreement by chance. That is, simply by knowing base rates or priors, observers are more likely to agree on behaviors that occur frequently. Chance-corrected statistics that control for base rates and guessing are more informative. Examples are Cohen's kappa (for categorical variables) and intra-class correlation (for continuous variables) [41]. In addition to average agreement or reliability across behavioral descriptors, it is informative to know the results for specific descriptors. Some behaviors are relatively easy to identify, others not. Because the reliability of manual measurement limits between-system agreement, more attention to reliability of behavioral descriptors would contribute to the success of machine learning work in automatic facial expression analysis and recognition.

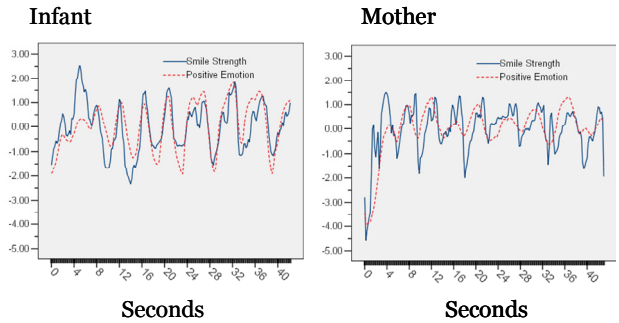

Fig. 5. Time series for AFA-measured lip-corner displacement and human-observer based ratings of positive affect in a mother-infant dyad. Data series for human observers are shifted by about ½ second to adjust for human reaction time. (©2006 ACM)

2.5 Concurrent validity for continuous measurement

Most efforts at automatic expression recognition have compared inter-method agreement (i.e., concurrent validity) for categorical descriptors. Facial actions and emotion expression can vary not just in type, however, but also in amplitude or intensity. For this reason, it is important to evaluate whether alternative measurement systems have concurrent validity for continuous measures of intensity. Our research group recently

examined inter-system precision for intensity contours by comparing CMU/Pitt Automatic Facial Image Analysis (AFA v.4) with continuous ratings of affective intensity and FACS intensity scoring of AU 12 by human observers. Lip-corner displacement in spontaneous smiles was measured by AFA at30 frames/second. We found high concurrent validity between the two methods (see Figure 5 for an example) [42, 43]. In other work, we found similarly high concurrent validity for intensity contours between AFA and facial EMG [40].

3. DYNAMICS

Both the configuration of facial features and the timing of facial actions are important in emotion expression and recognition. The configuration of facial actions (whether emotion-specified expressions or individual action units) in relation to emotion, communicative intent, and action tendencies has been the principal focus [4, 44, 45]. Less is known about the timing of facial actions, in part because manual measurement of timing is coarse and labor intensive [46]. Advances in computer-vision and graphics have made possible increased attention to the dynamics of facial expression [36, 47].

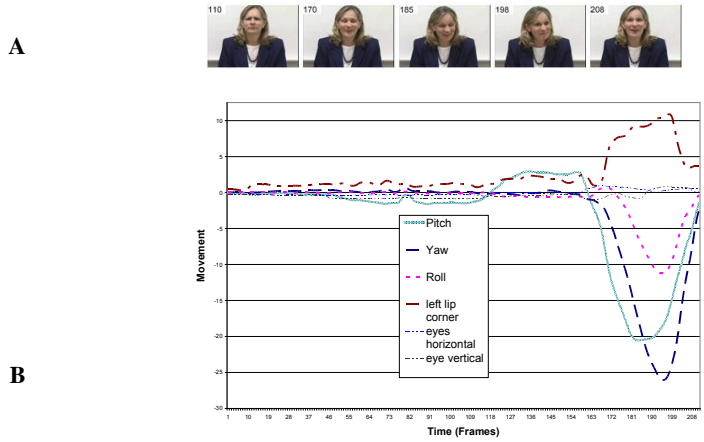

Fig. 6. Multimodal coordination of head motion, lip-corner displacement, and gaze in smiles of embarrassment. A: Selected frames from image sequence depicting embarrassment. B: Corresponding time series. Reprinted with permission from [50]. (©2004 IEEE)

