INTERPRETATION OF HISTORIC STRUCTURE FOR NON-INVASIVE ASSESSMENT USING EYE TRACKING

M. R. Saleem^{a,*}, A. Straus^b, R. Napolitano^a

^aPennsylvania State University, Department of Architectural Engineering, P.O.Box 16801, State College, PA, United States - (rakeh, nap)@psu.edu.

KEY WORDS: Eye tracking, Fixations, Damage detection, Cultural heritage inspection, Diagnostics, Historic structures, Decision making.

ABSTRACT:

With the aims of ensuring safety and decreasing maintenance costs, previous studies in bridge inspection research have worked to elucidate damage indicators and understand their correspondence to structural deficiency. During this process, understanding how an inspector looks at a structure comprehensively as well as how they localize on damage is vital to examining diagnostic bias and how it can play a role in the preservation and maintenance process. To understand human perception and assess the human-infrastructure interaction during the feature extraction process, eye tracking can be useful. Eye tracking data can accurately map where a human is looking and what they are focusing on based on metrics such as fixation, saccade, pupil dilation, and scan path. The present research highlights the use of eye tracking metrics for recognizing and inferring human implicit attention and intention while performing a structural inspection. These metrics will be used to learn the behavior of human eyes and how detection tasks can change a person's overall behavior. A preliminary study has been carried out for damage detection to analyze key features that are important for understanding human-infrastructure interaction during damage assessment. These eye tracking features will lay the foundation for human intent prediction and how an inspector performs inspection on historic structures for existing types of damage. In future, the results of this work will be used to train a machine learning agent for autonomous and reactive decision making.

1. INTRODUCTION

Infrastructure is continuously subjected to the combined effect of material aging, structural degradation, and environmental loads. To monitor this, on-site inspections are required by federal law to ensure the safety and longevity of infrastructure (Gee and Henderson, 2007). Presently, there are three modalities of inspection that are primarily carried out: 1) visual inspection by field inspectors, where they walk around a structure and identify features of concern; 2) pre-programmed flights for the unmanned aerial vehicle (UAV) to capture images of the structure for later analysis; 3) remote inspection by site inspector using remote-controlled UAVs (Dorafshan and Maguire, 2018). Reactive approaches to architecture inspection — at their simplest expression, allow the person or machine to automatically shift when related data require the inspector to investigate a particular feature or structural weakness initially unanticipated. Based on the second modality, little chance for reactive inspection and documentation as the UAV is on a predetermined path, limiting the feedback to the remote inspector. Based on the first and third modalities, which can be reactive, the inspector's knowledge is either dependent on their expertise doing a manual visual inspection or the quality and lag-time of the UAV's camera feed going back to the flight controller. To capture expert knowledge and human-infrastructure interaction, this work will consider the use of eye-tracking.

Eye-tracking technology is an objective tool that detects differences in viewing patterns. The application of eye-tracking as a means of evaluating human behavior has been established in many fields. It has been shown that eye tracking enables us to track and understand human eye movements, and in turn, this information can provide us with cues about a person's focal point and attention. Eye-tracking informs us about implicit human intent by telling where a person is looking, what caught their attention for long periods, and how they shift between different points in a visual scene. The two most crucial eye movements for this type of research are fixations and saccades, which tell researchers about human attention to a visible location (Borys and Plechawska-Wójcik, 2017).

Fixations are when eyes essentially stop scanning about the scene, holding the central foveal vision in place so that the visual system can take in detailed information about what is being observed. Saccades are rapid, ballistic movements of the eyes from one point of interest to another. Human perception is guided by alternating sequences of fixations and saccades (Purves et al., 2001). Due to the fast eye movement during a saccade, the image on the retina is poor quality; thus, information intake happens mainly during the fixation period. Both these elements are beneficial for uncovering patterns and understanding eye movements in natural tasks to see how people coordinate actions in real life. This, in turn, reveals cognitive/perceptual processes and limitations.

