Full-time Wearable Headphone-Type Gaze Detector

Hiroyuki Manabe
NTT DoCoMo, Inc.
3-5 Hikarinooka,
Yokosuka, Kanagawa, Japan
manabe@mml.yrp.nttdocomo.co.jp

Masaaki Fukumoto
NTT DoCoMo, Inc.
3-5 Hikarinooka,
Yokosuka, Kanagawa, Japan
fukumoto@mml.yrp.nttdocomo.co.jp

CHI 2006, April 22–27, 2006, Montréal, Québec, Canada.
ACM 1-59593-298-4/06/0004.

Abstract
A headphone-type gaze detector for a full-time wearable interface is proposed. It uses a Kalman filter to analyze multiple channels of EOG signals measured at the locations of headphone cushions to estimate gaze direction. Evaluations show that the average estimation error is 4.4° (horizontal) and 8.3° (vertical), and that the drift is suppressed to the same level as in ordinary EOG. The method is especially robust against signal anomalies. Selecting a real object from among many surrounding ones is one possible application of this headphone gaze detector.

Keywords
Gaze interface, EOG, wearable, headphone, Kalman filter

ACM Classification Keywords
H5.2.

Introduction
Full-time wearable devices are daily commodities, in which we wear wrist watches and bear audio players and cellular phones for example. The wearable interface suits these devices due to its features; the user can access the interface immediately, anywhere desired. For full-time wearable devices, the interface should be easy to wear, easy to use and not obstruct daily life. In this article, the "full-time wearable interface" is defined as an interface that the user can wear continuously

without obstructing daily life and can use easily and immediately whenever desired.

Since gaze represents what the user is paying attention to and is directly controlled by the user's will, it is often used for natural and hands-free interfaces. The most commonly-used approaches are limbus tracking, the cornea reflex method, and electro-oculography (EOG). However, all of them fail to yield full-time wearable interfaces. In image-based methods, the bottleneck is the sensor's position; this obscures user's field of view and its measurement range is limited. Fixing the sensor to the user's PC or desk limits the mobility of the user. EOG-based methods have the benefits of wide measurement range, easy to use, and not obscuring the user's view. However, existing methods are not cosmetically acceptable for daily use because electrodes must be attached near the eyes and frequent calibration is required to counter drift. Sensors can be fixed to caps in the image-based method [1] or to head bands in the EOG-based method [2]; but the problems mentioned above are not resolved.

Our goal is a full-time wearable gaze detector that does not obscure the user's view, has wide measurement range, and of course is cosmetically acceptable. Our idea is a headphone-type EOG-based gaze detector. Headphones are becoming a common full-time wearable device due to the popularity of digital audio players, and installing the electrodes into the headphones does not obscure the user's view. The headphone-type gaze detector is effective since headphones are often used and so gaze detection and a gaze interface can be accessed whenever desired. The issues to be overcome are drift, low SNR, and the separation of horizontal / vertical components. The proposed method builds a model of measured signals and gaze, extracts EOG from an electrode array, and estimates gaze position by using a Kalman filter.

EOG-based gaze detection

Since the human eyeball has a negative charge on the retina and a positive charge on the cornea, the electric field around the eye changes with eye movement. EOG-based gaze detection puts electrodes near the eye to detect the electric field (EOG), and estimates gaze direction. Many EOG-based gaze detectors have been used in interfaces [6,7]. The problem faced by EOG-based gaze detection, frequent periodic calibration, is caused by the presence of signal artifacts. These artifacts are drift, blink, overshoot, and small ripple [3]. Among them, drift is the most critical. Drift is a slow change in the DC component unrelated to eye movement. Although various factors have been advanced as causes of drift such as changing skin resistance, variable cornea-retinal potential due to light accommodation [4], hardware noise, and external noise, the primary cause is electrode polarization [5]. The common solution is to use an AC amplifier, but this gives only the velocity component of gaze position, and loses its absolute value. Moreover, using AC amplifier can not detect smooth pursuit. Several drift reduction or calibration methods have been proposed [3,6,7], but either they are not easy to use or they support only a few use cases.

Mounting the EOG electrodes on the headphone eliminates the appearance problem but the calibration or drift problem remains. Moreover, this placement raises two other issues: low SNR and poor separation of horizontal and vertical components. The SNR is low because the electrodes are far from the signal source. The separation problem is generally solved by placing horizontal electrodes on the right and left sides of the eye and vertical electrodes above and below the eye; such an arrangement is not possible with headphone mounting.

