

Mobility21

A USDOT NATIONAL
UNIVERSITY TRANSPORTATION CENTER

Carnegie Mellon University



Understanding and Guiding Pedestrian and Crowd Motion

Dongfang Yang (<https://orcid.org/0000-0001-9212-6804>)

Keith Redmill (<https://orcid.org/0000-0003-1332-1332>)

Ümit Özgüner (<https://orcid.org/0000-0003-2241-7547>)

FINAL RESEARCH REPORT - April 24, 2020

Contract # 69A3551747111

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. This report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. The U.S. Government assumes no liability for the contents or use thereof.

Contents

1	Introduction	3
2	Social Force Modeling of Vehicle-Pedestrian Interaction	6
2.1	Background	6
2.2	Problem Formulation	7
2.3	Pedestrian Motion Modeling	8
2.3.1	Fundamental Functions	8
2.3.2	Pedestrian Dynamics	10
2.3.3	Surrounding Pedestrian Effect	12
2.3.4	Destination Force	14
2.3.5	Vehicle Effect	15
2.3.6	Vehicle Motion	21
2.4	Conclusion	22
3	Dataset	23
3.1	Background	23
3.2	Related Works	26
3.3	Dataset	26
3.3.1	CITR Dataset	26
3.3.2	DUT Dataset	28
3.4	Trajectory Extraction	30
3.4.1	Video Stabilization	30
3.4.2	Vehicle and Pedestrian Tracking	32
3.4.3	Coordinate Transformation	32
3.4.4	Trajectory Filtering	32
3.5	Statistics	33
3.6	Conclusion	33
4	Model Calibration	36
4.1	Background	36
4.2	Calibration	36
4.2.1	Parameter Set	36
4.2.2	Scenarios and Dataset	37
4.2.3	Calibration Procedure	37
4.2.4	Genetic Algorithm	39
4.3	Result	40
4.3.1	Pedestrian-Pedestrian Interaction	40
4.3.2	Vehicle-Pedestrian Interaction	43

4.4	Conclusion	45
5	Applications	48
5.1	Background	48
5.2	Framework	49
5.3	Pedestrian Detection in Crowd	51
5.3.1	On-vehicle Sensors	51
5.3.2	Infrastructure Sensors	52
5.3.3	Sensor fusion	54
5.4	Transportation Efficiency in Shared Spaces	54
5.4.1	Scenario Prediction	55
5.4.2	Approaches	55
5.5	Conclusion	56
6	Conclusion	57
6.1	Conclusion	57
6.2	Future Work	57
A	Social Force Modeling	59
A.1	Validation of Pedestrian-Pedestrian Interaction	59
A.2	Validation of Vehicle-Pedestrian Interaction	60
A.3	Post-Simulation of Vehicle-Pedestrian Interaction	61
B	Research Products for this Project	62
B.1	Journal Publications	62
B.2	Conference Publications	62
B.3	Thesis	62
B.4	Datasets	62
	Bibliography	63

Chapter 1

Introduction

Future transportation system is expected to be in a hybrid form instead of the basic road-vehicle-passenger relationship that is still very common in most transportation systems. That is to say, future transportation will cover additional scenarios and different types of vehicles with partial or full automation. A very representative example is the first-mile and last-mile transportation (FLT), which provides the service that cannot be achieved by traditional transportation system because of how the geometrical structures were designed in certain areas, for example, large squares, university campuses, and old small streets. This type of service can be achieved by vehicles of smaller size such as golf carts, mobility scooters, and electric wheelchairs. This is especially helpful for people with limited mobility. One essential challenge of developing the systems of FLT is to make it capable of avoiding pedestrians, as pedestrians are very common in such shared space scenarios and the interaction between the vehicle and pedestrians is almost unavoidable.

Improving pedestrian safety has been regarded as a long-existing issue in the transportation system. A recent released report of Traffic Safety Facts by NHTSA in March 2019 [39] revealed that the percentage of pedestrian fatalities was increasing from 2008 to 2017, with 12% in 2008 but 16% in 2017, and this was happening in the background of the advanced driver-assistance systems (ADAS) being mature. This implies that the state-of-the-art technologies for protecting pedestrians are not good enough, or not as fast as the speed of development of other safety-related functions. Nowadays with the boost of automated driving technologies that are trying to totally remove human drivers, the pedestrian safety becomes increasingly important, as now it is the intelligent system on the vehicle that interacts with pedestrians but not the human driver any more. Therefore, the intelligent system needs to understand the behavior, or possibly the interactive and collective behavior of multiple pedestrians in shared space scenarios of the FLT applications, so that the system can effectively control the automated vehicle.

This project was conducted under the background of the Ohio State University's initiative of a network of "on demand automated vehicles", a.k.a., Smart Mobile Operation: OSU Transportation Hub (SMOOTH) [52, 16] since 2014. The SMOOTH demonstrated using a closed circuit of automated shuttles driving within the Ohio State University main campus and two selected stops within the outer campus. The automated shuttles have Vehicle to Vehicle (V2V) communication and are equipped with vulnerable road user detection technology, enabling them to function in pedestrian zones on campus. Contemporaneous similar projects include a European project CityMobil2 [8], which has been setting up a pilot platform for automated road transport systems that are made up of vehicles operating without drivers and providing service in areas of low or dispersed de-

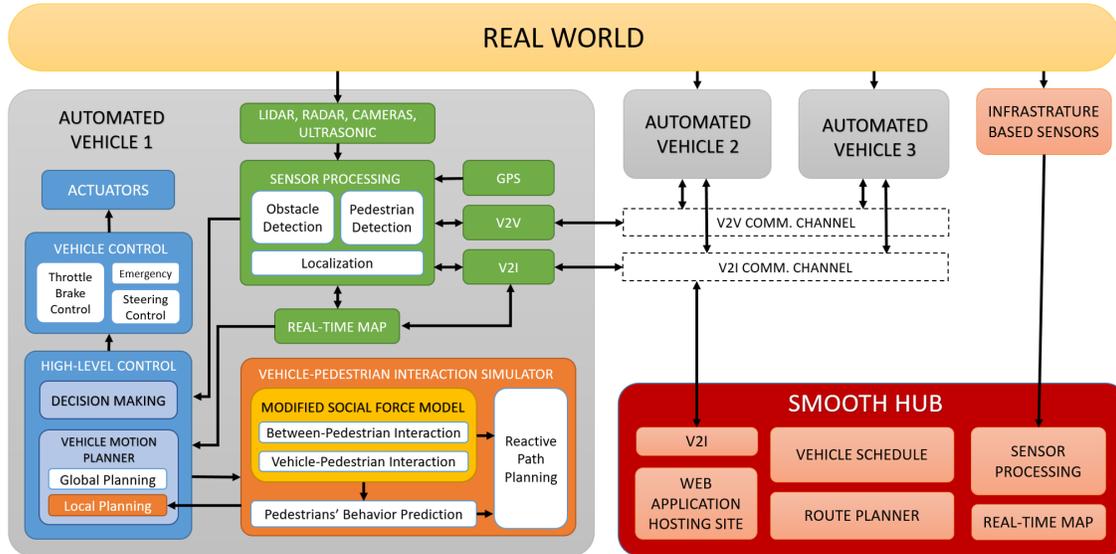


Figure 1.1: SMOOTH architecture and the integration of social force based vehicle-pedestrian interaction.

mand, and the LUTZ Pathfinder Project [50] under Transport Systems Catapult in UK, which has been developing driverless pods in public streets.

Fig. 1.1 shows the architecture of SMOOTH concept design that integrates an interactive pedestrian motion model. It primarily consists a transportation hub, several automated vehicles, and a mobile-first web-based application. The transportation hub is responsible for vehicle scheduling, route planning, infrastructure sensor data processing, communication, and web hosting. Three pilot automated vehicles of different sizes equipped with SMOOTH techniques are a motorized wheelchair, a single-person mobility scooter, and a four-passenger golf cart. The web-based application is available to all smart device, where users can log into their accounts and schedule an automated vehicle for service. The vehicle-pedestrian interaction module/simulator is designed to predict pedestrians' behavior, which is subsequently utilized by the automated vehicle to plan its motion and execute the control action. Specifically, the vehicle-pedestrian interaction simulator works according to the following steps. First, sensors such as LIDAR, radar, cameras, ultrasonic and GPS, and communication approaches such as V2V and V2I recognize the surrounding environment. Once the received data is processed, information including each pedestrian's position, velocity and potential destination can be identified and made available to the vehicle-pedestrian interaction simulator. Although in reality, it would be difficult to obtain such exact information due to occlusions and sensor noises. Here we assume that all pedestrian information is available because precise pedestrian detection is another research topic. With pedestrian information, pedestrian behaviors (possible moving trajectories) are simulated based on current configurations of both pedestrians and the vehicle. Using the simulation results, vehicle motion planner can determine more reliable and efficient local reactive decisions and perform real-time local path planning, which is a complement or an addition to global path planning. Both the global and local path plannings constitute the high-level controller, hence vehicle control is achieved and the corresponding action is executed.

The rest of the report will focus on the interactive pedestrian motion model, i.e.,

modeling, simulation, and application of multi-pedestrian interaction with vehicles. We modified the social force pedestrian motion model [22] by explicitly designing different types of vehicle effect on pedestrian behavior [56, 59, 60], which are detailed in chapter 2. We also established a dataset that captures the trajectories of the vehicle and pedestrians in fundamental vehicle-crowd interaction scenarios [57], which is described in chapter 3. Using the established dataset, the calibration of the proposed model [60] is demonstrated in chapter 4. A framework and several specific approaches to apply the proposed model into vehicle driving efficiency, along with several case studies [58] of vehicle-pedestrian interaction were proposed in chapter 5. Conclusions and future works are discussed in chapter 6.

Chapter 2

Social Force Modeling of Vehicle-Pedestrian Interaction

2.1 Background

The mathematical modeling of pedestrian motion began with pedestrian-pedestrian interaction, i.e., only the effect from surrounding pedestrians is considered. The pioneer work was the social force model [22], which was originally designed for crowd motion/flow simulation and analysis. It can reproduce typical crowd motion patterns such as lane formation and fluctuation, as summarized in a paper reviewing fundamental phenomena of pedestrian crowd [46]. Later on, mathematical modelings such as dynamic programming [24], discrete choice [2], cognitive method [37], linear trajectory avoidance [42], and heuristic with Voronoi diagram [54] were proposed. They were somehow derived from social force model but with more emphasis on computational efficiency and simplicity. These models rely on verified handcrafted mathematical rules to describe the interaction. In addition to the above rule-based models, neural network models that build on long-short term memory (LSTM) [1, 55, 20] have recently become an alternative choice of modeling pedestrian motion. Although the application of neural network models is more about motion prediction than simulation, essentially both rule-based models and neural network models share the same feature of addressing pedestrian motion.

Now, with vehicle considered, the effect of vehicle should be somehow incorporated into the pure pedestrian motion models. Some works proposed the modeling of vehicle influence on individual pedestrians [15, 53, 19]. They did not specifically consider multi-pedestrian interaction with vehicles, i.e., the effect of surrounding pedestrians was usually neglected, and the scenarios were usually restricted to crosswalks. The general vehicle-pedestrian interaction is more complicated than the interaction in restricted scenarios. Social force model is one of the options that can easily incorporate the effect of vehicle. This is because in social force model a pedestrian is regarded as a point mass agent, and dynamics of the agent is subject to a summation of individual effects from different sources, e.g., attraction to the destination, repulsion to surrounding pedestrians, etc. The effect of vehicle can be designed as an additional source into the summation, which is straightforward and effective.

Social force model has been continually improved and modified since its introduction in [22]. The original model only considered the repulsion and attraction of surrounding pedestrians, as well as the attraction of the destination. Later on, collision force was

added to account for extremely crowded situations such as emergency evacuation or pilgrimage [21]. In [61], the authors summarized and compared different specifications of the effect of surrounding pedestrians.

Different approaches have been attempted to add vehicle influence into social force model. Some works [3, 63, 62] modeled the vehicle influence as an additional force added on the pedestrian dynamics. The additional force not only considered the relative positions and velocities of the vehicle as in the effect of surrounding pedestrians, other vehicle features such as size, shape, and restriction on motion were also considered. To address the complexity introduced by vehicle influence, multi-layer models that contain social force were proposed [44, 3, 62, 26, 48, 41]. Social force was embedded into a layer usually referred to as behavioral layer. Above the behavioral layer, there is another layer mainly responsible for finding a global route or intermediate destinations that can guide the pedestrian motion. The social force of the behavioral layer is only responsible for adjusting local motion. For example, in [48], the proposed model consists of 3 different layers that are primarily responsible for scene context effect, local motion of the agents, and road conflicts, respectively. Vehicle influence does not only exist in the layer of social force, but part of the influence is also described in other layers in combination with specific approaches, for example, conflict resolving via 'shadow' method in [3], game theory in [48, 26], and long range and short range conflicts in [44, 41].

Our work, however, applied a single layer social force model. Instead of accounting for the complexity introduced by the interaction of different types of road users, the proposed model focuses on the effect of the vehicle itself inside the social force model. Moreover, the model aims to describe fundamental interaction scenarios of multi-pedestrian interaction with vehicle; hence the model doesn't consider any scene information, which is the main reason and the basis of applying multi-layer models.

The proposed social force model mathematically describes the pedestrian motion that is primarily affected by multiple surrounding pedestrians and a low speed vehicle. The model introduced a general vehicle-pedestrian interaction design that was validated by fundamental interaction scenarios of multiple pedestrians coming from different directions and interacting with the vehicle (front, back, and lateral interaction).

In the project, the proposed social model went through three versions. The major difference among these versions is the design of vehicle effect, with the first version being a simple elliptic effect [56], the second version using three-area effect [59], and the third and the latest version considering anisotropy [60]. Minor changes of the model such as the constraints of the speed/acceleration also exist in different version. This report will focus on the latest version of the proposed social force model, while the the design of the first two versions are summarized.

2.2 Problem Formulation

The fundamental problem is to design a social force based mathematical model that describes pedestrian motion in mixed traffic scenarios, primarily considering the effect of surrounding pedestrians and vehicles. The model should be able to generate future motion of the ego pedestrian based on the immediate status of the pedestrian. That is, given the current states of all interacting agents (positions and velocities of all surrounding pedestrians and vehicles), the associated model should output the next step's position and velocity of the ego pedestrian. With each individual pedestrian assigned with a model,

trajectories of all pedestrians can be generated by iteratively applying the models. The generated trajectories are also referred to as the predicted motion of the pedestrians.

Since multi-pedestrian interaction with a low speed vehicle is primarily considered in this work, the interaction scenarios are defined such that there are at least 5 pedestrians and the vehicle speed should be less than $4m/s$. The pedestrian number and the vehicle speed were empirically determined, which may vary in other circumstances. Multiple pedestrians are also referred to as crowd in this work. The low speed vehicle applies in most shared space scenarios, in which pedestrians and vehicles are mixed together to share the right of the road/space and the vehicle usually pays more attention to the pedestrians. The space layout is assumed to be empty, hence no scene information. This configuration releases the vehicle from being restricted in lanes so that a variety of vehicle maneuvers are available.

Mathematically, if we define, at time t , the state (position and velocity) of pedestrian i as $\vec{x}_t^i = (x_t^i, y_t^i, v_{x,t}^i, v_{y,t}^i)^T$, the state (position, speed, and orientation) of vehicle as \vec{x}_t^v , the model can be expressed as:

$$\vec{x}_{t+1}^i = f_i(\vec{x}_t^i, \{\vec{x}_t^{j \neq i}\}, \vec{x}_t^v) \quad (2.1)$$

This work does not focus on the generation of vehicle motion; hence the vehicle state $\{x_t^v, \forall t\}$ is assumed to be known all the time. It is either directly obtained from the trajectory dataset or intentionally synthesized. When intentionally synthesized, a kinematic bicycle model with a pure pursuit steering controller and a PID speed controller is applied to generate realistic vehicle motion.

2.3 Pedestrian Motion Modeling

2.3.1 Fundamental Functions

Some fundamental functions are described in this subsection, because they serve as essential components in the proposed social force pedestrian motion model. They are anisotropy functions and decaying functions. Anisotropy functions are used to describe different effect of the interacting agents from different directions. For example, a pedestrian right in front of the ego pedestrian obviously has bigger influence than a pedestrian on the left or right side of the ego pedestrian. Decaying functions are used to describe the different effect of different interaction distances. For example, a vehicle that is very far away from the ego pedestrian has merely no influence, while a vehicle that is very close to the ego pedestrian surely has large influence. Both types of functions have different specifications, which are selected based on the specific requirement of each component of the model.

Anisotropy Functions

Anisotropy functions take input as the angle between the ego pedestrian's walking direction and the direction to the target agent that is interacting with the ego pedestrian. The output of anisotropy function is a scalar ranging from 0 to 1, representing how the influence attenuates as the angle increases.

