Research Article

A System for Posting on an SNS an Author Portrait Selected using Facial Expression Analysis while Writing a Message

Taro Asada[1,*], Yuiko Yano[2], Yasunari Yoshitomi[1], Masayoshi Tabuse[1]

[1]*Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, 1-5 Nakaragi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan*
[2]*Nara Guarantee Corporation, 163-2 Horencho, Nara 630-8668, Japan*

ABSTRACT

We have developed a real-time system for expressing emotion as a portrait selected according to the facial expression while writing a message. The portrait is decided by a hair style, a facial outline, and a cartoon of facial expression. The image signal is analyzed by the system using image processing software (OpenCV). The system selects one portrait expressing one of neutral, subtly smiling and smiling facial expressions using two thresholds on facial expression intensity. We applied the system to post on a Social Network Service (SNS) a message and a portrait expressing the facial expression.

1. INTRODUCTION

Recently, Social Network Services (SNSs) have become very popular as communication tools on the Internet. A message, a static image, or a moving image can be posted on Twitter. However, information about our real emotions cannot be posted on Twitter while writing a message. As a result, a message might be misunderstood by a receiver. It was reported that such a misunderstanding on a message on LINE caused bullying at a school [1].

In this paper, we propose a real-time system for expressing three kinds of emotions as a portrait selected according to the facial expression while writing a message. For generating a portrait, a user selects one hair style and one facial outline before writing a message. The image signal is analyzed by our real-time system using image processing software (OpenCV) [2] and a previously proposed feature parameter (facial expression intensity) [3]. Then the system selects one cartoon expressing one of neutral, subtly smiling and smiling facial expressions using the facial expression intensity thresholds previously decided for these facial expressions. The portrait selection method used in this paper is based on a previously reported method [4].

2. PROPOSED SYSTEM AND METHOD

2.1. System Overview and Method Outline

In this system, a webcam moving image captured in real-time is analyzed via the following process.

The proposed method consists of (1) extracting the mouth area, (2) calculating the facial expression feature vectors, (3) determining the average facial expression intensity while writing a message, (4) selecting hair style and facial outline of portrait by a user, and (5) posting the message and an automatically selected portrait for the message on Twitter. The details of these steps are explained in the following subsections.

2.2. Mouth Area Extraction

The mouth area is selected for facial expression analysis because it is where the differences between neutral and smiling facial expressions appear most distinctly. We use the mouth area extraction method which is described in our reported research [5] and is briefly reviewed in this subsection.

First, moving image data while writing message are changed from RGB to YCbCr image data, and then the face area is extracted from the YCbCr image as a rectangular shape, and the lower 40% portion of the face area is standardized. Next, the mouth area is extracted from that portion.

2.3. Facial Expression Intensity Measurement

We use the facial expression intensity measurement method which is described in our reported research [5] and is reviewed in this subsection.

For the *Y* component of the selected frame, the feature vector for the facial expression is extracted for the mouth area. The extraction

Corresponding author. Email: t_asada@mei.kpu.ac.jp

is performed by using a Two-Dimensional Discrete Cosine Transform (2D-DCT) for each 8×8-pixel block. Two 8×8-pixel blocks at each of the left and right lower corners are not included for this measurement, because these blocks might cause errors due to the appearance of the jaw and/or neck line(s) there [5].

To measure the feature parameters of the facial expressions, we select low-frequency components from the 2D-DCT coefficients for use as the facial expression feature vector elements. However, the direct current component is not included. In total, 15 feature vector elements are obtained. The facial expression intensity is defined as the norm difference between the facial expression feature vectors of the reference and target frames. In this study, the first 20 continuous frames of mouth area data successfully extracted after the webcam recording begins are treated as reference frame candidates. The reference frame selection method is explained in detail in Shimada et al. [6].

2.4. Selecting Hair Style and Facial Outline of Portrait

For generating a portrait of a user, the user selects (1) one gender (male or female), (2) one hairstyle among seven ones for each gender (for female, see Figure 1) and (3) one facial outline among four ones for each gender, and registers them beforehand in the proposed system.

