Calculating the Accuracy of Eye Tracker Devices and the Impact of Fixation Algorithms on this Accuracy>

Prepared by Furkan Duman

Middle East Technical University Northern Cyprus Campus

Computer Engineering Department

01/01/2024

Contents

1	Purpose							
2	2 Factors in Data Quality							
3	Methods 3.1 Test Screen	4 5						
	3.1 Test Screen	6						
	3.3 Eye Data Focusing Rates	7						
	3.4 The method for obtaining test results	8						
4	First Test							
5	The Impact of Fixation Algorithms on Accuracy	11						
	5.1 Test With Fixation Algorithm	12						
	5.2 Comparing Test Results	13						
6	Future Work	19						

List of Figures

1	Test Screen	5				
2	Extract the raw data	6				
3	Function that finds which gaze points belong to which cell	7				
4	Shows which cell the eye data belongs to at different time intervals 7					
5	Accuracy Result	8				
6	Eye gaze rates for each time period-PART1	9				
7	Eye gaze rates for each time period-PART2	9				
8	Calculate the accuracy ratio	10				
9	Eye gaze rates with I-DT for each time period-PART1	12				
10	Eye gaze rates with I-DT for each time period-PART2	12				
11	Calculate the accuracy ratio with I-DT	13				
12	10s-15s	14				
13	20s-25s	15				
14	35s-40s	16				
15	120s-125s	17				
16	140s-150s	18				

1 Purpose

This study aims to explore the process of calculating the accuracy of eye-tracking devices. Through my research, I found that typically, the accuracy is determined after calibration tests by analyzing the correct gaze ratio obtained from focusing on specific areas. In my previous research, I couldn't find a dedicated system for testing eye-tracker accuracy, so I decided to develop my own solution to address this gap.

The other primary motivation for starting the project was the insufficient number of studies on the accuracy of fixation algorithms. Expressing the accuracy of fixation algorithms mathematically is believed to further facilitate the calculation of "system accuracy" for the EASIEST project. Therefore, this study takes a step towards a thorough understanding of the process of determining the accuracy of eye-tracking devices and how fixation algorithms impact this accuracy.

2 Factors in Data Quality

I am aware that designing the lighting system well is crucial to achieve accurate results during eye tracking. During the tests, I contributed to obtaining more precise and accurate data by directing all lighting systems towards the face of the person conducting the test. Additionally, to ensure a more stable experience, the test participant aimed to minimize head movements during the tests.

Kerem Keptiğ played a significant role in integrating the eye tracker device into my system. I thank him very much for integrating the current most seamless version of Webgezer into our system, and we tried to achieve the best efficiency with this configuration.

3 Methods

In this **flask** project, first I created an HTML for the test screen. To calculate the accuracy of eye tracking devices, I created a total of 48 cells and determined the individual coordinates for each cell. During the test, the participant looks at cells from the 1st to the 48th every 5 seconds. After completing the test, the raw dataset provides X, Y, and time values of eye data. Using these values, the determination of which eye data corresponds to which cell is calculated. Subsequently, for each 5-second time interval, the analysis of which cell the eye data belongs to is conducted, and based on this analysis, the accuracy rate is calculated.

When calculating the accuracy rate, it is crucial to consider the time intervals. For example, if the participant looks at the 1st cell and its neighboring cells (8th and 9th cells) between 0 and 5 seconds, this gaze is classified as a correct point. If the gaze goes elsewhere, those points are classified as incorrect points. The accuracy rate of the 1st cell between 0 and 5 seconds is then calculated by dividing the number of correct points by the total eye data. This process is repeated for each time interval.

Subsequently, the average accuracy value of the system is calculated by taking the arithmetic mean of the accuracy values calculated for all 48 time intervals.

Later on, I obtained fixation points from the raw data collected during the test using fixation algorithms. I observed how fixation algorithms had an impact on the accuracy of the system by analyzing these fixation points.

To make it easier to understand how often cells are looked at in specific time intervals, I made pie charts for each time period using the **matplotlib** library. These charts show visually how much attention cells get, helping us see how well the system works at different times.

3.1 Test Screen

1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	32	33	34	35
36	37	38	39	40	41	42
43	44	45	46	47	48	Finish

Figure 1: Test Screen

In this test display, there are a total of 48 cells.n. WebGazer usually starts recording eye data from the center of the screen, so when the eye data first arrives at the 1st cell, the system time is considered 0. Subsequently, every 5 seconds, the gaze shifts to the next cell. The entire test duration is calculated by multiplying 48 cells by 5 seconds, resulting in a total of 240 seconds.

At the end of the test, the raw data is transmitted to the system in the form of X, Y, and time metrics.