Three different types of anisotropies (linear, sinusoidal, and exponential) are used in

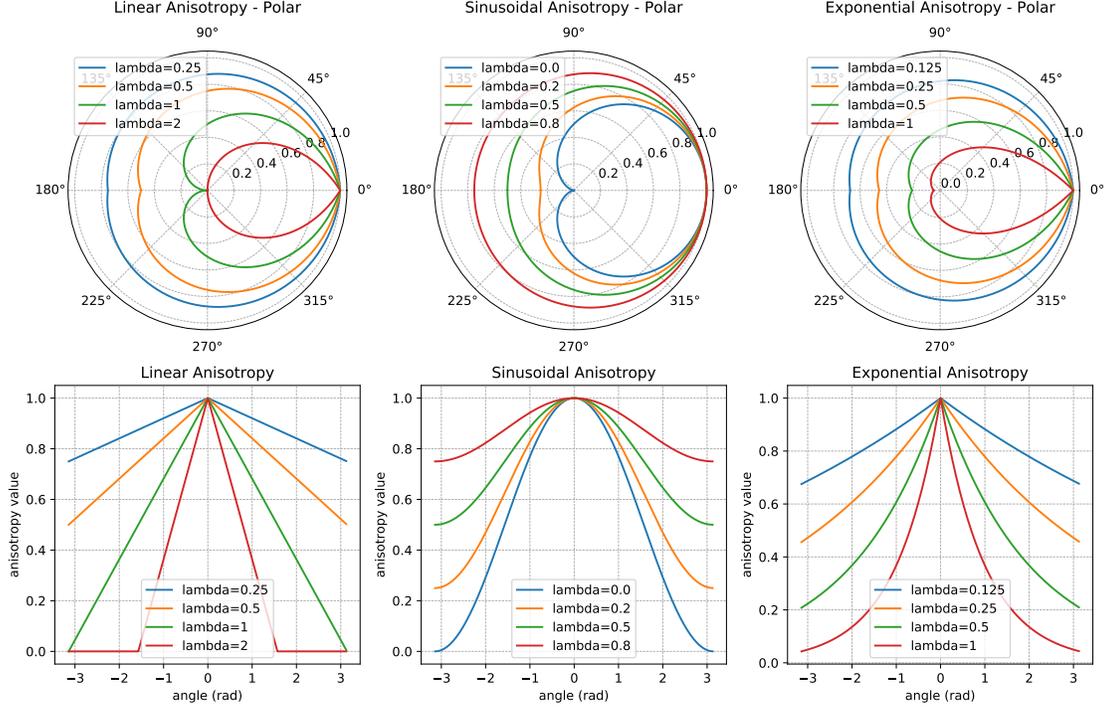


Figure 2.1: Illustration of different anisotropies. Columns from left to right: linear anisotropy, sinusoidal anisotropy, and exponential anisotropy. Different parameter values generate different anisotropies.

this work, as shown in figure 2.1, with the expressions:

$$\mathbb{A}_{lin}(\phi, \lambda) = \max\left\{1 - \lambda \cdot \frac{|\phi|}{\pi}, 0\right\} \quad (2.2)$$

$$\mathbb{A}_{sin}(\phi, \lambda) = \lambda + (1 - \lambda) \cdot \frac{1 + \cos|\phi|}{2} \quad (2.3)$$

$$\mathbb{A}_{exp}(\phi, \lambda) = \exp(-\lambda \cdot |\phi|), \quad (2.4)$$

where $\phi \in [-\pi, \pi]$ is a variable representing the interaction angle and λ is the parameter adjusting the anisotropy characteristics. The major difference among these anisotropies is the rate of attenuation at the angles near 0. For example, as $|\phi|$ increases from 0 to a certain angle (e.g. 90 degrees), exponential anisotropy attenuates very fast, but sinusoidal anisotropy attenuates relatively slow (see the second row in figure 2.1). This difference plays an important role in modeling pedestrian’s reaction to a target agent.

Decaying Functions

Decaying functions take input as the distance between the ego pedestrian to the target agent. The output is the magnitude of the influence, i.e., the force magnitude applied to the point mass dynamics. The magnitude decreases monotonically as the distance increases.

Exponential function, as shown in the left of figure 2.2, is a common option in most social force models, due to its simplicity and effectiveness:

$$f_{exp}(d, A, B) = A \exp(-Bd), \quad (2.5)$$

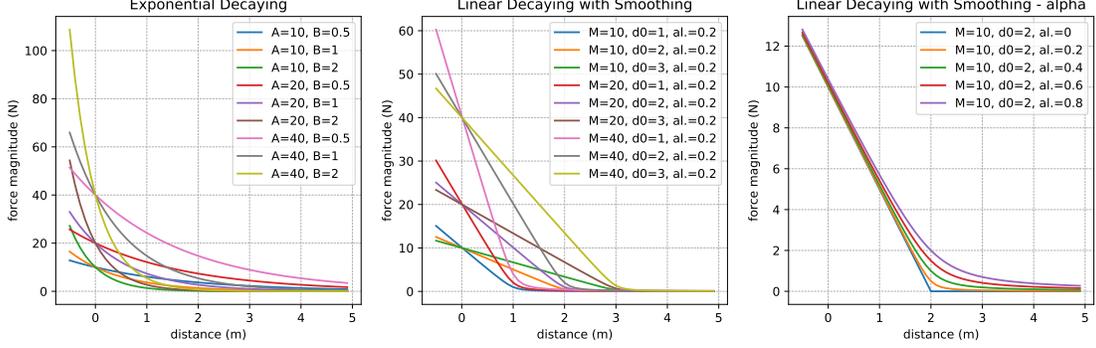


Figure 2.2: Illustration of decaying functions. Left: exponential decaying with different parameter values; Middle: linear function with smoothing with different parameter values while α is fixed; Right: linear function with smoothing with different parameter values while M and d_0 is fixed.

where d is a variable representing the distance between the ego pedestrian and the target agent, and A, B are parameters adjusting the characteristics of the decaying relationship.

Another type of decaying function, as shown in the middle of figure 2.2, describes a linear relationship with smoothness modification [12] on the point where the magnitude reaches zero:

$$f_{lm}(d, d_0, M, \sigma) = \frac{M}{2d_0} \cdot \left(d_0 - d + \sqrt{(d_0 - d)^2 + \sigma} \right), \quad (2.6)$$

where d is again a variable representing the distance, M is the force magnitude when the distance reaches zero (physical touch happens), σ is the parameter that modifies the smoothness around zero magnitude, and d_0 is a threshold distance where the magnitude almost reaches zero (equals zero if $\sigma = 0$).

The exponential function is effective in most situations, which has been demonstrated in various social force based models. In an exponential function, the rate of decaying is still exponential, which is helpful in some situations. For example, in the scenario of a vehicle approaching the ego pedestrian, when the distance from the ego pedestrian to the vehicle is very close, the exponential increase of the magnitude of the vehicle influence quickly drives the pedestrian away from the vehicle, which is normal due to the severe consequence of potential collision. However, the exponential relationship can be unrealistic in other situations. For example, the change of effect of a surrounding pedestrian from 4 meters to 2 meters prefers a more linear relationship than an exponential relationship.

2.3.2 Pedestrian Dynamics

Pedestrians are regarded as point mass agents in social force model. The motion of an agent is governed by Newtonian dynamics with the state of position x^i, y^i and velocity

v_x^i, v_y^i expressed as:

$$\begin{aligned}\dot{x}^i &= v_x^i \\ \dot{y}^i &= v_y^i \\ \dot{v}_x^i &= a_x^i = \frac{F_x^i}{m^i} \\ \dot{v}_y^i &= a_y^i = \frac{F_y^i}{m^i},\end{aligned}\tag{2.7}$$

where m^i is the mass of pedestrian i , $\vec{F}_t^i = (F_{x,t}^i, F_{y,t}^i)^T$ is the total force applied on the point mass. The above dynamics is discretized by a discretization time of Δt in this work. Therefore, a state-space vector $\vec{x}_t^i = (x_t^i, y_t^i, v_{x,t}^i, v_{y,t}^i)^T$ at time t is updated at every time step after calculating the total force $\vec{F}_t^i = (F_{x,t}^i, F_{y,t}^i)^T$ based on the immediate interaction status.

According to the definition of social force model, the total force \vec{F}_t^i is the summation of multi-source effect:

$$\vec{F}_t^i = \vec{F}_t^{i,ped} + \vec{F}_t^{i,veh} + \vec{F}_t^{i,des},\tag{2.8}$$

where $\vec{F}_t^{i,ped}$ is the pedestrian-pedestrian interaction force (effect of surrounding pedestrians), $\vec{F}_t^{i,veh}$ the vehicle-pedestrian interaction force (effect of vehicle), and $\vec{F}_t^{i,des}$ the destination force (effect of attraction of destination), all of which are detailed in the following sections.

Although each pedestrian is viewed as a point-mass agent, a virtual radius of R_i is considered when calculating the distance between the ego pedestrian and a target pedestrian. This allows two pedestrians to overlap a little bit with each other, which is regarded as the effect of pushing and squeezing. Therefore, the boundary distance between two pedestrians are defined as:

$$d_t^{ij} = |\vec{r}_t^{ij}| - R_i - R_j,\tag{2.9}$$

where $\vec{r}_t^{ij} := (x_t^j, y_t^j)^T - (x_t^i, y_t^i)^T$ is a vector that points from the ego pedestrian i to the target pedestrian j .

Constraints

Limits of velocity and acceleration are imposed on pedestrian motion. There is an absolute limit beyond which the pedestrian can never reach due to the physiological limit of human beings. In normal conditions, pedestrians don't reach the limit unless something emergent happens, e.g. a vehicle is approaching in a dangerous way. Pedestrians also tend to restrict the velocities and accelerations within a certain range in order to walk comfortably in free flow or to adapt particular situations, e.g., when the pedestrian density increases they naturally slow down. Therefore, the constraints applied on both the velocity and the acceleration are time-dependent. In this work, the constraints account for the vehicle-pedestrian interaction force $\vec{F}_t^{i,veh}$ and the reciprocal of nearby pedestrian density, i.e., the sparseness of nearby pedestrians S_t^i :

$$|\vec{v}_t^i| \leq v_{lim,t}^i(\vec{F}_t^{i,veh}, S_t^i)\tag{2.10}$$

$$|\vec{a}_t^i| \leq a_{lim,t}^i(\vec{F}_t^{i,veh}, S_t^i),\tag{2.11}$$

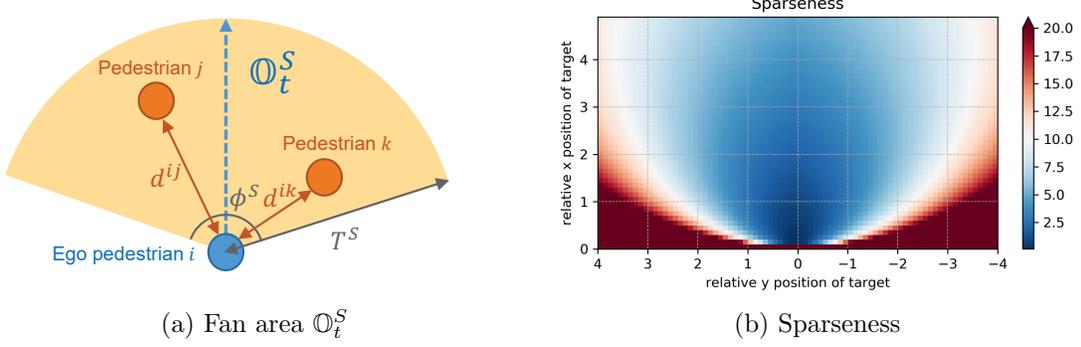


Figure 2.3: Sparseness and the corresponding fan area. The sparseness is plotted based on the calibrated parameters shown in table 4.1.

where

$$S_t^i := \min \left\{ \frac{d_t^{ij}}{\mathbb{A}_{lin}(\phi_t^{ij}, \lambda^S)} \right\}, \forall j \in \mathbb{O}_t^S. \quad (2.12)$$

$\phi_t^{ij} := \phi_{\langle \vec{v}_t^i, \vec{n}_{ij} \rangle}$ is the angle between the ego pedestrian's walking direction and the direction to the target agent, pedestrian j . λ^S is the anisotropy parameter. \mathbb{O}_t^S defines a fan area centered at the ego pedestrian's walking direction with a radius of threshold T^S and a field of view ϕ^S , which are illustrated in figure 2.3a. The sparseness value of a pedestrian depends on the relative position of the target, as illustrated in figure 2.3b.

In this work, 6 parameters are defined to model the limits of the velocity and the acceleration. Specifically, they are maximum velocity limit $v_{max} = 2.5m/s$, normal velocity limit $v_{nor} = 1.7m/s$, dense velocity limit $v_{den} = 0.3m/s$, maximum acceleration limit $a_{max} = 5m/s^2$, normal acceleration limit $a_{nor} = 2.5m/s^2$, and dense acceleration limit $a_{den} = 0.68m/s^2$. These values were determined according to the statistical findings in existing studies [6][31][36][30]. Both $v_{lim,t}^i$ and $a_{lim,t}^i$ are designed such that when the sparseness is small, pedestrians are restricted to small velocity and acceleration, but when the vehicle influence is large, the limits increase based on certain sparseness level. The limits never exceed the maximum values v_{max} and a_{max} . The relationships are expressed as:

$$v_{lim,t}^i(\vec{F}^{i,veh}, S_t^i) = \min(\beta_v^S \cdot \max(S_t^i - S_v^0, 0), v_{nor} - v_{den}) + v_{den} \\ + \min(\beta_v^F \cdot \max(|\vec{F}^{i,veh}| - F_v^0, 0), v_{max} - v_{nor}) \quad (2.13)$$

$$a_{lim,t}^i(\vec{F}^{i,veh}, S_t^i) = \min(\beta_a^S \cdot \max(S_t^i - S_a^0, 0), a_{nor} - a_{den}) + a_{den} \\ + \min(\beta_a^F \cdot \max(|\vec{F}^{i,veh}| - F_a^0, 0), a_{max} - a_{nor}), \quad (2.14)$$

where $\beta_v^S, S_v^0, \beta_v^F, F_v^0, \beta_a^S, S_a^0, \beta_a^F$, and F_a^0 are parameters that adjust the characteristics of the relationship. Figure 2.4 gives an example of the constraints on the velocity and the acceleration.

2.3.3 Surrounding Pedestrian Effect

The pedestrian-pedestrian interaction force accounts for the effect of all surrounding pedestrians. Therefore, for the ego pedestrian i , the total interaction force is the summation of

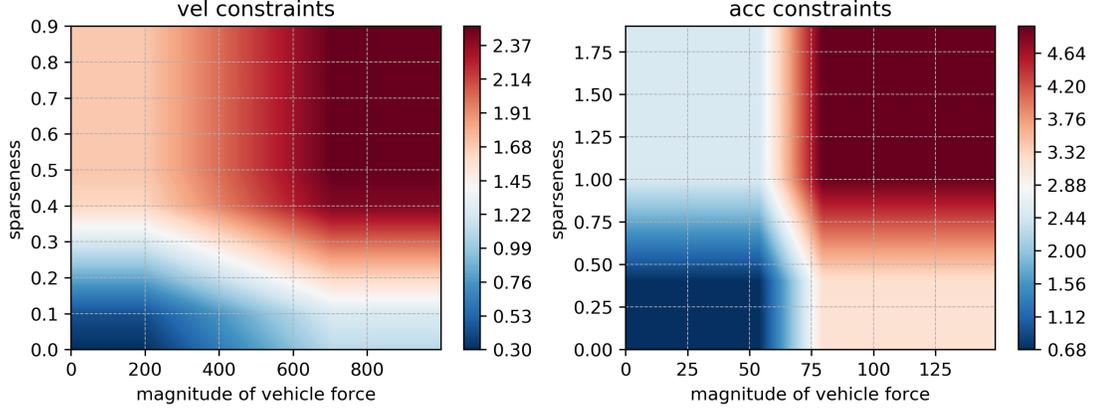


Figure 2.4: The constraints on the acceleration and velocity, which are plotted based on the calibrated parameters shown in table 4.1.

every individual interaction force. An individual interaction force is further divided into physical collision force $\vec{F}_t^{ij,col}$ and virtual interaction force $\vec{F}_t^{ij,vir}$:

$$\vec{F}_t^{i,ped} = \sum_{j \in \mathbb{Q}(i)} \left(\vec{F}_t^{ij,col} + \vec{F}_t^{ij,vir} \right), \quad (2.15)$$

where $j \in \mathbb{Q}(i)$ denotes the indexes of surrounding pedestrians that belong to ego pedestrian i .

Physical Collision Force

Physical collision force is effective only when the distance between two pedestrians is very close or the physical collision happens. This force allows the ego pedestrian to push the target pedestrian, especially in emergent situations such as a vehicle approaching in a dangerous way. It also describes extremely crowded situations as studied in [21]. The collision force is expressed as:

$$\vec{F}_t^{ij,col} = -\alpha^{col} \cdot \min\{d_t^{ij}, 0\} \cdot \vec{n}_t^{ij}, \quad (2.16)$$

where \vec{n}_t^{ij} is the unit vector pointing from the ego pedestrian i to the target pedestrian j , α^{col} is the parameter. The collision force is effective when the boundary distance d_t^{ij} is negative.

Virtual Interaction Force

Virtual interaction force makes the ego pedestrian to keep a certain 'social' distance to the target pedestrian. This is achieved by the combination of a repulsive force and a navigational force. Both of them are formulated as forces of which the magnitude and the direction calculated based on the temporal-spatial relationship between the ego and the target:

$$\vec{F}_t^{ij,vir} = \vec{F}_t^{ij,rep} + \vec{F}_t^{ij,nav}, \quad (2.17)$$

where

$$\vec{F}_t^{ij,rep} = -f_{lm}(d_t^{ij}, d_0^{rep}, M^{rep}, \sigma^{rep}) \cdot \mathbb{A}_{sin}(\phi_t^{ij}, \lambda^{rep}) \cdot \vec{n}_t^{ij} \quad (2.18)$$

$$\vec{F}_t^{ij,nav} = f_{lm}(d_t^{ij}, d_0^{nav}, M^{nav}, \sigma^{nav}) \cdot \mathbb{A}_{exp}(\phi_{v,t}^{ij}, \lambda^{nav}) \cdot \vec{n}_{\perp,t}^{ij}. \quad (2.19)$$

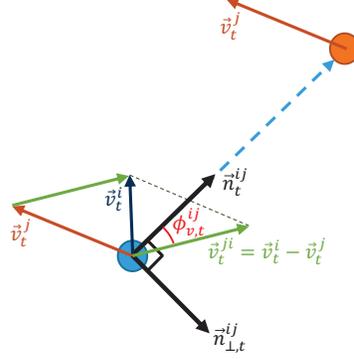


Figure 2.5: Illustration of navigational force. Blue circle indicates the ego pedestrian, red circle indicates the target pedestrian. The relative velocity (in blue color) represents the relative motion from the ego pedestrian to the target pedestrian.

As illustrated in figure 2.5, $\phi_{v,t}^{ij} := \phi_{\langle \vec{v}_t^{ij}, \vec{n}_t^{ij} \rangle}$ is the angle between the direction from ego pedestrian to the target pedestrian, \vec{n}_t^{ij} , and the direction of the relative velocity from pedestrian i to pedestrian j (in terms of pedestrian j 's coordinate), $\vec{v}_t^{ji} = \vec{v}_t^i - \vec{v}_t^j$. The unit vector $\vec{n}_{\perp,t}^{ij}$ is perpendicular to \vec{n}_t^{ij} . Here $\vec{n}_{\perp,t}^{ij}$ has two options (left or right side), depending on which side the $\phi_{v,t}^{ij}$ is.

