Analysis of Multiple Users' Experience in Daily Life Using Wearable Device for Facial Expression Recognition

Katsutoshi Masai, Yuta Itoh, Yuta Sugiura, Maki Sugimoto
Keio University, JST CREST
Yokohama, Japan
(masai,itoh,sugiura,sugimoto)@imlab.ics.keio.ac.jp

ABSTRACT
In this paper, we present a wearable facial expression recognition system that can analyse and enhance a daily experience. Our aim is to create a mindful experience in daily life by connecting the device with everyday objects and service. To this end, we made two prototypes that supports users to keep right side of emotions: 1) a text chatting system that automatically inserts an emoticon based on his/her facial expressions in the end of a comment a user typed, 2) a plant interface controlled by facial expressions. We also analysed multiple users' facial expressions while they played video games. We confirmed that visualization of sensor data from the device shows the possibility for estimating the transition of different facial expressions.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords
Affective Computing; Facial Expression; Social Interaction; Wearables;

INTRODUCTION
Facial expression plays an important role in daily life. When we communicate with others, we convey a lot of information through faces. Facial expression is one of the most reliable source of emotions as certain facial expressions are associated with basic emotions such as happiness, anger, surprise and so on [6]. Facial expression is affected by social contexts and it coveys social messages [3]. Facial expression can be made spontaneously or voluntary, and it conveys what people feel and think.

In order to record affective patterns in daily life, we developed the eyewear device that can recognize eight facial expressions using photo reflective sensors in our previous research [7]. The system is wearable so we can record facial expressions in daily life. Moreover, it is comfortable because the sensors are concactless unlike the other wearable methods such as [5], [8].

ACE '16, November 09-12, 2016, Osaka, Japan

DOI: http://dx.doi.org/10.1145/3001773.3014351

Figure 1. Creating mindful experience with wearable facial expression recognition devices.

Our device is not limited for recording and analysing one person's facial expressions. By using multiple devices at the same time, we can record multiple people's data during their interaction. Hence it is possible to consider two important aspects of facial expression, 1)the users' emotional aspects and 2)the social interaction between users with the eyewear devices. These pieces of information can be used for enhancing emotional experience in daily life (Figure 1). There are examples that enhance a daily experience from the facial expression information. Grubler et al. proposed the method of emotionally guided robot learning using a wearable EMG device that can measure facial expressions [4]. They used a trainer's emotional feedback from facial expressions to coach certain behaviors to a robot. Tsujita and Rekimoto developed the digital appliances to facilitate users to make smile faces so that they can be in a positive mood [9]. By connecting the device for facial expression recognition with everyday objects or services, we can open up new possibilities of a mindful daily experience.

In this paper, we describe the explorations to enhance daily interaction using the eyewear device. We propose two concepts with two applications. Also, considering the social function of

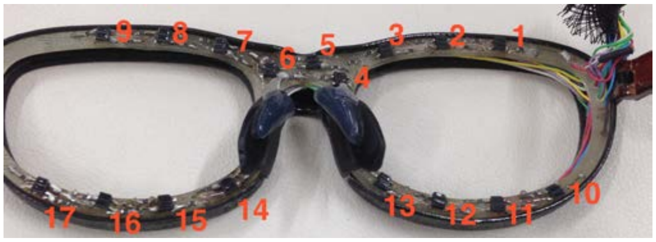

Figure 2. The eyewear device includes 17 sensors distributed on the board.

facial expression, we present the time-series visualization and analysis of facial expressions of two pairs of users when they played an video game against each other using the multiple devices we have developed.

WEARABLE FACIAL EXPRESSION RECOGNITION SYSTEM

In this section, we describe the eyewear prototype that can recognize facial expressions. The eyewear device has 17 photo reflective sensors placed in the front frame (Figure 2). Each sensor measures the distance between the sensor and skin surface on a wearer's face. The distance changes depending on different facial expressions. In the previous study, we applied a machine learning algorithm Support Vector Machine to the obtained sensor values in order to classify a wearer's facial expression into eight categories (neutral, happiness, disgust, anger, surprise, fear, sadness, contempt). More details can be seen in [7].

