To appear in Proc. of the 9th Intl. Conference on Human-Computer Interaction, New Orleans, LA, Aug, 2001

MODELING AND PREDICTION OF HUMAN DRIVER BEHAVIOR

Andrew Liu and Dario Salvucci*

MIT Man Vehicle Laboratory, Rm 37-219, 77 Massachusetts Ave., Cambridge, MA 02139
*Nissan Cambridge Basic Research, 4 Cambridge Center, Cambridge, MA 02142

ABSTRACT

Knowledge of the current and future driving context could facilitate the interaction between human driver and advanced driver assistance systems. A driver's intended actions (the future context) can be inferred from a number of sources, including the driver's current control actions, their visual scanning behavior, and the traffic environment surrounding them. In an approach similar to hidden Markov models, the intended actions (e.g., to turn or change lanes) are modeled as a sequence of internal mental states, each with a characteristic pattern of behavior and environmental state. By observing the temporal patterns of these features, it is possible to determine which action the drivers are beginning or intending to execute. This approach has been successfully demonstrated in a variety of simulated driving conditions for a wide range of driver actions including emergency maneuvers. In these studies, only the control actions of the driver (i.e., steering and acceleration actions) were used to infer the driver's state. We are presently exploring the use of the driver's visual scanning behavior as another source of information about the driver's state. Visual scanning behavior offers the additional advantage of prediction of driver actions since scanning generally takes place in areas ahead of the current car position.

1. INTRODUCTION

Our interest in modeling the behavior of humans, specifically human drivers, stems from our desire to improve the interaction with various types of automated systems. If these systems could recognize the human’s behavior and anticipate future actions, it could adjust its behavior to better suit the needs of the human operator. Driving behavior is the observable result of multiple levels of information processing and motor control (Michon, 1985; Boer, 1998) and it is becoming clear that models that capture both high-level cognitive processing and low-level operational control are needed. For practical applications, these models must capture the behavior of the overall population but also have facilities to adapt to a particular person or driver.

In this paper, we discuss an approach to modeling driver behavior in which we assume that the driver has a large number of internal cognitive states, each with its own associated control behavior and interstate transition probabilities (Pentland & Liu, 1999). The states of the model can be organized in a hierarchy to describe both short-term behaviors, such as passing or turning, and long-term behaviors, such as lane keeping. The control behaviors can be thought of as typical “operational-level” processes in the widely-adopted Michon (1985) driving model and the transition matrices as the cognitive components working at the “tactical-level” of driving. We will also examine whether driver eye movements, which have also been described with stochastic models, can be used within this human behavior modeling framework.

2. MARKOV DYNAMIC MODELS

To implement such an approach, we must make observations of the driver’s state, decide which model applies to the current state, and make a response based on that model. But the internal states of the driver are not directly observable, thus we must use an indirect estimation process on the observed behavior (e.g., steering or braking behavior in the case of driving). We have adapted the expectation-maximization methods developed for use with hidden Markov Models (HMMs) to perform this estimation task. Traditional HMMs used to model human behaviors such as speech (Rabiner & Juang, 1986) or gestures (Pentland, 1996) do not capture properties such as smoothness or continuity in their statistical framework. The small-scale structure of human behavior needs to be described by a set of dynamic models that are coupled together into a Markov chain. Unlike a simple multiple dynamic model approach in which the states have a fixed likelihood at each time step, the likelihood of any state in our Markov dynamic model (MDM) makes use of the estimate of the current internal state to adjust the transition probabilities.

In summary, the MDM describes how a set of dynamic processes must be controlled in order to generate the observed behavior, rather than describing the signal itself.

Since human behavior often changes over time due to greater experience or age effects, any model of the driver must also be capable of learning or adapting to these changes. For example, it is well known that driving behavior changes as drivers gain more experience. Thus it is necessary for the MDMs to be able to capture these changes over time. This is theoretically simple to implement since any classification of new behavior can be used as additional evidence to re-estimate the parameters of the model. Thus, MDMs should be capable of matching the behavioral changes in the driver over time, provided that the changes come about slowly.

