On the Roles of Eye Gaze and Head Dynamics in Predicting Driver's Intent to Change Lanes

Anup Doshi, *Student Member, IEEE*, and Mohan Manubhai Trivedi, *Life Fellow, IEEE*

***Abstract*—Driver behavioral cues may present a rich source of information and feedback for future intelligent advanced driver-assistance systems (ADASs). With the design of a simple and robust ADAS in mind, we are interested in determining the most important driver cues for distinguishing driver intent. Eye gaze may provide a more accurate proxy than head movement for determining driver attention, whereas the measurement of head motion is less cumbersome and more reliable in harsh driving conditions. We use a lane-change intent-prediction system (McCall *et al.*, 2007) to determine the relative usefulness of each cue for determining intent. Various combinations of input data are presented to a discriminative classifier, which is trained to output a prediction of probable lane-change maneuver at a particular point in the future. Quantitative results from a naturalistic driving study are presented and show that head motion, when combined with lane position and vehicle dynamics, is a reliable cue for lane-change intent prediction. The addition of eye gaze does not improve performance as much as simpler head dynamics cues. The advantage of head data over eye data is shown to be statistically significant ($p < 0.01$) 3 s ahead of lane-change situations, indicating that there may be a biological basis for head motion to begin earlier than eye motion during "lane-change"-related gaze shifts.**

***Index Terms*—Driver-assistance systems, driver behavior, driver intent inference, intelligent vehicles, machine vision, sparse Bayesian learning.**

I. Introduction

ADVANCED driver-assistance systems (ADASs) have the potential to save many lives by aiding drivers in making prompt safe decisions about driving maneuvers. Every year, traffic accidents result in approximately 1.2 million fatalities worldwide; without novel prevention measures, this number could increase by 65% over the next two decades [2]. In the U.S. alone, more than 43 000 fatalities are projected this year due to traffic accidents, with up to 80% of them due to driver inattention [3], [4]. To counter the effect of inattention, ADASs are designed to provide the driver ample warning time to impending dangerous situations and even assist the driver in reacting appropriately. The ADAS could thus prevent collisions and make roads safer.

Manuscript received March 17, 2008; revised March 6, 2009 and June 23, 2009. First published July 24, 2009; current version published September 1, 2009. This work was supported in part by a grant from University of California Discovery, Digital Media Innovations Program, and in part by the Volkswagen of America Electronics Research Laboratory. The Associate Editor for this paper was C. Laugier.

The authors are with the Laboratory for Safe and Intelligent Automobiles (LISA), University of California, San Diego, La Jolla, CA 92093-0403 USA (e-mail: andoshi@ucsd.edu; mtrivedi@ucsd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2009.2026675

The design of robust and practical intelligent assistance systems is an active field of research and development. In particular, the type and placement of sensors to achieve maximum performance has yet to be well understood. Several recent systems, including collision warning and brake support, backup warning systems, and lane-departure warning systems, use very specific environmental sensors that are associated with their application to augment the driver's awareness. However, the basis of all these systems is the sensors detecting the environment outside the vehicle, along with the vehicle dynamics.

Recent research has supported the incorporation of sensors looking inside the vehicle into these systems in a holistic manner [5], [6]. A major advantage of monitoring drivers is the ability to observe driver behavior and potentially infer driver intent. Such data could inform an ADAS on several matters.

1) *Awareness:* Is the driver aware of the pertinent surroundings in their environment? What is the focus of attention and distraction level?
2) *Intent:* Is the driver planning to move into a dangerous situation? What is the planned trajectory or possible movements?

Given the driver's attention patterns, it may be possible to infer whether the system should warn the driver and reduce false alarms. In this work, we will focus on the second application, i.e., how driver behavior can be used to predict future maneuvers.

In particular, we are interested in determining important driver cues for distinguishing intent to support future ADAS designs. In prior intent prediction research [1], [7], [8], *head dynamics*, which is a derivative of head pose, has been proposed as a pertinent cue. While robust monocular in-vehicle head pose estimation systems have been developed [9]–[11], it may be argued that head pose is not a sufficient estimate of true gaze. To derive precise gaze estimates, it follows that *eye gaze* should be included [12]. Unfortunately, there are several drawbacks with modern eye gaze estimators in vehicles, including the need to overcome lighting changes, shadows, occlusions, and potentially cumbersome stereo rigs or intrusive head-mounted cameras. Therefore, we are motivated to determine if eye gaze and head motion are useful intent predictors and, furthermore, which one (or combination) is the more informative cue.

In this paper, we use a lane-change intent-prediction system [1] to determine the relative usefulness of eye gaze and head dynamics data. Our comparative experiments are designed to distinguish the merits of the two cues and compare their importance. By determining the better cue, we hope to provide the basis for appropriate future designs of lane-change intent systems, as well as a foundation for interactive driver-assistance systems in general.

II. Driver Behavioral Cues

The analysis of driver behavior has long been a popular field of research in light of the potential for safety improvements. With respect to the particular maneuver of lane changes, the analysis of driver behaviors dates back at least 30 years.

Here, we present a summary of relevant research. We then present our methodology for determining driver behavior, in preparation for our comparative experiments.

A. Related Research in Lane Change Behavior Analysis

According to early research in the field, there is significant reason to believe that behavior analysis of the driver can lead to reliable predictions about lane-change intent. As described here, to safely decide to change lanes, a driver should have recently given some *attention* to the occupancy and state of the neighboring objective lane. According to Hoffman [13], "attention is free to move independent of the eyes, but eye movements require visual attention to precede them to their goal." Thus, by measuring driver behavior corresponding to a visual search, we can hope to capture obvious shifts in attention and thereby deduce lane-change intentions.

The time period 3 s ahead of the actual lane change was determined to be a critical time period during which the driver engages in a visual search to determine feasibility of lane change [14]. In the period several seconds prior to the lane-change maneuver, certain patterns emerge in the driver's visual search behaviors.

In fact, according to Tijerina *et al.* [15], there are specific eye glance patterns that take place in the period before a lane change. It was determined that, during left lane changes, there were 65%–85% chance of looking at the left mirror and 56%–67% chance of looking at the rearview mirror. Correspondingly, during right lane changes, drivers looked at the right mirror with 36%–53% probability and the rearview mirror with 82%–92% probability. Moreover, the mirror glance duration before lane change maneuvers lasted, on average, 1.1 s, varying between 0.8 and 1.6 s [3]. Mourant and Donohue observed that lengthy blind spot checks occurred only in conjunction with lane-change maneuvers; in lane-keeping situations, no such checks were performed by the drivers [14].

Bhise *et al.* [16] studied a series of naturalistic lane changes in real-world settings and discovered that most visual searches prior to lane changes involve head motions and relatively very few (5.4% in their study) involve eye glance alone. Furthermore, Robinson *et al.* [17] found a remarkably stable relationship between eye glance behavior and head movement behavior: In an experiment where a visual fixation was placed at 60°, the eyes moved first, and the head followed approximately 50 ms later. The relationship in which head movement immediately follows eye movement was found to be stable across all individuals in the experiment.

