Use Color, Sound and Vibration for An Appearance-Constrained Robot to Express Emotions

Sichao Song[1] and Seiji Yamada[2]

[1] Department of Informatics
The Graduate University for Advanced Studies (SOKENDAI)
sichaos@nii.ac.jp
[2] National Institute of Informatics
seiji@nii.ac.jp

Abstract: Many HAI and HRI researchers are now dedicating their efforts to study interactive modalities such as facial expressions, gestures and gaze for affective communication. These modalities make such communication between robots and individuals become natural and attractive. Although such humanoid-oriented research is promising, approaches are time and cost consuming which come along with many technical difficulties in many research studies. In order to tackle this issue, we alternatively turn our focus to other three simple but intuitive interaction modalities and their combinations, namely color, sound, and vibration. We conduct a structured study to evaluate the effects of such three modalities on a human`s emotional perception towards our ball-shaped robot 'Maru'. This results in a set of nine recommended expressions that other appearance-constrained robots may readily employ.

INTRODUCTION

People respond to robots similar to how they respond to people, especially if the robots communicate with people using the same body language and other nonverbal cues that people use. Thus, many researchers are now dedicating their efforts to studying interactive modalities such as facial expressions, gestures and gaze for affective communication. Unfortunately, these modalities are time and cost consuming. In addition, they are also restricted by their robot design. To tackle such limitations and make interaction design simple, low cost, and intuitive, we probe three modalities: color, sound, and vibration. Previous studies have shown their impact on a person's perception [3, 9, 8], but few works have comprehensively evaluated the effect of these modalities in scenarios involving affective communication with a social robot. Thus, this leads our research questions as to how the three modalities affect a human's emotional perceptions through expressions.

EXPRESSIONS DESIGN

Meta-Analysis

We first establish a set of basic expressions that represent the mappings between each single modality of color, sound, and vibration and each emotional category of relaxed, happy, sad, and angry. Further mixed-modality expressions were built upon these basic expressions.

Color

According to [1], green elicits the feeling of excitement, and red was stated as being tiring and depressive. Similarly, [5] supports the idea that the color green attained the most positive emotion. [4] claims that a strong color (especially red) puts the brain into a highly excited state and might induce a bad mood. Moreover, [6] reviews various studies on mapping emotional states onto colors. The studies suggest that white means peaceful, blue means depressed, and red means angry.

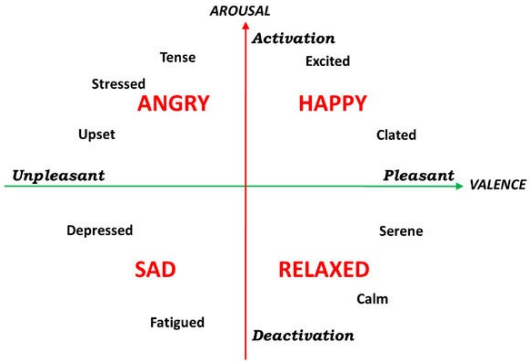


Figure 1. The circumplex model of affect.

Sound

[2] suggests that 1) when beep sounds with upward slopes (increasing intonation) are presented from a computer, people perceive the computer's attitude as showing "disagreement" regardless of the duration of the beeps. 2) when slower downward slopes (decreasing intonation) with a longer duration are presented, the computer's attitude is interpreted as "hesitation". Basically, "disagreement" can be interpreted as an emotion consisting of negative affection and a high level of arousal, while "hesitation" consists of a negative

Table 1. Assumptions of mappings between single modality and emotion, forming 12 basic expressions.

Emotion	Color(c)		Sound(s)		Vibration(v)	
relaxed	white	c1	flat beep sound	s1	mildly intense vibration	v1
happy	green	c2	flat beep sound (louder than s1)	s2	highly intense vibration (lower than v4)	v2
sad	blue	c3	falling beep sound	s3	low intense vibration	v3
angry	red	c4	rising beep sound	s4	highly intense vibration	v4

emotion with a low level of arousal. As a result, we claim that suggestions 1) and 2) also hold if "disagreement" and "hesitation" are replaced by the emotions "angry" and "sad".

Vibration

[7] and [9] infer that levels of vibration intensity are associated with different emotions. Hence, we decided to associate relaxed with a mildly intense vibration, happy with highly intense vibration (lower than that for angry), sad with low intense vibration, and angry with a highly intense vibration.

