Itchy Nose: Discreet Gesture Interaction using EOG Sensors in Smart Eyewear

Juyoung Lee[1] **Hui-Shyong Yeo**[2] **Murtaza Dhuliawala**[3] **Jedidiah Akano**[3]
Junichi Shimizu[4] **Thad Starner**[3] **Aaron Quigley**[2] **Woontack Woo**[1] **Kai Kunze**[4]

[1]KAIST, Daejeon, Republic of Korea [2]University of St Andrews, Scotland, United Kingdom
{ejuyoung, wwoo}@kaist.ac.kr {hsy, aquigley}@st-andrews.ac.uk
[3]Georgia Institute of Technology, Atlanta, Georgia, United States [4]Keio University, Yokohama, Japan
{murtaza.d.210, jakano, thad}@gatech.edu {jun.shimi, kai}@kmd.keio.ac.jp

ABSTRACT
We propose a sensing technique for detecting finger movements on the nose, using EOG sensors embedded in the frame of a pair of eyeglasses. Eyeglasses wearers can use their fingers to exert different types of movement on the nose, such as flicking, pushing or rubbing. These subtle gestures can be used to control a wearable computer without calling attention to the user in public. We present two user studies where we test recognition accuracy for these movements.

Author Keywords
Nose gesture; face gesture; subtle interaction; EOG; wearable computer; smart eyeglasses; smart eyewear;

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation (e.g. HCI): User interfaces - Input devices and strategies;

INTRODUCTION
Smart eyewear such as Google Glass, Snap's Spectacles, J!ns Meme and Microsoft HoloLens are becoming available. Yet interacting with these devices often requires actions such as swiping on the side of the device or finger tapping in mid-air. These actions may be considered disruptive, intrusive or socially unacceptable (i.e., distracting or rude) when socializing with others or during workplace interactions.

Some researchers advocate creating interactions where the use of the technology is obvious and self-explanatory to onlookers [6, 8]. Others seek to create interactions that are more subtle due to concerns of safety, awkwardness, or possible misinterpretation depending on the context of use [11, 17]. This research seeks to create control gestures which can be interpreted as typical nose touching movements. *We focus on social situations such as meetings where the user's wearable presents a notification, either through audio alerts or a head worn display, and the user would like to select one of a few options to respond to the notification without distracting others in the room.* We propose sensing based on electrooculography (EOG) sensors embedded in the J!ns Meme, a commercially available wearable computer that detects eye and head movements. Our technique allows the detection of finger flicking, pushing and rubbing on the nose. These actions can be easily disguised as scratching or rubbing one's nose. Health studies show these actions to be commonplace, with mouth/nose touches averaging 3.6 per hour [1].

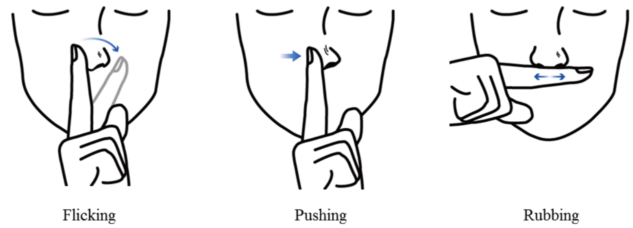

Figure 1: Three proposed input gestures.

RELATED WORK
Our work touches on several areas including EOG sensing and on-body interaction, especially on the face and nose area. The custom EOG goggles proposed by Bulling et al. [5] can efficiently recognize sequences of eye movements in real-time for interaction purposes. Using commercially available EOG glasses (J!ns Meme), Ishimaru et al. [10] identify four types of activities (typing, reading, eating and talking) at roughly 70% accuracy. Instead of eyewear, Manabe et al. attached EOG sensors to over-ear headphones [14] and in-ear earphones [15] to detect eye gestures for input and interaction. To our knowledge, using EOG to detect finger touching on the nose has been an unexplored area.

