

HUMAN-VEHICLE INTERFACES: THE POWER OF VEHICLE MOVEMENT GESTURES IN HUMAN ROAD USER COORDINATION

Malte Risto^{1*}, Colleen Emmenegger^{1*}, Erik Vinkhuyzen², Melissa Cefkin², Jim Hollan¹

¹UC San Diego, DesignLab
La Jolla, CA, USA

²Nissan Research Center Silicon Valley
Sunnyvale, CA, USA

{mristo, cemmenegeger, hollan}@ucsd.edu {erik.vinkhuyzen, melissa.cefkin}@nissan-usa.com

*Equal contributors

Summary: Autonomous vehicles will have to coordinate their behavior with human road users such as drivers and pedestrians. The majority of recently proposed solutions for autonomous vehicle-to-human communication consist of introducing additional visual cues (such as lights, text and pictograms) on either the car's exterior or as projections on the road. We argue that potential shortcomings in the visibility (due to light conditions, placement on the vehicle) and immediate understandability (learned, directive) of many of these cues make them alone insufficient in mediating multi-party interactions in the busy intersections of day-to-day traffic. Our observations of real-world human road user behavior in urban intersections indicate that movement in context is a central method of communication for coordination among drivers and pedestrians. The observed movement patterns gain meaning when seen within the context of road geometry, current road activity, and culture. While all movement communicates the intention of the driver, we highlight the use of movement as gesture, done for the specific purpose of communicating to other road users and give examples of how these influence traffic interactions. An awareness and understanding of the effect and importance of movement gestures in day-to-day traffic interactions is needed for developers of autonomous vehicles to design forms of human-vehicle communication that are effective and scalable in multi-party interactions.

INTRODUCTION

Imagine the following scenario: On a busy afternoon in a suburban community, a family has just finished their shopping at a local farmer's market and walk towards the crosswalk of a four-way stop intersection on the main street. A sedan that just had waited for another group of pedestrians to cross starts to inch forward. The father grabs his young son who is about to step into the street when the driver of the vehicle applies the brake and waves the family on. The father releases the grip on his son's shoulder and the family crosses. As they cross the mother raises her hand to thank the driver. Now imagine the driver of the car is reading the morning newspaper, which is covering the front window and obscuring the intersection and the family from view, as well as making it impossible for the family to see her. The car is one of those new autonomous vehicles (AVs). What is the family to do? How will the vehicle communicate that they can cross safely? Scenarios like this one illustrate the importance of research into the future of communication between AVs and human road users.

In urban road environments, action coordination between road users involves a broad range of parallel, nonverbal and often subtle forms of communication. Similarly, AVs will have to

coordinate their behavior with human road users. For this reason, developers of AVs must strive to reduce uncertainty regarding what the AV is about to do in traffic situations, if they want it to function safely and efficiently in the complex multimodal, multiparty road environment people negotiate every day. The range of actions road users communicate in traffic is limited and mainly refers to intended movement and coordinative actions vis-a-vis other road users, such as: going straight, turning, stopping, going ahead of someone, going to the left or right of someone, yielding, going next (Emmenegger et. al., 2016; Vinkhuyzen & Cefkin, 2016).

The majority of proposed solutions for AV-to-human communication consist of adding visual cues on either the car's exterior, such as LED light strips, LED light panels or as projections on the road in the form of laser light drawings (Habibovic et. al., 2016; Lundgren et. al., 2017; Matthews & Chowdhary, 2015; Nilsson et. al., 2015; Ackerman, 2016). For instance, a LED light strip or message board might say: "I'm waiting" or "Safe to cross." Designs in this solution space seem reasonable when envisioned in one-to-one situations (that is, a single AV communicating with a single human road user, such as a pedestrian preparing to cross a quiet intersection) and are often seen in public demonstrations. However, when viewed in the context of busy, more complex intersections with many pedestrians and many vehicles, the visibility and understandability of these signals are brought into question.

Human behavior in urban intersections shows movement in context is a central method of communication for coordination among cars and pedestrians (Müller, et. al., 2016; Emmenegger et. al., 2016). In contrast to front facing screens or light strips wrapped around a car's body, vehicle movement is visible by road users from different angles. Furthermore, in contrast to text or pictograms, vehicle movement is noticeable at a glance, allowing individual road users to assess the intention of several vehicles rapidly (Matthews, et. al., 2006). Finally, human road users are already familiar with vehicle movement, which improves understandability compared to patterns of blinking lights that have to be learned and memorized. Developers need to leverage the expressive power of purposeful vehicle motion to make the behavior of the AV transparent and predictable to human road users.

