

# CROSSMODAL PERCEPTION OF MISMATCHED EMOTIONAL EXPRESSIONS BY EMBODIED AGENTS

Yu Suk Cho, Ji He Suk and Kwang Hee Han

*Cognitive Engineering Lab, Department of Psychology, Yonsei University, Seoul, Korea*

**Keywords:** Embodied agents, Mismatched emotion, Emotional information.

**Abstract:** Today an embodied agent generates a large amount of interest because of its vital role for human-human interactions and human-computer interactions in virtual world. A number of researchers have found that we can recognize and distinguish between emotions expressed by an embodied agent. In addition many studies found that we respond to simulated emotions in a similar way to human emotion. This study investigates interpretation of mismatched emotions expressed by an embodied agent (e.g. a happy face with a sad voice). The study employed a 4 (visual: happy, sad, warm, cold) X 4 (audio: happy, sad, warm, cold) within-subjects repeated measure design. The results suggest that people perceive emotions not depending on just one channel but depending on both channels. Additionally facial expression (happy face vs. sad face) makes a difference in influence of two channels; Audio channel has more influence in interpretation of emotions when facial expression is happy. People were able to feel other emotion which was not expressed by face or voice from mismatched emotional expressions, so there is a possibility that we may express various and delicate emotions with embodied agent by using only several kinds of emotions.

## 1 INTRODUCTION

Expression and perception of emotions are very important in human-human interactions. A number of studies indicated that facial expressions of basic emotions can be recognized as the same meaning (Ekman, 1990). Emotion is vital for human-human interactions and human-computer interactions in virtual world. A recent study found that we can recognize and distinguish between emotions expressed by an embodied agent with similar accuracy to those of human emotional expressions (Bartneck, 2001). In addition we respond to simulated emotions in a similar way to human emotion (Brave et al., 2005). However there is a few researches focused on emotions of agent.

De Gelder and Vroomen (2000) have investigated how the combination of visual and audio elements of emotional expressions influence our perceptions. In this study despite participants being told to ignore the voice, the audio element of the emotional expression still influenced participants' perceptions. This result means crossmodal perception in visual and audio channel. Creed and Beale (2008) investigated crossmodal perception of emotional expressions by agent.

This study focused on perception of mismatched emotional expressions by the agent.

H1: Participants will perceive mismatched emotional expressions not depending on just one channel but depending on both visual and audio channels.

H2: Participants will perceive another emotion from mismatched emotional expressions; emotion expressed neither by face nor by voice.

## 2 METHOD

### 2.1 Participants

Participants were 45 (26 male and 19 female) undergraduate students enrolled in a psychology class.

### 2.2 Design and Measures

The study employed 4 (visual: happy, sad, warm, cold) X 4 (audio: happy, sad, warm, cold) within-subjects repeated measure design. The first factor was facial expressions of the agent (happy, sad, warm, cold) and the second factor was audio

expressions of the agent (happy, sad, warm, cold). The study was comprised of sixteen different conditions and participants were asked to rate the emotional expression in each condition using two seven-point semantic differential scales (happy-sad, warm-cold).

### 2.3 Materials

We made four facial animations that express four different emotions (happy, sad, warm, cold) with 3D MAYA and recorded female human voice in four moods (happy, sad, warm, cold). When developing the different emotional facial expressions, we used Ekman's (2004) study about expression and perception of emotions. For example, happy facial expression involved opening the mouth and raising the lip corners. We selected the sentence that was short and could be said in a variety of emotional voices for recording; 'It's eleven o'clock.' Facial animation and voice files were integrated using Flash CS3 Professional. These integrated animations included all possible combinations between the visual and audio expressions of emotions; four matched combinations ([happyface-happyvoice], [sadface-sadvoice], [warmface-warmvoice], [coldface-coldvoice]) and twelve mismatched combinations (e.g. [happyface-sadvoice]).

A pilot study was conducted to validate the agent's four emotional expressions and four emotional voices (see fig. 1). Participants included 12 undergraduate students. Participants viewed facial animations of the agent without the voice and evaluated the emotional expression using four five-point likert scales (happy, sad, warm, cold). In addition they evaluated the emotional expression of voices recorded by 2 females and 1 male.

The results of the pilot study indicated that participants perceived the emotional expression of agent as intended (happy face (happy  $M=4.5$ ), sad face (sad  $M=4.0$ ), warm face (warm  $M=3.7$ ), cold face (cold  $M=4.0$ )). As for the voice, we selected the voice perceived appropriately for almost emotions ( $M \geq 3.0$ ).

