

# Expressing Emotions through Color, Sound, and Vibration with an Appearance-Constrained Social Robot

Sichao Song  
The Graduate University for Advanced Studies  
(SOKENDAI)  
Department of Informatics  
2-1-2 Hitotsubashi, Chiyoda, Tokyo, Japan  
sichaos@nii.ac.jp

Seiji Yamada  
National Institute of Informatics  
2-1-2 Hitotsubashi, Chiyoda, Tokyo, Japan  
seiji@nii.ac.jp

## ABSTRACT

Many researchers are now dedicating their efforts to studying interactive modalities such as facial expressions, natural language, and gestures. This phenomenon makes communication between robots and individuals become more natural. However, many robots currently in use are appearance constrained and not able to perform facial expressions and gestures. In addition, although humanoid-oriented techniques are promising, they are time and cost consuming, which leads to many technical difficulties in most research studies. To increase interactive efficiency and decrease costs, we alternatively focus on three interaction modalities and their combinations, namely color, sound, and vibration. We conduct a structured study to evaluate the effects of the three modalities on a human's emotional perception towards our simple-shaped robot "Maru." Our findings offer insights into human-robot affective interactions, which can be particularly useful for appearance-constrained social robots. The contribution of this work is not so much the explicit parameter settings but rather deepening the understanding of how to express emotions through the simple modalities of color, sound, and vibration while providing a set of recommended expressions that HRI researchers and practitioners could readily employ.

## Keywords

Color; sound; vibration; emotional expression of a robot; multi-modality interaction; human-agent interaction (HAI).

## 1. INTRODUCTION

As Cynthia Breazeal [2] claimed, robots are actually a really intriguing social technology and have the ability to "push our social buttons." People respond to social media (robots in particular) 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, more and more researchers are now dedicating their efforts to studying human-like robots. Unfortunately, such robots are high cost and technically difficult to design. To date, a range of robots currently in use for applications such as search and rescue and domestic purposes (cleaning robots) are not designed to be anthropomorphic, meaning that their abilities in affective expression are highly restricted [7, 6, 33]. While such robots may not require rich expressivity, they do need to have certain abilities to show affection. [6, 7] addressed such an issue. For example, victims may perceive a rescue robot as "creepy" so that such a robot needs to convey affect to reduce their intense emotions. There is a significant challenge in finding what other effective interactive modalities these appearance-constrained robots can employ to support affective human-robot interaction.

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 [21, 36, 34]. However, few papers have comprehensively evaluated the effect of these modalities in scenarios involving affective communication with a social robot. Thus, this leads our research question as to how the three modalities affect a human's emotional perceptions through expressions.

In this work, we work through a structured process to reach our design. We survey a number of pieces of literature to form our fundamental assumptions, and further build our prototypical social robot "Maru" for an experiment. On the basis of the results, we recommend a set of nine expressions that can well express affection.

## 2. RELATED WORK

Affective communication for social robots has been discussed intensively in recent years. Various robots are now designed, from humanoids to androids, to be able to establish interactions with humans [3, 1, 16]. They have been intensively studied in various scenarios such as education [35, 14], autism therapy [5, 11], guidance [18], and driving support [17, 25]. For such social robots, natural language has been considered to be a significant interactive modality. However, the current state-of-the-art in natural language and related technologies is still far from satisfying [37]. Previous research has shown that over 80% of human communication is encoded in facial expressions and body movements [32]. Hence, non-verbal cues are viewed as essential affective communication methods [7]. For instance,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*HRI '17, March 06-09, 2017, Vienna, Austria*

© 2017 ACM. ISBN 978-1-4503-4336-7/17/03...\$15.00

DOI: <http://dx.doi.org/10.1145/2909824.3020239>

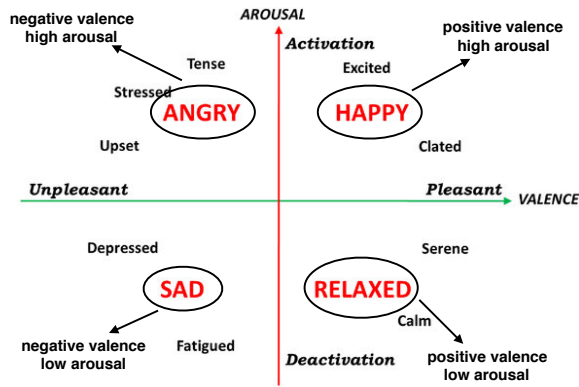


Figure 1: The circumplex model of affect.

facial expressions have been a popular mechanism for showing affection with robots as well as body movement [9, 8, 24, 31], posture [9, 8, 24, 31], and orientation [9, 8, 10, 28].

Although such interaction methods are natural and effective, they are limited by the embodiment of robots. As many currently-in-use robots are appearance-constrained, they do not have the abilities to provide social cues through modalities such as facial expression, gesture, and gaze [7, 6]. Thus, it is important to explore other interactive modalities that are simple, low cost, but intuitive.

