e2-MaskZ: a Mask-type Display with Facial Expression Identification using Embedded Photo Reflective Sensors

Akino Umezawa*
Yoshinari Takegawa*
Future University Hakodate
{g2217001,yoshi}@fun.ac.jp

Katsuhiro Suzuki
Katsutoshi Masai
Yuta Sugiura
Maki Sugimoto
Keio University
{katsuhirosuzuki,masai}@keio.jp
{sugiura,maki.sugimoto}@keio.jp

Yutaka Tokuda
Freelance
yutakamitsue@gmail.com

Diego Martinez Plasencia
Sriram Subramanian
University of Sussex
{D.Martinez-Plasencia,sriram}@sussex.ac.uk

Masafumi Takahashi
Hiroaki Taka
Keiji Hirata
Future University Hakodate
g2118022@fun.ac.jp
{taka,hirata}@fun.ac.jp

Figure 1: (Left) Face display mask and Face capture mask, (Center) The user's facial expression is reflected onto the avatar, (Right) Usage Image of e2-MaskZ

ABSTRACT

The goal of this research is to propose the e2-MaskZ, a mask-type display that changes the user's face to the face of an avatar. The e2-MaskZ is composed of a face-capture mask to recognize the facial expression, and a face-display mask to present the avatar that reflects the recognize expression of the system wearer. 40 photo reflective sensors are laid out across the entire surface of the face-capture mask, and the e2-Mask is made to learn the sensor data for each new facial expression.

CCS CONCEPTS

• **Human-centered computing** → **Systems and tools for interaction design**.

KEYWORDS

Augmented Face, Communication, Human Agent Interaction

*Both authors contributed equally to this research.

AHs '20, March 16–17, 2020, Kaiserslautern, Germany

ACM ISBN 978-1-4503-7603-7/20/03.
https://doi.org/10.1145/3384657.3385332

ACM Reference Format:
Akino Umezawa, Yoshinari Takegawa, Katsuhiro Suzuki, Katsutoshi Masai, Yuta Sugiura, Maki Sugimoto, Yutaka Tokuda, Diego Martinez Plasencia, Sriram Subramanian, Masafumi Takahashi, Hiroaki Taka, and Keiji Hirata. 2020. e2-MaskZ: a Mask-type Display with Facial Expression Identification using Embedded Photo Reflective Sensors. In *AHs '20: Augmented Humans International Conference (AHs '20), March 16–17, 2020, Kaiserslautern, Germany.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3384657.3385332

1 INTRODUCTION

The face is an important body part which decides the impression a person makes on others. Secord [9] asserts that, in face-to-face communication, each person makes assumptions about the other person's personality based on facial features. Kawanishi [5] showed that there is a positive correlation between pleasing facial features and an affinity with a person.

In contrast, researcher have explored a variety of wearable methods Head Mounted Displays (HMDs) to enable a virtual face manipulation in the context of face-to-face communication [1, 4, 7]. The method, however, requires all the interlocuters to wear HMDs to see essential for natural faec-to-face communication, it has been difficult to achieve with conventional HMD methods. If the actual face of the display wearer himself, which is seen by a large unspecified number of third parties, can be altered, it is thought that the impression and aura of the wearer that are conveyed to others can be manipulated more flexibly.

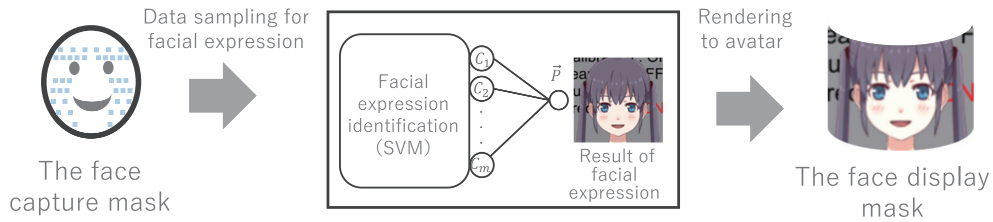

Figure 2: Overview of the e2-MaskZ workflow from face-capture to face-display

Therefore, we propose the e2-MaskZ (Figure 1 (left and center)), a mask-type display that replaces the user's face with that of an avatar. It should be stated that in this paper we define face-to-face communication as communication between two parties in the real world, such as in a medical examination (Figure 1 (right)), meeting or interview.

2 DESIGN AND IMPLEMENTATION

2.1 Work Flow

Figure 2 shows work flow of the e2-MaskZ. The e2-MaskZ is composed of a face-display mask and face-capture mask. The face-capture mask samples the facial expression data of the wearer and the face-display mask outputs the encoded face information as the avatar on the curved screen. To reproduce the convex profile of a human face, the face display mask utilizes an off-the-shelf flexible OLED display, Flexible Top Hat [8]. The photo reflective sensors laid out on the underside of the face capture mask. The system then identifies the facial expression based on the photo reflective sensor data and produces an avatar presenting the same facial expression[1]. The image of a generated avatar is displayed on the face display mask.

2.2 Facial Expression Identification Method

Hardware: We used photo reflective sensors, Genixtek Corporation's TPR-105 [2] to recognize the facial expression of a mask wearer [6]. As this miniature sensor has a small dimension of 2.7 x 3.2 x 1.4 mm (W x L x H), 40 units of it are densely arranged at sampling points on the face capture mask, as shown in Figure 1 (left). 40 photo reflective sensors were placed on the eyebrows, cheeks and around the eyes and mouth. We ensure that the sensors do not contact the eyes, nose and mouth of the mask wearer. Data from each sensor is collected using the Arduino Pro mini, then sent from the Arduino Pro mini to the PC by serial communication.

