

NOVEL METHODS IN TRAINING AUTONOMOUS
VEHICLES FOR URBAN ROADS

by

Rafael Vasquez

B.Sc. (Bachelor of Computer Science)

Ryerson University, Toronto, 2018

A Thesis

presented to Ryerson University

in partial fulfillment of the

requirements for the degree of

Master of Applied Science

in the Program of

Civil Engineering

Toronto, Ontario, Canada, 2020

© Rafael Vasquez, 2020

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A THESIS

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

NOVEL METHODS IN TRAINING AUTONOMOUS VEHICLES FOR URBAN ROADS

Rafael Vasquez
Master of Applied Science
Civil Engineering
Ryerson University, 2020

Abstract

This thesis presents the development and application of a novel platform to train autonomous vehicles (AV) for urban roads. Interactive and immersive virtual reality (VR) environments are developed for the collection of mobility preference, behaviour, movement, and orientation data. The resulting naturalistic data can be used directly to train AV control systems. This platform is exemplified in an end-to-end case study resulting in a multi-objective braking system which maximizes both pedestrian safety and passenger comfort. It begins with the development of an immersive VR pedestrian road-crossing environment and compilation of a unique, naturalistic dataset. A vehicle agent is then successfully trained against the dataset, learning a multi-objective brake control policy using deep reinforcement learning methods and reducing the negative influence on passenger comfort by half while maintaining safe braking operation. This platform offers the opportunity to simulate complex, human-in-the-loop scenarios AVs will inevitably face and train them for these scenarios.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Dr. Bilal Farooq for his role in my research and personal development. Having worked with him since before I intended to pursue graduate studies, he has acted as a mentor, providing direction and insight in each of my endeavours as well as countless opportunities to grow as a researcher and critical thinker.

I would also like to thank each of my colleagues in the Laboratory of Innovations in Transportation for their continued support. I would like to mention in particular Dr. Shadi Djavadian, with whom I collaborated on several occasions, for her guidance and encouragement throughout my time with the team.

Finally, I would like to express my appreciation for my family and friends and their love and care throughout my studies.

Rafael Vasquez

TABLE OF CONTENTS

Abstract	ii
Acknowledgements	v
List of Tables	vii
List of Figures	viii
I. Overview	1
1 Introduction	1
1.1 Motivation	1
1.2 Thesis Scope	2
1.3 Thesis Objective	3
1.4 Thesis Outline	3
2 Literature Review	5
2.1 Datasets	5
2.1.1 Non-Naturalistic	5
2.1.2 Naturalistic	7
2.2 Simulation	8
2.3 AV Training Techniques	10
2.4 Key Remarks	12
II. Methodology	15
3 Virtual Reality System	17
3.1 Virtual Reality as a Data Collection Tool	17
3.2 Virtual Immersive Reality Environment (VIRE)	18
3.2.1 Scenario Development (SD) System	19
3.2.2 Multi-Modal Traffic Micro-Simulation (MMTM) System	19
3.2.3 Virtual Environment Projection (VEP) System	19
3.2.4 Driving in VIRE	20
4 Deep Reinforcement Learning System	22
4.1 Policy Proximal Optimization (PPO)	23
4.2 Deep Deterministic Policy Gradient (DDPG)	24
4.3 Reward Function	24

III. Experiment and Results	27
5 Case Study	27
5.1 Naturalistic Data Collection	28
5.2 Training Scenario	31
5.3 Training Simulation Setup	32
5.4 Training Parameters	33
5.5 Results	35
6 Conclusion, Limitations and Future Work	38
6.1 Conclusion.....	38
6.2 Limitations	38
6.3 Future Work	39
Appendices.....	40
A. Should I give or should I take? Choice issues in automated vehicle control	41
B. Virtual Immersive Reality based Analysis of Behaviour Responses in Connected and Autonomous Vehicle Environment.....	46
References	55

List of Tables

Table 1: Comparison of Datasets	6
Table 2: Comparison of Simulation Platforms	8
Table 3: Levels of Adjustable Parameters	28
Table 4: PPO Configuration.....	34
Table 5: DDPG Configuration	34
Table 6: Reward Function Configurations.....	35

List of Figures

Figure 1: The Road-Crossing Scenario	2
Figure 2: Data Collection and Training Platform	15
Figure 3: Driving in VIRE	21
Figure 4: VIRE Pedestrian Physical Environment.....	29
Figure 5: VIRE Pedestrian-AV Virtual Environment.....	30
Figure 6: Waiting Time Frequency Distribution	31
Figure 7: Training Scenario	32
Figure 8: Cumulative Environment Reward	35
Figure 9: Profile of velocity, jerk, and actions in an example episode against the same road-crossing trial.....	37

I. Overview

1 Introduction

1.1 Motivation

Autonomous vehicles (AVs) and its technology has the potential to revolutionize or have a strong impact on many aspects of our civil society such as mobility, land use, individual behaviour and culture. Along with the continual advancement of the reinforcement learning paradigm for robotic systems and complex control tasks, the need for large and realistic training datasets is increasing. The quality of these datasets is central to the development of AVs, algorithms and tools need to be trained and evaluated before vehicles are deployed to physical environments. Most of these datasets are image-based and the annotated objects featured in them vary depending on the component of the AV system being developed and the learning approach used [1].

AV simulation platforms provide high-fidelity environments for data generation and to support the development, training, and validation of autonomous driving systems at a low cost. However, the traffic participants (pedestrians, cyclists, drivers) simulated in these platforms are confined to the underlying models used to describe them [2]. For example, some platforms include pedestrians in order to train vehicle agents to react to life-endangering situations that could occur in reality. However, pedestrian road-crossing behaviour is complex and can vary highly in aspects such as waiting time, pose, and path. These intricate behaviours can be difficult to simulate yet are crucial when considering pedestrian safety and the inevitable dangerous situations that AVs find themselves in.

The use of virtual reality (VR) as a data collection tool includes the capability to capture human movement, action and orientation data within highly customizable and well-controlled environments. Whether driving on or walking across a virtual road network, by providing respondents with a realistic and interactive experience, they can experience a sensation of immersion in the activities and their behaviour imitates that of the physical world [3], [4].

1.2 Thesis Scope

Consider a common pedestrian road-crossing situation with an approaching vehicle. Both road users must be able to anticipate what the other intends to do. Typically, the pedestrian will want to confirm that the driver is aware of them through eye contact, gestures, or by perceiving that the vehicle is slowing down. The driver should practice caution, confirming that the pedestrian is aware of their vehicle and, depending on who has the right-of-way, either prepare to react in the case that the pedestrian wrongly decides to cross or decelerate to let them cross safely.

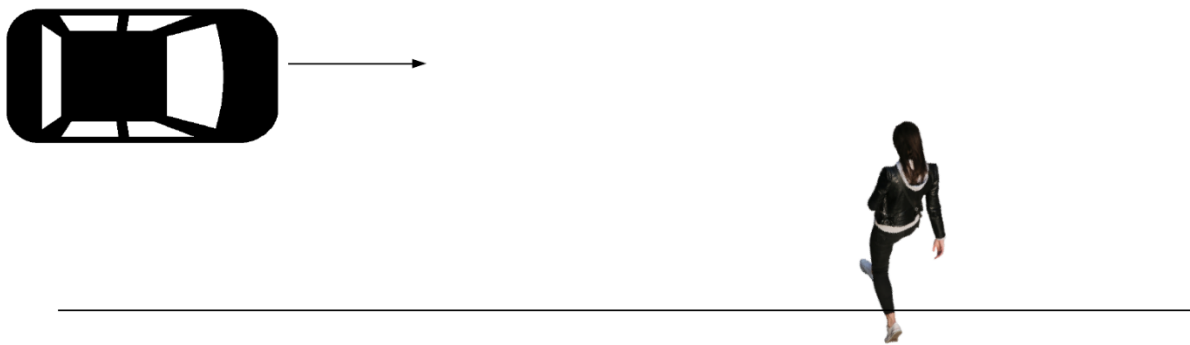


Figure 1: The Road-Crossing Scenario

The scope of this thesis is the road-crossing scenario in the absence of a human driver and the consideration of several environmental factors that affect a pedestrian’s road-crossing behaviour.

1.3 Thesis Objective

The first objective of this thesis is the development of a novel platform to collect and use naturalistic data towards the training of AVs for urban roads in simulation. The second objective is to illustrate the platform with an end-to-end case study in which pedestrian motion data is collected using VR and simulated directly in an AV training environment. The AV is trained against the collected data to learn a multi-objective policy for safe and comfortable brake control.

In conjunction, these objectives contribute to the body of research methods and provides examples for the creation of complex, human-in-the-loop mobility scenarios for naturalistic data collection along with the ability to prepare AVs.

1.4 Thesis Outline

Chapter 2 provides an examination of popular AV-related datasets (2.1), prevalent AV simulation platforms (2.2), and state-of-the-art AV training techniques (2.3).

Chapter 3 introduces VR as a data collection tool (3.1) and the novel Virtual Immersive Reality Environment platform (3.2).

Chapter 4 outlines the deep reinforcement learning methods (4.1, 4.2) and reward function (4.3) used in the case study.

Chapter 5 describes the data used in the case study (5.1), the training scenario, setup and parameters (5.2 – 5.4), and shows experimental results (5.5).

Chapter 6 provides final remarks along with a discussion of limitations and future work.

2 Literature Review

The platform presented considers two important elements for AV development: data and simulation. As such, the following review explores the current state of AV-related datasets as well as the simulation platforms used for training and testing. Additionally, recent techniques used to develop AV systems are reported and contrasted with the method presented in this thesis.

2.1 Datasets

As seen below, most datasets used for the purpose of training AV systems are comprised of image and video data. Algorithms with different objectives are typically trained and validated separately. As such, datasets tend to be captured and annotated for the purpose of training with a specific objective in mind (e.g. object detection and tracking) and there is a need for diversity among datasets in order to cover different scenarios and types of data to reduce the possibility of overfitting. For end-to-end systems, datasets must include sensory data such as vehicle control information. Lastly, some datasets are naturalistic which is to say they focus on human behaviour as opposed to hardware or the surrounding environment. In the context of driving, naturalistic data would include an individual's features while driving such as the drivers' gaze and attention [1]. Table 1 includes an overview of the features of the datasets reported in this section.

2.1.1 Non-Naturalistic

The PASCAL VOC dataset [5] was intended solely for object detection and classification and is comprised of images which include motorbikes, bicycles, cars and people. Although this was one of the first datasets relevant for AV systems, images tend to depict scenes that are not typically experienced in driving scenarios. Similarly, the KITTI Vision Benchmark [6] is an image database

Table 1: Comparison of Datasets

<i>Dataset</i>	<i>Images</i>	<i>LIDAR</i>	<i>Naturalistic</i>	<i>Weather</i>	<i>Time</i>
PASCAL VOC	✓				
KITTI Vision Benchmark	✓	✓			
BDD100K	✓	✓		✓	✓
Oxford RobotCar	✓	✓		✓	✓
Comma.ai	✓			✓	✓
LiVi-Set	✓	✓			
SHRP2	✓		✓		✓
Shanghai NDS	✓		✓	✓	✓
Changchun NDS	✓		✓	✓	

containing labeled driving scenes and is one of the most widely used datasets for automated driving [1]. It was compiled by equipping a vehicle with color and grayscale video cameras, a laser scanner and a GPS localization system and driving around a mid-size city in rural areas and on highways.