Three alternative modalities, namely color, sound, and vibration, have also been investigated particularly in HCI and psychology. For instance, a number of studies on the role of color and light [23, 22, 26, 27, 34, 15] have been carried out; a handful of studies, particularly on semantic-free utterances (SFU), have explored the design of sounds that allow emotion and intent expressions with machines [19, 21, 20, 37]; Vibration feedback has mostly been studied as an auxiliary means to support the communication of emotions [30, 36, 4]. However, no previous work discussed affective interaction through combinations of the three modalities. Such multi-modal approaches are important since there are currently no sound principles for expressing particular emotions through single modalities.

Particularly, [13] investigated different light behaviors (for a single LED) as a means to convey a system’s states to a user. Our approach is adapted and improved from their well-structured design process.

### 3. EXPRESSION DESIGN

The circumplex model of affect [29] is used to map emotions onto a valence-arousal space (see Figure 1). In this model, emotions can be represented at levels of valence and arousal. It can offer benefits to maximize diversity if we focus on four emotions, relaxed, happy, sad, and angry, since each emotion can be mapped onto a different quadrant of the valence-arousal space. Specifically, relaxed is of positive and low arousal, happy is of positive and high arousal, sad is of negative and low arousal, and angry is of negative and high arousal. Most emotions in the same quadrant are similar to each other but quite different from other quadrants. For instance, calm and serene are close to relaxed but distinct from happy. Therefore, we say that the four emotions can still represent participants’ perceptions to the greatest extent without introducing additional complexity if

more emotions are introduced. The remaining parameters regarding color, sound, and vibration modalities are selected on the basis of this model.

### 3.1 Meta-Analysis

We first surveyed a number of related pieces of literature to decide on a set of basic expressions that represent the mappings between each single modality and the emotions. Further, mixed-modality expressions were built upon these basic expressions.

#### 3.1.1 Color

According to [1], green elicits the feeling of excitement, and red was stated as being tiring and depressive. Similarly, [26] supports the idea that the color green attained the most positive emotion, and [22] claims that a strong color (especially red) puts the brain into a highly excited state and might induce a bad mood. Moreover, [27] reviews various studies on mapping emotional states onto colors. The studies suggest that white means peaceful, blue means depressed, and red means angry.

We found that mappings between emotional states and colors are conflictive on the basis of the above studies, which means that different authors may suggest different relationships between emotions and colors. Nevertheless, we selected the mappings that were supported by most researchers.

#### 3.1.2 Sound

Affective sounds, especially non-linguistic utterances (NLUs), are weighted much. For example, [19] suggests that, 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 and that, 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 emotion with a low level of arousal. As a result, we claim that both suggestions also hold if “disagreement” and “hesitation” are replaced by the emotions “angry” and “sad.”

#### 3.1.3 Vibration

Vibration is mostly investigated as an auxiliary modality for conveying affection in various HCI related studies. To our knowledge, no study uses single vibration modalities to communicate affection. [30] and [36] 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.

### 3.2 Pre-Design Session

We asked a panel of five researchers (members of our research group; one female) to discuss the selection of emotional expressions. None of them were familiar with our project before joining the session. The design session lasted for about 30 minutes. We asked them to comment on the parameters we currently decided on and give suggestions on expressing relaxed and happy through sound modality.

To be specific, we started by asking the researchers if they

**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
happy	green	c2	flat beep sound (louder than s1)	s2	highly intense vibration (lower than v4)
sad	blue	c3	falling beep sound	s3	low intense vibration
angry	red	c4	rising beep sound	s4	highly intense vibration

**Table 2: List of all 28 candidate expressions; 1 - 12 are basic expressions, 13 - 28 are mixed-modality expressions.**

1	c1	15	cs3
2	c2	16	cs4
3	c3	17	cv1
4	c4	18	cv2
5	s1	19	cv3
6	s2	20	cv4
7	s3	21	vs1
8	s4	22	vs2
9	v1	23	vs3
10	v2	24	vs4
11	v3	25	cvs1
12	v4	26	cvs2
13	cs1	27	cvs3
14	cs2	28	cvs4

were satisfied with our basic expressions. Although the researchers are all agreed with the mappings, they mentioned the possibility that some participants might not take these mappings as a ground truth. Thus, we decided to treat the basic expressions as assumptions but not as true relationships that are accepted by everyone. The researchers also suggested using a flat beep sound to express both relaxed and happy emotions. To differentiate between the two, the beep sound associated with happy was made louder since happy is of higher arousal than relaxation.

### 3.3 Candidate Expressions

Our set of candidate expressions consisted 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. Table 2 demonstrates all the candidate expressions, where Nos. 1 to 12 are basic expressions, and Nos. 13 to 28 are mixed-modality expressions.

