

# Detecting eye movements in dynamic environments

BRYAN REIMER

*Massachusetts Institute of Technology, Cambridge, Massachusetts*

and

MANBIR SODHI

*University of Rhode Island, Kingston, Rhode Island*

To take advantage of the increasing number of in-vehicle devices, automobile drivers must divide their attention between primary (driving) and secondary (operating in-vehicle device) tasks. In dynamic environments such as driving, however, it is not easy to identify and quantify how a driver focuses on the various tasks he/she is simultaneously engaged in, including the distracting tasks. Measures derived from the driver's scan path have been used as correlates of driver attention. This article presents a methodology for analyzing eye positions, which are discrete samples of a subject's scan path, in order to categorize driver eye movements. Previous methods of analyzing eye positions recorded in a dynamic environment have relied completely on the manual identification of the focus of visual attention from a point of regard superimposed on a video of a recorded scene, failing to utilize information regarding movement structure in the raw recorded eye positions. Although effective, these methods are too time consuming to be easily used when the large data sets that would be required to identify subtle differences between drivers, under different road conditions, and with different levels of distraction are processed. The aim of the methods presented in this article are to extend the degree of automation in the processing of eye movement data by proposing a methodology for eye movement analysis that extends automated fixation identification to include smooth and saccadic movements. By identifying eye movements in the recorded eye positions, a method of reducing the analysis of scene video to a finite search space is presented. The implementation of a software tool for the eye movement analysis is described, including an example from an on-road test-driving sample.

The locus of a driver's attention has been studied since the early days of the automobile (Trott, 1930, as cited in Goodman et al., 1997). With increasing numbers of in-vehicle devices and mobile connectivity, questions about the impact of these on driver safety are becoming increasingly urgent. When a driver's attention is focused away from the roadway, clearly some increase in risk is likely (Dingus, 2000). Current methods of analyzing the increase in risk are subjective and require tedious manual processing for identifying where a driver may be looking while driving. The principal causes of traffic incidents include driver distractions and inattention (Stutts & Hunter, 2003; Sussman, Bishop, Madnick, & Walter, 1985). A distraction is defined as any compelling reason for a driver to divert his or her attention away from the roadway. Inatten-

tion is the result of a noncompelling attention shift, such as fatigue (Brown, 1997; Nilsson, Nelson, & Carlson, 1997) or being "lost in thought" (Wang, Knippling, & Goodman, 1996). Many authors have investigated how distractions affect a driver's behavior and the potential implications for roadway safety (Goodman et al., 1997; Schreiner, 2000).

The effect of distractions on roadway safety can be studied by either epidemiological or behavioral analyses. The focus of epidemiological studies, such as Redelmeier and Tibshirani (1997) and Violanti (1998), is to estimate the increased risk associated with different types of distracting activities by examining incident data. Behavioral studies of drivers examine the impact of various distractions (McKnight & McKnight, 1991), hands-free cellular telephones (Harbluk, Noy, & Eizenman, 2002), and cognitive tasks (Sodhi, Reimer, & Llamazares, 2002). McKnight and McKnight focused on the perceptual and cognitive demands of driving. In this study, a subject's responses to videotaped driving scenes were used to gauge various levels of distraction. Harbluk et al. investigated the use of hands-free devices as a method for minimizing the level of biomechanical distraction associated with holding a cellular phone. Sodhi et al. used recorded eye positions to categorize a reduction in the standard deviation of subjects' scan paths during a computational task. Although based

---

Support for this project was provided by Joseph Coughlin at the MIT AgeLab, the NATO Undersea Research Center (La Spezia, Italy), the United States Department of Transportation's University of Rhode Island Transportation Center, and the United States Department of Transportation's New England University Transportation Center at the Massachusetts Institute of Technology. Correspondence concerning this article should be addressed to B. Reimer, AgeLab, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Room E40-291, Cambridge, MA 02139 (e-mail: reimer@mit.edu).

on limited populations, the studies by Harbluk et al. and by Sodhi et al. used eye tracking as a method of identifying the focus of a driver's attention. Although Stutts et al. (2003) used video recordings of the forward scene and the driver to identify types of distractions, the difficulties of manual coding and subjective interpretation are evident in the results, which can be interpreted to imply that cell phone use is not a major safety threat. Inasmuch as advances in sensors and mobile data collection systems make it easier to conduct more comprehensive behavioral studies, it is clear that corresponding advances in data analysis methods are required.

This article presents the development of a postrecording data analysis support system used to enhance the versatility of assessing on-road driver behavior, based primarily on eye position data. Saccades, fixations, and smooth movements are identified, using a set of precomputed models. Inferences about the interaction between the different eye movements and the external environment are then developed, using a combination of manual (video-based) and automated (model-based) analyses. An implementation of the system is described, and the feasibility of using this system to identify differences in a driver's behavior over two different road types is demonstrated with an example.

### Driver Distractions

The effect of driving distractions can be observed directly (by manual observers), indirectly (using automated data acquisition systems), or by a combination of both (Reimer, 2003). Manual observation studies involve subjective scores of a subject's responses to different tasks. Thus, manual observation limits the range and frequency of the conditions that can be assessed simultaneously to the bounds of human visual capacities. Indirect measurements of driving distraction involve condition monitoring of the subjects. Changes in the throttle, brake, or steering wheel position are simultaneously recorded as an indication of how a driver reacts to changes in the roadway's characteristics. For gaining insight into driving performance, biometrics typically monitored include heart rate, blood pressure, grip tightness, skin conductance, and eye movements. Whereas biometrics such as heart rate, electrocardiogram, blood pressure, and grip tightness may be recorded to indicate changes in a driver's stress level, eye tracking may be used individually or in conjunction with other data acquisition systems to measure how subjects focus their attention.

Although indirect monitoring requires additional processing to link the data to the subject's behavior, these systems provide the ability to capture fast and subtle changes that cannot be captured by manual observation. In the past, behavioral studies have commonly used combinations of manual, vehicular, cognitive, and biological measures to infer a driver's state (Healey, Seger, & Picard, 1999; McKnight & McKnight, 1991; Mourant & Rockwell, 1972; Nunes & Recarte, 2002; Summala, Nieminen, & Punto, 1996). More recent research, however, has successfully relied on eye movements alone to monitor and predict driver performance (Chapman & Underwood, 1999; Harbluk et al., 2002; Liu, 1998; Liu

& Salvucci, 2001; Sodhi et al., 2002). Chiang, Brooks, and Weir (2004) used a combination of eye movements and basic vehicular performance as a method of categorizing driver performance. Studies such as this, combining eye movement analysis and vehicular performance, are likely to become more popular as it becomes easier and more affordable to integrate eye-tracking technology into the vehicle. Correspondingly, methods of analyzing eye movements need to be developed that reduce the time requirements of current systems.

### Eye Movements and Attention

The link between eye movements and driver attention has been examined in several studies. The premotor theory provides evidence for the most direct link between eye movements and attention by suggesting "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, p. 569). The premotor theory is further supported by the sequential attention model, which specifies that attention is "directed to the specific location toward which the eyes will move prior to a saccadic eye movement" (Henderson, 1993, p. 43). Breitmeyer and Braun (1990) further clarified the relation by showing that saccadic movements lag attention shifts by about 100 msec. This model relating eye movements to shifts in attention is limited in one way, since attention can be redirected to a new location but the pursuing preprogrammed saccade can be canceled before the onset of motion (Liversedge & Findlay, 2000). Other studies (Chapman & Underwood, 1999; Land & Horwood, 1996; Liu & Salvucci, 2001; Underwood, Chapman, Crundall, Cooper, & Wallen, 1999) also have reported evidence of the link between driver attention and eye positions and movements. Despite this, eye movement studies are conducted largely under controlled conditions in which extraneous infrared sources, such as the sun, do not exist.

### Eye Movement Identification in a Dynamic Environment

Recorded eye positions form a picture of a subject's scan path (Buswell, 1935; Noton & Stark, 1971; Yarus, 1967). Traditional methods of analyzing eye movements recorded in a stationary visual environment focus largely on separating fixations from saccades and smooth movements, using point-to-point velocities, aggregation of consecutive points, and digital filtering (Salvucci, 1999; Salvucci & Anderson, 2001). In a dynamic environment, however, saccades and smooth movements provide additional information about a subject's scan path. Noise, head movements, and other recording errors make it difficult to locate individual eye movements exactly. To estimate the location of a subject's visual attention with greater precision, it is necessary to resolve the components of the scan path. For this, a method of dampening the recorded eye positions and identifying eye movements is required.

