Computation of Emotions

Peter Robinson
University of Cambridge Computer Laboratory
15 JJ Thomson Avenue
Cambridge CB3 0FD England
pr@cl.cam.ac.uk

ABSTRACT

When people talk to each other, they express their feelings through facial expressions, tone of voice, body postures and gestures. They even do this when they are interacting with machines. These hidden signals are an important part of human communication, but most computer systems ignore them. Emotions need to be considered as an important mode of communication between people and interactive systems. Affective computing has enjoyed considerable success over the past 20 years, but many challenges remain.

Categories and Subject Descriptors

H.5 [**Information Systems**]: Information interfaces and presentation—*User interfaces*; I.3.6 [**Computing Methodologies**]: Computer Graphics—*Interaction techniques*.

Keywords

Affective computing, emotions, human-computer interaction.

1. INTRODUCTION

Charles Darwin published *The expression of the emotions in man and animals* in 1872, exploring the role of emotional expression in communication between humans. Over a century later, Rosalind Picard observed that effective communication between people and computers also requires emotional intelligence; computers must have the ability to recognize and express emotions.

The study of affective computing has blossomed subsequently. This paper presents a summary of some of the challenges involved in affective computing, and illustrates them with examples drawn from work in the Computer Laboratory at the University of Cambridge[1]; it is not intended as a comprehensive survey. We have also produced a video, *The Emotional Computer* [6], which gives a light-hearted account of some of the projects[2].

[1]http://www.cl.cam.ac.uk/emotions/
[2]http://www.youtube.com/watch?v=whCJ4NLUSB8

ICMI'14, November 12–16, 2014, Istanbul, Turkey.
ACM 978-1-4503-2885-2/14/11.

2. RECOGNISING EMOTIONS

Although Darwin concentrated on facial features to convey emotions in Expression, he also mentions vocal sounds, other sounds, body posture and gesture, and physiological responses as further indications of emotion. All of these channels have been considered as ways of automatic monitoring emotion in humans, although these sensors used for some are more invasive than for others. Signals that can be monitored non-invasively are of particular interest for pervasive computing.

2.1 Facial expressions

People routinely express their mental states through their facial expressions and this is one of the clearest channels for communication. Inference from facial expressions has been studied using a variety of techniques but mostly restricted to six basic emotions. Recognising complex, cognitive mental states is more difficult, but probably more useful as part of general interaction with computer systems. We have developed a fully automatic system requiring no human intervention which operates in real-time [4]. Our Facial Affect Inference System uses a multi-level representation of the video input, combined in a Bayesian inference framework operating at four levels: facial feature points, FACS action units (AUs), gestures composed of several AUs, and mental states.

An evaluation considered six conditions together covering a tenth of the range of mental states. For a mean false positive rate of 4.7%, the overall accuracy of the system is 77%. The system also generalises well to faces not included in the training data.

We have also considered inference of continuous measures of valence (from negative to positive) and arousal (from passive to active) [1].

2.2 Non-verbal aspects of speech

The voice provides another significant channel for the expression of emotions. Features such as the pitch, energy and tempo can reveal a lot about the mood of the speaker. There are no characteristic features that indicate particular mental states but it is possible to distinguish between using two emotions using a small number of features, with a different set of features may be required to distinguish those emotions from others. Our approach [8] has been to calculate a large collection of about 170 features for each utterance. A training phase uses data mining to identify the features that separate each pair of emotional conditions. The operational phase then uses these pair-wise comparisons as preferences in a voting scheme to give an overall ranking.

Evaluation separated nine conditions with an accuracy of 70%, increasing to 83% if multiple winners were considered. The approach also generalises well to speakers other than those used for training, and even to other languages.

2.3 Body posture and gesture

The third natural channel for expression of emotions includes body posture and gesture. However, characteristics indicating the movement being considered and who is doing it must be discounted before it is possible to analyse how it is being done. Movement involves an individual bias, so the analysis is harder than for facial expressions or voice [2].

Our approach breaks complex motions down into a system of isolated elements whose dynamic cues can be used to distinguish affects. As with speech, pair-wise comparisons are used on individual motion segments, and each segment is classified using a majority vote. A complete motion is then classified by a majority vote of its component segments.

The method was tested on a corpus of about 1200 motion samples, representing roughly equal numbers of four expressions of four different actions. The average recognition rate of 81% is comparable to the rates achieved by human observers of similar data.

