3D Facial Geometry Analysis and Estimation Using Embedded Optical Sensors on Smart Eyewear

Nao Asano
Keio University
naoasano@imlab.ics.keio.ac.jp

Katsutoshi Masai
Keio University
masai@imlab.ics.keio.ac.jp

Yuta Sugiura
Keio University
sugiura@keio.jp

Maki Sugimoto
Keio University
sugimoto@ics.keio.ac.jp

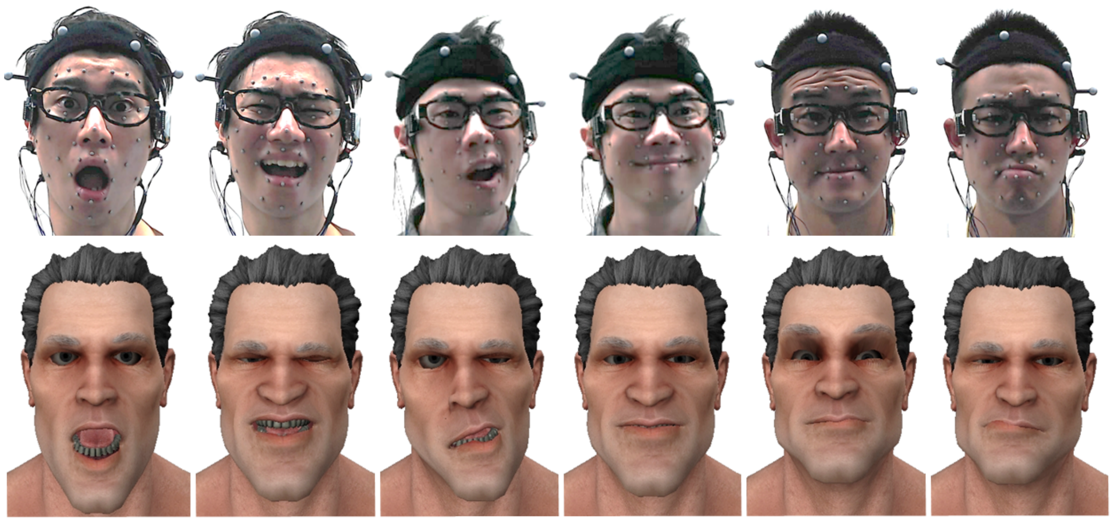

Figure 1: 3D Facial Geometry Examples estimated by the proposed system.

ABSTRACT

Facial performance capture is used for animation production that projects a performer's facial expression to a computer graphics model. Retro-reflective markers and cameras are widely used for the performance capture. To capture expressions, we need to place markers on the performer's face and calibrate the intrinsic and extrinsic parameters of cameras in advance. However, the measurable space is limited to the calibrated area. In this study, we propose a system to capture facial performance using a smart eyewear with photo-reflective sensors and machine learning technique. Also, we show a result of principal components analysis of facial geometry to determine a good estimation parameter set.

CCS CONCEPTS

• **Computer systems organization → Embedded systems**;

SIGGRAPH '18 Posters, August 12-16, 2018, Vancouver, BC, Canada

ACM ISBN 978-1-4503-5817-0/18/08.
https://doi.org/10.1145/3230744.3230812

KEYWORDS

Facial Performance Capture, Wearable Device

ACM Reference Format:
Nao Asano, Katsutoshi Masai, Yuta Sugiura, and Maki Sugimoto. 2018. 3D Facial Geometry Analysis and Estimation Using Embedded Optical Sensors on Smart Eyewear. In *Proceedings of SIGGRAPH '18 Posters.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3230744.3230812

1 INTRODUCTION

Common facial performance capture techniques use retro-reflective markers and multiple cameras to capture the expressions robustly [Williams 1990]. However, such systems need many markers on facial skin for every recording. In addition, they require the pre-calibration of internal and external camera parameters and we cannot move the cameras during capturing. Other methods use monocular cameras [Wu et al. 2016] and commodity RGB-D cameras [Bouaziz et al. 2013]. These methods capture faces with a small hardware setting, but the measurable space is still limited due to the viewing angle of the camera and occlusions.

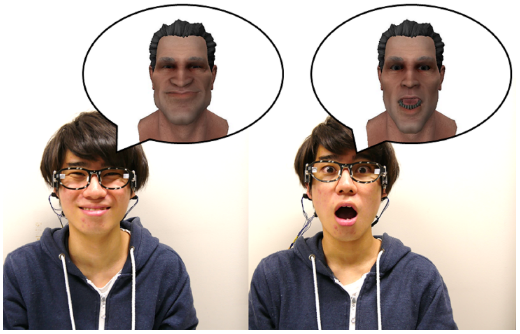

Figure 2: The concept of our facial performance capture system. The mocap marker positions are captured and learned during the training phase. After the training, our system is able to estimate facial geometry without mocap makers.

