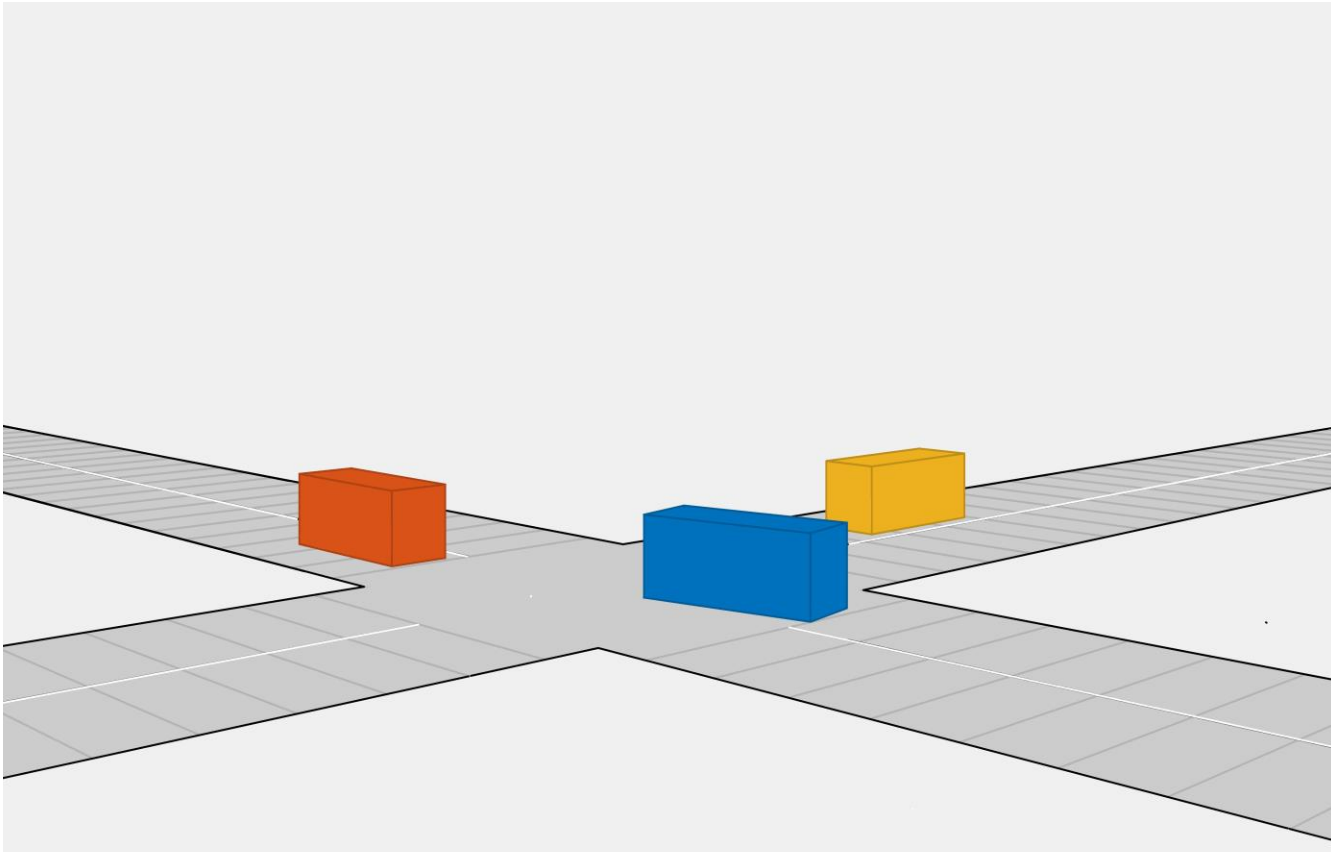
CHALMERS
UNIVERSITY OF TECHNOLOGY



Driver Glance Behaviour in Intersection Crashes: A SHRP2 Naturalistic Data Analysis

Master’s thesis in Master’s Program in Automotive Engineering

Tejas Chandran

Department of Mechanics and Maritime Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

MASTER’S THESIS IN MASTER’S PROGRAM IN AUTOMOTIVE ENGINEERING

Driver Glance Behaviour in Intersection Crashes: A SHRP2 Naturalistic Data Analysis

Tejas Chandran

Department of Mechanics and Maritime Sciences
Division of Vehicle Safety
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2019

Driver Glance Behaviour in Intersection Crashes: A SHRP2 Naturalistic Data Analysis
Tejas Chandran

Master's Thesis 2018:95
Department of Mechanics and Maritime Sciences
Division of Vehicle Safety
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: + 46 (0)31-772 1000

Cover:
Intersection sample generated using the Automated Driving System Toolbox package in MATLAB version 2017b.

Chalmers Reproservice / Department of Mechanics and Maritime Sciences
Göteborg, Sweden 2019-01-08

Driver Glance Behaviour in Intersection Crashes: A SHRP2 Naturalistic Data Analysis
Master’s thesis in Master’s Program in Automotive Engineering (MPAUT)
Tejas Chandran
Department of Mechanics and Maritime Sciences
Division of Vehicle Safety
Chalmers University of Technology

Abstract

Road traffic crashes are one of the biggest problems in modern history. Over recent years, several attempts have been made to gain an in-depth understanding of crashes and their causation. The introduction of unobtrusive data collection methods in the form of naturalistic driving studies has given traffic safety researchers an unprecedented level of insight into the vehicle and driver during these safety critical situations. This research was started with the aim of utilising this data from naturalistic driving studies to reconstruct crashes in intersections to gain a better understanding of the driver’s gaze behaviour.

This was done by first developing a toolbox within the software environment of MATLAB. The toolbox used the vehicle kinematic data from the naturalistic driving study SHRP2 or the Second Strategic Highway Research Project. The toolbox was designed to transform the driver’s gaze from vehicle coordinates to intersection coordinates. The target vehicle was approximated into the simulation environment with the help of the video data and manual annotation. Factors such the driver’s gaze directions, eyes on target and intersection gaze timings were obtained. Well known psychological models such as the Situational Awareness, SEEV model and the more recent Predictive Processing model were evaluated to understand the findings.

The results of the research showed that drivers in most of the events analysed had possibly seen and continued tracking the threat from the theoretical point of no return until the crash itself. These results were intriguing in that the drivers were not observed to engage in evasive manoeuvres until too late. A hypothesis was developed with the help of the predictive processing model to understand and explain this behaviour.

Keywords: Driver behaviour analysis, naturalistic driving studies, intersection, crash reconstruction, predictive processing, SHRP2

Contents

Preface

This thesis was carried out under the umbrella of the industry/academic research project Quantitative Driver Behaviour Modelling for Active Safety Assessment Expansion or QUADRAE (2016). The research was conducted at SAFER, Vehicle and Traffic Safety Centre at Chalmers University of Technology, Lindholmen, Göteborg, Sweden. The data used in the thesis research was provided by Volvo Car Corporation, Göteborg as a part of the QUADRAE project. The SHRP2 naturalistic data is maintained and processed by Virginia Tech Transportation Institute.

Acknowledgements

I would like to thank Jonas Bärgman, my examiner and supervisor, for his support and guidance through the course of my thesis. His insights and experience have been invaluable in my growth as a researcher and as an engineer. I would like to thank Emma Tivesten from Volvo Cars Corporation for providing the data used for this thesis. I would like to thank Esko Lehtonen for his inputs. I thank Erik Svanberg for helping with the extraction of data from the field operational test rooms. I thank Giulio Bianchi Piccinini for his support and encouragement during my time at SAFER. I thank everyone else at SAFER for the wonderful environment they created; it was a pleasure to work and learn from everyone there.

And finally, I thank my family and friends who have encouraged me and been the supporting hand during my master's degree.

Göteborg, 2019-01-08

Tejas Chandran

List of Figures

List of Tables

Notations

ADS	Automated Driving System
CAN	Controller Area Network
GPS	Global Positioning System
$IGRT$	Intersection Gaze Release Time
$LTAP/OD$	Left Turn Across Path – Opposite Direction
NDS	Naturalistic Driving Study
$NHTSA$	National Highway Traffic Safety Administration
PNR	Point of No Return
SA	Situational Awareness (model)
SCP	Straight Crossing Path
$SEEV$	Saliency, Effort, Expectancy, Value (model)
$SHRP2$	Second Strategic Highway Research Project
SV	Subject Vehicle
TV	Target Vehicle
$YPLL$	Years of Potential Life Lost

1 Introduction

1.1 Background

Road traffic crashes are currently one of the leading causes for *Year of Potential Life Lost (YPLL)*. YPLL is a metric that quantifies the impact of premature deaths, with an increased weight being given to younger individuals, thereby, allowing for an emphasis of the impact of premature deaths on the society and economy of a nation. While medical illness such as heart diseases and cancer statistically claim more lives, most of its victims are the older population. The use of YPLL allows for a better representation of the population (usually from 15-65 years of age) affected by crashes (Gardner & Sanborn, 1990). This is an important term in the analysis of road traffic crashes: in a study by the World Health Organization, it was found that approximately 1.2 million lives were lost to road traffic crashes in the year 2013, a staggering number which has remained fairly constant since 2007. The majority of these happen in lower and middle-income nations, costing them up to 3% of their GDP (World Health Organisation, 2015). In an attempt to curb this alarming number of casualties, vehicle and traffic research has been focusing on understanding the driver behaviours and factors that lead to crashes in order to develop systems that can recognize, warn and assist drivers in preventing and mitigating such events.

To understand crashes, and their underlying cause, a crash event can be broken down into three phases:

i. Pre-crash phase: The moments leading up to the crash, including the "*point of no return*" or the point in time beyond which the crash is inevitable. Evasive manoeuvres to avoid the crash or mitigate the severity of the crash after the point of no return can be performed during this phase. Many active safety systems such as Forward Collision Warning (FCW), and Autonomous Emergency Braking (AEB) are designed to recognize and react to prevent the event from evolving into the next phase (Battelle Memorial Institute, 2007; Brännström, Sjöberg, & Coelingh, 2008). It is the actions performed by the driver during this phase which is of particular importance in the field of crash analysis and prevention (Guo, Klauer, McGill, & Dingus, 2010).

ii. Crash phase: The crash is defined as '*Any contact with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated. Includes other vehicles, roadside barriers, objects on or off of the roadway, pedestrians, cyclists, or animals* (Dingus, et al., 2006, p. xviii) '.

iii. Post-crash phase: The ensuing events that occur as a consequence of the impact between the two vehicles, such as occupant movement and interactions with the vehicle structures.

Recent developments in active safety and automated driving systems have been directed at addressing the pre-crash phase to try to prevent the situation from evolving into a crash or else reduce the severity of the crash if the crash is inevitable. Studies have shown that driver error as the cause in more than 93% of all crashes whereas environmental and vehicle factors weighed in at 34% and 13% of crashes respectively (Treat, et al., 1979; Hendricks, Freedman, Zador, &

Fell, 2001). The staggering number of crashes because of driver error provokes the need to develop an in-depth understanding of the driver behaviours during the different stages of driving and most importantly in the moments leading up the crash. Several methods have been employed to gain an insight into this valuable data through computer simulations, track tests, naturalistic driving studies etc (Smith, Bjelkemyr, Bärgman, Johansson, & Lindman, 2009; Nobukawa, Barnes, Goodsell, & Gordon, 2009; Galpin, Underwood, & Crundall, 2009).

This research will deal extensively with naturalistic driving data and driver behaviour models. In the following sections, a brief literature review on naturalistic driving studies, intersection events, driver behaviour, crash reconstruction, and their relevance to this research will be discussed.

1.2 Naturalistic Driving Studies

To understand a driver's behaviour under various scenarios, it is important to understand their driving behaviours in their everyday driving situations. One means to achieve this is through Naturalistic Driving Studies (NDS), which is defined as *"A study undertaken to provide insight into driver behaviour during every day trips by recording details of the driver, the vehicle and the surroundings through unobtrusive data gathering equipment and without experimental control* (van Schagen, et al., 2011, p. 10-11)." The biggest advantage of NDS over other methods is that the majority of the drivers involved in the study drive their own vehicles, carrying out their everyday activities and routines. Further, as the measurement and monitoring devices (such as cameras) are unobtrusive, drivers soon get accustomed to them and forget about their presence. This allows a more *natural* capture of driving behaviour and errors, giving researchers the detailed insight into behaviours and actions under normal driving situations. This also allows a natural capture of the moments leading up to the crash (Neale, et al., 2002).

The 100 Car NDS was the first such study conducted over a 12-month period with 100 young volunteers and about 3,000,000km of driven. The study was devised as a method to capture the human factors involved in the decision making and threat response in safety-critical or crash-imminent situations. The drivers were provided with no special instructions and drove independently, without an experimenter present during the trials. These results have been very important in the development of active safety systems and have supported driver behaviour models developed under more *experimental* settings (Neale, et al., 2002; van Schagen, et al., 2011). The eUropean naturalistic Driving and Riding for Infrastructure and Vehicle safety and Environment or UDRIVE (Jansen, et al., 2017) and the US Second Strategic Highway Research Project or SHRP2 (Transportation Research Board, 2001) are examples of large-scale NDS conducted in recent years.

SHRP2 was developed in continuation to the *TRB Special Report 260: Strategic Highway Research: Saving Lives, Reducing Congestion, Improving Quality of Life* (2001). A special committee was created to evaluate the First Strategic Highway Research Project which was conducted in the late 1980s through early 1990s. The committee was created by the Transportation Research Board in response to the U.S. Congress's request to *"conduct a study to determine the goals, purposes, research agenda and projects, administrative structure, and fiscal needs for a new* *strategic highway research program* (Bulger & Crane, 1998)*.*" It concluded that a second study would be necessary to overcome the limited ability of existing resources and data for research and development to achieve the following goals: *"Renewal: Accelerating the Renewal of America's Highways; Safety: making a Significant Improvement in Highway Safety; Reliability: providing a Highway System with Reliable Travel Times; and Capacity: Providing Highway Capacity in Support of the Nation's Economic, Environmental, and Social Goals* (Transportation Research Board, 2001)". The SHRP2 was authorized in 2005 and set up to further address the following specific goals: understand driver interaction with factors such as the vehicle, traffic infrastructure, environment and roadway, and to assess the crash risk with respect to these factors and their interactions (Antin, 2011). *"The SHRP 2 Naturalistic Driving Study (NDS) was the largest and most comprehensive study of its kind ever undertaken* (Victor, et al., 2015)*."* The study was conducted with over 3000 volunteers aged between 16-98 and over 80,467,200km driven. Apart from vehicle kinematics such as speed, acceleration, and braking, the NDS also contained video data from *"Forward", "Rear"* and *"driver, face and hands",* as well as data from sensors such as radars. A total of approximately 1 petabyte of data containing more than 1 million hours of video and 3900 years of data was collected over a period of three years (Antin, 2011; Blatt, et al., 2015; Victor, et al., 2015). The SHRP2 NDS, with both crashes and near-crashes, is the data source used in this research. The version of SHRP2 used contains forward camera video, in addition to time series vehicle kinematics and glance annotations. Both the video data and kinematic data will be used together to reconstruct events of interest within the virtual environment of MATLAB.

1.3 Intersection Crashes and Near-Crashes

Although NDS contains vast troves of driving scenarios and data, it often contains very few crashes, as a result, there is a requirement for the use of other safety-critical events as a surrogate when evaluating driver behaviour. Studies have suggested that near-crashes can be used as a representative for crashes (Guo, Klauer, McGill, & Dingus, 2010). Hence, it is important to understand the difference between crashes and near-crashes. A near-crash can be defined as "*Any circumstance that requires a rapid evasive manoeuvre by the participant vehicle or any other vehicle, pedestrian, cyclist or animal to avoid a crash. A rapid evasive manoeuvre is defined as steering, braking, accelerating or any combination of control inputs that approaches the limits of the vehicle's capabilities* (Guo, Klauer, McGill, & Dingus, 2010, p. ii)". While crashes and near-crashes are different by definition, studies have shown that the factors leading to these safety critical situations are often similar. The difference between them largely depends on how the driver reacted and behaved. Using near-crashes to predict crashes and in risk assessment can lead to biased results. However, this bias can often and at least partially be accounted for during analysis. This significantly improves the precision of estimation for crashes and better development of driver models (Guo, Klauer, McGill, & Dingus, 2010; Bärgman, Lisovskaja, Victor, Flannagan, & Dozza, 2015; Bärgman, 2016).

By only investigating crash statistics in the USA for the year 2015, it was found that approximately 51% of all crashes where an injury was reported occurred at or near an intersection. In fact, these crashes account for nearly 29% of all fatal crashes (NHTSA, 2017). This has commonly been attributed to a variety of factors including the extremely high mental workload on the driver as s/he enters the intersection environment, which could include keeping track of vehicles, pedestrians and cyclists from multiple directions as well as understanding the intentions of these road users (Rumar, 1990; Tay & Rifaat, 2007; Bougler, Cody, & Nowakowski, 2008; Choi, 2010). Often, there are also a variety of other factors that can affect the situation such as, intersection design, traffic signals, environment and even the subject vehicle (Björklund & Åberg, 2005). Age was also found to be a contributing factor with older drivers being more prone to making errors and have relatively slower reaction times (Lord, Schalkwyk, I., & Chrysler, 2005; Tay & Rifaat, 2007; Romoser, Pollatsek, Fisher, & Williams, 2013). Most intersection crashes and near-crashes are understood to be the result of one or more of these factors.

Other factors include failure to comply with traffic regulations or even poor adoption of traffic rules. The behaviour the driver expects from other road users were also found to be a contributing factor. Drivers were found to expect other road users to conform with both formal and informal rules of the road environment (for example right of way) (Kulmala, 1990). It was found that drivers often prioritize other vehicles over pedestrians or cyclists (Björklund & Åberg, 2005). In a study conducted in Norway, it was found that many crashes occurred in intersections because the driver driving straight through the intersection expected the vehicle which was turning to yield but did not do so (Aust, Fagerlind, & Sagberg, 2012). Björklund & Åberg (2005) and Aust, Fagerlind, & Sagberg (2012) suggested that the drivers expect other drivers to conform to rules of the road even where there were no means to enforce the rules. In another study conducted in China at unsignalized intersections, it was found that drivers approached intersections without any intention of stopping but made decisions whether to yield or not based on their assumption of the other driver's intent (Liu, Lu, Wang, & Zhang, 2014). Further, Bärgman, Smith, & Werneke (2015) found drivers, who had the right of way, were often found to slow down to give additional time to allow the turning vehicle to pass through when they felt an encroachment was imminent. This study explains that this behaviour was linked to the driver's comfort zone, which is the spatiotemporal area around their vehicle within which they feel safe.

Among intersection crashes, a study by Tay & Rifaat (2007) found that X-junctions were prone to more severe crashes than T-junctions, with 1.18 times the fatality risk. Head-on and Head-to-side crashes were also observed to be more severe than head-to-rear types of crashes. The risk of fatal crashes was also found to be higher with an increase in the number of vehicles involved in the intersection crash; the more the vehicles, the greater the risk of multiple impacts. Head-on and head-to-side impacts, commonly observed in intersection crashes, result in higher crash severity and fatality risk due to the greater difference in kinetic energies and velocities between the two vehicles as compared to rear-end crashes. There is also a reduced area for the energy to dissipate in the side impacts, making the occupants in the struck car relatively more vulnerable to injuries (Tay & Rifaat, 2007). Analysing a report by NHTSA (2008), Choi (2010) found *inadequate surveillance* and *false assumption* were important types of driver errors associated with intersection crashes. In contrast, *aggressive driving* and *poor driver performance (poor directional control etc.)* were the leading cause of non-intersection crashes. It was found that inadequate surveillance could result in nearly six times higher likelihood of a crash in an intersection than other locations. The other factors identified in the study were *turned with an obstructed view, illegal manoeuvre, internal distraction* and *misjudgement of the gap or other's speed. Turned with an obstructed view* was found to be almost 335 times more likely to result in a crash in an intersection as compared to non-intersection (Choi, 2010).

