

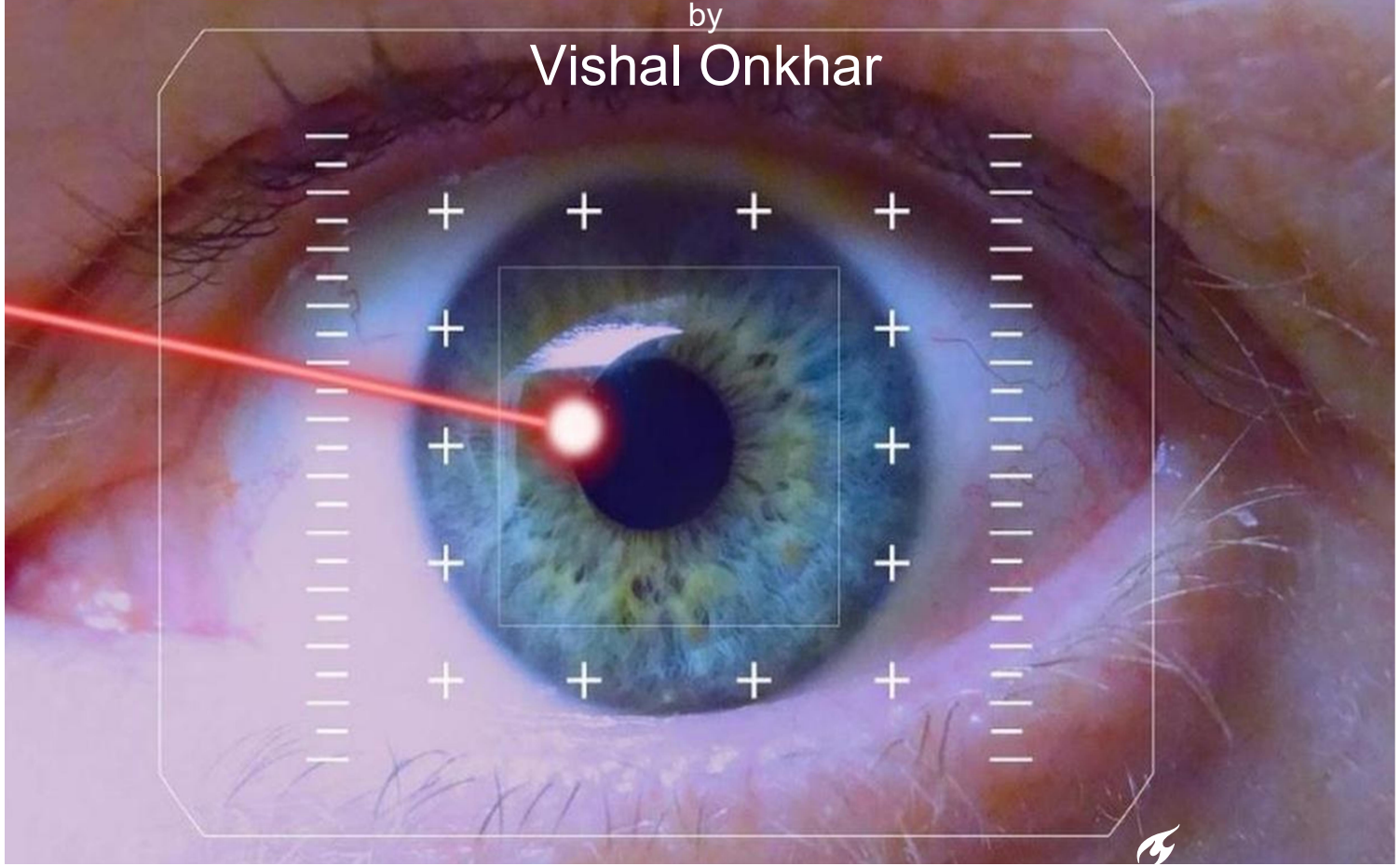


# MSc Thesis Report

Algorithmic detection of eye contact in driver-  
pedestrian interactions

by

Vishal Onkhar





# MSc Thesis Report

## Algorithmic detection of eye contact in driver-pedestrian interactions

by

Vishal Onkhar  
4741269

in partial fulfilment of the requirements for the degree of

**Master of Science**  
in Mechanical Engineering

at the Delft University of Technology

Supervisors: Dr.ir. J. C. F. de Winter  
Dr. P. Bazilinskyy  
Ir. J. C. J. Stapel

Project duration: February 2019 – February 2020

Thesis committee: Dr. ir. J. C. F. de Winter (Chair)  
Dr. ir. R. Happee  
Dr. P. Bazilinskyy  
Ir. J. C. J. Stapel



# Acknowledgement

As a mechanical engineer, the transition to pursue automotive human factors was both a challenging and rewarding one. It thrust a variety of topics such as statistics, psychology and computer programming at me, encouraging me to learn on the fly and ultimately helping me become a well-rounded engineer. This thesis also introduced me to the many opportunities in the rapidly growing fields of eye tracking and driver-pedestrian interaction, areas I would not have otherwise explored. My thesis journey would not have been possible without some people, so a round of heartfelt thanks is in order. First, I'd like to express my gratitude to my supervisor Dr. Joost de Winter for his guidance at every step of the way and his patience during times of slow progress. His faith in my abilities kept me motivated to successfully complete my thesis. Next, I'd like to thank my daily supervisor and tovarish Dr. Pavlo Bazilinskyy for proposing the idea of the project and our fruitful discussions over many games of chess. I owe a great debt of thanks to Jork Stapel who spent several hours of his time helping me set up and conduct my experiments, not to mention making sense of the data and mathematics. Further, I'm grateful to Reverend Lars Kooijman for always being available to help at short notice, his MATLAB coding expertise, and for offering his unerring advice in times of need. We have bonded over memes and common interests over the course of my thesis, and I hope our friendship endures into the future. I would be remiss if I did not mention the valuable contributions of my friends Anand Sudha, Arjun Anantharaman and Sparsh Sharma in carrying out my experiments and motivating me during my thesis. Thanks a lot, guys. I couldn't have done it without you. Furthermore, I'd like to thank all the participants who took time out of their busy schedules to help me conduct my research. Last but not least, I'm very grateful to my parents and family for their financial and emotional support during the long hours spent working on this thesis.



# Abstract

Pedestrians today are very vulnerable on urban roads. Clear communication between drivers and pedestrians is one way to reduce their plight. Non-verbal communication in particular plays an important role in road safety, and eye contact is a kind of non-verbal communication that has the potential to minimize on-road collisions. However, with the advent of automated vehicles, driver-pedestrian eye contact loses its meaning since there is no longer a driver. It is therefore useful to study and detect eye contact so that the knowledge obtained may be applied to automated vehicles of the future. To this end, the following research goals were adopted: (a) What is eye contact between a pedestrian and a driver in a car? How can eye contact be defined/operationalized using an algorithm?, (b) How accurate is the algorithm that operationalizes eye contact?, and (c) How is it possible to use two eye-trackers with inertial measurement units (IMUs) and pedestrian recognition in a Toyota Prius car to reconstruct the entire driver-pedestrian interaction through a 3-D animation? An indoor experiment, designed to resemble a driver-pedestrian interaction at a pedestrian crossing was conducted with 31 participants. Participants' (pedestrians') eyes were tracked using a Tobii Pro Glasses 2 eye-tracker and the researcher's (driver's) eyes were tracked using a Smart Eye Pro dx eye-tracker, both of which were synchronized. Participants' locations were also tracked using a stereo camera equipped with pedestrian detection capabilities. Pedestrians imagined that they were on a real road and performed six types of trials where they stood on / crossed from the left / right side curb in front of the stationary vehicle while either making eye contact or not making eye contact with the driver. The order of the trials was randomized, and each trial consisted of 3 repetitions of a driver-pedestrian interaction. If the driver and pedestrian were looking at each other at the same time there was eye contact, otherwise there was no eye contact. Significant differences in the percentages of eye contact between pedestrians standing on the left (median duration of 0.42 s) and the right (median duration of 0.54 s) were found. No significant differences in the percentages of eye contact between pedestrians crossing from the left (median duration of 1.23 s) and the right (median duration of 1.39 s) were found. Eye contact instants within trials were algorithmically detected by finding the angle between the 3-D gaze direction vectors of the driver and the pedestrian, and comparing it to an 'eye contact threshold'. Trials were classified as either involving eye contact or not involving eye contact based on their percentages of eye contact instants. The classification performance of the algorithm was quantified using two ground truths: (1) Imposed eye contact (in half of the trials, participants were instructed to make eye contact; in the other half, participants were instructed not to make eye contact), and (2) Manually annotated areas of interest (AOIs) from the Tobii Pro Glasses 2 showing pedestrian eye contact seeking. The algorithm's performance was found to be fair/poor and eye contact could be detected with an accuracy of 15-30%. A 3-D reconstruction of the driver-pedestrian interaction was achieved (in the form of an animation) by using the locations, head orientations and gaze directions of the driver and the pedestrian. This thesis provides objective measurements of driver-pedestrian eye contact and demonstrates how eye contact may be detected and reconstructed for use in automated vehicles of the future.



# Contents

1. Introduction .....	12
2. Methods .....	16
2.1. Experiment design and procedure .....	16
I. Participants.....	16
II. Equipment and software.....	16
III. Pilot test.....	18
IV. Design of the main experiment.....	19
V. Equipment setup.....	25
VI. Analysis.....	26
2.2. Gaze ray tracing/eye contact detection algorithm.....	30
3. Results .....	41
I. Participants and Tobii gaze data quality.....	41
II. Responses to post-experiment questionnaire.....	43
III. Definition of eye contact (Research Question 1).....	44
IV. Algorithm performance (Research Question 2).....	54
V. Reconstruction of interaction scenario (Research Question 3).....	63
4. Discussion.....	71
5. Conclusions.....	76
6. Recommendations.....	78
7. Bibliography.....	79
Appendix.....	82



# List of figures

<b>Figure 1:</b> <i>Concept of driver-pedestrian eye contact at a crossing conflict</i>	13
<b>Figure 2a:</b> <i>Eyes on a car</i>	14
<b>Figure 2b:</b> <i>Jaguar Eyes</i>	14
<b>Figure 3:</b> <i>Smart Eye Pro dx with 4 cameras mounted on the dashboard</i>	16
<b>Figure 4a:</b> <i>iDS UI-3060CP-C-HQ Rev. 2 stereo camera lens</i>	17
<b>Figure 4b:</b> <i>NexTorch high-intensity flashlight</i>	17
<b>Figure 5:</b> <i>Working of the Tobii Pro Glasses 2</i>	18
<b>Figure 6:</b> <i>Tobii's collection of gaze data</i>	18
<b>Figure 7:</b> <i>Pedestrian's role in the experiment</i>	20
<b>Figure 8:</b> <i>Pedestrian's perspective obtained via Tobii eye-tracking</i>	20
<b>Figure 9:</b> <i>Moments after the start of a trial</i>	23
<b>Figure 10:</b> <i>Driver's role in the experiment</i>	24
<b>Figure 11:</b> <i>Driver's perspective and Smart Eye tracking</i>	24
<b>Figure 12:</b> <i>Calibration of the Smart Eye using the 'checkerboard'</i>	26
<b>Figure 13a:</b> <i>Annotating events and times of interest (TOIs) in Tobii Pro Lab</i>	27
<b>Figure 13b:</b> <i>Drawing areas of interest (AOIs) in Tobii Pro Lab</i>	29
<b>Figure 14:</b> <i>Flowchart of eye contact detection algorithm</i>	31
<b>Figure 15:</b> <i>Regions of eye contact</i>	33
<b>Figure 16:</b> <i>Experiment layout</i>	33
<b>Figure 17a:</b> <i>Binary plot of eye contact for the ground truth of a standing trial</i>	37
<b>Figure 17b:</b> <i>Binary plot of eye contact for the algorithm's detections during the same standing trial</i>	37
<b>Figure 18a:</b> <i>Binary plot of eye contact for the ground truth of a crossing trial</i>	38
<b>Figure 18b:</b> <i>Binary plot of eye contact for the algorithm's detections during the same crossing trial</i>	38
<b>Figure 19a:</b> <i>Top view from the algorithmic reconstruction of a standing trial showing eye contact</i>	39
<b>Figure 19b:</b> <i>Top view from the algorithmic reconstruction of a standing trial showing the lack of eye contact</i>	39
<b>Figure 20a:</b> <i>Isometric view from the algorithmic reconstruction of a crossing trial showing eye contact</i>	40
<b>Figure 20b:</b> <i>Isometric view from the algorithmic reconstruction of a crossing trial showing the lack of eye contact</i>	40
<b>Figure 21:</b> <i>Data quality check using gaze sample percentages over entire recordings</i>	41



<b>Figure 22:</b> <i>Data quality check using gaze sample percentages over trials</i>	42
<b>Figure 23:</b> <i>Eye contact as a percentage of standing trial duration</i>	44
<b>Figure 24:</b> <i>Eye contact as a percentage of crossing trial duration</i>	45
<b>Figure 25:</b> <i>Standing trial durations or the time between a 'start' torch flash and an 'end' torch flash</i>	46
<b>Figure 26:</b> <i>Crossing trial durations or the time between a 'start' torch flash and an 'end' torch flash</i>	47
<b>Figure 27:</b> <i>Pedestrian eye contact seeking as a percentage of standing trial duration</i>	47
<b>Figure 28:</b> <i>Pedestrian eye contact seeking as a percentage of crossing trial duration</i>	48
<b>Figure 29a:</b> <i>Pedestrian eye contact seeking in a standing trial, involving 3 repetitions, each comprised of 1 visit</i>	49
<b>Figure 29b:</b> <i>Pedestrian eye contact seeking in a standing trial, involving 3 repetitions, each comprised of 1-2 visits</i>	49
<b>Figure 30a:</b> <i>Driver eye contact seeking in the same standing trial as Fig. 29a, involving 3 repetitions, each comprised of 1 visit</i>	50
<b>Figure 30b:</b> <i>Driver eye contact seeking in the same standing trial as Fig. 29b, involving 4 repetitions, each comprised of 1 visit</i>	50
<b>Figure 31a:</b> <i>Eye contact as the overlap between driver and pedestrian eye contact seeking in Fig. 29a and Fig. 30a</i>	51
<b>Figure 31b:</b> <i>Eye contact as the overlap between driver and pedestrian eye contact seeking in Fig. 29b and Fig. 30b</i>	51
<b>Figure 32:</b> <i>Pedestrian dwell times for standing trials</i>	52
<b>Figure 33:</b> <i>Pedestrian dwell times for crossing trials</i>	52
<b>Figure 34a:</b> <i>Pedestrian eye contact seeking in a crossing trial, involving multiple visits</i>	53
<b>Figure 34b:</b> <i>Driver eye contact seeking in the same crossing trial, involving one continuous visit</i>	54
<b>Figure 34c:</b> <i>Fragmented eye contact in the same crossing trial</i>	54
<b>Figure 35a:</b> <i>Eye contact binary plot (algorithm) for a standing trial using the theoretical eye contact threshold value</i>	55
<b>Figure 35b:</b> <i>Eye contact binary plot (AOI ground truth) for the same standing trial using the theoretical ground truth threshold value</i>	56
<b>Figure 36:</b> <i>ROC curves depicting algorithm performance for standing trials with eye contact</i>	57
<b>Figure 37:</b> <i>ROC curves depicting algorithm performance for crossing trials with eye contact</i>	58
<b>Figure 38:</b> <i>False positive rate curves depicting algorithm performance for standing trials with no eye contact</i>	58
<b>Figure 39:</b> <i>False positive rate curves depicting algorithm performance for crossing trials with no eye contact</i>	59
<b>Figure 40:</b> <i>Areas under the ROC curves showing algorithm accuracy for the standing trials with eye contact</i>	59
<b>Figure 41:</b> <i>Areas under the ROC curves showing algorithm accuracy for the crossing trials with eye contact</i>	60
<b>Figure 42a:</b> <i>Isometric view of a standing trial showing a moment when driver and pedestrian were looking directly at each other</i>	63



<b>Figure 42b:</b> <i>Isometric view of a standing trial showing a moment when driver and pedestrian were not looking at each other</i>	64
<b>Figure 43:</b> <i>Isometric view of a standing trial showing a moment when the driver was looking at the pedestrian but the pedestrian was probably looking at the left side-view mirror of the Prius</i>	64
<b>Figure 44:</b> <i>Isometric view of a standing trial showing a moment when the driver was looking at the pedestrian but the pedestrian was probably looking at the number plate of the Prius</i>	65
<b>Figure 45a:</b> <i>Isometric view of a standing trial showing a moment when driver and pedestrian were looking directly at each other</i>	65
<b>Figure 45b:</b> <i>Isometric view of a standing trial showing a moment when driver and pedestrian were not looking at each other</i>	66
<b>Figure 46:</b> <i>Isometric view of a crossing trial showing a moment when driver and pedestrian were looking directly at each other and the pedestrian was crossing from the left side</i>	66
<b>Figure 47:</b> <i>Isometric view of a crossing trial showing a moment when the driver was looking at the pedestrian but the pedestrian was probably looking at the number plate of the Prius, and the pedestrian was crossing from the left side</i>	67
<b>Figure 48:</b> <i>Top view of a standing trial showing a moment when driver and pedestrian were looking directly at each other</i>	67
<b>Figure 49:</b> <i>Top view of a standing trial showing a moment when the driver was looking at the pedestrian but the pedestrian was not looking at the vehicle (contrary to the provided instructions)</i>	68
<b>Figure 50:</b> <i>Top view of a standing trial showing a moment when the driver was looking at the pedestrian but the pedestrian was probably looking at the right side-view mirror of the Prius</i>	68
<b>Figure 51:</b> <i>Top view of a standing trial showing a moment when driver and pedestrian were looking directly at each other</i>	69
<b>Figure 52:</b> <i>Top view of a crossing trial showing a moment when driver and pedestrian were looking directly at each other and the pedestrian was crossing from the left side</i>	69
<b>Figure 53:</b> <i>Top view of a crossing trial showing a moment when the driver was not looking at the pedestrian and the pedestrian was probably looking at the right side-view mirror of the Prius while crossing from the right side</i>	70



# List of tables

<b>Table 1:</b> <i>Experiment scenarios or trials</i>	21
<b>Table 2:</b> <i>Tobii IV-T Fixation Filter settings</i>	29
<b>Table 3:</b> <i>Friedman test ANOVA table between the four types of standing trials</i>	41
<b>Table 4:</b> <i>Friedman test ANOVA table between the two types of crossing trials</i>	42
<b>Table 5:</b> <i>Friedman test ANOVA table between all types of trials</i>	42
<b>Table 6:</b> <i>Summary of responses to the post-experiment questionnaire</i>	43
<b>Table 7:</b> <i>Summary of actual values of the ground truth threshold</i>	57
<b>Table 8:</b> <i>Accuracy classification using area under an ROC curve</i>	60
<b>Table 9:</b> <i>Area under the curve (AUC) and TPR/FPR trade-off</i>	60
<b>Table 10:</b> <i>Algorithm performance at classifying all standing trials correctly as 'Eye Contact' or 'No Eye Contact'</i>	61
<b>Table 11:</b> <i>Algorithm performance at classifying all crossing trials correctly as 'Eye Contact' or 'No Eye Contact'</i>	61



# List of equations

<b>Equation 1:</b> <i>Calculation of the region of eye contact surrounding the driver's face</i>	28
<b>Equation 2:</b> <i>Total uncertainty is the theoretical 'eye contact threshold' for the algorithm</i>	32
<b>Equation 3:</b> <i>Total uncertainty (without the Tobii terms) is the theoretical 'ground truth threshold' for the Smart Eye in the ground truth</i>	32
<b>Equation 4:</b> <i>Calculation of typical eye contact durations for a standing pedestrian on the curb</i>	45
<b>Equation 5:</b> <i>Calculation of maximum eye contact durations for a crossing pedestrian on the road</i>	46
<b>Equation 6:</b> <i>Mean 'M' of actual ground truth thresholds for standing and crossing trials</i>	57
<b>Equation 7:</b> <i>Region of driver eye contact seeking in the ground truth</i>	57



# Abbreviations

EC	'Eye Contact'
NEC	'No Eye Contact'
L-S-EC	'Left Standing Eye Contact' trial
R-S-EC	'Right Standing Eye Contact' trial
L-S-NEC	'Left Standing No Eye Contact' trial
R-S-NEC	'Right Standing No Eye Contact' trial
W-EC	'Walking Eye Contact' trial
W-NEC	'Walking No Eye Contact' trial
L-W-EC	'Left Walking Eye Contact' half repetitions
R-W-EC	'Right Walking Eye Contact' half repetitions
L-W-NEC	'Left Walking No Eye Contact' half repetitions
R-W-NEC	'Right Walking No Eye Contact' half repetitions
IMU	Inertial measurement unit
TOI	Time of interest
AOI	Area of interest
ROS	Robot Operating System
P	Pedestrian
D	Driver
ROC	Receiver operating characteristics
TPR	True positive rate
FPR	False positive rate
AUC	Area under the curve

**Note 1:** 'Left/Right' sides refer to the standing/starting location of the pedestrian from the driver's/car's perspective

**Note 2:** 'Standing/Walking' refer to the pedestrian's action during a trial in the experiment

**Note 3:** 'Eye Contact/No Eye Contact' refer to the instructions provided to a pedestrian before a trial in the experiment

**Note 4:** 'Walking' has the same meaning as 'Crossing'



# 1 Introduction

There were 54 pedestrian deaths in the Netherlands in 2018, which was 8% of all traffic-related deaths in the country (CBS, 2018). Worldwide, 23% of deaths on the road are pedestrians, which amount to over 300,000 casualties, a disproportionately high figure (WHO, 2018). In the USA, pedestrian fatalities in 2018 were the highest they have been since 1990 and are increasing every year. In the last decade alone, there was a 35% increase in pedestrian deaths in the USA. Further, 35% of pedestrians in the USA were killed on local streets, and this number went up by 46% since 2008, suggesting a need for safer road crossings (GHSA, 2019; NHTSA, 2018). Most accidents involving pedestrians were found to occur during the act of street crossing (daSilva et al., 2004). Thus, pedestrians are vulnerable on urban roads and effective measures are needed to help them.

Various possible causes of the high number of pedestrian fatalities can be identified. First, pedestrians are largely ignored in the planning, design, and operation of roads (WHO, 2018). Another possible reason for these casualties is the rise in mobile phone usage by pedestrians while on the road (Nasar & Troyer, 2013). Yet other causes include population growth, which leads to more vehicles and pedestrians on the road, and a shift in sales from passenger cars to more massive alternatives like SUVs and light-duty trucks (GHSA, 2019).

In signalized crossings, traffic lights provide clear guidelines to road users on whether to yield or proceed. Traffic lights provide safety and reduce the need for direct communication between parties involved in a crossing conflict. At unsignalized crossings, however, people rely more on non-verbal cues to negotiate the right of way. These non-verbal signs have been shown to play an important role in the safe interaction of drivers and pedestrians (Farber, 2016; Habibovic et al., 2018). Although road crossing can occur even in the absence of nonverbal communication, pedestrians, in general, appreciate the acknowledgment from a driver that they have been noticed (Rothenbucher et al., 2015; Kooijman, 2018). This claim about the significance of non-verbal communication in a traffic setting is not without its skeptics, however. Some studies show that many pedestrians cross in front of a vehicle without any explicit signaling and use only its motion to arrive at their crossing decision (Moore et al., 2019; Dey & Terken, 2017). By extension, this finding would suggest that driver-pedestrian communication may not even be required for negotiating right of way. Close examination of the individual types of non-verbal communication would be useful to understand whether this claim is true or not.

Eye contact is one informal mode of communication on the road (Rasouli et al., 2017). Eye contact promotes a calm interaction between vehicle and pedestrian due to a feeling of safety arising from the knowledge of having been noticed (Lundgren et al., 2017). Literature also shows that it is a popular means of communication, with 84% of pedestrians and 34% of drivers seeking eye contact (Sucha et al., 2017); the difference can be explained by the unequal power distribution between the two during a crossing conflict. Over 50% of pedestrians were also observed trying to make eye contact with drivers in a 30 km/h speed limit zone (Schneemann & Gohl, 2016). Eye contact was also the most common reason mentioned by people for feeling safe to cross a street (Lundgren et al., 2017). Belief in the significance of eye contact in interactions on the road can also be found outside academic circles. This fact is evident from notices, flyers, pamphlets, posters, and billboards issued by traffic safety organizations all over the world. An example is shown in Figure 1.

Recent findings suggest that pedestrians misattribute to eye contact crossing decisions that were actually arrived at through vehicle kinematics. Apparently, in most situations, pedestrians cannot even see drivers, let alone make eye contact (AlAdawy et al., 2019). Apparently, despite the interior of approaching cars being dark and reflections on the windshield making it difficult to establish direct eye contact, pedestrians tend to seek information about the intention of the driver when the vehicle is close (Dey et al., 2019). Possibly, pedestrians think they establish eye contact when in fact they only gather information about the car's motion. More studies supporting these new arguments would be required to overturn the accepted beliefs about driver-pedestrian eye contact.

The concept of driver-pedestrian eye contact faces a unique challenge in the case of automated vehicles due to the absence of a human driver. It is important to understand the role eye contact currently plays in crossing conflicts and to detect eye contact reliably. This knowledge would benefit both the road users of today by defining eye contact



and its meaning in the traffic context, and future generations by allowing the implementation of eye contact between passengers and vulnerable road users (pedestrians and cyclists) in the decision-making processes of automated vehicles.

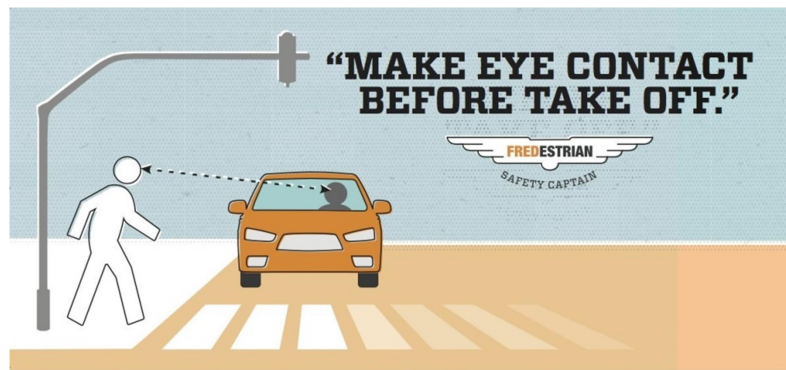


Fig. 1 – Concept of driver-pedestrian eye contact at a crossing conflict (CSP Public Affairs, 2018)

Posters like this issued by transportation departments illustrate the prevalent view about the importance of eye contact on the road.

There already exists some research which aimed to induce and study the effects of driver-pedestrian eye contact. Studies usually considered eye contact from one perspective – either the pedestrian's or the driver's. A few studies dealt with the pedestrian's end. The visual scanning behavior of pedestrians in a parking lot was studied using Tobii Pro Glasses 2, and it was found that pedestrians sought situational cues from salient features such as car wheels, headlights, and drivers to assess the intentions of nearby vehicles. It was inferred that eye contact was part of pedestrians' visual search strategy, and looking at the human driver was for the purpose of establishing eye contact (Hopmans et al., 2018). Pedestrians also demonstrated a greater willingness to cross the road upon establishing eye contact with a driver or even a virtual reality vehicle with interactive eyes, measured using a questionnaire and a button response, respectively (Lundgren et al., 2017; Chang et al., 2017). Examples of concepts of pedestrians making 'eye contact' with vehicles in studies can be seen in Figures 2a and 2b below.

More studies, however, dealt with the driver's end. In a PC-based eye-tracking study of participants gazing at images of cyclists from a driver's point of view, it was observed that the 'drivers' first sought eye contact before moving to other parts of the body, lingered more on faces than other body parts, and maintained their gaze longer if the image of the cyclist appeared to be seeking eye contact with them (Walker & Brosnan, 2007). In a related study, it was found that eye contact with pictures of cyclists slowed the decision-making processes of participants posing as drivers, increasing their risk of being unable to stop their vehicle in time to avoid collision (Walker, 2005). In a different study investigating oncoming vehicle behavior, drivers were found to brake 5 m earlier, brake smoother, and stop further away from the crossing when they made eye contact with a staged pedestrian standing on the curb (compared to when they did not) (Ren et al., 2016). Eye contact also made drivers more likely to stop and accept pedestrians' request to hitchhike, as opposed to when pedestrians did not seek eye contact (Snyder et al., 1974). However, another study demonstrated that drivers slowed down and stopped more often when pedestrians on a curb were not looking in their direction as opposed to those who were actively seeking eye contact (Katz et al., 1975). It is possible that these contradicting results were due to differences in the motives of the pedestrians as perceived by the drivers. Eye contact with a person requesting a ride may have had a positive effect on driver compliance to stop the vehicle, due to greater perceived trustworthiness of the hitchhiker. On the other hand, the lack of eye contact or gaze in their direction from pedestrians on the curb probably heightened drivers' fear of collision; possibly in an effort to lessen the consequences in case the former crossed without looking, drivers slowed down and stopped more frequently. Summarizing, most papers seem to agree that eye contact is beneficial to both drivers and pedestrians.



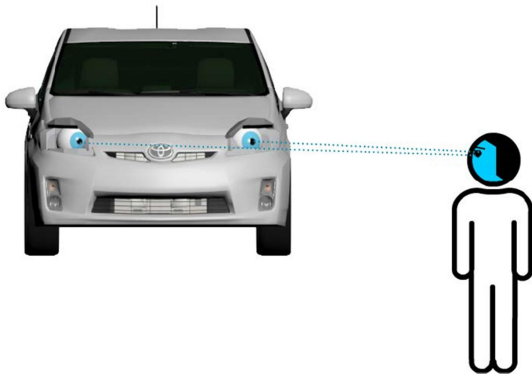


Fig. 2a – Eyes on a car (Chang et al., 2017)



Fig. 2b – Jaguar Eyes (Jaguar Land Rover, 2018)

Studies show that ‘eye contact’ with virtual and real vehicles has a similar effect on pedestrians as eye contact with a human driver.

As mentioned above, the literature leans towards the conclusion that eye contact plays a role in human interactions in traffic. However, in every one of the studies, either the driver’s or the pedestrian’s point of view during the crossing conflict was not considered, which led to uncertainty in whether eye contact occurred or not. Importantly, driver-pedestrian eye contact was not measured objectively, but only instructed to participants or assumed to occur. Reasons for this dearth in research likely include a recent interest among scientists for eye contact on the road only in the last decade, the unavailability of tools and technology for reliably detecting eye contact until the past few years, and recent rises in pedestrian fatalities prompting a need for mitigating measures. Thus, this vacuum in the literature presents the following research questions, which follow from the desire to define and detect eye contact:

**(a) What is eye contact between a pedestrian and a driver in a car? How can eye contact be defined/operationalized using an algorithm?**

No standard definition of eye contact exists in the literature on the subject of driver-pedestrian interaction, most probably because previous studies did not attempt to measure it objectively. Drivers and pedestrians looking in each other’s direction for arbitrarily defined durations of time such as 1 second or more have been assumed to include eye contact by some studies (Rasouli et al., 2017; Kotseruba et al., 2016). Further, metrics such as general head orientation and glance duration obtained from sources like direct observation, street camera recordings, and dash cams are inconclusive of the occurrence of eye contact due to limitations in their accuracy. An example is when a pedestrian is looking at a car to observe its motion, and this action is misconstrued as eye contact by an observer standing a few meters away from the intersection (Moore et al., 2019). Additionally, these techniques do not provide the perspectives of both parties on the crossing conflict.

Eye-tracking is a viable solution to this problem because it can pinpoint the driver’s and pedestrian’s gaze relatively well compared to the other methods mentioned above. Thus, it is suitable for the detection of eye contact by the synchronization of two eye trackers using an algorithm, which is the primary aim of this thesis. So far, there exists little to no research involving simultaneous tracking of drivers’ and pedestrians’ eyes during an interaction. This shortage is possibly because such applications require accurate and portable eye-trackers, and reliable head-mounted ones are a new technology, having only been around since 2010 (deVries, 2017). Perhaps the most closely related study is one that uses mutual situational awareness between drivers and pedestrians during a crossing conflict (Roth et al., 2016).

In this thesis, an experiment designed to resemble a driver-pedestrian interaction at a pedestrian crossing was performed with a randomized order of trials of ‘eye contact’ and ‘no eye contact’, which were then identified by an algorithm. A researcher acted as a driver in a stationary car, and participants acted as pedestrians standing/crossing in front of the vehicle. Both parties’ eyes were tracked using eye-trackers and their gaze rays were generated in 3-D. The two eye trackers in question are the Smart Eye Pro dx (a fixed, dashboard-mounted system installed in the car) to track the driver’s eyes, and a Tobii Pro Glasses 2 (a head-mounted, portable device worn by a person) to track the pedestrian’s eyes. This tracking and generating of gaze directions was dubbed ‘gaze ray tracing’. A second researcher facilitated the experiment and synchronized both eye-trackers by flashing a torch in view of both persons. Pedestrians were identified and their locations determined via a stereo camera. The head rotations of the pedestrian were also measured using a gyroscope in the Tobii Pro Glasses 2. The angle between the driver and pedestrian gaze rays was



compared against an 'eye contact threshold' to decide if eye contact occurred or not. A specific definition of driver-pedestrian eye contact in terms of time durations was also put forward using the measurements of eye contact from the experiments. By extension, the logic of this algorithm may be incorporated into automated vehicles for the detection of passenger-pedestrian eye contact in the future.

### **(b) How accurate is the algorithm that operationalizes eye contact?**

The performance of the algorithm would reveal if the identification of eye contact on the road using eye tracking is a viable option. One way to test the algorithm is to compare its output to established ground truths about eye contact (or its lack thereof). In the case of the experiments in this thesis, these ground truths are twofold:

- The instructions provided to participants before each trial (i.e. 'Make eye contact with the driver' or 'Do not make eye contact with the driver') along with the known order in which the trials were performed
- Eye contact determined by aggregating participants' gaze data falling within manually annotated areas of interest (AOIs) drawn around the driver's face, also called AOI hits

In order to determine driver-pedestrian eye contact in a trial, the following was needed:

- Pedestrian eye contact seeking via manually annotated AOIs (Ground truth)
- Pedestrian eye contact seeking gaze ray via the head-mounted Tobii eye-tracker (Undefined due to a lack of an earth frame of reference)
- Driver eye contact seeking via the gaze ray from the dashboard-mounted Smart Eye tracker (Ground truth)

The first and the third together (with some constraints) provided a ground truth of eye contact occurrence. The second and the third together provided the algorithm's detections of eye contact. The AOIs were drawn manually, but the gaze rays were to be determined automatically by the algorithm. A challenge was finding the gaze ray for the Tobii since it had a moving frame of reference, unlike the Smart Eye's frame of reference which was fixed. So, the former could not be measured and used directly like the latter. Therefore, the following were measured in addition in order to indirectly (and automatically) calculate the previously unknown gaze ray from the Tobii:

- Pedestrian's locations (via the stereo camera)
- Driver's head positions (via the Smart Eye)
- Pedestrian's head rotations (via the Tobii's gyroscope)
- Second researcher's locations (via the stereo camera, used for calibrating the gyroscope information)

The method by which pedestrian eye contact seeking was calculated by the algorithm is covered later in this report. Binary or square wave plots were used to compare the algorithm's detection capabilities with the ground truths. The 'eye contact threshold' was varied to find the best setting for the algorithm. True-positive and false-positive rates were calculated and the performance of the algorithm illustrated with the help of receiver operating characteristic (ROC) curves. The tolerance in the angular accuracy of the detections provides an idea of the algorithm's ability to distinguish between cases of eye contact and no eye contact.

### **(c) How is it possible to use two eye-trackers with inertial measurement units (IMUs) and pedestrian recognition in a Toyota Prius car to reconstruct the entire driver-pedestrian interaction through a 3-D animation?**

A reconstruction and visualization of eye contact in the crossing conflict would be useful to occupants of an automated vehicle as it is an indication of the vehicle's intentions. It also aids the understanding of the working of the algorithm and enables engineers to develop it using the visual feedback. Additionally, in the future, a reconstructed vehicle-pedestrian interaction (showing pedestrian attention in real-time) could be shared amongst a network of nearby, connected vehicles and lead to greater road safety. Thus, this thesis attempts to tackle the challenge of using the data obtained from the eye-tracking devices to recreate the interaction scenarios by tracing the gaze rays of the driver and the pedestrian. The generated animation shows: (1) The car, (2) Position of the driver inside the car, (3) Position of the pedestrian, (4) Gaze direction vector of the driver, and (5) Gaze direction vector of the pedestrian.



# 2 Methods

## 2.1 Experiment design and procedure

### I. Participants

Thirty-one participants (23 male, 8 female) took part in the experiment. Participants were recruited via social media, directly in person, and through referrals from existing participants. They were screened to exclude people who only wore spectacles; persons with perfect vision and those who wore corrective lenses were eligible to participate. If sighted persons possessed corrective lenses, they were requested to wear the latter on the days of the experiment in order to be able to participate. All the participants read and signed an informed consent form. The research was approved by the Human Research Ethics Committee of the Delft University of Technology. In Appendix A, a copy of the consent form can be found.

### II. Equipment and software

The experiment used a Tobii Pro Glasses 2 eye tracker, a Smart Eye Pro dx eye-tracker installed in a Toyota Prius intelligent vehicle, a stereo camera with integrated pedestrian detection also in the vehicle (Ferranti et al., 2019), and a NexTorch flashlight. The Tobii Pro Glasses 2 (also called the Tobii in this report) are portable, head-mounted glasses capable of recording videos from the participant's point of view, and tracking the gaze of the wearer's eyes at a sampling rate of 50 Hz (Tobii AB, 2015). The Tobii Pro Lab software was used to extract gaze data within defined times and areas, and to export the gaze data and metrics to Microsoft Excel, to enable subsequent analysis in MATLAB.

The Smart Eye Pro dx eye tracker (also called the Smart Eye in this report) consisted of a system of 4 cameras mounted on the dashboard of the car (as seen in Figure 3 below) that collected a researcher's (i.e. driver's) gaze data during the experiment at a frequency of 60 Hz (Smart Eye AB, 2018). The gaze data were exported as ROSBAG files in Linux Ubuntu and converted to the CSV format for analysis in MATLAB.





Fig. 3 – Smart Eye Pro dx with 4 cameras mounted on the dashboard

The cameras are positioned in such a way so that for any head pose, at least two of them can properly capture the driver's eyes, preferably from both the right side and the left.

The stereo camera was an iDS UI-3060CP-C-HQ Rev. 2 with two lenses. It was mounted in front of the rear-view mirror of the vehicle, facing forward. The single-shot detection technique was employed to detect pedestrians in the video and report their position and orientation with respect to the stereo camera at a sampling rate of 10 Hz (Liu et al., 2015). This data was included in the aforementioned ROSBAG file. The NexTorch flashlight was operated by a researcher as an instantaneous light source to synchronize the Tobii and Smart Eye trackers. These tools are pictured in Figures 4a and 4b below.



Fig. 4a – iDS UI-3060CP-C-HQ Rev. 2 stereo camera lens (iDS Imaging, 2014)



Fig. 4b – NexTorch high-intensity flashlight (NexTorch, 2011)

The torch was selected to be bright even in broad daylight and angled to be clearly visible to both the stereo camera and the Tobii's scene camera during the experiment.

The Tobii and the Smart Eye trackers both work on the principle of pupil center corneal reflection, that is, detection of reflected infrared light from the corneas of a user (Tobii AB, 2015; Smart Eye AB, 2012). They use infrared light sources to illuminate the person's eyes and cameras to capture images of the reflections of the illuminators. The artificial illumination is needed to minimize the effect of varying lighting conditions. From the obtained image, a vector is generated using the angle between the corneal reflections and the pupil. Then, using the direction of this vector, coupled with known geometrical features and distances in the reflections, the gaze direction of the person is calculated (as seen in Figures 5 and 6).



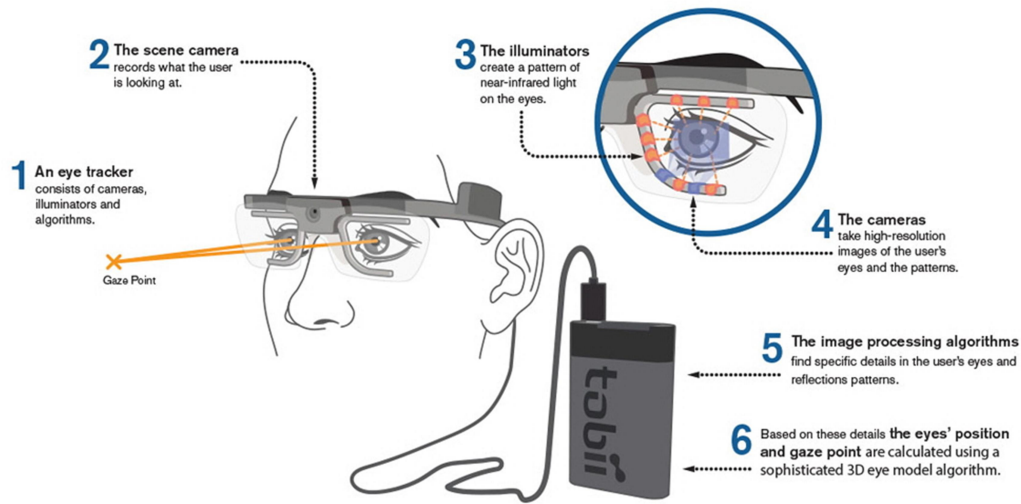


Fig. 5 – Working of the Tobii Pro Glasses 2 (Tobii AB, 2015)

The Tobii also provides a live feed of the wearer's view on a tethered device such as a laptop or a tablet computer.

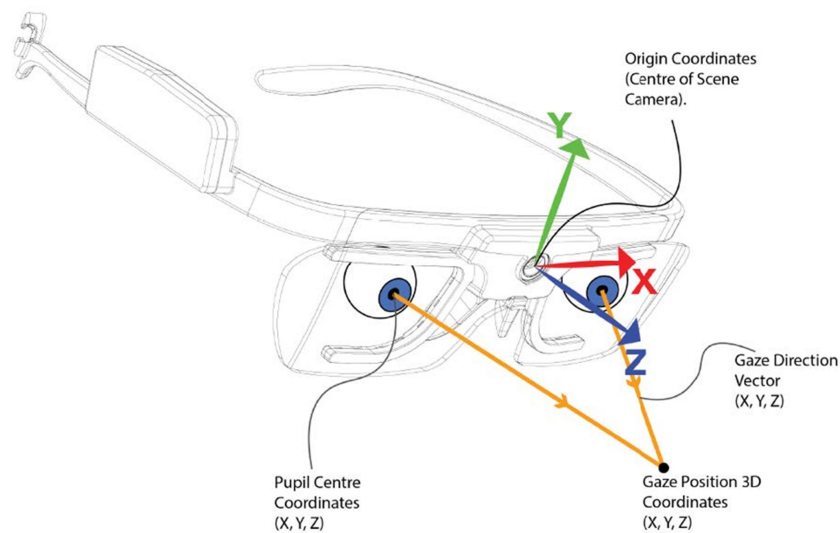


Fig. 6 – Tobii's collection of gaze data (Hryhoruk, 2019)

The coordinate system of the Tobii is shown here, the axes of which continuously move as the wearer's head moves.

### III. Pilot test

Initially, a pilot test was performed to determine if eye contact between drivers and pedestrians could be detected by synchronizing two eye trackers using a torch as a light source. No participants were recruited for the pilot study; instead, four researchers took part in it. One researcher acted as a driver seated in the stationary vehicle, two persons took turns being a pedestrian in front of the car or operating the torch, and one person operated the torch when the others were unavailable. Two combinations of eye-trackers were used, namely (1) Smart Eye (driver) with Tobii (pedestrian), and (2) Tobii (driver) with Tobii (pedestrian). The latter case was possible because two pairs of Tobii glasses were available. Both indoor and outdoor tests were conducted to assess if eye contact could be detected



using the eye-trackers. In the indoor trials, the 'pedestrian' stood at distances of 1 m, 2 m and 5 m from the front of the vehicle on an (imaginary) curb either to the left or to the right of the car. The location for the indoor trials of the pilot test was the garage on the ground floor of block F of the Cognitive Robotics department of the faculty of Mechanical, Maritime and Materials Engineering in the Delft University of Technology. In the outdoor trials, the 'pedestrian' additionally stood at distances of 10 m, 25 m and 50 m from the car. The outdoor trials in the pilot study were conducted on the public road 'Van der Maasweg' in Delft.

The pedestrian either waited at the curb and turned to make eye contact with the 'driver' in the vehicle or crossed the (imaginary) road while maintaining eye contact with the 'driver'. This interaction was repeated three times before changing to a different distance or to a different side. Each configuration of 'driver' and 'pedestrian' was defined as one trial or one scenario. Each scenario began and ended with both persons waiting for a flash from a torch controlled by another researcher.

The pilot test revealed that eye contact could indeed be detected indoors but was fragmented (as opposed to the desired continuous eye contact) due to limitations in the accuracy of the eye-tracking technology. Next, there was reason to expect that eye-tracking performance was dependent on eye color, due to differences of up to 25% between the gaze samples of the brown-eyed and blue-eyed researchers under similar testing conditions. Further, the Tobii eye-trackers were sensitive to background infrared light and so had very poor performance (i.e. single-digit gaze sample percentages) in sunny weather outdoors. This problem occurred despite fitting the Tobii with a tinted lens in an attempt to block IR radiation during the pilot study's outdoor trials. Additionally, it was observed that windshield glare interfered with the establishment of eye contact. Lastly, rapid head movements and dynamic conditions with moving users and moving targets deteriorated performance of the eye trackers.