Research suggests that motion can be used to communicate intent in drones (Szafir et. al., 2014), and mobile robots (Hoffman & Ju, 2014, Pennycook, 2012). Furthermore, interpersonal distance as the result of movement has been identified as an influence factor in non-verbal communication affecting people's expectations about others' behavior (Hall, 1969). Hoffman and Ju (2014) make a case for robot expressive motion in interaction with humans, arguing for taking the communicative power of movement into consideration early on in the design process.

The purpose of this study is to understand road user communication "in the wild." That is, to observe how drivers make use of movement in coordinating their behavior with other road users and to gain insights into other road users' understanding of movement in traffic interactions. An understanding of current patterns of behavior between vehicles and pedestrians will help guide the design of appropriate AV intention communication. Findings in this study support the notion that vehicle movement acts as communication in the coordination between road users. Furthermore people understand vehicle behavior as movement with intent that is used in the coordination process.

METHOD

Hutchins and colleagues (2013) point out that the meaning of actions can only be understood in the context of their occurrence because meanings are properties of larger units of activity. The meaning of the actions of a single driver cannot be understood from looking at the actions of that driver alone, nor in the actions and interactions of several road users. Rather, the meanings are embedded in the organization of the activity of the entire road system, including road users, road geometry and traffic control infrastructure. In order to understand the multiparty road user activity of multiple vehicles and people moving through intersections and across urban streets, we employed qualitative video analysis (Heath, Hindmarsh & Luff, 2010).

A variety of roadways and intersections were included in the study, each with a different roadway configuration, geometry and traffic control type, ranging from highly controlled four-way stops to completely uncontrolled middle of the street locations. Multiple cameras, either stationary or first person, recorded the environment from different perspectives. Stationary recordings were performed from locations adjacent to intersections. First person recordings used wearable or dashboard mounted cameras to record the point of view of pedestrians. Video feeds were synchronized and coded using ChronoViz, a software tool that aids visualization and analysis of multimodal time-coded data (Fouse, et. al., 2011). During the analysis, particular attention was paid to the way road users position their vehicles or their bodies with respect to road markings and traffic control devices such as stoplights and signs, and how they would change their behavior in the presence of other road users.



Figure 1. ChronoViz Session. Intersection includes side streets with stop signs and main roadway crosswalk with no stop sign, seen from two directions. Two pedestrians waited for SUV to pass before crossing. The timelines can be seen below the images. Within each of the top two timelines are two sets of annotations (top is movement, bottom is head position). The bottom timeline contains the movement annotations for three vehicles.

Figure 1 illustrates our approach in one ChronoViz session, with intersection recordings from two angles, and multiple timelines containing point and duration annotations recording the activity of two pedestrians (top for pedestrian A, center for pedestrian B) and multiple cars (bottom). Pedestrian annotations included movement (such as walk, run, speed, position) with respect to intersection configuration (such as curb, zebra crossing, intersection boundaries) and head position of each pedestrian (forward, down, left, right). In the bottom timeline, annotations

included vehicle motion (such as speed, direction) with respect to intersection configuration. Configuration elements varied depending on the roadway recorded.

In addition to the video analysis, semi-structured interviews were conducted to gain insight into how people perceive and understand road user activity. Twenty participants viewed eleven video snippets, selected from previously recorded videos, and asked to describe the activity in each of the videos. All participants were between 18-62 years of age and in possession of a California driver's license. After completion of the initial activity description of the snippets, participants were re-shown the same video snippets and asked to elaborate on their earlier comments. During both phases participants could re-watch each snippet as often as they wanted. Participant verbalizations were transcribed and analyzed for references to road user behavior.

RESULTS

We present descriptions of observed patterns in road user behavior, and how this behavior changed in the presence of other road users in the context of road geometry and traffic control infrastructure. Additionally, results from the semi-structured interviews are presented. The observed driver and vehicle behaviors included forward motion (vehicle roll or stop and pedestrian run, walk, etc.), hand gesture, head position, eye gaze, body posture and auditory signals (voice, horn, engine revving, etc.). Three vehicle movement patterns were observed repeatedly during the video analysis: advancing, slowing early and stopping short.