The repulsion force purely considers the distance between the ego pedestrian and the target pedestrian, which can be interpreted as a social rule of giving enough personal space when walking in a crowd. The linear decaying function with smoothing is applied with the assumption that the repulsion force is more of a linear relationship.

The navigation force is primarily designed for anticipating and avoiding potential collision. It considers both the positions and the velocities of two interacting pedestrians. The navigation is dependent on the relative motion (velocity) between the two interacting pedestrians. If the relative motion indicates that the ego pedestrian is moving toward the target pedestrian ($\phi_{v,t}^{ij}$ is around zero), then a navigation force with the direction of $\vec{n}_{\perp,t}^{ij}$ is generated to avoid the potential collision. The magnitude is calculated based on the distance. An anisotropy is applied to reduce the magnitude when the possibility of collision is becoming small due to a large $\phi_{v,t}^{ij}$.

2.3.4 Destination Force

Destination force assumes that each pedestrian has a desired walking speed in mind. The pedestrian tries to keep the desired speed as much as possible by generating the destination force:

$$\vec{F}_t^{i,des} = \beta^{des} \cdot k^{des} \cdot (\vec{v}_t^i - \vec{v}_t^{i,d}), \quad (2.20)$$

where k^{des} is a parameter that can be viewed as feedback gain for the destination force. The desired speed $\vec{v}_t^{i,d}$ is always pointing from the ego pedestrian to the destination and is updated at every time step.

$$\vec{v}_t^{i,d} = v_0^i \cdot \frac{\vec{x}_t^{i,des} - \vec{x}_t^i}{\sqrt{|\vec{x}_t^{i,des} - \vec{x}_t^i|^2 + (\sigma^{des})^2}}, \quad (2.21)$$

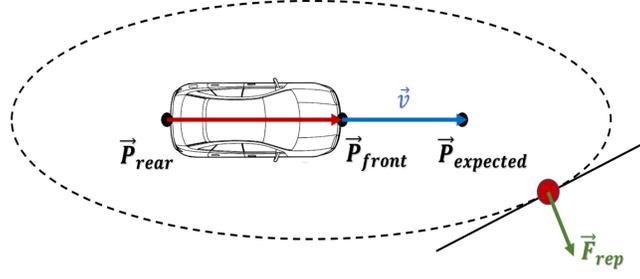


Figure 2.6: Illustration of the elliptic vehicle effect.

where the parameter v_0^i represents the desired speed and the parameter σ^{des} reduces the magnitude of desired speed as the pedestrian reaches the destination [12]. $\beta^{des} \in [0, 1]$ in equation 2.20 is a function of $|\vec{F}_t^{i,veh}|$ that adjusts the destination force when the effect of vehicle becomes large:

$$\beta^{des} = \max\{\min\{\frac{1}{F_2 - F_1} \cdot (|\vec{F}_t^{i,veh}| - F_2), 1\}, 0\}. \quad (2.22)$$

It decreases from 1 to 0 with two parameters F_1, F_2 as thresholds. This design allows the pedestrian to switch from reaching the destination to avoiding the collision with vehicle, which is the case in realistic situations.

It is necessary for the destination force to have a destination or temporal goal. In real-time application, it is estimated based on the historical pedestrian trajectory and the scenario layout. If the model is being evaluated based on recorded pedestrian trajectory data, the destination is usually available or can be estimated based on the entire recorded trajectory. If the model is used for the simulation of self-designed scenarios, the destination is usually pre-defined together with the scenario.

2.3.5 Vehicle Effect

Elliptic Effect

In the first version of our vehicle effect design, the vehicle effect is modeled as an elliptic shape. Fig. 2.6 shows the repulsive effect $\vec{F}_t^{i,veh}$ from the vehicle to the pedestrian. To create such an ellipse, we need two foci and a point on the boundary of the ellipse. The two points are defined as the rear point of the vehicle \vec{P}_{rear} , and the expected front point $\vec{P}_{expected}$. The distance between $\vec{P}_{expected}$ to \vec{P}_{front} is calculated based on the vehicle's current longitudinal velocity, so that the faster the vehicle, the longer the distance between two foci. The position of the pedestrian is used as the point on the boundary of the ellipse. The vehicle effect force is calculated by the following equation:

$$\vec{F}_t^{i,veh} = k_1 \exp(-k_2 b) \vec{n}, \quad (2.23)$$

where b is the length of semi-minor axis of the ellipse, k_1, k_2 are parameters, and \vec{n} is the force direction which is perpendicular to the tangent line of the ellipse at the position of the pedestrian. The detailed modeling and the simulation results that correspond to the elliptic vehicle effect can be found in [56].

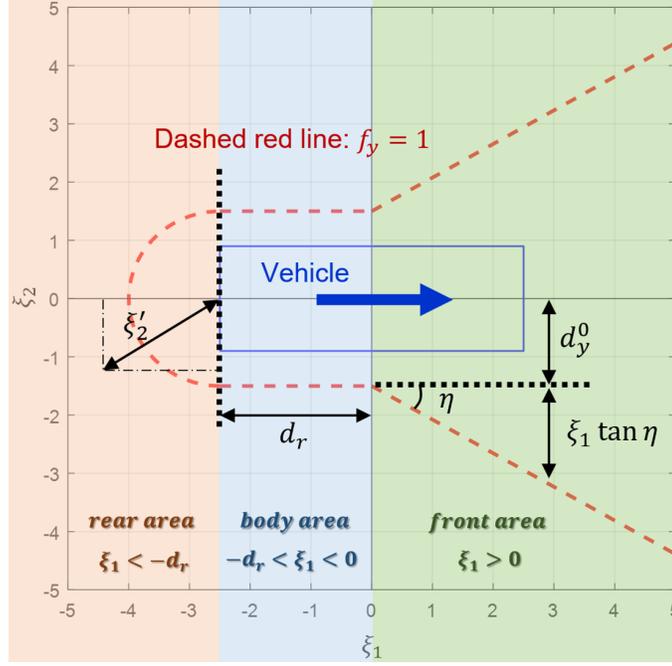


Figure 2.7: Illustration of the vehicle influence modeling

Three-area Effect

The second version of the vehicle effect divides the surrounding space of the vehicle into 3 areas: front area, body area, and rear area.

Considering the local coordinates with the origin at the center of vehicle and assuming that the vehicle is heading to the positive x-axis. The vehicle influence is described as the multiplication of longitudinal effect f_x , lateral effect f_y , and direction \vec{n} :

$$\vec{F}_t^{i,veh} = f_x \cdot f_y \cdot \vec{n}. \quad (2.24)$$

Let $\xi = (\xi_1, \xi_2)^T$ be the position of a pedestrian in the local coordinates. Different ξ characterize different areas: the front area ($\xi_1 > 0$), the body area ($-d_r < \xi_1 < 0$), and the rear area ($\xi_1 < -d_r$), as shown in figure 2.7.

In the *front area*, an outward angle η defines the major influence region, which is enclosed by the red dashed line in figure 2.7. Both f_x and f_y are modeled as follows:

$$f_x = \frac{1}{2d_x} \left(-(\xi_1 - d_x) + \sqrt{(\xi_1 - d_x)^2} \right) \quad (2.25)$$

$$f_y = A_y \cdot \exp(-b_y(|\xi_2| - d_y)) \quad (2.26)$$

$$d_x = d_x^0 + \alpha \cdot |v_v| \quad (2.27)$$

$$d_y = d_y^0 + \xi_1 \tan \eta \quad (2.28)$$

$$\vec{n} = \begin{bmatrix} \cos(\text{sign}(\xi_2) \cdot \zeta(\xi_1)) \\ \sin(\text{sign}(\xi_2) \cdot \zeta(\xi_1)) \end{bmatrix} \quad (2.29)$$

$$\zeta = \begin{cases} \frac{\pi}{2} - \Delta\zeta_m \cdot \frac{d_x - d_m - \xi_1}{d_x - d_m} & , \forall \xi_1 > d_m \\ \frac{\pi}{2} - \Delta\zeta_m \cdot \xi_1 & , \forall \xi_1 < d_m. \end{cases} \quad (2.30)$$

d_x is the look-ahead distance where the longitudinal effect vanishes, which is a function of the vehicle velocity $|v_v|$ with parameters d_x^0 and α , as shown in figure 2.8. f_y decreases

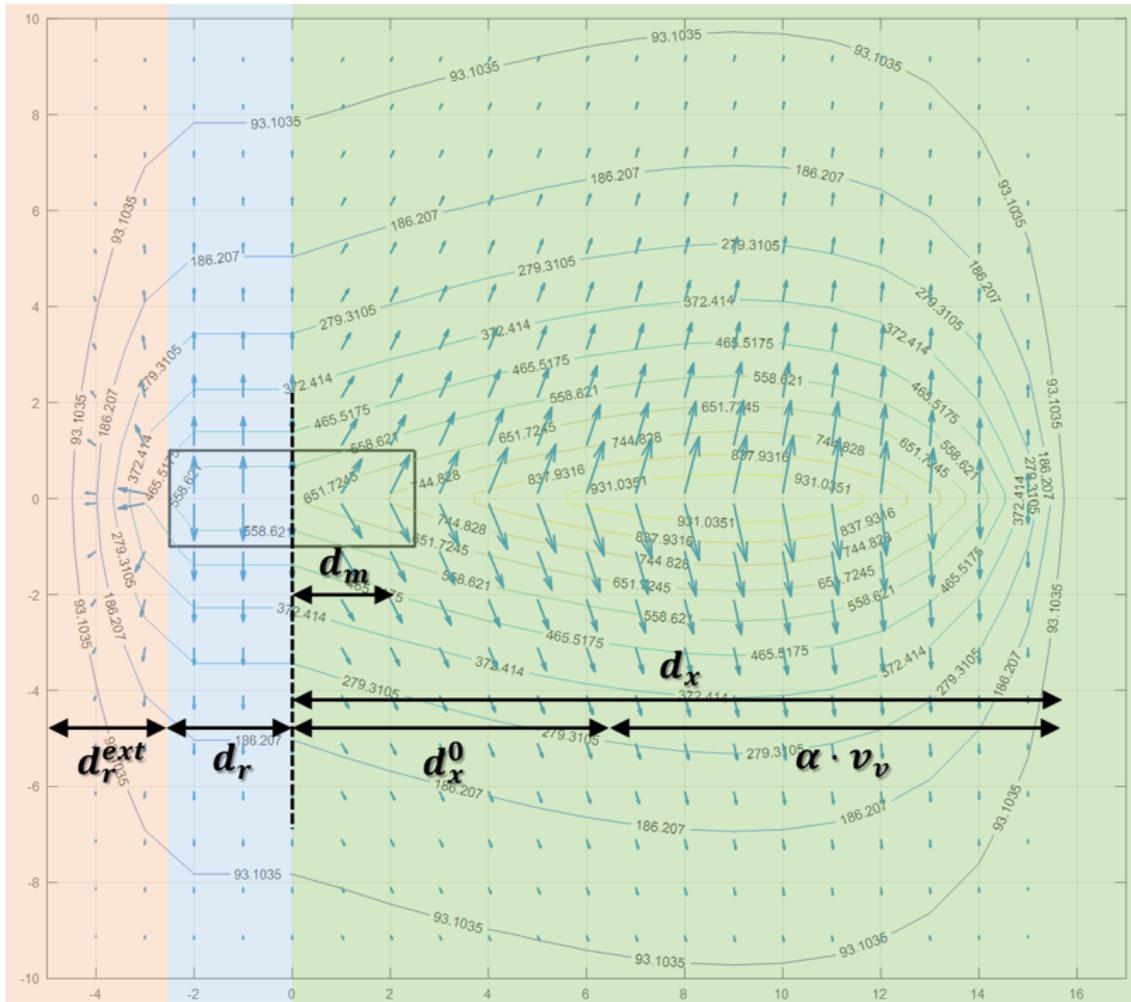


Figure 2.8: The magnitude (contours and arrow lengths) and direction (arrow directions) of vehicle influence with the modeling parameters

along the lateral direction. d_y is the distance where the lateral effect is constant ($f_y = A_y$). ζ is a small angle that changes the force direction \vec{n} , which is dependent on ξ_1 with parameters $\Delta\zeta_m$ and d_m . The change of force directions can be visualized in figure 2.8.

In the *body area*, $f_x = 1$ which means the longitudinal effect does not change. f_y decreases as the lateral distance increases. \vec{n} points to the lateral direction.

$$f_x = 1 \quad (2.31)$$

$$f_y = A_y \cdot \exp(-b_y(|\xi_2| - d_y^0)) \quad (2.32)$$

$$\vec{n} = \begin{bmatrix} \cos(\text{sign}(\xi_2) \cdot \frac{\pi}{2}) \\ \sin(\text{sign}(\xi_2) \cdot \frac{\pi}{2}) \end{bmatrix} \quad (2.33)$$

In the *rear area*, f_x is defined in a similar way as in the front area, with d_x being the vanishing point. f_y now decreases as the distance from the rear-center to the pedestrian position increases. \vec{n} points from the rear-center to the pedestrian.

$$f_x = \frac{1}{2d_r^{ext}} \left(\xi_1 + d_x + \sqrt{(\xi_1 + d_x)^2} \right) \quad (2.34)$$

$$f_y = A_y \cdot \exp(-b_y(|\xi'_2| - d_y^0)) \quad (2.35)$$

$$d_x = d_r + d_r^{ext} \quad (2.36)$$

$$\xi'_2 = \sqrt{(\xi_1 + d_r)^2 + \xi_2^2} \quad (2.37)$$

$$\vec{n} = \begin{bmatrix} \cos(\xi_1 + d_r) \\ \sin(\xi_2) \end{bmatrix} \quad (2.38)$$

Fig. 2.8 shows the contour plot as well as the force vectors with the proposed vehicle effect design. Fig. 2.9 provides a surface plot of the magnitude of the vehicle effect. All the modeling detail the the corresponding results can be found in [59].

Anisotropy-based Effect

In the third/latest version, we proposed a new design of vehicle-pedestrian interaction force. It considers the spatial-temporal relationship (relative positions and velocities) between the ego pedestrian and the vehicle. The vehicle's size and shape, as well as anisotropy, are also considered. In other words, different pedestrian orientations and velocities and different vehicle orientations of velocities create different combinations of vehicle-pedestrian interaction.

A virtual contour of the vehicle is defined as an extension of the vehicle actual contour and the consequence of slow motion. The virtual contour is illustrated in figure 2.10. An extension length l_e is added on based on original contour of the vehicle with length $l_r + l_f$ and width l_w . This can be conceptualized as a minimum distance or a buffer that the ego pedestrian wants to keep from the vehicle. d_x^0 is an extended length along the vehicle's orientation, which assumes that the pedestrian wants to keep a larger distance from the front bumper than from the rear bumper. $\alpha_x \cdot u_{veh}$ is another extended length along vehicle's moving direction, which is proportional to the vehicle longitudinal speed u_{veh} with the parameter α_x . The faster the vehicle, the longer the extension in front of the vehicle.

Once the virtual contour is available, an influential point P_t^{iv} on the contour is determined by finding the minimum distance d_t^{iv} from the ego pedestrian to the contour. Then,

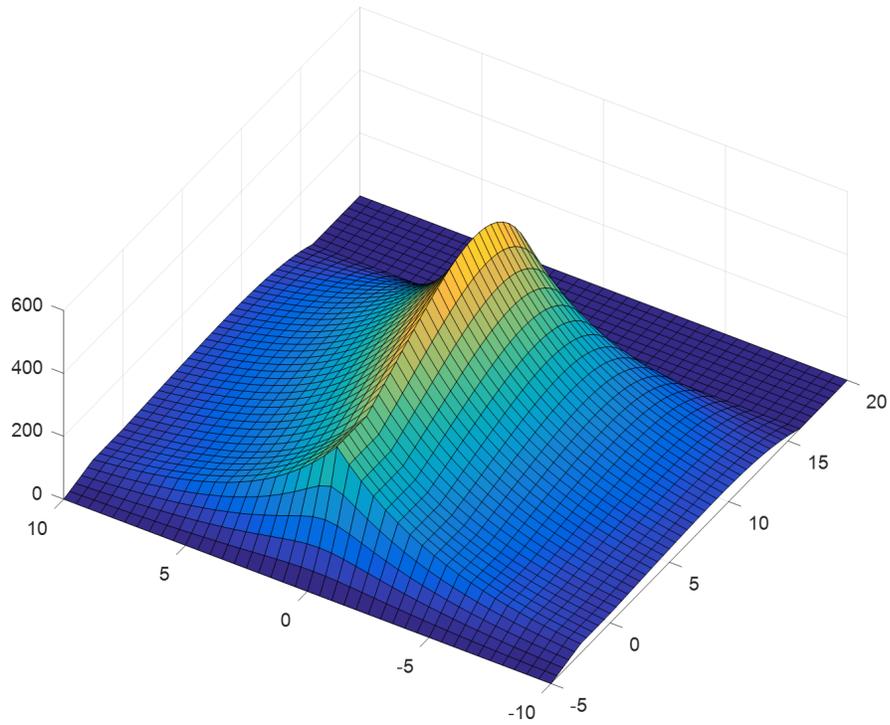


Figure 2.9: The surface plot of the vehicle influence magnitude

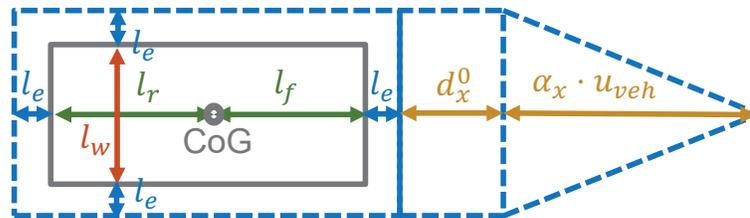


Figure 2.10: Vehicle's virtual contour (blue dashed line). A surrounding pedestrian needs to find the influential point (closest point to the virtual contour), and then calculates the vehicle-pedestrian interaction force.