Some valuable lessons were learned from the pilot test, which were incorporated into the main experiment. First, a greater number of participants with a greater variety of eye colors were required to ascertain whether eye contact could be reliably detected using synchronized eye trackers. Second, the main experiment would benefit from being conducted indoors, away from bright sunlight to avoid infrared effects and to prevent windscreen glare. This choice of such a setting would make the experiment's environment more controllable. The pitfall of the 'streetlight effect/bias' would be avoided because it was reasonable to assume that the experiment would retain its fundamental principles indoors, while minimizing extraneous light sources that made data collection harder. The other alternative was to schedule the experiment on overcast days with no rain (since the Tobii was not waterproof) while avoiding suboptimal lighting conditions, which proved to be infeasible. Third, randomization of the order of performed trials in the experiment would be required to counter learning effects. Fourth, trials where driver and pedestrian were not making eye contact would also be needed, to act as a reference to distinguish from eye contact trials. Fifth, trials had to involve a stationary vehicle to minimize dynamic conditions and due to safety and space constraints indoors. Finally, questionnaires would need to be administered before and after the experiment to obtain participant details and measure task involvement.

#### **IV. Design of the main experiment**

The main experiment was conducted indoors in the garage of the faculty of Mechanical, Maritime and Materials Engineering at the Delft University of Technology. The area was cordoned off to prevent interference from passersby and keep distractions out of view. The car was parked away from any windows and angled in such a way as to minimize the chance of onlookers and passersby being detected by the stereo camera and pedestrian detection system. At any given time, two researchers conducted the experiment; one person sat in the car as the driver and four persons played the part of the second researcher in shifts to operate the torch.



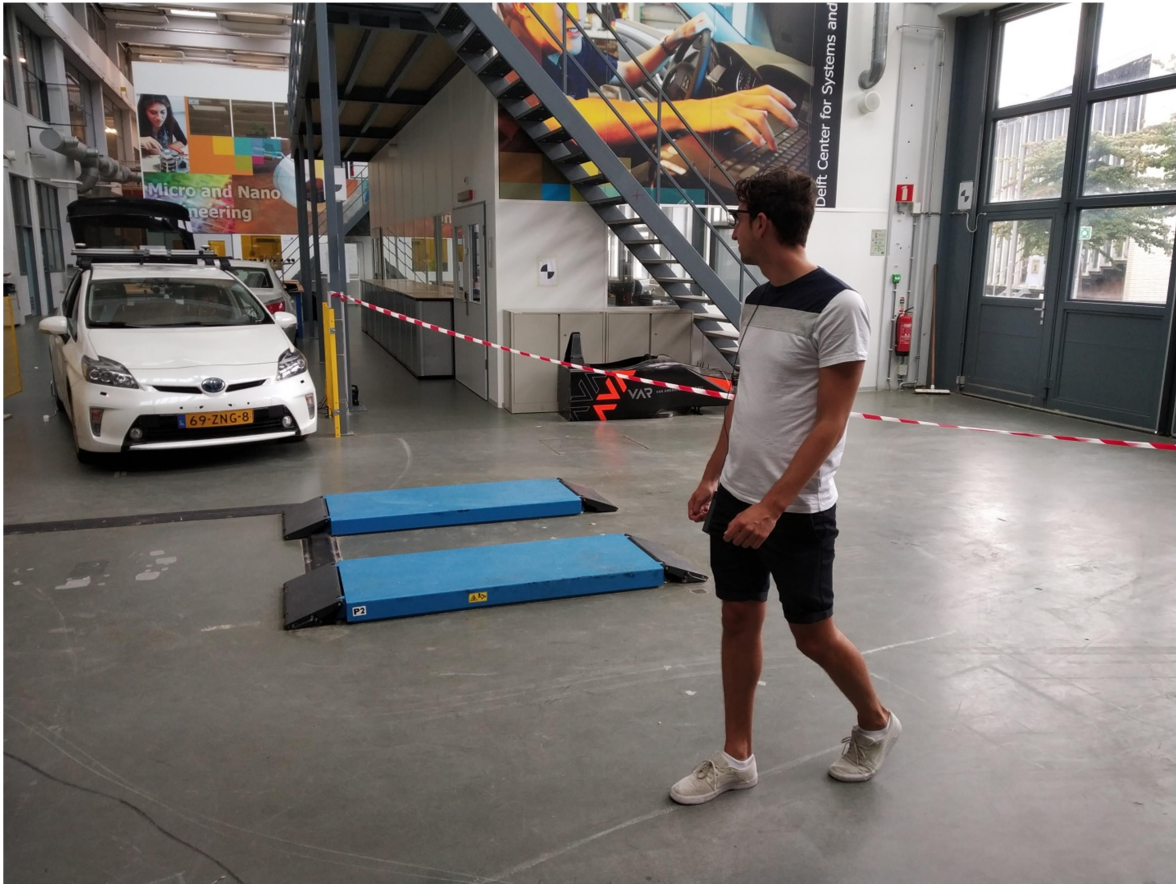


Fig. 7 – Pedestrian's role in the experiment

At this moment, the pedestrian is crossing from the left side of the vehicle at a longitudinal distance of 5 m from its front while maintaining constant eye contact with the driver.



Fig. 8 – Pedestrian's perspective obtained via Tobii eye-tracking



At this moment, the pedestrian is standing on the right side of the vehicle at a longitudinal distance of 5 m from its front and seeking eye contact with the driver (who is barely seen in the Tobii's recording but is visible to the pedestrian's eyes).

As the participants arrived, they were given an informed consent form to read and sign, which also contained a brief overview of the experiment. Next, participants filled out the pre-experiment questionnaire digitally. This questionnaire noted basic data such as height, sex, eye color, and eye correction, among others and may be found in Appendix B. The experiment, including the participant's task, was then orally explained in detail by a researcher.

Participants were told to imagine that they were at a real pedestrian crossing and that the car was approaching them while slowing down to a stop. They were informed that they would perform six types of trials while wearing the Tobii glasses (Table 1). They were also told that the trials involved standing/crossing in front of the stationary car while making/not making eye contact with the driver, and that each trial consisted of three repetitions, so that meant a total of 18 repetitions. Additionally, as a brief demonstration, a researcher assumed the role of the pedestrian and performed one standing and one crossing trial to give the participant an idea of what he/she had to do. Participants were aware that their eyes were being tracked and that their anonymous gaze data would be used to create an algorithm to detect driver-pedestrian eye contact. This information was present in both the consent form and repeated by the researcher orally. Participants did not need to remember the order in which the trials would be performed since a researcher would always tell them what type of trial was next. The participants only needed to remember what the current trial was and to perform it three times in a row before moving on to the next trial. An idea of participants' role and view in the experiment may be obtained by taking a look at Figures 7 and 8 respectively.

Table 1 – Experiment scenarios or trials

Eye-tracker combination (Pedestrian-Driver)	Scenario	Distance	Eye contact	Repetitions	Abbreviation
Tobii-Smart Eye	Right standing pedestrian	5 m	Yes	3	R-S-EC
Tobii-Smart Eye	Right standing pedestrian	5 m	No	3	R-S-NEC
Tobii-Smart Eye	Left standing pedestrian	5 m	Yes	3	L-S-EC
Tobii-Smart Eye	Left standing pedestrian	5 m	No	3	L-S-NEC
Tobii-Smart Eye	Crossing pedestrian	5 m	Yes	3	W-EC
Tobii-Smart Eye	Crossing pedestrian	5 m	No	3	W-NEC

After the demo and verbal explanation, participants were made to wear the Tobii Pro Glasses 2 and their gaze was calibrated before any readings were taken. In the meantime, the driver calibrated his own gaze in the car with the Smart Eye. While wearing the Tobii, the participants performed the six types of trials (each with three back-to-back repetitions) in a random order which included:

- Standing at a longitudinal distance of 5 m (from the front of the vehicle) and a lateral distance of 0.5 m (from the side of the vehicle) on an imaginary curb on the left/right of the stationary Prius while either making/not making eye contact with a researcher posing as the driver.
- Crossing back and forth (starting on the right-side curb from the driver's perspective) on an imaginary road in front of the stationary Prius at a longitudinal distance of 5 m (from the front of the vehicle) while either making/not making eye contact with the driver for the entire duration of the crossing.

The standing trials were intended to mimic a pedestrian waiting to cross at the curb. Participants were told to make eye contact with the driver for the same duration as they would when deciding to cross a road (typically, a few seconds). In the event of a trial involving no eye contact, participants were instructed to look at the vehicle for as long as they would when deciding to cross a road but to refrain from looking at the driver (also typically, a few seconds).

The crossing trials attempted to mimic a pedestrian in the act of crossing a street. During the crossing trials, participants were told to continuously make eye contact as they crossed (back and forth), or to avoid eye contact by looking somewhere else on the body of the car, depending on the type of trial. This instruction would ensure the occurrence of continuous eye contact which may not have happened otherwise, and which was needed to test the algorithm's detection capabilities in dynamic situations. In the crossing trials, the pedestrians always started from the right side of the car and walked to the left side and back. This choice was made for the sake of convenient analysis so that patterns of the ground truth data of different participants exported from the Tobii Pro Lab software were uniform.



Six different scenarios or trials were carried out by every participant, each with three repetitions. The standing scenarios were each their own separate trial, and the trials were carried out in random order. The crossing scenarios combined walking from the right and the left to allow convenient execution of each trial by participants, and these trials too were performed in random order. All trials began and ended with the researcher (driver) and participant (pedestrian) looking at a light signal from the torch operated by the second researcher. The switching on of the flashlight at the start and end were instantaneous events captured by both the Tobii camera and car's stereo camera, which allowed eye tracker synchronization in the post-processing stage. The first flash from the torch was considered the start of a trial and the second flash marked its end. In between these two flashes were the three repetitions of the driver-pedestrian interaction.

There was no exact demarcation between repetitions. One repetition in the standing trials was roughly turning to look at the other party, making/not making eye contact, and turning to look away. One repetition in the crossing trials was turning to look at the other person, walking back and forth once while making/not making eye contact, and turning to look away. The reason for having three repetitions was to average the gaze behavior of the pedestrian within each trial. The trials would be used in the analyses, and the values of metrics obtained from the trials would be divided by three to get their equivalent for one driver-pedestrian interaction. Figure 9 provides a photo of a trial in progress, for a clearer understanding of the experiment.

The driver's role in the experiment was similar to the pedestrian's, but with a small difference. During all the standing trials (including the ones where pedestrians did not seek eye contact), the driver sought eye contact with the pedestrian 3 times. During all the crossing trials (including the ones where pedestrians did not seek eye contact), the driver sought eye contact constantly as the pedestrian walked back and forth 3 times. An idea of the driver's role and view in the experiment may be obtained by viewing Figure 10 and 11 respectively. This choice regarding the driver's behavior during the so-called 'no eye contact' trials was made specifically for three reasons:

- To minimize the effect of the driver's (i.e. researcher's) gaze behavior on the outcome of the eye contact detection algorithm.
- To determine if just one party's (in this case, the pedestrian's) gaze behavior could influence whether eye contact was detected or not.
- To imitate (as best as possible) the gaze behavior of a real, responsible driver yielding for a pedestrian at a crossing.

Two recordings using the Tobii and Smart Eye were made per participant, one containing all four standing trials and the other containing both crossing trials. Both Tobii and Smart Eye trackers were recalibrated after every recording to counter any drift in the gaze data. After all the trials were completed in a random order, the participants filled out the post-experiment questionnaire digitally (Appendix C). This questionnaire measured responses to questions on concentration, task involvement, perceived task performance, clarity of instructions and feeling of immersion, among others, on a 7-point Likert scale (Kooijman, 2018). The randomization matrices for the scenarios may be found in Appendix D.





Fig. 9 – Moments after the start of a trial

As may be seen, eye-tracker synchronization using the torch at the start is over (performed by the researcher on the right) and the pedestrian (second from right) is standing on the left side of the vehicle and making eye contact with the driver (not seen in image).





Fig. 10 – Driver's role in the experiment

At this moment, the driver's eyes are following the pedestrian who is crossing from the left side of the vehicle.



Fig. 11 – Driver's perspective and Smart Eye tracking

At this moment, the driver's eyes are following the pedestrian (person on the right) who is crossing from the right side of the vehicle, and the torch operating researcher (person on the left) looks on.



## V. Equipment setup

The Smart Eye tracker and stereo camera in the vehicle were setup using the Smart Eye Pro software and Robot Operating System (ROS) packages on Ubuntu Linux. For each Smart Eye recording, the following steps were performed:

- The focuses of the four Smart Eye cameras on the dashboard were adjusted (if required) so that the driver's eyes were sharply visible in all cameras.
- A board printed with black and white squares, dubbed the 'checkerboard', was used to calibrate the Smart Eye tracker. This calibration was achieved by rotating the board in view of all four cameras until a sufficiently low margin of error in the detection of the position of the squares was obtained (see Figure 12). The acceptable error values were 0.3 pixels and below (on a scale starting from 0 and with no upper limit). The numbers were the average difference in the coordinates of the checkerboard's corner points.
- The calibration was verified to ensure that the desired level of accuracy was retained even on removing the checkerboard from view and positioning and rotating it differently.
- The positions of several known points inside the car in view of the four cameras were adjusted in the software (if required). These known points were already defined before the start of the experiment on the roof and pillars of the vehicle frame using beads. The distances to these points relative to the cameras were already known to a high level of accuracy, and the software used them to determine the positions and angles of the driver's head and eyes.
- Eye-tracking was enabled via the software after checking if the driver's eyes were detected correctly on the in-car display.
- The driver's gaze was calibrated using a different set of known points on the dashboard and side-view mirrors by having him look at these points and comparing the readings to the reference values.
- The processing, tracking and recording of the driver's gaze were initiated by running a pre-written C++ program in ROS before the start of every recording. These were low-level programs that prompted the Smart Eye to start collecting gaze data using its own algorithms. The gaze data was automatically bundled with the stereo camera video and pedestrian detection data and saved in a ROSBAG file.
- The standing/crossing trials were carried out and data collection was terminated after the end of the recording via ROS commands.

The C++ program used to track the driver's eyes via the Smart Eye was developed by members of the Department of Cognitive Robotics at the Delft University of Technology and is not publicly accessible. It was obtained for the purposes of the experiment via the courtesy of J. C. J. Stapel.





Fig. 12 – Calibration of the Smart Eye using the ‘checkerboard’

The ‘checkerboard’ was rotated about all three axes for a few seconds until the error value in the in-car display stabilized.

To set up the Tobii eye trackers, the glasses were first connected to a tablet computer via the wireless network of the glasses’ recording unit. Using the Tobii Pro Controller software, a study was created and participants added, identifiable only by their participation number in the experiment. Before each recording, the following was done:

- Once the participant wore the Tobii Pro Glasses 2, his/her eyes were calibrated using a calibration card with a printed bull’s-eye. The participants were asked to focus on the central black dot for a few seconds until the eye tracker calibrated their gaze.
- A recording was started via the software, the standing/crossing trials conducted, and the recording stopped.

The researcher posing as the driver controlled the Smart Eye recordings and the researcher operating the torch controlled the Tobii recordings. Since each participant provided two recordings on each device (i.e. one recording containing the four standing trials and another recording containing the two crossing trials), the Smart Eye and Tobii setup were done twice per participant. In total, four recordings (two of the pedestrian, two of the driver) per participant were collected.

## VI. Analysis

### i. Gaze quality check

The collected Tobii recordings contained videos of the experiment from the pedestrian’s point of view with his or her eye gaze overlaid on them. In Tobii Pro Lab, the gaze sample percentages for all recordings were checked to see if they were all above the minimum threshold of 60%, which they were. These values represented the percentage of the recording duration when the person’s gaze direction and gaze point could be found by the eye-tracker. The threshold was defined higher than the 45% value found in the literature (Hopmans et al., 2018), since the trials were performed indoors, away from direct sunlight, and not highly dynamic, so higher percentages were expected.



While it appeared that with this threshold, even a 40% loss of data was acceptable, this was actually not the case. Gaps in the gaze measurements usually occurred in between trials during the scenario changeover times, so 40% missing data in a recording did not necessarily mean a 40% loss of data during experimental trials. The scenario changeover times were the times between the end of one trial and the start of the next trial in a recording. These periods included the time it took the two researchers to look up separately which trial was next, confirm this information with each other, communicate this information to the participant, and the time it took the participant to move to a new location (if required, and as per the needs of the upcoming trial). Participants were free to look wherever they wanted during this time. More often than not, they gazed straight ahead at the transparent, sunlit doors/windows of the garage which led to a loss of gaze data due to interference from infrared light. Missing data was treated as a lack of eye contact between driver and pedestrian, as will become clear later on.

It was necessary to check how well the participants' eyes could be tracked by the Tobii at both the large scale (recordings) and the small scale (trials) since only together could a complete picture of the usability of the data be obtained. For example, high percentages in the recordings did not necessarily mean high percentages in the useful gaze data from the experiment (trials), and high percentages in the trials but low percentages in the recordings cast doubt on the reliability of the Tobii readings.

The Smart Eye was expected to be more robust in its ability to detect the driver's gaze (than the Tobii the pedestrian's) because of the use of four cameras to capture the driver's eyes and the shade from infrared light inside the vehicle. As such, there was no direct way to check the Smart Eye gaze quality, but the Smart Eye software showed uninterrupted detection of driver gaze direction on the in-car display during test runs conducted before the experiment. Further, the contents of the ROSBAG files for all recordings were inspected and showed almost no missing gaze data. This news also meant that the stereo camera detections of the pedestrian's location were also largely uninterrupted, since they were coupled with the Smart Eye data.

## ii. Trial 'start' and 'end' events and times of interest (TOIs)

Events at specific instants were added to the recordings by using the annotation functionality of the Tobii software. The instant the torch light was visible before a trial was marked as its 'start' event and the corresponding torch flash instant after the trial was marked as its 'end' event. The trial, consisting of three repetitions of a driver-pedestrian interaction, represented the period between a pair of torch flashes. Once all the events were annotated, times of interest (TOIs) were created for the duration of the trials (i.e. between relevant pairs of 'start' and 'end' events). A screenshot of Tobii Pro Lab can be found in Figure 13a, showing the TOI defining process.





Fig. 13a – Annotating events and times of interest (TOIs) in Tobii Pro Lab

The red circle enclosed in a white ring represents the participant's gaze point at the instant the recording is paused, and the coloured inverted triangles denote events, between pairs of which the custom TOIs are defined.

### iii. Dynamic areas of interest (AOIs)

For every recording, areas of interest (AOIs) were manually drawn around the driver's face at alternate frames (or in time steps of 40 ms) inside the parts of the TOIs when the driver was in the pedestrian's view. The constantly moving nature of the videos made the AOIs dynamic, and presented a challenge in keeping them locked on to the driver's face. Figure 13b shows the AOI drawing process. These areas of interest were all circles of diameter 150 pixels on a 1280 x 720 pixels recording. The size of the AOIs was decided in such a way that they covered the apparent size of the driver's head (with sufficient margin) at a maximum distance of 8 m from the pedestrian while keeping other enclosed objects to a minimum. Specifically, the 150 pixels diameter of the AOI translated to a real-world circular area which had a diameter of 0.91 m. This circle had approximately thrice the diameter of the driver's (i.e. researcher's) head, which was measured to be a sphere with a diameter of 0.3 m. This pixel-to-distance conversion scale was calculated by dividing the known width of the car in meters by its width in pixels in a Tobii Pro Glasses 2 recording. The pixel width of the car was measured at the height of the windshield when a pedestrian was directly in front of the vehicle on its longitudinal centerline. The obtained AOI size was deemed sufficient to account for any angular accuracy errors in the gaze direction measurements of the Tobii, whose accuracy tolerance was  $\pm 0.5^\circ$ . The calculations of these values are shown in Equation 1 below. The AOI subtended an angle of  $6.5^\circ$  from the pedestrian's point of view, as also seen further down in Figure 15. Figure 16 shows the layout of the experiment with the measurements used in the calculation of the maximum pedestrian distance from the driver.

$$\begin{aligned}
 l_{car} &= 4.48m \\
 b_{car} &= 1.7m \\
 h_{difference} &= 1m(assumed) \\
 \text{Maximum driver-pedestrian distance} &= \\
 r &= \text{ceil}(\sqrt{(5 + l_{car} / 2)^2 + (b_{car} / 2)^2 + h_{difference}^2}) = \text{ceil}(\sqrt{7.24^2 + 0.85^2 + 1^2}) = 8m \\
 \text{Pixel-to-distance ratio} = R &= \frac{b_{car}}{b_{car\_pixels}} = \frac{1.7}{280} = 0.00607m / \text{pixel} \\
 \text{AOI real-world diameter} = d_{aoi} &= R * d_{aoi\_pixels} = 0.00607 * 150 = 0.91m \\
 \text{Region of eye contact around driver's face} = \theta &= \frac{\text{arc}}{\text{radius}} = \frac{d_{aoi}}{r} = \frac{0.91}{8} = 6.5^\circ
 \end{aligned}$$

Eq. 1 – Calculation of the region of eye contact surrounding the driver's face

The equation is an approximation for small angles wherein an arc is assumed to be a straight line, and the letter 'l' denotes the length, 'b' is the breadth, 'h' is the height, and 'd' is the diameter.

The purpose of these AOIs was to collect and aggregate any gaze data that fell inside their bounds. Specifically, the number of video frames with gaze data landing within the AOIs was counted, and their total duration calculated. Since the AOIs were drawn around the driver's face, the captured gaze data was the 'eye contact seeking' behavior of the pedestrian. The underlying assumption was that if the pedestrian was looking at the driver's face, he or she was looking at the driver's eyes to establish eye contact. In other words, if a pedestrian's gaze falls inside the AOI, they are looking at the driver's face, and by extension, at the driver's eyes in an attempt to make eye contact. This belief is supported by the findings of some studies which show that people fixate more on the eye regions than other parts of a



face during facial perception (Perlman et al., 2009; Walker-Smith et al., 1977). Further, it is assumed that people look at faces only through the center of foveal vision (i.e. gaze center) and not via peripheral vision.

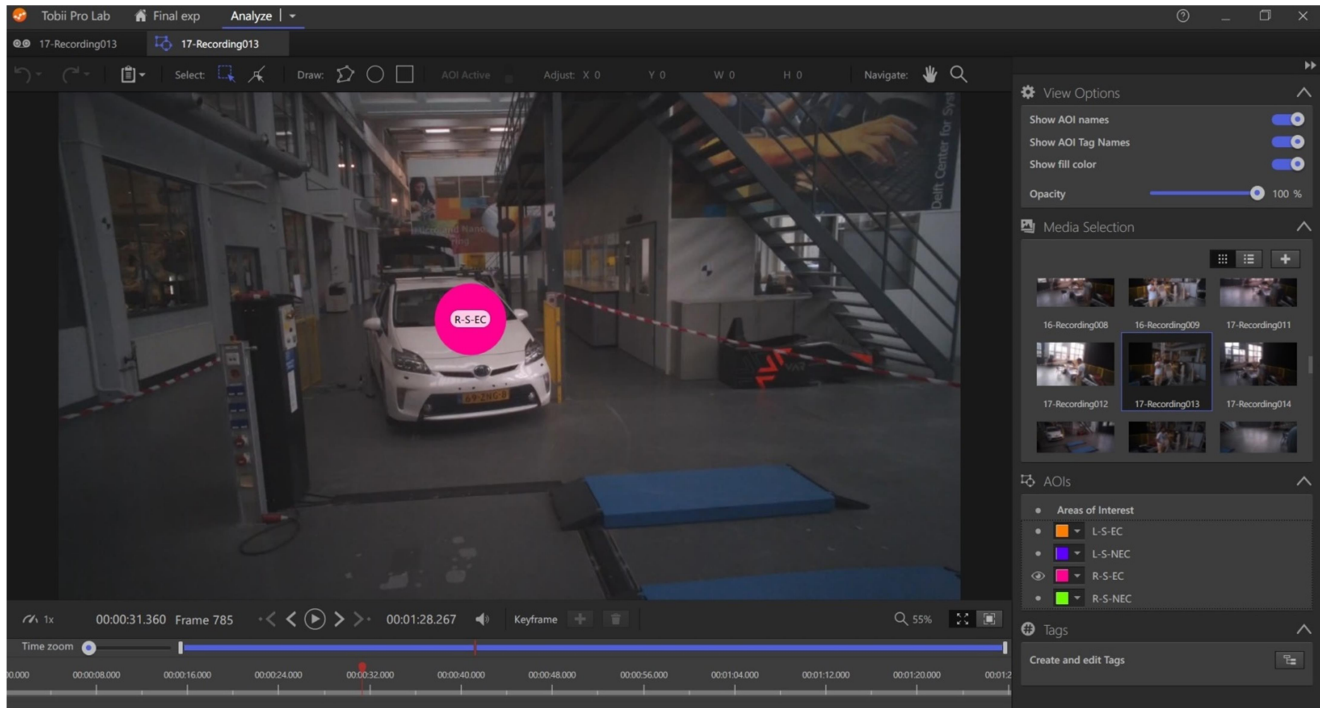


Fig. 13b – Drawing areas of interest (AOIs) in Tobii Pro Lab

The pink circle shows an AOI of 150 pixels diameter drawn around the driver's head to collect the pedestrian's eye contact seeking gaze data for the 'right side standing pedestrian with eye contact' scenario (R-S-EC, as seen on the label), and the coloured AOIs for other trials may be seen in the bottom-right of the image.

#### iv. Filtering and data export

As the raw gaze data proved to be noisy, an in-built filter in the Tobii Pro Lab software called the Tobii IV-T Fixation Filter was used after drawing the AOIs. As part of its preprocessing stage, the moving median technique of noise reduction was selected because it was less sensitive than the moving average technique to noise spikes of short durations typically found in raw gaze data.

After noise reduction, the Fixation Filter also had another function. This filter was a velocity-based filter that separated fixations from saccades using a defined velocity threshold (Olsen, 2012). In other words, if eye movement speed at a frame exceeded the velocity threshold of  $30^\circ/\text{s}$ , the frame was classified as part of a saccade; otherwise, it was classified as part of a fixation. If at least 3 consecutive frames had a lower velocity than the threshold, the frames were all labeled as part of the same fixation. A group of closely located fixations was classified as a visit. The filter settings are shown in Table 2 below.

This filter was ideal for static and mildly dynamic conditions but performed poorly in highly dynamic situations. Since only the pedestrians would be moving during the experiment and at low speeds, the IV-T Fixation Filter was chosen over its dynamic counterpart, the IV-T Attention Filter, after considering the manufacturer's recommendation (Tobii AB, 2018). The tradeoff was shorter periods of continuous eye contact but with greater gaze point accuracy. The filtering process did not alter the Tobii sampling rate (i.e. 50 Hz) but smoothed noise spikes, filled in small gaps, and identified if each frame was part of a saccade or a fixation.

Table 2 – Tobii IV-T Fixation Filter settings

Parameter	Value
Filtered variable	Eye movement angular velocity



Type	Low-pass
Cut-off velocity	30°/s
Velocity calculation window length	20 ms
Minimum fixation duration accepted	60 ms
Noise filtering technique	Moving median
Maximum gap interpolation	75 ms

Once the filter was applied in Tobii Pro Lab, the gaze data for all recordings were exported as Microsoft Excel files which contained pedestrians' gaze directions, eye contact AOI hits and gyroscope readings for head movements, among others. Additionally, the calculation of metrics such as total duration of eye contact seeking, mean duration of eye contact seeking visits (i.e. dwell time) were in-built in Tobii Pro Lab, and were also exported.

The gaze data from the Smart Eye was automatically filtered by its software, so there was no need to manually apply filters. The Smart Eye's filtered readings from the ROSBAG files were directly used in the analyses. The stereo camera's readings of pedestrian X, Y, Z coordinates (relative to itself) were also filtered automatically in ROS (using C++ code and a Kalman filter) before export and therefore used as is. In case of multiple or overlapping detections of the same pedestrian at some instants of the recordings, the detection with the highest confidence value was automatically chosen as the correct one.

## 2.2 Gaze ray tracing/eye contact detection algorithm

After verifying that the Tobii data had sufficient gaze samples to be usable, a MATLAB script was written to create an animated 3-D plot of eye contact (or lack thereof) between the driver and a pedestrian using their 3-D gaze direction vectors. This script may be found in Appendix G.

The basic principle (pseudo-code) to algorithmically detect eye contact, report detection performance, and reconstruct eye contact was as follows:

- Get driver (Smart Eye) and pedestrian (Tobii) gaze direction vectors, and pedestrian locations (stereo camera)
- Convert pedestrian locations into Prius IMU coordinate system
- Apply corrections for offsets in:
  - Stereo camera detections of pedestrian position
  - Tobii pedestrian gaze direction measurements
  - Smart Eye readings of driver gaze direction
- Compute angle between vectors
- Repeat for each instant of trial/scenario
- Check at which instants  $\text{angle} \leq \text{'eye contact threshold'}$
- Classify satisfying instants (frames) as 'eye contact'
- Classify non-satisfying instants/frames as 'no eye contact'
- Check if trial has high percentage of 'eye contact' instants (compared to other trials)
- If yes, classify trial as 'eye contact trial'
- If no, classify trial as 'no eye contact trial'
- Compare algorithm detections with ground truths (AOIs, experiment instructions)
- Report algorithm performance
- Reconstruct trial via animation
- Repeat for all trials/scenarios

A more detailed understanding of the working principle of the algorithm may be obtained by looking at the flow chart in Figure 14 below.



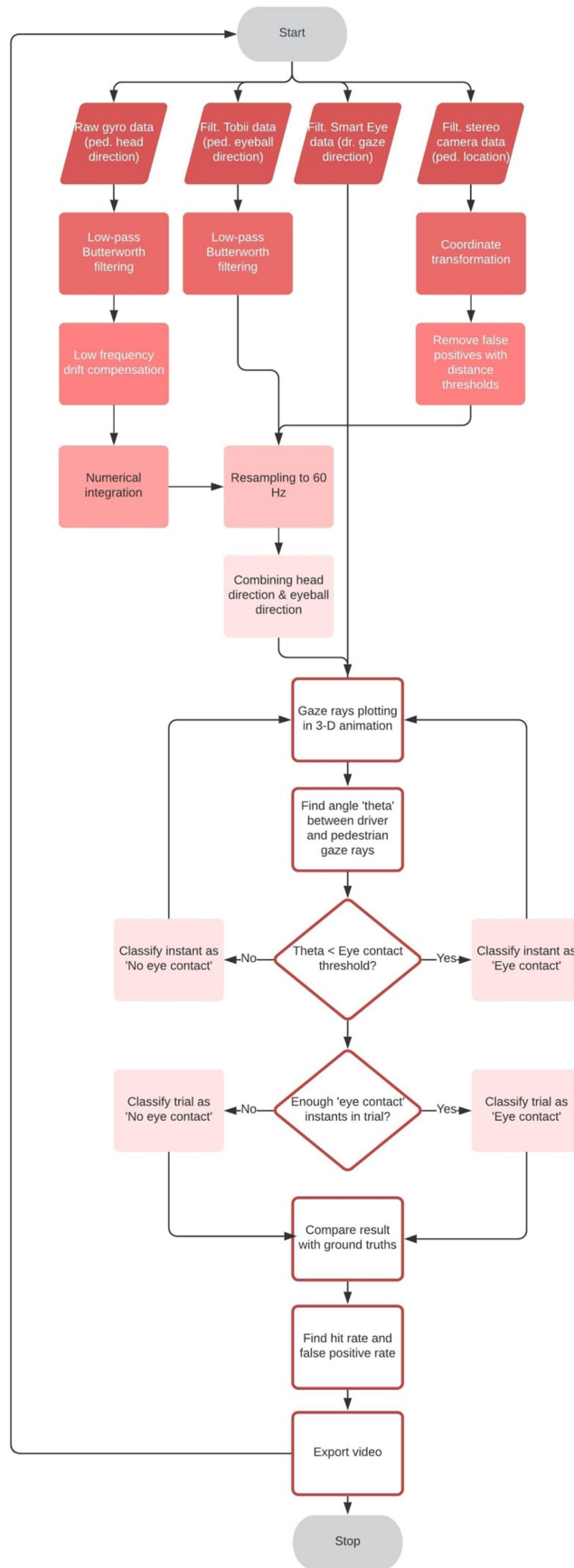


Fig. 14 – Flowchart of eye contact detection algorithm



First, theoretical calculations of the 'eye contact threshold' were made to find out the expected range of eye contact angles. To this end, the accuracy tolerances of the constituent measurements required to calculate the angle between gaze rays were compounded. Equation 2 shows how the uncertainty propagation formula (for summed measurements) was used for this purpose.

$$tol_{smarteye} = tol_{tobii} = tol_{stereo} = \pm 0.5^\circ$$

$$\text{Angular region of eye contact} = 6.5^\circ$$

$$tol_{aoi\_driver\_face} = tol_{pedestrian\_face} = \frac{6.5}{2} = \pm 3.25^\circ$$

$$\text{Theoretical 'eye contact threshold'} = \sqrt{3.25^2 + 3.25^2 + 0.5^2 + 0.5^2 + 0.5^2} = 4.68^\circ$$

Eq. 2 – Total uncertainty is the theoretical 'eye contact threshold' for the algorithm

The abbreviation 'tol' represents angular accuracy tolerance in a device or in the region of eye contact around a person's face.

The region of eye contact around the driver's face was up to  $6.5^\circ$  as found above using the Tobii AOIs. While no AOIs were drawn around the pedestrian's face, a region of eye contact of the same size was considered.

No 'trial threshold' or minimum percentage of eye contact instants in a trial for it to be classified as an 'eye contact trial' was set, because of the wide variability across participants and trials. Instead, trial eye contact percentages were compared with each other to determine which scenarios were 'eye contact' (higher percentages) and which were 'no eye contact' (lower percentages). In general, however, low values were observed because all trials contained periods of gazing away from the other party (in between the three repetitions) and because of inaccuracies in the gaze ray method of detecting eye contact, which are explained in a later section.

Second, theoretical calculations of a 'ground truth threshold' were made to obtain ballpark estimates of the angles required in the ground truth of eye contact. The reason a 'ground truth threshold' was needed because the Smart Eye software did not have a feature to generate hits inside moving AOIs (unlike the Tobii), and so driver eye contact seeking would need to be determined by other means. Moreover, there was no need for manual AOI annotation with the Smart Eye since it had a fixed frame of reference (unlike the Tobii).

Specifically, the angle between the line connecting the driver's gaze origin and the pedestrian's gaze origin (this line was dubbed the 'ideal driver gaze ray') and the driver's gaze direction vector (dubbed the 'actual driver gaze ray') was used to determine if the driver sought eye contact or not. On the driver's end, his gaze origin was the center point between his eyes (measured by the Smart Eye), and for the pedestrian, their gaze origin was assumed to be at their location (X, Y – as per the stereo camera) and at a height (Z) equal to 10 cm less than their height. The reason for choosing this Z-value was that the center point between the eyes of the pedestrian was assumed to be located 10 cm below the top of their head.

If the angle at an instant of the trial was less than or equal to the 'ground truth threshold', the driver was looking in the pedestrian's eyes, and if it was greater, he was not. This data was later combined with the AOI hits from the pedestrian's Tobii to obtain a ground truth of eye contact. Equation 3 shows how this 'ground truth threshold' was estimated.

$$tol_{smarteye} = tol_{stereo} = \pm 0.5^\circ$$

$$\text{Angular region of eye contact} = 6.5^\circ$$

$$tol_{pedestrian\_face} = \frac{6.5}{2} = \pm 3.25^\circ$$

$$\text{Theoretical 'ground truth threshold'} = \sqrt{3.25^2 + 0.5^2 + 0.5^2} = 3.33^\circ$$



Eq. 3 – Total uncertainty (without the Tobii terms) is the theoretical ‘ground truth threshold’ for the Smart Eye in the ground truth  
As before, the angular accuracy tolerances are denoted as ‘tol’.

Here, the uncertainty propagation formula was applied after removing the Tobii tolerance ( $tol_{tobii}$ ) and driver face tolerance ( $tol_{aoi\_driver\_face}$ ) terms. The use of AOIs to get pedestrian’s eye contact seeking (instead of using a gaze ray) meant that the Tobii terms no longer factored into the equation. Only the tolerances of the Smart Eye, the stereo camera and the region of eye contact around the pedestrian’s face were compounded.

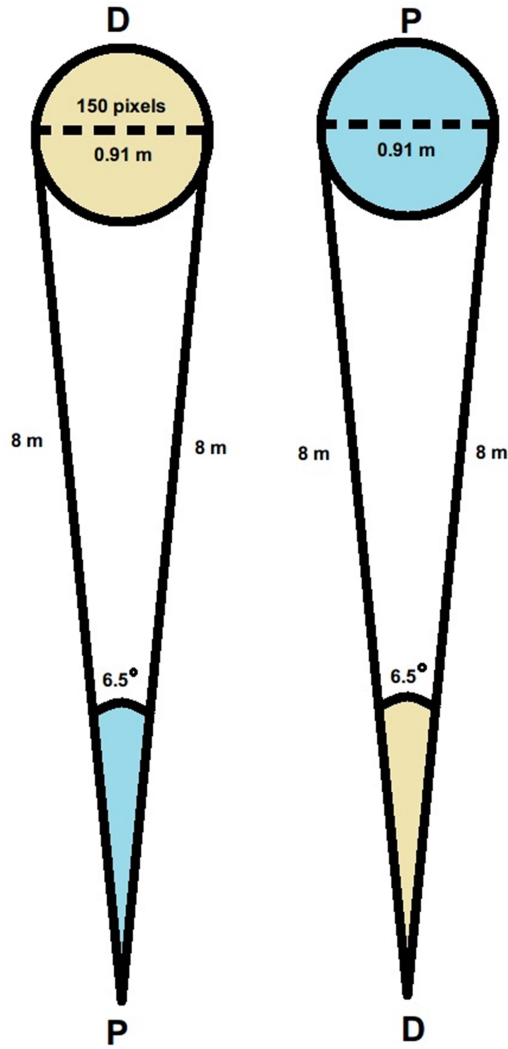


Fig. 15 – Regions of eye contact (Not to scale)

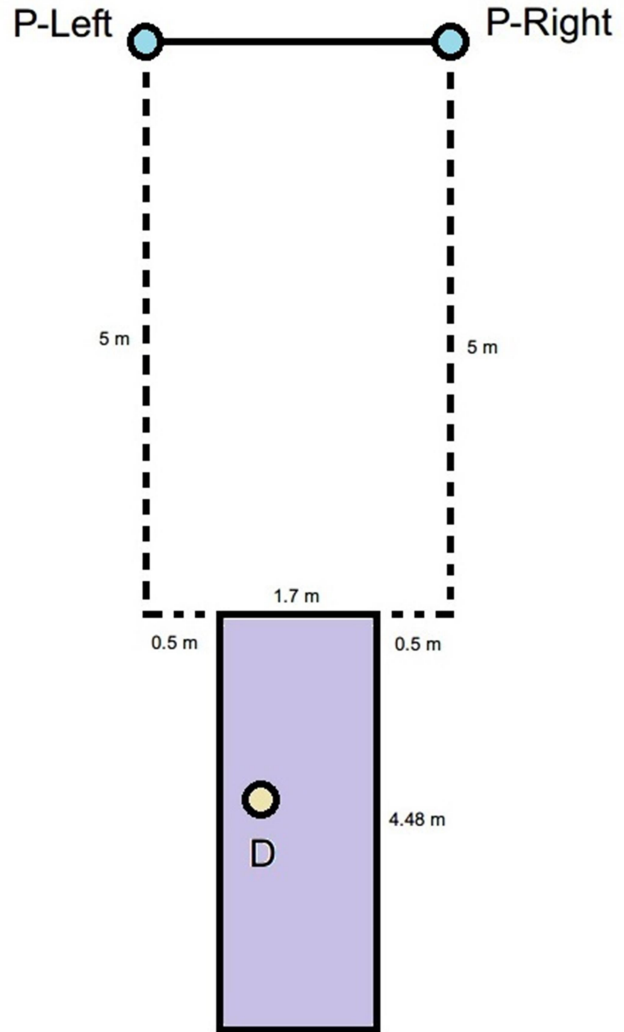


Fig. 16 – Experiment layout (Not to scale)

‘P’ denotes the location of the pedestrian, ‘D’ marks the location of the driver, and the maximum distance between ‘D’ and ‘P’ is worked out geometrically and rounded-up to the nearest integer. Here, the colour blue is used for the pedestrian, and the colour light brown for the driver. The lilac box denotes the car.

Using MATLAB code, the algorithm was implemented as follows:

#### i. Importing raw data from the sensors:

The required data were recorded at different sampling rates and with respect to different coordinate systems, owing to the different manufacturing specifications of the individual devices. The Smart Eye tracker collected driver gaze directions (and driver gaze origins) at 60 Hz, with an origin at the inertial measurement unit (IMU) of the car. The Tobii



Pro Glasses 2 collected pedestrian gaze directions at 50 Hz with a dynamic origin located at the scene camera of the glasses, positioned between the wearer's eyes. In other words, the origin of Tobii glasses moved as the wearer's head moved, making any readings of gaze direction relative and not absolute. The gyroscope installed in the Tobii glasses, however, measured at 95 Hz. The object detection capabilities of the stereo camera in the vehicle sampled data at 10 Hz with respect to an origin at its left lens. Example plots of some raw data may be found in Appendix F.

## **ii. Coordinate transformation from 'Stereo Camera' to 'IMU' coordinate system and data resampling**

Coordinate transformation in MATLAB was employed to bring the stereo camera detections to the IMU frame of reference. Translation followed by quaternion rotation accomplished this conversion. The reasons for choosing the Prius' IMU as the universal coordinate system were:

- It was fixed and unchanging unlike the Tobii's or a custom frame of reference like the driver's head.
- The transformation matrices to convert to the IMU were already known in ROS.
- All measurements by the Smart Eye were made with respect to the IMU, so it was straightforward to convert only stereo camera detections to IMU.

The Tobii measurements could not be directly converted to the IMU because transformation matrices were not available, and the Tobii was not stationary to allow their calculation. Resampling using the nearest-neighbor interpolation technique was used to bring the Tobii, its gyroscope, and the stereo camera to 60 Hz and synchronize it with the Smart Eye. Using the known time instants of the trial starts and trial ends in all the devices (determined using the torch flash instants), the data for the duration of the desired trials were extracted and used to generate the animation, as is explained further down.

## **iii. Removal of false positives in the detections of pedestrian's and torch operating researcher's locations**

False positives in the stereo camera's detections of the pedestrian location were eliminated using a distance threshold of 7.5 m. This distance was the maximum distance a pedestrian could be at from the stereo camera (not the driver, as calculated previously) during any trial since he/she always stood at a marked spot on the imaginary curb or walked along an imaginary crossing connecting two marked spots on the imaginary curb. The stereo camera was approximately 0.5 m from the driver's head (above and in front) in the direction of the pedestrian. So, the maximum distance of the pedestrian from the stereo camera was 0.5 m less than the maximum distance of the pedestrian from the driver. This threshold ensured that detections of the torch-holding experimenter, bystanders, passers-by, misclassified objects did not pollute the pedestrian location data, since they did not fall within the threshold.

A similar approach was used to eliminate false positives in the stereo camera's detections of the location of the torch-holding experimenter. His position was also necessary, in combination with the pedestrian's head's gyroscope data, to plot the pedestrian's gaze ray, as will be explained in a later paragraph. The same distance threshold of 7.5 m (from the stereo camera) as before was used, along with an additional lateral distance threshold equal to the car's total width on either side of the vehicle's centerline. This time, however, the 7.5 m distance threshold was the minimum distance the experimenter could be at during any trial, effectively separating his detections from that of the pedestrian. Since the experimenter nearly always stood directly in front of the vehicle, the lateral threshold excluded all detections to the sides or in the background, such as bystanders, passers-by and misclassified objects.

## **iv. 3-D rectangular box generation for the Prius**

A cuboid was constructed with the dimensions of the car to represent the car in a 3-D animation. The purpose of this animation is to recreate the driver-pedestrian interaction scenario. The cuboid could be drawn as a patch in the animation. Additionally, an image of the top view of the car was obtained from the internet and overlaid on the plots when viewed in the XY plane.

## **v. Gyroscope data, filtering and cumulative numerical integration**

The Tobii gyroscope measured head rotation rates (i.e. angular velocities in deg/s) about its X, Y and Z axes (Figure 6). Only the pedestrian's head yaw and head pitch were considered for plotting his/her gaze ray. These were rotations



about the Tobii gyroscope's Y-axis and X-axis respectively. Head roll (about the gyro Z-axis) was neglected since head tilt to the right/left side during tasks like standing at a curb or crossing a road was assumed to be small.

The gyroscope data from the Tobii Pro Glasses 2 was filtered using a Butterworth low-pass filter of order 2 and a cut-off frequency of 6 Hz. This filtering was performed to reduce high-frequency noise in the gyroscope measurements. The motivation for choosing these filter settings was because of the close match between raw and filtered signals, determined by visual inspection. The cut-off frequency value was chosen based on the findings of literature regarding the upper limits of human head-turning frequencies (Grossman et al., 1988; Kooijman, 2018). The filtered gyroscope data was cumulatively integrated to obtain head rotation angles in degrees.