The advantage of the device is the capability to track a wearer's facial expressions in the long term as it is wearable. It is comfortable because we used photo reflective sensors that is contactless and invasive. We designed the device as a form of eyewear because eyewear is common form and socially acceptable. In addition, it uses only low dimensional data that enables real-time classification. Therefore, it is a good fit for designing a mindful experience in daily life.

MINDFUL EXPERIENCE WITH THE FACIAL EXPRESSION RECOGNITION SYSTEM

Towards creating a mindful experience in daily life with the system, we made two concept prototypes that show possible applications. We believe people enjoying a mindful experience in daily life can keep on the right side of emotions. Two prototypes are aimed at support this idea. One helps people to be more expressive, and another one helps people to control their emotions. For those prototypes, the automatic classification of facial expressions is based on multi-class SVM. We also analysed the users' affective states and social interaction during the game play.

Application with Concept Prototype

Text Chatting System with Auto-emoticon

We developed the prototype of chatting system that enables the user to insert emoticon automatically based on the facial expression when users send messages to the others (Figure 3). This prototype aims at the realm between a video chat and a text chat, making use of social function of facial expression to enrich an emotional side of communication. With the prototype, the user can be expressive with nonverbal clue of facial expression during a chat, which can enhance their communication engagement.

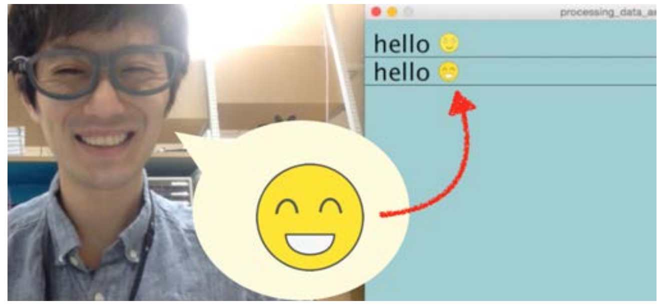

Figure 3. A text chatting system enables automatic emoticon insertion based on the facial expressions after a user typed.

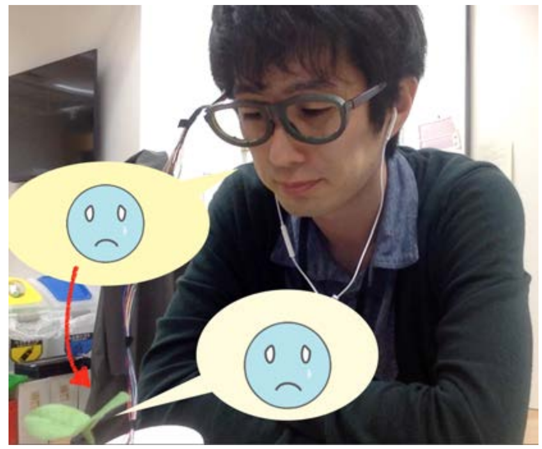

Figure 4. A plant robot responds to user's facial expressions.

Ambient Plant Interface

We hacked a plant robot Pekoppa manufactured by Sega so that it responds to a user's facial expression. For example, when a user shows happy expression, the plant robot flaps its leaves. As another example, when a user shows sad face, the plant robot bends its stem (Figure 4). The plant robot has two bio-metals that move its stem and leaves respectively. The plant is connected to Arduino that sends signals to each bio-metal based on the received data from the wearable system. With the system, a user can be aware of the mood naturally by looking at the plant. It allows the user to manage the emotional state more easily because facial expression is controllable and intentional expression can cause emotional state [2]. In addition, the movement of the plant can give empathy to the user because the plant always mirror the user's emotional state shown in the user's face. The empathy can make the user feel relaxed. Moreover, we can assume another scenario with this prototype. If an elder adult wears the device and the data is sent to the plant robot at remote family's home, the system can be used as a caring system for them, enabling natural conversation between them.