To further understand crashes at intersections, they are usually categorized as crossing path and non-crossing path crashes. Crossing path crashes occur when one vehicle cuts across the path of another, either from the lateral or opposite direction (Najm, Smith, & Smith, 2001). Crossing path crashes made up 25% of all crashes, and 44% of intersections crashes based on General Estimates System (GES) statistics for the year 2000 (Ragland & Zabyshny, 2003). Najm, Smith, & Smith (2001) further categorised crossing path crashes into five major types namely, a) Left Turn Across Path – Opposite Direction (LTAP/OD), b) Left Turn Across Path – Lateral Direction, c) Left Turn Into Path – merge conflict (LTIP) d) Right Turn Into Path – merge conflict, and e) Straight Crossing Path. Among these, SCP and LTAP/OD contributed to the majority of crossing path crashes at approximately 30% and 28% respectively (Najm, Smith, & Smith, 2001). This research will focus on these two crashes types, which can be understood through Figure 1 and Figure 2 below. The blue box represents the subject vehicle (SV) or the vehicle whose data is known through the NDS data and the red box represents the target vehicle (TV) whose information is estimated using forward camera feeds; this scheme will be followed through the rest of this document. All references made through text and figures are in perspective of the SV.

From a study perspective, LTAP/OD can be seen from two different contexts, the first where the SV performs the left turn across the path of the oncoming traffic and a second where the SV drives straight through the intersection towards the TV that is performing the left turn manoeuvre.

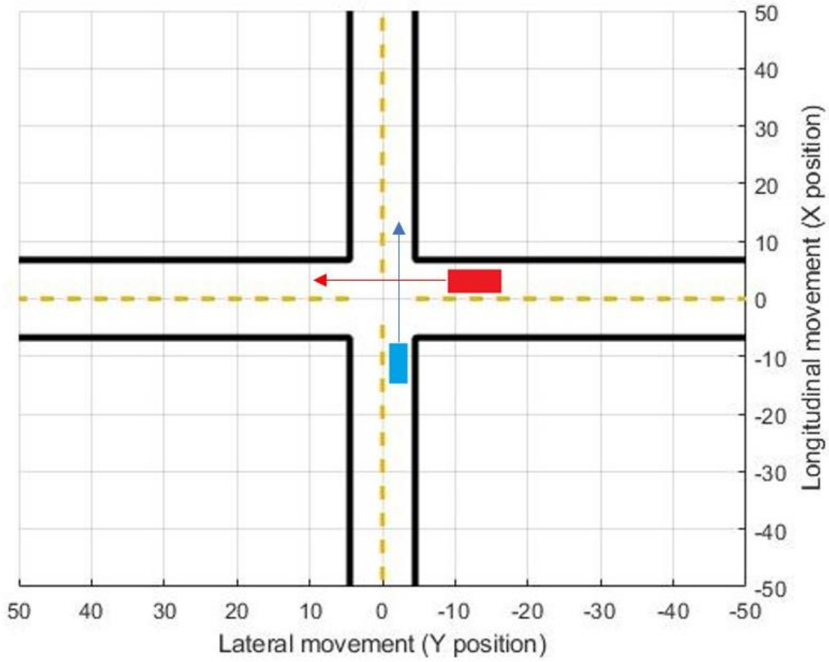


Figure 1: Representation of Straight Crossing Path (SCP)

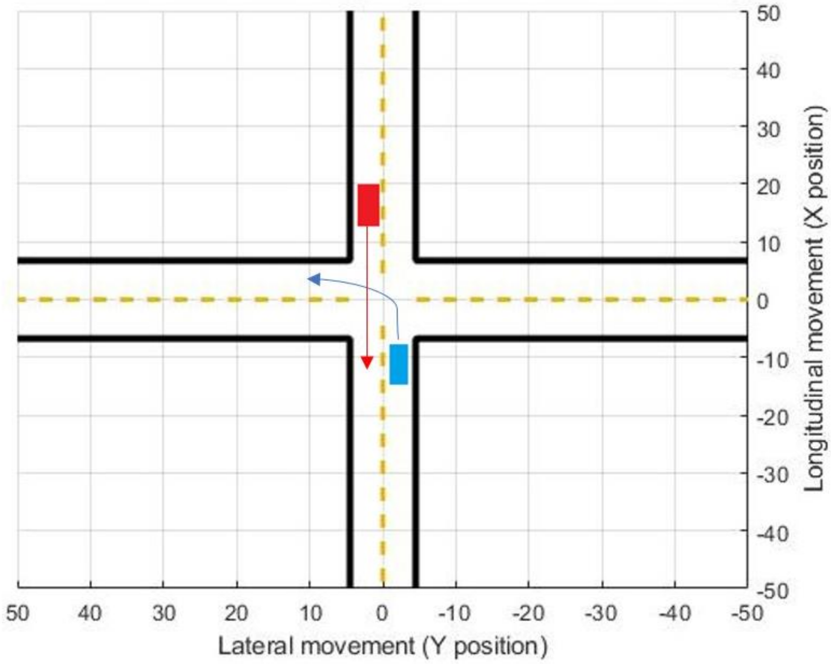


Figure 2: Representation of Left Turn Across Path/Opposite Direction (LTAP/OD)

Crashes at intersections are complex with many underlying factors. These include the driver, the environment, other road users and the driver's interpretation of all of these.

1.4 Glance Behaviour

As discussed in the previous section, studies have investigated and proposed the use of near-crashes as a surrogate for crashes in NDS analysis. However, a recent study by Seppelt et. al. (2017) using the 100 car NDS found crashes and near-crashes to have critical differences at intersections in the way the driver visually scans and processes the information. It was found that in near-crash events, drivers had a more consistent and balanced glance behaviour, with longer fixations on- and off-road objects, allowing vital information to be better processed by the driver (Seppelt, et al., 2017). This appears to be in line with the findings by Choi (2010) explored in the previous section.

Driving, in general, requires large amounts of mental resources, with the highest workload being on visual stimulus (Seppelt, et al., 2017). The driver has visual information from several channels to process simultaneously: from vehicle control (speed, lane) to awareness of other road users (pedestrians, cyclists, other vehicles, animals) and traffic markers (signals, road signs etc.). S/he must be able to use all this information to make decisions while driving. The amount of information at any time is dependent on the specific traffic scenario; the gaze and glance behaviour vary accordingly (Land, 2006). Land & Lee (1994) suggested that when driving through a bend, drivers were found to focus a lot more along the successive tangents of the curve they were driving along. However, if an additional stimulus is inserted, such as a cyclist, the driver would need to alternate his gaze between the preferred fixation point of the tangent to the cyclist multiple times. During each such switch on the object of fixation of the gaze, the information retrieved from the previous source is assumed to be kept "*on hold*". Thus, Land

(2006) suggested that the control of the vehicle is not just based on the object the driver is actively looking at, but also what s/he has stored in a buffer state from the previous instant. Driving in urban situations and intersections are even more complex. He explains that looking in the right direction is not enough. S/he must know what to look for before s/he can make a control decision. For example, whether to look at the direction of travel or the cyclist on the road. Land (2006) suggested this was done by comparing the relative position of the object/area of interest with its position stored in the driver's memory between each glance. In a study by Werneke & Vollrath (2014), it was observed that as drivers approached the intersection, they started scanning for gaps in the oncoming traffic. The authors suggested that the drivers' primary visual focus was on the traffic and hence gave less attention to the pedestrians; in contrast, pedestrians were more relevant when the driver was waiting at the intersection, leading to longer gazes in this direction. When accelerating from the intersection, the priority the driver gave towards the pedestrians was found to be dependent on the traffic density. This is a good link to previous studies, where the drivers adapt their gaze and the way they process information depending on the driving scenario at that moment (Summala, Pasanen, Räsänen, & Sievänen, 1996; Land, 2006; Werneke & Vollrath, 2012).

1.4.1 Psychological Models

Researchers have over the years tried to explain this behaviour in drivers by using human behaviour models. Among the many models developed, this text will cover four behaviour models, namely, SEEV or *S*aliency, *E*ffort, *E*xpectancy, *V*alue model (Wickens, Helleberg, Goh, Xu, & Horrey, 2001), NT-SEEV model - an extended version of the SEEV model (Wickens, McCarley, & Steelman-Allen, 2009), the SA or *S*ituational *A*wareness model (Endsley, 1987; 1995), and Predictive Processing models (Engström, et al., 2018).

The SEEV model, proposed by Wickens, Helleberg, Goh, Xu, & Horrey (2001) assumes that visual data is acquired by four factors. These four factors being: a) salient events or areas of interest in the visual field, where each of these events can be individually identified and captured by virtue of a visual detail such as the colour, size, and so on; b) the effort required to switch gaze between each of the areas of interest. Wickens, Hollands, Banbury, & Parasuraman (2015) suggested that the movement of the eye has a *"cost"*. The more extensive the search for an area of interest, the more fatigue it would cause, that is, it would cost more. The authors explain that humans become more hesitant when extended searches require neck rotation and/or even movement of the body. c) the expectancy to find information and a corresponding change in that area of interest. Senders (1964) suggested that this change is a function of time called bandwidth. Humans sample for information and expect to find information in the visual field. Wickens, Hollands, Banbury, & Parasuraman (2015) explains this by suggesting that we look where there is "action" and expect a certain change from them. In other words, the expectancy can be said to be dependent on the bandwidth of the visual field. For example, a driver would expect frequent changes to occur when driving through heavy traffic. d) and finally, how useful this information is to the observer. A person would act based on the relative importance of these areas of interest (Wickens, Hollands, Banbury, & Parasuraman, 2015). The authors explain this with an example of driving along a highway, where the road ahead has a low frequency of change whereas the billboards that pass by along the sides of the road have a higher rate of change. The road ahead, however, would have higher relative importance between the two.

Using this explanation of the SEEV model, consider the following example of a driver approaching an intersection. As the driver nears the intersection, s/he would scan the environment for the different areas of interest. These areas of interest can include the speedometer, the traffic lights, pedestrians, traffic from the opposite direction, the road in the direction s/he is heading and so on. As the number of areas of interests increases, more effort would be required from the driver, as these are all competing for the same resources/attention from the driver (Recarte & Nunes, 2000). Recarte & Nunes (2003) suggested this could be a reason for crashes where the driver "*looked but did not see*"; a term coined by Staughton & Storie (1977). As the driver shifts his/her attention between these areas of interests, s/he would then possibly make a mental model of which of these are more important. As s/he switches between them, the driver also has a certain expectancy in how they would behave, for example, the driver would expect the other vehicle to yield if s/he has the right of way. Lemonnier, Brémond, & Baccino (2015) explains this by comparing two scenarios, one where the driver had to yield/stop to another vehicle before passing the intersection and a second where the other vehicle yielded to allow the subject driver to pass. When the driver had to yield, s/he had a higher workload because the driver now looked further down the road for other vehicles behind the vehicle for whom s/he currently yielded. The driver would use this additional information to look for a suitable gap in the traffic. The authors explain that in the scenario where the other vehicle yielded for the subject driver, the driver could reallocate his/her attention onto the other parts of the road, such as the direction of intended travel, implying that the stopped vehicle is no longer as important.

Wickens, McCarley, & Steelman-Allen (2009) proposed an extended version of the SEEV model called the NT-SEEV, where NT denotes *Noticing Time.* The authors suggested that a person would notice a change faster if this visual event occurs closer to where the gaze is directed towards at that point in time. Applying this into a traffic scenario, a driver would be more likely to notice a flashing visual warning faster for forward collision if it is, say on the windshield or higher than if were lower down in say, the instrument cluster. In another example, it could be possible the driver would notice the threat of TV approaching from the road ahead sooner than s/he would if the TV was approaching from the left/right of the intersection. It would be interesting to see if this can be observed during this research.

The SA model (Endsley, 1987; 1995) is a very influential model used to understand human behaviour in dynamic decision-making environments. Endsley offered the general definition for SA as "*Situation awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future*(Endsley, 1995, p. 36)." This definition is broken by the author into three hierarchical levels, the lowest being *perception (level 1 SA),* followed by *comprehension (level 2 SA)* and at the highest-level *projection (Level 3 SA)*. In other words, "*What? So What? What now?* (Tenney & Pew, 2006) "

It was suggested that higher levels are affected directly by the input from lower levels. Thus, if one fails to notice an event, they cannot fully comprehend the situation (Wickens, Hollands, Banbury, & Parasuraman, 2015). The authors suggest unless one notices a change, they cannot know there was a change, and hence would not be able to understand the situation.

To understand SA with an example in the context of this research, consider a driver approaching an intersection to perform a left turn across path (opposite direction). At level 1 SA, the driver must be aware of their own vehicle's position, dynamics, other vehicles and their dynamics (Endsley, 1995). The driver would then use this information to understand the present situation they are in or level 2 SA. So, another driver approaching the intersection at a constant velocity, with no turn indicator blinking will most likely drive straight through the intersection. The driver can use this understanding of the current situation to make a prediction of the future state of the situation or level 3 SA. For instance, *if I slow down, the other driver can pass through and after which, I can perform the left turn safely*. Wickens (2000) argued that SA cannot address change instantly. He argues that in order to quickly adapt to change and prevent a crisis, a person would have to have predicted this negative outcome beforehand. In other words, the person would have to continue predicting the future state of the situation even if the current situation does not require it. For example, while driving, one does not expect to crash, but must be prepared to react appropriately should the need arise. Wickens (2000) suggested that this could be achieved by maintaining level 3 SA at all times.

This argument lays the introduction for the final behavioural model investigated in this research, that is, the predictive processing model (Engström, et al., 2018; Clark, 2013). This model proposes that the brain is a *statistical* organ which is constantly attempting to minimise the error between the predicted model and perceived information, both instantly and over time (Engström, et al., 2018). That is, as Clark (2013) explains, the brain is constantly trying to understand the world around it by matching the inputs from the different senses with a top-down expectation or prediction. He explains that these predictions reflect what the brain already knows and also does not know (uncertainties) in the environment. This is further explained by Engström et. al. (2018) that the brain then tries to minimise the deviation, if any, by two steps - *perception* and *action*. They explain that the brain would first update its prediction model based on the feedback it has been receiving from the sensory input and then acts to minimise this error, thus attempting to minimise the error. Thereby the authors suggest that in predictive processing, the brain tunes itself not only based on the sensory input but also its prediction of the sensory input.

One of the advantages of the predictive processing model is that it attempts to provide a unifying theory to understand how the brain behaves. The model not only gives an innovative approach to the way the human brain acts based on biological facts, but it also unifies the other behavioural models such as SA and SEEV within its scope (See Engström et. al. (2018)). This research will not, however, go through the specifics of the unifying theory itself, but will attempt to use this idea to explain and understand the driver behaviour in intersection crashes over the rest of this document.

1.4.2 Intersection Gaze Release Time

A final concept that is considered in this research is a metric known as Intersection Gaze Release Time (IGRT). IGRT was novel concept first introduced in a study conducted in Sweden, where the authors proposed that drivers entering the intersection analyse the situation by shifting their head/gaze across the different parts of the intersection for events of potential encroachment, or threats, and returns gaze back to the road ahead (or the future path) once they have successfully completed this diagnosis and have ruled out that there is no imminent encroachment of another vehicle (Smith, Bjelkemyr, Bärgman, Johansson, & Lindman, 2009). IGRT was then offered the definition by Jansen, et. al. (2017) as "*the time from when a driver looks towards an area of interest (e.g., an area of potential threat) the last time before he/she enters the encroachment zone, until the driver enters the encroachment zone* (Jansen, et al., 2017, p. 55)" where the encroachment zone is the area where the trajectory of the ego vehicle entering the intersection overlaps that of the target vehicle (Jansen, et al., 2017).

1.5 Crash reconstruction

A crash reconstruction is a method by which a researcher can gain an insight into a part or the whole crash to understand the probable causes and the sequence of events that led up to the eventual crash (Martinez, 1989). In other words, it allows the researcher to understand the entire pre-crash event, specific elements of the pre-crash such as glance or steering actions and sometimes the in-crash events such as occupant kinematics; vehicle kinematics; crash severity, etc. from historical data collected from the crash site, and when available, interviews with the involved drivers (Smith & Tsongos, 1986; Martinez, 1989; Martins, Ribeiro, Neves, & and Brito, 2016). This has long been used by law enforcement and traffic authorities for legal purposes. For example, it has been carried out by referring to witness statements, physical inspections, crash reports and logs from devices such as the Event Data Recorder (EDR) (Martinez, 1989). The National Highway Traffic Safety Administration (NHTSA) used the software WinSMASH to compute the delta-V using EDR data from the National Automotive Sampling System/Crashworthiness Data System (NASS/CDS) and for the crash tests performed on-site (Sharma, Stern, Brophy, & Choi, 2007).

With the development and success of NDS, data backed by video feeds are now available to researchers. This has significantly widened the horizon for crash reconstructions. The availability of glance locations in these databases is a core aspect of this thesis and the reconstruction performed as part of it. Whereas there exists several commercial software such as PC-Crash, WinSMASH etc, the toolbox that has been developed (detailed later in this report) is designed to also incorporate the glance locations of the driver. This glance information in the context of the crash itself can then be used to understand the behaviour and, hopefully, give an insight into the expectancy and the salient events that played a critical role in the events. With a limited number of cases in the NDS data that fit the requirements of the study, the toolbox developed was not highly automated and gives researchers the freedom to tweak and visualise each case individually. As Bärgman (2016) explains, care must be taken to not generalise these results extensively because of the small number of cases available.

2 Methodology

2.1 Data selection

The SHRP2 data obtained from Virginia Tech Transportation Institute (VTTI) contained 24251 cases of both crash and near-crash type events, of which 1011 events contained glance data. These events were then sorted to within the scope of this research, namely LTAP/OD and SCP. This was done by filtering *IncidentType1. Turn across path* was sorted as LTAP and *Straight crossing path* as SCP. LTAP events were individually reviewed by video analysis to eliminate all cases where the SV and TV were not approaching each other from opposite sides of the intersection as shown in Figure 2. Further, only events which contained the corresponding forward video feed was saved into the final MATLAB data file that was used for the rest of the study.

Table 1 lists the cases filtered out for straight crossing path. One case (marked with '*') could not be analysed within the toolbox due to a very low speed and longitudinal distance covered. The toolbox by default attempts to interpolate the data for a smoother playback of the simulation, further in detail in Section 2.3.2. The low speed and small distance travelled did not allow any suitable sampling time to be selected and hence had to be discarded.

Table 1: Straight crossing path cases

Total events available with glance data	1011
IncidentType1 – Straight crossing path	17
Event Type1 - Crash	17
EventType1- Near crash	0
Cases removed	
No suitable sampling time possible*	1
No forward camera video	1
Final count	**15**

Table 2 lists the filtering of LTAP/OD cases. Video analysis of the 23 turn across path cases was performed and seven were discarded as they were not turn across path in opposite direction or left turn. Of the remaining 15 cases, the SV performed the left turn manoeuvre in ten cases, while TV performs the left turn in the remaining five cases.

Table 2: Left turn across path / opposite direction

Total events available with glance data	1011
IncidentType1 – Turn across path	23
Event Type1 - Crash	23
EventType1- Near crash	0
Cases removed	
No forward camera video	1
Not LTAP/OD (video analysis)	6
Glance annotated as *No video available*	1
Final count	**15**

2.2 Estimation of vehicle position

Before the toolbox could be developed, the necessary equations of motion of the SV had to be derived to reconstruct its movement. This had to be done in such a way that it could be applicable to all cases using the available (onboard) vehicle kinematics data and didn't rely on GPS data. It must also be noted that GPS positions was not available in any of the cases analysed. The derived equations from each of the methods were then verified using basic two-dimensional plots and the automated driving system toolbox in MATLAB.

