

Generation of Facial Expression Emphasized with Cartoon Techniques Using a Cellular-Phone-Type Teleoperated Robot with a Mobile Projector

Yu Tsuruda, Maiya Hori, Hiroki Yoshimura, and Yoshio Iwai

Graduate School of Engineering, Tottori University
101 Minami 4-chome, Koyama-cho, Tottori, 680-8550 Japan

Abstract. We propose a method for generating facial expressions emphasized with cartoon techniques using a cellular-phone-type teleoperated android with a mobile projector. Elfoid is designed to transmit the speaker's presence to their communication partner using a camera and microphone, and has a soft exterior that provides the look and feel of human skin. To transmit the speaker's presence, Elfoid sends not only the voice of the speaker but also emotional information captured by the camera and microphone. Elfoid cannot, however, display facial expressions because of its compactness and a lack of sufficiently small actuator motors. In this research, facial expressions are generated using Elfoid's head-mounted mobile projector to overcome the problem. Additionally, facial expressions are emphasized using cartoon techniques: movements around the mouth and eyes are emphasized, the silhouette of the face and shapes of the eyes are varied by projection effects, and color stimuli that induce a particular emotion are added. In an experiment, representative face expressions are generated with Elfoid and emotions conveyed to users are investigated by subjective evaluation.

1 Introduction

Video-conferencing and videotelephony are used as tools for communication between people in remote areas. These systems use not only the voice of the speaker but also a video to make communication smoother. A technology called telepresence has been devised to increase the human presence by increasing the resolution of the video or presenting video to an immersive display. In terms of human presence, however, conventional telepresence systems are insufficient because they do not have dimensional information and tactile feedback of the speaker. Therefore, for communication with people in remote locations, robots that have human appearance have been developed. Some studies use a humanoid robot for transmission of human presence. In particular, teleoperated android robots, such as Geminoid F and Geminoid HI-1 [1], have appearances similar to an actual person, and were intended to transfer the presence of actual people. These humanoid robots have high degrees of freedom and can transfer the human presence. However, they are expensive and limited to a specific individual

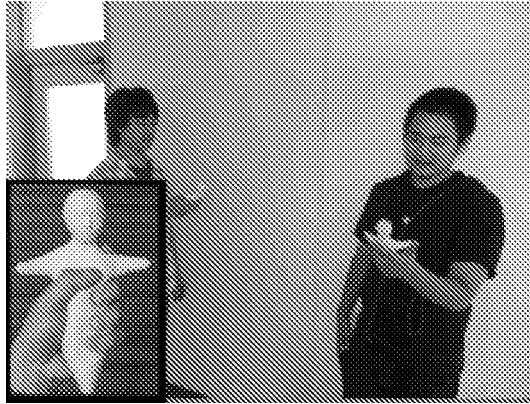


Fig. 1. Communication using Elfoid, which conveys the human presence to remote locations

target. A robot called Telenoid R1 has been developed to reduce the number of actuators and costs involved [2]. Telenoid is not limited to a specific individual target and is designed to appear and to behave as a minimalistic human at the very first glance. A person can easily recognize Telenoid as a human while the Telenoid does not appear to be gender- or age-specific. With this minimal design, Telenoid allows people to feel as if a distant acquaintance is next to them. Moreover, Telenoid's soft skin and doll-like body size make it easy to hold. However, it is difficult to carry by hand on a daily basis.

To allow daily use, a communication medium that downsizes Telenoid and has mobile-phone communication technology is now under development. The resulting Elfoid is easy to hold in the hand like a cellular phone as shown in Fig. 1. When we use such robots for communication, it is important to convey the facial expressions of a speaker and thus increase the modality of communication. In this research, facial expressions are generated using Elfoid's head-mounted mobile projector.