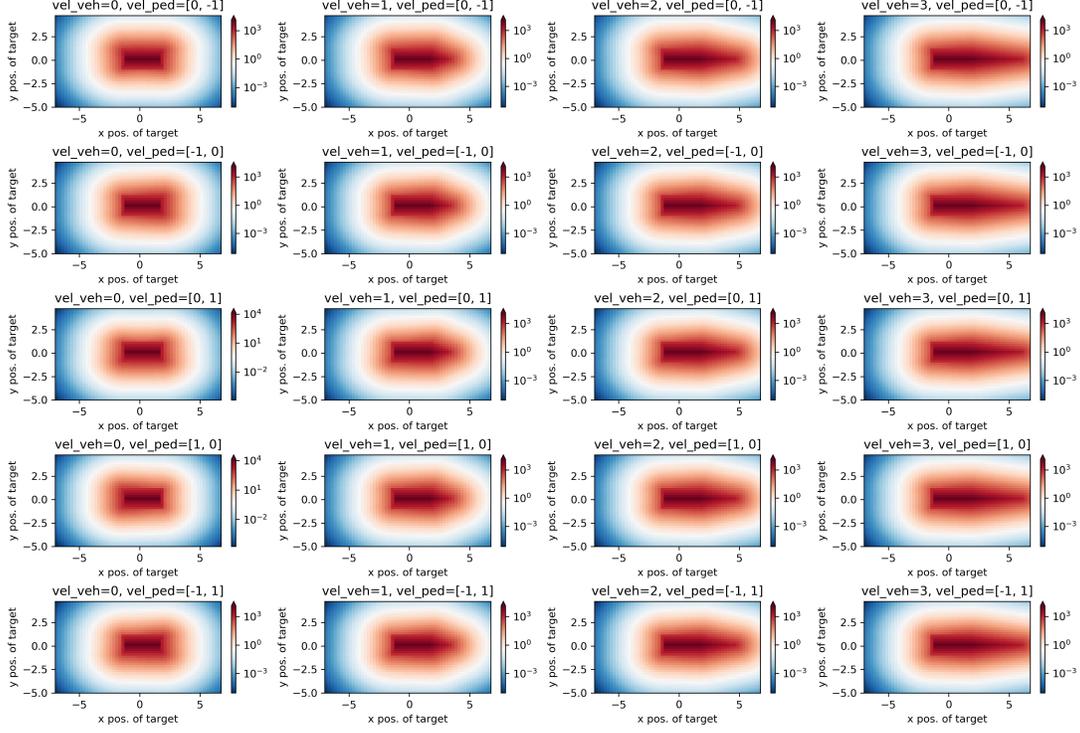


Figure 2.11: The magnitudes of vehicle influence under different conditions. Columns from left to right: vehicle longitudinal speed increases (0m/s, 1m/s, 2m/s, and 3m/s). Rows from top to bottom: pedestrian walking direction varies (with velocity vector of $[0, -1]$, $[-1, 0]$, $[0, 1]$, $[1, 0]$, and $[-1, 1]$ in Euclidean coordinate).

the vehicle-pedestrian interaction force is calculated by the following equation:

$$\vec{F}_t^{i,veh} = f_{exp}(d_t^{iv}, A^{veh}, b^{veh}) \cdot \mathbb{A}_{sin}(\phi_t^{iv}, \lambda^{veh}) \cdot \vec{n}_t^{vi}. \quad (2.39)$$

The direction of the vehicle-pedestrian interaction force is determined by \vec{n}_t^{vi} , which is a unit vector pointing from the influential point to the ego pedestrian. The magnitude applies an exponentially decaying function with parameters A^{veh} and b^{veh} , because as mentioned in the previous section, the pedestrian should be quickly driven away from the vehicle once getting close to the virtual contour. Anisotropy is considered in which $\phi_t^{iv} := \phi_{\langle -\vec{n}_t^{vi}, \vec{v}_t^i \rangle}$ and λ^{veh} is the parameter of the anisotropy. The anisotropy adjusts the magnitude of the vehicle influence based on the walking direction of the pedestrian with respect to the vehicle moving direction. For example, a pedestrian walking away from the vehicle should have less vehicle influence than a pedestrian walking toward the vehicle.

Figure 2.11 are heat maps that illustrate the magnitudes of vehicle-pedestrian interaction force in different situations. It compares different combinations of longitudinal vehicle speeds and different pedestrian walking directions. As vehicle speed increases, the influence area expands (primarily in the vehicle moving direction). Also, notice that the difference among the influence areas when the pedestrian walking direction changes (different rows).

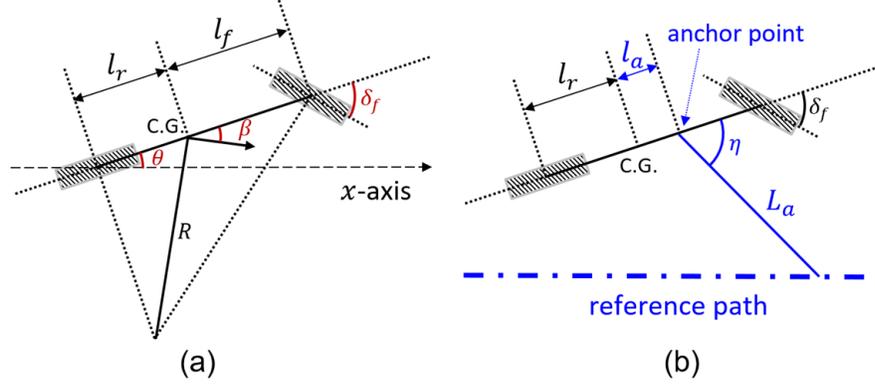


Figure 2.12: (a) Kinematic bicycle model; (b) Lane-following controller.

2.3.6 Vehicle Motion

We believe that vehicles should be considered more of mechanical systems instead of point-mass dynamics similar to pedestrians, because one major purpose of developing the proposed pedestrian motion model is to provide more information for automated systems on vehicles. Specifically, a common procedure (path planning, trajectory generation, and trajectory following) of automated driving is applied. The vehicle motion is generated by a kinematic bicycle model [40] with a pure-pursuit path tracking controller [28] that tracks a reference path.

The kinematic bicycle model has three assumptions: (a) the vehicle has planner motion; (b) both left and right wheels steer the same angle; and (c) there is no slip at both front and rear tires. It is applicable when the vehicle moves at relatively low speed, which is the case for most scenarios in this study. As illustrated in figure 2.12-(a), let $x_v \in \mathbb{R}^2$ and $\theta_v \in \mathbb{R}$ be the position and the orientation of the vehicle, the model can be described as:

$$\begin{aligned}
 \dot{x}_v^1 &= v_v \cos(\theta_v + \beta_v) \\
 \dot{x}_v^2 &= v_v \sin(\theta_v + \beta_v) \\
 \dot{v}_v &= f(u) \\
 \dot{\theta}_v &= \frac{v_v}{l_r} \sin \beta_v \\
 \beta_v &= \arctan \left(\frac{l_r}{l_f + l_r} \tan \delta_f \right)
 \end{aligned} \tag{2.40}$$

where v_v is the longitudinal speed, β_v is the velocity angle with respect to the vehicle center of gravity (C.G.), l_f , l_r are the distances from C.G. to the front wheel and the rear wheel, respectively, u is the longitudinal control action (brake/gas), and δ_f is the lateral control action (steering angle of the front wheel).

The pure-pursuit path tracking controller calculates the steering angle command δ_f to track the reference path:

$$\delta_f = -\tan^{-1} \left(\frac{(l_f + l_r) \sin \eta_v}{\frac{L_a}{2} + l_a \cos \eta_v} \right). \tag{2.41}$$

A proportional controller generates the acceleration/brake command to track the desired longitudinal speed. The reference paths are usually pre-defined for the simulation of self-designed scenarios. This allows us to test the performance of the proposed pedestrian

model in various vehicle-pedestrian interaction patterns. Note that in the process of calibrating model parameters (see chapter 4.2.3), ground truth vehicle trajectory is directly applied, because the primary focus in this work is pedestrian motion.

2.4 Conclusion

This chapter presents the design of the social force based model that can describe the pedestrian motion under vehicle influence. Three different versions of vehicle effect were discussed in this chapter. The last version proposed a straightforward yet efficient vehicle effect design. We tested and validated the proposed model based on several fundamental vehicle-pedestrian interaction scenarios. We first conducted qualitative evaluation by simulating these scenarios and visually inspecting the behavior of all pedestrians. Then, the model parameters were further calibrated using collected pedestrian trajectory data in the same scenarios. These will be presented in the next two chapters.

Chapter 3

Dataset

3.1 Background

This vehicle-crowd interaction (VCI) scenario has been drawing attention in recent years. Specific models [62][3][56][7] have been designed to describe the individual motion of a crowd in some specific situations where both interpersonal and vehicle-pedestrian interaction were differently considered. To either calibrate or train such models above and further evaluate their performance, providing ground truth trajectories of VCI is becoming increasingly important. However, to the best of authors' knowledge, there is no public dataset that covers VCI, especially in scenarios where interpersonal interaction is not negligible. To fill this gap, we built two VCI datasets. One (CITR dataset) focuses on fundamental VCI scenarios in controlled experiments, and the other (DUT dataset) consists of natural VCIs in crowded university campus.

In general, the approaches for modeling pedestrian motion in crowd can be classified in two categories. Traditionally, a rule-based model, e.g., social force models [61], is designed based on human experience and the parameters of the model are then calibrated by using ground truth pedestrian trajectories [62][9]. Recently, with the growing popularity of deep learning, long-short term memory (LSTM) networks have been applied to model this pedestrian motion [1][43] in the hope of taking advantage of the potential in deep neural networks, which heavily relies on pedestrian trajectory data. The requirement of ground truth pedestrian trajectories in both approaches confirmed the necessity of building more pedestrian/crowd trajectory dataset, especially in scenarios that have not been covered in existing ones. Existing dataset such as ETH [42] and UCY [29] only covers interpersonal interaction, which is not suitable for VCI. Stanford Drone Dataset [45] includes some vehicle trajectories, but the number of surrounding pedestrians is small so that there is little interpersonal interaction. This work aims to provide a new type of pedestrian trajectory dataset that can enrich the existing datasets, and meanwhile assists in solving pedestrian safety related problems in the areas of intelligent vehicles and intelligent transportation systems.

Unlike pure interpersonal interaction, VCI introduces more complexity. This complexity can be decomposed by separating vehicle influence from interpersonal influence and by identifying different types of vehicle influence on pedestrians. To this end, in our CITR dataset, controlled experiments were designed and conducted in a way that from interpersonal interaction scenarios to VCI scenarios, they can be pairwise compared so that separate effect, for example, the existence (or not) of a vehicle or the walking direction of

the crowd, can be identified and analyzed.

Some pedestrian motion models may consider personal characteristics, i.e., each pedestrian applies a model with a unique parameter set. CITR dataset provides such personality by assigning the same pedestrian always the same ID, hence more options are provided to researchers.

To supplement each other, in DUT dataset natural VCI data was constructed from a series of recordings of crowded university campus. A down-facing camera attached to a drone hovering above and far away from the ground was used as the recording equipment. Therefore, both the crowd and the vehicle are unaware of being observed, hence producing natural behavior. The DUT dataset can be used for final verification of VCI models or some end-to-end VCI modeling design.

Both CITR and DUT datasets applied a hovering drone as the recording equipment. This ensured the accuracy of the extracted trajectories by avoiding the issue of occlusion, a major deficiency if pedestrians are detected from the view of sensors mounted on moving vehicles or buildings.

The trajectories of individual pedestrians and vehicles were extracted by image processing techniques. Due to the unavoidable instability of the camera attached to a hovering drone (even with a gimbal system), the recorded videos were stabilized before further processing. A robust tracking algorithm (CSRT[32]) was then applied to automatically track pedestrians and vehicles, although the initial positions still have to be manually selected. In the last step, different Kalman filters were applied to further refine the trajectories of both pedestrians and vehicles. This design avoided tedious manual annotation as done in the ETH and UCY dataset [42][29], and possible imprecision of the tracking as done in the Stanford dataset [45].

In general, the contribution of the study can be summarized as follows:

- We built a new pedestrian trajectory dataset that covers both interpersonal interaction and vehicle-crowd interaction.
- The dataset includes two portions. One comes from controlled experiments, in which fundamental VCIs are covered and each person has a unique ID. The other comes from crowded university campus scenarios where the pedestrian reaction to a vehicle is completely natural.
- The application of a drone camera for video recording, a new design of tracking strategy, and the Kalman filters for refining trajectories made the extracted trajectories as accurate as possible.

In the rest of the chapter, section 3.2 reviews related dataset regarding pedestrian motion and vehicle-pedestrian interaction. Section 3.3 details the configuration of both CITR and DUT dataset. Section 3.4 describes the algorithm applied for trajectory extraction and the Kalman filters used for trajectory refinement. Section 3.5 shows some statistics of our dataset. Section 3.6 concludes this chapter and discusses possible improvement.

Table 3.1: Comparison with existing world coordinate based pedestrian trajectory dataset

Dataset Name	Scenarios	Pedestrian Density	Other Participants	Method of Annotation	FPS of Annotation	Amount of Trajectories	Camera Depression Angle (degrees)	From Pixel to World Coordinate	Video Resolution
ETH	campus, urban street	medium	no	manual	2.5	650	about 70-80	matrix file	720x576
UCY	campus, park, urban street	high, low	no	manual	interpolated	409	about 20-50	partially measurable	720x576
Stanford	campus	medium, low	cyclist, bus, golf cart, car	tracking + interpolation	29.97	3297	90	n.a.	595x326
CITR	specifically designed	medium	golf cart	CSRT tracker + initial annotation	29.97	340	90	measured	1920x1080
DUT	campus	high, low, medium	car	CSRT tracker + initial annotation	23.98	1793	90	measured	1920x1080
Town Center	urban street	medium	no	manual + tracking verification	25	2200	about 25-35	n.a.	1920x1080
Train Station	train station hall	high, medium	no	KLT keypoint tracker	varied	47866	about 40-50	n.a.	720x480

3.2 Related Works

Pedestrian dataset can be in general divided into two categories: world coordinate (WC) based dataset and vehicle coordinate (VC) based dataset. WC based dataset is usually applied to studies that need to consider interpersonal interaction, because the collective motion of pedestrians is clear, accurate enough, and easily accessible, while VC based dataset doesn't contain enough instances of interpersonal interaction. Popular WC based dataset includes UCY Crowds-by-Example dataset [29], ETH BIWI Walking Pedestrians dataset [42], Town Center dataset [5], Train Station dataset [66] and Stanford Drone dataset [45]. They have been widely used for crowd motion analysis, risk detection, and the calibration/training of various rule-based and learning-based pedestrian motion models [4]. The proposed dataset in this study aims to enrich the WC based dataset by incorporating the vehicle-crowd interaction. A comparison among the proposed and existing WC based datasets are shown in table 3.1. VC based dataset is usually used for single/multiple, but not too many, pedestrian detection and/or intention estimation from a mono camera mounted in front of the vehicle. A couple of datasets such as Daimler Pedestrian Path Prediction dataset [47] and KITTI dataset [17] provide vehicle motion information, hence the trajectories of both the vehicle and pedestrians in world coordinate can be estimated by combining vehicle motion and video frames. The estimated trajectories can serve as ground truth data for vehicle-pedestrian interaction but with little interpersonal interaction due to the limited number of pedestrians.

Some existing datasets also apply a down-facing camera attached to a hovering drone as the recording equipment. For example, in Stanford Drone dataset [45], the utilization of drone eliminated occlusion so that all participants (pedestrians, cyclists, cars, carts, buses) were individually tracked. Another dataset HighD [27], which focuses on vehicle-vehicle interaction on highway driving, also successfully demonstrated the benefit of using the hovering drone to remove occlusion.

3.3 Dataset

3.3.1 CITR Dataset

The controlled experiments were conducted in a parking lot near the facility of Control and Intelligent Transportation Research (CITR) Lab at The Ohio State University (OSU). Figure 3.1 shows the layout of the experiment area. A DJI Phantom 3 SE Drone with a down-facing camera on a gimbal system was used as the recording equipment. The video resolution is 1920×1080 with an fps of 29.97. Participants are the members of CITR Lab at OSU. During the experiments, they were instructed only to walk from one small area (starting points) to another small area (destinations). The employed vehicle was an EZ-GO Golf Cart, as shown in figure 3.2. 3 markers were put on top of the vehicle to help vehicle motion tracking, of which the vehicle position is calculated by geometry. The reason of using 3 markers is to reduce the tracking noise as much as possible.

The designed fundamental scenarios were generally divided into 6 groups, as shown in figure 3.3. They were designed such that by comparing pedestrian-only scenarios (pure interpersonal interaction) and VCI scenarios, the vehicle influence can be separated and analyzed. Therefore, except for the difference due to the existence (or not) of a vehicle, all other factors remain the same such as pedestrians' intention (starting point and destination), pedestrians' identity (who are these pedestrians), and environment layout

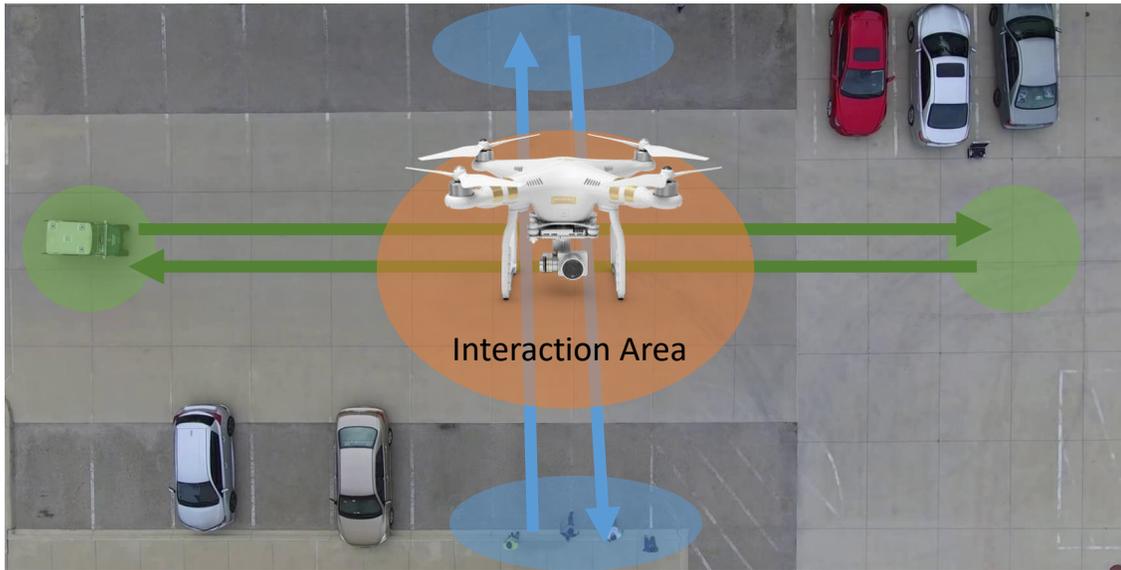


Figure 3.1: Layout of the controlled experiment area (a parking lot near CITR Lab at OSU). The vehicle (a golf cart) moves back and forth between two blue areas. Pedestrians move back and forth between two green areas. The interaction happens in the orange area, which is also the central area of the recording.

