Make-a-Face: A Hands-free, Non-Intrusive Device for Tongue/Mouth/Cheek Input Using EMG

Takuro Nakao
Keio University Graduate School of Media Design,
a8111140@keio.jp

Yun Suen Pai
Keio University Graduate School of Media Design,
yspai1412@gmail.com

Megumi Isogai
NTT Media Intelligence Laboratories
Japan
megumi.isogai.ks@hco.ntt.co.jp

Hideaki Kimata
NTT Media Intelligence Laboratories
Japan
hideaki.kimata.yu@hco.ntt.co.jp

Kai Kunze
Keio University Graduate School of Media Design,
kai.kunze@gmail.com

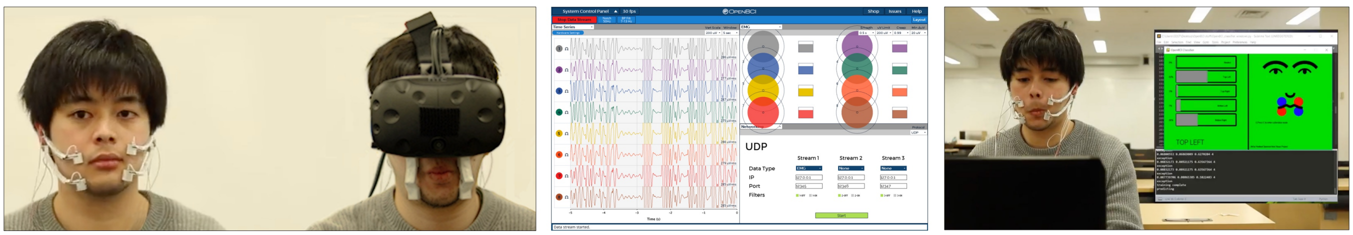

Figure 1: (a) Make-a-Face as a wearable and integrated with VR, (b) sensed raw electromyography data, and (c) Make-a-Face interface and gesture classification

ABSTRACT

Current devices aim to be more hands-free by providing users with the means to interact with them using other forms of input, such as voice which can be intrusive. We propose Make-a-Face; a wearable device that allows the user to use tongue, mouth, or cheek gestures via a mask-shaped device that senses muscle movement on the lower half of the face. The significance of this approach is threefold: 1) It allows a more non-intrusive approach to interaction, 2) we designed both the hardware and software from the ground-up to accommodate the sensor electrodes and 3) we proposed several use-case scenarios ranging from smartphones to interactions with virtual reality (VR) content.

CCS CONCEPTS

• **Computing methodologies** → **Virtual reality**; • **Hardware** → *Sensor devices and platforms*;

KEYWORDS

Hands-free, non-intrusive, electromyography, face input

ACM Reference format:

Takuro Nakao, Yun Suen Pai, Megumi Isogai, Hideaki Kimata, and Kai Kunze. 2018. Make-a-Face: A Hands-free, Non-Intrusive Device for Tongue/Mouth/Cheek Input Using EMG. In *Proceedings of SIGGRAPH '18 Posters, Vancouver, BC, Canada, August 12-16, 2018,* 2 pages.
https://doi.org/10.1145/3230744.3230784

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

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

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5817-0/18/08.
https://doi.org/10.1145/3230744.3230784

1 INTRODUCTION

We use more and more computing in everyday situations from IoT in our environment to wearable glasses like HoloLens or Intel's Vaunt. Interacting with them is usually not hands-free or if most interaction modes (speech etc.) can be considered disruptive or socially unacceptable especially when socializing. The hands and especially fingers are perfect tools for interactions as they contain some of the densest areas of nerve endings in the human body. The only other part that has similar areas is the human face, particularly the tongue [Coren 2003]. There is a lot of related work regarding tongue-based interactions. However, most solutions use a camera to capture input, require an item in the mouth, or are radar-based more suitable for fast gestures [Goel et al. 2015; Li et al. 2015; Niu et al. 2014; Sasaki et al. 2011]. There are also a couple of electromyography based solutions, yet they mostly have obtrusive setups with more electrodes (as they are aimed at the disabled or paralyzed) or they just aim at detecting tongue movements [Sasaki et al. 2016].

Therefore, we propose Make-a-Face, a hands-free wearable solution that detects the tongue, mouth, and cheek position based on electromyography (EMG) sensors. The use of EMG to detect tongue position has been explored before [Zhang et al. 2014], as well as using other means like face-mounted pressure sensors [Cheng et al. 2014]. However, these approaches do not include any hardware design and focused on tongue sensing only, making them a nonviable device for public use. Make-a-Face comes with its own custom-designed mounting device that attaches to the user's ears and can be combined with a face mask for increased subtlety and privacy. Since Make-a-Face focuses on detection on the lower half of the face, it does not obstruct the use of head-mounted display(HMD), and we also developed a VR-based solution where it is mounted to the HMD for ease of use. In addition to the pre-trained tongue/mouth/cheek gestures for each application scenario, the users also can train their custom gestures using Make-A-Face getting a confidence level as feedback on how well the custom gesture can be detected.