The MDM approach shares some similarities with modeling approaches based on cognitive architectures, such as ACT-R (Anderson & Lebiere, 1998), and specific models developed within these frameworks, such as the ACT-R driver model (Salvucci, Boer, & Liu, in press). Both approaches are computational models capable of predicting driver behavior. The approaches integrate both low-level perceptual and motor processes with high-level cognitive processes. Typically, the cognitive architecture relies on a central cognitive processor to maintain the mental model and direct the execution of lower-level behaviors. In MDMs, the cognitive process is captured in the transition probabilities connecting the individual models of a particular driving action. In cognitive architectures, the lower level processes are often encapsulated in a set of production rules which could capture the dynamics of the system specified by the dynamic models used in MDMs. Both approaches incorporate some form of learning into the model. This is done explicitly in cognitive architectures either with new rules, or implicitly by the strengthening or weakening of rules based on their usage. Similarly, MDMs can implicitly learn by continuously re-estimating their parameters using their own classifications of data. However, the approaches also have significant differences as well. Primarily, the functions of the models are quite different. The cognitive architectures have been used to simulate human behavior and predict behavior in novel conditions. The MDMs are to be used in real-time applications performing fast recognition of the driver state or prediction of behavior in the near-term. In addition, MDMs do not have the formalization of the cognitive processes such as with cognitive architectures, which have been developed from a theory of human cognition and perceptual-motor behavior. Instead, MDMs capture the cognitive processes within a probabilistic framework connecting the hidden cognitive states. There is never an explicit "model" of the cognitive processes at work.

3. EXPERIMENTAL IMPLEMENTATIONS

We have used the MDM approach in several experiments to test whether human driver behavior can be accurately recognized with MDMs in real-time and to determine what behavior must be observed to infer the driver's cognitive state. One study in a driving simulator (Pentland & Liu, 1999) looked at a number of general driving actions such as stopping, turning, changing lanes, and passing. The driving environment was an urban setting with autonomous traffic, buildings, and lane markings. Subjects performed various driving actions when prompted by text commands on the screen. The commands were generally completed in 5-10 seconds depending on the complexity of the action and surrounding environment. Their control of steering angle, steering velocity, car velocity, and car acceleration was recorded at 10Hz. The dynamic models used were specific to the simulated car (Nissan 240SX). Using three-state models estimated from the collective data (8 subjects, 20 minute sessions, about 600 different actions), we examined the ability of the models to classify the actions that occurred 2 seconds following the presentation of the text commands. This point generally occurs before any large changes in car position, velocity or acceleration occur. The mean recognition accuracy over all actions was $95\% \pm 3.1\%$. In comparison, Bayesian classification of the actual physical control data (measured accelerator, brake and steering wheel positions) were not statistically different from chance performance. The reason is that the pattern of brake taps, steering wheel movements and accelerator motions occur over a range of time scales and vary seemingly randomly, as the pattern depends on microevents in the environment, shifts in driver attention, and other disturbances. Only when these data are integrated by the dynamic models to obtain the state variables of the car motion do we begin to see the desired human control behavior.

Other variants of this approach have also been used with success for other driving maneuvers. Kuge et al (2000) constructed a HMM-based driver behavior recognition model to characterize emergency lane changes, normal lane changes, and lane keeping. The models were estimated from steering angle and steering angle velocity information only. The models were trained and tested on data collected in a driving simulator. The accuracy of continuous recognition of emergency lane changes with these models was 98% with a misclassification rate of 0% for normal lane changes and 0.3% for lane keeping. Again, correct recognition tended to occur within 0.5 – 0.7 seconds after the on-screen cue to begin the action.

