

From Cartoons to Robots: Facial Regions as Cues to Recognize Emotions

Tomoko Koda, Yuka Nakagawa, Kyota Tabuchi
Faculty of Information Science and Technology
Osaka Institute of Technology
Osaka, Japan
koda@is.oit.ac.jp

Zsafia Ruttkay
Moholy-Nagy University of Art and Design
Budapest, Hungary
ruttkay@mome.hu

Abstract—This paper introduces a preliminary result of a cross-cultural study on the facial regions as cues to recognize virtual agents’ facial expressions. We believe providing research results on the perception of cartoonish virtual agents’ facial expressions to HRI research community is meaningful in order to minimize the effort to develop robot’s facial expressions. The result implies 1) the mouth region is more effective in conveying the emotions of the facial expressions than the eye region, 2) there are cultural differences in using facial regions as cues to recognize cartoonish facial expressions between Hungary and Japan.

Keywords: *facial expression, cross-culture*

I. INTRODUCTION

Humanoid robots are increasing their importance and attention both in our everyday lives and in HRI research. Non-verbal behaviors, especially facial expressions, of humanoid robots take an important role in better human-robot interaction as they do in human-human interaction [1, 2]; however, creating facial expressions of robots requires complex mechanical movements. Most current humanoid robots have either mechanical or cartoonish faces with limited expressions [1, 2, 3] when their research focus is not on increasing realism of a humanoid robot [4] but on improving human-robot interaction. We believe providing research results on the perception of cartoonish virtual agents’ facial expressions to HRI research community is meaningful in order to minimize the effort to develop robot’s facial expressions. In this paper, we introduce an early result of a cross-cultural study on the facial regions as cues to recognize virtual agents’ facial expressions.

Ruttkay addresses the importance of designing facial expressions for virtual agents with a specific culture [5]. Cross-cultural study on recognition of avatar facial expressions drawn by Asian and Western designers suggested there are cultural differences in recognizing avatar’s facial expressions [6]. Yuki et al.’s study using emoticons and photorealistic human facial images suggests Americans tend to interpret emotions based on the mouth, while Japanese tend to focus on the eyes [7]. Research on human eye movements to interpret photo realistic human facial expressions shows East Asian participants in the study focused mostly on the eyes, but those from the West scanned the whole face [8].

Inspired by [6, 7], this study uses cartoonish avatar faces to find cultural differences in using eyes and mouth as cues to recognize facial expressions between Hungary and Japan.

II. EXPERIMENT

A. Design of Facial Expressions

Japanese designers designed three avatar faces, each with three facial expressions, namely, neutral, happy, and sad, by using CharToon [9], a design and animation tool for 2D cartoon faces. The three facial expressions followed FACS [10], and were pre-evaluated to make sure they convey the intended emotions (higher than 90% recognition accuracy). We then created 6 static expressions per avatar design by combining the eyes and mouths. The six combinations are “happy eyes and neutral mouth”, “happy eyes and sad mouth”, “sad eyes and happy mouth”, “sad eyes and neutral mouth”, “neutral eyes and happy mouth”, and “neutral eyes and sad mouth”. The total number of combined facial expressions is 18 (6 expressions x 3 avatars). Figure 1 shows examples of “sad eyes and happy mouth” and “happy eyes and sad mouth” expressions.

Next, we created 18 animations that start from the neutral expression of each avatar design and end with one of the combined expressions. Each animation lasts for 4 seconds, one second for the neutral expression, two seconds for the transition, and one second for the final expression.

B. Experiment Procedure

The experiment is being conducted on the web since December 2009. The 18 animations are shown using Adobe Flash Player. Subjects were asked to watch each animation that is shown randomly, and evaluate perceived emotions of the final expression of each animation using a 6-point Likert scale (6: very happy - 1: very sad).

Fifty-four Japanese and 21 Hungarian subjects participated in the experiment by invitation. The subjects reside in their native country and don’t have experiences in living in Western countries (for Japanese) or Asian countries (for Hungarian).



Figure 1. Examples of combined facial expression (left: “sad eyes and happy mouth”, right: “happy eyes and sad mouth”)

III. PRELIMINARY RESULTS AND DISCUSSION

The subjects’ answers of perceived emotions (6 point happy-sad scale) are shown in Figure 2. The combined expressions are categorized into four categories, “Happy_Eyes (HE)”, “Sad_Eyes (SE)”, “Happy_Mouth (HM)”, and “Sad_Mouth (SM)” by focusing on the shape of either the eye region or the mouth region. For example, the HE category includes “happy eyes and neutral mouth” and “happy eyes and sad mouth” expressions.

Firstly, we focus on the answers by facial categories (HE, SE, HM, and SM) regardless of the country. As shown in Figure 2, the HE category rated as happiest and the SM category as saddest in any other categories by both countries. The eye region (HE and SE) did not convey the intended emotions as effectively as the mouth region (HM and SM) did.