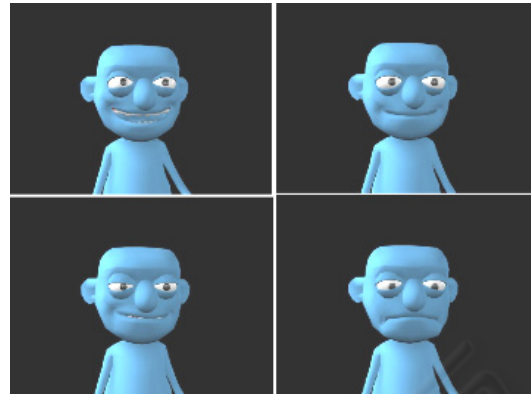


Figure 1: facial emotional expressions of agent (clockwise from top-left: happy, warm, sad, cold).

### 2.4 Procedure

Participants were initially asked to provide some basic demographic information about themselves including their age, gender. Agent was introduced to participants for friendly and comfortable feeling. Participants rated the emotional expression in each condition using four seven-point semantic differential scales (happiness-sadness, warmth-coldness, arousal-relax, pleasure-displeasure) and answered about an open question (how does the agent feel?). During the experiment, participants could see the animations as many times as they wanted.

## 3 RESULTS

### 3.1 Happiness-sadness

There was a significant main effect for the emotional faces and the emotional voices in animations with respect to happiness-sadness,  $F(3, 132) = 184.618$ ,  $p < 0.001$ ;  $F(2.305, 101.399) = 69.663$ ,  $p < 0.001$ . There was also a significant interaction between faces and voices  $F(9, 396) = 3.572$ ,  $p < 0.001$ . The perceived intensity of happiness has been influenced not by only one of two dimensions (a face and voice) but by combination of them. As ranging from 4.9 to 6.9, happy face has been more influenced by the voices than sad face. Warm and cold faces have no differences in happiness-sadness scale. However, It is noted that cold face with happy voice tends to be perceived as happy even though coldness and happiness are in different dimension of emotions.

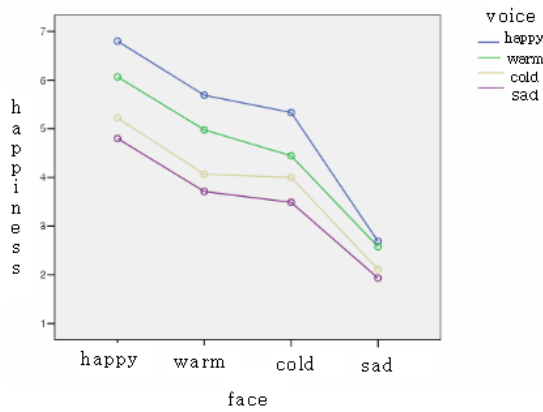


Figure 2: The perceived intensity of happiness.

### 3.2 Warmth-coldness

There was a significant main effect for the emotional faces and the emotional voices in animations with respect to warmth-coldness,  $F(3, 132) = 62.907, p < 0.001$ ;  $F(2.431, 106.974) = 28.972, p < 0.001$ . Warm face is rated as most warm whereas cold face as most cold. It is the same result as voice.

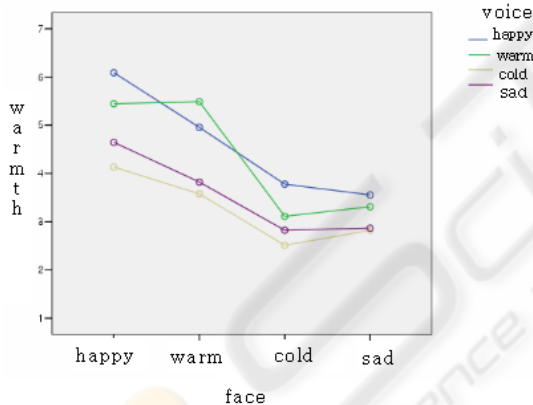


Figure 3: The perceived intensity of warmth.

There was also a significant interaction between faces and voices,  $F(6.590, 289.955) = 5.030, p < 0.001$ . The perceived intensity of warmth has been influenced not by only one of two dimensions (faces and voices) but by combination of them. Warm face has been more influenced by the voices than cold face.

### 3.3 Arousal, Pleasure and Open Question

We analyzed open question to examine the possibility that mismatched emotion of voice and

face may evoke another emotion; neither emotion expressed by face nor emotion expressed by voice. Many participants reported they felt ‘disappointed’ in sadface-warmvoice and sadface-coldvoice, ‘fatigued, bored’ in coldface-sadvoice and warmface-sadvoice, and ‘happy but pretending sad’ in happyface-sadvoice and happyface-coldvoice. We conducted MDS (multi dimensional scaling) to examine these results simply and graphically. MDS Stress and RSQ were 0.12, 0.927.

