Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 71–80
https://doi.org/10.54941/ahfe1008242

Design of a Simplified Humanoid Robot Face: Cultural Differences and Emotional Synchronization

Meina Tawaki[1], Ichi Kanaya[1], and Keiko Yamamoto[2]

[1]Graduate School of Integrated Science and Technology, Nagasaki University, Nagasaki, 852-8521, Japan

[2]Department of Design Engineering and Technology, Tokyo Denki University, Tokyo, 120-8551, Japan

ABSTRACT

Humanoid robots are expected to interact with people in everyday environments, where facial expressions serve as an important nonverbal interface for conveying emotions and social intentions. However, highly human-like appearances can cause discomfort, a phenomenon known as the Uncanny Valley. To address this issue, this study investigates simplified robot faces that prioritize readability, emotional comfort, and the artifact-like nature of robots over full human resemblance. We examined basic principles of facial expression design from three perspectives: simplified facial elements, cultural differences in emotion recognition, and emotional synchronization. Expressions were built from only two components, eyes and a mouth, each rendered as a simple geometric shape and systematically deformed by parameters such as openness and inclination. In a questionnaire experiment, university students in Denmark (n = 80) and Japan (n = 51) classified the perceived emotion of each expression using Ekman's six basic emotions plus two neutral categories. Happiness, anger, sadness, and neutral expressions were assigned to similar parameter configurations in both countries, suggesting that simplified geometric expressions may be interpreted consistently across these cultures. In contrast, disgust, fear, and surprise showed cultural differences, indicating that these emotions depend more on context. These findings suggest that existing emotion categories remain usable for highly reduced faces. We also introduce the concept of emotional synchronization, a shared emotional state between individuals that appears as matching facial expressions when two people feel the same way. An exploratory experiment using romance *manga* examined a method for observing emotional synchronization. Detecting such subtle expressions could ultimately allow robots to express and respond to emotions more naturally, making human–robot communication easier.

Keywords: Humanoid robot, Facial expression design, Uncanny valley, Cross-cultural emotion recognition, Emotional synchronization

INTRODUCTION

Robots emerged as objects of modern technology in the early 1960s, and robots used in production systems advanced rapidly in the 1970s. In the 1990s their use expanded from manufacturing to the service sector, with

Received October 15, 2026; Revised November 24, 2026; Accepted December 5, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

examples such as Honda's ASIMO and Sony's AIBO, and since the late 1990s research and development of humanoid robots has flourished (Doi, 2010; Inoue et al., 2004; Shirai and Asada, 2001). Many humanoid robots are now being built, and they may support future daily life in many respects.

As humanoid robots develop, the Uncanny Valley, proposed by the robotics researcher Masahiro Mori in 1970, has become an important consideration (Mori, 1970; Mori et al., 2012). As an artifact approaches human appearance and motion, human affinity toward it first increases. Beyond a certain degree of resemblance, however, affinity drops sharply and turns into aversion; this sudden drop is called the valley. It is said that the valley can be overcome only when a robot resembles a human so closely that it is recognized as one. To ensure that robots are safe and give a sense of security as artifacts, we argue that robot design should be based on "morphology" and "body theory" (Kawasaki, 2008). The former is a design for reassurance that prioritizes safety, and the latter a design for safety that prioritizes reassurance.

We regard a humanoid robot as an artifact that works in place of humans in human society and sometimes communicates with them. When a robot communicates with people, its form functions as one of its interfaces. A robot needs functions and a form suited to its role, so it does not follow that a robot must look exactly like a human to fit human living spaces.

Human communication is broadly divided into verbal and nonverbal communication. The former relies on the meaning of words; the latter relies on gestures, facial expressions, and gaze, as well as vocal tone and quality. When verbal and nonverbal messages conflict, nonverbal cues such as facial expression and tone of voice can dominate the impression that is formed (Mehrabian and Wiener, 1967), which shows how important appearance is. Among these cues, facial expressions carry a particularly large amount of information and are central to interpersonal communication: an expression is a sign of the other person's emotional state, and people use this sign as an interface for conveying intentions. When a robot has to communicate with people, its face is therefore the most important interface, and the design of the robot's face is a crucial element of robot design (Breazeal, 2003). Which expressions can be read from a face, and whether this is universal across the world, are thus central questions for this study.