Commercial eye-tracking systems reduce the reported noise by computing the subject's point of regard (POR) as

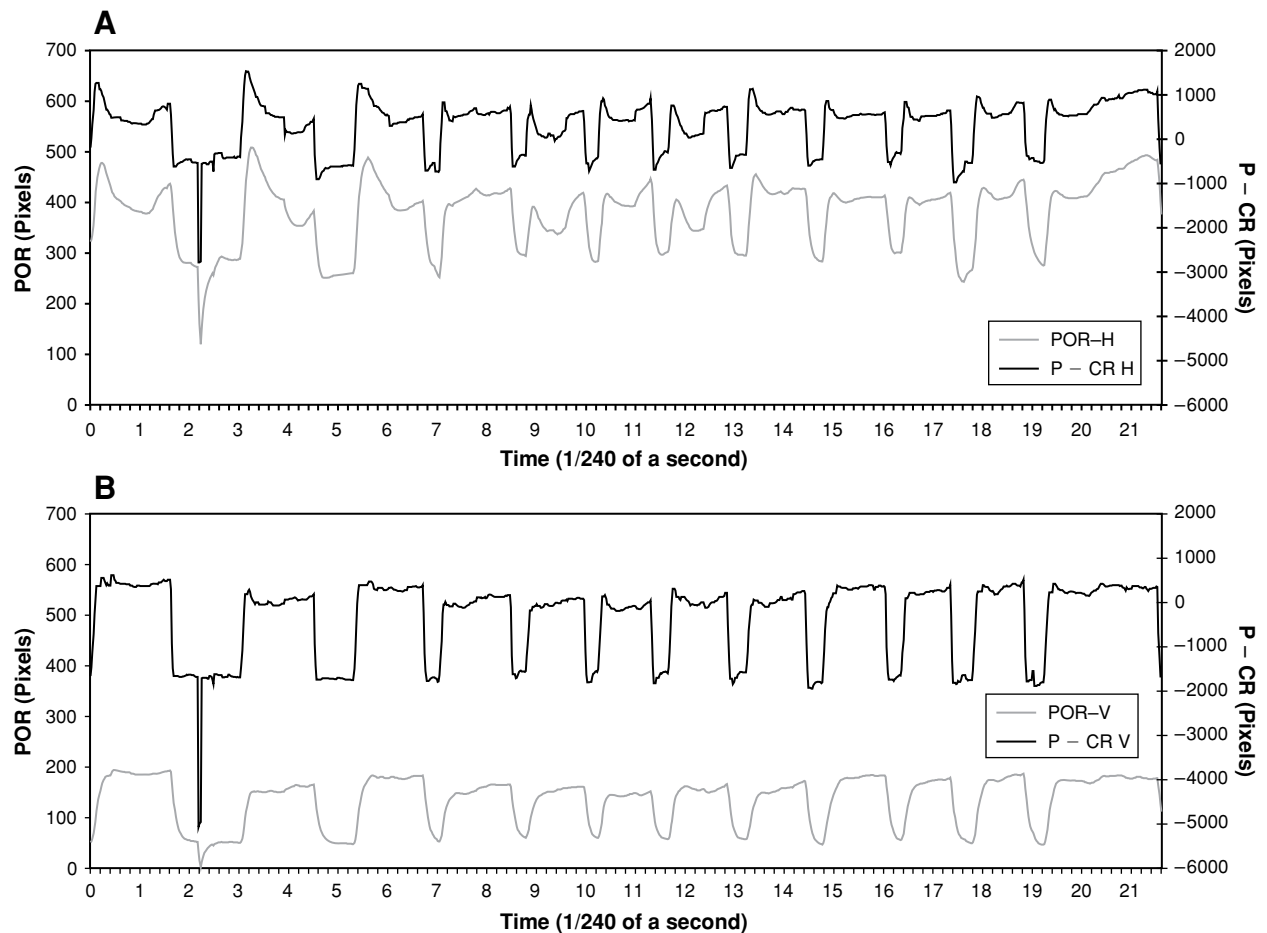
a running average of the eye's position over a predefined number of samples. Large fluctuations inconsistent with surrounding data are often dropped through a process largely unknown to the user. Thus, the POR is less dependent on slight fluctuations in the eye's position that may result from physiological nystagmus, sudden tremors in the eye, missing samples, or false samples of the pupil or corneal reflection from a secondary infrared source—that is, sunlight. As a secondary result of this averaging, saccades and smooth movements appear delayed from their actual position in time. Furthermore, the start and end of fixations are also averaged with surrounding movements. Although dampening causes discrepancies between raw eye positions and the calculated POR, without some averaging, fluctuations in the POR, visible in the scene video and data recording, would be too large to be useful in identifying an object of focus.

Therefore, the interpretation of a subject's scan path through the POR inherently overlooks small details present in the raw eye positions. As we begin to further understand the perception of visual information during saccadic movements (García-Pérez & Peli, 2001), the raw eye positions are critical for understanding details about the gaze trajectory.

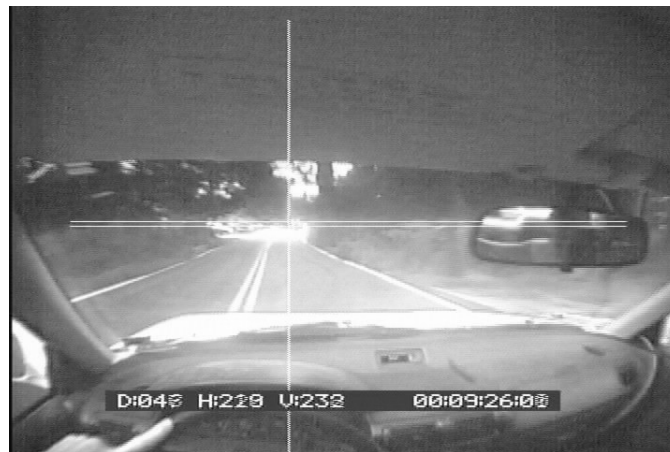
Thus, a subject's scan path (Noton & Stark, 1971) is best characterized through the most detailed pattern of viewing discernible. An example below will illustrate the difference between raw eye positions, interpreted through the centroid of the pupil minus the centroid of the corneal reflection of the eye camera ( $P - CR$ ) and the POR calculated on the basis of a running average of coordinates in the scene camera. In Figure 1, eye positions, represented as  $P - CR$  errors and the horizontal and vertical components of the POR, are plotted against time. The dampening and displacement of the POR signal with respect to the eye position, represented through the  $P - CR$  error, are clearly evident. Therefore, as compared with eye movements classified through  $P - CR$  errors, transitions between eye movements classified on the basis of the POR are harder to differentiate. The construct of these differences will be discussed later.

**Summarizing Subjects' Visual Behavior**

Consider the scene image in Figure 2, where the POR, indicated by the crosshairs, shows that the subject's POR is fixed on a region down the center of the roadway. What the image does not show, and this becomes apparent



**Figure 1.** A subject's gaze path, horizontal (A) and vertical (B) dimensions, represented as the pupil minus corneal reflection error ( $P - CR$ ) and point of regard (POR).



**Figure 2.** A frame of the scene video, recorded from the driver's perspective.

in later frames, is that cars appear parked in driveways near the road and an oncoming vehicle is approaching. Therefore, from this image alone, it is difficult to identify whether the subject's visual attention is focused down the road, unaffected by the oncoming vehicle. A comparison of the POR with subsequent frames provides more information: If a shift in attention to the oncoming car occurs, it would suggest that during the fixation before the shift, the subject was not monitoring the oncoming car at the center of the visual field. In more complex scenes, road signs, multiple oncoming vehicles, lead vehicles, objects on the side of the road, or devices inside the vehicle all compete for a subject's attention. These objects and devices may require short glances from the driver. One of the goals in identifying driver behavior patterns is to identify the subtasks that join together to make up the driving task, with the classification of subtasks largely dependent on the type of comparison that will result. A preclassification of complex tasks into different additive steps, similar to a GOMS analysis (Card, Moran, & Newell, 1983), can greatly facilitate this analysis. Since different tasks can be decomposed into different combinations of similar subtasks (Kieras, Wood, Abotel, & Hornof, 1995), behavior recognition relies on a multilevel model that reduces a problem into subproblems (task sets) and actions.

In any visually oriented search task, the process of solving and analyzing a problem will be composed of saccades that correlate with attention shifts and fixations/smooth movements that relate to the locus of perception. In a dynamic environment, fixations and smooth movements relate to the motion of information being perceived over different regions of the scan path. Thus, as an indication of a solution methodology during a visually oriented task, two methods of comparing different patterns of eye movements are possible. In the first method, inferences of driver behavior are correlated with measurable effects, such as the number of glances to a target or mean glance time, whereas the second method uses hidden Markov models (HMMs), a statistical model that assumes a hidden pro-

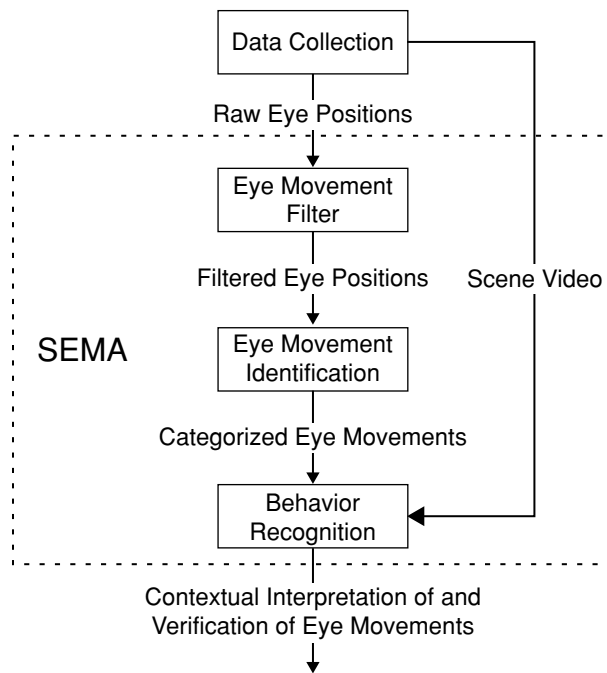
cess that generates observed behavior (Rabiner & Juang, 1986), to identify significant parameters, in addition to basic measurable effects associated with the eye movement patterns. In addition to computing basic summary statistics, an HMM can be used to infer the probability of future states (Rabiner & Juang, 1986).