#### **vi. Gyroscopic drift estimation and compensation**

Low-frequency gyroscope drift persisted in the gyroscope signal and could not be removed by high-pass filtering or linear de-trending in MATLAB. It was also amplified on the cumulative numerical integration of the gyroscope angular rate (deg/s) to obtain the angle (deg). This effect was countered by subtracting zero-drift rate obtained from reference measurements. These benchmarks were gyroscope readings taken over 3 minutes when the Tobii Pro Glasses 2 were either placed on a stationary table or swiveled on a rotating chair at a constant angular rate. The choice of 3 minutes for the benchmark recordings was for them to be comparable to the recordings from the experiments, which were also approximately 3 minutes long. The corrected head rotation angles were added to eyeball directions and used for dead-reckoning (odometry) of the pedestrian gaze ray, as will be elaborated in the next paragraph.

#### **viii. Addition of eyeball directions to head directions**

For greater accuracy in pedestrian gaze direction, the Tobii's readings of eyeball direction were added to the gyroscope readings of pedestrian head rotation. The former was the person's looking direction relative to their head position, while the latter was the person's facing direction. In other words, the eyeball direction measurements from the Tobii were the eye directions inside the wearer's head and were especially important when participants were gazing out of the corner of their eyes. Due to persistent high-frequency noise in the eyeball directions despite applying Tobii's IV-T Fixation Filter, the resampled eyeball directions were filtered once again in MATLAB using a low-pass Butterworth filter of order 2 and a cut-off frequency of 30 Hz, determined by trial-and-error.

#### **viii. 3-D plot of vehicle, persons and gaze rays**

With all data resampled to 60 Hz, the driver's gaze origin inside the Prius, the pedestrian's location, and both parties' gaze rays were plotted for every frame between the torch 'start' and torch 'end' instants of a trial. This plot reconstructed the driver-pedestrian interaction scenario by synchronizing all of the following data sources:

- Stereo camera detections of the pedestrian's and torch-operating researcher's locations
- Tobii gyroscope measurements of the pedestrian's head (facing) directions
- Tobii measurements of the pedestrian's eyeball (looking) directions
- Smart Eye readings of the driver's head positions and orientations
- Smart Eye readings of the driver's gaze directions

Separate animated plots were generated for each scenario in a recording. The Smart Eye measured the driver's gaze direction over time with reference to an IMU in the vehicle whose position was absolute and constant. This fact made mapping the driver's gaze ray straightforward. However, the Tobii measured the pedestrian's gaze direction relative to the center of its scene camera, located atop the wearer's nasal ridge, between the eyes. Clearly, due to the dynamic nature of the pedestrian's head during the experiment, this origin would not be fixed in 3-D space, thus making the gyroscope readings relative to an unknown initial head position and orientation.

To counter this effect, an instant when the pedestrian's head pose and eyeball directions were known was required. Using this known reference, it would be possible to determine the person's heading and gaze direction throughout the trial by dead reckoning (odometry) with the gyro measurements. A suitable solution was to use the torch 'start' instant of the trial when it was assumed the pedestrian was facing and looking directly at the flashlight held at the



experimenter's eye level (as per the instructions). Thus, for every frame, the pedestrian gaze ray was estimated and plotted by rotating the known gaze ray at the trial start by the corresponding angle obtained from the gyroscope after filtering and drift correction.

**ix. Classification of frames and trials as 'eye contact' or 'no eye contact'**

The vector dot product between the driver and pedestrian gaze rays was calculated to find the angle between them. In MATLAB, this angle was the angle between the rays in the common 2-D plane that contained them both (or one ray and the parallel of the other ray). If the angle was below the 'eye contact threshold' at any instant, that instant was marked as a '1' (denoting eye contact) and otherwise '0' (denoting no eye contact). The percentages of eye contact instants in all standing/crossing trials for a participant were calculated and sorted in descending order. The two highest (out of four, in the standing scenarios) and the higher (out of two, in the crossing scenarios) were classified as 'eye contact', and the others as 'no eye contact'. The classification was performed participant by participant instead of for all participants at once to minimize misclassification due to individual differences in eye contact seeking behavior. For example, since some participants sought less eye contact than others, it was possible that the formers' eye contact trials would be wrongly identified as no eye contact trials by the algorithm, if all participants' trials were classified at once.

**x. Classification performance of the algorithm**

To measure the performance of the algorithm, a couple of ground truths were used. First, the trial classification was compared with the instructions to the participant for that trial involving either making eye contact with the driver or not. Additionally, the instants of eye contact (or lack thereof) identified by the algorithm were compared with the instants of eye contact (or lack thereof) known from the ground truth via binary plots of driver-pedestrian eye contact. The ground truth was obtained by counting frames showing gaze data inside AOIs drawn around the driver's face, and the angles between the driver's actual gaze ray and his ideal gaze ray that fell below the 'ground truth threshold'. The former provided pedestrian eye contact seeking during the trial, and the latter the driver's. The logical AND of the two provided a ground truth for eye contact in that scenario. In other words, both driver and pedestrian were looking at the region around each other's head (ideally, at each other's eyes) at the same time. Any missing data in the ground truth (usually from the Tobii's end) was counted as instants of 'no eye contact' to ensure that the ground truth signal was unbroken. This decision was made after visual inspection of the plots of pedestrian eye contact seeking in the trials, which showed that missing data occurred primarily when the pedestrian was looking away from the driver at bright surfaces. Since such an instant would anyway be a 'no eye contact' instant, missing data was treated as such.

The individual eye contact seeking graphs are shown in the results section, and Figures 17a, 17b, 18a and 18b illustrate the driver-pedestrian eye contact during some trials.



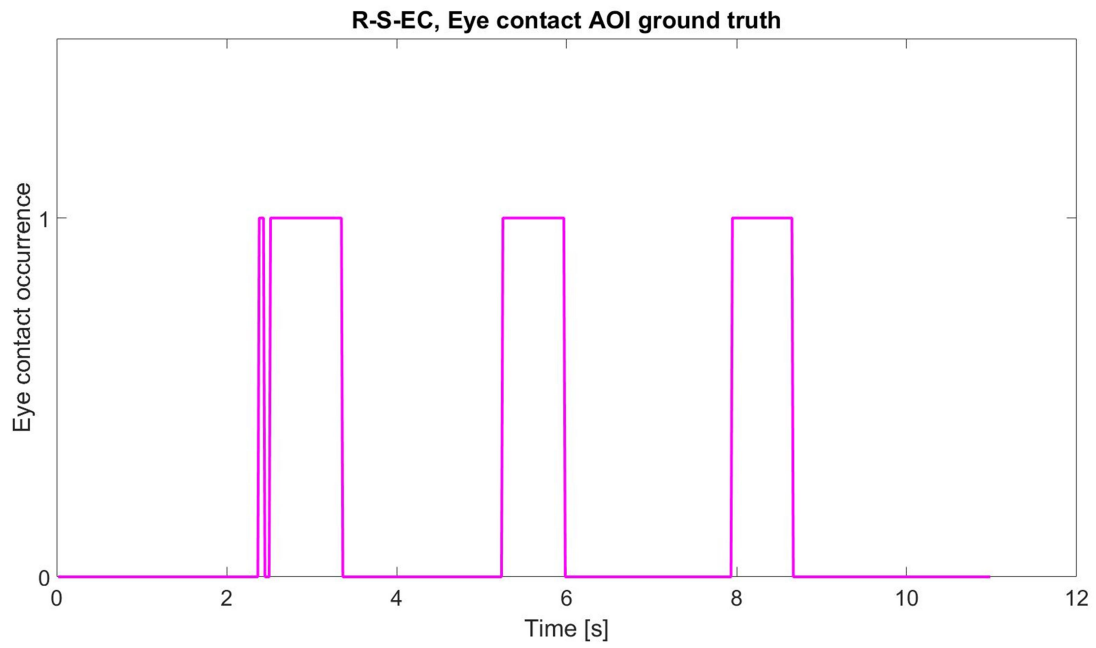


Fig. 17a – Binary plot of eye contact for the ground truth of a standing trial

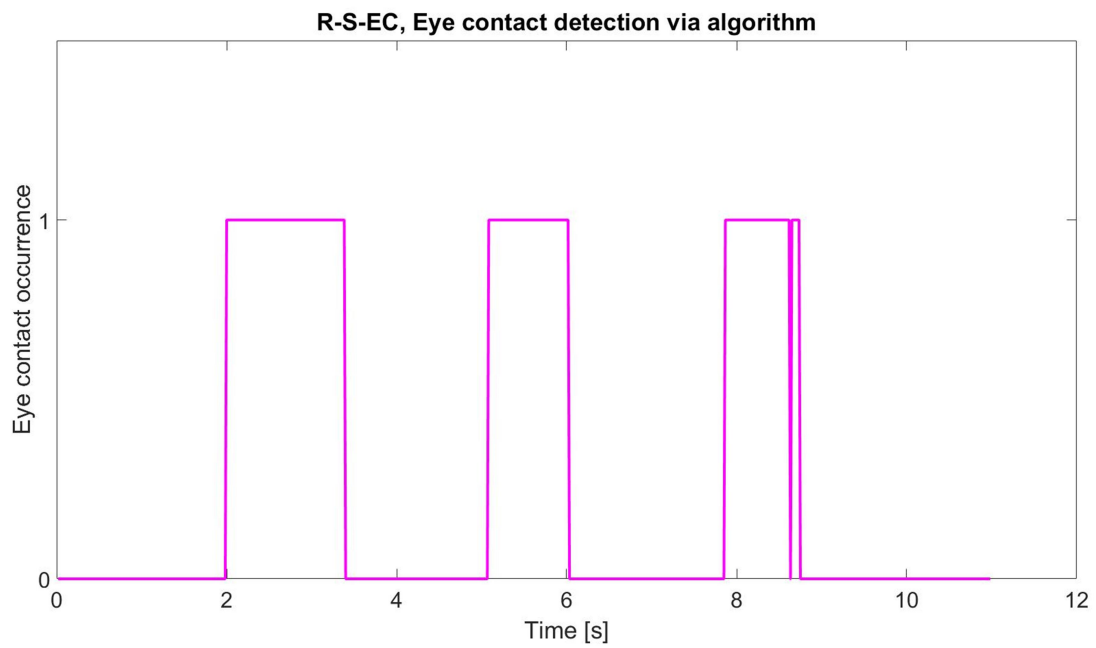


Fig. 17b – Binary plot of eye contact for the algorithm's detections during the same standing trial



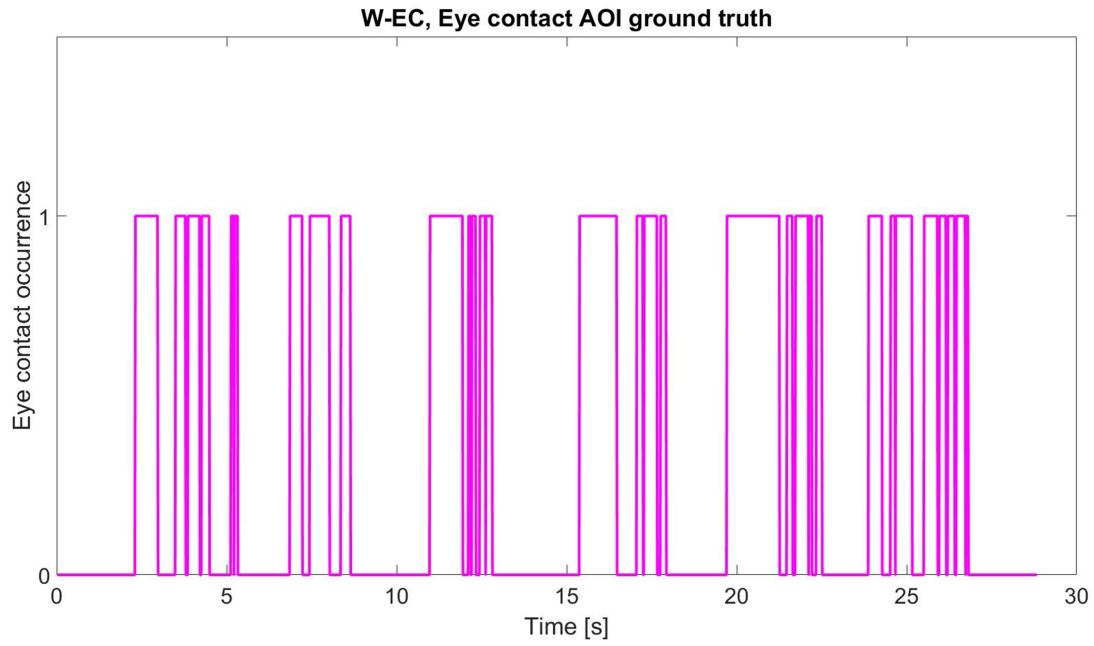


Fig. 18a – Binary plot of eye contact for the ground truth of a crossing trial

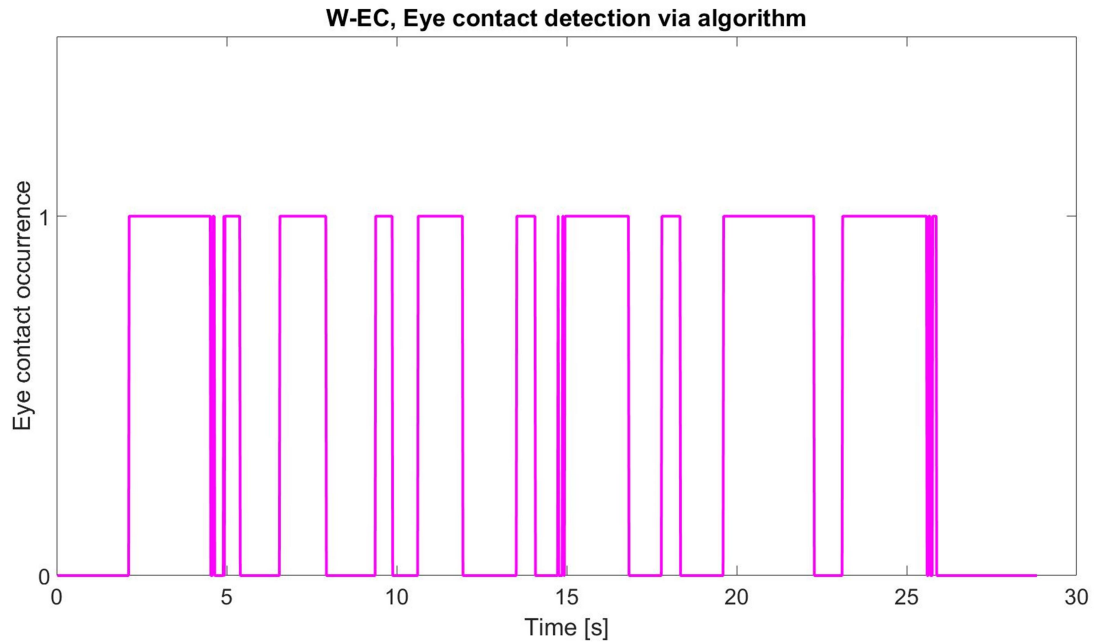


Fig. 18b – Binary plot of eye contact for the algorithm's detections during the same crossing trial

The binary plots for the algorithm and ground truth were square wave plots with '1' denoting eye contact and '0' denoting no eye contact for every instant of the trial. Hit rates and false-positive rates of the algorithm were calculated using the algorithm's detections and the ground truth for different 'eye contact thresholds' and a fixed 'ground truth threshold'. In other words, the 'ground truth threshold' (i.e. angle threshold between ideal driver gaze ray and actual driver gaze ray) was kept constant in the algorithm, while the 'eye contact threshold' (i.e. angle threshold between pedestrian gaze ray and actual driver gaze ray) was varied in steps to measure the algorithm's performance. These thresholds proved to be higher than their theoretical estimates, as will be explained later. Receiver operating



characteristics (ROC) curves were plotted to judge how well the algorithm worked, as may be seen in the results section.

#### xi. Video

The top view of the 3-D animated plots was written to an MPEG-4 video for convenient presentation. The top view specifically was chosen because it best illustrated the occurrence of eye contact and the angle between the gaze rays. The color blue was used for displaying the pedestrian's gaze ray and location, and the color red the driver's. Screenshots from a few videos are shown in Figure 19a, 19b, 20a and 20b below.

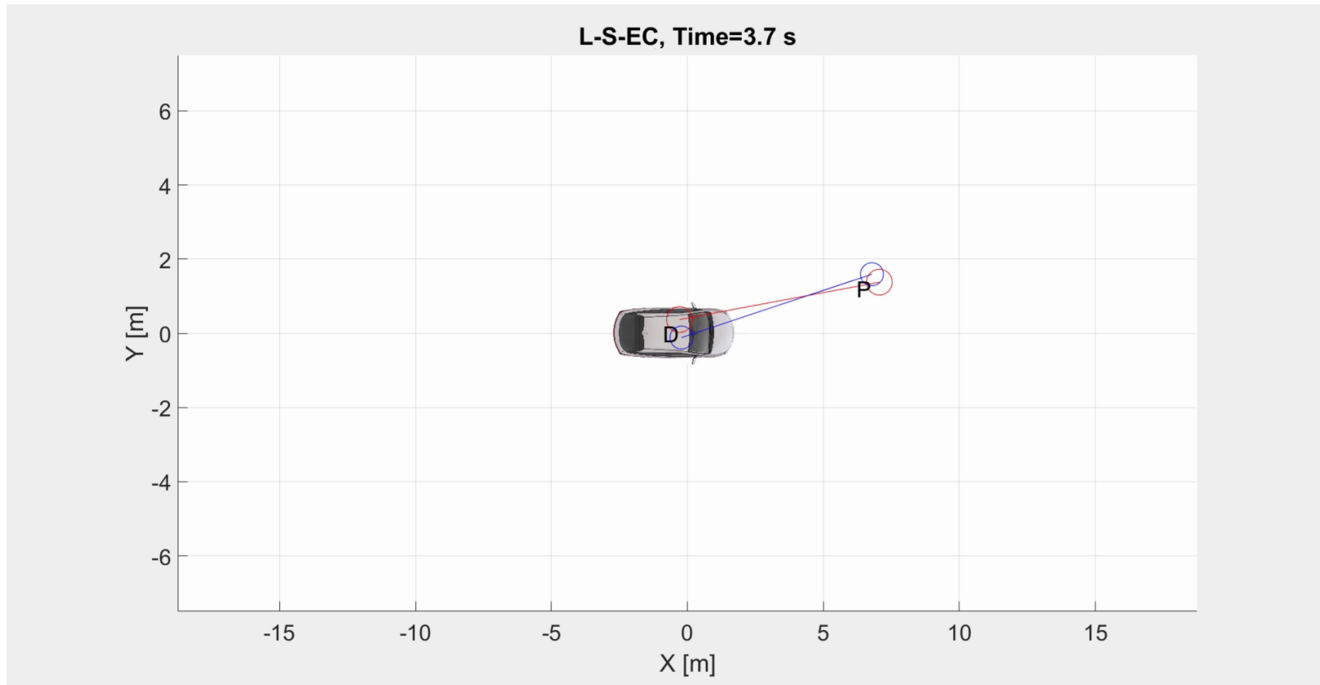


Fig. 19a – Top view from the algorithmic reconstruction of a standing trial showing eye contact

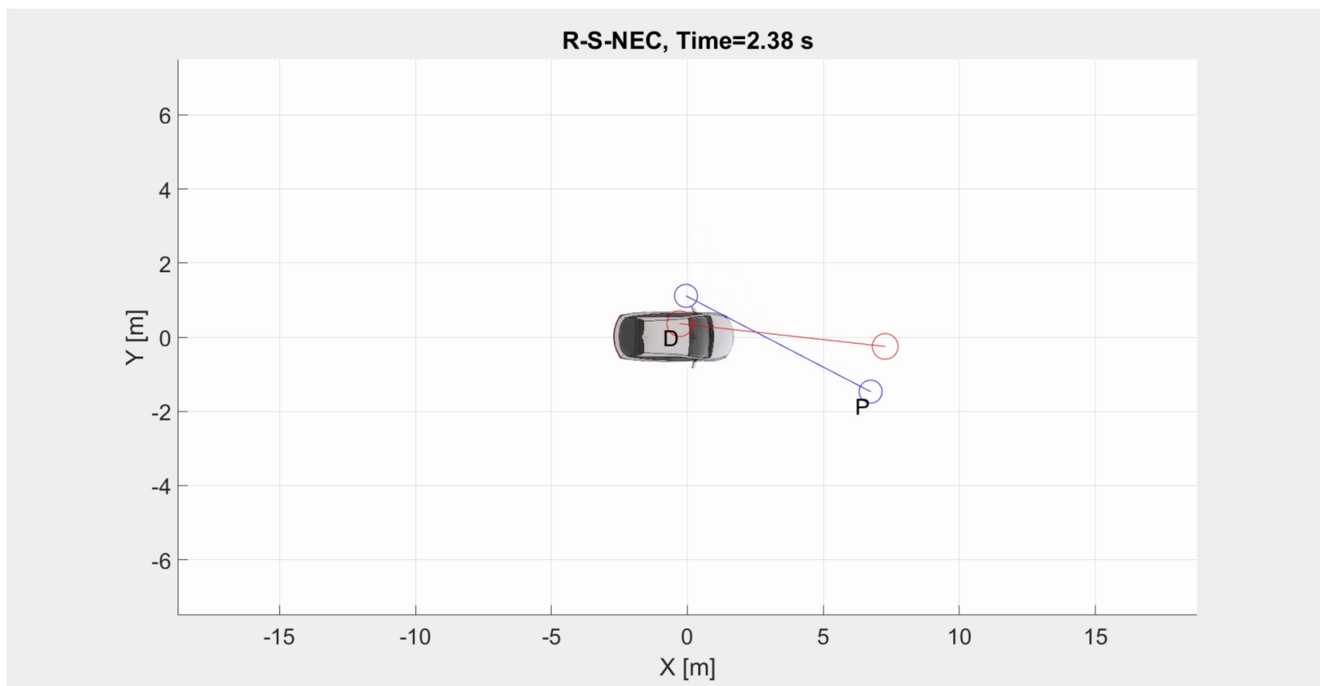




Fig. 19b – Top view from the algorithmic reconstruction of a standing trial showing the lack of eye contact

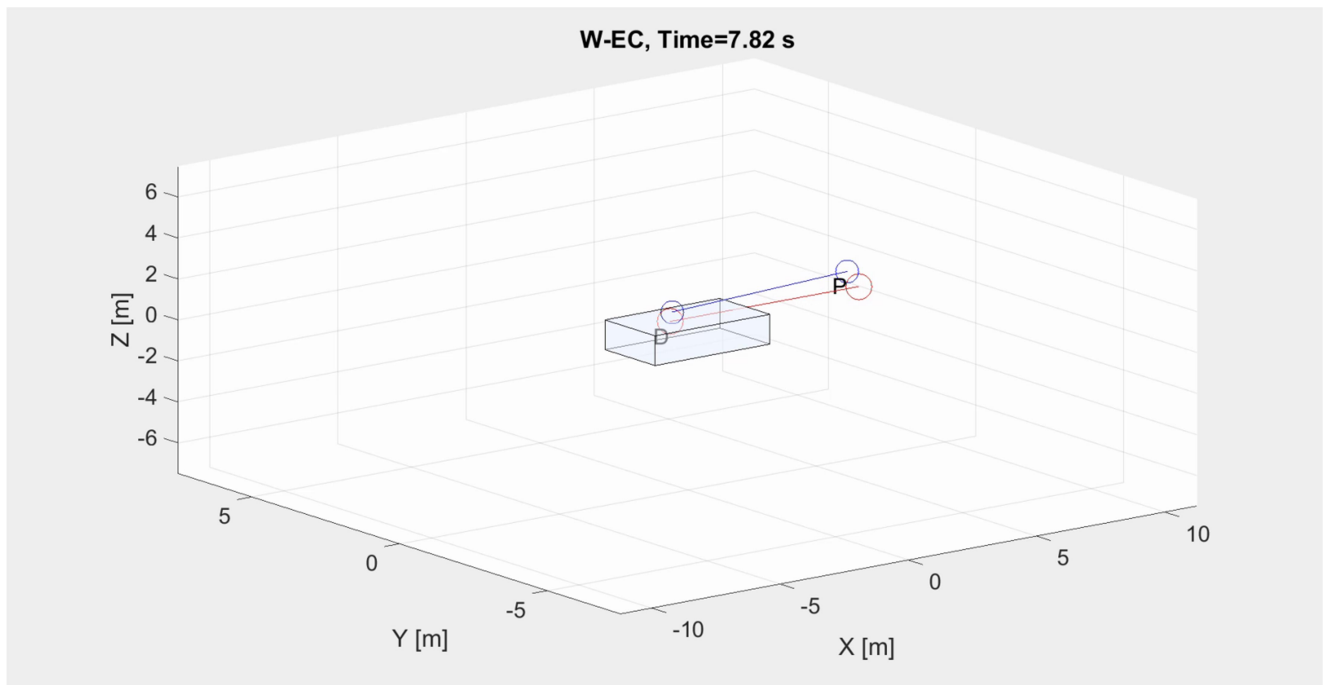


Fig. 20a – Isometric view from the algorithmic reconstruction of a crossing trial showing eye contact

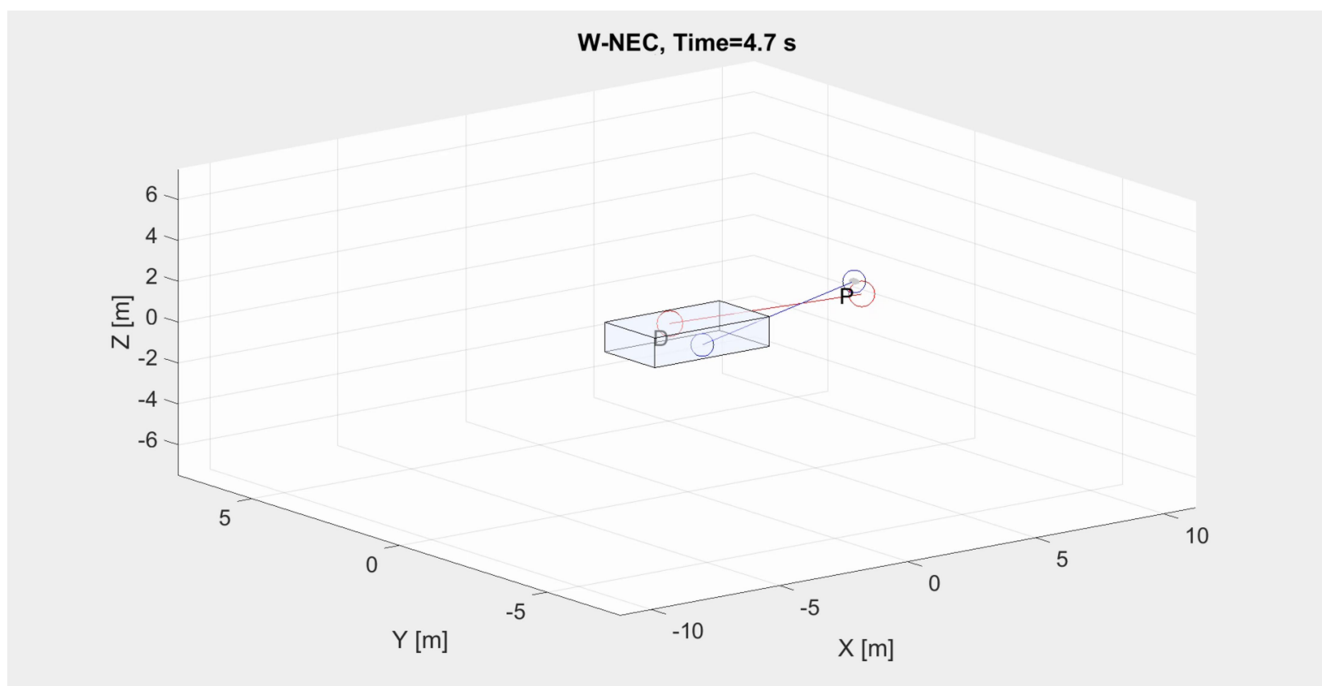


Fig. 20b – Isometric view from the algorithmic reconstruction of a crossing trial showing the lack of eye contact



# 3 Results

## I. Participants and Tobii gaze data quality:

Out of the total of 31 volunteers, it was observed that there was an overrepresentation of people with brown eyes (24 participants), males (23 participants), MSc students (22 participants) and Indian nationals (14 participants). Participants had a mean age of 24.84 years and a standard deviation of 2.28 years. They also had a mean height of 1.75 m with a standard deviation of 0.08 m.

The quality of the pedestrian gaze data from the Tobii Pro Glasses 2 was checked against the minimum requirements using visualizations of participant gaze sample percentages. These data quality checks were performed for both the whole durations of the recordings and for just the durations of the trials.

For comparing whole recordings, the recordings of all participants were segregated into 'standing' or 'crossing' categories as shown in Figure 21 below. The standing trials had a mean gaze availability of 85.29% (median 86%) with a standard deviation of 7.13%, and the crossing trials a mean gaze availability of 89.45% (median 90%) with a standard deviation of 5.38%. A Wilcoxon signed-rank test was used to compare the medians of the paired, non-normal data (i.e. standing trials vs. crossing trials) and it revealed a significant difference in the medians between standing and crossing trials ( $p < 0.001$ ) at the 5% significance level.

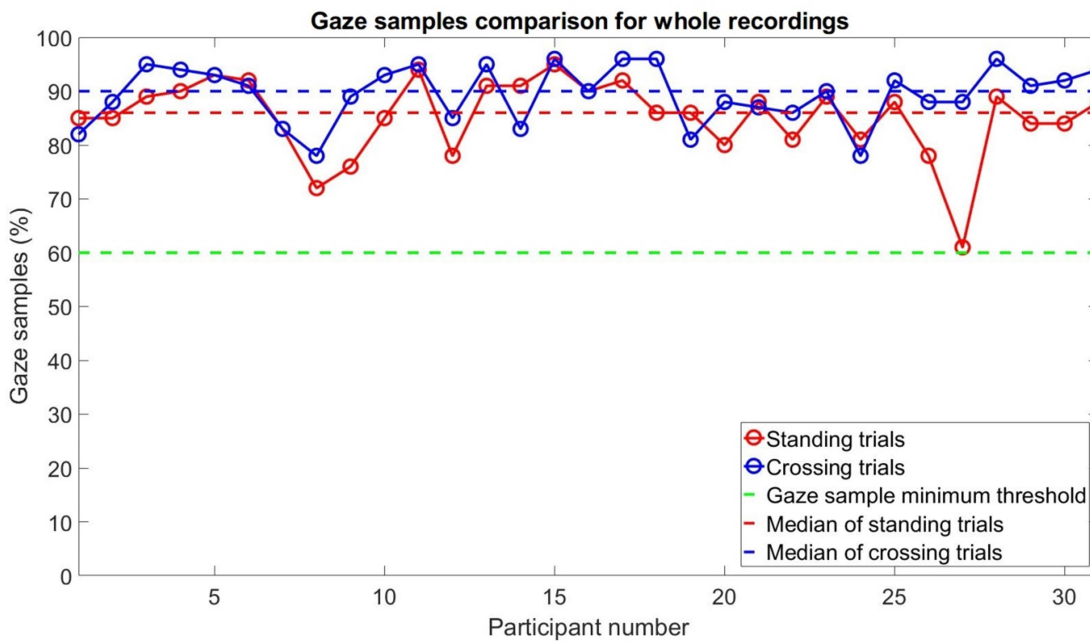


Fig. 21 – Data quality check using gaze sample percentages over entire recordings  
As seen, sufficient gaze data appeared to have been collected for every participant

For comparing individual trials, the trials were extracted from the recordings of all participants and separated based on the type of trial as seen in Figure 22. A Friedman test was used to compare the means of all four standing trials and it revealed no significant differences ( $p = 0.4862$ ) at the 5% significance level. The Friedman test's results for the standing trials are shown below in Table 3.

Table 3 – Friedman test ANOVA table between the four types of standing trials  
Here, 'SS' refers to sum of squares, 'dF' to degrees of freedom, and 'MS' to mean squares

Source	SS	dF	MS	Chi-square	Prob.>Chi-sq.
Columns	4.067	3	1.35556	2.44	0.4862
Error	145.933	87	1.67739		



Total	150	119			
-------	-----	-----	--	--	--

The Friedman test also revealed that there were no significant differences ( $p=0.0679$ ) in the means of the two types of crossing trials at the 5% significance level. These results for the crossing trials can be found below in Table 4.

Table 4 – Friedman test ANOVA table between the two types of crossing trials  
Here, 'SS' refers to sum of squares, 'dF' to degrees of freedom, and 'MS' to mean squares

Source	SS	dF	MS	Chi-square	Prob.>Chi-sq.
Columns	1.6667	1	1.66667	3.33	0.0679
Error	13.3333	29	0.45977		
Total	15	59			

However, on performing a third Friedman test comparing all the six types of trials at once, significant differences in the mean gaze sample percentages between standing and crossing trials were found ( $p<0.001$ ) at the 5% significance level. These results are shown in Table 5. These results together reveal that both at the level of the whole recordings and at the level of the individual trials, gaze data qualities were different for the standing and crossing trials.

Table 5 – Friedman test ANOVA table between all types of trials  
As before, 'SS' refers to sum of squares, 'dF' to degrees of freedom, and 'MS' to mean squares

Source	SS	dF	MS	Chi-square	Prob.>Chi-sq.
Columns	97.867	5	19.5733	27.96	<0.001
Error	427.133	145	2.9457		
Total	525	179			

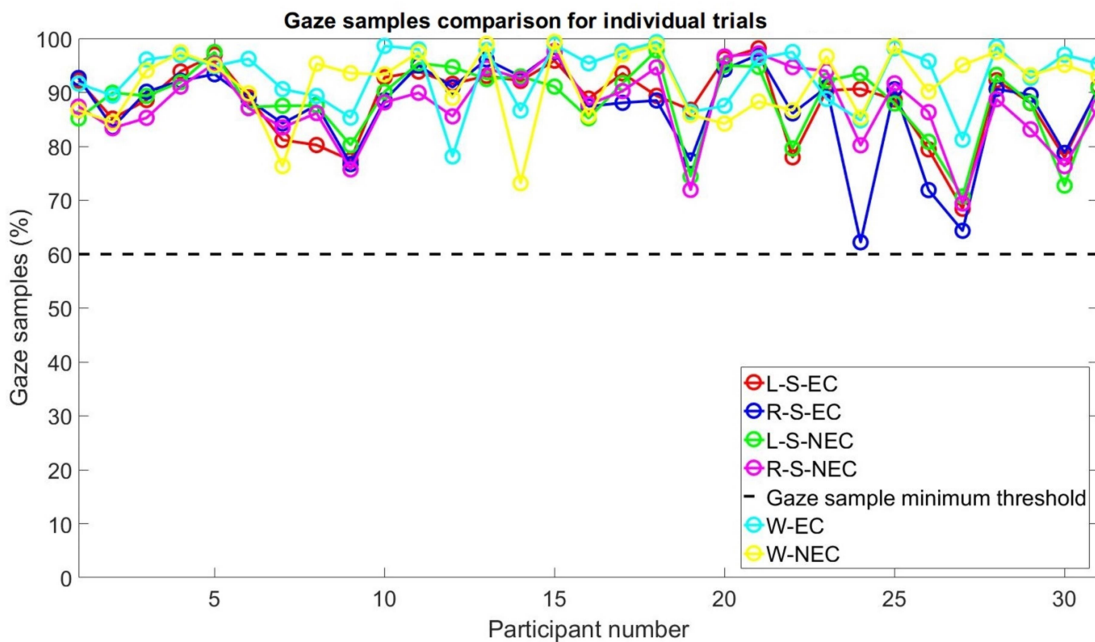


Fig. 22 – Data quality check using gaze sample percentages over trials  
Here too, all trials for all participants show sufficient gaze data

The statistical results above suggest that while there were no differences in gaze sample percentages for participants within the standing trials or within the crossing trials, there were differences between these two groups i.e. the crossing trials had higher gaze sample percentages than the standing trials.

The gaze sample percentages for all trials and recordings were above the minimum threshold of 60%, so no Tobii data was excluded. The high values (in both standing and crossing trials) suggested that interference from background



infrared light was successfully minimized. Further, almost all the lines plotting gaze samples for different trials had a similar trend. In other words, the valleys and peaks in the plots of different trials occurred at the same participant numbers.

Smart Eye gaze data from the driver was unavailable during the experiment of one participant, which meant that eye contact could not be determined for this person. This participant's data was excluded from the analyses, leaving 30 valid participants. The cause of the unavailability was attributed to a technical error in the Smart Eye software wherein it failed to transfer the buffered gaze data to ROS for exporting.

## II. Responses to post-experiment questionnaire:

A table summarizing the responses to the post-experiment questionnaire was also created to understand participants' experience of performing the experiment (Table 6). This questionnaire was interesting because it revealed the level of fidelity of the experiment, and the applicability of the results to real roads.

Table 6 – Summary of responses to the post-experiment questionnaire

Responses were recorded on a 7-point Likert scale, with 1 denoting 'Not at all', and 7 denoting 'Completely'. Responses of 1-3 were negative, 4 was neutral, and 5-7 were positive.

Question	Positive response (% of participants)	Negative response (% of participants)	Neutral response (% of participants)	Mean response score (1 to 7)
How well were you able to imagine a real traffic scenario at a real pedestrian crossing?	64.5	19.3	16.2	4.74
How well could you concentrate on the task during the experiment?	100	0	0	6.42
How well could you maintain eye contact during the experiment without looking elsewhere?	90.3	0	9.7	5.87
How well could you avoid eye contact during the experiment and look elsewhere?	100	0	0	6.29
How conscious were you of wearing eye tracking glasses while performing the experiment?	48.4	51.6	0	3.97
How much do you think the eye tracking glasses affected your natural road crossing behavior?	22.6	64.5	12.9	2.87
How realistic did the experiment feel compared to crossing a real road?	45.2	22.5	32.3	4.32
How involved were you during the experiment?	96.8	0	3.2	6.32
How clear were the instructions for the experiment?	96.8	0	3.2	6.68
How well do you think you followed the instructions for the experiment?	96.8	0	3.2	6.48

It was noted that 64.5% of respondents reported that they could imagine they were in a real traffic scenario while performing the experiment, compared to only 19.3% who did not. Additionally, it was observed that 45.2% of participants directly reported finding the experiment comparably realistic to crossing a real road, as opposed to only 22.5% who did not.

96.8% of respondents reported each of the following: that they were involved in the experiment, found the instructions clear, and did well to carry out the tasks (i.e. scores of 5, 6 or 7). All participants reported that they could avoid eye contact with the driver well but only 90.3% said they could maintain eye contact well. Responses varied on the feeling of consciousness while wearing the Tobii glasses (i.e. a mean score close to 4, and almost an equal number of positive and negative responses). It was noted that participants who wore contact lenses had a mean response of 2.5 for the feeling of consciousness while wearing the Tobii, compared to a mean response of 4.5 by those who did not use any form of eyesight correction.



### III. Definition of eye contact (Research Question 1):

The first research question of this thesis was:

- What is eye contact between a pedestrian and a driver in a car? How can eye contact be defined/operationalized using an algorithm?

To answer this, a definition of driver-pedestrian eye contact was framed. From the literature, it was already known that eye contact in such a traffic setting involved a driver and a pedestrian looking at each other's eyes at the same time. This could be translated as an overlap in the eye contact seeking behaviors/signals of the two parties. Figures 23 and 24 below illustrate the amount of eye contact that occurred between the 30 valid participants and the researcher in the various trials. These measurements of eye contact were obtained by combining the driver's eye contact seeking and the pedestrian's AOI-based ground truths (instead of the algorithm) as they provided a more accurate picture of eye contact occurrence.

#### i. Eye contact in the standing trials:

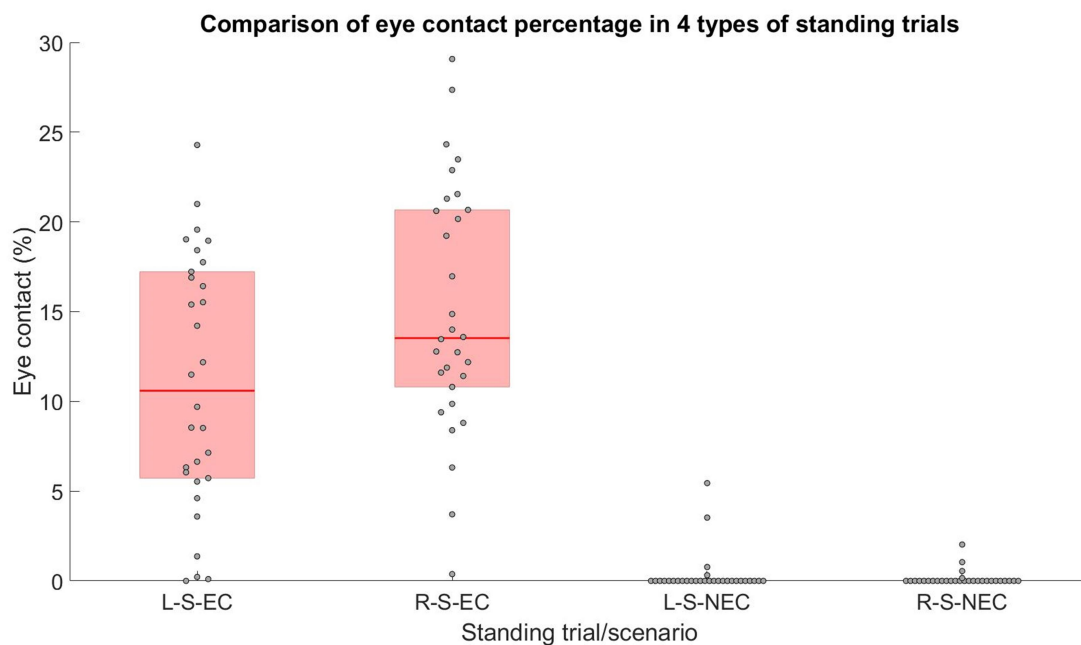


Fig. 23 – Eye contact as a percentage of standing trial duration

The median percentage of eye contact in a trial for the 'L-S-EC' (Left-Standing-Eye Contact) scenario was 10.59%, and for the 'R-S-EC' (Right-Standing-Eye Contact) scenario, it was 13.53%. These percentages were not as low as they seemed, as will be covered in the discussion of this report. A comparison of the eye contact percentage medians for 'L-S-EC' and 'R-S-EC' scenarios using the Wilcoxon signed-rank test revealed that the medians were unequal ( $p=0.027$ ) at the 5% significance level. This result meant that eye contact percentages were greater when the pedestrians stood on the right side of the car than when they stood on the left.

It was judged that eye contact was best expressed as a percentage of the standing trial duration (rather than as a percentage of total duration of available gaze data in a trial, or as a percentage of total fixation time in a trial) because the duration of a driver-pedestrian interaction on a road was more readily measurable (i.e. without eye-trackers) than the amount of available gaze data or total fixation time. Further, percentage measures were more generalizable than time durations, since durations might vary depending on the scenario.

Each repetition of the standing trials was designed to be similar to a real interaction on a road where both driver and pedestrian turned to look at each other briefly before negotiating right of way and then looking away. Thus, one-third of



the standing trial durations (i.e. the duration of one repetition) were a reasonable approximation of real driver-pedestrian interaction times on the road (since there were three repetitions per trial).

The boxplot of standing trial durations from Figure 25 showed that 'L-S-EC' interactions lasted for a median 11.88 s and 'R-S-EC' interactions for 11.97 s. Eye contact durations were obtained by multiplying corresponding values of eye contact percentage and trial duration, and dividing by three:

$$t_{eye\_contact\_left} = \frac{0.1059 * 11.88}{3} = 0.42s$$

$$t_{eye\_contact\_right} = \frac{0.1353 * 11.97}{3} = 0.54s$$

Eq. 4 – Calculation of typical eye contact durations for a standing pedestrian on the curb  
In other words, these durations represent eye contact in one repetition of a standing trial

These values are proposed as the typical duration of driver-pedestrian eye contact for a pedestrian standing on the curb and in a crossing conflict with an approaching vehicle.