For example, when a human inspector scans a structure for damage, they look at the structure dynamically to identify potential defects. While this is occurring, the inspectors are not aware of the electrical signals which cause the contractions in their eye muscles (squinting) and neck (turning) (Palinko et al., 2016). The human brain actively transmits signals to the sensory system through neurons and compensates for any new in-

^bSalamanca City Central School District, 50 Iroquois Drive, Salamanca, NY 14779 - astraus@salamancany.org

^{*} corresponding author

formation using feedback control and common-sense reasoning. While this is something a human can do within a blink of an eye, engineers must know the metrics that are responsible for sense-making and/or decision-making in order to understand the process more fully to quantify human biases and/or control intelligent, automated systems in the future.

In previous work, eye tracking has been used extensively for psychology (Mele and Federici, 2012), cognitive linguistics (Sagarra and Hanson, 2011) and product design (Eyeware, 2019). Recent developments in the eye tracking industry have enabled researchers to study efficient and accurate assessment of human visual attention (Brunyé et al., 2019). Eye movements can help understand a person's gaze pattern by showing where an individual is looking, the duration, and the pattern in which their attention switches from one location to another. Many studies have been conducted using either screenbased (Bylinskii et al., 2017); (Peters et al., 2015) or wearable devices (Hwang et al., 2013) for collecting eye tracking data and performing analyses for human intent prediction. Several researchers have explored the use of eye tracking technology to augment hazard recognition and examine worker activity for construction sites (Xu et al., 2019); (Jeelani et al., 2019). Furthermore, (Helmert et al., 2017) used eye tracking to analyze consumer behavior regarding visually suboptimal food items by altering design and colors to manipulate attention, cognitive process, and purchase decision making.

This paper aims to elucidate correlations between standard eyetracking metrics and human-based infrastructure inspection in consideration of previous work. To do this, a case study was carried out on the Palmer Museum of Art at Pennsylvania State University using Tobii eye-tracking technology. Statistical analysis of the experimental results shows the difference between the two inspection procedures and how they correspond to different damage. Documenting reactive inspection procedures will be necessary for future work in understanding cognitive and diagnostic biases in the human inspection process and capturing expert knowledge to automate specific inspection processes

2. EYE TRACKING METRICS

Eye trackers help assess a person's visual attention and human intent by analyzing their eye movements. Eye movements include where a person is looking, how long they are fixated, and their gaze pattern and come together to define how a person's attention switches between different elements. We will briefly explain some of the most common eye-tracking metrics used for research based on eye movements and pupillary responses (Sharafi et al., 2015):

- 1. **Fixations:** are periods when the eyes are stationary and stable. The average fixation time is 200-300 msec.
- Fixation duration: is the amount of time a person is fixated on a specific point while looking at an object or scene. Longer fixation duration can show a person's interest or lack of understanding about a particular task in an observation.
- 3. **Fixation count:** is the number of times a person looks back at a specific fixation location.
- 4. **Saccades:** are the rapid eye movements that occur between consecutive fixations. These are voluntary and can last about 30-50 msec on average. The directionality of saccades is useful in understanding the relationship between areas of interest (AOI).
- 5. Visit count: is the number of times a person looks back to the same fixation point in a specific AOI. The higher visit count indicates the significance of the AOI that caught more attention of the person.
- Visit duration: is the amount of time a person visits a specific AOI while looking at an object or scene. Longer visit duration can show a person's lack of understanding leading to long and repetitive visits.
- 7. **Scan path:** is the summation of fixations and saccades and represents a person's eye movement pattern.
- 8. Pupil gradient: indicates the dilation of pupils; this varies depending on the person's focus and lighting conditions. A change in a person's interest can be analyzed by pupil size as it indicates person's cognitive effort.

In addition to these metrics, several studies discuss how bias can be induced in measurement and can occur due to various reasons, regardless of a person's intent. These biases can significantly affect the decision-making of an inspector while enacting engineering judgment. These can include attentional bias, fixation bias, cognitive bias, and negativity bias. While these forms of bias will be critical to a comprehensive understanding of the human-infrastructure interface, they will not be examined in this work; this will be left to future studies.

