Understanding Visual Scanning Behavior in Driving: A Review and a New Perspective Using Statistical Pattern-Based Approach

Xingjian Ma
Department of Electrical and Computer Engineering
University of Florida, Gainesville, Florida, 32603
Email: xingjianma@ufl.edu

Minhee Kim
Department of Industrial and Systems Engineering
University of Florida, Gainesville, Florida, 32603
Email: mkim3@ufl.edu

Haolan Zheng
Department of Industrial and Systems Engineering
University of Florida, Gainesville, Florida, 32603
Email: haolan.zheng@ufl.edu

Wayne C.W. Giang
Department of Industrial and Systems Engineering
University of Florida, Gainesville, Florida, 32603
Email: wayne.giang@ise.ufl.edu

Word Count: 7224 words + 1 table (250 words per table) = 7474 words

Submitted [07/31/2023]

ABSTRACT

Eye tracking measures, such as gaze locations, play a vital role in understanding the visual attention of drivers, providing insights into an individual's driving skills and fitness to drive. Numerous studies have used eye-tracking data to analyze driving behaviors under various scenarios. Yet, these methods often either rely on aggregated or oversimplified representations of the eye-tracking data or require significant computational resources for long-term driving scenario analysis. To address these, this paper reviews existing eye-tracking data analysis methods and presents a new statistical pattern-based analytics approach. The proposed method extracts unique and representative visual scanning patterns of drivers by leveraging a novel time series data analytics technique, namely distance profile. The proposed method has several advantages in analyzing eye-tracking data from driving scenarios: (i) computational efficiency to analyze long-term driving scenarios; (ii) minimization of manual pre-processing required to simplify or summarize eye-tracking data; and (iii) pattern extraction to provide an intuitive and detailed understanding of driving behaviors. This new method seeks to complement and enrich conventional eye-tracking data analysis by offering a novel and detailed view of the data. A case study on the real-world eye-tracking data is further conducted, where we demonstrate the feasibility and applicability of the proposed approach, laying a strong foundation for future research in driving behavior analysis.

Keywords: Driving behavior, Eye movement, Eye-tracking data analysis, Visual scanning patterns, Pattern extraction, Scan path

INTRODUCTION

While driving is a routine activity that has been extensively researched, a large number of traffic fatalities (*1*), as well as the popularization of non-driving related activities and distractions in the vehicle have required ongoing efforts in understanding and facilitating safe driving practices. In various applications, including distracted driving (*2*), vehicle technology design and evaluation (*3*, *4*), and vehicle technology training (*5*, *6*), comprehending driver behavior is a crucial step in assessing driving performance in terms of safety and effectiveness.

One important tool in studying driver behavior is the driver's visual attention measured through eye tracking technology (*7–10*). Most driving studies use visual attention towards Areas of Interest (AOI), a representative area or object in the visual field, e.g., windshield, right mirror. AOIs can be specified before data collection or designated afterward to capture emergent items of interest and are typically created based on semantic information about the stimulus (*11*). AOIs provide various metrics such as the number of glances, glance durations, and percentage of time spent within each AOI. However, the application of AOIs may not always be optimal for analyzing driving tasks that involve dynamic scene changes. Specifying moving AOIs can be labor-intensive if conducted manually, or it may require specialized equipment or algorithms with limited applicability. Furthermore, there is a loss of data fidelity, as using large areas may not effectively capture the fine-grained details of gaze movements. For instance, considering the windshield as a single AOI could potentially fail to capture a driver's gaze transition from a preceding vehicle to a traffic sign, as such gaze transition occurs within this one AOI.

Driver visual attention scan paths are the particular sequence that the eyes travel over a scene that contains information about how people see. Scan paths may provide additional insights into the strategies that individuals use while driving. A few studies have examined driver scan paths, primarily using a sequence of AOIs. Braunagel et al. extracted a Driver-Activity Recognition (DAR) architecture using dynamic clustering and symbolic aggregate approximation patterns based on the sequence of AOIs (*12*). Navarro et al. examined different AOIs along scan paths to investigate sequences of visual scanning in manual and highly automated simulated driving (*13*). While several methods have been developed to handle scan paths without pre-defining AOIs in non-driving domains, such as education (*14*, *15*) and clinical medicine (*15*), the direct application of these methods to driving tasks presents a unique challenge. The reason is that processing large amounts of gaze position data from driving tasks can be computationally expensive, given that eye-tracking data tends to be longer than that from non-driving tasks.

To address these challenges, this paper presents a new statistical pattern-based analytics approach, designed to automatically extract meaningful visual scanning patterns of drivers that: 1) does not depend on pre-defined AOIs, 2) is computationally efficient, and 3) can be applied to understanding fundamental driving research questions. The extracted visual scanning patterns serve as valuable tools for researchers seeking to comprehend various driver behaviors influenced by factors like fatigue, gender, or age. By doing so, it can contribute to the training and education of drivers' safety awareness and enhance their driving performance. This paper provides a review of existing eye-tracking analysis methods and focuses on examining the technical aspects and feasibility of the newly proposed approach, along with its potential applications, without numerically comparing it with existing approaches. The paper aims to present an initial idea to showcase the viability of this innovative approach for validation and trial purposes to support the development of future methods for comparing and classifying drivers' visual attention behavior.

RELATED WORK

Visual Attention and Eye Movement Measures in Driving

There exists a rich literature on the analysis of visual attention and eye movements in driving tasks. Common visual attention analysis measures used in driving and their definitions are shown in **Table 1** (*4*). Fixation is a widely used measure in driving-related tasks that are characterized by their position (*10*, *16*) and duration (*10*, *17*). For example, Wang et al. conducted a study where they utilized the time to first fixation and first fixation duration to investigate the design of a head-up display (*18*). The transition time

(*10*), the distance between fixation, and dwell time (*18*) have been employed to examine the correlation between various fixations. Total glance duration and percent of glance duration are used to understand drivers' eye behavior before lane changing (*19*). The number of AOIs (*17*, *20*) or the frequency of AOI visits (*13*) is commonly employed to differentiate scan paths at the semantic level. Visual transition probability between different AOIs and visual stationary probability of different AOIs can also show how much attention the driver pays to different semantic information (*21*). Saccades, which is the brief movement of the eyes between fixations (*4*, *22*), have also been used to understand visual scanning behavior. Ahmad et al. examined the number and duration of saccadic eye movements to study the impacts of illuminance on eye gaze movement during driving (*23*).