## ii. Eye contact in the crossing trials:

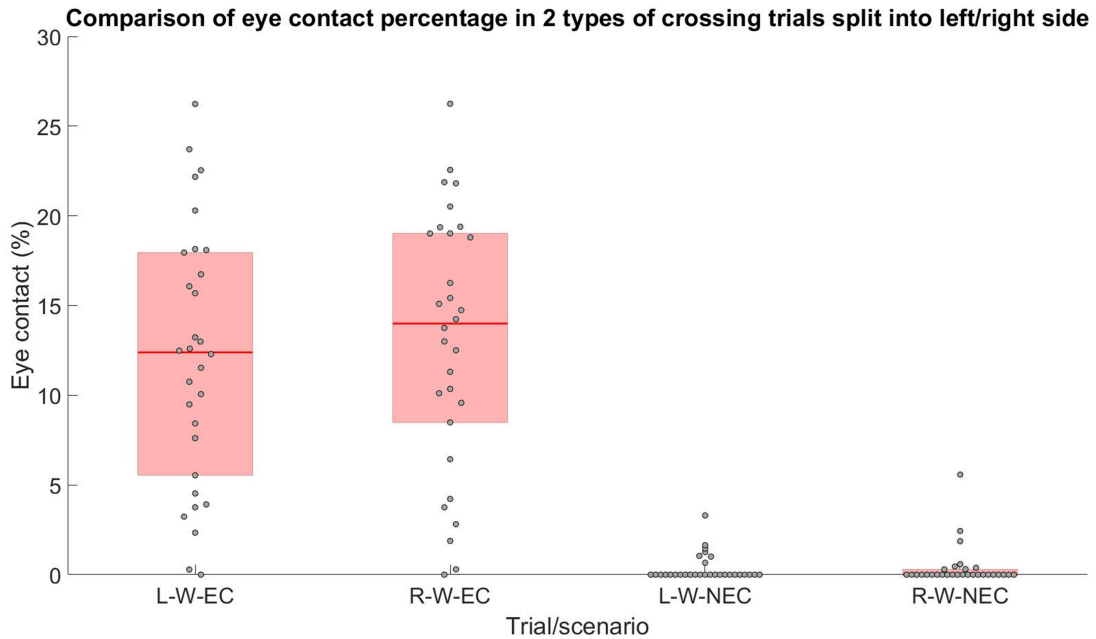


Fig. 24 – Eye contact as a percentage of crossing trial duration

To recap, the crossing trials consisted of three repetitions of walking back and forth in front of the vehicle. Hence, one repetition involved walking once from the right side of the vehicle to the left, followed by walking back once from the left side of the vehicle to the right. Participants were not obliged to look at the driver/vehicle while turning around in order to walk back. For the purposes of these results, the repetitions of the crossing trials were split into halves – the right side walking half and the left side walking half. This separation allowed the comparison of eye contact between the cases where the pedestrian started crossing from the right side of the vehicle, and when he/she started crossing from the left side of the vehicle.

For the left side crossing half-repetitions with eye contact (i.e. 'L-W-EC'), the median percentage of eye contact was 12.39% and for the right side crossing half-repetitions (i.e. 'R-W-EC'), it was 14%. Once again, these percentages seemed lower than they actually were, as will be explained in this report's discussion section. A comparison of the eye



contact percentage medians for 'L-W-EC' and 'R-W-EC' using the Wilcoxon signed-rank test revealed that there were no significant differences in the medians ( $p=0.5377$ ) at the 5% significance level. This result meant that eye contact percentages were the same irrespective of which side of the car people crossed from.

From Figure 26, the median crossing trial duration was 29.85 s. Since there were three half-repetitions each of walking from the left and the right:

$$t_{eye\_contact\_left\_walking} = \frac{0.1239 * 29.85}{3} = 1.23s$$

$$t_{eye\_contact\_right\_walking} = \frac{0.14 * 29.85}{3} = 1.39s$$

Eq. 5 – Calculation of maximum eye contact durations for a crossing pedestrian on the road  
In other words, these durations represent eye contact in half a repetition of a crossing trial

These numbers are proposed as the maximum values of the durations of driver-pedestrian eye contact for a pedestrian crossing a road. The reason is that participants were instructed to constantly seek eye contact as they crossed, an action that was reciprocated by the driver, and these behaviors probably do not occur often in a real interaction. Thus, it is likely that the durations of eye contact here are higher than those seen typically when a pedestrian crosses a public road in front of a car.

### iii. Duration of standing trials and crossing trials:

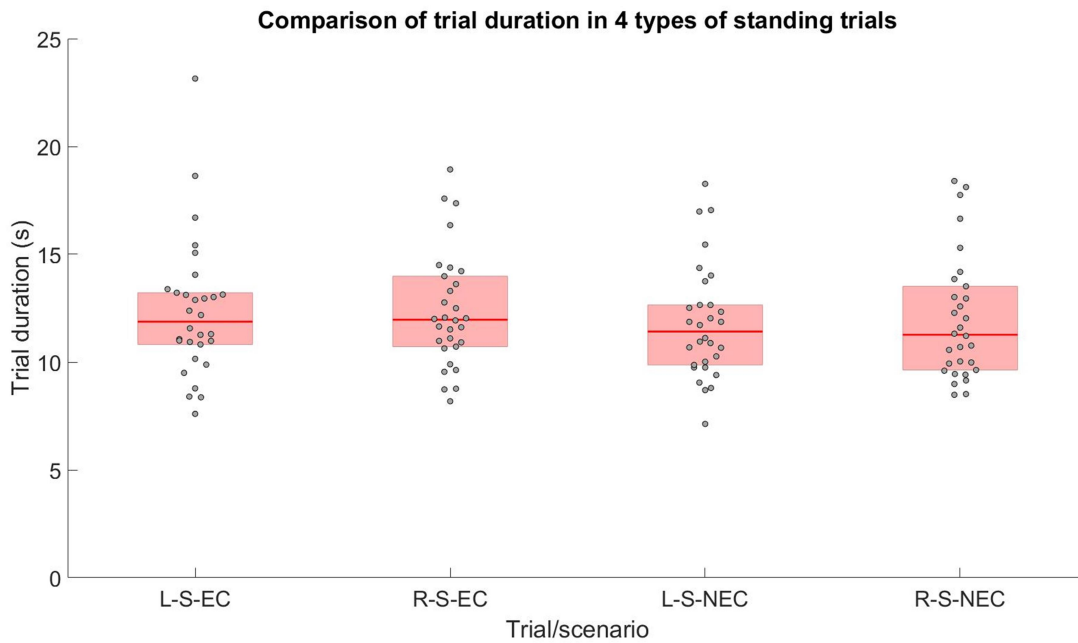


Fig. 25 – Standing trial durations or the time between a 'start' torch flash and an 'end' torch flash

A Friedman test revealed no significant differences in the means of the durations of the four standing trials ( $p=0.4373$ ) at the 5% significance level. Thus, the differences in eye contact percentage (in the standing trials) found earlier were probably not due to differing trial lengths.



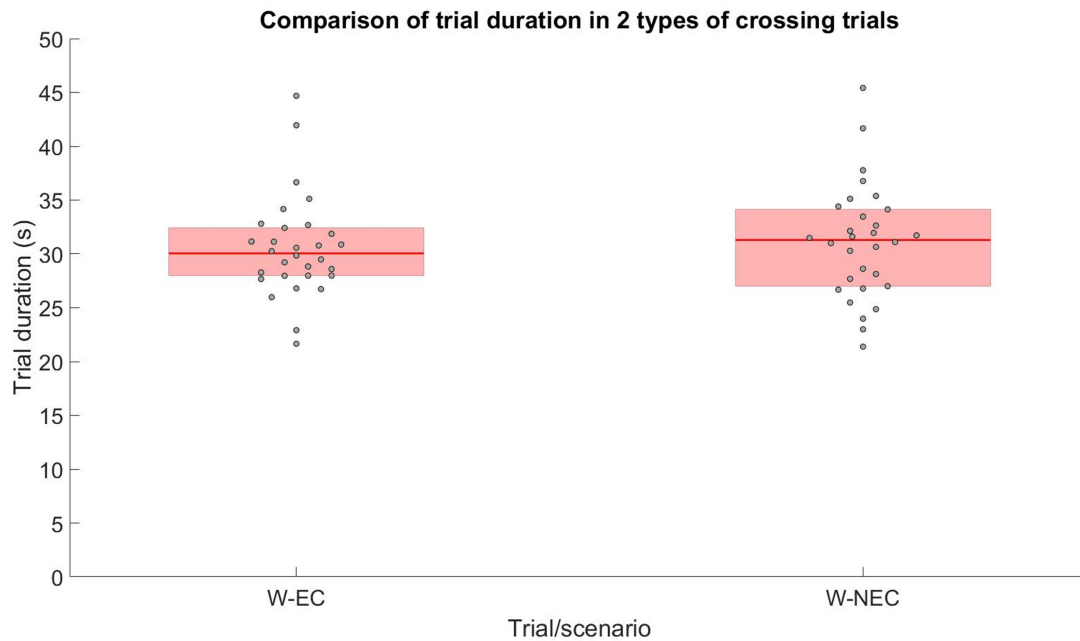


Fig. 26 – Crossing trial durations or the time between a ‘start’ torch flash and an ‘end’ torch flash

Similarly, a Friedman test revealed no significant differences in the means of the two crossing trials ( $p=0.3692$ ) at the 5% significance level.

The analysis of pedestrian eye contact seeking (obtained from the AOI-based ground truth) echoed the results of the eye contact analysis. In other words, when considering Tobii eye-tracking of the pedestrian alone (ignoring the driver) via AOI hits, similar results were obtained. Figure 27 and 28 show the eye contact seeking percentages of the 30 valid participants during the standing and crossing trials respectively.

**iv. Pedestrian eye contact seeking in the standing trials:**

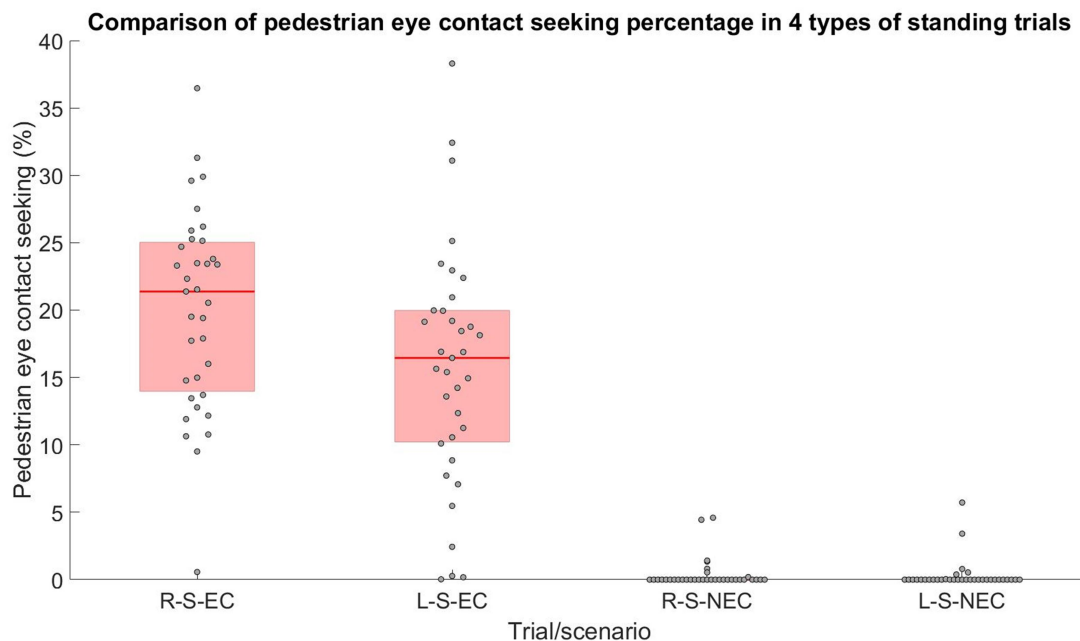


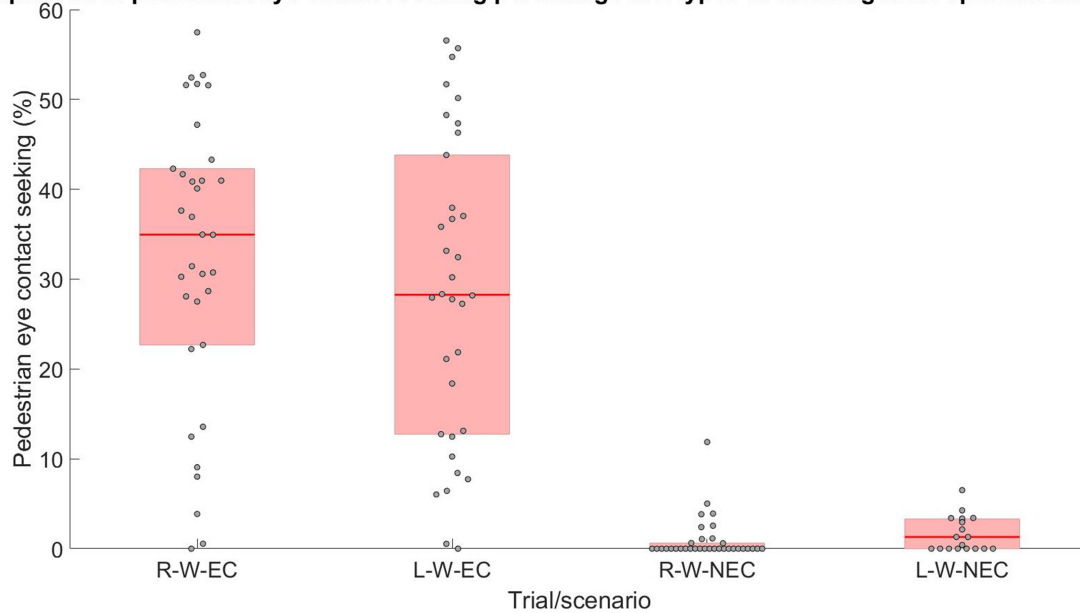
Fig. 27 – Pedestrian eye contact seeking as a percentage of standing trial duration



The median pedestrian eye contact seeking percentage for the 'R-S-EC' trial was 21.38%, and for the 'L-S-EC', it was 16.45%. The difference in the medians was found to be statistically significant by the Wilcoxon signed-rank test ( $p=0.0209$ ) at the 5% significance level. This result showed that pedestrian eye contact seeking percentages were greater when the pedestrian stood on the right side of the vehicle. Since there were no significant differences in standing trial durations, pedestrian eye contact seeking durations were also greater when they stood on the right side of the car. Also, pedestrians in the standing trials sought eye contact for approximately 1.5 times the duration of eye contact occurrence.

**v. Pedestrian eye contact seeking in the crossing trials:**

**Comparison of pedestrian eye contact seeking percentage in 2 types of crossing trials split into left/right side**



**Fig. 28 – Pedestrian eye contact seeking as a percentage of crossing trial duration**

As for the crossing scenarios, the median pedestrian eye contact seeking for right side half-repetitions ('R-W-EC') was 34.95%, and it was 28.26% for the left side half-repetitions ('L-W-EC'). No significant differences were found by the Wilcoxon signed-rank test ( $p=0.2216$ ) at the 5% significance level. Thus, pedestrian eye contact seeking percentage in the crossing trials did not differ based on the side they started walking from. Pedestrians also sought eye contact for approximately 2.5 times the duration of eye contact occurrence.

Pedestrian mean dwell time also provided an idea of how long participants' gaze stayed on the driver. It was the mean duration of all the eye contact seeking visits made by a pedestrian. To recap, an eye contact seeking visit was defined as a group of closely located eye contact seeking fixations. An eye contact seeking fixation was a collection of at least 3 back-to-back frames where the pedestrian's gaze fell inside an AOI and eye velocity was less than  $30^\circ/\text{s}$  in all frames. Most participants had 3 visits (one for each repetition) in the standing trials involving eye contact as seen in Figure 29a. However, it was observed that sometimes this number went up to 5 or got as low as 1. The former case can be seen in Figure 29b. The width of the bars (i.e. the eye contact seeking visits) represented the dwell time for the visits in a given trial.

**vi. Visits in driver and pedestrian eye contact seeking for the standing trials:**



The cases with fewer than 3 visits were possibly due to either a failure to detect by the Tobii or a failure to seek eye contact by the pedestrian. The cases with more than 3 visits were possibly due to noise or uncertainty that had slipped past the filters, or an eye blink which caused a short gap in the data which was not correctly filled / interpolated. This phenomenon may explain why some eye contact seeking repetitions were split into 2 (or more) visits, and why the gaps had durations of approximately 0.1 s.

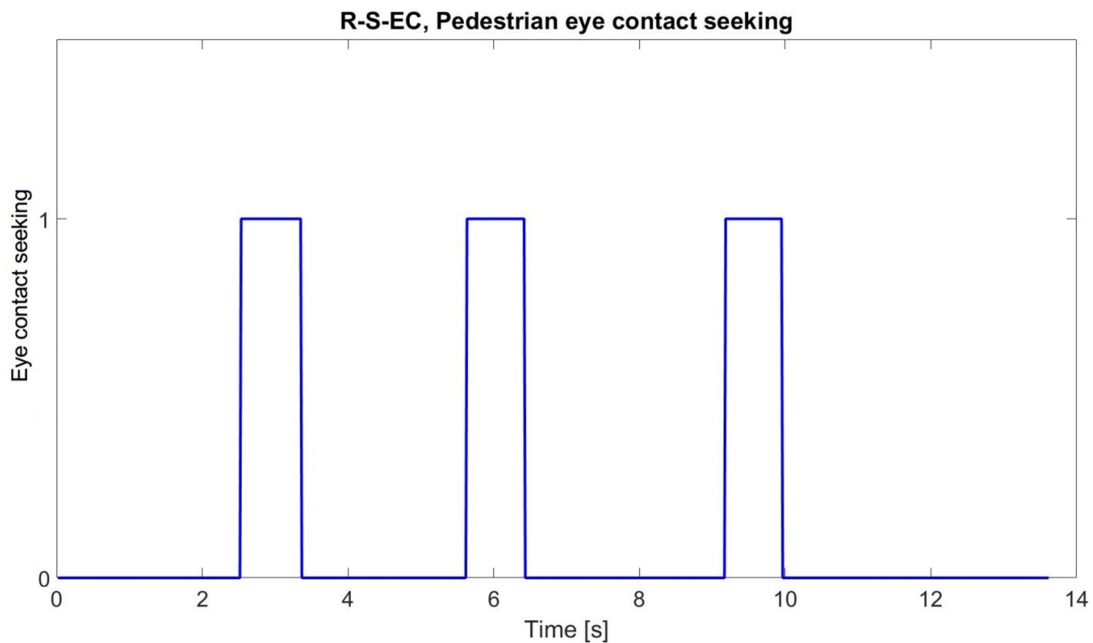


Fig. 29a – Pedestrian eye contact seeking in a standing trial, involving 3 repetitions, each comprised of 1 visit

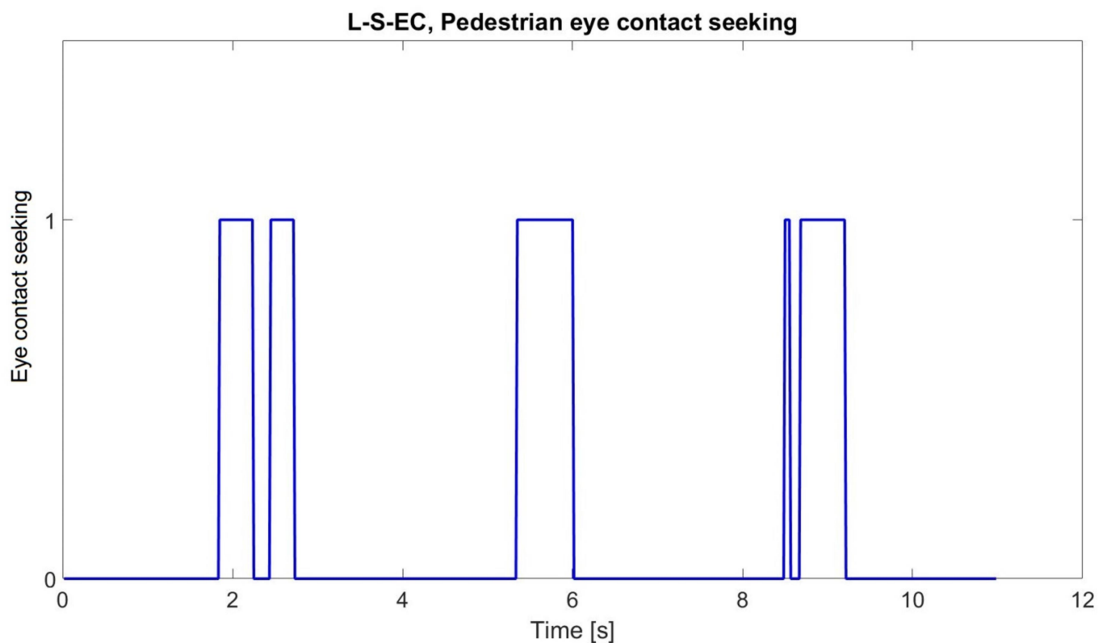


Fig. 29b – Pedestrian eye contact seeking in a standing trial, involving 3 repetitions, each comprised of 1-2 visits

Driver eye contact seeking in the standing trials, on the other hand, was steadier and rarely deviated from the ideal 3 visits, which was expected since the researcher carried out the same pre-defined behaviour each time, and the Smart



Eye was more robust than the Tobii in detecting gaze. Driver eye contact seeking in the same trials as above may be seen in Figure 30a and 30b correspondingly.

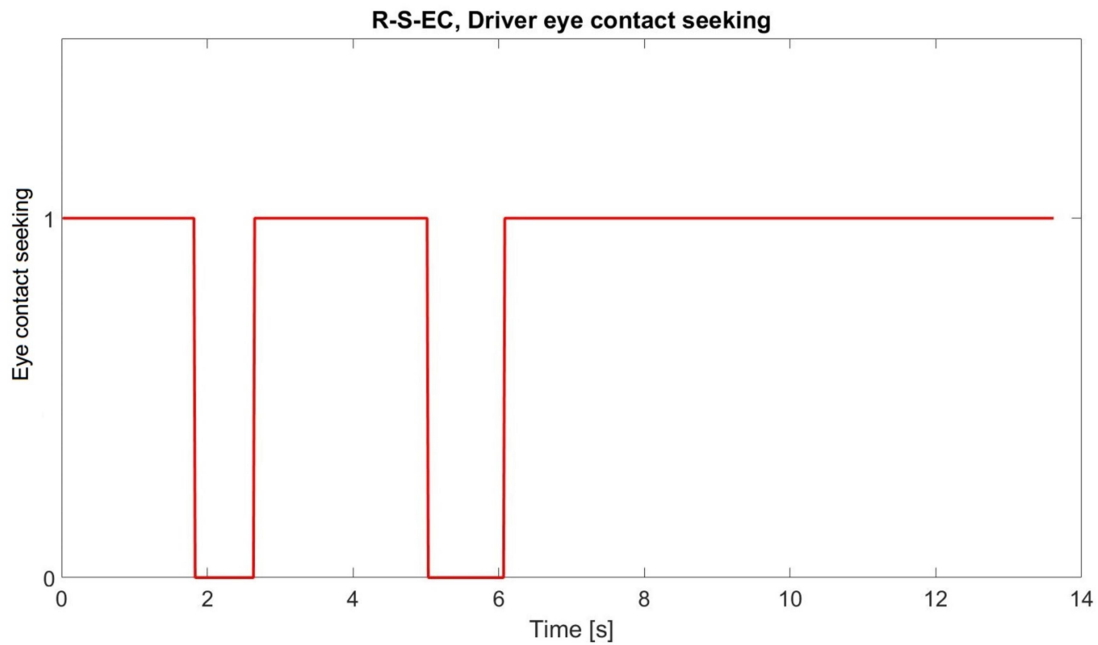


Fig. 30a – Driver eye contact seeking in the same standing trial as Fig. 29a, involving 3 repetitions, each comprised of 1 visit

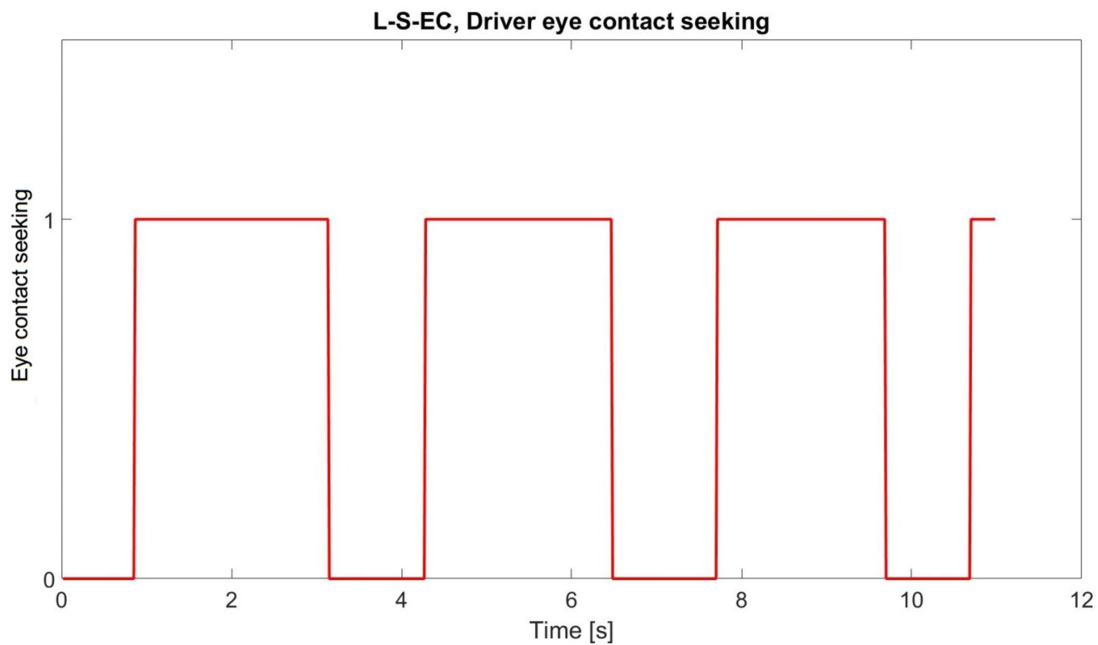


Fig. 30b – Driver eye contact seeking in the same standing trial as Fig. 29b, involving 4 repetitions, each comprised of 1 visit

These findings are interesting because they reveal that mismatch may be common between pedestrians' eye contact seeking and drivers' eye contact seeking (since it occurred for most participants in the standing trials). In Figures 31a and 31b, eye contact is shown to be the overlap of driver and pedestrian eye contact seeking.



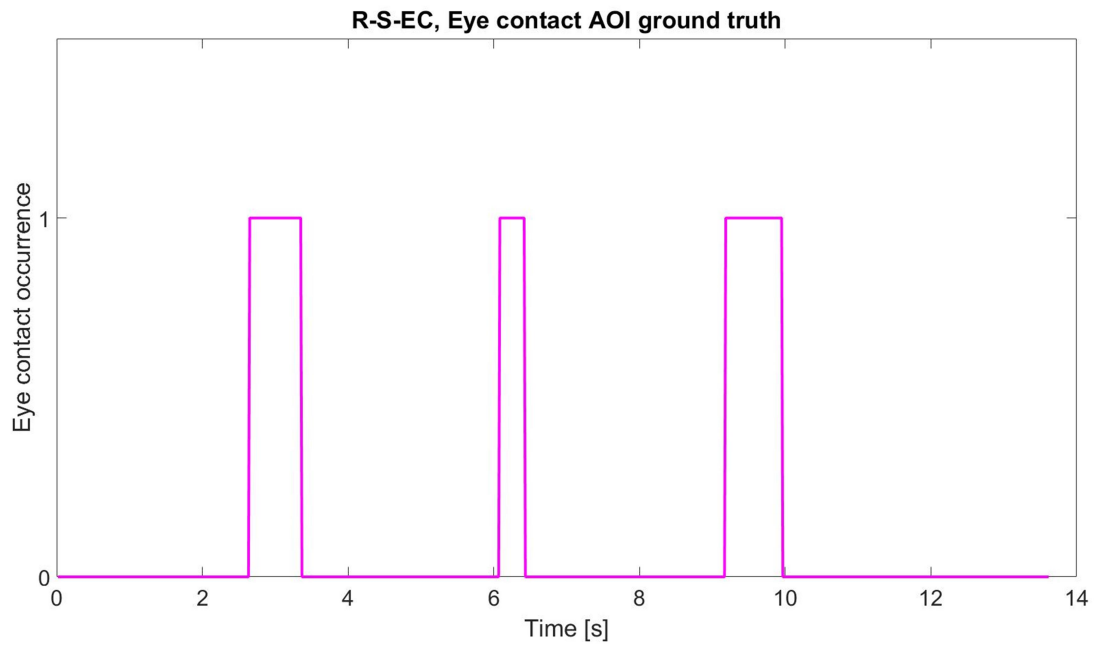


Fig. 31a – Eye contact as the overlap between driver and pedestrian eye contact seeking in Fig. 29a and Fig. 30a

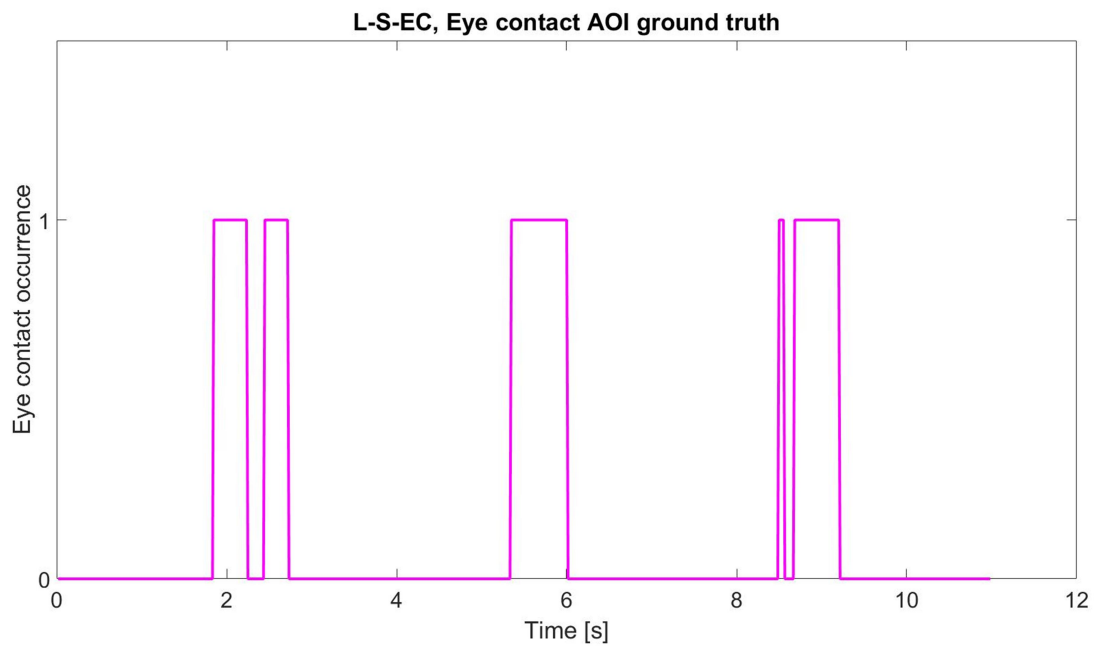


Fig. 31b – Eye contact as the overlap between driver and pedestrian eye contact seeking in Fig. 29b and Fig. 30b

**vii. Pedestrian dwell times and number of visits:**



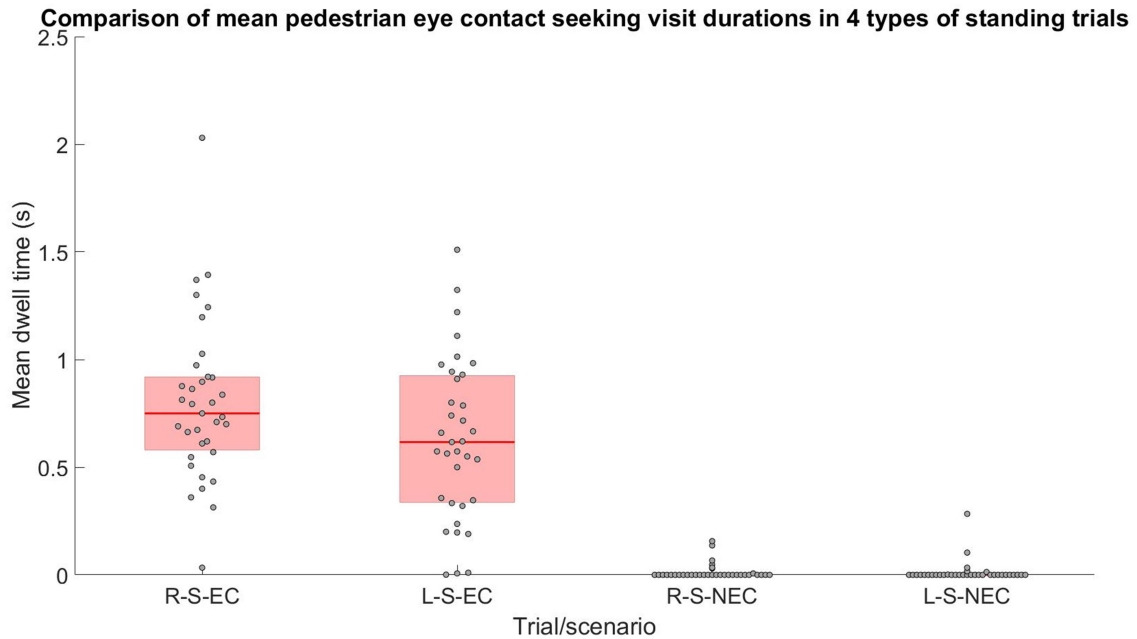


Fig. 32 – Pedestrian dwell times for standing trials

The Wilcoxon signed-rank test showed significant differences in the dwell time medians ( $p=0.0319$ ) between right and left sides in the standing trials at the 5% significance level (see Figure 32). However, no significant differences ( $p=0.8799$ ) in the number of eye contact seeking visits between the sides in the standing trials were found by the same test. These findings meant that pedestrian eye contact seeking visits lasted longer when pedestrians stood on the right of the car, but the number of eye contact seeking visits was the same for both sides.

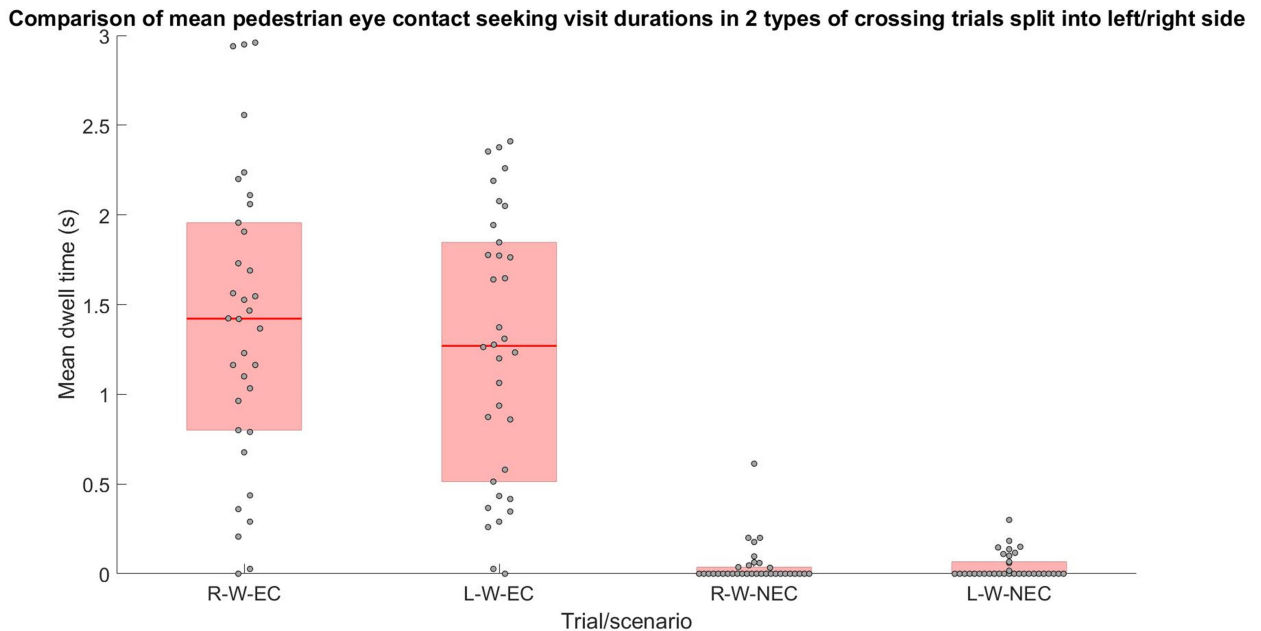


Fig. 33 – Pedestrian dwell times for crossing trials

Dwell times for both sides (i.e. right/left half-repetitions) of the crossing trials were compared, and the Wilcoxon signed-rank test showed no significant differences in the medians ( $p=0.1796$ ) between right and left sides at the 5% significance level (see Figure 33). Further, no significant differences ( $p=0.9913$ ) in the number of eye contact seeking



visits between the sides in the crossing trials were found by the same test. These findings meant that the duration of pedestrian eye contact seeking visits and the number of eye contact seeking visits were the same for both sides.

**viii. Visits in driver and pedestrian eye contact seeking for the crossing trials:**

The number of eye contact seeking visits by the pedestrian in the crossing trials was high, as expected. Values ranged from 7 to 11 per side (i.e. right or left), with some even as high as 17. The six crossings (three from the right and three from the left) performed by the pedestrian were seen as groups of closely spaced visits. The pedestrian's constant motion was the main reason for the fragmentation of his/her eye contact seeking into several visits; fixations did not remain close enough to each other for long when the pedestrian's position changed (Figure 34a). This observation also suggests that measured eye contact during the act of road crossing is broken up into small pieces, probably due to limitations in Tobii's eye-tracking technology in evenly tracking the crossing pedestrian's vestibulo-ocular movements to keep the driver as the visual target. The Smart Eye, on the other hands, proved to be capable of tracking the driver's smooth pursuit of the moving pedestrian, possibly because the driver's head moved very little compared to the pedestrian's (Figure 34b). Figure 34c shows eye contact as the overlap between driver and pedestrian eye contact seeking.

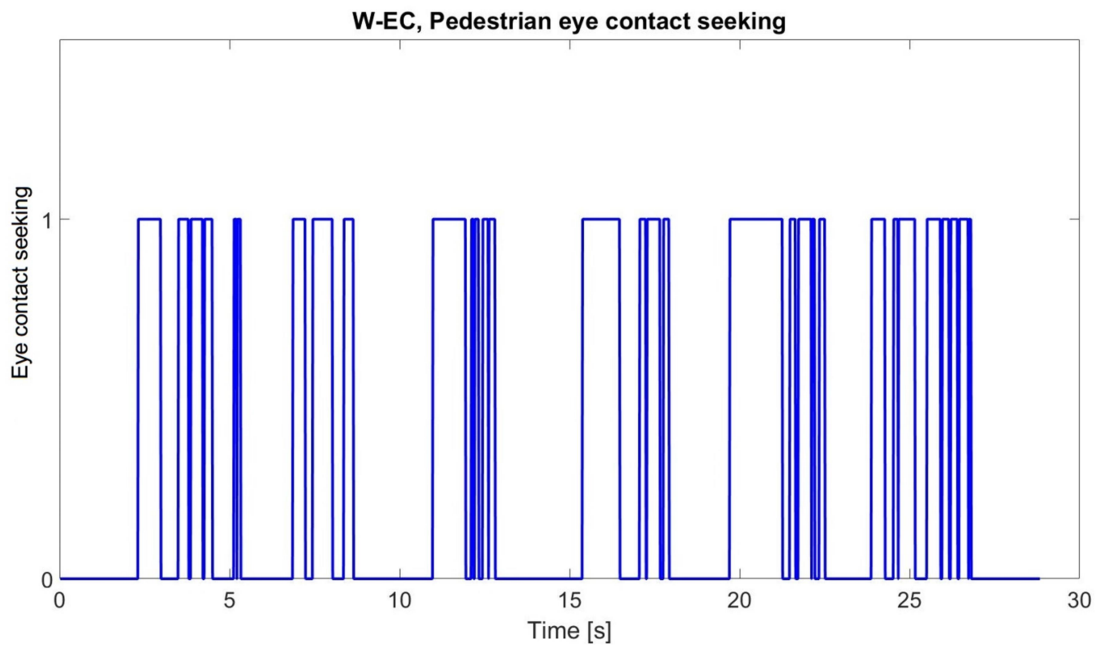


Fig. 34a – Pedestrian eye contact seeking in a crossing trial, involving multiple visits



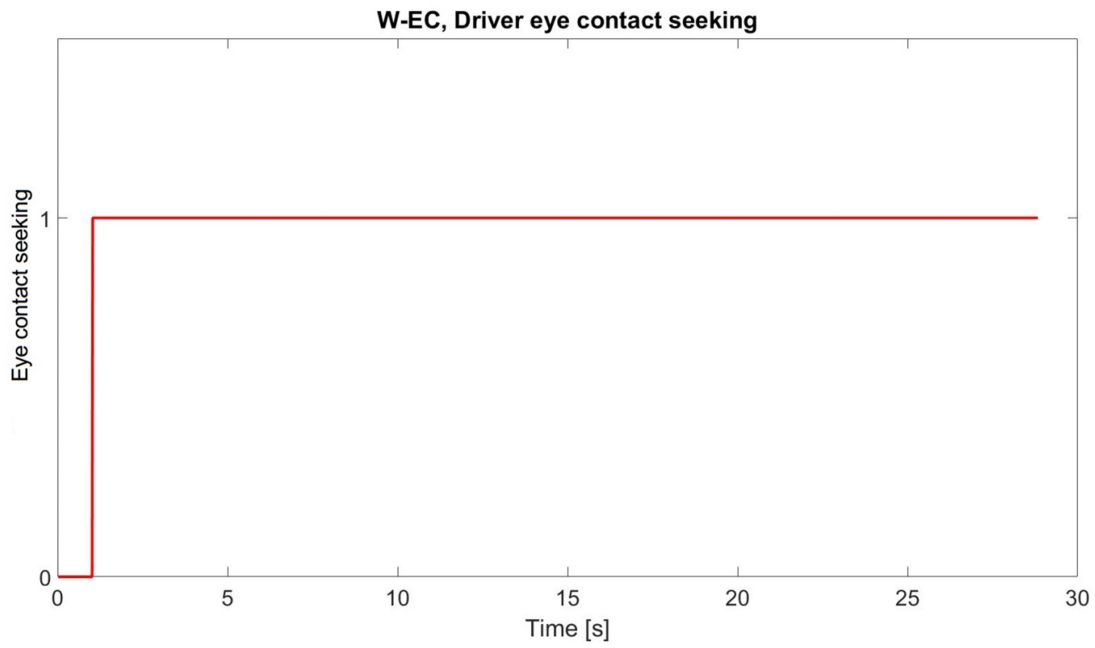


Fig. 34b – Driver eye contact seeking in the same crossing trial, involving one continuous visit

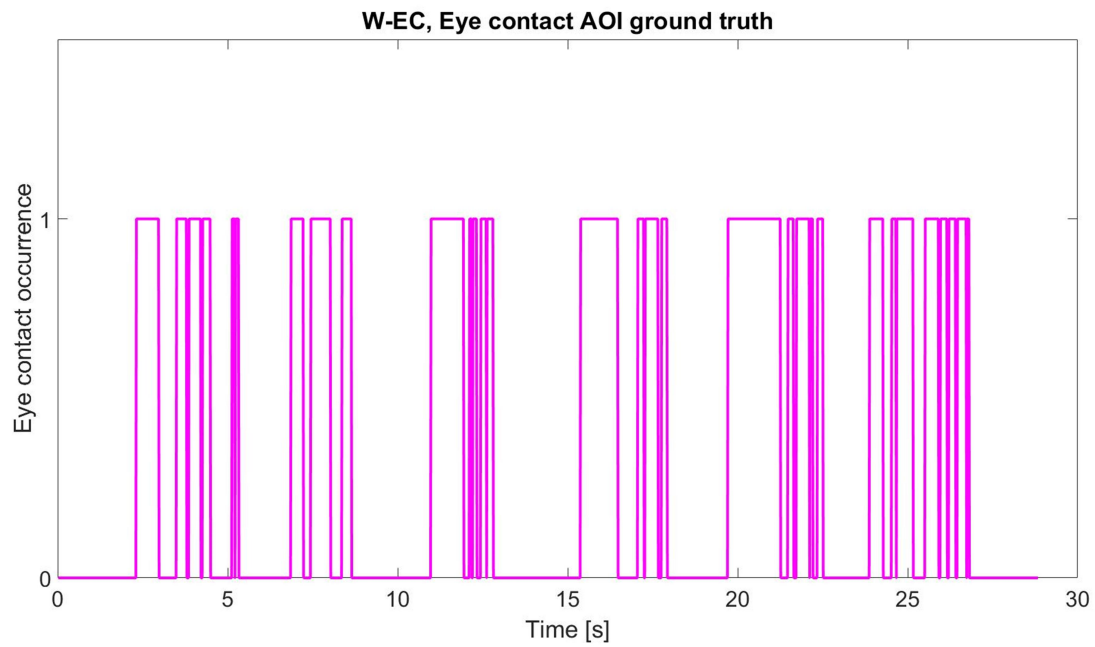


Fig. 34c – Fragmented eye contact in the same crossing trial

#### IV. Algorithm performance (Research Question 2):

The second research question of this thesis was:

- How accurate is the algorithm that operationalizes eye contact?

In relation to this, it was found that the actual 'eye contact threshold' and 'ground truth threshold' values were higher than the theoretical estimates. At the theoretical values of  $4.68^\circ$  and  $3.33^\circ$  respectively, there were almost no detected



instants of eye contact in either the algorithm's output or in the AOI-based ground truth, as may be seen below in Figures 35a and 35b. As explained before, the algorithm's output was obtained by combining the driver's gaze direction, the pedestrian's location, the pedestrian's eye direction, and the pedestrian's head direction to detect eye contact. The AOI-based ground truth used the driver's gaze direction and AOI hits from the pedestrian to detect eye contact.