2 SYSTEM DESCRIPTION

Make-a-Face uses the OpenBCI Cyton board with four activated channels placed on the user's face, specifically the orbicularis oris of the face, two above and two below the mouth. One reference electrode touches the ear lobe to reduce noise and the bias electrode touches another. The signals are processed with a notch filter of 60 Hz and bandpass filter of 7 to 13 Hz for further noise removal. The wires of each electrode are carefully lined through the 3D-printed mount that touches the electrodes to all the desired contact points. The Cyton board itself is powered by four 1.5V AA batteries and communicates wirelessly with our Windows or Mac computer via a universal serial bus (USB) dongle. The mount itself was designed using Fusion 3D to accommodate each of the mounted electrodes and fits onto most ear, with its design taking cues from Bluetooth-powered microphones or earphones. We also have another custom mount for VR headset, with our first prototype being for the HTC Vive. The mount hooks onto the cavity of the headset under the foam guard.

We then streamed the EMG signals via the user datagram protocol (UDP) to a Python program that further processes each of the signals as features for our machine learning algorithm. From these signals, we added four additional features; mean, standard deviation, maximum, minimum, and root mean square values to increase the accuracy. We then use the Logistic Regression algorithm to classify these features into five classes; neutral, top left, top right, bottom left, and bottom right, depending on the current tongue position and mouth deformation. The program interface displays the confidence level of each classification and a facial heat map for each gesture. From there, we further stream UDP data of the classified signals to corresponding paired devices, ranging from smart phones to VR displays.

3 USER EXPERIENCE

Based on our user experience, we found that Make-a-Face can be adapted to a wide variety of scenarios. We used Make-a-Face on everyday devices while performing activities like standing in a transit train and jogging to control our smart phone activity. To interact with a smart phone while being in transit, our device was especially useful for larger smart phones that occasionally require both hands to use, which can potentially be dangerous as one hand should be occupied on a handle at all times. For running, it was particularly useful to control our smart phone without the need to use our hands. We also developed a VR scenario that utilizes Make-a-Face for input and interaction, freeing our hands for other tracking and gesture activities. We believe it provides an especially interesting input modality for mobile VR and AR headsedts. One of the added benefits of Make-a-Face is due the mounting of the electrodes on the face instead of the interior of the mouth which can be unhygienic.

Figure 2: Make-a-Face used in various hands-free scenarios like (a) standing in a transit train with a mask, (b) running, (c) driving, and (d) gaming in VR

4 CONCLUSIONS

Make-a-Face provides the user with a hands-free input device that eliminates any intrusive or embarrassing properties of other available devices. In the future, we wish to perform an in-depth study of the accuracy of detection and the use of deep learning to allow calibration-free use cases.

REFERENCES

Jingyuan Cheng, Ayano Okoso, Kai Kunze, Niels Henze, Albrecht Schmidt, Paul Lukowicz, and Koichi Kise. 2014. On the tip of my tongue: a non-invasive pressure-based tongue interface. In *Proceedings of the 5th Augmented Human International Conference*. ACM, 12.

Stanley Coren. 2003. *Sensation and perception.* Wiley Online Library.

Mayank Goel, Chen Zhao, Ruth Vinisha, and Shwetak N Patel. 2015. Tongue-in-Cheek: Using Wireless Signals to Enable Non-Intrusive and Flexible Facial Gestures Detection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 255–258.

Zheng Li, Ryan Robucci, Nilanjan Banerjee, and Chintan Patel. 2015. Tongue-n-cheek: non-contact tongue gesture recognition. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*. ACM, 95–105.

Shuo Niu, Li Liu, and D Scott McCrickard. 2014. Tongue-able Interfaces: Evaluating Techniques for a Camera Based Tongue Gesture Input System. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility (ASSETS '14)*. ACM, New York, NY, USA, 277–278. https://doi.org/10.1145/2661334.2661395

Makoto Sasaki, Takayuki Arakawa, Atsushi Nakayama, Goro Obinata, and Masaki Yamaguchi. 2011. Estimation of tongue movement based on suprahyoid muscle activity. In *Micro-NanoMechatronics and Human Science (MHS), 2011 International Symposium on*. IEEE, 433–438.

Makoto Sasaki, Kohei Onishi, Dimitar Stefanov, Katsuhiro Kamata, Atsushi Nakayama, Masahiro Yoshikawa, and Goro Obinata. 2016. Tongue interface based on surface EMG signals of suprahyoid muscles. *ROBOMECH Journal* 3, 1 (2016), 9.

Qiao Zhang, Shyamnath Gollakota, Ben Taskar, and Raj P N Rao. 2014. Non-intrusive tongue machine interface. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 2555–2558.