Proceedings of the 2nd International Conference on Human-Computer Interaction
Prague, Czech Republic, August 14-15, 2014
Paper No. 136

Emotion-enhanced Mobile Chat – Conveying Emotions in Mobile Text Chat

Jackson Feijó, Thiago Valle, Wilson Prata
Nokia Technology Institute
Av. Torquato Tapajós, Manaus, Brazil
jackson.feijo@indt.org.br; thiago.valle@indt.org.br; wilson.prata@indt.org.br

Abstract **-** Great part of human communication is carried out nonverbally. All this information is lost in mobile text messaging. This work describes an attempt to augment text chatting in mobile phones by adding automatically detected facial expression reactions, such as a smile, to conversations. These expressions are detected using known image processing techniques. Known related work, concerning the investigation of non-verbal communication through text messaging are considered and distinguished from the present solution. The conception and implementation of a mobile phone application with the debated feature is described and user studies are narrated. Finally, context of application, conclusions and future works are also discussed.

Keywords**:** Affective user interfaces, Multimodal integration, Non-verbal communication, Dialogue acts.

1. Introduction

In early 1870s, scientific writings on nonverbal communication and behaviour were published in Charles Darwin's book The Expression of the Emotions in Man and Animals (Pease, 2004). Darwin argued that all mammals, both humans and animals, showed emotion through facial expressions. Stating questions such as: "Why do our facial expressions of emotions take the particular forms they do?" and "Why do we wrinkle our nose when we are disgusted and bare our teeth when we are enraged?" (Krauss, 2000)

Humans rely on visual cues when interacting within a social context, and in an online chatting scenario, the lack of sight due to physical remoteness interferes with the quality of social interactions. Since crucial non-verbal communication is comprised of facial gestures and expressions, text chat users are at a disadvantage when engaging in interpersonal communication

2. Problem Space

Non-verbal information conveyed in a conversation is often decisive for disambiguation. In (Neustaedter et al. 2012), long-distance relationships and video chat technology is well debated. In one of the interviews from this work, some subjects have described video as a way to see their partners' facial expressions and body language. In some cases, this helped avoid miscommunications.

"I always apparently sound pretty harsh when I'm talking …even when I'm joking it doesn't sound like I'm joking…I would sometimes upset her without even knowing I upset her and, of course, without intending." Quote from test subject interview in (Neustaedter et al. 2012)

3. Related Work

There has been significant work on the recognition and conveying of emotions in computer mediated communications.

3. 1. Video Chat

As minutiously described in (Neustaedter et al. 2012), the use of video chat successfully transmits consistent non-verbal elements in conversations. In this work, we will not consider video chat solutions as

related work, since we are trying to intervene in online text chat scenarios. This distinction is relevant due to the fact that video chat usually requires higher performance systems with more expensive hardware, faster networks, more processing power (compared to lower-end platforms with minimum requirements to exchange text through a server). Not only structurally, video-chatting constitutes a different social ritual (Neustaedter et al. 2012). In addition, video-chat may often convey extra imagery such as background rooms, make-up, clothing and hair style, as opposed to transmitting only processed facial expression, e.g. smile.

3. 2. Conveying Emotion in Text Chat

Several existing approaches have been trying to derive affect from written language on a semantic basis, making use of common sense knowledge (Bosma et al, 2004). The work in (Barros et al. 2012) is an example of approaches that automatically identify emotions in texts.

The work in (Derrick, 2013) attempts to detect deceptive chat-based communication using typing behaviour and message cues. It is an example of deriving non-verbal information in chat texts with a non-semantics approach. It hypothesizes that deception influences response time, word count, lexical diversity, and the number of times a chat message is edited. It focuses on detecting deception, unlike our work, which is attempting to convey emotions.

The works in (Bosma et al, 2004) and (Wang et al., 2004) distinguish themselves from the others because they attempt to integrate natural language and emotions by performing the recognition of these emotions based on physiological user input. They aim at collecting additional information on the user's emotional state by recording and analysing physiological feedback. Therefore, while interacting with the application, the user is monitored by means of bio sensors, measuring skin conductivity, heart rate, respiration and muscle activity (Bosma et al, 2004). Specifically in (Wang et al., 2004) alterations in the display of the text are used to communicate emotions (animated text). Our work differs from these because it relies solely on low-resolution mobile phones front cameras, not requiring extra hardware. Also, it restricts the recognition of emotions to processing images of faces through mobile phones.

The accessibility work in (Astler et al. 2011) approaches the present issue by building a computer vision system with facial recognition and expression algorithms to relay nonverbal messages to blind users. This project differs from our work by focusing on visually impaired audience. In addition, unlike our solution, robust processing power is required for demanding algorithms if compared to our software. Finally it is relevant to mention that it relies on extra hardware, for example, using a high speed IEEE-1394b camera.