Figure 5. Experimental setting

Figure 6. Facial expression (smile) made (left: during training phase, right: during the game)

Analysis of Facial Expressions during a video game

One of the key elements of an engaging experience is social interaction. It is important to analyze the experience from multiple people's views. In this experiment, we measured the facial muscle movements with the device when a pair of users played a video game against each other. Two pairs (four users) participated in the experiment. The video game was a crossover fighting game. Figure 5 shows the experimental setting. We used two cameras, one for recording the game screen and another for recording a pair of users' facial expressions.

The procedure of the experiment was as follows. 1) A pair of users wore the device respectively. They sat next to each other. 2) The users started playing the game. During the play, we recorded the sensor data and the video of the users' facial expressions and the game screen. 3) After their play of the game, we also recorded and collected the sensor data of their voluntary facial expressions as a reference.

For the analysis, we observed facial expressions of the users by watching the video of their play. From the video, we found that the collected training data is not appropriate to use for automatic classification of facial expressions because there is a critical difference between the facial expressions during the play and the ones during the training data collection phase (Figure 6) for most of the users. Hence we don't show prediction results of machine learning algorithm using data of training phase. Instead, we visualized the sensor values that are normalized so as the difference between sensor data at each frame and the data of neutral face can be in the range of -1 to 1 (Figure 7). If the normalized value is positive (negative), the distance between skin surface and the sensor is closer (further) than when the user shows neutral face. The sensor number corresponds to the number shown in Figure 2. As the sensor 1 didn't work when user C used, the visualization of the sensor value is blank. We also showed at the top of the figure the facial expressions we manually coded into three categories (neutral, positive and negative). These categories are what we were able to recognize from the video. We also added the process of the game (win/lose) to the figure. From the figure, we can see that there is a correspondence between the recognized expressions and the normalized sensor data. For example, The sensor values on the lower part was high when the users smiled. Also, the transition between different facial expressions is clear. Therefore, the normalized sensor data of the device is useful to analyze the time series of facial expression change for the purpose of quantifying facial expression information.

Overall, all users showed their smiles as a dominant expression during the play. It means the experience was enjoyable. Especially, the users except User D showed only two facial expressions (neutral and positive). Yet, they showed smiles in various contexts. For example, User A showed the smile both after she won and lost. User D showed negative expression soon after he made a smile, which may indicate that the smile is not positive. However, if the users had a big smile, those smiles look natural. The big smile has the higher sensor value than other smiles. Interestingly, in case a pair of the users had a smile synchronously, these smiles tended to be bigger than other smiles. The synchronized smiles are not always caused by the result of the game directly. When a pair of users talked based on the contents, they smiled synchronously. It means creating a conversation starter while users play games is an important aspect for their experiences.

LIMITATION AND FUTURE WORK

The coding of facial expressions requires professional skill, and so our manual coding may not be accurate. We didn't apply statical analysis to the data and only analyzed a video game played by two pairs of users using the visualization of sensor data and the recorded video. As a next step, we develop the algorithm to statistically analyze time series data from the sensors and compare the data with automatically coded facial expressions from recorded video using computer vision techniques. In addition, the analysis to the experience-based game where user moves and communicate with others is interesting because this type of contents is more immersive and getting popular. Understanding user's emotion during the play is a key for evaluating and designing the experience as affect is important to understand user's behaviors.