Properties of "headphone EOG"

In order to build the model needed, an experiment was performed. Four electrodes were attached to the head at the locations where headphone cushion would rest in normal use as shown in Fig. 1. They are called headphone electrodes and the EOG calculated from their output is called headphone EOG in this paper. Four regular EOG electrodes were placed in the ordinary EOG arrangement near the eyes. One earth and one reference electrode were also used for a total of 14 electrodes. DC amplifiers (bandwidth of DC–20Hz) were used, and 500ms median filter was applied to reject blink. These amplifiers and filters were also used in the other experiments described in this paper.

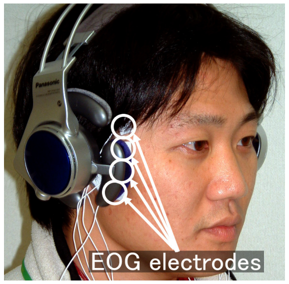

Fig. 1. Placement of headphone electrodes.

Drift

The differences between the reference and each electrode were measured when the subject's face was fixed and he was gazing at a fixed target after moving the gaze to the maximum extent both horizontally and vertically in sequence. Each signal is adjusted to 0 at time 0, as shown in Fig. 2. Since the EOG signal actually used is the difference between EOG electrodes, we note that the drift shown in Fig. 2 is not same as the regular EOG drift. The spikes are the remaining blink artifacts. The drift during 50-250sec. is of the same order as the dynamic range of ordinary EOG when moving gaze maximally. Moreover Fig. 2 also shows that the drift is electrode dependent. This means that the drifts of ordinary EOG and headphone EOG are of the same order and it is expectable to reduce the drift by statistical techniques.

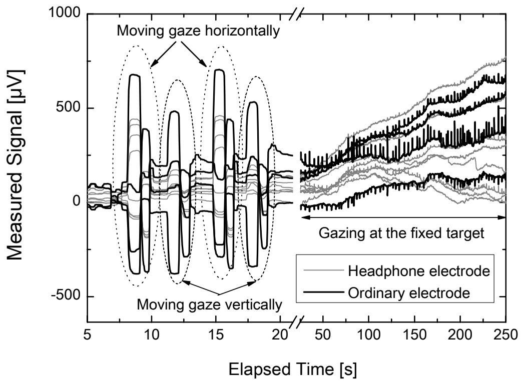

Fig. 2. Measured EOG signals.

Amplitude and horizontal / vertical components

Fig. 2 also provides information on headphone EOG. The dynamic range of horizontal (vertical) headphone EOG is about 50% (90%) smaller that of ordinary EOG. This means that the drift of headphone EOG corresponds to about 2–10 times of that of ordinary EOG. Furthermore, all headphone electrodes detect both horizontal and vertical components. Moreover, for the two headphone electrodes that are placed on either side of the head, the vertical components are in phase while the horizontal components have opposite phase.

For the two vertically-aligned headphone electrodes on the same side of the face, the horizontal components are in phase but have different amplitude; the vertical components exhibit greatly different gain. This analysis suggests that the horizontal headphone EOG should be calculated as the difference between electrodes on either side while the vertical headphone EOG should be taken as the sum of the two differences of each vertical pairs of electrodes on the same side to cancel the horizontal component. However, this simple processing for headphone EOG fails to completely separate the two components

Model

We modeled the measured EOG signals. Since it has been reported that EOG is basically proportional to gaze within the angular range of ±50° horizontally and ±30° vertically [7], the linear model is used. Let ***x*** be the gaze position, ***H*** the transform matrix, and ***e*** the noise (includes drift and DC offset). The difference of measured EOG signal is given as

$$\Delta \boldsymbol{EOG}_t = \boldsymbol{H}_t \cdot \Delta \boldsymbol{x}_t + \Delta \boldsymbol{e}_t \tag{1}$$

where

$$\begin{aligned} \Delta EOG_i(t) &= EOG_i(t) - EOG_i(t-1) \\ \Delta \boldsymbol{EOG}_t &= \{\Delta EOG_1(t), \cdots, \Delta EOG_n(t)\}^{\mathrm{T}} \end{aligned} \tag{2}$$

DC offset does not alter $\Delta \boldsymbol{e}_t$, so $\Delta \boldsymbol{e}_t$ is assumed to represent drift and other noise. Equation (1) and equation (3) can be solved by applying a Kalman filter.

$$\Delta \boldsymbol{x}_{t+1} = \boldsymbol{F}_t \cdot \Delta \boldsymbol{x}_t + \boldsymbol{G}_t \cdot \boldsymbol{w}_t \tag{3}$$

where $\boldsymbol{F}_t$, $\boldsymbol{G}_t$, $\boldsymbol{w}_t$ is a state transition matrix, driving matrix, and plant noise, respectively. The gaze position is given as $\Sigma \Delta \boldsymbol{x}_t$.