This article presents a novel method of filtering recorded eye positions and identifying eye movements. An overview of the semiautomated eye movement analysis (SEMA) system is provided in Figure 3. The algorithm is separated into three distinct steps, each encompassing distinct features often found individually in other methods of eye movement analysis. The proposed combination of methods presented in this article appears to be the first documented single multipass system that combines multiple methods of eye movement analysis. In Stage 1 (damping), recorded eye positions are cleaned to remove noise in the data recording. Stage 2 (eye movement identification) uses two thresholds and four duration minimums to automatically identify fixations, saccades, and smooth movements from the recorded eye positions. In the final stage (behavior recognition), a scene video is used to systematically evaluate each identified movement and to assign an associated corresponding behavior to the eye movement pattern. In summary, the model categorizes potential eye movements directly from raw eye positions. Information from the dynamic environment (scene video) is then used at discrete points to evaluate potential movements and to classify the allocation of visual attention to different stationary and moving stimuli.

## METHOD

### Filtering Recorded Eye Positions

Eye position data records are affected by three sources of error: falsely recorded positions, blinks, and missing data. The rate at which each type of error occurs is influenced by several factors, including the anatomy of the subject's eye, facial anatomy, the amount of extraneous infrared light, and the design of the eye-tracking equipment. In subjects



**Figure 3. Overview of data flow through the eye position filter, eye movement identification, and behavior recognition algorithms. SEMA, semiautomated eye movement analysis.**

with large eyelashes or small gaps between eyelids, the corneal reflection and/or pupil can be obscured from view during larger degree movements. This is most evident in bright environments where, as a secondary response, subjects with sensitive eyes often squint to reduce the amount of light entering the eye. Thus, eye movement studies in outdoor environments often have a significantly higher rate of missing samples, in comparison with laboratory conditions. Although the rate of missing samples can be reduced by careful subject selection and experimental controls, in many experimental scenarios, subjects and environments are difficult to select without introducing bias.

In outdoor environments, false positions are most often recorded as the result of extraneous high-intensity infrared light reflections. This reflection introduces error through a false corneal reflection. Another error usually associated with dark pupil eyetrackers occurs when the area around the eye appears darker than the pupil in the illuminated image. This is caused by a change in the brightness of ambient conditions that changes threshold levels for the eye image. Small errors in corneal reflections are difficult to detect, whereas large errors in pupil positions and corneal reflections away from the eyetracker's infrared illuminator, such as positions not on the surface of the eye, are easier to identify.

Two methods of identifying false eye positions have been considered. The first method finds false eye positions that, based on an incorrectly identified pupil or corneal reflection, are theoretically beyond the range of the eye camera. The second method identifies false eye positions that do not match interpolated values.

Blink errors are associated with the eye's being partially open or closed. Obviously, when the eye is closed, the eyetracker should not report a pupil or corneal reflection. False positions, however, are often recorded as reflections off the eyelids or the eye socket. False corneal reflections and/or pupil positions are also recorded when the eye is partially closed.

Errors due to missing data occur when either the corneal reflection or the pupil moves out of the range of the eye camera. The situation often occurs with saccadic movements to a short fixation on a peripheral stimulus. The off-axis movements are often brief and, therefore, not fully followed by a head movement to recenter the eye within the tracking range. Unlike blink errors, the saccadic movements in and out of the feasible region are easily mistakable for the eye's opening and closing. However, unlike blink errors, they are valid samples that may provide some information about the fixation points. Although blink errors result in a set of false position records, culminating in a period during which the eye is noted as being shut, missing data occur as false recordings (outliers) embedded in a set of feasible measurements. Thus, blink errors and missing data points can be easily misidentified, but they are important to distinguish for inference purposes.

To label samples that may be a result of one or more types of recording errors, a four-step filter-and-interpolation algorithm has been developed. This algorithm isolates samples that have a high probability of being false positions. Changes in the horizontal and vertical  $P - CR$  error are used as an indicator of a false eye position. Eye positions are categorized by the maximum displacement between time  $t$  and time  $t-1$  or  $t+1$ . Samples with a displacement faster than a saccade ( $900^\circ/\text{sec}$ ) are considered defective. For positions where  $t-1$  or  $t+1$  are not present in the data set, recording errors surrounding blinks and missing data are labeled defective. Extraneous samples unconnected to the surrounding scan path are labeled defective. Finally, linear interpolation is used to validate data over small ranges of missing eye positions.

The filtering algorithm uses 10 parameters, subdivided into three categories. System-dependent parameters relate to characteristics of the eye-tracking equipment. Theoretical parameters are based on assumed properties of eye movements, and subjective parameters are used to allow interactive control by the user. A complete list of parameters is detailed in Table 1. For most common head-mounted eye-tracking systems, the system-dependent parameters can easily be identified from the eye camera's specifications and recording software. However, for some systems, the eye camera's aspect ratio, *rel*, needs to be calculated. All applications of the eye movement identification algorithm must account for the variability of the eye camera's position in relation to the eye. The process of calculating the number of pixels in the eye camera that relate to a degree of movement in the visual field, *ppd*, is often implemented through the analysis of subjects' fixation patterns recorded during their performance of tasks constrained within known visual angles—that is, calibration points (Reimer, 2003). With experiments completed in situ, recording er-

**Table 1**  
**Parameters of the Eye Position Filtering Algorithm,**  
**Including Categories and Descriptions**

Category	Parameter	Description
System	$max_h$	The maximum feasible horizontal P – CR error that is within the range of the eye camera
	$max_v$	Same as $max_h$ , except for the vertical dimension (note that the horizontal and vertical coordinates are not equidistant)
	$ppd$	The calculated number of pixels in the eyetracker's coordinate system that approximates 1° of movement (note that this parameter is subject dependent)
	$rate$	The eyetracker's sampling rate (fixed by equipment)
	$rel$	The relationship between horizontal and vertical dimensions of each pixel in the eye camera (fixed by equipment)
	$run_a$	The number of samples averaged to form a point-of-regard measurement (adjustable in ISCAN's software)
	Theoretical	$max_c$
Subjective	$min_g$	The minimum number of consecutive good samples
	$max_i$	The maximum number of missing samples that can be filled through interpolation
	$widen$	The number of samples, before and after each set of missing data, that should be deleted

rors can largely influence data. Subjective filtering parameters provide user control over the level of accuracy with which the model assumes data are recorded.

The first subjective parameter,  $min_g$ , requires that a minimum number of contiguous good samples be present in the data set. Therefore, short extraneous samples unconnected to surrounding movements, which usually represent false eye positions and blink errors, are filtered from the data set. The value of this parameter often ranges from the minimum length of a saccade to that of a fixation, with the latter chosen in applications in which less secondary, manual correction is desired at the expense of failing to identify potential saccades. The number of samples marked for removal from the start and end of each data set,  $widen$ , provides a user with a method of eliminating false eye positions recorded as the subject blinks or as the eye moves in and out of the eyetracker's range. With eye-tracking systems recording at higher sampling rates, lower values of  $widen$  are more applicable. With the tested 240-Hz system, values greater than one are rarely considered. Finally, the maximum number of samples filled through interpolation,  $max_i$ , provides the user with some method of validating data over small missing segments. Unlike the previous parameters,  $max_i$  is best considered in relation to the overall accuracy of the identified eye movements; larger values presume a linear fit between adjacent points and require a higher degree of secondary verification to distinguish a smooth movement from a saccade and a fixation. Therefore, in most applications,  $max_i$  is limited to a few samples less than the minimum length of a saccade.

### Eye Movement Identification

Recorded eye positions are a discrete sample of a subject's eye movements that form the scan path (Buswell,

1935; Noton & Stark, 1971; Yarbus, 1967). One of the commonly used methods of analyzing eye position data is fixation identification, in which a subject's scan path is summarized by a series of fixations and intermediate periods (Salvucci, 1999). This is problematic, however, when the environment is dynamic. The analysis assumes (incorrectly) that fixations represent the only periods during which visual information is perceived (Salvucci, 1999). In a dynamic environment, subjects may also perceive information during smooth movements and, possibly, saccades (García-Pérez & Peli, 2001). Thus, fixation identification is, at best, a special case of eye movement identification. Furthermore, in experiments with unrestricted head movements, a scene video provides an indirect link between a set of recorded eye positions and the location of objects in the scene.

To identify the features of a subject's scan path with greater time resolution, an eye movement identification algorithm is proposed that combines two features previously used independently for automated fixation identification: point-to-point velocity and dwell time with manual video analysis. The analysis seeks to maximize the time domain accuracy of the identified eye movements by using raw, unaveraged, P – CR data for tracking eye position. Every sample is first validated by attempting a label assignment. If the probability of association of a label with the point is low, it is deemed unclassifiable or an error. In the analysis, unclassifiable points include positions identified as errors during the process of filtering and positions that fail to be classified as eye movements. A complete representation of a subject's scan path results with the labeling of unclassifiable eye positions: those that represent periods in which visual perception did not occur—that is, the eye was closed—or periods in which eye positions could not be associated with an eye movement.

As a precomputation to manual video analysis, an automated eye movement identification algorithm places a preliminary assignment to each eye movement recorded in the scanpath. Initially, the algorithm uses two velocity thresholds to differentiate fixations from saccades and microsaccades. Saccades are characterized first, so that the beginning and end of each movement, where the point-to-point velocity approaches that of the surrounding fixations and smooth movements, are correctly classified. Microsaccades are identified next on the basis of the velocity limits, as well as the total duration of the movement. Subsequently, the sojourn durations in a neighborhood are used to separate stationary eye positions from tracking movements, resulting in the identification of fixations. Smooth movements are identified from the unlabeled positions by clustering based on durations. Finally, the remaining points are labeled as errors. In addition to the parameters *ppd* and *rel*, defined in Table 1, the eye movement identification algorithm requires a set of subjective parameters as defined in Table 2 (Reimer, 2003).

Thresholds, defined as a fraction of a degree of visual angle, are converted to pixels through the parameter *ppd*. This confines the between-subjects variability to the measure *ppd*. Similarly, durations, defined in relation to a number of samples, are converted to the time domain through the sampling rate of the eyetracker, *rate*. Values for subjective parameters are provided in connection with the subsequent example; however, in a more general case, parameter values should be considered in relation to Type I and Type II errors, such as those that describe fixation and saccadic thresholds in Reimer (2003).