There is increasing evidence that people are responsive to the timing of facial actions [48]. Consider some examples. One, facial actions having the same configuration but different timing are interpreted differently. Whether a smile is perceived as genuine or false depends in part on how quickly it changes from neutral expression to peak. Within limits, smiles with slower onset times are perceived as more genuine [49]. Two, velocity and duration of the onset phase of spontaneous smiles are highly correlated; for deliberate

smiles, these components are uncorrelated. Indeed, spontaneous and deliberate smiles can be discriminated with nearly 90% accuracy from their dynamics alone [40]. Three, multimodal coordination of facial expression, direction of gaze, and head motion is a defining feature of embarrassment displays. Smile intensity increases and direction of gaze is lowered as the head pitches down and away from the partner; smile intensity decreases and gaze returns to frontal as head orientation comes back to a frontal orientation. An example is shown in Figure 6. Unless a classifier includes dynamic information, expressions cannot be accurately disambiguated.

Dynamics are critical as well to the very perception of subtle facial expressions. Subtle facial expressions that are not identifiable when viewed statically suddenly became apparent in a dynamic display [51] . An example is shown in Figure 7. Viewers were shown the same emotion expressions in one of four conditions. In the static condition, they saw only the peak expression; in multi-static, they saw each image in a video sequence from neutral to peak but with a visual mask inserted between each image. Multi-static gave viewers access to all of the images but the visual masks prevented the perception of motion. In the dynamic condition, they viewed the image sequence without the visual masks. In a first-last condition, they saw a movie sequence containing the first and last frames. Only in the latter two conditions in which motion information was afforded were expressions reliably recognized. This effect of motion was highly robust and was observed for all six basic emotions examined. An implication for machine perception is that recognition will be more difficult if performed independently for each frame, as is typically done. Dynamics are essential to human and machine perception of subtle facial expressions and to interpretation of almost all expressive actions.

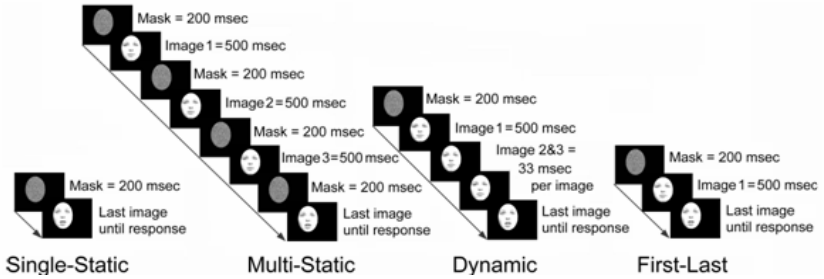

Fig. 7.. Four viewing conditions: single-static, multi-static, dynamic, and first-last. Single-static shows only the final image. Multi-static shows the image sequence, but inclusion of visual masks block perception of motion. Dynamic shows the same image sequence as multi-static but includes motion information. First-last shows first and last images and motion. With some individual exceptions, subtle expressions are visible only in the latter two conditions in which motion or dynamics is visible. From [52] (©2006 APS)

When observers view facial expressions of emotion, their own faces respond rapidly and involuntarily with specific facial actions. *Zygomatic major*, which pulls the lip corners obliquely in smiling, and *corrugator supercili,* which pulls the brows together and down in frowning, contract automatically within about 0.5 seconds following perception of happy and angry faces, respectively. While such effects have been demonstrated only

by facial electromyographic recordings (EMG) [53], automatic facial image analysis has recently shown strong concurrent validity with facial EMG across the range of visible movement [36]. The latter finding suggests that it is feasible using computer vision to precisely detect rapid and often low intensity increases in facial muscle activity corresponding to positive and negative emotion.

4. INDIVIDUAL DIFFERENCES

Stable individual differences in facial expression emerge early in development and by adulthood represent 25% or more of the variation in emotion expression [16, 54]. Individual differences include reaction range for positive and negative affect, response latency, and probability of conforming to display rules. In both individual and interpersonal contexts, facial expression is moderately stable over time periods from 4- to 12 months [16] and is comparable to what has been reported previously for self-reported emotion and personality. Stability is sufficiently robust that individuals can be recognized at far above chance levels based solely on their use of facial actions (which may be considered facial behavior signatures).