TABLE 1 Visual Attention Analysis Measures Widely Used in Driving Tasks (*4*)

Terms	**Definitions**	**Examples from literature**
Direction of gaze	The AOI to which the eyes are directed.	(*10*, *16*)
Dwell time	The sum of all consecutive fixations and saccades within the AOI between transitions to other AOIs.	(*18*)
Glance/Glance duration	The maintaining of visual gaze within an AOI, bounded by the perimeter of the AOI; comprised of at least one fixation and a transition to or from the AOI.	(*10*, *17*)
Glance frequency	The number of glances to an AOI within a sample interval where each glance is separated by at least one glance to a different AOI.	(*13*, *17*, *19*, *20*)
Glance location probability	The probability that the eyes are fixated at an AOI (or set of related AOIs) during a sample interval. This would be defined as the number of glances to an AOI divided by the number of glances to all AOIs in the sample.	(*21*)
Link value probability	The probability of a glance transition between two different locations. The link value probability between AOIs A and B is defined as the number of glance transitions from A to B plus the number of glance transitions from B to A divided by the total number of glance transitions between all pairs of locations in the sample interval time.	(*21*)
Total Eyes Off Road Time	The summation of all glance durations to all AOIs other than the road scene ahead during a sample interval.	(*20*)
Total glance time	The summation of all glance durations to an AOI (or set of related AOIs) during a sample interval.	(*19*)
Transition time	The duration of a transition.	(*10*)

Statistical methods and machine learning algorithms are often used to further analyze the frequency of various measures, find the statistical difference among subjects, or classify and categorize multiple subjects. Wang et al. examined the effect of animation and borders on several measures related to visual warnings through generalized linear models (*18*). Lethaus et al. performed a Markov analysis on the transition probabilities and percentage of fixations to explore the specific maneuvers employed by individuals during driving (*24*). Muttart et al. compared the glancing behavior of motorcyclists and car drivers at intersections using statistical methods (*25*). Deng et al. used the acquired environment data and measures as features, and applied the Random Forest (RF) algorithm to distinguish driving behaviors (*9*). Brishtel et al. used multiple measures as features to train the machine learning model to classify driving modes (*10*).

While these measures provide an overview of a driver's visual attention, they often do not provide a temporal dimension to the analysis. In the next subsection, we will focus on the scan paths, which will allow us to not only understand where the drivers are looking but also the sequence in which they process the visual information.

Scan Path Analysis

Scan paths, defined as the trajectories (paths) of the eyes when scanning the visual scene, stand out among various measures as a means to explain the dynamic sequential information of eye movement. To form a foundation for the proposed method, this subsection will review established scan path-based methods, including those not explicitly designed for driving tasks.

In real-world applications, it is crucial to accurately compare different scan paths and measure their similarities. This is often done by extracting diverse features, including vectors, directions, lengths, positions, and durations (*26*). Scan path-based methods can be broadly categorized into two main groups, depending on whether they simplify scan paths through discrete representations or not.

Discrete Representation-based Methods

Existing methods often simplify scan paths by transforming the fixation sequences into string-based (discrete) representations to achieve better computational efficiency. This can be done by defining discrete AOIs onto the stimulus space or using symbolic representation approaches (e.g., SAX representation) to assign each fixation to a character based on its location.

One commonly utilized method to compare simplified scan paths is the Levenshtein distance (*27*). However, this method has limitations regarding the ordinal information and the spatial position of fixations, as two spatially close fixations can be assigned to different characters. ScanMatch is a more advanced adaptation of Levenshtein distance using the Needleman–Wunsch algorithm (*28*), which addresses this issue and incorporates the duration of the fixations in aligning and comparing scan paths.

SubsMatch examines the frequency of exploratory eye movements and attention shifts, which is useful for comparing scan paths in interactive scenarios like simulated driving (*12*, *29*). SubsMatch utilizes the SAX representation where each fixation is assigned to a character based on the slice of equi-probabilistic data it falls into. The number of occurrences of each substring is stored in a hash table to compute the dissimilarity between two scan paths.

SubsMatch 2.0 adds machine learning techniques to SubsMatch, by classifying eye movements based on sequence-sensitive features extracted from scan paths (*30*). The frequencies of substrings are converted to the N-Gram feature and indexed to create large and sparse matrices of feature counts, which are normalized to represent feature occurrence frequencies. The method utilizes a support vector machine (SVM) for classification. Both SubsMatch and SubsMatch 2.0 were applied to the eye-tracking data from driving sessions to evaluate the fitness to drive.

A similar approach to SubsMatch 2.0 is MinHash (*31*). After aggregating N-Gram scanned subsequences in a dictionary, this method compares the minimal response of a set of hash functions over the dictionary, which involves repeated comparisons of randomly selected subsequences. The resulting frequencies are then utilized to estimate the Jaccard index, measuring the similarity between two scan paths. Unlike SubsMatch designed for dynamic scenarios, this approach is widely used in clustering and classifying documents.

Despite their usefulness, the main limitation of these methods lies in the process of discretizing fixations into characters as the exact spatial locations of fixations are not preserved. This implies that the performance of further analysis will be sensitive to the quality of discretization results which often require careful parameter tuning and might not be applicable across various types of eye-tracking data.

Continuous Representation-based Methods

In contrast to the methods that rely on discretization or string-based representations of scan paths, some approaches directly leverage fixation locations (e.g., X and Y coordinates). Their goal is to find an optimal mapping to match the closest neighboring fixations between two scan paths. The mapping results are then used to quantify the similarity between scan paths or identify patterns that are common or unique to specific groups.

Eyenalysis relies on a geometric representation of eye-tracking data (*32*). The method aims to establish mappings between fixations in one scan path and at least one fixation in another scan path, aiming to minimize the normalized sum of distances associated with all mappings. This summation serves as an overall similarity measure between two scan paths. This approach also captures fixation duration information. However, the main challenge is assigning proper weights to individual dimensions.

MultiMatch captures measures of shape, direction, and length of a scan path in addition to fixation position and duration (*26*). After being simplified through the clustering of fixations within a given directional threshold, the scan paths are temporally aligned based on their shape, which reduces the sensitivity to small temporal or spatial variations. Shape, length, position, and direction similarities are calculated and averaged over scan paths. MultiMatch provides a comprehensive evaluation of scan path similarity as each measure captures a unique component of the scan path. Yet, it remains challenging to determine which measure, or a set of measures, is most appropriate in a given scenario.

In literature, limited research has been dedicated to studying the spatial trajectory of individual fixations movement. In addition, most existing methods have high computational costs, making them unsuitable for long-term driving scenario analysis. There is a need for further research that specifically addresses these issues.

PROPOSED METHOD

Definition of Visual Scanning Patterns

As reviewed in the previous section, most of the existing studies have used aggregated or simplified measures directly obtained from the eye-tracking data. These measures are limited in providing fine-grained information about driving behaviors. The aggregation of information in these measures may mask some of the differences between groups of drivers that may be useful for classification and categorization. Moreover, while scan paths provide additional sequential information about the driver's gaze movements, many of the existing methods are limited in measuring the similarity between two scan paths in driving tasks, where the length of the scan path is relatively long, and a majority of the data represents routine driving without notable stimulus events.