### 3.4 Assumptions

We made assumptions to design our set of candidate expressions. On the basis of a survey of related literature, we first assumed a set of basic mappings between single

modalities and emotions. Further, we assumed that valid mixed-modality expressions should be combinations of basic expressions from the same emotion categories. The benefit of doing so was that the size of our design space was significantly reduced. We cut down the number of candidate expressions from the original number of 124 (allowing any combination of modalities) to the current number of 28.

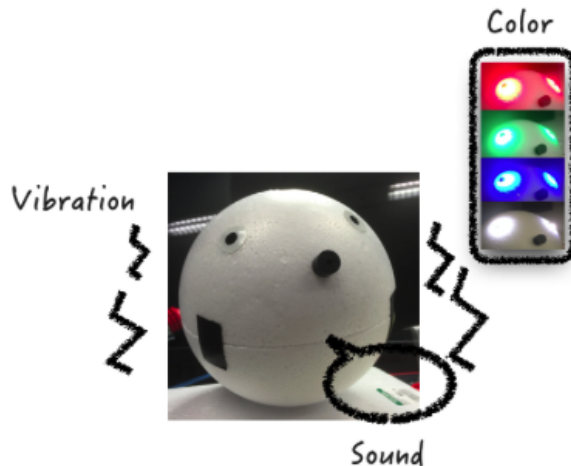
We are aware of the bias we introduce when such assumptions are made. However, since we do not take our assumptions as a ground truth for the participants, we claim that our results are minimally affected by the bias.

### 3.5 Maru the Robot

We built Maru as a prototypical social robot to carry out the user experiment. We applied a minimal design and intentionally made Maru’s embodiment and appearance simple while still having the attribute of anthropomorphism. Maru is made of two pieces of hollow, semi-spherical Styrofoam. 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 of color, sound, and vibration and their combinations.

## 4. 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. In



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

addition, none of them had experience in using or working with a robot.

### 4.1 Procedure

Maru was placed in front of the participants at a distance of about 50 cm (see Figure 3). The Arduino UNO board inside its body controlled all of its expressions. 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 (1-second-long expression followed by a 1-second pause; all modalities were synchronized) from the candidate set.

The expressions were randomized across participants. Each trial lasted for 10 seconds, and between each two trials, the participants had a 20-second 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 participants.

## 5. RESULTS

We summarized the answers from all 24 participants. Figure 4 provides an overview of the experimental results. For each candidate expression, the selection rate (SR), indicating how many participants perceived an expression as a particular emotion, was counted with regard to each of the four emotion categories. Because the total number of participants was 24 in this experiment, the value of the selection rate ranges from 0 to 24. For instance, c2 has a SR of 0 regarding the emotion of anger as no participant perceived c2 as angry when Maru expressed it, but it has SRs of 11, 11, and 2 with regard to relaxed, happy, and sad, respectively.

In addition to the SRs, we gathered participants’ subjective comments on our expression design through open questions in post-questionnaires. We believed that feedback from users would also be essential to our selection of expressions.

### 5.1 Criteria for Selecting Expressions

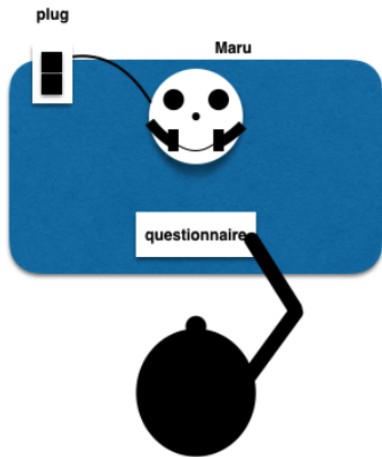


Figure 3: Experiment setting.

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 a 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 discarded expressions that were not in line with the participants’ comments. The reason for doing so is that we believe that the collected comments reveal the participants’ perceptions of our design. For example, the participants expressed their opinions through open questions such as “How do you think of the robot expressing emotions through color/sound/vibration.” Twenty-three out of the 24 participants answered the open questions. In Table 3, we summarize the representative viewpoints that were given by at least three participants. Nine expressions (Figure 4, shaded area) were selected as our recommended set of emotion expressions made through color, sound, and vibration. We now describe them with regard to each emotion category respectively.

#### 5.1.1 Relaxed

We recommend c1 for expressing a relaxed emotion. It also has the highest selection rate in the relaxed category (see Figure 5). Although both cv1 and cs2 also met criteria (1) and (2), they were discarded because of the participants’ comments (d) and (f) (see Table 3).