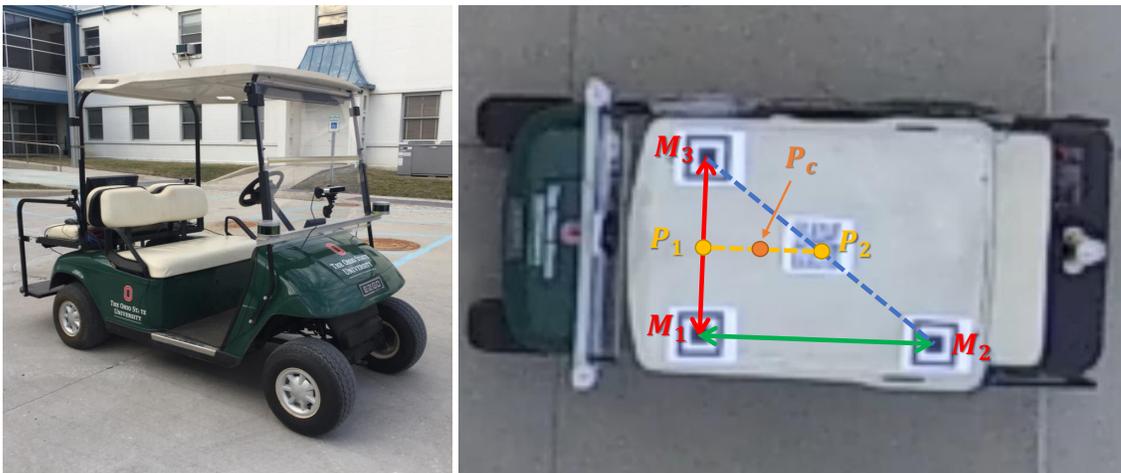


Figure 3.2: EZ-GO Golf cart employed in the experiments (left) and makers on top of the vehicle (right). In the vehicle tracking process, 3 markers (M_1, M_2, M_3) were continuously being tracked. By geometry, P_1, P_2 were calculated and recorded for vehicle orientation and P_c as the vehicle center position.

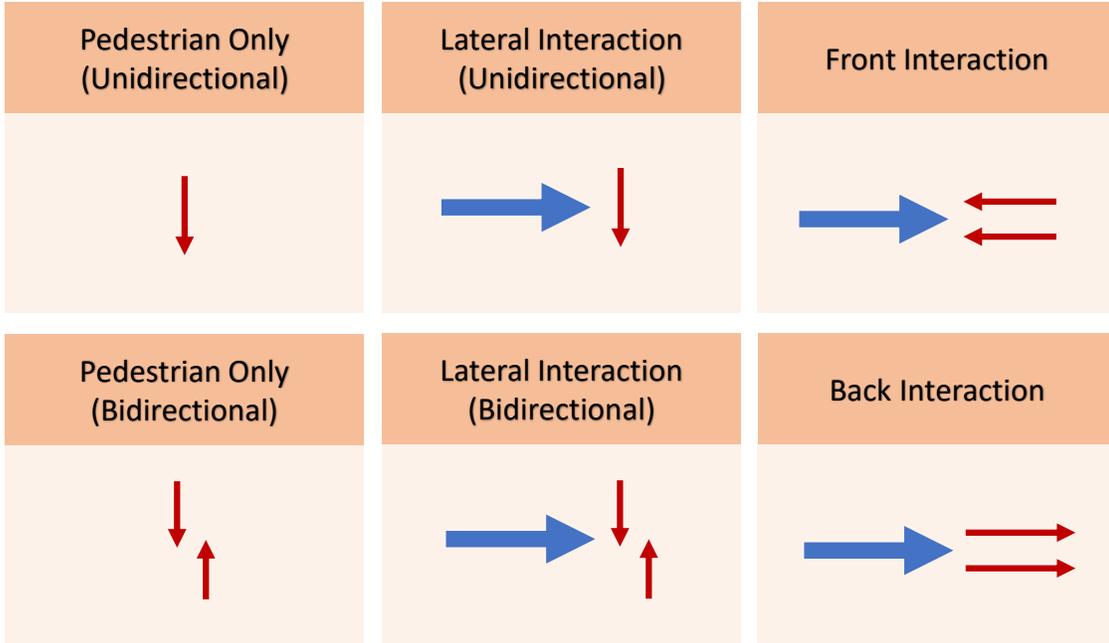


Figure 3.3: Designed scenarios of controlled experiments. Red arrows indicate the motion of pedestrians/crowd, while blue arrows indicate vehicle motion.

Table 3.2: Number of clips in each scenario of CITR dataset

Scenarios	Num. of clips
Pedestrian only (unidirectional)	4
Pedestrian only (bidirectional)	8
Lateral interaction (unidirectional)	8
Lateral interaction (Bidirectional)	10
Front interaction	4
Back interaction	4

(location, time period, weather, etc.). The scenarios of front, back, and side interactions help exploring typical VCIs which could guide the design of pedestrian motion models.

After processing, there are 38 video clips in total, which include approximate 340 pedestrian trajectories. The detailed information is presented in table 3.2.

3.3.2 DUT Dataset

The DUT dataset was collected at two crowded locations in the campus of Dalian University of Technology (DUT) in China, as shown in figure 3.4. One location includes an area of pedestrian crosswalk at an intersection without traffic signals. When VCI happens, in general there is no priority for either pedestrians or vehicles. The other location is a relatively large shared space near a roundabout, in which pedestrians and vehicles can freely move. Similar to CITR dataset, a DJI Mavic Pro Drone with a down-facing camera was hovering above the interested area as the recording equipment, high enough to be unnoticed by pedestrians and vehicles. The video resolution is 1920×1080 with an fps of 23.98. Pedestrians are primarily made up of college students who just finished classes and



Figure 3.4: Locations of DUT dataset. Upper: an area of crosswalk at an intersection without traffic signals. Lower: a shared space near a roundabout.

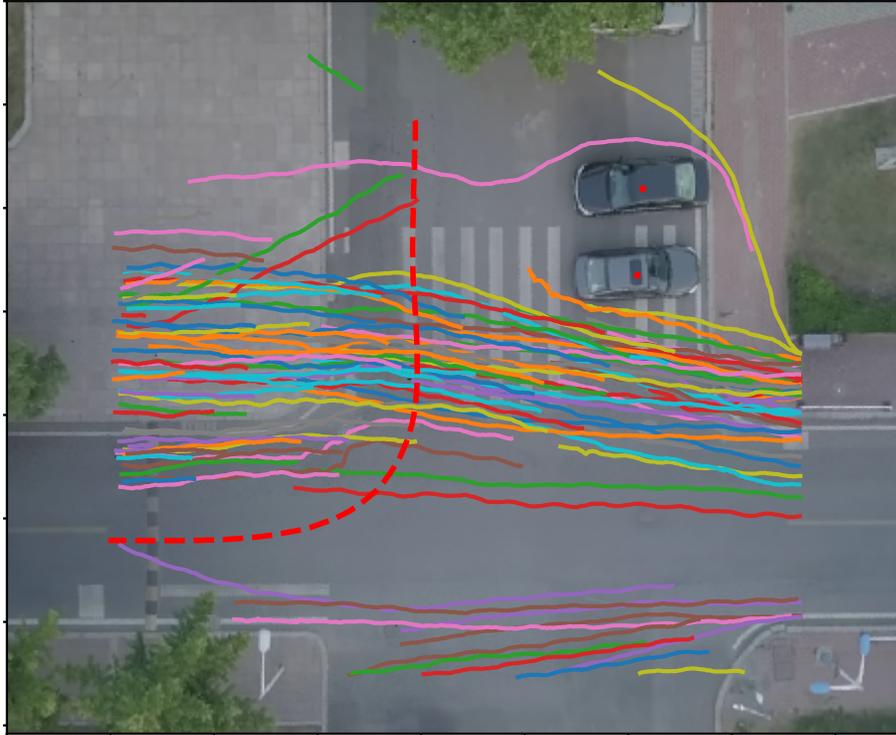


Figure 3.5: Trajectories of vehicles (red dashed line) and pedestrians (colorful solid lines) in a clip of the intersection scenario.

on their way out of classrooms. Vehicles are regular cars that go through the campus.

With this configuration, scenarios of DUT dataset consists of natural VCIs, in which the number of pedestrians varies hence introducing some variety of the VCI.

After processing, there are 17 clips of crosswalk scenarios and 11 clips of shared space scenarios, including 1793 trajectories. Some of the clips contains multiple VCIs, i.e., more than 2 vehicles interacting with pedestrians simultaneously, as in the lower picture in figure 3.4.

Figure 3.5 and 3.6 demonstrate the processed example trajectories of the DUT dataset.

3.4 Trajectory Extraction

Four procedures were done to extract the trajectories of both pedestrians and vehicles from the recorded top-view video.

3.4.1 Video Stabilization

First, the raw video was stabilized to remove the noise caused by unstable drone motion. This procedure applies several image processing techniques, which include scale-invariant feature transform (SIFT) algorithm for finding key-points, k-nearest neighbors (k-NN) for obtaining matches, and random sample consensus (RANSAC) for calculating perspective transformation between each video frame and the first video frame (reference frame). The detailed procedure is illustrated in algorithm 1.



Figure 3.6: Trajectories of vehicles (red dashed line) and pedestrians (colorful solid lines) in a clip of the shared space scenario.

Algorithm 1: Video Stabilization

Result: calibrated frames F_i^{cal}
 set 1st frame F_1 as reference F_{ref} ;
for each new frame $F_i, i = 2, 3, \dots$ **do**
 apply SIFT to find key-points in F_i and F_{ref} , separately;
 apply KNN to find matches;
 obtain good matches by removing matches that have long distance of pixel positions in F_i and F_{ref} ;
 apply RANSAC for the good matches to calculate the transformation matrix M_i from F_i to F_{ref} ;
 obtain F_i^{cal} by applying transformation M_i to F_i ;
end

3.4.2 Vehicle and Pedestrian Tracking

Once the video was stabilized, pedestrians and vehicles were automatically tracked by using Discriminative Correlation Filter with Channel and Spatial Reliability (CSR-DCF) [32]. In the tracking process, raw videos are partitioned into small clips, which contain separate and complete VCIs. Once pedestrians appear in the region of interest (ROI), the initial positions were manually given, hence initializing the trackers. When they exited the ROI, the trackers stopped. Due to the vehicle size, vehicle tracking was done by individually tracking either the 3 markers on top of the vehicle (CITR dataset) or four corners of vehicle (DUT dataset). Then, the vehicle position was calculated based on geometric relationship of these tracked points.

3.4.3 Coordinate Transformation

Pedestrian trajectories obtained in the previous step are in the coordinates of image pixels. A coordinate transformation operation is necessary to convert the trajectories from image pixels into actual scale in meters.

This can be done by either measuring the actual length of a relatively long reference line in the scene or measuring the distance between markers on top of the vehicle (if applicable). The assumption here is that, compared with the altitude of the hovering drone, the distance between the ground plane and the tracking plane (the plane of a pedestrian’s head or the vehicle’s top) is very small so that both planes can be treated as the same plane.

3.4.4 Trajectory Filtering

In the last step, Kalman filters [14] was applied to remove the noise and refine the trajectories. It is sufficient to use a linear Kalman filter with a point-mass model for pedestrian trajectories, in which the 2D velocity (in x and y axes) can be estimated. The state transition and measurement follows the equations:

$$\dot{x} = v + w_1 \tag{3.1}$$

$$\dot{v} = a + w_2 \tag{3.2}$$

$$y = x + v, \tag{3.3}$$

where position $x \in \mathbb{R}^2$ and velocity $v \in \mathbb{R}^2$ are the system state, $y \in \mathbb{R}^2$ is the measurement (recorded position), $w = [w_1^T, w_2^T]^T \sim N(0, Q)$ the state transition noise, and $v \sim N(0, R)$ the measurement noise.

When applying the Kalman filter, it is assumed that $a = 0$, which implies a constant velocity model.

Vehicle motion is somehow constrained, e.g., the lateral motion/velocity can not be abruptly changed. Therefore, an extended Kalman filter with a nonlinear kinematic bicycle

model was applied. The bicycle model follows:

$$\dot{x}_x = v \cos(\theta + \beta) + w_1 \quad (3.4)$$

$$\dot{x}_y = v \sin(\theta + \beta) + w_2 \quad (3.5)$$

$$\dot{\theta} = \frac{v}{l_r} \sin \beta + w_3 \quad (3.6)$$

$$\dot{v} = a + w_4 \quad (3.7)$$

$$\beta = \arctan \left(\frac{l_r}{l_f + l_r} \tan \delta_f \right) \quad (3.8)$$

$$y = [x_x, x_y]^T + v, \quad (3.9)$$

where x_x, x_y stands for the position, v is the longitudinal speed, β is the velocity angle with respect to the vehicle C.G., l_f, l_r are the distances from C.G. to the front wheel and the rear wheel, respectively, a is the longitudinal acceleration, δ_f is the steering angle of the front wheel, $w = [w_1, w_2, w_3, w_4]^T \sim N(0, Q)$ the state transition error, and $v \sim N(0, R)$ the measurement error.

At each step of the extended Kalman filter, the system is linearized at current state by calculating its Jacobian. It is assumed that both inputs $a = 0$ and $\delta_f = 0$.

3.5 Statistics

To give a more detailed description of the above dataset, the magnitude of pedestrian velocities (estimated by the Kalman filter) in all video clips were analyzed. The reason of analyzing velocity magnitude is that, pedestrian velocity is the most intuitive way of describing pedestrian motion, and, as argued in [4], if pedestrian trajectories are used to train neural network based pedestrian model, using pedestrian velocity (offset in motion at the next time step) is better than using absolute position, because different reference systems (how the global coordinates are defined) in different dataset usually cause incompleteness of training data.

Figure 3.7 and 3.8 show the distribution of the velocity magnitude for CITR dataset and DUT dataset, respectively. Table 3.3 presents the mean velocity magnitude and mean walking velocity magnitude. The walking velocity excludes the velocity magnitude that is less than $0.3m/s$, at which the pedestrian is considered as either standing or yielding to the vehicle instead of walking. The value of $0.3m/s$ was intuitively selected based on the shape of the histogram. It is obvious that, from the velocity distribution and the mean velocity results, the pedestrians in DUT dataset walk faster than the pedestrians in CITR dataset. The reason could be that, when conducting controlled experiments, as in the CITR dataset, pedestrians were more relaxed, while in the DUT dataset, pedestrians were in a little bit hurry because they just came out of classes. However, in general, the distribution and the mean velocity magnitude are in accordance with the preferred walking velocity in various situations [36].

3.6 Conclusion

Two dataset, experimentally designed CITR dataset and natural DUT dataset, were built in this study for pedestrian motion models that consider both interpersonal and vehicle-crowd interaction. The trajectories of pedestrians and vehicles were extracted by image

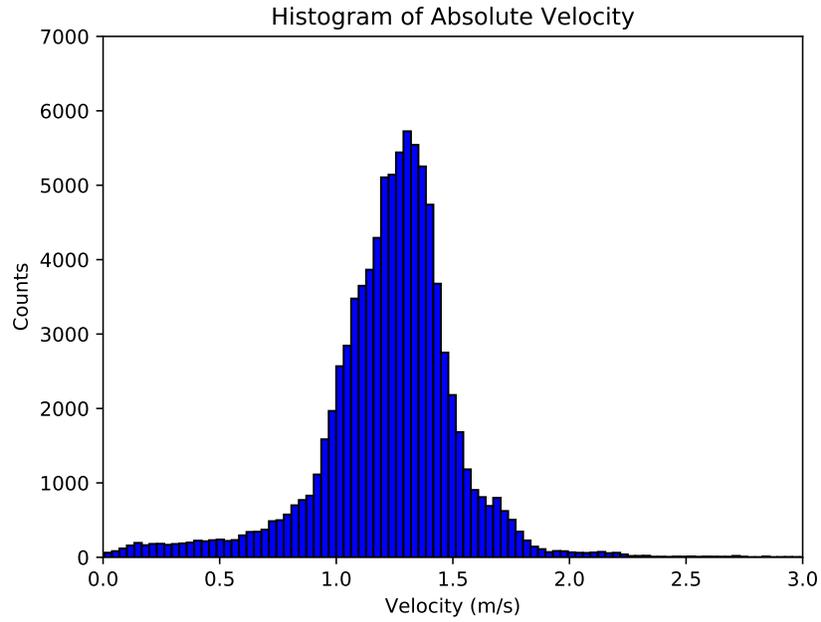


Figure 3.7: Distribution of velocity magnitude in CITR dataset

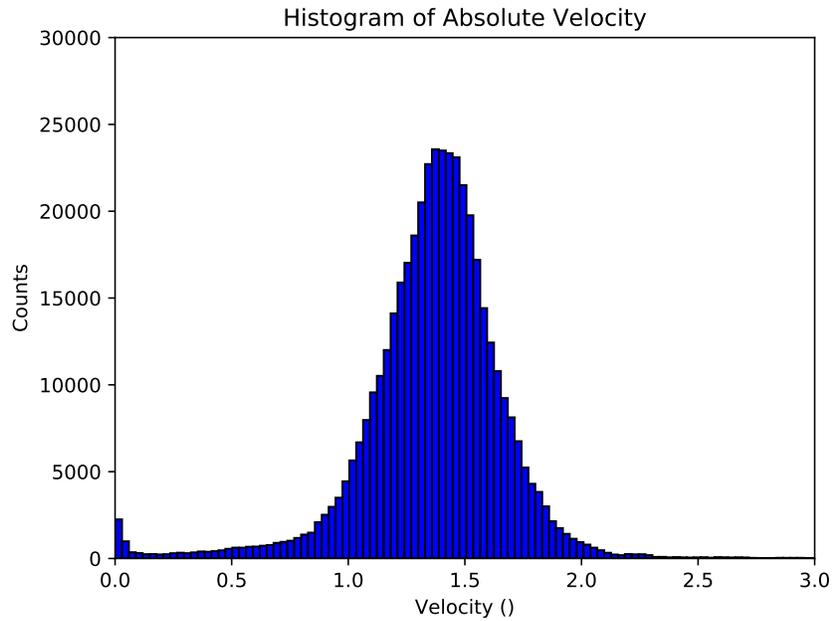


Figure 3.8: Distribution of velocity magnitude in DUT dataset

Table 3.3: Mean velocity magnitude

Dataset	Mean velocity	Mean walking velocity
CITR	1.2272	1.2435
DUT	1.3661	1.3825

processing techniques and refined by Kalman Filters. The statistics of the velocity magnitude validated the proposed dataset.

