

Towards the improvement of self-service systems via emotional virtual agents

Christopher Martin
School of Computing &
Engineering Systems
University of Abertay,
Bell Street, Dundee
c.martin@abertay.ac.uk

Leslie Ball
School of Computing &
Engineering Systems
University of Abertay,
Bell Street, Dundee
l.ball@abertay.ac.uk

Jacqueline Archibald
School of Computing &
Engineering Systems
University of Abertay,
Bell Street, Dundee
j.archibald@abertay.ac.uk

Lloyd Carson
School of Social & Health
Sciences
University of Abertay,
Bell Street, Dundee
l.carson@abertay.ac.uk

Affective computing and emotional agents have been found to have a positive effect on human-computer interactions. In order to develop an acceptable emotional agent for use in a self-service interaction, two stages of research were identified and carried out; the first to determine which facial expressions are present in such an interaction and the second to determine which emotional agent behaviours are perceived as appropriate during a problematic self-service shopping task. In the first stage, facial expressions associated with negative affect were found to occur during self-service shopping interactions, indicating that facial expression detection is suitable for detecting negative affective states during self-service interactions. In the second stage, user perceptions of the emotional facial expressions displayed by an emotional agent during a problematic self-service interaction were gathered. Overall, the expression of disgust was found to be perceived as inappropriate while emotionally neutral behaviour was perceived as appropriate, however gender differences suggested that females perceived surprise as inappropriate. Results suggest that agents should change their behaviour and appearance based on user characteristics such as gender.

Keywords: affective computing, virtual agents, emotions, emotion detection, HCI, computer vision, empathy.

1. INTRODUCTION

This paper describes research which contributes towards the development of an empathetic system which will detect and improve a user's affective state during a problematic self-service interaction (SSI) through the use of an affective agent. Self-Service Technologies (SSTs) are those which allow a person to obtain goods or services from a retailer or service provider without the need for another person to be involved in the transaction. SSTs are used in many situations including high street shops, supermarkets and ticket kiosks. The use of SSTs may provide benefits such as improved customer service (for example allowing 24 hour a day, 7 days a week service), reduced labour costs and improved efficiency (Cho & Fiorito, 2010). Less than 5% of causes for dissatisfaction with SST interactions were found to be the fault of the customer (Meuter et al., 2000; Pujari, 2004), indicating that there is a need for businesses and SST manufacturers to improve these interactions in order to reduce causes for dissatisfaction (Martin et al., unpublished). The frustration caused by a negative SSI can have a detrimental effect on a user's behavioural intentions towards the retailer, impacting the likelihood the user will continue doing

business with them in the future and whether they will recommend them to other potential users (Lin & Hsieh, 2006; Johnson et al., 2008). By adopting affective computing practices in SSI design, such as giving computers the ability to detect and react intelligently to human emotions and to express their own simulated emotions, user experiences may be improved (Klein et al., 1999; Jaksic et al., 2006; Wang et al., 2009).

Affective agents have been found to reduce frustration during human-computer interactions (HCIs) (Klein et al., 1999; Jaksic et al., 2006), therefore we are investigating their effectiveness at improving negative affective states in a SST user during a shopping scenario. We propose a system which will detect negative affective states in a user and express appropriate empathetic reactions using an affective virtual agent.

Two stages of research were identified. The purpose of stage 1 (reported in Martin et al., in press) was to investigate whether emotional facial expressions are present during SST use, to determine whether a vision-based emotion detector would be suitable for this system. The purpose of stage 2 (reported in Martin et al., unpublished) was

to gather opinions on the appropriateness of several emotional facial expressions which may be expressed by the virtual agent during a SSI.