This study also formulates and tests a hypothesis about emotional synchronization. We define emotional synchronization as a state in which person A and person B have the same emotion at the same time. For example, when A develops a liking for B, B feels the same toward A at the same moment. Whether the subtle expressions of everyday life differ across people is unclear. By subtle expressions we mean expressions that all count as smiles but carry different emotions, such as a happy smile, a lonely smile, or a smile that suppresses anger. Instead of reading such expressions one by one, we assume it is possible to read whether the expressions of two people are synchronized. We expect this to help design more approachable robots.

The goal of this study is to realize facial expressions by reducing facial elements, representing their shapes geometrically, and deforming them geometrically. Our final aim is to clarify the direction of robot design as

artifacts and to establish a foundational science of robot design, in order to achieve coexistence of humans and artifacts based on safety and reassurance.

RELATED WORK

Preference for face-like figures. Humans show a visual preference for stimuli with a face-like structure. In experiments using very simple line drawings, infants looked longer at face-like figures than at figures that do not look like faces, even when they had little or no experience of faces (Fantz, 1961). This supports the special status of faces and suggests that the preference is innate.

Models of emotion. Russell (1980) proposed the circumplex model of affect, in which emotions are arranged on a circle spanned by a pleasant–unpleasant valence axis and an arousal–sleepiness axis, and the intensity of an emotion is represented by the distance from the origin. Ekman's basic emotions (Ekman, 1992) are widely used for classifying facial expressions, and Ekman and Friesen (1971) reported their recognition across cultures.

Emotion elicited by facial expressions. Seeing another person's expression changes several components of emotion, such as evaluation, action tendency, physiological and motor responses, and subjective feeling, so an expression can be regarded as an emotion-eliciting stimulus. The emotion elicited is often assumed to be the same as the one expressed, for instance happiness from a happy face. Such "emotional assimilation" does not always occur, however, and the same expression can elicit different emotions depending on the situation, because expression processing is a kind of person perception that depends on the perceiver, the target, their interaction, and the context (Suzuki, 2022).

Cultural differences. A meta-analysis found an in-group advantage in emotion recognition (Elfenbein and Ambady, 2002). Japanese observers tend to weight the eyes more, and North Americans the mouth, when judging emotion (Yuki et al., 2007), and internal representations of expressions differ between East Asian and Western observers, especially for fear, disgust, and surprise (Jack et al., 2012). These findings are directly relevant to a face that consists only of eyes and a mouth.

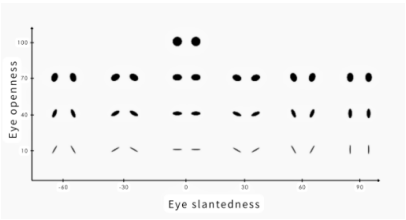
SIMPLIFIED FACE DESIGN

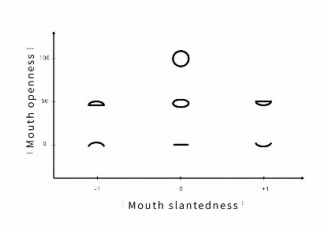
Rationale. Characters are the best example of artifacts that humans personify and empathize with. A character is made by personifying and deforming any concept, such as an animal, a tool, a natural phenomenon, or a country, and by simplifying and reducing its characteristic elements; this simplification is generally called characterization (McCloud, 1993). Rather than using robot faces that reproduce complex human expressions, we apply characterization to the robot face: the facial elements are reduced so that humans can personify the robot easily, while emotions can still be read.

Facial elements and base face. A face consists of two elements, the eyes and a mouth. We use the mouth rather than the nose because deformation of the mouth contributes much more to the formation of expressions than

deformation of the nose. Because deformation is considered, the geometric shapes of the eyes and the mouth are perfect circles, and they are arranged in an inverted triangle. Eyes normally have a pupil and a sclera, but we use only the pupil.