Chapter 4

Model Calibration

4.1 Background

This chapter presents the process of calibrating the parameters of the model proposed in chapter 2 using part of our established dataset in chapter 3. Previous works such as [29, 42] utilized pedestrian-only datasets to calibrate pedestrian-only models, which however demonstrated the advantage of using pedestrian trajectory data to improve pedestrian motion modeling. In terms of the calibration method, most works adopted genetic algorithm [24, 25, 61], evaluated based on the difference between the ground truth trajectories and simulated trajectories of the social force model. This method was proved to be effective in calibrating the parameters of pedestrian motion model. Our work also applied the genetic algorithm.

4.2 Calibration

4.2.1 Parameter Set

All parameters associated with the proposed model in chapter 2 are presented in table 4.1. They have been classified into 3 categories:

- Constant Parameters: the parameters that can be directly assigned based on statistics (e.g. using average pedestrian radius and mass) or ground truth (e.g. vehicle size).
- Ped2Ped Parameters: the parameters that are responsible for pedestrian-pedestrian interaction.
- Veh2Ped Parameters: the parameters that are responsible for vehicle-pedestrian interaction.

This study applied a two-step procedure for calibrating the parameters. First, the Ped2Ped parameters were calibrated and evaluated based on the trajectory data that does not contain vehicles. In this step the Veh2Ped parameters were fixed to arbitrary values, because whatever values of the Veh2Ped parameters are set, the vehicle force is always zero. Second, keeping the obtained Ped2Ped parameters fixed, Veh2Ped parameters were then calibrated and evaluated based on the data that contains vehicles. The reason of applying this configuration is that, if the Veh2Ped parameters were not fixed in the process of

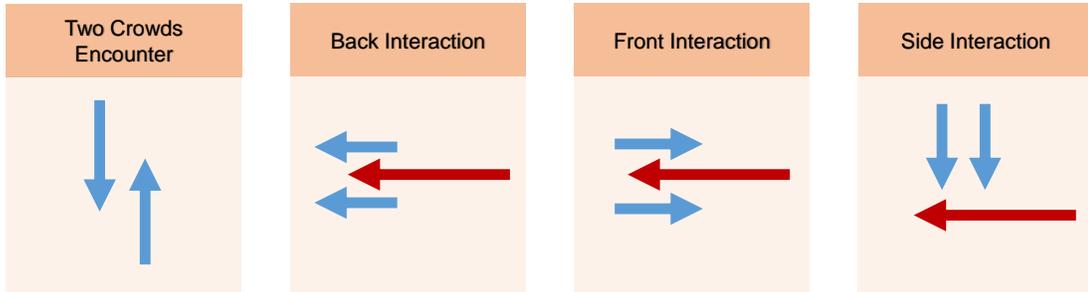


Figure 4.1: Scenarios: the first scenario was used for calibrating Ped2Ped parameters; the following three scenarios were used for calibrating Veh2Ped parameters.

calibrating Veh2Ped parameters, some Ped2Ped parameters might be modified in favor of obtaining better results on the data containing vehicle influence. We argue that the calibration cannot automatically differentiate between Ped2Ped parameters and Veh2Ped parameters.

4.2.2 Scenarios and Dataset

Four fundamental scenarios were used for the calibration, as shown in figure 4.1. The first scenario was used for pedestrian-pedestrian interaction. After the calibration, the model is expected to generate collective pedestrian behaviors such as lane formation and collision avoidance. Here only the scenario of bidirectional pedestrian motion is used, because the pedestrian motion patterns are similar in the scenarios that were used for vehicle-pedestrian interaction. The following three scenarios are used for calibrating vehicle-pedestrian interaction. They are fundamental interaction scenarios that consist of back interaction (vehicle coming behind the walking pedestrians), front interaction (vehicle coming in front of the walking pedestrians), and lateral interaction (vehicle coming from both sides of the walking pedestrians).

Trajectories of pedestrians and vehicles that correspond to the scenarios in figure 4.1 were collected by conducting controlled experiments. The experiments were conducted at an open space in a parking lot near the CAR-West facility at The Ohio State University, as shown in figure 4.2. Pedestrian motion, as well as vehicle motion, were recorded by a drone with a downward camera hovering above the experiment area. To obtain trajectories, positions were extracted by computer vision based tracking techniques, while velocities were reconstructed by Kalman filters. There are 80 pedestrian trajectories in total for pedestrian-pedestrian interaction, and 96 pedestrian trajectories in total for the vehicle-pedestrian interaction. Details about the dataset can be found in chapter 3.

4.2.3 Calibration Procedure

The calibration consists of manual calibration and data-driven calibration. In the manual calibration, a set of reasonable and acceptable parameters were obtained by trial-and-error, which was evaluated by visually inspecting the simulation results with the obtained parameters. Using the manually calibrated parameters as initial values, the genetic algorithm (GA) was then applied to further calibrate the parameters. The GA calibration was evaluated by the errors between the ground truth trajectories and the simulated tra-

Table 4.1: The list of calibrated parameters for the proposed model

Parameter	Calibrated Value	Category
R_i	0.27	Constant
m_i	80	Constant
l_r	1.2	Constant
l_f	1	Constant
l_w	1.2	Constant
β_v^S	3.9761	Ped2Ped
S_v^0	0.06566917	Ped2Ped
β_a^S	2.994062	Ped2Ped
S_a^0	0.39941	Ped2Ped
α^{col}	9825.125	Ped2Ped
d_0^{rep}	0.7801	Ped2Ped
M^{rep}	301.028	Ped2Ped
σ^{rep}	0.45971243	Ped2Ped
λ^{rep}	0.1	Ped2Ped (fixed)
d_0^{nav}	1.5892008	Ped2Ped
M^{nav}	410.875	Ped2Ped
σ^{nav}	0.41745	Ped2Ped
λ^{nav}	1	Ped2Ped (fixed)
T^S	3.665375	Ped2Ped
ϕ^S	121.39191	Ped2Ped
λ^S	1.87	Ped2Ped
v_0^i	1.394293	Ped2Ped
σ^{des}	1	Ped2Ped (fixed)
k^{des}	545.3125	Ped2Ped
β_v^F	0.001577598	Veh2Ped
F_v^0	199.3611	Veh2Ped
β_a^F	0.09775474	Veh2Ped
F_a^0	53.94855	Veh2Ped
l_e	0.2151011	Veh2Ped
d_x^0	0.510985	Veh2Ped
α_x	1.394358	Veh2Ped
A^{veh}	777.5852	Veh2Ped
b^{veh}	2.613755	Veh2Ped
λ^{veh}	0.3119132	Veh2Ped
F_1	199.7455	Veh2Ped
F_2	672.6487	Veh2Ped

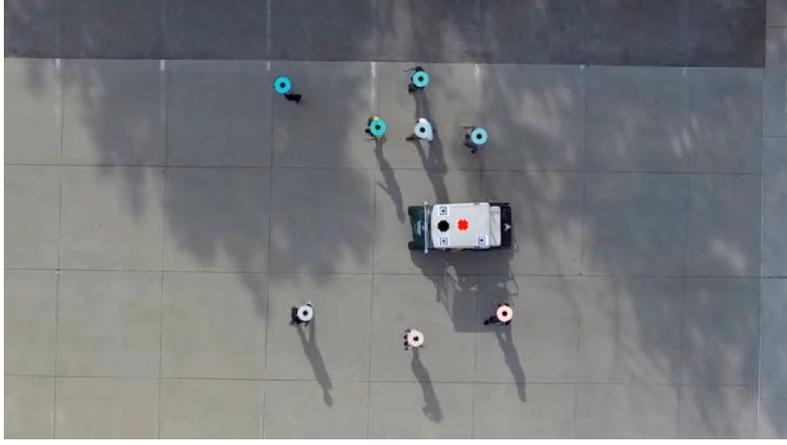


Figure 4.2: Illustration of the controlled experiment. Positions of both pedestrians and vehicles were extracted from the top-view video. Velocities were reconstructed by Kalman filters.

jectories. Ultimately, using the updated parameters, post-simulations were conducted to verify and validate the proposed model.

As mentioned in section 2.3.4, a destination is required for the social force model. In the manual calibration, the destinations of all pedestrians were pre-defined in the self-designed scenarios. In the data-driven calibration, a pedestrian’s destination was estimated based on the initial position $\vec{x}_0^{i,rec}$ and final position $\vec{x}_T^{i,rec}$ of the recorded trajectory. Specifically, for pedestrian-pedestrian interaction, the destination was estimated individually:

$$\vec{x}_t^{i,des} = \vec{x}_0^{i,rec} + \alpha^{rec} \cdot (\vec{x}_T^{i,rec} - \vec{x}_0^{i,rec}), \forall i \in \mathbb{S} \quad (4.1)$$

where \mathbb{S} refers to a specific scenario, and α^{rec} is a positive scalar, which was set as 1.5 in the calibration.

For vehicle-pedestrian interaction, since all pedestrians have similar motion, the destination was estimated based on the average of all initial positions and the average of all final positions:

$$\vec{x}_t^{i,des} = \vec{x}_0^{mean,rec} + \alpha^{rec} \cdot (\vec{x}_T^{mean,rec} - \vec{x}_0^{mean,rec}), \forall i \in \mathbb{S}. \quad (4.2)$$

4.2.4 Genetic Algorithm

Genetic algorithm (GA) [35] is a class of evolutionary algorithms that mimics natural selection. It is well suitable for finding the (near) optimal solutions to complex systems. The basic operators such as mutation, crossover, and selection introduce randomness, hence possibly overcoming the local minima problem.

In the GA calibration process, equation 2.1 was iteratively applied for each pedestrian to obtain simulated trajectories $\{\vec{x}_t^{i,sim}\}, \forall t \in \{1, \dots, T\}, \forall i \in \mathbb{S}$. Pedestrian’s initial state was set as the initial state of the recorded trajectory $\vec{x}_0^{i,sim} := \vec{x}_0^{i,rec}, \forall i \in \mathbb{S}$, while vehicle applied the whole recorded trajectory $\vec{x}_t^v := \vec{x}_t^{v,rec}, \forall t \in \{0, 1, \dots, T-1\}$.

The performance of a particular parameter set Θ was evaluated by a loss function comparing the simulated trajectories with the recorded trajectories.

Loss Function

A classical way of designing loss function is to maximize the likelihood of every point on the pedestrian trajectory. The assumption is that the error between the simulated trajectory and the recorded trajectory is Gaussian. If the log-likelihood is maximized, as presented in [9], it is equivalent to minimizing the mean square error:

$$\sigma_i^2(\Theta) = \frac{1}{T} \sum_{t=1}^T (\bar{x}_t^{i,rec} - \bar{x}_t^{i,sim})^2, \quad (4.3)$$

where Θ stands for the parameter set. This study applied mean square error as the loss (fitness) function of the genetic algorithm.

Initialization

The parameters were initialized as the manually calibrated parameters in the previous step. Lower bounds and upper bounds were added to ensure that, in the process of calibration, the parameters don't go beyond unrealistic values. For example, some parameters representing force magnitude should always be positive. And a parameter of the field of view obviously has an angle limit. The total number of population in the genetic algorithm was set to 200, which is sufficient for the calibration.

Implementation

The calibration and evaluation were conducted in MATLAB R2018b with Simulink. The Global Optimization Toolbox is used for executing the genetic algorithm. As an example of computation time, a typical GA calibration with 25 generations requires approximate 12 hours on an Intel(R) Core(TM) i7-4790 CPU @3.60GHz desktop computer. Simulation is done at the time interval of $\Delta t = 0.0334s$, which is equivalent to 29.97 frames per second (FPS), the same value as the FPS in the trajectory dataset.

4.3 Result

4.3.1 Pedestrian-Pedestrian Interaction

In this step, only Ped2Ped parameters, as shown in table 4.1, were calibrated. Some parameters that were manually calibrated in the previous step were fixed during the GA calibration. Specifically, λ^{ep} and λ^{nav} were fixed, because we think the associated anisotropies were good enough, and the fixation could also reduce the uncertainty of running the GA calibration. σ^{des} was fixed because it only reduces the desired velocity when a pedestrian is close to its destination, which does not affect the pedestrian motion too much if the pedestrian has not reached the destination.

Calibration

The GA calibration was executed for more than 30 generations. After 20 generations, both the best fitness and the mean fitness converged. The best fitness value of 1.00468 was obtained, which means the average error of all positions in all trajectories of all pedestrians is about 1 meter. This indicates the parameters were calibrated to some extent so that the

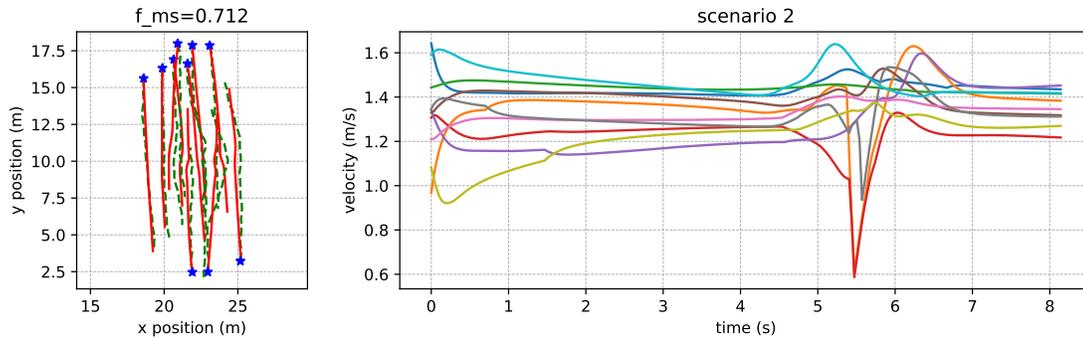


Figure 4.3: Validation of a selected scenario of pedestrian-pedestrian interaction. Left: A comparison between simulated trajectories (red solid lines) and recorded trajectories (green dashed lines). The asterisks indicate the initial positions of each pedestrian. f_{ms} shown in the title is the average of the mean square errors (as defined in equation 4.3) of all pedestrians. Right: The evolution of velocities of all pedestrians in this scenario. More results can be found in supplementary materials A.1.

model achieved its best performance based on the applied trajectory data. The obtained values of Ped2Ped parameters are presented in table 4.1.

Validation

The validation was done by simulating pedestrians with the same initial conditions as in the data used for calibration. This allows us to compare and analyze the simulated trajectories with recorded trajectories. Figure 4.3 plotted both types of trajectories as well as the evolution of velocities of all pedestrians from a selected pedestrian-pedestrian interaction scenario. The blue asterisks show the initial positions of all pedestrians. The red solid lines indicate the simulated trajectories, while the blue dashed lines indicate recorded ground truth trajectories. The results of more scenarios can be found in the supplementary materials A.1. The trajectory plotting shows that in general pedestrians are able to navigate around each other when there is potential conflict. The velocities demonstrate that pedestrians adjust their walking speed (slow down or accelerate) when interacting with each other.

Post-simulation

Post-simulations of pedestrians in more complex scenarios were conducted to further evaluate the model based on the calibrated parameters. In particular, a scenario was designed such that 4 groups of pedestrians interact with each other from 4 different directions. The screenshots of the simulation were displayed in figure 4.4. As the screenshots show (see the caption for detail description), all pedestrians were interacting as expected, even in the extremely dense situation. Figure 4.5 shows the trajectories and the evolution of velocities of all pedestrians in the simulation. The curved trajectories indicate the pedestrians navigate around to avoid potential collision. The multiple decreases in velocities indicate the pedestrians slow down to avoid collision.

It can be seen from the simulated pedestrian motion that each pedestrian is capable of navigating around any potential conflict with others.

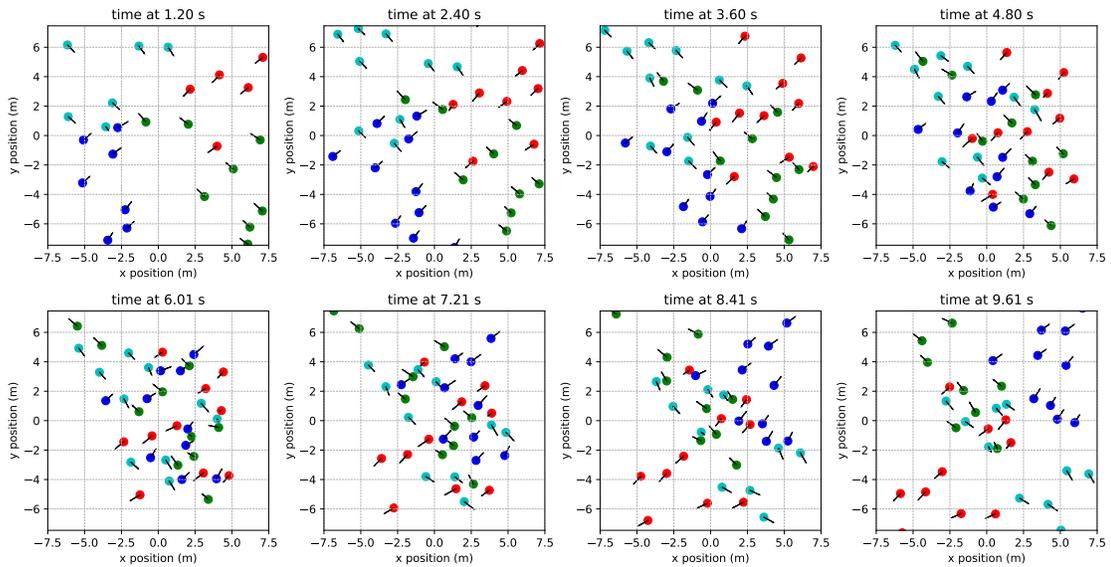


Figure 4.4: Simulation of 4 groups of pedestrians interacting from 4 different directions. Pedestrians in different groups are plotted in different colors. The small black arrows indicate the walking directions and the walking velocities (length of the arrow). Different groups were randomly initialized in different quadrants and were assigned a destination in the diagonal position, respectively. For example, the red group was initialized within the area of $x \in [0, 10]$ and $y \in [0, 10]$ and was assigned a destination at $[-10, -10]$.