The experiments of Land [18] corroborate these results in a real automotive setting. Upon examining the head movements and eye behavior of several drivers approaching intersections, the author found some remarkable tendencies in the drivers' behavioral patterns. When a decision to change gaze has subconsciously or "unthinkingly" been generated, eye and head movements, by "default," begin at nearly the same time. The eyes more quickly move than the head; however, the velocity of the head movement is a direct correlation of the magnitude of the gaze change (i.e., the head moves faster for a larger gaze change). These results indicate that head and eye movements are correlated under an unguided visual search, in a situation that is similar to the search prior to a lane change.

These results lead to the hypothesis that eye gaze and head pose can be reliable indicators of a driver's intent to initiate a lane change. Furthermore, it might be posited that head pose alone could be good enough, given that fixations tend to reliably draw head movements, along with eye-gaze changes.

Other studies more recently have included eye gaze measurements as a part of laboratory tests of driver fatigue [19], [20] or of simulated lane change events [21], [22]. Simulators though do not capture all the dynamics and variability of real-world environments [4]. Some real-world studies of driver behavior during lane changes have measured eye gaze by manually reducing data [3], [4], [15], [23], [24]. By doing so, they can ensure the reliability and accuracy of the eye-gaze data; these studies have shown some promise of using eye gaze as a cue for driver intent. There have also been real-world studies that relied on automatically detecting eye gaze, but their results were limited due to robustness issues, particularly with regard to occlusions from sunglasses and harsh lighting conditions [25]–[28]. Finally, there have been several studies that achieved promising results using just head motion as a cue for behavior prediction [1], [7], [8], [29].

As they involve testing more dangerous situations, we leave the in-depth study of lane-change behaviors in the presence of fatigue, distractions, and heavy traffic to future research; here, we assume that they play no more than a minor role in the process of changing lanes. To fully study the effects of these variables, simulator studies could be developed. In the following research, we investigate highway situations without heavy traffic or abnormal fatigue levels and distractions, in the interest of driver safety.

Henning *et al.* [30] examined the predictive power of "glances to the left mirror" prior to lane changes. They found that left mirror glances are a good predictor of intention to change lanes but should be combined with other indicators to reduce false alarms. The participants in the study were told about the goals of the study, however, potentially changing their behavior. In fact, the participants' near-ubiquitous usage of turn signals to indicate lane changes was significantly higher than turn indicator usage from naturalistic studies of drivers' lane-change behaviors [3]. As described here, the data used in our study come from experiments in which drivers were not told about the goals of the study, increasing the likelihood of naturalistic behaviors.

Prior studies have suggested that head pose may be better suited than eye gaze to infer lane-change intent [31] without providing any statistical evidence. This work significantly expands upon the analysis and extends those preliminary results, including a larger sample size and quantified significance tests, which finally allow concrete conclusions to be drawn. No other study that quantifiably compares the predictive power of each of these cues in determining a driver's intention to change lanes has been found. Additionally, we will propose a biological explanation for why head pose consistently has more predictive power than eye gaze earlier in time prior to a lane change.

B. Experimental Data Collection and Reduction

For this work, data were collected in a driving experiment with an intelligent vehicle testbed outfitted with a number of sensors detecting the environment, vehicle dynamics, and driver behavior. These data are the same data used in the lane change intent work by McCall *et al.* [1]. A lane-position detector and controller area network (CAN) bus interface provided most of the data related to the vehicle and surrounding environment. The lane detector was a camera-based lane detector based on the VioLET lane tracker [32], with the camera mounted on the top right of the windshield.

The main driver-focused sensor was a rectilinear color camera mounted above the radio facing toward the driver, providing 30 frame/s at 640 × 480 resolution. These data from this camera were collected and postprocessed to extract behavioral cues based on head pose and eye gaze, as described here.

The data set was collected from a naturalistic ethnographic driving experiment in which the subjects were not told that the objective was related to lane change situations. Eight drivers of varying age, sex, and experience drove for several hours each on a predetermined route. A total of 151 lane changes were found on highway situations with minimal traffic at speeds in the range of 55–75 mi/h. A total of 753 negative samples were collected, corresponding to highway "lane-keeping" situations.

1) Head Dynamics: To be invariant to illumination changes and independent of driver identity, head movement is estimated using optical flow and block matching, as described here.

To capture the essence of the head motion, optical flow vectors are calculated for each of the regions falling within the detected face region, which is found using the Viola/Jones face detector [33]. These vectors are calculated for each of the frames within the window specified. The vectors are integrated over time and are separate over space; the integrated values are input as features to the classifier. In this manner, any sort of rapid head movements will be captured, and the length of time and extent to which the head moved left or right will also be recorded. This methodology is based on that developed by McCall *et al.* [1] and proves to be a robust estimator of head motion. Other methods could also be used to estimate and derive dynamical cues from the true pitch and yaw of the driver's head [9].

Fig. 1 shows the head motion of a driver plotted along with the eye gaze patterns prior to lane changes. It can be seen that the driver's head motion increases before and during lane change maneuvers.

2) Eye Gaze: Because of the nature of the camera positioning, which was placed to minimize occlusion and distraction, automatic eye-gaze measurements were deemed too cumbersome and unreliable to collect. The data came from

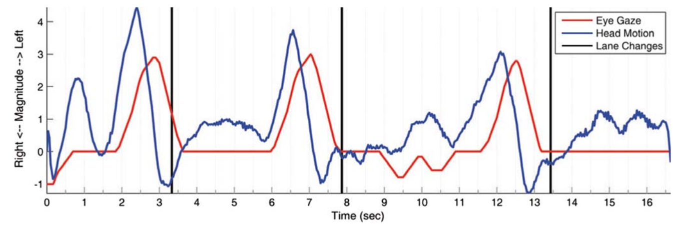


Fig. 1. Head motion and eye gaze positions prior to lane changes. (Black line) Lane changes. (Red line) Magnitude of the eye gaze shift left or right. (Blue line) Side-to-side head motion.

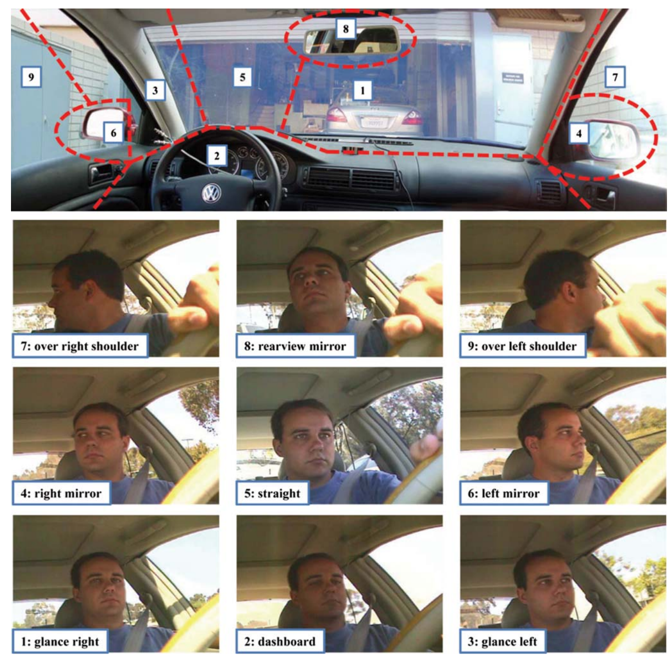


Fig. 2. (Top) Approximate distribution of eye-gaze location classifications for labeling purposes. (Bottom) Samples from the data set showing corresponding eye-gaze locations.

a monocular camera mounted near the radio controls at the center of the dashboard looking at an angle toward the driver. This angle was too obtuse for monocular eye gaze estimators such as [34] to reliably work. In an ideal world, a properly designed stereo or monocular eye-gaze system would provide robust data. Therefore, to approximate the ideal case and retain the best possible chance of getting reliable and accurate eye-gaze estimates, the data were manually reduced. The procedure followed was similar to those followed in the National Highway Traffic Safety Administration (NHTSA) lane-change and workload experiments [3], [4] to produce output that a real-world eye-gaze tracker would output in an optimal setting.

