



# Emotion Synchronization Method for Robot Facial Expression

Yushun Kajihara, Peeraya Sripiian<sup>(✉)</sup>, Chen Feng, and Midori Sugaya

Shibaura Institute of Technology, 3-7-5, Toyosu, Koto-Ku, Tokyo 135-8504, Japan  
peeraya@shibaura-it.ac.jp

**Abstract.** Nowadays, communication robots are becoming popular since they are actively used in both commercially and personally. Increasing empathy between human-robot can effectively enhance the positive impression. Empathy can be created by syncing human emotion with the robot expression. Emotion estimation can be done by analyzing controllable expressions like facial expression, or uncontrollable expression like biological signals. In this work, we propose the comparison of robot expression synchronization with estimated emotion based on either facial expression or biological signal. In order to find out which of the proposed methods yield the best impression, subjective impression rating is used in the experiment. From the result of the impression evaluation, we found that the robot's facial expression synchronization using the synchronization based on periodical emotion value performs the best and best suitable for emotion estimated both from facial expression and biological signal.

**Keywords:** Emotion estimation · Empathy · Robot facial expression

## 1 Introduction

The communication robot market is expanded since they are used actively in many sites such as commercial facilities, medical or nursing care facilities, or even use personally at home. In order to improve the acceptance of communication robots, many aspects are considered. Nonverbal behavior is one of the essential factors for enhancing communication for human-human. Many robots could communicate; however, a few robots employ nonverbal behavior in communication such as using various facial expressions. The positive impression could be increased when the robot's expression is synchronized with human emotion, Misaki et al. [1]. Recently Kurono et al. [2] compared emotion estimation using facial expression and biological signals and found that biological signals result in a better concordance with the subjective evaluation. Also, Sripiian et al. [3] compared subjective impression evaluation toward robots with expression based on emotions estimated from either source and found that impression like intellectual and is higher when the robot's expression is synchronizing with emotion estimated from the biological signal. The robot expression is shown only one time after the emotion is estimated for a certain period.

## 2 Background

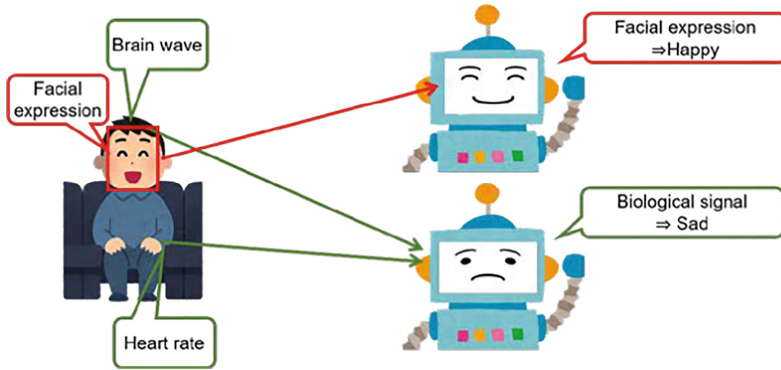
To express emotion on robots, Hirth et al. [4] developed a robot “ROMAN,” that can express six basic emotions consist of anger, disgust, fear, happiness, sadness, and surprise. The emotion state was calculated similarly to the method in Kismet-project [5]. However, the expressed emotion in the robot was not from the real understanding of human emotion at the time. There was no investigation for how human value robot’s emotion expression in the communication.

Emotion estimation is the process of identifying human emotion. Typically, the estimation can be done through observable expressions, such as facial expression that includes eyes, mouth, and facial muscles [6], or speech tone [7]. These expressions are carried by the somatic nervous system, which is a voluntary nervous system, hence, controllable by the sender. Meanwhile, emotion can also be estimated through unobservable expressions such as biological signals. The sender could not control biological signals such as heart rate and brain waves because it is driven by the automatic nervous system, which is an involuntary nervous system, or the unconscious mind.