There are noises that need to be considered for a future work. We observed that the sensor values changed when User A looked down at the beginning part. The change of head angle leads to change the sensor value even though the user keeps

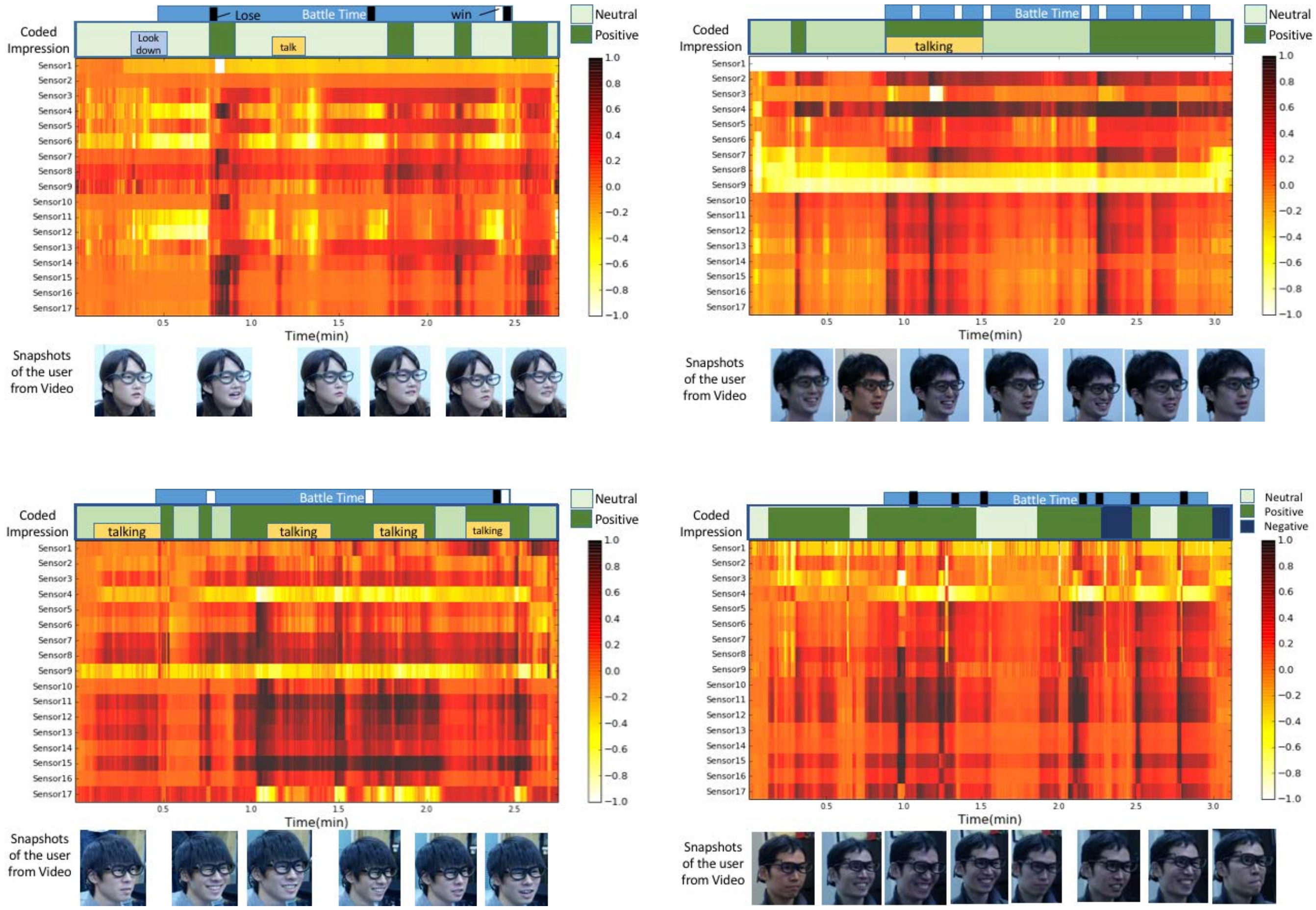

Figure 7. The flow of sensor value during the play (Left:First pair (top:User A,bottom: User B), Right:Second pair (top:User C,bottom: User D)

neutral expression. We also noticed that User B's sensor values on the lower part frequently fluctuated. It is because User B talked often not only to the opponent player but also to himself.

