W!NCE: Eyewear Solution for Upper Face Action Units Monitoring

Soha Rostaminia
University of Massachusetts Amherst
srostaminia@cs.umass.edu

Alexander Lamson
University of Massachusetts Amherst
alamson@umass.edu

Subhransu Maji
University of Massachusetts Amherst
smaji@cs.umass.edu

Tauhidur Rahman
University of Massachusetts Amherst
trahman@cs.umass.edu

Deepak Ganesan
University of Massachusetts Amherst
dganesan@cs.umass.edu

ABSTRACT

The ability to unobtrusively and continuously monitor one's facial expressions has implications for a variety of application domains ranging from affective computing to health-care and the entertainment industry. The standard Facial Action Coding System (FACS) along with camera based methods have been shown to provide objective indicators of facial expressions; however, these approaches can also be fairly limited for mobile applications due to privacy concerns and awkward positioning of the camera. To bridge this gap, W!NCE re-purposes a commercially available Electrooculography-based eyeglass (J!NS MEME) for continuously and unobtrusively sensing of upper facial action units with high fidelity. W!NCE detects facial gestures using a two-stage processing pipeline involving motion artifact removal and facial action detection. We validate our system's applicability through extensive evaluation on data from 17 users under stationary and ambulatory settings.

CCS CONCEPTS

• **Human-centered computing** → **Mobile devices**; *User studies*;
• **Computing methodologies** → **Neural networks**; • **Applied computing** → **Consumer health**.

KEYWORDS

FACS, Facial actions, EOG, J!NS MEME, Motion artifact, Eyeglasses, Wearable

ACM Reference Format:

Soha Rostaminia, Alexander Lamson, Subhransu Maji, Tauhidur Rahman, and Deepak Ganesan. 2019. W!NCE: Eyewear Solution for Upper Face Action Units Monitoring. In *2019 Symposium on Eye Tracking Research and Applications (ETRA '19)*, June 25–28, 2019, Denver, CO, USA. New York, NY, USA, 3 pages. <https://doi.org/10.1145/3314111.3322501>

1 INTRODUCTION

One of the crucial missing elements in wearable sensing is the ability to measure facial expressions in natural settings. While facial expressions can easily be monitored with cameras pointing

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).

ETRA '19, June 25–28, 2019, Denver, CO, USA

© 2019 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-6709-7/19/06.

<https://doi.org/10.1145/3314111.3322501>

Figure 1: J!NS MEME platform.

at the face, this is impractical in ambulatory settings and raises privacy issues in personal spaces. But alternate solutions have been hard to find — while wearable EEG [Emotiv 2018; MUSE 2015], electrodermal activity [Garbarino et al. 2014; PIP 2016], and respiration sensors [MD2K 2014] have all explored related topics like emotion detection, these provide relatively crude measures compared to the richness of facial expression sensing.

An intriguing new alternative that has emerged is wearable electrooculography (EOG). EOG electrodes have long been known to capture electrical potentials that exist between the cornea and ocular fundus. EOG electrodes are typically placed under, over and to the side of the eye as well as on the forehead, and the signal analyzed for tracking eye and eyelid movements. While such a setup is clearly too intrusive for normal use, recently available EOG-enabled eyeglasses, J!NS MEME [Jin 2018], use these electrodes in a much more unobtrusive manner by hiding three electrodes on the nose bridge (shown in Figure 1). The eyeglass is virtually identical to a normal pair of glasses except for the additional electronics and battery, and most importantly requires no cameras or electronics in the visual path of the user.

The availability of EOG electrodes on wearable glasses opens up a new avenue for sensing facial actions. In this work, we re-think what is considered signal and what is noise for EOG-based sensing — unlike traditional uses of EOG where the signal is eye movement and noise is all other activity, the information we need is contained in the noise. Our intuition is that the “noise” in the wearable EOG signal may carry useful information about facial expressions that are typically discarded in EOG processing pipelines.

The main challenge is that the light-weight design of electrodes in J!NS MEME results in impedance changes during head movement

Figure 2: Classification results of FACS detection: the confusion matrices show the performance for the stationary dataset, walking dataset, and the head-movement dataset.

actions and large body movements. These changes add unwanted motion artifacts that seep into the EOG signal and need to be removed if we want to capture facial expressions. In this work, we overcome this challenge by designing a motion artifact removal pipeline that leverages accelerometer and gyroscope information to remove the artifacts. Our experiments show that W!NCE can detect five distinct facial action units (i.e. blinks (AU45), brow lower (AU04), brow raiser (AU01), cheek raiser (AU06), nose wrinkler (AU09)) with a mean F1 score of 0.88 in stationary and 0.82 in ambulatory settings.