To address these problems, the proposed work will directly use gaze locations (continuous), instead of relying on subsequences of AOIs (discrete), and focus on "patterns" that appear throughout the scan path. Specifically, this work defines patterns as unique and representative visual scanning behaviors of drivers, e.g., checking the speedometer, gazing at the right side of the road, and seeking to identify and locate. Given that the gaze movements can be viewed as two-dimensional time series data (i.e., X and Y coordinates of gaze locations), the defined patterns correspond to short subsequences within the time series data. These subsequences should be significantly different from the common states of the entire sequence, e.g., looking at the forward center, and thus characterize distinct visual scanning behaviors.

The distinctiveness (uniqueness) of a subsequence can be straightforwardly calculated by Euclidean distances between this subsequence and all other subsequences in the entire time series. A subsequence whose sum of these Euclidean distances is large will be selected as a pattern. Nevertheless, given the considerable length of time series data in driving tasks, it poses a significant challenge to the scalability and computational efficiency of the method. In the next subsection, we will introduce novel time series data analysis concepts used in the identification of the defined patterns and addressing these computational issues.

Distance Profile

Our approach is built upon the algorithm designed for all-pairs-similarity-search of time series subsequence (*33*). The primary goal of this algorithm is to compute the dissimilarities (Euclidean distances) between all pairs of subsequences within a given time series.

Formally, a distance profile D_i is defined as a vector of the Euclidean distances between the ith query subsequence and each subsequence in the entire time series. Suppose $T_{i,m}$ is a subsequence of a time series T of length m starting from position i, i.e., $T_{i,m} = [t_i, t_{i+1}, \dots, t_{i+m-1}]$, where $T = [t_1, t_2, \dots, t_n]$. Given a query subsequence, we can "slide" a window of length m across T, to generate all $n - m + 1$ subsequences of length m to be compared with the query. We compute the Euclidean distances between the query subsequence (highlighted as red in **Figure 1**) and each subsequence of T (highlighted as orange in **Figure 1**) to form a distance profile D_i (last row in **Figure 1**). This process results in a collection of pairwise Euclidean distances for each query subsequence, forming a distance profile. The computational cost to obtain distance profiles can be significantly reduced by using the MASS algorithm, which efficiently produces all the distances between the query to the subsequences of an entire time series through the Fast Fourier Transform (FFT) (*34*).

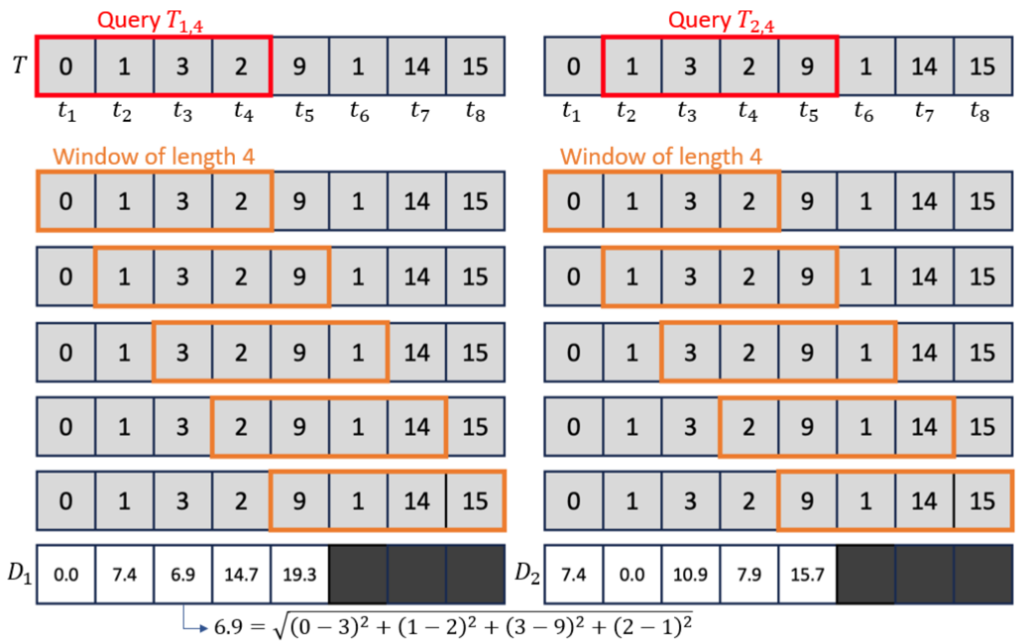


Figure 1 Schematic diagrams of distance profile computation. The red and orange boxes represent the subsequences to be compared. The last row represents the distance profile consisting of the Euclidian distances obtained from the comparison between the query and each subsequence.

In the context of the proposed approach, if the sum of the values in the distance profile of a given query subsequence is large, it indicates that this subsequence is unique (i.e., significantly different from other subsequences in the entire time series). For multi-dimensional time series data, such as gaze location data with X and Y coordinates, separate distance profiles can be created and then combined, accounting for all dimensions. While distance profiles have been studied in identifying meaningful patterns and anomalies within the time series data (*35*), such as pedestrian counting data or electrical power demand data, to the best of our knowledge, it has not been explored to study visual scanning patterns.

Visual Scanning Pattern Analysis

The previous section elaborates on how the distance profile can be used to identify unique visual scanning patterns (i.e., distinct subsequences) from the eye-tracking data. Here, we will explain how the proposed method leverages this to iteratively identify multiple patterns and their similar occurrences.

In each iteration, the first step is to find a prototypical pattern. This can be done by finding the query subsequence with the largest sum of the distance profile. The second step is to locate similar occurrences of the identified prototypical pattern. A small distance value between a subsequence and the prototypical pattern indicates that this subsequence could be another occurrence of the prototypical pattern. In other words, we locate small values in the distance profile of the prototypical pattern to find its similar occurrences. A subsequence whose distance from the prototypical pattern is below a certain threshold is considered as another occurrence of the same pattern. The threshold can be chosen using the quantile value of the distance profile (e.g., the 5th percentile of the distance profile). After identifying the pattern and its occurrences in the time series, we replace these segments with NA (i.e., not applicable) values before proceeding to the next iteration and repeating these steps. This is to avoid identifying another occurrence of the kth pattern as the $k+1$th pattern.

In the numerical study, we also observe that dynamically adjusting the threshold value over iterations results in better performance. Specifically, as we proceed from the patterns with larger movements to those with smaller movements, we can incrementally increase the percentile over iterations, e.g., linearly raising it from the 1st percentile to the 10th percentile over 10 iterations. As we proceed, the increasing threshold allows us to be more lenient in identifying similar occurrences of a given pattern.

NUMERICAL STUDY

To illustrate the application of this method to the driving data, we present the results drawn from a single participant's drive from an experiment. This section includes the visualization of the obtained patterns and the corresponding scan paths and presents derived insights into driving behaviors.