In a representative study, we found moderate stability over periods of about 25 months in mothers' facial expression. Mothers were observed during face-to-face interactions with their first and second infants when each was about 2 months of age [54]. Mothers' and infants' facial expression, gaze, and vocalization were coded on a 1-s time base with behavioral descriptors. Mothers' behavior was more positive with their second-born infant (a birth-order effect), yet moderately correlated at both time points. There was no correlation between expressive behavior of first- and second-born infants; they were independent. Despite differences between infants and in mothers' interactions with each, stability in mothers' affective expression was pronounced. Findings such as these suggest that it is important to consider facial expression as bounded by range of normal variation for each person moderated by context and within which deviations may be interpreted.

Sources of individual differences in emotion expression include temperament, personality, gender, birth order (as noted above), socialization, and cultural background [55-57]. Display rules are culturally specific prescriptions for when and how to show emotion in various contexts. In some cultures, for instance, children learn not to express anger; whereas in others, anger is considered important to self expression. Among traditional Japanese, for instance, anger is less likely to be shown outside the family than in the U.S. [58]. As another example, European-American and Chinese-American couples differ in proportion of positive and negative expressions, but not in autonomic reactivity or self- reported emotion, when discussing conflicts in their relationship [59]. These differences are consistent with relative value placed on emotion expression in each culture. In Western societies, expression of positive affect is emphasized. Specific emotion expressions may also be ritualized and culture specific. The tongue-bite display (see Figure 8) communicates embarrassment/shame in parts of India and south Asia but

not the U.S. Inferences about emotion become more reliable when individual differences are taken into account.

Of the various sources of individual differences, relatively little is known about the timing of facial actions. Available evidence suggests that individual differences exist here as well. Women, for instance, in addition to being more expressive than men, respond more quickly with *zygomatic major* contraction to stimuli intended to elicit positive affect [16]. When depressed, response latency is slower and more variable, which may make coordination of interpersonal timing more difficult to achieve [60]. As methods for measuring the timing of facial expression become more available, individual differences in timing will become better known.

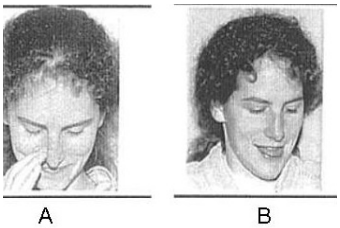

Fig. 8. Some expressions appear in all or almost all cultures. Others are culture specific (A and B, respectively). Examples here are for embarrassment. From [61]. (©1998 Taylor & Francis)

5. INTERPERSONAL REGULATION

Interpersonal regulation includes synchrony, reciprocity, and coordinated interpersonal timing. Synchrony, or coherence, refers to the extent to which individuals are moving together in time with respect to one or more continuous output measures, such as specific facial actions, affective valence, or level of arousal. Reciprocity refers to the extent to which behavior of one individual is contingent on that of the other. Both synchrony and reciprocity have proven informative in studies of marital interaction, social development, and social psychology. Figure 9 shows an example from mother-infant interaction [42, 43]. Facial features and head motion were tracked automatically by the CMU/Pitt automated facial image analysis system version 4 [36]. The time series plot shows displacement of mother and infant lip-corners during smiles. Note that while partners tend to cycle together, there is an alternating pattern in which mother and infant take turns in leading the dyad into shared smiling.

Coordinated interpersonal timing (CIT) is the extent to which participants in a social interaction match the duration of interpersonal pauses or floor switches [62]. Floor switches are pauses that occur between the time when one person stops speaking and another begins. Coordination of floor switches follows an inverted U-shaped function in relation to affective intensity and change with development. In clinical depression, for

instance, CIT becomes longer and less predictable [60] CIT has been studied most often with respect to vocal timing, but applies equally to facial expression and other modalities.

