



A Preliminary Experiment on the Estimation of Emotion Using Facial Expression and Biological Signals

Yuya Kurono, Peeraya Sripiam^(✉), Feng Chen, and Midori Sugaya

Shibaura Institute of Technology, 3-7-5,
Toyosu, Koto-ku, Tokyo 135-8548, Japan
peeraya@shibaura-it.ac.jp, doly@shibara-it.ac.jp

Abstract. Imagine the day that a robot would comfort you when you feel sad. In the field of artificial intelligence and robot engineering, there are many research regarding automatic classification of human emotion to enhance human-robot communication, especially for therapy. Generally, estimating emotions of people is based on information such as facial expression, eye-gazing direction, and behaviors that are expressed externally and the robot can observe through a camera and so on. However, there is some invisible information that cannot be expressed, or control not to express. In this case, it is difficult to estimate the emotion even if the analysis technologies are sophisticated. The main idea of this research is to compare the classified emotion based on two different sources: controllable and uncontrollable expression. The preliminary experiments show that our proposed method suggested that the classification of emotion from biological signals outperform the classification from facial expression.

Keywords: Emotion classification · Robotics · Biological signals · Facial expression · Sympathy · Feeling

1 Introduction

Nowadays, communication robots (social robots) are widely used at various sites such as commercial facilities and medical/nursing care facilities. According to the Ministry of Internal Affairs and Communications of Japan, the survey about communication robot usage shows that around 50% are actively willing to use the communication robot [1]. Breazeal et al. [2] proposed a robot with expressive acts to interact socially with humans. In their proposed system, the robot's gaze is changed according to the expression of human to maintain a regular social interaction. Also, many robots are build to have body elements which are similar to human and utilize those elements for their interaction [3, 4]. On the other hand, Hirth et al. [5] developed a robot, "ROMAN" that can express six emotions consisting of anger, disgust, fear, happiness, and sadness. However, there is little discussion on the emotion expression of the robot itself in correspondence to various context.

We focused on the feeling of "empathy", which is considered important in human-to-human communication, to enhance mutual reliability and relationship building [6].

Here, “empathy” is defined as synchronizing with the feelings of the opponent. It could be possible that the robot would mutually understand human emotion and empathize, similar to the communication between human. It was found that when a robot expresses emotion that is close to what human feels, an affinity relationship can be made from the empathy feeling. Misaki et al. [7] build a robot that its facial expression is synchronized with human emotion estimated from speech recognition. They used the SD method to evaluate and found that the positive items were significantly selected, hence, the participants had a positive impression on the robot because of empathy [7]. Haneda and Takeuchi [8] investigated how empathy between people and CG character could affect people. They found that when the participant’s mental state is “angry” and the CG character is also “angry,” the more helping behavior is observed as a result of emotion synchronization. Therefore, a friendly relationship could be established when the other person expresses emotions close to our own, even though he/she is a CG character [8]. However, there is not sufficient study regarding the accurate measurement of human emotions.

There is many studies that proposes the estimation of human emotion based on voice and facial expression. Ikeda et al. [9] classified emotion with facial expression and biological signals. Facial expression was derived through a computer vision method, and face recognition technique. They found that biological signals performed better in the estimation of genuine emotion. The emotion expressed through facial expression or voice tone is carried by the somatic nervous system which is a voluntary nervous system, could be controlled by the sender. On the other hand, biological signals such as heart rate and brain waves could not be controlled by the sender because it is driven by the autonomic nervous system, which is an involuntary nervous system, or the unconscious mind.

However, Ikeda et al.’s work did not describe whether the participant favor “controllable” emotion or “uncontrollable” emotion.

In this work, we propose a system for classifying emotions from controllable expression and uncontrollable expression and compared the classified emotions with the subjectively evaluated emotion in the experiment.

2 Literature Review

Various studies have performed human emotion estimation using biological signals based on Russell’s Circumplex Model of Affection [10]. The model, known as a structure to classify emotion, suggests that emotions are distributed into two dimensional circular space, arousal and valence dimensions. The vertical axis is represented by Arousal, and Valence represents the horizontal axis. The neutral valence and medium level of arousal are indicated with the center location of the circle.

