2008 IEEE Intelligent Vehicles Symposium
Eindhoven University of Technology
Eindhoven, The Netherlands, June 4-6, 2008

A Comparative Exploration of Eye Gaze and Head Motion Cues for Lane Change Intent Prediction

Anup Doshi and Mohan Trivedi

Abstract— **Driver behavioral cues may present a rich source of information and feedback for future Intelligent Driver Assistance Systems (IDAS). Two of the most useful cues might be eye gaze and head motion. Eye gaze provides a more accurate proxy than head motion for determining driver attention, whereas the measurement of head motion head motion as a derivative of pose is less cumbersome and more reliable in harsh driving conditions. With the design of a simple and robust IDAS in mind, we are interested in determining the most important driver cues for distinguishing Driver Intent. We use a lane change intent prediction system [1] 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 using real-world data 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 pose-based cues.**

I. INTRODUCTION

Intelligent driver assistance systems have the potential to save many lives by aiding drivers to make prompt, safe decisions about driving maneuvers. This year in the U.S. alone, over 43,000 fatalities are projected due to traffic collisions, with up to 80% of those due to driver inattention [2], [3]. To counter the effect of inattention, IDAS's could be designed to provide the driver ample warning time to impending dangerous situations, and even assist the driver in reacting appropriately. The IDAS could thus prevent collisions and make roads safer.

The basis of state-of-the-art IDAS systems today involve 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 [4], [5]. A major advantage of monitoring drivers is the ability to observe driver behavior and potentially infer driver intent.

We are interested in determining the important driver cues for distinguishing intent, in order to support future IDAS designs. In prior intent prediction research [1], [6], [7], *head motion* has been proposed as a pertinent cue. While robust monocular in-vehicle head pose estimation systems have been developed [8]–[10], it may be argued that head motion, as a derivative of pose, is not a sufficient estimate of true gaze. In order to derive precise gaze estimates, it follows that *eye gaze* should be included [11]. 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 experiment, we use a lane change intent prediction system [1] to determine the relative usefulness of eye gaze and head motion data. Our comparative experiment is 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.

Authors with LISA: Laboratory for Safe and Intelligent Automobiles, University of California, San Diego, http://cvrr.ucsd.edu/lisa.
{andoshi,mtrivedi}@ucsd.edu

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. NHTSA has most recently conducted studies of Driver Workload Metrics [3], including eye gaze as a proxy for driver workload. 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. Lane Changes and Driver Visual Search

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. The time period three seconds 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 [12]. In fact according to Tijerina et al. [13], there are specific eye glance patterns which take place in the period before a lane change. It was determined that during left lane changes, there were between 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 manuevers lasted on average 1.1 seconds, varying between .8 and 1.6 seconds [2]. Mourant and Donohue observed that lengthy blind spot checks occured only in conjunction with

978-1-4244-2569-3/08/$20.00 ©2008 IEEE.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on January 21, 2009 at 14:50 from IEEE Xplore. Restrictions apply.

