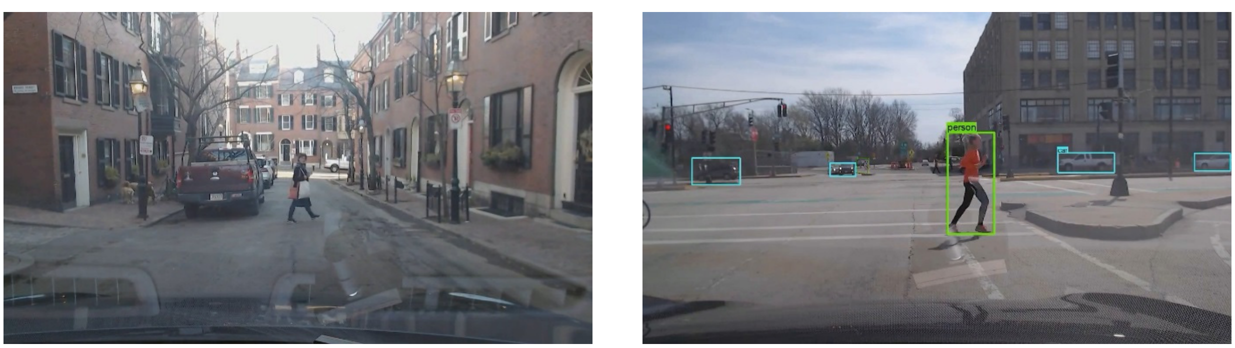
DYNAMICS OF PEDESTRIAN CROSSING DECISIONS BASED ON VEHICLE TRAJECTORIES IN LARGE-SCALE SIMULATED AND REAL-WORLD DATA

Jack Terwilliger[1], Michael Glazer[1], Henri Schmidt[1], Josh Domeyer[2], Heishiro Toyoda[2], Bruce Mehler[1,3], Bryan Reimer[1,3], Lex Fridman[1]

[1]Center for Transportation and Logistics, Mass. Institute of Technology, Cambridge, MA, USA
[2]Toyota Collaborative Safety Research Center, Ann Arbor, MI, USA
[3]MIT AgeLab & N.E. University Transportation Center, Cambridge, MA, USA
Email: fridman@mit.edu

Summary: Humans, as both pedestrians and drivers, generally skillfully navigate traffic intersections. Despite the uncertainty, danger, and the non-verbal nature of communication commonly found in these interactions, there are surprisingly few collisions considering the total number of interactions. As the role of automation technology in vehicles grows, it becomes increasingly critical to understand the relationship between pedestrian and driver behavior: how pedestrians perceive the actions of a vehicle/driver and how pedestrians make crossing decisions. The relationship between time-to-arrival (TTA) and pedestrian gap acceptance (i.e., whether a pedestrian chooses to cross under a given window of time to cross) has been extensively investigated. However, the dynamic nature of vehicle trajectories in the context of non-verbal communication has not been systematically explored. Our work provides evidence that trajectory dynamics, such as changes in TTA, can be powerful signals in the non-verbal communication between drivers and pedestrians. Moreover, we investigate these effects in both simulated and real-world datasets, both larger than have previously been considered in literature to the best of our knowledge.



(a) Example video frame of the forward roadway. (b) Pedestrian and vehicle detection via YOLO v3.

Figure 1. Example video frame and detection of vehicles and pedestrians from the MIT-AVT naturalistic driving dataset Fridman et al. (2017)

