Evolving HMI assessment toward a multi-modal vision in an increasingly automated driving experience

Bryan Reimer
Massachusetts Institute of Technology, AgeLab
77 Massachusetts Avenue, Room E40-209a
Cambridge, MA 02139
reimer@mit.edu

Abstract: *Today's voice enabled multi-modal interfaces are often integrated into components of a more traditional visual-manual centric HMI. However, when designs are unencumbered by prior components or current evaluation approaches, multi-modal design approaches may allow the driver to more effectively manage attention over time and space. Multi-modal design approaches may lead to different user experiences and different impacts on a driver's awareness of the roadway. The Advanced Human Factors for Automotive Demand (AHEAD) consortium is focused on the development of new scientific approaches to driver attention measurement based upon several core tenets of driver attention management. This talk will describe the results of several recently published assessments of voice-enabled systems that are now in production and new AHEAD-driven HMI assessment perspectives. As time allows, efforts to understand the interaction drivers have with automated driving systems will also be covered.*

Keywords: driver attention; mixed-mode and multi-modal interfaces; driver state detection; human centered vehicle automation.

Introduction

Multi-modal driver vehicle interfaces draw upon a range of input and output resources [1]. While studies have investigated a range of multi-modal characteristics (e.g. haptics, gesture, etc.), efforts in our laboratory and others have deeply considered demand characteristics of voice-based driver vehicle interfaces (e.g. [2, 3, 4, 5, 6]). These efforts have demonstrated that redundant and non-redundant information presented through visual displays is central to the design of many systems. Current demand evaluation frameworks [7, 8, 9, 10, 11] have largely centered on summary glance statistics such as mean single off-road glance time, total eyes off road time or detection response tasks that do not fully capture the unique characteristics of these modern vehicle interfaces. Given the complex demand characteristics of various implementations of voice-based driver vehicle interfaces and their impact on driver workload and driver attention, new methodological approaches are required to adequately assess and optimize system characteristics. Furthermore, the rapid automation of vehicle systems is resulting in a need for interfaces that provide moment-to-moment information on the status of automation, the vehicle's awareness of its surroundings, and the new functionality of controls (e.g. autopilot stalk) more traditional information and infotainment systems. Effective models are needed to guide developers as they strive to manage driver attention across the range of growing information sources. Real-time systems which measure driver state [12, 13, 14, 15] will increasingly offer opportunities to augment displays and other interfaces to support attentional focus.

The Advanced Human Factors for Automotive Demand (AHEAD) consortium

The MIT AgeLab, in collaboration with Denso Automotive and Touchstone Evaluations, Inc., formed the AHEAD consortium in 2013 to establish and optimize a multi-dimensional and integrated approach to assessing the demand of driver vehicle interfaces. The objective was to develop a new theoretical underpinning for demand assessment and create a set of empirical tools for applied evaluation.

With the support of Denso, Honda, Subaru and Jaguar Land Rover, a mission statement for the consortium was developed. AHEAD's strategic focus was initially directed at broadening scientifically valid perspectives and methodologies for the objective measurement of demand placed on drivers by in-vehicle systems and technologies during vehicle use, while considering the increasingly important role of attention support and management. A key goal was to evolve the measurement of demands on the driver forward from previous efforts (e.g. CAMP Driver Workload Metrics Project, Human Machine Interface and the Safety of Traffic in Europe (HASTE) Project, Adaptive Integrated Driver-vehicle InterfacE (AIDE) Integrated Project, and Advanced Driver Attention Metrics (ADAM) program [16, 17, 18, 19]), towards effectively supporting the assessment of demands from multimodal tasks and the assessment of advanced technologies that affect the safety of automotive HMIs. A conceptual evaluation system was formulated through an extensive review of prior work and identification of gaps including:

- Defining an assessment space that not only considers the classic visual-auditory-cognitive-psychomotor (VACP) dimensions, but that also considers spatial and temporal characteristics of a task. For example, demands of longer, multi-step voice-based interactions may not be realistically "measured" utilizing a ruler designed for relatively short visual-manual based interactions.

- Moving away from defining acceptable demand based upon fixed limits for each type of demand (or, alternatively, based upon a single specific task, such as the "classic radio tuning task"), and moving toward a framework in which demand can be optimized across all dimensions, i.e. visual, auditory, haptic, vocal, manual, etc. by taking into consideration the relative cost and benefit interactions of various input, output and processing modalities.
- Assessing interactions between secondary tasks and the broader operating environment, as opposed to solely focusing on tasks themselves. Moreover, the intent is to move the language of assessment from driver distraction to an emphasis on driver attention management and safe operation. The overall conceptualization aims to consider that conditions of under-load, over-load, and driving context may pull differently for task engagement, i.e. under-loaded drivers often initiate secondary tasks as a means of maintaining attention and technologies are rapidly evolving to help support driver attention.