3. METHODOLOGY

In this study, we used a screen-based eye-tracking system to understand the importance of inspector perception during reasoning tasks. The overall research workflow is shown in Fig. 1, which consists of three different phases: 1) Capture: eye tracking is used to capture data for documentation; 2) Understand: eye-tracking metrics will be analyzed to assist in understanding the collected data and latent features in that data set (intent, bias), 3) Analyze: these collected data can be analyzed to see how these metrics corresponds to human attention and focus

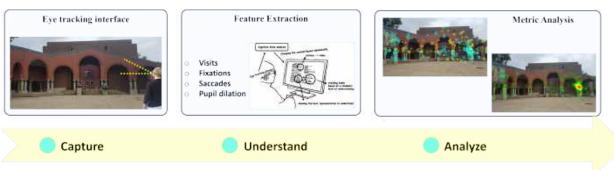


Figure 1. Overall research workflow of eye tracking



Figure 2. Experimental setup of Screen-based eye tracker

and get an intuition about implicit intent and bias. To this end, we will perform an exploratory study on cultural heritage museums and collect preliminary data to understand human interaction within a damaged infrastructure environment.

3.1 Device for measuring eye movement

To measure human eye movement and the pupillary response, we used the screen-based tracker Pro Nano by Tobii technology for eye tracking (Tobii Pro, 2014). Tobii Pro Nano (see Fig. 2)) is a second-generation screen-based tracker with high sampling frequency and low latency. The technical specifications are mentioned in Table 1. Screen-based trackers are significant for data collection with high accuracy and precision due to less head and body motion. In this study, we are trying to learn about human intent while doing damage detection on a structure for identifying and quantifying any potential damage. The traditional metrics for analyzing eye movements are based on fixation, saccades, and pupil diameter; however, Tobii Pro Lab (Tobii Pro AB, 2014) software, in addition to these, calculates visits, interval and event metrics on an area of interest (AOI). Table 2 discusses the metrics that Tobii pro Lab helps us analyze by providing quantitative measures for performance evaluation based on eye-gaze data.

4. CASE STUDY ON THE PALMER MUSEUM OF ART

A preliminary case study was conducted for damage assessment using screen-based eye trackers on Palmer Museum of Art, as shown in Fig. 3. The University Art Museum was built on October 7, 1972, located on the University Park campus in State College, Pennsylvania. The building is around 50 years old, and lack of maintenance led to some damage on the building's outer envelope. In this study, we examined the outer envelope

Table 1. Tobii pro nano technical specifications. (Tobii Pro, 2014)

Characteristics	Technical specifications
Eye tracking technique	Video-based pupil- and corneal reflection eye tracking with dark and bright pupil illumination modes
Sampling frequency	60 Hz
Precision	0.10° RMS in optimal conditions*
Accuracy	0.3° in optimal conditions*
Binocular eye tracking	Yes
System latency	1 frame (17ms)
Blink recovery time	1 frame (immediate)
Gaze recovery time	250ms

^{*} Using built-in filtering

and study each participant's behavior when they are informed to do an inspection. Some of the prominent damage are the result of environmental effects and they have caused such as cracks in the columns, efflorescence on the masonry, and biological growth in masonry joints as depicted in Fig. 3. In addition to these, dark spots (gypsum crust) are also prominent on the brick facade which could lead to structural deterioration. These damage are essential to diagnose to ensure proper restoration of the building and preservation.

To capture static eye movements in an indoor environment while looking at a computer screen, we analyzed multiple eye-tracking metrics (visit count, fixation, pupil dilation) for the three different types of damage described above to understand how an inspector behaves under these conditions. The data shown to participants was in the form of a video clip (8-minute duration), which covers sections of Palmer Museum's outer building envelope with no pre-processing. The participants were asked to view the building structure in the way an in-

Table 2. Different types of quantitative metrics for eye-tracking.