Advancing. In the presence of several vehicles in the stop sign controlled intersections, drivers were observed slowly rolling forward over designated stopping lines painted on the street even when another vehicle was still in the middle of the intersection. The left image of Figure 2 shows this prevalent phenomenon at one busy 4-way stop intersection. The vehicle labeled with a red '1' was first into the intersection. Car number 2 was next in line to enter the intersection, but was already significantly progressed into the intersection as vehicle 1 was exiting. Car number 3 was next, and entered the intersection just after vehicle 2 did. The black SUV, number 4 entered just after car number 3, leaving all four vehicles moving *inside* the intersection at one time. When a vehicle did not perform this motion gesture, they often lost their turn.

Slowing Early. Frequently, vehicles allow pedestrians to pass by reducing speed early and maintaining a slow speed, all while never making a complete stop for the pedestrian (see right image Figure 2). In the image, pedestrians walk comfortably across the intersection while the black van continues slowly moving forward.

Stopping Short. One behavior that was prominent in all recordings was drivers stopping their cars far short of the location required by law. The legal rules of the road state that drivers must "stop *at* the white limit line (a wide white line painted on the street) or before entering the crosswalk" (CA Driver Handbook, 2015). In our observations, the distance cars kept from the intersection was not the distance required by law when they were in the presence of nearby pedestrians. The distance to the painted line was often more than one car length. The positions of the cars relative to the painted lines of the crosswalks in Figure 3 show typical stopping positions of drivers when pedestrians are present in the crosswalks in front of them. In fact, of the three intersections controlled by stop signs (one 3-way, two 4-way), vehicles approaching the

intersections were much more likely to stop short when pedestrians were on or near the crosswalk in front of them (52.4% of cars stopped short), than when there were no pedestrians on or near the crosswalk in front of them (11.0% did not stop short).

When drivers did not stop significantly short of a crosswalk, pedestrians often demonstrated discomfort, showing stopping short is a social norm within the road user community. Figure 4 demonstrates such an interaction showing a pedestrian staring at a car that has stopped right at the crosswalk and changing his path slightly away from the car to the edge of the zebra crossing.



Figure 2. Advancing & Slowing Early Left, vehicles are labeled with red numbers indicating their turn taking order, showing cars *advance* into intersection, indicating they are ready for their turn, even before other cars are clear of intersection. All four cars were moving *inside* the intersection at the same time. Right, vehicle begins slowing at red line “1” and then more at line “2” and continues through without stopping.



Figure 3. Images of typical stopping short movement gestures Notice the distance of the three stopped vehicles from the painted lines as well as the positions of the pedestrians in the crosswalks.



Figure 4. Pedestrian discomfort Car approached and stopped quite close to the crosswalk while pedestrian crossing. Pedestrian changed course slightly right and stared at the vehicle for an extended period of time (indicated by the head mounted camera image on the right).

The observed movement patterns illustrate how drivers’ intentions are communicated through vehicle movement. We call these observed patterns “movement gestures.” The observations also suggest that road users are using movement gestures to coordinate their behavior. Results from the semi-structured interviews show that people from the general population describe pedestrian and vehicle behavior (and specifically in the context of the road geometry and signaling type), as purposely communicative. During the interviews, the terms “indicate” and “signal” were often used when observing particular vehicle or pedestrian movements. For example, a car slowing

down early or stopped short of a crosswalk is seen by the interview participant as communicating to the pedestrian in the video snippet that it is safe for them to cross: “[car] *stopped* so she [the pedestrian] knows it's safe.” These results offer further support that movement is understood as purposefully communicative in road users interaction.

DISCUSSION AND CONCLUSION

Our study results show that similar vehicle movement patterns can be observed in a variety of intersection and roadway arrangements. Drivers stop their cars far short of where they legally must, signaling they intend to wait for others to take the right of way. They roll their cars slowly forward indicating they will take the right of way next. Many of these patterns appeared more frequently in the presence of other road users, and did so not simply to follow traffic laws or for safety reasons. Additionally, when pedestrians were present in the absence of specific vehicle behaviors, they often appeared uncomfortable. These results suggest that movement gestures are deliberately used by drivers to communicate with other road users and that there is a common understanding of the meaning of them. Furthermore, interview results suggest that road users interpret some forms of movement as purposefully communicating a certain message. Developers must understand these established behaviors even if they do not want to use them, as AVs might incidentally perform movement gestures that trigger unexpected and possibly dangerous reactions from nearby drivers and pedestrians.