There are various works on detecting finger tapping on the body using acoustic sensing, such as the SkinPut [9] which determines the location of a finger tap by analyzing the sound that propagates through the skin. TapSkin [21] is able to achieve similar tapping detecting on the arm by using just the accelerometer and microphone in an off-the-shelf smartwatch. Bragi Dash [4] is a pair of consumer earphones that supports a simple tapping gesture on the face. However, these techniques rely on rather strong tapping, which would not be practical for the nose.

Serrano et al. [18] explore using hand to face input to interact with head-worn displays (HWDs). Due to its exploratory nature, their prototype uses a Vicon optical tracking system with infrared markers attached on users' fingers. FaceTouch [7] leverages touch-pads on the outer side of HMD for interaction with virtual content. Finally, Earput [13] explores touch interaction on the ear by placing electrodes around ear.

Instead of touching the face or ear, Stick it in your ear [3] explores using jaw movement for interaction purposes by using in-ear sensors. Similarly, Bitey [2] explores tooth click gestures for hands-free interface control. Palebrea Superioris [12] explores the design space of eyelid gestures. These works suggest that the human face can provide a rich medium for interaction with computing.

In terms of nose interaction, people occasionally use their nose to tap the screen of a phone (e.g., to replace finger or stylus), especially when the hand is covered with a glove (during winter) or the phone is too big for one-handed usage and the non-supporting hand is occupied. The Wall Street Journal [19] reports that people also use their nose to interact with smartwatches when their hands are occupied. Most related work focuses on this idea of using the nose to interact with a touchscreen, while our work focuses on detecting the finger gesturing on the nose itself. Snout [20] explores one-handed use of touch devices using the nose, and Nose-Tapping [16] presents an in-depth user study, identifying the main challenges and contributing to the design principles for nose-based interaction.

Figure 2: J!ns Meme and its EOG electrode placement.

DESIGN AND IMPLEMENTATION

The EOG sensors in the J!ns Meme are strategically placed around the nose: two on the nose pads and one on the nose bridge (Figure 2). We exploit these sensors to detect nose movement when it is being touched by a finger. We stream raw data from the J!ins Meme over Bluetooth to a remote computer for real-time processing and classification. We implemented the system in Python with Pygame for visualization and Scikit-learn for training machine learning classifiers. While our first intention was to use the nose as a joystick, we currently only support gestures that we can detect robustly: i) flick left/right ii) push left/right and iii) rubbing (Figure 1).

First we develop a simple heuristic-based system to perform a first-pass at real-time recognition during data collection. Our signal processing pipeline segments the data into one second windows. We run classification every 30ms. We perform two classifications on the data; one for classifying whether input has occurred and the other for classifying the type of input (e.g., flick, push or rubbing). We use 5 signals from the J!ns Meme: EOG left, right, horizontal, vertical, and the z-axis of gyroscope. Example signals during each of the 5 gestures

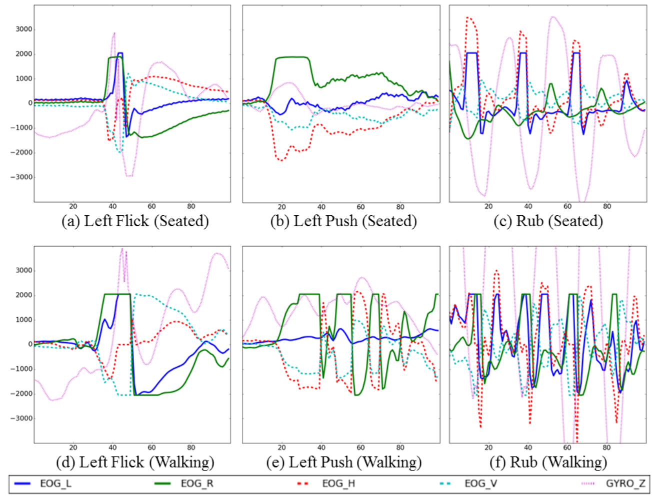

Figure 3: Example of signal patterns in both conditions.

can be seen in Figure 3. Note that the each channel of the EOG signals is derived from combinations of three EOG electrodes. We use the z-axis gyroscope as it can detect the small head rotation movement that occurs during a rubbing gesture. We extract 10 features for each signal, including 5 statistical features: (i) root mean square, (ii) mean, (iii) standard deviation, (iv) maximum and (v) minimum value of the segmented signal. The other 5 features include the number of (vi) positive and (vii) negative peak values (which are high on flick and rub), the numbers of values to cross a (viii) positive threshold and (ix) negative threshold and (x) the largest number of values that exceed these thresholds consecutively.