2. LITERATURE REVIEW

Several studies have focussed on having computers recognise emotions from facial expressions (Essa & Pentland, 1995; Tian et al., 2001; Valstar & Pantic, 2006; Kotsia & Pitas 2007; Kumano et al., 2009; Lucey et al., 2010; Satiyan & Nagarajan, 2010;). As reported in Martin et al. (in press), in stage 1 of this research we examined the facial Action Units (AUs) displayed during SSIs. AUs are descriptors for the movement of individual facial features which combine to form facial expressions (Kanade et al., 2000). Prototypical definitions of AU activation patterns have been identified for each of Ekman's (1987) basic emotions: anger, disgust, fear, happiness, sadness and surprise (Lucey et al., 2010). In order to develop an affect detection system, facial expressions which occur during SST interactions need to be identified so that a detector can be trained. We intend to use Haar-like feature extraction (Essa & Pentland, 1995; Kumano et al., 2009; Satiyan & Nagarajan, 2010) to detect facial expressions due to its ease of training and implementation.

Social and emotional intelligence, the abilities of an individual to deal rationally and effectively with social situations and the emotions of others (Salovey & Mayer, 1990, Pantic & Rothkrantz, 2003; Pantic, 2005), have been argued to be important for successful social functioning (Salovey & Mayer, 1990). Nass et al. (1994) showed that humans apply the same social rules to interactions with computers as they do to interactions with other humans. If people have similar social expectations for computers as they do for other humans, an interaction with a system whose perceived social or emotional intelligence falls below their expectations may prove unsatisfactory. Therefore it is important that a computer's perceived social and emotional intelligence is acceptable to users (Martin et al., in press; Pantic, 2005; Picard, 1997; Picard, 2001).

In order to progress towards the proposed empathetic self-service system, stage 2 of our research involved determining whether any emotional behaviours are found to be appropriate or inappropriate during a problematic SSI. Appropriate behaviours will be included in the abilities of our emotional agent while inappropriate behaviours will be excluded.

Virtual agents are personas involved in a HCI which interact with a user via a combination of a graphical representation, printed text and recorded

or synthesised speech. Affective virtual agents are those with affective abilities including affect detection and emotional expressiveness (Picard, 1998; Jaksic et al., 2006; Peters et al., 2006). The use of affective agents in HCIs has produced encouraging findings for the improvement of these interactions, including causing a user to interact longer with error-prone (Klein et al., 1999) or difficult tasks (Wang et al., 2009) and reducing frustration (Jaksic et al., 2006). While these findings indicate potential benefits of affective agents, the studies by Klein et al. (1999) and Jaksic et al. (2006) relied on a Wizard of Oz methodology, where the researchers manipulated the agent "behind the scenes" to simulate automated behaviour. The study by Wang et al. (2009) required users to manually self-report their emotional state. Our proposed empathetic system will be fully automatic.

The appropriateness of an agent's emotional behaviour for the situation in which it is used has been found to affect the agent's believability (Lim & Aylett, 2007), the degree of trust the user has in the agent (Cramer et al., 2010) and the user's stress level, with inappropriate emotional behaviour inducing stress (Becker et al., 2005). If our system is to be successful in reducing negative affective states in users during frustrating interactions, the emotional behaviour of the agent must be acceptable to users. Since appropriateness of emotional behaviour varies with situation, a study was carried out to evaluate a range of emotional facial expressions to determine whether any of them were acceptable in a problematic self-service shopping task. Acceptable emotions would be included in our empathetic system while inappropriate emotions would be excluded.

3. METHOD

In order to identify AUs which occur during SSIs, in stage 1 customers at a supermarket were filmed using self-service checkouts. Customers were approached when they were queuing and asked whether they would like to be filmed while they scanned their shopping. Those who agreed were directed to a self-service checkout where a video camera was set up on a tripod. The camera filmed the participant's head and shoulders for the duration of their interaction.

Thirty participants took part in the study (17 male, 13 female). In order to minimise disruption to customers and to increase the appeal of taking part, participants were not asked any questions or required to perform any additional task apart from read and sign the informed consent form. Although it may have been useful to gather demographic data, information about participants' technological

background and their current and previous experiences with SSTs, this study's main aim was simply to collect data on emotional facial expressions from as large a group of participants as possible, therefore the researchers wanted to maximise the appeal of taking part. The collected footage was viewed by a researcher and facial expressions were manually identified. The footage was divided into clips – short sequences of footage containing either a facial expression or a neutral face. Clips containing a neutral face were discarded and the remainder were analysed by the researcher in order to determine which AUs were present at peak emotional expression. Expressions exhibited as a result of interacting with a member of staff, companion or the researcher were discarded as they were not elicited by the SST device. It is possible that some AUs were exhibited by the participants but not observed by the researcher due to factors including the resolution of the recorded footage (320x240 pixels), the lighting in the scene and participant features such as hair styles, glasses, hats, scarves and facial hair.