In recent years, the means of estimating emotions based on biological signals have been actively studied. For example, the PAD model by Mehrabian et al. [8] that evaluated emotion by Pleasure (the degree of comfortableness to a particular event), Arousal (the degree of how active and bored one feels), and Dominance (how much control one has or how obedient one is). There are many studies that rely on Russell’s Circumplex Model of Affection [9], which related to Mehrabian’s PAD model. This model suggests that emotions are plotted on a circle on a two-dimensional coordinate axis; the Arousal axis and the Valence axis. The model has been widely used, for instant, Tanaka et al. [10] estimate emotion by associating brain waves to the Arousal axis and nasal skin temperature to the Valence axis. Accordingly, Ikeda et al. proposed a method to estimate emotion by correlating the value obtained from pulse sensors rather than nasal skin temperature with the Arousal axis [11]. They used the pNN50 calculated from pulse measurement. Figure 1 shows Russel’s Circumplex model of affection and emotion estimation used by works in [2, 3, 11].



**Fig. 1.** Emotion estimation on Russel’s circumplex model of affection



**Fig. 2.** The robot facial expression synchronization method. The emotion expressed on robot face is based on emotion estimated from biological signals (brain wave and heart rate) or from facial expression.

Kurono et al. [2] proposed the emotion classification method from biological signals and facial expression and compared with subjectively evaluated emotion. The emotion classification was based on the coordinate positioning of the Arousal and Pleasant axis on Russel's Circumplex Model of Affection. Although they found that biological signals performed better than facial expression, we concern that emotion mapping was not suitable. In many pieces of literature [10, 11], high arousal and high pleasant value are estimated as "Happy," while Kurono et al. [2] classified the emotion as "Surprise." Accordingly, low arousal and high pleasant are classified as "Happy," while other literature classified as "Relax." Therefore, the emotion mapping is corrected as shown in Fig. 1 in this work.

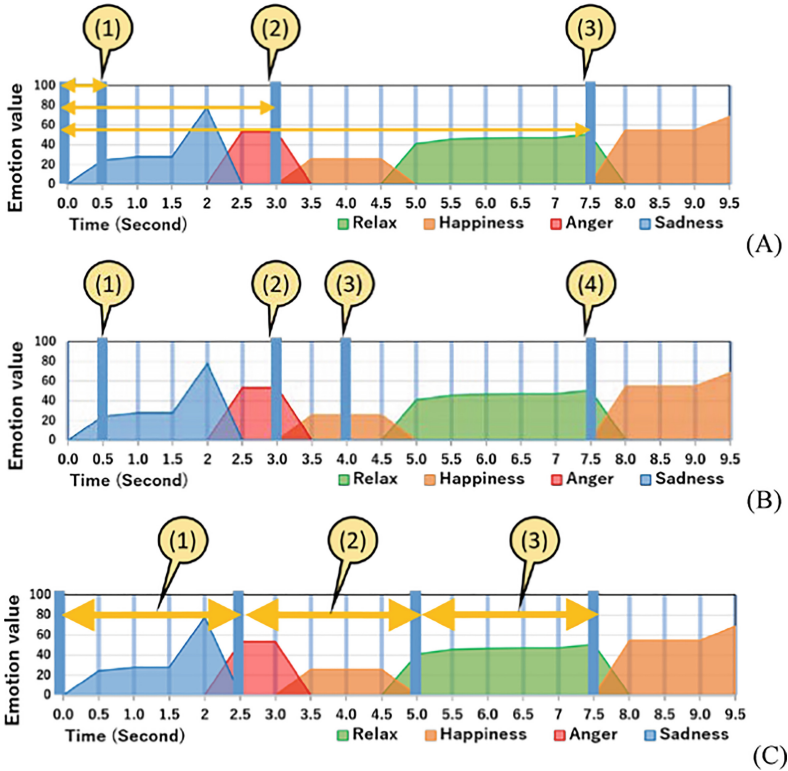
Also, the robot's facial expression did not synchronize in real-time with the user in Kurono et al. [2] Biological signals and facial expression are not static signals, and they tend to change many times within an event instantly. In order to achieve a real-time synchronization, we will verify the synchronization time interval for human facial expressions or biometric information about the timing to change the robot's facial expressions in time series.