Nine different gaze locations were derived from the procedure described in [4] as relevant to the task at hand: Looking Straight, Glancing Left or Right, Looking at Dash or Rearview Mirror, Looking at Left or Right Mirrors, and Looking Over Left or Right Shoulders. Sample images from each of these cases are shown in Fig. 2.

As described in Section III-A, data over a span of 1 s prior to the decision time are input to the lane-change intent system. The dynamics of the eye gaze over these time windows can be seen in Table I and Fig. 3. Table I shows the average amount of time spent looking in certain directions prior to

TABLE I

GLANCE DURATIONS FOR 3.0- AND 2.0-s-AHEAD DECISION TIMES. NOTE THE LACK OF A PATTERN IN THE EYE GLANCES LEFT OR RIGHT BEFORE LANE CHANGES, WHEREAS THE OVER-THE-SHOULDER LOOK IS MORE INDICATIVE OF A LANE CHANGE

Average amount of time(ms) spent looking...		*Straight*	*Glance Lt*	*Glance Rt*	*Lt Mirror*	*Rt Mirror*	*Over Lt*	*Over Rt*
Time window between 4.0sec and 3.0sec	Before Lane Change	573.5	49.1	51.1	129.0	23.3	68.1	44.4
	Before Lane Keeping	840.3	21.8	53.2	12.5	3.4	0.0	1.2
Time window between 3.0sec and 2.0sec	Before Lane Change	416.2	51.1	20.5	158.3	26.3	182.9	96.6
	Before Lane Keeping	842.6	22.3	53.4	11.7	3.4	0.0	1.3

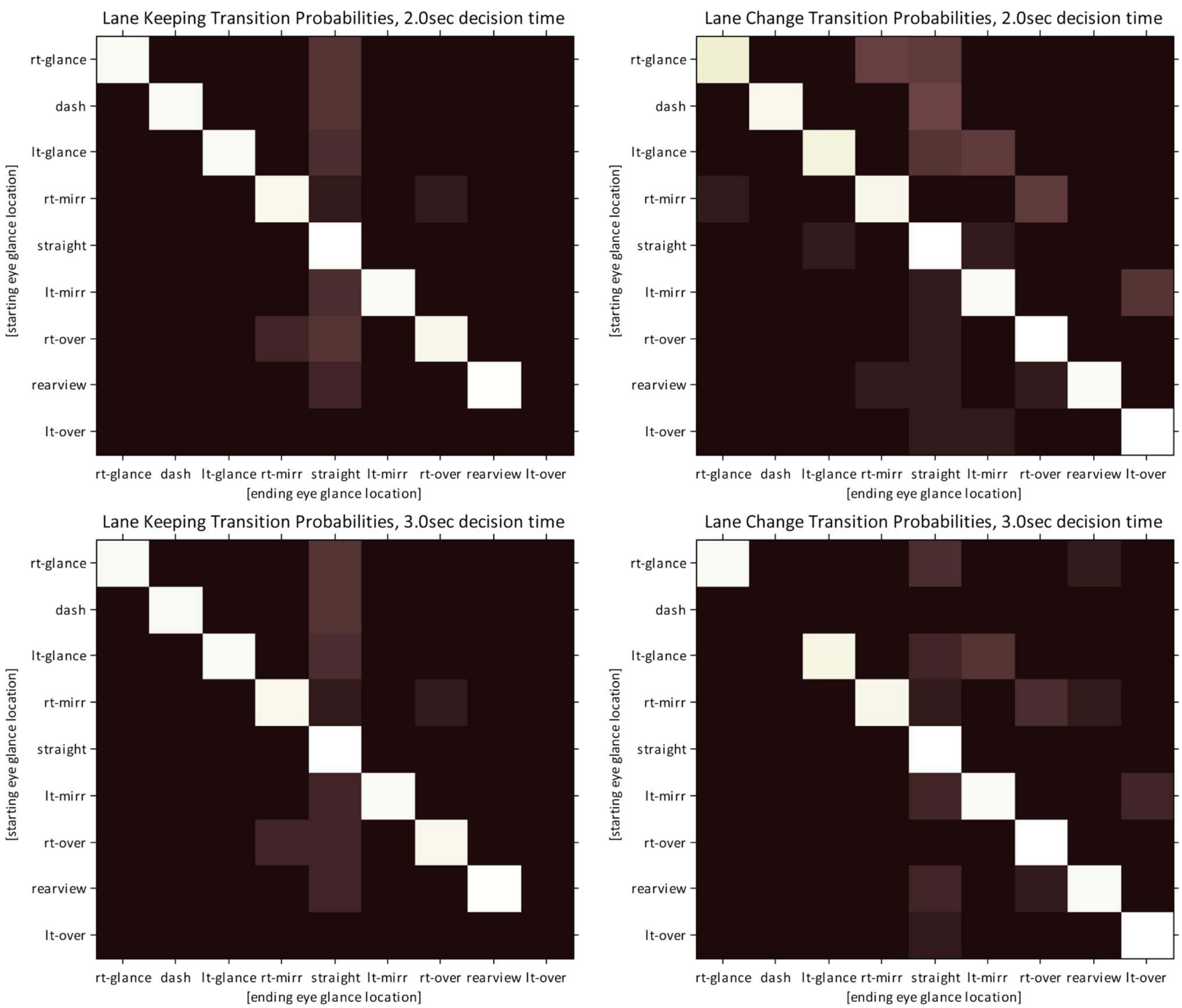


Fig. 3. Eye-glance transition probabilities for lane-keeping and lane-change situations used in this study (brighter squares indicate greater likelihoods; black indicates that no such data was found). During regular lane-keeping situations, the driver is likely to be looking straight; any other glances will also likely end up looking straight. Prior to lane changes, it is much more likely for a left or right mirror glance to transition to a corresponding over-shoulder glance.

lane changes, compared with normal lane-keeping situations. A significant difference is seen between lane change and lane-keeping situations, particularly in mirror and over-the-shoulder glances. Glance transition probabilities for each situation are shown in Fig. 3. For each decision time, the eye-glance behavior does not significantly change during lane-keeping situations, as expected. However, prior to lane changes, it is much more likely for mirror glances to end up in over-shoulder glances. Thus, to robustly represent the eye gaze over this time period, several different formats were input to the system, including raw eye-gaze classifications and the histogram of glances over the time window.