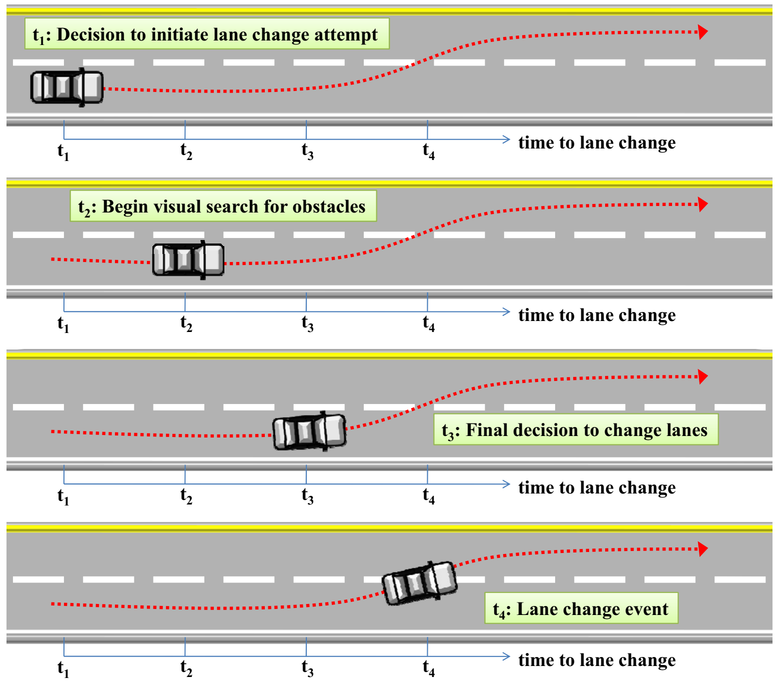

Fig. 1. Progression of a lane change attempt. Once the driver decides to initiate an attempt to change lanes (t_1), a visual search starting at time t_2 foretells the onset of a lane change maneuver. An IDAS should determine situational criticality by t_3 before the driver finally decides to start lateral movement.

lane change maneuvers; in lane keeping situations no such checks were performed by the drivers [12].

Several studies have shown that fatigue [14], traffic [15], and other cognitive and visual distractions [16] may have an effect on behavior prior to lane changes. We leave the in-depth study of these effects to future research, as they involve testing more dangerous situations. Simulator studies could be developed, however there are effects seen in real-world driving that cannot be reproduced in laboratory settings [2].

Bhise et al. [17] studied a series of naturalistic lane changes in real world settings, and discovered that most visual searches prior to lane changes involve head motions; only 5.4% of the searches involve eye glance alone. However Robinson also found a remarkably stable relationship between eye glance behavior and head movement behavior: In an experiment where a visual fixation was placed at 60 degrees, the eyes moved first and the head followed approximately 50 ms later. The relationship that head movement immediately follows eye movement was found to be stable across all individuals in the experiment.

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

Other studies more recently have included eye gaze measurements as a part of laboratory tests of driver fatigue [18], [19] or of simulated lane change events [20]. Simulators, though, do not capture all the dynamics and variability of real-world environments [3]. Some real-world studies of driver behavior during lane changes have measured eye gaze by manually reducing data [2], [3], [13], [21]. 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 as a proxy for fatigue, but their results were limited due to robustness issues, especially with regards to occlusions from sunglasses and harsh lighting conditions [22]–[25]. Finally, there have been several studies that achieved promising results using just head motion as a cue for behavior prediction [1], [6], [7], [26].

B. Data Collection and Reduction

1) Head Motion: In order to be invariant to illumination changes and independent of driver identity, head motion is estimated using block matching. The driver's head in the previous frame is matched to the current frame, giving a disparity which can be considered inter-frame motion. Figure 2 shows the head motion of a driver plotted along with the lane position of the vehicle. It can be seen that the driver's head motion increases before and during lane change maneuvers.

In order 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 [27]. These vectors are calculated for each of the frames within the window specified. The vectors are integrated over time, and separately over space; the integrated values are input as features to the classifier.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on January 21, 2009 at 14:50 from IEEE Xplore. Restrictions apply.

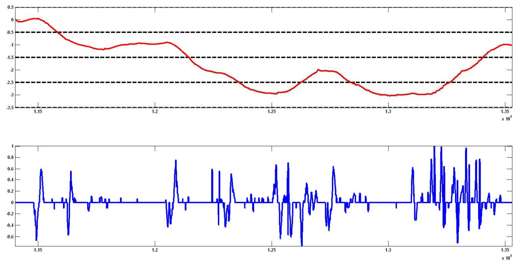

Fig. 2. Lane position and head motion detection results, from [1]. (*Top*) Black lines denote lane boundaries while the red line is the vehicle trajectory. (*Bottom*) Blue line represents side-to-side head motion.