2.5. Posting a Message and an Automatically Selected Portrait for the Message on Twitter

Facial expression intensity is measured using our previously discussed method [7]. A portrait is generated by the proposed system using a cartoon of facial expression, a hairstyle and a facial outline which are selected beforehand by the user. The cartoon is automatically selected by comparing the average value of facial expression intensity and two thresholds decided experimentally beforehand. Then the message and the selected portrait are posted on Twitter when the user presses a button of the proposed system. In this system, three portraits, expressing neutral, subtly smiling and smiling facial expressions, are used (see Figure 2).

Figure 1 | Selectable hair styles for female.

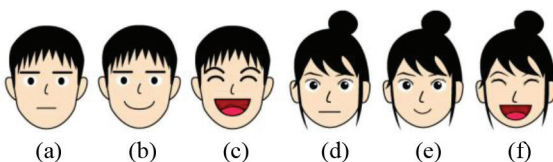   (a)   (b)   (c)   (d)   (e)   (f)

Figure 2 | Each example for male or female portrait expressing neutral (a and d), subtly smiling (b and e), and smiling (c and f) facial expressions.

3. EXPERIMENT

3.1. Conditions

The experiment was performed on a Dell XPS 9350 PC equipped with an Intel Core i7-6560U 2.2 GHz Central Processing Units (CPUs) and 8.0 GB of random access memory. The Microsoft Windows 7 Professional Operating System (OS) was installed on the PC and Microsoft Visual C++ 2008 and 2013 Express Editions were used as the development language. CoreTweet [6], as a library, and the Twitter API were used to post a message and portrait on Twitter.

First, the condition of the preliminary experiment [4] for a preparation of this experiment is reviewed in this subsection. We had performed experiments with one male (subject A in his 20s) and five females (subjects B–F in their 20s) under the two conditions listed below. As an initial condition in the experiment, the subjects had been instructed to maintain a neutral facial expression and face forward without speaking for about 5 s just after the start of the experiment. After the initial state of a neutral facial expression had been terminated, the subjects had been requested to intentionally respond with one of two types of facial expressions (Experiment P1, neutral; Experiment P2, big smile) and write a message, '友達から貰ったこのぬいぐるみかわいくない' (in Japanese), which means, 'This stuffed toy, a present from a friend, is pretty, isn't it?' Experiments P1 and P2 had been performed two times for each subject. In each experiment, facial expression intensity measurements had been performed for each subject during writing a message for 35 s and then the average facial expression intensity during writing that message had been calculated.

In this study, in order to distinguish between the neutral, subtly smiling and smiling facial expressions on the basis of average facial expression intensity while writing message, a threshold between neutral and subtly smiling expressions was set as the intermediate value of 12 average facial expression intensities of the six subjects for neutral (Experiment P1), and then one between subtly smiling and smiling expressions was set as that for big smile (Experiment P2).

In this study, three experiments (Experiments 1–3) were performed with two males (subject G in his 20 s, subject H in his 40 s) and one female (subject I in her 20 s). After the initial state of a neutral facial expression was terminated, the subjects were requested to intentionally respond with three types of emotions (Experiment 1, neutral; Experiment 2, subtly smiling; Experiment 3, smiling) and wrote a same message as that in Experiments P1 and P2. In each experiment, facial expression intensity measurements were performed for 35 s and then the average facial expression intensity for that message for each subject was calculated, after which both the message and the portrait expressing the facial expression while writing the message were posted on Twitter.

3.2. Results and Discussion

Table 1 shows the average of facial expression intensity under the condition of Experiment P1 and P2 for each subject. In Table 1, the values subtracted by the average of facial expression intensity for 10 s in the time range of 1–11 s just after starting the experiment are shown. The intermediate values of facial expression intensities

in Experiments P1 (neutral) and P2 (big smile) were 0.37 and 1.37, respectively (Table 1). Therefore, the two thresholds for distinguishing between the three types of facial expressions were respectively determined as 0.37 and 1.37. Thus, in the proposed system, a facial expression having an average facial expression intensity under 0.37 is judged to be a neutral facial expression, one 0.37 or higher and under 1.37 is judged to be a subtly smiling expression, and one 1.37 or higher is judged to be a smiling expression.

