Facial Expression Mapping inside Head Mounted Display by Embedded Optical Sensors

Katsuhiro Suzuki* **Fumihiko Nakamura*** **Jiu Otsuka***
Katsutoshi Masai* **Yuta Itoh*** **Yuta Sugiura***
Maki Sugimoto*
*Interactive Media Lab, Keio University
3-14-1 Hiyoshi, Kohoku, Yokohama, 223-0061 Japan
{katsuhirosuzuki | f.nakamura | jiu | masai | Itoh | sugiura | sugimoto}@imlab.ics.keio.ac.jp

ABSTRACT
Head Mounted Display (HMD) provides an immersive experience in virtual environments for various purposes such as for games and communication. However, it is difficult to capture facial expression in a HMD-based virtual environment because the upper half of user's face is covered up by the HMD. In this paper, we propose a facial expression mapping technology between user and a virtual avatar using embedded optical sensors and machine learning. The distance between each sensor and surface of the face is measured by the optical sensors that are attached inside the HMD. Our system learns the sensor values of each facial expression by neural network and creates a classifier to estimate the current facial expression.

Author Keywords
Virtual Reality, Facial Expression Recognition, Neural Network, Photo reflectivity, Wearable Sensing, Emotion.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Due to the appearance of HMD for consumers these days (e.g. Oculus Rift and PlayStationVR), it can be predicted that virtual reality (VR) is going to be widely used by individuals for various purposes such as gaming and communicating with people in remote locations in the near future. The advantage of VR is that it is possible for the user to have an immersive experience in the virtual world, as if the user is actually present in the virtual environment. Also, people who do not prefer to expose their face in the virtual world in terms of privacy can be able to communicate smoothly through the use of a virtual avatar. However, non-verbal communication cannot be expressed when user communicates in a virtual environment through an avatar. The listener may feel unnatural if the facial expression of the avatar is unchanged while the avatar talks with different kinds of emotions and intonations. In order for the avatar to imitate the facial expression of the user, it is necessary to estimate the current facial expression of the user, but camera-based methods cannot be used to recognize the facial expression in the situation because the HMD covers up the upper half of the face.

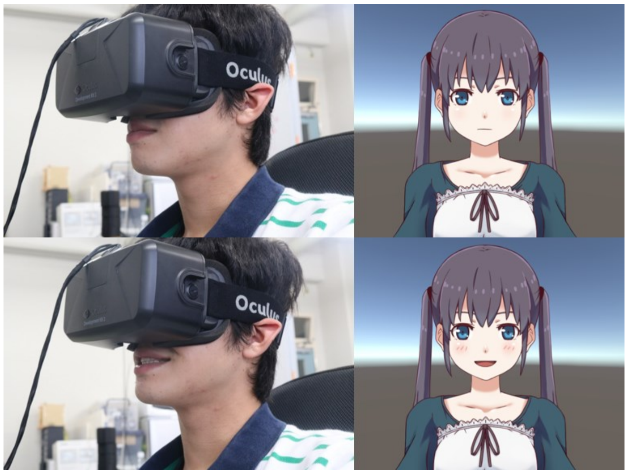

Figure 1. Facial expression reflects to the avatar (Top: Neutral, Bottom: Smile)

We propose a facial expression recognition technique using optical sensors that can be used even when wearing a HMD. The optical sensors can detect the distance between the sensors and skin surface. We estimate the facial expression of the user with the neural network by learning the sensor values that change depending on different facial expressions.

RELATED WORKS
Jocelyn et al. proposed a system to identify specific facial expression by sensing facial movement [3]. It is one of the first works that has applied facial expression recognition technique to a wearable device. Li et al.developed the system that can capture facial performance using strain gauges mounted on contact area to face and RGB-D camera attached to the HMD[1]. In their system, signals from strain gauges and RGB-D camera are corresponded to the 3D face

UIST'16 Adjunct, October 16-19, 2016, Tokyo, Japan
ACM 978-1-4503-4531-6/16/10.
http://dx.doi.org/10.1145/2984751.2985714

model. However, strain gauge requires stable contact to face due to the variation of distribution with the HMD position and head orientation. In order to solve this problem, we propose a robust facial expression recognition system with non-contact optical sensors. Masai et al. estimated the human facial expressions with an eyeglass type device that has a number of photo sensors [2]. This approach is also effective for HMDs because large materials cannot be attached inside.

FACIAL EXPRESSION MAPPING BY EMBEDDED OPTICAL SENSORS ON HEAD MOUNTED DISPLAY

An optical sensor can measure the distance from an object in front of the sensor by radiating infrared light and receiving the light that is reflected from the object. The fact that size of the sensor is small enough and it has a small power consumption allows multiple sensors to be attached in the interior of the HMD along with the flexible circuit board. This circuit board is used to measure the distance between the sensor and the surface of the skin, which slightly fluctuates as facial expression changes. The data obtained by the sensor can be used for estimating the current facial expression through the use of neural network.

Preprocessing

The range of sensor values varies by individual because facial geometry is different for each person. In order to counterbalance the difference, we collect calibration data by recording sensor values for neutral expression as a mean value, and values when creating various facial expressions to obtain maximum and minimum sensor values. We set a baseline value when the user makes a neutral face, and initialize the value each time the user wears the HMD. In this way, additional learning was not required as we normalize the sensor value based on the baseline value. The value of photo reflective sensor has a non-linear relationship to the distance between the sensor and surface. To achieve a linear relationship between these variables, the sensor value is compensated.

Training Data Collection by Mimicking Virtual Character

It is necessary to provide supervised data of each facial expression as an input for the learning phase. However, since it is difficult to operate a device in a state of wearing the HMD, we implemented an application which automatically collects the sensor data necessary for the classification of each emotion in the course of machine learning. The application shows an avatar which expresses several samples of facial expressions in order through the HMD. The user is asked to mimic each emotion which the avatar represents while wearing the HMD. The system waits for a moment to allow the user to change the facial expression.

Multi Class Classification of Facial Expression

In the learning phase, we input sensor values of each emotion and teaching signal of current facial expression of the user. Back propagation method is utilized for the three-layer perceptron network. The network outputs the degree of similarity between a given facial expression and the facial expressions that were provided for the teaching signals. In our case, facial expression which has the highest degree of similarity is given as a result for identifying the given expression.

Hardware

We designed and implemented a flexible printed circuit board that can be mounted on the interior of Oculus Rift DK2. We have installed 16 photo reflective sensors. 14 optical sensors are disposed around the eyes. In addition, 2 more optical sensors are attached to the bottom part of the HMD device to measure the movement of the cheek.

Figure 2. Inner of the HMD

CONCLUSION

We proposed a facial expression mapping technique between an avatar and user using embedded optical sensors and machine learning. By making use of the characteristics of optical sensors effectively, we can build the system without adding anything onto the exterior of the HMD

ACNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 16H05870.

REFERENCES

1. Li, H., Trutoiu, L., Olszewski, K., Wei, L., Trutna, T., Hsieh, P.-L., Nicholls, A., and Ma, C. Facial performance sensing head-mounted display. ACM Transactions on Graphics (Proceedings SIGGRAPH 2015) 34, 4 (July 2015).
2. Masai, K., Sugiura, Y., Ogata, M., Kunze, K., Inami, M., and Sugimoto, M. 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, 2016), 317-326.
3. Scheirer, J., Fernandez, R., and Picard, R. W. Expression glasses: A wearable device for facial expression recognition, 1999