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 the one developed by McCall et al. [1] and proves to be a robust estimator of head motion. Other methods can be used to estimate the true pitch and yaw of the driver's head [8].

2) Eye Gaze: Because of the nature of the dataset that was already collected, automatic eye gaze measurements were deemed too cumbersome and unreliable to collect. The data came from a monocular camera mounted near the radio controls, in the center of the dashboard looking at an angle toward the driver. This angle was too obtuse for monocular eye gaze estimators such as [28] to work reliably. In an ideal world, a properly designed stereo or monocular eye gaze system would provide robust data. Therefore in order to approximate the ideal case and retain the best possible chance of getting reliable and accurate eye gaze estimates, the data was manually reduced. The procedure followed was similar to those followed in the NHTSA lane-change and workload experiments [2], [3], 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 [3] 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 can be seen in Figure 3.

III. Lane Change Intent

Driver intent inference is a challenging problem, given the wide variety of potential behaviors and intents. To limit the scope of the problem, we examine simply 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], labeled the Driver Intent Inference System (DIIS). There are a number of other works in this field [29]–[31]; our current research will hopefully help inform the future design of these and other intent predictors when considering which inputs to include.

A. Driver Intent Inference System

The DIIS is a discriminative classifier to distinguish between two events: lane changing (either right or left) and lane keeping. The following classes of variables are available to the classifier: *Vehicle State Variables*, *Environment Variables*, and *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 CANBus network and are henceforth referred to as **Can** Data. *Environment Variables* collected include road curvature metric, heading, lateral lane position, lateral lane position 10m ahead, and lateral lane position 20m ahead; and are referred to as **Lane** Data. For more information on the process and methodology for acquiring the lane data, please see McCall et al. Finally, *Driver State Variables*, including the variables of particular interest in this research, namely **Head** motion and **Eye** gaze measurements, collected and preprocessed as described above.

Each of these variables, as a time series, is windowed to a length of 1 second 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. In order to find these important features and their relative weightings, the Relevance Vector Machine is employed as described below. The 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. Relevance Vector Machine

The Relevance Vector Machine (RVM) classifier used to train the DIIS is based on Sparse Bayesian Learning, developed by Tipping [32], [33] and implemented in [1]. The algorithm is a Bayesian counterpart to the popular Support Vector Machines (SVM); 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.

A detailed description of the RVM algorithm can be found in [32]; the specific algorithm used in these experiments is described in [33]. 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. The 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 described above. For example, at time t, the feature vector looks like

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

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on January 21, 2009 at 14:50 from IEEE Xplore. Restrictions apply.