3. EXPRESSING EMOTIONS

Humans routinely convey empathetic responses through involuntary facial mimicry, and this extends to human-robot interaction. An experiment showed that conversation between a participant and a robot is enhanced when the robot mimics the subject rather than moves randomly. This raises questions about the degree of human-likeness required in the appearance of robots that interact with humans. A further experiment investigated participants' empathy for robots shown in film clips, and the responses were directly correlated with human-likeness.

The technology has been tested using a high-fidelity robotic head which simulates movements disorders that might be encountered by trainee doctors [5], and we are currently looking at its use as an intervention for autism therapy.

4. APPLICATIONS

An early application of affective inference is monitoring cognitive load in command-and-control systems. Driving a car provides a good model for this, but it is difficult to construct repeatable experiments using real cars on real roads. Simulation allows controlled experiments, but fails to engage participants. Remote control of real vehicles retains control while enhancing emotional investment [3].

Autism Spectrum Conditions (ASCs) are neurodevelopmental conditions characterized by social communication difficulties and restricted and repetitive behaviour patterns. The European ASC-Inclusion project [7] aims to create and evaluate the effectiveness of an internet-based game platform, intended for children with ASCs and their carers. The platform combines several state-of-the art technologies in one comprehensive virtual world providing training through games, and including feedback from analysis of the player's gestures, facial and vocal expressions using a standard webcam and microphone. The game also includes text communication with peers and smart agents, animation, video and audio clips. One component of the game monitors the player's face while he or she is acting a particular emotion. Computer vision and machine learning are then used to infer the emotion depicted and report back, both assessing the player's performance and also suggesting changes to make it resemble a canonical performance more closely.

5. ACKNOWLEDGEMENTS

Many colleagues in the Computer Laboratory at the University of Cambridge have contributed to this work – notably Andra Adams, Shazia Afzal, Tadas Baltrušaitis, Ntombi Banda, Daniel Bernhardt, Ian Davies, Neil Dodgson, Yujian Gao, Vaiva Imbrasaitė, Rana el Kaliouby, Marwa Mahmoud, Tomas Pfister, Laurel Riek, Metin Sezgin and Tal Sobol Shikler. Simon Baron-Cohen and his colleagues in the Autism Research Centre have provided numerous technical insights. Kevin Hull produced and directed the *Emotional Computer* video. Parts of the work have been undertaken with support from the EPSRC, the European Commission, the Cambridge-MIT Institute, the Gates Cambridge Trust, Hanson Robotics, the Qualcomm Studentship, Thales Research & Technology, Toyota ITC, Vicon UK, the Neil Wiseman Fund and the University of Cambridge. ASC-Inclusion has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement number 289021.

6. REFERENCES

[1] T. Baltrušaitis, N. Banda, and P. Robinson. Dimensional affect recognition using continuous conditional random fields. In *IEEE Conference on Automatic Face and Gesture Recognition*, April 2013.

[2] D. Bernhardt and P. Robinson. Detecting affect from non-stylised body motions. In *Affective Computing and Intelligent Interaction*, pages 59–70, Lisbon, Portugal, 2007. Springer-Verlag.

[3] I. P. Davies and P. Robinson. Emotional investment in naturalistic data collection. In *Affective Computing and Intelligent Interaction*, Memphis, TN, Oct. 2011.

[4] R. el Kaliouby and P. Robinson. Real-time inference of complex mental states from facial expressions and head gestures. In *Computer Society Conference on Computer Vision and Pattern Recognition*, page 154, Washington, DC, 2004.

[5] L. D. Riek and P. Robinson. Synthesising head and facial movement disorders on android robots. In *International Symposium on Facial Analysis and Animation*, Edinburgh, Scotland, October 2010.

[6] P. Robinson, T. Baltrušaitis, I. P. Davies, T. J. Pfister, L. D. Riek, and K. Hull. The emotional computer. In *International Conference on Pervasive Computing*, San Francisco, CA, June 2011.

[7] B. W. Schuller, E. Marchi, S. Baron-Cohen, H. O'Reilly, P. Robinson, I. P. Davies, O. Golan, S. Fridenson, S. Tal, S. Newman, N. Meir-Goren, R. Shillo, A. Camurri, and S. Piana. ASC-Inclusion: Interactive emotion games for social inclusion of children with Autism Spectrum Conditions. In *Intelligent Digital Games for Empowerment and Inclusion*, May 2013.

[8] T. Sobol Shikler and P. Robinson. Classification of complex information: inference of co-occurring affective states from their expressions in speech. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(7):1284–1297, July 2010.