Software: Every time facial expression changes, the contours of the skin of the face change and simultaneously the distance between the photo reflective sensors and the skin of the face changes with each facial expression. Using this characteristic, we take the data from each photo reflective sensor as feature values and use Support Vector Machine (SVM) to construct a facial expression identification model. The input from each photo reflective sensor was first encoded as 10-bit data through the Arduino and then normalized between 0 and 1 on the PC. After this, the SVM (C = 10, gamma = 1.0) algorithm of rbf kernel was applied and the facial expression identification model was constructed. The 40 pieces of photo reflective sensor data that have been normalized are taken as feature values then the facial expression class is estimated by applying these to the facial expression identification model. To create the avatar's intermediate facial expression, described in section 2.3, the probability distribution of the feature values to belong to each class of facial expression was constructed in a database.

[1]Duet Display [3] is used to make the PC recognize an iPhone as an external display and present the created avatar on the iPhone. The organic EL display mounted onto Flexible Top Hat can perform mirroring of the image on the iPhone, by using the iOS application RoStyle, developed by Royole Corporation. This is used to make the image on the iPhone be displayed on the face display mask.

2.3 Output of Avatar

The user's facial expression is reflected onto the avatar based on the facial expression identification result explained in section 2.2. This method refers to the work of Suzuki et al. [10, 11]. In the case in which the facial expression to be reflected onto the avatar can be expressed by arbitrary n dimension facial expression parameters, the avatar's facial expression parameters $\vec{P}$ that are stored in the computer memory, such as the degree to which the left and right eyes are open, the width of the mouth, etc., can be composed by the formula below. In the formula below, $\vec{P_B}$ shows the straight-face facial expression parameters, $\vec{P_i}$ (i = 1, 2, ..., m) shows the m number of facial expression parameters besides the straight-face facial expression parameters, and $\vec{C} = (c_1, c_2, ..., c_m)$ shows the probability value of multiple-class classification of facial expression.

$$\vec{P} = \vec{P_B} + \sum_{k=1}^{n} c_i \times (\vec{P_i} - \vec{P_B})$$

Changing the avatar's expression by the above method makes it possible to express not only basic facial expressions, but also subtle ones, and also to output intermediate facial expressions that occur as facial expression is changing. We created an avatar 3D model by Live2D and rendered it with Unity3D Game Engine. Communication between the programs on the PC uses UDP and comprises transmission of facial expression identification results, keyboard input, etc.

3 FUTURE WORK

As future work, there is evaluation of facial expression identification accuracy, investigation into sense of agency, evaluation of the performance of avatars created on the e2-MaskZ (whether the wearer can think of the substituted avatar face as his/her own face) when wearing the e2-MaskZ, and a cognitive psychological experiment using the e2-MaskZ.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number JP19H04157.

REFERENCES

[1] Yuma Akaike, Satoshi Kanamaru, Jun Komeda, Kume Yuka, and Yutaka Arakawa. 2014. Effect of Augmented Reality Technology on Communication Ability. *IPSJ SIG Technical Report* 2014-CDS-10, 5 (May 2014), 1–8.

[2] LTD. AKIZUKI DENSHI TSUSHO CO. [n. d.]. Photo Reflector (Reflective) Type) TPR-105: Semiconductor AKIZUKI DENSHI TSUSHO-Electronic Components. Online Shopping. Retrieved March 24, 2018 from http://akizukidenshi.com/catalog/g/gI-03812/

[3] duet. [n. d.]. Duet Display. Retrieved Octorber 1, 2019 from https://ja.duetdisplay.com

[4] Saki Hagiwara and Kazutaka Kurihara. 2016. Development and Evaluation of a "Gaze Phobic KOMYUSHO" Support System using See-through HMD based on Social Welfare Approach. *Computer Software* 33, 1 (Jan. 2016), 52–62. https://doi.org/10.11309/jssst.33.1_52

[5] Chihiro Kawanishi. 1998. The effect of accuturacy motivation on face function in person perception. *The Japanese Journal of Psychology* 68, 6 (Feb 1998), 465–470. https://doi.org/10.4992/jjpsy.68.465

[6] 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)*. Association for Computing Machinery, New York, NY, USA, 317–326. https://doi.org/10.1145/2856767.2856770

[7] Hirotaka Osawa. 2014. AgencyGlass: Substitution of Emotional Labor with Anthropomorphization of Users. *IPSJ Interaction 2014* C6-2 (2014), 708–709.

[8] ROYOLE. [n. d.]. Flexible Top Hat. Retrieved October 30, 2019 from http://www.royole.com/jp/flexible-top-hat

[9] Paul Secord. 1958. Facial features and inference processes in interpersonal perception. *Person Perception and Interpersonal Behavior* (1958), 300–315.

[10] Katsuhiro Suzuki, Fumihiko Nakamura, Jiu Otsuka, Katsutoshi Masai, Yuta Itoh, Yuta Sugiura, and Maki Sugimoto. 2017. AffectiveHMD: Facial Expression Recognition and Mappingto Virtual Avatar Using Embedded Photo Sensors. *Transactions of the Virtual Reality Society of Japan* 22, 3 (Sep 2017), 379–389. https://doi.org/10.18974/tvrsj.22.3_379

[11] Katsuhiro Suzuki, Fumihiko Nakamura, Jiu Otsuka, Katsutoshi Masai, Yuta Itoh, Yuta Sugiura, and Maki Sugimoto. 2017. Recognition and mapping of facial expressions to avatar by embedded photo reflective sensors in head mounted display. In *2017 IEEE Virtual Reality (VR)*. IEEE, 177–185.