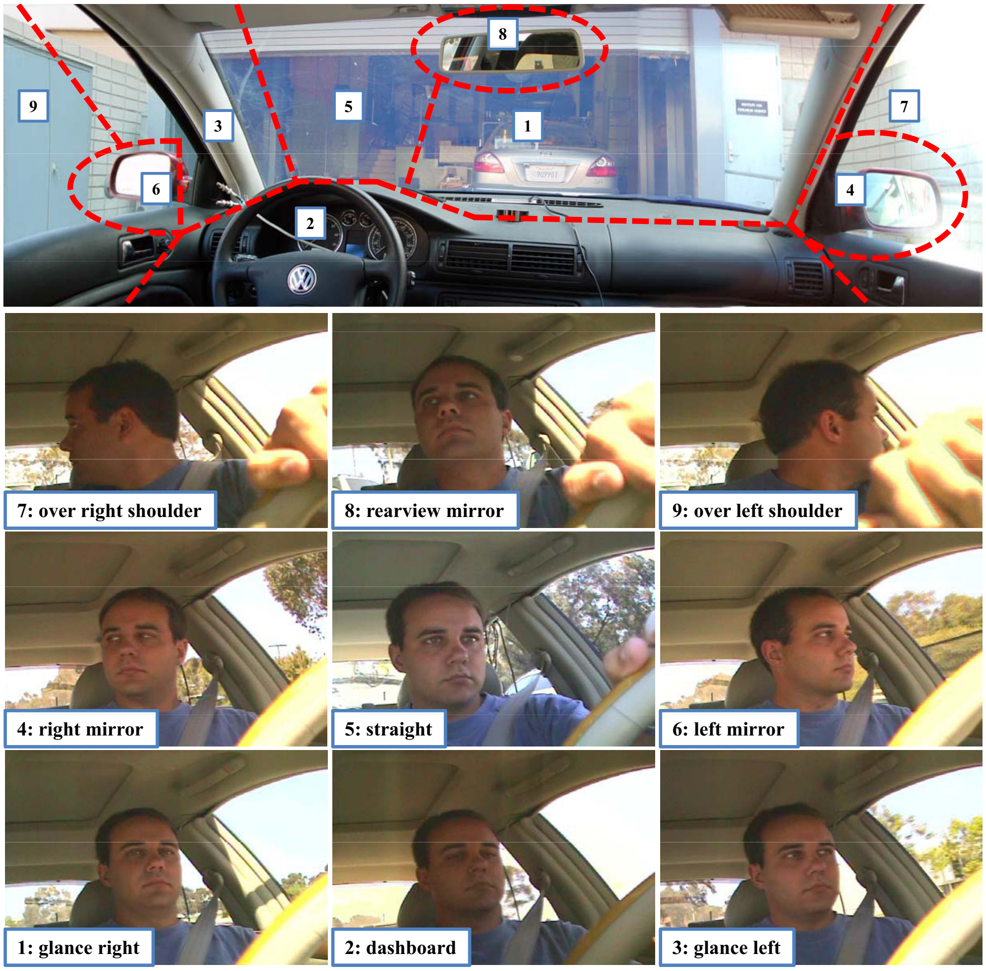

Fig. 3. (Top) Approximate distribution of eye gaze location classifications. (Bottom) Samples from dataset showing corresponding eye gaze locations.

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

IV. Experiments and Analysis

The dataset used was a naturalistic driving experiment in which the subjects were not told that the objective was related to lane changes. A total of 103 lane changes were found over six drivers on highway situations with minimal traffic. 538 negative samples were collected, of corresponding highway lane keeping situations.

In order 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 as well as the results in [1], we decided to obtain results for the decision time of 3 seconds. In this case data a window of one second prior to time was used to make the decision. The data was formatted as described above into a feature vector.

In order to counter the effects of scale in feature selection, each feature was renormalized to be zero mean and unit variance, where the mean and variance were estimated using the training data. The data was then sent through an RBF kernel as described in the SBL algorithm, with a kernel width of 0.5.

Data was split into training and testing datasets, in a ratio of 80%-20% for a 5-fold cross validation. 25 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 an ROC curve for the set of trials.

In order to judge the relative effects of Head and Eye data, various combinations of input features were tested, by including or discluding some subset of Head, Eye, Lane, or Can data from the feature vector. The results of these experiments are presented below.

As can be seen in Table I, as well as the comparison ROC curves using different input cues in Figure 4, we can make some general observations. It turns out "Eye" data basically has no effect on the performance of the detector. In fact the best performance occurs by using "Lane Can and Head" data, with "Lane Can Head Eye" barely below that. The reason for this dip may be noisiness in the eye data; especially since people tend to glance around much more than they shake their head.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on January 21, 2009 at 14:50 from IEEE Xplore. Restrictions apply.