BDD100K [7] was collected by crowd-sourcing videos from tens of thousands of riders. The dataset includes driving videos from four separate regions in the United States, and scenes vary in street types, weather conditions, and time of day, but did not include information about the drivers' themselves. The Oxford RobotCar Dataset [8] was collected by manually driving more than 1000km over a 16-month time period. The same route was traversed over 100 times during the period using 6 cameras, lidar, GPS and INS, capturing various environmental conditions, types of road users, and light settings. While conditions are labelled (e.g. sun, overcast, night, detour, poor

GPS) for simple collection of a particular condition, videos are not annotated and driving measurements are not included.

Comma.ai released a dataset [9] of videos totalling 7.25 hours of driving data along with several sensor measurements such as car speed, steering angle, and gyroscope, but is mainly restricted to highway driving while the LiVI-Set [10] includes point clouds and drivers' behaviours in various driving environments.

2.1.2 Naturalistic

Naturalistic datasets have a focus on the context in which a vehicle is being driven in, and the driver itself. The Strategic Highway Research Program (SHRP2) conducted a Naturalistic Driving Study (NDS) [11] collecting over 3000 volunteer participants' driving data during a 3-year period. It collected vehicle and sensor data and relative distances and speeds to surrounding vehicles while cameras captured the driver's face and hands to monitor facial expressions, hand maneuvers.

The Shanghai NDS (SH-NDS) [12] used the same SHRP2 NextGen data acquisition systems [11] to collect real-world driving data over a three-year period. Sixty driving participants were randomly sampled with a distribution of gender, age, and driving experience that represents the general Chinese driving population.

Zhu et al. collected real-world driving data of 84 drivers of diverse sociodemographic backgrounds and driving experiences in order to compare driving styles in an urban environment along a route in Changchun, China [13].

While most datasets for use in AV development are image and video based and include the vehicle’s surroundings and actions (e.g. throttle, steering), many fail to account for the social factors that influence driving behaviour such as age, gender, and experience.

2.2 Simulation

In both research and industry, simulation is used extensively for providing quick and efficient insight regarding concepts, strategies, and algorithms along with their shortcomings before real life testing. Both the safety concerns and the cost of training and testing AV technology are circumvented through the use of simulation. Typically, these platforms have the ability to realistically simulate vehicles’ instrumentation and are comprised of a software stack capable of translation to a physical AV system. Table 2 includes an overview of the features included in each of the platforms reported on in this section.

Table 2: Comparison of Simulation Platforms

<i>Platform</i>	<i>Sensors</i>	<i>Pedestrians</i>	<i>Weather</i>	<i>Time</i>	<i>Urban</i>	<i>Highway</i>
CARLA [14]	✓	✓	✓	✓	✓	
AutonoVi-Sim [15]	✓	✓	✓	✓	✓	✓
AirSim [16]	✓		✓	✓	✓	✓
TORCS [17]	✓		✓	✓		✓

TORCS [17] is a popularly used racing car simulator intended for the development of artificial intelligence (AI) agents for trajectory planning, control, and vision problems. The simulator features a simple damage model, collisions, tire and wheel properties, and aerodynamics and has been used in over 300 research papers as a benchmark for AI algorithms, however, it is not suitable

to train AVs for urban environments or situations which involve vulnerable road users such as pedestrians.

AirSim [16] is a high-fidelity physical simulation platform for data generation and training algorithms. It provides realistic environments, vehicle dynamics, multi-modal sensing and the capability to develop AV systems using deep learning, imitation learning, and reinforcement learning. While AirSim provides urban road environments, its focus is on realistic sensor simulation and does not include road users such as pedestrians or cyclists.

AutonoVi-Sim [15] is a high-fidelity simulation platform suited for the construction of complex traffic scenarios and includes other road-users such as cyclists and pedestrians. This inclusion allows for the exploration of situations that involve vulnerable road-users. The researcher can pre-program pedestrians' destinations within the road network and they will follow safe traffic rules to navigate to their goal. In order to simulate a jay-walking scenario, pedestrians can be setup to walk into the street in front the agent vehicle. However, based on video provided by AutonoVi-Sim, pedestrians walk across the road at a constant velocity and are unable to provide complex information regarding intention such as head pose (looking direction).

CARLA [14] is an open simulator for urban driving and supports training, prototyping, and validation of autonomous driving models and provides flexible specifications of sensor suits and environmental conditions. Like AutonoVi-Sim, CARLA also provides an implementation of pedestrians. They navigate the streets using a navigation map that conveys a location-based cost. The cost encourages pedestrians to use sidewalks and road crossings but allows them to cross a road at any point. While this approach is stochastic, it emphasizes only the time and starting point

of a crossing occurrence and not equally important aspects of pedestrian behaviour such as their head pose or varying velocities during the cross.

2.3 AV Training Techniques

Reinforcement learning (RL) is widely used to train an agent to interact with its surrounding environment. The agent learns by discovering which state-actions yield the most reward through trial-and-error. Deep reinforcement learning (DRL) has recently found success in bigger problems with more state-action pairs by using deep neural networks to approximate the state-action values. In both of these methods, a reward function directly determines the characteristics of the policy that the agent will adopt. Alternatively, inverse reinforcement learning (IRL) can be used for tasks in which the reward function cannot be predetermined or is not known. The goal in this approach is to find the reward function from sample data or expert demonstrations which best explain the behaviour presented.

Zhu et al. explored a reinforcement learning based throttle and brake control approach for a form of the autonomous vehicle following problem [18]. Using Neural Dynamic Programming as a model-free RL method, a control value corresponding to throttle or brake intensity is derived from a Markov Decision Process (MDP) framework. Using CarSim, a well-known vehicle dynamic simulator, it was found that while the agent vehicle could derive near optimal control commands for the autonomous following problem, slight vibrating exists in the control commands during the follow distance regulation process.

Vehicle-following was further explored in [19], using deep reinforcement learning and the SH-NDS data to present a framework for human-like autonomous car-following planning. The scenario the agent is learned in is described as a simple car-following environment, with a lead

vehicle controlled by external input while the following vehicle's speed, spacing, and velocity differences were initialized using SH-NDS data. Results include a model that was able to reproduce human-like car-following behaviour with higher accuracy than recent data-driven car-following models such as the Intelligent Driver Model, a variant of widely used car following models [20].

Gao et al. uses IRL to develop a car-following decision-making system for complex traffic conditions, trained by simulation in a highway environment [21]. The data used to find the reward function are obtained by recording the driving data of two experienced drivers in a driving simulator. The data pertaining to each driver are shown to have unique characteristics and the algorithm was verified to have captured these unique characters in a reward function for each driver.

Chowdhuri et al. presented a multi-modal multi-task learning approach to deep neural networks [22]. Multi-model learning combines different types of inputs while multi-task learning emphasizes training on side tasks. They introduced three distinct behaviour modes based on their own dataset which are used as a second type of input to the network, allowing for separate driving behaviours to form within a single multi-modal network. These modes are: a) Direct Mode, in which the car drives with few obstructions b) Follow Mode, in which the car follows a lead car in front of it, and c) Furtive Mode, in which the car drives slowly due to perceived boundaries such as bushes alongside the path. This approach inserts behavioural information directory into the network's process stream, allowing for parameter sharing between related modes.

Chae et al. proposed an autonomous braking system using DRL [23]. In the study, the agent's sensors receive information about the pedestrian's position and adapts the brake control to the state

change in a way that minimizes the chance of a collision. The system is based on a Deep Q-network (DQN). DQN is adapted from the RL method Q-learning [24] which iteratively searches for an optimal policy by using a Q-value function to calculate the expected sum of rewards for a state-action pair. Q-learning falls short in this case due to the continuous state space, which makes it impossible to find an optimal value of a state-action pair. Instead, DQN approximates the state-action value function using the deep neural network (DNN). This paper formulates the braking scenario as a discrete MDP model and defines a discrete action space using a set of four brake control intensities α_0 , α_{high} , α_{mid} , and α_{low} .

Chen et al. looked at pedestrian-rich environments, keeping in consideration the diverse human behaviours found in such environments [25]. Using DRL, a time-efficient navigation policy for multiagent collision avoidance is learned. In shaping the reward function, social norms were included in such a way that the agent would receive a negative reward for breaking a norm, what it should not do; rather than rewarded for successful navigation, or what it should do. This method allows the agent to learn a policy that respect social norms not explicitly found in the features of pedestrian movement.

2.4 Key Remarks

Most datasets used for the purpose of training AV systems have a focus on developing specific components of the system and are labelled for those tasks. However, the largest and most popular datasets are not naturalistic, they do not contain enough information to capture the social aspects of mobility scenarios. For an AV system to fully understand the intentions of other traffic participants, they must be able to consider social behaviour [26]. For example, there are numerous factors that affect pedestrian's crossing behaviour, including the presence of a driver in the vehicle

approaching [27]. Apart from on-board vehicle mechanisms (e.g. horn, signals, lights), communication between road users is often non-verbal and include facial expressions, eye contact, and gestures. As the utility of these types of communication become limited with the absence of a driver, there must be an emphasis on the other types of communication that occur between road users. As stated by Farber, these include anticipatory behaviour; predicting intention and movement based on the direction of a user's behaviour, as well as scheme formation; anticipating behaviour based on the users' characteristics (e.g. a child with a ball versus an elderly person with a walker) [28].

Similarly, most simulation platforms tend to focus on sensor suite customization and data generation, disregarding the intrinsic social and communicative dimensions of traffic situations. These platforms are only as good as the underlying models used to describe traffic participants [2] and this problem is highlighted with the absence of pedestrians in most platforms and unrealistic implementations where they are included.

Lastly, the RL paradigm and its variations are commonly used methods for AV behaviour development. In these cases, the type of data trained against and the formulation of the scenario and tasks are imperative. Some literature discretizes the action space despite vehicle control being best described as continuous, while others pay no attention to the passenger's experience [21]–[23].

The novel platform developed in this thesis addresses these issues. By compiling naturalistic datasets using realistic and immersive VR, the data can be directly used in simulation of complex traffic scenarios. These data capture participants' habits, instincts, and sociodemographic information and is used in conjunction with DRL to learn a policy for brake control expressed in

a continuous action space, avoiding discretization, and a multi-objective reward function formulated by considering two conflicting objectives: the pedestrian's perspective of the vehicle and the passenger's experience in the vehicle.

II. Methodology

In this thesis, a platform for developing AV systems using naturalistic data and reinforcement learning is presented. The platform is comprised of two main modules. Figure 2a. portrays the components used for data collection and is described in more detail in Chapter 3, while Figure 2b. illustrates the AV training module (Chapter 4) which uses the naturalistic data directly in a RL-based training environment.