Moreover, we didn't consider the relationships between a pair of users. There is the research that suggests the impression to a partner of the pair and the facial mimicry are correlated each other [1]. We plan to explore an experiment design considering the effect in the future.

We didn't use the sensor data of voluntary facial expressions because it is different from the data of natural ones. Therefore, in order to interrupt the sensor values during the video game, we had to code their expressions from the video. Toward capturing the natural expressions with our device, we need to develop a method to capture sensor data of natural expressions with emotional label. We plan to use automatically coded facial expressions and sensor values as a training data.

CONCLUSION

In this paper, we presented two concept prototypes based on wearable facial expression recognition system in order to create a mindful experience in daily life. The prototypes are a text-chatting system and a plant robot that can be controlled by user's facial expressions. We also presented the analysis of multiple users' facial expressions while they played a video game against each other using the wearable system for facial expression recognition. We confirmed that visualization of sensor data during their play of the video game is useful to understand the transitional change of their facial expressions. We found that the smiles a pair of users made synchronically tend to be bigger than a smile user made individually.

ACKNOWLEDGEMENT

This research is supported by 2016 Keio University Doctorate Student Grant-in-Aid Program and JST CREST.

REFERENCES

1. R. Buck, J. I. Losow, M. M. Murphy, and P. Costanzo. 1992. Social facilitation and inhibition of emotional expression and communication. *J Pers Soc Psychol* 63, 6 (Dec 1992), 962–968.
2. Walter B. Cannon. 1927. The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory. *The American Journal of Psychology* 39, 1/4 (1927), 106–124. `http://www.jstor.org/stable/1415404`
3. Alan J Fridlund. 1994. *Human facial expression: An evolutionary view*. Academic Press, San Diego, CA, US. 369 pages.
4. Anna Gruebler, Vincent Berenz, and Kenji Suzuki. 2012. Emotionally Assisted Human?Robot Interaction Using a Wearable Device for Reading Facial Expressions. *Advanced Robotics* 26, 10 (2012), 1143–1159. DOI: `http://dx.doi.org/10.1080/01691864.2012.686349`
5. Anna Gruebler and Kenji Suzuki. 2014. Design of a Wearable Device for Reading Positive Expressions from

Facial EMG Signals. *Affective Computing, IEEE Transactions on* 5, 3 (July 2014), 227–237. DOI: http://dx.doi.org/10.1109/TAFFC.2014.2313557

6. Dacher et al. Keltner. 2003. Facial expression of emotion. In *Handbook of affective sciences*. Oxford University, New York, NY, USA, 415–432.
7. Katsutoshi Masai, Yuta Sugiura, Masa Ogata, Kai Kunze, Masahiko Inami, and Maki Sugimoto. 2016. Facial Expression Recognition in Daily Life by Embedded Photo Reflective Sensors on Smart Eyewear. In *Proceedings of the 21st International Conference on Intelligent User Interfaces (IUI '16)*. ACM, New York, NY, USA, 317–326. DOI: http://dx.doi.org/10.1145/2856767.2856770
8. Jocelyn Scheirer, Raul Fernandez, and Rosalind W. Picard. 1999. Expression Glasses: A Wearable Device for Facial Expression Recognition. In *CHI '99 Extended Abstracts on Human Factors in Computing Systems (CHI EA '99)*. ACM, New York, NY, USA, 262–263. DOI: http://dx.doi.org/10.1145/632716.632878
9. Hitomi Tsujita and Jun Rekimoto. 2011. Smiling Makes Us Happier: Enhancing Positive Mood and Communication with Smile-encouraging Digital Appliances. In *Proceedings of the 13th International Conference on Ubiquitous Computing (UbiComp '11)*. ACM, New York, NY, USA, 1–10. DOI: http://dx.doi.org/10.1145/2030112.2030114