Monitoring Response Quality During Campimetry Via Eye-Tracking

Gustavo Vergani Dambros
Institute of Informatics
UFRGS
gustavo.vergani-dambros@student.uni-tuebingen.de

Judith Ungewiss
Vision Research
Aalen University
judith.ungewiss@hs-aalen.de

Thomas C. Kübler
Computer Engineering
University of Tübingen
thomas.kuebler@uni-tuebingen.de

Enkelejda Kasneci
Computer Engineering
University of Tübingen
enkelejda.kasneci@uni-tuebingen.de

Martin Spüler
Computer Engineering
University of Tübingen
spueler@informatik.uni-tuebingen.de

IUI'17 Companion, March 13-16, 2017, Limassol, Cyprus
ACM 978-1-4503-4893-5/17/03.
http://dx.doi.org/10.1145/3030024.3038268

Abstract
In a variety of use-cases, deriving information on user's fatigue is an important step for content adaptation. In this work, we investigate which eye-tracking related measures can predict the error rate (as a proxy of subject's fatigue) during a visual experiment. Data was collected during a 40 minutes campimetric task, where the user has to detect visual stimuli (i.e., dots) of different contrast. We found that eye-tracking measures can be used to train a machine learning model to predict the error rate of a user with an average correlation of 0.72 ± 0.17. The results show that this method can be used to measure the user's response quality.

Author Keywords
Fatigue; vigilance; campimetry; eye-tracking; pupil diameter; blink rate

ACM Classification Keywords
J.3 [Computer Applications]: Life and medical sciences; I.2 [Computing Methodologies]: Artificial Intelligence

Introduction
Knowledge on user's fatigue is essential for perceptual and behavioral studies. For example, a medical examination of the visual field is exhausting for patients as it involves repeatedly detecting and responding to hardly visible stimuli. The gain in accuracy and reliability with increased

Campimetry
Campimetry is the examination of the visual field on a flat surface (e.g., computer screen), where light are presented at different locations of the visual field. If a stimulus is perceived, the subject provides feedback by pressing a button [7, 8]. Through this procedure, the minimum perceivable stimulus-to-background contrast is determined at each location. Increasing the number of stimuli presented allows either for a finer resolution of the location grid, a higher number of different contrast levels or more retests.

Fatigue Indicators
Pupil size is determined by two neurophysiological reflexes, the pupillary light reflex regulating the amount of light entering the pupil, and accommodation, a change in the curvature of the lens.

test duration is antagonized by increasing fatigue, effectively limiting the achievable examination quality. The rate of response errors (i.e., missed detections that were above the perceptual threshold) increases with the duration of the measurement. We chose campimetry as a demo application for eye-tracking based fatigue detection, as it provides an objective performance indicator via the response error rate. However, once reliable indicators of fatigue are found, they are likely applicable to a broad range of similar tasks. For example, the fatigue level of a surgeon performing in virtual reality or subjects repeatedly viewing image trials might be of interest.

We study whether and how accurate the error rate, as a proxy of subject's fatigue, can be predicted by machine learning based solely on eye-tracking parameters. Thereby, the medical examination, but potentially also many other HCI devices based upon gaze input, could be enhanced to consider the subject's fatigue level as a parameter for their interaction. For example, subject response could be registered with a reliability or examinations could be terminated early.

Indicators of vigilance
Parameters of vigilance in eye-tracking data are relatively well studied. During campimetry and perimetry researchers have focused on fatigue waves, i.e., pupillary oscillations [2], although a wide variety of other parameters is available and can theoretically be measured through the same device - a camera directed at the subject's eye.

Henson and Emuh showed that it is possible to monitor vigilance during campimetry under photopic conditions using pupillometry [2]. In darkness, pupil size oscillations come in two flavors: Slow waves of dilatation and constriction of 4 to 40 seconds duration and an amplitude of up to 0.5 mm. Superposed fast inextensive oscillations, of 0.5 to 1 second duration and with amplitudes of usually below 0.1 mm, but they reach up to 0.3mm [4]. With decreasing vigilance the feedback loop that regulates the pupil diameter becomes unstable resulting in much larger pupillary oscillations [10] (up to an amplitude of 1 mm [2]).