2.2.1 Vehicle dynamics

Vehicle dynamics is the field of engineering that deals with the motion of a vehicle and the forces acting on it. Vehicle dynamics is extensively used within the field of active safety and autonomous driving systems to calculate and even predict the trajectories of vehicles. The key difference between vehicle dynamics and kinematics is in vehicle dynamics, the forces acting on the vehicle are considering during the calculations.

2.2.2 Vehicle kinematics

In this approach, the forces acting on the vehicle are not considered. The vehicle was assumed to be moving on a flat road and does not experience any aerodynamic drag. The effect of rolling friction was disregarded. Two separate methods were chosen to reconstruct the trajectory of the vehicle, the first of which was by considering a one-track or bicycle model and the second by performing cumulative trapezoidal numerical integration on the first and second derivatives of position - velocity and acceleration. Some of the other assumptions made for the rest of the analysis are as follows:

i. Length of the vehicle:
   a. Car: 4.7m

   - b. Pickup truck: 5.1m
   - c. Sports utility vehicle: 5.1m
   - d. Minivan: 5.1m
- ii. Width of the vehicle: 1.8m
- iii. Sampling frequency: 15Hz

The video data from the forward camera was recorded at 15 frames per second, whereas all data from the CAN was at 10Hz. In order to match the frequency of the two, all data was interpolated to 15Hz, and will be detailed further in Section 2.3.

2.2.2.1 Method 1

Consider the one-track kinematic model for lateral motion shown in Figure 3.

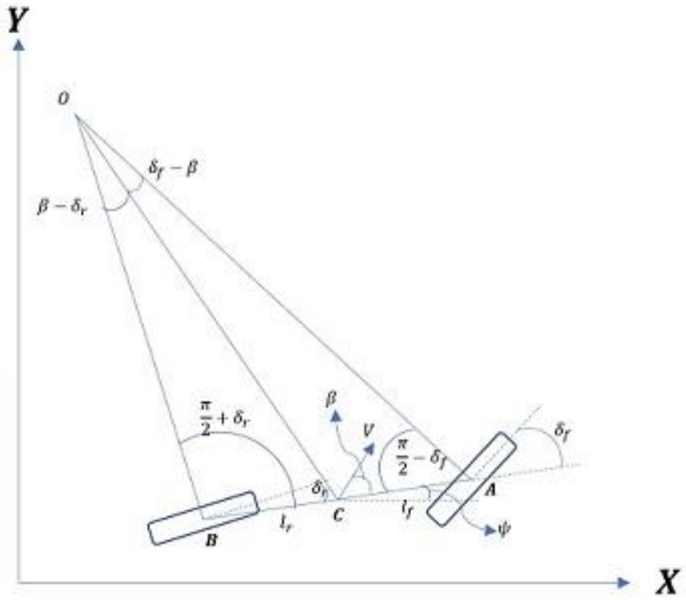


Figure 3: One-track model adapted from (Rajamani, 2011)

Where,

δ_f and δ_r are the front and rear wheel steering angles respectively.
β is the slip angle
C is the centre of gravity (c.g.) of the vehicle
l_f and l_r are the distance from the front and rear axles to c.g. respectively.
Where, length of vehicle $L = l_f + l_r$
ψ is the orientation or heading of the vehicle
V is the velocity of the vehicle at c.g.
A and B are the front and rear axles of the vehicle
O is the instantaneous rolling centre

It is assumed that the steering angles of the left and right wheels are equal and can be represented by a single wheel at A. The rear wheel steering angle assumed to be zero such that the vehicle is steered through only the front wheel. OC represents the turning radius R. This model was then simplified using basic trigonometry to obtain the following equations:

$$\beta = \tan^{-1}\left(\frac{l_f \cdot \tan\delta_r + l_r \cdot \tan\delta_f}{l_f + l_r}\right) \tag{1}$$

Or

$$\beta = \tan^{-1}\left(\frac{l_r \cdot \tan\delta_f}{l_f + l_r}\right) \tag{2}$$

Further, the overall equations of motion become,

$$\dot{X} = V \cdot \cos(\psi + \beta) \tag{3}$$

$$\dot{Y} = V \cdot \sin(\psi + \beta) \tag{4}$$

$$\dot{\psi} = \frac{V \cdot \cos\beta}{l_f + l_r} \cdot (\tan\delta_f) \tag{5}$$

Here, δ_f, V and ψ are data available from the SHRP2 database which can be used as the inputs for the calculation of β, X and Y.

2.2.2.2 Method 2

An alternative method to calculate the positions was devised as δ_f was not available for many cases. Further, the varying sampling frequency between each case resulted in the loss of important information when the data was interpolated to fill the missing steering wheel angles. This is because the driver can make multiple steering adjustments in different directions within one second (if the data is at 1Hz frequency). This is more likely during evasive manoeuvres. The position of the vehicle in the intersection was calculated by integrating the derivatives of position, i.e., speed and acceleration in the X and Y directions.

Consider the intersection shown in Figure 4 below.

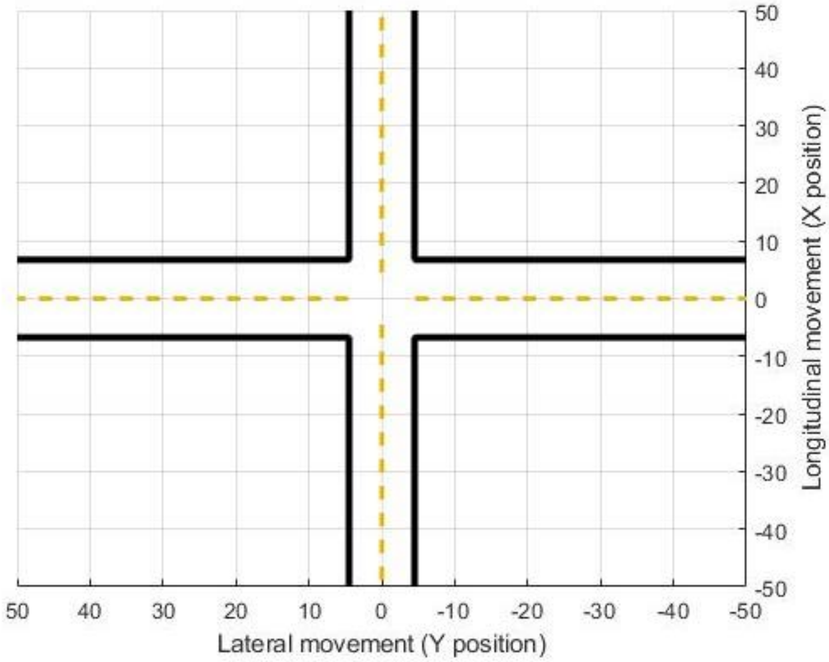


Figure 4: Reference intersection used for this study

This intersection is used as the reference for the calculations of X and Y positions of the vehicle with respect to the centre, such that the centre of the intersection always lies at the origin, (0,0).

The database contained speed from both GPS and controller area network (CAN). The GPS speed, however, was sampled at a low frequency (1Hz), compared to the CAN data which was 10Hz for most cases. The MATLAB function *cumtrapz* or cumulative trapezoidal numerical integration was used to perform the integration of the variables. The highest derivative of position available (acceleration) was integrated over uniformly spaced intervals of time to obtain the speed. This calculated speed had to be offset by the constant of integration to obtain values of speed comparable to the speed from the CAN. Acceleration is the change in speed over time, however, if the speed is fairly constant, the change in speed is negligible, and thus a very low acceleration. By adding the constant of integration, the computed speed can be offset to a value comparable to the actual speed. This constant of integration is calculated by first selecting a point on the speed profile where a relatively stable state is observed, and an exhaustive set of values are iterated to obtain an offset that has the lowest cost difference between the actual speed and the integrated speed. The cost was calculated by the absolute sum of all numerical values of each iteration of offset speed and its difference computed against the cost of the network speed. The value with the lowest cost difference was chosen. Figure 5 below shows the speed profiles from CAN data and integrated acceleration, the various iterations performed and the corresponding best fit.

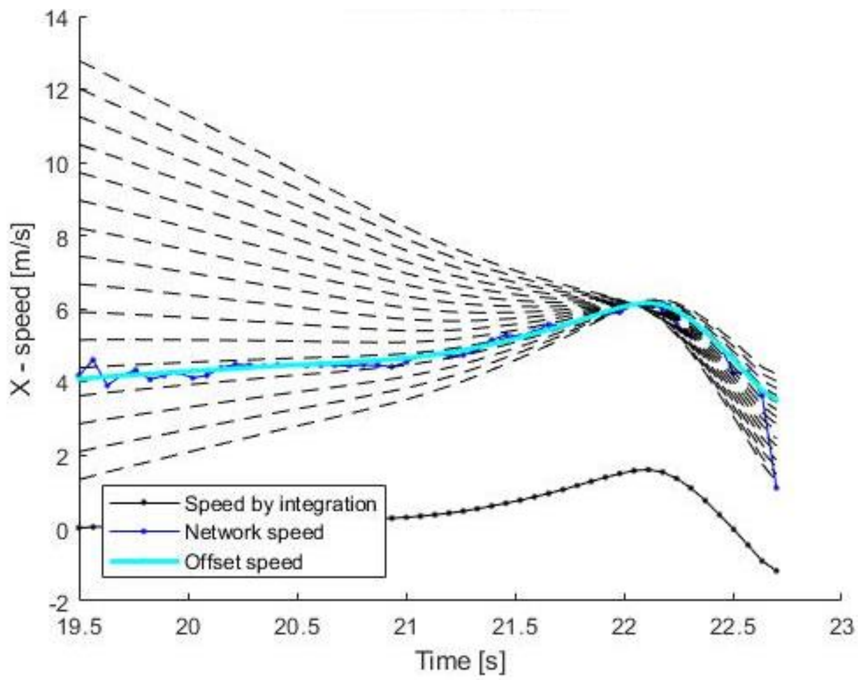


Figure 5: Offset adjusting the integrated variables

The integrated speed and CAN speed are integrated once again to obtain the position. It is assumed that the offset between the two is negligible as they had been matched in the previous iteration.

The difference between the method used in X and Y positions is as follows:

i. To calculate the longitudinal position, the accelerometer reading and the speed from CAN is used. The acceleration is first integrated to speed, offset adjusted to match CAN speed, following which the two are integrated to obtain the position.

$$\dot{x}_{acc} = \int_{t_0}^{t_1} \ddot{x} \cdot dt + C \tag{6}$$

Where $\dot{x}_{acc}$ is speed obtained by integrating acceleration $\ddot{x}$.
C is the constant of integration
t_0 is the start of the event
t_1 is the end of the event

Or, in MATLAB,

$$\dot{x}_{acc} = cumtrapz(Time, \ddot{x} + C) \tag{7}$$

ii. To calculate the lateral position about the centre of the road, the lateral acceleration from the accelerometer and the yaw rate from gyroscope are used.

$$y_{acc} = \int_{t_0}^{t_1} \int_{t_0}^{t_1} \ddot{y} \cdot dt \cdot dt + C \tag{8}$$

Or,

$$y_{acc} = cumtrapz\big(time, cumtrapz(time, \ddot{y} + C)\big) \tag{9}$$

And similarly,

$$yaw = cumtrapz(time, yaw\ rate) \tag{10}$$

The yaw is then used to calculate the lateral position as follows.
Consider Figure 6 shown below,

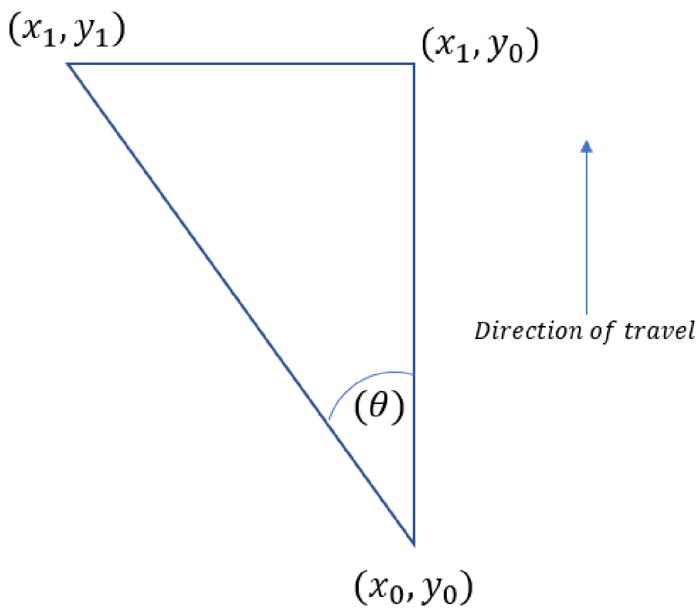


Figure 6: Calculating the change in lateral position

Where, x_0, y_0 are the coordinates at time t_0
x_1, y_1 are the coordinates at the next time step.
θ is the yaw angle.

Using trigonometric relations, y_1 can be obtained as Equation (11) below,

$$y_1 = \left((x_1 - x_0) \cdot \tan\theta\right) + y_0 \tag{11}$$

The best visual representation of the lateral movement between the two methods is then chosen as the lateral position.

Figure 7 below shows the trajectory of the vehicle using the calculated coordinates.

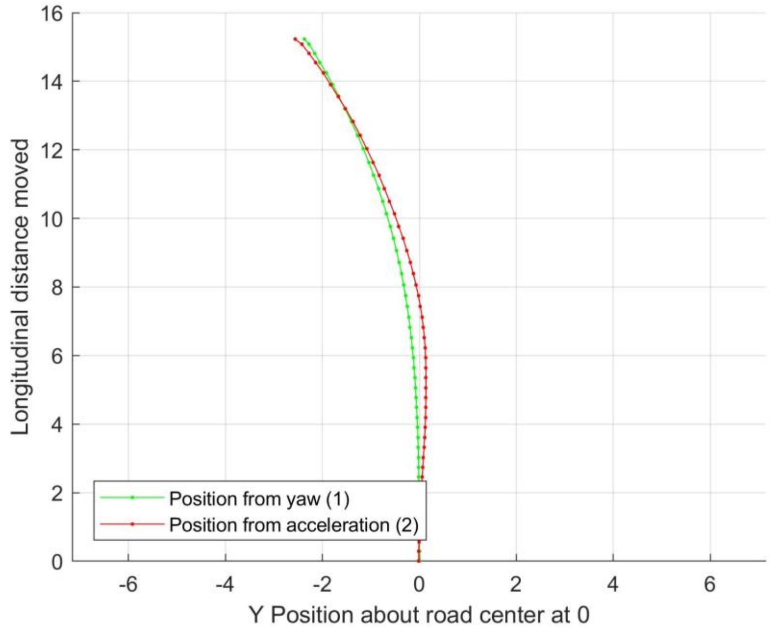


Figure 7: Trajectory of vehicle

2.3 Toolbox development

With the coordinates of the vehicle in the intersection calculated, the toolbox that will be used to recreate the intersection event could be developed. An important function of the toolbox is to enable the visualisation of the driver's glance locations at each interval of time, in relation to the vehicle's position in the intersection. The toolbox was designed in such a way as to perform the following distinct tasks:

i. Selection of data range and processing of selected data
ii. 3D reconstruction and comparison with the video feed
iii. Incorporating glance data
iv. Introduction of the target vehicle

The software implementation flow chart for the toolbox operation is shown below in Figure 8 and Figure 9. Figure 8 shows the operation of the main algorithm used for event selection, data selection, resampling and computing vehicle coordinates with respect to the intersection. Figure 9 details the working of the 3D reconstruction, bird's eye view of the event and finally, implementation of glance data into the event.