Ikeda et al. [9] proposed a method of estimating emotion using involuntary biological signals. The proposed method uses brain waves and heart rate for estimating the actual emotion of the participant, which is the result of the unconscious mind. From their series of experiments, they found that more accuracy could be derived by including biological signals in the emotion estimation other than using the observable emotion alone. Also, they found that there was a correlation between parts of the

personality and the accuracy of the estimated results. To estimate emotion, they use the pNN50 method for pulse analysis [11], combining with Y-axis of Russel's Circumplex Model. pNN50 measurement is a time domain method for Heart Rate Variability (HRV), which analyzes the physiological phenomenon of the oscillation in the interval between consecutive heartbeats. From the combination of the pNN50 pulse interval and the awakening degree, the "joy" emotion can be defined when the awakening degree is 0 or more, and pNN50 is 0.3 or more.

The International Affective Picture System (IAPS) [12] has been widely used to study emotion and attention in psychological research. IAPS is an international image database developed by the Center for the Study of Emotion and Attention: CSEA, University of Florida. IAPS is a large set of standardized, emotionally-evocative, internationally-accessible, color photographs that includes contents across a wide range of semantic categories. The images set is accompanied by a detailed list of the average rating of the emotions elicited by each picture. Such an average rating procedure for the IAPS is based on the assumption that emotional assessments are identifiable by the three dimensions: valence, arousal, and dominance [13]. Each image is registered with discrete values of 1 to 9 for Arousal dimension (**Sleepiness-Arousal**) and the Valence dimension (**Misery-Pleasure**) dimension which corresponds to both dimensions in the Russell's Circumplex Model of Affection.

The rating procedure for IAPS is done through many subjective evaluation experiments. The participants are asked to rate how pleasant/unpleasant, how calm/excited and how controlled/in-control they felt when looking at each picture. The Self-Assessment Manikin (SAM) [14] is used for this rating procedure. The SAM is a non-verbal pictorial assessment technique that directly measures the feeling of pleasure, arousal, and dominance that are associated with a person's affective reaction to a variety of stimuli. This type of non-verbal assessment, although easy to administer and quick, it is highly correlated with ratings obtained using the verbal or lengthier semantic differential scale.

3 Proposed Method

Figure 1 shows the proposed system. Two types of input data will be retrieved from the participant; the **controllable** expression and the **uncontrollable** expression. The system will estimate emotion based on these input data, and send the estimated emotion to the robot to be displayed. Here, we use facial expression as the controllable expression and biological signals: heart rate and brain wave, are used as the uncontrollable expression. Since biological signals are controlled by autonomic nerves which are involuntary nervous systems, it is considered difficult to change them arbitrarily. In this paper, we focus on the measurement and comparison of biological signals and the facial expression to subjective evaluation of emotion (See Fig. 1 left square).

3.1 Emotion Estimation Using the Biological Signals

We use the method for the estimation of emotion from the biological signals proposed by Ikeda et al. [9]. This method estimates eight types of emotion from the data

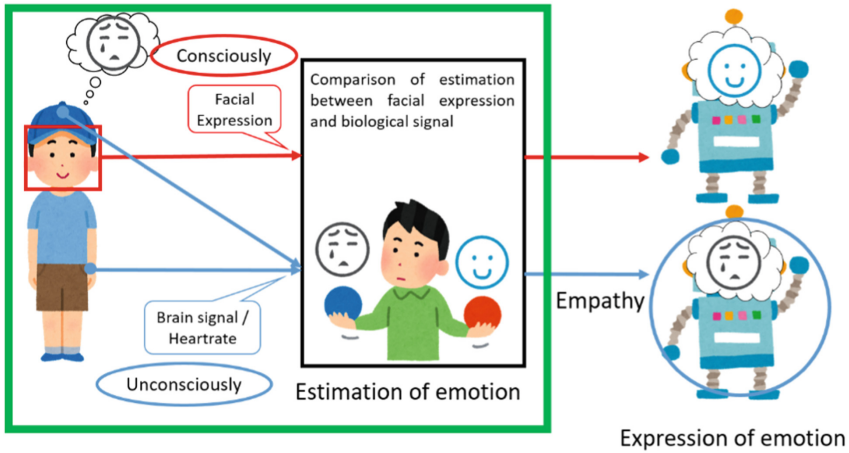


Fig. 1. Overview of the system

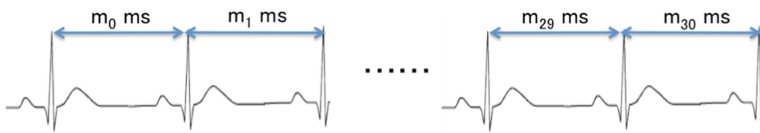


Fig. 2. pNN50, Pulse interval

calculated from brain waves and heart rate. A heart-shaped pulse sensor [15] by World Famous Electronics Ilc. is used to measure a heart rate in this work. The value at the pulse over a one-minute interval, or pNN50, is used to estimate emotion. The pNN50 is defined using time domain measurement of HRV measures.