III. LANE-CHANGE INTENT PREDICTION

Driver intent inference is a challenging problem, given the wide variety of potential behaviors and intents. To limit the scope of the problem, we simply examine the driver's intent to change lanes at a particular time in the near future. We base our experiments on the lane change intent system developed by

McCall *et al.* [1], which was labeled "Driver Intent Inference System (DIIS)." It is important to note that we are using this system as a basis for our comparative study about various input cues. There are a number of other works in this field [24], [30], [35]–[37]; our current research will hopefully help inform the future design of these and other intent predictors when considering which inputs to include.

A. DIIS

The DIIS is a discriminative classifier for distinguishing between two events, i.e., lane changing (either right or left) and lane keeping. The following classes of variables are available to the classifier: 1) *vehicle state variables*; 2) *environment variables*; and 3) *driver state variables*.

Vehicle state variables include gas pedal position, brake pedal depression, longitudinal acceleration, vehicle speed, steering angle, yaw rate, and lateral acceleration; these are derived from the vehicle's CAN bus network and are henceforth referred to as "**Vehicle** data." *Environment variables* collected include road curvature metric, heading, lateral lane position, lateral lane position 10 m ahead, and lateral lane position 20 m ahead and are referred to as "**Lane** data." More information on the process and methodology for acquiring the lane data is given by McCall *et al.* [1]. Finally, *driver state variables*, including the variables of particular interest in this work, i.e., head or **Head** dynamics, and **Eye**-gaze measurements, are collected and preprocessed, as previously described.

Each of these variables, as a time series, is windowed to a length of 1 s prior to the chosen decision time. They are then concatenated into a large feature vector, from which a small subset of useful features should be chosen to determine the intent. To find these important features and their relative weightings, a relevance vector machine (RVM) is employed as described here. This classifier outputs a class membership probability, which can then be thresholded to determine a true positive and false positive rate for the predicted lane-change intent.

B. RVM

The RVM classifier used to train the DIIS is based on sparse Bayesian learning (SBL), which was developed by Tipping [38], [39] and implemented in [1]. The algorithm is a Bayesian counterpart to the popular support vector machines (SVMs); it is used to train a classifier that translates a given feature vector into a class membership probability. RVMs, in particular, use a parameterized prior to prune large feature vectors and facilitate a sparse representation of the data.

The basic form of the RVM classifier is given as follows:

$$y(\mathbf{x}) = \sum_{i=1}^{M} w_i K(\mathbf{x}, \mathbf{x}_i) \tag{1}$$

where $\mathbf{x}$ is the input feature vector, w_i is the learned model weight, and $K(\cdot,\cdot)$ is a kernel function. Output $y(\mathbf{x})$ then represents the probability that $\mathbf{x}$ belongs to a particular class.

For our purposes, the feature vector for each example $\mathbf{x}_i$ includes temporal blocks of each of the input cues previously described. For example, at time t, the feature vector looks like

$$\begin{aligned}\mathbf{x}(t) = [&\text{LateralPos}(t), \ldots, \text{LateralPos}(t-N+1)\\ &\text{SteeringAngle}(t), \ldots, \text{SteeringAngle}(t-N+1)\\ &\text{EyeHistogram}(1), \ldots, \text{EyeHistogram}(9); \text{etc.}]\end{aligned} \tag{2}$$

where N represents the number of past values of each variable that have internally been stored; we selected N such that the feature vector represented a 1-s-long sliding window of data.

A detailed description of the RVM algorithm can be found in [38]; the specific algorithm used in these experiments is described in more detail in [39]. The RVM has been shown to be quite effective in predicting lane change intent [1]; thus, for our comparative study, we have selected it as a baseline, although other methods could be used.

Indeed, several advantages of this methodology motivate the use of RVMs over other algorithms, such as SVMs and hidden Markov models (HMMs). The RVM is designed to sift through large feature sets and obtain a sparse representation of the data, which is particularly useful in this application in identifying a small set of useful features. Multimodal data from various sets of sensors can thereby be easily combined, with the discriminating cues from each modality automatically chosen by the RVM. The resulting sparse representation allows for quick computation and classification in real-time and real-world conditions with limited hardware.

The SBL methodology is general enough to consider cases, such as in our experiment, where there are relatively few training examples, compared with the number of features. As opposed to SVMs, RVMs also provide a theoretical framework for determining class membership probabilities. This allows the user to tune the decision boundary to achieve desired performance in a principled manner. Finally, by including the windowed time series of cues in the final feature vector and applying the kernel function, the RVM is capable of determining nonlinear temporal relationships between features, eliminating the need for HMMs.

IV. Experiments and Analysis

In this section, we describe the analysis procedures and results from our experiments.

To predict lane change intent, the classifier needs to be trained for a particular decision time, with a given window of data prior to that time. Based on the prior research and the optimal results in [1], we decided to obtain results for two decision times of 2 and 3 s prior to the lane change. By detecting a lane change this far in advance, an ADAS would be able to warn the driver in time for the driver to be able to safely react [1], [21], [37]. In each case, data in a window of 1 s prior to that decision time were used to make the decision. The data were formatted as previously described into a feature vector.

To counter the effects of scale in feature selection, each feature was renormalized to be of zero mean and unit variance, where the mean and variance were estimated using the training

TABLE II
TRUE POSITIVE AND FALSE POSITIVE RATES FOR A FIXED THRESHOLD ($T = 0$) FOR THE 3-S-AHEAD DECISION TIME

				TPR	FPR
Lane	Vehicle			47.97%	2.27%
Lane	Vehicle	Head		79.46%	0.66%
Lane	Vehicle		Eye	61.08%	1.82%
Lane	Vehicle	Head	Eye	75.68%	0.86%
		Head	Eye	71.08%	0.97%
		Head		70.95%	0.99%
			Eye	46.76%	1.88%

data. The data were then sent through a radial basis function kernel, as described in the SBL algorithm, with a kernel width of 0.5.

Data were split into training and testing data sets, in a ratio of 80%–20% for a fivefold cross validation. Five such randomized trials were conducted. Since the output of the classifier was a class membership probability, the decision threshold was varied across the range of probabilities to obtain a receiver operating characteristic (ROC) curve for the set of trials.

To judge the relative effects of Head and Eye data, various combinations of input features were tested, by including or excluding some subset of Head, Eye, Lane, or Vehicle data from the feature vector. The results of these comparative experiments for the 3- and 2-s-ahead decision times are presented here.

The data for the 3-s-ahead decision time was collected in a window between 4 and 3 s in advance of the lane change. As can be seen in Table II and the comparison ROC curves using different input cues in the top left of Fig. 4, we can make some general observations. It turns out that "Eye" data basically has no effect on the performance of the detector. In fact, the best performance occurs by using "Lane & Vehicle & Head (LVH)" data, with "Lane & Vehicle & Head & Eye (LVHE)" barely below that. As discussed here, the reason for this dip may be noisiness in the eye data particularly since people glance around much more than they move their head. However, in all cases, the performance of the 2.0-s detector is much better.

Data for the 2-s-ahead decision time were collected in a window between 3 and 2 s in advance of the lane change. Once again, the relative effects of Eye data are not great, as seen in Table III and the right half of Fig. 4. We can note that adding Head data to Lane & Vehicle improves performance better than adding Eye data to Lane & Vehicle. However, the best performance is achieved by using all four sources of data.