3.2 Eye Tracker Data Analysis and Find the Cells

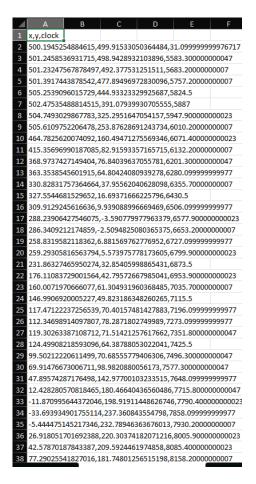


Figure 2: Extract the raw data

Raw data can contain thousands of rows for 240 seconds of data. I take all of these gaze points as x, y and clock, just like in the EASIEST project.

Then, I find which cell each point belongs to, according to the cell coordinates I calculated before. I use the following function for this.

Figure 3: Function that finds which gaze points belong to which cell

3.3 Eye Data Focusing Rates

After analyzing the X, Y, and time values associated with eye movements, the system calculates the distribution of eye data across each time interval, determining the percentage of data points allocated to each cell.

```
0s - 5s: {1: 0.8, 2: 0.066666666666666667, 9: 0.1333333333333333333}

5s - 10s: {9: 0.21428571428571427, 2: 0.7142857142857143, 10: 0.07142857142857142}

10s - 15s: {10: 0.35714285714285715, 11: 0.42857142857142855, 3: 0.21428571428571427}

15s - 20s: {4: 0.5, 11: 0.25, 18: 0.125, 26: 0.125}
```

Figure 4: Shows which cell the eye data belongs to at different time intervals

This image illustrates which cells the eye data corresponds to during specific time intervals. Typically, these data are calculated for up to 240 seconds, but I'm including this image briefly for explanation.

3.4 The method for obtaining test results

Then, the data is analyzed to determine which cells should be looked at during specific time intervals, and their respective rates are calculated. Subsequently, the arithmetic mean of these rates is obtained, resulting in the average accuracy calculation.

```
 \text{Average Correct Area Viewing Ratio} = \frac{\text{Correct Area Viewing Ratio}_1 + \ldots + \text{Correct Area Viewing Ratio}_{48} }{48}
```

As I said before, correct look rates are calculated at each time interval, and since there are 48 cells, the average is taken by dividing by 48.

For a simpler explanation, I only showed the data up to the 20th second in the figure above. I will now calculate the accuracy rate for this data.

```
[0.8666666666666667, 0.9285714285714286, 0.5714285714285714, 0.75]
Accuracy Result: 0.779166666666667
```

Figure 5: Accuracy Result

For instance, the value 0.866666666666667 in this graph is calculated by summing the accuracy ratios of the 1st cell and its neighboring cells (neighbors for the 1st cell are the 8th and 2nd cells) between the 0th and 5th seconds. As an example, the last element in the list with a value of 0.75 is calculated as the total value of the 4th cell and its neighboring cells (ratios from the 4th cell are 0.5, and from the 11th cell, 0.25).

And the average accuracy was calculated by taking the average of these values.

I want to say again, I used only 4 time intervals for easier understanding, but the system has a total of 48 time intervals.

4 First Test

Kerem Keptiğ volunteered for the first system test, and I appreciate his participation. Thank you!

After setting up the necessary environment, the test started. At the beginning of the test, the calibration test of WebGazer yielded a result of 92%. In our test, the calculated value for accuracy was 87.788%.

```
0s - 5s: {1: 0.941764705882355, 8: 0.058823529411764705}
5s - 10s: {1: 0.4708247768227682, 2: 0.47085570685470}, 9: 0.02564102564102564, None: 0.132962476823247685, 2: 0.470855742685714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.0087285714285, 9: 0.008728572957782, None: 0.0087457627118644, 11: 0.00847457627118644, 10: 0.02872772813559734, 9: 0.0087285727711865, 8: 0.440728727118644, 9: 0.0087457627118644, 11: 0.00847457627118644, 10: 0.028727727118574, 9: 0.0087287728711865, 9: 0.4407287118574, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 11: 0.00847457627118644, 10: 0.0287277727727777, 9: 0.00874758421085, 9: 0.4912870175418574, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874757627118644, 9: 0.00874767627118644, 9: 0.00874767627118644, 9: 0.00874767627118644, 9: 0
```

Figure 6: Eye gaze rates for each time period-PART1

Figure 7: Eye gaze rates for each time period-PART2

It was found that the eye data looked at which cells and at what rate in which seconds. Now, using this data, it will be calculated to what extent the eye data is looking at the right area.

Figure 8: Calculate the accuracy ratio

After calculating the correct gaze ratio for each time interval, the average accuracy was computed by taking the mean of these values. The average accuracy was determined to be 87.79%.