4. Emotion Enhanced Mobile Chat

Our solution attempts to intervene in mobile online text chat by detecting users' emotional reactions during the conversations. Initially, we implemented the chat application itself to be integrated with the emotion detection feature.

4.1. Mobile Chat Application

The mobile chat application was developed for Windows Phone, using the XMPP protocol (Web 1) (formerly and also known as Jabber) as not only it is defined in open standards but it is also widely deployed across the Internet.

This standard is used by very popular chat platforms like Google Talk, Messenger and Facebook Chat, which made it easier for test subjects to use our solution using their own everyday accounts.

4.2. Mobile Emotion Detection Software

In this work, we are trying to detect emotions only when new messages have arrived, in order to increase the probability of a facial expression to be a reaction to the most recent message. This means that real-time emotions detection performed by demanding algorithms is not necessary.

Due to the discrete nature of this detection scenario, the image processing can be executed on the server side. This is important because lower end devices, with low memory and low processing power can be used with good performance.

In, (Krishna et al., 2005). uses a Principal Component Analysis (PCA) algorithm for facial recognition. Our system uses a commercial solution (Rekognition; Orbeus, Inc.)(Web 2) for face and emotion detection.

4.3. System Structure

Figure 1 shows a simple example of a use case of the application.

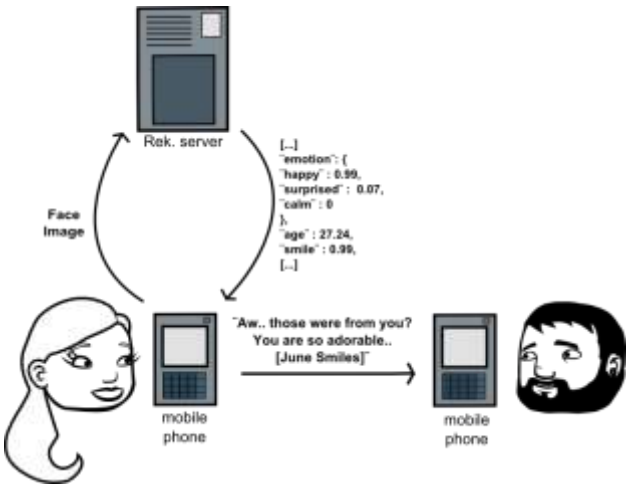

Fig. 1. System structure.

As shown in Figure 2, the user chats from the mobile, using our application. As soon as a conversation starts, the face images start being collected, converted to base64, and sent via http to the server to be processed. The server sends back a JSON feed, which is parsed to retrieve emotions status. The emotions results range from 0 to 1, providing the detected level of "happy", "surprised" and "calm". If the level one of these emotions are above a pre-determined threshold, the software automatically adds "<name-of-user> smiles" or "<name-of-user> is surprised", to the conversation thread.

4.4. Performance Analysis

We have performed a few tests concerning the position of the phone in relation to the users. We tested front and front-low. This second, more likely to happen when users interact with mobile devices. In Table 1 we can see both positions work well for the emotion detection process we are using.

Face Image	Capturing Position – Front or Front-low	Smile result (0.0 – 1.0)
	Front-low	0.53

	Front-low	1
	Front	0.69
	Front	0.25
	Front	1

4.5. Screenshot

Figure 3 is a screenshot of the application, with an on-going dialogue. It is relevant to mention that no actual face images are exchanged. This consolidates the social ritual of the text dialoguing, when people are able to have conversations unmindful of make-up, hair and overall facial aesthetics status. There is also a clear demonstration of when the "emotions feature" is on or off.

5. Interviews and Experiments

Initially, we observed that the test subjects tended to change their facial expressions for no particular reason, in order to check with their chat partners whether the emotion detection accurately worked. Because of that, we decided to implement a preliminary educational usability test phase, in order to get test subjects habituated with the software.

For the sake of productivity, we recruited test subjects, chat partners, who are used to talk to each other using mobile text chat.

We referred to some inquiries in the work by (Wang et al., 2004) when developing the questionnaire. The subjects were asked to answer the questions after using both modes of chat. See Table 2.

Question	EF	Non-EF
I became very involved in the conversation	4.6	3.7
I enjoyed using the system	4.8	3.0
It was easy to understand my partner's emotions	4.0	2.7
I emulated some of my expressions	1.2	0
I felt more connected to my chat partner	4.9	2.3
I would use this application in my everyday chat conversations.	3.8	3.2

6. Conclusions and Future Work

We developed an approach to the integration of emotions to mobile text chat. The motivation is to enrich mobile text chat by conveying frequently lost non-verbal communication. This benefits users by relay information that is valuable to disambiguate dialogues.