Figure 2 shows the R-R Interval. RR_i denotes the time from the i^{th} to the $i + 1^{\text{st}}$ R peak. \overline{RR} is the average interval, giving n intervals in total. pNN50 can be obtained using the following equation;

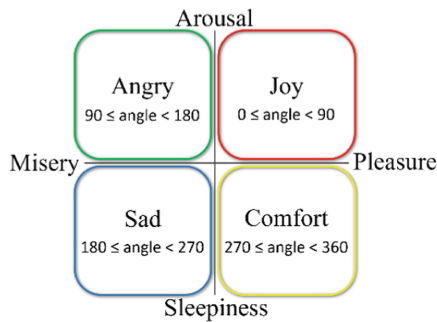


Fig. 3. Emotion estimation model

$$\text{pNN50} = P(|RR_{i+1} - RR_i| > 50 \text{ ms}) \quad (1)$$

Generally, since RR_i is presumed to have a certain fluctuation due to the influence of respiration and blood pressure, pNN50 is set to a pleasant state [16] when it is 0.3 or more. Therefore, we associate pNN50 with the horizontal axis or the **Pleasure dimension** in Russell's Circumplex Model of Affection by setting pNN50 at 0.3 as an origin point.

For brain wave measurement, we use NeuroSky™ MindWave Mobile. The sensor from this electroencephalograph measures EEG/ECG signal and calculates the level of the attention and meditation and return the value from 0 to 100. In this work, the output value is positioned on the vertical axis or the **Arousal dimension** on Russell's Circumplex Model of Affection (Fig. 3). Therefore, the Arousal dimension at the origin or zero means that the value of attention is equal to meditation from the brain wave measurement.

Hence, the combination of pNN50 and the brain wave measurement can be set to (x, y) coordinate on Russell's Circumplex Model of Affection to estimate emotion. In summary, the **joy** emotion can be estimated when the arousal value is a positive value and pNN50 is 0.3 or more.

4 Experiment

Before proceeding to the comparison with robot expression, we conduct a preliminary experiment for validation of the estimated emotion. The objective of this experiment is to find out that the estimated emotion from a different source (controllable expression/uncontrollable expression), which matches more with the participant's emotion (answered by the questionnaire).

During the whole experiment, the participant is asked to wear an electroencephalograph and a pulse sensor. In each trial, the participant is shown an image that will induce one of the four basic emotions; **Joy, Angry, Sad, Comfort**. The participant will fill in the questionnaire regarding his/her emotion toward the shown image. After the experiment, we calculated emotions derived from two sources; biological sensors and the facial expression and analyzed the concordance rate between estimated emotion and subjective evaluation (Fig. 1).

4.1 Subjects

Two male students (20–25 years old) voluntarily participated in the experiment with consent.

4.2 Stimuli

Ten color pictures were taken from the International Affective Picture System [14]. The pictures ranged from somewhat negative (low valence and medium arousal) to neutral (medium valence and low arousal) to somewhat positive (high valence and medium arousal). The pictures represented various content. The order of presentation is random.

4.3 Devices

- Pulse sensor from World Famous Electronics llc. and NeuroSky™ MindWave Mobile are used for biological signals measurement.
- Omron's OKAO™ Vision facial image analysis software is used for the measurement of facial expression.

4.4 Procedure

Figure 4 shows overview of the evaluation procedure. The participant is asked to wear a brain wave sensor and a pulse sensor during the whole experiment. OMRON's OKAO™ Vision is set on a table in front of the participant to detect the participant's facial expression. After begin retrieving all input data from all sensors, the experimental procedure is as follows.

1. The participant stays still (Rest) for 30 s.
2. A standby image (Symbol) is presented on the screen for 10 s.
3. The stimulus image is presented for 15 s.
4. The participant evaluates the stimulus image using SAM [14] on a paper questionnaire.

The above procedure is repeated until all ten stimulus images were evaluated.

4.5 Emotion Classification

We divide the input data into two types; the controllable expression and the uncontrollable expression. The controllable expression is the expression that human has control over, for example, facial expression. Meanwhile, the uncontrollable expression is origin from the unconscious mind, such as heart rate and brain waves.