Data Set and Data Collection

In order to evaluate the feasibility of the proposed approach, eye-tracking data from a previous driving simulator study (*5*) was used. The data from this experiment examined the effects of different types of adaptive cruise control (ACC) training protocols on driving performance and visual attention allocation during simulated drives for younger and older drivers. Two types of training protocols were used: basic training, consisting of text content similar to what is found in a typical owner's manual, and comprehensive training, consisting of the basic training content with the addition of information about the driver's roles and responsibilities and a description of the tasks that are needed while driving with ACC. A short description of the methods used in the paper is provided below for context, with additional details found in (*5*).

Participants

A total of 40 participants (n=20 younger adults, ages 18-26; n=19 older adults, ages 65+) took part in the simulator experiment. All participants were screened to ensure no cognitive impairment or other medical comorbidities; had normal or normal-to-corrected vision and had a valid license and drove within the last three months. Participants in both age groups were randomly allocated to a training condition (basic vs. comprehensive) so that there were roughly equal participants in each group. The research project received ethical approval from the UF Institutional Review Board (IRB#: 201801988).

Apparatus

The experiment was conducted in a Realtime Technologies Inc. (RTI; Royal Oak, MI) RDS-2000 fixed-base driving simulator. The simulator consisted of a full vehicle body, with a 180-degree field of view display in front of the vehicle, an additional display positioned behind the vehicle, and two displays for each of the side mirrors. A custom-developed ACC system with a fixed headway distance (2.2 seconds based on Toyota's Safety Sense system middle setting) was used in the simulated drives.

A Tobii Pro 2 (Tobii Technology AB, Sweden) head-mounted eye-tracker was used to capture visual attention. These glasses feature two eye cameras per eye and a full HD scene camera, providing an 82° horizontal and 52° vertical field-of-view. The eye-tracking system recorded data at a rate of 50 Hz and captured the front scene video at 25 frames per second. Eye-tracking data from the glasses were captured using Ergoneers' D-Lab software (Ergoneers, Germany) at 60 Hz. D-Lab allowed for the automatic detection of AOIs through the use of fiducial markers located in the real-world (**Figure 2**). The software would detect these markers using computer vision, and AOIs could be defined relative to the positions of the markers.

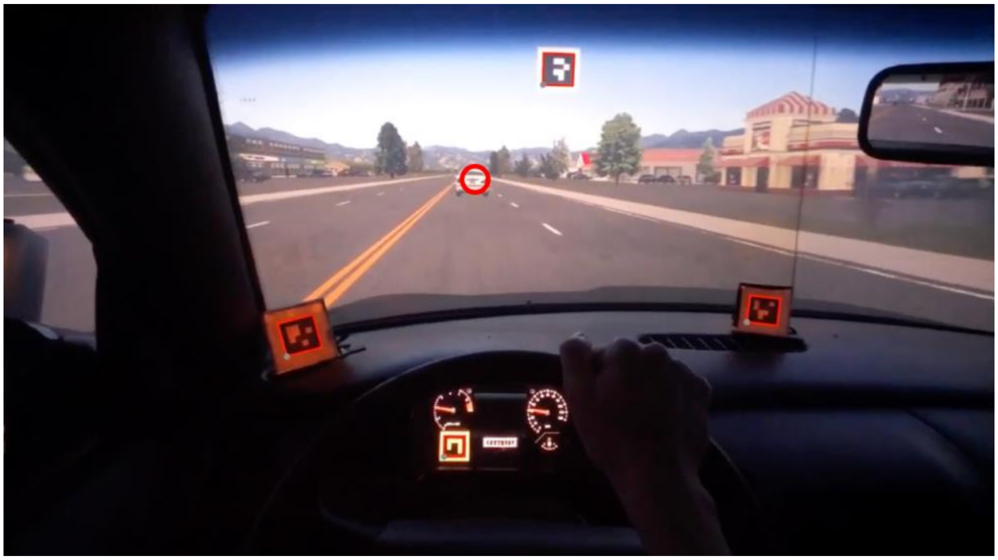


Figure 2 The gaze that falls on the AOI of the windshield is automatically detected. The three fiducial markers at the top and in the center locate the AOI of the windshield, and the fiducial markers at the bottom locate the AOI of the dashboard.

Driving Scenarios and Visual Attention Allocation

During the experiment, participants completed a total of seven experimental drives along a four-lane road of approximately 3.7 km with multiple intersections. Two of the drives were manual drives with no ACC, two were with ACC, and one was a “failure” drive where the operational domain of the ACC was exceeded (a tire was placed in the path of travel that would not be detected by the ACC), and finally, there were two post-failure ACC drives. A lead vehicle was present during each of the drives, which traveled at speeds that varied throughout the drive but were also tethered to the participants’ vehicle to ensure that a lead vehicle was always present. The speed limit changed throughout the course of the drive, requiring participants to monitor for speed limit changes through signage and adjust their ACC’s set speed accordingly. Participants were also expected to monitor the traffic lights at intersections, as the ACC is not able to adjust its speed in response to changes to the lights. Thus, within each of the drives, participants were expected to visually scan the environment for a number of important external cues, and the degree to which they scanned for these events was hypothesized to depend on their understanding of the limitations and capabilities of the ACC.

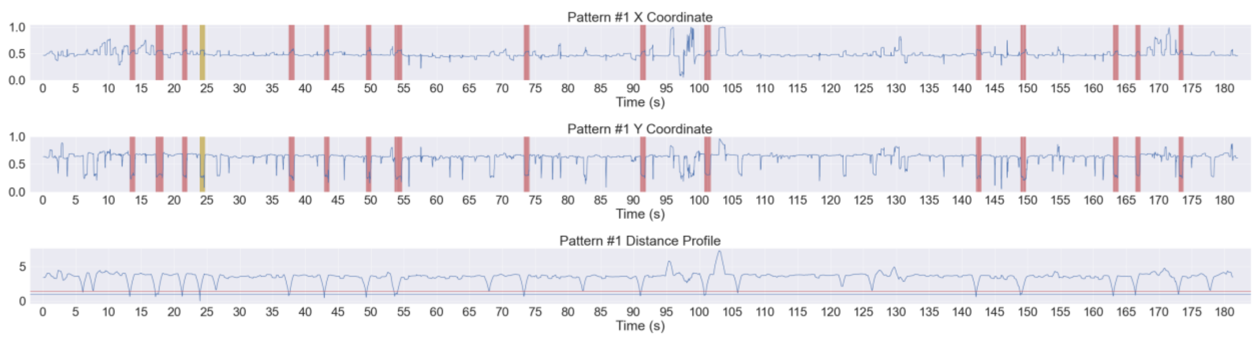
Data Set Characteristics

The data collected from the Tobii Pro Glasses 2 and D-Lab included various variables associated with the pupil position of the left and right eyes, gaze direction of the left and right eyes, gaze position, and corresponding AOI information. For our analysis, we focused on specific variables about gaze position and corresponding AOI information, which encompassed the coordinates of the gaze position on the X and Y axes, as well as whether the fixation fell within the designated AOI regions, including the windshield, dashboard, rearview mirror, left side mirror, and right side mirror areas. The X (Y) coordinate of the gaze location is normalized between 0 and 1, representing a range from left (bottom) to right (top). These variables corresponded to the eye-tracker’s estimates of gaze location within the video footage captured by the scene camera. The scene camera was located on the glasses and would represent the participants' current field of view. Thus, the variables represent the location of the gaze relative to the participant's field of view (e,g, relative to their head position).

