Emotion through Intel RealSense

CS2951K Final Project Proposal

Ning Hou (nhou), Lee Painton, Eric Rosen

February 26, 2015

1 Research question

Affect display is the combination of facial, gestural and vocal cues by which persons consciously or unconsciously communicate emotion. In their recommendations on affective multimodal HCI, Pantic et al [3, p.3] suggest the use of an artificial neural network when infering affect with the rationale that some subtle cues are difficult to pick up via conventional analysis. We are curious if this rationale holds; namely given a set of modes whether a multimodal bayes filter with analytically tuned parameters or a neural network provides more accurate results.

2 Significance

Reliably determining user affect is an open problem in HCI and part of a field called affective computing. The development of affect sensitive intelligent agents would computers to interact more effectively with humans in tasks where emotion has an impact, learning or driving for example. Part of the problem with infering affect is based in technological limitations which we hope to address using Intel's RealSense technology. Properly trained, emotionally aware agents might even be able to pick up subtle shifts in affect that would elude the average human observer. There has also been work in using affect aware agents as learning companions [1] and affect recognition in children with autism spectrum disorder [2].

3 Methodology

We plan to formulate the problem at the highest level as both a multimodal bayes filter and a neural network where the objects being filtered on are affective states treated as points in a multi-dimensional space where the dimensions are factors of affect (e.g. anxiety-confidence, boredom-facination). We will also compile a corpus of training and testing data by interviewing 10-20 subjects about topics designed to elicit emotional experiences and recording the results using the RealSense device. If after a few trials this fails to yield useful data we will revert to a backup plan, generating data by recording individuals watching samples of audio-visual media designed to elicit emotional responses.

1. Set up method of obtaining facial expressive, gestural and vocal data from Intel RealSense.
2. Construct parallel Bayes filter and neural network models for determining affect.
3. Solve the decision problem using Bayes filter and neural network.
4. Test the model by user input using data set.
5. Adjust and formalize the model based on results.

4 Related Work

Kapoor et al [1] describes a theoretic framework used to describe affective states. Our work borrows from this idea but is generalized rather than focused on the activity of learning.

5 Schedule

Date	TODO
2/26 - 3/5	Finalize theoretic framework and experiment design
3/6 - 3/12	Program initial models and test with false data
3/13 - 3/19	Design interview script and post interview survey, find interview subjects and schedule
3/20 - 3/26	Have at least 10 subjects interviewed with collected data or move to backup plan
3/27 - 4/2	Test data on models and compare
4/2 - 4/7	Checkpoint presentation
4/8 - 4/14	Collect more data as needed; tweak models
4/15 - 4/21	Prepare results
4/23 - 4/28	Final presentation

6 Responsibilities

Everyone will have involvement with all parts of the project. We will assign work based on what is needed to meet milestones.

References

[1] Ashish Kapoor, Selene Mota, and Rosalind W Picard. Towards a learning companion that recognizes affect. In *AAAI Fall symposium*, pages 2–4, 2001.

[2] Changchun Liu, Karla Conn, Nilanjan Sarkar, and Wendy Stone. Affect recognition in robot assisted rehabilitation of children with autism spectrum disorder. In *Robotics and Automation, 2007 IEEE International Conference on*, pages 1755–1760. IEEE, 2007.

[3] Maja Pantic, Nicu Sebe, Jeffery F Cohn, and Thomas Huang. Affective multimodal human-computer interaction. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 669–676. ACM, 2005.