Assessing Visual Attention Using Eye Tracking: Position, Velocity, and Acceleration as Indicators of Focus

Student: Ioana Munteanu. Mentor: Dr. Stephen J. Guy

Undergraduate Research Scholarship. Computer Science Department. University of Minnesota Twin Cities

INTRODUCTION

Artificial Intelligence has played a major role in recruiting workers for the past few years in terms of matching candidates with job descriptions and enhancing diversity.[1] But what about predicting how skilled a worker is? For the research project, I was interested in answering the question "Between beginner and experienced surgeons, how do their pupils differ in terms of position, velocity, and acceleration?" The results will be used in future machine learning projects that could potentially determine if a surgeon is experienced or not based on the observations that I am going to make, and thus facilitate the hiring process. Previous studies on this topic analyzed the maximum pupil size, the change frequency in pupil size and predictability of pupil change, the eye gaze rate, fixation and saccades duration, and designated areas of interest in the subject's visual field.[2]

PROCEDURE

Data was gathered by Dr. Sofia Lyford-Pike and her team. 18 surgeons at the University of Minnesota performed surgeries while using Pupil Lab's mobile eye-tracking glasses, as shown in Figure 1. For the purpose of our experiment, this number is appropriate, since it generates statistically significant results in a previous study.[3] The glasses recorded the surgeries and registered the x and y coordinates of the pupil over time, as well as the confidence level of each operation.

Figure 1: *Eye-tracking glasses*
1. world camera
2. nose support
3. adjustable eye cameras
4. USB-C connector clip

Source: https://docs.pupil-labs.com/core/hardware/

Each surgeon was assigned an experience level. Residents were assigned an integer between 0 and 6 based on their residency year, while the staff was assigned level 7. Two groups were considered: Juniors (level 0 – 4) and seniors (level 5 – 7). Each surgery had two parts: the scrub-in part and the surgery itself. For each of them, the visual attention was assessed by extending a previous study that analyzed the velocity of the pupils of junior and senior surgeons.[4] My hypotheses are that senior surgeons are more focused than junior surgeons and that all surgeons are more focused while they perform the surgery than when they scrub in.

POSITION AS MEASURE OF FOCUS

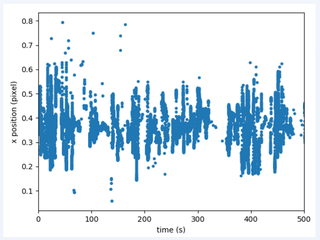


Figure 2: *Pupil Position During the Scrub In*

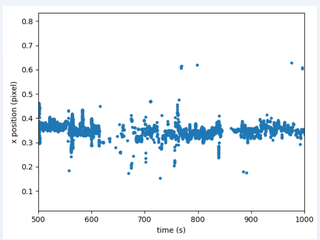


Figure 3: *Pupil Position During the Surgery*

As expected, the movement of the pupil oscillates more in Figure 2 than in Figure 3, which shows that the surgeon is more focused during the surgery than during the scrub in.

VELOCITY AS MEASURE OF FOCUS

The graph of velocity over time was plotted in Figure 4. For each surgeon, the surgery was divided into intervals of 100 milliseconds. The maximum velocity of each interval was registered. If the maximum value was less than a given threshold, then we would call it a focused interval. Otherwise, it would be a non-focused interval. A percentage of focus was calculated by dividing the focused intervals over the total number of intervals.[4] Multiple velocity thresholds were tested by comparing the percentages of focus in juniors and seniors using a t-test. The results are registered in Figure 5.

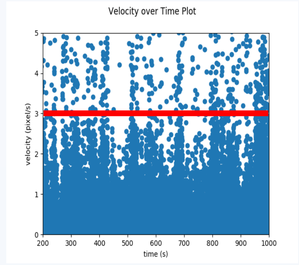


Figure 4: *Velocity in the middle of the surgery. The red line is the threshold = 3 pixel/s*