Methods
Nine subjects (4 female and 5 male, aged 20-32) participated in the experiment. During the campimetric task they were asked to press a button in response to a perceived light stimulus. A central fixation cross was presented in between. Stimuli were presented at twelve different contrast levels and appeared at three screen locations (0°, 4.67°), (-4.07°, -2.34°) and (4.07°, -2.34°) and at the center (but only for $\sim$3 % of trials to motivate looking at the fixation cross). Additionally, catch-trials were inserted, i.e., trials with a stimulus contrast so high that an attentive subject has to perceive it (positive catch-trial), or stimuli that were not visible at all (negative catch-trial). In the following, we refer to a *false negative* as a positive catch-trial that was not reacted to, and a *false positive* as a negative catch-trial (i.e., no stimulus presented), for which the participant pressed the response button.

Eye movements and pupil diameter were recorded by means of an SMI RED 250 at a frequency of 250 Hz. A chin rest stabilized the position of the head 60 cm from the screen. In total 1,488 stimuli were presented, 25 % of which were positive and 25 % negative catch-trials. Each stimulus was shown for 200 ms followed by a 1300 ms response window. To increase the likelihood of fatigue, the experiments were conducted in the afternoons.

Data processing and prediction
Blink rate, blink duration, average pupil diameter and vergence were z-score normalized to the two first minutes

Blink rates of healthy individuals range between 5 and 15× per minute [1]. Blink rate, duration, and lid closure speed reflect the level of fatigue and sleepiness. Light fatigue is associated with an increase in blink frequency, sleepiness with an increase in blink duration [5]. **Vergence angle** of both eyes is connected to vigilance in a not necessarily straight-forward way: Heterophoria has been found to grow with fatigue and when performing an unfamiliar task [3]. Some authors report that the vergence system is particularly fragile with fatigue [9].

Subject	Correlation	RMSE
NC	0.86	0.030
AB	0.83	0.026
PZ	0.82	0.010
LR	-0.17	0.059
JD	0.55	0.014
HE	-0.60	0.014
LA	0.38	0.018
MO	0.10	0.053
PA	0.36	0.019
Mean	0.35	0.028
Std	0.50	0.018

Table 1: Correlation and root-mean-square error between catch error rate and prediction for each subject with a complete or incomplete recording.

of data as a baseline. Linear interpolation was used to fill blinks and tracking losses for all non-blink related indicators. Vergence was filtered by a moving window filter selecting the 75%-percentile.

The error rate in the response to catch-trials was predicted by a leave-one-subject-out cross-validation, in which data of 8 subjects was used to train the prediction model and the model was tested on the remaining subject. This procedure was repeated for each subject. For training the prediction model, we use support vector regression with linear kernel. To evaluate the quality of the prediction we used Pearson's correlation coefficient (CC) and the root-mean-squared error (RMSE), as those are good performance indicators [6].

Results

As some of the subjects finished the experiment early due to abortion or technical difficulties, we analysed in two ways, using either all data, or only data from subjects with a complete recording. For the group with complete recordings an average correlation of 0.72 was obtained, while the average correlation was only 0.35 when also including the incomplete data.

More detailed results for all subjects can be found in Table 1, where the correlation and root-mean-square error between the error rate and the prediction is shown. Table 2 shows the values when using only data from the subjects with complete recordings. Figure 2 shows some of the fatigue indicators derived from the eye-tracking signal. Table 3 shows the weight of a linear support vector regression trained on data from the 5 subjects with complete data.

Discussion

Individual indicators are spiky and depend highly on individual baselines. For example, blink behaviour is highly

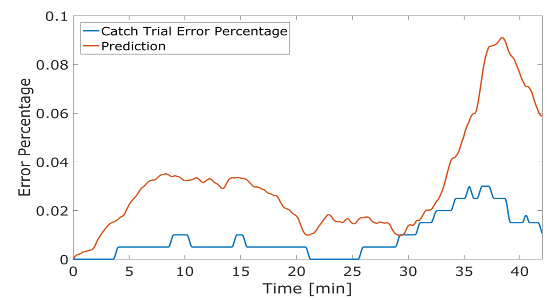

Figure 1: Error rate and prediction for subject AB

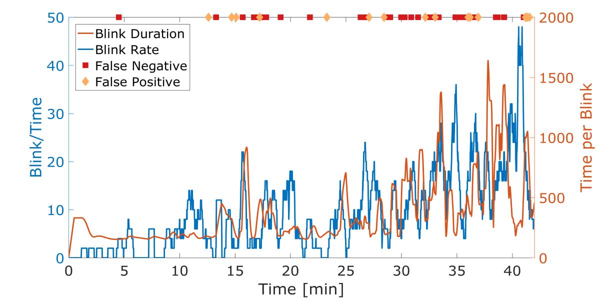

Figure 2: Blink rate and blink duration for subject AB

predictive for many subjects, however the experiment procedure implied a certain optimal blink time that may mask this effect for other subjects. By training a regression model we were to able to combine multiple indicators to obtain a more stable error-rate prediction across subjects. An analysis of the coefficients obtained by the regression model revealed that blink duration, vergence and fatigue waves are the most important features for prediction. This work shows that features obtained by eye-tracking can be used for a reliable prediction of the error-rate and thereby can be used to measure response quality in a task (i.e.,