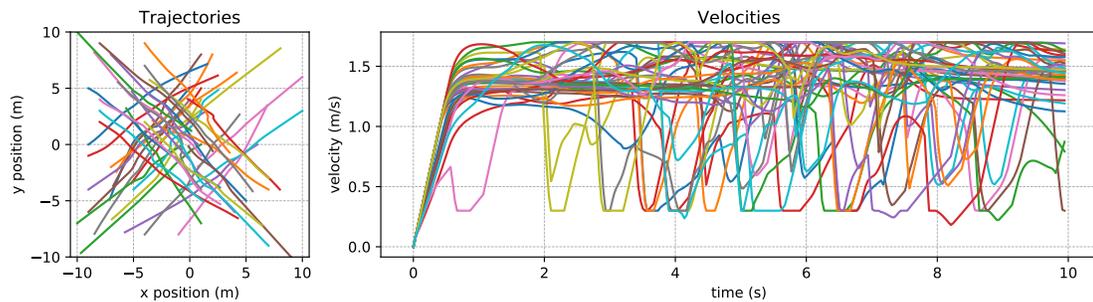


Figure 4.5: Trajectories and velocities of 4-group interaction. Left: Trajectories of all pedestrians in the simulation. Right: The evolution of velocities of all pedestrians in the simulation.

4.3.2 Vehicle-Pedestrian Interaction

After obtained the Ped2Ped parameters, the Veh2Ped parameters, as shown in table 4.1 were calibrated by using the data that includes vehicles. This step did not change the values of Ped2Ped parameters.

Calibration

The GA calibration was executed for more than 30 generations. Similarly, after about 20 generations, both the best fitness and the mean fitness converged. The best fitness value of 4.1918 was obtained. The obtained parameter values are shown in table 4.1. The best fitness value for Veh2Ped parameters is larger compared with the best fitness value obtained for Ped2Ped parameters. This is because the vehicle-pedestrian interaction is much more complex than the pedestrian-pedestrian interaction, so it is reasonable that the fitness value is larger. Complex interaction may require a considerably large amount of data for calibration. But in any case, the convergence of both mean fitness and best fitness indicates that the model achieved its best performance based on the calibration data. Therefore, we can still conclude that the model performance was improved to some extent.

Validation

The model was still calibrated by simulating pedestrians with the same initial conditions as in the data used for calibration. Different from pedestrian-pedestrian interaction, vehicle motion was added in the simulation by using the ground truth vehicle trajectories. Figure 4.6, 4.7, and 4.8 show 3 selected scenarios, which correspond to back interaction, front interaction, and lateral interaction, respectively. More results can be found in supplementary materials A.2. According to the simulated trajectories, pedestrians are able to avoid the vehicle from different directions. The velocities indicate that pedestrians may slow down or accelerate to avoid the potential collision with the vehicle. There is still a certain degree of error between the simulated trajectories and recorded trajectories, which can be explained by the following reasons:

- Different types of fundamental interactions (back, front, and lateral interactions) may need a different parameter set, which implies that one general model may not be sufficient to describe the vehicle-pedestrian interaction.
- Our assumption of homogeneous pedestrians (every pedestrian applies the same parameter set) limits the model performance. Even for the same pedestrian in the same situation, the pedestrian behavior could also be affected by his or her inner thought (e.g. in a rush or not).
- The limited amount of data for calibration could also cause the error.

Post-simulation

Post-simulations were conducted to further evaluate the performance of the model. Scenarios of 3 types of fundamental vehicle-pedestrian interactions were designed and simulated: back interaction, front interaction, and lateral interaction. The vehicle motion was simulated by using a pure-pursuit controller tracking a pre-defined path, as described

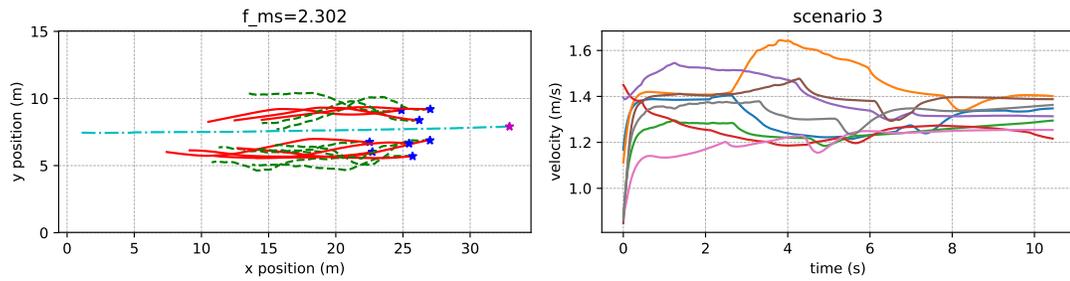


Figure 4.6: Validation result of a selected scenario of back interaction. Left: A comparison between simulated trajectories (red solid lines) and recorded trajectories (green dashed lines). The cyan dash-dotted line represents the ground truth trajectory of the vehicle motion. The asterisks indicate the initial positions of each participant. f_{ms} shown in the title is the average of the mean square errors (as defined in equation 4.3) of all pedestrians. Right: The evolution of velocities of all pedestrians in this scenario. More results can be found in supplementary materials A.2.

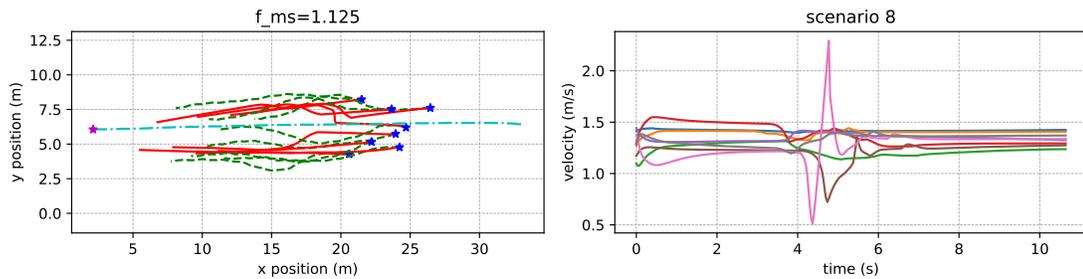


Figure 4.7: Validation result of a selected scenario of front interaction. The notation is the same as in figure 4.6.

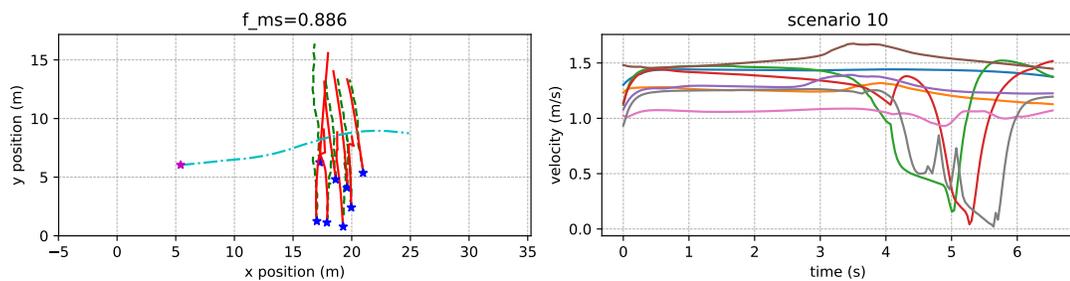


Figure 4.8: Validation result of a selected scenario of lateral interaction. The notation is the same as in figure 4.6.

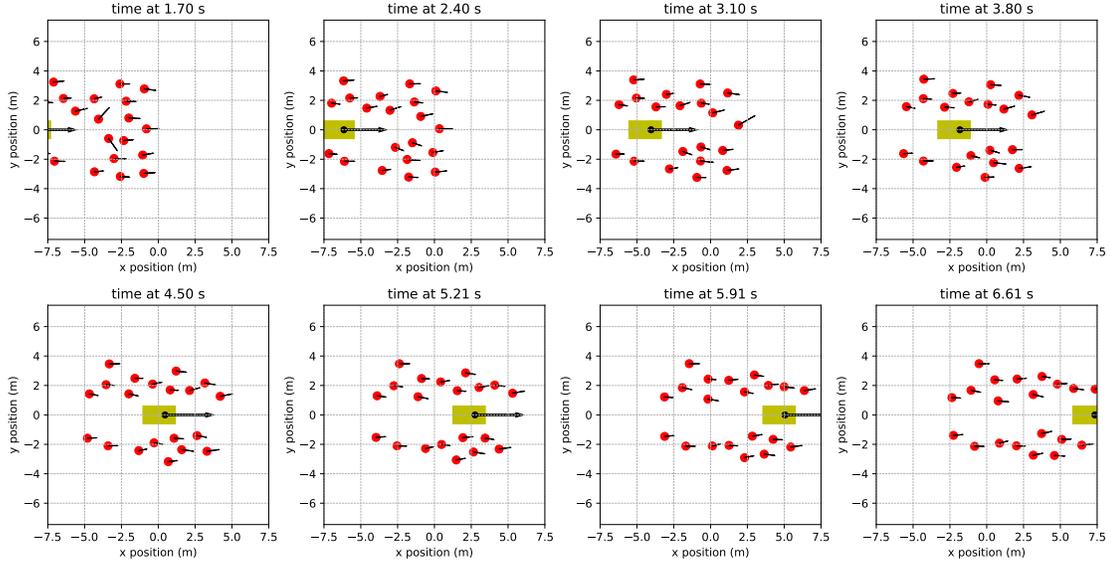


Figure 4.9: Simulation snapshots of back interaction: A slowing-moving (3m/s, which is the average speed in the dataset) vehicle (yellow box) was moving from negative x axis to positive x axis. Pedestrians (red circles) were randomly initialized within the area of $x \in [-9, -3]$ and $y \in [-3, 3]$ and was assigned a destination at $[20, 0]$. The small black arrows indicate the walking directions and the walking velocities (length of the arrow) of pedestrians, while similarly the big black arrow indicates the orientation and the velocity of the vehicle.

in section 2.3.6. The simulation results are shown in figures 4.9, 4.10, and 4.11, respectively. Details of the simulation configuration can be found in the figure captions. The pedestrians in the simulation are able to avoid the collision with the vehicle from different directions. Notice that in a relatively crowded situation, as shown in figure 4.11, there are a couple of instances of small overlap among pedestrians. This is exactly what we expected because the pedestrians should be able to push others if a vehicle is approaching in a dangerous way. Trajectories and velocities of the above simulation can be found in the supplementary materials A.3, in which pedestrian behavior such as slowing down and accelerating to avoid the vehicle can be identified.

Overall, although the calibration of vehicle-pedestrian interaction did not generate a very good fitness value, the post-simulation still validated the performance of the proposed model.

4.4 Conclusion

This chapter describes how the proposed social force model was evaluated and validated by both the simulation and calibration of fundamental vehicle-pedestrian interaction scenarios (back, front, and side interaction).

In this work, each pedestrian is assumed to be associated with a homogeneous pedestrian motion model (same values of parameters), because we were looking for a general model that can describe general pedestrian motion under vehicle influence. The proposed model is able to describe pedestrian motion in fundamental vehicle-pedestrian interac-

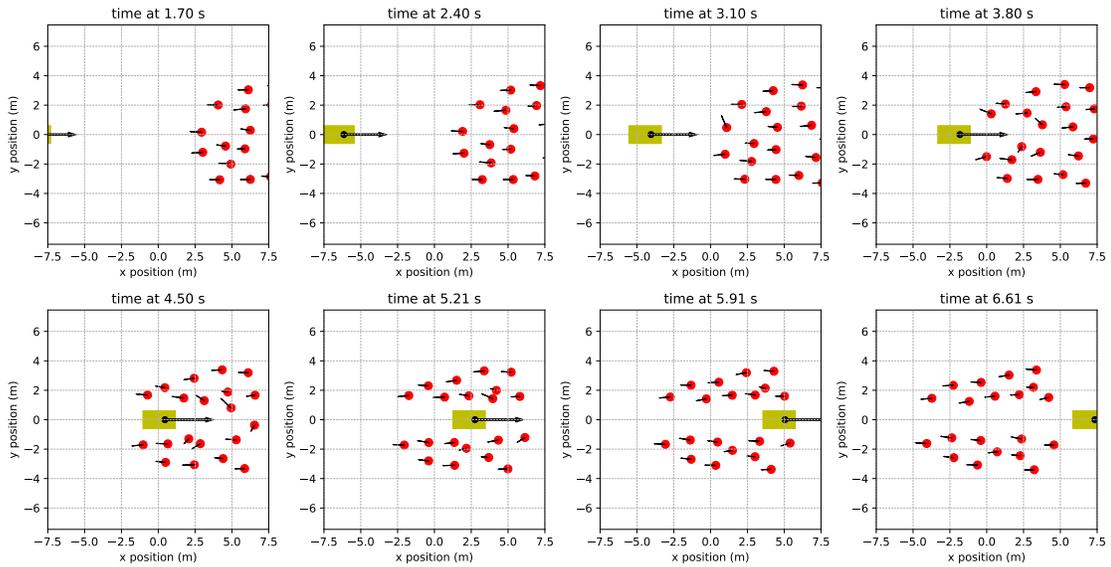


Figure 4.10: Simulation snapshots of front interaction: The configuration is the same as in figure 4.9, except the pedestrians were initialized within $x \in [5, 11]$ and $y \in [-3, 3]$ and was assigned a destination at $[-20, 0]$

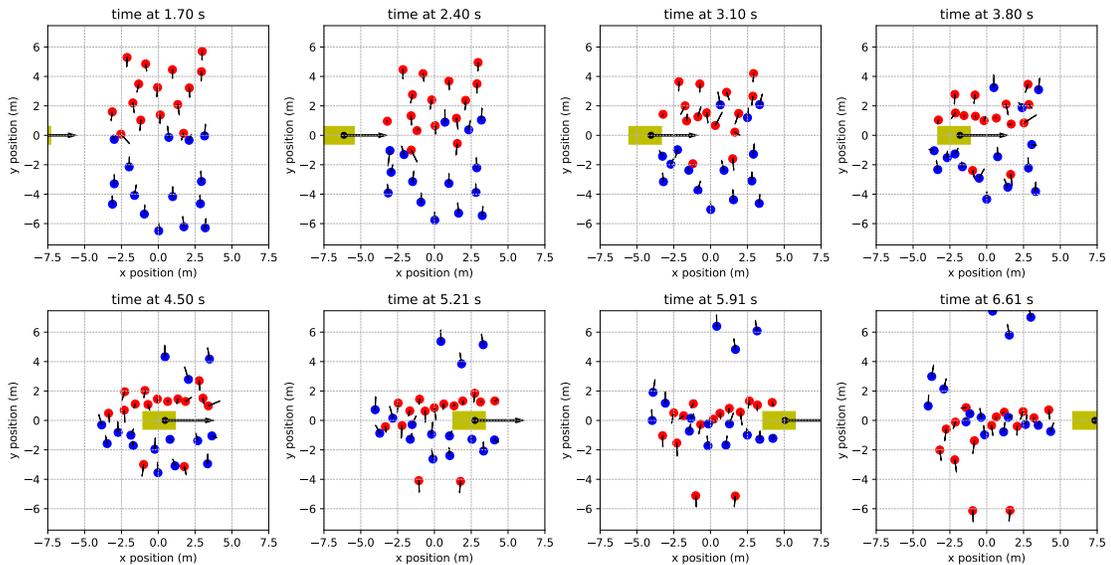


Figure 4.11: Simulation snapshots of lateral interaction: The configuration is the same as in figure 4.9, except there were two groups of pedestrians. One group (red circles) was initialized within $x \in [-3, 3]$ and $y \in [2, 8]$ with a destination at $[0, -20]$. The other group (blue circles) was initialized within $x \in [-3, 3]$ and $y \in [-8, -2]$ with a destination at $[0, 20]$.

tion scenarios, as demonstrated in the simulation. Pedestrian-pedestrian interaction is in general good. However, a certain amount of error exists in vehicle-pedestrian interaction, hence indicating that a homogeneous model is not enough for describing detailed pedestrian behavior under vehicle influence.

Chapter 5

Applications

5.1 Background

The intention of crowd pedestrians plays an important role for autonomous vehicles or intelligent systems in transportation. This intention is especially critical to shared space scenarios that involve crowd pedestrians and autonomous vehicles. One of the most common issues that relies on the crowd intention is how to improve the transportation efficiency in shared spaces when autonomous vehicles traverse the shared spaces that are partially or mostly occupied by crowd pedestrians. To do this, it is necessary to evaluate the status of individual pedestrians as well as their interactions with other pedestrians and the autonomous vehicles. This evaluation requires simultaneous handling of pedestrian motion, vehicle action, and the driving efficiency of the vehicle.

The proposed social force pedestrian model can be applied to this type of problems. This chapter proposed a framework that includes the pedestrian motion model. The framework aims to (a) detect individual pedestrian's state in the crowd via multiple sensors, (b) predict crowd pedestrians' motion given the driving strategy of the autonomous vehicles, and (c) evaluate the vehicle driving efficiency based on the scenario simulation, which eventually contributes to the transportation efficiency in shared spaces. A multi-sensor strategy was introduced for accurately detecting and estimating the individual pedestrian's state. Initial results of pedestrian detection on each separate sensor are presented. The social force based pedestrian model was combined with a vehicle model to predict and evaluate the scenario. Based on the prediction and evaluation, the driving efficiency was consequently adjusted. The approaches to address several types of driving efficiency problems were discussed. Two case studies were conducted to improve two types of driving efficiencies. The first case is an offline approach which predicts the desired passing time for an autonomous vehicle to pass through a crowd of pedestrians. The second case is an online approach which actively adjusts the vehicle's control action to pass through a crowd of pedestrians.