**Calculating the average eye position and rate of change.** When eye positions are tracked using raw P – CR errors, averaging is required to reduce variability. Although various moving point windows can be considered, only models that encompass an equal number of values before and after the current position dampen the recorded signal without introducing a time shift (lag). A bidirectional average can encompass any odd number of points. Averages encompassing a larger number of samples reduce a model’s ability to detect a sudden change in eye position—that is, the start of a saccade. Likewise, averages computed upon a smaller number of points are likely to provide too little dampening—that is, fail to reduce fluctuations from physiological nystagmus. Thus, an

optimal length average for eye movement identification is the minimum number of samples required that reduces fluctuations due to physiological nystagmus without overpowering small changes in eye position. In any case, the number of coordinates included in any average should remain substantially less than the POR calculated for presentation in the scene video of commercial systems.

With most 60-Hz head-mounted eye-tracking systems, three-point averages are likely to be suitable for most applications. However with a 240-Hz eye-tracking system, a three-point average retains too much variability. Therefore, this model, developed for a 240-Hz system, is defined using five-data-point averaging: two samples immediately preceding the current point, the current point, and two samples immediately after. Future implementations of the eye movement analysis algorithms, based on faster or slower recordings systems, are likely to readdress window size. Furthermore, to reduce the influence of missing and defective samples, the model excludes any calculated averages that would include a missing/filtered sample. The process of averaging data and calculating the rate of change has three steps. First, the average eye positions are independently calculated in horizontal and vertical dimensions for a given sample. Second, maintaining the independence between the horizontal and the vertical dimensions, the displacement in rectilinear direction is computed from the average eye positions. Finally, these orthogonal directions are combined to obtain the rate-of-change vector for each sample, using the parameters *rel* and *rate*.

**Initial threshold of eye position data.** The recorded eye positions are classified into saccades and fixations by comparing the calculated rate of change for a sample against two thresholds. Eye positions with a rate of change greater than *ppd/sac* are considered candidates for saccades, with positions exhibiting a rate of change below *ppd/fix* considered potential fixations. The thresholds *sac* and *fix* represent a fraction of a degree of movement per sample with respect to the saccadic threshold and the fixation threshold, respectively. Since the fractions *fix* and *sac* remain constant between subjects, intersubject variability is confined to *ppd*. Furthermore, it is possible to select the parameters *fix* and *sac* on the basis of the features of an experiment—that is, stimulus velocity and dispersion—to improve the separation of saccades and fixations from smooth movements.

**Table 2**  
**Parameters of the Eye Movement Identification Algorithm**

Parameter	Description
<i>min<sub>f</sub></i>	Minimum duration for a fixation (samples)
<i>min<sub>s</sub></i>	Minimum duration for a smooth movement (samples)
<i>min<sub>b</sub></i>	Minimum duration for a long saccade (samples)
<i>min<sub>l</sub></i>	Minimum duration for a short saccade (samples)
<i>sac</i>	Minimum rate of change for samples to be considered as a potential component of a saccade (fraction of a degree of visual angle)
<i>fix</i>	The maximum rate of change for a position to be considered as a potential component of a fixation (fraction of a degree of visual angle)

**Eye movement composition: Identifying fixations and saccades.** Using the labeled eye position data, duration minimums are now used to compose the classified components into eye movements. Saccades are identified first, for several reasons. Since they are of shorter duration than fixations, errors misclassifying samples as fixations when they are actually saccades have a disproportional affect on the saccadic durations. When a period with  $min_b$  or more potential components of a saccade has been sequentially identified, the entire subset is classified as a large saccade. Similarly, a microsaccade, which is a movement of less than  $1^\circ$  of visual angle (Young & Sheena, 1975), is identified, where the number of continuous samples is below  $min_b$  but equal to or above  $min_1$ . Samples identified as saccades (micro or large) are no longer considered as potential components of fixations, which are identified as a set of  $min_f$  contiguous data labeled as potential fixation components.

Although the actual values of  $min_b$  and  $min_1$  may vary according to the desired outcome of the model, the minimum duration for saccadic movements is most often considered to be 30 msec (Palmer, 1999). Microsaccades may be embedded in fixations to recenter the eye as it naturally drifts away from a point of fixation. Corrective saccades may occur during smooth movements or larger saccades to correct any over-/undershoot caused by ballistic eye movements. A low minimum duration,  $min_1$ , may not result in the identification of corrective saccades. A driver's minimum fixation duration is considered to be between 50 and 100 msec (Recarte & Nunes, 2000; Salvucci & Anderson, 2001; Unema, Dornhoefer, Steudel, & Velichkovsky, 1999).

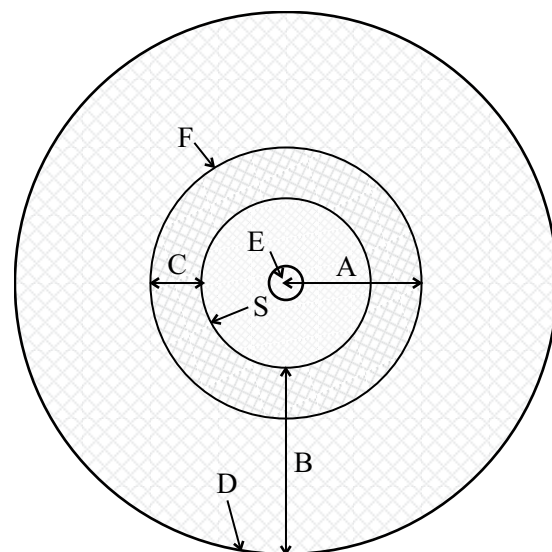
For in situ driving data collected using an ISCAN, Inc. (Burlington, MA) 240-Hz head-mounted eyetracker, a common starting value for  $fix$  is 2.5, and for  $sac$  4. In such cases, eye position samples are identified as potential fixations when the rate of change of eye positions is below four tenths of a degree per sample, and potential saccades are identified when a rate of change is determined to be greater than a quarter of a degree per sample. Because the velocity profiles of both types of movements overlap during the ramp-up and ramp-down periods of a saccade, some eye positions may be classified as both a potential fixation component and a potential saccade component. With precedence given to saccades, the eye movement identification algorithm assumes the latter. Comparisons of the classified eye movements and the scene video can be used to adjust the selected classification, thresholds, and durations and then recompute the model.

**Smooth movement identification.** Smooth movement identification uses a minimum duration,  $min_s$ , for all non-zero eye positions that remain unassigned after saccadic and fixation identification. The nonnull eye positions that remain unassigned after smooth movement identification are first considered for membership in neighboring fixations. If no neighboring fixation exists, the position is considered unclassifiable. This second stage classification forms the final correction for poor recordings by increasing the length of fixations to accommodate samples

that violate the fixation threshold. This assumes that the recording errors that surround a fixation are best represented as part of a neighboring fixation. Finally, the eye movement identification algorithm concludes the analysis by labeling all unclassifiable positions, blinks, poor recordings, and samples removed through other methods of filtering and averaging as null movements.

When the velocity profile of a set of recorded eye positions changes repeatedly over a short duration between the velocities associated with a saccade and that of a fixation, the set of positions is best identified as a component of a smooth movement. No minimum duration for a smooth movement has been reported in the literature. The maximum velocity of smooth movements, however, has been shown to be approximately  $100^\circ/\text{sec}$  (Palmer, 1999). The maximum velocity overlaps with that of a saccade; thus, it is assumed that the maximum point-to-point velocity of a smooth movement will occur continuously for a number of samples greater than  $min_1$ .

Figure 4 represents the point-to-point velocity between two eye position samples at time  $t$  and time  $t-1$ , with the center point representing the bidirectional average position at time  $t-1$ . Thus, the variability between samples due to physiological nystagmus is contained within Region E. A sample is classified as a potential component of a fixation if it remains within Region A, the area bounded by Region F, the fixation threshold ( $fix$ ). Similarly, a potential component of a saccade is classified if the position at time  $t$  falls within Region B, bounded by  $S$ , the saccadic threshold ( $sac$ ), and the maximum speed of a saccade,  $max_c(D)$ . Thus, movements in Region C are classified as



**Figure 4.** Parameters used for processing point-to-point velocity of recorded eye positions. A is the region of movement of the fixation, B is the extent of a saccade, and C is the overlap during this period. The thresholds used are fixation threshold (F), saccadic threshold (S), and maximum point-to-point velocity of a saccade (D), and the variability caused by the physiological nystagmus region is shown as E.



potential components of both a saccade and a fixation. The eye movement identification algorithm classifies a fixation where the rate of change of  $min_f$  or more contiguous samples fall within Region A. A corrective or short saccade is where  $min_i$  to  $min_o$  contiguous samples fall within Region B. A large saccade (attention shift) is classified where  $min_o$  or more samples fall within Region B. Finally, smooth movements are considered where  $min_s$  contiguous samples fall between Regions A and B.

**Validation of the proposed movement identification.** Given the individual differences and the variability of recording conditions, a general method of validation is difficult and will depend on the algorithms used, the environment, and the test individuals. Therefore, the manual verification step often implemented during behavior recognition provides a method of verifying and adjusting results according to the dynamic nature of the environment recorded through the scene video. Two applications of increasing complexity have been provided in Reimer (2003) as an attempt to illustrate the proposed algorithms. The first experiment instructed subjects to follow a predefined target, so that in the absence of extraneous sojourns, an exact representation of the scan path

could be reconstructed and an analysis of saccadic length computed. Subsequently, an experiment involving simulated driving was presented to illustrate system accuracy under more complex laboratory conditions. The results indicate that under laboratory conditions, eye movement identification provides basic movement recognition with a high degree of accuracy.