In this case, as opposed to the 3-s-ahead case, the addition of eye data does improve the performance of the overall detector. The improvement is slight but noticeable, whereas, in the 3-s-ahead data, the eye data had negligible effect. The progression of a lane change attempt could then be tipped off first by head movements 3 s before the lane change; then, closer to the 2-s-ahead threshold, eye-movement data would become a useful additional input. This result is corroborated by the statistical significance tests presented here.

1) Statistical Significance: This study is intended to build upon the results of McCall *et al.* [1]; hence, the data used in this study comprise the usable data obtained from that study, as previously described. As this strategy involved a limited population of eight drivers, it is important to determine the statistical significance of the results before drawing conclusions.

To determine the statistical significance of the results, we performed an analysis of variance (ANOVA). The data under test are the intent prediction confidences, which were output as the result of three different lane change intent classifiers. For each driver, the output confidences for every lane change example of that driver are averaged together. We compare this sample population of average prediction confidences across three different classifiers: 1) Lane & Vehicle & Eye (LVE); 2) LVH; and 3) LVHE. The analysis of the confidence outputs allows judgment on whether the results were significantly different between the eye-gaze- and head-pose-based classifiers, as well as when both cues are included.

Using this analysis, we can conclude, in a statistically significant way ($p < 0.01$), that, 3 s before a lane change, the head-pose-based classifier is more confident than the eye-gaze-based classifier. Table IV shows the p values for various pairs of one-way ANOVA calculations. In the case of 3 s prior to the lane change, both LVHE and LVH are significantly more confident than LVE.

Two seconds prior to the lane change, the results are not as significant, although the trend ($p = .10$) is similar. This might indicate that the eye gaze cue provides better information about upcoming lane change intentions closer to the actual event. The head dynamics, however, seem to achieve a better result overall and are certainly significantly better than eye dynamics at an earlier point.

One may call into question the assumption under ANOVA that the data are drawn from a Gaussian distribution. We thus additionally employ a nonparametric test that requires no assumptions about the distribution of the data. We present the results of the paired two-sided WSR test in Table IV. The resulting p values are more or less in line with the results of the ANOVA test; therefore, we can reasonably draw similar conclusions based on either of these tests.

Thus, while it may not seem to be a large population, the conclusion may be drawn that the patterns are so consistent and dramatic that the trends for the 3-s decision time are statistically significant. We can reliably conclude based on the ANOVA analysis that, for most drivers, head dynamics have significantly more predictive power than eye dynamics 3 s prior to an intended lane change.

A. Discussion

In some sense, it is surprising that the addition of head dynamics by itself does as well as or better than eye gaze since eye gaze would inherently seem to include the motion of the head in its estimate. After observing the data, there seem to be two major causes for the lack of influence held by the eye-gaze data.

Closer analysis of the patterns of head movement and eye gaze movement from Fig. 1 is shown in Fig. 5. Data suggest that, in this movement pattern and in many others in the data set, head motion actually starts before the eye-gaze movement. In other words, the gaze shift began with an initial head

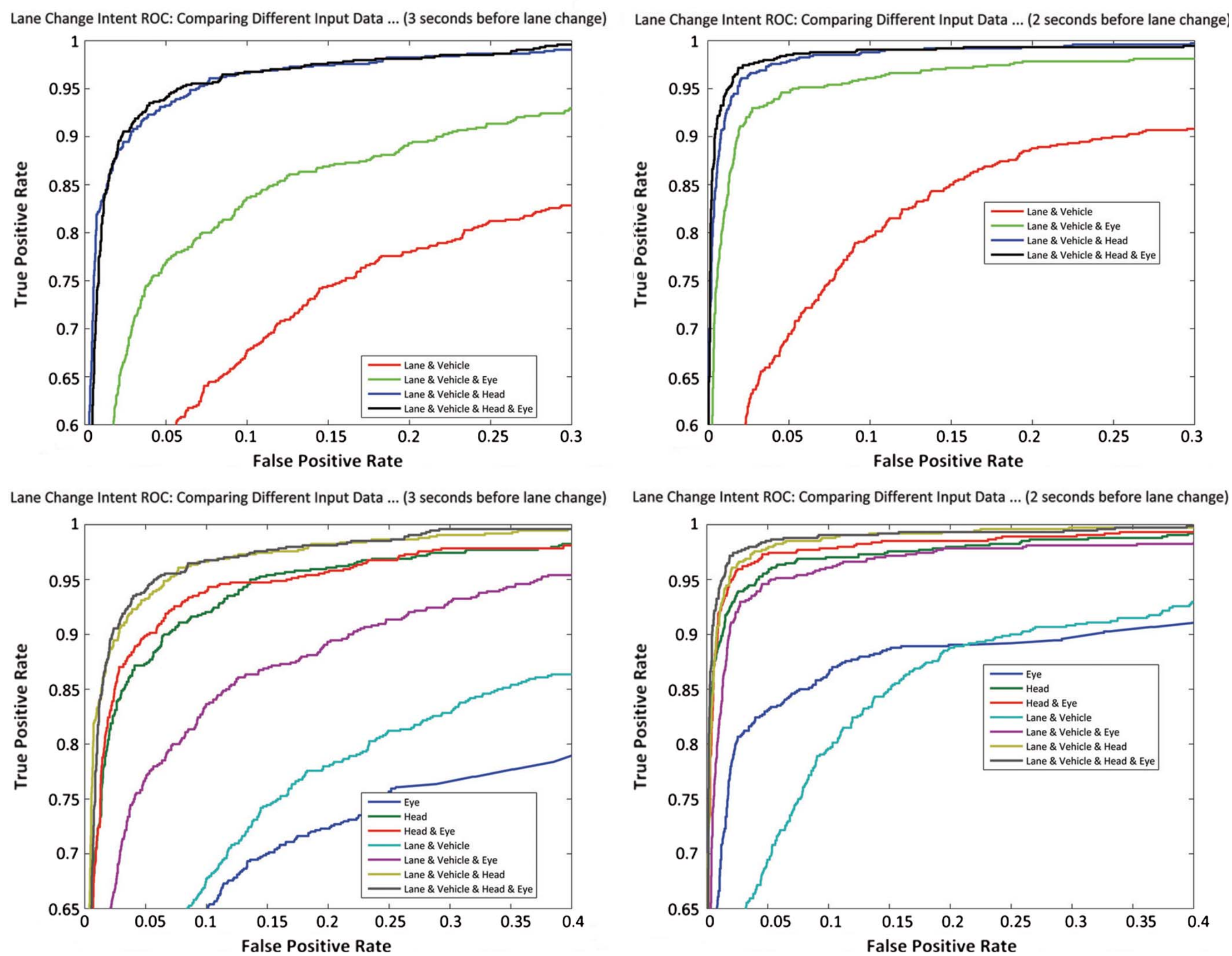


Fig. 4. ROC comparing different input data, i.e., 3- and 2-s decision times. The figures represent the same data, comparing the output of the classifier using various sets of inputs and times. The top figures show that the addition of Eye data to Lane & Vehicle improves performance but not as much as the addition of Head data. When using all four sets of inputs, the results are more or less the same as those without Eye data. Similar patterns are seen between the (left) 3-s and (right) 2-s decision times. All the data are shown for comparison in the bottom figures.