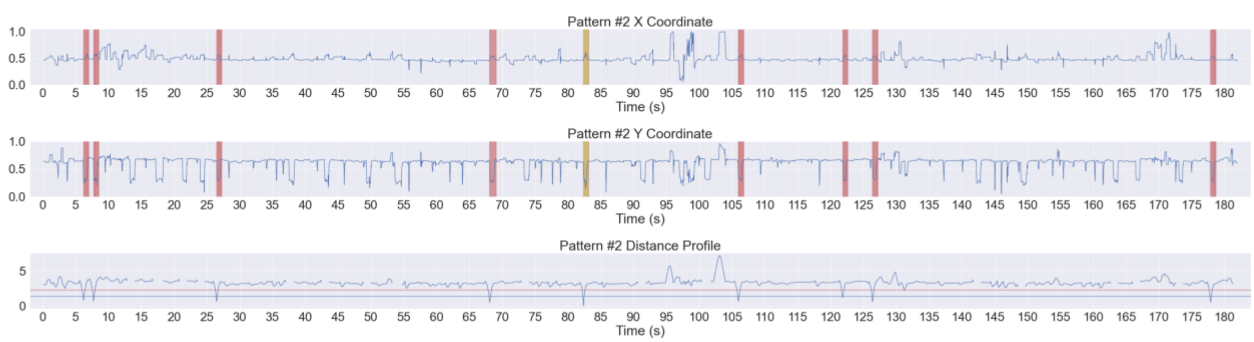
Although it would be ideal to have gaze positions in relation to the real-world coordinates, gathering such data without calibration and anchoring to physical objects may be challenging and more prone to errors. Hence, we initially evaluated our approach using data within the scene-camera space. The benefits and limitations of this approach will be discussed in the discussion section.

Visual Scanning Pattern Identification

In this and the following subsections, the results are based on an older driver who received basic training and performed manual driving without ACC. The proposed method is applied to identify 10 distinct patterns from the eye-tracking data. The sequence length is set to 0.7 seconds. The identification of 10 patterns and their occurrences takes 7.4 seconds on an Intel Core i7-1265U 1.80-GHz processor with 16-GB RAM. For the first two patterns, **Figures 3** (a) and (b) illustrate the locations of the prototypical patterns (yellow), with their multiple occurrences (red) and distance profiles (last row). The results show that the proposed method successfully identifies distinct scanning behaviors along with their similar occurrences throughout the entire sequence. However, it is observed that both Patterns #1 and #2 represent similar behaviors (i.e., a small jump in the X coordinate and a large drop in the Y coordinate, indicating a down-right scan path), suggesting that the threshold is underestimated. Note that if an identified prototypical pattern has fewer than 5 occurrences, this pattern is discarded to ensure that each pattern represents a repetitive behavior of a driver. **Figure 4** shows the final segmentation results after identifying and locating 10 patterns. We can see the diversity of visual scanning behaviors and each pattern's multiple occurrences.



(a)



(b)

Figure 3 Illustration of the locations of the identified prototypical pattern (yellow) and its occurrences (red) and the distance profile of the pattern (last row). (a) Pattern # 1 and (b) Pattern #2. The blue horizontal line in the distance profile plot shows the threshold to identify multiple occurrences.

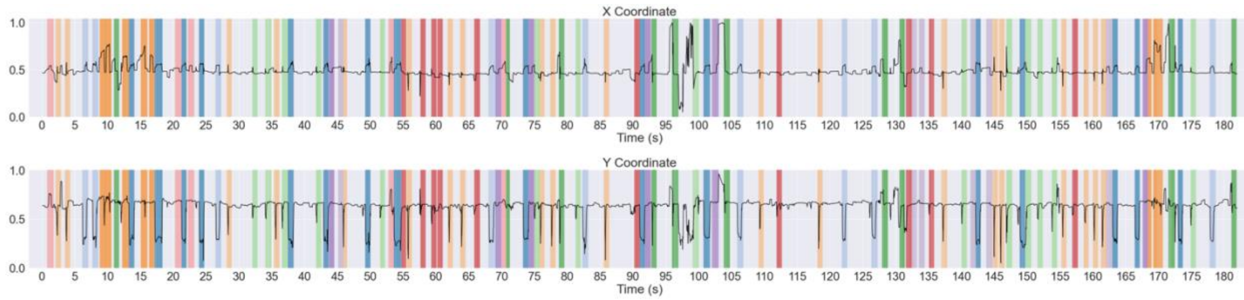


Figure 4 Segmentation results using 10 patterns identified by the proposed method. Each pattern is represented by a different color.

Visual Scanning Pattern Visualization

Figure 5 provides the visualization of each prototypical pattern. **Figures 6** (a) and (b) represent the visual scanning behaviors corresponding to Pattern #1 and #3, respectively. We can easily interpret that Pattern #1 corresponds to checking the speedometer (dashboard) and Pattern #3 corresponds to observing the scenery on the right side (buildings, trees, etc.). Although Pattern #4 displays a scan path visualization similar to Patterns #1 and #2, it is observed that its fixation duration is significantly shorter than those of Patterns #1 and #2. Most of the patterns indicate the driver's gaze shifting within the windshield AOI. It is worth noting that using the existing AOI-based methods, such gaze transitions within the windshield AOI would not be individually captured. Moreover, Patterns #9 and #10 indicate that the proposed method can capture and distinguish visual scanning behaviors that cover very small areas. This highlights the method's effectiveness in detecting fine-grained and localized gaze behaviors in the eye-tracking data. In essence, all extracted patterns encompass diverse eye movements and fixations occurring across and within the AOIs.