EXPERIMENT 1: SEATED

We imagine Itchy Nose interactions will be most useful for a user responding to notifications while attending meetings or other seated, face-to-face interactions. Thus, we first focus on collecting data while users are in a seated position. We recruited seven volunteers from a local university, including students and researchers, aged between 24 and 30 (M:26.9, SD:1.9). We asked the participants to sit on a chair and watch the display while wearing J!ns Meme. We demonstrate the 5 gestures (flick, push and rub) and allow the participants to experiment with our classifier running in real-time. At this stage, the classifier was based on simple heuristics. During the experiment, the display prompts the user with a random gesture and counts down three seconds before the data collection starts. We ask the participants to perform the gesture repeatedly with one second pauses between each iteration until the gesture is successfully recognized by the system, whereas unsuccessful recognition was confirmed manually by the experimenter. Each successful gesture was collected 10 times, resulting in 50 successful trials for the 5 gestures.

Results

As expected, we could not obtain satisfactory accuracy using our simple heuristic approach. Only 62% of attempts could be detected on the first trial, as shown on Table 1. We then used the collected data for post-hoc evaluation, training and evaluating a random decision forest (RDF) classifier. We train and

Participant	Number of trials until success				
	1	2	3	4	4<
1	42	8	0	0	0
2	33	10	4	1	2
3	31	15	4	0	0
4	31	7	2	4	6
5	27	9	7	4	3
6	31	15	3	0	1
7	21	12	6	5	6
Total	217(62%)	78(22%)	29(8%)	18(5%)	18(5%)

Table 1: Number of trials per participant for seated gestures using heuristic approach.

		Predicted (UD)					Predicted (UI)				
		(a)	(b)	(c)	(d)	(e)	(a)	(b)	(c)	(d)	(e)
Actual	Left Flick (a)	154 (96%)	5	0	0	2	149 (93%)	3	6	0	3
	Left Push (b)	2	82 (95%)	0	2	0	8	69 (80%)	0	9	0
	Right Flick (c)	0	0	132 (98%)	1	2	2	0	130 (96%)	3	0
	Right Push (d)	1	0	5	84 (93%)	0	1	3	6	80 (89%)	0
	Rub (e)	4	0	1	0	148 (97%)	4	0	1	0	148 (97%)

Table 2: Random forest classifier results for seated gestures in user dependent (UD) and user independent (UI) tests.

test the gestures in isolation. All of the gestures triggered in the real-time data collection are used, including the gestures mis-recognized by the heuristic recognizer. The first second of each gesture was segmented manually, and the RDF recognizer makes a forced choice decision on this window of data.

To explore the difficulty of the recognition task, we test our system for user dependent, user independent, and user adaptive situations, always maintaining independent training and test sets. For user dependent tests, we run 5-fold cross-validation using just the data produced by one person. For user independent tests, we use leave-one-user-out evaluation (train with n-1 users and test on the remaining user). For user adaptive tests, we simulate the situation where we have created a user independent classifier and that we collect data for a new user to personalize the models for them. For now, we are interested in the best performance possible in this situation, so we assume all the data collected for a given user is available to us (minus the test set). We use 10-fold cross-validation over all the users' data to approximate the user adaptive situation. **The average accuracies are 96.1% (STD 5%) user dependent, 93.2% (STD 8%) user independent, and 95.8% (STD 4%) user adaptive.** Confusion matrices for the user dependent and user independent tests are shown in Table 2.