Emotion Classification From the Uncontrollable Expression

The input data from a brain wave sensor and a pulse sensor is positioned on x,y coordinate of Russell's Circumplex Model of Affection. The classification method is adapted from Ikeda et al. [9], which the emotions are classified into eight types. In order to correlate the estimated emotion with OMRON's OKAO™ Vision output, we classify the emotion into four types. The figure shows the proposed adapted model for emotion classification. As mentioned in Sect. 3.1, the vertical axis is represented by the value from the brain wave sensor, and the horizontal axis is represented by the value calculates from a pulse sensor.

Emotion Classification From the Controllable Expression

In our work, we classify controllable emotion from the participant's facial expression. The device's output is classified into five kinds of emotion: Neutral, Happy, Sorrow, Furious, Surprise at the intervals of 0.2 to 0.3 s. For the comparison in our work, we regroup the emotions classified by the device into four types of emotion using the proposed classification table in Table 1.

Table 1. The emotion classification table

Classified emotion	Joy	Angry	Sad	Comfort
Output emotion from OKAO™ Vision	Happiness Surprise	Furious	Sorrow 1/3 Neutral	2/3 Neutral

Emotion Evaluation Method

Given X as a type of emotion (joy, angry, sad and comfort) occurring at each second i , we can obtain the total presented emotion during the stimulus exposure time EM_X as follows.

$$EM_X \sum_{i=0}^{i=N} X_i \quad (2)$$

Here, the stimulus exposure time is 15 ms, therefore $N = 15$. We take the largest value of EM from the four types of emotion as the presented emotion for further evaluation.

4.6 Subjective Evaluation of Emotion Using SAM

Subjective evaluation for emotion is performed using SAM (Self-Assessment Manikin) [14], a non-verbal pictorial assessment technique used to measure pleasure, arousal, and dominance. After finishing each stimulus exposure, the participant is asked to rate how he/she felt for the stimulus in terms of pleasure and arousal, using nine levels of facial expression of the manikin illustration. The pleasure dimension has the leftmost picture indicating lowest pleasure (unhappy) and the rightmost illustration indicated the highest pleasure (happy). Accordingly, the leftmost illustration in the arousal dimension indicates the lowest arousal (calm) and the rightmost illustration indicates the highest arousal (excitement). With values on each dimension, we can position the pleasure rating onto X-axis and the arousal rating onto Y-axis in Russell's Circumplex Model of Affection to derive subjectively evaluated emotion toward a stimulus.

5 The Analysis of the Result

To analysis the result, we calculate the concordance rate (%), C , as follows;

$$C = \frac{N_{Match}}{10} \cdot 100 \quad (3)$$

Here, N_{Match} is the number of images (out of ten) where the classified emotion is the same as the subjective emotion. The concordance rates for each participant are described in Table 2.

From Table 2, the concordance rate of participant 1 is rank higher for facial expression and vice versa for participant #2. Although the result does not agree, we can

imply that the biological signals gave a rather high (around 50%) concordance rate in both participants, whereas the facial expression resulted in as low as 20% concordance rate for participant 2.

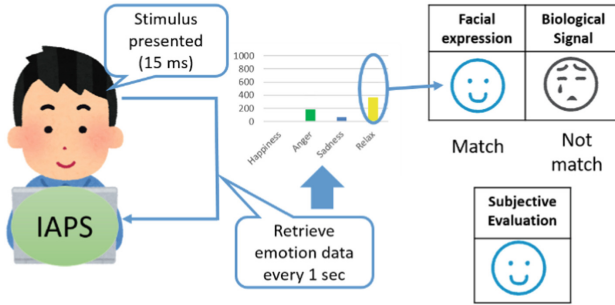


Fig. 4. Overview of the evaluation procedure

Figure 5 shows the transition of each type of emotion from the biological signals and facial expression of participant 2 from the presentation of stimulus 10. The strength of each emotion is a normalized value of 0 to 100. Using facial expression in the classification, we can observe that the classified emotion of “sad” stays higher than other types of emotion throughout 15 ms exposure time.

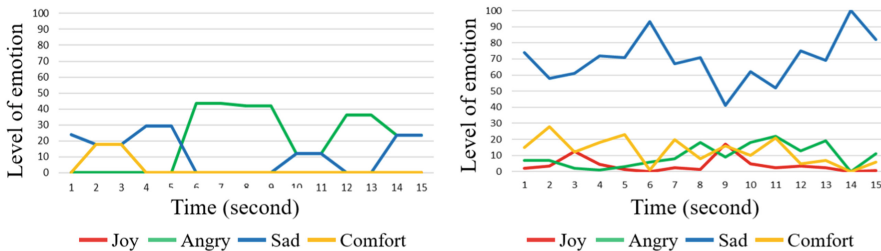


Fig. 5. Left: Classified emotions from the biological signals of trial #10 for participant #2. Right: Classified emotions from facial expression of trial #10 for participant #2.