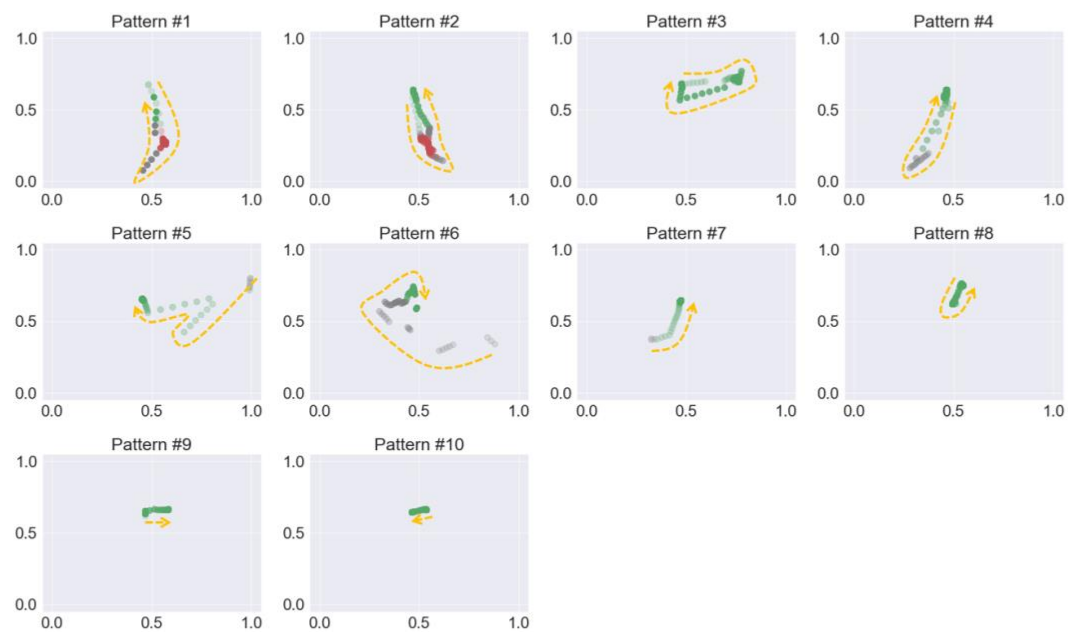


Figure 5 Scan paths of the identified 10 prototypical patterns. The color represents the AOIs (red: dashboard, green: windshield, grey: invalid). The opacity visualizes the chronology of eye movements, which gradually decreases from the initial gaze location to the final gaze location. The yellow dashed line shows the overall trajectory and direction of each prototypical pattern.

(a) (b)

Figure 6 Example of the scanning behavior shown in (a) Pattern #1, the driver is checking the speedometer and (b) Pattern #3, the driver is observing the building on the right side.

Average Visual Scanning Pattern

Figure 7 visualizes multiple occurrences of Patterns #2 and #3. The results for Pattern #1 are similar to those for Pattern #2, yet they are omitted here, as there are a total of 16 occurrences of Pattern #1. We can see that multiple occurrences of a pattern exhibit similarities with slight variations. Note that different occurrences of the same pattern may vary in length as we group consecutive subsequences assigned to the same pattern into one occurrence of the pattern.

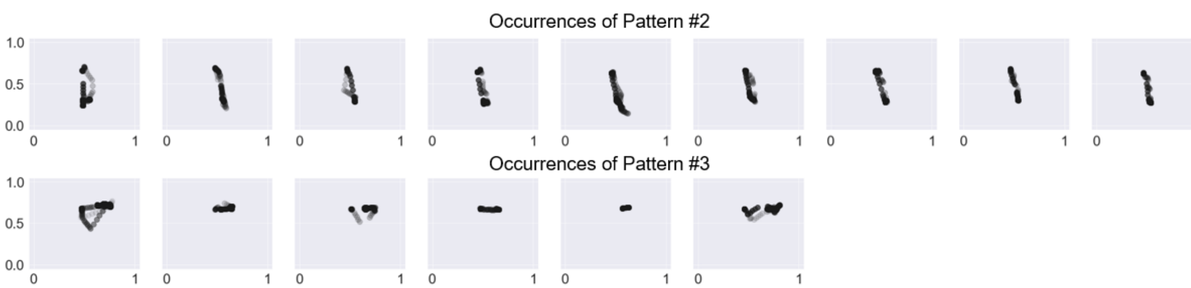


Figure 7 Multiple occurrences of Patterns #2 and #3. The opacity visualizes the chronology of eye movements, the opacity gradually decreases from the initial gaze location to the final gaze location.

To obtain the average of each pattern and its occurrences, Dynamic Time Warping (DTW) method (*36*) is used. DTW employs a “warping” operation to improve alignment between the varying length sequences. **Figure 8** shows the resulting average behavior of each pattern.

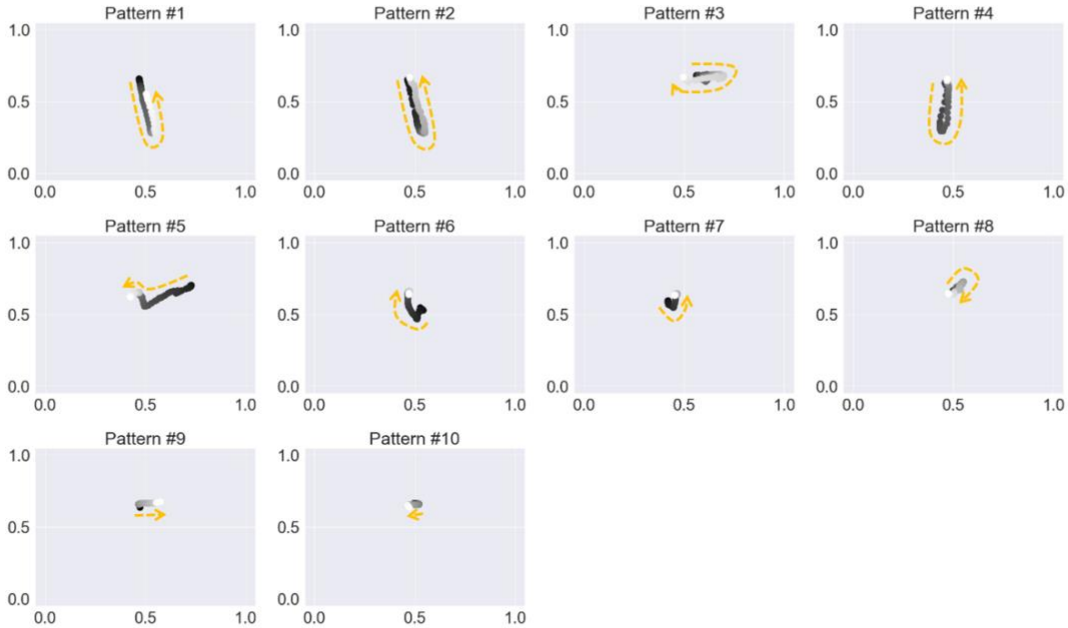


Figure 8 The average scan paths of the identified 10 patterns and their occurrences. The color visualizes the chronology of eye movements from the initial gaze location (black) to the final gaze location (white). The yellow dashed line shows the overall trajectory and direction of each pattern.

These results show the average patterns do differ slightly from the prototypical patterns identified previously. For example, manual inspection of Pattern 6 revealed that multiple occurrences that were classified as Pattern #6 actually vary in terms of gaze transition direction. These results suggest the necessity of a more systematic approach to determine the threshold for grouping similar occurrences or further post-processing in pattern identification.