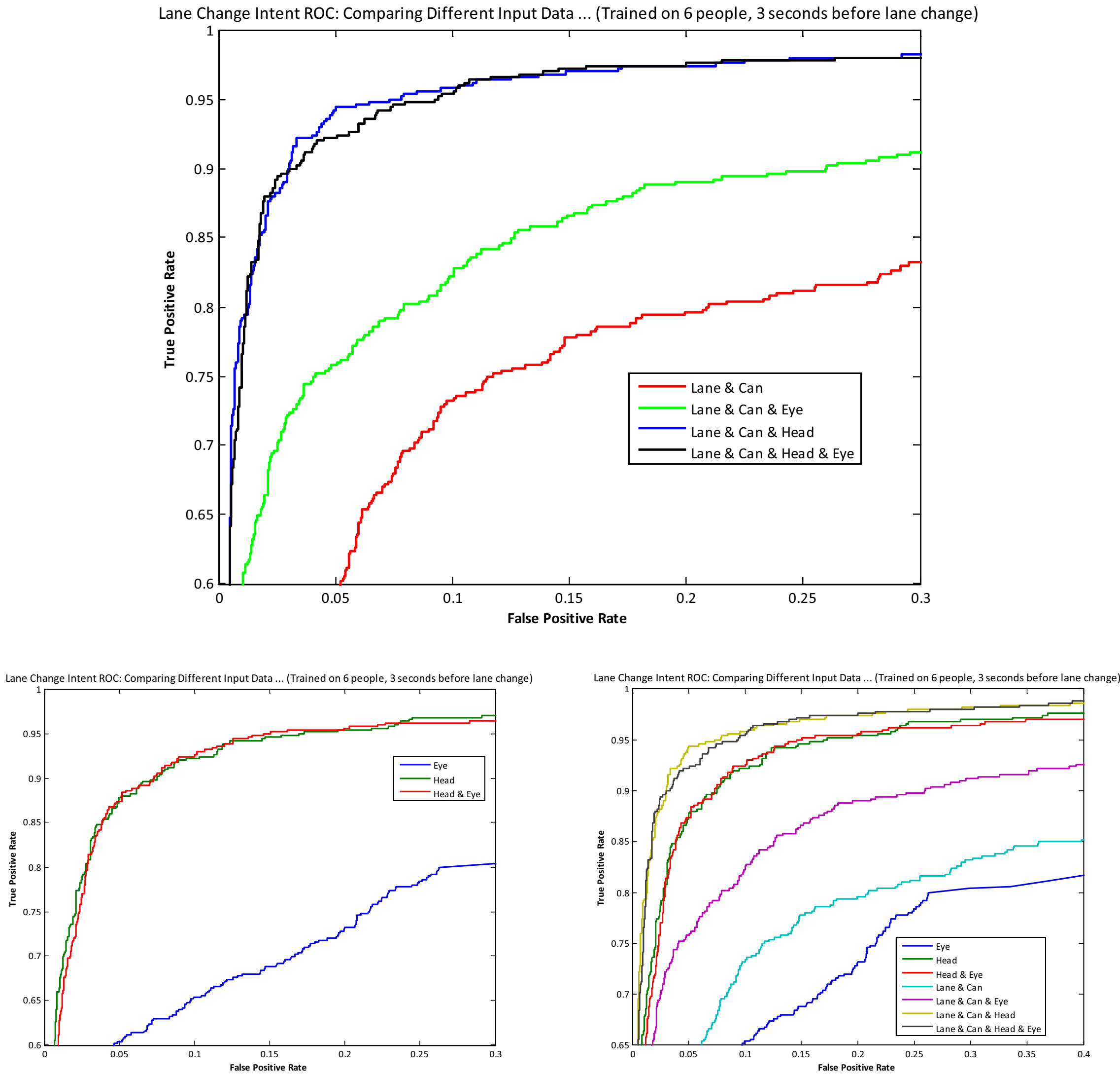

Fig. 4. ROC Comparing Different Input Data, 3-sec Decision Time. The figures represent the same data, comparing the output of the classifier using various sets of inputs. The top figure shows that the addition of Eye data to Lane & Can improves performance, but not as much as the addition of Head data. When using all four sets of inputs, the results are more or less the same as without the Eye data. The Eye data has little added information over Head data - as seen on the bottom left figure. All the data are shown for comparison in the bottom right.

