App-Based Saccade Latency and Directional Error Determination Across the Adult Age Spectrum

Hsin-Yu Lai, *Member, IEEE*, Gladynel Saavedra-Peña, Charles G. Sodini, *Fellow, IEEE*, Thomas Heldt, *Senior Member, IEEE*, and Vivienne Sze, *Senior Member, IEEE*

Abstract—*Objective:* We aid in neurocognitive monitoring outside the hospital environment by enabling app-based measurements of visual reaction time (saccade latency) and directional error rate in a cohort of subjects spanning the adult age spectrum. *Methods:* We developed an iOS app to record subjects with the frontal camera during pro- and anti-saccade tasks. We further developed automated algorithms for measuring saccade latency and directional error rate that take into account the possibility that it might not always be possible to determine the eye movement from app-based recordings. *Results:* To measure saccade latency on a tablet, we ensured that the absolute timing error between on-screen task presentation and the camera recording is within 5 ms. We collected over 235,000 eye movements in 80 subjects ranging in age from 20 to 92 years, with 96% of recorded eye movements either declared good or directional errors. Our error detection code achieved a sensitivity of 0.97 and a specificity of 0.97. Confirming prior reports, we observed a positive correlation between saccade latency and age while the relationship between directional error rate and age was not significant. Finally, we observed significant intra- and inter-subject variations in saccade latency and directional error rate distributions, which highlights the importance of individualized tracking of these visual digital biomarkers. *Conclusion and Significance:* Our system and algorithms allow ubiquitous tracking of saccade latency and directional error rate, which opens up the possibility of quantifying patient state on a finer timescale in a broader population than previously possible.

Hsin-Yu Lai and Gladynel Saavedra-Peña are with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, USA.

Charles G. Sodini is with the Department of Electrical Engineering and Computer Science, Institute for Medical Engineering and Science, and the Microsystems Technology Laboratory, Massachusetts Institute of Technology, USA.

Thomas Heldt is with the Department of Electrical Engineering and Computer Science, Institute for Medical Engineering and Science, and the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: thomas@mit.edu).

Vivienne Sze is with the Department of Electrical Engineering and Computer Science and the Research Laboratory of Electronics, Massachusetts Institute of Technology, USA.

Digital Object Identifier 10.1109/TBME.2021.3112007

I. INTRODUCTION

IT REMAINS challenging to track neurodegenerative disease progression objectively, accurately, and frequently. Current assessments of neurodegenerative diseases are subjective and sparse, and standard neurocognitive and neuropsychological test batteries require a trained specialist to administer and score [1], [2]. Additionally, these tests demand significant patient time and cooperation, and can therefore be influenced by a patient's level of attention and comfort with the clinical setting [3]. Quantitative, objective, and frequent assessments may mitigate the effects of individual physician's clinical acumen and patient fatigue when determining neurodegenerative disease progression.

Assessment of eye movement is a promising candidate for such a quantitative and objective test. First, eye movements are readily observable. Second, their neural pathways involve several brain regions, and they might hence be affected by degenerative processes affecting various brain centers [4]. For example, Huntington's disease and progressive supranuclear palsy directly affect oculormotor pathways. As a result, clinical eye movement assessments are key to diagnosing and tracking these diseases.

In the context of neurodegenerative disease assessment and progression monitoring, pro- and anti-saccade visual reaction tasks are often used challenge tests [5], [6]. In the pro-/anti-saccade tests, a subject is asked to look towards/away from a visual stimulus. An anti-saccade task, in particular, requires a person to inhibit a natural reflexive eye movement towards the stimulus and initiate an eye movement in the opposite direction of the stimulus. Thus, it requires more cognitive processing than a pro-saccade task [7], [8]. Because these tasks demand cognitive abilities which can be affected by neurodegenerative diseases, two saccadic eye movement features were observed to be significantly different between healthy subjects and patients: saccade latency (visual reaction time) and directional error rate (the proportion of eye movements towards the wrong direction) [9]–[12]. However, these features are commonly measured with dedicated infrared cameras and chinrests, which limits the measurements to the doctor's office or the neurophysiological

Fig. 1. Age distribution of subjects with single or multiple recording sessions.
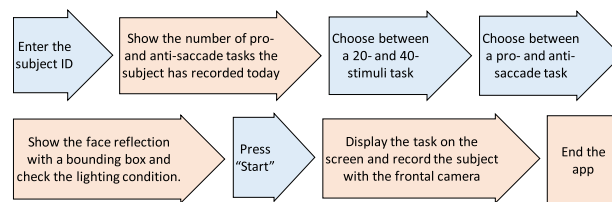

Fig. 2. The flow of the app. Blue arrows request the input from the subject. Orange arrows denotes the response of the app.
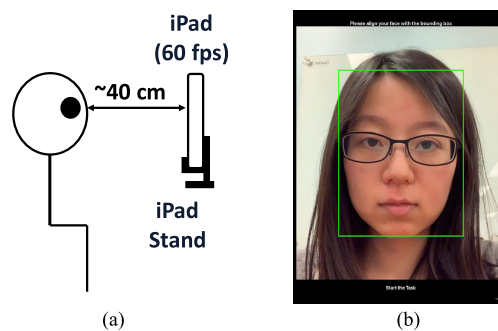

Fig. 3. (a) Recording setup; (b) before showing the task on the screen, the app displays the face of the subject with a bounding box. If the distance measurement from the camera to the subject's face is accessible (i.e. between 30 and 50 cm), the box will turn green. If the automatically detected ISO is greater than 1000, a warning will be shown to guide the subject to move to a better-illuminated place.

laboratory. In our previous work [13]–[15], we showed that we can accurately and robustly determine saccade latency from recordings obtained with a smartphone camera.

In this work, we developed an app to display the pro-/anti-saccade tasks on a tablet computer while recording a subject's eye movements with the built-in camera. We present an automated processing pipeline to determine pro-/anti-saccade latency and directional error rate, thus enabling ubiquitous recording of these neurological digital biomarkers. With this novel recording platform and pipeline we collected over 6,800 videos and over 235,000 individual eye movements from 80 subjects across the adult age spectrum.

II. MATERIALS

A. Recruitment

To study the responses of subjects of different ages to pro- and anti-saccade tests, we recruited 80 self-reported healthy adult subjects, ranging in age from 20 to 92 years.

Video recording of volunteers was approved by MIT's Committee on the Use of Humans as Experimental Subjects (protocol # 1711147147), and informed consent was obtained from each participant before recording. A subject could choose to participate either once or for multiple recording sessions. Each recording session consists of three pro-saccade trials and three anti-saccade trials, where each trial is saved as a separate video [15]. For subjects who chose to participate in a single session, each trial consists of a set of 40 visual stimuli. Subjects who chose to participate in multiple recording sessions were still asked to provide three pro- and three anti-saccade trials per session though were given a choice of either 20 or 40 stimuli per trial. The number of single and multiple recording sessions, grouped by decades of age is shown in Fig. 1.

B. App Design

In our previous work [15], we displayed the visual reaction task on a laptop and recorded the subjects with an iPhone. Synchronization of the recording and task display was achieved through a second screen that mirrored the laptop screen and was recorded alongside the subject's response [15]. Given the elaborate set-up, the recording was limited to our laboratory setting. Our goal here was to allow for ubiquitous recording and hence for subjects to record themselves in the comfort of their homes or offices. We therefore developed an iOS app so subjects could record themselves with the frontal (i.e. selfie) camera as the tasks were displayed on the screen. While the app can run on iPhones, our platform of choice was the iPad (Generation 2 and 3) for their larger dimensions and hence larger angular gaze amplitudes ($\sim$12.7 degrees at a distance of 40 cm to the camera).