DISCUSSION

In this section, we will outline the limitations of the proposed approach and several potential topics for future research. The first limitation of this paper is that our approach was demonstrated on eye-tracking data from a single drive from one participant without comparative studies between different drivers or drives. The focus of this study was to demonstrate how our method is able to extract meaningful visual scanning patterns from the data. The next step will be to apply this method across multiple drivers and different situations.

Furthermore, our case study also provided evidence that the pattern identification performance of the proposed method is sensitive to the threshold parameter. Further work is required for a more systematic approach to determine the threshold parameter considering the experiment's design, dataset, and research questions. Also, the validity of the patterns is a human judgment right now, this is certainly a limitation when we have a large number of patterns that need to be validated. To support these efforts, we are developing a methodology to test the validity of these patterns after their extraction to determine whether the behaviors represented by these patterns genuinely represent particular kinds of drivers.

In addition, in the case study, eye-tracking data was collected using a wearable eye tracker, which collects eye movement information relative to the scene camera. Thus, in this current analysis, we relied on gaze positional data that was relative to the head rather than the world. Transforming the gaze position to be relative to the vehicle will likely improve the accuracy of the method, but our current data analysis results show that the proposed approach is robust even while using position data that is relative to the head.

Finally, it is important to note that there are various factors affecting pattern extraction based on scan paths including the visual scanning environment, diverse driving characteristics, distinct driving states, and disparities among visual scanning devices. This poses a challenge to the generalization of this method for broader studies. Therefore, future research endeavors should focus on exploring ways to mitigate or eliminate these limitations when making comparisons among different scenarios.

Future development of our proposed approach will involve the use of extracted patterns for the classification and detection of driving behaviors. This includes the classification of behaviors that exhibit similarity at the AOI level. The potential utility of these classifications may surpass existing methods for behavior classification and detection. Pattern extraction based on scan paths can benefit from incorporating other classification approaches, such as computer vision, to discern subtle distinctions among different patterns. Additionally, a suitable method for analyzing the degree of similarity between two scan paths will be explored. For example, the similarity can be quantified based on the frequency or locations of the shared patterns.

CONCLUSION

This work reviews existing eye-tracking data analysis methods and presents a novel approach for extracting driving visual scanning patterns. The proposed method identifies distinctive patterns by obtaining the distance profile of the scan path. By directly analyzing the scan path, this new method reveals the overall trends, semantics, and intricate differences in eye movement, surpassing the AOI-based feature extraction method. The numerical study elaborates on the visualization of the obtained patterns and the scan paths of corresponding prototypical and average patterns and presents derived insights into driving behaviors.

ACKNOWLEDGMENTS

This work was supported by the University of Florida, ISE Research Thrust Pilot Project (RTPP) Program and we acknowledge the University of Florida's Institute for Disability, Activity, Participation, and Technology (IDAPT) for the support of the driving tasks. We acknowledge using ChatGPT version 3.5 to revise the grammar and language of early drafts during the writing process.

AUTHOR CONTRIBUTIONS

The authors confirm their contribution to the paper as follows: study conception and design: M. Kim and W. Giang; data collection: H. Zheng and W. Giang; analysis and interpretation of results: X. Ma, M. Kim, and W. Giang; draft manuscript preparation: X. Ma, M. Kim, and W. Giang. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. National Center for Statistics and Analysis. *Early Estimate of Motor Vehicle Traffic Fatalities in 2022*. Publication DOT HS 813 428. National Highway Traffic Safety Administration, 2023.

2. Carter, P. M., C. R. Bingham, J. S. Zakrajsek, J. T. Shope, and T. B. Sayer. Social Norms and Risk Perception: Predictors of Distracted Driving Behavior Among Novice Adolescent Drivers. *Journal of Adolescent Health*, Vol. 54, No. 5, 2014, pp. S32–S41. https://doi.org/10.1016/j.jadohealth.2014.01.008.

3. National Highway Traffic Safety Administration. Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices. Department of Transportation, 2014.

4. SAE International. *Definitions and Experimental Measures Related to the Specification of Driver Visual Behavior Using Video Based Techniques*. Publication J2396_201705. SAE International, 2017.

5. Zheng, H., J. R. Mason, S. Classen, and W. C. W. Giang. Pilot Study: Effect of Roles and Responsibility Training on Driver's Use of Adaptive Cruise Control between Younger and Older Adults. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 94, 2023, pp. 53–66. https://doi.org/10.1016/j.trf.2023.01.023.

6. Carney, C., J. G. Gaspar, and W. J. Horrey. Longer-Term Exposure vs Training: Their Effect on Drivers' Mental Models of ADAS Technology. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 91, 2022, pp. 329–345. https://doi.org/10.1016/j.trf.2022.09.017.

7. Green, P. Chapter 4: Where Drivers Look While Driving (and for How Long). In *Human Factors in Traffic Safety*, Lawyers & Judges Publishing Company, Inc., p. 550.

8. Zahabi, M., Y. Wang, and S. Shahrampour. Classification of Officers' Driving Situations Based on Eye-Tracking and Driver Performance Measures. *IEEE Transactions on Human-Machine Systems*, Vol. 51, No. 4, 2021, pp. 394–402. https://doi.org/10.1109/THMS.2021.3090787.

9. Deng, Q., J. Wang, K. Hillebrand, C. R. Benjamin, and D. Soffker. Prediction Performance of Lane Changing Behaviors: A Study of Combining Environmental and Eye-Tracking Data in a Driving Simulator. *IEEE Transactions on Intelligent Transportation Systems*, Vol. 21, No. 8, 2020, pp. 3561–3570. https://doi.org/10.1109/TITS.2019.2937287.

10. Brishtel, I., S. Kraus, T. Schmidt, J. R. Rambach, I. Vozniak, and D. Stricker. Classification of Manual Versus Autonomous Driving Based on Machine Learning of Eye Movement Patterns. Presented at the 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Prague, Czech Republic, 2022.

11. Blascheck, T., K. Kurzhals, M. Raschke, M. Burch, D. Weiskopf, and T. Ertl. Visualization of Eye Tracking Data: A Taxonomy and Survey: Visualization of Eye Tracking Data. *Computer Graphics Forum*, Vol. 36, No. 8, 2017, pp. 260–284. https://doi.org/10.1111/cgf.13079.

12. Braunagel, C., D. Geisler, W. Rosenstiel, and E. Kasneci. Online Recognition of Driver-Activity Based on Visual Scanpath Classification. *IEEE Intelligent Transportation Systems Magazine*, Vol. 9, No. 4, 2017, pp. 23–36. https://doi.org/10.1109/MITS.2017.2743171.

13. Navarro, J., O. Lappi, F. Osiurak, E. Hernout, C. Gabaude, and E. Reynaud. Dynamic Scan Paths Investigations under Manual and Highly Automated Driving. *Scientific Reports*, Vol. 11, No. 1, 2021, p. 3776. https://doi.org/10.1038/s41598-021-83336-4.