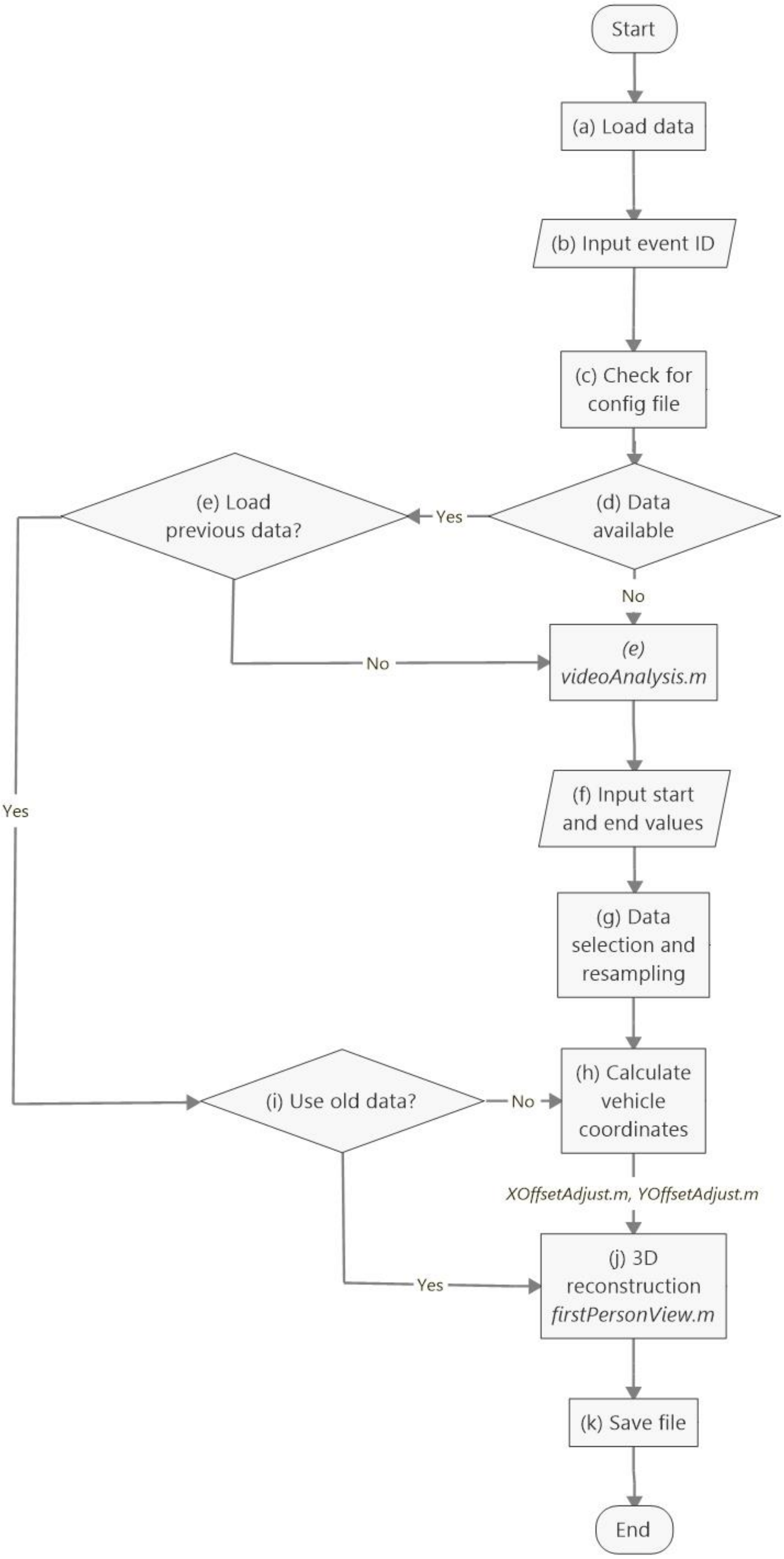


Figure 8: Algorithm depicting the operation of the toolbox

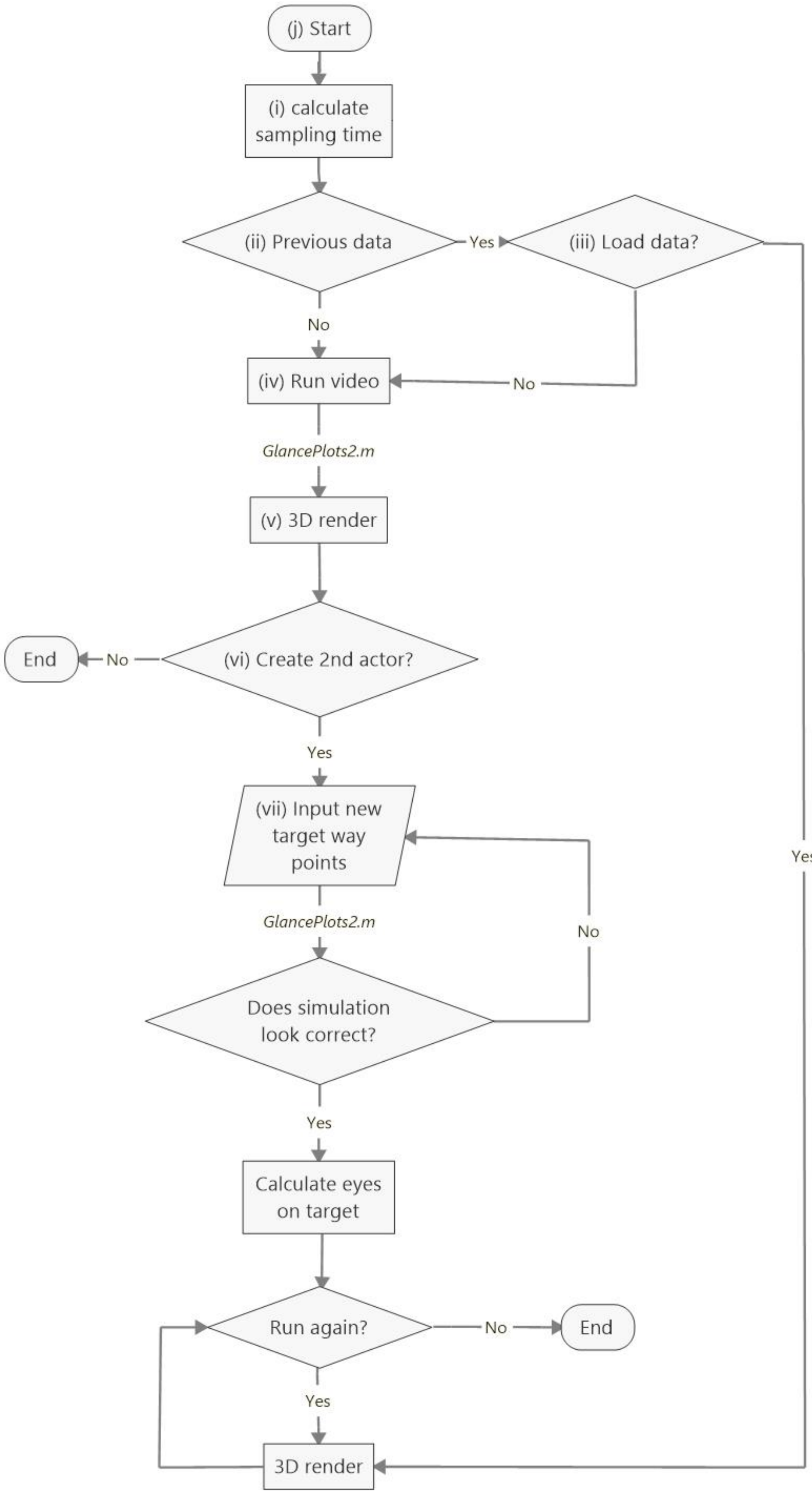


Figure 9: Algorithm of reconstruction (expanded block (j) of the first flow chart)

2.3.1 Selection of data range and processing

The algorithm first checks for the availability of results from a previous analysis. The design is such that a user can skip one or multiple steps of the analysis, when previous results are available, to visualise only that part of results needed. The available data by default contains 30 seconds of vehicle data, video feed and glance data. However, the part of the data relevant to the crash is only a small portion of this data. In order to select only the relevant data range, the user (of the toolbox) first chooses the start and end time of the event by reviewing the video. The video player allows a frame by frame manipulation to make the data selection easier. The analysis range was selected between a) the point at which the TV is first clearly visible in the video feed or the point at which the SV appears to enter the intersection, whichever of the two occurs first. The latter is determined by using one or more visual cues available from the video feed such as stop lines, corners, breaks in the road barriers, crossings etc. to b) one frame before the crash, as visible from the video feed. The algorithm is set to take one timestep or $0.1s$ earlier than the user input as the crash point. It then selects the vectors within this range which will be used through the rest of the analysis. It also resamples all the vectors to a frequency of 15Hz to match the video frame rate. If speed from CAN is not available, the GPS speed data is chosen instead.

Once the data range is chosen, the position of the vehicle at each time step is then calculated as explained in Section 2.2.2.2. Multiple trials offset adjustment may be performed until a satisfactory fit is obtained. The lateral movement could be compared by visually approximating the movement of the SV in the video. The longitudinal movement was compared to cues such as road corners and road centres. At the end of this step, all necessary data to begin the 3D reconstruction is calculated.

2.3.2 Reconstruction of event

The core of the 3D rendering of the toolbox works on the autonomous driving toolbox available in MATLAB 2017a and later versions. The ADS toolbox allows researchers to create a variety of driving scenarios, including various road profiles, traffic situations, vehicle properties and also incorporate other road users such as pedestrians, cyclists and other vehicles. The properties listed in Table 3 were used to generate the driving scenario in the toolbox.

Table 3: Scenario generation properties

Property	**Value**
Length of road	$100m$
Width of road	$9m$
Intersection centre	[0,0,0]
Vehicle length	As selected in Section 2.2.2
Vehicle width	$1.8m$
Camera height from road	$0.85m$

The 3D scenario generated using ADS is played at the same framerate as the video data available. A first-person view, a bird's eye view and the video from the front camera is played one frame at a time, side by side to allow easy comparison between the real forward video and the virtual camera viewpoint. The ADS toolbox interpolates the data further to produce a smoother playback. However, this "feature" causes a discrepancy in sizes of all the vectors. To overcome this, an exhaustive set of sampling times in steps of 0.001s is tried from 0.1s to 0s. Suitable sampling time is chosen when the size of kinematic vectors generated by the ADS toolbox is equal to the input vector size, i.e. the ADS toolbox does not need to resample the input data at this chosen sampling time. This is not a desired feature from the toolbox for this type of use.

2.3.3 Glance Location

Once the sampling time has been selected, two subfunctions are called in to generate the field of view of driver glances on the bird's eye plot. The cone for the field of view is generated by an approximation based on the direction and object in focus.

2.3.3.1 Looking ahead

There are 5 major categories in the forward 180^0 where it can be assumed the driver is looking at an object in the distance, namely, *forward, left windshield, left window, right windshield* and *right window.* A visual range of 100m is assumed in all these cases (this distance was never a limitation for this application, as the vehicles at all times were much closer than 100m in the simulations). Table 4 lists the angles for the fields of view used while Figure 10 gives a pictorial representation of the same. All angles are in perspective of the driver's location in the car.

Table 4: Angular range of different glance locations with respect to longitudinal axis

Direction of glance	Angle 1	Angle 2
Forward	$15°$	$-15°$
Left windshield	$-15°$	$-45°$
Left window	$-45°$	$-90°$
Right windshield	$15°$	$45°$
Right window	$45°$	$90°$

2.3.3.2 All other areas of interest

All other glance locations in the database are directed either to an object within the vehicle or correspond to cases where the eye data is not available. Glances within the car have been allocated a close range (approximated to 1m) and do not take into consideration peripheral vision. This was done so that only glances outside the car could return true if the driver's field of vision crosses the target vehicle.

The field of view cone rotates with the same yaw as the vehicle when as it goes through the intersection, in other words, the principal axis of the driver is always parallel to the principal or longitudinal axis of the car, as seen in Figure 11.

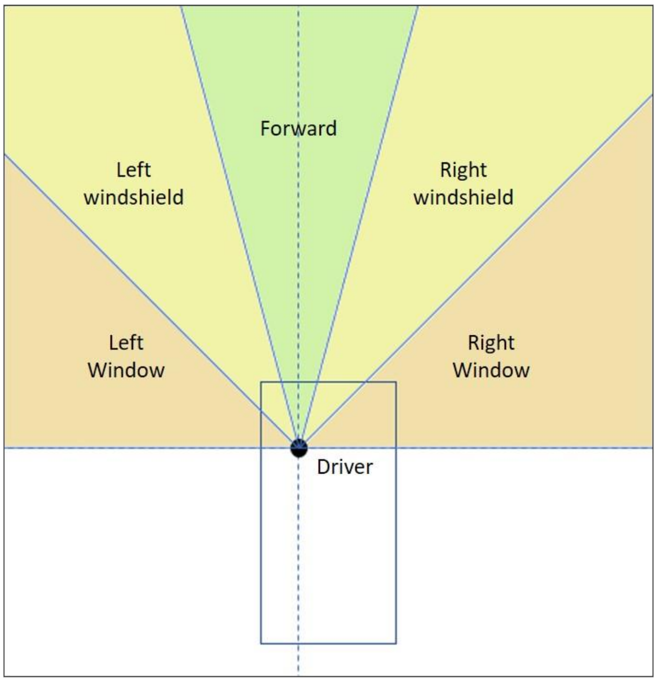


Figure 10: Angular ranges of glance locations. (Vehicle dimensions are not to scale)

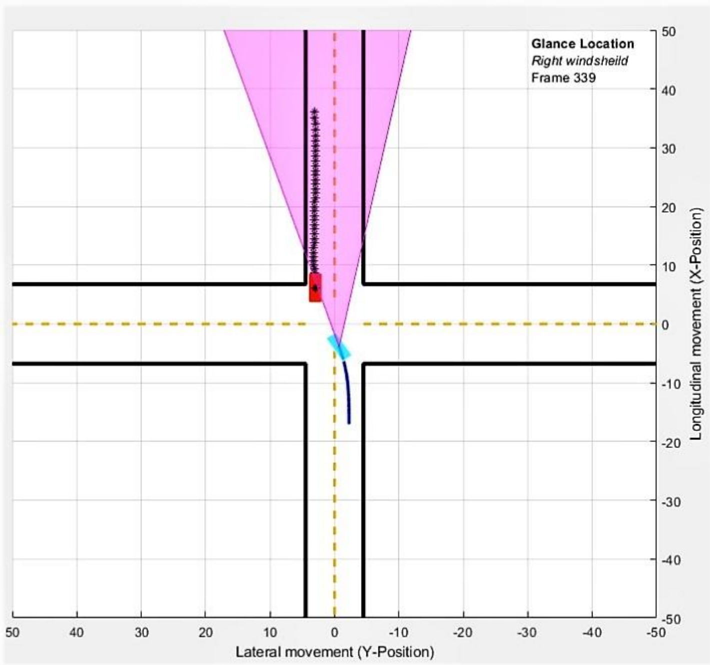


Figure 11: Field of view changing with the orientation of the vehicle turning in the intersection

2.3.4 Target vehicle input

Once the vehicle trajectory and glance location have been successfully plotted, the approximate coordinates of the TV position in the intersection are manually input to the bird's eye plot using MATLAB command *ginput*. This is done by a trial and error method by visually comparing the video feed against the plot. The video is played frame by frame, and an input of the approximate location of the target vehicle is requested on the first and last frame and every fifth frame in between. The data is then interpolated to a 15Hz frequency to obtain the approximated coordinates of the TV with respect to the SV. The velocity of the TV is calculated as the change in longitudinal position per unit time, shown in Equation (12) below,

$$velocity_{TV} = \frac{\Delta x_{TV}}{\Delta time} \tag{12}$$

This process can be repeated until a "close-enough" visual match between the real forward view and the virtual (3D) forward view is obtained. This is a key step in the reconstruction of the kinematics.

2.3.4.1 Eyes on target

Once the target vehicle's trajectory has been accepted as satisfactory by the user, the simulation is run in the background to determine all the events where the driver looks at the target vehicle. The target vehicle is assumed to be a circle with a radius 3m, thus allowing for a slightly larger margin for detection (allowing for compensating some inaccuracies in glance coding and vehicle positions). The coordinates of the circle are calculated using Equation (13) and (14)

$$x_{circle} = radius \cdot \cos\theta + x_{TV} \tag{13}$$

$$y_{circle} = radius \cdot \sin\theta + y_{TV} \tag{14}$$

Where, $\theta = 0\ to\ 2\pi$

If any point on the circumference of this circle lies within the field of vision triangle, it returns true. The *inpolygon* function in MATLAB is used for this, and a resultant vector of logical 1 and 0 is obtained, with 1 as a point lies within the field of view. These results are then automatically saved into a configuration file. This configuration file, with extension **.mat*, allows users to skip parts of the overall analysis in the future.

2.4 Further processing and analysis of data

In a further manipulation of the results, all values of eyes on target corresponding to glance annotations of *No eyes visible, no eyes visible – eyes off road, no video* and *other* are replaced by $NaNs$ to prevent false representation of results.

The point of no return used in this study is defined as *that point of time beyond which a crash is inevitable if the driver does not actively decelerate or brake at a constant, maximum rate equal to* $7\ m/s^2$*, assuming no other collision avoidance manoeuvre is performed during this period of braking.* The point of no return is calculated by finding the point of intersection between the speed profile of the entire event and a line of slope $7(m/s^2)$ passing through the crash point, seen in Figure 12 below. The line was constructed by using the equation of a line of slope m passing through the y axis at point c, given by Equation (15),

$$y = m \cdot x + c \tag{15}$$

Where x and y are the x and y axis coordinates, corresponding to event time and velocity of the SV respectively.

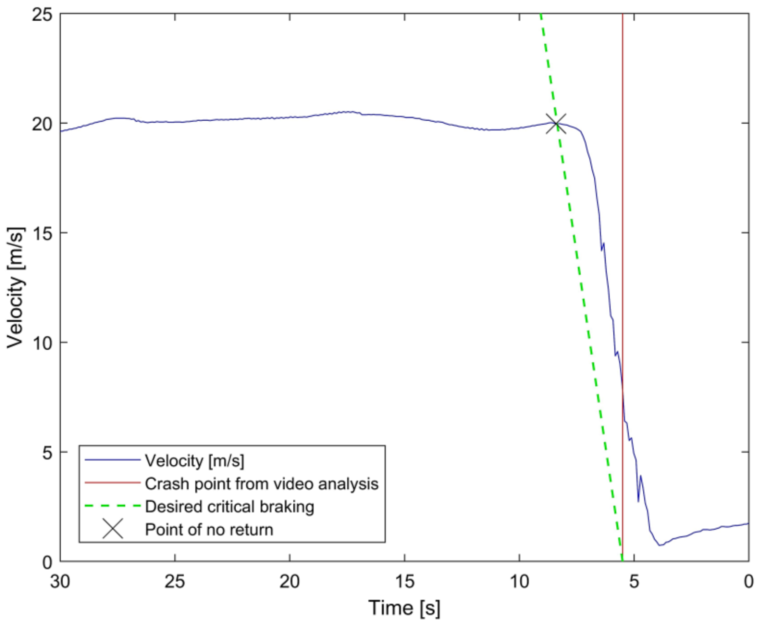


Figure 12: Point of no return calculated over the entire event

Further, in the context of this study, the Intersection Gaze Release Time (IGRT) is calculated with respect to the point of no return and the point of crash. IGRT is defined as *the time from when the driver last looks at the target vehicle to the point of no return, or crash point, respectively*. This means that if the driver was looking at the TV at the time of interest (point of no return or crash point, respectively), the IGRT would be 0, whereas if the driver never sees the TV in the seconds leading up to the time of interest, the IGRT is set to infinite.

The glances location and their distribution were then computed by weighting in the number of cases at each timestep. The distribution of the glance location was computed by first generating a vector containing all events of a given incident type. This vector was adjusted so that all events at end at the same point in time or the crash point. For each timestep, the sum of each glance location with all events was calculated. Their ratio of occurrence for each timestep was calculated by dividing this sum with the number of events containing glance data for that timestep, shown in Equations (16), (17) and (18) expressed as MATLAB script.

```
glanceLocationCount = (sum(glanceLocation==i))                 (16)

glanceLocationNotNaN = sum(~isnan(glanceLocation))             (17)

glanceLocationCount =
(glanceLocationCount./glanceLocationNotNaN)*100                (18)
```

Where,

`i` corresponds to the different glance locations annotated from 1 to 19,
`glanceLocation` is a vector of all events with glance locations adjusted to end at the crash point.
`sum` returns a row vector of the sum of each column of `glanceLocation`.

This is then plotted as a stacked bar plot, using a *jet* colour scheme for consistency. A similar approach was used to plot the distribution of the eyes on target.

Another set of stacked bar plots for the distribution of glance locations and eyes on target were generated but now offset to the point of no return. The point of no return was set to the middle column of an empty vector whose length was approximately three times the length of the longest event. This column was used as a reference point to adjust the remaining data points. The vector was trimmed to fit at both ends; vectors for glance data and eyes on target were generated with this vector as the reference. They were then plotted similarly to the previous.

3 Results

The toolbox provided some automation of crash reconstruction for the different events, but at the same time flexibility to treat each case uniquely. There were 15 events of SCP type and 15 events of LTAP/OD, in which the SV performed the left turn in 10 events and drives straight through in 5 events. The following sections will discuss the results of the toolbox's effectiveness in the crash reconstruction and insights into the driver behaviour.

3.1 Vehicle kinematics

Figure 13 and Figure 14 below shows an example of vehicle kinematics in the longitudinal direction. The kinematics computed in the longitudinal direction from speed and acceleration were, in general, found to have a good fit in all cases.

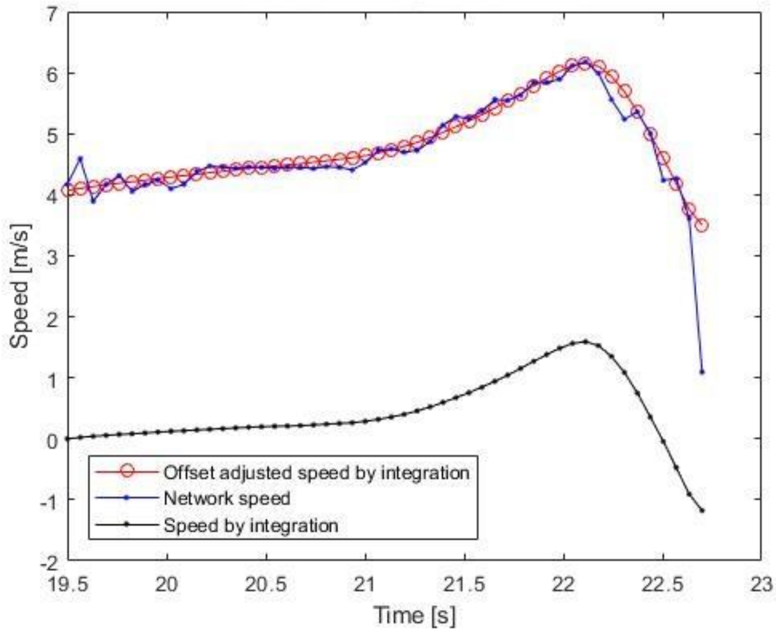


Figure 13: Comparison between longitudinal speed and integrated longitudinal speed.

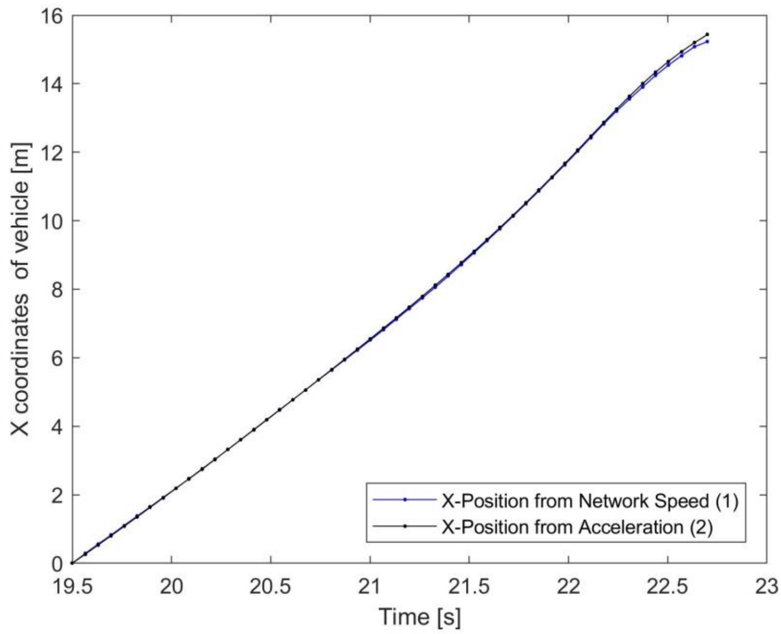


Figure 14: Comparison between the longitudinal positions

Figure 15 shows an example of the lateral movement of the vehicle as it moves in the longitudinal direction, computed with the toolbox.