Approaching design and assessment from the perspective of attention management and the support of safe operation provides a perspective that can encompass future efforts that more holistically consider the demands on the driver, active safety systems, and other higher order forms of automation as a whole. Thus, while the effort was initially targeted at meeting the needs for the evaluation of demand in the context of modern day, largely manually-controlled vehicles, a vision was formed that building upon appropriately conceptualized theoretical underpinning of driver attention management could later support evolution of the models and measurement approaches for use in more highly automated vehicles.

Human centered vehicle automation and the Advanced Vehicle Technology (AVT) consortium

Every day, automobiles equipped with increasingly sophisticated autonomous driving technologies are cruising onto the world's roadways. There is a remarkable amount already known about the capabilities of these technologies, even in the face of unpredictable driving situations. But in addition to the road, these systems must interact with something far more complex: the driver, who is rapidly becoming a part-time passenger in his or her own car. With this change, the role of display technology within the vehicle must evolve in parallel to support the driver / operator.

The MIT AgeLab, in collaboration with Touchstone Evaluations, Inc. and Agero, founded the Advanced Vehicle Technology (AVT) consortium in 2015 to develop a new and deeper understanding of how drivers leverage vehicle automation, driver assistance technologies, and the range of in-vehicle and portable technologies for connectivity and infotainment. The AVT Consortium's membership now includes Delphi, Liberty Mutual Insurance, Jaguar Land Rover, Autoliv, and Toyota.

Of special interest are the fleeting, yet critical, moments when control transfers from driver to vehicle and back again. By assessing the driver's gaze, hand placement, body posture and position, drowsiness, emotional state, and more; and then combining those data with vehicle telemetry and secure geographical location information; researchers are assembling a more complete picture of how people and semiautonomous vehicles work together—or don't. Furthermore, deeper insight is expected to evolve on the types of activities drivers choose to engage in under different levels of control (e.g. how does attention allocation change under L0, L1, L2 and L3). As the work evolves, it will develop deeper insight into how drivers leverage different display technologies (e.g. HUD, etc.) to accomplish everyday driving / non-driving related tasks.

This work is leveraging an increasing pool of naturalistic data collected from a growing fleet of vehicles and drivers: Tesla Models S & X, and MIT owned vehicles (2017 Volvo S90s, 2016 Range Rover Evoques, 2014 Mercedes CLAs and 2014 Chevrolet Impalas) to objectively and subjectively characterize the behavioral effects outlined. The effort has a specific emphasis on leveraging modern "Big Data" analytical approaches including machine learning, and computer vision to rapidly reduce terabytes of collected data into actionable knowledge.

Conclusion

As the automotive ecosystem continues to move towards an increasingly diverse vehicle fleet that involves greater levels of automation in-vehicle, external displays will have a pivotal role in informing other drivers and road users of vehicle intent and related situational awareness. AHEAD is producing an HMI evaluation framework that will support vehicle manufacturers and suppliers as they continue to develop manually driven vehicles and transition towards system designs that consider the drivers periodic use of higher level automated driving features. This framework, is grounded in safety efficacy and strives to move beyond previous efforts to robustly model attention. AVT is providing a critically needed view into the relationship developing between drivers and automation (L1 & L2) that aims to provide a stronger and more robust protective bubble around the vehicle. Furthermore, the effort is collecting insight into how moment-to-moment lateral and longitudinal control assistance is being leveraged on the road. Looking forward, the AVT effort offers an opportunity to begin assessing how drivers adjust (over months and years) to automation that may forever change how attention is devoted to traditional vehicle displays while AHEAD addresses the need for new display concepts that support the rapid acquisition of situational awareness of the road environment.

Acknowledgment

Support for the content of this talk was provided by the Advanced Human Factors for Automotive Demand (AHEAD) consortium, the Advanced Vehicle Technology (AVT) consortium, US DOT's Region I New England University Transportation Center at MIT, Toyota's Collaborative Safety Research Center (CSRC), the Santos Family Foundation, the Insurance Institute for Highway Safety, and the Toyota Class Action Settlement Safety Research and Education Program. The views and conclusions being expressed are those of the authors, and have not been sponsored, approved, or endorsed by Toyota or plaintiffs' class counsel.