TABLE III
TRUE POSITIVE AND FALSE POSITIVE RATES FOR A FIXED THRESHOLD ($T = 0$) FOR THE 2-s-AHEAD DECISION TIME

				TPR	FPR
Lane	Vehicle			58.38%	2.18%
Lane	Vehicle	Head		88.51%	0.72%
Lane	Vehicle		Eye	78.92%	0.83%
Lane	Vehicle	Head	Eye	87.30%	0.39%
		Head	Eye	81.35%	0.39%
		Head		83.38%	0.33%
			Eye	75.14%	1.74%

TABLE IV
STATISTICAL SIGNIFICANCE TESTS (ONE-WAY ANOVA). A p VALUE OF LESS THAN 0.05 INDICATES THAT THE TWO POPULATIONS UNDER TEST HAVE SIGNIFICANTLY DIFFERENT MEANS. VARIOUS COMBINATIONS OF LANE (L), VEHICLE (V), HEAD (H), AND EYE (E) CUES ARE COMPARED AGAINST EACH OTHER. $\Delta IPC - Means$ SHOWS THE DIFFERENCE BETWEEN THE AVERAGES OF EACH GROUP'S INTENT PREDICTION CONFIDENCES. p VALUES FROM ANOVA AND THE WILCOXON SIGNED RANK (WSR) TEST ARE SHOWN

3 second Decision Time				
Comparing...		*ΔIPC-Means*	*ANOVA p-value*	*WSR p-value*
LVE	LVH	-0.438	0.00492	0.00781
LVE	LVHE	-0.423	0.00518	0.01563
LVH	LVHE	0.016	0.87432	0.54688
2 second Decision Time				
Comparing...		*ΔIPC Means*	*ANOVA p-value*	*WSR p-value*
LVE	LVH	-0.195	0.10465	0.03906
LVE	LVHE	-0.135	0.22351	0.19531
LVH	LVHE	0.060	0.52869	0.46094

motion; then, the eye shift followed later. This behavior may seem counterintuitive; however, such a pattern turns out to be consistent with a specific biological model of attention shifts.

According to an experimental study analyzing the relationship between eye and head movements by Zangemeister and Stark [40], such early head movement with respect to the overall gaze shift occurred mainly in gaze shifts of large amplitude, gaze shifts with predictable targets, and/or very rapid shifts. The study identified various other models of eye–head movement, including early eye movements in situations with small amplitude shifts or unknown target location.

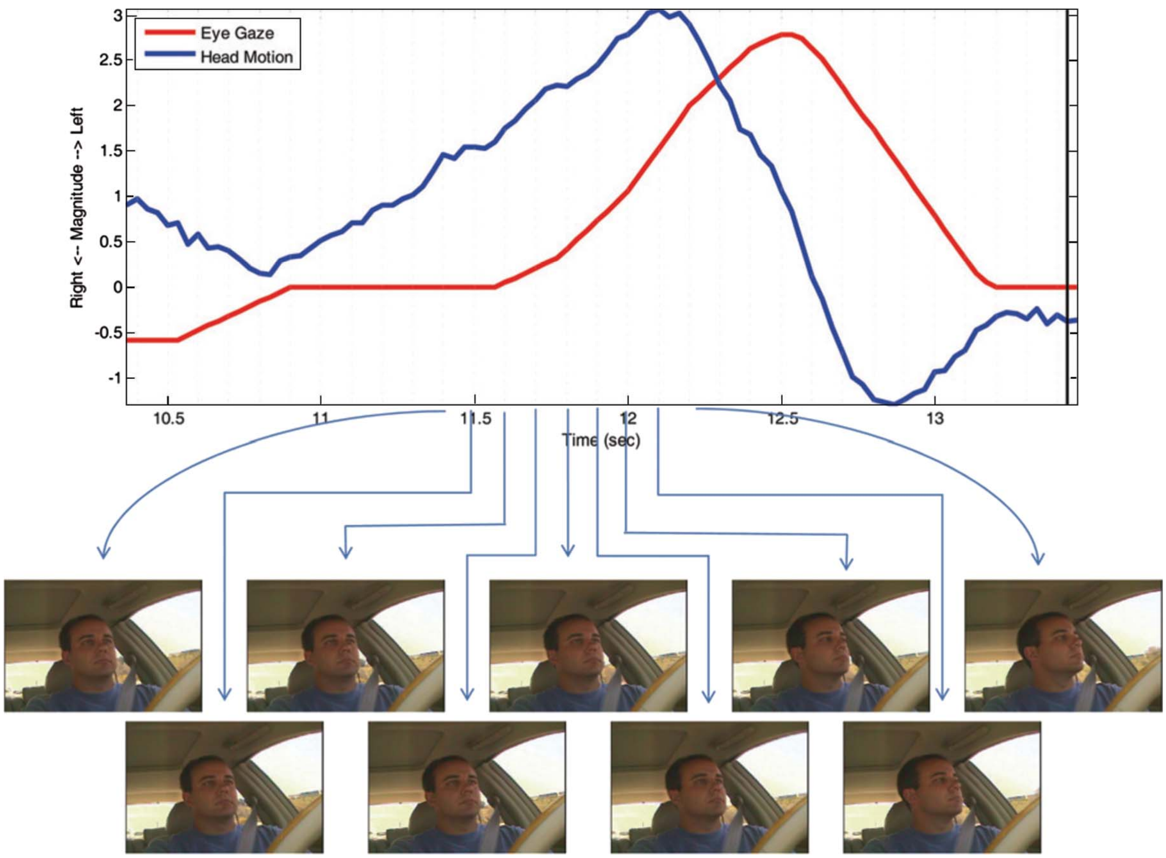


Fig. 5. Head motion and eye-gaze positions prior to a lane change. (Black line) Lane change. (Red line) Magnitude of the eye gaze shift left or right. (Blue line) Side-to-side head motion. The driver's image from various points is shown at the corresponding time. Note that the slight head motion visibly starts before the eye-gaze shift occurs.

The task-oriented gaze shift associated with lane change glances certainly falls into the former category, where the driver has a premeditated gaze target and thus prepares the gaze shift with a preliminary head movement. By capturing this movement, we can predict the intention of the driver earlier than if we wait for the eye-gaze shift.

This leads us to the conclusion that, in situations where there are tasks requiring large gaze shifts to predetermined locations prior to the execution of the task, head motion can be used to predict the onset of the gaze shift earlier than eye motion can. The data previously presented indeed support this; head motion is shown to have significantly more predictive power than eye gaze 3 s prior to the task, but the difference is not as significant 2 s prior to the task.

A secondary factor in the performance of the eye gaze data is that the amount of eye movement also varies between drivers, whereas head motion may be more consistent. With a limited data set, it is difficult to train a classifier to adapt to each driver's own style of eye glancing prior to lane changes. Head pose movements, however, occur in a more telling manner across the population; this pattern extends from the general results of Bhise *et al.* [16]. This property makes head dynamics a more reliable metric for inferring driver intent.