### 3 Proposed Method

In order to achieve the final goal of creating empathy between humans and robots, we aim to develop the method proposed in [2, 3] by investigating methods for emotion synchronization with robot expression in this paper. Methods for emotion synchronization are proposed as follows;

1. Synchronization based on cumulative emotion value.
2. Synchronization based on one shot of emotion value.
3. Synchronization based on periodical emotion value.

All of the above methods are described in detail in the next section. In this work, we perform an experiment that compares subjective evaluation toward the robot's facial



**Fig. 3.** Proposed methods for synchronization of estimated emotion with robot expression

expression, synchronized with the emotion estimated from biological signals or facial expression using one of the three proposed methods. Before starting the experiment, the participant will answer a questionnaire regarding personal interest and knowledge in the robot according to Okada and Sugaya's findings [12]. Also, a questionnaire about self-control or nonverbal skill is conducted before the experiment. Finally, we use the SD method [13] for subjective's impression evaluation toward the robot expression, similar to [2, 3].

Biological signals are measured from brain waves and heart rate, while facial expression is taken from a camera. The synchronization of robot facial expression is depicted in Fig. 2.

## 4 Proposed Method

We propose three synchronization methods for robot expression. Figure 3 illustrates each proposed method accordingly.

#### 4.1 Synchronization Based on Cumulative Emotion Value

This method is based on Kurono et al.'s work [2]. As illustrated in Fig. 3 (A), the emotion is estimated by taking the cumulative of emotion value (cumulative from starting time to a particular time) for each emotion at the interval of 0.5 s, 3 s, and 7.5 s. These intervals are taken from [2, 3] since they yield an appropriate result for emotion classification in the experiment.

From Fig. 3 (A), at 0.5 s (1), shows that emotion "Sadness" is observed while other emotions are zero, so the robot will show "Sad" expression at this point. Meanwhile, at 3 s, the cumulative value of Anger is 106, and sad is 158, while other emotions are still zero, so the robot will show "Sad" expression. Finally, at 7 s, each emotion has a cumulative value of Happiness is 78, Anger is 106, Sadness is 158, and Relax is 278 accordingly. At this point, the robot will show "Relax" expression.

#### 4.2 Synchronization Based on One Shot of Emotion Value

This method show robot emotion based on the emotion value occurring at a particular timing. For an instant, at (1) the emotion "Sad" is shown on the robot's face, at (2), the emotion "Anger" is shown, at (3), the emotion "Happy" is shown, and at (4), the emotion "Relax" is shown. Figure 3 (B) illustrates this method.

#### 4.3 Synchronization Based on Periodical Emotion Value

As shown in Fig. 3 (C), the robot emotion is expressed based on the maximum cumulative value of a defined period. The example shows that for every 2.5 s, the emotion value is calculated. For example, (1) calculates cumulative emotion occurring from 0.0 to 2.5 (The robot would express the emotion "Sad") (2) calculates cumulative emotion occurring from 2.5 to 5.0 (The robot would express the emotion "Happy"). Finally, (3) calculates cumulative emotion occurring from 5.0 to 7.5 (The robot would express the emotion "Relax").

### 5 Experiment

We conducted a preliminary experiment by presenting an 80 s video clip that evokes the emotion of "Happy" to evaluate which of the three proposed synchronization method yield a better subjective evaluation.

#### 5.1 Subjects

Three students (two males and one female) age range from 18–21 participated in the experiment with a given consent.

## 5.2 Stimuli

The stimulus is an 80 s length of the video sequence, composed from 3 to 4 small video clips manually selected from annotated video clips database LIRIS-ACCEDE [14]. All selected small video clips have high scores in emotional and high alertness; therefore, they evoke “Happy” emotion. It is the largest video database currently in existence interpreted by an extensive population using induced emotional labels. The three synchronization methods are tested with all participants in a random manner. We prepared a total of 8 video sequences as stimuli.

## 5.3 Procedure

Before the experiment, the participants have to answer the pre-experiment questionnaire. The participant is asked to wear a brain wave sensor and a pulse sensor during the whole experiment. OMRON’s OKAO<sup>TM</sup> Vision is set on a table in front of the participant to detect the participant’s facial expression. Figure 4 shows a photo taken from the experiment. After begin retrieving all input data from all sensors, the experimental procedure is as follows.