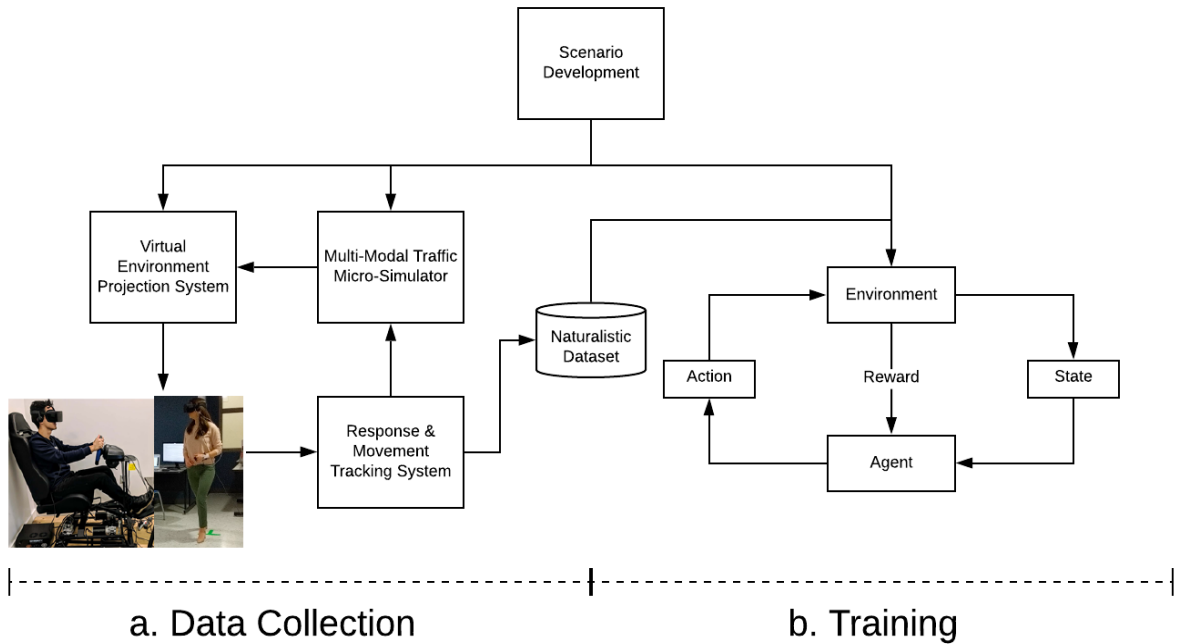


Figure 2: Data Collection and Training Platform

The case study (Chapter 5) demonstrates this platform. Naturalistic data is collected and used directly toward the learning of a multi-objective autonomous braking system for urban roads. Pedestrian motion data is compiled with the Virtual Immersive Reality Environment [29] and is

used to simulate a pedestrian walking across a road in front of an approaching vehicle agent. The simulated pedestrian can cross at any given time depending on the trial data used. By placing the vehicle agent in this uncertain situation, it learns to smoothly control its velocity while reacting safely as necessary in circumstances where danger is imminent. With a concisely defined reward function, a policy that maximizes passenger comfort while minimizing the chance of an accident occurring is learned. This policy for brake control is learned using two reinforcement learning methods i.e. Proximal Policy Optimization and Deep Deterministic Policy Gradient which are defined in the following sections.

3 Virtual Reality System

3.1 Virtual Reality as a Data Collection Tool

In transportation, Stated Preference (SP) experiments are considered to be a valuable tool used to evaluate behavioural responses towards forthcoming services and technologies. They present the respondents with framed, hypothetical scenarios in order to accurately assess their choice preferences [30]. However, there are situations in which SP surveys fail to realistically represent the exact choice scenario [31]. This criticism is prevalent in forthcoming AV related scenarios where the respondent has never had the opportunity to experience or interact with an AV, and therefore must base their preference on their own perception of the technology. In SP experiments, the use of visualization and simulation has shown to improve realism [31], allowing an opportunity to utilize VR technology for experiments regarding emerging technologies such as AVs.

Major benefits of using virtual environments include the abilities to curate detailed and complex situations and analyze relevant data. Respondents are able to act naturally as well as make irreversible decisions that are difficult or expensive to recreate in the real world. An Immersive Head Mounted VR environment (IHMVR) is a VR display device that uses an optical system to directly present virtual scenes received by the display and works with the human brain to produce a strong sense of immersion [32]. Virtual environments which allow for movement may preserve the perception-action coupling that is critical in examining many visual timing skills [33].

In this context, VR has been used several times before. Sobhani et al. studied the effects of multi-tasking (walking and using smartphone) on adults' street crossing performance by using a simulated intersection crossing task programmed in an IHMVR with an integrated audio and sensors which allowed close emulation of real word behaviour of pedestrians [34]. Neider et al.

used a high-fidelity street crossing simulator and showed that naturalistic cell phone conversations impair crossing performance and increase crash rates [35]. Similarly, Stavrinou et al. demonstrated significant costs to simulated crossing performance while conversing on a cell phone [36], [37].

3.2 Virtual Immersive Reality Environment (VIRE)

Coupling the technologically advanced Head Mounted Display (HMD) with VR, the Laboratory of Innovations in Transportation (LITrans) introduced the VIRE platform [29] for a range of transportation experiments that are highly realistic, immersive, interactive and have the capability to collect additional information about the motion, orientation, and actions of the respondents. VIRE not only addresses the realism issue in SP experiments, but also enables researchers to design complex experiments while collecting naturalistic data. Through the use of VIRE, respondents can experience transportation technologies and scenarios realistically. By projecting the traffic simulation directly to their eyes, participants are able to experience and interact with fixed simulation objects (e.g. AVs, cyclists, etc.) in real-time. VIRE is also capable of collecting data relevant to study such as the participant's gaze, the position of the participant and other simulated objects within the virtual environment, the participant's and simulated objects' speeds to name a few. In this thesis, the following hardware was used to run VIRE:

- Oculus Rift with motion and touch sensors
- Intel Core i7 8th Generation Processor
- NVIDIA GeForce 1080 Graphic Card

Ultimately, VIRE is a mix of software and hardware systems working together. The data collection module outlined in Figure 1a. is comprised of four main components.

3.2.1 Scenario Development (SD) System

The theoretical experiment design is translated and developed into a 3D scenario. VIRE based in Unity, an open-source physics and gaming engine [38]. The scenario is created using 3D models of the buildings, vehicles, persons, cyclists, trees, roads, and other components necessary. Programmatically, roads are coded as a collection of links and nodes and vehicles, persons, cyclists are mapped to travel among these. This system is capable of importing already available 3D resources and incorporating them into the virtual environment.

3.2.2 Multi-Modal Traffic Micro-Simulation (MMTM) System

VIRE currently supports the large-scale simulation of vehicles, person, and cyclists. Vehicular movement is simulated using the Intelligent Driver Model [20]. This model realistically simulates typical behaviour of AVs, avoiding collisions and maintaining a safe gap and stopping distance from other road users. For simulated pedestrians within an experiment, VIRE adapts the social force model [39] to incorporate the multi-modal nature of the simulation. Implemented models and their parameters within VIRE are easy to change to specific requirements of the experiments. Objects are programmed to move within Unity's provided waypoint system. By defining starting, in-between, and destination coordinates (e.g. the coordinates of two different road nodes as well as the links in between), the object will smoothly move toward each point in the manner that the underlying implemented model describes. Implemented models and their parameters are easy to change to specific requirements of the experiments via user interface.

3.2.3 Virtual Environment Projection (VEP) System

Unity supports VR and animation along with commercially available HMDs. Currently, the Oculus Rift is used for projection but can support other HMDs such as the HTC Vive and Microsoft

HoloLens. The projection system takes the environment developed by the SD and incorporates dynamically moving objects, whose positions are provided by the MMTM system. VIRE is also capable of generating and associating audios to 3D objects, for example a car passing by can make the appropriate noise for the respondent to hear through the HMD.

3.2.4 Driving in VIRE

VIRE has been used in several contexts [29], [40]–[43] thus far. Recently, driving hardware such as a steering wheel, acceleration and braking pedals, and a motion simulator rig was incorporated for vehicle-based experiments and generalizability. With the inclusion of this hardware, participants can realistically drive in virtual environments under various environment and traffic conditions. Alternatively, respondents can experience riding in an AV as a passenger, with the option of multi-tasking. The controllers included in the Oculus system can be mapped to any virtual object of choice and used to make the respondent interact with the item. Examples include allowing the rider to read a virtual newspaper or play a game on a virtual smartphone. Past experiments have included the collection of physical data such as participants' gaze and driving behaviours as well as virtual data such as the participants' coordinates, collisions and near-collisions, routes taken, and total trip times. Figure 3 depicts both a respondent sitting in the motion simulator rig and the virtual environment the respondent is experiencing via the HMD. The case study contained in this thesis does not make use of VIRE's driving capabilities or driving data. The hardware coupled with VIRE for driving includes the following:

- DOF Reality H3 Consumer Motion Simulator Platform
- Thrustmaster T150 Force Feedback Racing Wheel



Figure 3: Driving in VIRE

4 Deep Reinforcement Learning System

The training module of the platform (Figure 1b.) is formalized as a basic RL problem. RL is a process where a learning agent interacts with its environment to achieve a goal. It does so continually and receives a reward for the actions it takes, which the agent tries to maximize over time [44]. More specifically, at each time step t , the learning agent observes the environment state s_t ($s_t \in S$), where S is the set of all possible environment states. An action a_t is selected from the set of available actions the agent can take. The environment then moves to state s_{t+1} resulting from the implemented action and a reward $r_{t+1} \in R$ associated with the transition (s_t, a_t, s_{t+1}) is communicated to the learning agent. Upon receiving the immediate reward at each state transition, the agent forms a policy that maps the current state to an optimal action for future use. The agent attempts to maximize the expected discounted return defined as $R_t = \sum_{t=1}^{\infty} \gamma^{t-1} * r_t$ where γ is a discount factor in $(0, 1)$ to determine the importance of future rewards. Deep RL (DRL) has found success in bigger problems with more state-action pairs by using deep neural networks to approximate the state-action values.

In the case study’s multi-objective autonomous braking system (Chapter 5), the vehicle drives along a straight road using the learned policy towards the pedestrian. At the beginning of each training episode, the vehicle is driving at an initial velocity v_{veh} from an initial position $(vehpos_x, vehpos_y)$ and the pedestrian is located at an initial position $(pedpos_x, pedpos_y)$. A “safe box” surrounds the pedestrian’s crossing location. If the vehicle is within this boundary while the pedestrian is crossing, it is considered an accident. There are four events that can end an episode:

- *Accident*: the vehicle is within the safe box at the time of pedestrian crossing the road
- *Pass*: the vehicle passes the pedestrian before the pedestrian has begun to cross
- *Cross*: the pedestrian finishes crossing the road
- *Stop*: the vehicle comes to a complete stop

4.1 Policy Proximal Optimization (PPO)

PPO [45] is an on-policy algorithm used for environments with either discrete or continuous action spaces. It is a simplification of Trust Region Policy Optimization (TRPO), proposing to use a clipped surrogate loss function to limit the likelihood ratio of old and updated policies $r_t(\theta) = \pi_\theta(a_t|s_t)\pi_{\theta_{old}}(a_t|s_t)$ where π_θ is a stochastic policy. In TRPO, the surrogate function is maximized subject to a constraint on the size of the policy update. To avoid excessively large policy updates, PPO introduces a penalty to changes to the policy that move the probability ratio away from 1. To include this penalty, the following objective was developed:

$$L^{CLIP}(\theta) = \hat{\mathbb{E}}_t \left[\min(r_t(\theta) \hat{A}_t, \text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon) \hat{A}_t) \right]$$

where epsilon is a small hyperparameter. The second term $\text{clip}(r_t(\theta), 1 - \epsilon, 1 + \epsilon)$ modifies the surrogate objective in order to remove the incentive for moving r_t outside of the interval $[1 - \epsilon, 1 + \epsilon]$. Then, taking the minimum of the clipped and unclipped objective so that the final objective is a lower bound on the unclipped objective. This results in the inclusion of changes in probability ratios only when it makes the objective worse:

$$\max_{\theta} \hat{\mathbb{E}}_t \left[\frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)} \hat{A}_t \right]$$

4.2 Deep Deterministic Policy Gradient (DDPG)

DDPG was developed by combining the actor-critic approach found in DPG and insights learned from DQN’s recent success [46]. It is a model-free, off-policy actor-critic algorithm that was proven to learn competitive policies for tasks using low-dimension observations such as cartesian coordinates. It follows a standard reinforcement learning setup as previously described. Unlike PPO, DDPG is inspired by Q-learning, using the greedy policy $\mu(s) = \arg \max_a Q(s, a)$ to find the optimal action a for a given state s . Instead of directly implementing Q learning, a modified target network which creates a copy of the actor and critic networks $Q'(s, a|\theta^{Q'})$ and $\mu(s|\theta^{\mu'})$ for calculating the target values. The weights of these target networks are then updated, slowly tracking the learned networks which constrains them to change slowly and improves the stability of learning: $\theta' \leftarrow \tau\theta + (1 - \tau)\theta'$ with $\tau \ll 1$. It also implements a finite sized replay buffer \mathcal{R} which stores transitions sampled from the environment according to the exploration policy in a tuple (s_t, a_t, r_t, s_{t+1}) . This buffer allows the algorithm to benefit from learning across a set of uncorrelated transitions by sampling a minibatch uniformly at each timestep to update the actor and critic.