2 W!NCE PROCESSING PIPELINE

In this section, we provide a brief overview of the working of W!NCE. Our proposed method has two main computational stages — the first stage is motion artifact removal by leveraging IMU information, and the second stage is a classification stage where a Convolutional Neural Network (CNN) is used to detect facial action units (for details please refer to the full paper [Rostaminia et al. 2019]).

2.1 Stage 1: Motion Artifact Removal

Our approach is to learn a model that uses the readings of the IMU sensor embedded in the temples of the eyeglass to estimate the corresponding motion artifact on the EOG signals. For learning the IMU to noise mapping function, we design a non-linear Neural Network regression model. At a high level, this model takes all the six axes (accelerometer + gyroscope) as input and simultaneously predicts the four axes of the EOG signal resulting from the motion.

2.2 Stage 2: Facial Action Detection

Once motion artifacts are removed and a significant EOG change is detected, we run our CNN model to detect instances of action units. Our CNN model contains two convolutional layers and one fully connected layer.

3 PERFORMANCE

We present a brief evaluation of W!NCE on data collected from 17 users, 9 male and 8 female. The participants were asked to perform each action 20 times while wearing J!NS MEME for each stationary,

head-movement, and walking settings (over 1700 upper face action units and 8300 blinks). We used this data to perform experiments testing W!NCE’s effectiveness.

Table 1 shows the precision, recall and F1 score over all 17 participants for stationary, walking, and head-movement datasets. We show the results of blink classification separately from other facial actions since blinks use a shorter window size compared to the other classification tasks.

Table 1: W!NCE performance (over 1700 upper face action units and 8300 blinks). We see strong results across the board including scenarios that involve motion artifacts due to head movement and walking.

Action	Setting	Precision	Recall	F1 Score
FACS	Stationary	0.88	0.88	0.88
	Walking	0.84	0.83	0.83
	Head-movement	0.82	0.80	0.81
Blinks	Stationary	0.96	0.94	0.95
	Walking	0.90	0.91	0.90
	Head-movement	0.91	0.89	0.90

In a stationary scenario, we have an F1 score of 0.88 for FACS detection and this dips a bit to 0.83 and 0.81 for the more noisy walking and head movement scenarios. The F1 score for blink detection is higher — we get 0.95 for the stationary dataset and 0.9 for the more noisy walking and head movement datasets. Overall, these results are very promising and show that facial gestures can generally be accurately detected across a range of scenarios.

Figure 2 shows the confusion matrices of FACS detection for the different datasets. While the results are good for most of the FACS actions, “cheek raiser” seems to be confused with other classes more frequently. This is because we generally find that the signal is a bit weaker for the cheek raiser action compared to other actions due to the distance of the action from the electrodes.

ACKNOWLEDGMENTS

This work was partially supported by grants and awards from the National Institutes of Health (1U54EB020404-01) and the National Science Foundation (1815347 and 1839999).

REFERENCES

2018. JINS MEME Eyewear. <https://jins-meme.com/en/>.
- Emotiv. 2018. Emotiv: Mobile EEG Brainwear. <https://www.emotiv.com/>
- M. Garbarino, M. Lai, D. Bender, R. W. Picard, and S. Tognetti. 2014. Empatica E3 – A wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. In *2014 4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH)*. 39–42. <https://doi.org/10.1109/MOBIHEALTH.2014.7015904>
- MD2K. 2014. MD2K: NIH Center of Excellence on Mobile Sensor Data-to-Knowledge. <https://md2k.org/>
- MUSE. 2015. MUSE: The Brain Sensing Headband. <http://www.choosemuse.com/>
- PIP. 2016. PIP: The Stress Management Device. <https://thepip.com/>
- Soha Rostaminia, Alexander Lamson, Subhansu Maji, Tauhidur Rahman, and Deepak Ganesan. 2019. W!NCE: Unobtrusive Sensing of Upper Facial Action Units with EOG-based Eyewear. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 1, Article 23 (March 2019), 26 pages. <https://doi.org/10.1145/3314410>