References

1. Reimer, B., L. Angell, D. Strayer, L. Tijerina, B. Mehler, "Evaluating Demands Associated with the Use of Voice-Based In-Vehicle Interfaces," *Proceedings of the 60th Annual Meeting of the Human Factors and Ergonomics Society*, Washington, DC, September 19-23, 2016.
2. Chiang, D. P., A. M. Brooks, and D. H. Weir, "Comparison of Visual-Manual and Voice Interaction with Contemporary Navigation System HMIs," *SAE 2005 World Congress & Exhibition*, Warrendale, PA: SAE International SAE Technical Paper 2005-01-0433, 2005.
3. Shutko, J., K. Mayer, E. Laansoo, and L. Tijerina, "Driver workload effects of cell phone, music player, and text messaging tasks with the Ford SYNC voice interface versus handheld visual-manual interfaces" *SAE Technical Paper 2009-01-0786*, 2009.
4. Mehler, B., D. Kidd, B. Reimer, I. Reagan, J. Dobres, and A. McCartt, "Multi-modal assessment of on-road demand of voice and manual phone calling and voice navigation entry across two embedded vehicle systems," *Ergonomics*, vol. 59, no. 3, pp. 344-367, 2016.
5. Reimer, B., B. Mehler, I. Reagan, D. Kidd, and J. Dobres, "Multi-Modal Demands of a Smartphone Used to Place Calls and Enter Addresses during Highway Driving Relative to Two Embedded Systems," *Ergonomics*, 2016.
6. Mehler, B., B. Reimer, J. Dobres, J. Foley, and K. Ebe, "Additional Findings on the Multi-Modal Demands of "Voice-Command" Interfaces," *SAE Technical Paper 2016-01-1428*, 2016.
7. EEC, "Commission Recommendation of 2a December 1999 on safe and efficient in-vehicle information and communication systems: A European statement of principles on human machine interface", *Official Journal of the European Communities L 19/64,25.1.2000*, European Commision DGXIII, Brussels, 1999.
8. Japan Automobile Manufacturers Association (JAMA), "Guideline for In-Vehicle Display Systems, Version 3.0. (2004)," http://www.jama-english.jp/release/release/2005/In-vehicle_Display_GuidelineVer3.pdf, 2004. Retrieved 8/28/2016.
9. ESoP, "European Statement of Principles on the Design of Human Machine Interaction (ESoP, 2005): Draft", European Commission Information Society and Media Directorate-General-G4 ICT for Transport, 2005.
10. Driver Focus-Telematics Working Group, "Statement of Principles, Criteria and Verification Procedures on Driver Interactions with Advanced in-Vehicle Information and Communication Systems" Version 2.0: Alliance of Automotive Manufacturers, 2006.
11. National Highway Traffic Safety Administration, "Visual-Manual NHTSA Driver Distraction Guidelines for in-Vehicle Electronic Devices" (Docket No. NHTSA-2010-0053), Washington, DC, U.S. Department of Transportation National Highway Traffic Safety Administration (NHTSA), 2013.
12. Ji, Q., Z. Zhu, and P. Lan, "Real-time nonintrusive monitoring and prediction of driver fatigue," I*EEE transactions on vehicular technology*, vol. 53, no. 4, pp. 1052-1068, 2004.
13. Bergasa, L. M., J. Nuevo, M. A. Sotelo, R. Barea, and M. E. Lopez, "Real-time system for monitoring driver vigilance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 1, pp. 63-77, 2006.
14. Liang, Y., M. L., Reyes, and J. D. Lee, "Real-time detection of driver cognitive distraction using support vector machines," *IEEE transactions on intelligent transportation systems*, vol. *8, no.* 2, pp. 340-350, 2007.
15. Fridman, L., L. Lee, R. Reimer, and B. Mehler, "A Framework for Robust Driver Gaze Classification," *SAE Technical Paper 2016-01-0177*, 2016.
16. Carston, O., N. Merat, N. Janssen, E. Johannson, M. Fowkes, and K. Brookhuis, "HASTE Final Report, Contract No. GRD1/2000/25361 S12.319626", Human Machine Interface adn the Safety of Traffic in Europe (HASTE) Project, 2005.
17. Östlund, J., B. Peters, B. Thörslund, J. Engström, G. Markkula, A. Keinath, D. Horst, S. Mattes, and U. Föhl, " Driving performance assessment -- methods and metrics," EU Project AIDE, IST - 1 - 507674-IP, D2.2.5, 2005.
18. Angell, L. S., Auflick, J., Austria, P. A., Kochhar, D., Tijerina, L., Biever, W., et al., "Driver Workload Metrics Task 2 Final Report" (DOT HS 810 635). Washington, DC, : U.S. Department of Transportation National Highway Traffic Safety Administration. 2006.
19. Mattes, S. and A. Hallen, "Surrogate Distraction Measurement Techniques: The Lane Change Test (Chapter 8)", In Regan, M. A., J. D. Lee, K. L. Young, Driver Distraction: Theory, Effects, and Mitigation, Boca Raton, Florida: CRC Press. 2009.