4.3 Reward Function

As the most important component of our system, the reward function determines the behaviour of the brake control. We must formulate it such that it accurately describes the characteristics the system’s behaviour should ultimately exhibit. Given the multi-objective nature of our system, our task in defining the reward function becomes more difficult; we are attempting to optimize two possibly conflicting factors, avoiding dangerous situations with pedestrians and providing the

passenger with a comfortable experience. To this end, we propose a reward function which captures both safe and comfortable braking behaviour:

$$r_t = -(\eta v_{veh_t}) \mathbf{1}(S_t = accident)$$

$$-(\beta v_{veh_t})$$

$$-(\mu a_t) \left| \frac{d^2 v_{veh}}{dt^2} \right|$$

$$\eta, \beta, \mu > 0$$

where at time step t , v_{veh_t} is the velocity of the vehicle and a_t is the value of the brake action chosen and $\mathbf{1}(x = y)$ has a value of 1 if the statement inside is true and 0 otherwise. The first term is included from the reward function proposed by Chae et al., describing the penalty given to the agent should in the case of an accident [25]. The penalty is proportional to the velocity of the vehicle, reflecting the severity of the accident and therefore encouraging the vehicle to slow down in the case that an accident is unavoidable. The second and third terms are our own contributions. The second encourages the vehicle to use caution and prevents the agent from abusing the stochastic nature of our dataset and unintentionally learning to take the risk of driving passed the pedestrian before they begin to cross the road. The final term is a function of the jerk of the vehicle at time step t . Mathematically, jerk is the second derivative of velocity with respect to time. Physically, it is felt as the increasing or decreasing force on the body [47]. The absence of jerk has been found to be an important characteristic of comfortable driving [48], [49] and thus, our reward function aims to minimize this force, penalizing the vehicle as a function of the jerk and the

intensity of the brake action. The constants η , β , and μ are weight parameters to control the trade-off between each objective.

III. Experiment and Results

5 Case Study

As shown previously, one of the focuses of some AV simulation platforms is the inclusion of pedestrians in order to train agents to react to life-endangering situations that could occur in reality. However, pedestrian road-crossing behaviour is complex and can vary highly in aspects such as waiting time, pose, and path. These intricate behaviours can be difficult to simulate yet carries with it one of the highest priorities when considering pedestrian safety and the inevitable dangerous situations that AVs find themselves in.

Another important aspect of transferable autonomous braking systems is having a realistic formulation of the problem during training. This includes accurately representing both the objectives and the mechanism used to reach those objectives. Intuitively, the braking control mechanism of a vehicle resides within a continuous (real-valued) action space. By avoiding the discretization of the action space similar to the method found in [25], we can consider more complex driving behaviours which capture reality more accurately, such as the consideration of passenger comfort.

While pedestrian safety is a top priority, passenger comfort is an important and often overlooked factor in the acceptance and successful implementation of AVs [50]. Naturalness and perceived safety of the AV's driving manoeuvres are two major influences on a rider's comfort. Thus, designing an objective function which accounts for this is crucial for training an acceptable AV braking system.

In this case study, naturalistic pedestrian road-crossing data collected by the VIRE platform is used to train a multi-objective continuous autonomous braking system in simulation.

5.1 Naturalistic Data Collection

VIRE was recently used to capture pedestrians' movements in a dense urban environment under varying traffic conditions as well as their general interaction and behaviour towards AVs [41]. A road-crossing environment was developed in which a respondent can stand at the side of a two-lane road with vehicles driving by from one or both directions and cross the virtual street when there is an opportune moment without interrupting traffic flow. In the environment, simulated vehicles are implemented using the Intelligent Driver Model [20] while simulated pedestrians use the social force model [39]. Various conditions in the environment could be changed easily by the researcher carrying out the experiment, choosing from various levels of parameters described in Table 3.

Table 3: Levels of Adjustable Parameters

Variable	Levels		
Speed Limit (km/h)	30	40	50
Lane width (m)	2.5	2.75	3
No. of braking levels	1	2	3
Minimum Gap (s)	1	1.5	2
Flow rate (veh/hr)	530	750	1100
Road type	1-way	2-way	2-way with median
Traffic Automation Status	Human Driven	Mixed Traffic	Fully Automated
Time of day	Day		Night
Weather	Clear		Snowy

Running was prohibited because the respondent may hit some physical obstacle in the laboratory or trip themselves with the wires. Figure 4 shows a respondent participating in the experiment, while Figure 5 shows the virtual environment the participant is seeing and interacting with through the HMD and sensors.



Figure 4: VIRE Pedestrian Physical Environment

From a total of 160 respondents participated in the experiment with a varied demographic which included young teenagers, university students and professionals. The data collected by VIRE for each trial during these experiments include coordinates corresponding to both the position of the respondent and the direction in which they were facing recorded at a time interval of approximately 0.1 seconds. With a total of 2463 individual road-crossing trials, the dataset was found to vary in

waiting times, velocities, and completion times. Figure 6 presents the frequency distribution of waiting times in the dataset. Since each respondent was asked to complete multiple trials, their habits and instincts are captured in the data, giving more naturalism and value to the dataset. We are able to use the coordinate data to reproduce the movement and pose of the respondent in each trial in simulation directly using Unity.



Figure 5: VIRE Pedestrian-AV Virtual Environment

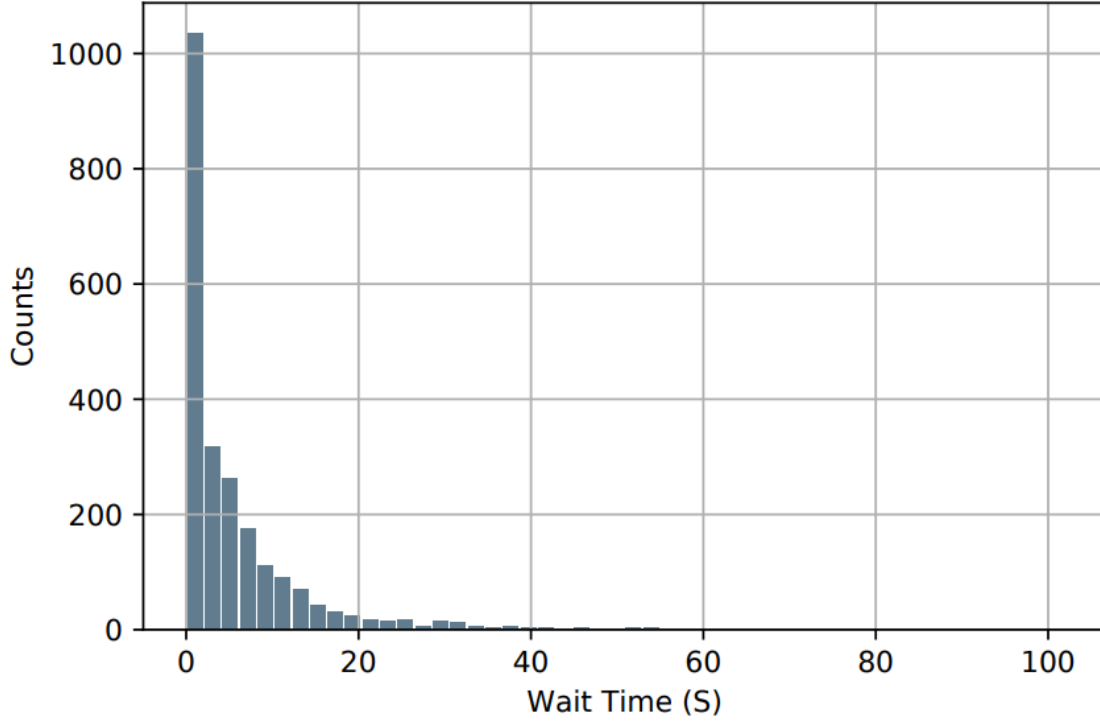


Figure 6: Waiting Time Frequency Distribution

5.2 Training Scenario

One of the biggest challenges for autonomous vehicles is safely handling the risks involved with pedestrians who are considering crossing the road. Acting too cautiously with abrupt and jerky decelerating results in discomfort for the passenger and disruption of traffic flow. Acting aggressively and not decelerating quickly enough can result in an accident. The optimal behaviour results from the ability to recognize the intention of the pedestrian while maintaining safe and comfortable control of the vehicle's speed. By detecting a pedestrian's initial movements (e.g. edging towards the sidewalk curb, looking left and right) and considering their relative position, the vehicle should be able to decelerate safely and smoothly. In our autonomous braking system, we assume perfect sensor measurements and pedestrian detection. However, stochasticity in the

measurements can be included in the future improvements of this work. The vehicle moves at speed v_{veh} from a position $(vehpos_x, vehpos_y)$ and is provided the relative position to the pedestrian $(pedpos_x - vehpos_x, pedpos_y - vehpos_y)$, the looking direction of the pedestrian $(head_x, head_y, head_z)$, and the velocity of the pedestrian v_{ped} , at each time step. Using these observations along with its own velocity, the vehicle decides whether to apply brake, a continuous value between $[0, 1]$. Figure 7

provides a depiction of the proposed autonomous braking system and scenario.

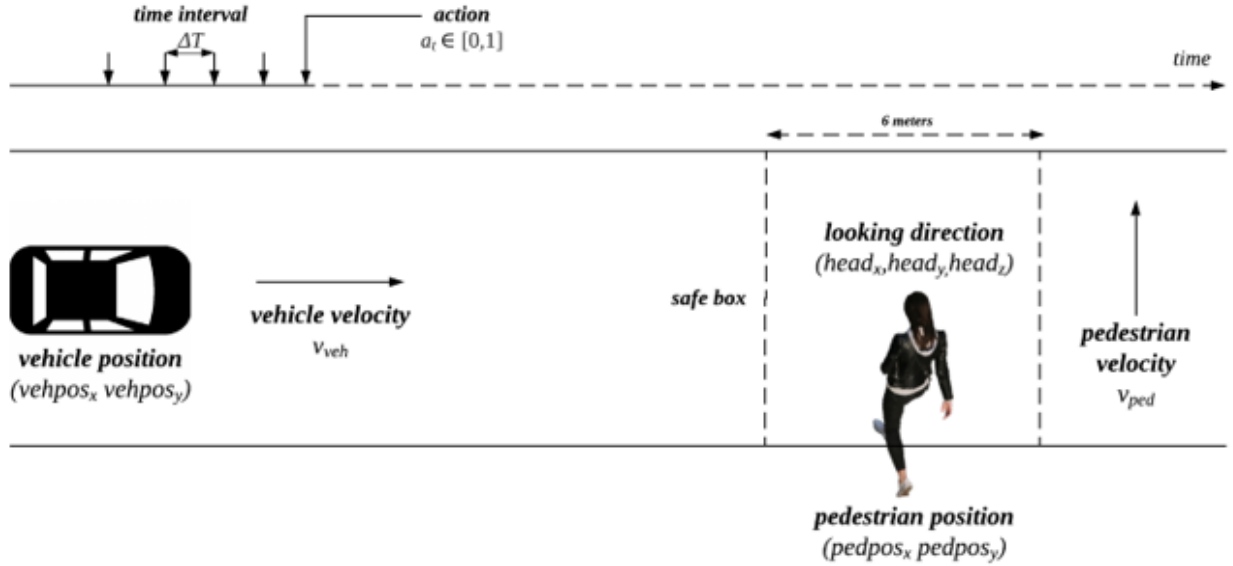


Figure 7: Training Scenario

5.3 Training Simulation Setup

For training, we used the game engine platform Unity along with its open-source plugin ML-Agents Toolkit [38] to train the PPO agent. We designed an environment in which our pedestrian dataset that was also collected using Unity, is easily transferable. As described in Section 3.2.2, the coordinates for each crossing trial in the dataset can be used to move a simulated pedestrian

within a virtual environment. In this case, a simulated pedestrian is programmed to move across a virtual road according to the coordinates fed to the MMTM from the dataset at each timestep. Simultaneously, those coordinates are also passed as environmental observations to the learning agent as per the ML-Agents plugin. In each episode, the simulated pedestrian is situated at the same initial position, but the pedestrian crossing trial used is chosen at random from a subset of the dataset each episode. The rest of the dataset is used to evaluate the model. Trials vary in waiting times, crossing velocities and head movements, helping to train our system for generality. The vehicle is initially located 160 meters away from the pedestrian and has an initial velocity of $v_{init} = 11.11\text{m/s}$ (40 km/h). The pedestrian’s safe box size is 3 meters wide, covering any situation in which the vehicle is within 3 meters of the virtual pedestrian.

