

# From Cartoons to Robots Part 2: Facial Regions as Cues to Recognize Emotions

Tomoko Koda, Tomoharu Sano  
Faculty of Information Science and Technology  
Osaka Institute of Technology  
1-79-1 Kitayama, Hirakata city  
Osaka, 573-0196 Japan  
81-072-866-5182  
koda@is.oit.ac.jp

Zsafia Ruttkay  
Media Institute Creative Technology Lab  
Moholy-Nagy University of Art and Design  
1121 Budapest, Zugligeti ut 9-25  
Hungary  
36-20-253-8697  
ruttkay@mome.hu

## ABSTRACT

This paper reports 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 Japanese weighed facial cues more heavily in the eye regions than Hungarians, who weighed facial cues more heavily in the mouth region than Japanese.

## Categories and Subject Descriptors

H.5.1 [Multimedia Information Systems] Animations, Evaluations/methodology

## General Terms

Human Factors

## Keywords

facial expression, cross-culture

## 1. 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.

Recent psychological research shows evidences that Asians tend to focus on the eyes and Westerns on the mouth when recognizing real human facial expressions [5, 6]. In order to apply this finding to the case of virtual avatars, we conducted a web-based survey to confirm the following hypothesis. In cartoonish faces, Japanese weigh facial cues in the eye regions more heavily than Hungarians, who weigh facial cues in the mouth region more heavily than Japanese. We conducted the first experiment using Japanese designed avatars in 2009. The result confirmed the hypothesis [7]. This paper reports our second experiment using Hungarian designed avatars.

## 2. EXPERIMENT

### 2.1 Design of Facial Expressions

Hungarian designers designed five avatar faces, each with three facial expressions, namely, neutral, happy, and sad, by using CharToon [8], a design and animation tool for 2D cartoon faces. The three facial expressions followed Ekman's FACS [9], 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 30 (6 expressions x 5 avatars). Figure 1 shows examples of "happy eyes and sad mouth" and "sad eyes and happy mouth" expressions.

Next, we created 30 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.

### 2.2 Experiment Procedure

The experiment was conducted on the web from November to December 2010. The 30 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 Japanese and 41 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).

### 3. RESULTS AND DISCUSSION

The subjects' answers of perceived emotions (6 point happy-sad scale) are shown in Figure 2. The 30 combined expressions are categorized into six categories, "happy eyes and neutral mouth (HN)", "happy eyes and sad mouth (HS)", "neutral eyes and happy mouth (NH)", "neutral eyes and sad mouth (NS)", "sad eyes and happy mouth (SH)", "sad eyes and neutral mouth (SN)" by focusing on the shape of the eye and mouth region.

Analysis by two-way ANOVA indicates significant cultural differences in the perceived emotions between Japan and Hungary in the all combined expressions, namely, HN (Japan=3.41 (n=50), Hungary=3.22 (n=41),  $F=6.88$ ,  $p<0.01$ ), HS (Japan=2.59 (n=50), Hungary=2.09 (n=41),  $F=47.20$ ,  $p<0.01$ ), NH (Japan=4.86 (n=50), Hungary=4.62 (n=41),  $F=8.67$ ,  $p<0.01$ ), NS (Japan=2.65 (n=50), Hungary=2.12 (n=41),  $F=68.90$ ,  $p<0.01$ ), SH (Japan=3.78 (n=50), Hungary=4.23 (n=41),  $F=19.89$ ,  $p<0.01$ ), SN (Japan=2.28 (n=50), Hungary=2.59 (n=41),  $F=17.09$ ,  $p<0.01$ ).

Japanese rated the perceived emotions of HN and HS, where the eyes convey happiness, as significantly happier than Hungarians did, and the perceived emotions of SH and SN, where the eyes convey sadness, as significantly sadder than Hungarians did. On the contrary, Hungarians rated the perceived emotions of SH, where the mouth conveys happiness, as significantly happier than Japanese did, and the perceived emotion of HS and NS, where the mouth conveys sadness, as significantly sadder than Japanese. NH expression had a significant difference in the opposite direction from the hypothesis.

In addition to cultural differences, the result suggests important design issues. The expressions with happy eyes (HN and HS) are rated as sadder than the expression with a happy mouth (NH and SH). This indicates that regardless of the country, the mouth region more effectively conveys the intended emotions than the eye region. 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.

We conducted two experiments between Hungary and Japan, the first experiment used Japanese designed expressions, and the second used Hungarian designed expressions. The results of the second experiment showed that Japanese weighed facial cues in the eye regions more heavily than Hungarians, who weighed facial cues in the mouth region more heavily than Japanese in the case of avatar expressions designed by Hungarians.



Figure 1. Examples of combined facial expressions (left: "happy eyes and sad mouth", middle: "sad eyes and happy mouth", right: "sad eyes and sad mouth")

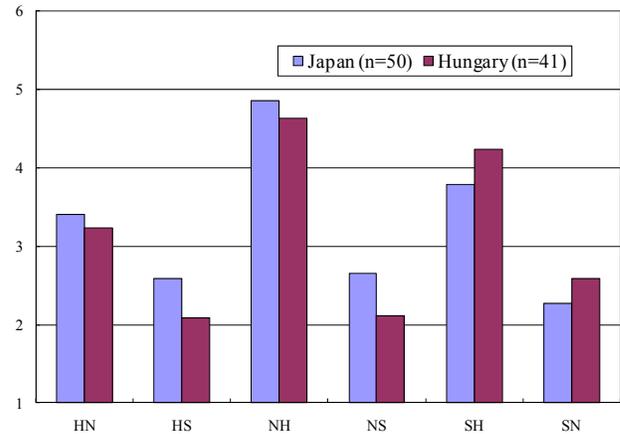


Figure 2. Perceived emotions of combined expressions shown by country (The scale indicates 1: very sad to 6: very happy. All Japan-Hungary pairs :  $p<0.01$ )

### 4. ACKNOWLEDGMENTS

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.

### 5. REFERENCES

- [1] C. Breazeal, "Emotion and sociable humanoid robots", *International Journal of Human-Computer Studies*, Vol. 59(1-2), pp. 119-155, 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] 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.
- [6] R. Jack, Blais, E, C. Scheepers, P. G. Schyns, R. Caldara, "Cultural Confusions Show Facial Expressions are Not Universal", *Current Biology*, 19 (18), pp. 1543—1548, 2009.
- [7] Koda, T., Ruttkay, Z., Nakagawa, Y., Tabuchi. K. Cross-cultural Study on Facial Regions as Cues to Recognize Emotions of Virtual Agents. *Culture and Computing*, LNCS Vol. 6259, pp.16-27, Springer-verlag, 2010.
- [8] Z. Ruttkay, H. Noot, "Animated CharToon faces", 1st international symposium on non-photorealistic animation and rendering, pp. 91-100, 2000.
- [9] P. Ekman, W.V. Friesen, "Facial action coding system: A technique for the measurement of facial movement", Consulting Psychologists Press, Palo Alto, 1978.