A limitation of this study is that the majority of observations were made in Southern California, with its own driving culture. Driving culture influences road user interpretation of driver behavior (Shor, 1964). Research is needed to identify the use of movement gestures within different driving cultures. Image-processing techniques could improve the speed of the identification and classification of behavioral patterns (Shen et. al., 2014), so more cultures could be observed.

REFERENCES

- Ackerman, E. (2016, Nov 13). Drive.ai Solves Autonomous Cars' Communication Problem. *IEEE Spectrum*. Retrieved from <http://spectrum.ieee.org/cars-that-think/transportation/self-driving/driveai-solves-autonomous-cars-communication-problem>
- Emmenegger, C., & Risto, M., Bergen, B., Norman, D. & Hollan, J. (2016, July). *The Critical Importance of Standards for the Communication Between Autonomous Vehicles and Humans*. Poster presented at the Automated Vehicles Symposium 2016, San Francisco, CA.
- Fouse, A., Weibel, N., Hutchins, E., & Hollan, J. D. (2011). ChronoViz: a system for supporting navigation of time-coded data. *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 299-304). ACM.
- Habibovic, A., Andersson, J., Nilsson, M., Lundgren, V. M., & Nilsson, J. (2016). Evaluating interactions with non-existing automated vehicles: three Wizard of Oz approaches. *Intelligent Vehicles Symposium (IV), 2016 IEEE* (pp. 32-37). IEEE.
- Hall, E. T. (1969). *The hidden dimension*. New York, NY: Anchor Books.
- Heath, C., Hindmarsh, J., & Luff, P. (2010). Video in qualitative research. Sage Publications.

- Hoffman, G., & Ju, W. (2014). Designing robots with movement in mind. *Journal of Human-Robot Interaction*, 3(1), 89-122.
- Hutchins, E., Weibel, N., Emmenegger, C., Fouse, A., & Holder, B. (2013). An integrative approach to understanding flight crew activity. *Journal of Cognitive Engineering and Decision Making*, 7(4), 353-376.
- Lundgren, V. M., Habibovic, A., Andersson, J., Lagström, T., Nilsson, M., Sirkka, A., ... & Saluäär, D. (2017). Will There Be New Communication Needs When Introducing Automated Vehicles to the Urban Context? In *Advances in Human Aspects of Transportation* (pp. 485-497). Springer International Publishing.
- Matthews, M., & Chowdhary, G. V. (2015). Intent Communication between Autonomous Vehicles and Pedestrians. In *Robotics: Science and Systems 2015 - Workshop on Model Learning for Human-Robot Communication conference*. July 13-17. Sapienza University of Rome, Italy.
- Matthews, T., Forlizzi, J., & Rohrbach, S. (2006). *Designing a glanceable peripheral display*. Technical Report UCB/EECS-2006-113. Retrieved from <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2006/EECS-2006-113.html>
- Müller, L., Risto, M., & Emmenegger, C. (2016). The social behavior of autonomous vehicles. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp '16* (pp. 686–689). New York, NY, USA: ACM Press. DOI: 10.1145/2968219.2968561
- Nilsson, M., Thill, S., & Ziemke, T. (2015). Action and intention recognition in human interaction with autonomous vehicles. In *"Experiencing Autonomous Vehicles: Crossing the Boundaries between a Drive and a Ride" workshop in conjunction with CHI2015*.
- Pennycooke, N. (2016). *AEVITA: Designing Biomimetic Vehicle-to-Pedestrian Communication Protocols for Autonomously Operating & Parking On-Road Electric Vehicles*. Masters thesis. Massachusetts Institute of Technology (pp: 1-184).
- Shen, W., Deng, K., Bai, X., Leyvand, T., Guo, B., & Tu, Z. (2014). Exemplar-based human action pose correction. *IEEE transactions on cybernetics*, 44(7), 1053-1066.
- Shor, R. E. (1964). Shared Patterns of Nonverbal Normative Expectations in Automobile Driving. *The Journal of Social Psychology*, 62(1), 155–163. DOI: 10.1080/00224545.1964.9919512
- State of California Department of Motor Vehicles. (2015). *California Driver Handbook*. Retrieved from www.dmv.ca.gov/web/eng_pdf/dl600.pdf
- Szafir, D., Mutlu, B., & Fong, T. (2014). Communication of intent in assistive free flyers. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction* (pp. 358-365).
- Vinkhuyzen, E., & Cefkin, M. (2016). Developing Socially Acceptable Autonomous Vehicles, Proceedings of EPIC 2016, Ethnographic Praxis in Industry Conference (Aug 29 - Sep 01), Minneapolis, MN.