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 [see Table 4 and Figure 9(a), significantly above 25%,  $p < 0.001$ ], while results for the other three emotions were not (happy: n.s.; sad:

Table 3: Summary of representative comments from participants. Numbers in parenthesis indicate number of participants who gave comments.

a	It was difficult to recognize the happy emotion.(7)
b	Color is the most important modality for expressing affection.(8)
c	Using vibration alone was confusing.(4)
d	Vibration conveys negative emotions, and a highly intense vibration especially conveys angry.(6)
e	Rising/falling sounds were easily recognized as angry/sad, but flat sounds were difficult to interpret.(7)
f	It was difficult to recognize the relaxed emotion when sound was used.(3)
g	Using multiple modalities is better understandable than using a single modality alone.(6)

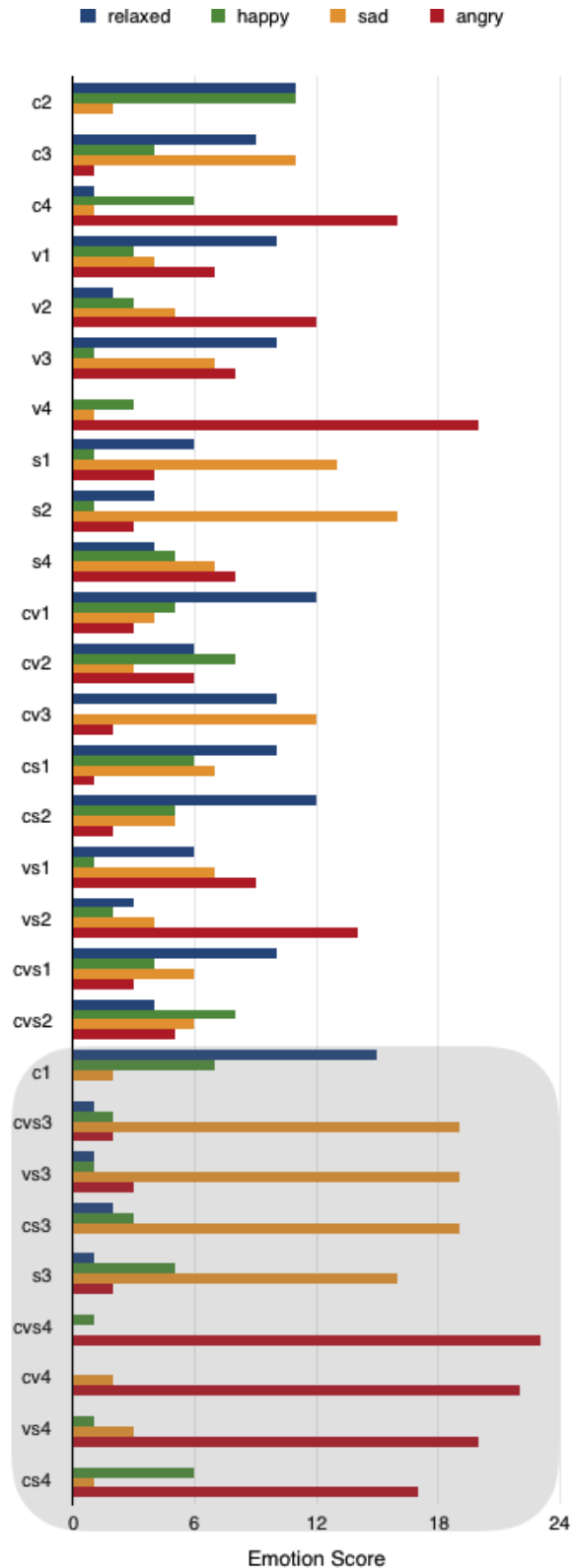


Figure 4: Overview of our results. Selection rate (SR), ranging from 0 to 24, indicates how many participants mapped expressions onto emotions; shaded area shows nine recommended expressions.

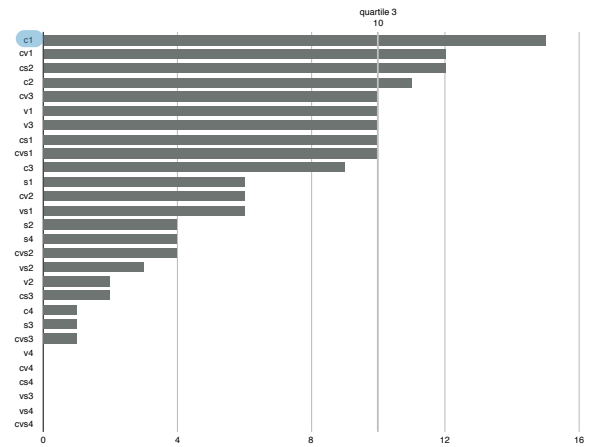


Figure 5: Selection rates (SR) for relaxed emotion category.

Table 4: Post-hoc tests with Bonferroni correction on c1; see Figure 9(a).

	p(relaxed)	p(happy)	p(sad)	p(angry)
c1	p<0.001, significantly above 25%	n.s.	n.s.	p<0.01, significantly under 25%

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).