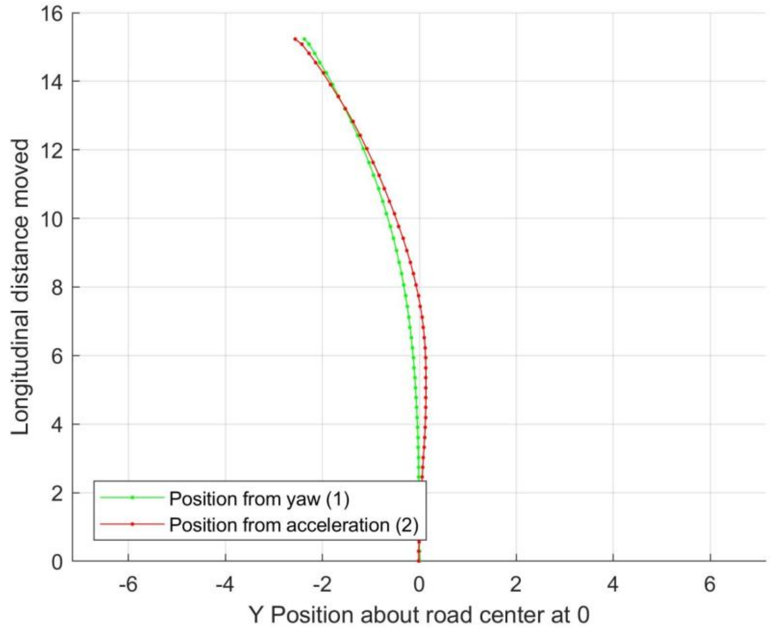


Figure 15: Comparison between lateral positions calculated form yaw rate and lateral acceleration

These computed values were then used to reconstruct the event and verified from a first-person (driver) "3D" view and a bird's eye view perspective, shown in Figure 16 and Figure 17. The movement of the SV was found to be visually similar to what was observed in the video, traced in Figure 18. One case contained faulty yaw data, which resulted in the vehicle rotating rapidly on a straight road. The analysis length for this event had to be chosen appropriately to capture as much data as possible, without causing a false data capture. SV movements in the 29 other events were found to visually match the video data to a high degree, especially considering its exact location with respect to the intersection was approximated/estimated through manual annotation on the birds-eye view.

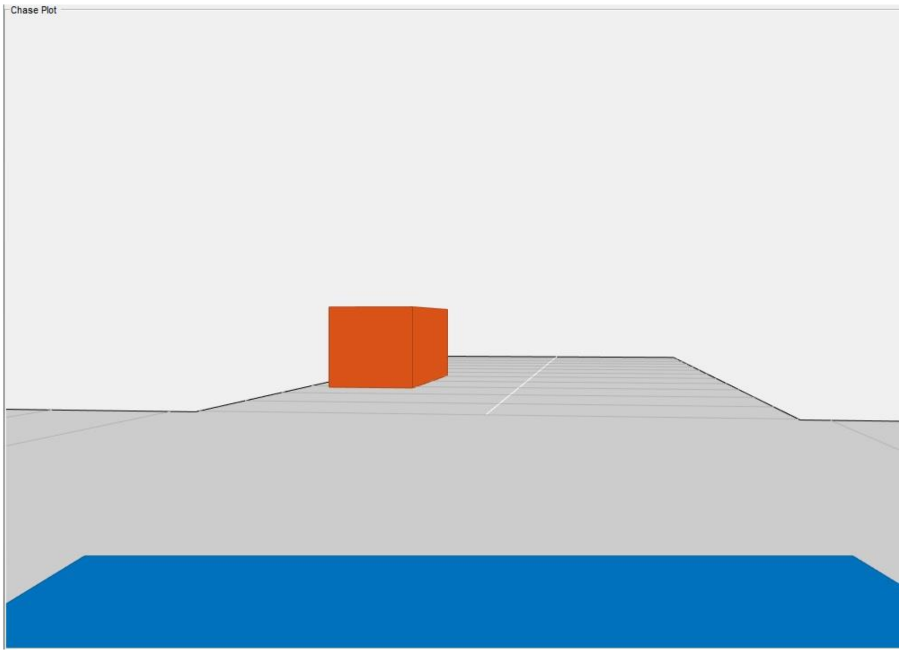


Figure 16: First person view using the toolbox

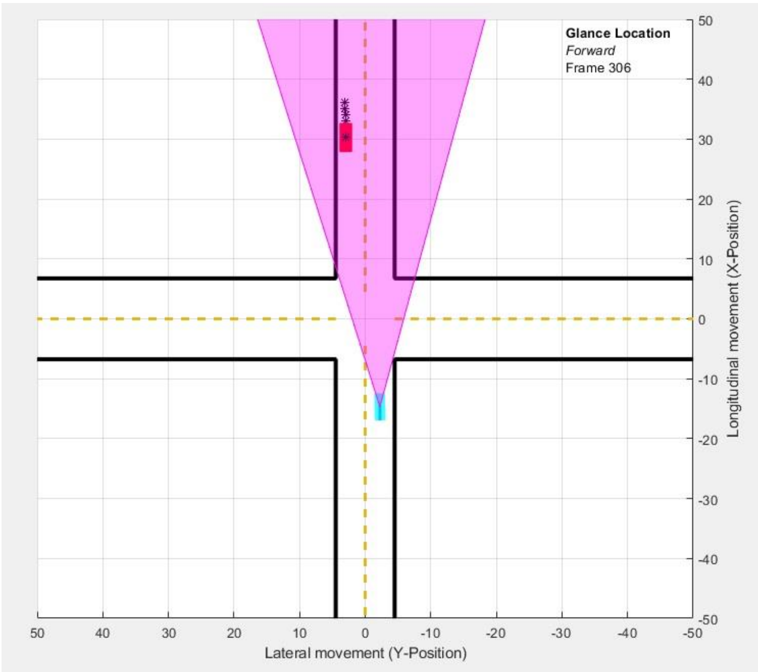


Figure 17: Bird's eye view of the same event

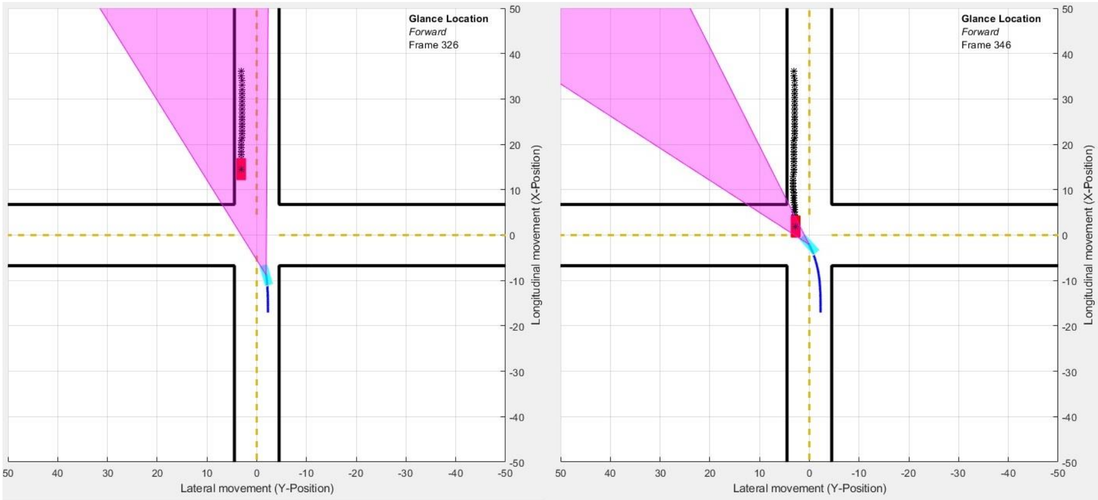


Figure 18: Bird's eye plot of the rest of the event

The TV inputs were tried over several iterations until the closest possible representation was obtained. Since the position of the TV was entirely approximated by reference to the video feed, a general behavioural match was aimed for, rather than a perfect trajectory path. A major drawback in the toolbox was identified during the TV trajectory estimation. The nature of the ADS toolbox is such that all vector lengths must be the same. If any of the vector lengths related to the kinematics of the TV or SV is smaller, the toolbox by default chooses the smallest vector size and performs the simulation. Attempts to match vector lengths by the *NaN*s also proved futile as the toolbox disregarded *NaN*s. This was especially true for TV position inputs as the toolbox is designed to take TV position inputs from only that frame when the user first sees the TV. Hence, if the user first sees the TV 1 second before the crash in a 3-second-long simulation, the first user input of the TV occurs at 1 second before crash. MATLAB then simulates only 1 second of the event, ignoring the 2 seconds preceding it. Using a fixed point or keeping the vehicle stationary at one point also resulted in the same outcome.

In order to overcome this, a very small movement was provided to the TV by calculating a position of 0.001m backward in both directions for each "missing" timestep. When the simulation was performed, however, this "small" movement appeared to be very significant and resulted in odd TV movements. The toolbox does correct this movement to the expected trajectory from the start of the user input. Further, if the points between two user inputs of the approximate TV position was too close to each other, it often led to repeated coordinates points during the interpolation. The toolbox, however, requires each coordinate or waypoint to be unique and consequently crashes. Multiple trials had to be performed in order to ensure that this odd TV movement did not impede the toolbox's ability to recreate the event satisfactorily. The bird's eye plot, however, was not affected by these problems and hence captured *eyes on target* and TV trajectory successfully.

Since the driver's eye gaze direction was defined in this research as absolute (with respect to the intersection/world), as listed in Table 4, it resulted in a field of view cone completely missing the TV even though it could be understood that the driver changed their gaze direction upon identifying the threat in that direction. Figure 19 below shows an example situation where this problem was identified.

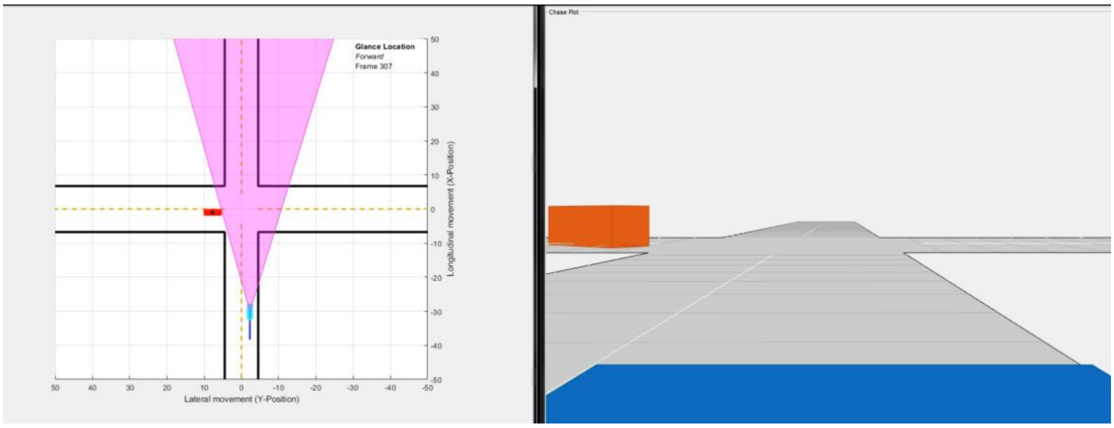


Figure 19: Example of the field of view cone failing to capture TV

The margin by which the field of view cone missed the TV was very small and was identified in only three cases. This was manually corrected by carefully reviewing the video with glance annotations and the bird's eye plots.

The results obtained from the simulations performed with the toolbox, despite having a very high level of approximation and assumption of the driver's intent with respect to vehicle movement, was able to reconstruct the event and glance locations with a satisfactory level of detail as shown in Figure 18.

3.2 Glance Analysis

3.2.1 Distribution of glance locations

Figure 20 and Figure 21 show the distribution of the glance locations as a function of time for SCP and LTAP/OD events. The calculated percentage of events is a function of the number of events in the at each time step. Time at $t = 0s$ corresponds to the crash point. Figure 22 lists the legend for the colour schema used to represent the different glance locations in all figures with glance annotations.

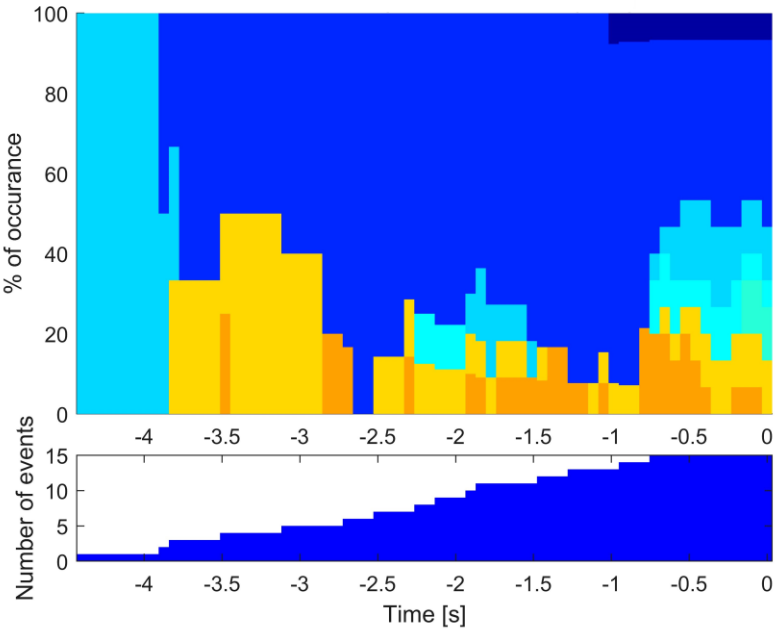


Figure 20: Distribution of glance locations for SCP events. Time at 0s corresponds to the crash point.

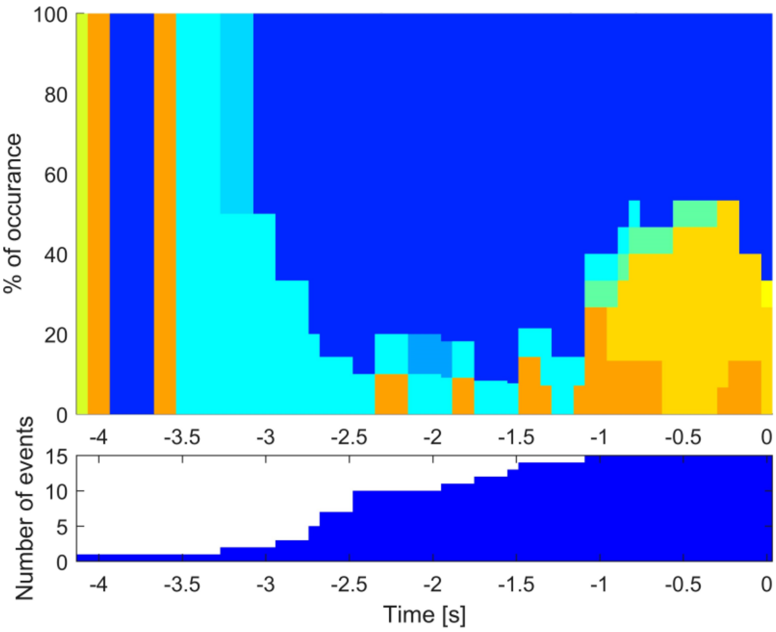


Figure 21: Glance locations for all LTAP/OD events. Time at 0s corresponds to the crash point.

Figure 22: Legend of different glance locations

The gaze direction was found to be predominantly forward in both the SCP and LTAP/OD events. SCP events have an equal distribution of events where the driver switch towards either the left or the right side, indicating that they possibly noticed the threat arriving from the respective direction, although very late. In LTAP/OD, however, the drivers switch towards the right side which is the direction of the oncoming traffic as they perform the left turn manoeuvre. This is further evident when the LTAP/OD events are analysed from the two perspectives, where the majority of the cases involving the left turn manoeuvre by the SV has the driver looking towards the right side in the final second leading to the crash.

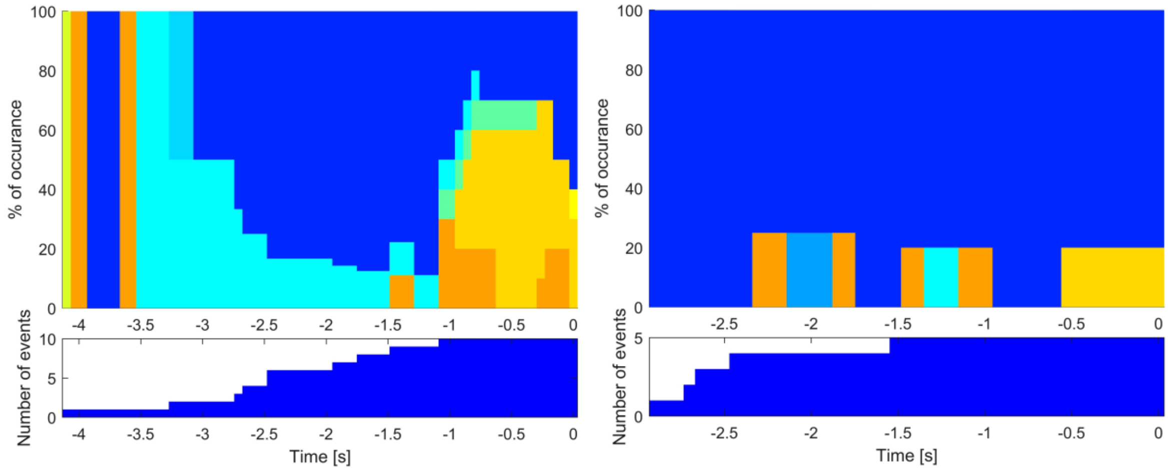


Figure 23: Glance locations by different LTAP scenarios. Left: SV performs left turn manoeuvre; Right: TV performs left turn manoeuvre. Time at 0s corresponds to the crash point

When the distribution of the *eyes on target* is plotted for the different events, it appears that a larger proportion of drivers had “seen” (gaze in the same general direction as) the threat in LTAP/OD events sooner than the drivers involved in an SCP incident. Figure 24 and Figure 25 shows the distribution of eyes on target for the different event types. This is possibly because the target vehicle is within the driver’s field of view much sooner than in SCP. The SV is the striking vehicle in only 5 events, with the SV being struck from either the left or the right side in 5 events respectively, reflected in Figure 20. There was only one LTAP/OD event where the driver never saw the TV at all, where the driver was using a cell phone while driving through the intersection. This event had the SV passing through the intersection at about 90kph at night, but through a green light – the turning TV was thus at fault. It is highly likely that the driver of the SV expected the TV, if any, to yield to the red light, while he/she passed through the intersection.