2 Related Work

In communication between people, it is important to convey the emotions of the speaker. There have been a considerable number of studies on basic human emotions [3–6]. Ekman et al. [6] defined basic facial expressions for anger, disgust, fear, happiness, sadness and surprise. This shows that the facial expressions of emotion are not culturally determined but are universal across all human cultures and are thus biological in origin. In communication, these emotions are important in efforts to increase the modality.

Studies [7–9] have used a communication robot to convey emotions. However, it is difficult for communication robots to generate facial expressions like a human face. To overcome this problem, the studies displayed colors with motions of

a robot for communication between humans and robots. Elfoid also cannot generate facial expressions like a human face because it has a compact design that cannot be activated intricately. That is, since a priority of Elfoid is portability, the modality of communication is less than that of Telenoid. For this reason, it is necessary to convey emotions some other way.

In this research, facial expressions are generated using Elfoid's head-mounted mobile projector. If the speaker's facial movements estimated employing conventional face-recognition approaches are accurately regenerated with Elfoid, the human presence can be conveyed. However, even if a captured face image is projected directly, details of facial expression cannot be conveyed because the resolution of projection is low. According to the facial action coding system (FACS) [6], which describes relationships between emotion and facial movement, movements around the mouth and eyes play important roles. However, some emotions cannot be conveyed even if the movements around the mouth and eyes are emphasized [10]. In this study, facial expressions are emphasized using cartoon techniques [11]. It is widely recognized that cartoons have a strong advantage in expressing emotions and feelings.

3 Generation of Facial Expression Using Elfoid with a Projector

3.1 Elfoid Characteristics

Elfoid is used as a cellular phone for communication as shown in Fig.1. To convey the human presence, Elfoid has the following functions.

- Elfoid has a body that is easy to hold in the hand.
- Elfoid's design is recognizable at first glance to be nothing more than a human and is capable of being interpreted equally as male or female, old or young.
- Elfoid has a soft exterior that provides a feeling of human skin.
- Elfoid is equipped with a camera and microphone.

Additionally, a mobile projector is mounted in Elfoid's head and a facial expression is generated by projecting images from within the head as shown in Fig.2.

The procedure of the proposed method is as follows. First, individual facial images are captured by a camera mounted within Elfoid. Next, facial expression is recognized using a conventional method such as the use of a point distribution model [12][13]. Effects that induce a particular emotion are added to the image. Finally, the generated image is projected on the face of Elfoid from within.

3.2 Generation of Facial Expressions Using Cartoon Techniques

In this study, facial expressions are emphasized using cartoon techniques [11]. It is widely recognized that cartoons have a strong advantage in expressing emotions and feelings. According to FACS [6], which describes relationships between

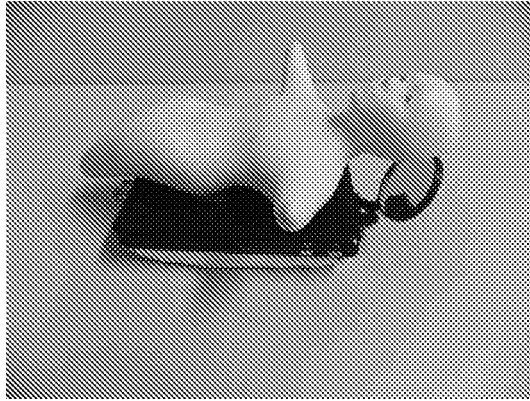


Fig. 2. A mobile projector is mounted in Elfoid's head

emotion and facial movement, features around the mouth and eyes play important roles. In this study, facial expressions associated with the speaker showing basic emotions defined by Ekman et al. [6] are generated according to the following rules.

- Angry expression
Eyes are moved to the center. Corners of the mouth are lowered. Complexion is changed to red.
- Disgusted expression
Only one eye is moved to the center. Only one corner of the mouth is raised. Complexion is changed to blue-green.
- Fearful expression
Eyes are moved to the center and each inside is raised. Mouth is opened. Complexion is changed to blue.
- Happy expression
Shapes of eyes are varied so that the positions of the cheeks are raised. Corners of the mouth are raised. Cheek color is associated with blushing.
- Sad expression
Eyes are moved to the center and each inside is raised. Corners of the mouth are lowered. Symbols of tears are added to the eyes.
- Surprised expression
Eyes are opened wide. Movement of the mouth is increased.

