Affect Decoding Measures and Human-Computer Interaction

L.D. Riek, S. Afzal, and P. Robinson
Computer Laboratory, University of Cambridge, Cambridge, UK, lr314@cl.cam.ac.uk

Introduction

Affect measurement has been identified as a critical issue for Human-Computer Interaction (HCI) [4]. There is an established body of work in the communication science and psychology literature considering an individual's ability to decode affect. However, this area has not yet been applied to HCI. Decoding affect is defined as "the ability to sense, perceive accurately, and respond appropriately to one's personal, interpersonal, and social environment" [2]. Objective measurement of affect decoding is relevant for a variety of HCI applications, such as designing and selecting distance collaboration tools, analysis of behavioral video, and explicitly describing observed nonverbal behavior during dyad and group interaction studies. We are used to adapting interfaces for users with disabilities [6]; it is equally important that systems that support cooperative work should adapt to their users' difficulties with decoding affect. It is therefore necessary to be able to measure an individual's affective decoding abilities.

Measurement Methods for Decoding Affect[1]

A wide range of methods are available to quantitatively measure how people decode affect. These methods usually involve evaluation of one or more channels of affective communication, such as body gesture, facial expression, vocal behavior, or other nonverbal cues. Each method has its advantages and drawbacks; indeed, when one reviews such methods it is useful to note the internal consistency score of the method. We describe four methods that measure the affect decoding ability of subjects presented with non-static stimuli. The stimuli presented are typically video and/or audio of people expressing various levels of affect.

Profile of Nonverbal Sensitivity Test (PONS) [11]

PONS measures the accuracy with which subjects infer the nonverbal cues of emotionally laden content acted out by a female expresser via face, body, and vocal channels. The test is available in both full and short versions. The full-length test consists of 220 2-second audio clips, video clips, or both. After each clip, subjects are given a multiple choice question and asked to select the correct portrayal. People with higher PONS scores are rated as more interpersonally sensitive by supervisors and peers, have more well-adjusted personalities, and are rated as better job performers as clinicians and teachers [5].

Communication of Affect Receiving Ability (CARAT) [3]

CARAT is a standardized method for measuring how subjects decode spontaneously generated facial expressions. First, stimulus subjects are videotaped while viewing 32 emotionally evocative slides (i.e., laughing children or injured animals). Test subjects then view these recorded segments and are asked to choose which slide the stimulus subject was viewing. The total score is a measure of the test subjects' emotional decoding skill.

Empathic accuracy standard-cue methodology [7]

This test measures the "empathic accuracy" of subjects as they watch videos of three female patients interacting (individually) with a therapist. Subjects are shown and/or played 30 15s video clips. After each clip, they are asked to rate if the patient was feeling something and, if so, to describe the feeling in a single sentence. Following this, eight independent raters then compare the subjects' response in tandem with the video and measure accuracy on a three-point scale. An emphatic accuracy score is then generated [7].

Empathy Quotient (EQ) [1]

EQ is another measure of empathy and is a short, easy to use and score evaluative method. Subjects are given an 80-question, 4-point Likert scale pencil-and-paper test. EQ has 40 questions that probe emotional empathy ("It doesn't bother me too much if I am late meeting a friend."), and 20 filler questions. On the empathic behavior questions subjects receive 1 point for a mild response and 2 points for a strong response. The empathy questions are evenly balanced between "strongly agree" and "strongly disagree" responses to avoid bias. EQ has been shown to be significantly lower in adults with Asperger Syndrome or high-functioning autism compared with controls. Further, women tend to score significantly higher on EQ than men [1].

Applications to HCI

Methods for measuring affective decoding ability are applicable to a variety of tasks relevant to the HCI community. First, for Computer-Supported Collaborative Work (CSCW) designers, the ability to design systems based on the affective decoding scores of their target user population could greatly inform their decision-making. Second, when HCI practitioners are selecting subjects for user trials, it can be very important to know their subjects' ability to decode affect. Finally, for scientists who analyze behavioral video data, knowledge of their own affective decoding skills can be helpful in both determining their level of expertise and in explicitly describing how they derive their conclusions.

