Vision Research 118 (2016) 48–59

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Pupil size dynamics during fixation impact the accuracy and precision of video-based gaze estimation

Kyoung Whan Choe [a], Randolph Blake [a,b], Sang-Hun Lee [a,*]

[a] Department of Brain and Cognitive Sciences, Seoul National University, Seoul 151-742, Republic of Korea
[b] Vanderbilt Vision Research Center and Department of Psychology, Vanderbilt University, Nashville, TN 37240, USA

ARTICLE INFO

Article history:
Received 11 July 2014
Received in revised form 8 December 2014
Available online 9 January 2015

Keywords:
Eye tracking
Fixation monitoring
Pupillometry
Pupil decentration
Artifact correction
Ellipse vs centroid pupil tracking

ABSTRACT

Video-based eye tracking relies on locating pupil center to measure gaze positions. Although widely used, the technique is known to generate spurious gaze position shifts up to several degrees in visual angle because pupil centration can change without eye movement during pupil constriction or dilation. Since pupil size can fluctuate markedly from moment to moment, reflecting arousal state and cognitive processing during human behavioral and neuroimaging experiments, the pupil size artifact is prevalent and thus weakens the quality of the video-based eye tracking measurements reliant on small fixational eye movements. Moreover, the artifact may lead to erroneous conclusions if the spurious signal is taken as an actual eye movement. Here, we measured pupil size and gaze position from 23 human observers performing a fixation task and examined the relationship between these two measures. Results disclosed that the pupils contracted as fixation was prolonged, at both small (<16 s) and large (~4 min) time scales, and these pupil contractions were accompanied by systematic errors in gaze position estimation, in both the ellipse and the centroid methods of pupil tracking. When pupil size was regressed out, the accuracy and reliability of gaze position measurements were substantially improved, enabling differentiation of 0.1° difference in eye position. We confirmed the presence of systematic changes in pupil size, again at both small and large scales, and its tight relationship with gaze position estimates when observers were engaged in a demanding visual discrimination task.

1. Introduction

Video-based eye trackers estimate gaze positions by inferring the center of the pupil from sampled video images of the eye (Merchant, Morrissette, & Porterfield, 1974; Young & Sheena, 1975). Because they are noninvasive, easy to use, and robust, particularly compared to the alternative method relying on magnetic search coils (Collewijn, van der Mark, & Jansen, 1975; Robinson, 1963), video-based eye trackers are widely used for monitoring eye movements and for enforcing strict fixation in behavioral and neuroimaging experiments on humans. Despite its merits and popularity, however, this technique has a potentially serious drawback: it can generate spurious eye movement signals up to several degrees in visual angle (Drewes, Masson, & Montagnini, 2012; Ivanov & Blanche, 2011; Kimmel, Mammo, & Newsome, 2012; Wyatt, 2010) mainly because pupil centration changes as pupil size changes (Charlier, Behague, & Buquet, 1994; Walsh, 1988; Wildenmann & Schaeffel, 2013; Wilson, Campbell, & Simonet, 1992; Wyatt, 1995; Yang, Thompson, & Burns, 2002). Although this drawback was recognized by the inventors themselves when they first described the video-based eye tracking method (p. 314 of Merchant, Morrissette, & Porterfield, 1974), relatively little attention was paid to the problem until very recently, when it was highlighted in a series of papers by Wyatt (1995, 2010).

Recent studies have characterized the basic relationship between pupil size and gaze position measurements by explicitly evoking changes in pupil size by variations in light intensity (Drewes, Masson, & Montagnini, 2012; Ivanov & Blanche, 2011; Kimmel, Mammo, & Newsome, 2012; Wyatt, 2010), a reasonable strategy since the pupillary reflex is highly predictable with minimal variation across individual observers. But these studies capture only part of the problem arising from pupil size changes, for modulations in pupil size can also arise from endogenous factors, including arousal (Bradshaw, 1967; Henson & Emuh, 2010; Hess & Polt, 1960) and task-related cognitive demands (de Gee, Knapen, & Donner, 2014; Hess & Polt, 1964; Kahneman & Beatty, 1966; Nassar et al., 2012), that are bound to occur in studies using even simple tasks. Thus, it is important to learn the relationship

* Corresponding author.
E-mail addresses: kywch@snu.ac.kr (K.W. Choe), randolph.blake@vanderbilt.edu (R. Blake), visionsl@snu.ac.kr (S.-H. Lee).

http://dx.doi.org/10.1016/j.visres.2014.12.018
0042-6989/

between pupil size and gaze position measurements under situations where endogenous factors may be influencing pupil size dynamics and, hence, measurements of gaze control. This motivated the current study, which aims to characterize endogenously driven changes in pupil size and to examine the relationship of those changes with video-based gaze position measurements from a relatively large sample of observers while they performed two different tasks each with its own unique demands. Part of the results have been published previously elsewhere (Choe, Blake, & Lee, 2014).

2. Materials and methods

2.1. Observers

A total of 23 paid volunteers (11 females, 12 males; aged 18–36 years), who were recruited by online posting, all participated in both Experiment 1 (Section 2.3) and Experiment 2 (Section 2.8) in this study after giving informed consent, in accordance with the guidelines and approval of the Institutional Review Board at Seoul National University. None of the participants reported any history of reading problems or symptoms of abnormal vision. All participants were naïve to the purpose of the study.