The cartoon techniques that are used in this study are shown in Fig. 3. The movements around the mouth and eyes are emphasized. Moreover, color stimuli that induce a particular emotion are added.

3.3 Generation of Projection Images for Elfoid

In this study, projection images are generated to backproject Elfoid's face with consideration of Elfoid's material. To generate a facial expression with Elfoid,

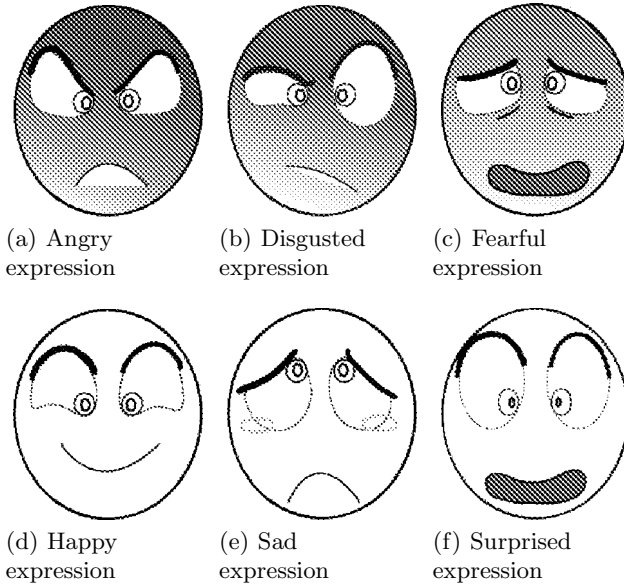


Fig. 3. Facial expressions conveyed by cartoon techniques

calibration with respect to both projection position and projection color is needed. As for the projection color, color variation after projection is investigated by projecting various colors in advance. As for the projection position, three-dimensional positions on the surface of Elfoid are investigated by projecting an image that has a number of known two-dimensional positions. The correspondence between the three-dimensional positions on Elfoid's surface and two-dimensional positions in the projection image can be obtained by this calibration. Facial expressions are generated on the surface of Elfoid using the characteristics described in 3.2, and a projection image is generated using the result of the calibration.

4 Experiment

To investigate the emotions that are derived from each facial expression, we conducted experiments with Elfoid.

4.1 Generation of Facial Expressions with Elfoid

In an experiment, facial expressions associated with the speaker showing basic emotions defined by Ekman et al.[6] are generated. The emotions are anger, disgust, fear, happiness, sadness, and surprise. Figure 4 shows the generated projection images. An animation is generated varying from a normal expression

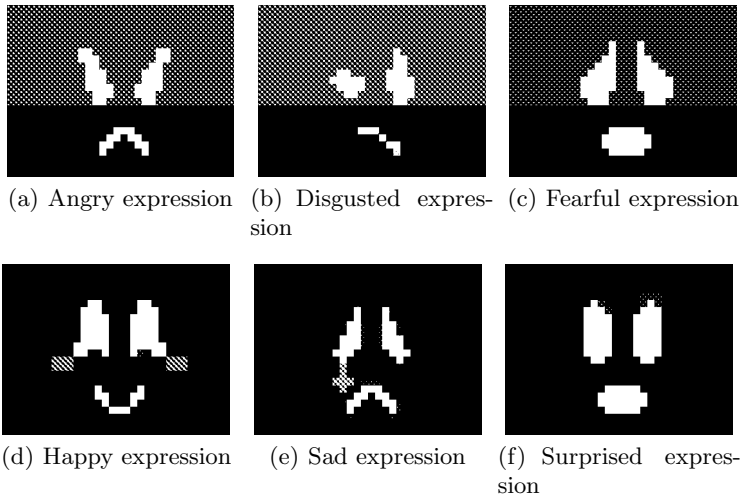


Fig. 4. Projection images

to the generated expression using morphing technology. Figure 5 shows the facial expressions generated with Elfoid.