Table 2 shows the average of facial expression intensity under the condition of Experiments 1–3 for each subject. The values in Table 2 were calculated in the same manner as that for Table 1. Figures 3–5 show changes in facial expression intensity of mouth area and mouth image at the characteristic timing point for each subject. As shown in Tables 2 and 3, our proposed system distinguished among the three types of facial expressions (neutral, subtly smiling, smiling) at the accuracy of (8/9). Our system misjudged as a smiling facial expression for a subtly smiling one of subject I (Tables 2 and 3).

Table 1 | The average of facial expression intensity under the condition of Experiment P1 (neutral) and P2 (big smile) [4]

Subject	A		B		C	
1st or 2nd	1	2	1	2	1	2
Neutral	0.38	0.44	0.36	0.17	−0.4	0.51
Big smile	1.48	3.1	0.29	1.01	0.68	3.4

Subject	D		E		F	
1st or 2nd	1	2	1	2	1	2
Neutral	0.4	0.34	0.16	0.38	0.66	−0.06
Big smile	0.45	1.85	0.83	1.25	12.96	11.98

Table 2 | The average of facial expression intensity under the condition of Experiments 1 (neutral), 2 (subtly smiling) and 3 (smiling)

Subject	G	H	I
Exp. 1 (smiling)	0.13	0.07	0.26
Exp. 2 (subtly smiling)	0.87	1.26	1.78
Exp. 3 (smiling)	10.36	1.83	7.9
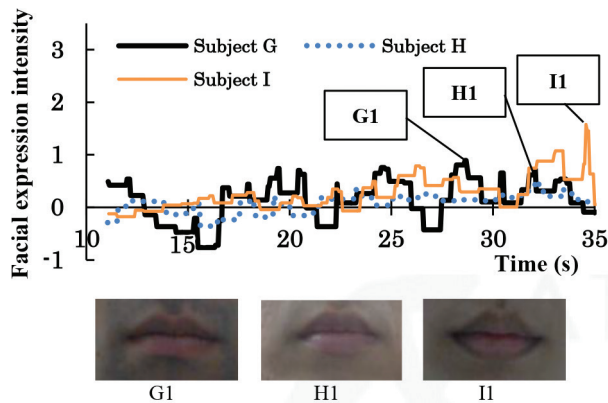
Figure 3 | Changes in facial expression intensity of mouth area for subjects G, H and I (upper graph). Mouth images are shown for three moments during Experiment 1 (neutral) (G1, H1, I1), as indicated on the graph (lower images).
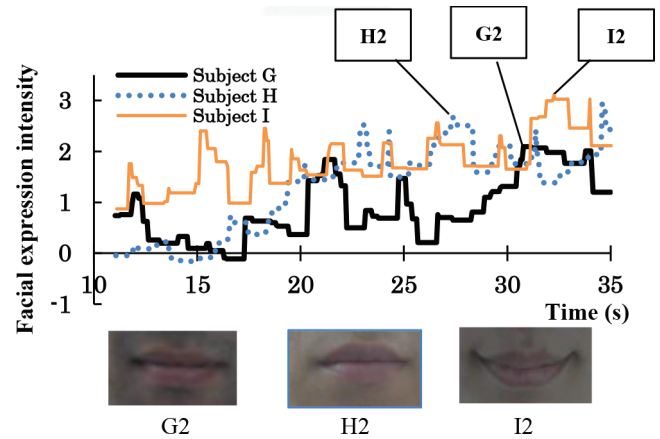
Figure 4 | Changes in facial expression intensity of mouth area for subjects G, H and I (upper graph). Mouth images are shown for three moments during Experiment 2 (subtly smiling) (G2, H2, I2), as indicated on the graph (lower images).
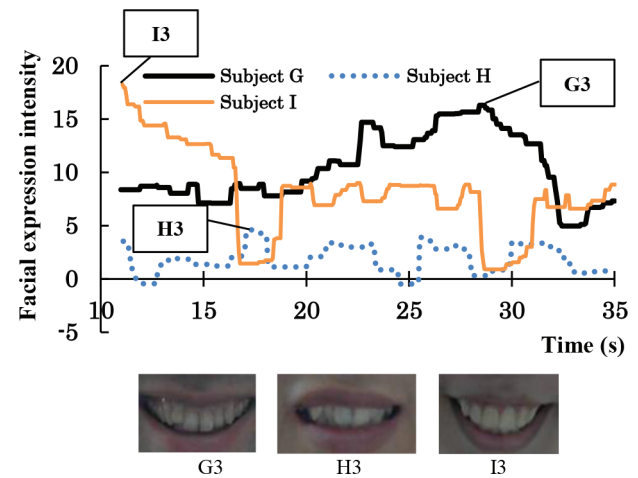
Figure 5 | Changes in facial expression intensity of mouth area for subjects G, H and I (upper graph). Mouth images are shown for three moments during Experiment 3 (smiling) (G3, H3, I3), as indicated on the graph (lower images).