The flow of the app is shown in Fig. 2. The app first obtains the subject's ID and then reminds the subject of the number of pro- and anti-saccade tasks they have performed the same day. Subsequently, subjects are prompted to select the number of stimuli they wish to perform (20 or 40) and whether they would like to perform a pro- or anti-saccade task.

To minimize the influence of environmental conditions on the quality of the recordings, we initially asked subjects to position the iPad at a distance of 30 to 50 cm. In subsequent releases for iPads with depth-sensing capability, the app senses the distance from the subject and provides visual feedback so the subject can position the iPad within the desired target distance (Fig. 3). Besides distance, the app also guides the subject to position themselves in proper lighting conditions, and the subject will be asked to move to a brighter location if the automatically detected ISO is greater than 1000.

When the subject is ready to perform the task, they start the recording, and a count-down will be displayed so the subject can begin to focus their attention on the pro-/anti-saccade task. The task will then be displayed at 60 frames per second while the frontal camera simultaneously records the subject's face at
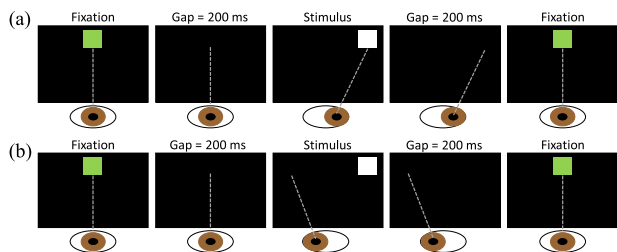
Fig. 4. (a) Pro-saccade task: Look toward the stimulus. (b) Anti-saccade task: Look away from the stimulus.
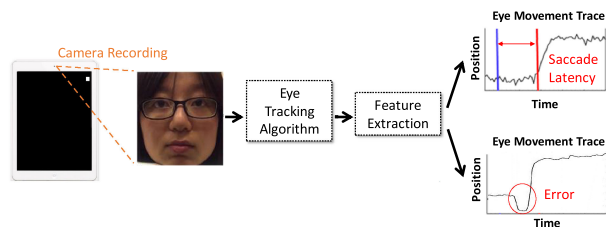

Fig. 5. The measurement pipeline includes the tablet-based video recording, an eye tracking algorithm, a saccade-latency measurement algorithm, and an error detection algorithm.

60 frames per second. After the task is completed, a detailed set of data files is saved for each recording, including the actual video recording, the timestamps of each recorded frame, the timestamps of each frame displayed on the screen, and a text file containing information about the recording system (iOS version, iPad generation), the distance of the iPad to the subject (when available), and the recorded ISO value at the beginning of the recording.

To acquire accurate saccade latency measurements, it is crucial to synchronize the task display on the iPad screen and the recording from the iPad camera. We detailed and evaluated the synchronization in Appendix. By requiring the ISO to be less than 1000, we showed that we can bound the absolute synchronization error to be within 5 ms.

C. Task Design

In this work, we implemented two commonly studied tasks in the literature, namely a gap-pro-saccade and a gap-anti-saccade task [10], [12], [16]. Both tasks start with a fixation period. During the fixation period (1 s), a fixation point (green square) is shown at the top center of the screen (as shown in Fig. 4). We chose the fixation point to be presented at the top to prevent occlusion from the eyelids. Subjects were instructed to look at the fixation point during this period. The fixation is followed by a 200-ms gap period, where the fixation point disappears and the screen stays black. After the gap period, a stimulus (white square) is presented on either left or right side of the screen. If a subject is performing a pro-/anti-saccade task, the subject is instructed to move their eyes towards/away from the stimulus as quickly and accurately as possible. This stimulus period will last for 1.2 s and be followed by another 200-ms gap period. This sequence of "fixation-gap-stimulus-gap" will repeat for 20 or 40 times, with half of the stimuli presented to the right of the fixation point and half to the left in randomized order.

III. METHODS

A. Measurement Pipeline Requirements

In this section, we discuss our measurement pipeline as shown in Fig. 5. Building on our prior work [13]–[15], we used iTracker-face to estimate the gaze of the subject (Fig. 6). The inputs to iTracker-face include the cropped face, as determined by the Viola-Jones algorithm [17], and a face grid indicating the face position. The outputs of iTracker-face are the (x, y)-coordinates of the estimated gaze position on the screen in the
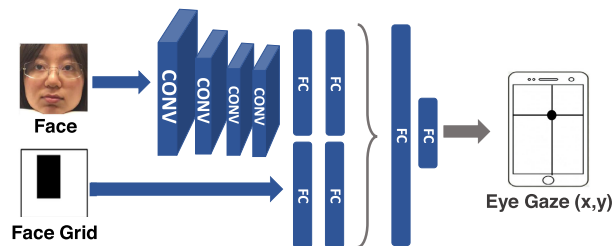
Fig. 6. Convolutional neural network architecture used by iTracker-face. "CONV" stands for convolutional layers and "FC" stands for fully connected layers. The details of the architecture can be found in [13].

unit of centimeters. To attain our horizontal eye movement trace, we retain the x-coordinate of the gaze position across frames. As discussed in Section II, we have synchronized the camera recording with the task display. We use the timestamps of the screen frames to acquire the time when each stimulus appears. With the stimulus presentation time and the eye movement trace, we can determine saccade latency and detect eye-movement errors.

In [15], we measured saccade latency by fitting a hyperbolic tangent ($\tanh$) to a fixed window of the eye movement trace, from 100 ms before to 500 ms after the stimulus presentation, and determined the saccade onset as the time when the best model fit exceeded 3% of the maximal saccade amplitude. Saccade latency was then computed as the time difference between the saccade onset and the time when the stimulus presented. A major benefit of using this model-based approach is that it provides an automated signal-quality quantification by means of the normalized root-mean-square error (NRMSE) between the model fit and the eye-position trace. We marked a trace as unusable if its NRMSE was greater than 0.1 [15]. In addition, since the output of iTracker-face is in the unit of centimeters, we normalized the trace to the expected saccade amplitude using the best fit model. Since our saccade onset determination is scale invariant, the measured saccade latency is invariant to this normalization.

In the current study, we expanded upon our initial study cohort in [15] by specifically including self-reported healthy subjects across the adult age spectrum. Consequently, we observed a larger heterogeneity in saccadic eye-movement patterns that necessitated revisions to our previously established processing pipeline.

To allow for latency measurements from subjects with slower response times, we needed to increase the window of fit for
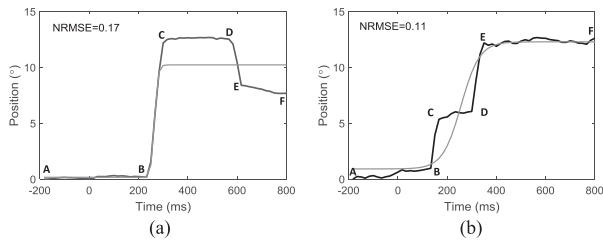
Fig. 7. Examples where `tanh` cannot be fitted to the entire trace: (a) gaze returning (b) hypometric saccade [4], [6]. As one can see, to find the saccade latency, the window where we fit a `tanh` model should be from A to D.

the `tanh` model from 200 ms before to 800 ms after the stimulus presentation. However, we noticed that by expanding the window, it is more likely to capture a subject's eye movements back toward the center position (Fig. 7 a). Additionally, subjects may perform a series of hypometric saccades in which the initial saccadic movement does not reach the final position and a second saccade is made to correct for this undershoot (Fig. 7 b). Correct identification of hypometric saccades is of relevance since an increased incidence of hypometric saccades is associated with certain neurodegenerative pathologies [6], [11]. The single `tanh` model cannot fit well to these traces if we use a fixed window to determine latency values. To determine saccade latency, we need to allow for an adaptive window of fit for the `tanh` model to identify the initial saccadic movement to be fitted. We also note that we cannot convert the unit of the eye-movement trace from centimeter to degree using the best fit `tanh` model. As a result, we needed to revise our method to normalize the trace.