Subject	Correlation	RMSE
NC	0.85	0.033
AB	0.78	0.027
PZ	0.79	0.016
LR	0.42	0.056
JD	0.78	0.018
Mean	0.72	0.030
Std	0.17	0.016

Table 2: Correlation between catch error rate and prediction, and root-mean-square error for each subject with a complete recording.

Feature	Coefficient
BD	0.1000
BR	-0.0495
APD	-0.0239
PV	-0.0409
Vrg	0.1267
FW	0.0629

Table 3: Coefficients for each feature obtained by linear regression trained with the data from the 5 subjects with complete recording. To make coefficients comparable, all features where z-score normalized. Abbreviation of the features: blink duration (BD), blink rate (BR), average pupil diameter (APD), Pupil variability (PV), Vergence (Vrg), fatigue waves (FW)

campimetric exermination). Nonetheless, it is challenging to identify which fatigue has a bigger influences in the increase of the error rate, either the mental or eye fatigue.

In future work, we will apply our fatigue prediction method to different tasks where no easily measurable performance indicator is available. There are certain preconditions to be met when generalizing to other experimental setups that might restrict the availability of individual indicators. For example, the vergence angle requires binocular tracking and fatigue waves have only been found in scotopic and mesopic conditions [2]. Generally, pupil diameter related measures require extensive normalization if the brightness varies during the experiment. We believe that a reliable and robust fatigue prediction via eye-tracking will enable intelligent devices to adapt to the cognitive state and current abilities of the user and thereby to interact in a more efficient and productive way.

REFERENCES

1. Giuseppe Barbato, Gianluca Ficca, Giovanni Muscettola, Mariateresa Fichele, Michele Beatrice, and Franco Rinaldi. 2000. Diurnal variation in spontaneous eye-blink rate. *Psychiatry Research* 93, 2 (2000), 145–151.
2. David B Henson and Thomas Emuh. 2010. Monitoring vigilance during perimetry by using pupillography. *Investigative Ophthalmology & Visual Science* 51, 7 (2010), 3540–3543.
3. Herbert Kaufmann and Heimo Steffen. 2012. *Strabismus* (4th ed.). Georg Thieme Verlag, Stuttgart.
4. Otto Lowenstein, Richard Feinberg, and Irene E Loewenfeld. 1963. Pupillary movements during acute and chronic fatigue a new test for the objective evaluation of tiredness. *Investigative Ophthalmology & Visual Science* 2, 2 (1963), 138–157.
5. Robert Schleicher, Niels Galley, Susanne Briest, and Lars Galley. 2008. Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? *Ergonomics* 51, 7 (2008), 982–1010.
6. Martin Spüler, Andrea Sarasola-Sanz, Niels Birbaumer, Wolfgang Rosenstiel, and Ander Ramos-Murguialday. 2015. Comparing metrics to evaluate performance of regression methods for decoding of neural signals. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 1083–1086.
7. Enkelejda Tafaj, Thomas C Kübler, Jörg Peter, Wolfgang Rosenstiel, and Martin Bogdan. 2011. Vishnoo – An open-source software for vision research. In *24th International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 1–6.
8. Enkelejda Tafaj, Christian Uebber, Janko Dietzsch, Ulrich Schiefer, Martin Bogdan, Martin Bogdan, and Wolfgang Rosenstiel. 2010. Mobile and fast detection of visual field defects for elderly drivers as a necessary input into driver assistance systems for mobility maintenance. *GMM-Fachbericht-AmE 2010-Automotive meets Electronics* (2010).
9. Marine Vernet and Zoï Kapoula. 2009. Binocular motor coordination during saccades and fixations while reading: a magnitude and time analysis. *Journal of Vision* 9, 7 (2009), 2–2.
10. Barbara Wilhelm, Henner Giedke, Holger Lüdtke, Evelyn Bittner, Anna Hofmann, and Helmut Wilhelm. 2001. Daytime variations in central nervous system activation measured by a pupillographic sleepiness test. *Journal of Sleep Research* 10, 1 (2001), 1–7.