**An example of eye movement identification.** The methods discussed above are used to filter a set of raw eye positions, as shown in Figure 5. These were recorded during in situ driving, using a 240-Hz head-mounted eyetracker manufactured by ISCAN Inc., and represent a variety of different eye movements. The locations of the start and end of each eye movement have been determined by the eye movement identification algorithm (Figure 6). Some data errors are clearly visible in the plot (Figure 5). The large spikes correspond to blink errors, recorded when the eye appears to be closed. The filtered eye positions demonstrate how the algorithm removes poorly recorded eye positions from the data set. The average eye position plots (Figure 6) “smooth” fluctuations between samples without affecting the time of onset of different movements. The identified eye movements are observed

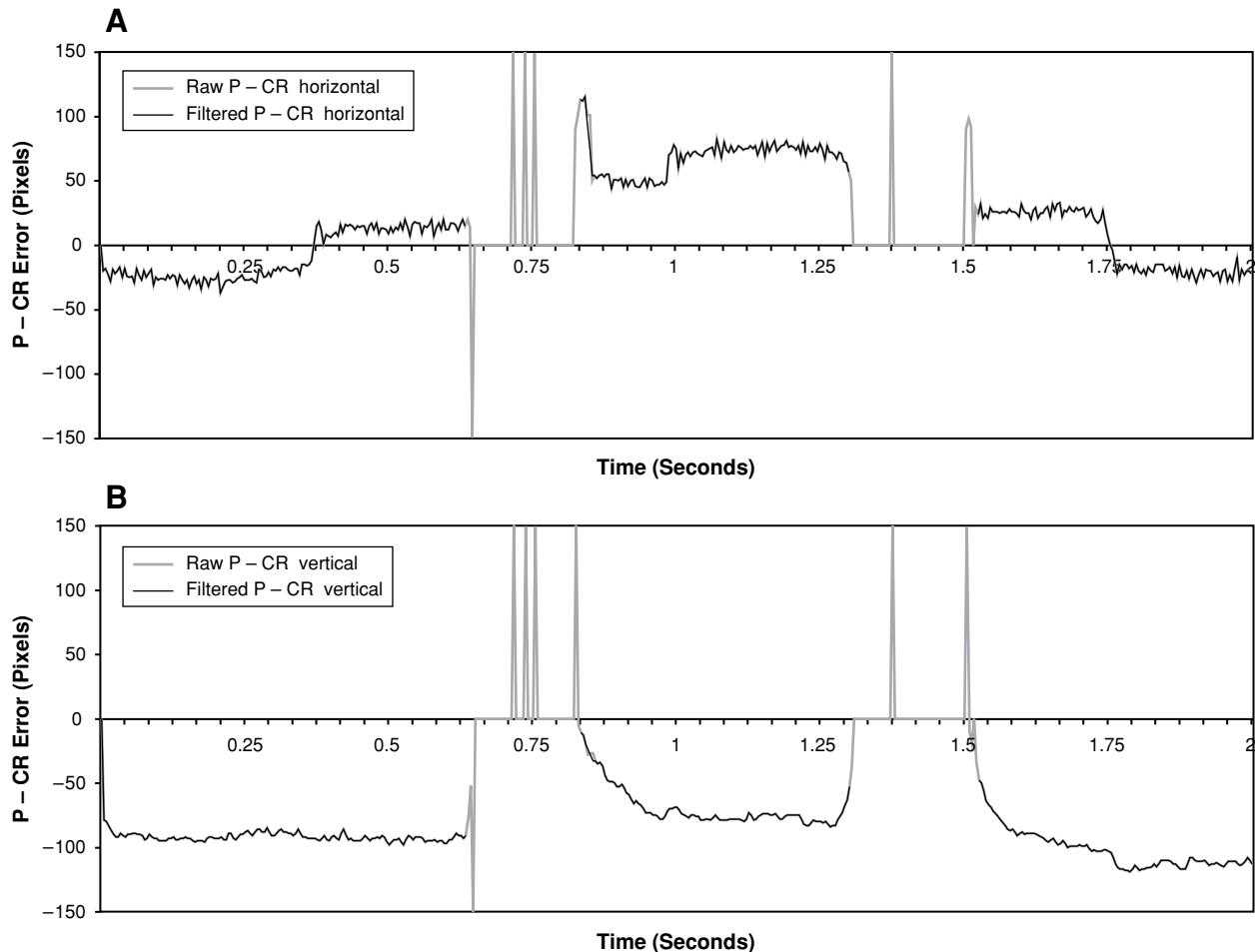


Figure 5. Raw and filtered eye positions in the horizontal (A) and vertical (B) dimensions.

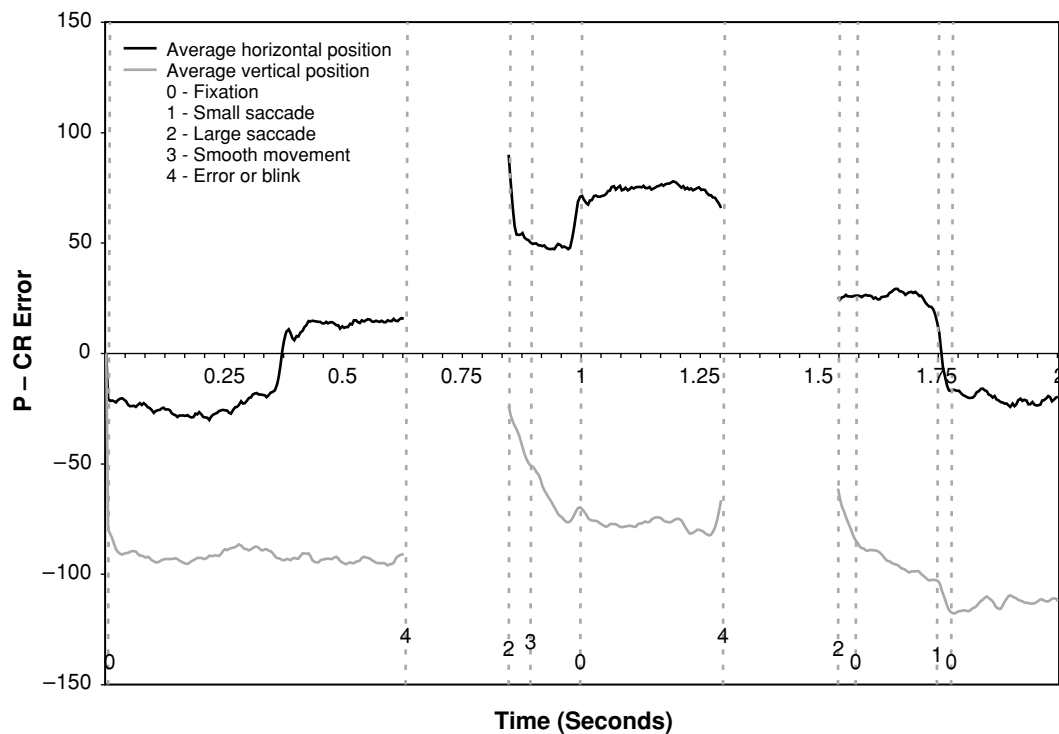


Figure 6. Average horizontal and vertical eye positions computed from the filtered data (Figure 5), with the identified eye movements.

in reference to the scan paths from which they were classified. Included are representations of a long and a short saccade, a smooth movement, a fixation, and errors or blinks. The difference in velocity between a smooth movement and a saccade are observed just after three quarters of a second. The high-velocity, saccadic movement appears to end in the horizontal dimension well before it ends in the vertical. The lower velocity observed in the trailing end of the vertical motion, characterized as a smooth movement, represents a subject's tracking a vertically moving stimulus. In the following section, the identified eye movements are manually verified, using information contained within the scene video and superimposed POR.

**Manual video-based correction.** The eye movement identification algorithm results in a first-pass identification of the composition of the scan path including smooth movements. In the second stage, (manual) identification uses features in a corresponding scene recording to verify and, if necessary, reclassify the distant movements. The classification compares each identified eye movement with the corresponding POR in the recorded scene video. Objects of focus and surrounding frames provide the final verification of the proposed analysis. Therefore, the differentiation of smooth movements from fixations involving drift is easily interpreted on the basis of an evaluation of a discrete number of frames from the scene video. The second-pass manual filter is implemented in parallel with behavior recognition, the manual linking of scan path components with task-related constructs. In fu-

ture implementations, automated processing of the scene video, being developed separately, will be linked with the eye movement identification algorithms, thereby reducing the number of steps currently computed manually.

### Behavior Recognition

The contextual interpretation of eye movements is necessary to obtain information about the effect of distractions on driving performance. Consequently, after eye movements have been identified, it is necessary to examine these with reference to the conditions under which they were generated. In addition to the eye position data, a second data source, the scene video, contains the dynamic visual field from the subject's perspective. The duration and transition between different visual stimuli, as noted from the scene video, can enable a comparison of the behavior patterns between or within subjects.

In general, eye movements identified through P - CR errors are directly related to the POR recorded as an overlay in a scene video of a head-mounted eyetracker. With rapid shifts in the scene video's POR, it is difficult to summarize a subject's behavior visually while simultaneously viewing the recorded scene video at normal or slow speeds. With a precise synchronization between the identified eye movements and the scene video through a video overlay of the recording time, as displayed in Figure 2, it is possible to make inferences about a subject's behavior by using individual frames of the scene video corresponding to periods surrounding the start and end of each identified eye movement.

The finite search space, therefore, reduces manual computations from traditional frame-by-frame scene analysis.