To recap, the 'ground truth threshold' was the limit for the angle between the driver's actual gaze ray and the driver's ideal gaze ray in the ground truth. The driver's actual gaze ray represented where he was looking and his ideal gaze ray represented where he should have been looking. Thus, the driver's ideal gaze ray was simply the line joining the driver's eyes and the pedestrian's eyes. If the angle was lower than the 'ground truth threshold', the driver was seeking eye contact with the pedestrian, and otherwise, he was not. The 'eye contact threshold' was limit for the angle between driver's gaze ray and pedestrian's gaze ray in the algorithm. If the angle was lower than the 'eye contact threshold', the driver and pedestrian were looking at each other at the same time, and eye contact was occurring, and otherwise, no eye contact was happening.

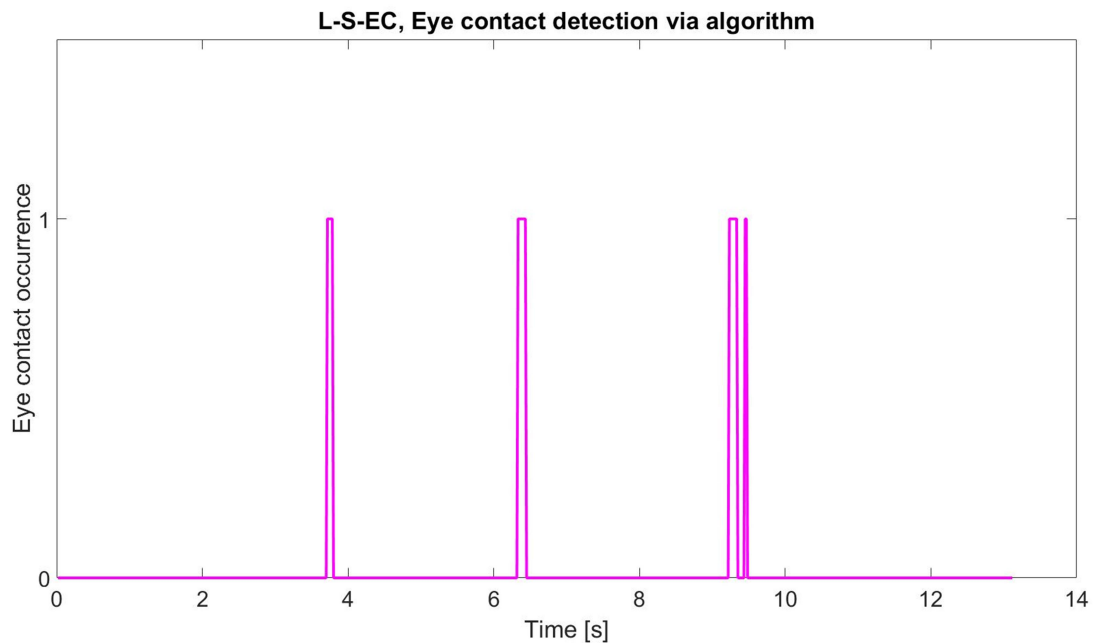


Fig. 35a – Eye contact binary plot (algorithm) for a standing trial using the theoretical eye contact threshold value  
This plot shows the three repetitions but almost no sustained detections of eye contact.



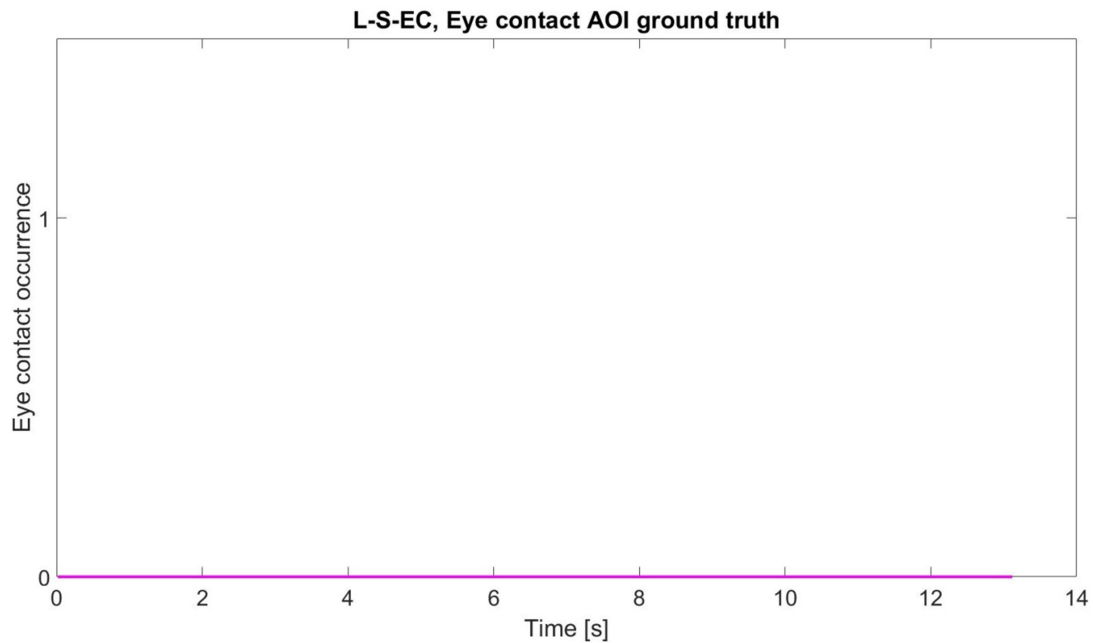


Fig. 35b – Eye contact binary plot (AOI ground truth) for the same standing trial using the theoretical ground truth threshold value  
This plot shows no detections of eye contact.

Possible reasons for this observation were the simplification of the kinematics in the algorithm (e.g. neglecting pedestrian head roll), and insufficient accuracy of the Tobii, and unexpected errors in measurement.

The actual values for the ground truth threshold were determined manually by visual inspection of the binary plots of eye contact (obtained from the AOI-based ground truth) for each participant. These values were the minimum angles at which the entire durations of the driver's eye contact seeking attempts in all standing/crossing trials were captured. They were found by increasing the ground truth threshold angles until no further increase in the duration of eye contact in any standing/crossing trial was observed. Since the pedestrian's eye contact seeking was in the form of AOI hits and did not change, the increase in eye contact duration was due to an increase in driver eye contact seeking duration.

The actual values of the ground truth threshold across participants for the standing/crossing trials are summarized in Table 7 below.

Table 7 – Summary of actual values of the ground truth threshold

Threshold	General scenario	Mean (°)	Std. Dev. (°)
Ground truth	Standing	23.63	11.96
Ground truth	Crossing	12.17	3.68

The ground truth thresholds for the crossing trials were lower and had less variability than those of the corresponding standing trials perhaps because they involved smooth pursuit (of the pedestrian) for the driver (using the Smart Eye) with gradual changes in gaze direction. The standing trials, on the other hand, had short periods of steady gaze but also periods of looking away from the pedestrian (which involved rapid head turns). It is possible that this behavior led to more frequent peripheral eye contact seeking (while turning towards and away from the pedestrian) and sharp changes in gaze direction that led to higher ground truth thresholds.

Bearing this information in mind, a common ground truth threshold of  $15^\circ$  (0.26 rad) was fixed for all trials and participants. Not only was this value close to the average of the mean ground truth thresholds of the crossing and standing trials as shown in Equation 6,



$$M = \frac{(23.63 + 12.17)}{2} = 17.9^\circ$$

Eq. 6 – Mean ‘M’ of actual ground truth thresholds for standing and crossing trials

it also subtended an arc of approximately 2 m when the pedestrian was at a maximum distance of 8 m from the driver as can be seen in Equation 7

$$a = r * \theta = 8 * 0.26 = 2.08 \text{ m}$$

Eq. 7 – Region of driver eye contact seeking in the ground truth

The symbol ‘a’ represents the arc, ‘r’ the radius, and ‘θ’ the angle

An approximation for small angles has been made, where an arc is assumed to be a straight line

This meant that in the ground truth, an allowance of approximately  $\pm 1$  m was offered between the pedestrian’s actual location and the driver’s gaze point in 3-D space, which was acceptable. Once the ground truth threshold was set, the eye contact threshold was varied between 0-180° in steps of 5°, hit rates and false positive rates calculated, and ROC curves plotted as shown below. These curves provided an idea of the algorithm’s capabilities at detecting eye contact instants.

The hit rate or true-positive rate was defined as the number of ‘eye contact’ instants correctly detected by the algorithm divided by the total number of ‘eye contact’ instants in the ground truth. The false positive rate was defined as the number of falsely detected ‘eye contact’ instants by the algorithm divided by the total number of ‘no eye contact’ instants in the ground truth.

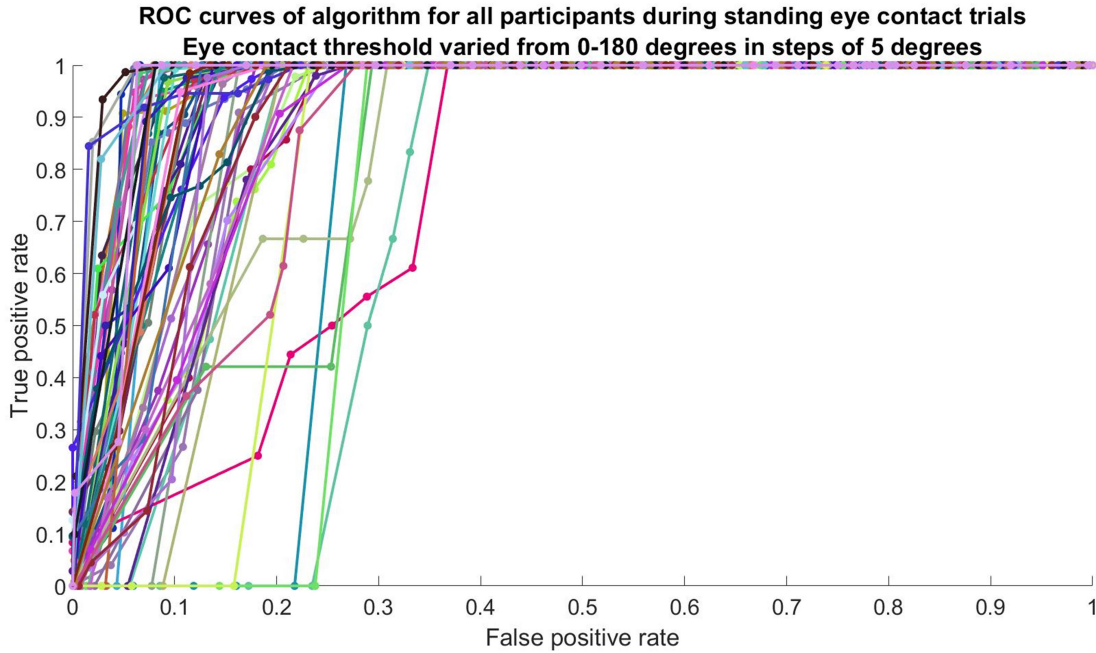


Fig. 36 – ROC curves depicting algorithm performance for standing trials with eye contact



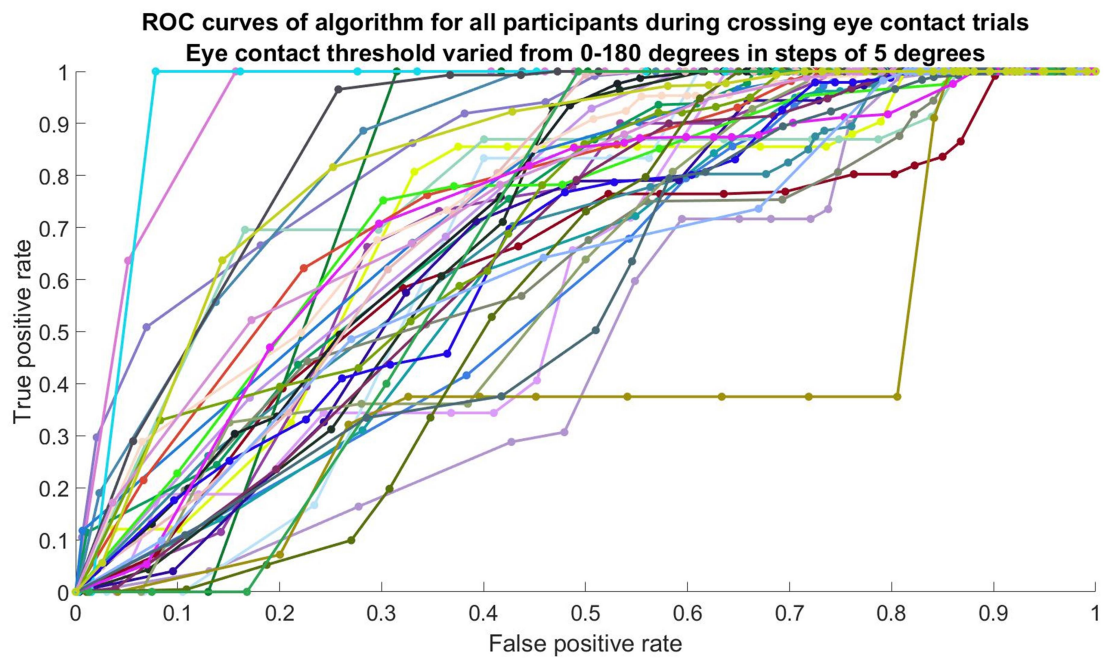


Fig. 37 – ROC curves depicting algorithm performance for crossing trials with eye contact

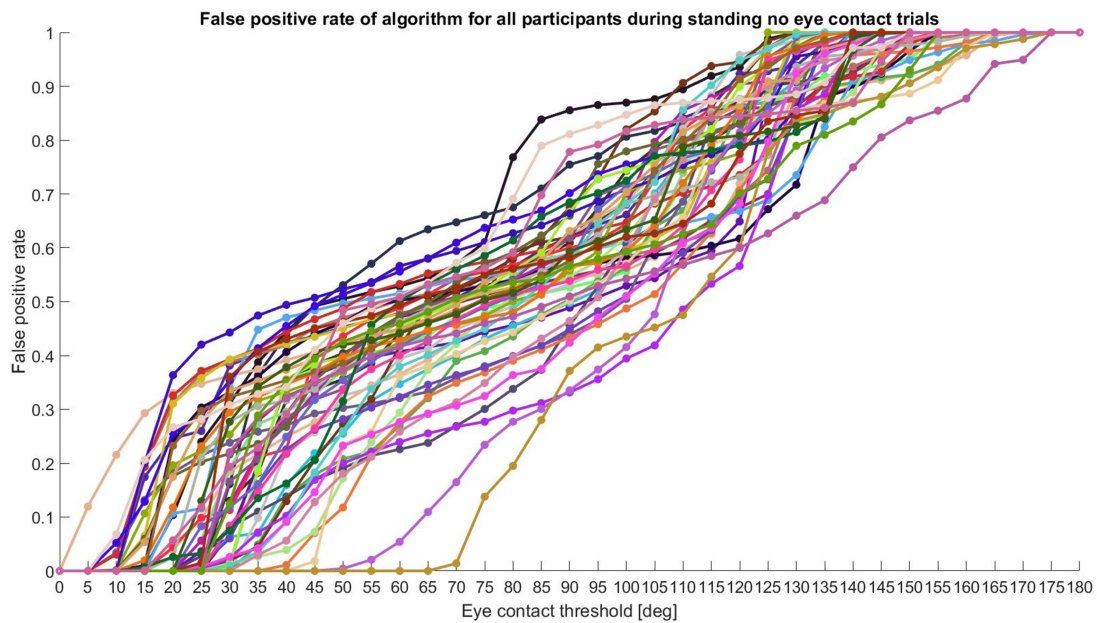


Fig. 38 – False positive rate (FPR) curves depicting algorithm performance for standing trials with no eye contact



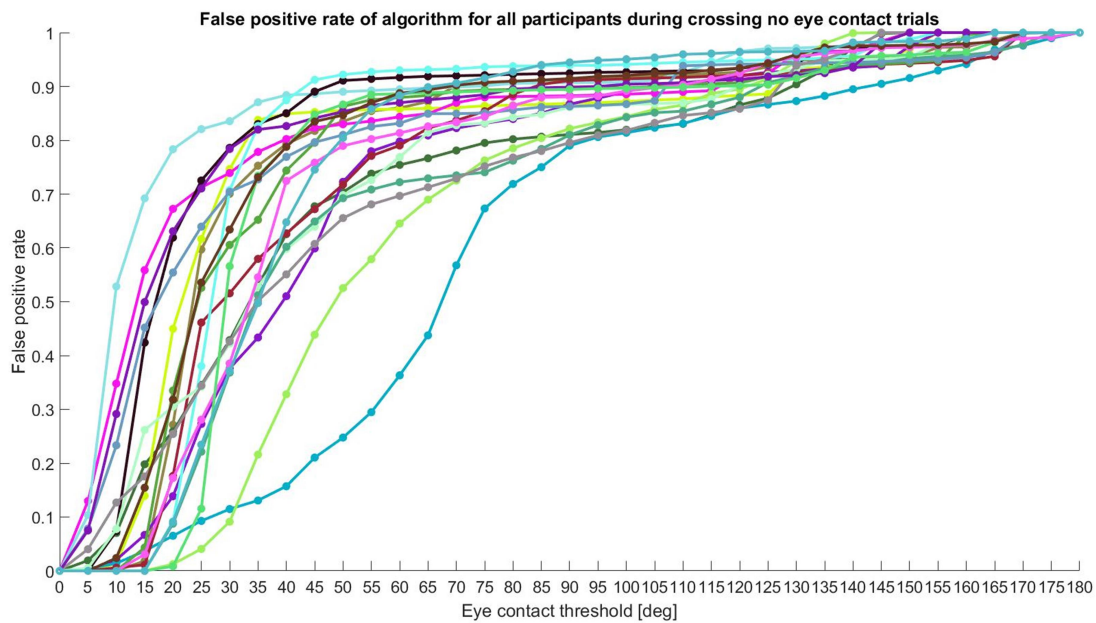


Fig. 39 – False positive rate (FPR) curves depicting algorithm performance for crossing trials with no eye contact

Clearly, the algorithm performed better in the standing scenarios than in the crossing ones. This was attributed to simpler static conditions versus relatively complex dynamic conditions. For example, pedestrian head roll likely played a smaller role in the standing trials than the crossing trials, so the negative effect on algorithm performance when it was neglected was greater for the latter scenarios.

The standing trials showed a high ratio of hit rate to false alarm rate for the eye contact scenarios (as shown in Figure 36) and a gentle increase in false positives with eye contact threshold (as seen in Figure 38). The crossing trials, on the other hand, showed a lower ratio of hit rate to false positive rate (shown in Figure 37), combined with a steep increase in false detections with eye contact threshold (see Figure 39). The area under the ROC curves was calculated as a measure of algorithm accuracy and displayed in Figure 40 and 41 below.

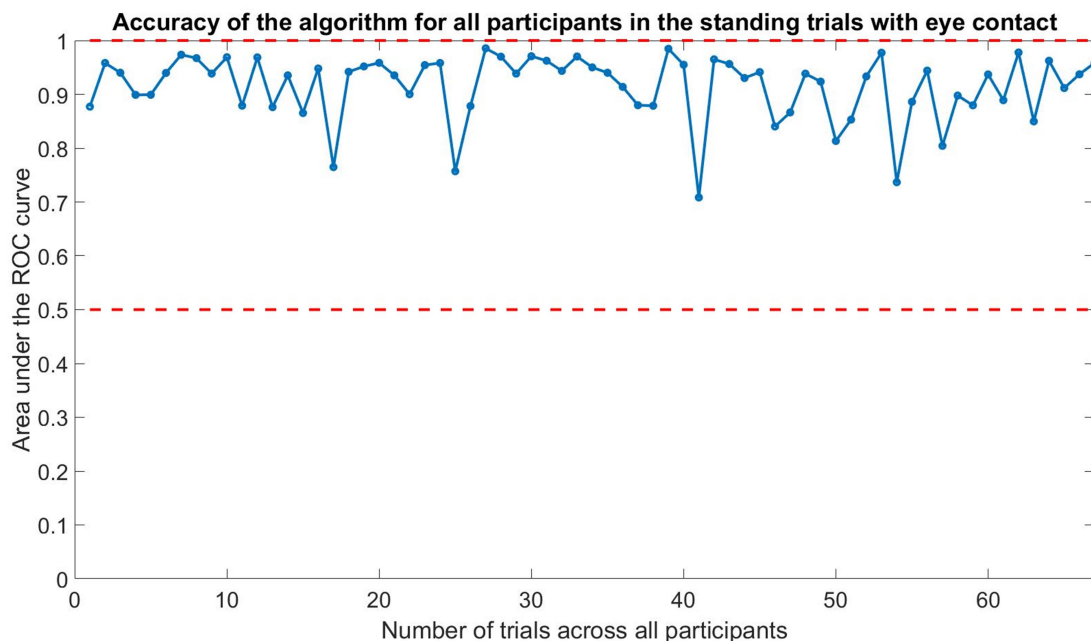




Fig. 40 – Areas under the ROC curves showing algorithm accuracy for the standing trials with eye contact

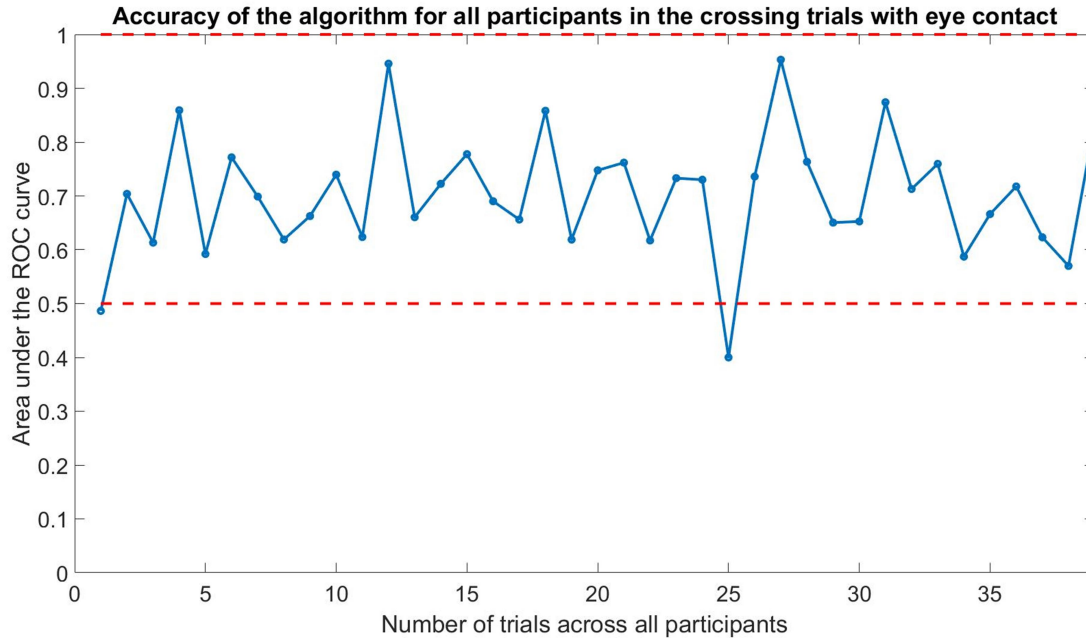


Fig. 41 – Areas under the ROC curves showing algorithm accuracy for the crossing trials with eye contact

In ROC curves, an area of 0.5 denoted the failure/inability of the algorithm to distinguish between eye contact and no eye contact. If the area was less than 0.5, then the algorithm was making inverse/opposite classifications. An area of 1 meant a perfect classification. The various categories of algorithm performance are shown below in Table 8 (Tape, 1990).

Table 8 – Accuracy classification using area under an ROC curve

Area under the ROC curve	Accuracy	Grade
0.9 – 1	Excellent	A
0.8 – 0.9	Good	B
0.7 – 0.8	Fair	C
0.6 – 0.7	Poor	D
0.5 – 0.6	Fail	F

Table 9 – Area under the curve (AUC) and TPR/FPR trade-off

General scenario	Mean area under the curve (AUC)	TPR/FPR trade-off
Standing	0.91	20°
Crossing	0.70	25°

As shown in Table 9, the mean area under the curve (AUC) for the standing scenarios showed an excellent accuracy at correctly detecting instants of eye contact and that of the crossing scenarios was fair. From the ROC curves, by visual inspection, the optimal trade-off between TPR and FPR was at eye contact thresholds of 20° (from a range of 15°–25°) for the standing trials and 25° (from a range of 20°–30°) for the crossing trials. In general, the optimal range for best results from the algorithm was 15°–30°.

The algorithm was also capable of classifying whole trials as ‘eye contact’ or ‘no eye contact’ using the percentages of eye contact instants within the trials themselves. From each recording, the trials were extracted and sorted in the descending order of their eye contact percentages. It was found that there was no universal percentage value that could reliably separate ‘eye contact’ trials from ‘no eye contact’ trials. The reason for this was a combination of poor accuracy in estimating gaze rays for the algorithm, and wide variability in successful eye contact making between participants. Therefore, the sorted trials were separated into two halves, with the upper half of higher percentages



classified as 'eye contact' trials and the lower half as 'no eye contact' trials. This exercise was repeated for the optimal range of 'eye contact threshold' values (15°–30°) shown earlier and the classification results compared with the ground truth that was instructions to participants before each trial. If all the classified trials matched the ground truth, the corresponding box for the participant and 'eye contact threshold' was ticked as shown below. The trial classification performance is displayed below in Tables 10 and 11.

Table 10 – Algorithm performance at classifying all standing trials correctly as 'Eye Contact' or 'No Eye Contact'

Participant number	Ability to separate standing EC/NEC trials at eye contact threshold			
	15°	20°	25°	30°
1	✓	✓	✓	✓
2				
3	✓	✓	✓	✓
4				
5			✓	✓
6			✓	
7		✓	✓	✓
8				
9 (No data)				
10				
11	✓			✓
12		✓	✓	✓
13		✓	✓	✓
14				
15				
16				
17	✓			
18			✓	✓
19	✓	✓	✓	✓
20				
21		✓	✓	✓
22				
23				
24				
25				
26				
27				
28		✓	✓	✓
29				
30			✓	✓
31				

Clearly, the 'eye contact threshold' value for best trade-off between TPR and FPR in the standing scenarios was not the same as that for most correct classifications of trials. The latter was found to be 25°, with preference being given to lower of the two angles with the most successful classifications. The most probable reason for the number being rather low (i.e. correct classifications of all trials for only 12 out of 30 valid participants) was that the simplified kinematics in the algorithm led to comparable eye contact percentages (false positives) in the 'no eye contact' trials, making distinguishing them from the 'eye contact' trials more difficult. Appendix E contains a sister table with the trial eye contact percentage values for the standing trials.

Table 11 – Algorithm performance at classifying all crossing trials correctly as 'Eye Contact' or 'No Eye Contact'

Participant number	Ability to separate crossing EC/NEC trials at eye contact threshold			
	15°	20°	25°	30°
1	✓	✓		
2			✓	
3	✓	✓	✓	✓
4	✓	✓	✓	✓
5	✓			
6	✓	✓	✓	✓



7	✓	✓	✓	✓
8				
9 (No data)				
10	✓			
11	✓	✓		✓
12	✓	✓	✓	✓
13				
14	✓	✓	✓	✓
15				
16	✓	✓	✓	✓
17	✓	✓	✓	✓
18	✓	✓	✓	
19				
20	✓	✓	✓	✓
21	✓	✓	✓	✓
22			✓	✓
23				
24	✓	✓	✓	
25			✓	✓
26				
27				
28	✓	✓	✓	✓
29		✓	✓	
30	✓	✓		
31	✓	✓	✓	✓

For the crossing trials too, the ‘eye contact threshold’ for most correct trial classifications was 15°, which was different from the value for optimal TPR/FPR trade-off. The reason for the best classification being achieved at the lowest of the four threshold values was that the negative effect of simplified kinematics grew as the angle increased (i.e. more false positives), especially in the crossing trials. At 15°, the algorithm correctly classified the crossing trials of 19 out of 30 valid participants. As to why the classification was overall better in the crossing scenarios than the standing ones, it was because the algorithm only needed to separate 2 types of the former, while it needed to separate 4 types of the latter. Appendix E presents a similar table showing the percentages of eye contact in the crossing trials.

Overall, the performance of the eye contact detection algorithm was fair bordering on poor, taking into account the classification performance of instants within trials and of whole trials themselves. Its performance may be summarized as follows:

- In the standing trials involving eye contact, the algorithm distinguished between almost all instants of eye contact and no eye contact (Figure 36).
- In the standing trials involving no eye contact, the algorithm generated many false positives of eye contact instants, which in turn led to many trial misclassifications (Table 10).
- In the crossing trials involving eye contact, the algorithm detected most instants of eye contact but at the cost of many false positives (Figure 37).
- In the crossing trials involving no eye contact, the algorithm had fewer false positives of eye contact instants, which in turn led to the correct classification of most trials (Table 11).



## V. Reconstruction of interaction scenario (Research Question 3):

The third research question of this thesis was:

- How is it possible to use two eye-trackers with inertial measurement units (IMUs) and pedestrian recognition in a Toyota Prius car to reconstruct the entire driver-pedestrian interaction through a 3D animation?

As an answer to that, it was demonstrated that eye contact between a driver and a pedestrian could be recreated and visualized using the gaze rays or gaze direction vectors in a 3-D animated plot. In the isometric view, the animation plotted the car as a cuboid, and the heads of the driver and pedestrian as red (driver) and blue (pedestrian) spheres labelled as 'D' and 'P' respectively, and their estimated 3-D gaze points also as spheres. Their gaze directions were plotted as red/blue lines joining the centre of their heads to the corresponding gaze points. While the gaze direction vectors were accurately known (at least for the Smart Eye), the gaze points in 3-D space were not, because a measurement of the intersection of gaze rays from the person's left and right eyes was required. This data was available but unreliable, since neither eye-tracker produced accurate readings of 3-D gaze point beyond short distances of 1.5 m (Borys, 2019). This fact was attributed to portable eye tracking technology still being in its infancy and not yet suited for studying medium- to long-range activities. Therefore, the gaze rays in the animation were terminated at a distance equal to the distance between the driver and pedestrian, and the ends were assumed to be the 3-D gaze points. The screenshots in Figures 42a to 47 are from the 3-D animations of different trials from different participants and illustrate how the driver-pedestrian interaction scenario was recreated.

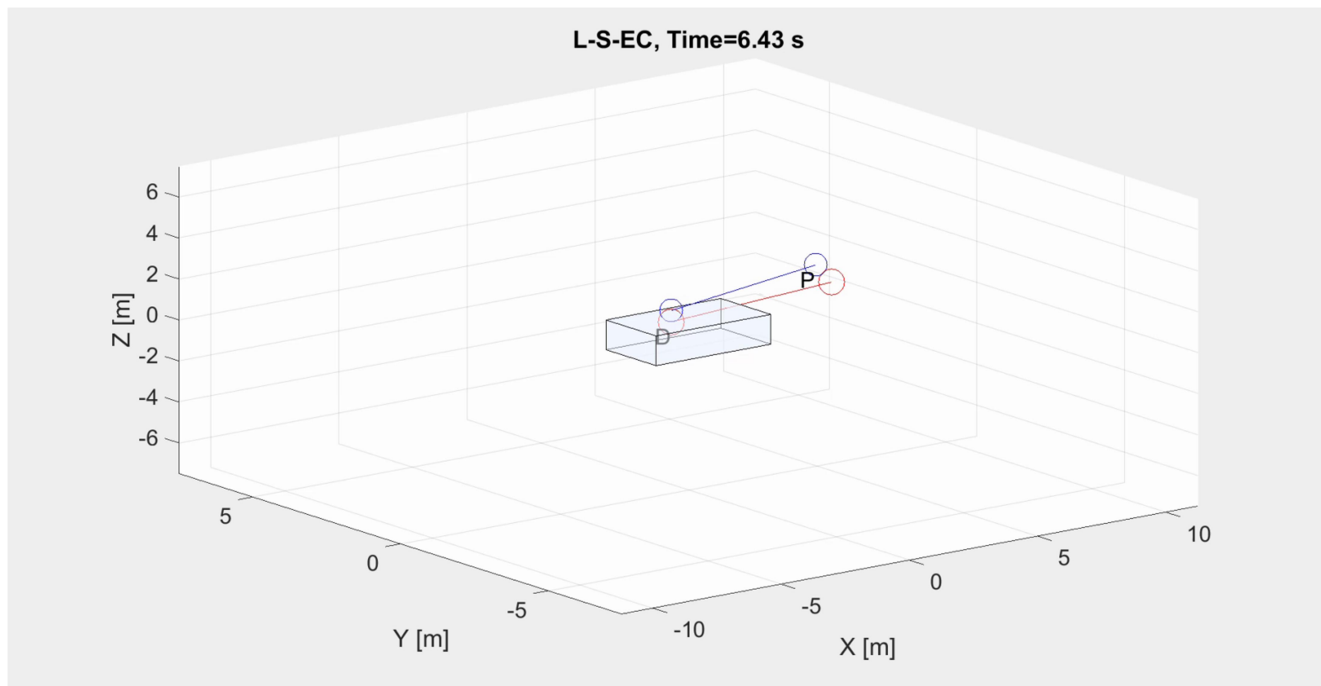


Figure 42a – Isometric view of a standing trial showing a moment when driver and pedestrian were looking directly at each other



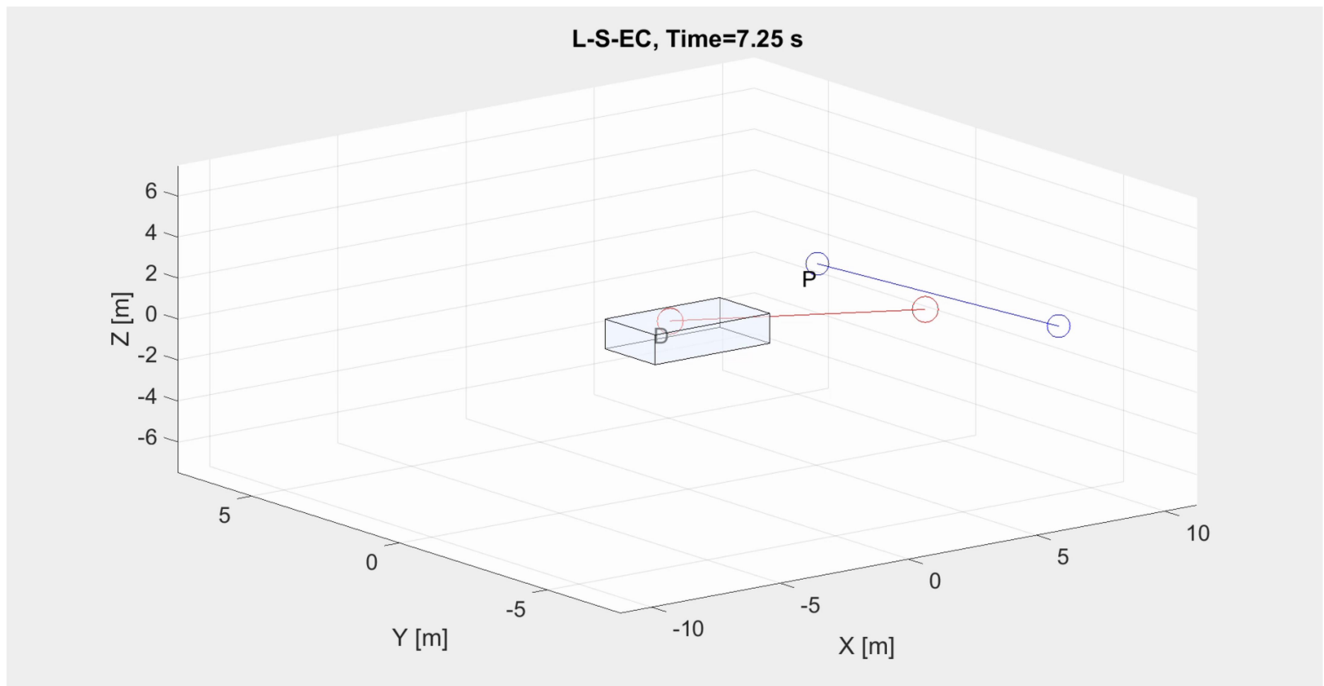


Figure 42b – Isometric view of a standing trial showing a moment when driver and pedestrian were not looking at each other

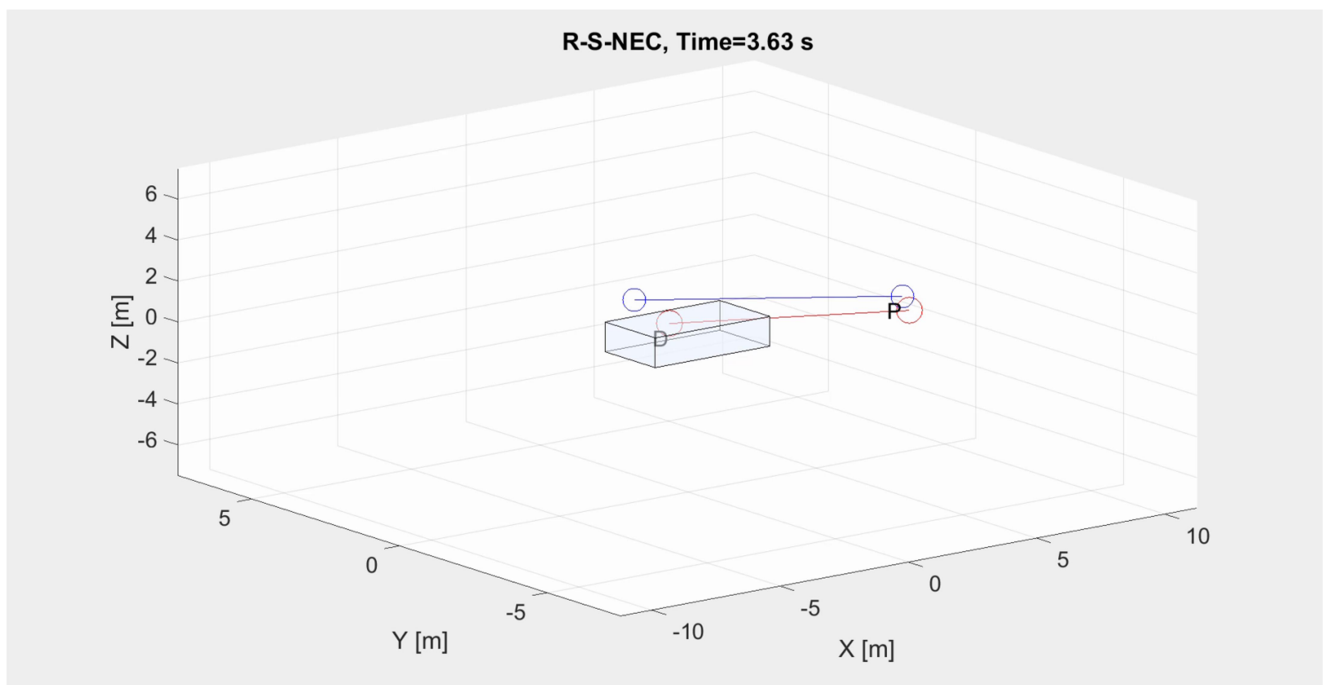


Figure 43 – Isometric view of a standing trial showing a moment when the driver was looking at the pedestrian but the pedestrian was probably looking at the left side-view mirror of the Prius



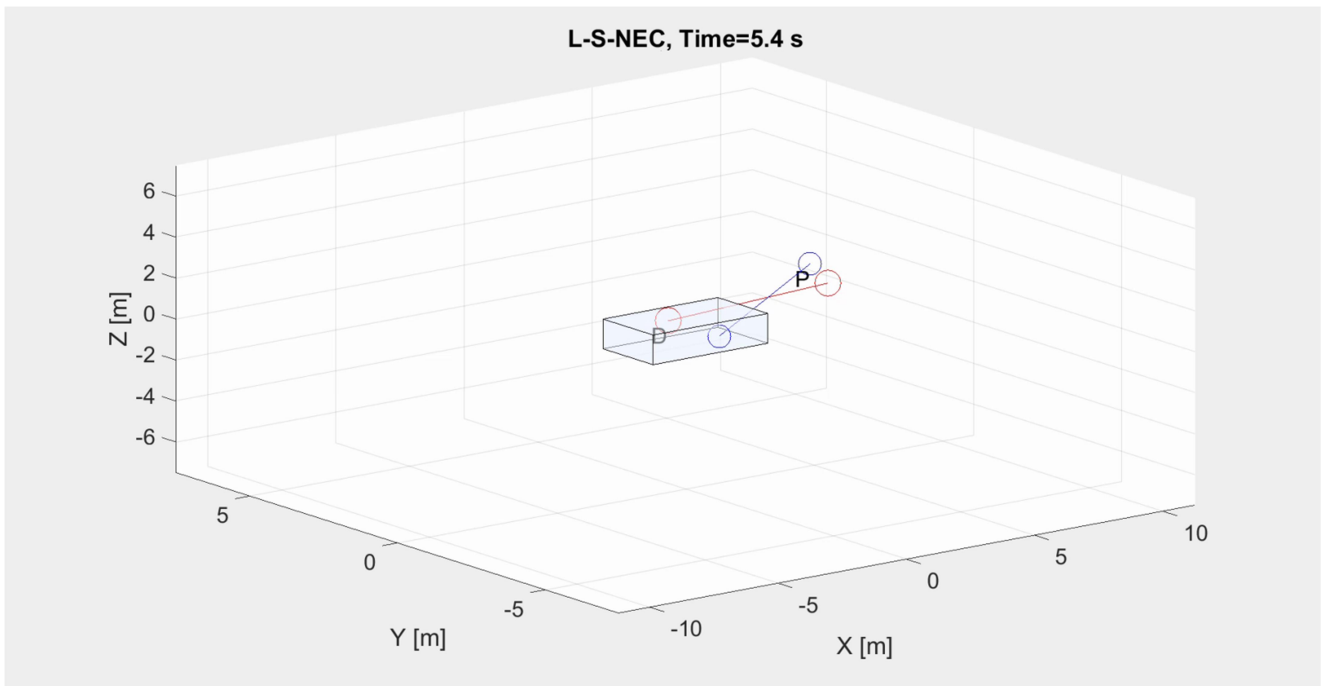


Figure 44 – Isometric view of a standing trial showing a moment when the driver was looking at the pedestrian but the pedestrian was probably looking at the number plate of the Prius

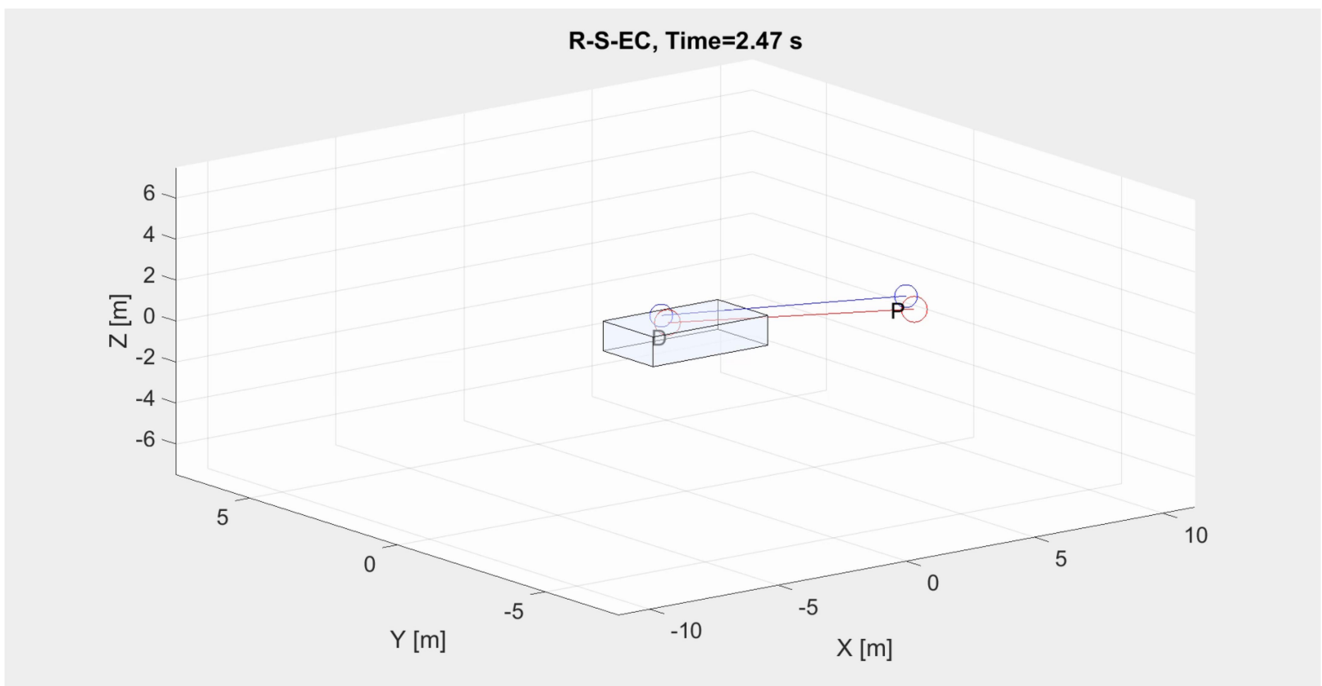


Figure 45a – Isometric view of a standing trial showing a moment when driver and pedestrian were looking directly at each other