Implementation

The implemented system used 4 headphone electrodes on each side, and 2 electrodes for earth and reference, see Fig. 1. The 16 horizontal EOGs and 36 vertical EOGs were then calculated. The horizontal EOGs correspond to all combinations on either side (4 x 4), and the vertical EOGs to the combinations of the 2 sides using 2 electrodes selected from one side (${}_4C_2$ x ${}_4C_2$). Gaze position is estimated every 32ms. ***H*** can be estimated by moving gaze horizontally and vertically. $\boldsymbol{R}_0$ (the initial value of $\boldsymbol{R}_t$ which is the covariance matrix of $\Delta \boldsymbol{e}_t$) is estimated by fixing gaze. We note that using multiple electrodes increases the risk of signal anomalies such as those caused by electrode detachment. Accordingly, the system has a function of updating $\boldsymbol{R}_t$ dynamically to improve robustness against signal anomalies.

Evaluations

Three evaluations were performed. One confirmed the suppression of the drift, the second checked the robustness against electrode disconnection, and the third checked accuracy of gaze estimation.

Drift suppression

Fig. 3 shows that the drift of ordinary EOG, one example of headphone EOG, and estimated gaze position by proposed method. In Fig. 3, each EOG / gaze position is normalized by its change when moving gaze to the maximum extent horizontally or vertically at the beginning of the experiment (0–20sec.). After that the gaze is fixed on the target. In proposed method, the drift is suppressed to the same low value as ordinary EOG. Considering that the drift of headphone EOG is 2–10 times than that of ordinary EOG, the proposed method reduces the drift by factors ranging from 2 to 10. Fig. 3 also indicates that the proposed method does not need calibration for periods of at least several minutes.

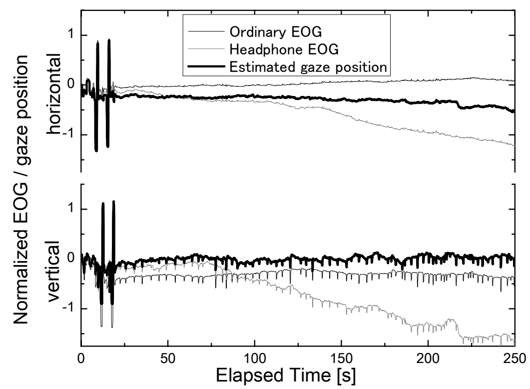

Fig. 3. Suppression of drift.

Robustness against cable disconnection

Fig. 4 shows 2 estimated horizontal gaze positions using fixed $\boldsymbol{R}_t$ and using adaptively updated $\boldsymbol{R}_t$ when an electrode cable is disconnected to replicate the detachment of an electrode from the skin. $\boldsymbol{R}_t$ was estimated as the differential components between $\Delta\boldsymbol{EOG}_t$ and its projection to the plane which horizontal and vertical components of $\boldsymbol{H}_t$ stretched, was considered as noise component. The effective projection coefficients were restricted. At first disconnection, both estimated positions change strongly. After some disconnections, however, estimated position using adaptively updated $\boldsymbol{R}_t$ recovers its original value, unlike that with the fixed $\boldsymbol{R}_t$. This suggests that adaptively $\boldsymbol{R}_t$ updating allows gaze can be estimated correctly in the event of the detachment of one electrode.

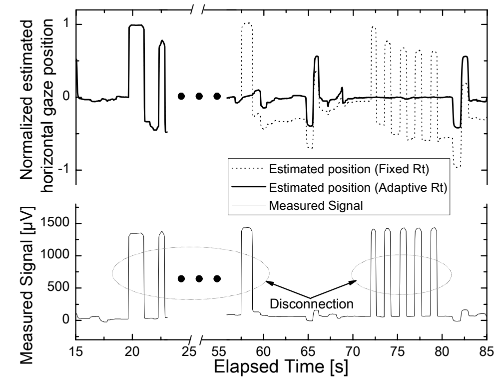

Fig. 4. Dynamic updating of the covariance matrix of noise.