Quantitative measures	Definitions
Fixations based metrics	
Fixation count	The total number of fixations made within the AOIs
Fixation duration	The total time for which an individual has fixated the gaze for an AOI
Average fixation duration	The average duration of the fixations within each AOI
Visit based metrics	
Visit counts	The number of visits that are made within the AOI
Visit duration	The total time an individual has visit an AOI
Average visit duration	The average duration of the visits within each AOI
Eye movement types	The type of eye movement classified by the fixation filter for gaze data i.e., fixation, saccade or unclassified
Pupil diameter	The estimated size of the pupils when viewing a stimulus
Mapped gazed point	The combined assisted and manually assisted gaze point coordinates
Assisted mapping gaze point score	Similarity score of the assisted mapping gaze points
Scan path	Spatial arrangement of the sequence of fixations during visualization of the stimuli
Heat map	Display heat map of an image where the fixation values are represented as color-coded, the highest point colored in red (default color) represents high fixation count



Figure 3. Palmer Museum of Art at University Park

spector would look for any potential damage.

For this study, we considered two participants' data. The participants were advanced students in architectural engineering with prior knowledge about different types of building damage. Participant 1 is a third-year undergraduate male in his early twenties. This participant has no prior working experience but general knowledge of diagnostics and damage assessment. Participant 2 is a first-year Ph.D. male in his mid-twenties. This participant has prior experience developing convolutional neural networks to detect cracks in concrete bridge piers automatically. Pro Nano was calibrated using a 9-point calibration technique in Tobii Pro Lab software to ensure accuracy and precision without any need for additional pre-processing on the collected video data which could lead to bias. The participants were given a prompt to do a building diagnostic and assess the structure if any damage caught their attention.

5. RESULTS AND ANALYSIS

5.1 Data processing and fixation classification

Since the viewing patterns vary widely among different participants and research fields as well as since the characteristics in the visual scene under observation vary greatly, it becomes difficult to determine the parameter values to measure that are optimal for most use case scenarios in eye tracking research. Tobii pro lab features I-VT (Velocity-Threshold Identification) fixations, attention, and raw filter (Olsen, 2012) in addition to custom filter which can be tuned depending on the settings from the the users to be able to use it in either default or custom settings for fixation classification. When selecting the I-VT filter parameters, one needs to ensure that the settings are as optimal according to the specific situations at hand. One of the challenges in doing visual inspection is that there is a lot of head and body movement that can affect the data collection. However, for an indoor experiment, its the opposite—less head and

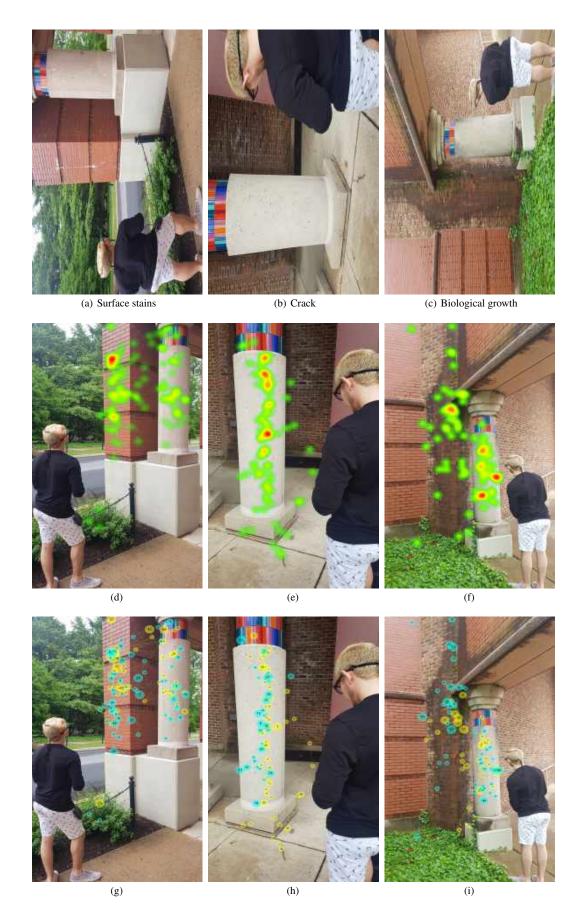


Figure 4. Heatmaps and gaze plots for different sections of building structure; (a-c): Original images; (d-f): Aggregated heatmaps; (g-i): Aggregated gaze plots—'Yellow' indicates participant-1 and 'Cyan' indicates participant-2

body movement leads to more stable gaze samples which can yield better results. For those different cases, we must choose the attention or fixation filter that fits the given experimental design requirements.

