Aurora

Performance Report for Practical Use Case Scenario

2024-08-23

Smart Eye
Research Instruments

This report presents key performance indicators for the Aurora when utilized in a practical setting mirroring real-world use cases. For additional information on the diagrams and the terminology used in them, see the appendix at the end of the document.

1 Measurement Setup

Table 1.1

Camera Type	*Framerate*
Aurora	60

Table 1.2

Screen Dimensions	*Screen Diagonal*	*Mean Target Separation (X-Axis)*[1]	*Mean Target Separation (Y-Axis)*[1]
0.531 x 0.298 m	0.609 m/24 in	0.056 m	0.056 m

[1] Note: *Target Separation* only takes into account the targets that the subjects actually looked at (which is not all of the targets displayed in the world model).

The data collection was performed in a dimly lit room. The single window present was covered using blinds and ceiling lights were on. The subject was illuminated from the top/front with a maximum solar irradiance of 0.2 W/m^2. The screen was positioned 65 cm in front of the subject with its upper edge level with the subject's eyes. The tracker was mounted below the screen.

Before targets were shown, the subject was instructed to look around. When the subject was done looking around and had returned to a neutral head position, targets started being presented one at a time as black dots on a light gray background. No instructions on movement were given while targets were being displayed.

Figure 1.1 and 1.2 illustrate the locations of the targets that the subjects were asked to look at and the head position in relation to the setup. This is a 3D representation and not what the subject actually saw. In total a subject looked at 45 different targets. The data presented in this report is calculated exclusively in the time frame during which a subject has been requested to look at a certain target. No data is presented for activity in between, i.e while switching targets.

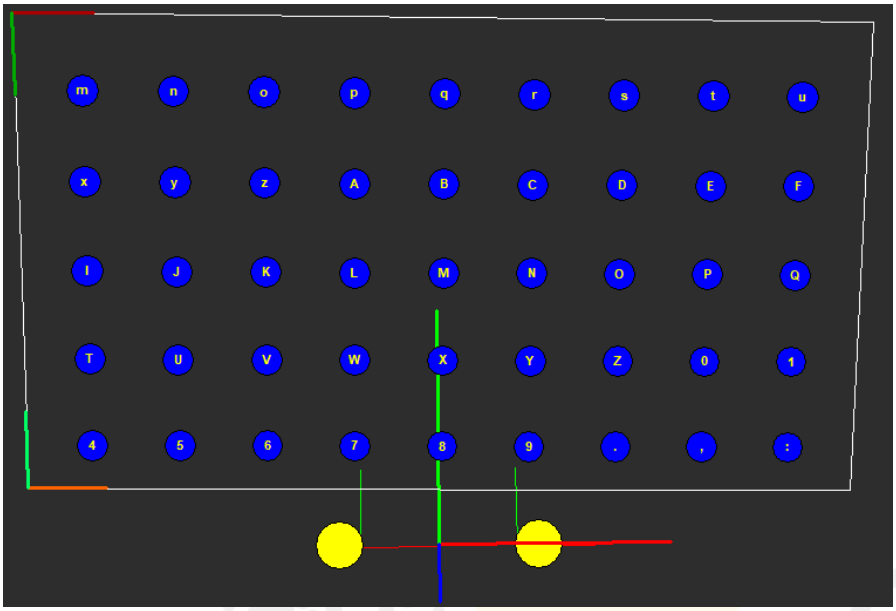

Figure 1.1: A wireframe visualization of the screen in the measurement setup that the subjects are looking at. The blue spheres are locations where targets can be displayed for the subject to look at.