5.4 Training Parameters

The neural network implementation for PPO consists of a fully-connected feed-forward network with three hidden layers and the hyperparameter configuration shown in Table 4.

Our implementation of DDPG includes two hidden, fully connected layers with 256 and 128 nodes for the actor and critic networks respectively. Its configuration is shown in Table 5.

A total of three models were trained: two of the models include the passenger comfort objective, one trained using PPO and the other with DDPG to compare the efficiency of their learning. A third model was trained with PPO but did not include the passenger comfort objective in the reward function in order to contrast their action trajectories. These models are outlined in the Table 6.

Table 4: PPO Configuration

Batch size	64
Buffer size	10240
Hidden Units	256
Time Horizon	1024
Learning Rate	10^{-3}
Gamma	0.99
η	0.1
β	0.01
μ	0.01

Table 5: DDPG Configuration

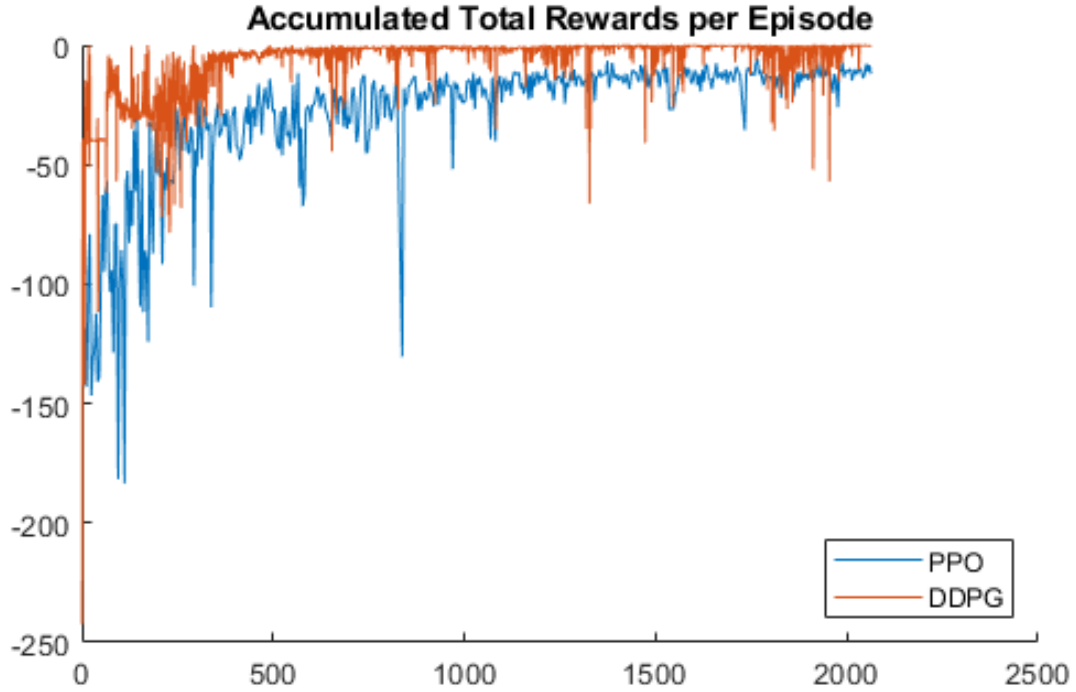
Replay Buffer Size	10240
Minibatch Size	128
Gamma	0.99
Tau	10^{-3}
Actor Learning Rate	10^{-4}
Critic Learning Rate	10^{-4}
η	0.1
β	0.01
μ	0.01

Table 6: Reward Function Configurations

<i>Algorithm</i>	<i>Passenger Comfort</i>
PPO_1	✓
PPO_2	✗
$DDPG$	✓

5.5 Results

Figure 8 shows a plot of the total accumulated rewards spanning every episode during training using PPO_1 and DDPG. It can be observed that DDPG not only converges much earlier than PPO_1 , but also consistently attains high total reward thereafter.

**Figure 8: Cumulative Environment Reward**

With regards to braking behaviour, each model produced reported an absence of any accidents with the pedestrian. This was to be expected as the vehicle was given ample time to come to a

complete stop in all cases. However, it is more important to note the braking profiles produced by the agents trained with PPO₁ and PPO₂. Figure 9 shows a detailed trajectory of brake actions for each agent for one example case in which the pedestrian completes a crossing after approximately 12 seconds. Figure 6 (a) shows the velocity of the vehicle over time, as it approaches the pedestrian. The jerks produced by both vehicles is presented in (b) while (c) shows the brake action values applied. Both agents begin their deceleration at the same time but execute actuation differently.

It is clear that the agent trained using PPO₂ decelerated steeply, using higher brake action values to do so and in doing so, produced relatively severe jerk. Alternatively, the agent trained using PPO₁ has a curved decent in velocity as a result of its use of weaker brake action values applied at the beginning part of its deceleration and strengthening as the vehicle nears the pedestrian. Most importantly, the mean jerk force produced is halved, implying a much more comfortable experience for those in the vehicle.

This example shows that while both agents exhibit safe braking operation by avoiding a collision, it is clear which one provides a better travelling experience. The inclusion of vehicle jerk in the reward function results in more comfortable ride for passengers while maintaining reasonable safe braking operation.

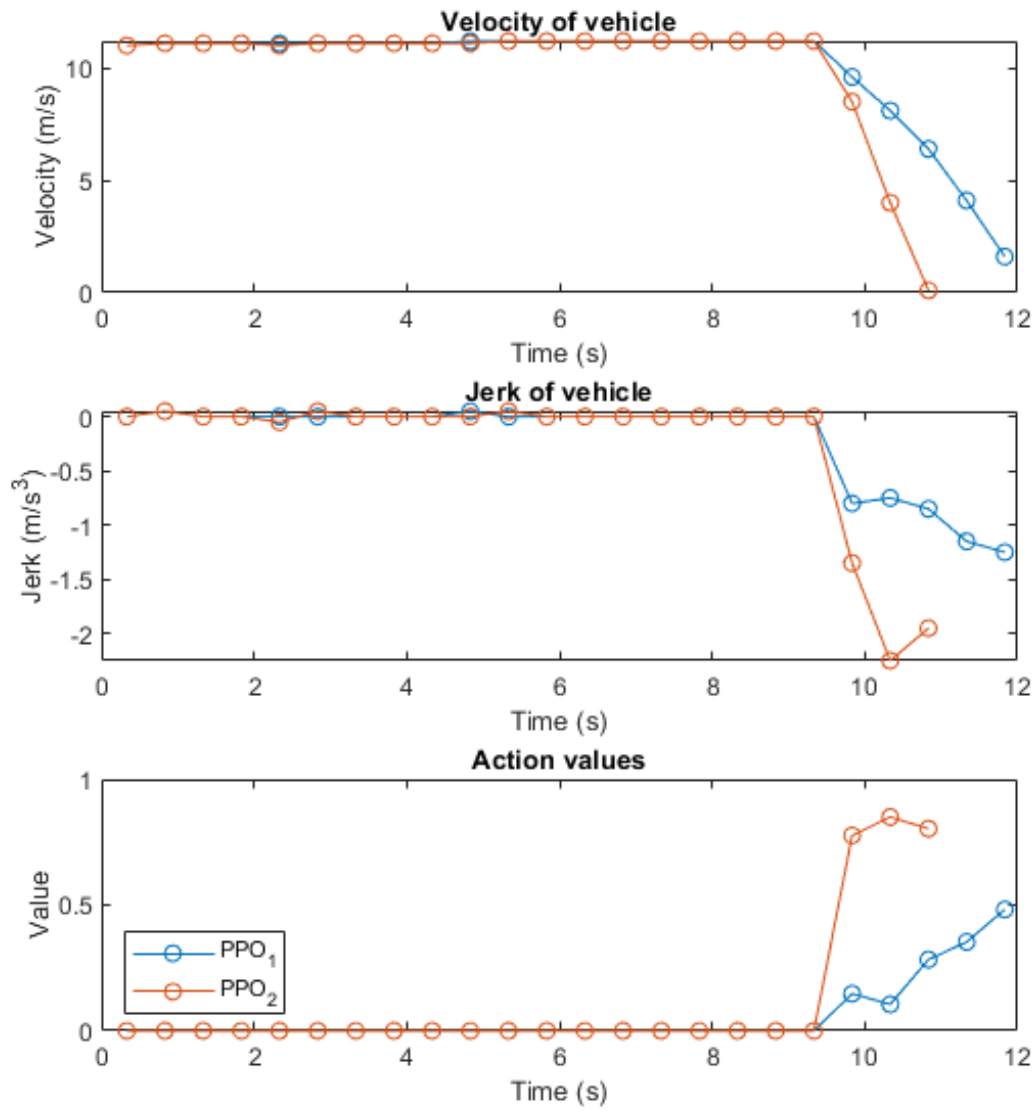


Figure 9: Profile of velocity, jerk, and actions in an example episode against the same road-crossing trial

6 Conclusion, Limitations and Future Work

6.1 Conclusion

The largest and most popular datasets are not naturalistic and do not contain enough information to capture the social aspects of mobility scenarios while AV simulation platforms disregard the intrinsic social and communicative dimensions of traffic situations. For an AV system to fully understand the intentions of other traffic participants, they must be able to consider social behaviour [26]. Addressing these issues, this thesis has focused on presenting a novel method for training AVs in simulation using naturalistic data. This work presents the components of a novel VR platform for the development and simulation of traffic scenarios with the capability of collecting naturalistic data and use that data directly in an RL-based training environment to train AVs. This platform was demonstrated in an end-to-end case study in which pedestrian crossing data is collected and used to train a multi-objective autonomous braking system while establishing the significance of an objective which maximizes vehicle passenger comfort. It was shown that after training, the system was able to reduce the negative influence on passenger comfort (i.e. jerk force) by half while maintaining a safe braking policy.

6.2 Limitations

In this study, the Oculus Rift is used. It was determined the Oculus Rift S' tracking system have a translational accuracy of 1.66 ± 0.74 mm for the HMD and 4.36 ± 2.91 mm for the controllers, while the rotational accuracy of the system was $0.34 \pm 0.38^\circ$ for the HMD and $1.13 \pm 1.23^\circ$ for the controller. While there is minimal error in tracking a respondent's movement in a virtual environment, this should be accounted for when training AV systems using the presented method [51].