Sensitivity to behavioral cues is a key factor for efficient coordination and collaboration. For example, a group of non-collocated architects collaborating on a design task involves interpersonal sensitivity. Each architect's ability to understand the affect of their colleagues can greatly impact the success of the task. By pre-testing the architects, a technology designer or selector can better compensate for the strengths and weakness of the group at large. In this instance, the use of the PONS test might prove helpful to provide a multi-channel score with which to work.

Another application of these methods is for video analysis of natural corpora. Affective computing [9] methods often involve dealing with naturally collected data. Labeling this data is important as it serves as the ground-truth for validation of the developed computational techniques. However, training annotators to do this task is often very costly, particularly if they are not skilled at recognizing nonverbal affect. It may be beneficial to use measures such as CARAT, Empathic Accuracy, or EQ to pre-screen annotators, as well as to indicate where training is required. Annotator scores can also serve as reliability indicators for the labeled data.

Proceedings of Measuring Behavior 2008 (Maastricht, The Netherlands, August 26-29, 2008)
Eds. A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, G. Smit, and P.H. Zimmerman

More generally, for scientists studying affect among dyads or groups, familiarity with measures such as PONS can be very informative for explicitly describing observed nonverbal behavior as well as for communicating how such descriptions were derived.

References

1. Baron-Cohen, S., Wheelwright, S. (2004). The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders* **34:2**, 163-175.
2. Bernieri, F. J. (2001). Toward a Taxonomy of Interpersonal Sensitivity. In J. A. Hall & F. Bernieri (Eds.). *Interpersonal Sensitivity: Theory and Measurement* pp. 3-20. Mahwah: Earlbaum.
3. Buck, R. (1979). Measuring Individual Differences in the Nonverbal Communication of Affect: The Slide-Viewing Paradigm. *Human Communication Research* **6**, 47-57.
4. Crane, E. A., Shami, N. S., Peter, C. (2007). Let's Get Emotional: Emotion Research in Human Computer Interaction. *In Proc. of Conference on 2007 ACM Conference Computer-Human Interaction (CHI)*.
5. Hall, J. Bernieri, F. J., Carney, D. R. (2005). Nonverbal and Interpersonal Sensitivity. In Jinni, A. Harrigan, J. A., Rosenthal, R., and Scherer, K. R. (Eds.) *The New Handbook of Methods in Nonverbal Behavior Research,* pp. 267-269. Oxford: Oxford University Press.
6. Hughes, G., Robinson, P. (2006). Lecture Adaptation for Students with Visual Disabilities Using High-Resolution. *In Proc. of 2006 ACM Conference on Computers and Accessibility (ASSETS)*.
7. Ickes, W., Marangoni, C., Garcia, S. (1997). Studying Empathic Accuracy in a Clinically Relevant Context. In W. Ickes (Ed.), *Empathic Accuracy*, pp. 282-310. New York: Guilford Press.
8. Jinni, A. Harrigan, J. A., Rosenthal, R., Scherer, K. R. (2005). *The New Handbook of Methods in Nonverbal Behavior Research*. Oxford: Oxford University Press.
9. Picard, R.W. (1997). *Affective Computing*. Cambridge: The MIT Press.
10. Riggio, R. E. (2006). Nonverbal Skills and Abilities. In V. Manusov, V. and Patterson. M. (Eds.), *The SAGE Handbook of Nonverbal Communication*, pp. 79-87. Thousand Oaks: Sage Press.
11. Rosenthal, R.E., Hall, J. A., DiMatteo, M. R, Rogers, P. L., and Archer, D. (1979). *Sensitivity to Nonverbal Communication: The PONS Test.* Baltimore: Johns Hopkins University Press.

[1]For a more thorough overview of these methods, as well as a description of many others, please see *The SAGE Handbook of Nonverbal Communication* [10] and *The New Handbook of Methods in Nonverbal Behavior Research* [8].