1. The participant stays still (Rest) for 30 s for baseline measurement.
2. One of the video clips (80 s) is presented on the screen as the stimulus.
3. During the video clip presentation, the robot changes its facial expression according to the synchronization method, estimated emotion from either facial expression or biological signal.
4. The participant is asked to evaluate the impression of the robot during that trial using the SD method.
5. Repeat steps 2 to 4 until all video clips are presented.

## 5.4 Subjective Evaluation

To evaluate the participant’s impression on robot expression, we utilized the 12 adjective pairs that compose the Japanese property-base adjective measurement method [15, 16]. The impression rating on adjective pairs is done using Osgood’s Semantic Differential (SD) method [13], which is usually used to measure opinions, attitudes, and values on a psychometrically controlled scale. Similar to [3], we use the three attributes “Intimacy,” “Sociability,” and “Vitality” and selected four corresponding property-based adjectives for each attribute.



Fig. 4. The photo of experiment.

## 6 Results and Discussion

Figure 5 shows the average results of an impression evaluation questionnaire of the robot. From average impression evaluation scores, grouping into three main attributes. It can be implied that the robot’s facial expression synchronized with facial expression is rated with a higher impression score in “Vitality” and “Intimacy” attributes for most of the methods. Overall, robot expression that synchronizes with the participant’s facial expression appears to result in an overall higher impression rating score when the synchronization method B (moment) and C (5.0 cycle) were used.

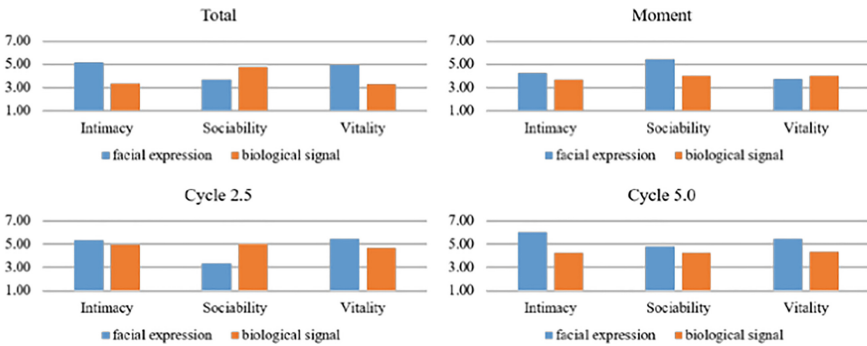
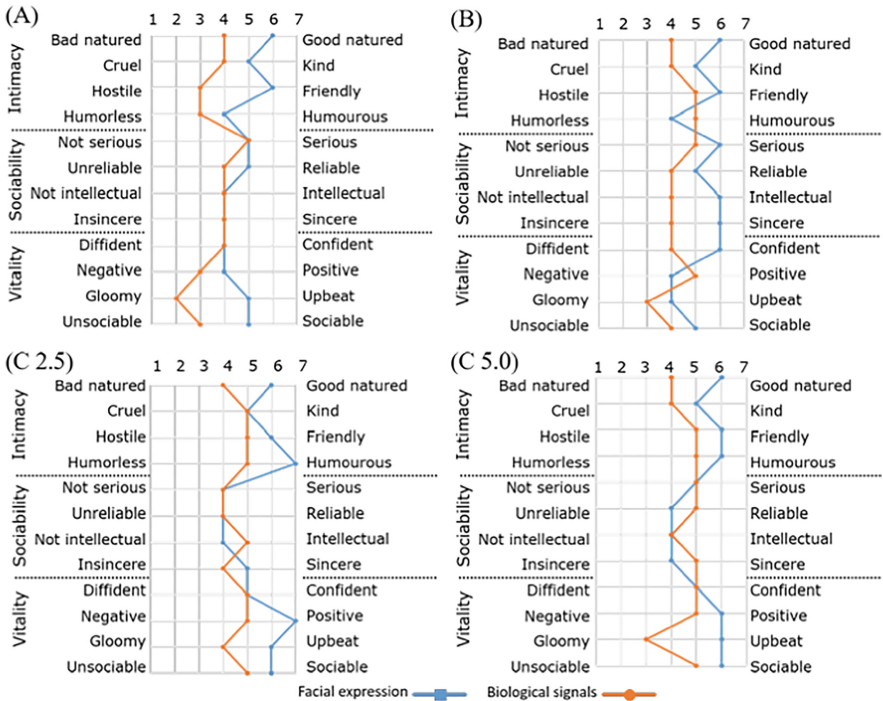