Parameters. Expressions are made by deforming the circles. For the eyes, changing the flatness of the circle expresses opening wide or narrowing the eyes, and changing the inclination expresses raised or lowered outer corners. For the mouth, changing the flatness expresses closing or opening the mouth, and tilting or bending the circle expresses raised mouth corners or a downturned mouth. We call these two kinds of parameters *openness* and *inclination*. For the eye designs, we created 19 different patterns. For the mouth designs, we created 7 different patterns. By combining these variations and adding 38 patterns without a mouth, we obtained a total of 171 expression patterns (Figures 1 and 2).





Figures 1 and 2: Geometric parameters: eye and mouth parameters. Facial expressions are created by adjusting openness and inclination.

EXPERIMENT 1: CROSS-CULTURAL PERCEPTION

Purpose and method. The purpose was to reduce the elements of a face and to find out which factors allow people to read emotion from it, and whether expressions, as a means of conveying emotion, differ across cultures. We created expressions by varying the facial elements with the parameters above and asked participants by questionnaire which emotion they perceived in each expression.

Emotion categories. Because the survey was designed for cross-cultural use, we used Ekman's six basic emotions (happiness, anger, sadness, disgust, fear, and surprise) to classify the perceived emotion. We added two neutral categories: neutral (+), which does not give the partner an unpleasant feeling, and neutral (−), which does. Neutral expressions were included because robots with static faces are considered.

Participants. The questionnaire was administered to 80 students at Aalborg University in Denmark and 51 students at the University of Nagasaki in Japan. Denmark was chosen for three reasons: modern European design often originates in Germanic-language regions, Denmark is one of the regions where modern design is highly favored, and its population is relatively homogeneous compared with the Netherlands or Sweden.

Participants selected, for each expression, the emotion that applied to them from the eight options shown in the upper right corner of the form

(Figure 3). Each participant answered all nine sheets. The questionnaires were carried out in English (Denmark) and in Japanese (Japan).

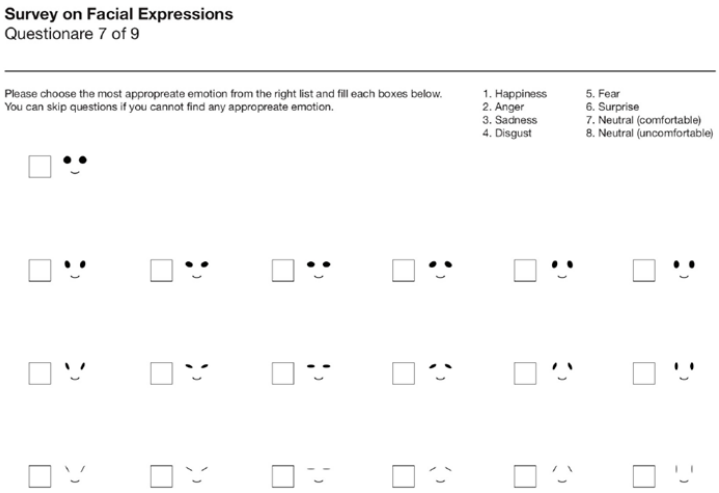


Figure 3: Survey Form: The questionnaire consisted of nine sheets, each showing 19 different expressions (171 expressions in total).

RESULTS OF EXPERIMENT 1

Figure 4 compares the two countries. For happiness, anger, sadness, and neutral (+) and (−), the mean parameter values were close between the two countries. For disgust, fear, and surprise, they differed.