### 5.1.2 Happy

No expressions met our selection criteria in the happy category (see Figure 6). Although the four expressions had SRs above the third quartile, post-hoc tests showed that none of the four expressions were iconic [see Table 5 and Figure 9(b)]. This meets comment (a) suggesting the difficulty of recognizing the happy emotion.

### 5.1.3 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 7). We conducted multinomial tests for the top six expressions that met criteria (1). Table 6 and Figure 9(c) show the results of post-hoc tests with Bonferroni correction, indicating that all the six expressions also met criteria (2).

We further discarded s2 and s1 because of comment (e). All of the remaining 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.

### 5.1.4 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 8). Similar with the sad category, all of the top six expressions were iconic in the angry category [see Table 7 and Figure 9(d)].

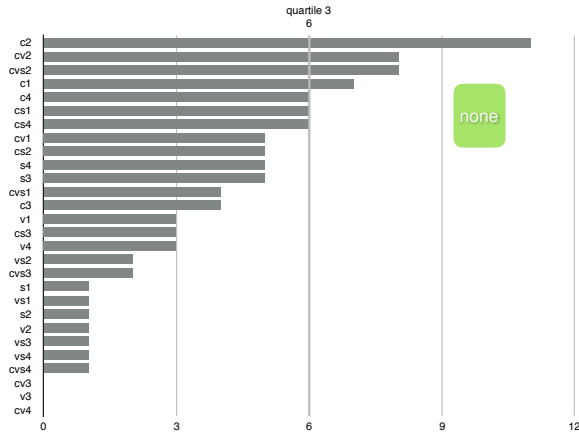


Figure 6: Selection rates for happy emotion category.

Table 5: Post-hoc tests with Bonferroni correction on c1; see Figure 9(b).

	p(relaxed)	p(happy)	p(sad)	p(angry)
c2	n.s.	n.s.	n.s.	p<0.05, significantly under 25%
cv2	n.s.	n.s.	n.s.	n.s.
cvs2	n.s.	n.s.	n.s.	n.s.
c1	p<0.001, significantly above 25%	n.s.	n.s.	p<0.05, significantly under 25%

Further, we discarded v4 and c4 because of comments (c) and (g). 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.

## 6. DISCUSSION

On the basis of our analysis results, we recommend a set of nine emotional expressions that can well convey affection (see Figure 4, shaded area). For a relaxed emotion (positive valence and low arousal), we suggest c1. For sad (negative valence and low arousal), we suggest four expressions, cvs3, vs3, cs3, and s3. For angry (negative valence and high arousal), we suggest four expressions, cvs4, cv4, vs4, and cs4. Unfortunately, we did not find any good expressions for happy (positive valence and high arousal) emotion. A possible reason may be that the assumptions we made for the happy category are inadequate. It is also possible as we find that people tend to perceive highly intensive expressions as negative emotion rather than positive emotions.

We conclude from our findings that expressions made through the three modalities 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. On the basis of our findings, we offer six suggestions as general design guidelines:

- I. It is better to use expressions that contain color modality;
- II. When expressing sadness, a falling sound is strongly recommended;

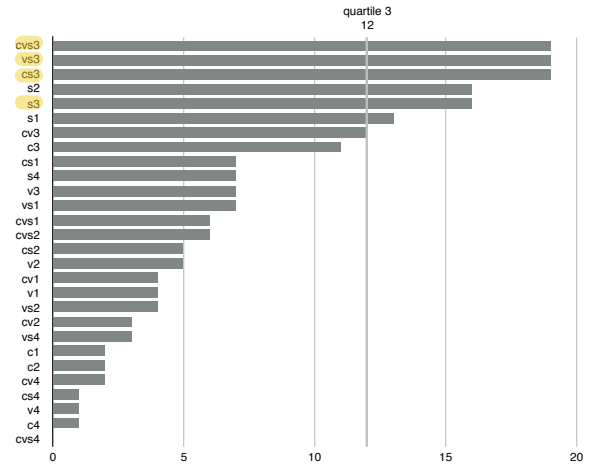


Figure 7: Selection rates for sad emotion category.

Table 6: Post-hoc tests with Bonferroni correction on cvs3, vs3, cs3, and s3; see Figure 9(c).

	p(relaxed)	p(happy)	p(sad)	p(angry)
cvs3	p<0.05, significantly under 25%	n.s.	p<0.001, significantly above 25%	n.s.
vs3	p<0.01, significantly under 25%	p<0.01, significantly under 25%	p<0.001, significantly above 25%	n.s.
cs3	n.s.	n.s.	p<0.001, significantly above 25%	p<0.01, significantly under 25%
s3	p<0.05, significantly under 25%	n.s.	p<0.001, significantly above 25%	n.s.

- III. When expressing anger, a rising sound and highly intense vibration are strongly recommended;
- IV. Use multiple modalities rather than a single modality, if possible;
- V. It is better not to use vibrations for positive emotions.
- VI. It is much easier to express negative expressions rather than positive expressions.