INTRODUCTION

As experienced human drivers, we take for granted our ability to reason about pedestrians' movements, intents, mental models, and conflict resolution dynamics. As pedestrian, vehicle passengers, and vehicle drivers, we quickly develop the necessary perceptual capabilities such as foresight into whether a pedestrian is likely to cross the street and the ability to communicate with pedestrians in explicit, non-verbal ways. As an illustration, consider a situation in which someone is driving through a bustling street in downtown Boston. The driver spots a pedestrian on the sidewalk in the middle of a city block walking towards the curb. She notices the pedestrian is looking in her direction. The pedestrian pauses, but then the driver decelerates. The pedestrian then jaywalks (crosses outside a crosswalk) across the street in front of the vehicle. While banal, this example encourages us to ask: (1) to what extent did the driver's influenced the pedestrian's decision to cross and (2) how the driver was able to reason about the interaction. To design vehicle automation that operates safely and efficiently in urban environments with an awareness of pedestrians, we will need answers to the above questions. In this paper, we investigate (1) the relationship between vehicle trajectories and pedestrian crossing decisions and (2) people's ability to update their estimates of a vehicle's time to arrival (TTA) when vehicles accelerate.

Previous work has recorded the TTA between vehicles and pedestrians at the moment pedestrians begin to cross the street. In 1953, Moore (1953) first showed evidence that speed and distance influence when pedestrians decide to cross and in 1955, Cohen et al. (1955) began investigating TTA. More recently, Brewer et al. (2006) found that 85% gap acceptances (i.e., instances where pedestrians choose to cross) fall between 5.3 and 9.4 seconds. While these studies have provided valuable information and models about real-world crossing behavior, to design robust safety systems and vehicle automation, it's important to understand how dynamics of trajectories, as opposed to a static notion of TTA, relate to pedestrian decision making.
In order to understand pedestrian-vehicle interaction in greater depth we investigated behaviors both in a dynamic real-world environment and through simulation that considers dynamic trajectories. We perform our analysis on two large-scale datasets.

METHODS

Large-Scale Naturalistic Data Analysis

Vehicle kinematics data originated from an approximately 200,000 mile subset of the MIT-AVT naturalistic driving dataset (Fridman, 2018; Fridman et al., 2017). The dataset includes data from Greater Boston area drivers in vehicles equipped with automation technologies throughout medium (1 month) and long-term (over a year) observations. This dataset contains video, vehicle kinematics, and various messages from a vehicle's systems. The video data include 720p 30fps video of (1) the forward roadway, (2) the driver's face, and (3) the instrument cluster. Vehicle kinematic data include odometer, speedometer, and steering angle information recorded via the vehicle's CAN bus diagnostic port as well as GPS and IMU data collected via an installed data collection system. Signals from a vehicle's computer include previously mentioned kinematic data, whether the brake was activated when a forward collision warning occurred etc.; a data collection system recorded these signals via a CAN diagnostic port. For this study, we used video of the forward roadway, vehicle kinematics, and GPS data. To ensure the integrity of our analyses, all data were synchronized to video frames of the forward roadway. See Fig. 1 for examples of forward roadway video frames and detections of pedestrians and vehicles.