In behavioral science literature, time- and frequency domain analyses have emphasized issues of quasi-periodicity in the timing of expressive behavior and bidirectional influence with respect to amplitude (see, for instance, [63]). Lag-sequential and related hidden Markov modeling have been informative with respect to the dynamics of discrete actions and individual and dyadic states [64]. Recent work with dampened oscillator models considers regulation of changes in velocity and acceleration [65]. Most approaches assume that time series are stationary. This assumption may not always hold for behavioral data. Boker [66] identified "symmetry breaks," in which the pattern of lead-lag relationships between partners abruptly shifts. Failure to model these breaks may seriously compromise estimates of mutual influence.

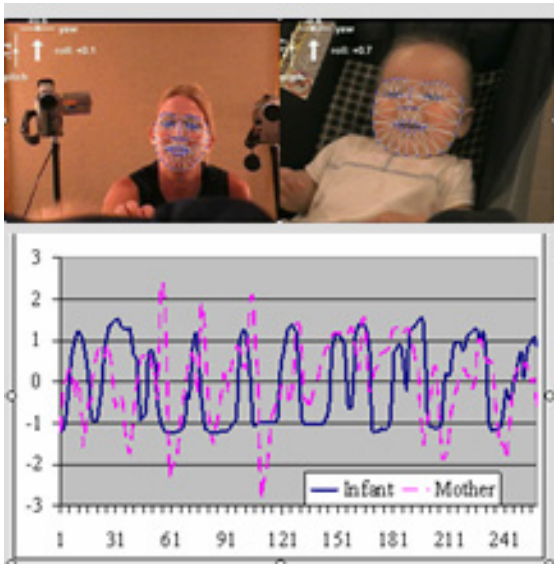

Fig. 9. Example of interaction analysis. Synchrony and reciprocity of smiling between mother and infant. Source: [42, 43]. (©2006 ACM)

6. CONCLUSION

Emotions are species-typical patterns that evolved because of their value in addressing fundamental life tasks [22]. They are central to human experience, yet largely beyond the comprehension of contemporary computer interfaces. Human-centered computing seeks to enable computers to unobtrusively perceive, understand, and respond appropriately to human emotion, to do so implicitly, without the need for deliberate human input. To

achieve this goal, it is argued [13] that we forgo the notion of "emotion recognition" and adopt an iterative approach found in human-human interaction. In our daily interactions, we continually make inferences about other people's emotions – their intentions, action tendencies, appraisals, other cognitions, and subjective feelings – from their expressive behavior, speech, and context. The success of human-centered computing depends in part on its ability to adopt an iterative approach to inference. Computing systems are needed that can automatically detect and dynamically model a wide range of multimodal behavior from multiple persons, assess context, develop representations of individual differences, and formulate and test tentative hypotheses though the exchange of communicative signals. Part of the challenge is that the computer becomes an active agent, in turn influencing the very process it seeks to understand. Human emotions are moving targets.

7. ACKNOWLEDGMENTS

Portions of this work were supported by NIMH grant R01 MH51435 and NSF HSD grant 0527397 to the University of Pittsburgh and Carnegie Mellon University.

8. REFERENCES

1. Lazarus, R.S.: Emotion and adaptation. Oxford, NY (1991)
2. Ekman, P.: Emotions revealed. Times, New York, NY (2003)
3. Keltner, D., Haidt, J.: Social functions of emotions at multiple levels of analysis. Cognition and Emotion **13** (1999) 505-522
4. Darwin, C.: The expression of the emotions in man and animals (3rd Edition). Oxford University, New York, New York (1872/1998)
5. Ekman, P., Rosenberg, E. (eds.): What the face reveals. Oxford, New York (2005)
6. Pantic, M., Patras, I.: Dynamics of facial expressions: Recognition of facial actions and their temporal segments from profile image sequences. IEEE Transactions on Systems, Man, and Cybernetics, Part B **36** (2006) 443-449
7. Tian, Y., Cohn, J.F., Kanade, T.: Facial expression analysis. In: Li, S.Z., Jain, A.K. (eds.): Handbook of face recognition. Springer, New York, New York (2005) 247-276
8. Pantic, M., Rothkrantz, M.: Automatic analysis of facial expressions: The state of the art. IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000) 1424-1445
9. Davidson, R.J., Ekman, P., Saron, C.D., Senulis, J.A., Friesen, W.V.: Approach-withdrawal and cerebral asymmetry: Emotional expression and brain physiology I. Journal of Personality and Social Psychology **58** (1990) 330-341
10. Levenson, R.W., Ekman, P., Friesen, W.V.: Voluntary facial action generates emotion-specific autonomic nervous system activity. Psychophysiology **27** (1990) 363-384
11. Watson, D., Tellegen, A.: Toward a consensual structure of mood. Psychological Bulletin **98** (1985) 219-235