B. Saccade Normalization

We make three assumptions to convert the unit of the eye-movement traces from the iTracker-face generated centimeter to degree. First, we assume that subjects were looking at the fixation point during the fixation period. Second, we assume that subjects did not overshoot their gaze. Finally, we assume that during the stimulus period, subjects either (a) did not move their eyes at all, (b) gazed at the stimulus, or (c) gazed at the opposite position of the stimulus.

With these assumptions, we normalize the trace as follows. First, to simplify the algorithms, we flip the trace if needed so that positive excursions correspond to eye movements in the correct direction. We then smooth the eye-movement trace with a Savitzky-Golay filter [18], [19] (of order 3 and frame length 5) to make the final normalization more robust to noise. Subsequently, we determine two reference points to scale and shift the eye-movement trace. Our first reference point is set as the starting gaze position of a trace, that is 200 ms before the stimulus presentation. With the second assumption, our second reference point is either the maximum or the minimum value of the smoothed trace, depending on whether the subject makes a correct saccade, a corrected error, or an uncorrected error. Scaling and shifting coefficients can be found by shifting the first reference point to zero degree and scaling the second to either the final expected amplitude (12.7 degrees) or the negative amplitude (−12.7 degrees). We emphasize that the amplitude

conversion is based on the expected distance of the subject to the iPad and is hence not an accurate estimate of the actual gaze amplitude. While our determination of directional error and saccade onset are independent of saccade amplitude other pathognomonic eye movement features that depend on accurate saccade amplitude information are currently not resolved.

More precisely, we consider three scenarios. Operating on the output of iTracker-face, if the difference between the maximum value and the starting gaze position is greater than 0.2 cm, we assume that the subjects have made a correct saccade or a corrected error, and we scale the second reference point to the positive expected amplitude value. If the difference between the maximum value and the starting gaze position is smaller than 0.2 cm but the absolute difference between the minimum value and the starting gaze position is greater than 0.2 cm, we assume that the subjects have made an uncorrected error and we scale the second reference point to the negative expected amplitude value. If neither of these criteria is met, we assume that the subjects have made only subtle eye movements or that the eyes were occluded.

In the first two scenarios, we find the scaling and shifting coefficients from the smoothed trace and normalize the original trace using these coefficients. One key observation of this normalization is that after normalization, traces with the same shape will become identical. This characteristic ensures that if the saccade-latency measurement algorithm and the error-detection algorithm are designed using this normalized trace, the algorithms will be scale-and-shift-invariant. That is, eye movement features are measured only based on the shape of a trace. In the third scenario, we noticed on visual inspection of the video recordings that the sizes of the eye movement were often comparable with noise and subtle head movement. To account for such observations, we label such traces as "dLS" (declared Low Signal) to acknowledge the fact that we are uncertain whether there is an actual eye movement even by visualizing the original videos. Traces labeled dLS will be excluded from the saccade-latency measurement and the error-detection algorithm.

C. Adaptive Windowing and Saccade Latency Measurement

Returning to the examples in Fig. 7, during the period from A to B and C to D, the subject's eyes are fixated. During the period from B to C, the subject performed a correct saccade, in the sense that the eyes moved in the correct direction. As a result, the proper window of fit is the first sequence of fixation, directionally correct eye movement, and fixation. This period can be identified using the velocity of the gaze. We estimate gaze velocity by computing the first-order derivative of the Savitzky-Golay filtered trace to avoid amplifying high-frequency noise.

We then classify a sequence of time instances as a correct saccade period if the velocity values cross 30 degrees/s, as an incorrect saccade period if the velocity values cross −30 degrees/s, and as a fixation period otherwise. When there is more than one correct saccade period, we will fit our model to the one that first crosses a third of the amplitude. Fig. 8 shows that by choosing the window of fit to be the period associated with the sequence of fixation, directionally correct eye movement, and fixation, we can
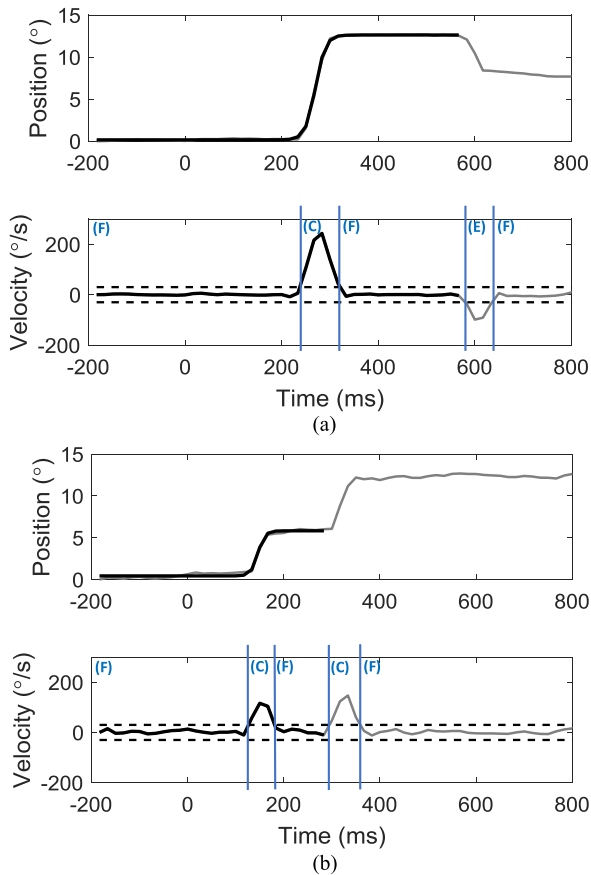
Fig. 8. Tanh fitting example: (a) gaze returning (b) hypometric saccade. The top panels show the eye movement traces obtained from iTracker-face after normalization. The dark lines show the fitted hyperbolic tangent models. The bottom panels show the velocity of the eye movements and the velocity threshold (the dash lines). With such a threshold, we label different parts of the trace as fixation (F), correct saccade (C), or error saccade (E). The window of fit is chosen as the first "fixation(F)-correct saccade(C)-fixation(F)" period that crosses a third of the amplitude.

fit the $\tanh$ model to traces with multiple transitions and measure their saccade latencies. We compared the previously described fixed-window approach with the adaptive-window approach and observed that the proportion of saccades with a NRMSE > 0.1 dropped from 17% to 3% with the adaptive-window approach. Hence, by moving to the adaptive-window approach, we were able to compute significantly more latencies with this improved saccade-latency measurement algorithm.

D. Directional Error Detection

In the clinical literature, a directional error is defined as an initial eye movement in the wrong direction [16]. Manual annotation is often involved in the determination of these directional errors [10], [11]. Because such clinical studies have traditionally relied on specialized environments and eye-tracking equipment, including use of chinrests, infrared illumination, and research-grade cameras, there were usually comparatively few traces collected per subject and the traces tended to be clean. As a result, manual annotation of traces is possible in these cases. In contrast, to enable collection of large amounts of data, we
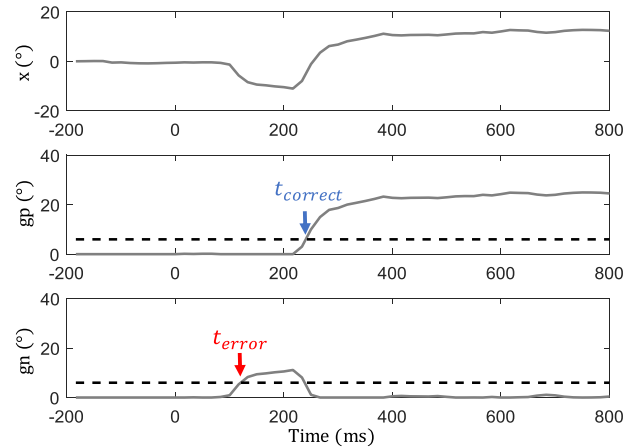

Fig. 9. Error detection example. The top panel shows the x coordinates of the iTracker-face output over time (x_t). The middle and the bottom panel show gp_t and gn_t. The dashed line indicates the threshold T. When gp_t and gn_t cross the threshold T, $t_{correct}$ and t_{error} are detected, respectively. In this case, since $0 < t_{error} < t_{correct}$, an error is detected.

use consumer-grade cameras and do not use a chinrest. As a result, we obtained significantly many more traces, though some were affected by glares or head movements. Our goal is thus to reject poor recordings and develop an accurate and robust error detection algorithm.