Secondly, we focus on the answers between Hungary and Japan. Japanese rated the HE category as significantly happier than Hungarians regardless of the shape of the mouth (Hungary=2.60 (n=21), Japan=2.99 (n=54), $p<0.01$). While Hungarians rated the SM category as significantly sadder than Japanese regardless of the shape of the eyes (Hungary=2.40, Japan=2.62, $p<0.01$). There were no significant differences in the rating of the SE category (Hungary=3.51, Japan=3.31, $p=0.11$), and the HM category (Hungary=4.76, Japan=4.87, $p=0.17$).

In conclusion, the above results imply overall tendencies that the mouth region is more effective in conveying the emotions of the facial expressions than the eye region in the cartoonish faces used in this experiment. Cultural differences are found in expressions with “happy eyes” and “sad mouth” in the same direction as indicated in [7, 8]. The Japanese subjects reacted more dynamically toward happiness to the shapes of the eyes in “happy eyes”, while Hungarians toward sadness on the shapes of the mouth in “sad mouth”. However, data gathering and analysis of the web experiment is still in progress. Thus, we would like to hold our conclusion until we complete our data analysis with more balanced number of subjects from Hungary.

In addition to cultural differences, the result suggests important design issues. The current result that “happy eyes” are rated as “less happier” than “sad eyes” indicates that the “happy eyes” should be more carefully designed. Although perceived emotions of the “happy eyes” using static images of

the eye region in the pre-evaluation test had 90% accuracy, we should be more careful in designing eyes when they are used solely apart from other facial regions. Further study should use wide variety of designs from both countries.

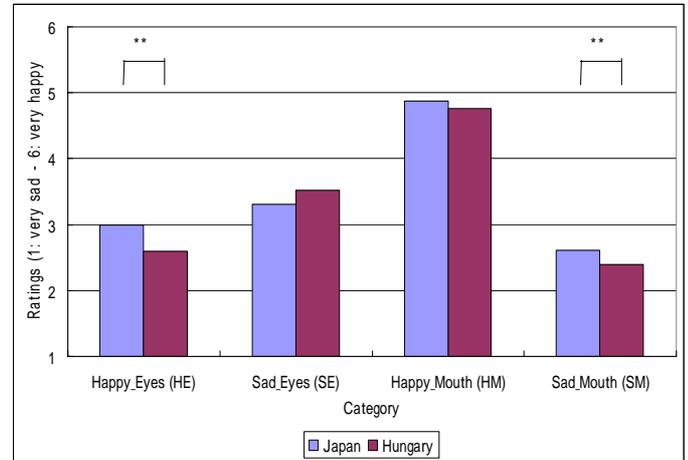


Figure 2. Perceived emotions of combined expressions (The scale indicates 1: very sad to 6: very happy. Number of subjects: Japan (n=54), Hungary (n=21). ** indicates $p<0.01$)

ACKNOWLEDGMENT

This research is supported by a Grant-in-Aid for Scientific Research (C) 20500196 (2008-2010) from the Japan Society for the Promotion of Science. CharToon was developed by Zsolt Ruttkay and Han Noot.

REFERENCES

- [1] C. Breazeal, “Emotion and sociable humanoid robots”, *International Journal of Human-Computer Studies*, Vol. 59(1-2), pp. 119-155, July 2003.
- [2] C. Breazeal, “Sociable Robots.” *Journal of the Robotics Society of Japan*, Vol. 24(5), pp. 591-593, 2007.
- [3] H. Ishiguro, T. Ono, M. Imai, T. Maeda, T. Kanda, R. Nakatsu, “Industrial Robot”, *International Journal*, Vol. 28(6), pp. 498-504, 2001.
- [4] H. Ishiguro, S. Nishio, “Building artificial humans for understanding humans”, *Journal of Artificial Organs*, vol.10(3), pp. 133-142, 2007.
- [5] R. Ruttkay, “Cultural dialects of real and synthetic emotional facial expressions”, *Journal of AI & Society*, Vol.24(3), Springer London, pp. 307-315, 2009.
- [6] T. Koda, T. Ishida, M. Rehm, E. Andre, “Avatar culture: cross-cultural evaluations of avatar facial expressions”, *Journal of AI & Society*, Vol.24(3), Springer London, pp.237-250, 2009.
- [7] M. Yuki, W.W. Maddux, T. Masuda, “Are the windows to the soul the same in the East and West?”, *Journal of Experimental Social Psychology*, Vol. 43, pp. 30-311, 2007.
- [8] R. Jack, Blais, E. C. Scheepers, P. G. Schyns, R. Caldara, “Cultural Confusions Show Facial Expressions are Not Universal”, *Current Biology* (in press).
- [9] Z. Ruttkay, H. Noot, “Animated CharToon faces”, 1st international symposium on Non-photorealistic animation and rendering, pp. 91-100, 2000.
- [10] P. Ekman, W.V. Friesen, “Facial action coding system: A technique for the measurement of facial movement”, Consulting Psychologists Press, Palo Alto, 1978.