TABLE I

TRUE POSITIVE AND FALSE POSITIVE RATES FOR A FIXED THRESHOLD (T=0) FOR A 3 SECOND-AHEAD DECISION TIME

				TPR	FPR
Lane	Can			49.40%	2.95%
Lane	Can	Head		79.20%	1.08%
Lane	Can		Eye	64.80%	1.68%
Lane	Can	Head	Eye	79.20%	1.16%
		Head	Eye	69.80%	1.64%
		Head		71.80%	1.46%
			Eye	51.40%	2.35%

A. Discussion

In some sense it is surprising that the addition of head motion data by itself does as well as or better than eye gaze, since eye gaze would inherently include the gaze of the head in its estimate. There are several potential causes for the lack of influence held by the eye gaze data.

One factor influencing the eye gaze information could be the inherent noisiness of the data. A slight change in gaze may not be as indicative of intentions as a motion of the head. Furthermore, the amount of eye movement could vary between drivers. Head pose movements have been shown to occur in a more telling manner across the population [17]. Finally, the timing of eye glances could happen at different moments prior to the lane change compared to the head motions. Further research is required to determine the noisiness, consistency, and timing of the eye gaze data.

Based on these initial results, given the choice between the two cues, head motion might then be considered as a better indicator to use for lane change intent inference. Eye gaze

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on January 21, 2009 at 14:50 from IEEE Xplore. Restrictions apply.

may still be useful, however, for driver workload and distraction studies. For intent analysis, behavioral information derived from head movement being more important than eye gaze data, and robust systems may be designed using just head motion along with lane and CAN data.

V. Concluding Remarks

We have presented a comparative study of the influence of eye gaze and head motion on driver behavior and intent prediction with respect to lane change maneuvers. Intent prediction was carried out using a discriminative classifier based on Sparse Bayesian Learning, where various combinations of features were used to train a classifier given labeled naturalistic driving data. We found that in general eye gaze was not as informative as head motion in helping determine the correct prediction of whether a driver would change lanes. Head motion together with lane and CAN data serves as a very good indicator of lane change intent.

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 much as simpler head pose-based cues.

Future studies could include examining other windows of time to determine if eye gaze and head movement are more or less useful at other points before a lane change. Other 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.

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 Transactions on Intelligent Transportation Systems*, Sept. 2007.

[2] S. Lee, , C. Olsen, and W. W. Wierwille, "A comprehensive examination of naturalistic lane changes," *Report DOT HS 809702, NHTSA, U.S. Department of Transportation*, March 2004.

[3] L. Angell, J. Auflick, P. Austria, D. Kochhar, L. Tijerina, W. Biever, T. Diptiman, J. Hogsett, and S. Kiger, "Driver workload metrics task 2 final report," *Report DOT HS 810635, NHTSA, U.S. Department of Transportation*, Nov 2006.

[4] M. M. Trivedi, T. Gandhi, and J. C. McCall, "Looking in and looking out of a vehicle: Computer vision-based enhanced vehicle safety," *IEEE Transactions on Intelligent Transportation Systems*, January 2007.

[5] M. M. Trivedi and S. Y. Cheng, "Holistic sensing and active displays for intelligent driver support systems," *IEEE Computer*, May 2007.

[6] S. Y. Cheng and M. M. Trivedi, "Turn intent analysis using body-pose for intelligent driver assistance," *IEEE Pervasive Computing, Special Issue on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 28–37, October-December 2006.

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

[8] E. Murphy-Chutorian, A. Doshi, and M. M. Trivedi, "Head pose estimation for driver assistance systems: A robust algorithm and experimental evaluation," *IEEE Conf. Intelligent Transportation Systems*, Sept 2007.

[9] 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," *CMU-RI-TR-04-10*, Feb 2004.