The study carried out during stage 2 required participants to attempt a computer-based self-service shopping task. The task was similar to most online shopping experiences involving adding items to a virtual shopping basket. Twenty-six participants took part (14 males, 12 females) with an average age of 29.3 years. Ethical approval was obtained from the University of Abertay School of Computing and Engineering Systems Research Ethics Sub-Committee. The study had a within subjects design with each participant being subjected to the same conditions. Participants interacted with a computer program which required them to complete a shopping task, similar to most online shopping scenarios which utilise a “shopping basket”. Participants received on-screen instructions to add various items from a list into a virtual shopping basket. A 3D animated agent was present on the screen during the task, displaying a looped animation of emotionally neutral behaviour. Participants were asked to attempt to complete all tasks, with the option of clicking an onscreen “quit” button if they felt they could not complete a task.

In order to simulate a real-world SST interaction where an event makes it impossible for a user to achieve their goal without the assistance of a member of staff, the program purposefully obstructed the user, making the task impossible to complete. This was achieved by causing the program to add incorrect items to the shopping basket after a preset number of actions had been carried out by the user. If participants asked the researcher for assistance they were advised to click the “quit” button if they felt that they were unable to complete the task. After the quit button

was pressed, participants were presented with a screen displaying an array of seven animated faces, each displaying either anger, disgust, fear, happiness, sadness, surprise or the emotionally neutral behaviour expressed during the task. The faces were arranged randomly and animations lasted for four seconds before being repeated.

Participants used on-screen 5-point Likert scales to record how appropriate they felt each emotional expression was in the current situation. The points of each Likert scale were labelled, in order, as “completely inappropriate”, “slightly inappropriate”, “neither appropriate nor inappropriate”, “slightly appropriate” or “completely appropriate”. Participants were then presented with an on-screen questionnaire gathering details of their age and gender as well as a freeform writing section where they could suggest changes they would like to see made to the agent in order to make it more acceptable.

The emotionally expressive agent used in this study was developed by Sloan et al. (2011). Its emotional behaviours consisted of animations of the six basic universal emotions (Ekman, 1987). These animations had previously been validated (Sloan et al., 2010) to confirm that each reliably portrayed the emotion it purported to. The basic universal emotions were chosen for this study as each has a corresponding facial expression which is universally produced and recognised by people of all cultures (Ekman, 1987). Humans arguably have a much wider range of emotions (e.g. rage, frustration and contentment), however these do not have universally recognised facial expressions and were therefore unsuitable for this study.

4. RESULTS

In the stage 1 research, 30 interactions were filmed and analysed, with an average length of 302 seconds. The activation of 16 different AUs was observed in the video clips. A chi-square goodness-of-fit test was carried out to determine whether the frequencies of occurrences for each AU deviated significantly from the expected distribution. This was found to be the case ($\chi^2(15, n = 213) = 261.066, p < 0.001$), with AU 23 and AU 24 occurred at least twice as frequently as any other AU (Figure 1).

In the stage 2 study, results from agent appropriateness Likert scales were combined from 5 to 3 categories: Inappropriate, Neither Appropriate Nor Inappropriate and Appropriate (Figure 2). Chi-Square goodness-of-fit tests were carried out to assess responses to each of the animated expressions. Appropriateness ratings for the ‘disgust’ and emotionally neutral facial

expressions were found to deviate significantly from the expected distribution.