Accuracy of gaze estimation

Accuracy of gaze estimation by proposed method was measured using 15 targets which were placed on a 5 (horizontal) by 3 (vertical) grid. The interval of the grid was 20°. The estimated position was calibrated first by asking the subjects to gaze at one of four corner targets located at ±50° horizontal and ±30° vertical. The subjects were asked to first gaze at the center target, then move and fix their gaze for at least 5sec. on the target which were randomly selected. Estimation error between the target and estimated gaze position (in degrees) was calculated as the RMS average during the 5 second period; the results (averaged) of 6 subjects are shown in Fig. 5. The horizontal (vertical) radius of the ellipse represents the horizontal (vertical) error for the target. Overall (average) estimation error is 4.4° horizontal and 8.3° vertical. The vertical error exceeds the horizontal error for all targets. The error increases with distance from the center. Moreover, the vertical error of lower targets is larger than that of upper targets. When gazing at a lower target, we found the initial estimated vertical position was often correct but it moved upwards with time. This rarely happened when gazing at the upper target.

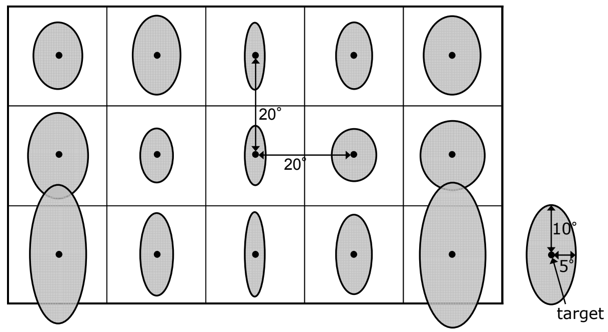

Fig. 5. Estimation error for each target.

Possible applications

This system can be used as a simple controller for many daily use devices or applications, such as audio players. It can also be used as a selector that allows the user to choose surrounding objects. When the gaze detector is supplemented with a video camera and a wireless communication device and the surrounding objects have identifying tags like QR codes, the user can get information about the object of interest simply by gazing it.

Discussion and Conclusion

We proposed the headphone-type gaze detector as a full-time wearable interface. It uses an electrode array and processes measured signals by a Kalman filter. Tests confirmed that proposed method suppresses the drift to the same low level as ordinary EOG, is robust against signal anomalies, and estimates gaze with an error of 4.4° (horizontal) and 8.3° (vertical). With the goal of improving the accuracy, the simplest approach is to use more electrodes to suppress the drift because proposed method is based on statistical processing. Moreover, since the estimation error is also considered to be caused by several factors such as nonlinearity between EOG and gaze, mismatch with the model, and error in estimating $\boldsymbol{R}_t$, building a more accurate model and/or applying an existing robust estimation method are future tasks. We also need to develop headphones that include the electrodes and amplifiers.

Reference

[1] nac image technology, EMR-8: http://www.eyemark.jp/lineup/EMR-8/EMR-8b.html.

[2] Lims technology, NeuroGate: http://lims.co.kr/eng/neurogate.htm.

[3] Kawasaki K., Tamura T.: Automated measurement of the electro-oculogram for clinical use, Doc. Ophthalmol. Vol.66 (1987), 85-94.

[4] Shaviv, B. D.: The design and improvement of an eye controlled interface, Research report, 2002, from http://www.cs.sunysb.edu/~vislab/projects/eye/Reports/report/report.pdf.

[5] Gu J. J., Meng M., Cook. A., Faulkner M. G.: A study of natural eye movement detection and ocular implant movement control using processed EOG signals, Proc. IEEE Conf. Robotics & Automation (2001), 1555-1560.

[6] Patmore W. D., Knapp B. R.: Towards an EOG-based eye tracker for computer control, Proc. ACM Conf. Assistive technologies (1998), 197-203.

[7] Chen Y., Newman S. W.: A human-robot interface based on electrooculography, Proc. IEEE Conf. Robotics & Automation (2004), 243-248.

[8] Bonnet P. Buzenac V. Baylou P., Najim M., Paty J.: EOG segmentation using Kalman and hysteresis filters, Proc. IEEE Conf. Eng. Med. Biol. (1992), 2570-2571.