It was found that the projection is seamless in spite of the complicated surface of Elfoid.

4.2 Subjective Evaluation of Emotional Conveyance

In this experiment, an emotion conveyed to users is investigated by subjective evaluation. Six facial patterns were presented to 28 subjects in random order. Each subject was asked to rate a facial pattern for six emotions: anger, disgust, fear, happiness, sadness, and surprise. Each emotion was rated from 1 (not conveyed at all) to 6 (conveyed extremely strongly). The results of the subjective evaluation process for each facial expression are shown in Fig. 6. The items in these figures are the average score and standard variation of the subjective evaluation. Dunnett's test [14] is used to compare the average scores. Dunnett's test is a multiple-comparison procedure used to compare each of a number of treatments with a single control. The emotions conveyed when generated facial expressions are presented are discussed as follows.

– Angry expression

Figure 6(a) shows the emotions conveyed when the generated angry expression is displayed. The highest score is observed for the emotion “anger”; the average score of the subjects is 5.25, and Dunnett's test is performed with reference to this score. Dunnett's test indicates a significant difference between the score for “anger” and the scores for all other emotions. Therefore, the angry expression of Elfoid efficiently conveys “anger”.

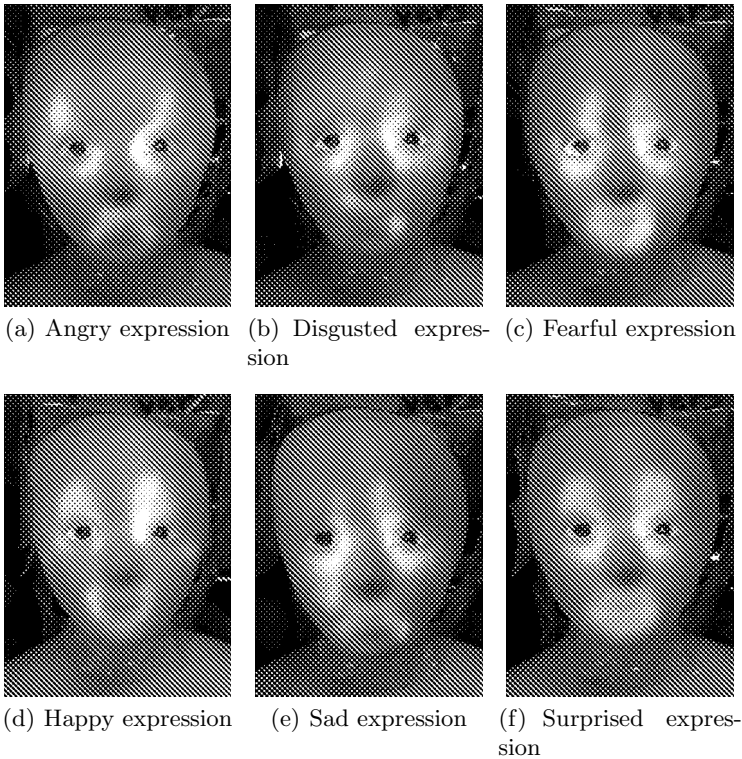


Fig. 5. Facial expressions generated with Elfoid

– Disgusted expression

Figure 6(b) shows the emotions conveyed when the generated disgusted expression is displayed. The highest score is observed for the emotion “disgust”; the average score of the subjects is 4.32, and Dunnett’s test is performed with reference to this score. Dunnett’s test indicates a significant difference between the score for “disgust” and the scores for all other emotions. Therefore, the disgusted expression of Elfoid efficiently conveys “disgust”.

