Smooth Eye Movement Interaction using EOG Glasses

Murtaza Dhuliawala
Georgia Tech
Atlanta, Georgia, U.S.
murtaza.d.210@gatech.edu

Juyoung Lee
KAIST
Daejeon, Republic of Korea
ejuyoung@kaist.ac.kr

Junichi Shimizu
Keio University
Yokohama, Kanagawa, Japan
jun.shimi@kmd.keio.ac.jp

Andreas Bulling
Max Planck Institute
Saarbrücken, Germany
bulling@mpi-inf.mpg.de

Kai Kunze
Keio University
Yokohama, Kanagawa, Japan
kai.kunze@gmail.com

Thad Starner
Georgia Tech
Atlanta, Georgia, U.S.
thadstarner@google.com

Woontack Woo
KAIST
Daejeon, Republic of Korea
wwoo@kaist.ac.kr

ABSTRACT

Orbits combines a visual display and an eye motion sensor to allow a user to select between options by tracking a cursor with the eyes as the cursor travels in a circular path around each option. Using an off-the-shelf Jins MEME pair of eyeglasses, we present a pilot study that suggests that the eye movement required for Orbits can be sensed using three electrodes: one in the nose bridge and one in each nose pad. For forced choice binary selection, we achieve a 2.6 bits per second (bps) input rate at 250ms per input. We also introduce Head Orbits, where the user fixates the eyes on a target and moves the head in synchrony with the orbiting target. Measuring only the relative movement of the eyes in relation to the head, this method achieves a maximum rate of 2.0 bps at 500ms per input. Finally, we combine the two techniques together with a gyro to create an interface with a maximum input rate of 5.0 bps.

CCS Concepts

•**Human-centered computing** → **Interaction design; Ubiquitous and mobile computing;** *Interaction devices; Interaction techniques; Gestural input; Interaction design process and methods; Interface design prototyping;* User interface design; Empirical studies in ubiquitous and mobile computing;

Keywords

Eye tracking; gaze interaction; wearable computing

1. INTRODUCTION AND MOTIVATION

ICMI'16, November 12–16, 2016, Tokyo, Japan

http://dx.doi.org/10.1145/2993148.2993181

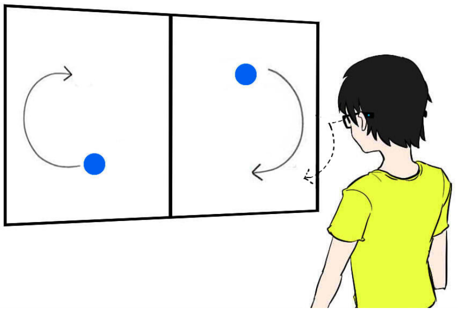

Figure 1: Eye Orbits: the user tracks an orbiting cursor with the eyes to select an option whose icon is (optionally) rendered in the center of the orbit.

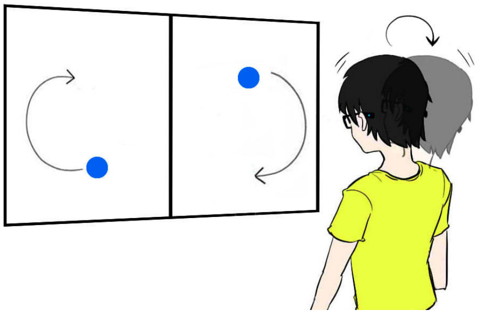

Figure 2: Head Orbits: the user to fixates on the center of rotation and move the head synchronously with the orbiting cursor.

Eye-based interfaces have long been used for communication by people with paralysis due to Amyotrophic Lateral Sclerosis (ALS) and traumatic injury. Eye interfaces have also been used for decades for virtual reality and augmented reality headsets. If a lower power and unobtrusive eye interface could be constructed which fit properly for a large part of the population, one can imagine their use for wearable computers now becoming available. With MP3 players and Bluetooth headsets embedded in sunglasses, eye gestures could be used to play, pause, or otherwise control the user's music, or an incoming phone call might be sent to voice mail with a gesture with the eyes while a different gesture answers the call.