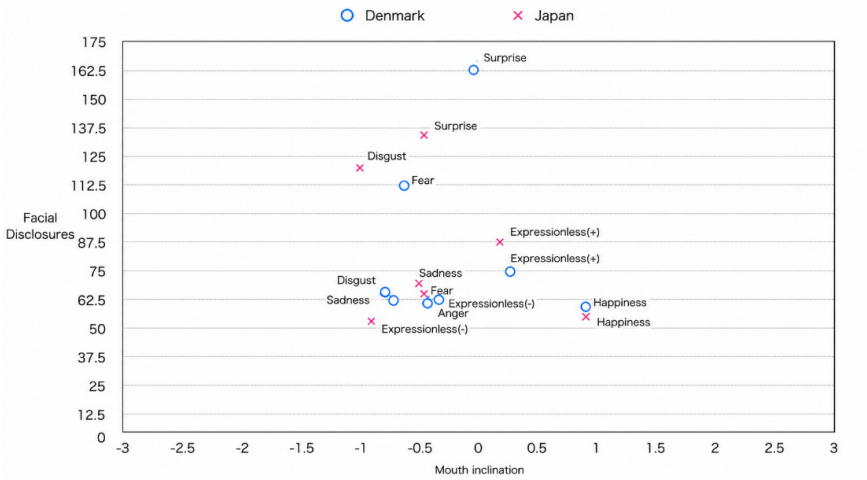


Figure 4: Scatter plot of the mean parameter values for each emotion in Denmark (blue circles) and Japan (red crosses): Each marker represents the mean of the expressions that participants selected for one emotion. The vertical axis is the sum of eye openness and mouth openness, and the horizontal axis is the inclination of the mouth.

Because the mean values were close for several emotions, a common expression design may be usable in Denmark and Japan for these emotions.

EXPERIMENT 2: EMOTIONAL SYNCHRONIZATION

Purpose. No cultural differences were found in how expressions are formed between Japan and Denmark for several emotions, but differences in the way emotions are expressed, such as in emoticons, still exist between Western countries and Japan (Yuki et al., 2007). The subtle expressions of daily life remain unclear. Instead of reading them individually, we examined whether the expressions of two persons are synchronized (Figure 5).

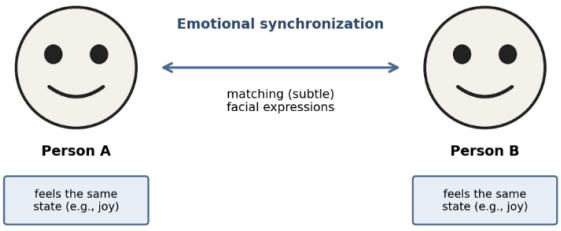


Figure 5: Concept of emotional synchronization: a shared emotional state between two people appears as matching facial expressions.

Material. We asked several participants to read the same *manga* that features multiple characters and to mark the panels in which the feelings of characters A and B "synchronized". The *manga* had to satisfy five conditions: (1) it is a romance *manga*, so that emotional change between characters and synchronization are easy to identify; (2) the characters have rich facial expressions; (3) an English translation has been published, so the survey can be repeated abroad; (4) romance begins within the first three volumes, so that participants can keep their concentration; and (5) the quality of the drawing is consistently high. *High School Debut* (Kawahara 2004) satisfied these conditions.

Participants and procedure. Ten students of a character design department took part, because they are used to visual expression and are strongly interested in character design. Each received questionnaire sheets covering volumes 1 to 3, with serial numbers assigned to every panel of the *manga*. The sheet asked for gender, age, and, for each scene in which they felt that emotions were synchronized, the volume, page, panel, and character names. We explained the concept of emotional synchronization and asked participants to answer from the panels, not from the flow of the story.

Results. Nine of the ten participants responded (six women and three men). All panels were given serial numbers, and panels selected by two or more participants were tallied (Figure 6). Panel 1852 received the most responses, followed by panels 2341 and 2622. In all three panels, the heroine, Haruna, confesses her feelings, and many participants judged that the emotions of the two characters were synchronized in these scenes.

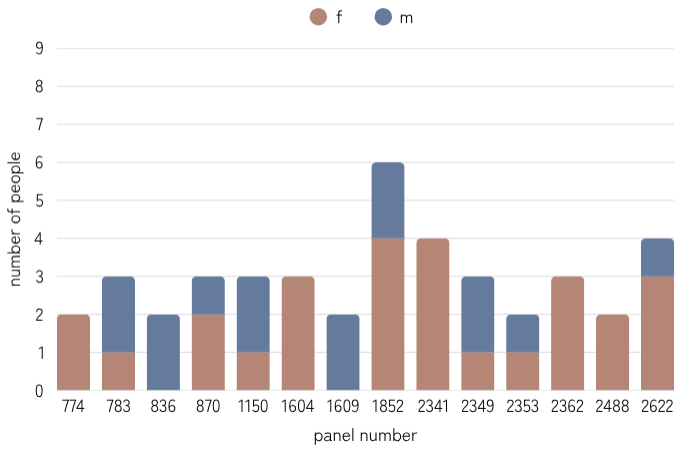


Figure 6: Panels in which two or more participants reported that the emotions of the two characters were synchronized: Each bar shows the number of respondents for one panel. The horizontal axis is the serial number of the panel, and the vertical axis is the number of respondents. Bar colors indicate participant gender (f, female; m, male).