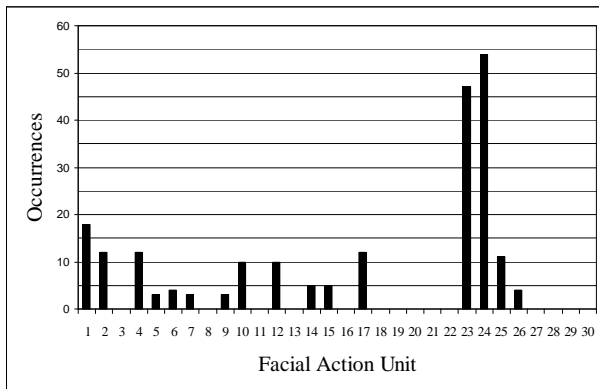


Figure1: Occurrences of AUs observed in SST users

Emotions were rated as Neither Appropriate Nor Inappropriate by a minority of participants. The majority of participants rated emotions as either Inappropriate or Appropriate. When responses for males and females were examined individually, it was found that the distribution for surprise in females varied significantly from the expected distribution ($\chi^2(2, n = 12) = 6.5, p < 0.04$), indicating that females found surprise inappropriate.

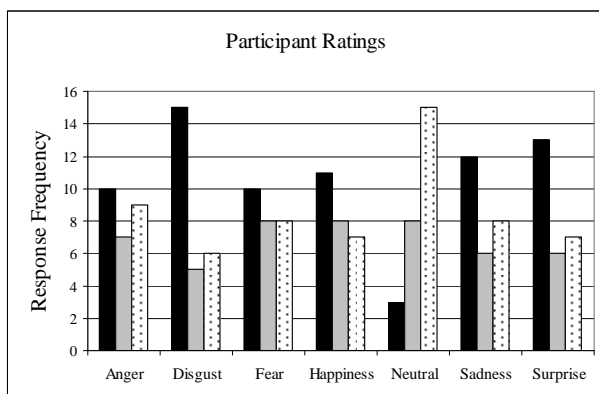


Figure2: Combined appropriateness responses to emotional expressions. Black = inappropriate, grey = neither, white = appropriate

The frequency of each rating was compared between males and females. Significant variations were found in the distributions of each rating between males and females, indicating that:

- More females rated surprise as inappropriate.
- More males rated surprise as appropriate.
- More males rated fear as inappropriate.
- More females rated fear as neither.
- More females rated happiness as neither.
- More females rated sadness as neither.

Participants made several suggestions about improvements which could be made to the agent, the majority of which related to the agent's appearance, including comments suggesting the agent should change its age and gender based on the user's age and gender (although it was not specified whether these should be changed to match or contrast the user's gender and age).

5. DISCUSSION

In stage 1, a collection of video clips was gathered showing people spontaneously expressing facial expressions during a SSI. AU 23 and AU 24 were observed with particular frequency. When displayed simultaneously, these AUs have been linked to the experience of anger (Kotsia & Pitas, 2007; Lucey et al., 2010; Rozin et al., 1999) while AU 23 alone is also associated with this emotion (Carrol & Russell, 1997 ; Schmidt et al., 2010).

We propose that the appearance of these AUs in the collected footage indicates a negative emotional state occurring in these SST users. These AUs were used to train a Haar-like feature detection system, however the performance of the system made it unsuitable for use.

The results of stage 2 indicate that, for an affective agent in a SSI, disgust was rated as an inappropriate facial expression with significantly higher frequency than expected while an emotionally neutral expression was rated as appropriate with significantly higher frequency than expected. This suggests that disgust is an unsuitable emotion to be expressed during a problematic SST interaction while an emotionally neutral expression may be preferable to other basic emotions.

It was found that females rated surprise as inappropriate with significantly higher frequency than expected. Significant differences were found between males' and females' appropriateness ratings for fear, happiness, sadness and surprise, suggesting that an agent's behaviour should vary, depending on the user's gender, in order to be perceived as appropriate. Participants also suggested that the agent should be able to change its age and gender depending on that of the user, although participants did not specify whether they would prefer an agent of their own or opposite gender or their own or a different age.

The results suggest that none of the basic emotions were considered appropriate by all participants. This could be due to participants not wanting to see more complex emotions (frustration, impatience, concern etc.), participants wanting to see less extreme manifestations of the basic

emotions or participants not wanting to be presented with any agent emotions during an SSI. Complex emotions may be difficult to implement as there is a lack of literature on reliably recognised facial expressions for these emotions.