Eye interfaces have faced many difficulties, however. Eye trackers often require controlled illumination making mobile eye tracking difficult due to the interference of the sun. In addition, computer vision based eye trackers often require significant power both due to the camera and computation involved; furthermore, embedding eye trackers in eyeglasses is difficult due to the size of battery required.

Another difficulty is that the eyes are only under semi-conscious control; they are always moving slightly which can lead to imprecision in target selection. Since the fovea encompasses two degrees of visual arc, there is additional imprecision in knowing at which point the user is looking. To complicate the issue further, eye tracking interfaces often require calibration to align eye fixation points with graphics on a computer screen, and often this calibration must be repeated each time the user dons the interface. For pointing based eye interfaces, interaction designers typically trade time to improve accuracy. For example, interfaces that use the eyes as a mouse pointer typically require the user to dwell on a target for selection. The result is an interface which is slow, has a high perceived workload, and leads to discomfort.

More recently, several interfaces exploit smooth pursuit eye motion, where the eye tracks a target on the screen for selection [21, 18, 20]. These interfaces are quite promising as they rely on relative motion, not absolute position, for selection. Tracking of an object can be less tiring than fixating at one spot, and eye tracker calibration might be eliminated with interfaces that depend only on relative motion.

We adopt the Orbits method of eye gaze selection which has shown to have high accuracy and to be robust to false positives [3], and we make several contributions to this method. First, we demonstrate that the Orbits method can be performed with an unobtrusive, low cost, low power eye movement sensor embedded in a pair of eyeglasses. Specifically, we use the Jins MEME [7] which employs three electrodes to sense eye movement through electrooculography (EOG): one electrode on the bridge of the nose and one on each of the nosepads of the eyeglasses. Since only relative movement is used (Figure 1), calibration is not required; in development the authors could often hand one Jins MEME between them for rapid testing. Secondly, we extend the Orbits method by combining it with head motion to create a new interaction technique we call Head Orbits (Figure 2). Finally, we introduce information transfer rate as a metric for these interfaces.

2. RELATED WORK: EYE GAZE

Vidal et al. proposed to use smooth pursuit eye movements for calibration-free gaze interaction [18, 19]. Instead of being based on gaze location like conventional gaze interaction techniques, their technique correlated pursuit movements with objects dynamically moving on the interface. Similar techniques followed for eye tracker calibration [13], user authentication using PINs [1] and text entry [10]. More recently, Esteves et al. introduced a hands-free and calibration-free interaction technique for smartwatches employing pursuits on orbital targets rotating on the interface along radial trajectories [3]. Kangas et al. compared different feedback modalities (visual, auditory, haptic, none) in a continuous adjustment technique for smooth pursuit gaze tracking and found clear user preferences for haptic and audio feedback [8]. In another work they compared smooth pursuit based widgets and one time-based dwell widget for adjusting a continuous value [16]. The circular smooth pursuit widget was found to be about equally efficient as the dwell based widget. Khamis et al. conducted a field study on spontaneous gaze-based interaction with a public display using Pursuits [9]. Jalaliniya et al. propose a technique to detect a user's object of interest by analyzing eye movements while presenting a visual stimulus moving horizontally or vertically [6]. Finally, Schenk et al. presented SPOCK, a gaze interaction technique based on smooth pursuits for static user interfaces [14]. While all of these works demonstrated the flexibility and appeal of pursuits for natural and calibration-free interaction, they relied on video-based eye tracking – either remote or head-mounted.