– Fearful expression

Figure 6(c) shows the emotions conveyed when the generated fearful expression is displayed. The highest score is observed for the emotion “fear”; the average score of the subjects is 3.75, and Dunnett’s test is performed with reference to this score. Dunnett’s test indicates that there was no significant difference between the scores for “disgust” and “sadness”. Therefore, the fearful expression of Elfoid does not convey “fear” accurately. This may be because the shape of the eyes in the fearful expression gives a negative impression.

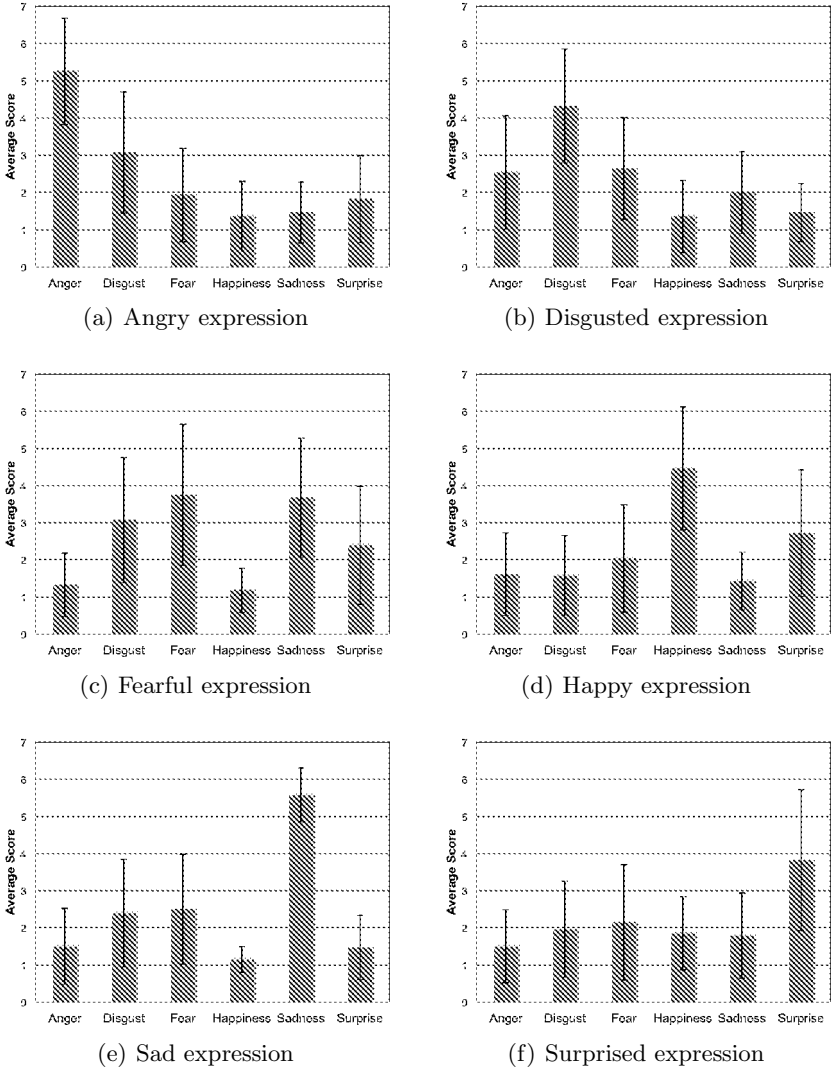


Fig. 6. Results of evaluation

– Happy expression

Figure 6(d) shows the emotions conveyed when the generated happy expression is displayed. The highest score is observed for the emotion “happiness”; the average score of the subjects is 4.46, and Dunnett’s test is performed with reference to this score. Dunnett’s test indicates a significant difference between the score for “happiness” and the scores for all other emotions. Therefore, the happy expression of Elfoid efficiently conveys “happiness”.