Many different algorithms have been used to identify the eye movements such as saccades or smooth pursuits (Komogortsev et al., 2010), and depending on which kind of eye movement is of interest, the different classification algorithms are more or less suitable. In this study, we have selected fixation filter as we are more interested in understanding where a person gaze is getting fixated and what grabs their focus while assessing a structure. The I-VT fixation classification algorithm is a velocity based classification algorithm (Salvucci and Goldberg, 2000) based on the velocity of the directional shifts of the eye measured in visual degrees per second ($^{\circ}/s$). If it is above a certain threshold for the I-VT filter, it is classified as a saccade and below it is seen as part of a fixation.

5.2 Experimental results

For this study, the data was collected from the two participants described above. Fig. 4 shows a comparison of heatmaps and gaze plots for different parts of the structure. They provide fast, qualitative information about how each inspector viewed different components of the structure. Each heatmap (see Fig. 4(d)-4(f)) shows the fixations of individual participants and helps in understanding the overall correlation of their gaze; here, red indicates long fixation periods, where green represents short fixation periods. Similarly, the gaze plot (see Fig. 4(g)- 4(i)) indicates the participant's eye patterns and path while looking at the structure by moving their eyes from one element to another. The yellow color gaze trail indicates Participant 1 data, and the cyan color gaze trail indicates Participant 2 data collected through eye trackers. The numbers in the circles represent the order of the gaze from one fixation to another. It is observed that the gaze plot of both participants varies, which could indicate differences in their diagnostic reasoning capabilities and understanding of the test stimulus while looking for damage. Participant 1 seems to scan the overall structure with a stable gaze close to potential damage and defects in a structure, while Participant 2's gaze path indicates a thorough and somewhat less concise assessment. In Fig. 4(g), it can be seen that Participant 1 is fixing their eye gaze at white spots of efflorescence and minor cracks in the column while Participant 2's gaze indicates some randomness indicating difficulty in finding damage. Similarly, Fig. 4(h) indicates both participants got their attention fixed at the black spots (gypsum crusts and other surface staining) and on significant cracks running vertically across the column. Lastly, Fig. 4(i) indicates a portion of the structure that contains biological growth (algae) on the brick wall and a lot of black surface stains. These defects indicate water leakage that results in creep and water infiltration into the building facade. Both participant gaze indicates their fixations on the algae and some minor cracks present on the column supporting the girder. It is depicted that Participant 2 is thoroughly observing the structure while Participant 1 gaze points are only where the defects have occurred.

Further, it can be argued that these results have some bias due to the quality of the video data collected, and some room is present for bias due to noise and jitters. While these results indicate both participants' fixations and gaze directions how they shifted from one point in time to another, it is necessary to perform metric analysis. To better make sense of the data, bias, or



Figure 5. AOI labeling on visual stimuli for damage category

engineering judgment that was occurring, additional analysis was carried out considering fixations and visits in detail, as described in the next section.

5.3 Metric analyses and discussion

For more rigorous analysis, one pair of the structural columns on the front of the Museum were taken into further consideration. The structural columns, as shown in Fig. 5, were marked into three main regions denoting damage typology: crack, white surface staining, and black surface staining. The three AOIs were segmented as potential damage regions for inspection and the metrics under consideration are eye fixation and number of visits. Fixation and visit metrics were analyzed in detail corresponding to the marked AOIs for both participants. The analytical results of fixations count and duration as well as visit count and duration are shown in Fig. 6 and Fig. 7, respectively.