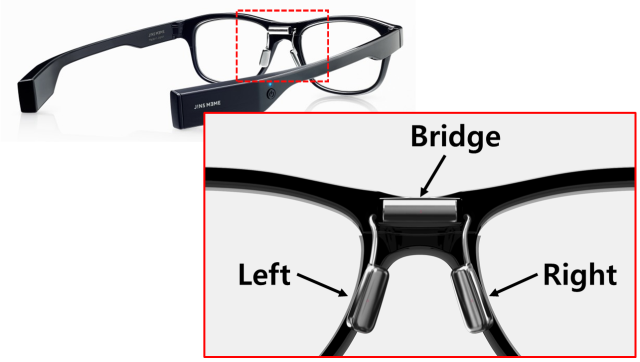

Figure 3: Jins MEME and its electrodes: Bridge, Left, Right

3. RELATED WORK: HEAD GESTURES

Recognition of implicit head gestures has been used to infer social interactions and mental states [2, 12], but here we focus on using head movements for intentional, hands-free operation. Google Glass uses head gestures for acknowledging and dismissing notifications [17], and Hansen et al. [5] combine gaze interaction with wristworn devices. Several works use head gestures for interacting with head mounted displays for people with disabilities [4] or to enhance other input modalities such as eye gestures [11, 15].

4. ORBITS ON JINS MEME

In the original Orbits paper, Esteves et al. [3] used correlation between the x and y movements of the orbiting cursor displayed on a screen and the x and y movements of the user's eye observed by the eye tracker. We wish to use the

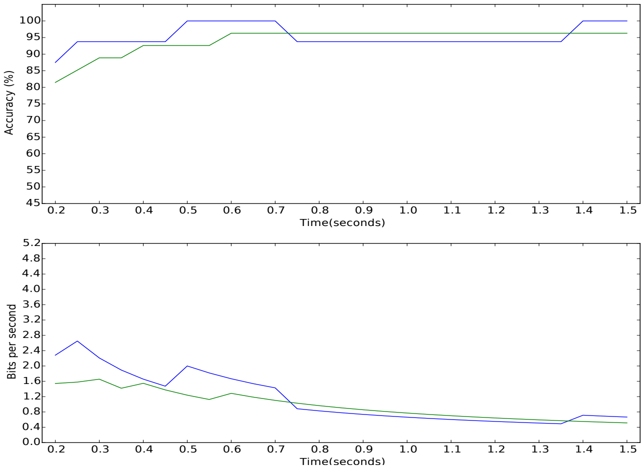

Figure 4: Eye Orbits for two users: Decision time versus accuracy (top) and bits per second (bottom)

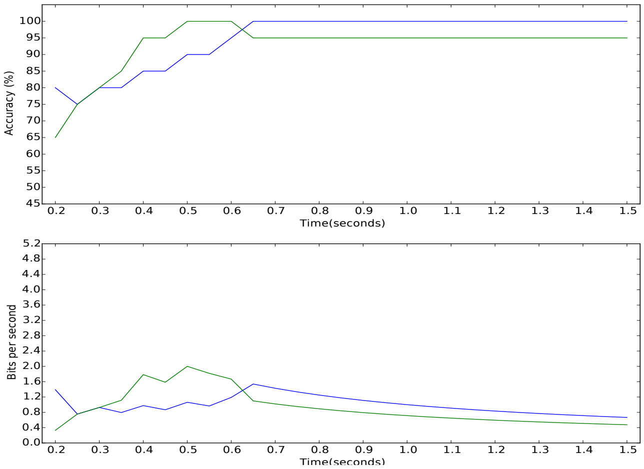

Figure 5: Head Orbits for two users: Decision time versus accuracy (top) and bits per second (bottom)

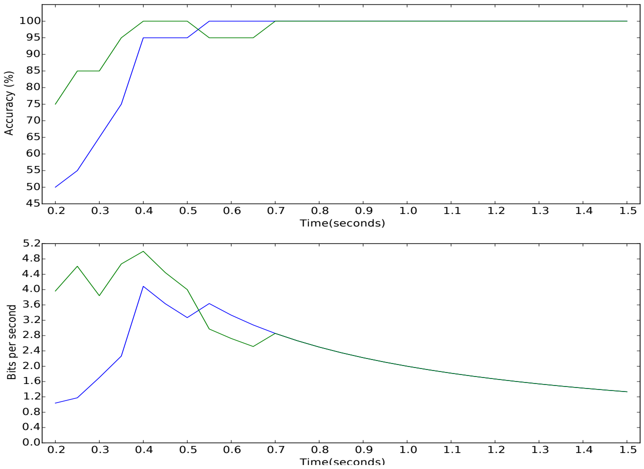

Figure 6: Hybrid Orbits for two users: Decision time versus accuracy (top) and bits per second (bottom)