Fig. 5. Comparison of average (N = 3) impression evaluation scores of robot expression that synchronize with emotion estimated from facial expression and from biological signal. The impression evaluation scores are grouped into three attribute: intimacy, sociability, vitality.

However, the impression evaluation score is subjective. So we further investigate the individual impression rating score. Figure 6 shows the results of an impression evaluation questionnaire of the robot of the participant#2. For this participant, the robot’s facial expression synchronized with facial expression is rated with a higher impression score in “Vitality” attributes for most of the methods. So, we look at the emotional value of the participant#2 (Fig. 7) for more in-depth analysis. It was found that the emotion “Sad” and “Anger” are estimated frequently. Meanwhile, there are many times that neutral facial

expression is observed, which is estimated as “Relax” emotion, hence, the robot express “Relax” on its screen. In addition, the synchronization based on periodical emotion value (for every 2.5 s) results in a higher average score of impression rating when estimate emotion from facial expression. Similarly, the same synchronization method C, for every 5 s, results in a higher average score of impression rating when estimate emotion from biological signals.



**Fig. 6.** Comparison of impression rating of participant#2 on the robot’s expressed emotion estimated from facial expression and Biological signals. (A) shows the impression rating for emotion synchronized based on cumulative emotion value. (B) shows the impression rating result for emotion synchronized based on one shot of emotion value. (C) shows the impression rating result for emotion synchronized based on periodical emotion value, 2.5 s and 5.0 s, respectively.

From the experiment results, it may be possible to imply that the pleasant emotion like “Happy” and “Relax” is related to “Intimacy” attribute in the robot impression. Also, during the experiment, many unpleasant emotions are evoked by looking at the video clips even though we manually picked “Happy” emotion video from the database as the stimuli. This could be because all of our participants are Japanese. Therefore, cultural differences or language barriers could occur because some of the video clips contain English dialogues or events that are mutually understandable only in western culture. It could be assumed that the stimuli may not be suitable for Japanese participants.

For the cumulative emotion value used in the synchronization method A, it appears that if one of the emotion is frequently estimated, the result would be biased toward



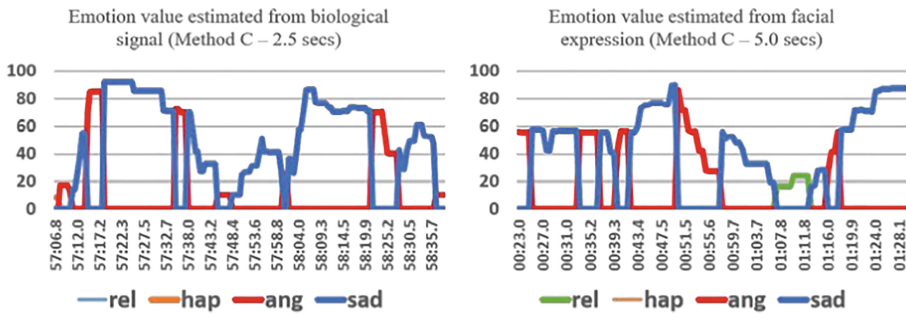


Fig. 7. Emotion value estimated from biological signal and facial expression for participant#2

that emotion. This could fix the robot’s facial expression toward only one type of emotion. Therefore, “Intimacy” and “Vitality” attributes are rated lower than other methods. Meanwhile, we observed from the participants’ free comments that the robot expression changes too quickly when synchronized with method B. Therefore, almost all items in the impression rating is given a rather low score for this method.

Based on these results, we consider that it is better to use the synchronization method C for robot facial expression with emotion estimated from both facial expression and biological signals.

