Eye movement identification within dynamic environment

Bryan Reimer
Massachusetts Institute of Technology, AgeLab
Corresponding author: reimer@mit.edu

Manbir Sodhi
University of Rhode Island, Kingston

Abstract

With increasing commute times and advances in wireless and computing technology, a variety of tasks compete for drivers' attention. The visual, auditory, biomechanical or cognitive requirements of the tasks are difficult to identify. Drivers identify and prioritize the requirements of each primary, driving, and secondary tasks. When attention is devoted to a secondary task, a reduction in the level of attention devoted to the roadway results. In-vehicle device designers would like to minimize the level of distraction associated with operating in-vehicle technologies. Simulation is often used as a tool to identify these differences. However, with a simulated environment questions can arise as to how results relate to actual driving conditions.

The process of identifying the focus of a driver's attention in a real driving environment is difficult. Eye movements, closely correlated with visual attention, are possibly useful in indicating the locus of visual perception. Thus, recorded eye positions form a guide for studying the effect of in-vehicle technologies on a driver's behavior. In the past, the analysis of eye positions recorded in a dynamic environment relied on manual identification of visual attention from a recorded scene video. This paper presents current research on the development of a new methodology and toolset for the analysis of eye positions recorded in a dynamic scene, including a method of summarizing the allocation of a subject's attention.

Introduction

Automobile driving is one of the more complex dynamic environments people face every day. Drivers are often affected by a variety of different distractions which are comprised of four interrelated components: visual, auditory, biomechanical, and cognitive (Ranney, Mazzae, Garrott, & Goodman, 2000). Various levels of each component are associated with distracting activities. Unavoidable distractions originate both inside and outside the vehicle, whereas in-vehicle devices present additional distractions that a driver may choose to become involved with. Each distraction that a driver becomes involved with may have an associated safety implication. Obviously, vehicular safety can be improved through the minimization of the number, level and type of distractions affecting a driver ((Ranney et al., 2000) and (Serafin, Wen, Paelke, & Green, 1993)). Manual observations and data acquisition systems have been used to study the effects of in-vehicle devices (Reimer, 2003).

Techniques based upon manual observation of a driver's behavior offer the most direct link to a driver's actions. However, the low sample rate and limited detail describing each observation makes this method difficult to implement in studies with complex interactions between behaviors. Data acquisition systems overcome these limitations by increasing the sampling rate, allowing observation of more subtle changes in the operator's behavior. However, when compared to manual methods, a concise relation between various measures and the actions that a driver is believed to be involved with are difficult to develop. Eye tracking, a single-stream method of data acquisition, reduces the complexity involved with the correlation of multiple parameters by focusing on the assumed relationship between eye movements and attention. Typical eye tracking applications sample the eyes' position at a rate between 30 - 240 Hz, and consequently slight changes in a subject's attentional focus can be related to the observations in a scene recording.

Sodhi et al. (2002) uses eye tracking to identify behavior patterns for a variety of common driving tasks. In (Sodhi et al., 2002), eye positions are recorded using a head-mounted eye tracker in the subject's own vehicle navigating a 30 - 40 minute course down rural roads. The manual analysis of the recorded eye positions, as discussed in Sodhi, Reimer, and Llamazares (2002), identifies differences in the standard deviation for eye positions recorded during cognitive tasks when compared to controls. Differences in the glance durations to positions on or off the road were also identified for a radio changing and rear view mirror task but not for an odometer-checking task. Although effective, the manual analysis is not feasible for processing large data sets. To further extend the use of eye tracking as a method of characterizing behavior changes in a dynamic environment, a greater degree of automation of the analysis is required. This paper presents a flexible approach for identifying eye movements from eye positions recorded in a dynamic environment. Details on the assumed relationship of eye movements and attention are provided. Then pupil / corneal reflection eye trackers are discussed, highlighting a method for identifying saccades, smooth movements and fixations based upon environments identified through the displacement between the pupil and corneal reflection. Finally, a description of an interface to map the identified movements to observations from the scene image is illustrated with an example of eye movement data recorded from an on-road driver.