signals captured by the electrodes on the Jins MEME (Figure 3), specifically, on the left nosepad (L), right nosepad (R), and bridge of the nose (B). From these we calculate both a horizontal signal $(L-R)$ and a vertical signal $(B-\frac{L+R}{2})$. Since we expect to observe circular motion, we convert these signals to polar coordinates, r and θ. To establish a center of rotation for this calculation, we average the vertical and horizontal signals over time. This process is started with two samples and is continued for a maximum 100 samples, maintained in a circular buffer. Surprisingly, this method is robust even with only a few samples with which to estimate the center. When collecting data over USB, the Jins MEME provides a consistent frame rate of 30 samples per second. For convenience, we also record the position of the orbiting cursor at 30Hz in r and θ.

For prompts, we present right and left orbiting cursors, as seen in Figure 1, that travel clockwise and remain 180° out of phase. In this manner both the motion and the absolute angle of the prompts are most distinct from each other. For each trial, our goal is to determine if the user's eye movement is more consistent with the right or left orbits in as short a time as possible. To determine the shortest time required to make an accurate decision, we compare varying length time windows of eye movement with the orbiting cursors. The first θ value in the eye movement data is matched to the closest θ value for the left and right orbits within $\frac{5}{30}$ths of a second of the correct time stamp. This step helps compensate for clock skew and small offsets in the user's head angle. If any match is within 10°, the pointer into the graphical orbit data is updated to match the entry with the minimal difference in the Jins MEME data. Otherwise, the pointer is advanced only by one. This process continues for each subsequent data point returned by the Jins MEME up to the desired window of time. For each step where the compared θ values are within 10°, we consider them a match. The orbit with the most matches to the eye data is considered to be correct.

5. HEAD ORBITS ON JINS MEME

Head Orbits takes advantage of the vestibulo-ocular reflex to elicit smooth eye motion relative to the moving head as the eyes attempt to remain fixed on a target (Figure 2). Even though there is significant head motion, the Jins MEME successfully detects this eye motion. For discriminating between the user following a left or right orbit, the same comparison method can be used as with Eye Orbits. Because the user is fixating his eyes on the center of the orbit and moving the head, the relative eye movement will be 180° out of phase with the orbiting cursor. Thus, while the same algorithm is used as with Eye Orbits, the decision is flipped. For situations where all four user responses are possible (Left and Right Eye Orbit and Left and Right Head Orbits), we compare the difference of the maximum and minimum values of the gyro recorded during each Orbits comparison window. Orbits detected where the gyro difference is over a threshold are labeled as Head Orbits. In the Jins MEME, the gyro is located in the pod over the ear, which provides surprisingly stable data.

6. EXPERIMENTS

For each trial, we display the first animation frame of a left and right orbit to the user. The cursor for the left orbit is at the bottom-most position while the cursor for the right is at the top-most position. We ask the user to look at the left or right cursor and then start tracking the movement of the cursor. The participant is expected to track the movement of the cursor until we signal that the trial is over. We ask that the participant re-start the trial if they indicate they made a mistake (e.g., tracked the wrong target, are not looking at the cursor when it starts orbiting, etc.). Each trial lasts three seconds, and the orbits are set to one revolution per second. The diameter of the orbits on the screen is 5.75 inches. We request the participant stay approximately 8 inches from the screen, though we did not enfocce that constraint. In more anecdotal testing, we found that the methods work within a wide range of distances from the screen.

Both subjects performed 20 random trials for each of the eye, head, and hybrid orbit situations. Trials were balanced left versus right and head orbits versus eye orbits, as appropriate. The matching algorithm was run off-line on time windows ranging from 200ms to 3 seconds from the start of the motion of the cursors. The algorithm is forced to decide between a left or right orbit.