For each emotion, we selected all of the expressions that met our criteria rather than pick only the best one. 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. 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.

Our assumptions on basic mappings between each single modality and emotion formed our set of basic expressions, on the basis of which we further designed the mixed-modality expressions. Noticeably, we never made these assumptions as a ground truth for the participants. In other words, we

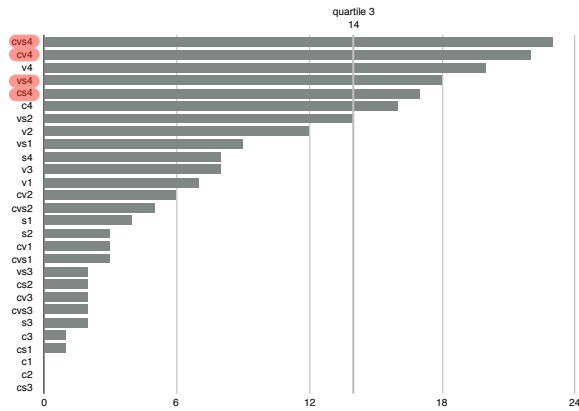


Figure 8: Selection rates for angry emotion category.

Table 7: Post-hoc tests with Bonferroni correction on cvs4, cv4, vs4, and cs4; see Figure 9(d).

	p(relaxed)	p(happy)	p(sad)	p(angry)
cvs4	p<0.01, significantly under 25%	p<0.05, significantly under 25%	p<0.01, significantly under 25%	p<0.001, significantly above 25%
cv4	p<0.01, significantly under 25%	p<0.01, significantly under 25%	n.s.	p<0.001, significantly above 25%
vs4	p<0.01, significantly under 25%	n.s.	n.s.	p<0.001, significantly above 25%
cs4	p<0.01, significantly under 25%	n.s.	n.s.	p<0.001, significantly above 25%

did not interpret if the participants correctly recognized an expression. In fact, our purpose for introducing assumptions to the design process was to reduce the size of our design space (we successfully cut down the number of candidate expressions from 124 to 28). We only wanted to find the expressions that can well convey emotions. As the selected expressions all met the assumptions we made, we believe that the bias caused by these assumptions was not strong enough to affect the experimental results. Thus, we claim that our design process was appropriate and our recommendation of the expressions is valid.

The generality of our results may be restricted due to the design of our robot Maru. Maru was built to have two eyes to gain the attribute of anthropomorphism, and LEDs are attached behind the eyes. As a result, expression through color modality is achieved by Maru “blinking” in the eyes. There is thus the possibility that our results (particularly expressions with regard to color modality) depend on the face-like appearance of the robot. However, because of humans’ intrinsic mechanism of anthropomorphizing things [12], we say that the generality of our findings is minimally affected.

This work can be further explored. For example, we restricted our set of candidate expressions to avoid too large a design space. We made our assumptions on the mappings between single modalities and emotions, and we further designed mixed-modality expressions on the basis of these assumptions. As a result, many other combinations of modalities are out of discussion. How expressions through

those conflicting modalities affect a human’s emotional perceptions is unknown. Further exploration may reveal an interesting phenomenon.

When we evaluated our candidate expressions, we only considered the combinations of basic parameters (color for color modality, pitch variation for sound modality, and level of intensity for vibration modality). However, explicit values of such parameters have not been studied thoroughly. In addition, many other parameters such as duration could also affect people’s emotional perceptions towards the expressions. Thus, it is important to further explore the setting of parameters on the basis of our current findings.

## 7. 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 emotions. 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).

Although most current studies on affective communication are focused on interaction modalities such as facial expressions, natural language, and body gestures, their limitations cannot be ignored. They in general suffer from high cost and technical difficulties. In addition, many of the present robots are designed without the capability to make natural interactions, especially through facial expressions and human-like body gestures. The three modalities, in comparison, have the advantage of being simple, low cost, and intuitive. This paper suggests multi-modal approaches for affective communication for an appearance-constrained robot, for instance, a rescue and search robot or a domestic-use cleaning robot. We believe that designers in HRI as well as other related fields will find our results beneficial.

## 8. ACKNOWLEDGMENTS

This study was partially supported by JSPS KAKENHI “Cognitive Interaction Design” (No.26118005). The authors would like to thank Yushan Pan for his valuable suggestions to this work.

## 9. REFERENCES

- [1] [https://www.ted.com/talks/cynthia\\_breazeal\\_the\\_rise\\_of\\_personal\\_robots](https://www.ted.com/talks/cynthia_breazeal_the_rise_of_personal_robots), 2008 (accessed September 3, 2016).
- [2] [https://www.ted.com/talks/cynthia\\_breazeal\\_the\\_rise\\_of\\_personal\\_robots](https://www.ted.com/talks/cynthia_breazeal_the_rise_of_personal_robots), 2010 (accessed September 7, 2016).
- [3] <https://www.ald.softbankrobotics.com/en/cool-robots/nao>, accessed September 6, 2016.
- [4] F. Arafsha, K. M. Alam, and A. El Saddik. Design and development of a user centric affective haptic jacket. *Multimedia Tools and Applications*, 74(9):3035–3052, 2015.