5 The Impact of Fixation Algorithms on Accuracy

We know that fixation algorithms improve accuracy. However, I couldn't find a study on the specific rates at which they enhance accuracy. Therefore, I aimed to conduct this study.

During my tests, I employed the I-DT algorithm, one of the most well-known dispersion-based fixation algorithms. The aim was to evaluate its impact on the accuracy of the eye-tracking system.

For the I-DT algorithm, I utilized the raw eye-tracking data from Kerem Keptiğ, which initially had an accuracy of 87.73%. Then, I calculated the corresponding cells for the resulting fixation points. By analyzing the time intervals, I determined which cells each fixation point belonged to. I calculated its corresponding cell and, based on the results, the I-DT algorithm improved the accuracy of the data from 87.73% to 90.25%."

5.1 Test With Fixation Algorithm

Figure 9: Eye gaze rates with I-DT for each time period-PART1

```
150s - 155s: {30: 0.35443037974683544, 24: 0.20253164556962025, 31: 0.4430379746835443}
155s - 160s: {31: 0.3176470588235274, 24: 0.09411764705882353, 32: 0.5294117647058824, 25: 0.058823529411764705}
160s - 165s: {35: 0.4225352112676056, 25: 0.056338028164914086, 26: 0.070427253521126761, 33: 0.45070427253521127}
165s - 170s: {33: 0.3581851381851851851875187518517, 24: 0.185185185187, 34: 0.4691358024601358, 27: 0.160847832716047832}
170s - 175s: {34: 0.3222222222222224, 27: 0.111111111111111, 35: 0.4333333333333, 28: 0.08888888888889, 42: 0.0444444444444446}
175s - 180s: {35: 0.1518987341772152, 28: 0.21518987341772155, None: 0.05063291139240506, 29: 0.189873417721519, 36: 0.3924050632911392}
180s - 180s: {35: 0.15254237288135594, 29: 0.15254237288135594, 37: 0.62711864406677966, 29: 0.16979661016949153}
185s - 190s: {37: 0.28723406425531915, 30: 0.20212765957446807, 38: 0.51063829787234064}
190s - 195s: {30: 0.0111111111111111, 38: 0.466666666666666, 39: 0.43333333333333, 32: 0.05555555555555555, 31: 0.03333333333333}
195s - 200s: {31: 0.020618556701030927, 39: 0.463917525773199, 40: 0.5154639175257731}
200s - 200s: {40: 0.035615789473684, 39: 0.05563157894736842, 31: 0.05263157894736842, 41: 0.5789473684210527, 34: 0.013157894736842105}
205s - 210s: {34: 0.03571428571428571, 41: 0.2857142857142857, 40: 0.047619047619047610, 42: 0.5476190476190477, 35: 0.03571428571428571, 48: 0.0476190476190476190476190476190477, 35: 0.108427052631578947368421052, 32: 0.1842105263157894, 36: 0.308952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952380952
```

Figure 10: Eye gaze rates with I-DT for each time period-PART2

In these results, we can observe that fixation algorithms have eliminated some points with very low gaze rates (saccade points). To illustrate this more clearly, I will present the corresponding graphs.

Figure 11: Calculate the accuracy ratio with I-DT

5.2 Comparing Test Results

To visually demonstrate how fixation algorithms improve accuracy, I created pie charts comparing data with applied fixation algorithms (90.25% accuracy) to data without fixation algorithms (87.73% accuracy) for each time interval from 0s to 240s.

The charts display the rate of each cell in its respective segment. If there is no cell number written above a rate, it means that point is not present on the screen and will be considered as a NULL point.

Since there are a total of 48 time intervals, instead of showing all time intervals, specific rates are displayed. The graphs for all time intervals will be uploaded to GitHub along with the code.

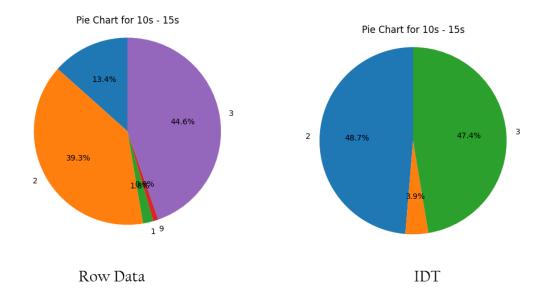


Figure 12: 10s-15s

It is necessary to focus on the 3nd cell and its neighboring cells between 10 seconds and 15 seconds. I-DT identified approximately 14% saccade points and increased accuracy by 9% in this time period.