7. RESULTS

Figure 4 graphs decision time versus accuracy for the two participants for Eye Orbits. Accuracy is surprisingly high at 200ms and reaches a maximum at 600ms for each. We also calculate information transfer rate, following the typical metric used in the brain computer interface literature to compare techniques. Bits conveyed per decision can be calculated by the formula

$$B = log_2(N) + Plog_2(P) + (1 - P)log_2(\frac{1 - P}{N - 1})$$

where N is the number of targets and P is the probability that each target is hit. Bit rate is calculated by dividing by the amount of time per trial. While it is unreasonable to expect a user to go from one pair of orbits to the next immediately and to expect correction from errors would be optimally efficient, we can use this bit rate as a maximum for comparison between different conditions. For example, the bottom graph of Figure 4 plots decision time versus bit rate for each of the participants. The first participant achieves a surprising 2.6 bits per second (bps) input rate at 250ms per input. Figure 5 shows the accuracy and bitrate achieved by the participants for Head Orbits. The Head Orbits method achieves a maximum rate of 2.0 bps at 500ms per input. Figures 6 shows the accuracy and bit rate curves for the hybrid of Eye and Head Orbits. The combined forced four choice experiment shows a maximum input rate of 5.0 bps at 400ms per input.

8. DISCUSSION

During testing, significant noise was observed when the participant moved, as is expected with EOG systems. However, once the user was stationary and tracking an orbit, the system showed low amplitude but consistent signals. In addition, the authors could easily hand the device between themselves without requiring calibration.

Such a system may have more promise than it first appears. Users with certain disabilities like ALS expect to be stationary when providing input to a computer. A game system that uses 3D shutter glasses and a flat panel display could make being stationary to use the eye interface part of the game mechanism. Mobile users are often stationary during times they desire hands-free, silent control of their devices, such as when attending a meeting and triaging incoming notifications. By using LEDs mounted around the frame of the lenses of a pair of eyeglasses plus three electrodes in a nose bridge, Eye Orbits could be an inexpensive and low power way to interact silently with eyewear computing. Adding a low power gyro would add the capability for simultaneous use of Eye & Head Orbits.

One difficulty lies in how obvious, and how tiring, the large eye gestures are for Eye and Head Orbits. While this experiment used large movements of the eyes, anecdotal testing showed that much smaller Eye and Head Orbits could still be discriminated with high accuracy.

9. CONCLUSION AND FUTURE WORK

The experiments above use off-line data to explore what maximum information transfer rate might be obtained with Eye, Head, and Hybrid Orbits approaches and forced choice binary decisions. In an on-line system, feedback to the user might improve accuracy and allow the user to decline either choice. For example, as the user follows an orbit, the system may calculate its certainty that each orbit is being followed by comparing the percentage of samples that are being matched with that orbit within a window of the most recent samples. This metric might be visualized by an expanding green circle at the center of the corresponding orbit. When a certain threshold is reached, the expanding circle's circumference matches the orbit and that option is selected. If the user sees that an improper option is being selected, the user can just fixate his eyes at the center of the circle and watch the circle shrink. After a certain amount of time, neither option is selected. In addition to experimenting with real-time interfaces and visualizations, future work will increase the number of participants, experiment with optimal sizes and speeds of orbits, and expand the number of prompts per trial in an effort to determine the maximum bit rate possible with these techniques. One hope is that these interface techniques, the Jins MEME, and a VR headset might be combined to allow a silent, hands-free, and calibration-free method of interacting with virtual environments.

10. REFERENCES

[1] D. H. Cymek, A. C. Venjakob, S. Ruff, O. H.-M. Lutz, S. Hofmann, and M. Roetting. Entering pin codes by smooth pursuit eye movements. *Journal of Eye Movement Research*, 7(4):1–11, 2014.