As mentioned in Section III-C, we exclude the traces labeled dLS, since we cannot distinguish between saccadic eye movements and noise/head movements. Out of the remaining traces, we noticed that a typical error trace shows a period of fixation followed by a directionally incorrect eye movement (top panel of Fig. 9). Since our goal is to detect such a change, we developed our algorithm based on the change detection literature [20]. In particular, we extended the cumulative sum (CUSUM) algorithm [21] for our purposes.

We first assume that our measured eye movement trace x_t at time t is composed of an eye movement θ_t and an additive measurement noise ϵ_t. We then use a recursive least square filter to estimate the eye movement $\hat{\theta}_t$ according to

$$\hat{\theta}_t = \lambda \hat{\theta}_{t-1} + (1-\lambda)x_t. \tag{1}$$

The residual error then becomes $\hat{\epsilon}_t = x_t - \hat{\theta}_t$. If there is neither a positive trend nor a negative trend in x_t, $\hat{\epsilon}$ will be centered around zero. As a result, when we consider the cumulative sum of the residual error $s_t = s_{t-1} + \hat{\epsilon}_t$, s_t will be centered around zero as well. However, if there is a negative trend in x_t (Fig. 9), s_t will become progressively more negative. We can then use a threshold to determine whether s_t is sufficiently negative for ϵ_t to be unlikely to just represent additive measurement noise.

To distinguish between correct saccades and incorrect saccades, we define two separate variables for s_t: $gn_t = \max\{gn_{t-1} - \hat{\epsilon}_t, 0\}$ and $gp_t = \max\{gp_{t-1} + \hat{\epsilon}_t, 0\}$. That is, gn_t accumulates negative trends and gp_t accumulates positive trends. As a result, when gn_t and gp_t cross the pre-determined threshold, we detect an incorrect and a correct saccade, respectively. To apply the definition of a directional error as an initial eye movement towards the wrong direction, we detect
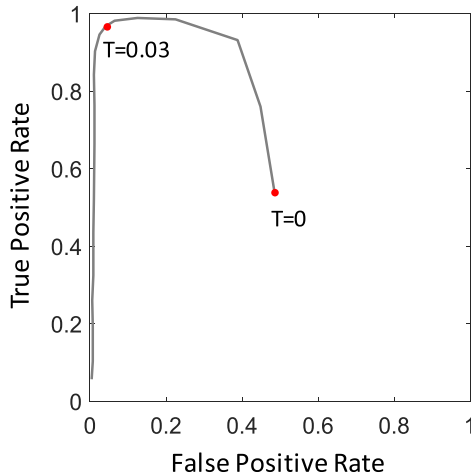
Fig. 10. The true positive rate and the false positive rate as we increased the error detection threshold T from 0 to 0.1. We chose $T = 0.03$ as our final threshold to achieve a sensitivity of 0.97 and a specificity of 0.97.

a directional error if gn_t crosses the pre-determined threshold after 0 ms and before gp_t crosses its threshold.

Here, we chose to scale the threshold with respect to the estimated (corrected) saccade amplitude. We notice that if there is no error in a trace, gn_t will be around zero while gp_t will approximate the amplitude of the saccade (Fig. 9). When there is an error, gp_t will approximate the amplitude of the corrected saccade. On the other hand, when there is an uncorrected directional error, gn_t will approximate the amplitude of the saccade. As a result, we approximate the (corrected) saccade amplitude by $max_t\{gp_t, gn_t\}$. We further observe that if the saccade amplitude before the normalization is sufficiently large, the saccade will be less affected by head movement and noise. Thus, we can consider lowering the threshold to detect smaller directional errors. On the other hand, if the original saccade amplitude is closer to the size of the head movement and noise, the threshold needs to be sufficiently large to avoid artifacts from being detected. Recall that in Section III-C, we scale the trace and shift it to normalize it from centimeters to degrees. We can use the scaling coefficient (denoted as B in **Algorithm**) as a metric to evaluate the size of the original saccade amplitude. If B is small (< 8), it means that the original amplitude is large and the threshold could be smaller. If B is large (≥ 8), we will use a fixed threshold. Here the value 8 can be considered as a hyperparameter that we can tune. The final threshold is $max_t\{gp_t, gn_t\} \cdot \min\{B, 8\} \cdot T$. To determine the threshold T, we asked four subjects to perform six anti-saccade tasks of 40 stimuli each. Two expert annotators reviewed the videos and annotated the directional errors. Out of the $4 \cdot 6 \cdot 40 = 960$ saccadic eye movements, there were only two disagreements between the annotators which were resolved after these two disagreements were reviewed together. With the annotated data set at hand, we swept the threshold T and determined the true positive and false positive rates for detecting a directional error (Fig. 10). When the threshold is lower than the noise level, gp_t and gn_t may cross the threshold due to noise rather than a saccadic eye movement. That is, gp_t may be equally likely to

---

Algorithm: Directional Error Detection.

input : $x = [x_1, \ldots, x_N], B, x_1$ is chosen to be the first instance after the stimulus presentation, B is the scaling coefficient in the saccade normalization

output : $t_{error}, t_{correct}$ (An error is only detected if the first element in t_{error} is smaller than the first element in $t_{correct}$.)

parameter: λ, T

for *round=0:1* **do**

    $\hat{\theta} = x_1, t_{error} = [], gn = [0], gp = [0];$

    for *t=2:N* **do**

        $\hat{\theta} = \lambda\hat{\theta} + (1 - \lambda)x[t];$

        $\hat{\epsilon} = x[t] - \hat{\theta};$

        $gn.append(\max\{gn[t - 1] - \hat{\epsilon}, 0\});$

        $gp.append(\max\{gn[t - 1] + \hat{\epsilon}, 0\});$

        if *round==1* **then**

            if $gn[t] > A \cdot T$ **then**

                $t_{error}.append(t);$

                $gn[t] = 0;$

                $\hat{\theta} = x[t];$

            end

            if $gp[t] > A \cdot T$ **then**

                $t_{correct}.append(t);$

                $gp[t] = 0;$

                $\hat{\theta} = x[t];$

            end

        end

    end

    $A = \min\{8, B\} \cdot \max\{gp, gn\};$

end

---

cross the threshold as gn_t. Recall that we only detect a trace as a directional error if gn_t crosses the threshold before gp_t. As T goes to zero, the true positive rate and the false positive rate go to 0.5. On the other hand, if the threshold is too large, the amplitude of an incorrect saccade may be smaller than the threshold and the error may not be detected. When T is larger than the noise level but smaller than the amplitude of an error, we can get high sensitivity and specificity. By choosing $T = 0.03$, we can achieve a sensitivity of 0.97 and a specificity of 0.97 for detecting a directional error.