All tests have relatively high accuracy, which provides some confidence that the recognition problem is manageable. However, both push gestures have relatively poor accuracy in user independent testing. Eliminating these two gestures would raise system accuracy significantly. Alternatively, perhaps users new to the Itchy Nose system could be asked to provide some push gestures for calibration, thereby increasing accuracies to an acceptable level. Encouraged by the results,

		Predicted (UD)					Predicted (UI)				
		(a)	(b)	(c)	(d)	(e)	(a)	(b)	(c)	(d)	(e)
Actual	Left Flick (a)	57 (95%)	0	3	0	0	48 (80%)	10	1	0	1
	Left Push (b)	3	55 (92%)	0	1	1	9	42 (70%)	0	4	5
	Right Flick (c)	1	0	57 (95%)	1	1	1	2	52 (87%)	4	1
	Right Push (d)	0	0	4	56 (93%)	0	0	6	11	41 (68%)	2
	Rub (e)	1	1	0	0	58 (97%)	2	8	0	0	50 (83%)

Table 3: User dependent (UD) and user independent (UI) results for walking conditions in Experiment 2.

we attempt another experiment assuming a more difficult scenario: user input while walking.

EXPERIMENT 2: WALKING

We recruited six participants from the same local university as the previous experiment, aged between 26 and 32 (M:27.8, SD:2.6). Participants were prompted to perform gestures while walking on a track created in the lab. Unlike the first experiment, no real-time feedback was provided and gestures were attempted only once for each trial. The five gestures were presented five times in random order for a total of 25 gestures per session. Each participant performed two sessions, resulting in 10 examples of each gesture per participant.

Results

We tested the collected data in the same manner as experiment 1. Average accuracies in the walking condition are 94.3% (STD 5) user dependent, 85.0% (STD 14) user adaptive, and 77.7% (STD 14) user independent.

User dependent accuracy is high, comparable to the seated experiment. However, the accuracy difference between the user dependent and user independent tests is much larger than seated condition. This result suggests that walking creates more variance in how the gestures are performed. The accuracies of the five gestures in user adaptive testing were 81.7% left flick, 73.3% left push, 86.7% right flick, 83.3% right push, and 96.7% rubbing, further supporting the idea that walking increases variability in how the gestures are performed. As with the seated condition, push gestures tend to be less well recognized than flick and rub throughout all testing. Interestingly, rub improves the most between user independent and user adaptive testing while walking. In fact, rubbing maintains approximately 97% accuracy across most conditions tested.

DISCUSSION, LIMITATIONS AND FUTURE WORK

The results suggest that the five gestures can be distinguished well when the user is walking. The second experiment suggests that some users may have to train the system in order to get sufficiently high recognition rates for a good user experience. More experimentation with features and the variability between subjects is needed to reduce the error rate while walking. However, one could imagine smart eyewear where the Itchy Nose interface is only enabled when the user is relatively still (as measured by an embedded gyroscope). This approach may be more reasonable than it first appears. While on-the-go a user may be less concerned with how using the interface is perceived and be willing to use a head worn touchpad or other interface. The Itchy Nose interface may be reserved for those times the user is in face-to-face meetings, which are generally held while seated or standing.

Eliminating the push gestures would improve recognition rates while still allowing enough gestures to select between three options when responding to a notification. Rubbing is the best recognized gesture, and we are currently pursuing experiments which suggest that continuously monitoring for the rubbing gesture would have few false positives per hour while retaining a high recognition rate, even while walking. Thus, the rubbing gesture might be used to initiate an interaction with the wearable, whereas left and right flicks and pushes might be used to navigate an interface. This set of five gestures would be sufficient to control many wearable interfaces, such as the timeline-based interface of Google Glass.

The EOG sensor in the J!ns Meme was originally designed for detecting eye movements, but here we re-purpose it to detect finger gesturing on the nose. One worry is that eye movement will cause false positives. However, we are finding that the nose gestures have a much higher EOG signal magnitude than eye movements, and a simple threshold is sufficient to avoid false triggering.

Originally, we struggled to detect vertical gestures (flick and push up/down) with satisfactory accuracy. If this limitation can be overcome, the input can be more powerful, essentially transforming the nose into a joystick controller. It could then also support more directions and gestures (circling).

In the future, we seek to test the social acceptability and noticeability of various gestures, the influence of fit and nose shape on recognition rates, nose fatigue, and cleanliness issues. We also expect to collect more training data and test the recognition system in-the-wild, where we will have participants wear a live recognition system the entire day and prompt them using SMS to attempt various gestures at random intervals.