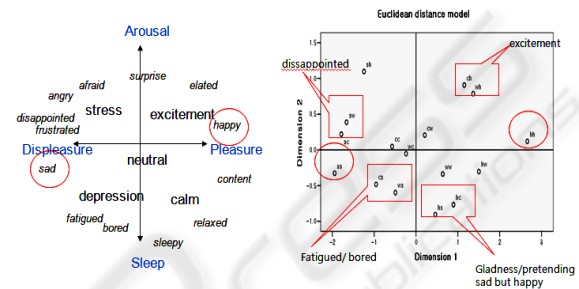


Figure 4: Circumplex model of affect and MDS graph.

In figure4, the left graph is circumplex model of affect (Russell, 1980) and the right one is our 16 conditions expressed by MDS. Dimension1 can be explained with pleasure-displeasure and dimension2 can be explained with arousal-relax. Two graphs are very alike. The position of happyface-happyvoice condition is similar to the position of happiness in Russell’s circumplex model of affect and sadface-sadvoice is matched for sad in Russell’s circumplex model of affect.

However, we can’t confirm other conditions’ emotion with comparing two graphs, so we display corresponding open question results to the right graph. In spite of that the position of each condition is the result of seven-point semantic differential scales and emotion at the square is the result of open question, we can notice that each emotion is very similar to Russell’s emotion when comparing positions in each graph. Mismatched emotional expressions are not perceived queerly but may be interpreted another emotion.

## 4 CONCLUSIONS

The results of this study suggest that people perceive agent’s emotions not depending on just one channel but depending on both visual and audio channel. Additionally facial expression (happy face vs. sad face) makes a difference in influence of two channels. That is, audio channel has more influence

in interpretation of emotion when facial expression is happy; according to the voice, participants perceived more various emotions when facial is happy. This result can be explained with cultural effect. In East Asian countries such as Japan, China, and Korea, where people are more collectivistic and interdependent, it is more important for emotional expressions to be controlled and subdued, and a relative absence of affect is considered important for maintaining harmonious relationships (Heine et al., 1999). This tendency is stronger in negative emotional expressions; they tend to avoid expressing their negative emotion freely. Because of this cultural characteristic, participants may take account of emotional information from both channels to discover the true emotional state of others and this may be for perception of agent's emotion.

In the pilot study, when presented with cold face participants have reported feeling little happy. However, it has found that combined cold face with happy or warm voice is perceived as a little happy or warm. It seems that this is because people do not always express their own emotions on their faces clearly. However cold expression has been perceived, people also have reflected voice containing particular emotion and identified emotion expressed.

The result from open question and MDS indicates that when facial expressions were not congruent with, participants were able to feel other emotion which was not expressed by face or voice. They recognized certain emotions such as being bored, fatigued, disappointed by combining two channels. This finding suggests that it is possible to combine one facial expression and other vocal emotion in order to represent different kinds of emotions by agents. In other words, every emotion agents intend to convey needs not producing because agent may express various emotions by using only several kinds of emotions.

However, there are a few limitations in the study. First, one limitation is that the agent in this experiment is unlike human appearance. Therefore, participants may be influenced as recognizing the emotion expressed by agents and feel the emotion a little differently. Second, coldness and warmth seem to be more ambiguous than basic emotions. Because warm facial expression is similar to happy facial expression, it was difficult for us to produce warm and happy facial expression distinctively in the experiment. Future research should use basic emotions and find which emotion would evoke if particular basic emotions merge.

## REFERENCES

- Bartneck, C., 2001. Affective expressions of machines. *Extended Abstracts of CHI'01:Conference on Human Factors in Computing Systems*, 189–190.
- Bickmore, T., 2003. Relational Agents: Effecting Change through Human. Computer Relationships. Ph.D. Thesis, Department of Media Arts and Sciences, Massachusetts Institute of Technology
- Brave, S., Nass, C., Hutchinson, K., 2005. Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human-Computer Studies*, 62 (2), 161-178.
- Creed, C., Beale, R., 2008. Psychological responses to simulated displays of mismatched emotional expressions. *Interacting with computers*, 20, 225-239
- De Gelder, B., Vroomen, J., 2000. The perception of emotions by ear and by eye. *COGNITION AND EMOTION*, 14 (3), 289-311.
- Ekman, P., Davidson, R.J., Friesen, W. V., 1990. The Duchenne smile: emotional expression and brain physiology. *Journal of Personality and Social Psychology* 58 (2), 342–353.
- Ekman, P., 2004. Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life. Henry Holt & Co.
- Heine, S. J., Lehman, D. R., Markus, H. R., Kitayama, S., 1999. Is there a universal need for positive self-regard. *Psychological Review*, 106, 766– 794.
- Russell, J. A., 1980. A circumplex model of affect. *Journal of personality and social psychology*, 39(6), 1161.