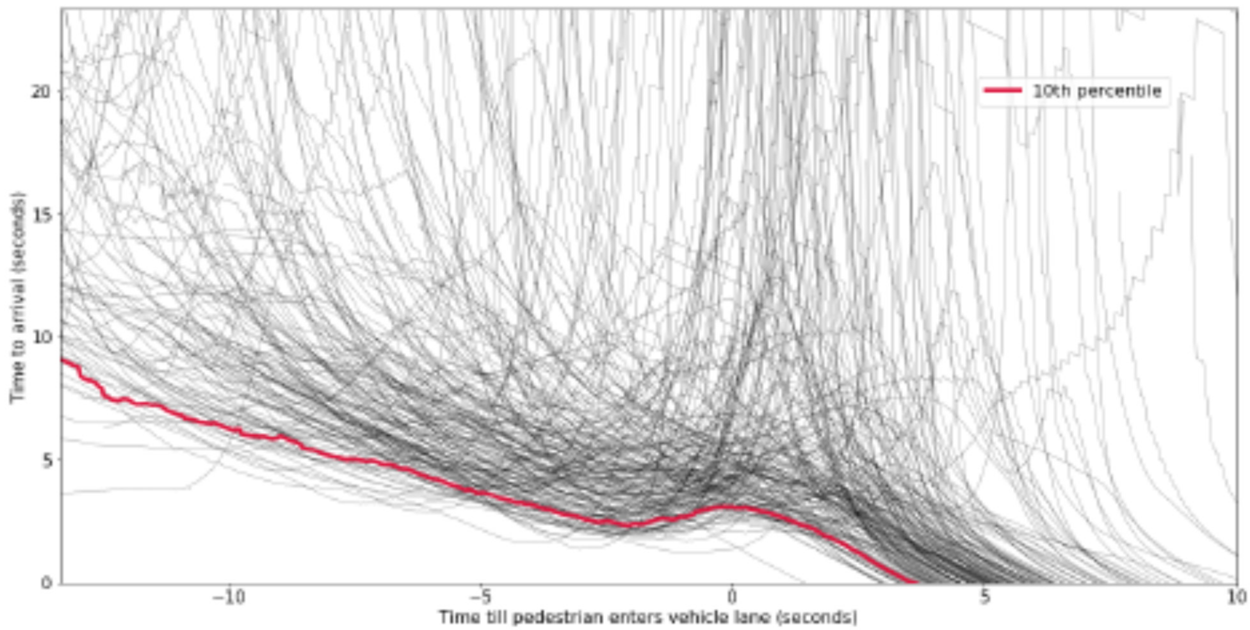


Figure 2. TTA $\frac{\text{velocity}}{\text{distance}}$ in seconds. All trajectories are aligned on the frame the pedestrian entered the path of the oncoming vehicle.

In order to study how vehicle kinematics influence pedestrian behavior at intersections, we needed to extract and annotate instances of short interactions between drivers and pedestrians. Below, we outline a pipeline which involves (1) a kinematics-based filter which excludes most highway driving (2) a computer vision approach which extracts situations in which pedestrians likely crossed the street, (3) a manual filter which selects only those interactions that fall within a set of study criteria, and (4) a manual annotation tool for labeling crossing-related events (e.g., entering the roadway, entering the path of the approaching vehicle, etc.) and pedestrian body language (e.g., head orientation, hand-waving, walking, standing, etc.). Note that the order of pipeline ensured that more costly steps operate over the least amount of data.

Pedestrian Detection. In order to extract sections of driving in which pedestrians likely crossed the street, we, first, processed the remaining forward roadway video using YOLO v3 (Redmon and Farhadi, 2018), a real-time visual object detection system. In the context of computer vision, object detection is the problem of classifying and localizing (via bounding boxes) multiple objects in an image. There are several practical advantages to YOLO v3, (a) YOLO v3 is a deep learning based architecture which does not require manually crafted image features, (b) YOLO v3 can process video 4x faster than comparable alternatives (at 30fps on modern consumer hardware)(Redmon and Farhadi, 2018), and (c) we were able to detect the presence of more object classes than just pedestrian, which provides value for future related research.

We then extracted 30-second video clips of the detected pedestrian crossings: 20 seconds prior to the frame with a crossing pedestrian and 10 seconds after it. If two videos overlapped, we combined them into one video.

Manual Annotation of Crossing Event Characteristics. In order to label crossing-related events and pedestrian body pose, we manually annotated the videos using a custom OpenCV/Python tool. All annotations were of or relative to the lead pedestrian. Body pose included (a) whether a pedestrian's head was oriented toward or away from the driver or whether it was oriented down, (b) whether the pedestrian was standing, walking, or running, (c) whether the pedestrian waved at the vehicle. Crossing events included (a) when the pedestrian entered the roadway, i.e. when the pedestrian stepped onto the roadway (b) when the pedestrian entered the paths of the ego-vehicle, which may occur after the pedestrian steps onto the road (c) when the pedestrians exited the path of the ego vehicle, (d) when the pedestrian exited the roadway and (e) when the vehicle crossed the path the pedestrian took to cross, i.e. the point where the pedestrian's and the vehicle's paths crossed. Features of the intersection included (a) whether the intersection occurred at a stop light, (b) whether the intersection included a zebra crossing, (c) whether the pedestrian was jay-walking.