Considering fixation duration and count (see Fig. 6), Participant 1 inspected the structure overall with all the three regions marked as defects. The average fixation duration of participant 1 showed greater attention towards the black staining and slightly less but significant fixation on the cracks in the columns in comparison to the white staining that occurs at the top support of the columns. Participant 2 spent more time on the crack damage and black surface stains with longer fixations and little to no time on white surface change defects. The time distribution shows an average fixation of 30 sec for crack damage, 6.5 sec for black surface stains due to gypsum crust damage, and 21.5 sec for white stain damage due to calcium carbonate leeching out the surface, respectively. Similarly, Fig. 7 shows visit metrics and visit duration for the two participants. As a reminder to the reader, the difference in fixation and visitation is that fixation can be any point in the visual scene while visit corresponds to annotated AOIs under consideration. It can be deduced from the results that Participant 1 visited white surface stains due to calcium carbonate leeching many times for longer durations; paired with the gaze plot (see 4(h)), which shows more gaze points at the top and the back-and-forth movement

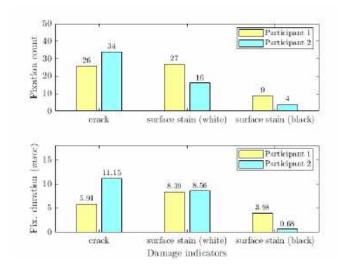


Figure 6. Comparison of fixation count Vs. total fixation duration per AOI

of eyes, this data indicates that the participant was trying to compare the damage regions. Alternatively, Participant 2 spent more time on crack damage with an average visit duration of 1.03 sec for each fixation and did not show evidence of comparison between AOI.

These results confirm the qualitative plots above and show how an inspector performs their background biases assessments (i.e., someone who has worked explicitly with damage inspection focused on thorough assessment versus someone who has no prior work experience struggles to assess the critical damage and focus on irrelevant data. Thus, the background can alter the focus of an inspector, which can cause bias during an inspection. Further investigation into how training and previous experience affect fixation during an inspection will be vital information to have for both training new inspectors and training robots in a way that will minimize bias.

6. CONCLUSION

In this paper, we study and analyze eye-tracking metrics to identify human-gaze patterns and their correspondence to visual attention while performing the structural inspection of a cultural heritage building. A pilot study was conducted using a test structure to understand how inspectors with diverse backgrounds perform inspection and analyze their performance using fixation and visit metrics. From the results and analysis, we confirm how the performance of an inspector can be biased depending on their prior work knowledge and experience, which can mislead their focus point. This prior work knowledge and experience is crucial for engineers as they rely on the inspector's abilities while making an engineering judgment and perform decision-making. Therefore, understanding how training and previous knowledge prune inspector abilities and affect fixation is essential for future trainees and robots in an attempt to minimize bias.

This pilot study serves as the basis of preliminary data, which can be further analyzed to build a multi-agent model to train drones for inspection and diagnostics of structure. It will help in making the heritage documentation procedure autonomous, less-intrusive, and accurate to detail. In the future, we will extend the number of participants and conduct experiments on a

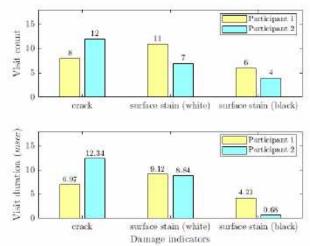


Figure 7. Comparison of visit count Vs. total visit duration per AOI

small-scale structure with a trained model for damage detection to understand human-infrastructure interaction better.

ACKNOWLEDGEMENTS

The authors of this work would like to thank the undergraduate student, Mr. Joseph Yousef who helped in collecting experimental data and providing their thoughts on the pilot study and the PSU Office of Physical Plant and Mr. Anand Swaminathan for providing information about locations on campus where we could perform our initial experiments. We are grateful for their continued help with these projects.

REFERENCES

Borys, M., Plechawska-Wójcik, M., 2017. Eye-tracking metrics in perception and visual attention research.