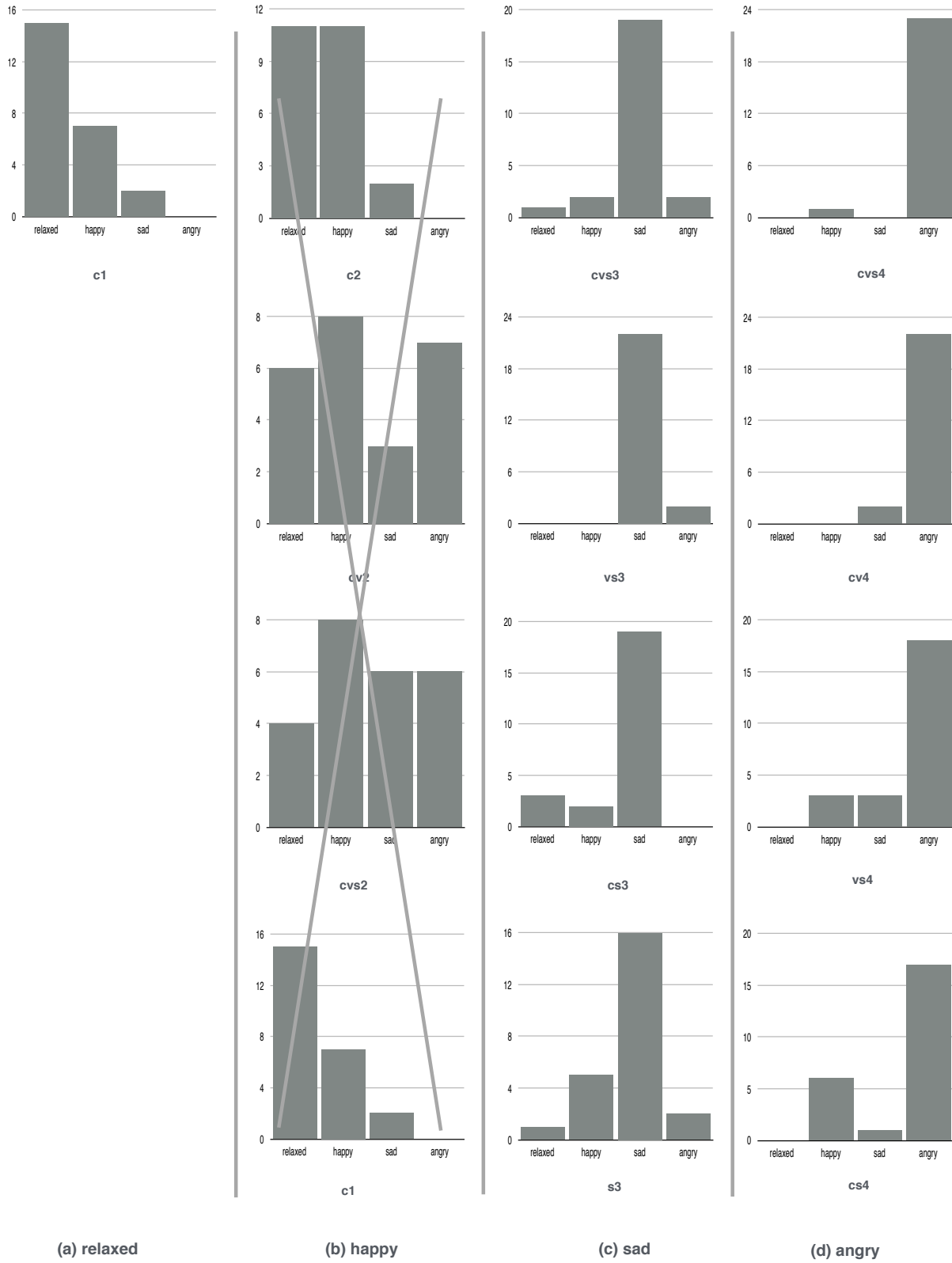


Figure 9: Selection rates vs. emotions regarding each expression. All expressions except for four under happy emotion category are recommended.