2 RELATED WORKS

There are several methods for recognizing facial expressions using wearable devices instead of cameras such as [Gruebler and Suzuki 2014]. Masai et al. developed a system to classify basic facial expressions using a smart eyewear with photo-reflective sensors and machine learning [Masai et al. 2016]. They measure the proximity between the skin surface on a face and the eyewear with the sensors. Their method allows the classification of facial expressions in many situations. However, the previous works with wearable devices did not try to estimate geometry change in various facial expressions.

3 SYSTEM DESIGN

We present a facial performance capture system using a smart eyewear with embedded photo-reflective sensors. In the training phase, we obtain the marker positions correlating to facial geometry from the marker-based motion capture system and reflection intensity information from the sensors of the eyewear. We convert sensor values into the distance and reduce the dimension of the marker positions by applying principal component analysis (PCA) . Then, we generate regression models representing the relationship between the two. In the estimation phase, using the regression models, we estimate the 3D position of the markers representing the facial expression shape of the user from the sensor values. We can retarget the facial expression to the 3D model (NaturalPoint CohenExpRig) based on the estimated marker position.

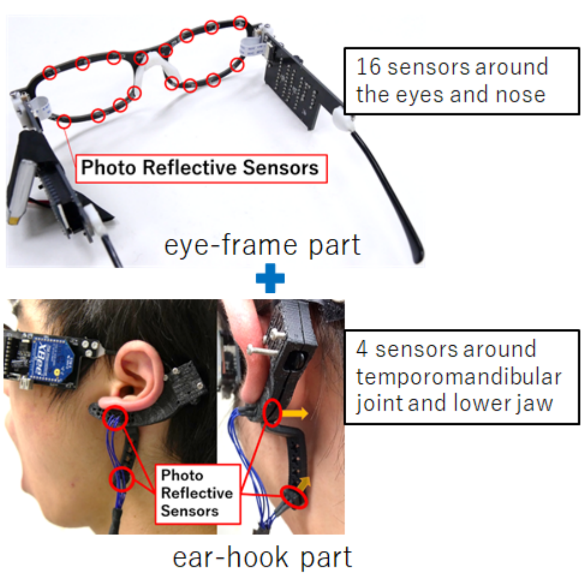

Figure 3: Hardware design of the proposed system. The estimation accuracy around the mouth can be improved by additional sensors measure displacement around the jaw.

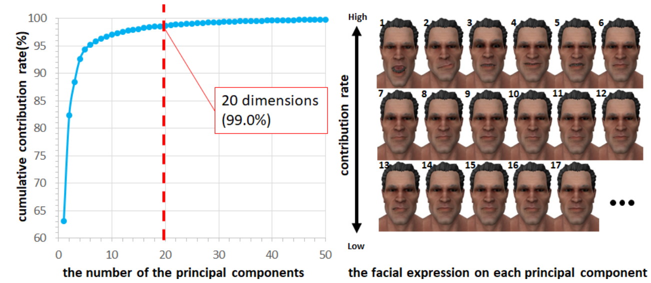

Figure 4: A result of PCA of mocap markers (37 markers, in total 111 dimensions) on various facial expressions. The result shows 20 principal components keeps 99% of the marker information. By dimension reduction of the facial geometry parameters, while keeping the information as much as possible, we can estimate it by simple embedded photoreflective sensors (20 sensors in total) efficiently.

4 CONCLUSIONS

In this study, we proposed a facial performance capture system using photo-reflective sensors placed on a smart eyewear. Our performance capture system estimates 3D points on the skin surface using regression models. The system allows us to estimate 3D facial geometry with a simple wearable device.

ACKNOWLEDGMENTS

This research was partially supported by JST CREST (JPMJCR14E1) and JSPS KAKENHI (16H05870).

REFERENCES

Sofien Bouaziz, Yangang Wang, and Mark Pauly. 2013. Online Modeling for Realtime Facial Animation. *ACM Trans. Graph.* 32, 4, Article 40 (July 2013), 10 pages. https://doi.org/10.1145/2461912.2461976

A. Gruebler and K. Suzuki. 2014. Design of a Wearable Device for Reading Positive Expressions from Facial EMG Signals. *IEEE Transactions on Affective Computing* 5, 3 (July 2014), 227–237. https://doi.org/10.1109/TAFFC.2014.2313557

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. https://doi.org/10.1145/2856767.2856770

Lance Williams. 1990. Performance-driven Facial Animation. In *Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '90)*. ACM, New York, NY, USA, 235–242. https://doi.org/10.1145/97879.97906

Chenglei Wu, Derek Bradley, Markus Gross, and Thabo Beeler. 2016. An Anatomically-constrained Local Deformation Model for Monocular Face Capture. *ACM Trans. Graph.* 35, 4, Article 115 (July 2016), 12 pages. https://doi.org/10.1145/2897824.2925882