DISCUSSION

Common interpretation of core emotions. Happiness, anger, sadness, and neutral (+) and (−) expressions were assigned to similar parameter values in Denmark and Japan. A plausible reason is that these emotions can be signaled by a few salient features, such as the bend of the mouth and the inclination of the eyes, which survive extreme simplification. This suggests that existing emotion categories remain usable for simplified expressions and supports a common expression design for these emotions in at least Denmark and Japan. Because Denmark was chosen as a Germanic-language, design-oriented region, the same design may also work in other Germanic-language countries, although this remains a hypothesis to be tested.

Context-dependent emotions. Disgust, fear, and surprise differed between the two countries. In realistic faces these emotions rely on subtle cues, such as nose wrinkling, raised eyebrows, or a specific combination of wide eyes and an open mouth, that our two-element faces cannot show. Without those cues, observers may rely on culturally learned conventions or on context. Two explanations should be tested: East Asian and Western observers weight the eyes and the mouth differently (Yuki et al., 2007) and form less similar representations of these emotions (Jack et al., 2012); and Japanese participants may be familiar with *manga* conventions for such expressions (McCloud 1993). For a robot used across cultures, simplified faces therefore seem suitable for the core emotions, whereas for disgust, fear, and surprise the face should be supported by other channels such as speech, gesture, or situational context. This study measured recognition, not comfort, so the claim that simplified faces avoid the Uncanny Valley still has to be verified with affinity ratings.

Emotional synchronization. Responses differed among individuals. The panels selected most often were scenes in which the heroine confesses, which suggests that scenes with a clear emotional turning point are easier to perceive as synchronized. This may reflect different understanding of the conventions of *Shojo manga*, the genre of the material. The result may also depend on the characteristics of the participants, who were character design students, so the survey should be expanded to other groups. Nevertheless, the method gave us a way to examine how subtle expressions between two persons are read, and, because synchronization between characters could be read, the method may be applicable to robot face design. Because people spontaneously mimic others' expressions (Dimberg et al., 2000) and emotions spread between people (Hatfield et al., 1993), a robot that detects subtle changes in its partner's expression and responds with a matching one could foster synchronization.

Limitations. Only two countries and only university students were studied, stimuli were static, and the response was a forced choice among eight labels. The absence of a large difference does not prove equivalence, so equivalence tests and larger samples are needed. The *manga* experiment used a small sample (nine responses) and a stylized medium.

FUTURE RESEARCH PLAN

We plan to proceed in four phases (Table 1). The first two extend and refine the cross-cultural findings, the third widens the synchronization study, and the fourth moves toward implementation on robots.

Table 1: Planned research phases.

Phase	Objective	Method	Expected Outcome
1	Extend Experiment 1 to other regions	Repeat the survey in other Germanic-language and non-Germanic regions; use equivalence tests; map parameters to a valence–arousal space (Russell 1980)	Finer cultural differences and common design rules
2	Handle disgust, fear, and surprise	Add contextual cues (gesture, color, speech); compare localized and common designs	Guidelines for context-dependent emotions
3	Widen the synchronization study	Use participants outside design departments and in other countries with the English edition; analyze gender and individual differences	A method usable for cultural comparison
4	Implement on robots	Animate the parameters; measure recognition and affinity; detect subtle expressions and test a robot that responds	Evidence on readability, comfort, and synchronization

In the long term, we aim to use these findings for the face design of humanoid robots, so that people of different cultures can read the robot's emotions and feel at ease with it, as a step toward a foundational science of robot design.