E. Directional Error Rate Estimation

In the clinical literature, directional error rate is often reported as the proportion of directional errors, though it is not usually discussed whether noisy traces are excluded from such calculation. Given the use of special-purpose equipment and optimized environmental conditions in clinical research studies, such recordings may have very few noisy traces. Without a chinrest and a controlled laboratory setup, we obtained more noisy traces. We carefully identified the causes of these noisy traces: glares, head movements, eyelids drooping. Many of these causes could be reduced with more careful instruction. However, even with careful instruction, it is hard to eliminate all these causes, due to the nature of the much more relaxed and varying recording environment and the large number of recordings. As a
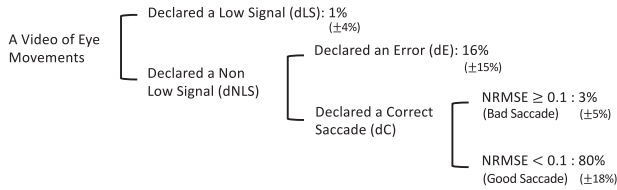
Fig. 11. Breakdown of saccades into declared low signal (dLS) and declared non low signal (dNLS). Breakdown of dNLS into declared error saccades (dE) and declared correct saccades (dC). Breakdown of dC into good saccades and bad saccades.

result, it is important to define a directional error rate that takes these noisy traces into consideration.

An eye movement was either declared a correct saccade (dC), declared a directional error (dE), or declared a low signal (dLS). If we estimate the directional error rate $P(E)$ as the proportion of directional errors out of all the traces $P(dE)$, we might significantly underestimate the directional error rate in recordings with a lot of eye movements in the dLS category. To estimate the directional error rate that accounts for such recordings, we estimate the directional error rate in a recording as #dE/(#traces-#dLS), as shown in Eq. (2). The question then arises under which conditions this estimation approximates the directional error rate.

Under the assumption that
- $P(dE|C, dNLS) \approx 0, P(dC|E, dNLS) \approx 0$,
- $P(dLS|E) \approx P(dLS|C)$,

where E denotes directional errors, C denotes correct saccades, $dNLS$ denotes the traces that are not declared as "Low Signal" (as shown in Fig. 11), we can estimate the directional error rate as

$$
\begin{aligned}
\frac{P(dE)}{1 - P(dLS)} &= P(dE|dNLS) \\
&= P(dE|E, dNLS)P(E|dNLS) \\
&\quad + P(dE|C, dNLS)P(C|dNLS) \\
&\approx P(E|dNLS) \\
&\approx P(E).
\end{aligned} \tag{2}
$$

The first assumption states that the false positive and the false negative are essentially zero. As discussed in Section III-D, our error detection algorithm achieved a sensitivity of 0.97 and a specificity of 0.97. Therefore, the first assumption is indeed met. The second assumption states that a correct saccade is equally likely to be dLS as an error saccade. Since our determination of dLS is simply based on the size of the trace and the amplitude of a saccade is generally much larger than the threshold for the determination of dLS, this condition is met as well. Therefore, it is reasonable to estimate the directional error rate as #dE/(#traces-#dLS). (Note that since our current algorithm achieves similar sensitivity and specificity and that a correct saccade is more likely to occur than a directional error saccade, this estimation tends to overestimate the directional error rate. However, by keeping both sensitivity and specificity close to one, this overestimation is minimized.)

IV. DATA ANALYSIS

With our system, we have collected 6823 videos and 236900 eye movements from 80 subjects across the adult age spectrum. With the saccade latency and directional error determinations, we labeled the traces as in Fig. 11. We observe that in videos with a substantial number of dLSs, subjects' eyes were often partially occluded due to eyelid droop. Videos with a large number of bad saccades tend to contain more head movements. As a result, the number of dLSs and bad saccades indicates whether a subject recorded themselves properly. We therefore discard a video if more than half of the saccades are dLSs or bad saccades. After discarding the videos with too many dLSs and bad saccades, we retained 6787 videos and 235520 eye movements from 80 subjects. Out of the remaining videos, we calculated the mean (standard deviation) of the proportions of each label in a video. There are 1% (4%) of dLSs and 3% (5%) of bad saccades. That is, on average, 96% of the saccades are good saccades or declared directional errors.

With these data, we can analyze the responses of eye movement features in different age groups (Fig. 1). This is important because it gives us a baseline when we compare the results with data from patients. We calculate the mean saccade latency and directional error rate for each individual and then compute the mean and standard error of the individual mean saccade latencies and directional error rates per age group. As a result, the mean of an age group is not biased towards those subjects who provided more recordings. To evaluate the correlation between age and eye movement features, we compared our result with [22], [23], where data were collected from specialized equipment (DC electrooculography with a head rest) in a controlled environment. We notice that [22] defined an anticipatory saccade as any saccade (including errors) with latency < 90 ms. To evaluate how changing this threshold may affect the result, we show the data with and without this anticipation threshold (Fig. 12).

Several observations are worth noting. First, since anti-saccade tasks are more complex and require more cognitive processing [9]–[11], the mean anti-saccade latency and anti-saccade directional error rate in every age group is larger than the corresponding mean pro-saccade latency and pro-saccade directional error rate. Moreover, we see that the saccade latency is positively correlated with age, whereas the correlation between directional error rate and age is not significant. These observations are in agreement with the data by Muñoz *et al.* [22], though the actual saccade latency values in our study tend to be lower than those reported by Muñoz *et al.* Our results suggest that our measurement system and processing pipeline can identify similar trends as shown in the clinical literature. Another observation is that the definition of anticipatory saccades affects the measured pro-saccade latency and directional error rate. On one hand, this observation is reasonable, since pro-saccade tasks are much easier to perform and directional errors tend to be caused by anticipation. On the other hand, while there is no consistent definition of anticipatory saccades in the literature, our observation highlights that they should be carefully defined.

In addition, with the accessibility to sizable data, we can study individual distributions, instead of only reporting the population mean as in most clinical literature. We analyzed the mean
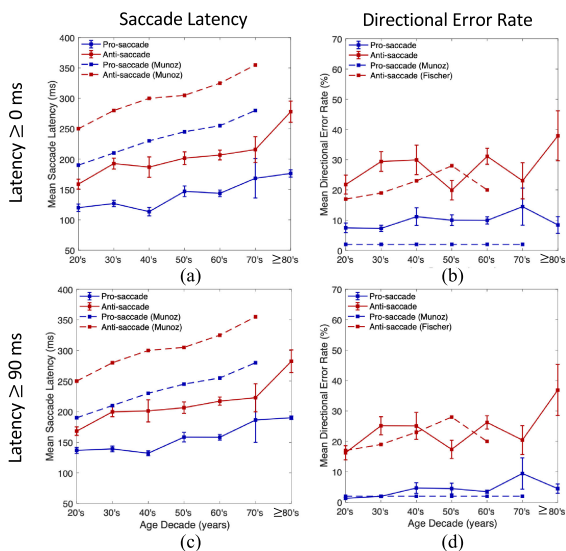
Fig. 12. Eye movement features as a function of age with saccades > 0 ms: (a) mean saccade latency (b) mean directional error rate, and with saccades > 90 ms: (c) mean saccade latency (d) mean directional error rate. The bars showed one standard error.

pro-saccade latency of each subject in seven age groups and chose from each age group the subject with the median mean pro-saccade latency as the representative subject. In Fig. 13, we showed example saccade latency distributions of these representative subjects. We observe that there are significant intra- and inter-subject variations in saccade latency across our study cohort, which suggests that aggregated results may lose the information encoded in individual distributions.