The NHTSA Lane Change Study [3] and the Intelligent Vehicle Highway System review [41] both led to the hypothesis that measuring head dynamics may be enough to detect distinctive behavioral cues prior to a lane change. Having confirmed that hypothesis, it can be argued that the relative ease of capturing head motion information, compared with eye gaze in vehicles, outweighs the advantages of adding eye data to head, lane, and vehicle data. Given the choice between the two cues, head pose can be considered as a better and earlier indicator to use for lane-change intent inference.

Eye gaze may still be useful, however, for driver workload and distraction studies. For intent analysis, behavioral information derived from head motion is more important than eye-gaze data, and robust systems may be designed using just measurements of driver head dynamics, along with lane and vehicle data.

V. Concluding Remarks

We have presented a comparative study of the influence of eye gaze and head movement dynamics on driver behavior and intent prediction with respect to lane change maneuvers. Intent prediction has been carried out using a discriminative classifier based on SBL, where various combinations of features were used to train a classifier, given labeled naturalistic driving data. We have found that, in general, eye gaze was significantly not as informative as head motion $(p < 0.01)$ in helping determine the correct prediction of whether a driver would change lanes 3 s prior to the lane change. Head motion, together with lane and vehicle data, serves as a very good indicator of lane-change intent, and we have discussed a biological reason why head pose is actually an earlier indicator than eye gaze.

With the design of simple robust intelligent driver assistance systems in mind, we have thus attempted to determine the important driver cues for distinguishing driver intent. The addition of eye gaze is relatively cumbersome and potentially unreliable in harsh conditions and does not improve performance as early nor as robustly as do simpler head dynamics cues.

Future studies could further examine the effects of distractions and fatigue on these behavioral cues prior to lane change events, potentially in a driving simulator or controlled environment [45], along with the effects of external vehicles on the driver's motivations to change lanes [46].

ACKNOWLEDGMENT

The authors would like to thank their team members at the Laboratory for Intelligent and Safe Automobiles and the reviewers for their insightful comments and assistance.

REFERENCES

[1] J. C. McCall, D. Wipf, M. M. Trivedi, and B. Rao, "Lane change intent analysis using robust operators and sparse Bayesian learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 3, pp. 431–440, Sep. 2007.

[2] M. Peden, R. Scurfield, D. Sleet, D. Mohan, A. A. Hyder, E. Jarawan, and C. Mathers, Eds., *World Report on Road Traffic Injury Prevention*. Geneva, Switzerland: World Health Org., 2004.

[3] S. E. Lee, C. B. Olsen, and W. W. Wierwille, "A comprehensive examination of naturalistic lane changes," Nat. Highway Traffic Safety Admin., U.S. Dept. Transp., Washington, DC, Rep. DOT HS 809702, Mar. 2004.

[4] L. Angell, J. Auflick, P. A. Austria, D. Kochhar, L. Tijerina, W. Biever, T. Diptiman, J. Hogsett, and S. Kiger, "Driver workload metrics task 2 final report," Nat. Highway Traffic Safety Admin., U.S. Dept. Transp., Washington, DC, Rep. DOT HS 810635, Nov. 2006.

[5] M. M. Trivedi, T. Gandhi, and J. C. McCall, "Looking in and looking out of a vehicle: Computer vision-based enhanced vehicle safety," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 1, pp. 108–120, Mar. 2007.

[6] M. M. Trivedi and S. Y. Cheng, "Holistic sensing and active displays for intelligent driver support systems," *Computer*, vol. 40, no. 5, pp. 60–68, May 2007.

[7] S. Y. Cheng and M. M. Trivedi, "Turn intent analysis using body-pose for intelligent driver assistance," *Pervasive Comput.—Special Issue Intell. Transp. Syst.*, vol. 5, no. 4, pp. 28–37, Oct.–Dec. 2006.

[8] J. C. McCall and M. M. Trivedi, "Driver behavior and situation aware brake assistance for intelligent vehicles," *Proc. IEEE*, vol. 95, no. 2, pp. 374–387, Feb. 2007.

[9] E. Murphy-Chutorian, A. Doshi, and M. M. Trivedi, "Head pose estimation for driver assistance systems: A robust algorithm and experimental evaluation," in *Proc. IEEE Int. Transp. Syst. Conf.*, Sep. 2007, pp. 709–714.

[10] S. Baker, I. Matthews, J. Xiao, R. Gross, T. Ishikawa, and T. Kanade, "Real-time non-rigid driver head tracking for driver mental state estimation," Robot. Inst., Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. 04-10, Feb. 2004.

[11] Y. Zhu and K. Fujimara, "Head pose estimation for driver monitoring," in *Proc. IEEE Intell. Vehicles Symp.*, 2004, pp. 501–506.

[12] M. Smith and H. Zhang, "Safety vehicles using adaptive interface technology (task 8): A literature review of intent inference," in "SAVE-IT Phase 1 Final Report," Volpe Nat. Transp. Syst. Center, Cambridge, MA, 2004. [Online]. Available: http://www.volpe.dot.gov/hf/roadway/saveit/

[13] J. E. Hoffman, "Visual attention and eye movements," in *Attention*, H. Pashler, Ed. London, U.K.: Psychol., 1998, ch. 3.

[14] R. R. Mourant and R. J. Donahue, *Mirror Sampling Characteristics of Drivers*. Warrendale, PA: Soc. Automotive Eng., 1974. no. 740964.

[15] L. Tijerina, W. R. Garrott, D. Stoltzfus, and E. Parmer, "Eye glance behavior of van and passenger car drivers during lane change decision phase," *Trans. Res. Rec.*, vol. 1937, pp. 37–43, 2005.

[16] V. D. Bhise, J. Meldrum, D. Jack, G. Troell, D. Hoffmeister, and L. Forbes, *Driver Head Movements in Left Outside Mirror Viewing*. Warrendale, PA: Soc. Automotive Eng., 1981. no. 810761.

[17] G. H. Robinson, D. Erickson, G. Thurston, and R. Clark, "Visual search by automobile drivers," *Hum. Factors*, vol. 14, no. 4, pp. 315–323, 1972.

[18] M. F. Land, "Predictable eye–head coordination during driving," *Nature*, vol. 359, no. 6393, pp. 318–320, Sep. 1992.

[19] Y. Liang, M. L. Reyes, and J. D. Lee, "Real-time detection of driver cognitive distraction using SVM," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 340–350, Jun. 2007.

[20] Q. Ji and X. Yang, "Real-time eye, gaze, and face pose tracking for monitoring driver vigilance," *Real-Time Imaging*, vol. 8, no. 5, pp. 357–377, Oct. 2002.

[21] D. D. Salvucci and A. Liu, "The time course of a lane change: Driver control and eye-movement behavior," *Trans. Res. Part F: Traffic Psychol. Behavior*, vol. 5, no. 2, pp. 123–132, Jun. 2002.

[22] H. Zhou, M. Itoh, and T. Inagaki, "Detection of lane change intention via analyses of eye and head movement behaviors," in *Proc. SICE Symp. Syst. Inf.*, 2006, pp. 231–236.

[23] E. C. B. Olsen, S. E. Lee, and W. W. Wierwille, "Eye glance behavior during lane changes and straight-ahead driving," *Trans. Res. Rec.*, vol. 1937, pp. 44–50, 2005.