Brunyé, T. T., Drew, T., Weaver, D. L., 2019. A review of eye tracking for understanding and improving diagnostic interpretation. *Cognitive Research: Principles and Implications*, 4(7). https://doi.org/10.1186/s41235-019-0159-2.

Bylinskii, Z., Borkin, M. A., Kim, N. W., Pfister, H., Oliva, A., 2017. Eye fixation metrics for large scale evaluation and comparison of information visualizations.

Dorafshan, S., Maguire, M., 2018. Bridge inspection: human performance, unmanned aerial systems and automation. *Journal of Civil Structural Health Monitoring*, 8(3), 443–76. https://doi.org/10.1007/s13349-018-0285-4.

Eyeware, 2019. Understanding eye tracking how it can work for you: Definitions, metrics, and applications.

Gee, K. W., Henderson, G., 2007. Highway bridge inspections. https://www.transportation.gov/testimony/highway-bridge-inspections.

Helmert, J. R., Symmank, C., Pannasch, S., Rohm, H., 2017. Have an eye on the buckled cucumber: An eye tracking study on visually suboptimal foods. *Food Quality and Preference*, 60, 40-47. https://doi.org/10.1016/j.foodqual.2017.03.009.

- Hwang, B., Jang, Y.-M., Mallipeddi, R., Lee, M., 2013. Probing of human implicit intent based on eye movement and pupillary analysis for augmented cognition. *International Journal of Imaging Systems and Technology*, 23(2), 114-126. https://doi.org/10.1002/ima.22046.
- Jeelani, I., Albert, A., Han, K., Azevedo, R., 2019. Are Visual Search Patterns Predictive of Hazard Recognition Performance? Empirical Investigation Using Eye-Tracking Technology. *JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT*, 145(1). https://10.1061/(ASCE)CO.1943-7862.0001589.
- Komogortsev, O. V., Gobert, D. V., Jayarathna, S., Koh, D. H., Gowda, S. M., 2010. Standardization of Automated Analyses of Oculomotor Fixation and Saccadic Behaviors. *IEEE Transactions on Biomedical Engineering*, 57(11), 2635-2645. https://10.1109/TBME.2010.2057429.
- Mele, M., Federici, S., 2012. Gaze and eye-tracking solutions for psychological research. *Cognitive Processing*, 13, 261-265. https://doi.org/10.1007/s10339-012-0499-z.
- Olsen, A., 2012. The Tobii I-VT Fixation Filter. 57(11), 21. http://www.vinis.co.kr/ivt $_filter.pdf$.
- Palinko, O., Rea, F., Sandini, G., Sciutti, A., 2016. Eye Tracking for Human Robot Interaction. 327–328. https://doi.org/10.1145/2857491.2888590.
- Peters, J. R., Surana, A., Bertuccelli, L., 2015. Eye-tracking metrics for task-based supervisory control.
- Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O., , Williams., S. M., 2001. Types of eye movements and their functions. *Neuroscience*.
- Sagarra, N., Hanson, A. S., 2011. Eyetracking Methodology: A User's Guide for Linguistic Research. *Studies in Hispanic and Lusophone Linguistics*, 4(2), 543–556. https://doi.org/10.1515/shll-2011-1113.
- Salvucci, D. D., Goldberg, J. H., 2000. Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the 2000 Symposium on Eye Tracking Research amp; Applications*, Association for Computing Machinery, New York, NY, USA, 71–78.
- Sharafi, Z., Shaffer, T., Sharif, B., Guéhéneuc, Y.-G., 2015. Eye-tracking metrics in software engineering. 2015 Asia-Pacific Software Engineering Conference (APSEC), 96–103.
- Tobii Pro, 2014. Tobii pro nano.
- Tobii Pro AB, 2014. Tobii pro lab. Computer software.
- Xu, Q., Chong, H.-Y., chao Liao, P., 2019. Exploring eyetracking searching strategies for construction hazard recognition in a laboratory scene. *Safety Science*, 120, 824-832. https://doi.org/10.1016/j.ssci.2019.08.012.