
Ashish Tawari\textsuperscript{1}, Member, Sayanan Sivaraman\textsuperscript{2}, Member, Mohan Manubhai Trivedi\textsuperscript{1}, Fellow, IEEE, Trevor Shannon\textsuperscript{2} and Mario Tippelhofer\textsuperscript{2}

Abstract—This paper details the research, development, and demonstrations of real-world systems intended to assist the driver in urban environments, as part of the Urban Intelligent Assist (UIA) research initiative. A 3-year collaboration between Audi AG, Volkswagen Group of America Electronics Research Laboratory, and UC San Diego, the driver assistance portion of the UIA project focuses on two main use cases of vital importance in urban driving. The first, Driver Attention Guard, applies novel computer vision and machine learning research for accurately tracking the driver’s head position and rotation using an array of cameras. The system then infers the driver’s focus of attention, alerting the driver and engaging safety systems in case of extended driver inattention. The second application, Merge and Lane Change Assist, applies a novel probabilistic compact representation of the on-road environment, fusing data from a variety of sensor modalities. The system then computes safe and low-cost merge and lane-change maneuver recommendations. It communicates desired speeds to the driver via Head-up Display, when the driver touches the blinker, indicating his desired lane. The fully-implemented systems, complete with HMI, were demonstrated to the public and press in San Francisco in January of 2014.

I. INTRODUCTION

In 2012, there were 5.6 million police-reported motor vehicle crashes in the United States with over 33,000 fatalities, a 3.3-percent increase from the previous year \cite{1}. Fatalities in urban crashes alone increased by 4.9-percent. The urban driving environment presents an array of challenges to the driver. Navigation and path planning are made more difficult due not only to the “urban canyon,” but also local traffic states, construction, other factors - like events as well as other road users such as pedestrians, cyclists, and other vehicles.

The Urban Intelligent Assist project was formulated to address these challenges, to help drivers during in an increasingly urbanized world. A 3-year research collaboration between Audi AG, VW Electronics Research Laboratory, and 3 leading California universities, the goals of the project include enhancement of the drivers convenience, comfort, and safety in the urban driving environment. Various universities have worked on applications including routing, parking prediction, and driver assistance tasks. In this work, we detail two driver assistance applications, which observe the driver, and the on-road environment.

Driving tasks involve decision making in three different time-scales: strategic, tactical and/or critical. These timescales have been proposed by prior researchers \cite{2}, \cite{3}, \cite{4} in driver modeling. The strategically-planned maneuvers are associated with long-term time scale, minutes or hours of prior planning, and are often motivated by destination goal and sometimes comfort as well such as route planning, parking choice etc.

Tactical, or short-term, timescales are on the order of seconds, and encompass many successive critical operations. Tactically-planned maneuvers are usually motivated by an uncomfortable situation, or occasionally by a recently modified destination goal of the driver, for example, lane changes, turns, stops, upcoming exit etc. Finally, the critical or operational timescale, on the order of hundreds of milliseconds, is the shortest possible timescale for human interaction and are a generally a result of a driver’s desire to remain safe in following the rules of the road (posted speed limit, road curves, etc.), while carefully operating the vehicle within its limits. Each of these time-scales presents its own research challenges. The tactical and operational time scale timescales are of particular interest to us, as in former, the driver still has time to react to the feedback from an assistance system and in later, the system can intervene appropriately to provide autonomous assistance.

In this paper, we report recently accomplished project and its findings including system design, results and demonstrations. In particular, we present two components of the system, called as, 1) Drive Attention Guard and 2) Merge/Lane Change Recommendation system. Attention Guard provides early detection of driver distraction by continuously monitoring driver and surround traffic situation. It further provides proactive countermeasures that maintain the ego-vehicle a safe following distance from lead vehicle and stay within marked lane lines. Merge/Lane Change Recommendation system monitors full surround of the ego-vehicle using range sensors to not only detect that a car is in the blind spot, but also if a vehicle approaching or being approached is entering the space in the adjacent lane. The system continuously adjusts predictions about the surround vehicles’ trajectory in order to assist driver during stress inducing situations like lane change and merging to a highway.
Each system has been implemented on two prototype vehicles; an instrumented Audi A8 at University of California San Diego, and an instrumented Audi A6 at the Volkswagen Group of America Electronics Research Laboratory in Belmont, California.

II. RELATED WORK AND MOTIVATION

In the driving context, challenges lies in the development of a system which is robust, reliable and can operate continuously within the desired design specification. Another important aspect is how to efficiently relay the outcomes and the suggestions to the driver in order to maximize the utility and gain driver’s trust. A good Human-Machine-Interface (HMI) is an essential design requirement. From commercial point of view, these systems need to be non-contact, non-invasive and be cost effective and yet not compromise aesthetics of the vehicle. In this section, we discuss some of the recent works with above design criteria and are relevant to the two proposed systems. Each system uses several modules such as head pose tracking, vehicle detection, lane detection etc. Detailed literature review of each module is beyond the scope of this paper; however, we will point out to relevant survey works with above design criteria and are relevant to the two proposed systems. Each system uses several modules such as head pose tracking, vehicle detection, lane detection etc.