6. CONCLUSION & FUTURE WORK

This paper describes two stages in the development of an empathetic system which will detect a negative affective state in a SST user and attempt to improve this state via empathetic feedback from an affective agent. Facial expressions of self-service checkout users were analysed and it was found that AUs commonly linked to anger were observed with particular frequency. A study was carried out to determine the appropriateness of emotional facial expressions displayed by an affective agent in response to a problematic SSI. Results indicate that disgust was perceived as inappropriate and emotionally neutral behaviour as appropriate by all participants. Gender differences were also observed, suggesting that agents may be considered more appropriate if they alter their behaviour based on the user's gender.

In order to continue development of the proposed empathetic system, the researchers will investigate appropriate emotions to be displayed in problematic SSIs, gender differences in perception of appropriate affective behaviour and the development of an acceptable affect detection system.

7. REFERENCES

- Becker, C., Prendinger, H., Ishizuka, M. Wachsmuth, I. (2005) Evaluating affective feedback of the 3D agent max in a competitive cards game. In Tao, J., Tan, T., Picard, R. (eds), *Affective Computing and Intelligent Interaction*. Springer Berlin / Heidelberg.
- Carrol, J. M., Russell, J. A. (1997) Facial Expressions in Hollywood's Portrayal of Emotion. *Journal of Personality and Social Psychology*, 72(1), pp. 164–176.
- Cho, H., Fiorito, S. S. (2010) Self-service technology in retailing. *The Case of Retail Kiosks, Symphonia. Emerging Issues in Management*, 1, pp. 43-56.
- Cramer, H., Goddijn, J., Wielinga, B., Evers, V. (2010) Effects of (in)accurate empathy and situational valence on attitudes towards robots. *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* '10), Nara, Japan, 2– 5 March, pp. 141-142. IEEE Press, Piscataway, NJ, USA.
- Ekman, P. et al. (1987) Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, 53(4), pp. 712–717.
- Essa, I. A., Pentland, A. P. (1995) Coding, Analysis, Interpretation, and Recognition of Facial Expressions. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 19(7), pp. 757–763.
- Jaksic, N., Branco, P., Stephenson, P., Encarnação, L. M. (2006) The effectiveness of social agents in reducing user frustration. *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, Montréal, Canada, 22–27 April, pp. 917-922. ACM, New York, NY, USA.
- Johnson, A. J., Pepperman K., Quist, L. (2008) The role of emotion in selfservice technologies: how can the U-Scan system be improved to positively influence users' emotions and shopping experiences. [online] Available: <http://74.125.155.132/scholar?q=cache:6zCQBXL0EjIJ:scholar.google.com/+%22self+service%22+checkout+anger&hl=en&assdt=2001>. (Accessed 15 December 2010).
- Kanade, T., Cohn, J. F., Tian, Y. (2000) Comprehensive Database for Facial Expression Analysis. *Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, Grenoble, France, 28–30 March, pp. 46-53. IEEE Press, Piscataway, NJ, USA.
- Klein, J., Moon, Y., Picard, R. W. (1999) This computer responds to user frustration. *CHI '99 Extended Abstracts on Human Factors in Computing Systems*, Pittsburgh, PA, USA, 15–20 May, pp. 242-243. ACM, New York, NY, USA.
- Kotsia, I., Pitas, I. (2007) Facial expression recognition in image sequences using geometric deformation features and support vector machines. *IEEE transactions on Image Processing*, 16(1), pp. 172–187.
- Kumano, S., Otsuka, K., Yamato, J. (2009) Pose-invariant facial expression recognition using variable-intensity templates. *International Journal of Computer Vision*, 83(2), pp. 178–194.
- Lim, M. Y., Aylett, R. (2007) Feel the difference: a guide with attitude! In Pelachaud, C., Martin, J., André, E., Chollet, G., Karpouzis, K., Pelé, D. (eds), *Intelligent Virtual Agents*. Springer Heidelberg / Berlin, Germany.