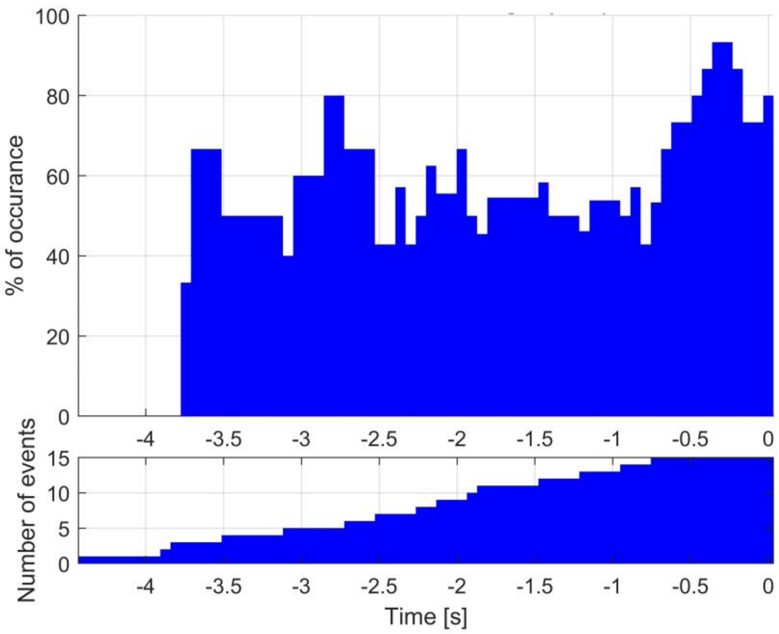


Figure 24: Eyes on Target in SCP events. Time at 0s corresponds to the crash point.

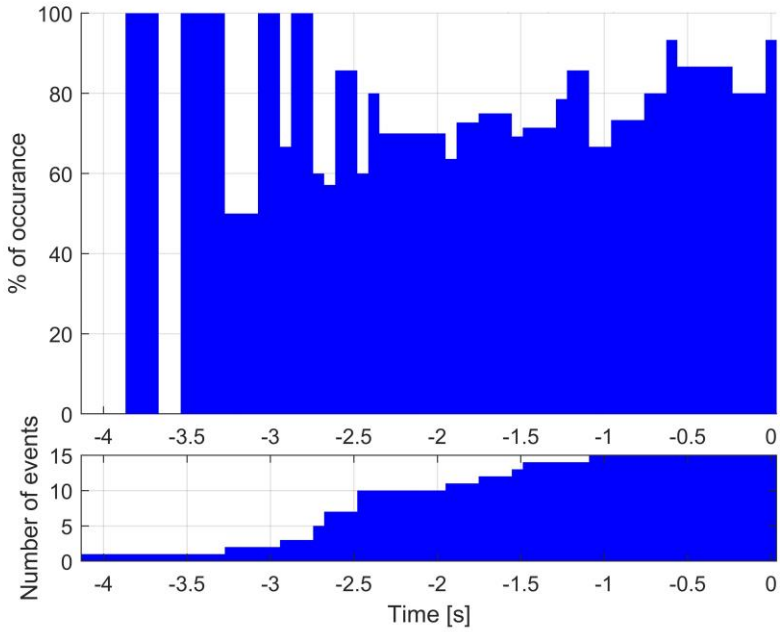


Figure 25: Eyes on target in LTAP/OD events. Time at 0s corresponds to the crash point.

When the glance locations were offset to the point of no return a large proportion of the drivers were looking forward in SCP events as compared to the LTAP/OD. Very few drivers had the TV within their field of view at the point of no return, as a result, in these cases, the crash would have been unavoidable, this can be seen in Figure 26 and Figure 27. More in detail in Section 3.2.2.1. Time at $t = 0s$ in the figures corresponds to the time of no return.

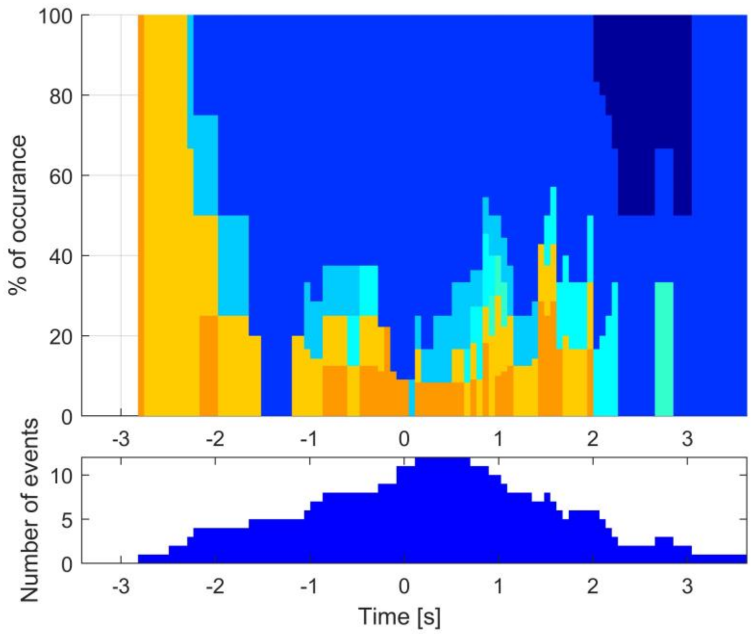


Figure 26: Glance distribution adjusted at the point of no return for SCP. Time at 0s corresponds to point of no return.

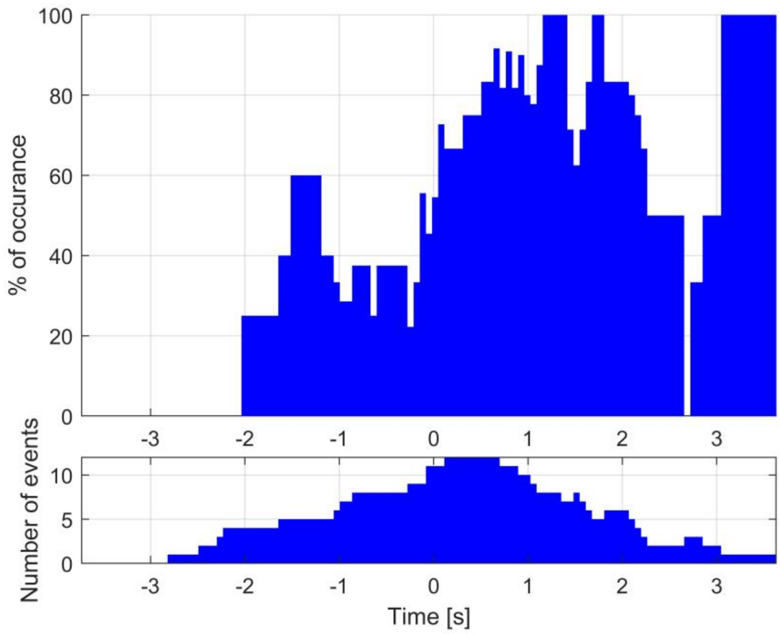


Figure 27: Eyes on target adjusted at point of no return for SCP. Time at 0s corresponds to point of no return.

In LTAP/OD, however, it appeared that close to 90% of the drivers had spotted the threat by the point of no return, with gaze directed primarily in the forward direction, followed by the forward right, presumably while engaged in the act of steering left. Figure 28 and Figure 29 depicts this behaviour.

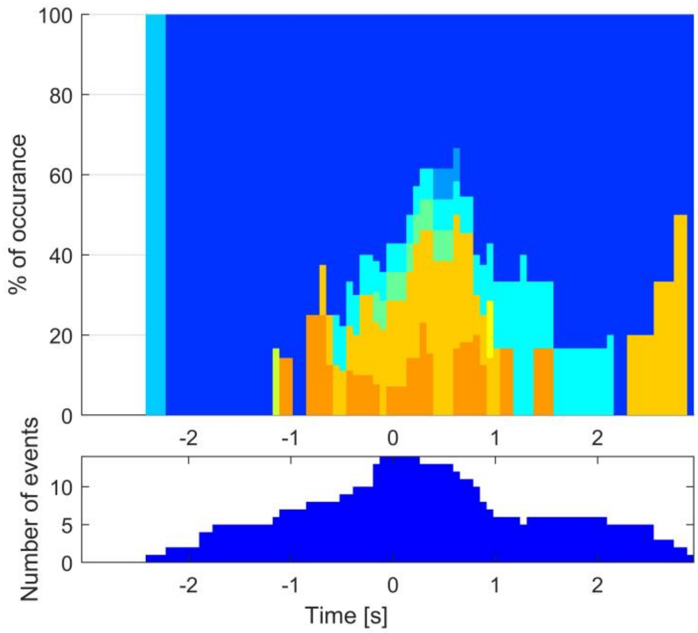


Figure 28: Glance distribution adjusted at the point of no return for LTAP/OD. Time at 0s corresponds to point of no return.

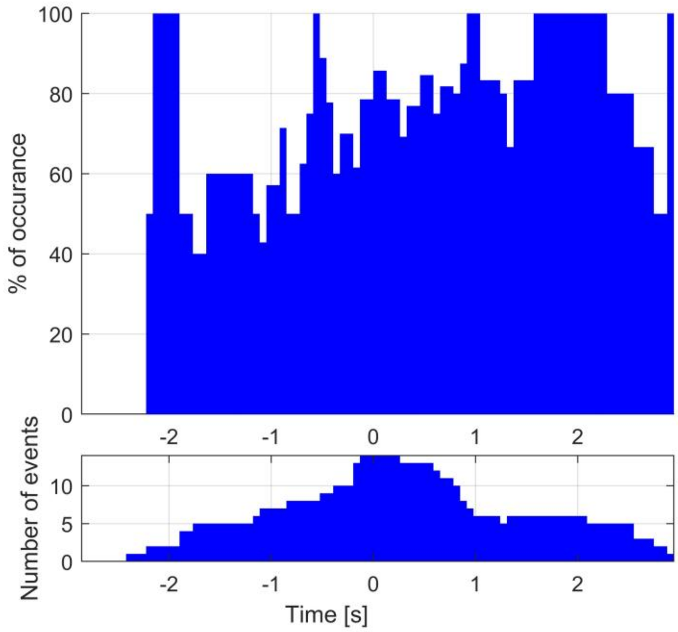


Figure 29: Eyes on target adjusted at the point of no return for LTAP/OD. Time at 0s corresponds to point of no return.

3.2.2 Intersection gaze release time

3.2.2.1 SCP events

The drivers did not see the threat in 67% of the SCP cases at the point of no return, as a result of which collision would have been unavoidable in these cases. The drivers in all those cases were observed to be either accelerating or moving at a constant speed after the point of no return. These cases are beyond the scope of this research as they cannot be attributed to behaviour error. In 20% of the SCP events, the driver had possibly seen the threat and was looking at the TV at the point of no return. However, 2 out the 3 such events involve the SV running a red light, in both of which cases, the SV was the striking car. There was only one case observed where the driver was seen shifting gaze between different areas of interest, TV on the left to pedestrians on the right while approaching the intersection. Figure 30 below shows the velocity profile and the driver's gaze on target. This is perhaps an example of where the driver had *expected* a certain behaviour from the different salient events s/he assessed in the intersection, however, did not successfully adapt to it. The light blue line represents the *eyes on target* annotated as 1 (0 eyes off target), the dashed green line shows the desired rate of braking to have avoided the crash.

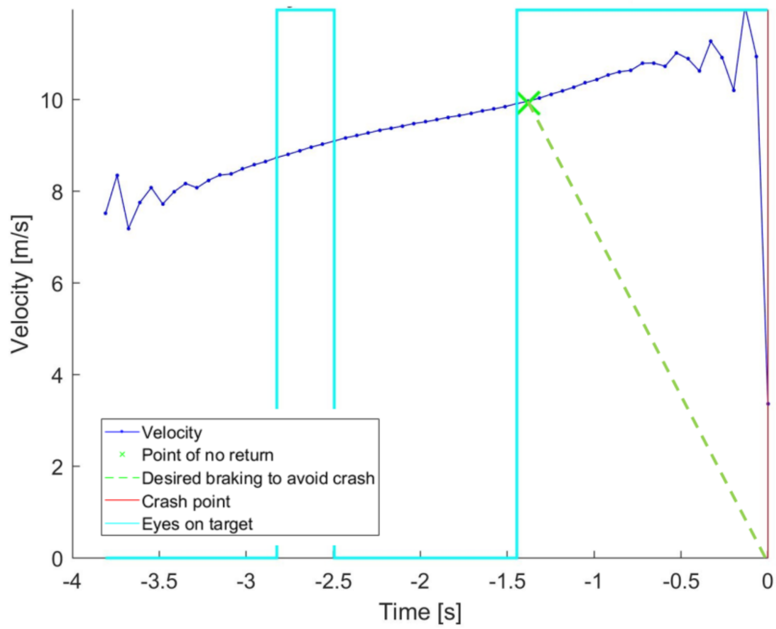


Figure 30: Example of an event with eyes on target as a function of velocity (SCP event)

Further video analysis shows that in 40% of the SCP events, the SV or the TV's field of view was obstructed by a static occlusion, such a parked car (five events) or road dividers (one event) until too late. This could be a possible cause for the large number of events where the never sees the threat until past the point of no return.

The drivers had seen the TV in 80% of the SCP events just before the event entered the crash phase (the vehicles collided). The event where the driver was distracted by a *cell phone* is the only event where the driver never sees the TV even at the point of crash. In the remaining 2 events where the driver appeared to have a positive value of IGRT, the IGRT is a very small where the driver appears to have taken eyes off the target just 0.2s before the crash and is annotated as *No eyes visible*. It likely the driver had been looking at the direction of the TV and could be assumed as the driver was looking at the TV while the event entered crash phase. In one of the two events, the driver had seen the TV, but it is possible that they were expecting to pass the TV (even after jumping the red light) but had started braking when they realised that crash was inevitable, nearly 1.5s *after* the point of no return. Table 5 and Table 6 below shows the distribution of IGRTs at the point of no return and crash respectively for the SCP events.

Table 5: IGRT at the point of no return for SCP

Scenario	Number of cases
Driver never sees threat (IGRT = inf)	10
Driver is looking at the threat (IGRT = 0)	3
IGRT > 0	2

Table 6: IGRT at crash for SCP

Scenario	Number of cases
Driver never sees threat (IGRT = inf)	1
Driver is looking at the threat (IGRT = 0)	12
IGRT >0, but very small	2

3.2.2.2 LTAP/OD events

LTAP/OD had a very different result for IGRT at the point of no return, with 93% (14/15) of the drivers looking at the TV at the point of no return. Only one event stands out in that the driver was observed to scan the environment, where the driver transitions his/her gaze from the target to the intersecting road and back at the target, seen in Figure 31 below. An interesting aspect of this event is that the driver began evasive braking at a rate similar to maximum critical braking defined in this study, however, s/he did so too late. The driver was not observed to be in violation of any traffic rules but actively scanned the environment for any other threats before zeroing in the on the TV. The video lacked sufficient resolution to determine if the TV had the turn indicators blinking during the event. An evasive steering action was performed by the driver when s/he realised the collision is imminent, which is surprisingly not observed in most events where the driver has seen the threat before the point of no return.

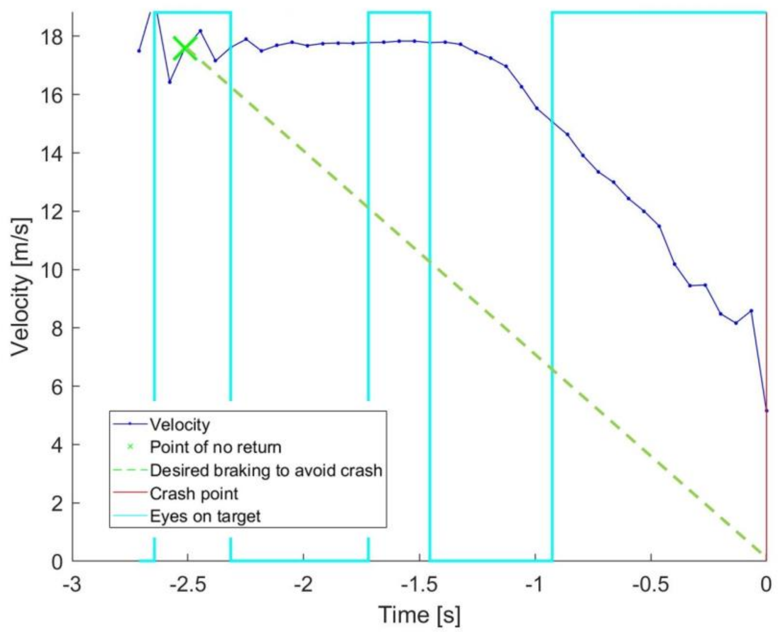


Figure 31: Example of an event with eyes on target as a function of velocity (LTAP/OD event)

The IGRT at crash was mostly zero, drivers never released gaze in all but one event in which the driver was found to look away from the TV to the road ahead. Video analysis of this event reveals this action was possibly performed in an attempt to assess if there were any other possible conflicts with the other vehicles approaching the intersection. The driver was observed to release their gaze from the TV about 0.5s before the crash and approximately 1s after the point of no return. It could be assumed the driver realised that the crash with the primary TV approaching from the left is inevitable and takes pre-emptive measures to avoid conflict with a second TV approaching from the right, this can be observed in Figure 23 (right).

Table 7: IGRT at the point of no return for LTAP/OD

Scenario	Number of cases
Driver never sees threat (IGRT = inf)	2
Driver is looking at the threat (IGRT = 0)	9
IGRT > 0	4

Table 8: IGRT at crash for LTAP/OD

Scenario	Number of cases
Driver never sees threat (IGRT = inf)	0
Driver is looking at the threat (IGRT = 0)	14
IGRT > 0	1

3.2.3 Other time dependent factors

The velocities at the point of no return and just before crash point (two timesteps or 0.13s before the crash) was then compared and is shown in Figure 32. The median velocity at the point of no return is higher in SCP than in LTAP/OD, a similar observation can be made just before the crash.

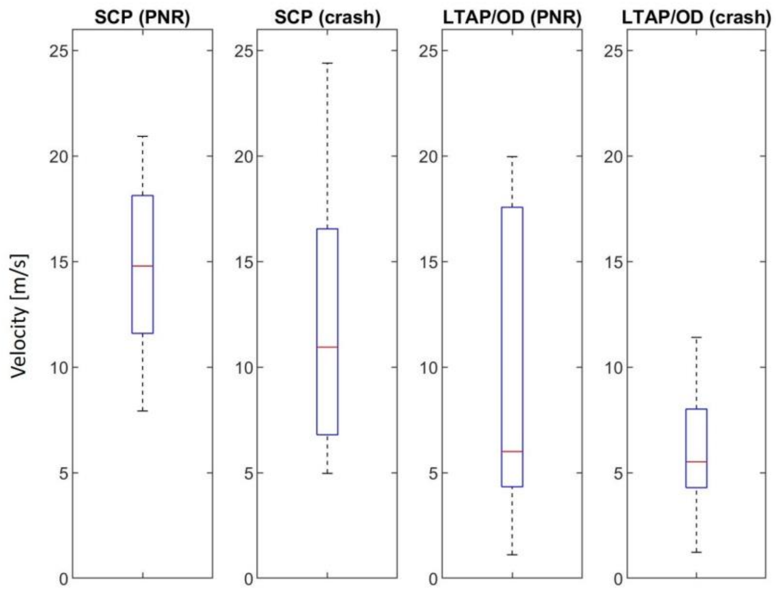


Figure 32: Comparison of velocities at the point of no return and crash point

Comparing the change in velocity from the point of no return to the velocity before the crash, it becomes quite clear that a larger proportion of drivers continued accelerating or at the same velocity in SCP as compared to LTAP/OD. Figure 33 shows the change in velocity from the point of no return to 0.2 seconds before the crash for each case. A negative velocity difference indicates a decrease in the velocity of the SV from the point of no return to the crash point, while a positive velocity difference means the driver was accelerating. Figure 34 below shows the boxplots comparing the same. It can be seen from both these figures that a larger number of drivers either did not change in velocity or continued accelerating in the SCP events.