Driver’s focus of attention is intrinsically linked to eye-gaze and head pose [5], [6], [7]. Therefore, eye or head tracking technologies have been extensively used for visual distraction detection. Robust and reliable eye gaze estimation using non-contact remote systems is very challenging in the real-world driving conditions. Even though precise eye gaze is a better indicator of driver’s attention, head pose can still provide good cues for driver distraction [8] estimation. It is also argued that gaze away from the road with detectable head deviation is, with respect to safety concern, more severe than without head deviation [8]. In search of a reliable technology, therefore, in recent years, a lot of advancements have been made in robust, accurate and continues head pose estimation and tracking [9]. Readers are encouraged to refer to a good overview of different head pose estimation algorithms by Murphy-Chutorian and Trivedi [10].

Ahlstrom et al. [11] investigated a distraction warning system which utilizes eye-gaze and when not available, head pose to detect the distraction event. A rule based system is used to regulate a fixed time length (2 sec) ‘attention buffer’ - an indicator of driver’s attentive state and the trigger for a warning. Although the study does not provide generalizable conclusions due to limited number of participants, it shows a positive trend towards improved visual behavior. The system, however, does not take the surround context into consideration and uses heuristics (e.g. minimum speed threshold) to avoid false alarms (e.g. in urban areas). A systematic review of various driver distraction detection systems can be found in [12]. In our work, specifically targeted for urban areas, we incorporate holistic features including vehicle state and surround context to assess driver attentive state.

Next, we focus on the studies related to driver assistance derived from lane position and surround vehicles dynamics. Complex maneuvers such as lane changes and merges, which require the driver to maintain an awareness of the vehicles and dynamics in multiple lanes, are of particular concern. Until recently, decision-making for driver assistance lane changes has fit a binary decision paradigm, the systems based on fundamental on-road perception answering a yes/no question. Many decision systems for lane changes have focused on when a lane change is infeasible, as well as how to execute the lane change or merge.

III. TESTBED DESIGN AND ARCHITECTURE

In order to develop the future Urban Intelligent Assist system, a complete vehicular context is required. This complete vehicular context includes the driver, vehicle surround and vehicle state. At the backbone of this research is the creation of a human-centered intelligent vehicle that captures
this complete contextual data and interacts with its occupants to assist the driver in critical (stressful) situations that exist during urban driving.

The challenges in the design of such testbed lie in selection of sensor modalities and sensor placements. From commercial point of view, they need to be unobtrusive and should not compromise aesthetics. From the research point of view, the data need to be collected from on-road naturalistic driving conditions. Since they present the actual real-world scenarios and set realistic requirements for the development of a robust and reliable system. Hence the sensor placement should not alter normal driving behavior, for example, sensor obstructing driver’s view for safe maneuvering.

A. Hardware Architecture

Built on a 2011 Audi A8, the automotive testbed has been outfitted with extensive auxiliary sensing for the research and development of advanced driver assistance technologies. The goal of the testbed buildup is to provide a near-panoramic sensing field of view for experimental data capture. Figure 1 shows the different sensing modal as detailed below:

1) Internal Vision: Three cameras are placed monitoring driving head and foot behavior - two cameras on the A-pillar and near the rear view mirror observing drive’s head movements and another camera pointing down observing foot movements. They capture face view in color video stream at 30fps and foot view monochrome infrared stream at 25fps.

2) External Vision: For looking out at the road, the UIA experimental testbed features a single forward-looking camera, captured at 25Hz. In this study we use the camera for lane marker detection and lane tracking.

3) Radar: For tracking vehicles on the sides of the ego-vehicle, we employ two medium-range radars, which have been installed behind the rear-side panels on either side of the vehicle. The radars are able to detect and track vehicles as they overtake the ego vehicle on either side.

4) Lidar: The UIA testbed features two lidar sensors, one facing forward and one facing backward. We use these sensors for detecting and tracking vehicles, as well as detecting obstacles such as guardrails and curbs. The lidars provide high fidelity sensor information, and are able to estimate parameters such as vehicle length, width, and orientation, as well as position and velocity.

5) GPS and Vehicle Dynamic Sensors: Vehicle state - steering and pedal profile are measured through CAN bus interface and for the precise vehicle localization a GPS with inertial measurement unit is installed in the back of the vehicle trunk.

Currently, the experimental testbed features robust computation in the form of a dedicated PC for development, which taps all the available data from the on-board vehicle systems. Sensor data from the radars and lidars are fused into a single object list, with complete object tracking and re-identification handled by sensor-fusion.