- Lin, J., Hsieh, P. (2006) The role of technology readiness in customers' perception and adoption of self-service technologies. *International Journal of Service Industry Management*, 17(5), pp. 497–517.
- Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I. (2010) The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. *Computer Vision and Pattern Recognition Workshops, 2010 IEEE Computer Society Conference on*, San Francisco, CA, USA, 13–18 June, pp. 94-101. IEEE Press, Piscataway, NJ, USA.
- Martin, C. J., Archibald, J., Ball, L. Carson, L. (in press) Towards an affective self-service agent. *Third international conference on intelligent human computer interaction*, Prague, Czech Republic, 29–31 August 2011. Springer-Verlag.
- Martin, C., Ball, L., Archibald, J., Carson, L. (unpublished) Improving self-service technology with emotionally expressive affective agents. *ASE/IEEE International Conference on Social Computing*, Amsterdam, The Netherlands, 3–6 September 2012. IEEE Press, Piscataway, NJ, USA.
- Meuter, M. L., Ostrom, A. L., Roundtree, R. I., Bitner, M. J. (2000) Selfservice technologies: understanding customer satisfaction with technology-based service encounters. *Journal of Marketing*, 64(3), pp. 50–64.
- Nass, C., Steuer, J., Tauber, E. R. (1994) Computers are social actors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Independence*, Boston, MA, USA, 24–28 April, pp. 72-78. ACM, New York.
- Pantic, M. (2005) Affective Computing. In Pagani, M. (ed), *Encyclopedia of Multimedia Technology and Networking*. Idea Group Reference, Hershey, PA.
- Pantic, M., Rothkrantz, L. J. M. (2003) Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE*, vol. 91(9), pp. 1370-1390.
- Picard, R. W. (1997) *Affective Computing*. MIT Press, Cambridge, MA, USA.
- Picard, R. W. (2001) Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Transactions On Pattern Analysis And Machine Intelligence*. 23(10) pp. 1175–1191.
- Pujari, D. (2004) Self-service with a smile? Self-service technology (SST) encounters among Canadian business-to-business. *International Journal of Service Industry Management*, 15(2), pp. 200–219.
- Rozin, P., Lowery, L., Imada, S., Haidit, J. (1999) The CAD triad hypothesis: a mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology*. 76(4), pp. 574–586.
- Satiyan, M., Nagarajan, R. (2010) Recognition of facial expression using Haar-like feature extraction method. *International Conference on Intelligent and Advanced Systems (ICIAS)*, Kuala Lumpur, Malaysia, 15-17 June, pp. 1-4. IEEE Press, Piscataway, NJ, USA.
- Salovey, P., Mayer, J. D. (1990) Emotional intelligence. *Imagination, Cognition, and Personality*. 9, pp. 185–211.
- Schmidt, R. M., Ruth, C. R., Walters, C. M., Robinson, T. B., Shiyonov, P. A., Goodyear, C. D., Funke, M. A., McIntire, L. K. (2010) Facial expressions of emotion and the assessment of performance. *Air Force Research Laboratory, 711th Human Performance Wing, Human Effectiveness Directorate, Wright-Patterson Air Force Base, OH, USA*.
- Sloan, R. J. S. (2011) Emotional avatars: choreographing emotional facial expression animation. Thesis, (PhD) University of Abertay, Dundee.
- Sloan, R. J. S., Robinson, B., Scott-Brown, K., Moore, F., Cook, M. (2010) Choreographing emotional facial expressions. *Computer Animation and Virtual Worlds*, 21, pp. 203-213.
- Tian, Y., Kanade, T., Cohn, J. F. (2001) Recognizing action units for facial expression analysis. *IEEE Transactions On Pattern Analysis And Machine Intelligence*, 23(2), pp. 97–115.
- Valstar, M., Pantic, M. (2006) Fully automatic facial action unit detection and temporal analysis. *Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, New York, NY, USA, 17-22 June, pp. 149-156. IEEE Piscataway, NJ, USA.
- Wang, C., Chen, G., Liu, C., Liu, B. (2009) Design an empathic virtual human to encourage and persuade learners in e-learning systems. *Proceedings of the First ACM International Workshop on Multimedia Technologies for Distance Learning*, Beijing, China, 23 October, pp. 27-32. ACM, New York, NY, USA.