12. Cohn, J.F., Elmore, M.: Effects of contingent changes in mothers' affective expression on the organization of behavior in 3-month old infants. Infants Behavior and Development **11** (1988) 493-505
13. Schmidt, K.L., Cohn, J.F.: Human facial expressions as adaptations: Evolutionary perspectives in facial expression research. Yearbook of Physical Anthropology **116** (2001) 8-24
14. Gottman, J., Levenson, R., Woodin, E.: Facial expressions during marital conflict Journal of Family Communication **1** (2001) 37-57
15. Cohn, J.F., Campbell, S.B.: Influence of maternal depression on infant affect regulation. In: Cicchetti, D., Toth, S.L. (eds.): Developmental perspectives on depression. University of Rochester Press, Rochester, New York (1992) 103-130
16. Cohn, J.F., Schmidt, K.L., Gross, R., Ekman, P.: Individual differences in facial expression: Stability over time, relation to self-reported emotion, and ability to inform person identification. International Conference on Multimodal User Interfaces, Pittsburgh, PA (2002) 491-496
17. Bartlett, M.S., Littlewort, G., Frank, M., Lainscsek, C., Fasel, I., Movellan, J.: Fully automatic facial action recognition in spontaneous behavior. IEEE International Conference on Automatic Face and Gesture Recognition, Vol. FG 2006, Southampton, England (2006) 223-228
18. Valstar, M.F., Pantic, M., Ambadar, Z., Cohn, J.F.: Spontaneous vs. posed facial behavior: Automatic analysis of brow actions. ACM International Conference on Multimodal Interfaces, Banff, Canada (2006) 162-170
19. Lucey, S., Matthews, I., Hu, C., Ambadar, Z., De la Torre, F., Cohn, J.F.: AAM derived face representations for robust facial action recognition. Seventh IEEE International Conference on Automatic Face and Gesture Recognition, Southampton, UK (2006) 155-160
20. Boker, S.M., Cohn, J.F., Matthews, I., Ashenfelter, K., Spies, J., Brick, T., Deboeck, P., Covey, E., Tiberio, S.: Coordinated motion and facial expression in dyadic conversation. NSF Human and Social Dynamics Principal Investigators Meeting, Washington, DC. (2006)
21. Cohn, J.F., Ambadar, Z., Ekman, P.: Observer-based measurement of facial expression with the Facial Action Coding System. In: Coan, J.A., Allen, J.B. (eds.): The handbook of emotion elicitation and assessment. Oxford University Press Series in Affective Science. Oxford University, New York, NY (In press)
22. Ekman, P.: An argument for basic emotions. Cognition and Emotion **6** (1992) 169-200
23. Ekman, P., Friesen, W.V.: Unmasking the face: A guide to emotions from facial cues. Prentice-Hall Inc., Englewood Cliffs, NJ (1975)
24. Ekman, P., Friesen, W.V.: Facial action coding system. Consulting Psychologists Press, Palo Alto, CA (1978)
25. Kanade, T., Cohn, J.F., Tian, Y.: Comprehensive database for facial expression analysis. Fourth IEEE International Conference on Automatic Face and Gesture Recognition, Vol. 4, Grenoble (2000) 46-53
26. Pantic, M., Sebe, N., Cohn, J.F., Huang, T.S.: Affective multimodal human-computer interaction. ACM International Conference on Multimedia (2005) 669-676
27. Pantic, M., Rothkrantz, M.: Toward an affect-sensitive multimodal human-computer interaction. Proceedings of the IEEE **91** (2003) 1371-1390
28. Keltner, D., Ekman, P.: Facial expression of emotion. In: Lewis, M., Haviland, J.M. (eds.): Handbooks of emotions. Guilford, New York (2000) 236-249
29. Scherer, K.R.: What does facial expression express? In: Strongman, K.T. (ed.): International Review of Studies on Emotion, Vol. 2. John Wiley & Sons Ltd. (1992) 138-165