– Sad expression

Figure 6(e) shows the emotions conveyed when the generated sad expression is displayed. The highest score is observed for the emotion “sadness”; the average score of the subjects is 5.57, and Dunnett’s test is performed with reference to this score. Dunnett’s test indicates a significant difference between the score for “sadness” and the scores for all other emotions. Therefore, the sad expression of Elfoid efficiently conveys “sadness”.

– Surprised expression

Figure 6(f) shows the emotions conveyed when the generated surprised expression is displayed. The highest score is observed for the emotion “surprise”; the average score of the subjects is 3.82, and Dunnett’s test is performed with reference to this score. Dunnett’s test indicates a significant difference between the score for “surprise” and the scores for all other emotions. However, since the average score for “surprise” is low, the surprised expression needs to be improved.

5 Conclusion

We proposed a method for generating facial expressions emphasized with cartoon techniques using Elfoid with a mobile projector. In experiments, facial expressions were generated by backprojecting facial patterns to Elfoid’s face. Five facial expressions, but not a fearful expression, can be conveyed as the intended emotion. In the case of the fearful expression, negative emotions, such as “sadness” and “disgust”, are conveyed co-instantaneously with “fear”. In future work, we will implement an entire system to convey emotions with Elfoid according to the results of facial recognition studies.

Acknowledgment. This research was supported by the JST CREST (Core Research for Evolutional Science and Technology) research promotion program “Studies on cellphone-type tele-operated androids transmitting human presence”.

References

1. Asano, C.B., Ogawa, K., Nishio, S., Ishiguro, H.: Exploring the uncanny valley with geminoid HI-1 in a real-world application. In: Proc. Int’l Conf. of Interfaces and Human Computer Interaction, pp. 121–128 (2010)
2. Ogawa, K., Nishio, S., Koda, K., Balistreri, G., Watanabe, T., Ishiguro, H.: Exploring the natural reaction of young and aged person with Telenoid in a real world. *Jour. of Advanced Computational Intelligence and Intelligent Informatics* 15(5), 592–597 (2011)
3. Arnold, M.B.: *Emotion and Personality: Psychological aspects*. Columbia University Press (1960)

4. Plutchik, R.: *Emotion: A Psychoevolutionary Synthesis*. Harper & Row (1980)
5. Izard, C.E.: *The Psychology of Emotions*. Springer (1991)
6. Ekman, P., Friesen, W.: Facial action coding system: A technique for the measurement of facial movement. Consulting Psychologists Press (1978)
7. Sugano, S., Ogata, T.: Emergence of mind in robots for human interface - research methodology and robot model. In: *IEEE Int'l Conf. Robotics and Automation*, pp. 1191–1198 (1996)
8. Hiraiwa, A., Hayashi, K., Manabe, H., Sugimura, T.: Alter-ego interface technology. *NTT Technical Review* 1(8), 72–76 (2003)
9. Ariyoshi, T., Nakadai, K., Tsujino, H.: Effect of facial colors on humanoids in emotion recognition using speech. In: *Int'l Workshop on Robot and Human Interactive Communication*, pp. 59–64 (2004)
10. Hori, M., Takakura, H., Yoshimura, H., Iwai, Y.: Generation of facial expression for communication using elfoid with projector. In: Schwenker, F., Scherer, S., Morency, L.-P. (eds.) *MPRSS 2012. LNCS*, vol. 7742, pp. 27–34. Springer, Heidelberg (2013)
11. Thomas, F., Johnston, O.: *The Illusion of life: Disney animation*. Disney Press (1995)
12. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 23(6), 681–685 (2001)
13. Saragih, J.M., Lucey, S., Cohn, J.F.: Deformable model fitting by regularized landmark mean-shift. *Int'l Jour. of Computer Vision* 91(2), 200–215 (2011)
14. Dunnett, C.: New tables for multiple comparisons with a control. *Biometrics* 20(3), 482–491 (1964)