

Designing Evolving Computer Agent Capable of Emotion Recognition and Expression

Rahadian Yusuf, Shuyao Wang, Ivan Tanev, Katsunori Shimohara

Socioinformatics Laboratory, Doshisha University, Kyotanabe Campus, Kyoto, Japan
{yusuf2013, wang2013, itanev, kshimoha}@sil.doshisha.ac.jp

Abstract

Emotion cannot be separated from Affective Computing. This research is focusing on recognizing emotion of a single user using an evolving computer agent. The computer agent evolves itself by interacting with the user. It also gradually learns the user's unique features through the interactions.

Introduction

A computer agent is software that can be used to help in communication between computer and user. The idea of computer agent is to improve usability and user interface, therefore it can help users in operating computers.

The agent to be developed should be able to understand the mood and general emotion of a user, and should be able to express and evoke a better emotion from the user. This form of Affective Computing would help users in operating computers, and should help users in reducing the stress in using computers.

There are many approaches in developing a computer agent that is capable to recognize a user's emotion. This research is focusing on nurturing the interactions between user and computer agents, and evolving the computer agent to be able to recognize a user's emotion better. This research assumes that there will be a single user interacting with the computer agent.

Evolve through Interactions

The usual approach in recognition of emotions is by training an artificial intelligence using numerous images. This approach has several drawbacks, such as the generalization of facial expressions. Another drawback is the usage of still images or photographs, while a stream of video and time-related information is necessary for real-time interactions.

The usual approach also often uses actors as a model to provide expression samples (Ekman 2003). An indirect method sometimes are used (e.g. by requesting the actor to imagine a specific event to trigger an emotion), but the

actor might behave differently from the expected expression, thus might not yield the proper expression for emotion sample.

Assuming that each user has unique features and unique ways in expressing emotions, an approach to recognize their emotions using a computer agent that evolves along with interactions became feasible.

By learning a specific user, the agent might be able to recognize user's unique ways in expressing emotion. This kind of personal approach would be helpful in improving the quality of service in using PC, as well as capturing a feature that might be dominant for that specific user.

Features

There are many features that might relate to emotions, namely facial expressions and user's activity. As for facial expressions, there are many researches that pointed out the importance of eyes, eyebrows, and mouths in emotion detection (Ekman 2003). There are also other features that can be extracted from a user that might relate to emotion, such as proximity and stillness/activity (Pentland 2008).

When a user is operating PC, the distance between the user and the screen might relate about user's engagement towards their current activity. User's movements (head shakes, body tilts, etc.) might also relate to their current nervousness and/or eagerness. By implementing these two features in addition of facial expressions detection, user's unique features might be obtained and can be used for emotion recognition.

Evolution and Self Learning

As the computer agent will interact with the user during operating PC, the computer agent will evolve itself, adapting to a better recognition of user's features.

The interactions to be had should be simple, such as a yes/no question whenever there are some drastic changes on user's features, and also when the user consciously selects their current emotion (similar with updating status).

Based on this interaction, a learning mechanism can be implemented for the computer agent to evolve gradually, until it can recognize most of user's features and also on the emotion based on the features obtained.

Proposed System

Figure 1 below illustrates the diagram of the proposed system.

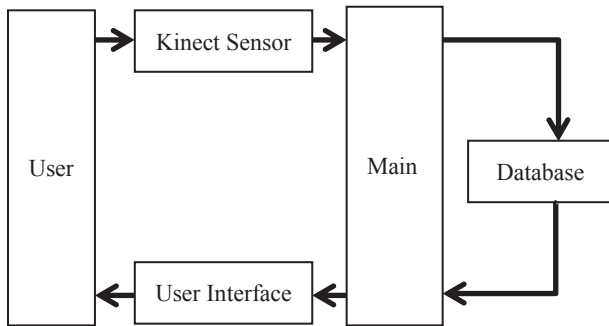


Figure 1 – System Diagram

Computer agent will extract features using Microsoft Kinect, as the device has the capability of capturing depth and color information, along with multi-array speaker that can be used for future researches (multi-modal feature extractions).

Computer agent will evolve using genetic programming, with an offline process during down time (such as when the user is not actively in front of the computer). Features that are being used consisted of head pose, facial muscles condition, proximity and stillness/activity of user.

User's extracted features will be stored on a database, and to be used to evolve the classifier.

Features are extracted in real-time, and with user's interaction (i.e. if the user sets the emotional status) the system will analyze the data few seconds before and after the input. Extra data regarding user's current emotional status will be stored as well for further evolutions, and when there are some major changes in the feature the system will interact again with the user.

Interactions

According to several researches (Ekman 2003, Pentland 2013, Ledoux 1996) there are several features related to emotion, such as facial muscles conditions (namely of head pose, eyes, eyebrows, mouth, and lips) and general proximity and activity. These features will be used for this research.

The agent would start with zero knowledge regarding the user's features, but always capturing said features in real time. Whenever there is a significant change on the features, the agent would assume there is some emotional changes and remind the user to update current emotional state (one of the option is 'neutral' state), while guessing user's current emotional state using stored information (starting condition is zero knowledge).

Another interaction is done from the user to the agent by deliberately selecting user's current emotional state. User can select or update their emotional state and force the agent to capture and store the information.

After emotional state setting by the user (deliberately or after being reminded by the agent), the agent would record the captured features from relevant timeslot into a database. During downtime, the agent will learn about the correlation between user's features and user's emotion (unique to each user) and adapt itself using the stored information, so that it can offer a better response to the user.

By using interaction approach, the user can still use the PC normally, and only during several situations would interact with the agent. The agent also captures user's features without the user fully realized (thus reducing other noise factors such as user's nervousness).

Conclusions and Future Works

The system designed can be improved greatly by implementing multi-modal feature extraction, such as by detecting user's voice. There are times when user utters a sudden groan or excitement, and these spikes of sound wave can be detected and used as feature for the system.

Another improvement can also be gained by refining the feature extraction methods, in order to obtain robust and high-fidelity features. Currently there are problems such as the fluctuated position data from the sensor.

Currently there are only few interactions: selecting limited variety of emotional state and agent reminding the user to update. A better method of interactions can be done by implementing a conversational approach, and adjusting the conversation/words according to the user.

Different learning mechanism or method is also possible. Genetic programming still has a long processing time, and improvements on the data structure might help as well.

References

- Pentland, A. 2008. *Honest Signals*. MIT Press.
- Ledoux, J. 1996. *The Emotional Brain*. Touchstone.
- Ekman, P. 2003. *Emotions Revealed*. Times Books.
- Picard, R.W. Affective Computing. MIT Media Laboratory Perceptual Computing Section Technical Report no. 321.
- Alfaro-Cid, E., Sharman, K., and Esparcia-Alcazar, A.I. 2006. Evolving a Learning Machine by Genetic Programming. in IEEE Congress on Evolutionary Computation, Canada.
- Song, M., Shinomori, K., and Zhang, S. 2010. The Contributions of Facial Features on the Smiling and Angry Expressions Judgment and Its Implications for the Computer Recognition System. in International Conference on Information and Automation.
- Picard, R.W. 2010. Affective Computing: From Laughter to IEEE. IEEE Transactions on Affective Computing, Vol 1, No. 1.