30. Izard, C.E., Dougherty, L.M., Hembree, E.A.: A system for identifying affect expressions by holistic judgments. Instructional Resources Center, University of Delaware, Newark, Delaware (1983)
31. Ekman, P., Friesen, W.V., O'Sullivan, M.: Smiles when lying. Journal of Personality and Social Psychology **54** (1988) 414-420
32. Cohn, J.F., Ekman, P.: Measuring facial action by manual coding, facial EMG, and automatic facial image analysis. In: Harrigan, J.A., Rosenthal, R., Scherer, K. (eds.): Handbook of nonverbal behavior research methods in the affective sciences. Oxford, New York (2005) 9-64
33. Ekman, P., Friesen, W.V., Hager, J.C. (eds.): Facial action coding system. Research Nexus, Network Research Information, Salt Lake City, UT (2002)
34. Rosenthal, R.: Conducting judgment studies. In: Harrigan, J.A., Rosenthal, R., Scherer, K.R. (eds.): Handbook of nonverbal behavior research methods in the affective sciences. Oxford, NY (2005) 199-236
35. Oster, H., Hegley, D., Nagel, L.: Adult judgments and fine-grained analysis of infant facial expressions: Testing the validity of a priori coding formulas. Developmental Psychology **28** (1992) 1115-1131
36. Cohn, J.F., Kanade, T.: Use of automated facial image analysis for measurement of emotion expression. In: Coan, J.A., Allen, J.B. (eds.): The handbook of emotion elicitation and assessment. Oxford, New York, NY (in press)
37. Eibl-Eibesfeldt, I.: Human ethology. Aldine de Gruvier, NY, NY (1989)
38. Sayette, M.A., Cohn, J.F., Wertz, J.M., Perrott, M.A., Parrott, D.J.: A psychometric evaluation of the Facial Action Coding System for assessing spontaneous expression. Journal of Nonverbal Behavior **25** (2001) 167-186
39. Fridlund, A.J., Sabini, J.P., Hedlund, L.E., Schaut, J.A., Shenker, J.J., Knauer, M.J.: Audience effects on solitary faces during imagery: Displaying to the people in your head. Journal of Nonverbal Behavior **14** (1990) 113-137
40. Cohn, J.F., Schmidt, K.L.: The timing of facial motion in posed and spontaneous smiles. International Journal of Wavelets, Multiresolution and Information Processing **2** (2004) 1-12
41. Fleiss, J.L.: Statistical methods for rates and proportions. Wiley, New York, New York (1981)
42. Messinger, D.S., Chow, S.M., Koterba, S., Hu, C., Haltigan, J.D., Wang, T., Cohn, J.F.: Continuously measured smile dynamics in infant-mother interaction. Miami (2006)
43. Ibanez, L., Messinger, D., Ambadar, Z., Cohn, J.F.: Automated measurement of infant and mother interactive smiling. American Psychological Society (2006)
44. Russell, J.A., Fernandez-Dols, J.M. (eds.): The psychology of facial expression. Cambridge University, Cambridge, United Kingdom (1997)
45. Ekman, P.: Facial expression and emotion. American Psychologist **48** (1993) 384-392
46. Cohn, J.F.: Automated analysis of the configuration and timing of facial expression. In: Ekman, P., Rosenberg, E. (eds.): What the face reveals. Oxford, New York (2005) 388-392
47. Theobald, B.J., Cohn, J.F.: Facial image synthesis. In: Sander, D., Scherer, K.R. (eds.): Oxford companion to affective sciences: An encyclopedic dictionary for the affective sciences. Oxford University Press, NY (In press) xxx-xxx
48. Edwards, K.: The face of time: Temporal cues in facial expressions of emotion. Psychological Science **9** (1998) 270-276
49. Krumhuber, E., Kappas, A.: Moving smiles: The role of dynamic components for the perception of the genuineness of smiles. Journal of Nonverbal Behavior **29** (2005) 3-24
50. Cohn, J.F., Reed, L.I., Moriyama, T., Xiao, J., Schmidt, K.L., Ambadar, Z.: Multimodal coordination of facial action, head rotation, and eye motion. Sixth IEEE International