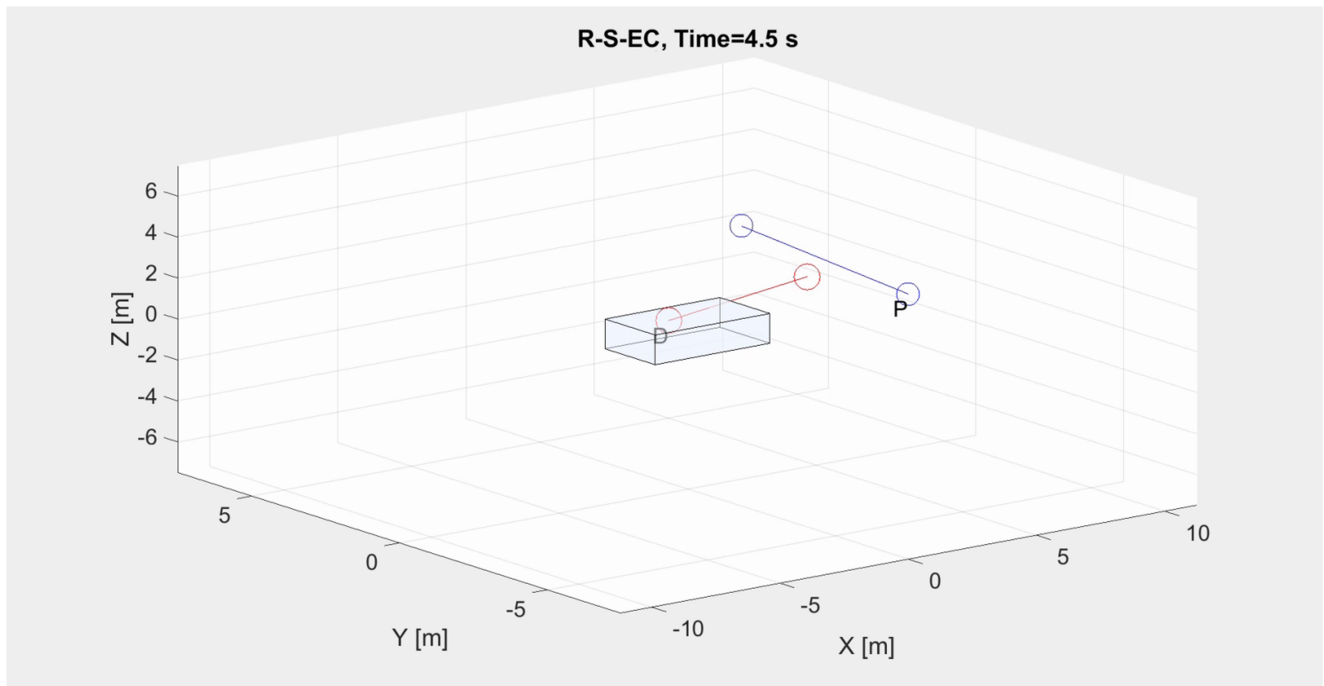


Figure 45b – Isometric view of a standing trial showing a moment when driver and pedestrian were not looking at each other

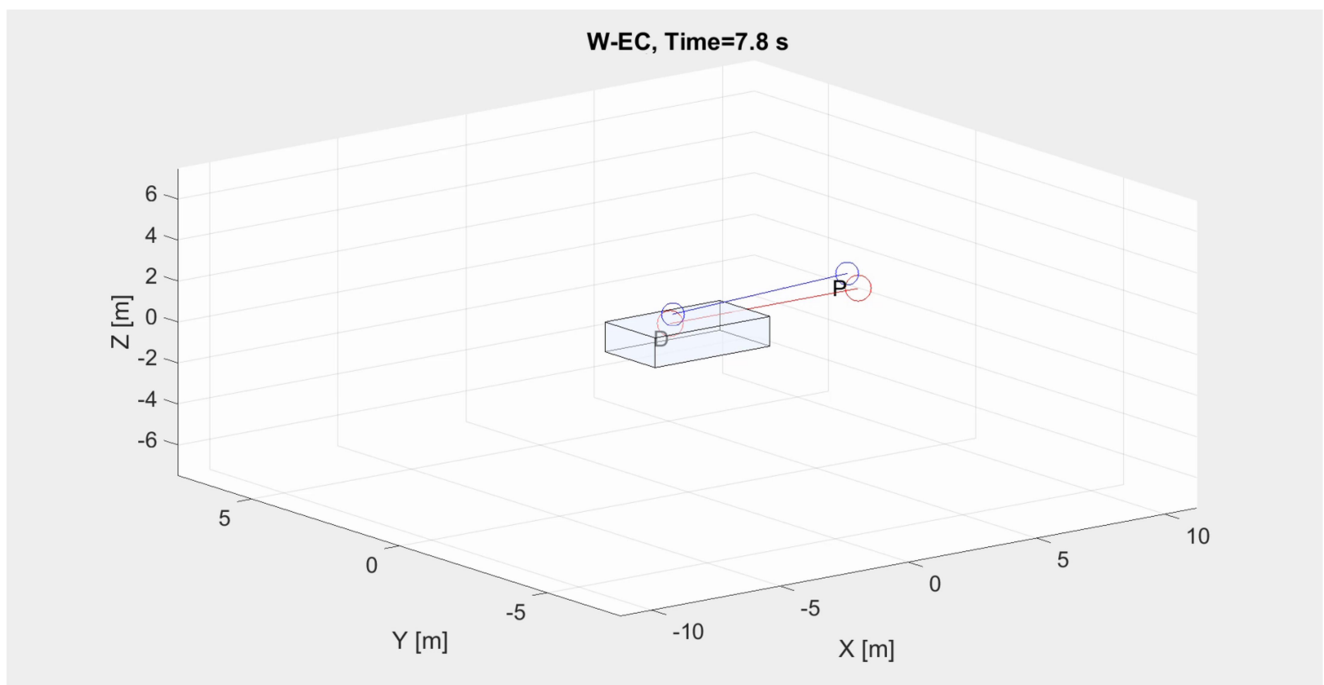


Figure 46 – Isometric view of a crossing trial showing a moment when driver and pedestrian were looking directly at each other and the pedestrian was crossing from the left side



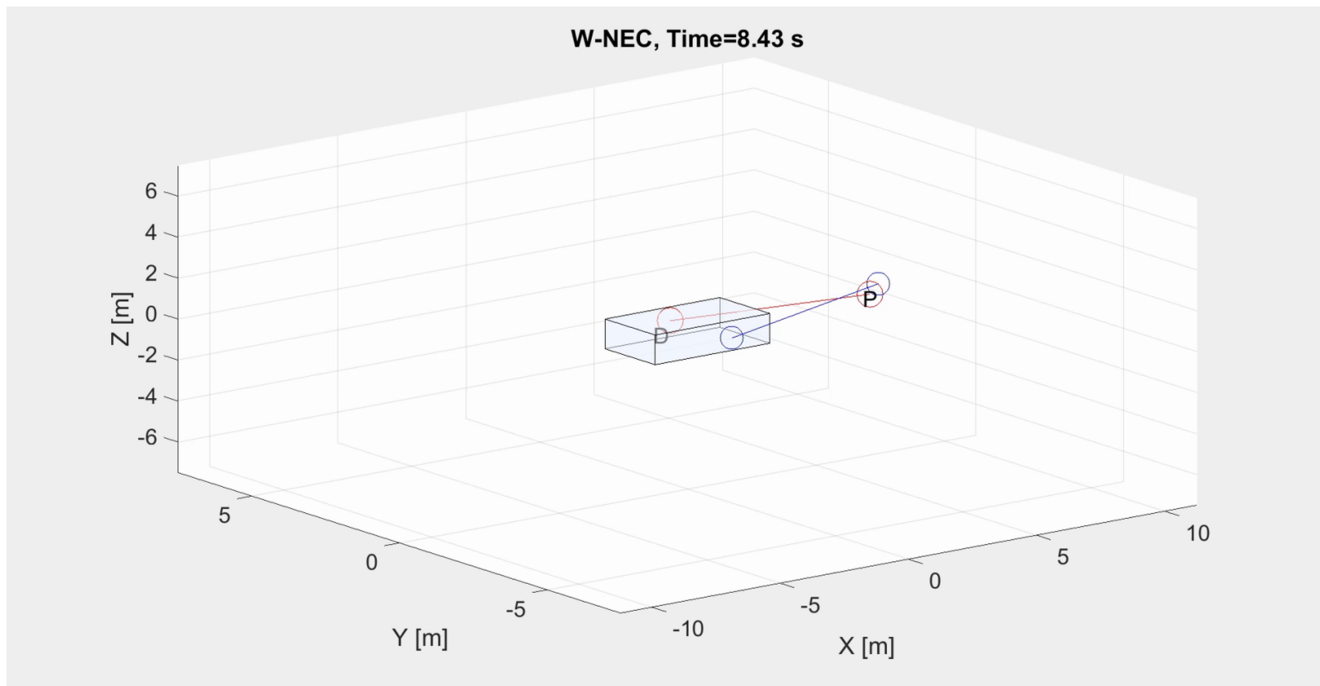


Figure 47 – Isometric view of a crossing trial showing a moment when the driver was looking at the pedestrian but the pedestrian was probably looking at the number plate of the Prius, and the pedestrian was crossing from the left side

Of the three orthographic views of the 3-D plot, the top view best illustrated the gaze rays of the two parties and the occurrence of eye contact. The primary reason for this was that the height difference ( $Z$ ) between the heads of the driver and pedestrian was relatively small compared to its lateral ( $Y$ ) and longitudinal ( $X$ ) counterparts, allowing a sufficiently good idea of the gaze interaction to be obtained from just viewing the  $XY$  plane. Therefore, side and front views of the 3-D animation were not generated. An image of the top view of the Toyota Prius was scaled and superimposed on the plot for added realism and to provide perspective for the driver's position inside the vehicle. The screenshots in Figures 48 to 53 provide an idea of the top view of the interaction of different participants in different scenarios.

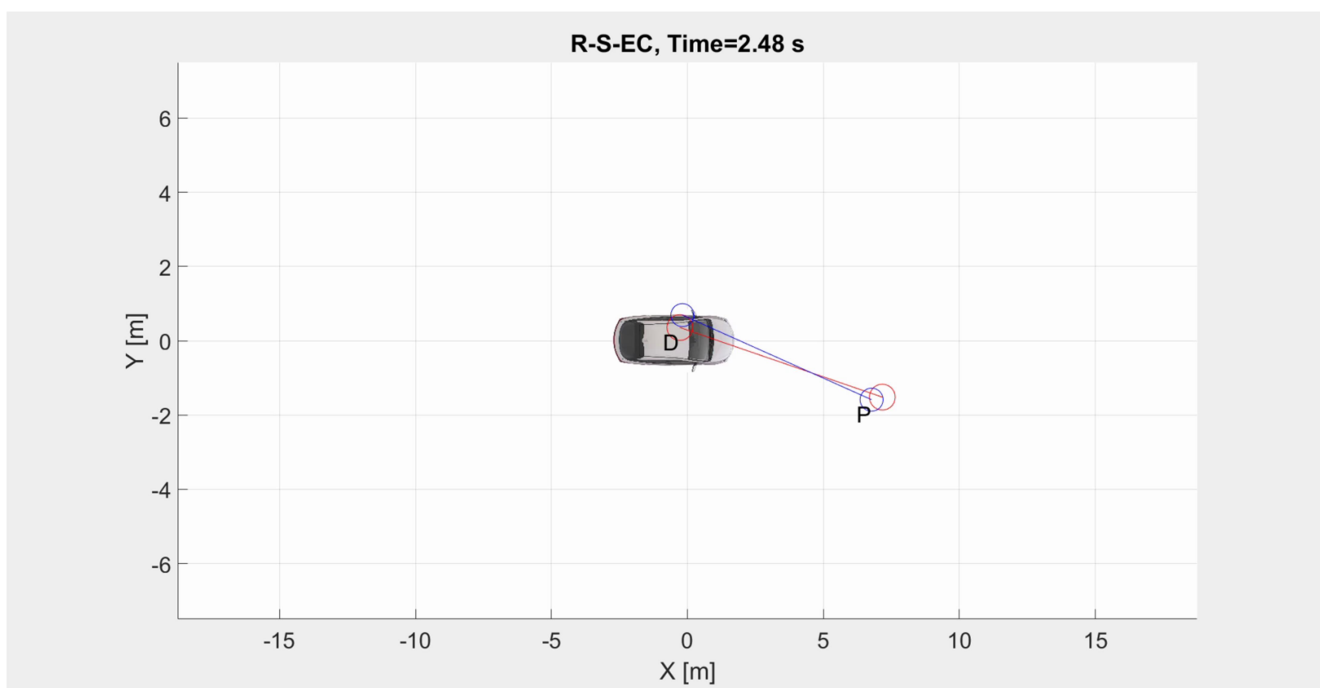


Figure 48 – Top view of a standing trial showing a moment when driver and pedestrian were looking directly at each other



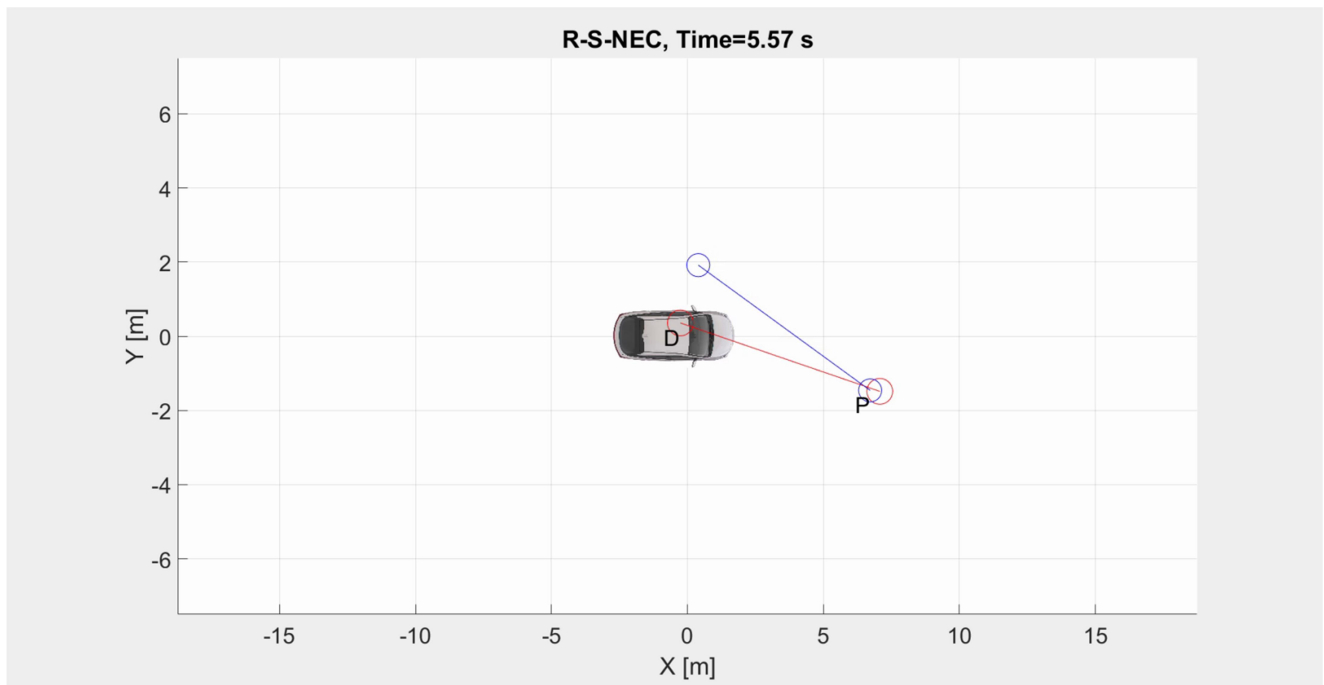


Figure 49 – Top view of a standing trial showing a moment when the driver was looking at the pedestrian but the pedestrian was not looking at the vehicle (contrary to the provided instructions)

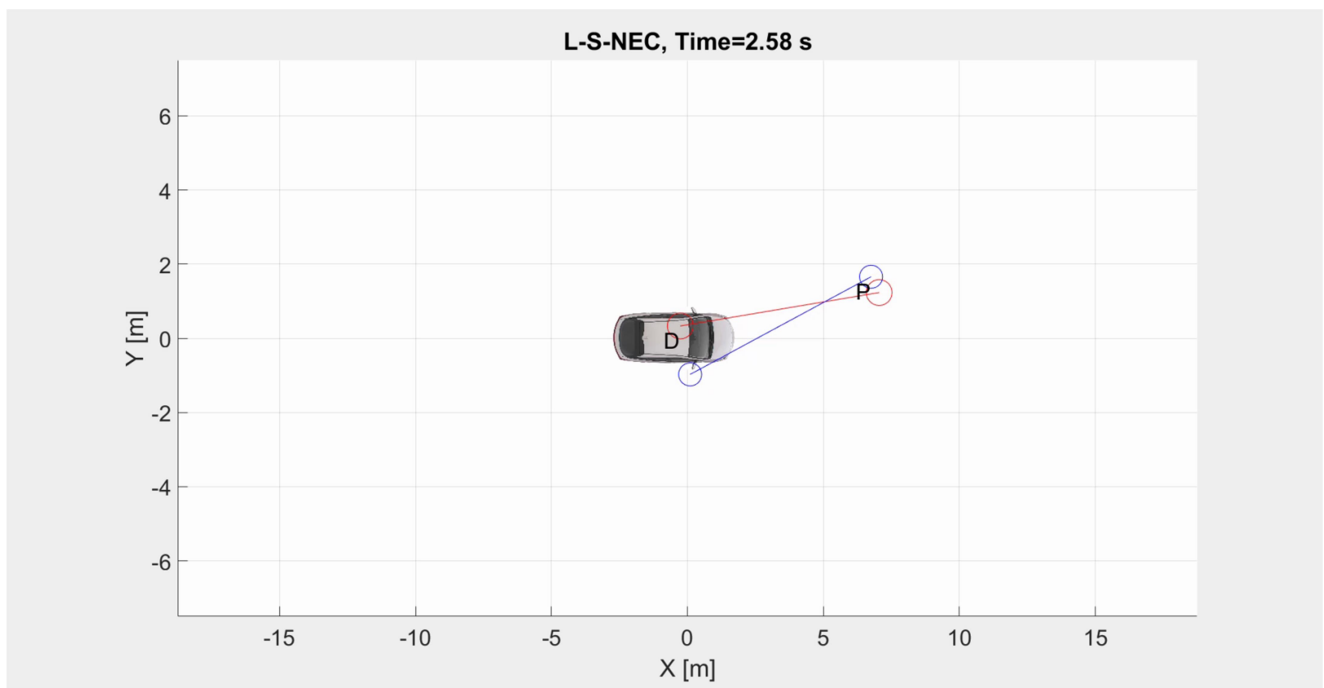


Figure 50 – Top view of a standing trial showing a moment when the driver was looking at the pedestrian but the pedestrian was probably looking at the right side-view mirror of the Prius



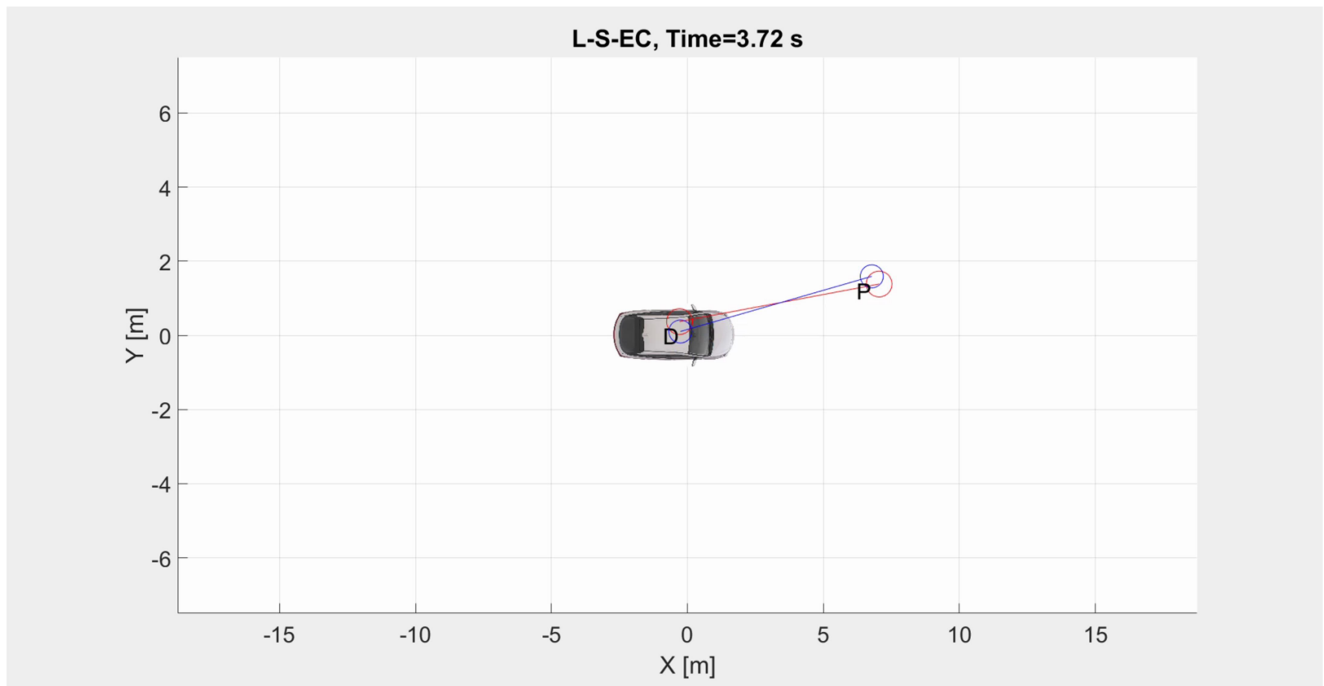


Figure 51 – Top view of a standing trial showing a moment when driver and pedestrian were looking directly at each other

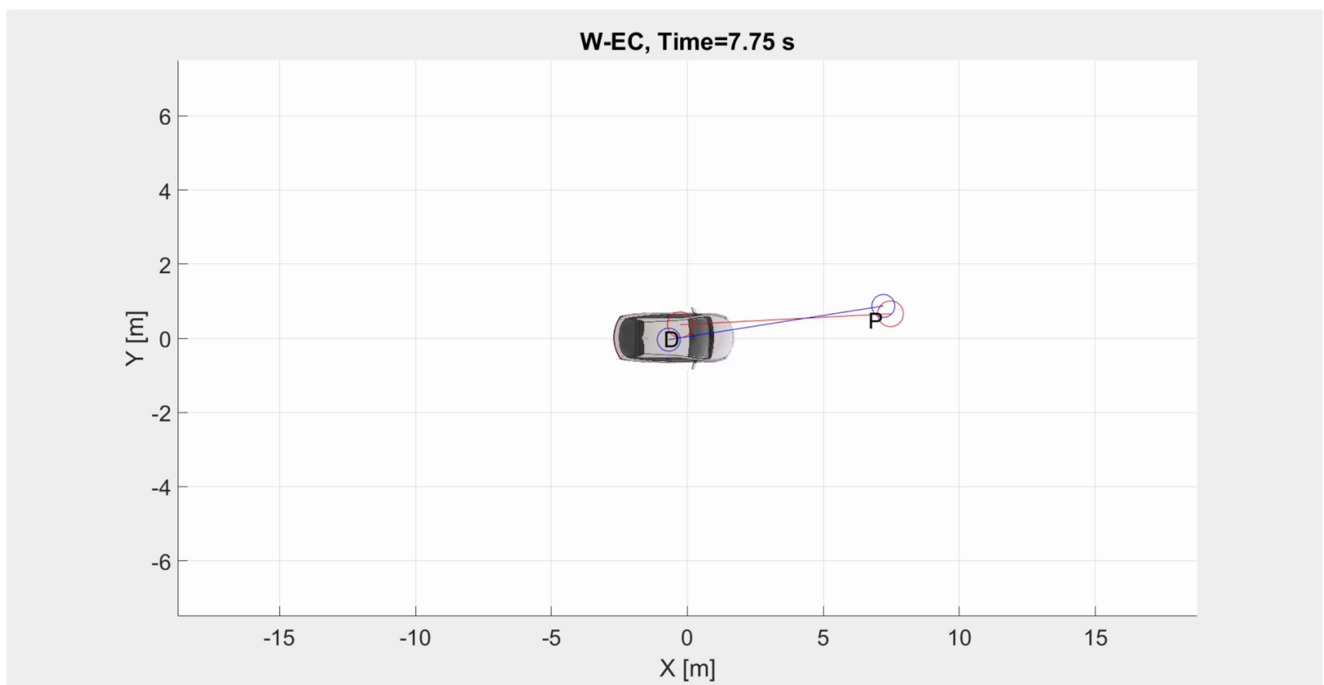


Figure 52 – Top view of a crossing trial showing a moment when driver and pedestrian were looking directly at each other and the pedestrian was crossing from the left side



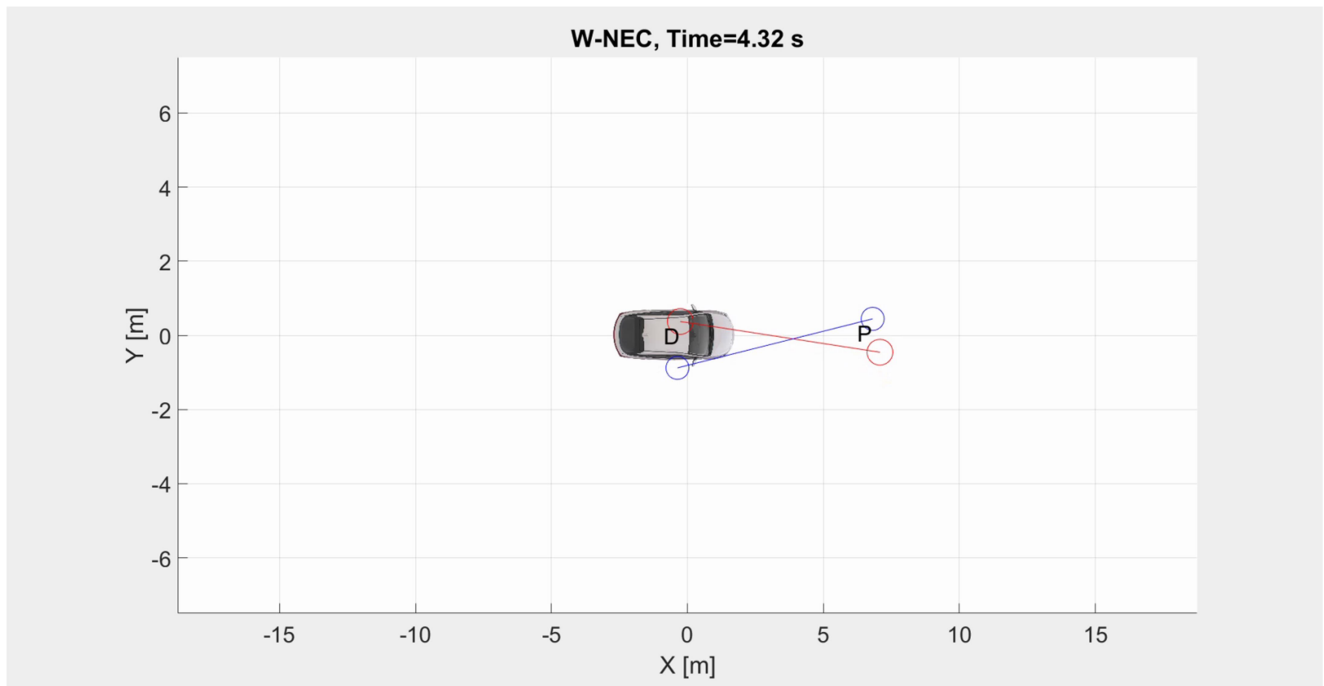


Figure 53 – Top view of a crossing trial showing a moment when the driver was not looking at the pedestrian and the pedestrian was probably looking at the right side-view mirror of the Prius while crossing from the right side



# 4 Discussion

The aims of this thesis were to define driver-pedestrian eye contact, develop an algorithm to automatically detect driver-pedestrian eye contact, measure the performance of the algorithm, and reconstruct the driver-pedestrian interaction as a 3-D animation. In the results section, it was shown that for a pedestrian standing on the curb, eye contact was typically made for 0.42-0.54 s, and for a pedestrian crossing the road, eye contact lasted for a maximum of 1.23-1.39 s. Algorithm detection of eye contact instants was achieved by finding the angle between the 3-D gaze rays of the driver and the pedestrian, and comparing it with an 'eye contact threshold'. Trials from the experiment were classified as 'eye contact' or 'no eye contact' by comparing the percentages of their eye contact instants within the duration of the trials. The overall performance of the algorithm was determined by comparing its results against the ground truths of pre-trial instructions and AOI-based eye contact detections. It was found that the classification performance was fair/poor. Individual instants of eye contact were more frequently correctly detected in the standing trials than the crossing trials, but whole trials were more frequently correctly classified if they were the crossing trials than if they were standing trials. Reconstruction of the driver-pedestrian interaction was achieved in the form of an animated 3-D plot showing the car, the driver, the pedestrian, and the gaze rays of both persons.

## *i. Generalizability of the eye contact results to public roads:*

It must be said that it is possible that the results of this experiment are not entirely representative of typical driver-pedestrian interactions on most Dutch roads. Aside from the reasons of the car being stationary, the road being imaginary, and the 'pedestrians' knowing the objective of the research, there were other factors at play. For example, as observed in the pilot study, eye color likely influenced gaze data quality, and cultural and gender differences among participants, their young ages, and the overrepresentation of some groups probably affected the measured durations of eye contact seeking. The reasons for the overrepresentation were threefold: (1) a greater number of males (than females) studying/working in the vicinity of the experiment, leading to a greater likelihood of males being recruited, (2) a higher number of Indian nationals (than other nationalities) as acquaintances to the researchers and studying/working near the venue of the experiment, leading to a greater chance of Indians being recruited, and (3) cancellations by would-be-participants, leading to last-minute replacements. Due to this bias, there were widely varying numbers of persons in each group upon classification based on eye color, gender and country. So, meaningful (i.e. high-power) statistical analyses could not be performed to compare the groups.

For a pedestrian standing on the curb, it is possible that eye contact seeking durations are shorter in countries where pedestrians usually have right of way (e.g. high-income countries like The Netherlands), as opposed to countries where that is not the case in practice (e.g. low-income countries like India). Shorter durations of eye contact seeking might also occur in internationals (from low-income countries) living in the Netherlands due to an acclimatization to receiving right of way. This difference in eye contact seeking duration may be due to greater levels of certainty in the pedestrians that vehicles will yield for them, thereby reducing the need for confirmation via eye contact. Almost all the participants in this experiment were residents of the Netherlands. Therefore, the durations of eye contact seeking (and therefore, eye contact durations, assuming unchanging driver gaze behavior) proposed for a standing pedestrian in the results section are conservative estimates and are likely to be higher in other parts of the world. Nevertheless, this claim would benefit from the results of repeat experiments performed in different nations.

For a pedestrian crossing a road, it is probable that eye contact seeking durations are shorter worldwide than the values proposed in the results, assuming the same road width. Pedestrians likely do not seek eye contact with drivers for the entire duration of their crossing (as they were instructed to do in the experiment). Once again, it is speculated that pedestrian eye contact seeking durations (and therefore, eye contact durations, assuming unchanging driver gaze behavior) while crossing are shorter in pedestrians from countries where pedestrians usually have right of way, than in places where they do not.

Since all the participants in the study were young people, it is also likely that they sought eye contact for less time than older people, due to the latter's reduced cognitive and visual abilities, and lower walking speeds (Fitzpatrick et al.,



2006; Montufar et al., 2007). This conjecture points to eye contact for a standing pedestrian lasting a little longer on the road compared to this experiment, since pedestrians will be from all age groups. The addition of older persons would conceivably also increase eye contact durations during the act of road crossing, but perhaps not to the extent of the values in this experiment.

That being said, the results of this experiment might be applicable to the streets of student towns like Delft, with its sizeable population of young internationals. This claim is slightly supported by the fact that more participants found the experiment realistic than those who did not. Further, since almost all participants reported that they were involved in the experiment and followed the instructions well (backed by video evidence from the Tobii and stereo camera), their behavior was probably at least partly realistic. Still, the high number of neutral responses to the question asking how realistic the experiment was suggests that the realism (of the experiment and participants' behavior) is lower than it is apparent from the questionnaire responses. Additionally, the fact that almost half of the participants were conscious of wearing the Tobii glasses and having their eyes tracked suggests their behavior may not have been entirely realistic.

The cause for the wide variation in responses to the question about the feeling of consciousness while using the Tobii is suspected to be due to: (1) participants wearing the Tobii glasses (when they normally do not wear glasses, or wear regular spectacles which made the Tobii feel unfamiliar in comparison), (2) the knowledge that their eyes were being tracked, and (3) a combination of both. This conjecture follows from the fact that participants were a mixed group comprising of persons who did not use any form of eyesight correction and people who wore contact lenses (who also wore spectacles at other times). The difference in the mean responses of these two groups suggested that awareness of their eyes being tracked influenced participants' feeling of consciousness and possibly also their gaze behavior during the experiment. Also, the fact that the mean response of the persons who wore contact lenses (i.e. those who were accustomed to wearing spectacles) was itself not very low suggested that wearing the Tobii glasses did not feel like wearing ordinary spectacles.

If more naturalistic results are desired i.e. results that could be directly applied to real-life scenarios, more accurate portable eye-trackers and sensors would be required in an outdoor experiment involving a moving vehicle (i.e. dynamic conditions). In the current experiment, the 'ground truth threshold' and 'eye contact threshold' values were higher than their theoretical estimates by a considerable margin. While the primary reason appears to be the use of simplified kinematics in the algorithm, it also suggests that the Tobii eye-tracker is not very accurate at measuring gaze direction at distances greater than 1.5 m (Borys, 2019). So it is likely that the tolerance in the Tobii measurements is greater than the quoted  $\pm 0.5^\circ$ , which would justify the need for more accurate eye-tracking solutions. Further, if dynamic conditions are present, GPS measurements of the car's position and higher report rates of pedestrian location than the current 10 Hz by the stereo camera would be required to cope with the rapidly changing distances. Additionally, image recognition techniques in the front-facing camera of the pedestrian's head-mounted eye tracker would be beneficial to automatically identify the object of the pedestrian's attention. Finally, high-precision gyroscopes or IMUs would be necessary to better record the head accelerations of the moving pedestrian. With these tools, it would also be possible to extend the algorithm's scope to other vulnerable road users (VRUs) such as cyclists, whose interactions with vehicles are more dynamic than that of pedestrians.

Ideally, the pedestrian's gaze direction would need to be determined in real-time via the vehicle itself without requiring pedestrians to wear eye-trackers. This task could potentially be accomplished using a combination of pedestrian head orientations and the detection of salient objects in the 3-D world. Using these sources of information, it would be possible to calculate where the pedestrian is looking at any given time during the interaction. While 100% certainty in the pedestrian's gaze point / eye contact seeking may never be achieved, they could be reported with confidence values in real-time. With this step, outdoor controlled experiments could be substituted by real data from interactions with pedestrians on public roads, moving the state-of-the-art on detection of eye contact forward and making it applicable to automated vehicles of the future.

## ***ii. Streetlight effect in the experiment's design:***

Moving on to the design choices for the main experiment, care was taken to avoid the pitfall of the streetlight effect. The poor gaze data quality achieved in the pilot study meant that some compromises had to be made in order to avoid repeating past mistakes in the main experiment. These included performing all the trials indoors to minimize



interference from background infrared light, and using a stationary vehicle to minimize dynamic conditions. The 'indoor trials' constraint also necessitated the 'stationary car' constraint, since it was not possible to have a moving vehicle inside the building for a plethora of reasons. These reasons included insufficient space to drive and maneuver the vehicle, the presence of multiple obstacles leading to a high chance of collision, and safety risks to participants, researchers and passers-by alike. Since the indoor trials were chosen after failed attempts outdoors, the choice was not attributed to the streetlight effect on the researchers' part. Further, the experiment retained its fundamental principles (in terms of recreating a driver-pedestrian interaction involving eye contact), which further reduced the likelihood of streetlight effect in the design of the experiment.

### ***iii. Tobii gaze data quality:***

In the trials, the higher gaze sample percentages of the crossing trials (as compared to the corresponding standing trials) were attributed to there being less time in between different scenarios and within scenarios to look at bright surfaces like windows. In other words, it is probable that due to breaks during the standing trials to allow for scenario changeover and the downtime between repetitions of each scenario, participants were more likely to look at sources of infrared radiation since they were free to look wherever they wanted. This behavior likely prevented their gaze from being detected in those times and led to slightly lower gaze sample percentages in the standing trials. Other possible reasons were that the crossing trials involved more variable eye movements, less constant camera images, and large vertical accelerations (due to walking) which led to a greater loss of gaze data.

It was also observed that the plots of gaze samples within individual trials had similar trends. This fact likely meant that within each participant, differences in the gaze data quality between trials were more likely due to different gaze behaviors exhibited by the person rather than changes in the environment or the method of conducting the experiment. Further, considering that the experimental method was standardized and conducted indoors (albeit with many windows in view), the variation in gaze data quality between participants was probably due to a combination of individual differences between people (eye shape, eye color, gaze behavior etc.) and outdoor infrared levels.

### ***iv. Percentage of eye contact in the trials:***

The percentages of eye contact in the standing trials and crossing trials shown in the results section were not really as low as they appeared. To begin with, the Tobii data had gaze sample losses of 10-15% on average, which were treated as instants lacking eye contact seeking by the pedestrians (and by extension, as lacking driver-pedestrian eye contact). Further, after factoring in inaccuracies in the algorithm leading to some eye contact instants not being detected and the nature of the task in the experiment, context was provided to the low percentages, as explained below.

In the standing trials, 10-15% of the trial duration being eye contact was expected, given that a large part of the trials involved looking away from the driver, looking at the torch, and head turning. Also, the driver too was performing these actions so there was likely to be a mismatch in when the two parties looked at each other. This mismatch is expected to be common even in real driver-pedestrian interactions on public roads. Further, eye contact was not the same as eye contact seeking. Pedestrian eye contact seeking in the standing trials was 15-25% of the trial duration, which meant that the pedestrian was attempting eye contact with the driver for around 1.5 times as long as eye contact lasted. To put it differently, there were many instants when the pedestrian was seeking eye contact but eye contact was not happening because the driver was looking elsewhere or because the algorithm failed to detect.

Similarly, in the crossing trials, the percentages of eye contact were 10-15% of the trial duration. The crossing trials involved three repetitions of crossing back and forth in front of the vehicle. Each repetition was split into two halves (crossing from the right of the car, and crossing from the left of the car), and the percentages of eye contact shown in the results section were for these half-repetitions. Thus, the percentages, when put in perspective, were not that small. Further, pedestrian eye contact seeking for these half-repetitions was 25-35% which showed that pedestrians sought eye contact for most of the trial, when periods of looking away while turning around (on reaching the other end of the imaginary road), looking at the torch, and head turning were subtracted.



**v. *Right-left differences in eye contact duration:***

As per the measurements, eye contact lasted longer when pedestrians stood on the right side of the car, when compared to the left. Pedestrian eye contact seeking durations and dwell times were also measured to be higher (with significant differences) on the right side than on the left. However, no significant differences were found in the number of eye contact seeking visits between the sides in the standing trials. This information meant that pedestrians' eye contact seeking visits were about the same in number irrespective of the side, but measured longer when pedestrians were to the right of the car. It was probably the increased length of these eye contact seeking visits that caused increased durations of pedestrian eye contact seeking and driver-pedestrian eye contact. This finding also suggested that driver eye contact seeking may not have been the cause for the differences in eye contact duration.

These differences were also unlikely to be due varying trial lengths since no significant differences between the right side and left side standing trial durations were found. Further, there were no significant differences in gaze data quality between the right and left sides, so differing extents of data loss as the cause was also unlikely. It was also improbable that eye contact behavior itself (more specifically, pedestrian eye contact seeking behavior) was different depending on the side the pedestrian stood on. It was more likely that the differences were introduced in the measurement of eye contact duration (and pedestrian eye contact seeking duration). These statements followed from the facts that: (1) the imaginary road in the experiment was only 1 m wider than the car itself, which made it similar to the width of a single lane road, (2) the vehicle-pedestrian interaction was symmetrical about the centerline of the imaginary road, and so the right side and left side interactions could be assumed to be identical, and (3) the trials were performed in random order. Thus, given the circumstances, it was unlikely that the participants exhibited a common behavior that made them seek eye contact for longer only when they stood on the right of the car.

One plausible explanation for the differences might be researcher bias towards one side while drawing the AOIs. To be more specific, it is possible that the researcher unintentionally drew AOIs at more frames on the right side than on the left. These extra AOIs were probably located at the times when the pedestrian was turning to look at the driver and when turning away from the driver. The greater number of AOIs could have led to a greater number of AOI hits (due to peripheral eye contact seeking), which translated to longer durations of eye contact seeking visits.

Another possible explanation is geometrical in nature. The pedestrian was closer to the driver when he/she stood on the left of the car, so the apparent size of the driver's head on the left would have been larger compared to the right. This would mean that the driver's head was bigger relative to the AOI size (which was constant for both sides – a circle of diameter 150 pixels) for a pedestrian on the left. Thus, the AOI on the right covered a larger region outside the driver's head than the AOI on the left. Thus, it might have been more difficult to obtain an AOI hit (in the Tobii's software) on the left, which in turn led to shorter dwell times. The converse would have applied when the pedestrian stood on the right. A greater distance to the driver would have led to a smaller apparent size of the driver's head, which meant a smaller head relative to the AOI size, which in turn meant a larger region covered outside the driver's head, easier AOI hits and longer dwell times.

No differences in eye contact duration between the sides (i.e. right/left half-repetitions) were present for crossing pedestrians. This result may be because researcher bias (towards one side) while drawing the AOIs and differing apparent sizes of the driver's head did not have a major effect, since all the crossing trials involved both sides (i.e. starting from one side and ending on the other).

In the future, the issue with the apparent size of the driver's head in relation to the AOI could potentially be solved by having different sized AOIs depending on the side. Researcher bias could perhaps be mitigated by having multiple researchers separately draw the AOIs, followed by comparing their annotations and using the best combination from different sources. In this way, the reliability and repeatability of the annotation method could also be measured.

**vi. *Detection performance of the algorithm in the context of AVs:***

Algorithmic detection of eye contact is important for automated vehicles of the future because it could be used in their control logic or in an eHMI. The knowledge that a pedestrian has seen the approaching AV or has made eye contact



with the passenger(s) would allow the AV to make a decision on whether to yield or not, while communicating its decision to the pedestrian and other nearby vehicles, all in real-time. Thus, a first step to achieving this goal would be to detect eye contact between drivers and pedestrians.

The reasons for the fair/poor performance of the algorithm at detecting eye contact instants and trials are believed to be the neglecting of pedestrian head roll and poor accuracy of the Tobii eye-tracker beyond short distances. These would explain why the actual values of the 'ground truth threshold' and 'eye contact threshold' were so much higher than their theoretical counterparts. Missing data being treated as instants of 'no eye contact' is unlikely to have significantly affected algorithm performance, since most of the data loss occurred when the pedestrian was looking away from the driver and at bright surfaces. At such instants, the angle between driver and pedestrian gaze rays would have been quite high, leading the algorithm to classify the instants as 'no eye contact' anyway. That said, it is possible that the small amount of missing data when the pedestrian was looking at the driver/vehicle (if any) may have caused a marginal increase in the number of false positives in the algorithm.

A look at the ROC and false positive rate (FPR) curves of the standing and crossing trials showed that eye contact thresholds for some participants needed to be almost zero in order to eliminate false positives in eye contact instants. This observation indicated that false positives could not be entirely removed from the algorithm without seriously compromising correct detections. This result was especially evident in the crossing trials due to the presence of dynamic conditions, which made it harder to separate eye contact from gazing at the vehicle (in the pedestrian data).

In the context of an automated vehicle, it was important to know when false positives were completely removed because a false detection of eye contact arguably posed a greater danger to a pedestrian than a failure to detect eye contact. In the latter case, the automated vehicle might treat the situation as one where a pedestrian was going to cross the road while being unaware of the approaching AV. Here, pedestrian body kinematics alone could still be used by an automated vehicle to safely yield for the person. However, in the former case, the situation might appear to the AV as one where a pedestrian saw the approaching vehicle as he/she was going to cross the road. Here, a collision could occur due to conflicting information provided by false positives in eye contact and pedestrian kinematics. In other words, the AV might (wrongly) assume that the pedestrian would slow down or stop at the curb because eye contact was (falsely) detected, and allocate less priority to the pedestrian's trajectory that predicted a collision. Here, if the AV decided not to yield, the consequences would be serious.



# 5 Conclusions

In conclusion, this master's thesis achieved the three goals it set out to achieve. First, the lack of objective measurements for driver-pedestrian eye contact in the literature was addressed by conducting an indoor experiment designed to resemble an interaction on a road and by quantitatively measuring eye contact using the novel method of synchronizing two eye-trackers. Two inertial measurement units and a stereo camera with pedestrian detection capabilities were also used in tandem with the eye-trackers.