- [5] M. Begum, R. W. Serna, D. Kontak, J. Allspaw, J. Kuczynski, H. A. Yanco, and J. Suarez. Measuring the efficacy of robots in autism therapy: How informative are standard hri metrics?. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 335–342. ACM, 2015.
- [6] C. L. Bethel and R. R. Murphy. Affective expression in appearance constrained robots. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 327–328. ACM, 2006.
- [7] C. L. Bethel and R. R. Murphy. Survey of non-facial/non-verbal affective expressions for appearance-constrained robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(1):83–92, 2008.
- [8] C. Breazeal. Social interactions in hri: the robot view. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 34(2):181–186, 2004.
- [9] C. L. Breazeal. *Designing sociable robots*. MIT press, 2004.
- [10] A. Bruce, I. Nourbakhsh, and R. Simmons. The role of expressiveness and attention in human-robot interaction. In *Robotics and Automation, 2002. Proceedings. ICRA'02. IEEE International Conference on*, volume 4, pages 4138–4142. IEEE, 2002.
- [11] P. S. Dehkordi, H. Moradi, M. Mahmoudi, and H. R. Pouretmad. The design, development, and deployment of roboparrot for screening autistic children. *International Journal of Social Robotics*, 7(4):513–522, 2015.
- [12] S. R. Fussell, S. Kiesler, L. D. Setlock, and V. Yew. How people anthropomorphize robots. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, pages 145–152. ACM, 2008.
- [13] C. Harrison, J. Horstman, G. Hsieh, and S. Hudson. Unlocking the expressivity of point lights. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1683–1692. ACM, 2012.
- [14] D. Hood, S. Lemaignan, and P. Dillenbourg. When children teach a robot to write: An autonomous teachable humanoid which uses simulated handwriting. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 83–90. ACM, 2015.
- [15] I. Infantino, G. Pilato, R. Rizzo, and F. Vella. I feel blue: Robots and humans sharing color representation for emotional cognitive interaction. In *Biologically Inspired Cognitive Architectures 2012*, pages 161–166. Springer, 2013.
- [16] H. Ishiguro. Interactive humanoids and androids as ideal interfaces for humans. In *Proceedings of the 11th international conference on Intelligent user interfaces*, pages 2–9. ACM, 2006.
- [17] N. Karatas, S. Yoshikawa, and M. Okada. Namida: Sociable driving agents with multiparty conversation. In *Proceedings of the Fourth International Conference on Human Agent Interaction*, pages 35–42. ACM, 2016.
- [18] Y. Kato, T. Kanda, and H. Ishiguro. May i help you?: Design of human-like polite approaching behavior. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 35–42. ACM, 2015.
- [19] 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.
- [20] T. Komatsu and S. Yamada. How does the agents’ appearance affect users’ interpretation of the agents’ attitudes: Experimental investigation on expressing the same artificial sounds from agents with different appearances. *Intl. Journal of Human-Computer Interaction*, 27(3):260–279, 2011.
- [21] 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.
- [22] R. Küller, B. Mikellides, and J. Janssens. Color, arousal, and performance? a comparison of three experiments. *Color Research & Application*, 34(2):141–152, 2009.
- [23] B. Manav. Color-emotion associations and color preferences: A case study for residences. *Color Research & Application*, 32(2):144–150, 2007.
- [24] L. Moshkina and R. C. Arkin. Human perspective on affective robotic behavior: A longitudinal study. In *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1444–1451. IEEE, 2005.
- [25] Y. Nakagawa, K. Park, H. Ueda, and H. Ono. Driving assistance with conversation robot for elderly drivers. In *International Conference on Universal Access in Human-Computer Interaction*, pages 750–761. Springer, 2014.
- [26] 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.
- [27] N. A. Nijdam. Mapping emotion to color. *Book Mapping emotion to color?(2009)*, pages 2–9, 2009.
- [28] I. R. Nourbakhsh, C. Kunz, and T. Willeke. The mobot museum robot installations: A five year experiment. In *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, volume 4, pages 3636–3641. IEEE, 2003.
- [29] J. Posner, J. A. Russell, and B. S. Peterson. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(03):715–734, 2005.
- [30] 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.
- [31] M. Scheeff, J. Pinto, K. Rahardja, S. Snibbe, and R. Tow. Experiences with sparky, a social robot. In

- Socially Intelligent Agents*, pages 173–180. Springer, 2002.
- [32] J. Scheirer and R. Picard. Affective objects. *MIT Media lab Technical Rep.*, 524, 2000.
- [33] A. Singh and J. E. Young. A dog tail for utility robots: exploring affective properties of tail movement. In *IFIP Conference on Human-Computer Interaction*, pages 403–419. Springer, 2013.
- [34] 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.
- [35] F. Tanaka and S. Matsuzoe. Children teach a care-receiving robot to promote their learning: Field experiments in a classroom for vocabulary learning. *Journal of Human-Robot Interaction*, 1(1), 2012.
- [36] S. ur Réhman and L. Liu. ifeeling: Vibrotactile rendering of human emotions on mobile phones. In *Mobile multimedia processing*, pages 1–20. Springer, 2010.
- [37] S. Yilmazyildiz, R. Read, T. Belpeame, and W. Verhelst. Review of semantic-free utterances in social human–robot interaction. *International Journal of Human-Computer Interaction*, 32(1):63–85, 2016.