Eye movements and attention

A basic link between attention, the ability to concentrate on a task, and eye movements is established in the premotor theory which suggests "that overt orienting through eye movements and covert orienting through attentional movements are controlled by closely related mechanisms and that eye movements normally follow attentional movements" (Palmer, 1999). Henderson (1993) provides further verification of the premotor theory with the sequential attention model. The sequential attention model describes attention as being "directed to the specific location toward which the eyes will move prior to a saccadic eye movement." Breitmeyer and Braun (1990) provides a more numeric relation by suggesting that a saccadic movement lags behind an attention shifts by about 100ms. This model is limited when attention is directed to and from a location without a saccade being executed, such as the case where attention shifts are faster then the delayed response of a preprogrammed saccade (Liversedge & Findlay, 2000).

The link between eye movements and attention has been used to study task execution in a stationary environment. Just and Carpenter (1976) investigate the sequence and duration of fixations observed during different cognitive tasks. Results show a relationship between activity of the central processor and the locus, duration and sequence of eye fixations. Salvucci (1999) looked at developing attention-based modeling based upon information that subjects used in solving static scene problems. Recorded eye positions were used as a method of identifying the length of time required to process information, the pattern used for acquiring information and what information was forgotten and needed additional focus. Finally, Recarte, Nunes, Lopez, and Recarte (1999) investigates the link between eye movements and mental imagery, visual imagery, image manipulation and verbal tasks on a driver. Recarte et al. (1999) compares fixation patterns recorded during secondary tasks and normal driving situations. Results show differences in the pattern of fixations, long and short, during the imagery manipulation task and a pattern of short fixations during the verbal tasks. However, models that use fixation identification, cannot be used in dynamic environments where saccades and smooth movements must be identified to form a clear picture of the information a subject perceives.

Eye tracking equipment and methods of identifying eye positions

A variety of different techniques have been used for tracking human eye movements. Eye tracking methods such as Electro-Oculography and contact lenses based methods are too invasive and difficult to use in dynamic environments such as in a subject's vehicle (Glenstrup & Engell-Nielsen, 1995). Systems dependent on one feature such as the pupil, corneal reflection or iris-scleral boundary (limbus) can only be used with some degree of head stability (Young & Sheena, 1975). Unrestricted head movements that occur in on-the road driver behavior studies, are related to the gaze path through a technique that calculates a subject's point-of-regard (POR), within a scene camera fixed to the subjects head. The process uses the pupil and corneal reflections, two features of the eye that "move differently as a function of head motion and eye rotation, such that pure eye rotation can be deduced from them." (Young & Sheena, 1975) Head-mounted eye tracking systems record the relation between a subject's pupil and corneal reflection. Then through a predefined calibration with an image recorded by a scene camera fixed to the subject's head, the pupil and corneal reflection is related to the POR.

Young and Sheena (1975) details the theory used to calculate the POR from the pupil and corneal reflection. Two examples are used from Young and Sheena (1975) to illustrate the process. If a subject stares at a light, an observer will see the corneal reflection aligned with the subject's pupil. Furthermore, if the subject stares at a set of different colored lights, the POR can be determined by which light source appears reflecting on the center of the pupil. The concept is further defined mathematically through the radius of curvature of the cornea that is less than that of the eye. When the eye moves, the displacement of the corneal reflections is about half the distance in the opposite direction of the eye. Since the displacement is measured relative to the optic axis, or the center of the pupil and "the subject's angle of gaze with respect to the light source is approximately proportional to the distance between the image of the light source and the center of the pupil." (Young & Sheena, 1975) A single light source and a pre-defined calibration can be used to calculate the POR based upon the error between the center of the pupil and the center of the corneal reflection (P-CR). In addition, eye positions calculated based upon rotational movement of the pupil and corneal reflection reduce the effect of slight linear movements of the recording or illuminating equipment in relation to the eye, increasing the accuracy of the POR (Razdan, 2002).

Traditional methods of analyzing eye movements focus largely on separating fixations from saccades and smooth movements, based upon velocity, aggregation of consecutive points and digital filtering (Salvucci, 1999). In a dynamic environment saccades and smooth movements provide additional information about a subject's visual perception. Smooth movements represent periods where a subject perceives visual information from moving objects and shows that some visual information is perceived during saccades. Sauter, Martin, Renzo, and Vomscheid (1991) use a Kalman filter to identifying saccades and smooth movements from eye positions recorded in a dynamic environment. However, the method appears too limited for use in real-world situations where the eye positions are sampled at rates lower than the tested 1000 Hz and a representation of the POR is required. Thus, a method modeling based upon changes in the relationship between the pupil and corneal reflection has been used to accurately estimate eye movements.