Eye contact durations of 0.42 s and 0.54 s were measured for pedestrians standing on a curb to the left and to the right of a car, respectively. It was proposed that these measurements would be lower than values obtained for real interactions on public roads in the Netherlands and worldwide. The presence of older participants was expected to increase eye contact durations, as were countries where pedestrians did not usually receive right of way. Significant differences in the amounts of eye contact between the two sides were attributed to researcher bias in drawing the AOIs and differing apparent sizes of the driver's head relative to the AOIs.

Eye contact durations of 1.23 s and 1.39 s were measured for pedestrians crossing a road from the left and from the right of a car, respectively. It was proposed that these measurements would be higher than values obtained for real interactions on public roads. The instructions provided to participants to seek eye contact for the entire duration of their crossing were the likely cause for the high measurements of eye contact duration. No significant differences in the amounts of eye contact between the two sides were found.

This thesis also demonstrated that driver-pedestrian eye contact could be detected automatically by using an algorithm to find the angle between the 3-D gaze rays of the two parties at every instant of a trial in the experiment. Angles were compared with an 'eye contact threshold' to determine if eye contact occurred at an instant or not. Trials were also classified as either involving eye contact or not, based on their percentage of eye contact instants.

Second, the performance of the algorithm at correctly classifying instants and trials involving eye contact was measured. This assessment was made by comparing the outputs of the algorithm with the following ground truths: (1) instructions provided to participants before each trial, and (2) eye contact detections using manually annotated AOIs. The overall classification accuracy of the algorithm was determined to be fair/poor, and gaze data losses were found to be low. The optimal range of 'eye contact threshold' angles for best performance by the algorithm was found to be 15–30°. In other words, the algorithm could detect driver-pedestrian eye contact with an accuracy of 15–30°.

In the trials involving a standing pedestrian seeking eye contact, almost all eye contact instants were correctly identified but at the cost of many false positives in the trials where the pedestrian was not seeking eye contact. This trade-off led to poor classification of the trials themselves (correct for 12 participants out of 30 valid ones). For the standing trials, best detection of eye contact instants occurred at an 'eye contact threshold' of 20°, and best detection of eye contact trials at 25°.

In the trials involving a crossing pedestrian seeking eye contact, most eye contact instants were correctly identified, along with many false positives. In the crossing trials involving a crossing pedestrian who was not seeking eye contact, there were fewer false positives. This trade-off led to good classification of the trials themselves (correct for 19 participants out of 30 valid ones). For the crossing trials, best detection of eye contact instants occurred at an 'eye contact threshold' of 25°, and best detection of eye contact trials at 15°.

Third, the driver-pedestrian interaction was reconstructed as a 3-D animation showing the positions of the car, the driver, and the pedestrian, along with their gaze rays of the persons. This animation was realized by using the readings of multiple devices: (1) driver gaze directions from the Smart Eye, (2) driver head positions from the Smart Eye, (3) pedestrian locations from the stereo camera, (4) pedestrian head directions from the Tobii's gyroscope, and (5) pedestrian eyeball directions from the Tobii. Isometric and top views of the interaction in the animation were exported as videos for the purposes of communication.

Finally, from the findings of this master's thesis, some concluding remarks may be made as follows:



- Eye contact on the road lasts likely lasts atleast 0.4-0.5 s and at most 1.23-1.39 s.
- Driver-pedestrian eye contact can indeed be detected algorithmically but there is room for improvement in the accuracy and reliability.
- There is potential to apply the detection of eye contact in the decision-making and eHMI modules of automated vehicles to improve road safety.
- Reconstruction of driver-pedestrian interactions is achievable and has scope for application in vehicle-to-vehicle and vehicle-to-infrastructure communication in automated vehicles.

It is hoped that the findings in this report are beneficial to future research attempting to detect driver-pedestrian eye contact for the purpose of application in automated vehicles.



# 6 Recommendations

It must be said that this master's thesis was merely an initial foray into measuring driver-pedestrian eye contact, and was therefore by no means a comprehensive study of the subject. Its aims were primarily to provide numerical definitions of eye contact on the road, and to determine if it was possible to detect and reconstruct driver-pedestrian eye contact reliably. That being said, there is much scope for expansion from the base that is this piece of research:

- A follow-up experiment of eye contact detection conducted outdoors on a real road (under favorable lighting conditions for the eye trackers – late afternoons or cloudy/overcast days with no precipitation) and with minimal instructions to participants (to avoid influencing their natural road scanning/crossing behavior). Light intensity measurements using a photometer are advised to counter any bias due to different lighting conditions outside during the experiment. The use of high-precision external gyroscopes or IMUs for the Tobii (to obtain absolute head directions) and the consideration of all possible rotations is recommended. Moving vehicle trials with different approach speeds, approach directions, and yielding conditions would also be beneficial. GPSs for the car and pedestrian detection at high report rates would be helpful for dynamic trials. Further, pedestrian willingness to cross may be measured using a button response, and questionnaires targeting eye contact (its importance, its perceived duration, its perceived frequency of occurrence) may be administered. Repeat experiments of the above in different countries and with a wider variety of participants would also be useful. These efforts will provide an understanding of eye contact in naturalistic on-road scenarios, and make the results more applicable to pedestrian safety around automated vehicles.
- Real-time detection/reconstruction of interaction situations showing eye contact occurrence, and possibly 3-D gaze points (accurate calculations of the intersection point of the right and left eye gaze rays) in place of gaze directions. This data would also allow the mapping of gaze onto objects at different depths in a 3-D simulation/environment, and aid in studies of how road users perceive depth. Image recognition techniques could be applied to video from the head-mounted eye-tracker to also achieve this on the pedestrian's end. Alternatively, pedestrians' foci of attention could be tracked from the vehicle itself (perhaps using head orientations, and without requiring them to wear eye-trackers) to indirectly obtain pedestrian gaze directions. As an extension, real-time 'eye' tracking of passengers and pedestrians in 3-D may be implemented in automated vehicles and shared amongst a network of nearby, connected vehicles for better communication and safety on the road.
- An extension of this experiment and algorithm to cover other VRUs such as bicyclists. A well-controlled experiment may offset any accuracy issues in the gaze ray method of eye contact detection due to the increased dynamicity of a bicycle-vehicle interaction. Especially in a country like the Netherlands, where the number of cyclists are high, the safety benefits of research on eye contact by cyclists will be considerable.
- Creation of eye contact datasets by having recruited pedestrians and cyclists wear portable eye-trackers like the Tobii Pro Glasses 2 during their daily commute. Driver eye contact seeking duration and frequency may be gleaned from this. Researchers could also be enlisted to drive around cities in the Netherlands in the Toyota Prius with Smart Eye tracking while collecting stereo camera recordings of pedestrians' eye contact seeking behavior.

Research on driver-pedestrian eye contact has serious implications for the future of traffic. Streets filled with automated vehicles capable of detecting pedestrians' gaze / eye contact seeking attempts, communicating their intentions and data from the interaction amongst themselves, and gesturing to pedestrians implicitly / explicitly is a conceivable vision of the future. The end goal is seamless and safe interaction between VRUs and automated vehicles, despite 'eye contact with a driver' no longer being possible. An end goal that perhaps entails substituting (the soon-to-be obsolete) driver-pedestrian eye contact with an alternative (e.g. passenger-pedestrian eye contact, eye contact eHMI) that serves the same or a superior role to road users and is functionally indistinguishable from or superior to driver-pedestrian eye contact.



# 7 Bibliography

- [1] Centraal Bureau voor de Statistiek (CBS). (2018). Overledenen; doden door verkeersongeval in Nederland, wijze van deelname. Retrieved from: <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/71936ned/table?ts=1552322329668>
- [2] World Health Organization (WHO). (2018). Global status report on road safety. Retrieved from: [https://www.who.int/violence\\_injury\\_prevention/road\\_safety\\_status/2018/en/](https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/)
- [3] Governor's Highway Safety Association (GHSA). (2019). Pedestrian traffic fatalities by state: 2018 preliminary data. Retrieved from: <https://www.ghsa.org/resources/Pedestrians19>
- [4] National Highway Traffic Safety Administration (NHTSA). (2018). 2017 fatal motor vehicle crashes: overview. Retrieved from: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812603>
- [5] Farber, B. (2016). Communication and communication problems between autonomous vehicles and human drivers. In: Maurer, M., Gerdes, J., Lenz, B., Winner, H. (eds) Autonomous Driving. Springer, Berlin, Heidelberg
- [6] Rasouli, A., Kotseruba, I. & Tsotsos, J.K. (2017). Agreeing to cross: How drivers and pedestrians communicate. 2017 IEEE Intelligent Vehicles Symposium (IV), 264 – 269
- [7] Kotseruba, I., Rasouli, A. & Tsotsos, J. K. (2016). Joint attention in autonomous driving (JAAD). arXiv preprint. arXiv: 1609.04741 [cs.RO]
- [8] Hopmans, B., Wesdorp, D., Visscher, J. & de Vlam, V. (2018). Visual scanning behavior in a parking lot. Bachelor end project, TU Delft
- [9] Walker, I. & Brosnan, M. (2007). Drivers' gaze fixations during judgements about a bicyclist's intentions. Transportation Research Part F: Traffic Psychology and Behavior, 10(2), 90 – 98
- [10] Lundgren, V. M. et al. (2017). Will there be new communication needs when introducing automated vehicles to the urban context? In: Stanton, N., Landry, S., Di Bucchianico, G., Vallicelli, A. (eds) Advances in Human Aspects of Transportation. Advances in Intelligent Systems and Computing, vol 484. Springer, Cham
- [11] Ren, Z., Jiang, X. & Wang, W. (2016). Analysis of the influence of pedestrians' eye contact on drivers' comfort boundary during the crossing conflict. Procedia Engineering 137, 399 – 406
- [12] Snyder, M., Grether, J. & Keller, K. (1974). Staring and compliance: A field experiment on hitchhiking. Journal of Applied Social Psychology, 4(2), 165 – 170
- [13] Chang, C., Toda, K., Sakamoto, D. & Igarashi, T. (2017). Eyes on a car: An interface design for communication between an autonomous car and a pedestrian. AutomotiveUI '17: Proceedings of the 9th ACM International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 65 – 73
- [14] Sucha, M., Dostal, D. & Risser, R. (2017). Pedestrian-driver communication and decision strategies at marked crossings. Accident Analysis and Prevention 102, 41 – 50
- [15] Rothenbücher, D. et al. (2015). Ghost Driver: A platform for investigating interactions between pedestrians and driverless vehicles. AutomotiveUI '15: Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 44 – 49
- [16] Kooijman, L. (2018). The detection of pedestrian crossing behaviour. Master thesis, TU Delft. Retrieved from: <https://repository.tudelft.nl/islandora/object/uuid%3A9cd6529f-039e-44ec-9451-72c18174e3e0?collection=education>
- [17] Habibovic, A. et al. (2018). Communicating intent of automated vehicles to pedestrians. Frontiers in Psychology 9: 1336



- [18] Tobii AB. (2015). How do Tobii eye trackers work? Retrieved from: <https://www.tobii.com/learn-and-support/learn/eye-tracking-essentials/how-do-tobii-eye-trackers-work/>
- [19] Smart Eye AB. (2018). Smart Eye Pro Manual, Revision 8.2
- [20] Liu, W. et al. (2015). SSD: Single Shot MultiBox Detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds) Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science, vol 9905. Springer, Cham
- [21] Smart Eye AB. (2012). Smart Eye Pro User Manual, Revision 551. Retrieved from: [https://www.kostech.net/admin/bbs/down.php?code=manual\\_tm&idx=6463&no=1](https://www.kostech.net/admin/bbs/down.php?code=manual_tm&idx=6463&no=1)
- [22] Olsen, A. (2012). The Tobii I-VT Fixation Filter algorithm description. Retrieved from: <https://www.tobii.com/siteassets/tobii-pro/learn-and-support/analyze/how-do-we-classify-eye-movements/tobii-pro-i-vt-fixation-filter.pdf>
- [23] Tobii AB. (2018). When do I use the IV-T Attention Filter? Retrieved from: [https://connect.tobii.com/s/article/When-do-I-use-the-I-VT-Attention-filter?language=en\\_US](https://connect.tobii.com/s/article/When-do-I-use-the-I-VT-Attention-filter?language=en_US)
- [24] Grossman, G. E. et al. (1988). Frequency and velocity of rotational head perturbations during locomotion. *Experimental Brain Research*, 70(3), 470 – 476
- [25] Jaguar Land Rover. (2018). The virtual eyes have it. Retrieved from: <https://www.jaguarlandrover.com/2018/virtual-eyes-have-it>
- [26] CSP Public Affairs [@CSP\_News]. (2018). Pedestrians: follow the advice of Fred Edestrian and make eye contact with all drivers before putting your fragile body in the path of an oncoming vehicle. Which means you can't be looking at your phone and the driver at the same time (hint hint) [Tweet]. Twitter. Retrieved from: [https://twitter.com/CSP\\_News/status/984485458594836480](https://twitter.com/CSP_News/status/984485458594836480)
- [27] Nasar, J. L. & Troyer, D. (2013). Pedestrian injuries due to mobile phone use in public places. *Accident Analysis and Prevention* 57, 91 – 95
- [28] daSilva, M. P., Smith, J. D. & Najm, W. G. (2004). Analysis of pedestrian crashes. Department of Transportation, National Highway Traffic Safety Administration (NHTSA). Retrieved from: <http://www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/Crash%20Avoidance/2003/DOTHS809585.pdf>
- [29] Moore, D., Currano, R., Strack, G. E. & Sirkin, D. (2019). The case for implicit external human-machine interfaces for autonomous vehicles. *AutomotiveUI '19 Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 295 – 307
- [30] Dey, D. & Terken, J. M. B. (2017). Pedestrian interaction with vehicles: roles of explicit and implicit communication. *AutomotiveUI '17 Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct*, 158 – 162
- [31] AlAdawy, D. et al. (2019). Eye contact between pedestrians and drivers. *Proceedings of the 10th International Driving Symposium on Human Factors in Driver Assessment*, 301 – 307
- [32] iDS Imaging. (2014). UI-3060CP Rev. 2. Retrieved from: <https://en.ids-imaging.com/store/ui-3060cp-rev-2.html>
- [33] NexTorch. (2011). NexTorch myTorch 3AAA Smart LED Flashlight. Retrieved from: <http://www.nextorch.com/en/index.php/products/myTorch-3AAA-Smart-LED-Flashlight>
- [34] Tape, T. G. (1990). The area under an ROC curve. University of Nebraska Medical Center. Retrieved from: <http://gim.unmc.edu/dxtests/roc3.htm>
- [35] Dey, D., Walker, F., Martens, M. & Terken, J. (2019). Gaze patterns in pedestrian interaction with vehicles: towards effective design of external human-machine interfaces for automated vehicles. *AutomotiveUI '19 Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 369 – 378



- [36] de Vries, S. (2017). Using a wearable eye-tracking device on bicyclists to explore the possibility of measuring motorcyclist eye movements. Master thesis, University of Twente. Retrieved from: [https://essay.utwente.nl/73485/1/devries\\_MA\\_BMS.pdf](https://essay.utwente.nl/73485/1/devries_MA_BMS.pdf)
- [37] Hryhoruk, A. (2019). Case number 00467567: Origin w.r.t which X, Y, Z gaze points in 3-D are measured. Tobii Pro Customer Support. Retrieved from: <https://connect.tobiiipro.com/>
- [38] Katz, A., Zaidel, D. & Elgrishi, A. (1975). An experimental study of driver and pedestrian interaction during the crossing conflict. *Human Factors*, 17(5), 514 – 527
- [39] Walker, I. (2005). Signals are informative but slow down responses when drivers meet bicyclists at road junctions. *Accident Analysis and Prevention*, 37(6), 1074 – 1085
- [40] Schneemann, F. & Gohl, I. (2016). Analyzing driver-pedestrian interaction at crosswalks: a contribution to autonomous driving in urban environments. 2016 IEEE Intelligent Vehicles Symposium (IV), 38 – 43
- [41] Perlman, S. B. et al. (2009). Individual differences in personality predict how people look at faces. *PLoS ONE*, 4(6): e5952
- [42] Walker-Smith, G. J., Gale, A. G. & Findlay, J. M. (1977). Eye movement strategies in face perception. *Perception*, 6(3), 313–326
- [43] Ferranti, L. et al. (2019). SafeVRU: A research platform for the interaction of self-driving vehicles with vulnerable road users. 2019 IEEE Intelligent Vehicles Symposium (IV), 1660–1666
- [44] Roth, M., Flohr, F. & Gavrilă, D. M. (2016). Driver and pedestrian awareness-based collision risk analysis. IEEE Intelligent Vehicles Symposium (IV), 454–459. Retrieved from: [http://gavrila.net/iv16\\_driver\\_ped.pdf](http://gavrila.net/iv16_driver_ped.pdf)
- [45] Borys, A. (2019). Case number 00472212: Measurement tolerances of Tobii Pro Glasses 2. Tobii Pro Customer Support. Retrieved from: <https://connect.tobiiipro.com/>
- [46] Fitzpatrick, K., Brewer, M. A. & Turner, S. (2006). Another look at pedestrian walking speed. *Transportation Research Record: Journal of the Transportation Research Board*, 1982(1), 21–29
- [47] Montufar, J., Arango, J., Porter, M. & Nakagawa, S. (2007). Pedestrians' normal walking speed and speed when crossing a street. *Transportation Research Record: Journal of the Transportation Research Board*, 2002(1), 90–97



# Appendix

## A. Consent form

### Consent form for participants Research title: “Eye contact between a driver and a pedestrian”

**Researchers:**

Vishal Onkhar, MSc student  
[V.Onkhar@student.tudelft.nl](mailto:V.Onkhar@student.tudelft.nl)

Dr. Pavlo Bazilinskyy - Supervisor  
[P.Bazilinskyy@tudelft.nl](mailto:P.Bazilinskyy@tudelft.nl)

Dr. ir. Joost de Winter - Supervisor  
[J.C.F.deWinter@tudelft.nl](mailto:J.C.F.deWinter@tudelft.nl)

**Location of the experiment:**

Hangar of the 3mE building

**Introduction:** Before agreeing to participate in this study, it is important that the following is read and understood.

**Purpose of the experiment:** The purpose of this experiment is to collect data about eye contact between a pedestrian and a driver in a stationary car. In this experiment, you are the 'pedestrian'. Based on the collected data, the researchers will develop a computer algorithm that can automatically detect eye contact between a driver and a pedestrian.

**Duration:** Your participation in this study will last about 60 minutes.

**Procedures:** First, you will be asked to complete a questionnaire regarding your basic demographic characteristics. After completing the questionnaire, you will be asked to put on the Tobii Pro Glasses 2. This is a portable head-mounted eye tracker that looks like normal glasses. The Tobii does not obscure your vision. Before starting with the experiment, an experimenter will calibrate the Tobii, which will take about 3 minutes.

Next, you will perform about 50 trials of standing/walking in front of the car (Toyota Prius). You will receive instructions from the experimenter to stand/walk in a particular direction and to either make eye contact or NOT make eye contact with the 'driver' in the car. The car is parked and will not move during the experiment. After the experiment, you will complete a brief questionnaire about your experiences.

Using the Tobii, we will capture your forward field of view using a camera, and record where you are looking. Furthermore, the car is equipped with cameras and pedestrian recognition. The experimenter is also equipped with an eye tracker.

**Risks and discomforts:** There are no higher risks than the risks of regular walking. The Tobii does not obstruct your vision. If, for any reason, you feel uncomfortable at any point during the experiment, do not hesitate to remove the glasses and walk back to the experimenter.

**Anonymity:** All the data collected in this study will be kept anonymous. Throughout the study, you will be identified by a participant number only. The data may be shared with other researchers or stored in a public repository, but only in an anonymous form. We will ensure that in any publications, you will not be identifiable, for example, by blurring your face where appropriate.

**Right to refuse or withdraw:** Your participation is voluntary, and you may refuse to participate or discontinue your participation at any time, without negative consequences.

**Questions:** If you have any questions concerning this study, you may contact one of the researchers (contact details at the top of this document).



I have read and understood the information provided above. I give permission to process the data for the purposes described above. I voluntarily agree to participate in this study.

Name:

.....

Signature:

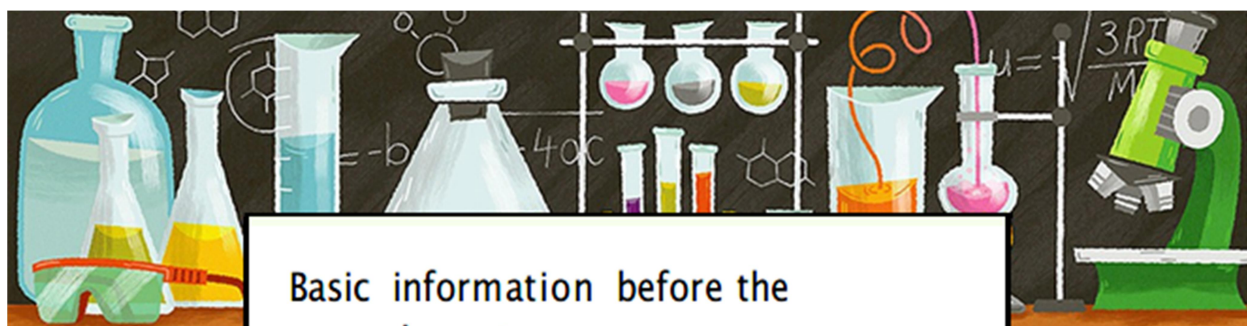
.....

Date:

..... / ..... / .....

## B. Pre-experiment questionnaire





## Basic information before the experiment

5 minutes

\* Required

Participation number \*

Your answer

---

Age \*

Your answer

---

Sex \*

☐ Male

☐ Female

☐ Other: \_\_\_\_\_

Height (m) \*

Your answer

---

Occupation \*

☐ BSc student

☐ MSc student

☐ PhD candidate

☐ Postdoc

☐ Employee

☐ Intern

☐ Unemployed



Home country \*

Your answer

---

Driver's license \*

☐ Yes

☐ No

How often did you drive in the last 12 months? \*

☐ Everyday

☐ Few times a week

☐ Few times a month

☐ Few times a year

☐ Never

How often did you walk to a destination in the last 12 months? \*

☐ Daily

☐ Few times a week

☐ Few times a month

☐ Few times a year

☐ Never

Eye color \*

☐ Brown

☐ Blue

☐ Green

☐ Hazel

☐ Grey

☐ Amber

Eye correction \*

☐ None

☐ Contact lenses



**Right eye error (0 if no correction)\***

Your answer \_\_\_\_\_

**Left eye error (0 if no correction)\***

Your answer \_\_\_\_\_

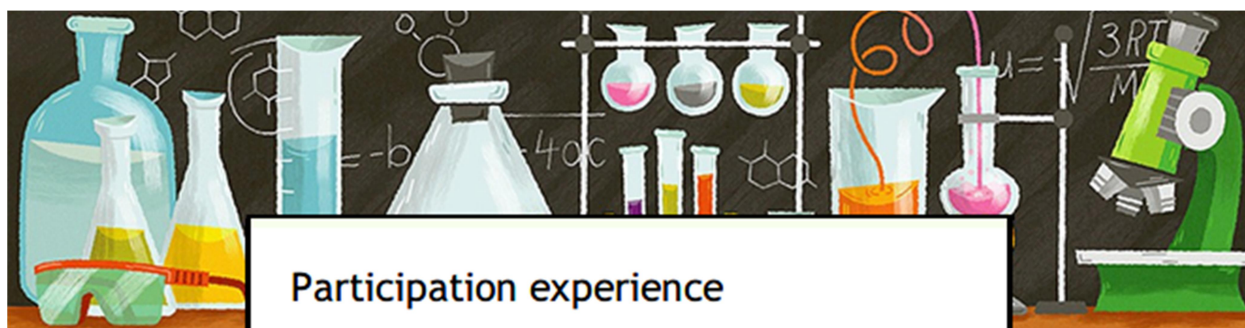
**SUBMIT**

Never submit passwords through Google Forms.

Google

## C. Post-experiment questionnaire





## Participation experience

5 minutes

\* Required

Participation number \*

Your answer

How well were you able to imagine a real traffic scenario at a real pedestrian crossing? \*

1 2 3 4 5 6 7

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

Completely

How well could you concentrate on the task during the experiment? \*

1 2 3 4 5 6 7

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

Fully

How well could you maintain eye contact during the experiment without looking elsewhere? \*

1 2 3 4 5 6 7

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

Perfectly

How well could you avoid eye contact during the experiment and look elsewhere? \*

1 2 3 4 5 6 7

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

Perfectly

How conscious were you of wearing eye tracking glasses while performing the experiment? \*

1 2 3 4 5 6 7

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

Fully



How much do you think the eye tracking glasses affected your natural road crossing behavior? \*

1 2 3 4 5 6 7

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

Totally

How realistic did the experiment feel compared to crossing a real road? \*

1 2 3 4 5 6 7

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

Completely

How involved were you during the experiment? \*

1 2 3 4 5 6 7

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

Totally

How clear were the instructions for the experiment? \*

1 2 3 4 5 6 7

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

Crystal

How well do you think you followed the instructions for the experiment? \*

1 2 3 4 5 6 7

Not at all

☐ ☐ ☐ ☐ ☐ ☐ ☐

Perfectly

SUBMIT

Never submit passwords through Google Forms.

This content is neither created nor endorsed by Google. [Report Abuse](#) - [Terms of Service](#)

Google Forms



## D. Trial randomization matrices

These matrices present the random order in which the standing/crossing trials were performed within the standing/crossing recordings, and the order in which the recordings themselves were performed for the 31 participants.

Crossing trials			Standing trials				Recordings	
1	W-NEC	W-EC	1	L-S-EC	R-S-EC	R-S-NEC	L-S-NEC	1 Standing Crossing
2	W-EC	W-NEC	2	L-S-NEC	R-S-NEC	R-S-EC	L-S-EC	2 Standing Crossing
3	W-NEC	W-EC	3	R-S-EC	L-S-NEC	R-S-NEC	L-S-EC	3 Crossing Standing
4	W-EC	W-NEC	4	R-S-EC	L-S-EC	R-S-NEC	L-S-NEC	4 Crossing Standing
5	W-EC	W-NEC	5	R-S-EC	R-S-NEC	L-S-EC	L-S-NEC	5 Standing Crossing
6	W-NEC	W-EC	6	L-S-NEC	L-S-EC	R-S-NEC	R-S-EC	6 Crossing Standing
7	W-NEC	W-EC	7	L-S-EC	R-S-NEC	L-S-NEC	R-S-EC	7 Standing Crossing
8	W-EC	W-NEC	8	L-S-EC	L-S-NEC	R-S-NEC	R-S-EC	8 Standing Crossing
9	W-EC	W-NEC	9	L-S-EC	R-S-NEC	L-S-NEC	R-S-EC	9 Standing Crossing
10	W-NEC	W-EC	10	R-S-EC	L-S-NEC	L-S-EC	R-S-NEC	10 Crossing Standing
11	W-EC	W-NEC	11	L-S-EC	L-S-NEC	R-S-EC	R-S-NEC	11 Crossing Standing
12	W-EC	W-NEC	12	R-S-NEC	L-S-NEC	L-S-EC	R-S-EC	12 Standing Crossing
13	W-EC	W-NEC	13	R-S-EC	L-S-EC	R-S-NEC	L-S-NEC	13 Standing Crossing
14	W-NEC	W-EC	14	R-S-NEC	L-S-EC	R-S-EC	L-S-NEC	14 Crossing Standing
15	W-NEC	W-EC	15	L-S-NEC	L-S-EC	R-S-NEC	R-S-EC	15 Standing Crossing
16	W-NEC	W-EC	16	L-S-EC	R-S-NEC	R-S-EC	L-S-NEC	16 Crossing Standing
17	W-EC	W-NEC	17	R-S-EC	L-S-NEC	R-S-NEC	L-S-EC	17 Crossing Standing
18	W-EC	W-NEC	18	L-S-NEC	R-S-NEC	L-S-EC	R-S-EC	18 Crossing Standing
19	W-EC	W-NEC	19	L-S-EC	L-S-NEC	R-S-EC	R-S-NEC	19 Standing Crossing
20	W-EC	W-NEC	20	R-S-NEC	L-S-NEC	R-S-EC	L-S-EC	20 Crossing Standing
21	W-NEC	W-EC	21	L-S-NEC	R-S-NEC	R-S-EC	L-S-EC	21 Standing Crossing
22	W-EC	W-NEC	22	R-S-EC	R-S-NEC	L-S-EC	L-S-NEC	22 Standing Crossing
23	W-EC	W-NEC	23	R-S-NEC	L-S-NEC	L-S-EC	R-S-EC	23 Crossing Standing
24	W-NEC	W-EC	24	L-S-EC	R-S-NEC	L-S-NEC	R-S-EC	24 Crossing Standing
25	W-EC	W-NEC	25	L-S-EC	L-S-NEC	R-S-EC	R-S-NEC	25 Crossing Standing
26	W-NEC	W-EC	26	L-S-EC	R-S-EC	R-S-NEC	L-S-NEC	26 Standing Crossing
27	W-EC	W-NEC	27	L-S-NEC	R-S-NEC	R-S-EC	L-S-EC	27 Standing Crossing
28	W-NEC	W-EC	28	R-S-EC	L-S-NEC	R-S-NEC	L-S-EC	28 Crossing Standing
29	W-EC	W-NEC	29	R-S-EC	L-S-EC	R-S-NEC	L-S-NEC	29 Crossing Standing
30	W-EC	W-NEC	30	R-S-EC	R-S-NEC	L-S-EC	L-S-NEC	30 Standing Crossing
31	W-NEC	W-EC	31	L-S-NEC	L-S-EC	R-S-NEC	R-S-EC	31 Crossing Standing



## E. Trial eye contact seeking percentages

Participant number	Eye contact percentages in the standing trials, ordered as shown in the randomization matrix, for an 'eye contact threshold' of:												Standing trials											
	15°				20°				25°				30°											
1	30.49	8.74	0.43	0.00	38.92	28.39	10.36	0.00	41.51	31.61	23.95	0.00	43.03	39.66	31.07	16.13								
2	0.00	29.29	13.47	0.12	0.00	33.02	15.97	1.49	0.00	34.75	18.95	11.58	6.70	36.22	20.62	23.66								
3	4.86	0.00	0.94	2.81	19.86	0.00	3.90	6.67	26.53	0.00	9.67	12.11	31.53	13.02	18.41	20.35								
4	0.00	0.00	17.44	0.00	10.73	8.05	24.62	4.41	25.61	17.73	25.98	9.88	30.62	34.15	38.12	11.29								
5	0.00	0.00	0.00	0.00	5.48	0.00	0.00	0.00	12.53	3.65	10.90	0.00	21.28	6.04	23.55	0.51								
6	0.00	0.00	0.00	0.00	0.00	10.99	0.00	0.00	0.00	14.46	1.37	7.84	21.64	19.73	6.32	15.99								
7	0.00	0.00	0.00	3.20	13.28	4.24	5.38	7.74	21.63	10.50	14.66	20.37	25.24	15.10	23.38	27.78								
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.63	0.00	0.00	5.68	23.66	0.00	0.00	19.13	25.38								
9 (No data)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00								
10	18.25	0.00	0.00	13.18	24.52	0.00	4.18	18.24	27.38	2.75	7.97	21.45	29.85	8.06	9.96	23.82								
11	19.31	1.03	4.14	0.00	31.12	24.30	11.90	10.67	36.74	30.34	25.56	11.59	38.76	33.14	37.72	33.99								
12	0.00	0.00	0.00	18.53	0.00	0.00	11.21	25.20	0.00	0.00	18.33	28.13	0.00	0.00	25.91	31.60								
13	0.00	0.44	0.00	0.00	9.57	5.92	0.00	0.00	24.64	11.84	5.61	0.00	31.57	16.01	12.38	1.96								
14	19.33	16.16	9.81	0.00	32.57	27.90	15.06	0.00	37.14	32.32	22.10	0.00	39.27	35.82	32.18	1.24								
15	0.00	1.13	10.67	8.54	0.00	4.92	19.63	14.62	0.28	21.06	22.33	16.50	2.35	23.96	25.32	21.85								
16	2.53	6.51	24.96	0.00	11.47	31.03	30.02	0.00	28.67	35.82	34.55	1.38	32.55	39.27	36.82	16.61								
17	29.29	5.98	12.86	17.92	34.60	17.61	25.57	25.41	36.87	20.27	28.64	27.95	39.30	21.76	35.38	32.15								
18	0.00	0.00	0.00	4.50	0.00	0.00	0.00	21.77	0.00	0.00	1.85	28.23	1.44	0.00	25.89	31.08								
19	0.74	0.00	1.90	0.00	11.39	1.58	7.79	0.00	28.11	3.31	14.88	1.06	30.77	5.68	20.59	6.03								
20	20.03	0.00	2.75	2.46	36.36	2.52	13.44	16.55	42.03	21.06	16.03	19.15	44.27	27.68	20.15	25.99								
21	0.00	0.00	1.71	0.00	0.00	0.00	10.16	10.48	0.00	8.29	14.81	12.94	0.00	11.80	20.81	19.92								
22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	15.06	0.00	0.00	0.00	30.56	1.00	0.44	0.00								
23	27.61	5.12	13.83	4.63	29.47	20.37	21.13	7.76	31.55	27.89	23.23	15.87	35.15	31.15	25.44	19.81								
24	25.56	1.96	5.34	1.61	30.31	11.78	17.48	8.82	34.13	23.16	27.90	20.85	36.07	29.44	33.42	27.39								
25	2.02	0.00	0.00	0.00	9.85	0.00	18.72	23.24	23.25	0.00	24.58	29.73	25.98	8.41	28.21	32.30								
26	14.37	0.00	0.00	0.70	23.75	0.00	0.55	2.60	28.94	0.00	3.19	3.30	35.83	0.00	8.49	7.51								
27	0.00	15.78	0.00	2.06	0.00	23.44	0.00	11.84	0.00	24.84	0.00	24.20	0.00	26.09	5.66	31.02								
28	19.67	0.00	0.00	0.00	27.98	0.00	0.00	0.51	29.92	0.00	0.81	2.18	31.86	0.00	2.60	8.32								
29	0.00	0.13	0.00	0.00	0.00	14.21	0.00	0.00	0.00	23.35	12.96	1.48	18.76	25.63	27.71	36.00								
30	0.00	0.00	0.00	0.00	2.49	0.00	0.00	0.00	29.70	0.00	0.30	0.00	34.21	22.04	26.56	12.52								
31	20.61	18.43	0.10	7.14	26.68	20.13	5.67	17.02	28.24	22.63	11.83	19.16	30.82	28.44	19.35	22.12								

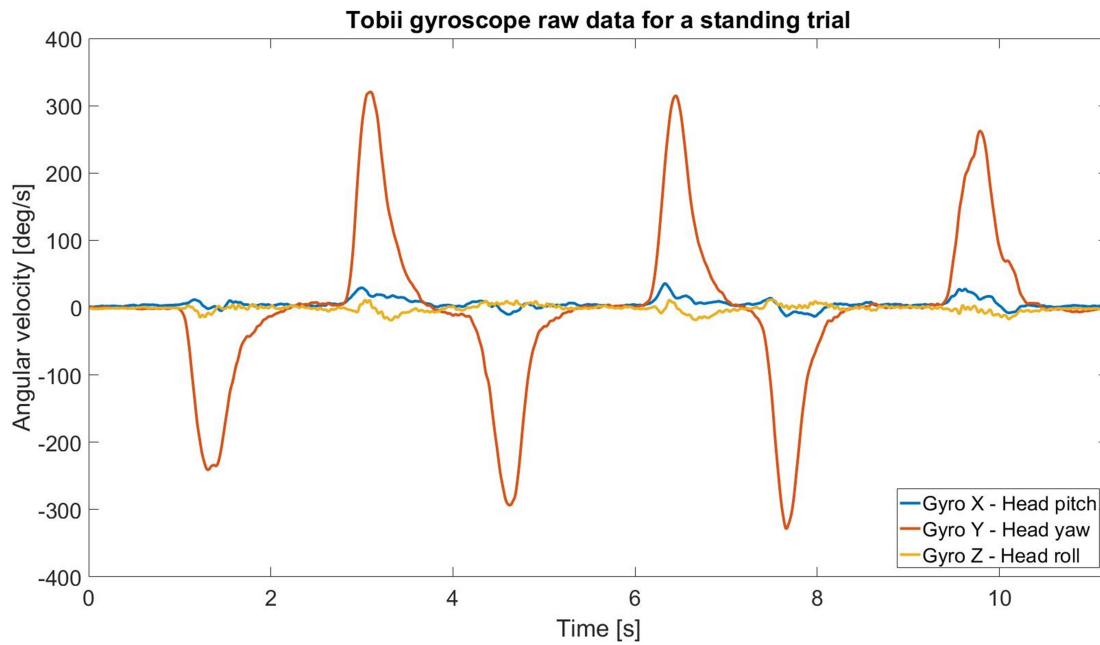


Crossing trials

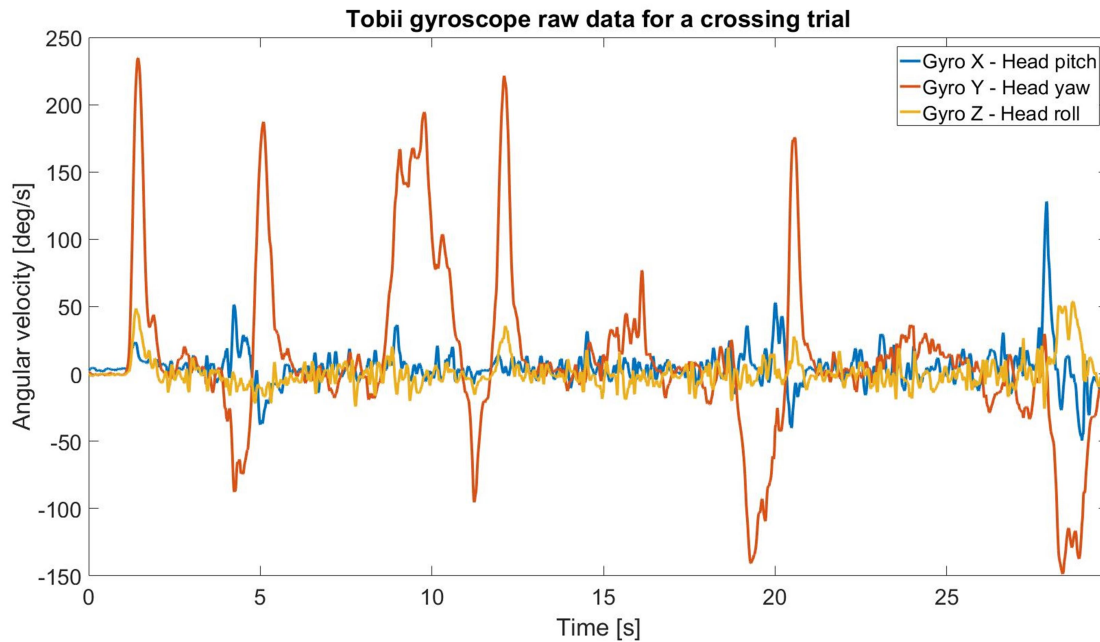
Participant number	Eye contact percentages in the crossing trials, ordered as shown in the randomization matrix, for an 'eye contact threshold' of:									
	15°			20°			25°			30°
1	1.85	25.07		27.15	39.45		59.71	43.86	70.05	55.98
2	10.16	19.80		21.60	25.92		35.21	34.54	39.53	42.87
3	12.16	17.71		17.20	30.02		24.47	46.07	29.75	51.31
4	39.90	0.13		60.81	17.40		69.28	30.28	74.60	40.75
5	34.70	13.94		41.73	44.97		47.01	61.66	53.42	74.66
6	6.68	14.16		13.87	29.10		27.30	47.35	37.60	60.66
7	0.21	38.68		3.02	53.75		6.56	62.28	10.32	66.92
8	7.90	28.12		22.92	41.12		35.77	46.36	46.61	49.25
9 (No data)	0.00	0.00		0.00	0.00		0.00	0.00	0.00	0.00
10	55.87	61.96		67.25	65.89		71.23	70.98	73.91	73.16
11	5.05	0.00		17.87	12.97		30.15	31.69	55.12	41.91
12	26.41	3.68		41.32	6.52		48.30	9.28	56.95	11.48
13	43.11	69.21		52.89	78.33		61.82	82.07	67.67	83.55
14	1.13	15.63		17.63	30.92		46.13	46.32	51.56	58.81
15	42.41	39.88		61.92	47.97		72.52	55.89	78.59	62.15
16	0.00	40.08		1.25	52.89		4.07	62.48	9.13	71.53
17	50.12	28.44		65.38	38.33		75.43	50.72	80.64	59.68
18	14.29	0.00		45.41	9.19		50.93	38.05	58.05	70.74
19	14.07	26.96		20.67	32.80		24.54	39.54	33.97	44.82
20	25.14	8.20		40.51	16.46		48.28	27.90	53.73	33.74
21	26.17	42.07		30.54	48.55		34.20	58.18	42.23	63.95
22	2.97	4.34		20.75	33.49		57.52	52.57	67.07	60.56
23	24.95	41.24		29.48	47.66		33.73	51.05	38.54	54.78
24	0.00	12.02		8.76	22.96		22.14	30.05	36.83	35.73
25	43.45	45.19		52.93	55.43		64.34	63.95	71.51	70.43
26	17.54	17.37		25.43	25.30		34.41	29.66	42.61	34.65
27	7.44	49.94		16.74	63.09		30.49	71.03	49.43	78.38
28	3.08	6.49		17.24	27.04		28.08	41.47	38.53	55.18
29	0.00	0.00		8.96	0.84		35.04	11.56	52.70	56.59
30	30.07	15.44		40.42	31.81		50.76	53.56	57.38	63.40
31	0.00	25.59		9.10	38.05		23.45	54.05	37.17	66.95



## F. Example raw data plots

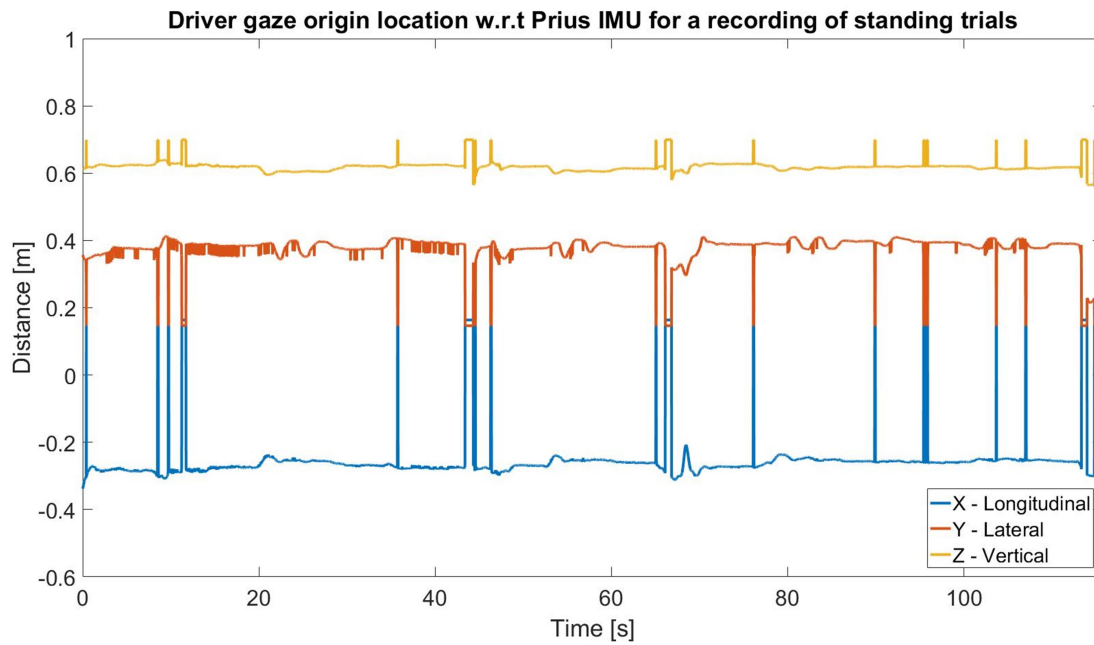


For gyro Y, the sharp drops (valleys) represent the pedestrian (standing on the left of the car) turning his/her head to look at the driver. The sharp rises (peaks) represent the pedestrian turning his/her head to look away from the driver. Eye contact likely occurs in the flat sections in between valleys and peaks; the pedestrian's head remains steady during these times

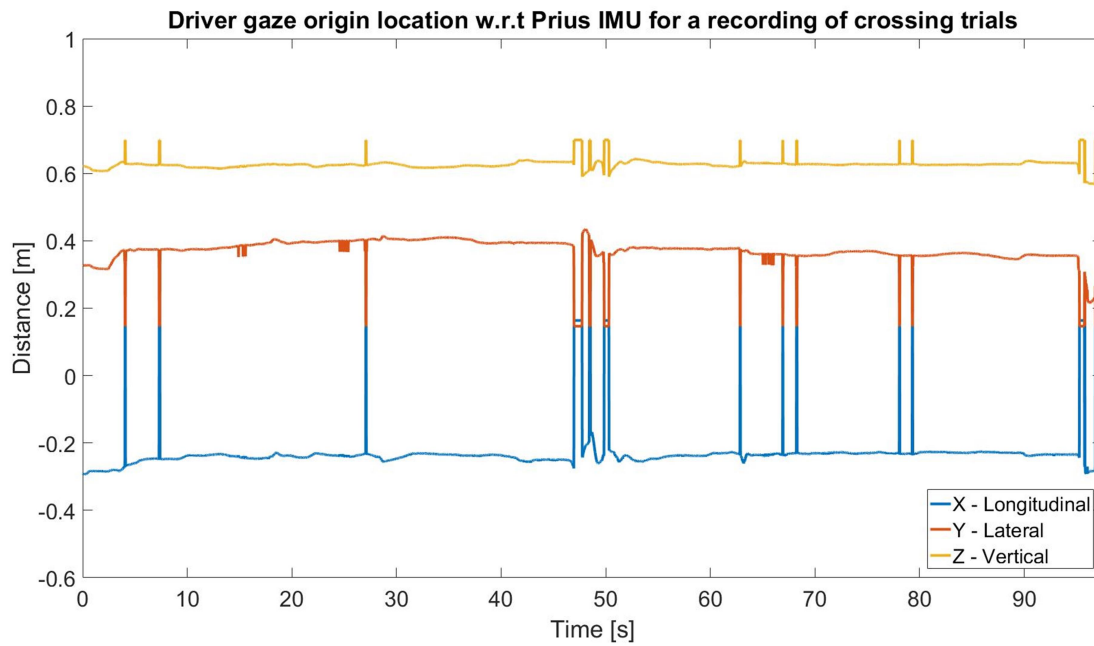


For gyro Y, sharp drops (valleys) followed immediately by sharp rises (peaks) or vice-versa represent the pedestrian turning around to walk back upon reaching the other side of the road. The shallower but wider valleys and shorter but wider peaks likely mean gentler turns by the pedestrian.