Simulator Experiment

Our simulator experiment tested people's ability to estimate TTA under conditions when (1) vehicles approached at a constant velocity and (2) vehicles approached while decelerating. This experiment was designed and conducted to supplement the large-scale real-world dataset of pedestrian crossing in order to analyze the nuance of vehicle trajectory dynamics as they relate to pedestrian crossing decision. In real-world data, we cannot control either the pedestrians nor the vehicles, but simple observe and analyze the kinematics of both. In the virtual environment, we can control the vehicle trajectory and observe its effect on the pedestrian crossing decision.

Design

In this experiment, we tasked participants with estimating the time to arrival (TTA) of a vehicle: or specifically, with estimating when a vehicle would reach a white line painted across a virtual road. In each trial, after traveling some distance, the vehicle disappeared before reaching the white line. This forced participants to estimate the TTA based on prior kinematic information. We define the ground truth TTA as the time between the moment a vehicle disappeared and the time it would arrive at the white line. We measured participants' estimated TTA by asking them to press the spacebar on their computer when they thought the vehicle would reach the white line – estimated TTA is thus the time between the moment the vehicle disappeared and the time a participant pressed the spacebar.

RESULTS

Large Scale Naturalistic Data Analysis

We now illustrate the characteristics of vehicle trajectories found "in the wild" in situations where pedestrians chose to cross. Specifically, we show (1) evidence that temporal dynamics influence pedestrian decision-making, and (2) results convergent with (Petzoldt, 2014) which suggest that, while pedestrians use TTA when deciding whether or not to cross, they underestimate the TTA at higher velocities.

(1) In Fig. 2 we show 284 vehicle trajectories (TTA over time) relative to the moment a lead pedestrian entered the path of the vehicle. While it may appear redundant to plot TTA over time, because vehicles accelerate/decelerate as they approach, in order to accurately estimate the time they have to cross, a pedestrian must update their estimates over time.

We see a trend, 34% of drivers slow the vehicle such that the time to collision increases before the pedestrian steps in front of their vehicle. Here, TTA refers to a simple linear extrapolation of vehicle kinematics, i.e. $\frac{\text{velocity}}{\text{distance}}$. To normalize the data, we align each trajectory on the frame in which an annotator determined a pedestrian entered the path of the oncoming vehicle. Though we are unable, with these data, to ask the counterfactual "what if the driver had not slowed down?", these data suggest that, in real-world situations, pedestrians tend only to cross when vehicles slow down such that the time the pedestrian has to cross increases.

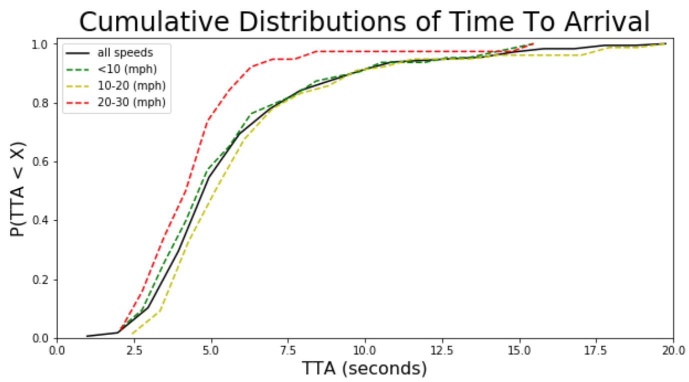


Figure 3. TTA $\frac{\text{velocity}}{\text{distance}}$ in seconds at the moment a pedestrian entered the path of the oncoming vehicle.