## 7 Conclusion and Future Work

We proposed three methods for robot facial expression synchronization with estimated emotion from facial expression or biological signals. A preliminary experiment was performed to investigate the best suitable methods that gave the highest impression rating toward the robot expression. From the result of the impression evaluation, the robot’s facial expression synchronization using the synchronization based on periodical emotion value performs the best, hence suitable for emotion estimated both from facial expression and biological signal.

There are several considerations during the experiment. For instant, the emotion-induced video database may not be suitable for Japanese participants due to cultural differences. The number of participants is low. Also, a non-verbal evaluation index (SAM [17]) could be used toward to robot in addition to the SD method. In the future, the main experiment could be performed with more participants, using more suitable stimuli and collect subjective evaluation from more post-experiment questionnaires.

## References

1. Misaki, Y., Ito, T., Hashimoto, M.: Proposal of human-robot interaction method based on emotional entrainment. In: HAI Symposium (2008). (in Japanese)
2. Kurono, Y., Sripian, P., Chen, F., Sugaya, M.: A preliminary experiment on the estimation of emotion using facial expression and biological signals. In: Kurosu, M. (ed.) HCII 2019. LNCS, vol. 11567, pp. 133–142. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-22643-5\\_10](https://doi.org/10.1007/978-3-030-22643-5_10)

3. Sripian, P., et al.: Study of empathy on robot expression based on emotion estimated from facial expression and biological signals. In: The 28th IEEE International Conference on Robot & Human Interactive Communication, New Delhi, India (2019). IEEE
4. Hirth, J., Schmitz, N., Berns, K.: Emotional architecture for the humanoid robot head ROMAN. In: IEEE International Conference on Robotics and Automation, pp. 2150–2155. IEEE (2007)
5. Breazeal, C., Scassellati, B.: A context-dependent attention system for a social robot. *rn* **255**, 3 (2003)
6. Ekman, P., Friesen, W.V.: Facial Action Coding System: Investigator's Guide. Consulting Psychologists Press, Palo Alto (1978)
7. Nwe, T.L., Foo, S.W., De Silva, L.C.: Speech emotion recognition using hidden Markov models. *Speech Commun.* **41**(4), 603–623 (2003)
8. Mehrabian, A.: Basic Dimensions for a General Psychological Theory: Implications for Personality, Social, Environmental, and Developmental Studies. Oelgeschlager, Gunn & Hain, Cambridge (1980)
9. Russell, J.A.: A circumplex model of affect. *J. Pers. Soc. Psychol.* **39**(6), 1161 (1980)
10. Tanaka, H., Ide, H., Nagashima, Y.: An attempt of feeling analysis by the nasal temperature change model. In: SMC 2000 Conference Proceedings, 2000 IEEE International Conference on Systems, Man And Cybernetics. 'Cybernetics Evolving to Systems, Humans, Organizations, and Their Complex Interactions' (cat. no. 0. 2000. IEEE
11. Ikeda, Y., Horie, R., Sugaya, M.: Estimate emotion with biological information for robot interaction. In: 21st International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES-2017), Marseille, France, pp. 6–8 2017
12. Okada, A., Sugaya, M.: Interaction design and impression evaluation of the Person and the active robot. In: Human Computer Interaction (HCI), pp. 1–6 (2016). (in Japanese)
13. Osgood, C.E.: Semantic differential technique in the comparative study of cultures. *Am. Anthropol.* **66**(3), 171–200 (1964)
14. Baveye, Y., et al.: LIRIS-ACCEDE: a video database for affective content analysis. *IEEE Trans. Affect. Comput.* **6**(1), 43–55 (2015)
15. Hayashi, F.: The fundamental dimensions of interpersonal cognitive structure. *Bull. Fac. Educ. Nagoya Univ.* **25**, 233–247 (1978)
16. Hayashi, R., Kato, S.: Psychological effects of physical embodiment in artificial pet therapy. *Artif. Life Robotics* **22**(1), 58–63 (2017). <https://doi.org/10.1007/s10015-016-0320-7>
17. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* **25**(1), 49–59 (1994)