V. DISCUSSION

The neural circuits involved in generating eye movements can be affected by neurodegenerative diseases. In particular, pro-/anti-saccade latency and directional error rates have been shown in the clinical literature to be significantly different between healthy subjects and patients with certain neurodegenerative conditions such as Alzheimer's disease and Parkinson's disease [9]–[11]. Thus, such eye movement features may be promising candidates for tracking disease progression. However, these features are commonly measured in special, somewhat artificial environments and with special-purpose infrared-illuminated cameras, which limits broad accessibility and repeat measurements to track neurodegenerative disease progression longitudinally. In this work, we present, validate, and use an iOS application to enable such data collection. Additionally, we present algorithms for measuring saccade latency, determining directional errors, and calculating directional error rate that takes into account the possibility that it might not always be possible to determine the eye movement from home-based recordings.

Recording setup: In our previous work, we showed that instead of a special-purpose camera, we can measure saccade latency using a smartphone camera. The recording setup, nevertheless, required a laptop to display the task, a screen synchronized with the laptop to be placed behind the subject, and a researcher to record both the subject's eye movement and the

synchronized screen using the back camera of an iPhone. Due to these requirements, the recording setup was not sufficiently flexible for a subject to take recordings on their own in their homes or offices, which limits the possibility of using such a system to flexibly and ubiquitously monitor neurocognitive decline or disease progression. In this work, we designed an iOS app to record a subject with the frontal camera of an iPad while the subject is following a task shown on the screen. There are two challenges to achieve this goal.

First, unlike in the clinical setup and in our previous work where an expert researcher takes recordings of a subject, our app needs to guide the subject to record themselves at a proper distance to the camera and in a well-lit environment. To resolve this first challenge, before recording a subject, the app displays the subject on the screen and guides the subject to align their face with a bounding box shown on the screen. With such guidance, most subjects were recorded at an appropriate distance. To ensure the environment is well-lit, the app also asks the subject to move to a better-illuminated environment if the measured ISO is greater than 1000.

Second, the camera recording and the task displayed on the screen need to be well-synchronized to obtain accurate saccade latency. This can be challenging as most applications (e.g., video chatting) only require the synchronization error to be unnoticeable by a human (i.e., < 80 ms). With careful app design and evaluation of the synchronization error, we show that we can restrict the absolute timing error to be within 5 ms, which is well within the standard deviation of a subject's saccade latency distribution.

Algorithm design: In our previous work, we measured pro-saccade latency by combining a deep convolutional neural network for gaze estimation with a model-based approach for saccade onset determination that also provides automated signal-quality quantification and artifact rejection. Here, we also include an anti-saccade task and extend our measurement pipeline to measure the associated saccade latencies and directional error rates.

Since eye movements are now recorded outside of a clinical environment, our first observation is that in cases where eye movements are too small in amplitude or when the eyes are occluded, the eye movement signals can be smaller than noise. In these cases, we cannot tell the direction of the eye movement either from the trace or from the original video. As a result, we cannot classify these traces into a correct or a directional erroneous eye movement and cannot determine the saccade onset. We show that we can identify these traces using the raw output of iTracker-face, label these traces as "dLS"s (declared low signal), and exclude them from the saccade latency measurement and directional error detection. If a trace is not dLS, we then normalize the trace to the final expected amplitude (12.7 degrees or −12.7 degrees). With this normalization, the saccade-latency measurement algorithm and the error-detection algorithm are scale-and-shift invariant. That is, we do not need the amplitude measurement to be exact (which is affected by the distance between the subject and the camera). However, the distance between the subject and the camera still needs to be in an appropriate range to ensure that the recorded eye movments are sufficiently large and that the number of dLSs can be minimized.
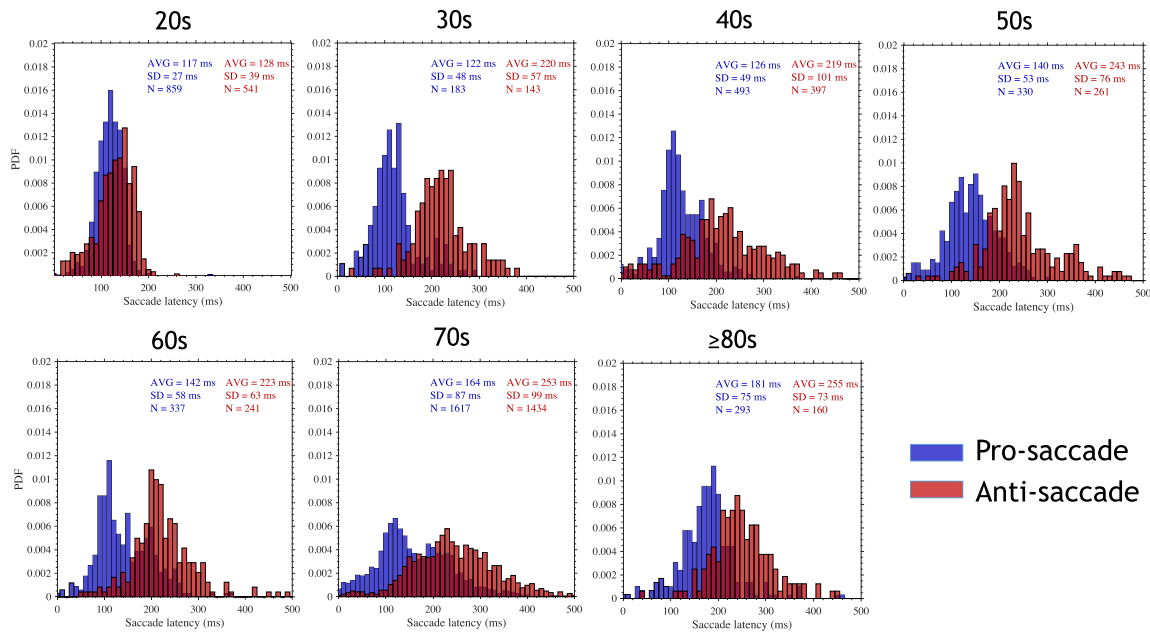
Fig. 13. Representative normalized distributions, shown as probability density functions (PDFs), of pro-saccade (blue) and anti-saccade (red) latencies for each decade in age of the study population. Subjects whose mean pro-saccade latency is the median of the corresponding age group were chosen to represent each group. No censoring was applied to eliminate anticipatory saccades. AVG: average latency; SD: standard deviation; N: number of eye movements.

Our second observation is that, since we now implement both pro- and anti-saccade tasks and that anti-saccade latencies are usually larger than pro-saccade latencies, we need to increase the size of the window where we fit our $\tanh$ model. However, by doing so, we also increase the potential of including more than one saccade movement in the window. For example, subjects may make a hypometric saccade or return their gaze towards the center of the screen. Being able to measure saccade latency from these traces is crucial, especially when these eye movements indicate a certain phenotype. For instance, patients with Parkinson's disease may make more hypometric saccades [6], [11] than patients age-matched controls. Our previous saccade latency measurement algorithm cannot measure latencies from these traces since a $\tanh$ model with a fixed window cannot fit well on these traces. In this work, we show how we can find the appropriate windows of fit for these traces and thus enable saccade latency measurement. By doing so, we keep 96% of the traces to be either a good saccade (the saccade with NRMSE $\leq$ 0.1) or an error saccade, which is much more than 82% of the traces to be either a good saccade or an error if we use a fixed window.

Our third observation is that, to detect directional error is the same as to detect a change in the negative direction in an eye-movement trace. We extend the CUSUM algorithm for this purpose and show that our error detection algorithm can achieve a sensitivity of 97% and a specificity of 97%. Our final observation is that, given the absence of infrared illumination, high-speed cameras, and chinrests, there may be more dLSs and bad saccades (saccades with NRMSE > 0.1) in recordings where the subject did not record themselves properly or had several head movements. If we still estimate the directional error rate as the proportion of directional errors out of all the saccades as in the clinical literature, when there are a lot of dLSs, we

may underestimate the directional error rate. As a result, after discarding undesirable recordings (recordings with more than half of the saccades being dLSs or bad saccades), we define the directional error rate as the proportion of directional errors excluding dLSs and show that this definition is a reasonable approximation for the directional error rate used in the clinical literature. The accuracy of such an approximation depends on the number of false positives and the number of false negatives introduced by the directional error detection algorithm. Therefore, the accuracy of this approximation can be improved if one can further improve the 0.97 and 0.97 sensitivity and specificity we have achieved.