[24] M. Itoh, K. Yoshimura, and T. Inagaki, "Inference of large truck driver's intent to change lanes to pass a lead vehicle via analyses of driver's eye glance behavior in the real world," in *Proc. SICE Annu. Conf.*, Sep. 2007, pp. 2385–2389.

[25] X. Liu, F. Xu, and K. Fujimura, "Real-time eye detection and tracking for driver observation under various light conditions," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2002, pp. 344–351.

[26] R. Smith, M. Shah, and N. da Vitoria Lobo, "Determining driver visual attention with one camera," *IEEE Trans. Intell. Transp. Syst.*, vol. 4, no. 4, pp. 205–218, Dec. 2003.

[27] J. P. Batista, "A real-time driver visual attention monitoring system," in *Proc. Iberian Conf. Pattern Recog. Image Anal.*, 2005, pp. 200–208.

[28] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea, and M. E. Lopez, "Real-time system for monitoring driver vigilance," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 63–77, Mar. 2006.

[29] S. Y. Cheng, S. Park, and M. M. Trivedi, "Multi-spectral and multi-perspective video arrays for driver body tracking and activity analysis," *Comput. Vis. Image Underst.*, vol. 106, no. 2/3, pp. 245–257, May 2007. DOI:10.1016/j.cviu.2006.08.010.

[30] M. J. Henning, O. Georgeon, and J. F. Krems, "The quality of behavioral and environmental indicators used to infer the intention to change lanes," in *Proc. 4th Int. Driving Symp. Human Factors Driver Assessment, Training, Vehicle Des.*, Jul. 2007, p. 231.

[31] A. Doshi and M. M. Trivedi, "A comparative exploration of eye gaze and head motion cues for lane change intent prediction," in *Proc. IEEE Intell. Vehicles Symp.*, Jun. 2008, pp. 49–54.

[32] J. C. McCall and M. M. Trivedi, "Video based lane estimation and tracking for driver assistance: Survey, system, and evaluation," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 20–37, Mar. 2006.

[33] P. Viola and M. J. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2001, pp. 511–518.

[34] J. Wu and M. M. Trivedi, "Simultaneous eye tracking and blink detection with interactive particle filters," *EURASIP J. Advances Signal Process.*, vol. 2008, no. 11, pp. 1–17, 2008.

[35] N. Kuge, T. Yamamura, O. Shimoyama, and A. Liu, "A driver behavior recognition method based on a driver model framework," in *Proc. Soc. Automotive Eng. World Congr.*, 2000, pp. 469–476.

[36] N. Oliver and A. P. Pentland, "Graphical models for driver behavior recognition in a smartcar," in *Proc. IEEE Intell. Vehicles Symp.*, 2000, pp. 7–12.

[37] D. D. Salvucci, H. M. Mandalia, N. Kuge, and T. Yamamura, "Lane-change detection using a computational driver model," *Hum. Factors*, vol. 49, no. 3, pp. 532–542, Jun. 2007.

[38] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Sep. 2001.

[39] M. E. Tipping and A. C. Faul, "Fast marginal likelihood maximisation for sparse Bayesian models," in *Proc. 9th Int. Workshop Artif. Intell. Statist.*, C. M. Bishop and B. J. Frey, Eds., pp. 3–6, Jan. 2003.

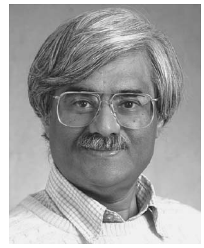
[40] W. H. Zangemeister and L. Stark, "Types of gaze movement: Variable interactions of eye and head movements," *Exp. Neurol.*, no. 77, pp. 563–577, 1982.

[41] P. Finnegan and P. Green, "The time to change lanes: A literature review," Univ. Michigan Intell. Vehicle-Highway Syst., Ann Arbor, MI, Tech. Rep. 90-13, Sep. 1990.
[42] J. C. McCall, O. Achler, M. M. Trivedi, P. Fastrez, D. Forster, J. B. Haue, J. Hollan, and E. Boer, "A collaborative approach for human-centered driving assistance systems," in *Proc. IEEE Int. Transp. Syst. Conf.*, Oct. 2004, pp. 663–667.
[43] R. R. Mourant and C. G. Grimson, "Predictive head-movements during automobile mirror sampling," *Perception Motor Skills*, vol. 44, no. 1, pp. 283–286, Feb. 1977.
[44] D. D. Salvucci, "Inferring driver intent: A case study in lane-change detection," in *Proc. Human Factors Ergonomics Soc. 48th Annu. Meeting*, 2004, pp. 2228–2231.
[45] A. Doshi and M. M. Trivedi, "Investigating the relationships between gaze patterns, dynamic vehicle surround analysis, and driver intentions," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2009, pp. 887–892.
[46] A. Doshi and M. M. Trivedi, "Head and gaze dynamics in visual attention and context learning," in *Proc. CVPR Workshop VCL*, Jun. 2009, pp. 77–84.

Anup Doshi (S'03) received the B.S. and M.S. degrees in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, in 2005. He is currently working toward the Ph.D. degree with the Laboratory for Intelligent and Safe Automobiles (LISA), University of California, San Diego (UCSD), La Jolla.

His research interests include computer vision and machine learning for applications in driver-assistance and intelligent vehicles, human-computer interfaces, and surveillance.

Mr. Doshi is a recipient of the UCSD Jacobs Fellowship (2005–2008) and the University of California Transportation Center (UCTC) Doctoral Dissertation Grant (2009–2010).



Mohan Manubhai Trivedi (LF'09) received the B.E. degree (with honors) from the Birla Institute of Technology and Science, Pilani, India, and the Ph.D. degree from Utah State University, Logan.

He is currently a Professor of electrical and computer engineering and the Founding Director of the Computer Vision and Robotics Research Laboratory, University of California, San Diego (UCSD), La Jolla. He has established the Laboratory for Intelligent and Safe Automobiles (LISA), UCSD, to pursue a multidisciplinary research agenda. He and his team are currently pursuing research in active vision, visual learning, distributed intelligent systems, human body modeling and movement analysis, multimodal affect analysis, intelligent driver assistance, semantic information analysis, and active safety systems for automobiles.

Prof. Trivedi served as the Editor-in-Chief of the *Machine Vision and Applications* journal from 1996 to 2003 and has served on the Editorial Boards of several other journals. He is currently an Associate Editor of the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. He served as a Program Chair for the 2006 IEEE Intelligent Vehicles Symposium (IV 2006) and will serve as the General Chair for IEEE IV 2010 in San Diego. He was the recipient of the Distinguished Alumnus Award from Utah State University, Pioneer (Technical Activities) and Meritorious Service Awards from the IEEE Computer Society, and a number of "Best Paper" Awards. One of his Ph.D. students won the *"2008 Best Dissertation Award"* from the IEEE Intelligent Transportation Society. He is a Fellow of the *SPIE*. He serves on the Executive Committees of the University of California Digital Media Innovation Program and of the California Institute for Telecommunication and Information Technologies [Cal-IT2] as the Leader of the Intelligent Transportation and Telematics Layer, UCSD. He regularly serves as a consultant to industry and government agencies in the U.S. and abroad. He has given over 50 keynote/Plenary talks. He is serving as an Expert Panelist for the Strategic Highway Research Program (Safety) of the National Academy of Sciences.