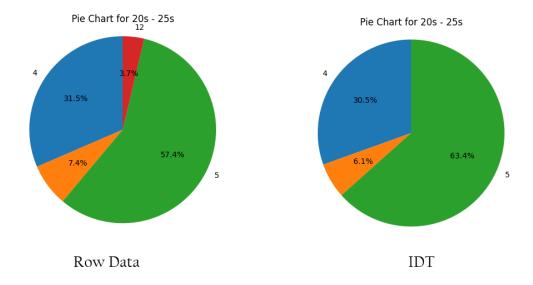


Figure 13: 20s-25s

It is necessary to focus on the 5nd cell and its neighboring cells between 20 seconds and 25 seconds. In this time interval, I-DT has classified the gaze belonging to the 12th cell as a saccade with a rate of 3.7%. Additionally, the rate for the NULL cell has decreased from 7.4 to 6.1.

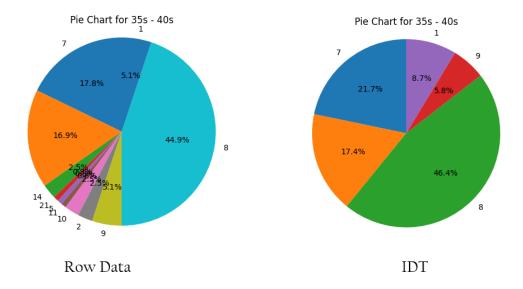


Figure 14: 35s-40s

It is necessary to focus on the 8nd cell and its neighboring cells between 35 seconds and 40 seconds. During this time interval, I-DT has eliminated numerous saccade points. In this time interval without I-DT, a total of 11 points are observed, while with I-DT, saccade points are eliminated, reducing the total gaze points to 5.

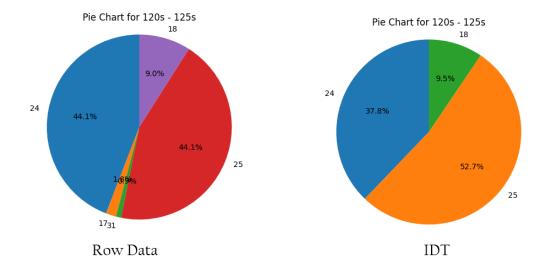


Figure 15: 120s-125s

It is necessary to focus on the 25nd cell and its neighboring cells between seconds and seconds. I-DT has identified and eliminated cells and 31 as saccade points during this time interval.

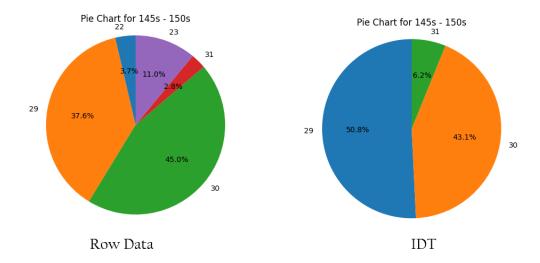


Figure 16: 140s-150s

Between the 140th and 150th seconds, it is expected to focus on cell 30 and its neighbors (cells 23, 29, 37, and 31). Although I-DT has reduced the viewing ratio of cell 30 from 45% to 43.1%, the accuracy of cell 30 and its neighboring cells is considered for calculating the system's accuracy. Since cell 22, found in the raw data, is eliminated, the success rate in this time interval has increased from 96.3% to 100%.

This interval is one of the most illustrative examples of why I include neighboring cells in the accuracy calculation. As we know, one of the biggest problems with eye trackers is the constant tremor of the cursor on the screen. This can sometimes make it appear as if we have shifted our focus slightly from where we are actually looking. Due to this effect and the small size of each cell I created, neighboring cells have been included in the accuracy calculation.

6 Future Work

- I have decided to transform the system I developed into a web-based platform, allowing us to test the accuracy of any eye-tracking device online.
 The aim is to create a web-based website that performs accuracy calculations for eye-tracking devices.
- By making the system web-based, we enable users to easily access and test
 the accuracy of their eye-tracking devices online. This platform will serve
 as a website where users can conveniently perform accuracy calculations
 for their eye-tracking devices, making the testing process more accessible
 and user-friendly.
- We've witnessed how fixation algorithms improved the accuracy. In the current initial test, achieving an accuracy increase from 87.3% to 90.25% is a great result. For now, the rate of improvement is calculated as 3.32%. However, with future tests, an average value will be determined.
- I want to introduce a new perspective to calculate the overall accuracy of the EASIEST project. Instead of using the accuracy of WebGazer, I believe we should calculate the accuracy of fixation points generated by the fixation algorithm and combine this result with the accuracy of the ML model.

For instance, if we plan to use an ML model with 90% accuracy in the upcoming period, we can estimate the system accuracy for the tests in this report as $90.25\% \times 90\% = 81.225\%$.