Table 3 | Generated portraits for subjects G and I under the condition of Experiments 1–3

Exp.	1 (Neutral)		2 (Subtly smiling)		3 (Smiling)	
Subject	G	I	G	I	G	I
Portrait						

4. CONCLUSION

Herein, we have developed a real-time system for expressing three emotions (neutral, subtly smiling, smiling) as the portraits generated according to the facial expression while writing a message. We applied the system to the posting on Twitter of both a message and a portrait. The improvement of the system for a subtly smiling judgement is one of the next targets.

CONFLICTS OF INTEREST

ACKNOWLEDGMENTS

REFERENCES

[1] K. Takeuchi, Sumaho Jidai ni Taiou Suru Seitoshidou-Kyouiku Soudan. Tokyo, Honnomori Shuppan, 2014 (in Japanese).

[2] Open CV. Available from: http://opencv.org/ (accessed 13 December 2017).

[3] T. Asada, Y. Yoshitomi, R. Kato, M. Tabuse, J. Narumoto, Quantitative evaluation of facial expressions and movements of persons while using video phone, J. Robot. Netw. Artif. Life 2 (2015), 111–114.

[4] T. Asada, Y. Yano, Y. Yoshitomi, M. Tabuse, Portrait selection system using facial expression analysis while writing messages on SNS, Proceedings of the 2018 Human Interface Symposium, Ibaraki, Japan, 2018, pp. 219–222 (in Japanese).

[5] Y. Yoshitomi, T. Asada, K. Mori, R. Shimada, Y. Yano, M. Tabuse, Facial expression analysis and its visualization while writing messages, J. Robot. Netw. Artif. Life 5 (2018), 37–40.

[6] R. Shimada, T. Asada, Y. Yoshitomi, M. Tabuse, Real-time system for horizontal asymmetry analysis on facial expression and its visualization, Proceedings of the 2017 International Conference on Artificial Life and Robotics, Oita, Japan, 2017, pp. 120–123.

[7] T. Asada, K. Mori, R. Shimada, Y. Yoshitomi, M. Tabuse, Facial expression analysis in real time while writing messages, Proceedings of the 2017 Human Interface Symposium, Osaka, Japan, 2017, pp. 469–472 (in Japanese).

[8] CoreTweet. Available from: https://github.com/CoreTweet/ (accessed 20 December 2017).

Authors Introduction

Dr. Taro Asada

He received his B.S., M.S. and PhD degrees from Kyoto Prefectural University in 2002, 2004 and 2010, respectively. He works as an Associate Professor at the Graduate School of Life and Environmental Sciences of Kyoto Prefectural University. His current research interests are human interface and image processing. HIS, IIEEJ member.

Dr. Yasunari Yoshitomi

He received his B.E., M.E. and PhD degrees from Kyoto University in 1980, 1982 and 1991, respectively. He works as a Professor at the Graduate School of Life and Environmental Sciences of Kyoto Prefectural University. His specialties are applied mathematics and physics, informatics environment, intelligent informatics. IEEE, HIS, ORSJ, IPSJ, IEICE, SSJ, JMTA and IIEEJ member.

Ms. Yuiko Yano

She received her B.S. degree from Kyoto Prefectural University in 2018. She works at Nara Guarantee Corporation.

Dr. Masayoshi Tabuse

He received his M.S. and PhD degrees from Kobe University in 1985 and 1988 respectively. From June 1992 to March 2003, he had worked in Miyazaki University. Since April 2003, he has been in Kyoto Prefectural University. He works as a Professor at the Graduate School of Life and Environmental Sciences of Kyoto Prefectural University. His current research interests are machine learning, computer vision and natural language processing. IPSJ, IEICE and RSJ member.