6.3 Future Work

There are several directions that can be taken with this work. Firstly, the presented case study can be built upon to configure more complex pedestrian situations such as incorporating multiple pedestrians along the road. Coupled with the introduction of a throttle action for the vehicle, passenger comfort and pedestrian safety can be further evaluated. The physics-based Unity environment also allows for the introduction of different weather conditions and their adverse effects on the vehicle such as loss of traction. Additionally, data collected with VIRE can be used to train specific mechanisms found in AVs. For example, the pedestrian-crossing data presented in the case study can be used toward modeling pedestrian path and pose as well as developing and evaluating pedestrian intention and trajectory prediction techniques. Regarding driving-centered studies, VIRE can be employed to collect information about driving styles, route choices, reaction times, and the effects of distractions on drivers. The road-crossing environment used in this study can also be modified to include more variation and realism, thus inducing more naturalism in the respondent's behaviour. The inclusion of more vehicle-types, varying curb heights and bicycle lanes are just a few more features that can affect the way a pedestrian crosses the road in an urban environment.

Lastly, introducing methods for interpretable RL such as [52], [53] can generate human-readable and verifiable policies as well as discover multiple ways to perform the same task in complex environments. This could help to understand the correlation between inputs in socially-driven traffic situations such as pedestrian crossing, as well as ensure that policies learned for these safety-critical scenarios are consistent with desired properties.

Appendices

One of the major components of the novel platform presented in this thesis is VIRE, the system used for data collection. This system was built in such a way that it could be used as a generic tool for curating environments and scenarios in which a participant can immerse themselves and interact with naturally. While the case study presented above emphasizes the collection and use of motion and orientation data captured using an original VIRE environment, the additional subsequent publications showcase the collection and use of preference and behaviour data in choice preference and behaviour analysis studies.

A. Should I give or should I take? Choice issues in automated vehicle control

Shadi Djavadian

Rafael Vasquez

Bilal Farooq

Introduction

The advent of vehicle automation has provided new opportunities to mitigate traffic, increase productivity, and reduce emission. Despite these advantages, automated vehicle (AV) face some deployment and adoption challenges. An important factor that has major effect on the effectiveness of AVs is driver's willingness to give control to the system or take it back when required. Another associated challenge is driver's mental workload and situational awareness in the operational safety of automated driving. In the case of automation, drivers become passive observers as opposed to information processors resulting in decrease in situational awareness of drivers and increase in accident rate (Hirose et al, 2015). To gain further insights into behavioral responses of drivers to automation, in this study a stated preference laboratory experiment is designed employing Virtual Reality Immersive Environment (VIRE) driving simulator (Farooq et al., 2018).

The aim is to answer following research questions:

- Under what conditions drivers are more willing to give partial or full control to the automated vehicle?
- Under what conditions drivers have higher situational awareness and ability to safely take back control from automated vehicles?

Background

Online stated preference surveys (SP) have been widely used in the recent years to explore the factors influencing travelers' acceptance of automated vehicles (Becker & Axhausen, 2017).

However, online based SP surveys lack realism especially when it comes to new and upcoming technologies. Driving simulators and field experiments have also been used to investigate situational awareness and mental workload of drivers when it comes to safely operating automated vehicles (Stapel et al., 2017). However, they also have their own limitations e.g. lack of interaction with environment, controlled environment. As a result, virtual reality based stated preference surveys have been gaining interest among researchers to study behavioural responses to AVs, since it allows users to form a visual image of the innovative alternative as opposed just purely mental image (Djavadian et al., 2019).

Methodology

To answer the earlier posed research questions, we design and conduct laboratory experiments using our in-house VIRE driving simulator (Farooq et al., 2018). Fig. 1 presents examples of the VIRE setup. To increase users comfort level and bring additional realism to the experiments, a driving rig is used. The laboratory experiment is divided into three sections: a) pre-experiment questionnaire, b) actual experiment, b) post-experiment questionnaire. The pre-experiment questionnaire collects socio-demographic information of users, their prior driving experience and habits, and participants' prior knowledge and experience with regards to AV. During the actual experiment, drivers are asked to drive from their assigned origins to their assigned destinations under different conditions (e.g. congestion level, network familiarity, multi-tasking, incidents, etc.). This provides us the opportunity to investigate users' behavioral response, situational awareness and willingness to take/give control to AV under different conditions. The post-experiment questionnaire collects data on participants' experience during the experiment (mental work load, reason for selecting one option over another, motion sickness, etc.).

In this study, the VIRE is set up in such a way that it provides participants the option of fully giving control to AV en-route to their destination, allowing us to explore under what conditions users are more willing to give up control to AV. Similarly, while being driven to their destinations by AV users may be asked to take over control from AV in order to investigate their situational awareness. Different multi-tasking options are implemented in the simulation and are presented to the users (e.g. reading, playing games). This allows us to examine responses of users to each option and investigate the effect on their willingness to take/give control of AV along with their level of situational awareness. Galvanic skin response sensors are used to measure emotional arousal of the users. Virtual pedestrians are also introduced in the simulation to measure situational awareness of users and their mental workload.

Based on the data gathered from the experiments we conduct qualitative and quantitative analysis (e.g. logit choice model) of the results. It is expected that the results from this study will provide us with insights related to what extend and under what circumstance drivers are willing to give full control to the automated vehicles. In addition, project results will allow us to draw conclusions as to what group of population (age, gender, driving experience, income, etc.) will be more adaptable to automation. Furthermore, the results will provide us with understanding of drivers' mental workload and situational awareness when it comes to safely operation AVs.



(a) Take over control

(b) Multitasking while driven by AV

Fig. 1 Participants in VIRE driving simulator

References

- Becker, F. & Axhausen, K. W. (2017). Literature review on surveys investigating the acceptance of automated vehicles. *Transportation*, 44(6), pp 1293-1306.
- Djavadian, S., Farooq, B., Vasquez, R., & G. Yip (2019). Analysis of Behavioural Responses in Connected and Autonomous Vehicle Environment. *Accepted for publication in IATBR 2018 conference book, titled: Mapping the Travel Behavior Genome*.
- Farooq, B., Cherchi, E., & Sobhani, A. (2018). Virtual Immersive Reality for Stated Preference Travel Behavior Experiments: A Case Study of Autonomous Vehicles on Urban Roads. *Transportation Research Record*, 0361198118776810.
- Hirose, T., Kitabayashi, D., & Kubota, H. Driving Characteristics of Drivers in a State of Low Alertness when an Autonomous System Changes from Autonomous Driving to Manual Driving. *SAE Technical Paper*, 2015-01-1407.
- Stapel, J., Mullakkal babu, F., Happee, R. (2017) Driver Behaviour and Workload in an On-road Automated Vehicle. In proceedings of Road Safety & Simulation International Conference, The Hague, NL, October 17-19, 2017.

B. Virtual Immersive Reality based Analysis of Behaviour Responses in Connected and Autonomous Vehicle Environment

Shadi Djavadian

Bilal Farooq

Rafael Vasquez

Grace Yip

Motivation

The advent of connected autonomous vehicles (CAVs) at high penetration rate of 70% (Calvert et al, 2017) offers a vast range of new options to manage traffic flows, reduce congestion, increase road safety and benefit the environment without a need for centralized traffic control center or adding new infrastructures. There are many safety and sustainability benefits to this technology, such as having the ability to avoid collisions without requiring human reaction time or to smooth traffic flow patterns (dubbed “green driving”) to reduce excess energy consumption or emissions. While the emphasis has been on immediate benefits like collision avoidance or adaptive cruise control, there is an opportunity to consider more advanced decentralized traffic control designs to fully leverage the technology.

Literature Overview

One major application of the early generation of Connected Vehicle (CV) has been the dynamic traffic navigation, which provides real time travel time information and status to the driver. Studies have shown that route guidance based on CV can reduce travel time (Yang & Recker, 2006; Park & Lee, 2008; Katan et al, 2012). However, in the case of high congestion levels and adverse conditions the above-mentioned studies showed that using route guidance based on CV and without any coordination actually results in the increase in travel time and congestion. The coordination can be done by information sharing through a vehicular ad hoc network (VANET)

among vehicles (V2V) (Du et al, 2014); through a centralized route information server (V2C) (Yamashita et al, 2005); or via a road side infrastructures (V2I) (Claes et al, 2011). Under coordinated route guidance, vehicles share their route information with each other, this way they are aware of each other's actions. Thus, they can anticipate congestion on a particular route before choosing that route. It is possible for them to choose an alternative route and thus reducing congestion. The results from Claes et al. (2011), Du et al. (2014), and Yamashita et al. (2005) have all showed that the proposed distributed coordinated outperforms traditional independent selfish-routing mechanism (everyone choosing their fastest route without coordination)—reducing the expected travel time for both overall system and individual vehicles. An important factor that has major influence on the effectiveness of coordinated route guidance is drivers' willingness to share information and cooperate with each other. However, as shown in previous research an important factor that has major influence on the effectiveness of coordinated route guidance is drivers' willingness to share information and cooperate with each other. However, as also shown in previous research (Djavadian et al., 2014; Knoop et al., 2011, Netten et al, 2006), drivers are not always willing to consider and comply with information they receive or share information. Knoop et al. (2011) observed that drivers are much more receptive to mandatory systems as oppose to voluntary systems. Djavadian et al. (2014) observed that drivers are more willing to comply when they are well informed and well rewarded. Mandatory (prescriptive) information high compliance rate whereas voluntarily (descriptive) information low compliance rate unless rewarded. Netten et al., (2006) showed that drivers are more willing to comply when they can acknowledge the gain, for instance, they are more willing to comply with an advice that instructs them to accelerate rather than to an advice that instructs them to slow down.

Proposed Connected and Automated System

To increase the efficiency of route guidance system using CAVs and address the non-compliance nature of the drivers, we proposed a dynamic route guidance based on a network of intelligent intersections using level 5 (Gasser & Westhoff, 2012) connected “autonomous” vehicles (CAV), where full cooperation and coordination can be expected (Rahnamaei et al., 2017). Under the proposed framework, drivers are no longer decision makers within certain designated urban area that recurrently suffers from congestion. Automated vehicles are guided through the network from origin to destination in such a way that the proposed system optimizes network throughput. We adapted the dynamic end-to-end routing principle, Intermediate System to Intermediate System (IS-IS) dynamic link state routing, from computer networking (Tanenbaum & Wetherall, 2010) where packets (in our case CAVs) need to be routed from source (origin) to host (destination) without having a full centralized view of the network, and with unreliable intermediate nodes (intersections, link capacity, etc.). Similar to Internet protocol that routing is done by routers, in this study routing is done by intersections using dynamically updated routing table, where use of traffic light signal is replaced by vehicle to infrastructure communication. The proposed framework for end-to-end dynamic routing (E2EDR) in CAVs environment is presented in **Figure 1**. As can be seen from **Figure 1** each link periodically collects vehicles’ states using V2I and communicates them to downstream Intelligent Intersections (**I2**). The intelligent intersections have the capability of sensing, communicating, estimating and actuating. They estimate link travel times using information collected from links and then communicate them to their neighbouring upstream intersections. Each intersection using information received from its neighbour is able to create full dynamic impedance matrix of the network. Using the impedance matrix, the intelligent intersection creates a dynamic routing table.

Once a CAV arrives at an intersection it communicates its destination to the intersection, and the intelligent intersection using the routing table guides the CAVs to the next intersection that minimizes travel time to the CAV's destination.