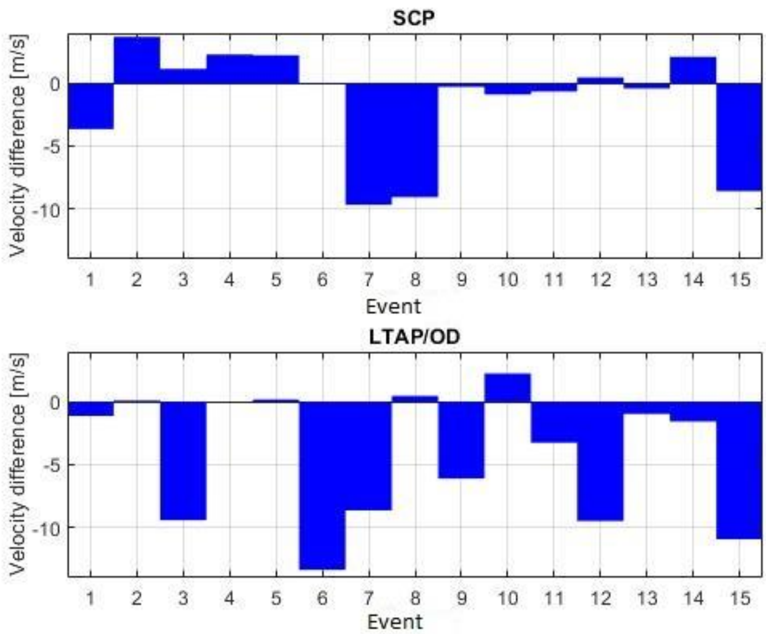


Figure 33: Change in velocity from point of no return to crash for each case

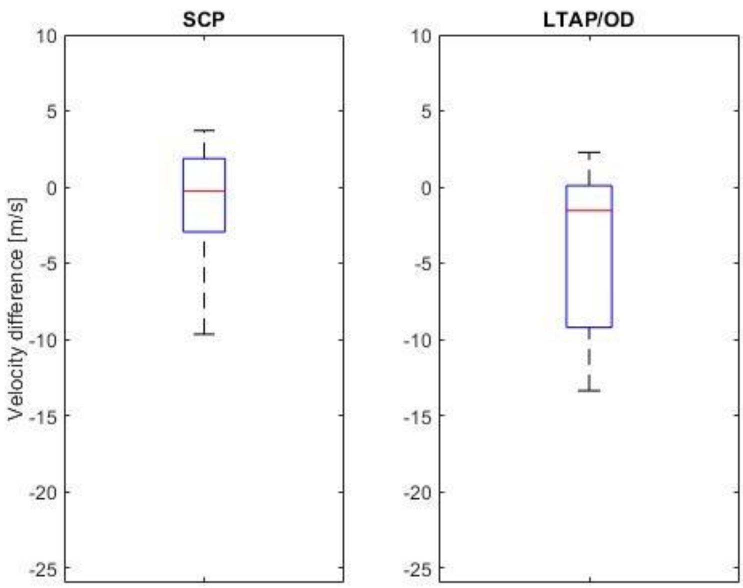


Figure 34: Change in velocity from point of no return to just before the crash

he results obtained above are interesting in that it could be possible that because the drivers in SCP events appear to have seen the threat much later than the drivers in the LTAP/OD cases, they did not react adequately nor prepare to for the crash scenario and approached the intersection at a constant or increasing velocity. Figure 27 and Figure 29 show a large difference in the number of cases where the driver spots the threat before or at the point of no return. This is further seen in Figure 24 and Figure 25 where a larger proportion of drivers see the threat much closer to the crash itself in SCP as compared to with LTAP/OD.

4 Discussion

Studies in the recent years (Werneke & Vollrath, 2012; 2014; Lemonnier, Brémond, & Baccino, 2015) have worked with driver behaviour models such as the SEEV (Wickens, Helleberg, Goh, Xu, & Horrey, 2001; Wickens, McCarley, & Steelman-Allen, 2009), Situational Awareness (SA) (Endsley, 1987; 1995; Wickens, et al., 2003)and more recent predictive processing model (Clark, 2013; Engström, et al., 2018). Many of these studies had been performed with the help of volunteers in simulators and test tracks, for instance, Smith, Bjelkemyr, Bärgman, Johansson, & Lindman (2009) made use of simulator and volunteer data in understanding the driver glance behaviour. With the introduction of naturalistic driving studies, vehicle and traffic safety researchers now have an unprecedented level of information into the various vehicle kinematics and driver behaviours that lead to safety critical situations. Seppelt, et. al. (2017) used the 100-car naturalistic data to gain valuable insight into the driver glance behaviour in real-world crashes and near-crashes.

NDS has also allowed researchers to bridge the gap in results from an experimental environment and the real-world driver data. The unobtrusive nature of naturistic data played a key role in closing this gap. However, this comes with the limitation that high precision eye tracking tools for saving or tracking eye movements of drivers in real time cannot be executed satisfactorily. But the availability of the driver face camera video has allowed researchers to develop a general understanding of the direction in which the driver directs their gaze during the entirety of the event. Careful video analysis by researches has helped break down camera data into glance annotations. These annotations do not provide direct means to determine whether the driver had noticed or observed the threat before, during or after the critical event, however, only serve as an indication that the driver had directed their gaze in that approximate direction.

This research aimed at building a link between the glance annotations available in the NDS and object of interest in the direction of driver's gaze, that is, the target vehicle. This was by developing the toolbox within the MATLAB environment where a method to transform driver glance annotations and vehicle kinematics from the SHRP2 into a virtual environment. This allowed the visualisation of the driver's gaze and glance behaviour. As detailed in the Sections 2.3.2, 2.3.3 and 2.3.4, the method used in this research was an estimation based on the time-series data and the forward camera feed from the NDS. While the actual kinematics of the vehicle including speed, acceleration and yaw used data were recorded in real-time during the event, other variables such as the location of the SV in the intersection, the TV during the event reconstruction and the driver's glances made extensive use of visual cues.

The behaviour of the SV in the virtual world was considered satisfactory for the scope of this research. The SV was found to move in a manner similar to what was observed in the real-world video. The TV, however, with several iterations of trial and error eventually recreated the movement of the TV comparable to the video footage. The TV movement was limited in the sense that its movement could be reproduced only from the time that it first appears on the video feed. This was limited by the field of view of the camera, the resolution and clarity of the feed and finally, the actual orientation of the SV in the video with respect to the TV. It would not be correct try to estimate the position of the TV before it appears, even when there is no other visual occlusion because of possible variations in velocity, lane and orientation. This is especially true for SCP events where the TV is not visible to the field of view of the camera until the TV is very close to the SV. Further, the low resolution and clarity of the video feed made it difficult to judge precisely when the TV first appears in the video or even its exact location. This required careful analysis of the video to ensure the TV is correctly identified as early and as accurately as possible. Night time and poor weather posed additional challenges where headlamps from several oncoming vehicles made it difficult to ascertain which ones belonged to the TV. Rain and snow resulted in distorted images and lights in the video.

Once the TV's movements were successfully reconstructed, the reconstruction was carefully compared to the real-world video and glance annotations to ensure glance location, TV position and the field of view were in sync. Because of the nature geometry of the field of view defined in the study, there were instances which resulted in narrow misses in what was most likely *eyes on target*. The use of approximation radius around the TV proved very helpful in avoiding this problem. In a real-world scenario, this field of view would not have been geometrically bound as such, and it is likely that the peripheral view in addition to the foveal vision could have aided the driver in identifying and assessing the threat.

In general, the toolbox provided a good understanding and insight into the different areas of interest for the driver during the event. An interesting finding of this study was the striking number of events where the driver had looked at the threat at or before the point of no return and continued tracking the object continuously until the crash. Understanding this behaviour by applying the predictive processing model discussed by Engström et. al. (2018) is interesting. It appeared that the drivers had *looked* at the TV, *but failed to see* it as a threat; a term introduced by Staughton & Storie (1977) to explain the large number of crashes with drivers having looked at the target but failing to act. Drivers approaching the intersection, develop a pattern in which they predict the situation would play out, modulating the controls of their vehicle based on an error in predicted pattern and the actual pattern. Considering the driver had been looking in the direction the TV in 93% of the LTAP/OD cases, it is likely they failed to gauge the error between the predicted pattern of how the events will play out against the reality.

This is relevant in context of this study because as proposed by Smith, Bjelkemyr, Bärgman, Johansson, & Lindman (2009), it was expected that the driver would change the direction of glance towards the centre of the intersection or the road ahead when they had understood/determined that the TV is no longer a threat. As a result, it could have been interpreted as a case of misjudgement or poor decision by the SV. However, the IGRT at crash was zero in nine out of fifteen cases of LTAP/OD events and predominantly zero at the point of no return for the same, meaning the driver continued looking at the TV from the point of no return to crash in most of the cases. Comparing this with the change in velocities between the point of no return to just before the crash raises an interesting question - *At* *which point in time did the driver fully perceive the threat as the collision was imminent?*

To explain this, the following hypothesis was derived in an attempt to understand this behaviour. Michon (1985) explains that control when driving is hierarchal in nature, consisting of three levels namely, Strategic, Tactical and Operational controls. Strategic level or the highest level correspond to long-term goals of driving. This includes factors such as choice of route and motivational factors. At the tactical level or current goals, there are factors like headway distance or decision to overtake. And finally, at an operational level are control decisions such as pressing the brake pedal or steering the vehicle. Putting all the three into a single example could be seen as follows: consider a driver on their way to work. At a strategic level, they make the decision to follow their usual route because of the low traffic volumes. As they approach an intersection, they decide at a tactical level, which would be *I will turn left here*. And finally, at the operational level, the driver modulates the brake and accelerator, and turns the steering to carry out the tactical level objective.

Analysing this further, it can be said that at the tactical level, there is a top-down (Casson, 1983) expectation, or motivated by an overall goal, which is to turn left and drive safely through the intersection. At the operational level, however, the controlling actions are performed based on the sensory inputs of one's own trajectory and position, that is, bottoms-up processing (Casson, 1983). Driving itself is often a highly automatized process for most people, this is especially true when driving through familiar routes (Reason, 1984; Logan, 1983). This was further corroborated by Charlton & Starkey (2011; 2013) through volunteer tests where the participants became familiar to a certain route by driving through it several times over the period of the experiment. To quote participant feedback from their experiments - "*'I found myself going into auto, not paying attention to what I was doing'*, *'Feels very normal, just like the drive home; thinking mostly about food'* (Charlton & Starkey, 2011, p. 467)". That is, over time, the process of driving becomes so automated that a driver might not be aware of the decisions or choices they make during their trip (Endsley, 1995).

Thus, the top-down expectation becomes very strong and dominant over time. The hypothesis presented is that if this top-down expectation is so strong, could it be possible that even though the brain is receiving bottoms-up feedback that something is not right, the brain overrides or pushes away this feedback? In other words, when a driver is driving through the same intersection they have driven on every day, they would expect that '*today's trip is like every other day'*. So, when there is a small deviation from this, the brain fails to completely comprehend this deviation due to the stronger expectation of '*driving through like every other day'*. The predictive processing model explains that the brain makes an array of predictions from the sensory inputs it is receiving and compares these against what is actually happening to identify and tune its model to minimise this error. However, when the stronger motivation is '*I will drive through like I do every day*', even though there is a bottoms-up input which informs the brain '*Something is wrong!'* the brain does not completely acknowledge the input. This could be possibly the reason that majority of the drivers appeared to have seen the TV and continue tracking the TV all the way to the crash itself, with little to no evasive actions.

To further understand this reasoning, consider Figure 35, which is the same event previously presented in Figure 17 and Figure 18. In this LTAP/OD crash, the SV approached the intersection on the dedicated left turn lane. The traffic signal was green throughout the event. The driver slows down on approaching the intersection, allowing a first oncoming vehicle to pass. The driver, however, fails to notice the second vehicle approaching the crossing several meters behind the first vehicle as s/he proceeds with the left turn. Even though left turn manoeuvre can be assumed as a safe, legal manoeuvre, the driver failed to notice this approaching second vehicle despite looking in that direction while beginning the left turn manoeuvre.

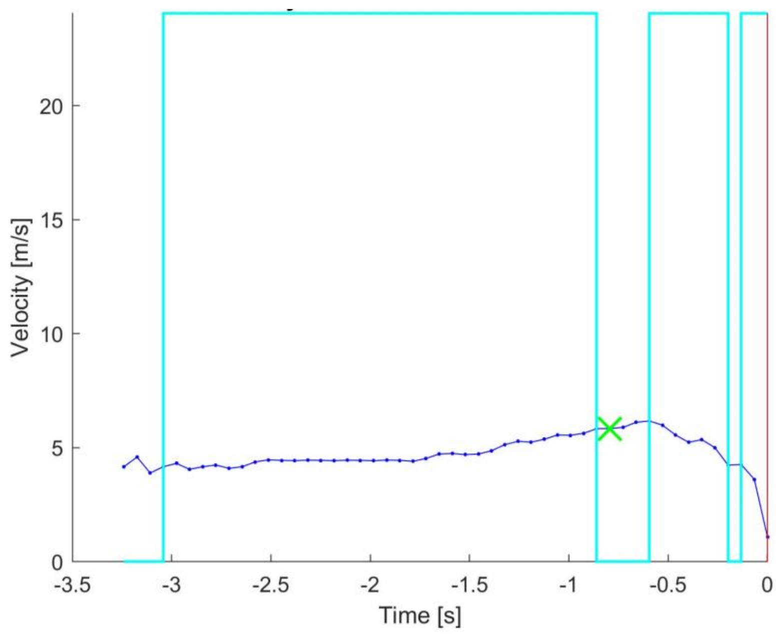


Figure 35: Was the driver looking at the TV throughout the incident? (Legend see Figure 31)

This leads to the reasoning that perhaps the driver failed completely to perceive the rate at which the TV was approaching when s/he starts performing the left turn manoeuvre. A possible inference is that as the first vehicle passes, the driver quickly switched between the green light, possibly clear path ahead and his/her approach velocity into the intersection, making a mental model of how s/he would expect his/her safe passage through the intersection to happen. In this dynamic situation, it is likely that the driver failed to acknowledge that the TV was approaching at a higher rate than s/he had anticipated. However, it was observed that as soon as the driver realised the error in the initial prediction or expectation against the actual scenario, the driver brakes, resulting in only a graze against the TV.

Another example where such the driver was observed as looking forward with his eyes on the road but fails to perceive the threat is shown in Figure 36 below.

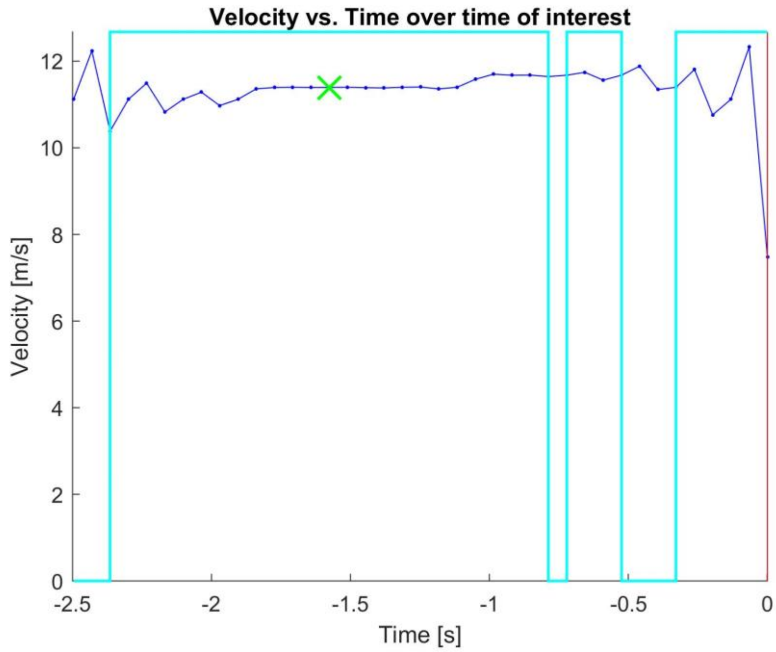


Figure 36: Looking at the TV but fails to see TV or red light. (Legend see Figure 31)

The final event summary from the database is as follows-

> *“Subject vehicle is travelling in the left lane of an undivided roadway in an urban area. As subject approaches a signalized intersection, he begins tapping his fingers on the steering wheel, possibly distracted by the radio. Even though the subject's eyes are on the road ahead, he seems cognitively absent and does not react to the red light at the intersection. As subject begins to pass through the intersection, he appears to notice V2 crossing the intersection from the left on a green light, but still fails to decelerate for the red light. Once the subject driver realizes that impact can occur, he brakes hard but is unable to avoid striking the right side of V2. Both vehicles skid before stopping on the right side of the intersecting roadway.”*

While the event description states the subject was possibly distracted due to an auditory excitation, he fails to recognise and register multiple salient events, the red light, TV and even his own rate of approach into the intersection. Figure 36 shows that the driver could have been looking at the TV for over 90% of the time from when the TV appears in the field of view. It is likely that as the driver starts moving from a first traffic signal which just turned green, he assumed that the following one would also be green. With this expectation in mind, he approaches the intersection expecting a safe passage through. He glances at TV assuming it would yield for him to pass. However, as the event progresses, the error in predicted behaviour of the TV becomes evident to the driver, albeit too late.

Yet another case to support this reasoning had the following incident report-

> *“The subject driver travels on an undivided road. V2 exits a parking lot on the subject driver's left and travels straight across the subject vehicle's path. The subject driver sounds the horn and brakes but cannot avoid colliding with the right rear side of V2.”*

Looking at the velocity and *eyes on target* plot for this event in Figure 37, it is seen that the driver was, in fact, accelerating as s/he approaches the TV.

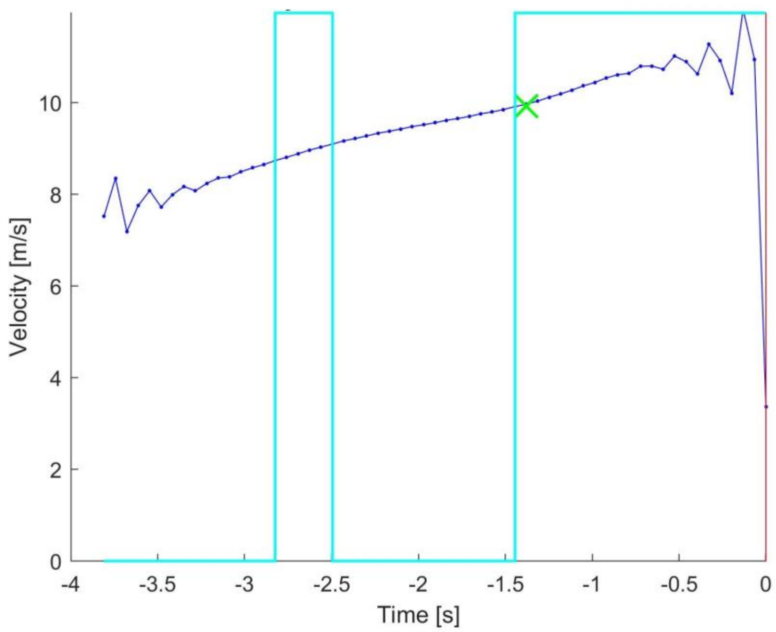


Figure 37: Driver sees the TV but fails to perform control manoeuvres. (Legend see Figure 31)

From Figure 38, it was observed that driver notices the TV, transitions his/her gaze towards the right side for over a second, before returning gaze back towards the TV. The driver, however, does not perform any evasive manoeuvre, although it was documented that s/he honked. The driver possibly expected the TV to yield for SV to pass as the SV had priority. Having decided it would not be a threat, the driver glances towards the right. When the driver realises this event had not evolved as s/he had expected, s/he honks to alert the TV, still expecting the TV to yield. However, the driver realizes far too late that the event did not evolve in the manner predicted.

In each of these cases, the driver had possibly seen the threat long before the point of no return but failed to act. The SV was the striking vehicle in each of these events. Further, the drivers did not appear to understand the deviation in predicted behaviour, and in extension, was unable to compare or update their prediction model against the real-time data. Thus, the brain considers itself to already be *in tune* with the environment, consequently leading to a failure of perception.