B. Software architecture

Large amount of data from above mentioned multi-modal multi-sensory network coming through different channels USB, CAN, FlexRay and Ethernet bus system, requires seamless synchronized capture. We utilize EB Assist ADTF\(^1\) (Automotive Data Time Triggered Framework) software development tool. ADTF provides a graphical user interface to deploy and connect various software components as shown in Figure 2. Besides this the framework offers tools for for real-time data playback, data handling, processing and visualization in the lab as well in the test car.

IV. LOOKING-IN LOOKING-OUT (LiLo) FRAMEWORK

A human-centric UIA system should be capable of maintaining dynamic representations of the external world surrounding the vehicle, the state of the vehicle itself, and state of the driver. The LiLo framework provides the ability to simultaneously correlate the dynamic contextual information of the vehicle interior and the vehicle exterior. In following sections, we present two systems developed in this framework using computer vision, signal processing and machine-learning techniques.

A. Driver Attention Guard

We develop a head pose and its dynamic based attention monitoring system. The aim is to mitigate driver distraction effects using active assistance. The holistic integration of driver, vehicle and surround object states is proposed in the design of the Attention Guard system. There are following three components of the system, continuously monitoring inside driver state and outside environment state along with ego vehicle parameters:

1) CoHMEt (Continuous Head Movement Estimator): A calibrated multi-camera setup is used for reliable and continuous head monitoring. Each perspective is processed by a commercially available head pose tracker and their outputs are fused at later stage, as proposed in [9], to get the final estimate with respect to car reference frame. Readers are encourage to refer [9],

\(^1\)http://automotive.elektrobit.com/home/driver-assistance-software/eb-assist-adtf.html

2) FOVEA (Field of View Estimation for Attention) Monitoring: Head pose and dynamics are then used to estimate driver’s field of view [15], in particular, to determine whether the driver is paying attention to the front or not. Furthermore, a confidence measure to determine the quality of the estimate, is calculated based on the head pose tracking accuracy provided by the commercial system. This measure allow the system to determine whether or not to engage the assistance.

3) Attention Buffer Modulation: Attention buffer reflects driver’s attentive state in driving task. Number of studies have shown that higher percentage of eyes-off-road duration, either due to single long glances or accumulated glance duration, is directly proportional to the increased likelihood of the occurrence of safety critical events [1]. Hence we use the total eyes-off-road duration (time duration driver not looking road ahead) as a measure of distraction. Based on the suggestion of Alliance of Automobile Manufacturers (AAM) in designing HMI, 2 sec worth baseline attention buffer is used which is decreased if the driver looks away from front road or increased otherwise until it reaches the 2 s buffer size. Depletion rate is dependent on front vehicle (if any) dynamics in order to never exceed time-to-collision. When the buffer is fully empty the driver is considered distracted and the assistance is provided by maintaining the ego-vehicle position within lane markings and following distance to the lead vehicle.

The full system has been implemented in C++ and runs in real-time. It is tested in controlled test track along with another vehicle to simulate the front vehicle behavior; further testing under naturalistic driving condition is ongoing. Hope is to make the system and the driver response (if alert) non-differentiable. This is in the direction of positive HMI, where driver is not overwhelmed by the frequent warnings but only when absolutely needed and/or assisted with a smooth active control.

B. Merge Recommendations for Driver Assistance

In this work, we make recommendations for lane change and merge maneuvers. The recommendations consist of recommended accelerations and timings to execute the maneuver, specifying how and when to change lanes or merge into traffic, a problem that features a high number of variables. There are myriad combinations of surrounding vehicles, lane configurations, obstacles, and dynamics to consider.

We base our computation on an intermediate, compact probabilistic representation of the on-road environment, the Dynamic Probabilistic Drivability Map [16]. This compact representation allows us to compactly encode spatial, dynamic, and legal constraints into one probabilistic representation, which we use to compute the timings and accelerations necessary to execute a maneuver. For detail algorithm descriptions, we encourage reader to refer [17]. Here, we only mention the main characteristics, which are pertinent to the discussion.

The DPDM consists of an array of cells, each carrying a probability of drivability $P(D)$, as well as the position, dimensions, and dynamics of any tracked vehicles or obstacles that lie within the cell’s boundaries. Figure 3 shows an example of the drivability map around the ego-vehicle.

We solve for the lowest-cost recommendation to get into the adjacent lane, by formulating the problem as a dynamic programming solution over the DPDM. Dynamic programming breaks down a large problem into a series of inter-dependent smaller problems. We use the DPDM to decompose the task into a discrete number of computations, computing the cost of accelerating to each possible cell location within 5 car-lengths of the ego-vehicle.