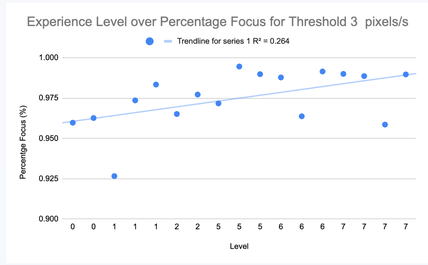


Figure 5: Linear Regression of *Experience Level over Percentage Focus for Threshold 3 pixel/s. R = 0.5138*

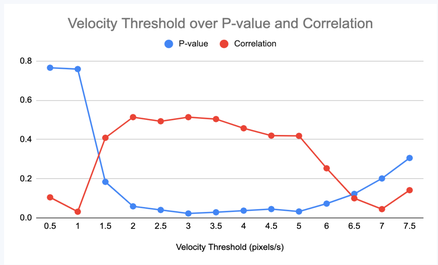


Figure 6: *Velocity Threshold Over P-value and Correlation. The optimal threshold is 3 pixel/s, since it maximizes* ***R = 0.5138*** *and minimizes* ***p-value = 0.0226***

The two samples are random and independent, but not normally distributed. For a confidence level of 95%, $\alpha = 0.05$ and the p-value = 0.0226 < α. The 95% confidence interval of the difference of the mean of percentage of focus of juniors and that of seniors is (-0.0366, -0.0005). Since all the values in the interval are negative, we can conclude that the mean of percentage of focus of juniors is less than the percentage of focus of seniors.

ACCELERATION AS MEASURE OF FOCUS

The same procedure is performed for acceleration in order to find the ideal threshold. The p-value is optimized for the threshold of 500 pixel/s^2, as shown in Figure 6. In this case, the two samples consisting of the percentage of focus junior and senior surgeons respectively, are normally distributed, as shown by the Q-Q plot in Figure 7. Therefore, the acceleration threshold gives more statistically significant results than the velocity threshold.

Figure 7: *Acceleration Threshold Over P-value and Correlation. The optimal threshold is 500 pixel/s, since it maximizes* ***R = 0.5621*** *and minimizes* ***p-value = 0.0358***

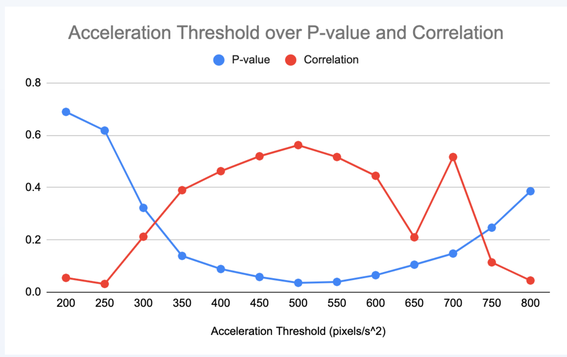
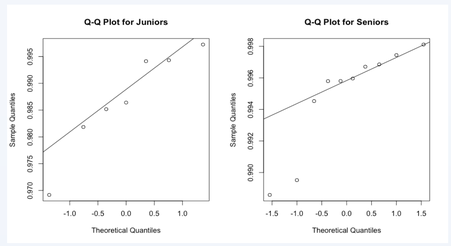


Figure 8 (below): *Q-Q plots for the two groups. Since the distribution is following the straight line, we can conclude that it is normal.*



COMPARISON BETWEEN VELOCITY AND ACCELERATION

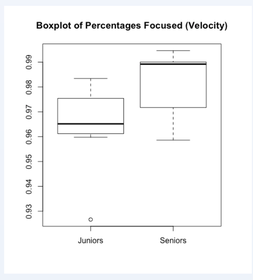


Figure 9: *Boxplot of Percentage Focused calculated using velocity. The median of seniors is greater than the median of juniors. Variances are similar, with the variance in seniors slightly bigger.*

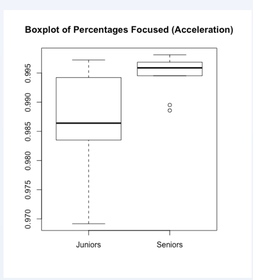


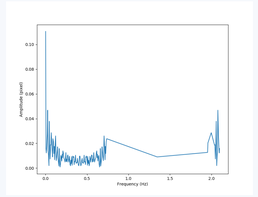
Figure 10: *Boxplot of Percentage Focused calculated using acceleration. The median of seniors is a lot greater than the median of juniors. There is a lot more variance in juniors than in seniors.*