And thus, to summarize the hypothesis, driving is a highly automatized task for most people (Reason, 1984; Logan, 1983), driven strongly by short- and long-term goals. Over time, these goals, or expectations become stronger or dominant to enable this automaticity in driving. On the rare occasion, when something occurs out of the ordinary, the top down expectation could be stronger than the bottoms up input which alerts something is not happening as expected. This could be a possible explanation for why drivers who appear to have seen the threat fail to acknowledge it as a threat.

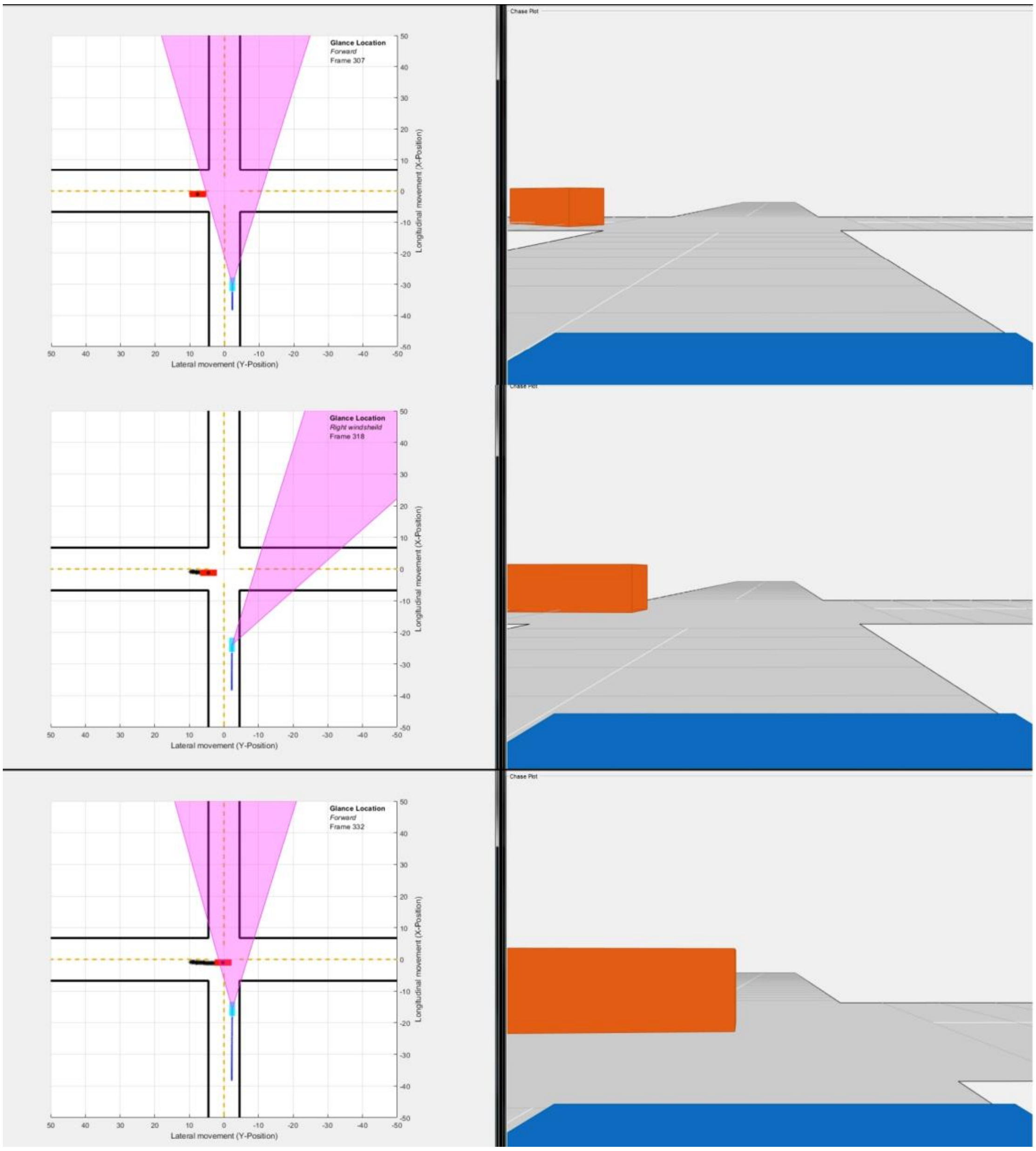


Figure 38: Glances over the whole event

It would be interesting to validate this hypothesis against near-crashes and understand the difference, if any, between them. Seppelt et. al (2017) observed a striking difference in the manner in which the drivers switched gaze in near-crashes and crashes, proposing that the way the drivers switched between gaze locations provided drivers in near-crashes the valuable information that helped avert the crash. Further, it would be interesting to analyse the looming in these cases and understand if it had any effect in the way the drivers perceived the threat. The rate at which the TV was approaching the SV would be most evident through looming; looming is the rate at which the projected image grows in size on the retina of the eye as the object moves closer to the subject (Raviv & Joarder, 2000). Markkula, Engström, Lodin, Bärgman, & Victor (2016) found that drivers involved in rear-end crashes in the SHRP2 NDS did not react until looming reached a threshold. This is a possible direction for future research in intersection crashes.

5 Conclusions

Driving is a complex task where a driver has to pay attention to not only their status in the environment but also that of the other vehicles and road users. Although complex, the repetitive nature of driving in most of our lives makes this a highly automatized process. Over time, with increased experience and familiarity of routes and traffic, we develop expectations. The human brain is a complex computer that not only uses what it has learnt over time but also information from the world around it to understand and comprehend what it is happening to minimise entropy. Many researchers, based on biological evidence, believe that the brain is constantly trying to be in a state of synchronisation with the world around it. This is said to be done by using the inputs from all five senses and developing an array of predictions for the future states of these inputs. These predictions are based on both the sensory inputs and the expectation of the situation itself. The brain smoothens this model to the real-time information by acting to minimise the deviation between the prediction and real-time data.

However, with repetitive or highly automatized driving patterns, the expectations or short-term and long-term goals could eventually become so strong that in the rare event that there is a deviation from ordinary, the brain could easily overlook the feedback from sensory input. As a result of which, the brain does not have enough time to correct its internal model to match the real-time information, leading to a lapse in the way a person responds in the situation. While this hypothesis could be an explanation for the crashes investigated within the scope of this research, the limited number of cases instigates the need for further, in-depth studies into the gaze behaviour at intersections..

On a final note, the toolbox developed in this research is possibly among the first of its kind attempts at reconstructing the gaze behaviour of the driver in intersections using naturalistic driving data. The findings of this research show promise in continuing the development of such toolboxes or other similar methods to gain a further and deeper understanding of driver gaze and glance behaviour. With the vast troves of naturalistic driving studies and soon, data from commercial driver monitoring systems, research into driver behaviour and development of safety systems stand to gain. After all, every life saved is a step in the right direction.

6 References

Antin, J. F. (2011). *Design of the in-vehicle driving behavior and crash risk study: in support of the SHRP 2 naturalistic driving study.* Transportation Research Board.

Aust, M. L., Fagerlind, H., & Sagberg, F. (2012). Fatal intersection crashes in Norway: Patterns in contributing factors and data collection challenges. *Accident Analysis and Prevention*, 782-791.

Bärgman, J. (2016). *Methods for Analysis of Naturalistic Driving Data in Driver Behaviour Research.* Göteborg: Chalmers University of Technology.

Bärgman, J., Lisovskaja, V., Victor, T., Flannagan, C., & Dozza, M. (2015). How does glance behavior influence crash and injury risk? A 'what-if' counterfactual simulation using crashes and near-crashes from SHRP2. *Transportation Research Part F, 35*, 152-169.

Bärgman, J., Smith, K., & Werneke, J. (2015). Quantifying drivers' comfort-zone and dread-zone boundaries in left turn across path/opposite direction (LTAP/OD) scenarios. *Transportation Research Part F 35*, 170-184.

Battelle Memorial Institute. (2007). *Evaluation of the Volvo Intelligent Vehicle Initiative Field Operational Test, Final Report Version 1.3.* Washington, D.C.: U.S. Department of Transportation.

Björklund, G. M., & Åberg, L. (2005). Driver behaviour in intersections: Formal and informal traffic rules. *Transportation Research Part F, 8*, 239-253.

Blatt, A., Pierowicz, J., Flanigan, M., Lin, P., Kourtellis, A., Lee, C., . . . Richard, C. (2015). *Naturalistic Driving Study: Field Data Collection No. SHRP 2 Report S2-S07-RW-1.* Washington, D.C.: Transportation Research Board.

Bougler, B., Cody, D., & Nowakowski, C. (2008). *California intersection decision support: a driver-centered approach to left-turn collision avoidance system design.* University of California, Berkeley.

Brännström, M., Sjöberg, J., & Coelingh, E. (2008). A situation and threat assessment algorithm for a rear-end collision avoidance system. *Intelligent vehicles symposium, IEEE*, 102-107.

Bulger, T., & Crane, S. (1998). TRANSPORTATION EQUITY ACT FOR THE 21ST CENTURY (TEA 21). *TDM Review, 6(3).*

Casson, R. W. (1983). Schemata in cognitive anthropology. *Annual review of anthropology, 12(1)*, 429-462.

Chalmers University of Technology. (2016, November 11). *Quantitative Driver Behaviour Modelling for Active Safety Assessment Expansion (QUADRAE).* Retrieved from Chalmers University of Technology: http://www.chalmers.se/en/projects/Pages/Quantitative-Driver-Behaviour-Modelling-for-Active-Safety.aspx

Charlton, S. G., & Starkey, N. J. (2011). Driving without awareness: The effects of practice and automaticity on attention and driving. *Transportation Research Part F 14*, 456-471.

Charlton, S. G., & Starkey, N. J. (2013). Driving on familiar roads: Automaticity and inattention blindness. *Transportation Research Part F 19*, 121-133.

Choi, E.-H. (2010). *Crash Factors in Intersection-Related Crashes: An On-Scene Perspective (No. HS-811 366).* Washington, DC: National Highway Traffic Safety Administration.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences, 36(3)*, 181-204.

Dingus, T. A., Klauer, S., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., . . . Knipling, R. (2006). *The 100-Car Naturalistic Driving Study, Phase II – Results of the 100-Car Field Experiment.* Washington, DC: National Highway Traffic Safety Admin. (NHTSA).

Endsley, M. R. (1987). *A methodology for the measurement of situation awareness (NOR DOC 87-83).* Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37(1)*, 32-64 .

Engström, J., Bärgman, J., Nilsson, D., Seppelt, B., Markkula, G., Piccinini, G. B., & Victor, T. (2018). Great expectations: a predictive processing account of automobile driving. *Theoretical issues in ergonomics science, 19(2)*, 156-194.

Galpin, A., Underwood, G., & Crundall, D. (2009). Change blindness in driving scenes. *Transportation Research Part F 12*, 179–185.

Gardner, J. W., & Sanborn, J. S. (1990). Years of Potential Life Lost (YPLL)—What Does it Measure? *Epidemiology, 1(4)*, 322-329.

Guo, F., Klauer, S. G., McGill, M. T., & Dingus, T. A. (2010). *Evaluating the Relationship Between Near-Crashes and Crashes: Can Near-Crashes Serve as a Surrogate Safety Metric for Crashes?* Washington D.C.: National Highway Traffic Safety Administration.

Hendricks, D. L., Freedman, M., Zador, P. L., & Fell, J. C. (2001). *The relative frequency of unsafe driving acts in serious traffic crashes.* Washington, D.C.: National Highway Traffic Safety Administration.

Jansen, R. J., Lotan, T., Winkelbauer, M., Bärgman, J., Kovaceva, J., Donabauer, M., . . . Nes, N. v. (2017). *Interactions with vulnerable road users.* UDRIVE Deliverable 44.1. EU FP7 Project UDRIVE Consortium. Retrieved from https://doi.org/10.26323/UDRIVE_D44.1

Kulmala, R. (1990). Driver behaviour at urban junctions with the right-hand rule. *Proc., 3rd ICTCT Workshop*.

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in retinal and eye research, 25(3)*, 296-324.

Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature, 369(6483)*, 742-744.

Lemonnier, S., Brémond, R., & Baccino, T. (2015). Gaze behavior when approaching an intersection: Dwell time distribution and comparison with a quantitative prediction. *Transportation research part F: traffic psychology and behaviour, 35* , 60-74.

Liu, M., Lu, G., Wang, Y., & Zhang, Z. (2014). Analyzing drivers' crossing decisions at unsignalized intersections in China. *Transportation Research Part F, 24*, 244-255.

Logan, G. D. (1983). Automaticity, Resources, and Memory: Theoretical Controversies and Practical Implications. *Human Factors, 30(5)*, 583–598.

Lord, D., Schalkwyk, v., I., S. L., & Chrysler, S. (2005). *Reducing Older Driver Injuries at Intersections Using More Accommodating Roundabout Design Practices.* Texas Transportation Institute .

Markkula, G., Engström, J., Lodin, J., Bärgman, J., & Victor, T. (2016). A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies. *Accident Analysis and Prevention 95*, 209–226.

Martinez, J. E. (1989). A primer on automobile accident reconstruction (No. 890631). *SAE Technical Paper*.

Martins, J., Ribeiro, R., Neves, P., & and Brito, F. (2016). Accident Reconstruction Using Data Retrieval from Crash-Test Video Images. *SAE Technical Paper 2016-01-1464*.

Michon, J. A. (1985). A critical view of driver behavior models: what do we know, what should we do? *Human behavior and traffic safety*, 485-524.

Najm, W., Smith, J. D., & Smith, D. L. (2001). *Analysis of crossing path crashes (No. DOT-VNTSC-NHTSA-01-03).* John A. Volpe National Transportation Systems Center (US).

Neale, V., Klauer, S., Knipling, R., Dingus, T., Holbrook, G., & Petersen, A. (2002). *The 100 car naturalistic driving study, Phase I-experimental design (No. HS-809 536,).* Washington, D.C.: U.S. Department of Transportation, National Highway Traffic Safety Administration.

NHTSA. (2008). *National Motor Vehicle Crash Causation Survey: Report to Congress.* Washington, D.C.: National Highway Traffic Safety Administration.

NHTSA. (2017). *Traffic Safety Facts 2015.* Washington, D.C.: National Highway Traffic Safety Administration.

Nobukawa, K., Barnes, M., Goodsell, R., & Gordon, T. (2009). Reconstruction of Vehicle Trajectories for Intersection Conflict Analysis Using Vehicle-Based Sensors. *International Association for Vehicle System Dynamics*.

Ragland, D. R., & Zabyshny, A. A. (2003). *Intersection decision support project: Taxonomy of crossing-path crashes at intersections using GES 2000 data.* UC Berkeley: Safe Transportation Research & Education Center.

Rajamani, R. (2011). *Vehicle dynamics and control.* Springer Science & Business Media.

Raviv, D., & Joarder, K. (2000). The visual looming navigation cue: a unified approach. *Computer Vision and Image Understanding, 79(3)*, 331-363.

Reason, J. T. (1984). Absent-mindedness and cognitive control. *Everyday memory, actions and absentmindedness*, 113-132.

Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of experimental psychology: Applied, 6(1), 31.*

Recarte, M. A., & Nunes, L. M. (2003). Mental workload while driving: effects on visual search, discrimination, and decision making. *Journal of experimental psychology: Applied, 9(2), 119.*

Romoser, M. R., Pollatsek, A., Fisher, D. L., & Williams, C. C. (2013). Comparing the glance patterns of older versus younger experienced drivers: Scanning for hazards while approaching and entering the intersection. *Transportation research part F: traffic psychology and behaviour, 16*, 104-116.

Rumar, K. (1990). The basic driver error: late detection. *Ergonomics, 33:10-11*, 1281-1290.

Senders, J. W. (1964). The human operator as a monitor and controller of multidegree of freedom systems. *EEE Transactions on Human Factors in Electronics, (1)*, 2-5.

Seppelt, B. D., Seaman, S., Lee, J., Angell, L. S., Mehler, B., & Reimer, B. (2017). Glass half-full: On-road glance metrics differentiate crashes from near-crashes in 100-Car data. *Accident Analysis and Prevention*, 48-62.

Sharma, D., Stern, S., Brophy, J., & Choi, E.-H. (2007). An overview of NHTSA's crash reconstruction software WinSMASH. *Proceedings of the 20th International Technical Conference on Enhanced Safety of Vehicles.*
Smith, K., Bjelkemyr, A., Bärgman, J., Johansson, B., & Lindman, M. (2009). *IVSS Intersection accidents: Analysis and Prevention Final Report - Volume 1 (Main report).* Göteborg, Sweden.
Smith, R. A., & Tsongos, N. G. (1986). Crash Phase Accident Reconstruction (No. 860209). *SAE Technical Paper*.
Staughton, G. C., & Storie, V. J. (1977). *Methodology of an in-depth accident investigation survey (No. TRRL Lab Rept. 762).* Wokingham, Berkshire United Kingdom : Transport and Road Research Laboratory (TRRL).
Summala, H., Pasanen, E., Räsänen, M., & Sievänen, J. (1996). Bicycle accidents and drivers' visual search at left and right turns. *Accident Analysis & Prevention, 28(2)*, 147-153.
Tay, R., & Rifaat, S. M. (2007). Factors Contributing to the Severity of Intersection Crashes. *Journal of Advanced Transportation, 41(3)*, 245-265.
Tenney, Y. J., & Pew, R. W. (2006). Situation Awareness Catches On: What? So What? Now What. *Reviews of Human Factors and Ergonomics 2(1)*, 1–34.
Transportation Research Board. (2001). *TRB Special Report 260: Strategic Highway Research: Saving Lives, Reducing Congestion, Improving Quality of Life.* Washington, D.C.: Transportation Research Board.
Treat, J., Tumbas, N., McDonald, S., Shinar, D., Hume, R., Meyer, R., . . . Castellan, N. (1979). *Tri-Level study of the causes of traffic accidents: Final Report.* Washington D.C.
van Schagen, I., Welsh, R., Backer-Grøndahl, A., Hoedemaeker, M., Lotan, T., Morris, A., . . . Winkelbauer, M. (2011). *Towards a large scale European Naturalistic Driving study: final report of PROLOGUE: Deliverable D4.2.* Leidschendam, The Netherlands: SWOV Institute for Road Safety Research.
Victor, T., Dozza, M., Bärgman, J., Boda, C.-N., Engström, J., Flannagan, C., . . . Markkula, G. (2015). *Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk (No. SHRP 2 Report S2-S08A-RW-1).* Washington, D.C.: Transportation Research Board.
Werneke, J., & Vollrath, M. (2012). What does the driver look at? The influence of intersection characteristics on attention allocation and driving behavior. *Accident Analysis and Prevention, 45*, 610-619.
Werneke, J., & Vollrath, M. (2014). How do environmental characteristics at intersections change in relevance for drivers before entering an intersection: analysis of drivers' gaze and driving behavior in a driving simulator study. *Cognition, Technology and Work, 16*(2), 157-169.
Wickens, C. (2000). The trade-off of design for routine and unexpected performance: Implications of situation awareness. *Situation awareness analysis and measurement*, 211-225.
Wickens, C., Helleberg, J., Goh, J., Xu, X., & Horrey, W. J. (2001). *Pilot task management: Testing an attentional expected value model of visual scanning.* Savoy, IL: UIUC Institute of Aviation Technical Report.
Wickens, C., Hollands, J., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology & human performance.* Taylor & Francis.

Wickens, C., McCarley, J., & Steelman-Allen, K. (2009). NT-SEEV: A model of attention capture and noticing on the Flight Deck. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 53, No. 12)*, 769-773.

Wickens, C., McCarley, J., Thomas, L., Foyle, M. I., Goodman, A., & Hooey, B. L. (2003). Attention-Situation Awareness (A-SA) Model. *NASA Aviation Safety Program Conference on Human Performance Modeling of Approach and Landing with Augmented Displays*, 189-225.

World Health Organisation. (2015). *Global status report on safety.* World Health Organisation.