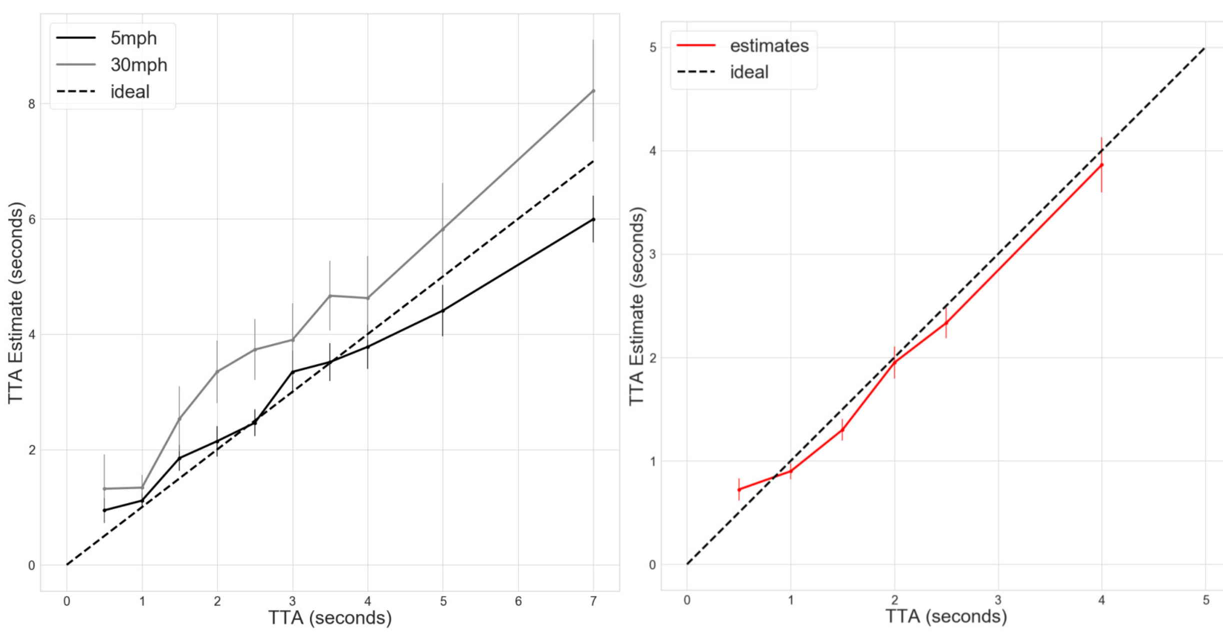


Figure 4. (Left) Participants' estimates of TTA of a vehicle traveling at constant velocities: 5 mph (solid black) and at 30 mph (solid grey). Participants overestimated TTA when the vehicle traveled at 30 mph. Participants underestimated TTA when the vehicle traveled at 5 mph when the car was far away. (Right) Participants' estimates of TTA of a vehicle traveling at a constant velocity and then decelerating.

(2) In Fig. 3 we show the empirical cumulative distributions of TTA at the moment crossing pedestrians entered the path of the oncoming vehicle N=195 (we removed cases where

TTA was greater than 20). Performing a Kolmogorov-Smirnov test between each category of vehicle speed indicates a significant difference between when pedestrians cross the street in cases where vehicles traveled between 10-20 mph and cases where vehicles traveled between 20-30 mph (D-statistic=0.15, p¡0.05). The test does not indicate significant differences between any other pair of vehicle speed categories see Table 1.

Table 1. Kolmogorov-Smirnov Test Table

Samples	D-Statistic	p-value
<10 mph & 10-20 mph	0.15	0.375
<10 mph & 20-30 mph	0.22	0.145
10-20 mph & 20-30 mph	0.30	0.011*

Results of a Kolmogorov-Smirnov Test between each pair of the three vehicle speed categories.

These results, taken from unconstrained real-world situations, provide strong supplementary evidence, that pedestrians base their decision of when to cross on TTA. We find, surprisingly, at higher speeds, pedestrians enter the lane with less time than at lower speeds. According to (Petzoldt, 2014), pedestrians overestimate the TTA at higher speeds – a result consistent with other literature (Hancock and Manster, 1997) (Sidaway et al., 1996). We note that (Petzoldt, 2014) did not find evidence that overestimating TTA influenced gap acceptance. The Petzoldt (Petzoldt, 2014) study was conducted in a lab setting and the differences between our findings and theirs may be the result of their participants becoming aware of and correcting for their tendency to overestimate the TTA in a predictable environment.