Eye movements

Recorded eye positions form a complete picture of the path of the eyes movement. In high noise environments, such as in a vehicle, the saccades, smooth movements and fixations are more difficult to identify. To reduce noise in the recorded P-CR error, eye tracking systems, such as the ISCAN ETL-500, represent the POR as a running average of the eye's position over a predefined number of samples. The averaging reduces environmental influences, the effect of individual measurements and the effect of physiological nystagmus. Dampened representation of changes in eye position recorded through the POR overlook sudden shifts caused by saccades and smooth movements and make the movements appear delayed in the POR. Effects of the averaging are observed in short fixations that appear more as smooth movements from averaging with saccades at the start and end. The larger the number of samples used to form an average eye position, the greater the differences between the appearance of the POR in the scene video and the shift in eye position. However, to reduce fluctuations in the POR as recorded in a scene video some averaging is needed to increase the ease with which objects of focus can be identified.

To depict the differences between raw eye positions and the POR, an example, illustrating a driver changing the radio station, is considered through the gaze path presented in Figure 1. In Figure 1 the horizontal and vertical POR is plotted over time with the actual shifts in eye position, represented through the P-CR error. The dampening and displacement of the POR signal is clearly evident through the lack of spikes and dwells when compared to the P-CR error. Using the P-CR errors saccades, smooth movements and fixations are clearly identifiable.

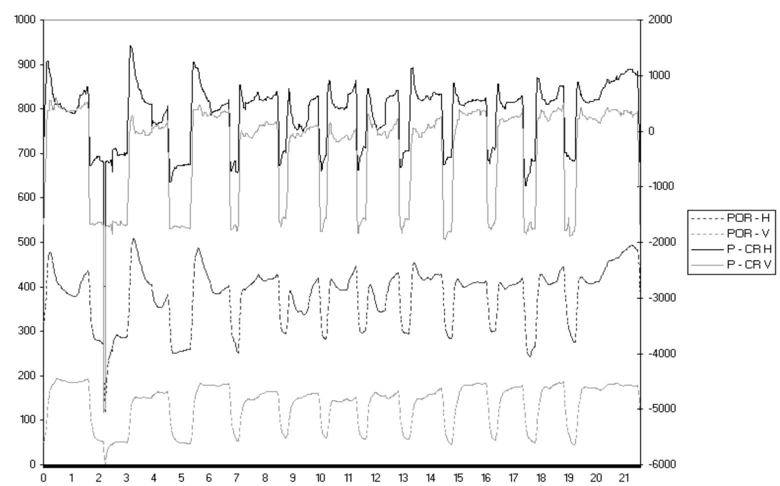

Figure 1. A subject's gaze path represented with the P-CR error and POR for the horizontal and vertical positions

Semi-automated eye movement identification

A semi-automated eye movement identification (SEMA) is system proposed to help process eye movement data, assisting in the identification of saccades, smooth movements, fixations and errors. The P-CR data is used for eye movement identification, and this is done after recorded positions are filtered to remove both positions that abut missing data and positions that are infeasible

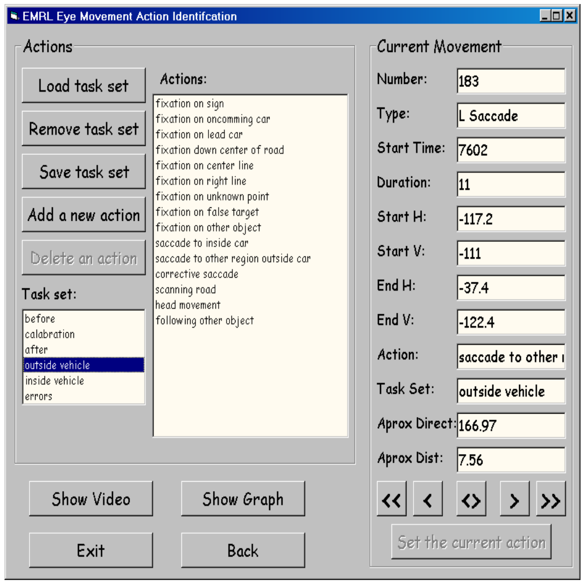

Figure 2. The state identification interface

according to an assumed maximum velocity of the eyes movement. Further samples are considered as a component of the error if they do not satisfy a minimum continuous set size. Interpolation is then used to replace short sections of missing eye positions by assuming, that for short periods of time, the eye moves in a linear fashion. Eye movements are identified from the recorded eye positions by first using velocity thresholds to separate potential components of fixations and saccades. Second, duration minimums are used to identify saccades and fixations from the potential components. Using another duration minimum, smooth movements are classified for sets of positions not identified as part of either a saccade or fixation. Then positions not classifiable into any of the three categories are considered as part of the error component. A small amount of averaging is used to improve identification by reducing any noise that remains after filtering.