We used a queue-based traffic simulation to test the efficiency of the proposed route guidance on Queen's Quay network in downtown Toronto. **Figure 2** presents Downtown Toronto core, with Queen's Quay network being highlighted in red. The reason for selecting Toronto Downtown core is to test the proposed algorithm on a maximum to high demand network. For this study, it was assumed that the vehicles operate manually before entering the downtown core, and once they are in the downtown core they are required to switch to auto-pilot.

Similar to compliance of the drivers, an important factor that has a major effect on the effectiveness of the dynamic CAV routing is drivers' acceptance of giving out control fully or partially to the system (Hoogendoorn et al, 2014; Kulmala, 2010; De Vos and Hoekstra, 1997). Therefore, there is a strong need to investigate the behavioral response of drivers in CAVs environment and to the best of our knowledge as also stated by Calvert et al, (2017), no empirical studies have been done in an experimental setting to understand this problem. The focus of our work here is to address this key gap.

Behavioural Analysis

We systematically investigate the travellers' reactions and possible adaption to E2EDR system? To answer the research questions, we use a Bayesian D-efficient Stated Preference (SP) experiment design (Yang et al, 2014). We design and conduct multiuser laboratory pilot experiment to evaluate behavioural responses of drivers in autonomous connected vehicle environment using our in-house Virtual Immersive Reality Environment (VIRE) simulator (Farooq & Cherchi, 2016, 2017). For this study, the current VR simulator is being modified to

allow experiments where users can drive in the VR simulator using steering wheel, thus using it as a base case. **Figure 3** presents the VR simulator that will be used for this study.

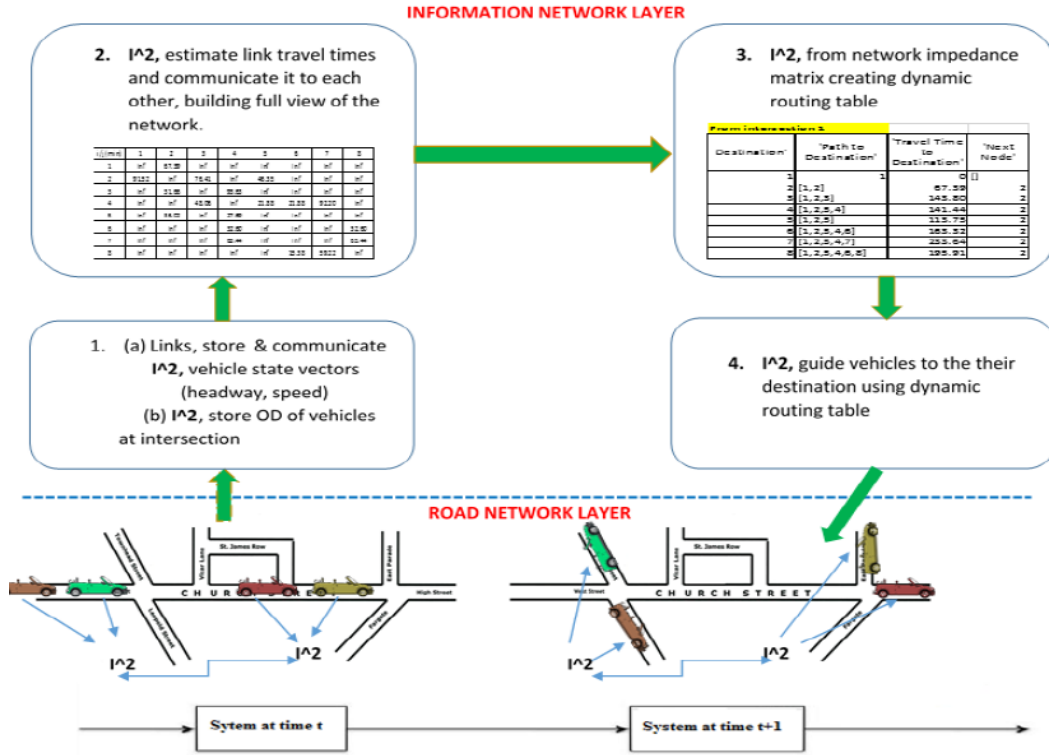


Figure 1: End-to-End dynamic route guidance in CAV Environment

To find the prior estimates of the model parameters a pilot study with 10 participants will be conducted. Once the parameters are estimated the full experiment will be conducted with group of 100 participants. Two distinct scenarios will be considered for the laboratory experiment. Under the scenario A, users will be asked to driver from point X to point Z, choosing their preferred route and have the full control of the vehicle. Under the scenario B, users will be asked to go from point X to Z, but this time, the vehicle will be automatically guided through the road network. We will repeat the scenario B three times with the variable values dictated by the initial D-efficient design. The example of some of the variables that will be investigated are: comfort level and multi-tasking

while driving, difference between CAVs chosen path and manually chosen path, ease of driving under different traffic congestions. The laboratory experiments will allow us to evaluate the responses of the users and their reaction to driving in connected autonomous vehicle environment. For the case of CAVs experiments, the VR will simulate use of smartphone, reading a book for users, this will further allow them to compare their experience of driving CAV to manual car. In combination with SP, we will collect socio-economic characteristic of each user and will provide them attitude based questionnaire. Often diffusion process is used to predict market for a new product or innovation, however, the problem with diffusion models is that they do not take into account the socio-economic characteristics of users and their discrete choice sets. As such in recent years hybrid models that combine diffuse models with advance choice models have been gaining interest among researchers. For example, Jensen et al. (2017) used joint discrete choice model and diffuse model to predict the potential for electric vehicles. In our study, similar to Jensen et al. (2017) we will be using joint discrete choice model and diffuse model.

The project results will provide us with insights on to what extend and under what circumstance drivers are willing to give full or partial control to the automated vehicles. In addition, project results will allow us to draw conclusion as to what group of population (age, gender, driving experience, income, etc.) will be more accepting of the automated routing system. The project results will be of interest to car manufacturers, since it will allow them to target right group of population thus designing the connected autonomous vehicle environment in such a way that will have maximum driver acceptance increasing market share. Also, the results will be of benefit to traffic information providers, traffic management centers, city planners and municipalities allowing them to target right audience, apply CAVs routing in situations that will results in more acceptance and efficient use of transportation system.

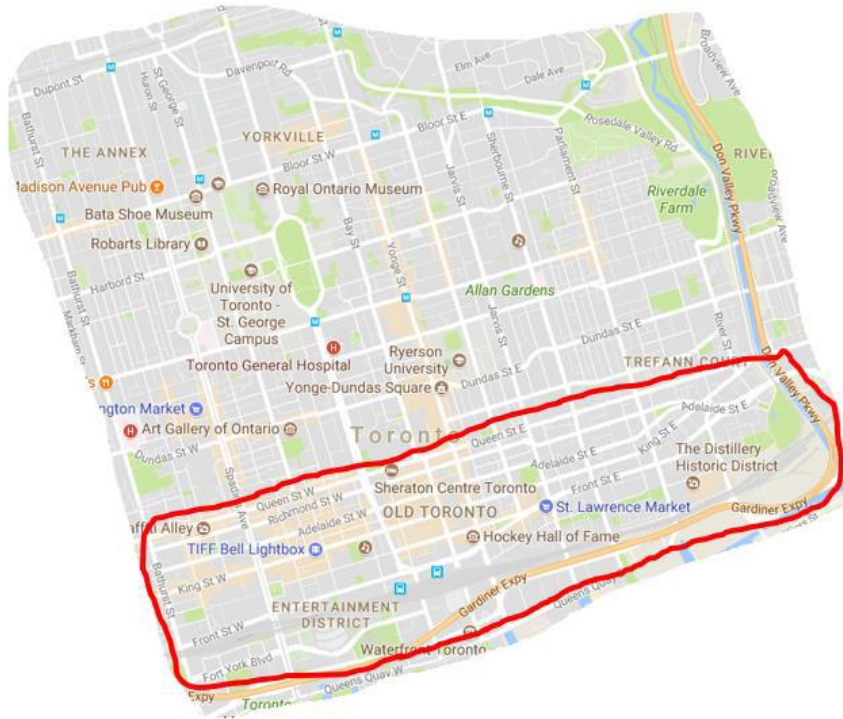


Figure 2: Queen's Quay area located in downtown Toronto (Google maps, 2017)



Figure 3: Virtual Immersive Reality Environment (VIRE) simulator

References

- Calvert, S.C., Schakel, W.J., van Lint, W.C., (2017). Will Automated Vehicles Negatively Impact Traffic Flow? *Journal of Advanced Transportation* 2017. <https://doi.org/10.1155/2017/3082781>
- Claes, R, Holvoet, T., Weyns, D., (2011). A Decentralized Approach for Anticipatory Vehicle Routing Using Delegate Multiagent Systems. *IEEE Transactions on Intelligent Transportation Systems* 12(2), 364-373.
- De Vos, A., Hoekstra, W., (1997). Behavioural Aspects of Automatic Vehicle Guidance (AVG); Leaving the Automated Lane.
- Djavadian, S., Hoogendoorn, R.G., van Arem, B., Chow, J.Y.J., (2014). Empirical evaluation of drivers' route choice behavioral responses to social navigation. *Transportation Research Record* 2423, 52- 60.
- Du, L., Han, L., Li, X.Y., (2014). Distributed coordinated in-vehicle online routing using mixed strategy congestion game. *Transportation Research Part B* 67, 1-17.
- Hoogendoorn, R., van Arem, B., Hoogendoorn, S., (2014). Automated Driving, Traffic Flow Efficiency and Human Factors: A literature Review. *Proceedings of the 93rd Annual Meeting of Transportation Research Board*. 12-16 January 2014, Washington D.C, USA.
- Farooq, B., Cherchi, E., (2017). Virtual Immersive Reality Environment (VIRE) for Disruptive Technology Choice Experiments. *Proceedings of International Choice Modelling Conference*.
- Farooq, B., Cherchi, E., (2016). Immersive Virtual Reality Experiments for Analyzing Autonomous Vehicle Demand. Transportation Research Board Invited Speaker at Automated Vehicles Symposium 2016, San Francisco, 18-22 July 2016.
- Gasser, T., Westhoff, D., (2012). BAST-study: Definitions of Automation and Legal Issues in Germany. In *Proceedings of the 2012 road vehicle automation workshop*. Automation Workshop.
- Jensen, A.F., Cherchi, E., Mabit, S., de Dios Ortuzar, J., (2016). Predicting the Potential Market for Electric Vehicles. *Transportation Science*. 51(2), 427-440
- Katan, L., Mousavi, M., Far, B., Harschnitz, C., Radmanesh, A., Saidi, S., (2012). Microsimulation of the Potential Impacts of Vehicle-to-Vehicle Communication (V2V) in Disseminating