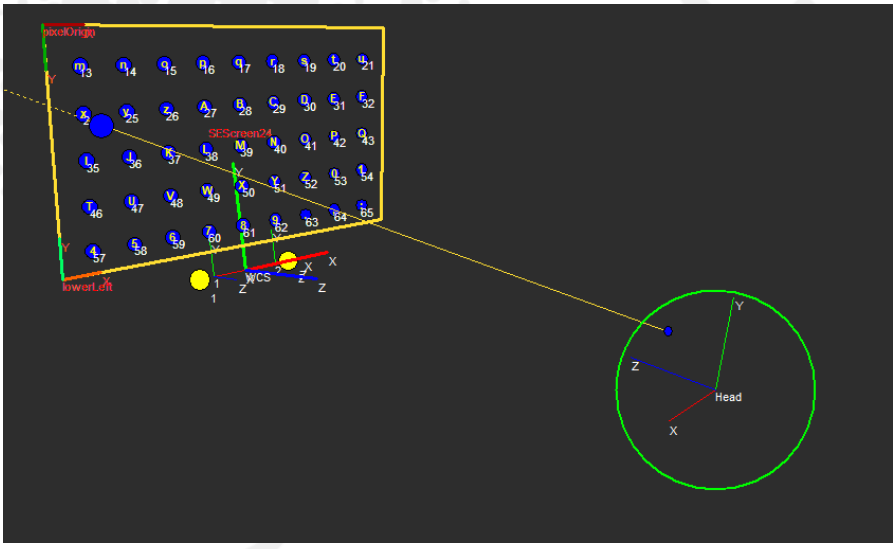

Figure 1.2: An angled view of a wireframe visualization of the measurement setup. The green sphere represents the subjects head position and the red line represents the gaze direction.

2 About The Data

Table 2.1

Total	*No Glasses*	*With Glasses*	*Lenses*	*Beard*	*Makeup*	*Long Hair*	*Male*	*Female*
38	30	8	0	17	10	9	27	11

In total 38 subjects participated in the recordings. Table 2.1 presents a summation of different characteristics that the subjects have, such having a beard or wearing glasses. A single subject can have multiple characteristics, meaning that the summation of the different characteristics can be greater than the total number of subjects. The subjects were not specifically asked about their ethnicity. Based solely on visual observation, approximately 16% are of east asian descent and the remaining 84% are of european origin. The participants did not use a chin rest, were asked to sit comfortably, and not prompted to adjust their seating during the test. Each subject looked at each target for a maximum of 2 seconds (when collecting samples).

Table 2.2

Blue Eyes	*Gray Eyes*	*Green Eyes*	*Brown Eyes*	*Other*	*Unknown*
15	1	9	6	3	4

The eye colors of the subjects are presented in table 2.2. *Other* specifies that the subject had an eye color that did not fit into the other categories or was hard to identify. *Unknown*, on the other hand, specifies that we have no data on what eye color the subject has.

Table 2.3

Best Subject (Smallest Error)	*All Subjects Mean*
0.408°	0.636°

Table 2.3 shows the best case and what the average *median mean error* over all subjects is. The *median mean error* is the mean error for each target, and then the

median over all targets for that subject. An error is the difference, in degrees, between the target and the gaze intersection of the infinite 2D plane which lies in tangent to the target point.

Figure 2.1, 2.2, and 2.3 show the number of targets (y-axis), as a percentage, for which the *mean error* is below a certain threshold (x-axis). The figures are split into different categories based on characteristics, with a shared plot of the distribution of errors for all subjects combined.

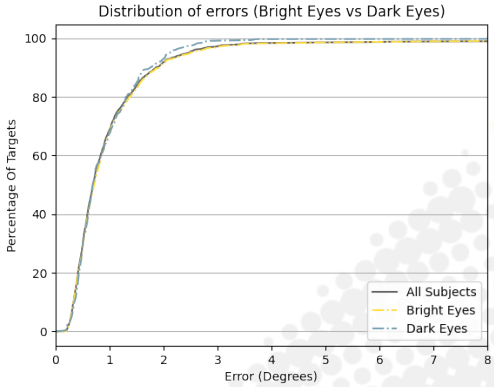

Figure 2.1

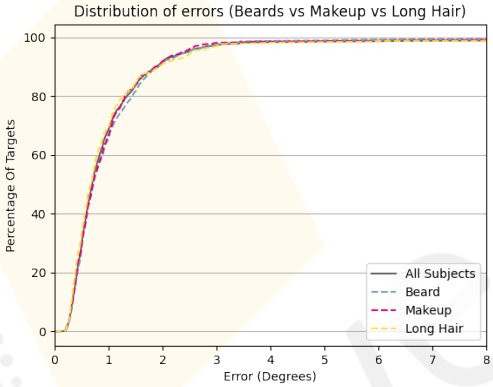

Figure 2.2

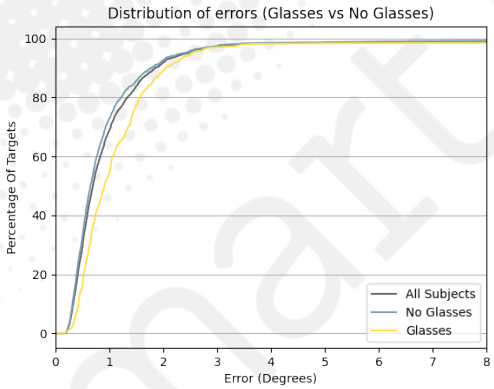

Figure 2.3

For Smart Eye tracking systems *gaze quality* is a value that indicates how confident the system is with its gaze calculation. The quality value is expressed on a range from 0.0 to 1.0, where 1.0 indicates that the system is very confident and 0.0 indicates that the system has no confidence about the calculated gaze direction. It is essential that the *gaze quality* value is taken into account before using any data. For this test, any data with a quality lower than 0.5 was filtered away. *Gaze availability* is used to indicate the proportion of frames for which the system provided gaze with a quality of over 0.5.