With eye movements so identified, it is possible to map each movement to an object of focus as recorded in the scene video. An interface, illustrated in Figure 2 and the associated media player illustrated in Figure 3, have been developed for this purpose. The mapping process involves three stages leading to identification of a particular eye movement segment: identification, relating and mapping. The identification stage requires that an operator take an eye movement and correctly locate, in the scene video, the corresponding recording time. The point of focus is subsequently identified through the POR recorded in the scene video, and is related to one of the predefined sets of tasks and actions. Finally, a relationship is formed between the eye movement and action by mapping the selected action to the eye movement. In general, fixations are then defined as periods when visual information is perceived from stationary objects, smooth movements as periods where visual information is perceived from eye positions moving with a related target. Saccades are considered during periods where the eye position moves rapidly between two points of focus. In summary, if a defined set of actions is associated with a subject's eye movements, eye positions recorded over different tasks provide a link to use in comparing changes in the subject's behavior between the tasks.

Figure 3. The state identification video review interface

An example

Consider a set of eye positions, P-CR errors, representing a driver's point of gaze traveling down a portion of North Road in South Kingston, RI. North road is a sparsely populated rural road with a speed limit of 25 MPH, a low traffic density, and limited peripheral stimuli such as billboards, road signs etc. The recorded positions are related to 688 eye movements and then mapped to associated actions. A summary of the classification is provided in Table 1. The summary includes: the number of times each state was visited, the average time spent in each state, the standard deviation of the time spent in each state and the percentage of the total number of movements each state represents. Transition matrixes are used to illustrate the probability of shifting between states. An example is illustrated in Table 2 for transitions from all states to states involving glances outside of the vehicle. A full transion matrix for this data is twice as large, including transitions to inside vehcile and error states in addition to the outside of the vehicle states. The state ID, as defined in Table 1, is in the first column and row of the table.

Conclusions

This paper describes a method of analyzing eye positions recorded in a dynamic environment. Based on an overview of the filtering process and identification of saccades, smooth movements, fixations and errors, a method of categorizing attention shifts classified through the identified eye movements is presented. The tools provided in this paper illustrate methods of identifying and studying individual components of human thought for tasks completed in a dynamic environment. A software interface for relating identified eye movements to the POR in a recorded scene video is illustrated with an example of a driver's actions summarized.

The methodology developed here for identifying the components of a subject's eye movements in dynamic environments is easily extended to permit comparisons between or within subjects. Differences observed between subjects operating similar tasks or between an individual operating different tasks will allow designers to better fit the requirements of tasks to the methods used for their operation. The applications of the methods to a sample data track shows that the tool is useful for categorizing on-road driver behavior. In summary, this research forms the basis for a

Id	Task set	Action	Number	Average	Std	Percent
				(samples)	(samples)	
1	Outside	Fixation on sign	10	188.6	157.62	1.37
2		Fixation on oncoming car	6	90.67	69.37	0.82
3		Fixation on center of road	112	99.69	65.11	15.38
4		Fixation on center line	15	99	58	2.06
5		Fixation on right line	40	90.93	54.07	5.49
6		Fixation on unknown point	24	65.17	31.25	3.30
7		Fixation on other object	19	94.79	86.42	2.61
8		Saccade to inside	11	19.18	6.68	1.51
9		Saccade to outside	70	8.33	3.5	9.62
10		Corrective saccade	165	12.2	4.55	22.66
11		Scanning road	13	20.15	8.09	1.79
12		Head movement	1	22		0.13
13	Inside	Fixation on rear view mirror	7	52	18.25	0.96
14		Fixation on left mirror	1	39		0.13
15		Fixation on dash (speed / odometer)	5	57.4	12.76	0.69
16		Fixation on other region inside	3	46.67	34.27	0.41
17		Saccade to outside	3	14.67	4.04	0.41
18		Saccade to inside	2	6.5	0.71	0.27
19		Corrective saccade	3	5.33	0.58	0.41
20		Head movement	5	12.8	7.22	0.69
21	Errors	Possible blink	149	44.43	11.73	20.47
22		Recording error (unknown)	2	24	4.24	0.27
23		Recording error (sun)	31	98.77	111.10	4.26
24		Unclassifiable point	16	3.31	2.75	2.2
25		Part of previous movement	15	8	5.71	2.06