Candidate Expressions

Our set of candidate expressions consists of basic expressions (expressions through one single modality) and mixed-modality expressions (expressions through multiple modalities). Table 1 shows 12 basic expressions. Each of them was assigned with a unique code, for example a white color expression was assigned with "c1", and a falling beep sound was assigned with "s3". On the basis of them, we further designed 16 mixed-modality expressions. Specifically, each mixed-modality expression was a combination of two or three basic expressions from the same emotion category. Their names were decided by mixing codes of combined modalities followed by a number indicating which emotion category of the basic expressions they belonged to. For instance, cvs1 is a mixed-modality expression that consists of three basic expressions, c1, s1, and v1.

Maru the Robot

We built Maru to carry out the user experiments. Four LEDs (white, green, blue, and red) are assembled behind each of its eyes. In addition, a speaker is used to generate beep sound cues, and a vibration motor is attached to the inner body to produce vibration cues. An Arduino UNO board is programmed to control the robot. Figure 2 shows Maru and how it expresses emotions through the three modalities color, sound, and vibration and their combinations.

EXPERIMENT

Twenty-four Japanese in total (12 males, 12 females) ranging from 20 to 39 years old (M = 29.09, SD = 5.90) were recruited for the experiment. All of them were native Japanese speakers with a certain amount of knowledge on English.

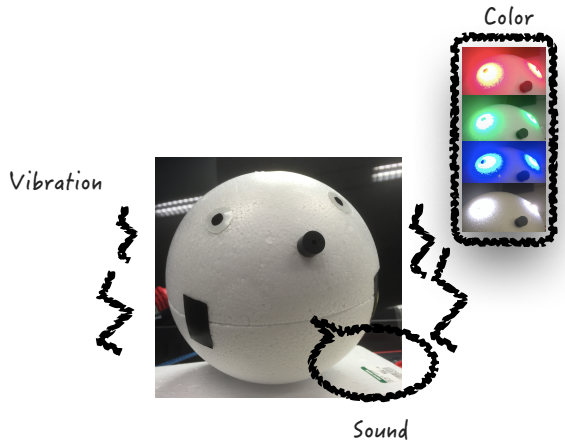


Figure 2. Maru and its expressions made through color, sound, and vibration.

Procedure

Maru was placed in front of the participants at a distance of about 50 cm. Before the experiment started, the experimenter briefly explained the purpose and setting of the study. The participants were required to complete a short pre-questionnaire consisting of demographic information and questions regarding experience with robots. After finishing the pre-questionnaire, the experimenter started the experiment. In total, 28 trials were conducted for each participant, where in each trial, Maru repeatedly performed a single expression from the candidate set. The order of expressions was counterbalanced across participants. Each trial lasted for 10 seconds, and between each two trials, the participants had a 20-seconds pause to select one emotion out of the four (relaxed, happy, sad, and angry) that they believed Maru had just expressed. After all the trials were completed, the experimenter ended the experiment and thanked the participant.

RESULTS

Criteria for Selecting Expressions

We analyzed the candidate expressions with regard to the four emotions separately. For the evaluation, we first introduced two criteria to select good expressions: 1) an expression must have a strong interpretation regarding an emotion (selection rate in the top quartile, or in other words above the third quartile), and 2) an expression must be iconic, meaning that it has only one dominant perception among the four emotions. For instance, an expression is ambiguous and not desirable if the participants perceive it as more than one emotion.We assessed the iconic-ness for each candidate expression that meets criteria (1). To evaluate, we used one-sample

tests of proportions with multinomial test. For each test that was significant, we further conducted post-hoc multinomial tests with Bonferroni correction for multiple comparisons. Because of the four emotion categories, the hypothesized probability that each emotion would be chosen at random regarding an expression was set to one-fourth (25%, which is the probability of a random guess). After a selection based on the two criteria, we further discarded expressions that were not in line with the participant' comments (gathered from open questions in the post-questionnaire).

Relaxed

We recommend c1 for expressing a relaxed emotion. It also has the highest selection rate in the relaxed category (see Figure 3). Although both cv1 and cs2 also met criteria (1) and (2), they were discarded because of the participants' comments "Vibration conveys negative emotions, and especially that high intense vibration conveys angry" and "It was difficult to recognise relaxed emotion when sound was used".

A multinomial test indicated a significance difference in c1 ($p<0.01$). Post-hoc tests with Bonferroni correction suggest that the result for relaxed was significant (significantly above 25%, $p<0.001$), while results for the other three emotions are not (happy: n.s.; sad: n.s.; angry: significantly under 25%, $p<0.01$). In addition, the selection of c1 met our assumptions as we assumed the mapping between the color white and relaxed emotion (see Table 1).