- Warning Information under High Incident Occurrence Conditions. *Int. J. ITS Res.* 10: 137-147
- Knoop, V. L et al. 2011, 'Individual Travelers' Advice: System Setup, Measures, and Expected Results' in *Transportation Research Board Annual Meeting 2011* Paper #11-1043.
- Kulmala., R (2010). Ex-ante assessment of the safety effects of intelligent transport systems. *Accid Anal Prev* 42:1359–1369. doi:10.1016/j.aap. 2010.03.001
- Lee, J., Park, B., (2008). Evaluation of Vehicle Infrastructure Integration (VII) based route guidance strategies under incident conditions. *Transportation Research Record: Journal of the Transportation Research Board.* No. 2086, pp. 107-114.
- Netten, B., Wilmink, I., Hogma, J., Hegyi, A., Wang, W., Baan, J., Passchier, I., (2011). Dynamax InCar: system, concepts, research questions, hypothesis, simulation and evaluation. TNO Report
- Rahnamaei, N., Djavadian, S., Farooq, B., (2017). Dynamic Routing of Connected Vehicles in Origin to Destination Urban Network. Ryerson University Laboratory of Innovations in Transportation. (Working paper)
- Tanenbaum, A.S., Wetheral, D., (2010). *Computer Networks*. Harlow, Essex: Pearson Education, 2010. Print.
- Yamashita, T., Izumi, K., Kurumatani, K., Nakashima, H., (2005). Smooth traffic flow with a cooperative car navigation system. In *Proc. 4th International Joint Conference on Autonomous agents and multiagent systems*, 478–485.
- Yang, X., Recker, W., (2006). Modeling Dynamic Vehicle Navigation in a self-organizing, peer to-peer, distributed traffic information system. *Journal of Intelligent Transportation Systems*. Volume 10, Issue 4.
- Yange, F., Chen, L., Cheng, Y., Luo, X., Ran, B., (2014). An Empirical Study of Parameter Estimation for Stated Preference Experimental Design. *Mathematical Problems in Engineering 2014*. <http://dx.doi.org/10.1155/2014/292608>

References

- [1] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, “A Survey of Autonomous Driving: Common Practices and Emerging Technologies,” Jun. 2019.
- [2] J. E. Stellet, M. R. Zofka, J. Schumacher, T. Schamm, F. Niewels, and J. M. Zollner, “Testing of Advanced Driver Assistance Towards Automated Driving: A Survey and Taxonomy on Existing Approaches and Open Questions,” in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 1455–1462.
- [3] Animesh, Pinsonneault, Yang, and Oh, “An Odyssey into Virtual Worlds: Exploring the Impacts of Technological and Spatial Environments on Intention to Purchase Virtual Products,” *MIS Quarterly*, vol. 35, no. 3, p. 789, 2011.
- [4] C. Jennett *et al.*, “Measuring and defining the experience of immersion in games,” *International Journal of Human-Computer Studies*, vol. 66, no. 9, pp. 641–661, Sep. 2008.
- [5] M. Everingham, L. van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The Pascal Visual Object Classes (VOC) Challenge,” *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, Jun. 2010.
- [6] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for autonomous driving? The KITTI vision benchmark suite,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3354–3361.
- [7] F. Yu *et al.*, “BDD100K: A Diverse Driving Video Database with Scalable Annotation Tooling.”
- [8] W. Maddern, G. Pascoe, C. Linegar, and P. Newman, “1 year, 1000 km: The Oxford RobotCar dataset,” *The International Journal of Robotics Research*, vol. 36, no. 1, pp. 3–15, Jan. 2017.
- [9] E. Santana and G. Hotz, “Learning a Driving Simulator,” Aug. 2016.
- [10] Y. Chen *et al.*, “LiDAR-Video Driving Dataset: Learning Driving Policies Effectively,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5870–5878.
- [11] A. Blatt *et al.*, *Naturalistic Driving Study: Field Data Collection*. Washington, D.C.: Transportation Research Board, 2014.
- [12] M. Zhu, X. Wang, A. Tarko, and S. Fang, “Modeling car-following behavior on urban expressways in Shanghai: A naturalistic driving study,” *Transportation Research Part C: Emerging Technologies*, vol. 93, pp. 425–445, Aug. 2018.

- [13] B. Zhu, Y. Jiang, J. Zhao, R. He, N. Bian, and W. Deng, "Typical-driving-style-oriented Personalized Adaptive Cruise Control design based on human driving data," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 274–288, Mar. 2019.
- [14] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," *arXiv preprint arXiv:1711.03938*, 2017.
- [15] A. Best, S. Narang, L. Pasqualin, D. Barber, and D. Manocha, "AutonoVi-Sim: Autonomous Vehicle Simulation Platform with Weather, Sensing, and Traffic Control," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018, pp. 1161–11618.
- [16] S. Shah, D. Dey, C. Lovett, and A. Kapoor, "Airsim: High-fidelity visual and physical simulation for autonomous vehicles," in *Field and service robotics*, 2018, pp. 621–635.
- [17] B. Wymann, E. Espié, C. Guionneau, C. Dimitrakakis, R. Coulom, and A. Sumner, "Torcs, the open racing car simulator," *Software available at <http://torcs.sourceforge.net>*, vol. 4, no. 6, 2000.
- [18] Q. Zhu, Z. Huang, Z. Sun, D. Liu, and B. Dai, "Reinforcement Learning based Throttle and Brake Control for Autonomous Vehicle Following."
- [19] M. Zhu, X. Wang, and Y. Wang, "Human-like autonomous car-following model with deep reinforcement learning," *Transportation Research Part C: Emerging Technologies*, vol. 97, pp. 348–368, Dec. 2018.
- [20] A. Kesting and M. Treiber, "Calibrating Car-Following Models by Using Trajectory Data," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2088, no. 1, pp. 148–156, Jan. 2008.
- [21] H. Gao, G. Shi, G. Xie, and B. Cheng, "Car-following method based on inverse reinforcement learning for autonomous vehicle decision-making," *International Journal of Advanced Robotic Systems*, vol. 15, no. 6, Nov. 2018.
- [22] S. Chowdhuri, T. Pankaj, and K. Zipser, "MultiNet: Multi-modal multi-task learning for autonomous driving," in *Proceedings - 2019 IEEE Winter Conference on Applications of Computer Vision, WACV 2019*, 2019, pp. 1496–1504.
- [23] H. Chae, C. M. Kang, B. Kim, J. Kim, C. C. Chung, and J. W. Choi, "Autonomous Braking System via Deep Reinforcement Learning," Feb. 2017.
- [24] C. J. C. H. Watkins and P. Dayan, "Q-learning," *Machine Learning*, vol. 8, no. 3–4, pp. 279–292, May 1992.

- [25] H. Chae, C. M. Kang, B. Kim, J. Kim, C. C. Chung, and J. W. Choi, "Autonomous braking system via deep reinforcement learning," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 2017, pp. 1–6.
- [26] A. Rasouli, I. Kotseruba, and J. K. Tsotsos, "Understanding Pedestrian Behavior in Complex Traffic Scenes," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 1, pp. 61–70, Mar. 2018.
- [27] M. M. Ishaque and R. B. Noland, "Behavioural Issues in Pedestrian Speed Choice and Street Crossing Behaviour: A Review," *Transport Reviews*, vol. 28, no. 1, pp. 61–85, Jan. 2008.
- [28] B. Färber, "Communication and communication problems between autonomous vehicles and human drivers," in *Autonomous driving*, Springer, 2016, pp. 125–144.
- [29] B. Farooq, E. Cherchi, and A. Sobhani, "Virtual Immersive Reality for Stated Preference Travel Behavior Experiments: A Case Study of Autonomous Vehicles on Urban Roads," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 50, pp. 35–45, Dec. 2018.
- [30] E. Cherchi and D. A. Hensher, "Workshop Synthesis: Stated Preference Surveys and Experimental Design, an Audit of the Journey so far and Future Research Perspectives," *Transportation Research Procedia*, vol. 11, pp. 154–164, 2015.
- [31] Z. Patterson, J. M. Darbani, A. Rezaei, J. Zacharias, and A. Yazdizadeh, "Comparing text-only and virtual reality discrete choice experiments of neighbourhood choice," *Landscape and Urban Planning*, vol. 157, pp. 63–74, Jan. 2017.
- [32] Y. Wang *et al.*, "Development of an immersive virtual reality head-mounted display with high performance," *Applied optics*, vol. 55, no. 25, pp. 6969–6977, 2016.
- [33] H. Wu, D. H. Ashmead, and B. Bodenheimer, "Using immersive virtual reality to evaluate pedestrian street crossing decisions at a roundabout," in *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, 2009, pp. 35–40.
- [34] A. Sobhani, B. Farooq, and Z. Zhong, "Immersive head mounted virtual reality based safety analysis of smartphone distracted pedestrians at street crossing," in *Road Safety and Simulation conference*, 2017, no. 1–12.
- [35] M. B. Neider, J. S. McCarley, J. A. Crowell, H. Kaczmariski, and A. F. Kramer, "Pedestrians, vehicles, and cell phones," *Accident Analysis & Prevention*, vol. 42, no. 2, pp. 589–594, 2010.
- [36] D. Stavrinos, K. W. Byington, and D. C. Schwebel, "Distracted walking: Cell phones increase injury risk for college pedestrians," *Journal of Safety Research*, vol. 42, no. 2, pp. 101–107, Apr. 2011.

- [37] D. Stavrinou, K. W. Byington, and D. C. Schwebel, "Effect of Cell Phone Distraction on Pediatric Pedestrian Injury Risk," *PEDIATRICS*, vol. 123, no. 2, pp. e179–e185, Feb. 2009.
- [38] A. Juliani *et al.*, "Unity: A general platform for intelligent agents," *arXiv preprint arXiv:1809.02627*, 2018.
- [39] S. Sahaleh, M. Bierlaire, B. Farooq, A. Danalet, and F. Hänseler, "Scenario analysis of pedestrian flow in public spaces," in *12th Swiss Transport Research Conference*, 2012, no. CONF.
- [40] A. Sobhani and B. Farooq, "Impact of smartphone distraction on pedestrians' crossing behaviour: An application of head-mounted immersive virtual reality," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 58, pp. 228–241, Oct. 2018.
- [41] A. Kalatian and B. Farooq, "DeepSurvival: Pedestrian Wait Time Estimation in Mixed Traffic Conditions Using Deep Survival Analysis," *arXiv preprint arXiv:1904.11008*, 2019.
- [42] S. Djavadian, B. Farooq, R. Vasquez, and G. Yip, "Virtual Immersive Reality based Analysis of Behavioral Responses in Connected and Autonomous Vehicle Environment," *arXiv preprint arXiv:1901.07151*, 2019.
- [43] S. Djavadian, R. Vasquez, and B. Farooq, "Should I give or should I take? Choice issues in automated vehicle control," *6th International Choice Modelling Conference*, 2019.
- [44] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [45] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.
- [46] T. P. Lillicrap *et al.*, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.
- [47] D. Eager, A.-M. Pendrill, and N. Reistad, "Beyond velocity and acceleration: jerk, snap and higher derivatives," *European Journal of Physics*, vol. 37, no. 6, p. 065008, Nov. 2016.
- [48] H. Bellem, B. Thiel, M. Schrauf, and J. F. Krems, "Comfort in automated driving: An analysis of preferences for different automated driving styles and their dependence on personality traits," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 55, pp. 90–100, May 2018.
- [49] H. Bellem, T. Schönenberg, J. F. Krems, and M. Schrauf, "Objective metrics of comfort: Developing a driving style for highly automated vehicles," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 41, pp. 45–54, Aug. 2016.

- [50] F. Hartwich, M. Beggiato, and J. F. Krems, “Driving comfort, enjoyment and acceptance of automated driving – effects of drivers’ age and driving style familiarity,” *Ergonomics*, vol. 61, no. 8, pp. 1017–1032, Aug. 2018.
- [51] T. A. Jost, B. Nelson, and J. Rylander, “Quantitative analysis of the Oculus Rift S in controlled movement,” *Disability and Rehabilitation: Assistive Technology*, pp. 1–5, 2019.
- [52] A. Hayat, U. Singh, and V. P. Namboodiri, “InfoRL: Interpretable Reinforcement Learning using Information Maximization,” *arXiv preprint arXiv:1905.10404*, 2019.
- [53] A. Verma, V. Murali, R. Singh, P. Kohli, and S. Chaudhuri, “Programmatically interpretable reinforcement learning,” *arXiv preprint arXiv:1804.02477*, 2018.