[2] R. El Kaliouby and P. Robinson. Real-time inference of complex mental states from facial expressions and head gestures. In *Real-time vision for human-computer interaction*, pages 181–200. Springer, 2005.

[3] A. Esteves, E. Velloso, A. Bulling, and H. Gellersen. Orbits: Enabling gaze interaction in smart watches using moving targets. In *UIST*, pages 457–466, 2015.

[4] J. Gray, P. Jia, H. H. Hu, T. Lu, and K. Yuan. Head gesture recognition for hands-free control of an intelligent wheelchair. *Industrial Robot: An International Journal*, 34(1):60–68, 2007.

[5] J. P. Hansen, H. Lund, and F. Biermann. Wrist - worn Pervasive Gaze Interaction. In *ETRA*, pages 57–64, 2016.

[6] S. Jalaliniya and D. Mardanbegi. Eyegrip: Detecting targets in a series of uni-directional moving objects using optokinetic nystagmus eye movements. In *CHI*, pages 5801–5811, New York, NY, USA, 2016. ACM.

[7] Jin CO., LTD. JINS MEME ES. https://jins-meme.com/en/products/es/, 2015.

[8] J. Kangas, O. Špakov, P. Isokoski, D. Akkil, J. Rantala, and R. Raisamo. Feedback for smooth pursuit gaze tracking based control. In *Augmented Human*, pages 6:1–6:8, New York, NY, USA, 2016. ACM.

[9] M. Khamis, F. Alt, and A. Bulling. A field study on spontaneous gaze-based interaction with a public display using pursuits. In *Ubicomp*, pages 865–874, 2015.

[10] O. H.-M. Lutz, A. C. Venjakob, and S. Ruff. Smoovs: Towards calibration-free text entry by gaze using smooth pursuit movements. *Journal of Eye Movement Research*, 8(1):2, 2015.

[11] D. Mardanbegi, D. W. Hansen, and T. Pederson. Eye - based head gestures. In *ETRA*, pages 139–146, 2012.

[12] L.-P. Morency, C. Sidner, C. Lee, and T. Darrell. Head gestures for perceptual interfaces: The role of context in improving recognition. *Artificial Intelligence*, 171(8):568–585, 2007.

[13] K. Pfeuffer, M. Vidal, J. Turner, A. Bulling, and H. Gellersen. Pursuit calibration: Making gaze calibration less tedious and more flexible. In *UIST*, pages 261–270, 2013.

[14] S. Schenk, P. Tiefenbacher, G. Rigoll, and M. Dorr. Spock: A smooth pursuit oculomotor control kit. In *Extended Abstracts*, CHI, pages 2681–2687, New York, NY, USA, 2016. ACM.

[15] O. Špakov. Enhanced Gaze Interaction Using Simple Head Gestures. In *Ubicomp*, pages 205–210, 2012.

[16] O. Špakov, P. Isokoski, J. Kangas, D. Akkil, and P. Majaranta. Pursuitadjuster: an exploration into the design space of smooth pursuit–based widgets. In *ETRA*, pages 287–290. ACM, 2016.

[17] T. Starner. Project glass: An extension of the self. *Pervasive Computing, IEEE*, 12(2):14–16, 2013.

[18] M. Vidal, A. Bulling, and H. Gellersen. Pursuits: Spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. In *Ubicomp*, pages 439–448, 2013.

[19] M. Vidal, A. Bulling, and H. Gellersen. Pursuits: Spontaneous eye-based interaction for dynamic interfaces. *ACM SIGMOBILE Mobile Computing and Communications Review*, 18(4):8–10, 2015.

[20] M. Vidal, K. Pfeuffer, A. Bulling, and H. Gellersen. Pursuits: eye-based interaction with moving targets. In *Extended Abstracts CHI*, pages 3147–3150, 2013.

[21] R. Yonetani, H. Kawashima, T. Hirayama, and T. Matsuyama. Gaze probing: Event-based estimation of objects being focused on. In *International Conference on on Pattern Recognition*, pages 101–104. IEEE, 2010.