Conference on Automatic Face and Gesture Recognition, Vol. FG'04, Seoul, Korea (2004) 645-650
51. Ambadar, Z., Schooler, J., Cohn, J.F.: Deciphering the enigmatic face: The importance of facial dynamics to interpreting subtle facial expressions. Psychological Science **16** (2005) 403-410
52. Ambadar, Z., Cohn, J.F., Reed, L.I.: All smiles are not created equal: Timing characteristics and interpretation of spontaneous smiles. (2006)
53. Dimberg, U., Thunberg, M., Grunedal, S.: Facial reactions to emotional stimuli: Automatically controlled emotional responses. Cognition and Emotion **16** (2002) 449–471
54. Moore, G.A., Cohn, J.F., Campbell, S.B.: Mothers' affective behavior with infant siblings: Stability and change. Developmental Psychology. **33** (1997) 856-860
55. Oster, H., Camras, L.A., Campos, J., Campos, R., Ujiee, T., Zhao-Lan, M., Lei, W.: The patterning of facial expressions in Chinese, Japanese, and American infants in fear- and anger-eliciting situations. Poster presented at the International Conference on Infant Studies, Providence, RI (1996)
56. Matsumoto, D., Willingham, B.: The thrill of victory and the agony of defeat: spontaneous expressions of medal winners of the 2004 Athens olympic games. Journal of Personality & Social Psychology **91** (2006) 568–581
57. Camras, L.A., Chen, Y.: Culture, ethnicity, and children's facial expressions: A study of European American, mainland Chinese, Chinese American, and adopted Chinese girls. **6** (2006) 103–114
58. Markus, H.R., Kitayama, S.: Culture and the self: Implications for cognition, emotion, and motivation. Psychological Review **98** (1991) 224-253
59. Tsai, J.L., Levenson, R.W., McCoy, K.: Cultural and temperamental variation in emotional response. Emotion **6** (2006) 484-497
60. Zlochower, A.J., Cohn, J.F.: Vocal timing in face-to-face interaction of clinically depressed and nondepressed mothers and their 4-month-old infants. Infant Behavior and Development **19** (1996) 373-376
61. Haidt, J., Keltner, D.: Culture and facial expression: Open-ended methods find more expressions and a gradient of recognition. Emotion and Cognition **13** (1999) 225-266
62. Jaffe, J., Beebe, B., Feldstein, S., Crown, C.L., Jasnow, M.: Rhythms of dialogue in early infancy. Monographs of the Society for Research in Child Development **66** (2001)
63. Cohn, J.F., Tronick, E.Z.: Mother-Infant face-to-face interaction: Influence is bidirectional and unrelated to periodic cycles in either partner's behavior. Developmental Psychology **34** (1988) 386-392
64. Cohn, J.F., Tronick, E.Z.: Mother infant interaction: The sequence of dyadic states at three, six, and nine months. Developmental Psychology **23** (1987) 68 77
65. Chow, S.M., Ram, N., Boker, S.M., Fujita, F., Clore, G.C.: Emotion as a thermostat: representing emotion regulation using a damped oscillator model. Emotion **5** (2005) 208-225
66. Boker, S.M., Xu, M., Rotondo, J.L., King, K.: Windowed cross–correlation and peak picking for the analysis of variability in the association between behavioral time series. Psychological Methods **7** (2002) 338-355