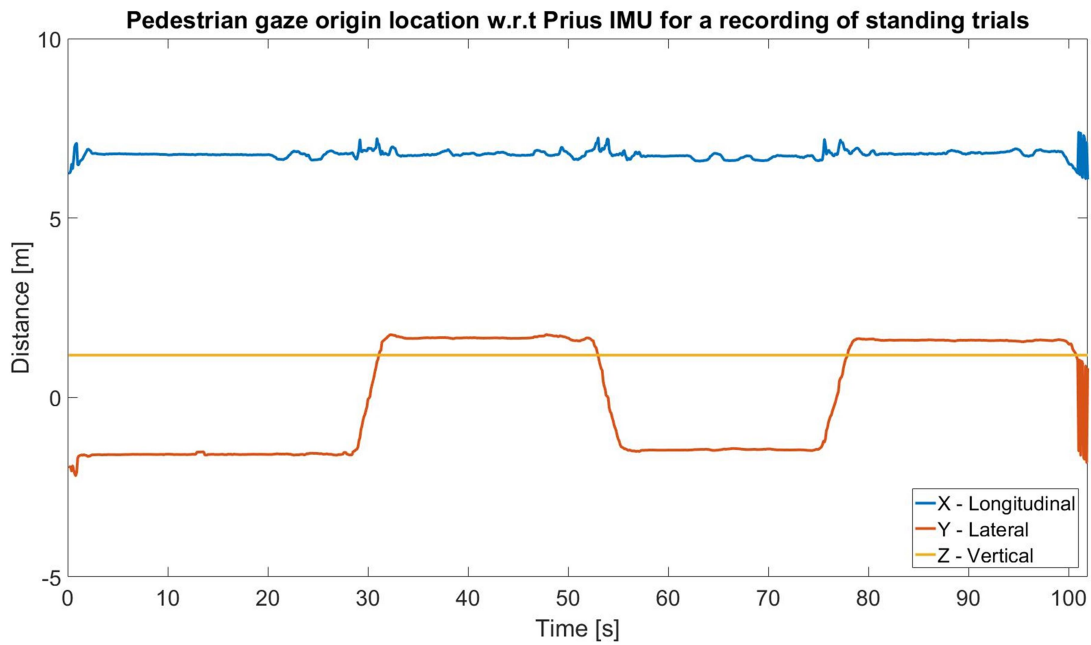


Driver gaze origin changed very little since the driver was seated. The most change (relatively) was seen in the lateral direction, due to the driver moving his head slightly to get a better view of the pedestrian. The clusters of spikes were likely due to sudden (rapid) head movements during scenario changeover (i.e. checking cables, checking Smart Eye output, checking randomization matrix). Isolated spikes were likely noise in the Smart Eye detections of driver gaze origin.

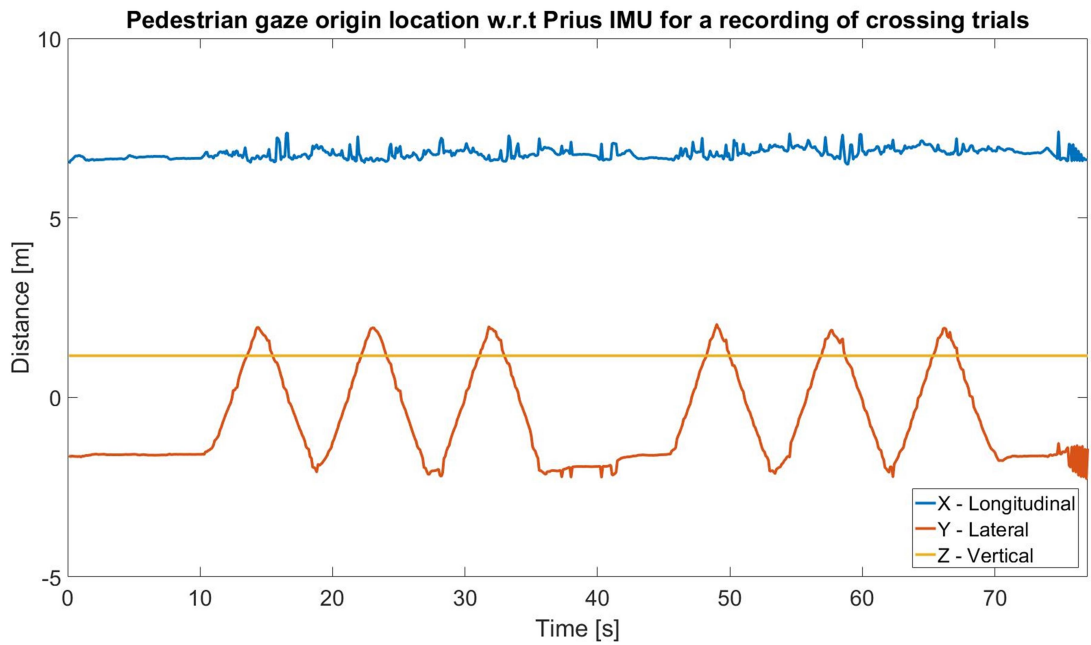


Driver gaze origin changed very little since the driver was seated. The minimal head movements suggest that the driver used smooth pursuit eye movements to follow the crossing pedestrian throughout the trial. The cluster of spikes in the middle of the recording was likely due to rapid head movements during scenario changeover. Isolated spikes are possibly noise in the measurements.





The change in lateral position of the pedestrian is attributed to scenario changeover, when the pedestrian is asked to relocate to the other side of the road (if necessary). This claim is easily verified using the randomization matrix.



The repeated change in lateral position of the pedestrian is attributed to the pedestrian crossing back and forth in front of the car. Atypically, this participant spent almost no time before turning back on reaching the other side of the road. This fact is evident from the pointed peaks of lateral position.



## G. MATLAB script of algorithm

```
%% Tobii-Smarteye eye contact and gaze ray tracing
% MATLAB script by Vishal Onkhar
% MSc Thesis 2019-2020
close all;
clear all;
clc;

% eye_cont_percent=zeros(31,4);

for participant_number=1:31

    scenario_choice=1;

    if participant_number==9 % No Smart Eye data for this participant

        continue;

    end

    %% User input
    % participant_number=input('Participant number? 31 - Max\n');
    % Enter in MATLAB command window
    % scenario_choice=input('Type of scenario? 1 - Standing or 2 - Crossing\n');
    % Enter in MATLAB command window
    tic; % Start timer

    if scenario_choice==1

        standing=1;
        crossing=0;
        scenario_string='S';

    elseif scenario_choice==2

        standing=0;
        crossing=1;
        scenario_string='W';

    end

    %% Read from data files
    rp_smarteye='../Final experiment/Smarteye data';
    % Relative path to smarteye data folder
    s_smarteye=dir(fullfile(rp_smarteye, '*.csv'));
    % Obtain all .csv files in folder
    n_smarteye={s_smarteye.name};
    % Obtain all .csv filenames in folder

    x_smarteye=~cellfun(@isempty, strfind(n_smarteye, strcat(sprintf('%02d', participant_number), '- ', scenario_string, '- Gazedata')));
    % Find matching smarteye data file
    f_smarteye=xlsread(fullfile(rp_smarteye, n_smarteye{x_smarteye}));
    % Read matching data file

    rp_tobii='../Final experiment/Tobii data/Pedestrians';
    % Relative path to tobii data folder
    s_tobii=dir(fullfile(rp_tobii, '*.xlsx'));
    % Obtain all .xlsx files in folder
```



```

    n_tobii={s_tobii.name};
% Obtain all .xlsx filenames in folder

x_tobii=~cellfun(@isempty, strfind(n_tobii, strcat(sprintf('%02d', participant_number), '-
', scenario_string, '-Gazedata')));
    % Find matching tobii data file
    [f_tobii, t_tobii]=xlsread(fullfile(rp_tobii, n_tobii{x_tobii}));
% Read matching data file

    rp_stereo='../Final experiment/Smarteye data';
% Relative path to stereo camera data folder
    s_stereo=dir(fullfile(rp_stereo, '*.csv'));
% Obtain all .csv files in folder
    n_stereo={s_stereo.name};
% Obtain all .csv filenames in folder

x_stereo=~cellfun(@isempty, strfind(n_stereo, strcat(sprintf('%02d', participant_number), '-
', scenario_string, '-Stereo')));
    % Find matching stereo data file
    f_stereo=importdata(fullfile(rp_stereo, n_stereo{x_stereo}), ':');
% Read matching data file and delimit by colon

    n_gyro_zero_drift='../Final experiment/Drift/Stationary_Table_Gyro_Data.xlsx';
% Relative path to gyro zero drift data
    f_gyro_zero_drift=xlsread(n_gyro_zero_drift);
% Read zero drift data file

    n_gyro_rot_drift='../Final experiment/Drift/Rotating_Chair_Gyro_Data.xlsx';
% Relative path to gyro rotating drift data
    f_gyro_rot_drift=xlsread(n_gyro_rot_drift);
% Read rotating drift data file

    n_basicdata='Basic information before the experiment (Responses).xlsx';
% Pre-experiment questionnaire responses file
    f_basicdata=xlsread(n_basicdata);
% Read questionnaire #1 responses

    n_prius_top_view='../Prius 3d mesh/top-view-cropped.jpg';
% Relative path to prius top view image for animation
    f_prius_top_view=imread(n_prius_top_view);
% Read image file

    if standing==1 && crossing==0

        n_trials='Randomized standing trials.xlsx';
        % Orders of performed standing trials in all recordings
        n_se_start_times='Smarteye standing trial start times.xlsx';
        % Torch start instant times w.r.t smarteye for standing trials
        n_se_end_times='Smarteye standing trial end times.xlsx';
        % Torch end instant times w.r.t smarteye for standing trials

    elseif standing==0 && crossing==1

        n_trials='Randomized crossing trials.xlsx';
        % Orders of performed crossing trials in all recordings
        n_se_start_times='Smarteye crossing trial start times.xlsx';
        % Torch start instant times w.r.t smarteye for crossing trials
        n_se_end_times='Smarteye crossing trial end times.xlsx';
        % Torch end instant times w.r.t smarteye for crossing trials

    end

```



```

[~,f_trials,~]=xlsread(n_trials);
% Read the trial orders
f_se_start_times=xlsread(n_se_start_times);
% Read the smarteye trial start times
f_se_end_times=xlsread(n_se_end_times);
% Read the smarteye trial end times

%% Hardcoded data
smarteye_hz=60;
% Smarteye recording frequency (Hz)
tobii_hz=50;
% Tobii recording frequency (Hz)
gyro_hz=95;
% Tobii gyro sampling rate (Hz)
ts=2;
% Time (s) column number in smarteye data file
prius_dim=[4.48;1.7;1.45];
% Prius dimensions (m) as L,B,H
imu=[0;0;0];
% Universal origin IMU with X forward, Y left, Z up
prius_cent=[0;0;0.205];
% Prius centroid location (m) w.r.t the IMU as X,Y,Z
gox=34;
% Gaze origin x (m) column number in smarteye data file
gdx=70;
% Filtered gaze direction x (ix) column number in smarteye data file
event=10;
% Torch event column in tobii data file
gdirlx=17;
% Gaze direction left x in tobii data file
aoi_start=39;
% AOI start column in tobii data file
ht=4;
% Height column number in basic data file
median_tb_height=1.79;
% Median height of torch bearers (m)
max_ped_dist=7.5;
% Approx. pedestrian distance from stereo camera rounded up to nearest 0.5
gyro_x=43;
% Gyro X column in tobii data file
no_of_rot=100;
% 100 chair rotations performed to test tobii gyro drift
tobii_centre=[0,0,1];
% Pedestrian looking straight ahead i.e. x=0 (left), y=0 (up), z=1 (forward) in
tobii's own frame of reference
no_of_reps=3;
% Number of repetitions of each trial/scenario

%% Coordinate transformation matrices
quat=[0.52547,0.47357,-0.46474,0.53257];
% Quaternion rotation with X,Y,Z,W in ros but changed to W,X,Y,Z for matlab
trans=[0.1351,0.67706,-0.24641];
% Left stereo camera to IMU translation

smarteye_start_times=f_se_start_times(participant_number,:);
% Trial torch start times (s) w.r.t recording start
smarteye_end_times=f_se_end_times(participant_number,:);
% Trial torch end times (s) w.r.t recording start
ped_heights=f_basicdata(:,ht);
% Pedestrian heights (m)
median_ped_height=median(ped_heights);
% Median pedestrian height (m)

```



```

cur_ped_height=ped_heights(participant_number);
% Current pedestrian height (m)
eye_height=cur_ped_height-0.1;
% Pedestrian eye height (m), assuming 10 cm from top of head to eyes
scenarios=f_trials(participant_number,:);
% Order of trials or scenarios in one recording
no_of_trials=length(scenarios);
% Number of trials or scenarios in one recording

%% Obtaining the correct AOIs depending on the trial
if crossing==1 && standing==0

    aoi_columns=str2double(strrep(strrep(scenarios,'W-EC','2'),'W-NEC','3'));
% Converting random order of scenarios into numbers
    aoi_columns=[aoi_columns(1),aoi_columns(1)-2,aoi_columns(end),aoi_columns(end)-
2]; % Correct AOI column order
    aoi_columns=aoi_columns+aoi_start;
% Correct AOI column numbers

elseif standing==1 && crossing==0

    aoi_columns=str2double(strrep(strrep(strrep(strrep(scenarios,'L-S-EC','0'),'L-
S-NEC','1'),'R-S-EC','2'),'R-S-NEC','3'))+aoi_start;
% Random trial orders to correct AOI column numbers

end

aoi_end=aoi_start+3;
% AOI end column number in tobii data file
l_smarteye=length(f_smarteye);
% Number of smarteye data points
l_stereo=length(f_stereo.txtdata);
% Length of stereo camera messages
l_tobii=length(f_tobii);
% Length of tobii messages

%% Assuming if eyes are not found i.e. -1 then there is no eye contact i.e. 0
for i=aoi_start:aoi_end

    f_tobii(f_tobii(:,i)==-1,i)=0;
% Replace -1s with 0s for pedestrian eye contact seeking binary plot

end

%% Find trial start and end instants using the torch flash event
torch_start=zeros(1,no_of_trials);
% Pre-allocation of array to hold trial start times
torch_end=zeros(1,no_of_trials);
% Pre-allocation of array to hold trial end times
count1=1;
% Loop counter variable

for i=1:l_tobii

    if ~isempty(strfind(t_tobii{i,event},'-Start'))

        torch_start(count1)=i-1;
% Subtracting one to account for title row

        elseif ~isempty(strfind(t_tobii{i,event},'-End'))

```



```

        torch_end(count1)=i-1;
    % Subtracting one to account for title row
    count1=count1+1;
    % Updating loop variable

end

end

%% Find pedestrian and torch bearer locations from stereo camera data
ped_locations=zeros(10000,3);
    % In the IMU frame of reference with X front, Y left, Z up
torch_bearer_locations=zeros(10000,3);
    % In the IMU frame of reference with X front, Y left, Z up
count1=1;
    % Loop counter variable
count2=1;
    % Loop counter variable
i=1;
    % Loop skip variable

while i<=l_stereo

    if strcmp(f_stereo.textdata(i,2),'person') &&
str2double(f_stereo.textdata(i+10,2))<max_ped_dist
        % Only accept class ID 'person' within 7.5 m w.r.t stereo camera

ped_left=str2double([f_stereo.textdata(i+8,2),f_stereo.textdata(i+9,2),f_stereo.textdata(i+10,2)]); % Ped. pos. w.r.t stereo camera with X right, Y down, Z front
        cur_ped_loc=quatrotate(quat,ped_left+trans);
    % Ped. pos. w.r.t IMU (also try rotate first, translate second)
        ped_locations(count1,:)=cur_ped_loc;
    % Aggregated pedestrian positions
        count1=count1+1;
    % Updating loop variable
        i=i+10;
    % Updating skip variable

        elseif strcmp(f_stereo.textdata(i,2),'person') &&
str2double(f_stereo.textdata(i+10,2))>max_ped_dist &&
abs(str2double(f_stereo.textdata(i+8,2)))<prius_dim(2)
        % Only accept class ID 'person' beyond 7.5 m w.r.t stereo camera and within twice
the car's width

tb_left=str2double([f_stereo.textdata(i+8,2),f_stereo.textdata(i+9,2),f_stereo.textdata(i+10,2)]);
        % Torch bearer position w.r.t left stereo camera with X right, Y down, Z front
        cur_tb_loc=quatrotate(quat,tb_left+trans);
    % Torch bearer pos. w.r.t IMU (also try rotate first, translate second)
        torch_bearer_locations(count2,:)=cur_tb_loc;
    % Aggregated torch bearer positions
        count2=count2+1;
    % Updating loop variable
        i=i+10;
    % Updating skip variable

end

    i=i+1;
    % Updating skip variable

```



```

end

ped_locations=ped_locations(any(ped_locations,2),:);
% Removing rows of zeros
torch_bearer_locations=torch_bearer_locations(any(torch_bearer_locations,2),:);
% Removing rows of zeros

%% Find absolute smarteye start and end times
tns=ts+1;
% Time (ns) column number in smarteye data file
ped_locations(:,3)=eye_height-0.5*prius_dim(3)+prius_cent(3);
% Eye height Z (m) relative to IMU since stereo camera reading is inaccurate
smarteye_rec_start_time=f_smartheye(1,ts)+f_smartheye(1,tns)*10^-9;
% Recording absolute start time w.r.t smarteye
smarteye_rec_end_time=f_smartheye(end,ts)+f_smartheye(end,tns)*10^-9;
% Recording absolute end time w.r.t smarteye
stereo_hz=(smarteye_rec_end_time-smarteye_rec_start_time)/length(ped_locations);
% Stereo camera sampling frequency (Hz) = Recording duration/Number of readings
smarteye_start_times=smarteye_start_times+smarteye_rec_start_time;
% Trial torch absolute start times w.r.t smarteye
smarteye_end_times=smarteye_end_times+smarteye_rec_start_time;
% Trial torch absolute end times w.r.t smarteye
ped_locations=imresize(ped_locations,[1_smartheye,3],'nearest');
% Nearest-neighbour interpolation to upsample pedestrian position from 10 hz to 60 hz

%% Assuming torch is held at eye height
torch_bearer_locations(:,3)=median_tb_height-0.1-0.5*prius_dim(3)+prius_cent(3);
% Torch bearer eye height Z (m) relative to IMU since stereo camera reading is
inaccurate
torch_bearer_locations=imresize(torch_bearer_locations,[1_smartheye,3],'nearest');
% Nearest-neighbour interpolation to upsample torch position from 10 hz to 60 hz

%% Find trial start and end instants in the smarteye data
smarteye_start_indices=zeros(1,no_of_trials);
% Trial start indices for smarteye
smarteye_end_indices=zeros(1,no_of_trials);
% Trial end indices for smarteye
count1=1;
% Trial start counter
count2=1;
% Trial end counter

for i=1:1_smartheye

    smarteye_abs_time=f_smartheye(i,ts)+f_smartheye(i,tns)*10^-9;
% Absolute time w.r.t smarteye

    if count1<=no_of_trials && smarteye_abs_time>=smarteye_start_times(count1)
% If the trial has started

        smarteye_start_indices(count1)=i;
% Storing the trial start index
        count1=count1+1;
% Updating loop variable

    end

    if count2<=no_of_trials && smarteye_abs_time>=smarteye_end_times(count2)
% If the trial has ended

```



```

        smarteye_end_indices(count2)=i;
% Storing the trial end index
        count2=count2+1;
% Updating loop variable

    end

end

pedestrian_start_locations=ped_locations(smarteye_start_indices,:);
% Pedestrian locations at trial starts
pedestrian_end_locations=ped_locations(smarteye_end_indices,:);
% Pedestrian locations at trial ends
torch_bearer_start_locations=torch_bearer_locations(smarteye_start_indices,:);
% Torch bearer locations at trial starts
torch_bearer_end_locations=torch_bearer_locations(smarteye_end_indices,:);
% Torch bearer locations at trial ends

%% Data file column numbers
goy=gox+1;
% Gaze origin y column number in smarteye data file
goz=goy+1;
% Gaze origin z column number in smarteye data file
gdy=gdx+1;
% Filtered gaze direction y (jy) column number in smarteye data file
gdz=gdy+1;
% Filtered gaze direction z (kz) column number in smarteye data file
gdirlx=gdirly+1;
% Gaze direction left y in tobii data file
gdirlz=gdirly+1;
% Gaze direction left z in tobii data file
gdirrx=gdirly+1;
% Gaze direction right x in tobii data file
gdirry=gdirrx+1;
% Gaze direction right y in tobii data file
gdirrz=gdirry+1;
% Gaze direction right z in tobii data file
gyro_y=gyro_x+1;
% Gyro Y column number in tobii data file
gyro_z=gyro_y+1;
% Gyro Z column number in tobii data file

%% Draw 3D rectangular box for plotting the car
% euler_angles=[0;0;0];
% Euler rotation angles
% c1=cos(euler_angles(1));
% s1=sin(euler_angles(1));
% c2=cos(euler_angles(2));
% s2=sin(euler_angles(2));
% c3=cos(euler_angles(3));
% s3=sin(euler_angles(3));
% rotation_matrix=[c1*c2,-c2*s1,s2;c3*s1+c1*s2*s3,c1*c3-s1*s2*s3,-c2*s3;s1*s3-
c1*c3*s2,c3*s1*s2+c1*s3,c2*c3]';
% Rotation matrix
% prius_x_vert=0.5*prius_dim(1)*[-1 1 1 -1 -1 1 1 -1]';
% X vertices
% prius_y_vert=0.5*prius_dim(2)*[1 1 1 1 -1 -1 -1 -1]';
% Y vertices
% prius_z_vert=0.5*prius_dim(3)*[-1 -1 1 1 1 1 -1 -1]';
% Z vertices
% prius_faces=[1,2,3,4;5,6,7,8;4,3,6,5;3,2,7,6;2,1,8,7;1,4,5,8];
% Cuboid faces
% prius_vert=zeros(3,8);

```



```

%
% for i=1:length(prius_vert)
%
%
prius_vert(1:3,i)=rotation_matrix*[prius_x_vert(i);prius_y_vert(i);prius_z_vert(i)]+rotation_matrix*prius_cent;
    % Rotation and translation of prius vertices
%
% end

%% Countering gyro drift
raw_zero_drift=f_gyro_zero_drift(:,gyro_x-2:gyro_z-2);
    % Zero drift rate in deg/s
raw_zero_drift(any(isnan(raw_zero_drift),2),:)=[];
    % Removal of rows with nan
zero_drift=(1/gyro_hz)*cumtrapz(raw_zero_drift,1);
    % Zero drift in deg
zero_drift_rate=zero_drift(end,:)/length(zero_drift);
    % Zero drift rate in deg/frame

% raw_rot_drift=f_gyro_rot_drift(:,gyro_x-2:gyro_z-2);
    % Rotational drift rate in deg/s
% raw_rot_drift(any(isnan(raw_rot_drift),2),:)=[];
    % Removal of rows with nan
% rot_drift=(1/gyro_hz)*cumtrapz(raw_rot_drift,1);
    % Rotational drift in deg
% l_rot_drift=length(rot_drift);
    % Length of rotational drift test
% rot_drift_rate=rot_drift(end,:)/l_rot_drift;
    % Rotational drift rate in deg/frame
% rot_drift_rate(2)=rot_drift_rate(2)-360*no_of_rot/l_rot_drift;
    % Correction for drift rate y i.e. yaw

%% Butterworth filter for noisy gyro data
% % [b,a]=butter(4,[0.15,10]/gyro_hz/2,'bandpass');
% % Setting low cut (0.5 hz) and high cut (10 hz) frequencies
[b,a]=butter(2,6/gyro_hz/2,'low');
    % 6 hz cut-off frequency, 2nd order filter based on Grossman et al. and Kooijman,
Lars
% % [d,c]=butter(4,0.5/gyro_hz/2,'high');
% % 0.5 hz cut-off frequency, 4th order filter
toc;
    % Stop timer

%% 3D animated plot and angle between gaze rays
% mycells=zeros(74,no_of_trials);
% % Pre-allocate array for hit rate and false positive rate
% count=1;
% % Loop counter variable

% for threshold=0:5:180
% % Loop through different angle thresholds for ROC curves
if standing==1 && crossing==0

    threshold=15;
    % Set gaze ray angle threshold (15)

elseif crossing==1 && standing==0

    threshold=30;
    % Set gaze ray angle threshold (30)

```



```

end

aoi_threshold=15;
% Set AOI angle threshold at 15 deg

for k=1:no_of_trials
% Same as length(smarteys_start_indices) and length(smarteys_end_indices)

    l_trial=length(smarteys_start_indices(k):smarteys_end_indices(k));
% Length of the trial in frames sampled at 60 hz

    %% Pedestrian eye-in-head directions
    trial_ped_gaze_dir_left=f_tobii(torch_start(k):torch_end(k),gdirlx:gdirly);
% Tobii pedestrian eye gaze direction left
    trial_ped_gaze_dir_left(any(isnan(trial_ped_gaze_dir_left),2),:)=[];
% Removing rows with nan

    trial_ped_gaze_dir_left=imresize(trial_ped_gaze_dir_left,[l_trial,3],'nearest');
% Resampling to 60 hz
    trial_ped_gaze_dir_right=f_tobii(torch_start(k):torch_end(k),gdirrx:gdirry);
% Tobii pedestrian eye gaze direction right
    trial_ped_gaze_dir_right(any(isnan(trial_ped_gaze_dir_right),2),:)=[];
% Removing rows with nan

    trial_ped_gaze_dir_right=imresize(trial_ped_gaze_dir_right,[l_trial,3],'nearest');
% Resampling to 60 hz

    [d,c]=butter(2,30/tobii_hz/2,'low');
% 2nd order, 30 hz cut-off frequency found using the trial and error method

    trial_ped_gaze_dir=(trial_ped_gaze_dir_left+trial_ped_gaze_dir_right)/2;
% Mean of left and right eye gaze directions
    trial_ped_gaze_dir(:,2)=0;
% Setting vertical axis component i.e.y=0
    trial_ped_gaze_dir=filtfilt(d,c,trial_ped_gaze_dir);
% Low-pass filtering
    trial_ped_gaze_dir=normr(trial_ped_gaze_dir);
% Row-wise normalization
    rep_tobii_centre= repmat(tobii_centre,l_trial,1);
% Duplicating reference gaze direction (straight ahead) for matrix multiplication
    cross_prod=cross(rep_tobii_centre,trial_ped_gaze_dir,2);
% Vector cross product between reference and mean eye direction
    dot_prod=dot(rep_tobii_centre,trial_ped_gaze_dir,2);
% Vector dot product between reference and mean eye direction
    eye_angles=atan2d(sqrt(sum(cross_prod.^2,2)),dot_prod);
% Pedestrian eye-in-head angles w.r.t to straight ahead reference

    %% Gyro filtering and resampling
    rawgyro=f_tobii(torch_start(k):torch_end(k),gyro_x:gyro_z);
% Reading gyro data from tobii data file
    rawgyro(any(isnan(rawgyro),2),:)=[];
% Removing rows with nan
    l_rawgyro=length(rawgyro);
% Number of gyro data points
    steps_gyro=1:l_rawgyro;
% Same as 1:length(gyro)
    filtgyro=filtfilt(b,a,rawgyro);
% Low-pass filtering
    gyro=(1/gyro_hz)*cumtrapz(filtgyro,1);
% Integrating gyro data to get angles in degrees (also try detrending and using
rawgyro)

```



```

%
cum_rot_drift=[rot_drift_rate(1)*steps_gyro;rot_drift_rate(2)*steps_gyro;rot_drift_rate
(3)*steps_gyro]';
% Rotational drift in deg for each instant of the trial

cum_zero_drift=[zero_drift_rate(1)*steps_gyro;zero_drift_rate(2)*steps_gyro;zero_drift_
rate(3)*steps_gyro]';
% Zero drift in deg for each instant of the trial

    if standing==1 && crossing==0 && ~isempty(strfind(scenarios{k},'L'))

%         cum_rot_drift(:,2:3)=-cum_rot_drift(:,2:3);
% Switch signs for left side correction

        cum_zero_drift(:,2:3)=-cum_zero_drift(:,2:3);
% Switch signs for left side correction
        eye_angles=-eye_angles;
% Switch signs for left side eye-in-head angles

    elseif crossing==1 && standing==0

%
ped_eye_cont_seeking1=imresize(f_tobii(torch_start(k):torch_end(k),aoi_columns(2*k)),[l
_trial,1],'nearest');
%
ped_eye_cont_seeking2=imresize(f_tobii(torch_start(k):torch_end(k),aoi_columns(2*k-
1)),[l_trial,1],'nearest');

        ped_eye_cont_seeking=imresize((-
f_tobii(torch_start(k):torch_end(k),aoi_columns(2*k))+f_tobii(torch_start(k):torch_end(
k),aoi_columns(2*k-1))),[l_trial,1],'nearest');
% Combining hits from different AOI columns into one while differentiating them
by sign
        sign=ped_eye_cont_seeking(ped_eye_cont_seeking~=0);
% Removing zeros

        if isempty(sign)
% Failsafe in case of no AOI hits

            sign=(-1).^(0:2*no_of_reps-1)';
% Dividing trial into six equal chunks with alternating sign corrections

        end

        corr=imresize(sign,[l_rawgyro,1],'nearest');
% Sign correction for left-right zero drift in crossing trials
        cum_zero_drift(:,2:3)=bsxfun(@times,cum_zero_drift(:,2:3),corr);
% Applying the sign corrections
        corr=imresize(sign,[l_trial,1],'nearest');
% Sign correction for left-right eye direction in crossing trials
        eye_angles=eye_angles.*corr;
% Applying the sign corrections

    end

%     gyro=gyro-cum_rot_drift;
% Rotational drift correction

    gyro=gyro-cum_zero_drift;
% Zero drift correction

```



```

        gyro=imresize(gyro,[l_trial,3],'nearest');
    % Resampling gyro data from 95 hz to 60 hz

%% Fast fourier transform and bode plot of gyro data
%   gyro_fft=fft(rawgyro);
%   Fast fourier transform
%   figure;
%   plot(gyro_fft);
%   title('Gyro data in frequency domain: FFT');
%   xlabel('Frequency (Hz)');
%   ylabel('Amplitude');
%   gyro_mag=20*log10(abs(gyro_fft));
%   Magnitude in dB
%   l_gyro=length(gyro_fft);
%   df=gyro_hz/l_gyro;
%   fs=df*(0:l_gyro-1)';
%   Stepped frequency
%   figure;
%   loglog(fs,gyro_mag);
%   Bode magnitude plot
%   title('Bode magnitude plot of gyro data: Log-log');
%   xlabel('Frequency (Hz)');
%   ylabel('Magnitude');

%% Pedestrian and driver positions during the trial

    tbl=torch_bearer_start_locations(k,:);
% Torch bearer locations in the trial i.e. assumed same as torch location

    trial_ped_loc=ped_locations(smarteye_start_indices(k):smarteye_end_indices(k),:);
% Pedestrian locations in the trial

    if standing==1 && crossing==0

ped_eye_cont_seeking=imresize(f_tobii(torch_start(k):torch_end(k),aoi_columns(k)),[l_trial,1],'nearest');
% 50 hz tobii AOI hit data upsampled to 60 hz to match smarteye

        if ~isempty(strfind(scenarios{k},'R'))

            trial_ped_loc(trial_ped_loc(:,2)>0,:)=[];
% Removing false positives
            trial_ped_loc=imresize(trial_ped_loc,[l_trial,3],'nearest');
% Resampling to 60 hz

            elseif ~isempty(strfind(scenarios{k},'L'))

                trial_ped_loc(trial_ped_loc(:,2)<0,:)=[];
% Removing false positives
                trial_ped_loc=imresize(trial_ped_loc,[l_trial,3],'nearest');
% Resampling to 60 hz

            end

        end

        psl=trial_ped_loc(1,:);
% Pedestrian location at trial start i.e. same as pedestrian_start_locations(k,:)

    trial_dr_gaze_orig=f_smarteye(smarteye_start_indices(k):smarteye_end_indices(k),gox:goz

```



```

);
    % Driver gaze origin (m) w.r.t IMU

trial_dr_gaze_dir=f_smarteye(smarteye_start_indices(k):smarteye_end_indices(k),gdx:gdz)
;
    % Driver gaze direction (unit vector) w.r.t IMU

    %% Driver eye contact seeking and eye contact occurrence
    norms=sqrt(sum((trial_ped_loc-trial_dr_gaze_orig).^2,2));
% Distance between driver and pedestrian
    rep_norms=repmat(norms,1,3);
% Replicating for matrix multiplication
    cross_prod=cross(trial_ped_loc-trial_dr_gaze_orig,rep_norms.*trial_dr_gaze_dir-
trial_dr_gaze_orig,2);
    % Vector cross product of gaze rays
    dot_prod=dot(trial_ped_loc-trial_dr_gaze_orig,rep_norms.*trial_dr_gaze_dir-
trial_dr_gaze_orig,2);
    % Vector dot product of gaze rays
    dr_eye_cont_seeking=atan2d(sqrt(sum(cross_prod.^2,2)),dot_prod)<=aoi_threshold;
% Angle between driver gaze point and pedestrian location, under threshold for eye
contact

    % eye_cont_grnd_trth1=ped_eye_cont_seeking1 & dr_eye_cont_seeking;
    % Left
    % eye_cont_grnd_trth2=ped_eye_cont_seeking2 & dr_eye_cont_seeking;
    % Right

    eye_cont_grnd_trth=ped_eye_cont_seeking & dr_eye_cont_seeking;
% Eye contact ground truth using AOIs
    trial_timesteps=(1:l_trial)/smarteye_hz;
% Cumulative times of frames in the trial

    % eye_cont_percent(participant_number,2*k-
1)=100*sum(eye_cont_grnd_trth1)/l_trial; % Left
    %
    eye_cont_percent(participant_number,2*k)=100*sum(eye_cont_grnd_trth2)/l_trial; % Right

    %% Driver and pedestrian eye contact seeking binary plots
    % figure;
    % plot(trial_timesteps,abs(ped_eye_cont_seeking),'-b','LineWidth',3);
    % Absolute value to convert -ls in crossing trials
    % title(strcat(scenarios(k),' Pedestrian eye contact seeking'));
    % xlabel('Time [s]');
    % ylabel('Eye contact occurrence');
    % ylim([0,1.5]);
    % set(gca,'ytick',[0,1]);
    % figure;
    % plot(trial_timesteps,dr_eye_cont_seeking,'-r','LineWidth',3);
    % title(strcat(scenarios(k),' Driver eye contact seeking'));
    % xlabel('Time [s]');
    % ylabel('Eye contact occurrence');
    % ylim([0,1.5]);
    % set(gca,'ytick',[0,1]);
    %% Eye contact ground truth binary plot
    % figure;
    % plot(trial_timesteps,eye_cont_grnd_trth,'-m','LineWidth',3);
    % title(strcat(scenarios(k),' Eye contact AOI ground truth'));
    % xlabel('Time [s]');
    % ylabel('Eye contact occurrence');
    % ylim([0,1.5]);
    % set(gca,'ytick',[0,1]);

```



```

%% 3D reconstruction of driver-pedestrian eye interaction scenario
figure('units','normalized','outerposition',[0 0 1 1]);
% Setting figure boundaries
F=struct('cdata',cell(1,l_trial),'colormap',cell(1,l_trial));
% Pre-allocating structure to store individual frames of animation
angles=zeros(l_trial,1);
% Pre-allocating array to store angles between gaze rays
eye_cont=zeros(l_trial,1);
% Pre-allocating array to store eye contact occurrence
% gaze_point_z=zeros(l_trial,1);
% Pre-allocating array to store Z value (m) of pedestrian gaze point
for i=1:l_trial

    go=trial_dr_gaze_orig(i,:);
% Instantaneous driver gaze origin
    pl=trial_ped_loc(i,:);
% Instantaneous pedestrian location
    ped_dist=norms(i);
% Instantaneous distance between driver and pedestrian
    gp=tbl-psl;
% Translation of origin and point of rotation to pedestrian start location from
IMU
    gp=gp*ped_dist/norm(gp);
% Correcting distance of pedestrian gaze point
    theta=gyro(i,2)+eye_angles(i);
% Tobii y-axis gyro data i.e. pedestrian head yaw + eye-in-head direction
    alpha=gyro(i,1);
% Tobii x-axis gyro data i.e. pedestrian head pitch

    % theta=-theta;
% To rotate about -ve z-axis
    % gp=[cosd(theta),-
sind(theta),0;sind(theta),cosd(theta),0;0,0,1]*gp')'+psl;
% Rotation about +ve z-axis followed by translation of origin back to IMU
    % gp=[cosd(theta),-sind(theta),0;sind(theta),cosd(theta),0;0,0,1]*gp';
% Rotation first about +ve z-axis (of IMU)
    % gp=[cosd(alpha),0,sind(alpha);0,1,0;-
sind(alpha),0,cosd(alpha)]*gp')'+psl;
% Rotation second about +ve y-axis (of IMU)

    gp=[cosd(alpha),0,sind(alpha);0,1,0;-sind(alpha),0,cosd(alpha)]*gp';
% Rotation first about +ve y-axis (of IMU)
    gp=[cosd(theta),-sind(theta),0;sind(theta),cosd(theta),0;0,0,1]*gp')'+psl;
% Rotation second about +ve z-axis (of IMU)
    % gaze_point_z(i)=gp(3);
% Pedestrian gaze point Z value
    ped_gaze_ray=[pl;gp];
% Pedestrian gaze ray vector
    gdv=trial_dr_gaze_dir(i,:);
% Driver gaze direction unit vector
    dr_gaze_ray=[go;ped_dist*gdv];
% Driver gaze ray vector

    %% Angle between gaze rays and eye contact
    angles(i)=atan2d(norm(cross(pl-gp,ped_dist*gdv-go)),dot(pl-gp,ped_dist*gdv-
go));
% Instantaneous angle between gaze rays
    if angles(i)<=threshold
% Check if the instant is eye contact
        eye_cont(i)=1;
% Record instant if yes
    end
end

```



```

%% 3D animation
plot3(dr_gaze_ray(:,1),dr_gaze_ray(:,2),dr_gaze_ray(:,3),'-
ro','MarkerSize',28);
% Measured driver->pedestrian eye gaze ray
hold on;
image('CData',f_prius_top_view,'XData',[-prius_dim(1)/2-0.5,prius_dim(1)/2-
0.5],'YData',[-prius_dim(2)/2,prius_dim(2)/2]);
% Overlay prius top view image in plot
plot3(ped_gaze_ray(:,1),ped_gaze_ray(:,2),ped_gaze_ray(:,3),'-
bo','MarkerSize',25);
% Measured pedestrian->driver eye gaze ray
hold off;
view(0,90);
% Top view of 3D plot

% %
text(dr_gaze_ray(:,1),dr_gaze_ray(:,2),dr_gaze_ray(:,3),{'D','P*'},'VerticalAlignment',
'top','HorizontalAlignment','right');
% %
text(ped_gaze_ray(:,1),ped_gaze_ray(:,2),ped_gaze_ray(:,3),{'P','D*'},'VerticalAlignmen
t','top','HorizontalAlignment','right');

text(dr_gaze_ray(1,1),dr_gaze_ray(1,2),dr_gaze_ray(1,3),'D','VerticalAlignment','top','
HorizontalAlignment','right');
% Marking the driver's position in the plot

text(ped_gaze_ray(1,1),ped_gaze_ray(1,2),ped_gaze_ray(1,3),'P','VerticalAlignment','top
','HorizontalAlignment','right');
% Marking the pedestrian's position in the plot
title(strcat(scenarios(k),' Time=',num2str(round(i/smarte_ey_hz,2)),' s'));
% Scenario and trial time
xlim([-2.5*max_ped_dist,2.5*max_ped_dist]); % 2.5 x
ylim([-max_ped_dist,max_ped_dist]); % 1 x
zlim([-max_ped_dist,max_ped_dist]); % 1 x
xlabel('X [m]');
ylabel('Y [m]');
zlabel('Z [m]');
grid on;

% legend('Driver gaze ray','Pedestrian gaze ray');
% Legend causes plot lag

% [xs,ys,zs]=sphere(50);
% Generate a sphere
% rs=0.125;
% Radius of the sphere
% xs=xs*rs+trial_ped_loc(i,1);
% ys=ys*rs+trial_ped_loc(i,2);
% zs=zs*rs+trial_ped_loc(i,3);
% Option 1:
% fvc=surf2patch(xs,ys,zs);
% Create a patch using the sphere parameters
%
patch('Faces',fvc.faces,'Vertices',fvc.vertices,'FaceColor','none','EdgeColor',[0.8,0.8
,0.8]);
% Plot the 3D patch
% Option 2:
% surface(xs,ys,zs,'FaceColor','none','EdgeColor',[0.8,0.8,0.8]);
% Create a surface using the sphere parameters
%
patch('Faces',prius_faces,'Vertices',prius_vert,'FaceColor',[0.94,0.97,1],'FaceAlpha',
0.5);
% Plot the 3D surface

```



```

        h=findobj('FontName','Helvetica');
    % Obtain plot font object
        set(h,'FontSize',24);
    % Increase font size
        drawnow;
    % Flush current frame to display (fixes frame leak into next animation)
        F(i)=getframe(gcf);
    % Grab the current frame

    %         pause(1/smarteye_hz);
    % Use 0.05 for a smooth animation

    end

    %     eye_cont_percent(participant_number,k)=100*sum(eye_cont)/l_trial;

    %     tpr=nnz(eye_cont_grnd_trth&eye_cont)/nnz(eye_cont_grnd_trth);
    % Hit rate
    %     fpr=nnz(~eye_cont_grnd_trth&eye_cont)/nnz(~eye_cont_grnd_trth);
    % False positive rate

    %     mycells(2*count-1,k)=tpr;
    % Store true positive rates
    %     mycells(2*count,k)=fpr;
    % Store false positive rates

    %% Pedestrian gaze point Z value plot
    %     figure;
    %     plot(trial_timesteps,gaze_point_z,'-k','LineWidth',3);
    %     title(strcat(scenarios(k),' Z value of pedestrian gaze point'));
    %     xlabel('Time [s]');
    %     ylabel('Z [m]');
    %     ylim([0,3]);

    %% Angle between gaze rays plot
    %     figure;
    %     plot(trial_timesteps,angles,'-g','LineWidth',3);
    %     title(strcat(scenarios(k),' Angle between gaze rays'));
    %     xlabel('Time [s]');
    %     ylabel('Angle [deg]');

    %% Algorithmic eye contact detection binary plot
    %     figure;
    %     plot(trial_timesteps,eye_cont,'-m','LineWidth',3);
    %     title(strcat(scenarios(k),' Eye contact detection via algorithm'));
    %     xlabel('Time [s]');
    %     ylabel('Eye contact occurrence');
    %     ylim([0,1.5]);
    %     set(gca,'ytick',[0,1]);

    %% Turn animation into a video
    %     vidname=strcat('EyeContactAnimation-',num2str(participant_number),'-
',scenarios{k});
    % Participant number and scenario name
    %     vidname=strcat(num2str(participant_number),'-',scenarios{k});
    %     vid=VideoWriter(vidname,'MPEG-4');
    % MP4 format for video
    %     vid.FrameRate=24;
    % Use 'smarteYE_hz' for real time
    %     vid.Quality=100;
    % Maximum quality

```



```

        open(vid);

        for i=1:l_trial

            writeVideo(vid,F(i));
% Append to video frame by frame

        end

        close(vid);

    end

end
% count=count+1;
% Updating loop variable
%
% end
%
% mycells=round(mycells,4);
% Rounding TPR and FPR values to four decimal places
h=findobj('FontName','Helvetica');
% Obtain plot font object
set(h,'FontSize',24);
% Increase font size

%% Notes
% Tobii gyro x is pedestrian head pitch or nodding
% Tobii gyro y is pedestrian head yaw or looking sideways, left is +ve and right is -ve
% Tobii gyro z is pedestrian head roll or shaking

```