6 Discussion

The images used as stimuli in this experiment was selected with arousal normative values were in the medium range. We assumed that there is a possibility that images presented did not induce strong emotion, therefore, there is no significant change in emotional value observed with the biological signals. When further analysis is performed, we found that most of the facial expression is classified as “sad.” In our work, we include “neutral” expression when weight for the “sad” classification, therefore, it may affect the result.

In the experiment, there are possibly other uncontrollable factors that could affect the expressed emotion of the participant. For example, boredom from the experiment could result in many “neutral” or “sad” facial expression. Hence, the extremely high classification for “sad” is observed in Fig. 5 Right. We assumed from the experiment result that the biological signals could provide a more genuine emotion classification. Hence, a better impression may be achieved when further applied the classified emotion on the robot expression.

In the future, more participants are necessary for more reliability of the data analysis. Experiment procedure could be improved such as selection of stimuli, presented time, record time and so on. Stimuli that could induce more explicit emotion could be used for more distinct facial expression.

Table 2. The concordance rates of participant 1 and 2, comparison between emotion classified by facial expression and emotion classified by biological signals.

	Participant 1	Participant 2
Facial expression	60%	20%
Biological signals	40%	50%

7 Conclusion

We proposed the emotion classification method from biological signals and facial expression and compared with subjectively evaluated emotion in this research. For classification procedures, biological signals and facial expression are separately calculated to be positioned in the coordinate of Russell’s Circumplex Model of Affection. Similarly, we positioned the participant rating of arousal and pleasant level in the same model of affection for comparison. The concordance value for emotion classified from biological signal and emotion classified from facial expression are compared. Although there is no significant difference observed because of the limited number of participants, it can be implied that emotion classified with biological signals resulted in high concordance for all participants. The finding from this work will lead to the development of the system for the next experiment, to finally compare with the emotion expressed in the robot.

References

1. Ministry of Internal Affairs and Communication. <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h27/html/nc241350.html>. Accessed 15 Feb 2019
2. Breazeal, C., Scassellati, B.: A context-dependent attention system for a social robot. *mn* **255**, 3 (1999)
3. Komatsu, T., Kuki, N.: Investigating the contributing factors to make users react toward an on-screen agent as if they are reacting toward a robotic agent. In: RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication, pp. 651–656. IEEE (2009)

4. Mutlu, B., Osman, S., Forlizzi, J., Hodgins, J., Kiesler, S.: Perceptions of ASIMO: an exploration on co-operation and competition with humans and humanoid robots. In: Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction, pp. 351–352. ACM (2006)
5. Hirth, J., Schmitz, N., Berns, K.: Emotional architecture for the humanoid robot head ROMAN. In: 2007 IEEE International Conference on Robotics and Automation, pp. 2150–2155. IEEE (2007)
6. Satoshi, U., et al.: Cognitive Science in Communication 2: Empathy (Iwanami lecture). Iwanami Shoten (2014). in Japanese
7. Misaki, Y., Ito, T., Hashimoto, M.: Proposal of human-robot interaction method based on emotional entrainment. In: HAI Symposium (2008), in Japanese
8. Haneda, T., Takeuchi, Y.: Study of empathy feeling of CG facial expression (2003)
9. 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)
10. Russell, J.A.: A circumplex model of affect. *J. Pers. Soc. Psychol.* **39**, 1161 (1980)
11. Ewing, D.J., Neilson, J., Travis, P.: New method for assessing cardiac parasympathetic activity using 24 hour electrocardiograms. *Heart* **52**, 396–402 (1984)
12. Lang, P.J., Bradley, M.M., Cuthbert, B.N.: International affective picture system (IAPS): technical manual and affective ratings. In: NIMH Center for the Study of Emotion and Attention, pp. 39–58 (1997)
13. Osgood, C.E.: The nature and measurement of meaning. *Psychol. Bull.* **49**, 197 (1952)
14. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* **25**, 49–59 (1994)
15. World Famous Electronics Ilc. <https://pulsesensor.com/pages/code-and-guide>. Accessed 15 Feb 2019
16. de Carvalho Abreu, E.M., de Souza Alves, R., Borges, A.C.L., Lima, F.P.S., de Paula Júnior, A.R., Lima, M.O.: Autonomic cardiovascular control recovery in quadriplegics after handcycle training. *J. Phys. Therapy Sci.* **28**, 2063–2068 (2016)