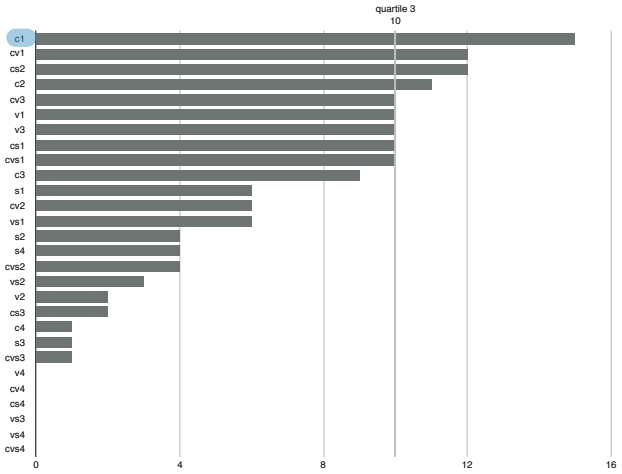


Figure 3. Selection rates for relaxed emotion category.

Happy

No expressions met our selection criteria in the happy category (see Figure 4). Although the four expressions had scores above the third quartile, post-hoc tests showed that none of the four expressions were iconic. This meets participants' comment (a) suggesting the difficulty of recognizing the happy emotion.

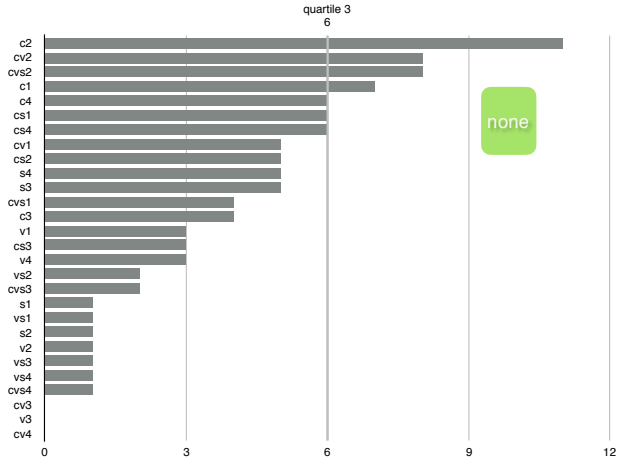


Figure 4. Selection rates for happy emotion category.

Sad

We recommend cvs3, vs3, cs3 and s3 for expressing a sad emotion. They had the top five selection rates in the sad category except for s2 (see Figure 5). We conducted multinomial tests for the top six expressions that met criteria (1). The results of post-hoc tests with Bonferroni correction indicats that all the six expressions also met criteria (2).

We further discarded s2 and s1 because of the participants' comments "Rising/falling sounds were easily recognised as angry/ sad, but flat sounds were difficult to interpret". All of the rest four expressions formed our recommend expressions for sad, which consist of basic expressions that are mapped to the sad emotion (c3: blue color; v3: low intense vibration; s3: falling beep sound). This also met our assumptions.

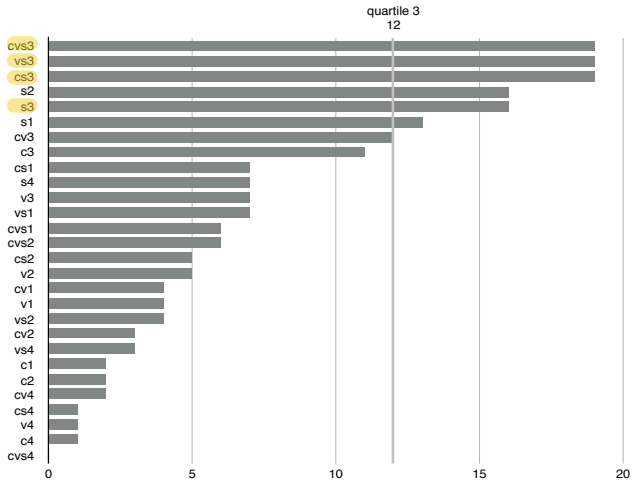
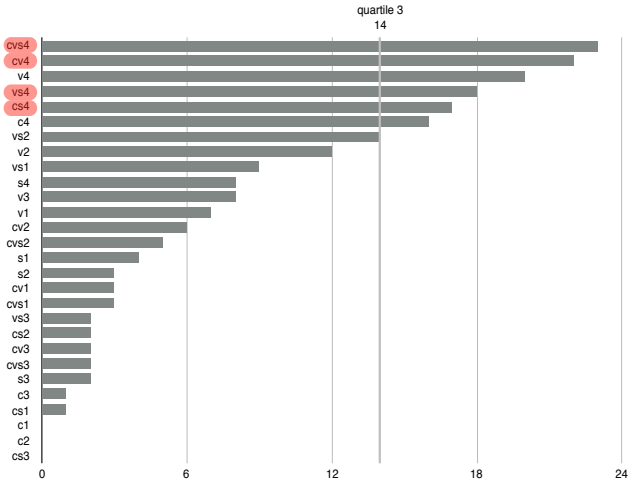


Figure 5. Selection rates for sad emotion category.