CONCLUSION

Itchy Nose uses finger movements on the nose to command a wearable computer. With recognition accuracies in the mid-90% range, Itchy Nose may allow users to respond to notifications quickly without distracting nearby colleagues.

ACKNOWLEDGMENTS

This work is partly supported by JST Sakigake Grant No. JP-MJPR16D4.

REFERENCES

1. Alonso, W. J., Nascimento, F. C., Shapiro, J., and Schuck-Paim, C. Facing ubiquitous viruses: When hand washing is not enough. *Clinical Infectious Diseases 56*, 4 (02 2013), 617–617.
2. Ashbrook, D., Tejada, C., Mehta, D., Jiminez, A., Muralitharam, G., Gajendra, S., and Tallents, R. Bitey: An exploration of tooth click gestures for hands-free user interface control. MobileHCI (2016).
3. Bedri, A., Byrd, D., Presti, P., Sahni, H., Gue, Z., and Starner, T. Stick it in your ear: Building an in-ear jaw movement sensor. UbiComp/ISWC Adj. (2015).
4. Bragi Dash. `https://www.bragi.com/thedash/`.
5. Bulling, A., Roggen, D., and Tröster, G. Wearable eog goggles: Eye-based interaction in everyday environments. CHI EA (2009).
6. Ens, B., Grossman, T., Anderson, F., Matejka, J., and Fitzmaurice, G. Candid interaction: Revealing hidden mobile and wearable computing activities. UIST (2015).
7. Gugenheimer, J., Dobbelstein, D., Winkler, C., Haas, G., and Rukzio, E. Facetouch: Enabling touch interaction in display fixed uis for mobile virtual reality. UIST (2016).
8. Harrison, C., and Faste, H. Implications of location and touch for on-body projected interfaces. DIS (2014).
9. Harrison, C., Tan, D., and Morris, D. Skinput: Appropriating the body as an input surface. CHI (2010).
10. Ishimaru, S., Kunze, K., Uema, Y., Kise, K., Inami, M., and Tanaka, K. Smarter eyewear: Using commercial eog glasses for activity recognition. UbiComp Adj. (2014).
11. Jones, M., Robinson, S., Pearson, J., Joshi, M., Raju, D., Mbogo, C. C., Wangari, S., Joshi, A., Cutrell, E., and Harper, R. Beyond "yesterday's tomorrow": Future-focused mobile interaction design by and for emergent users. *PUC* (Feb. 2017).
12. Jota, R., and Wigdor, D. Palpebrae superioris: Exploring the design space of eyelid gestures. GI (2015).
13. Lissermann, R., Huber, J., Hadjakos, A., and Mühlhäuser, M. Earput: Augmenting behind-the-ear devices for ear-based interaction. CHI EA (2013).
14. Manabe, H., and Fukumoto, M. Full-time wearable headphone-type gaze detector. CHI EA (2006).
15. Manabe, H., Fukumoto, M., and Yagi, T. Conductive rubber electrodes for earphone-based eye gesture input interface. *PUC* (Jan. 2015).
16. Polacek, O., Grill, T., and Tscheligi, M. Nosetapping: What else can you do with your nose? MUM (2013).
17. Rico, J., and Brewster, S. Usable gestures for mobile interfaces: Evaluating social acceptability. CHI (2010).
18. Serrano, M., Ens, B. M., and Irani, P. P. Exploring the use of hand-to-face input for interacting with head-worn displays. CHI (2014).
19. Apple Watch Users Discover Another Way to Go 'Hands Free'. `https://www.wsj.com/articles/nosy-apple-watch-users-discover-another-way-to-go-hands-free-1451077454`.
20. Zarek, A., Wigdor, D., and Singh, K. Snout: One-handed use of capacitive touch devices. AVI (2012).
21. Zhang, C., Bedri, A., Reyes, G., Bercik, B., Inan, O. T., Starner, T. E., and Abowd, G. D. Tapskin: Recognizing on-skin input for smartwatches. ISS (2016).