We note that our system tells the driver not just whether he can merge, but when and how to merge. The end result is a recommendation to the driver, detailing when and how to merge into highway traffic. The current implementation computes the recommended acceleration, which can be presented to the driver as a highlighted recommended range of speeds in the instrument cluster and Heads-Up Display (HUD).

The full system has been implemented in C++ and runs in real-time. Multiple HMI concepts for relaying the maneuver recommendations to the driver via HUD have also been prototyped, with on-road and user interactivity studies in progress. Figure 6 shows some example HUD mock-ups. The goal is to relay information to the driver in an intuitive, non-distracting manner. In the current implementation, the driver initiates the HUD merge/lane visualization by using the blinker, and the system delivers recommendations for the desired adjacent lane, left or right. Further applications and directions for this body of research, such as wireless connectivity and cooperative driving, are explored in [18].
V. EXPERIMENTAL RESULTS AND ON-ROAD TESTING

The Urban Intelligent Assist project held a public showcase in San Francisco, CA on January 9 and 10, 2014. During this event, real on-road systems were demonstrated to members of the public and press, participants given the opportunity to sit behind the wheel and experience prototype implementations of lane and merge assist, driver attention guard, and research-grade prototype routing and parking applications.

We showcase some examples of the on-road systems working live on the road. Figure 4 shows a screen grab of the lane change assistance system. The blue area of the figure is a bird’s-eye visualization of raw sensor and vehicle tracking data. The driver has indicated that he would like to change to the right lane.

In the first frame, the system shows yellow, and shows a rearward arrow, to indicate that the driver should slow down to execute the maneuver. In the next frame, a few seconds later, the driver has slowed down to 60mph, and the system shows green to indicate that the right lane change is now safe.

Distributed camera framework, CoHMEt, was critical for the success of robust and continuous driver monitoring needed in the Attention Guard system. We reported a thorough study [9] analyzing different camera configurations including number of cameras and their positioning. Figure 5 shows improved operational range of the system over 140-degree using two cameras with proper positioning. From the head pose, driver’s field of view is inferred next. For frontal (driving direction) and non-frontal gaze detection, the system shows very high accuracy. For seven zones (over left-shoulder, left mirror, front, rear-view mirror, right mirror, over right-shoulder and center console) glances itself, the system reached close to 95% classification accuracy [15].

For testing the full functionality of the Attention Guard system, we require driver to be distracted, looking away from the driving direction. For obvious reasons, this is not done on public roads. A pilot test, however, is successfully conducted in a controlled test track environment equipped with lane markings and a lead vehicle to simulate front vehicle behavior. Figure 6 shows a snapshot of testing phase, when driver is not paying attention to the driving direction. Figure 7 shows what are seen by the vehicle sensors - the inside (driver) and the outside context (lead vehicle behavior). The Attention Guard system recognizes the distracted driver, the attention buffer going to zero (red signal), and activates the lane keep assist and adaptive cruise control system for steering and braking assistance, needed for the distracted driver.

The experiences of the technology media, present in the event, about the above two systems are available in online news media forums and blogs [19], [20].

VI. CONCLUSIONS

Urban streets and arteries roads present unique challenges with congested traffic condition and other road users like pedestrians, bicyclist etc. With fast urbanization around the globe, Urban Intelligent Assist project’s aim is to make urban day-to-day driving safer and stress-free. Towards this goal, we presented Driver Attention Guard and Merge/Lane Change Recommendation systems. The systems are installed in a novel testbed, based on 2011 Audi-A8 and Audi-A6 model, with unique and innovative sensing technology. The systems performed real-time and were tested in naturalistic on-road driving conditions as well as in controlled environment. Further testing in different driving conditions, weather conditions as well as for longer period are needed to make the proposed prototype system, a commercial product. To further increase the scope of such systems in understanding in-vehicle driver activities, we recommend to include other body parts analysis e.g. hand in conjunction with head analysis [21]. We also suggest to include driver behavior analysis for early intent prediction [22] for the future assistance and recommendation system.
Fig. 6: Controlled testing of Attention Guard system in Candlestick Park, San Francisco, CA. (a) Driver looking away from the forward driving direction and (b) ego-vehicle (white) coming to stop slowly while depending upon lead vehicle behavior.

Fig. 7: Attention Guard: A functional illustration in a controlled experiment in a driving test track. The driver is interacting with infotainment system, not looking toward forward driving direction. The blue screen visualizes the surround context seen by the car - a lead vehicle (in red) and the ego-vehicle (in black). The depleted attention buffer (red signal) shows the distracted driver state with high confidence (blue signal).

ACKNOWLEDGMENT
The authors would like to thank the support from UC Discovery and industry partners especially Audi and VW Electronic Research Laboratory. We thank other university partners, USC and California PATH, Berkeley, for the useful discussions and collaboration. We also thank LISA colleagues for their valuable assistance.

REFERENCES