2.2. Apparatus and eye tracking setup

Stimuli were presented in a dimly lit room on a gamma-linearized 22-inch CRT monitor (Totoku CV921X CRT monitor) operating at vertical refresh rate of 180 Hz and a spatial resolution of 800 × 600 pixels. Stimuli were generated using MATLAB (MathWorks) in conjunction with MGL (http://justingardner.net/mgl) on a Macintosh computer. Observers viewed the monitor at a distance of 90 cm while their binocular eye positions were sampled at 500 Hz by an infrared eye tracker (EyeLink 1000 Desktop Mount, SR Research; instrument noise, 0.01° RMS; Fig. 1A). The LED illuminator and camera (broken-line boxes in Fig. 1A) were positioned side by side, at a distance of 65 cm from the observer (broken line with arrow ends in Fig. 1A), and angled toward the observer's face to insure that infrared light illuminated both eyes and was being reflected from both eyes and imaged on the camera sensor.

An observer sat on a height-adjustable chair with his/her head supported by a forehead and chin rest (HeadSpot, UHCOTech), which were, together with the monitor, mounted on a height-adjustable table (Fig. 1B). To minimize body and head movements that compromise the quality of eye tracking measurements, the following procedure was applied. First, an observer was given enough time to find a comfortable arrangement of the chair, table, forehead, and chin rest by adjusting the heights of those devices. Second, the lower part of the head was harnessed by wrapping a memory-foam cushion around the neck such that the cushion's ends were tightly secured to the head post and the sides of the chin. Third, the upper and middle part of the head was constrained by fastening a wide buckled cotton strap over the forehead, the head post, and the lower back of the head. To mitigate discomfort associated with tight head fixation, baby-proofing cushion tapes were attached on the contact surfaces of the chin-rest.

The eye tracker was calibrated using the built-in five point calibration routine (HV5), not only at the beginning of each daily session but also whenever the observer was disengaged from a previously calibrated head positioning setup. During a session, the observer was allowed to take as many breaks as desired, disengaging from the eye tracking setup and moisturizing the eyes using disposable artificial tears as needed. Eye tracking signals were acquired in a 'pupil-corneal reflection (P-CR)' mode, and the pupil center was estimated using the ellipsoid fitting method, which is known to be robust to pupil occlusion by the eyelids. To check the possibility that the relationship between pupil size and gaze position is dependent on pupil tracking methods, we also collected data using the centroid method, an alternative method of pupil center estimation. The results using the ellipsoid and centroid methods did not differ (see Appendix A for details).

2.3. Experiment 1: visually guided saccade task

Observers performed a visually guided saccade task (Tse, Baumgartner, & Greenlee, 2010) by fixating their gaze on a target that appeared at three different positions on the monitor. An experimental run consisted of two alternating blocks of eye tracking measurements. In 'prolonged-fixation (PR)' blocks, a central gray (30 cd/m^2) dot (0.12° in diameter) was presented as a fixation target (FT) for 16 s against dark (3 cd/m^2) background. In 'short-lived fixation (SL)' blocks, which lasted for 31 s, the position of the FT was updated at 1 Hz, appearing either in the left (−0.12°) or in the right (+0.12°) side of the center of the monitor. The position of the FT was determined by an m-sequence (31 trials with base of 2 and power of 5), making location order unpredictable over time with zero autocorrelation. Every run started and ended with a PR block, and contained 5 SL blocks, resulting in a total of 251 s (16 s × 6 PR blocks plus 31 s × 5 SL blocks) for one single run. Each observer performed a single run.

2.4. Preprocessing of eye tracking data

The EyeLink system estimates gaze position and pupil area using built-in proprietary software and provides those estimates to end users in a digitized format called 'EDF.' In this file format

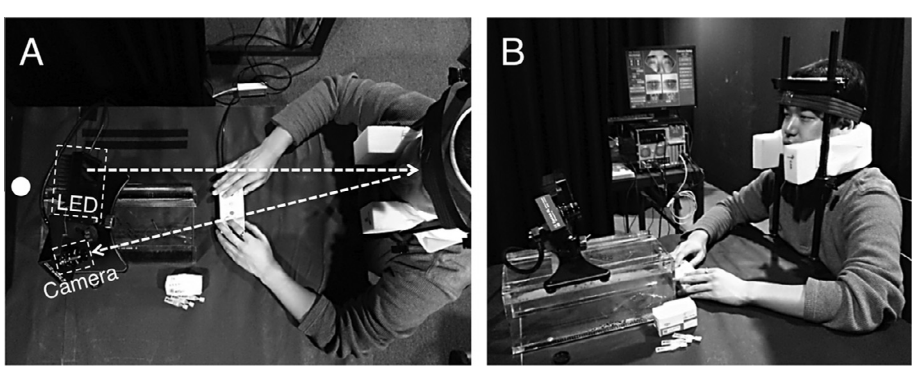

Fig. 1. Eye tracking experimental setup with an observer. (A) Locations of the LED illuminator and camera relative to the head and the display. The white dot demarcates the center of the screen. The dashed line arrows indicate the mean trajectory of the LED light projected onto and reflected from the eyes. (B) Setup for minimizing body and head motion. The setup included height-adjustable chair and table, a forehead and chin rest, a memory-foam cushion around the neck, and a buckled cotton strap over the forehead (see Section 2.2 for details).