Oliver (2000) developed a set of MDMs based on actual on-road driving data and examined the recognition performance of models estimated from various sets of vehicle (e.g., velocity, steering angle) and environmental (e.g., lateral lane position or relative position of other vehicles) parameters. She found that recognition accuracy was greatly improved when both vehicle and environmental data were used. Recognition with only vehicle data was quite good for certain maneuvers such as stopping, starting, and passing but poor for lane changes. The addition of environmental data improved the overall recognition performance since it provides additional context for discriminating between similar maneuvers such as passing and lane changes. Similar to the other studies described here, the models, on average, were able to recognize the maneuvers 1 second before any significant change (20%) in the car or contextual signals. Recognition performance was somewhat lower overall than the performance found in the previously described simulator studies. This is not surprising since there are many factors in the real world that could influence the execution of the driving action.

4. DRIVER EYE MOVEMENTS

Eye movements are another source of information about the cognitive state of a person. There are numerous examples illustrating changes in eye fixation patterns with different mental states (e.g., Yarbus, 1967; Stark & Ellis, 1981). This is not really surprising as the context of the task determines the salience of different features in the visual scene. It is important to note that patterns of eye movements are somewhat idiosyncratic— that is, the actual gaze patterns may be quite different from person to person — but it seems that the fixation locations in the scene are generally similar. Driver eye movements also exhibit different patterns of fixation behavior for different driving tasks such as lane keeping (e.g., Mourant & Rockwell, 1971), curve negotiation (e.g., Land, 1994), and car following (e.g., Veltri, 1995). Very often the gaze location tends to be ahead of the car by 2-5 seconds, thus providing for the possibility of actually predicting upcoming actions. Eye movements also change with different tasks involving instruments inside the car, such as with a moving map display system (Antin, 1990).

Stark and Ellis (1981) were one of the first to model eye movements as a Markov process, showing that eye movements were well modeled by a first-order Markov process. The location of a current fixation is dependent only on the previous location. Thus the pattern of eye fixations is captured in the set of conditional probability matrices which can be empirically measured. For a first order Markov process, an m x m matrix, M_1, contains the probabilities of a fixation in Region R_i being followed by a fixation in region R_j. If only a few entries in a row have high transition probability, then there is a high probability of passing through a cyclic sequence. This approach provides a quantitative method for modeling eye movements and statistically differentiating between models.

5. EXPERIMENTAL ANALYSIS OF EYE MOVEMENTS

The Markov analysis was applied to driver eye movements in the context of lane keeping/curve negotiation and car following to determine if characteristic patterns in driver behavior could be identified (Liu, Veltri, and Pentland, 1998). The underlying assumption was that driving could be considered as a combination of “basis” or one-action situations and that the Markov transition matrix for driving over a period of time could be predicted by a linear combination of the Markov transition matrices characterizing those basic situations. A first-order Markov analysis of the data was performed on gaze data collected in a simulated single lane roadway. The results of that analysis found statistically different transition matrices characterizing the different driving tasks, suggesting that the additive model of driver fixation behavior was reasonable. To restate this in terms of our MDM framework, the driver must transition between the two cognitive states since the control of eye movements by multiple cognitive processes can only occur serially. The observed gaze data merely reflect the information acquisition needs of these processes.

Oliver (2000) also incorporated gaze data into the MDMs used in her experiment and found that they generally slightly improved the recognition performance of the models for lane changing. The gaze data were obtained with video post-processing and only incorporated information about the driver’s glances to the various mirrors in the car or whether the driver was looking to the right or left. Considering how drivers systematically view the external scene, recognition performance could probably be significantly improved by improving the resolution of the gaze data such that features of the external visual scene (e.g., tangent points in curves or near and far points ahead of the vehicle) could be identified. We hope to be able to perform such an analysis on data from a recent empirical study of driver behavior in a simulated highway environment (Salvucci & Liu, in preparation). We captured steering data and eye movement behavior during normal lane keeping and curve negotiation as well as lane changing. Since the experiment was run in a simulator it is relatively easy to recover the gaze locations in terms of visual scene features. Our preliminary analysis of the lane changing data show that significant changes in the predominant gaze location on the roadway correlate with the onset of the maneuver. Furthermore, there is an increase in the frequency of gazes to the inside mirror approximately one second before the lane change. We hope that these changes can be captured in the MDM to improve the speed and accuracy of recognizing lane changes.