CONCLUSION

This study aimed to clarify the direction of humanoid robot face design. We described cultural differences in reading simplified expressions. The eyes and the mouth, the minimum elements that can be recognized as a face, were deformed from a perfect circle by the parameters of openness and inclination, and questionnaires asking participants to select an emotion category were administered in Denmark and Japan. The results suggest that existing emotion categories remain usable for simplified expressions. Happiness, anger, sadness, and neutral (+) and (−) were assigned to similar parameter values in both countries, so a common expression design can probably be used at least in Japan and Denmark, and possibly in other Germanic-language countries. In contrast, disgust, fear, and surprise differed between the two countries, suggesting that these emotions depend more on context and cannot be conveyed by two-element faces alone. We also examined emotional synchronization, reading whether the expressions of two persons are synchronized instead of reading subtle expressions individually. In the manga experiment, participants most often reported synchronization in scenes where the heroine confesses her feelings, although responses varied among individuals. The experiment gave us a method for examining how subtle expressions between two persons are read. This method may serve as a means of comparing Japan and other countries. Future work will repeat the surveys in other regions to find finer differences in expression perception and apply them to humanoid robot face design.

ACKNOWLEDGMENT

This work was supported by JST SPRING, Grant Number JPMJSP2172, and JSPS KAKENHI Grant Number 25H01228.

REFERENCES

Breazeal, C. (2003). Emotion and sociable humanoid robots. International Journal of Human-Computer Studies, 59(1–2), pp. 119–155.

Dimberg, U., Thunberg, M. and Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. Psychological Science, 11(1), pp. 86–89.

Doi, S. (2010). Research on Robot Face Design for Humanoids: Elemental Design for Facial Expression Composition Methodology, Master's thesis, Graduate School of Engineering, Osaka University (in Japanese).

Ekman, P. (1992). An argument for basic emotions. Cognition and Emotion, 6(3–4), pp. 169–200.

Ekman, P. and Friesen, W.V. (1971). Constants across cultures in the face and emotion. Journal of Personality and Social Psychology, 17(2), pp. 124–129.

Elfenbein, H.A. and Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. Psychological Bulletin, 128(2), pp. 203–235.
Fantz, R.L. (1961). The origin of form perception. Scientific American, 204(5), pp. 66–72.
Hatfield, E., Cacioppo, J.T. and Rapson, R.L. (1993). Emotional contagion. Current Directions in Psychological Science, 2(3), pp. 96–99.
Inoue, H., Kanade, T., Anzai, Y. and Sena, H. (2004). Creation of robotics. Tokyo: Iwanami Shoten (in Japanese).
Jack, R.E., Garrod, O.G.B., Yu, H., Caldara, R. and Schyns, P.G. (2012). Facial expressions of emotion are not culturally universal. Proceedings of the National Academy of Sciences, 109(19), pp. 7241–7244.
Kawahara, K. (2004). Koko debut [High School Debut]. Tokyo: Shueisha.
Kawasaki, K. (2008). Words of design: "Ro". AXIS, 135, pp. 86–87 (in Japanese).
McCloud, S. (1993). Understanding comics: the invisible art. New York: HarperCollins.
Mehrabian, A. and Wiener, M. (1967). Decoding of inconsistent communications. Journal of Personality and Social Psychology, 6(1), pp. 109–114.
Mori, M. (1970). The uncanny valley. Energy, 7(4), pp. 33–35 (in Japanese).
Mori, M., MacDorman, K.F. and Kageki, N. (2012). The uncanny valley [from the field]. IEEE Robotics & Automation Magazine, 19(2), pp. 98–100.
Russell, J.A. (1980). A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), pp. 1161–1178.
Shirai, Y. and Asada, M. (2001). Robots becoming familiar. Osaka: Osaka University Press (in Japanese).
Suzuki, A. (2022). Emotion elicitation by facial expression images: a perspective from facial expression (person) perception research. Japanese Journal of Research on Emotions, 29(2–3), pp. 58–63 (in Japanese).
Yuki, M., Maddux, W.W. and Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. Journal of Experimental Social Psychology, 43(2), pp. 303–311.