Multi-modal Continuous Human Affect Recognition

Haolin Wei *, David S. Monaghan*, Noel E. O'Connor *
Insight Centre for Data Analytics
haolin.wei2@mail.dcu.ie, david.monaghan@dcu.ie, Noel.OConnor@dcu.ie

Abstract

Automatic human affect analysis has attracted increasing attention from the research community in recent years. This extended abstract gives an overview of the current affect recognition work carried out at Dublin City University.

1. Motivation

Understanding human affective state is indispensable for human-human interaction and social contact. As computer systems have become part of our daily lives, machines that can understand human emotions could be potentially useful for Human Computer Interaction (HCI), customer services, call centers, E-learning, intelligent autonomous vehicles, games and entertainment.

2. Problem Statement

The majority of datasets and studies reported in the Affect Computing literature have used unrealistic scripted acted emotions. Such acted or induced data cannot model natural interaction sufficiently as expressions produced are too pronounced and are seldom encountered in more realistic data. In addition, almost all datasets are captured using controlled environments with fixed blank backgrounds and constant lighting. Systems trained on such data are very likely to fail if used to understand the subtle complex nature of spontaneous interactions in real-world applications.

3. Related Work

Given the practical and theoretical importance of the affective computing field, lots of research have been conducted towards automatic emotion and affect recognition. Various multi-modal spontaneous datasets have been created [2]. However there still lack of continuous annotated spontaneous dataset consists video, audio and depth signals. Early research focused on recognising basic emotions from visual and vocal cues [1] while recent trends have shifted towards recognising continuous and dimensional affective state [3].

4. Research Question

Could the depth modality increase the accuracy and robustness of the affective recognition system?

5. Hypothesis

This research will demonstrate that, the dataset captured in an unconstrained environment could generalise the complexity of expressive behaviours found in real-world settings and the use of additional depth modality could improve the recognition result by providing more robust facial landmark detection and additional appearance features.

6. Proposed Solution

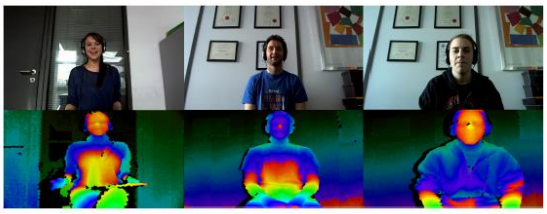

Figure 1: Sample screenshot from captured affect dataset

In this research, we propose a robust affect analysis and detection system for real-world applications. A state of art multi-modal affect dataset is first collected (See Figure 1). The datset consists of video, audio and depth signals. The spontaneous affective states are elicited using a three way debate scenario in an unconstrained environment with various lighting conditions and backgrounds. Secondly, the geometric and appearance facial features along with audio features and additional depth features extracted from visual and vocal cues will be used to train an regression model to continuously predict the affect value.

7. Evaluation

The evaluation will be performed on each modality separately and on the fusion of video and audio modalities as well as video, audio and depth modalities. The evaluation will quantify the system error based on a manually annotated ground-truth.

References

[1] H. Gunes and M. Piccardi. Bi-modal emotion recognition from expressive face and body gestures. *Journal of Network and Computer Applications*, 30(4):1334 – 1345, 2007. Special issue on Information technology.

[2] H. Gunes and B. Schuller. Categorical and dimensional affect analysis in continuous input: Current trends and future directions. *Image and Vision Computing*, 31(2):120–136, 2013.

[3] S. Petridis and M. Pantic. Audiovisual discrimination between laughter and speech. In *IEEE International Conference on Acoustics, Speech and Signal Processing, 2008. ICASSP 2008.*, pages 5117–5120, March 2008.

*Acknowledgement EU FP7-ICT-281123 REVERIE