**Semiautomated Eye Movement Analysis**

A semiautomated eye movement analysis (SEMA) system (Reimer, 2003) has been developed for eye movement identification and behavior recognition using P – CR errors. Saccades, smooth movements, fixations, and errors can be identified on the basis of specified parameters. The system integrates file handling, data/video relation, eye position filtering, eye movement identification, and behavior recognition into one application. Given the appropriate theoretical, system, and subjective parameters, SEMA first automates the filtering and eye movement identification algorithms. It then provides an operator with an interface to assist in linking behaviors observed in the scene video to identified eye movements. The automation of the filtering and eye movement identification algorithms offers the flexibility to adjust subjective parameters rapidly and to rerun the filtering and eye movement identification algorithms. Therefore, SEMA provides a method of manually selecting the best filtering and identification models for each particular data set. SEMA outputs include the number of recorded eye positions, the number of null samples before filtering, and the number of null samples after filtering. Similarly, the eye movement identification algorithm provides a summary of the number and duration, in seconds, of each movement type identified, including saccades, smooth movements, fixations, and nulls.

The process of linking a set of recorded data to a scene recording is completed through the interfaces, illustrated in Figures 7 and 8. The eye movement navigation interface (Figure 7) assists an operator in browsing through the video data with tools to define, manage, and link actions

to an active eye movement. The media player (Figure 8) complements the eye movement navigation interface by providing a utility by which to locate the first frame in the scene video that corresponds to the current eye movement displayed in the navigation interface. An operator can then interpret the subject’s point of focus, using the navigation tools in the media player to view surrounding frames of the scene video, and then skip to the start of the next eye movement.

In summary, SEMA is composed of four components: an eye position filtering component, a rule-based eye movement identification module, a scene video and data file relationship module, and an interactive behavior recognition tool. Behavior recognition consists of three components: identification/verification of eye movements in the scene video, manual relation to behaviors, and mapping. In the identification stage, an operator takes a classified eye movement and locates it in the scene video, using the controls provided in the eye movement navigation interface. The media player then allows the surrounding periods in the scene video to be considered in the linking of an eye movement to a defined task set and action.

**An Example**

Eye positions recorded, using a 240-Hz head-mounted eyetracker, from a driver navigating two roadways are used to demonstrate the utility of SEMA. Two data sets, consisting of 36,043 and 35,003 eye positions, 150.18 and 145.84 sec, were recorded, for a total of 71,046 samples. The first set of data was recorded on a rural road with an average of 3.33 oncoming cars per minute and a speed limit of 25 MPH. The second set of data was recorded on a rural road with an average of 16.66 oncoming cars per minute and a speed limit of 35 MPH.

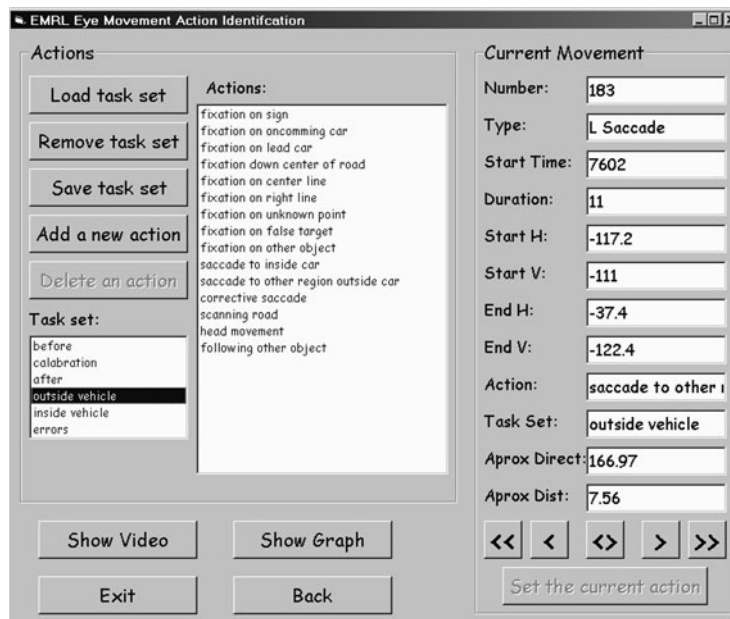


Figure 7. The eye movement navigation interface.



Figure 8. The eye movement navigation video review interface.

**Summary of the analyses.** Using the parameters listed in Table 3, the filter and eye movement identification algorithms are applied to the collected data. The eye movement identification algorithm identified 1,368 eye movements (688 in Road Segment 1 and 680 in Segment 2) from the 71,046 eye positions. The eye movements were manually separated into two categories: those associated with actions outside the vehicle and those associated with actions inside the vehicle. Not all the eye movements could be segregated into these categories. When the operator was unable to categorize a particular eye movement, it was labeled as an error. A finer resolution of the categories is summarized in Table 4. In total, the manual labeling of the data took approximately 3.5 h and represents a reduction of the analysis of 8,880 frames to 1,368 positions associated with the start of each new eye movement.

Figure 9 graphically represents the results of the analysis of an arbitrary 12-sec record of the eye movements. Large shifts in eye position are easily identifiable in the figure. These shifts indicate transitions between task categories. Breaks in the scan path may be due to blinks or misrecorded data that are often found in scan paths recorded outdoors with head-mounted optics.

**Table 3**  
Values Used for the Filtering and  
Eye Movement Identification Algorithms

Parameter	Value	Parameter	Value
$max_i$	5	$min_f$	24
$min_g$	24	$min_s$	12
$max_c$	900	$min_b$	8
$ppd$	10.83	$ppn_1$	6
$rel$	1.62	$sac$	4
$run_a$	10	$fix$	2.5
$widen$	1		

**Comparing the data sets.** Statistics for the eye movements of each category were compared for the two data sets. This was done in order to assess differences in the driver's behavior over the two segments of roadway. As a reminder, Segment 1 was a roadway with a speed limit of 25 MPH, and Segment 2 was a roadway with a speed limit of 35 MPH. Table 5 summarizes the components of the two scan paths presented in Table 4. The goal of this comparison was to demonstrate the utility of eye movement analysis in facilitating the analysis of recorded eye positions.

In Table 5, 16 out of the 242 fixations are associated with in-vehicle tasks in Segment 1, as compared with 31 out of 287 fixations in Segment 2. In the first segment, this translates to an average of one glance inside the vehicle every 9.39 sec, as compared with one glance every 4.70 sec in the second segment. A similar comparison can be done for noncorrective saccades. In the first segment, there are 86 noncorrective saccades, as compared with 159 in the second segment. This implies that the drivers shifted their attention, on average, every 1.74 sec in the first segment, as compared with every 0.92 sec in the second segment. Statistical tests on the data for quantifying the significance of the differences are shown in Tables 6 and 7.

Using the methods and algorithms presented in this article, it is also possible to compute a state transition matrix. This matrix represents the probability of transitioning from any state to any other state. If the entire data record is included in the analysis, this matrix is analogous to the transition matrix of a Markov chain—that is, the sum of each row's probabilities is unity. An illustration is provided in Table 8.

**Discussion.** From the data (Table 4) and statistics (Tables 6 and 7), differences in driving patterns for the two road segments can be determined. From Table 6, the

**Table 4**  
**Summary of the Classification of Eye Movements Into Those Related to Actions Outside the Vehicle, Those Related to Actions Inside the Vehicle, and Errors**

ID	Action	Section 1				Section 2			
		No.	Avg. for Samples	SD for Samples	%	No.	Avg. for Samples	SD for Samples	%
Actions Outside the Vehicle									
1	Fixation on sign	10	188.60	157.62	1.37	2	82.00	21.21	0.28
2	Fixation on oncoming car	6	90.67	69.37	0.82	23	157.35	235.45	3.19
3	Fixation on center of road	112	99.69	65.11	15.38	102	95.09	68.06	14.17
4	Fixation on center line	15	99.00	58.00	2.06	17	101.59	58.94	2.36
5	Fixation on right line	40	90.93	54.07	5.49	46	96.04	80.38	6.39
6	Fixation on unknown point	24	65.17	31.25	3.30	1	39.00		0.14
7	Fixation on other object	19	94.79	86.42	2.61	40	62.63	29.62	5.56
26	Fixation on lead car					25	69.64	33.63	3.47
8	Saccade to inside	11	19.18	6.68	1.51	14	23.36	7.70	1.94
9	Saccade to outside	70	8.33	3.5	9.62	134	8.67	5.01	18.61
10	Corrective saccade	165	12.20	4.55	22.66	119	7.63	2.63	16.53
11	Scanning road	13	20.15	8.09	1.79	1	15.00		0.14
12	Head movement	1	22.00		0.13	4	15.75	7.89	0.56
Actions Inside the Vehicle									
13	Fixation on rear-view mirror	7	52.00	18.25	0.96	12	84.00	80.93	1.67
14	Fixation on left mirror	1	39.00		0.13	8	106.75	137.59	1.11
15	Fixation on dash	5	57.40	12.76	0.69	5	108.40	46.07	0.69
16	Fixation on other region	3	46.67	34.27	0.41	6	72.17	37.40	0.83
17	Saccade to outside	3	14.67	4.04	0.41	2	12.50	6.36	0.28
18	Saccade to inside	2	6.50	0.71	0.27	9	8.78	5.09	0.13
19	Corrective saccade	3	5.33	0.58	0.41	8	5.875	0.64	1.11
20	Head movement	5	12.80	7.22	0.69	3	17.67	2.89	0.42
27	Scanning					1	19.00		0.14
Errors									
21	Possible blink	149	44.43	11.73	20.47	114	40.34	10.07	15.83
22	Recording error (unknown)	2	24	4.24	0.27	6	58.5	61.70	0.83
23	Recording error (sun)	31	98.77	111.10	4.26	4	132.50	42.47	0.56
24	Unclassifiable point	16	3.31	2.75	2.20	7	4.42	2.15	0.97
25	Part of previous movement	15	8.00	5.71	2.06	7	6.71	2.21	0.97

Note—Avg., average duration; SD, standard deviation of duration; %, percentage of total movements represented by each state.

number of glances directed inside the vehicle and outside were significantly different at  $\alpha = .05$  for the two road segments tested. The number of glances outside the vehicle was larger for the road segment with a higher traffic density, as would be expected. Mean durations would be expected to account for this upward shift in the number of glances. In Table 7, differences in glance patterns for a number of tasks are compared for the two road segments. Of these, when driving on the higher traffic density road, the driver devoted more time to monitoring oncoming cars and to the car ahead. There is (statistically) no difference in the glance patterns directed toward road markings, such as the centerline and the right-lane marking. The differ-

ence in durations of glances directed toward the signs is significant at  $\alpha = .05$ , as are the durations for glances toward the center of the road and toward other objects.