The following are heatmaps showing the mean error for each target on the screen. A heatmap represents the screen itself; the squares are arranged in the same way as the targets are on the screen in real life (see image of the measurement setup). In figure 2.4 the mean error for each target, averaged over all subjects in the respective category, is illustrated. Also illustrated is the best mean error measured per target. The availability across the entire screen is practically 100%, there are a very small number of targets that drop to at most 95% for certain features.

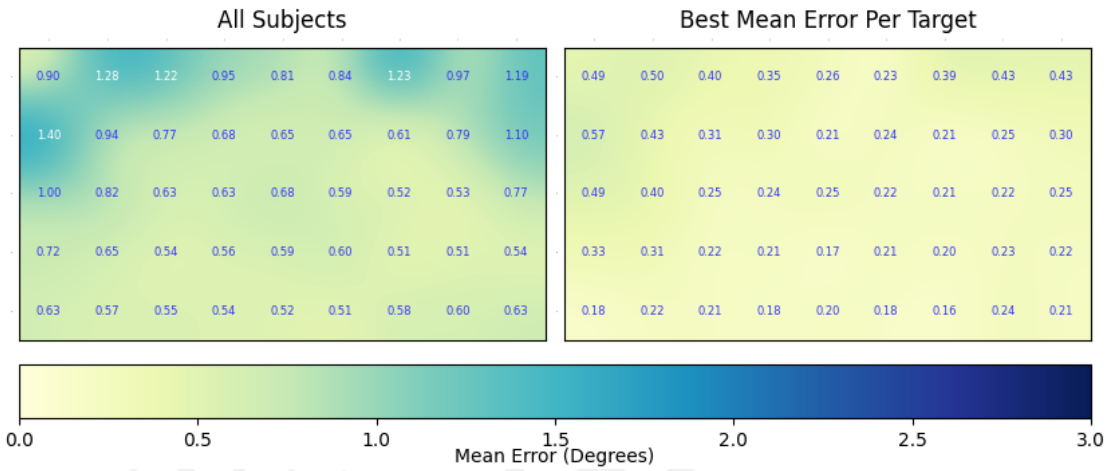

Figure 2.4

3 Conclusion

The purpose of this report was to test the performance of the Aurora in ways that applied researchers ideally would like to use it, in flexible situations and with diverse populations The report shows that even with a diverse dataset of N=38, with only the room lighting controlled, the average gaze error for 45 target points on a 24 inch screen is as low as 0.636 degrees. There are of course individual differences with the best participants displaying a median mean error of 0.408 degrees and those with glasses showing higher mean errors. Hair, beard, eye-color and makeup did not significantly affect the results and were consistent across the different points on the screen. The report takes raw data to evaluate from all participants but these values could be improved further by stricter exclusion criteria, using a chin rest, and filtering data during fixation classification. The present report shows the Smart Eye Aurora 60 Hz to be at par with state-of-the-art eye-trackers and shows robust and consistent performance in real life testing situations that can be generalized globally.

4 Appendix

The data used in the report is taken from the data file located in the folder for the precise analyses for the set of recordings. In this section there will be a more detailed explanation of what the figures represent and the measures that were used.

4.1 Distribution Plots (Errors)

These plots show the distribution of errors in degrees. To generate this plot the individual values for mean errors for each measured target for all subjects were used. The y-axis shows the cumulative error distribution (the percentage of errors that are under the error size for each x value).

4.2 Median Mean Error

Median mean error is a measure that shows the median value of the mean errors. For every target that the subject looks at a mean error is calculated. The median of all target mean errors for each subject is then calculated.

4.3 Bar Chart

4.4 Heatmaps: Mean Error/Gaze Availability per target on screen

The heatmap is divided into 45 boxes which each represent one of the 45 targets on the screen. These values are then interpolated to create the colored heatmaps shown. This gives a good approximation for any point on the screen.