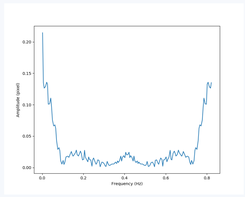
	Velocity		Acceleration	
	Junior Surgeons	Senior Surgeons	Junior Surgeons	Senior Surgeons
Mean	0.9641	0.9826	0.9869	0.9949
Median	0.9652	0.9892	0.9864	0.9959
Standard Deviation	0.0185	0.0129	0.0096	0.0033
IQR	0.0141	0.0142	0.0107	0.0020

Figure 11: *Table showing quantitative data about the two groups, using the velocity and the acceleration thresholds. In both cases, the mean and the median of percentages of focus of seniors is greater than those of juniors. The standard deviation of the juniors is bigger than that of seniors in both cases, which means that there is a lot more variance.*

FAST FOURIER TRANFORM AS MEASURE OF FOCUS

The Fourier Transform is a mathematical function that decomposes a function of time, into the frequencies that make it up. The Fast Fourier Transform was applied on the velocity graphs during the scrub in and during the surgery, aiming to notice relations and differences between the results.





Figures 12 and 13: *Velocity during the scrub in (left) and surgery (right) after the Fast Fourier Transform. In order to see what frequencies make it up, we need to look at the x coordinate of the spikes.*

One of the assumptions of the Fast Fourier Transform method is that the data is evenly distributed over time. However, one issue encountered in comparing equal numbers of sample points is that the lengths of time of the intervals were not equal, since data was selected only if the confidence was greater than 60%.

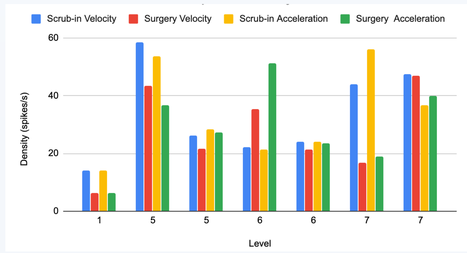


Figure 14: *Graph of Density of Spikes after Fast Fourier Transform was applied to the graphs of velocity and acceleration both during the scrub in and the surgery.*

Density was calculated as number of spikes over time. No correlation between the density in the scrub in and the surgery was observed. We notice that the density is very small for the junior surgeon, but there is not enough data to support this observation.

CONCLUSIONS AND FUTURE WORK

Altogether, there is a correlation between the velocity and acceleration of a surgeon's pupils and how much experience they have: senior surgeons have a higher percentage of focus than junior surgeons. Using the method to determine the ideal threshold, a neural network could be trained in the future to determine if a surgeon is a junior or a senior. Additionally, the plot of the position of the pupils over time shows that surgeons are more focused during the surgery than during the scrub in. On the other hand, no specific difference could be observed between the two using the Fast Fourier Transform, since the data does not satisfy the assumption that it is evenly distributed over time. In order to have statistically significant results, more relevant data should be collected, and samples should be bigger and equal for the two groups.

REFERENCES

1. Whittaker M., Alper M., Bennett C., Hendren S., Kaziunas L., Mills M., MorrisR., Rankin J., Rogers E., Salas M., West S. November 2019. *Disability, Bias, and AI*. 16 p.
2. Puckett Y., Baronia B. *Technical Report of Successful Deployment of Tandem Visual Tracking During Live Laparoscopic Cholecystectomy Between Novice and Expert Surgeon.*
3. Tien T, Pucher PH, Sodergren MH, Sriskandarajah K, Yang G-Z, Darzi A (2015) Differences in gaze behaviour of expert and junior surgeons performing open inguinal hernia repair. 405–413 p.
4. Chaffee I., Beck J., Wang K, Lyford-Pike S., Guy S. 2019 Analyzing Pupil Trajectories in the Operating Room. Velocity as a measure of focus and ability across surgical residency levels.

ACKNOWLEDGEMENTS

Research was supported by Undergraduate Research Scholarship. Special thanks to Sofia Lyford-Pike, Jenna Van Beck, and Zach Chavis.