The intent of this analysis is to illustrate the methodology; it does not purport to present conclusions about the nature of drivers' glance patterns. The statistics shown above need to be supported by additional testing with a larger set of subjects than is shown here. Additional testing is required to form general conclusions on the topics discussed. Although the methods are computationally intensive, it is now possible to use such analyses for the evaluation of subjects in complex environments. The ability to compute a transition matrix can also facilitate safety

**Table 5**  
**Summary of the Total Number of Classified States by Task Set and Type of Movement**

Task Set	Section 1			Section 2		
	Fix.	Sac.	Tot.	Fix.	Sac.	Tot.
Outside	226	246	486	256	267	528
Inside	16	8	29	31	19	54
Error			213			138
Total	242	254	728	287	286	720

**Table 6**  
**Comparison of the Number of In-/Out-of-Vehicle Glances During the Two Road Segments**

Task Set	Fixation		Saccade		Total	
	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>	$\chi^2$	<i>p</i>
Outside vehicle	7.19	.007	5.74	.012	7.46	.006
Inside vehicle	5.36	.021	4.93	.026	8.28	.004
Error					20.07	.000

Note—With  $df = 1$  and  $\alpha = .05$ ,  $\chi^2(2) = 3.841$ .

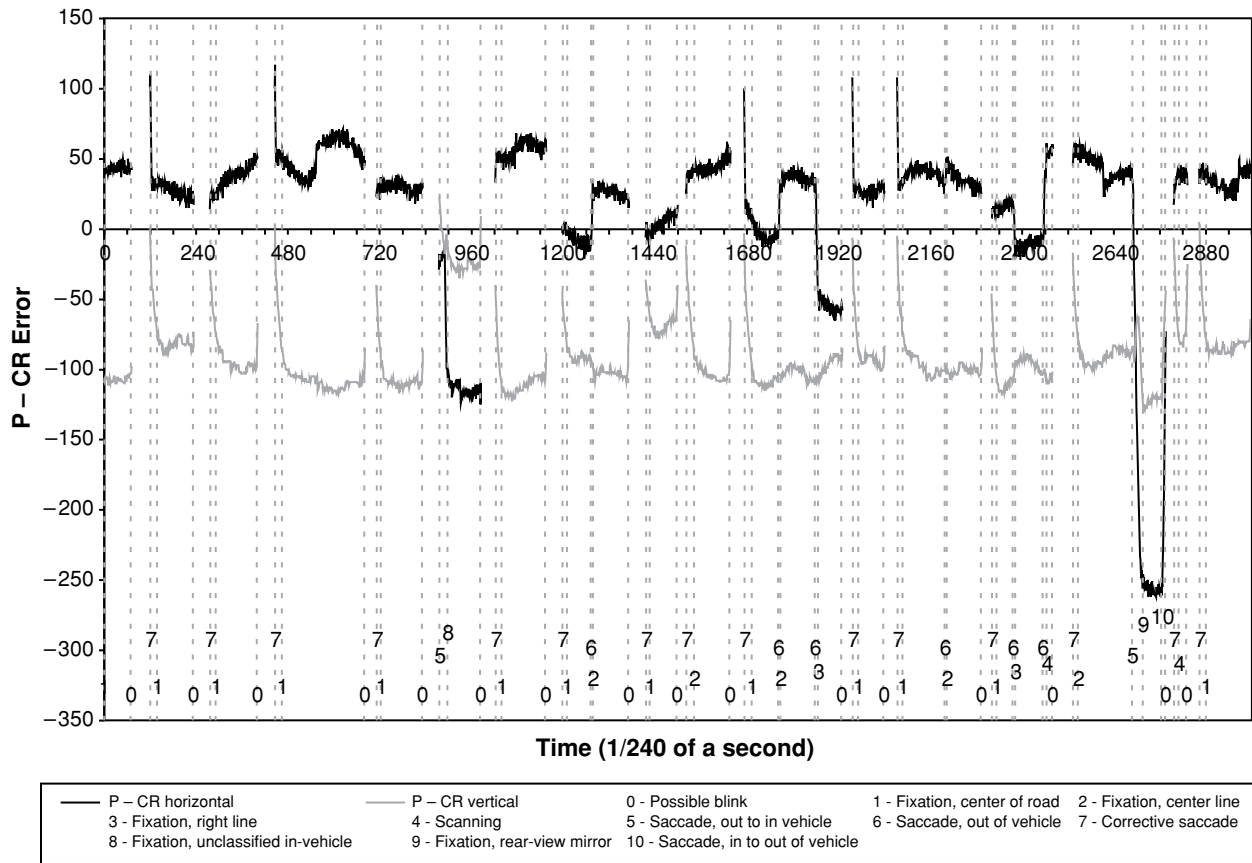


Figure 9. Twelve seconds of a driver’s scan path, with the state of each eye movement and the mapped action identified.

calculations for in-car devices, on-road distractions, and so forth. As was stated earlier, the entire analysis took 3.5 h of computation/analysis effort, and any subsequent analysis is likely to be done with less effort, which is a necessary support for large-scale testing with a diverse group of subjects and conditions.

**CONCLUSION**

This article has presented a method for analyzing eye positions recorded in a dynamic environment with a head-mounted eyetracker. This methodology presented in this

article clearly extends previous research in three distinct ways. It extends previous methods of analysis of eye movements, which have been based on fixation identification in a static scene, to dynamic environments. Second, automated identification of eye movements provides a “key” to facilitate the systematic manual categorization of scene video. Finally, the publication of associated algorithms provides researchers in areas outside basic oculomotor research with a scheme for systematically analyzing eye movement data. The processing is divided into three stages: filtering raw eye positions, classifying eye movements, and mapping eye movements to hypothesized actions.

The unique model presented for filtering recorded eye positions related to poor measurements provides input to a first-pass model whose focus is on identifying all the categories of eye movements that relate to shifting and maintaining attention. So that an individual’s behavior over all periods of time is accounted for, the overall representation of a subject’s scan path includes a classification of all unidentifiable periods. A method systematically evaluating the target of focus and manually mapping each eye movement to a hypothesized action has been proposed and illustrated through a software interface, SEMA.

The models developed for eye movement identification provide a unique and flexible methodology that can

**Table 7**  
Comparison of the Number of Glances Associated With Each In-Vehicle Task During the Two Road Segments

Action	$\chi^2$	<i>p</i>
Sign	6.56	.010
Oncoming car	8.50	.004
Down center of road	4.59	.032
Center line	0.00	.999
Right line	0.01	.938
Other object	5.82	.016
Lead car	23.27	.000
Unknown object	25.54	.000

Note—With *df* = 1 and  $\alpha$  = .05,  $\chi^2(2)$  = 3.841.

**Table 8**  
**Transition Matrix for Roadway Section 1 ( $\times 10^{-2}$ )**

	1	2	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	23	24	25	26	27	
1								10		10											80					
2										17												83				
4								2	24	5	1											62		6		
5								7	27													66				
6								8	40													44		8		
7									46	4												46		4		
8								5	26	11												58				
9													46		18	18					9					9
10	3	6	25	14	13	19	11					1										3		1	4	
11	5	1	52	3	18	5	6					4		1												5
12			8					8	31	8													37		8	
13																										
14																		29			14					57
15																										100
16																						20	20	20		40
17																							33			34
18			33																							33
19																										34
20																										
21																										
23			1					1	1	96																20
24																										1
25			6		3					6	26	3														3
26																										53
27			20			20	7						6													6
																										7
																										81
																										7

be used in identifying differences in a series of subjects' behaviors. This represents the first documented model to address the need to gather complex eye position data from subjects in a dynamic environment in which unobtrusive methods of data recording are required. Examples of such environments would include drivers in their own vehicles interacting with various forms of technology, clinicians interacting with patients, and individuals involved in decision strategies in the home, at work, or at the store. Thus, eye movement analysis provides a key element in the further development of more complex in situ studies.

Although SEMA provides the methodology to be used in developing an understanding of the differences in drivers' behaviors, the manual correction for fixation and smooth movement separation and for the mapping process is too time consuming to be generally applicable. Future applications focusing on the automation of scene processing will reduce current time requirements and expand the model's applicability to larger data sets. Furthermore, with the development of automated scene processing, additional information from the scene may be considered in the classification of eye movements, while the manual requirements of eye movement analysis are decreased further, the primary goal of the current developments. The resulting ability to model large data sets will allow for a more precise understanding of state duration and transition rates, which could form the basis for identifying changes in subjects' scan paths under different levels of cognitive and physical load.