Table 1: Actions associated with each eye movement.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	1							0.1		0.1		
2										0.17		
3								0.02	0.24	0.05	0.01	
4								0.07	0.27			
5								0.08	0.4			
6									0.46	0.04		
7								0.05	0.26	0.11		
8												
9	0.03	0.06	0.24	0.14	0.13	0.19	0.11				0.01	
10	0.05	0.01	0.52	0.03	0.18	0.05	0.06				0.04	
11			0.08					0.08	0.31	0.08		
12												
13												
14												
15												
16												
17			0.33									
18												
19												
20												
21			0.01					0.01	0.01	0.97		
22										0.5		
23			0.06		0.03				0.06	0.26	0.03	
24												0.06
25			0.2			0.2	0.07				0.2	

Table 2: The first part of the transition matrix for section one, representing the transitions to outside of the vehicle actions.

more detailed characterization of driver behavior and the building block for automation of methods for identifying changes in driver behavior.

Bibliography

Breitmeyer, B., & Braun, D. (1990). Effects of fixation and attention on saccadic reaction time. In R. Groner, G. d'Ydewalle, & R. Parham (Eds.), *From eye to mind. information acquisition in perception. search, and reading* (p. 71-79). Amsterdam: North-Holland.

Glenstrup, A., & Engell-Nielsen, T. (1995). *Eye controlled media: Present and future state.* Unpublished master's thesis, University of Copenhagen.

Henderson, J. (1993). Visual attention and saccadic eye-movements. In G. d'Ydewalle & J. van Rensbergen (Eds.), *Perception and cognition. advances in eye-movement research* (p. 37-50). Amsterdam: North-Holland.

Just, M., & Carpenter, P. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*, 441-480.

Liversedge, S., & Findlay, J. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Science*, *4*, 6-14.

Palmer, S. (1999). *Vision science photons to phenomenology.* Cambridge, MA: MIT Press.

Ranney, T., Mazzae, E., Garrott, R., & Goodman, M. (2000). *Driver distraction research: past, present and future* (Report). National Highway Traffic Safety Administration.

Razdan, R. (2002, Burlington, MA). Phone interview, ISCAN.

Recarte, M., Nunes, L., Lopez, R., & Recarte, S. (1999). Effect of two different mental tasks on visual search behavior while driving. In A. Gale, I. Brown, C. Haslegrave, & S. Taylor (Eds.), *Vision in vehicles VII* (p. 215-224). Amsterdam, North Holland: Elsevier.

Reimer, B. (2003). *Automated methods of detecting driver distractions.* Unpublished doctoral dissertation, University of Rhode Island, Industrial and Manufacturing Engineering.

Salvucci, D. (1999). *Mapping eye movements to cognitive processes.* Unpublished doctoral dissertation, Carnegie Mellon University.

Sauter, D., Martin, B., Renzo, N. D., & Vomscheid, C. (1991). Analysis of eye tracking movements using innovations generated by a kalman filter. *Medical and Biological Engineering and Computing*, 63-69.

Serafin, C., Wen, C., Paelke, G., & Green, P. (1993). Car phone usability: A human factors laboratory test. In *Proceedings of the human factors and ergonomics society 37th annual meeting* (p. 220-224). Santa Monica CA: Human Factors and Ergonomics Society.

Sodhi, M., Reimer, B., Cohen, J., Vastenburg, E., Kaars, R., & Kirschenbaum, S. (2002). On-road driver eye movement tracking using head-mounted devices. In *Etra 2002: Eye tracking research and applications symposium.* New Orleans, LA.

Sodhi, M., Reimer, B., & Llamazares, I. (2002). Glance analysis of driver eye movements to evaluate distraction. *Behavior Research Methods, Instruments and Computing*, *34*(4), 529-538.

Young, L., & Sheena, D. (1975). Methods and designs: Survey of eye movement recording methods. *Behavior Research Methods and Instruments*, *7*(5), 397-429.