Age and eye movement features: With the improvement in our measurement pipeline, we took 6,823 recordings from 80 subjects ranging in age from 20 years to 92 years, a significantly larger number compared to our previous work – around 500 recordings from 29 subjects mostly in their 20's and 30's, and most other work collected just a few recordings from each subject [11], [16], [24]. Moreover, we have 43 subjects with multiple recording sessions compared to 11 subjects in our previous work. Even after discarding undesirable recordings, we retained 6,787 recordings and 235,520 eye movements from 80 subjects.

As in the literature, we observe that anti-saccade latency and directional error rate tend to be larger than pro-saccade latency and directional error rate, respectively. Across the age range, we also observe that saccade latency is positively correlated with age while a strong relationship between directional error rate and age is not apparent. This observation also matches the observation in prior work [22], [23]. Although our saccade latency values are smaller than values reported in [22], [23], our values are within the range of latency values reported in the clinical literature [10], [16], [25], [26]. Several hypotheses may help explaining why

our values may be smaller. First, our recording setup is less constrained. As mentioned in [3], recording subjects in dedicated environments may affect a subject's cognitive awareness. Second, the amplitude and the distance in our experiments are smaller than the setup in [22]. It was studied that a smaller saccade amplitude and a smaller distance may result in a smaller reaction time [4], [27] Third, with longitudinal recordings, some latency may become smaller due to learning effects. Last but not least, we also have fewer subjects in the 70's and 80's than in other age brackets. While one of the three subjects in the 70's has latency values much closer to the values reported in the literature, two other subjects have smaller latency values.

We also observe that the definition of an anticipatory saccade may significantly affect the measured pro-saccade latency and directional error rate. While the definition is not consistent across the clinical literature, our observation suggests that a more careful investigation into the effect of picking a latency threshold for anticipatory saccades on mean saccade latency is warranted. Some investigations designed tasks to avoid anticipatory saccades [28], [29], for example, by randomizing the length of the fixation period or by including more positions where a stimulus can be presented. However, we suspect that these modifications may result in an increased directional error rate. Since we aim to design and validate our directional error detection algorithm in this work, we did not implement either of these modifications. Nevertheless, it is worth analyzing how these modifications may affect saccade latency and directional error rate.

Last but not least, we show that with multiple recordings from each subject, we can study individual saccade latency distributions. We observe that there is significant intra- and inter-subject variability in these distributions. This observation suggests that such distinctive differences within and across subjects is lost if we were only to report a single summary statistic (mean or median) for each subject or across each age group. Our pipeline fundamentally enables the collection and the analysis of a large number of measurements to characterize the distributional characteristics for each subject.

VI. CONCLUSION

In this work, we developed, validated, and deployed an app to allow for robust determination of pro- and anti-saccade latencies in a visual reaction task. Additionally, we extended our previously reported signal processing pipeline to automatically detect low-signal recordings that should not be further analyzed and also identified directionally erroneous eye movements. With this platform in place, we collected over 235,000 eye movements from 80 self-reported healthy volunteers ranging in age from 20 to 92 years, an order of magnitude more measurements than presented in our previous work. We observed that pro- and anti-saccade latency is positively correlated with age whereas the relationship between directional error rate and age is not significant. Moreover, we observed notable intra- and inter-subject variability across participants, which highlights the need to track eye-movement features in a personalized manner. By enabling app-based saccade latency measurements and directional error rate determination, our work paves the way to use these digital biomarkers to aid in the quantification of neurocognitive decline and possibly from the comfort of the patient's home.
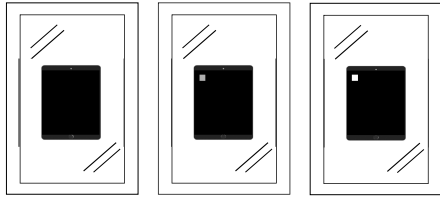

Fig. 14. Example for determining $\hat{r}_i$, the time when the i-th stimulus appears. In this example, the first stimulus appears in recording frame 85 at $\hat{r}_1 = 832.157$ ms.
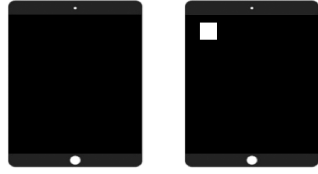

Fig. 15. Example for acquiring s_i, the time when the i-th stimulus presents on the screen. Picture 11 is a black image, and Picture 13 is the image with a left stimulus. The first stimulus shows up when Picture 13 is displayed. As a result, in this example, $\hat{s}_1 = 832.392$ ms.

APPENDIX

In this appendix, we detail how we bound the error associated with saccade latency determination using the app. The accuracy of the saccade latency determination depends on the following four timings:

- the real timing of the i-th stimulus presentation time s_i,
- the estimated timing of the i-th stimulus presentation time $\hat{s}_i$ from the screen timestamps,
- the real timing of the i-th saccade onset o_i,
- the estimated timing of the i-th saccade onset $\hat{o}_i$.

The i-th saccade latency is $o_i - s_i$ whereas the estimated i-th saccade latency is $\hat{o}_i - \hat{s}_i$. Therefore, the accuracy of the i-th saccade latency estimation is $(o_i - s_i) - (\hat{o}_i - \hat{s}_i) = (o_i - \hat{o}_i) - (s_i - \hat{s}_i)$. We can define $D_i^s := s_i - \hat{s}_i$ as the error in the screen timestamps for the i-th stimulus. Similarly, we have $o_i - \hat{o}_i = D_i^r + D_i^t$ where D_i^r is the error in the recording timestamps for the i-th saccade and D_i^t is the error introduced by the $\tanh$ fitting algorithm. The accuracy of the saccade latency estimation then becomes $D_i^r - D_i^s + D_i^t$. While D_i^t was evaluated in [15] to be close to zero, $D_i := D_i^r - D_i^s$ is affected by queued access to the processor clock and is related to our app design.

To estimate D_i, we designed the following experiment. We placed the device in front of a mirror and ran a 40-saccade task. With the mirror, we can identify the recording frame in which each of the 40 stimuli appears first. In Fig. 14, for example, the first stimulus began to appear in Frame 85. With the 40 frame indices and the associated recording timestamps, we can translate these indices into time instants $\hat{r}_i$ (ms), $i = 1, \ldots, 40$. In Fig. 14, $\hat{r}_1 = 832.157$ ms. Similarly, from the screen timestamps, we can obtain the time $\hat{s}_i$ when the i-th stimulus is shown on the screen. Fig. 15, shows $\hat{s}_1$ to be 832.392 ms. (The timestamps were acquired from "mach absolute time clock," which starts at an arbitrary point. More information can be found in Apple Developer Document.)
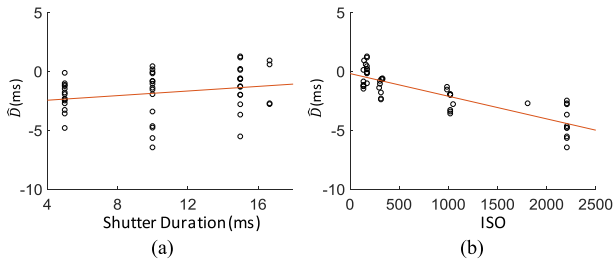
Fig. 16. The estimated synchronization error as a function of (a) shutter duration and (b) ISO. Each dot denotes one recording.