14. McIntyre, N. A., and T. Foulsham. Scanpath Analysis of Expertise and Culture in Teacher Gaze in Real-World Classrooms. *Instructional Science*, Vol. 46, No. 3, 2018, pp. 435–455. https://doi.org/10.1007/s11251-017-9445-x.

15. Castner, N., E. Kasneci, T. Kübler, K. Scheiter, J. Richter, T. Eder, F. Hüttig, and C. Keutel. Scanpath Comparison in Medical Image Reading Skills of Dental Students: Distinguishing Stages of Expertise Development. Presented at the ETRA '18: 2018 Symposium on Eye Tracking Research and Applications, Warsaw Poland, 2018.

16. Lappi, O., J. Pekkanen, and T. H. Itkonen. Pursuit Eye-Movements in Curve Driving Differentiate between Future Path and Tangent Point Models. *PLoS ONE*, Vol. 8, No. 7, 2013, p. e68326. https://doi.org/10.1371/journal.pone.0068326.

17. Li, X., R. Schroeter, A. Rakotonirainy, J. Kuo, and M. G. Lenné. Effects of Different Non-Driving-Related-Task Display Modes on Drivers' Eye-Movement Patterns during Take-over in an Automated Vehicle. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 70, 2020, pp. 135–148. https://doi.org/10.1016/j.trf.2020.03.001.

18. Wang, Y., J. Song, X. An, S. Ma, H. Li, D. Wang, and Z. Yang. Head-Up Display Graphic Warning to Support Collision Avoidance: Effect of Graphic Animation and Border on Driving Behavior and Eye Movement Pattern. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2677, No. 5, 2023, pp. 636–652. https://doi.org/10.1177/03611981221134628.

19. Olsen, E. C. B., S. E. Lee, and W. W. Wierwille. Eye Glance Behavior during Lane Changes and Straight-Ahead Driving. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1937, No. 1, 2005, pp. 44–50. https://doi.org/10.1177/0361198105193700107.

20. Olsen, E. C. B., S. E. Lee, and B. G. Simons-Morton. Eye Movement Patterns for Novice Teen Drivers: Does 6 Months of Driving Experience Make a Difference? *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2009, No. 1, 2007, pp. 8–14. https://doi.org/10.3141/2009-02.

21. Jeong, H., Z. Kang, and Y. Liu. Driver Glance Behaviors and Scanning Patterns: Applying Static and Dynamic Glance Measures to the Analysis of Curve Driving with Secondary Tasks. *Human Factors and Ergonomics in Manufacturing & Service Industries*, Vol. 29, No. 6, 2019, pp. 437–446. https://doi.org/10.1002/hfm.20798.

22. Volkmann, F. C., L. A. Riggs, K. D. White, and R. K. Moore. Contrast Sensitivity during Saccadic Eye Movements. *Vision Research*, Vol. 18, No. 9, 1978, pp. 1193–1199. https://doi.org/10.1016/0042-6989(78)90104-9.

23. Ahmad, A., S. A. Rosli, and A.-H. Chen. Eye Tracking System Measurement of Saccadic Eye Movement with Different Illuminance Transmission Exposures during Driving Simulation. Presented at the 2022 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, Malaysia, 2022.

24. Lethaus, F., and J. Rataj. Do Eye Movements Reflect Driving Manoeuvres? *IET Intelligent Transport Systems*, Vol. 1, No. 3, 2007, p. 199. https://doi.org/10.1049/iet-its:20060058.

25. Muttart, J. W., L. R. Peck, S. Guderian, W. Bartlett, L. P. Ton, C. Kauderer, D. L. Fisher, and J. E. Manning. Glancing and Stopping Behavior of Motorcyclists and Car Drivers at Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2265, No. 1, 2011, pp. 81–88. https://doi.org/10.3141/2265-09.

26. Dewhurst, R., M. Nyström, H. Jarodzka, T. Foulsham, R. Johansson, and K. Holmqvist. It Depends on How You Look at It: Scanpath Comparison in Multiple Dimensions with MultiMatch, a Vector-Based Approach. *Behavior Research Methods*, Vol. 44, No. 4, 2012, pp. 1079–1100. https://doi.org/10.3758/s13428-012-0212-2.

27. Noton, D., and L. Stark. Scanpaths in Saccadic Eye Movements While Viewing and Recognizing Patterns. *Vision Research*, Vol. 11, No. 9, 1971, pp. 929-IN8. https://doi.org/10.1016/0042-6989(71)90213-6.

28. Cristino, F., S. Mathôt, J. Theeuwes, and I. D. Gilchrist. ScanMatch: A Novel Method for Comparing Fixation Sequences. *Behavior Research Methods*, Vol. 42, No. 3, 2010, pp. 692–700. https://doi.org/10.3758/BRM.42.3.692.

29. Kübler, T. C., E. Kasneci, and W. Rosenstiel. SubsMatch: Scanpath Similarity in Dynamic Scenes Based on Subsequence Frequencies. Presented at the ETRA '14: Eye Tracking Research and Applications, Safety Harbor Florida, 2014.

30. Kübler, T. C., C. Rothe, U. Schiefer, W. Rosenstiel, and E. Kasneci. SubsMatch 2.0: Scanpath Comparison and Classification Based on Subsequence Frequencies. *Behavior Research Methods*, Vol. 49, No. 3, 2017, pp. 1048–1064. https://doi.org/10.3758/s13428-016-0765-6.

31. Geisler, D., N. Castner, G. Kasneci, and E. Kasneci. A MinHash Approach for Fast Scanpath Classification. Presented at the ETRA '20: 2020 Symposium on Eye Tracking Research and Applications, Stuttgart Germany, 2020.

32. Mathôt, S., F. Cristino, I. D. Gilchrist, and J. Theeuwes. A Simple Way to Estimate Similarity between Pairs of Eye Movement Sequences. *Journal of Eye Movement Research*, Vol. 5, No. 1, 2012. https://doi.org/10.16910/jemr.5.1.4.

33. Yeh, C.-C. M., Y. Zhu, L. Ulanova, N. Begum, Y. Ding, H. A. Dau, D. F. Silva, A. Mueen, and E. Keogh. Matrix Profile I: All Pairs Similarity Joins for Time Series: A Unifying View That Includes Motifs, Discords and Shapelets. Presented at the 2016 IEEE 16th International Conference on Data Mining (ICDM), Barcelona, Spain, 2016.

34. Mueen, A. The Fastest Similarity Search Algorithm for Time Series Subsequences under Euclidean Distance., 2022.

35. Yeh, C.-C. M., N. Kavantzas, and E. Keogh. Matrix Profile VI: Meaningful Multidimensional Motif Discovery. Presented at the 2017 IEEE International Conference on Data Mining (ICDM), New Orleans, LA, 2017.

36. Dynamic Time Warping. In *Information Retrieval for Music and Motion*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 69–84.