Angry

Four expressions, cvs4, cv4, vs4, and cs4, are recommended for expressing an angry emotion. These

are the top five scores in the angry category except for v4 (see Figure 6). Similar with sad category, all of the top six expressions were iconic in the angry category.
Further, we discarded v4 and c4 because of the participants' comments "Using vibration alone was confusing" and "c2 11 Using multiple modalities is better understandable cv2 8 than using single modality alone". All of the four recommended expressions consisted of basic expressions that were mapped to the angry emotion (c4: red color; v4: highly intense vibration; s4: rising beep sound), which again met our assumptions.



DISCUSSION

We conclude from our findings that expressions made though the three modalities of color, sound, and vibration are capable of conveying various emotions. In particular, the color modality, among the three, was found to be the most important for communicating affection. Sound and vibration are considered to show a certain bias toward some emotions. For example, falling and rising sounds can be interpreted as sad and angry, and vibrations convey negative emotions in general. Thus, we can offer five suggestions as general design guidelines: 1) it is better to use expressions that contain color modality; 2) when expressing sadness, a falling sound is strongly recommended; 3) when expressing anger, a rising sound and highly intense vibration are strongly recommended; 4) use multiple modalities rather than a single modality, if possible; 5) it is better not to use vibrations for positive emotions.

For each emotion, we selected all of the expressions that met our criteria rather than pick only one best expression. This is because there might not be one best expression that holds for everyone. Instead, we offer a set of good expressions so that variety and flexibility are promised. For example, a practical issue could be of the various designs of robotic platforms. A robot may not be able to perform expressions through all the three modalities, especially vibration. Thus, for designers who would apply our findings to their projects, we suggest that they start with choosing the expressions that have the highest selection rates while meeting their hardware configurations and that they further adjust their choices on the basis of the performance.

CONCLUSION

In this paper, we discussed our research on how to express emotions with a social robot through three modalities: color, sound, and vibration. We worked through a structured design process to carefully select a set of expressions that can well convey four emotions: relaxed, happy, sad, and angry. Our study began with a broad literature survey followed by a pre-design session. This resulted in a set of 28 candidate expressions that consist of one or multiple modalities. The results suggest, in total, nine best expressions that can well convey relaxed, sad, and angry emotions, while no expression can be recommended for the happy emotion. We claim that our findings possess good diversity, as each of the four emotions was mapped onto a particular quadrant of a valence-arousal space (the circumplex model).

ACKNOWLEDGEMENT

This study was partially supported by JSPS KAKENHI "Cognitive Interaction Design" (No.26118005).

REFERENCE:

1. http://news.mit.edu/2008/nexi-0409 — accessed 03.09.2016.
2. T. Komatsu. Toward making humans empathize with artificial agents by means of subtle expressions. In International Conference on Affective Computing and Intelligent Interaction, pages 458–465. Springer, 2005.
3. T. Komatsu, S. Yamada, K. Kobayashi, K. Funakoshi, and M. Nakano. Artificial subtle expressions: intuitive notification methodology of artifacts. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pages 1941–1944. ACM, 2010.
4. R. Kuller, B. Mikellides, and J. Janssens. Color, arousal, and performancea–a comparison of three experiments. Color Research & Application, 34(2):141–152, 2009.
5. K. Naz and H. Helen. Color-emotion associations: Past experience and personal preference. In AIC 2004 Color and Paints, Interim Meeting of the International Color Association, Proceedings, volume 5, page 31. Jose Luis Caivano, 2004.
6. N. A. Nijdam. Mapping emotion to color. (Eds.): Book Mapping emotion to colora (2009, edn.), pages 2–9, 2009.
7. S. U. Réhman and L. Liu. Vibrotactile emotions on a mobile phone. In Signal Image Technology and Internet Based Systems, 2008. SITIS'08. IEEE International Conference on, pages 239–243. IEEE, 2008.
8. M. V. Sokolova and A. Fernández-Caballero. A review on the role of color and light in affective computing. Applied Sciences, 5(3):275–293, 2015.
9. S. ur Réhman and L. Liu. ifeeling: Vibrotactile rendering of human emotions on mobile phones. In Mobile multimedia processing, pages 1–20. Springer, 2010.