6. CONCLUSIONS

In this paper, we have described an approach for modeling the behavior of drivers such that we can also perform real-time recognition. It is hoped that these models will be useful in improving the interaction between humans and intelligent automated systems. We briefly described the similarities and differences in our efforts to use cognitive architectures to model human driver behavior. Finally, we described past and current efforts to implement our MDM framework in simulated driving tasks and discussed the possibilities of using gaze information to further improve the recognition performance of the MDMs.

7. ACKNOWLEDGMENTS

We thank Prof. Sandy Pentland for his guidance with HMMs and in developing the modeling approach.

REFERENCES

Anderson, J. R., and Lebiere, C. (1998). *The atomic components of thought*. Hillsdale, NJ: Erlbaum.

Antin, J.F., Dingus, T.A., Hulse, M.C., and Wierwille, W.W. (1990). An evaluation of the effectiveness and efficiency of an automobile moving-map navigation display. *International Journal of Man-Machine Studies*, 33, 581-594.

Boer, E.R, and Hoedemaeker, M. (1998). Modeling driver behavior with different degrees of automation: A hierarchical framework of interacting mental models. In *Proceedings of the 17th European Annual Conference on Human Decision Making and Manual Control*, Valenciennes, France.

Kuge, N., Yamamura, T., Shimoyama, O., and Liu, A. (2000). A driver behavior recognition method based on a driver model framework. In *Proceedings of the 2000 SAE World Congress*, Detroit, MI.

Land, M.F., and Lee, D.N. (1994). Where we look when we steer. *Nature*, 369, 742-744.

Liu, A. (1998). What the driver's eye tells the car's brain. In: Underwood, G., (Ed.), *Eye Guidance in Reading and Scene Perception* (pp. 431-452). Oxford: Elsevier Science.

Liu, A., and Pentland, A. (1997). Towards real-time recognition of driver intentions. In *Proceedings of the IEEE Intelligent Transportation Systems Conference* (pp. 236-241). Boston, MA.

Liu, A., Veltri, L., and Pentland, A.P. (1998). "Modelling changes in eye fixation patterns while driving." In A.G. Gale et al., (Eds.), *Vision in Vehicles VI* (pp. 13-20). Amsterdam: Elsevier.

Michon, J. A. (1985). A critical view of driver behaviour models: What do we know, what should we do?, In: Evans, L. & Schwing, R.C., (Eds.), *Human Behaviour and Traffic Safety*. New York: Plenum Press.

Mourant, R.R. and Rockwell, T.H. (1970). Mapping eye movement patterns to the visual scene in deriving: an exploratory study. *Human Factors*, 12(1), 81-87.

Oliver, N. (2000). Towards perceptual intelligence: Statistical modeling of human individual and interactive behaviors. Doctoral Thesis, Program in Media Arts and Sciences, Massachusetts Institute of Technology, Cambridge, MA.

Pentland, A., and Liu, A. (1999). Modeling and prediction of human behavior. *Neural Computation*, 11, 229-242.

Salvucci, D.D., Boer, E.R., and Liu, A. (in press). Towards an integrated model of driver behavior in a cognitive architecture. *Transportation Research Record*.

Salvucci, D.D., and Liu, A. (in preparation). Control and monitoring during highway driving.

Stark, L. and Ellis, S.R. (1981). Scanpaths revisited: Cognitive models direct active looking. In: D.F. Fisher et al. (Eds.), *Eye Movements: Cognition and Visual Perception* (pp. 193-226). Hillsborough, NJ: Erlbaum.

Veltri, L. (1995). Modelling eye movements in driving. Master's Thesis, Dept. of Electrical Engineering & Computer Science, Masachusetts Institute of Technology, Cambridge, MA.

Yarbus, A. (1967). *Eye Movements and Vision*. New York: Plenum Press.