Since the stimulus appearing on the screen would be captured by the next camera frame and the time difference between two frames is $\frac{1000}{60}$ ms in a 60-fps recording, with the errors D_i^r and D_i^s in the timestamps, we have $\hat{r}_i + D_i^r - \frac{1000}{60} < \hat{s}_i + D_i^s \leq \hat{r}_i + D_i^r$. That is, $\hat{r}_i - \frac{1000}{60} < \hat{s}_i - D_i \leq \hat{r}_i$ where D_i is exactly the error in the saccade latency estimation introduced by the synchronization error between the screen and the recording.

From the recording timestamps, we can only find one time instant $\tilde{r}_i(D)$ as a function of a variable D that satisfies $\tilde{r}_i(D) - \frac{1000}{60} < \hat{s}_i + D \leq \tilde{r}_i(D)$. In other words, if each recording timestamp is denoted as t_j where j denotes the frame index as in Fig. 14, then $\tilde{r}_i(D) := \min\{t_j | t_j \geq s_i + D\}$. We now have defined a function $\tilde{r}_i(D)$ of D for each stimulus. Our goal is to estimate the synchronization error $\hat{D}$ of a recording using all these functions. If our estimated synchronization error $\hat{D} = D_i$ for all i, we will have $\tilde{r}_i(\hat{D}) = \hat{r}_i$. As a result, we can then define

$$\hat{D} = \arg\min_D \sum_i |\tilde{r}_i(D) - \hat{r}_i|. \tag{3}$$

With careful app design, we can ensure $\sum_i |\tilde{r}_i(\hat{D}) - \hat{r}_i| = 0$. That is, we achieve $\hat{D} = D_i$ where the estimation of the synchronization error is correct and is constant throughout each recording.

We observed that an iOS camera changes its shutter duration and ISO based on the lighting condition, which may affect the accuracy of the recording timestamps. We showed in Fig. 16 that the shutter duration does not affect the synchronization error while ISO is positively correlated with the absolute value of the synchronization error. As a result, we set the shutter duration to 16 ms, which is close to the maximum duration $1000/60$ ms in a 60-fps recording, to allow for adequate light. To bound the absolute synchronization error to be within 5 ms, we restrict the ISO values to be less than 1000 by asking the subject to move to a brighter environment if the automatically determined ISO exceeds 1000.

REFERENCES

[1] S. Hoops *et al.*, "Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease," *Neurology*, vol. 73, no. 21, pp. 1738–1745, 2009.

[2] A. Mitchell, "A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment," *J. Psychiatr. Res.*, vol. 43, no. 4, pp. 411–431, 2009.

[3] National Academies of Sciences, Engineering, and Medicine, *Harnessing Mobile Devices for Nervous System Disorders: Proceedings of a Workshop*. Washington, DC, USA: The National Academies Press, 2018.

[4] R. Leigh and D. Zee, "saccadic system," *The Neurol. Eye Movements*. Oxford, U.K.: Oxford Univ. Press, 2015, ch. 4, pp. 169–288.

[5] S. Tabrizi *et al.*, "Biological and clinical manifestations of Huntington's disease in the longitudinal track-hd study: Cross-sectional analysis of baseline data," *Lancet Neurol.*, vol. 8, no. 9, pp. 791–801, 2009.

[6] T. Anderson and M. MacAskill, "Eye movements in patients with neurodegenerative disorders," *Nature Rev. Neurol.*, vol. 9, no. 2, pp. 74–85, 2013.

[7] D. Munoz and S. Everling, "Look away: The anti-saccade task and the voluntary control of eye movement," *Nature Rev. Neurosci.*, vol. 5, no. 3, pp. 218–228, 2004.

[8] J. E. McDowell *et al.*, "Neurophysiology and neuroanatomy of reflexive and volitional saccades: Evidence from studies of humans," *Brain Cogn.*, vol. 68, no. 3, pp. 255–270, 2008.

[9] R. Shafiq-Antonacci *et al.*, "Spectrum of saccade system function in Alzheimer's disease," *Arch. Neurol.*, vol. 60, no. 9, pp. 1275–1278, 2003.

[10] T. Crawford *et al.*, "Inhibitory control of saccadic eye movements and cognitive impairment in Alzheimer's disease," *Biol. Psychiatry*, vol. 57, no. 9, pp. 1052–1060, 2005.

[11] U. Mosimann *et al.*, "Saccadic eye movement changes in Parkinson's disease dementia and dementia with Lewy bodies," *Brain*, vol. 128, no. 6, pp. 1267–1276, 2005.

[12] S. Garbutt *et al.*, "Oculomotor function in frontotemporal lobar degeneration, related disorders and Alzheimer's disease," *Brain*, vol. 131, no. 5, pp. 1268–1281, 2008.

[13] H.-Y. Lai *et al.*, "Enabling saccade latency measurements with consumer-grade cameras," in *Proc. IEEE Int. Conf. Image Process.*, 2018, pp. 3169–3173.

[14] G. Saavedra-Peña *et al.*, "Determination of saccade latency distributions using video recordings from consumer-grade devices," in *Proc. IEEE Eng. Med. Biol. Conf.*, 2018, pp. 953–956.

[15] H.-Y. Lai *et al.*, "Measuring saccade latency using smartphone cameras," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 3, pp. 885–897, Mar. 2020.

[16] S. Rivaud-Péchoux *et al.*, "Mixing pro- and antisaccades in patients with parkinsonian syndromes," *Brain*, vol. 130, no. 1, pp. 256–264, 2006.

[17] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2001, pp. 511–518.

[18] M. Nyström and K. Holmqvist, "An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data," *Behav. Res. Methods*, vol. 42, no. 1, pp. 188–204, 2010.

[19] A. Savitzky and M. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, 1964.

[20] F. Gustafsson, "On-line approaches," in *Adaptive Filtering and Change Detection*. ch. 3, Chichester, U.K.: Wiley, 2001, pp. 55–87.

[21] E. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1/2, pp. 100–115, 1954.

[22] D. Munoz *et al.*, "Age-related performance of human subjects on saccadic eye movement tasks," *Exp. Brain Res.*, vol. 121, no. 4, pp. 391–400, 1998.

[23] B. Fischer, M. Biscaldi, and S. Gezeck, "On the development of voluntary and reflexive components in human saccade generation," *Brain Res.*, vol. 754, no. 1-2, pp. 285–597, 1997.

[24] Q. Yang *et al.*, "Specific saccade deficits in patients with Alzheimer's disease at mild to moderate stage and in patients with amnestic mild cognitive impairment," *Age*, vol. 35, no. 4, pp. 1287–1298, 2013.

[25] J. Holden *et al.*, "Prodromal alzheimer's disease demonstrates increased errors at a simple and automated anti-saccade task," *J. Alzheimer's Dis.*, vol. 65, no. 4, pp. 1209–1223, 2018.

[26] C. Bonnet *et al.*, "Eye movements in ephedrone-induced parkinsonism," *PLoS One*, vol. 9, no. 8, pp. 1–8, 2014.

[27] Q. Yang, M. Bucci, and Z. Kapoula, "The latency of saccades, vergence, and combined eye movements in children and in adults," *Investigate Ophthalmol. Vis. Sci.*, vol. 43, pp. 2939–2949, 2002.

[28] A. Boxer *et al.*, "Saccade abnormalities in autopsy-confirmed frontotemporal lobar degeneration and Alzheimer's disease," *Arch. Neurol.*, vol. 69, no. 4, pp. 509–517, 2012.

[29] S. Hopf *et al.*, "Age dependent normative data of vertical and horizontal reflexive saccades," *PLoS One*, vol. 13, no. 9, 2018, Art. no. e0 204008.