[10] Y. Zhu and K. Fujimara, "Headpose estimation for driver monitoring," *IEEE Intelligent Vehicles Symposium*, 2004.

[11] M. Smith and H. Zhang, "Safety vehicles using adaptive interface technology (task 8): A literature review of intent inference," *SAVE-IT Phase 1 Final Report, http://www.volpe.dot.gov/hf/roadway/saveit/*, 2004.

[12] R. Mourant and R. Donahue, "Mirror sampling characteristics of drivers," *Society of Automotive Engineers*, no. 740964, 1974.

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

[14] R. Hanowski, W. W. Wierwille, S. A. Garness, and T. A. Dingus, "A field evaluation of safety issues in local/short haul trucking," *Proceedings of the IEA 2000/HFES 2000 Congress*, 2000.

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

[16] M. A. Recarte and L. M. Nunes, "Effects of verbal and spatial-imagery tasks on eye fixations while driving," *Journal of Experimental Psychology: Applied*, vol. 6, no. 1, pp. 31–43, 2000.

[17] V. Bhise, J. Meldrum, D. Jack, G. Troell, D. Hoffmeister, and L. Forbes, "Driver head movements in left outside mirror viewing," *Society of Autmotive Engineers*, no. 810761, 1981.

[18] Y. Liang, M. L. Reyes, and J. D. Lee, "Real-time detection of driver cognitive distraction using svm," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 340–350, June 2007.

[19] 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.

[20] D. D. Salvucci and A. Liu, "The time course of a lane change: Driver control and eye-movement behavior," *Transportation Research Part F: Traffiv Psychology and Beahvior*, vol. 5, no. 2, pp. 123–132, June 2002.

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

[22] X. Liu, F. Xu, and K. Fujimura, "Real-time eye detection and tracking for driver observation under various light conditions." in *Proceedings of IEEE Intelligent Vehicles Symposium*, June 2002, pp. 18–20.

[23] R. Smith, M. Shah, and N. da Vitoria Lobo, "Determining driver visual attention with one camera," *IEEE Transactions on Intelligent Transportation Systems*, vol. 4, pp. 205–218, December 2003.

[24] J. P. Batista, "A real-time driver visual attention monitoring system," *Iberian Conference on Pattern Recognition and Image Analysis*, 2005.

[25] L. Bergasa, J. Nuevo, M. Sotelo, R. Barea, and M. Lopez, "Real-time system for monitoring driver vigilance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, pp. 63–77, March 2006.

[26] S. Y. Cheng, S. Park, and M. M. Trivedi, "Multi-spectral and multi-perspective video arrays for driver body tracking and activity analysis," *Computer Vision and Image Understanding* in press, 2007, doi:10.1016/j.cviu.2006.08.010.

[27] P. Viola and M. J. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE Conference on Computer Vision and Pattern Recognition*, 2001.

[28] J. Wu and M. M. Trivedi, "Simultaneous eye tracking and blink detection with interactive particle filters," *EURASIP Journal on Advances in Signal Processing*, Oct 2007.

[29] N. Kuge, T. Yamamura, O. Shimoyama, and A. Liu, "A driver behavior recognition method based on a driver model framework," *Proceedings of the Society of Automotive Engineers World Congress*, 2000.

[30] N. Oliver and A. Pentland, "Graphical models for driver behavior recognition in a smartcar," *IV*, 2000.

[31] D. D. Salvucci, H. M. Mandalia, N. Kuge, and T. Yamamura, "Lane-change detection using a computational driver model," *Human Factors*, vol. 49, no. 3, June 2007.

[32] M. E. Tipping, "Sparse bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.

[33] M. E. Tipping and A. C. Faul, "Fast marginal likelihood maximisation for sparse bayesian models," in *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*, C. M. Bishop and B. J. Frey, Eds., Jan 3-6 2003.

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on January 21, 2009 at 14:50 from IEEE Xplore. Restrictions apply.