Obviously, additional data are required to form general conclusions about the effect of road conditions on driver behavior. Yet the example provided illustrates some of the detailed comparisons possible through the use of

eye movement analysis. The methodology is easily extended to studying general problems of interaction within static or dynamic scenes. In conclusion, this article has described the development of eye movement analysis and has provided an overview of a unique eye-movement-based modeling as a tool that can be used to further understand detailed differences in human behavior. The developed model for eye movement identification considers raw eye positions as indications of a subject's state. In the application being considered (driving), the information for objects in the scene video is not available and, thus, is not used during the classification. Although it may be possible to extract dynamic features from the scene video automatically with sufficiently powerful computational resources, our experience is that, for now, an analysis of this nature is computationally intractable and best done manually. In summary, the goal of our development is to reduce the labor required for manual processing of eye movement data in the context of a dynamic scene. In the context of this methodology, scene information is used for behavior recognition and to validate identified eye movements.

**LIMITATIONS**

Although SEMA provides a novel approach to the identification and classification of eye movements, some obvious limitations to the proposed algorithms must be considered. First, the methodology was developed around an ISCAN head-mounted eyetracker. Although the algorithms developed should easily be extended to other eye-tracking systems, this has not been tested. Second, this system was designed for the analysis of eye position data collected in situ in driving studies, where extraneous infrared light

from the sun often results in high noise levels. Being the first documented system of this kind, no previous methods outside of full manual analysis exist with which to benchmark the results. To accommodate for varying degrees of noise, often observed through changes in sunlight intensity and individual differences, changes in subjective parameters are designed to facilitate a wide range of user preferences. Therefore, guidelines for parameter selection must be identified for individual experiments.

## REFERENCES

- 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* (pp. 71-79). Amsterdam: North-Holland.
- BROWN, I. D. (1997). Prospects for technological countermeasures against driver fatigue. *Accident Analysis & Prevention*, **29**, 525-531.
- BUSWELL, G. T. (1935). *How people look at pictures: A study of the psychology of perception in art*. Chicago: University of Chicago Press.
- CARD, S. K., MORAN, T. P., & NEWELL, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Erlbaum.
- CHAPMAN, P., & UNDERWOOD, G. (1999). Looking for danger: Drivers' eye movements in hazardous situations. In A. G. Gale, I. D. Brown, C. M. Haslegrave, & S. P. Taylor (Eds.), *Vision in vehicles-VII* (pp. 225-232). Amsterdam: Elsevier.
- CHIANG, D. P., BROOKS, A. M., & WEIR, D. H. (2004). On the highway measures of driver glance behavior with an example automobile navigation system. *Applied Ergonomics*, **35**, 215-223.
- DINGUS, T. (2000). *Driver distraction: New features, new tasks, new risks* (Presentation). Available at [www-nrd.nhtsa.dot.gov/PDF/nrd-13/DingusT\\_ppt.pdf](http://www-nrd.nhtsa.dot.gov/PDF/nrd-13/DingusT_ppt.pdf).
- GARCÍA-PÉREZ, M. A., & PELI, E. (2001). Intrasaccadic perception. *Journal of Neuroscience*, **21**, 7313-7322.
- GOODMAN, M. [J.], BENTS, F. D., TIJERINA, L., WIERWILLE, W. W., LERNER, N., & BENEL, D. (1997). *An investigation of the safety implications of wireless communications in vehicles* (Rep. DOT HS 808-635). Washington, DC: United States Department of Transportation, National Highway Traffic Safety Administration.
- HARBLUK, J. L., NOY, Y. I., & EIZENMAN, M. (2002). *The impact of cognitive distraction on driver visual behaviour and vehicle control* (Rep. TP 13889 E). Ottawa: Ergonomics Division, Road Safety Directorate and Motor Vehicle Regulation Directorate.
- HEALEY, J., SEGER, J., & PICARD, R. (1999). Quantifying driver stress: Developing a system for collecting and processing bio-metric signals in natural situations. *Biomedical Sciences Instrumentation*, **35**, 193-198.
- HENDERSON, J. M. (1993). Visual attention and saccadic eye movements. In G. d'Ydewalle & J. van Rensbergen (Eds.), *Perception and cognition: Advances in eye movement research* (pp. 37-50). Amsterdam: North-Holland.
- KIERAS, D. E., WOOD, S. D., ABOTEL, K., & HORNOF, A. (1995). GLEAN: A computer-based tool for rapid GOMS model usability evaluation of user interface designs. In *Proceedings of the Eighth Annual Symposium on User Interface Software and Technology* (pp. 91-100). Pittsburgh: ACM SIGGRAPH.
- LAND, M., & HORWOOD, J. (1996). The relations between head and eye movements during driving. In A. G. Gale, I. D. Brown, C. M. Haslegrave, & S. P. Taylor (Eds.), *Vision in vehicles-V* (pp. 153-160). Amsterdam: North-Holland.
- LIU, A. (1998). What the driver's eye can tell the car's brain. In G. Underwood (Ed.), *Eye guidance in reading, driving, and scene perception* (pp. 431-452). Oxford: Elsevier.
- LIU, A., & SALVUCCI, D. (2001). Modeling and prediction of human driver behavior. In M. J. Smith, G. Salvendy, D. Harris, & R. J. Koubek (Eds.), *Usability valuation and interface design: Cognitive engineering, intelligent agents, and virtual reality* (pp. 1479-1483). Mahwah, NJ: Erlbaum.
- LIVERSEGE, S., & FINDLAY, J. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, **4**, 6-14.
- McKNIGHT, J., & McKNIGHT, A. S. (1991). *The effect of cellular phone use upon driver attention* (Report). Landover, MD: National Public Services Research Institute.
- MOURANT, R. R., & ROCKWELL, T. H. (1972). Strategies of visual search by novice and experienced drivers. *Human Factors*, **14**, 325-335.
- NILSSON, T., NELSON, T. M., & CARLSON, D. (1997). Development of fatigue symptoms during simulated driving. *Accident Analysis & Prevention*, **29**, 479-488.
- NOTON, D., & STARK, L. (1971). Eye movements and visual perception. *Scientific American*, **224**, 34-43.
- NUNES, L. [M.], & RECARTE, M. A. (2002). Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Pt. F: Traffic Psychology & Behaviour*, **5**, 133-144.
- PALMER, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.
- RABINER, L. R., & JUANG, B. H. (1986). An introduction to hidden Markov models. *IEEE ASSP Magazine*, **3**, 4-16.
- RECARTE, M. A., & NUNES, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology: Applied*, **6**, 31-43.
- REDELMEIER, D. A., & TIBSHIRANI, R. J. (1997). Association between cellular-telephone calls and motor vehicle collisions. *New England Journal of Medicine*, **336**, 453-458.
- REIMER, B. L. (2003). *Automated methods of detecting driver distractions*. Unpublished doctoral dissertation, University of Rhode Island.
- SALVUCCI, D. D. (1999). *Mapping eye movements to cognitive processes*. Unpublished doctoral dissertation, Carnegie Mellon University.
- SALVUCCI, D. D., & ANDERSON, J. R. (2001). Automated eye-movement protocol analysis. *Human-Computer Interaction*, **16**, 39-86.
- SCHREINER, L. M. (2000). *An investigation of the effectiveness of a strobe light as an imminent rear warning signal*. Unpublished master's thesis, Virginia Polytechnic Institute and State University, Blacksburg.
- SODHI, M., REIMER, B. [L.], & LLAMAZARES, I. (2002). Glance analysis of driver eye movements to evaluate distraction. *Behavior Research Methods, Instruments, & Computers*, **34**, 529-538.
- STUTTS, J. [C.], FEAGANES, J., RODGMAN, E., HAMLETT, C., MEADOWS, T., REINFURT, D., ET AL. (2003). *Distractions in everyday driving* (Report). Chapel Hill: University of North Carolina, Highway Safety Research Center.
- STUTTS, J. C., & HUNTER, W. W. (2003). Driver inattention, driver distraction and traffic crashes. *ITE Journal*, **73**, 34-36, 43-45.
- SUMMALA, H., NIEMINEN, T., & PUNTO, M. (1996). Maintaining lane position with peripheral vision during in-vehicle tasks. *Human Factors*, **38**, 442-451.
- SUSSMAN, E. D., BISHOP, H., MADNICK, B., & WALTER, R. (1985). Driver inattention and highway safety. *Transportation Research Record*, **1047**, 40-48.
- UNDERWOOD, G., CHAPMAN, P., CRUNDALL, D., COOPER, S., & WALLEN, R. (1999). The visual control of steering and driving: Where do we look when negotiating curves? In A. G. Gale, I. D. Brown, C. M. Haslegrave, & S. P. Taylor (Eds.), *Vision in vehicles-VII* (pp. 245-252). Amsterdam: Elsevier.
- UNEMA, P., DORNHOEFER, S., STEUDEL, S., & VELICKOVSKY, B. M. (1999). An attentive look at drivers' fixation durations. In A. G. Gale, I. D. Brown, C. M. Haslegrave, & S. P. Taylor (Eds.), *Vision in vehicles-VII* (pp. 335-344). Amsterdam: Elsevier.
- VIOLANTI, J. M. (1998). Cellular phones and fatal traffic collisions. *Accident Analysis & Prevention*, **30**, 519-524.
- WANG, J.-S., KNIPLING, R. R., & GOODMAN, M. J. (1996). The role of driver inattention in crashes: New statistics from the 1995 Crashworthiness Data System. In *Proceedings of 40th Annual Meeting of the Association for the Advancement of Automotive Medicine* (pp. 377-392). Barrington, IL: AAAM.
- YARBUS, A. L. (1967). *Eye movements and vision*. New York: Plenum.
- YOUNG, L. R., & SHEENA, D. (1975). Survey of eye movement recording methods. *Behavior Research Methods & Instrumentation*, **7**, 397-429.