Our approach differed from the related research for it did not attempt to extract affection from the semantics of the text and did not use any extra hardware to collect images or physiological signals. Finally, our work is devoted to relaying emotions on mobile text chat scenarios, relying on lower end devices with front cameras.

We implemented a mobile text chat application based on the XMPP protocol, easily connectable with popular chat platforms, with the feature of detecting users' emotions automatically and transmitting them to chat partners.

We also tested the detection of the facial expressions in two different positioning of the mobile phone front camera – directly in front of the user and in a lower front. The system provided good results in both positions

During the experiments and interviews we found that people perceived their partners' reactions as an accomplishment or a reward to what they just posted. This has included some sort of pressure in the engagement of the conversation, in comparison to plain text chat. In some cases, subjects have mentioned that their partners inserted more humour then usual to the dialogue, in order to receive the automatically detected smiles. During the preliminary trials, we found that people tended to try to foresee their chat partner's reactions, creating frustration when no smiles when detected or causing great sense of accomplishment when it did.

From the interviews, which were conducted individually and in groups, we were able to observe the rise of the sense of connection that chat partners experienced with one another.

In future versions of this work, we wish to investigate the performance of the automatic conveying of emotions in mobile text chat when users are on the move. We wish to investigate how to image capturing and processing will work when users are walking, standing on moving subway cars, etc.

We are also presently executing the integration of synchronization algorithms to optimize the posting of messages, reading speed, facial reactions and emotion detection, to be published in later papers.

References

Argyle, M. Salter, U. Nicholson, H. Williams, M.Burgess, P. The communication of inferior and superior attitudes by verbal and non-verbal signals. British Journal of Social Psychology, 9:222–231, 1970.

Astler, D., Chau, H., Increased Accessibility to Nonverbal Communication through Facial and Expression Recognition Technologies for Blind/Visually Impaired Subjects. ASSETS '11 The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility. Pages 259-260. ACM New York, NY, USA 2011.

Barros, L., Rodríguez, P., Ortigosa, A. Emotion Recognition in Texts for User Model Augmenting. INTERACCION '12 Proceedings of the 13th International Conference on Interacción Persona-Ordenador. Article No. 45. ACM New York, NY, USA 2012.

Bosma, W., Andre, E. Exploiting emotions to disambiguate dialogue acts. IUI '04 Proceedings of the 9th international conference on Intelligent user interfaces. Pages 85-92. ACM New York, NY, USA 2004.

Derrick, Douglas C., Meservy, Thomas O., Jenkins, Jeffrey L., Burgoon, Judee K., Nunamaker,Jay F., Jr. Detecting Deceptive Chat-Based Communication Using Typing Behavior and Message Cues. ACM Transactions on Management Information Systems (TMIS) TMIS Homepage archive. Volume 4 Issue 2, August 2013. Article No. 9. ACM New York, NY, USA.

H. Liu, H . Lieberman, and T. Selker. A model of textual affect sensing using real-world knowledge. In W. L. Johnson, E. Andre, and J. Domingue, editors, International Conference on Intelligent User Interfaces, pages 125–132, Miami, Florida, USA, January 2003.

Krauss, R .M., Chen, Y., and Chawla, P. (2000). Nonverbal behavior and nonverbal communication: What do conversational hand gestures tell us? Advances in Experimental Social Psychology, 1(2), 389-450.

Little, G. Black, J. Panchanathan, S. Krishna, S. A. wearable face recognition system for individuals with visual impairments. In 7th Annual ACM Conference on Assistive Technologies, Pages 106–113, 2005.

Neustaedter,C., Greenberg,S. Intimacy in Long-Distance Relationships over Video Chat. CHI '12 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Pages 753-762. ACM New York, NY, USA 2012

Pease B., Pease A. (2004). The Definitive Book of Body Language. New York, NY: Bantam Books.

Power, Mary R., "Chapter 11: Non-verbal communication" (1998). Working Through Communication.Paper 12.

Wang, H., Predinger, H., Igarashi, T. Communicating Emotions in Online Chat Using Physiological Sensors and Animated Text. CHI EA '04 CHI '04 Extended Abstracts on Human Factors in Computing Systems. Pages 1171-1174. ACM New York, NY, USA, 2004.

Web sites:

Web-1: XMPP Standards Foundation, http://www.xmpp.org consulted 20 Mar. 2014.

Web-2: Rekognition APIs and SDK, https://rekognition.com consulted 20 Mar. 2014.