Simulator Experiment

We now illustrate the results of how our participants were able to estimate TTA when a vehicle was traveling at a constant velocity and when a vehicle was decelerating.

In Fig. 4 (left), we show evidence that people overestimate TTA of vehicles traveling at higher velocities. The plot shows the ground truth TTA (x axis) vs. participants' estimates of TTA (y axis). The dashed black line (x=y) shows what an ideal estimator would look like. Estimates above the dashed line are over estimates; estimates below the dashed line are under estimates. This data agrees with (Petzoldt, 2014) that vehicle speed influences TTA estimates. This suggests the source of our findings from naturalistic study (that pedestrians enter the lane sooner under less TTA when vehicles are traveling at high speeds) is based on the perceptual bias – to overestimate TTA when vehicles are traveling at high speeds.

In Fig. 4 (right), we show that people are sensitive to changes of speed and are able to rapidly update their estimates of the kinematics of oncoming vehicles. As in the previous plot, this plot shows the ground truth TTA (x axis) vs. participants' estimates of TTA (y axis). This demonstrates that, as expected, people are able to rapidly update their estimates of the kinematics of oncoming vehicles. This result provides grounds for interpreting our findings that drivers alter their trajectories as they approach pedestrians as a non-verbal signal, which pedestrians may use to infer the intent of drivers.

CONCLUSION

As more of the driving task becomes automated, we must deepen our knowledge of how pedestrians react to trajectories of human-driven vehicles. Closing this knowledge gap is important for developing both effective autonomous motion planning algorithms and communication protocols in a mixed fleet that includes vehicles controlled both by humans and machines.

Here we have shown evidence that (1) in real-world situations pedestrian decision-making is biased – they tend to give themselves less time when vehicles travel at faster speeds, (2) dynamics of vehicle trajectories, namely increases in TTA, appear to serve as signals that it is safe to cross, and (3) that people can update their estimates of TTA as vehicles change speed. While these results provide a pragmatic conclusion, that automated technology ought to account for human bias to overestimate TTA at higher speeds, they also motivate the need to further study of dynamic trajectories in order to understand pedestrian-driver interactions at a more nuanced level.

ACKNOWLEDGMENT

This work was in part supported by the Toyota Collaborative Safety Research Center. The views and conclusions being expressed are those of the authors and do not necessarily reflect those of Toyota.

REFERENCES

M.A. Brewer, K. Fitzpatrick, J.A. Whitacre, and D. Lord. Exploration of pedestrian gap-acceptance behavior at selected locations. *Transportation research record,* 1982(1):132-140, 2006.

J. Cohean, E. Dearnaley, and C. Hansel. The risk taken in crossing a road. *Journal of the Operational Research Society*, 6(3): 120-128, 1955.

L. Fridman. Human-centered autonomous vehicle systems: Principles of effective shared autonomy. *CoRR*, abs/1810.01835, 2018 URL https://arxiv.org/abs/1810.01835.

L. Fridman, D. E. Brown, M. Glazer, W. Angell, S. Dodd, B. Jenik, J. Terwilliger, J. Kindelsberger, L. Ding, S. Seaman, H. Abraham, A. Mehler, A. Sipperley, A. Pettinato, L. Angell, B. Mehler, and B. Reimer. MIT autonomous vehicle technology study: Large-scale deep learning based analysis of driver behavior and interaction with automation. *CoRR*, abs/1711.06976, 2017. URL https://arxiv.org/abs/1711.06976.

P.A. Hancock and M. Manster. Time-to-contact: More than tau alone. *Ecological Psychology*, 9(4):265-297, 1997.

R.L. Moore. Pedestrian choice and judgment. *Journal of the Operational Research Society*, 4(1):3-10, 1953.

T. Petzoldt. On the relationship between pedestrian gap acceptance and time to arrival estimates. *Accident Analysis & Prevention*, 72:127-133, 2014.

J. Redmon and A. Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.

B. Sidaway, M. Fairweather, H. Sekiya, and J. Mcnitt-Gray. Time-to-collision estimation in a simulated driving task. *Human factors*, 38(1):101-113, 1996.