On-vehicle pedestrian detection is primarily achieved by a combination of different sensors, i.e., monocular camera, LiDAR, and stereo camera. Monocular vision [13, 18, 11] is the primary source of pedestrian detection because it provides the texture of pedestrians. LiDAR pointcloud[34] and stereo vision[65][38] provide depth information hence usually being used as supplementary source to the monocular vision. On-vehicle sensors provide instant detection results of pedestrians in the neighborhood of the autonomous vehicle. However, when the crowd density is high enough, it is generally difficult for on-vehicle

sensors to detect all individual pedestrians due to massive occlusions. Nowadays, with the commercialization and the decreasing prices of unmanned aerial vehicles (UAVs), it is possible to use UAVs with downward facing aerial cameras as infrastructure sensors hovering above the interested area so that the individual pedestrians can be more easily detected. This is also the case for the cameras mounted on infrastructures of certain height. Therefore, we propose a multi-sensor pedestrian detection strategy that relies on both infrastructure sensors and the on-vehicle sensors to handle the massive occlusion problems.

Scenario prediction provides necessary information for adjusting the driving efficiency of autonomous vehicles traversing the crowd in shared spaces. The prediction requires analyzing the interactive behavior of both the crowd pedestrians and the autonomous vehicles. Several studies have inspected this interaction mechanism[3][62][59]. Due to the complexity of the interaction mechanism, scenario simulation[56] is an effective approach to address the driving efficiency problem. Assuming the correctness of pedestrian detection in the previous stage and the validity of the vehicle-crowd interaction mechanism, analyzing the simulation results gives useful information for the improvement the driving efficiency.

For the rest of the chapter, section 5.2 presents the proposed overall framework. Section 5.3 describes the multi-sensor strategy and the methods for pedestrian detection on individual sensors with the corresponding initial results. Section 5.4 details the approach to evaluate and improve the transportation efficiency in shared spaces, with a simulation case study for the estimated time to pass through the crowd. Lastly, conclusions and future work are discussed.

5.2 Framework

Figure 5.1 shows the proposed overall framework. First, for an area of interest, individual pedestrians in the crowd are detected via both on-vehicle sensors and infrastructure sensors. Second, for the subject vehicle, the detection results are fused with the results from nearby vehicles (if they exist) and the results of the infrastructure sensors. The current pedestrian states, i.e., the positions and orientations of individual pedestrians at the current time step, are determined by combining the fused detection results with the high precision map. Next, the autonomous vehicle plans an initial driving strategy based on current pedestrian states. Both the driving strategy and the pedestrian states are sent to the scenario prediction module. The output of the scenario prediction is fed back to the autonomous vehicle so that the driving strategy can be updated based based on a specific driving efficiency objective, which is generated according to the transportation efficiency objective in shared spaces. Finally, the actual motion of the autonomous vehicle physically affects the area of interest.

Transportation efficiency objectives in shared spaces should be pre-specified as an input or a criteria for the framework. The transportation efficiency objective is then translated into the vehicle driving efficiency objective, because the autonomous vehicle is the primary and the most influential participant in the scenario.

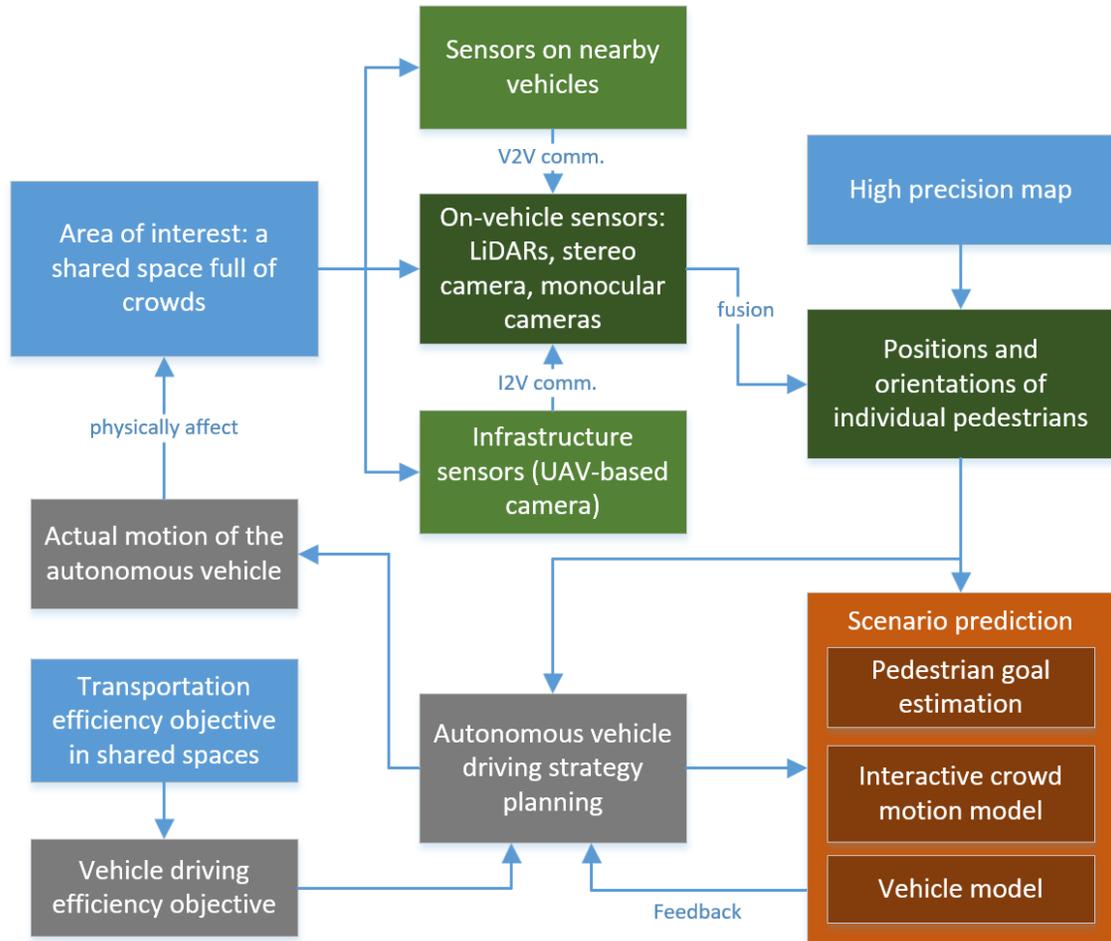


Figure 5.1: The proposed framework to improve transportation efficiency.

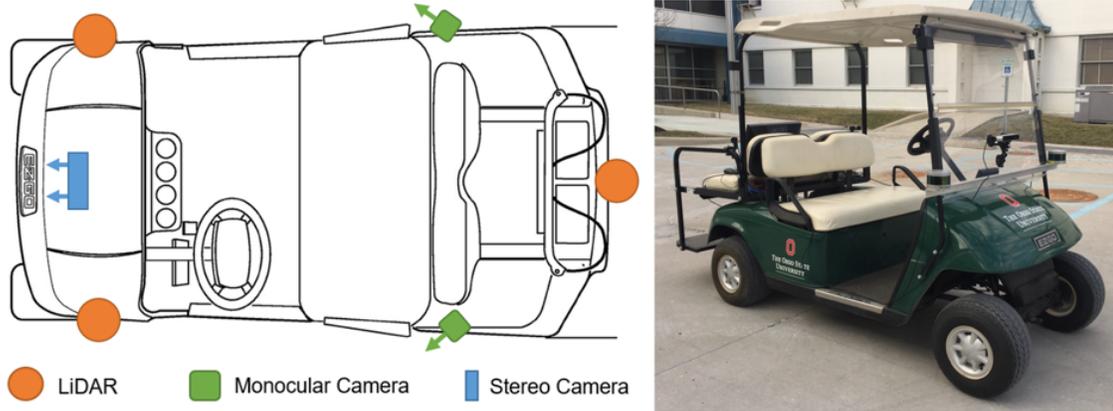


Figure 5.2: Experimental vehicle (right) and the configuration of on-vehicle sensors (left).

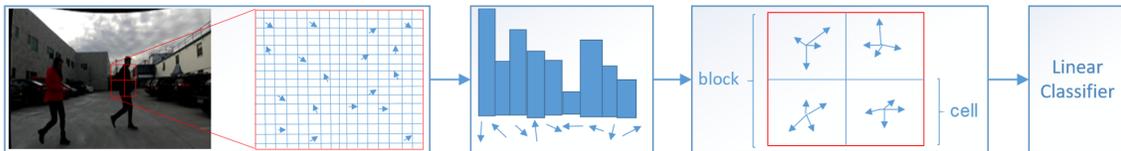


Figure 5.3: The pedestrian detection process of a monocular camera. Step 1: calculate pixel gradients, i.e., magnitude and direction in 8×8 cells. Step 2: each pixel votes for its cell gradient orientation depending on its gradient magnitude. Step 3: concatenate cell histograms into blocks of 2×2 that describe the final HOG feature vector of the whole image. Step 4: classify resulting HOG feature vector into pedestrians and other.

5.3 Pedestrian Detection in Crowd

5.3.1 On-vehicle Sensors

An E-Z-GO golf cart is used as our experimental vehicle. The vehicle contains a front facing stereo camera, two front/side facing monocular cameras, and three LiDAR (Light Detection And Ranging) sensors as shown in figure 5.2.

Monocular camera

Monocular camera vision comes from two FLEA 3 GigE Vision cameras. It also comes from either one of the channels of the stereo camera. Our study uses the approach in [10], which relies on extracting a Histogram of Oriented Gradient (HOG) features from the image, followed by a linear classifier using Support Vector Machines (SVMs). The overall detection process is illustrated in figure 5.3. Those methods usually work, specifically for pedestrians, as HOG features are robust against illumination and small local pose differences due to the fact that pixel gradients are normalized locally within blocks in the image. Adding texture information using Local Binary Pattern (LBP) descriptors to HOG features is a well-known method of detecting shapes and textures in the image feature space and could be applied to our system for better detection accuracy [51].

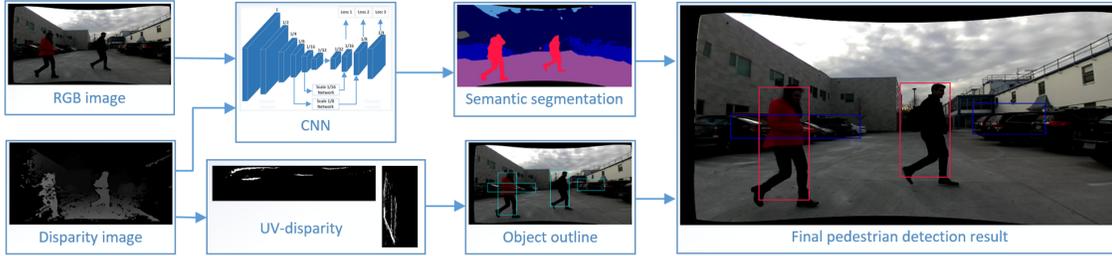


Figure 5.4: Stereo vision based pedestrian detection. The top row shows the process of semantic segmentation while the bottom row shows the process of UV-disparity map based object detection.

Stereo camera

A ZED 2K stereo camera is used primarily for detecting pedestrians in front of the autonomous vehicle. As shown in figure 5.4, the detection task is divided into two parallel processes: UV-disparity map based object detection, and the semantic segmentation of a monocular vision based on a convolution neural network (CNN). The U-disparity map can be used to detect the ground plane and find the upper and lower edges of the object while the V-disparity map can be used to find the left and right edges. Once the objects are identified, they are compared with the semantic segmentation result, which is achieved by ICNet [64]. Pedestrians can be identified by combining the semantic segmentation result and the objects found using the UV-disparity map.

LiDAR

Three Velodyne VLP-16 LiDAR sensors are used in conjunction to form a single 3D point cloud. Each LiDAR unit has 16 vertical layers covering a $\pm 15^\circ$ vertical field of view and a 360° horizontal field of view with a 100 meter range. The point cloud data is received from each LiDAR at 10 Hz and then translational and rotational offsets are applied before combining the point clouds to properly account for their different mounting locations. The translational offset is measured manually and the rotational offset is measured by the extrinsic rotational calibration method presented in [33].

The pedestrian detection method is similar to that used in [34] in regards to the LiDAR data being used in conjunction with monocular camera data. The overall process is illustrated in figure 5.5. The ground plane is first removed from the combined point cloud using the ground plane extraction algorithm from [23]. After the ground plane removal, object segmentation is performed as also done in [23]. The objects found in the point cloud after segmentation guide the camera-based pedestrian detection by providing regions of interest and narrowing the search space.

5.3.2 Infrastructure Sensors

Infrastructure sensors could be any combination of dedicated cameras mounted on nearby buildings, regular surveillance cameras, and downward facing aerial cameras mounted on UAVs. This section only focuses on UAV-based infrastructure sensors, as the detection methods apply similarly to others. In this study, a DJI Phantom 3 SE with a built-in camera, as shown in figure 5.6, is used as the UAV-based infrastructure sensor.

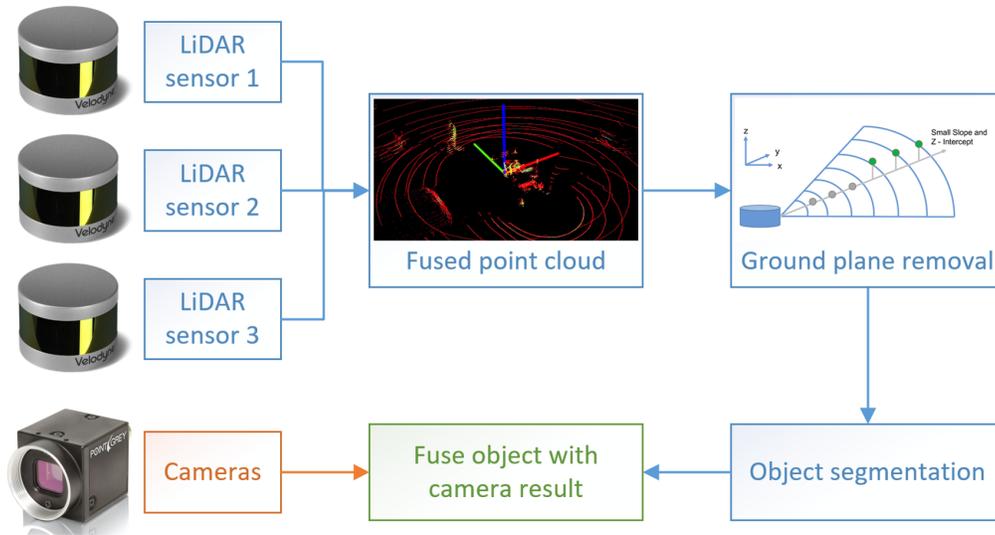


Figure 5.5: The flowchart of LiDAR based pedestrian detection.



Figure 5.6: The DJI[®] Phantom 3 SE unmanned aerial vehicle (UAV).

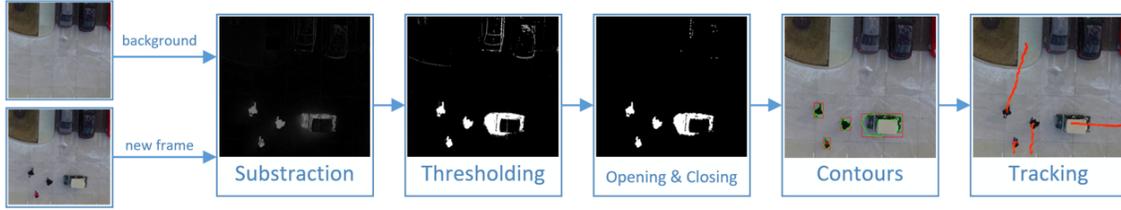


Figure 5.7: Pedestrian detection of UAV-based infrastructure camera. In the scenario, 3 pedestrians are walking in shared space while the vehicle tries to traverse.

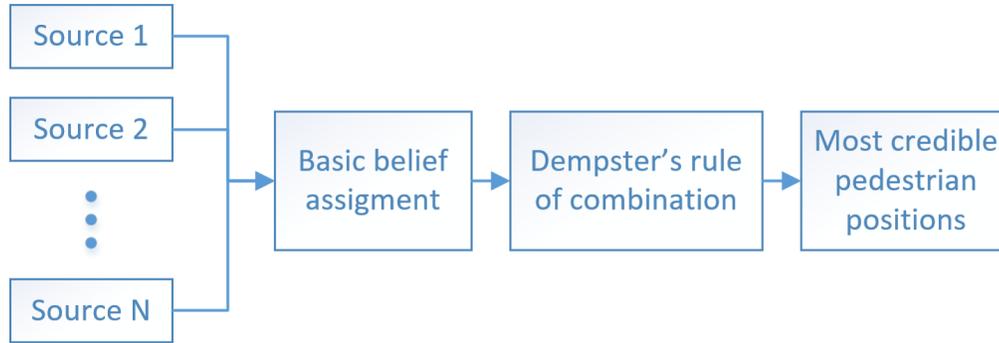


Figure 5.8: The flowchart of the Dempster-Shafer theory (DST).

Since the UAV is part of the infrastructure, it is safe to assume that the background of the area of interest is known. Individual pedestrian detection is done based on each new frame calibrated and subtracted with respect to the background. For each background-removed frame, a series of image processing operations (thresholding, opening, and closing) are applied and the contours and bounding boxes of all objects are then generated. Using the contours, positions of individual pedestrians can be easily determined. The above detection process is illustrated in figure 5.7.

5.3.3 Sensor fusion

Once we have performed the pedestrian detection for both on-vehicle sensors and UAV-based infrastructure sensors, the next step is to fuse the detection results. The purpose of sensor fusion is to exploit the complementary and redundant characteristics of the sensors for increasing the reliability and accuracy of the pedestrian detection. The Dempster-Shafer theory (DST) is applied for the sensor fusion task, which combines the sources of evidence while avoiding counter-intuitive results [49]. Figure 5.8 shows how multiple sources of evidence are processed by DST. First, the basic belief assignment (BBA) is done based on multiple sources. Then the most credible pedestrian positions are generated by applying Dempster's rule of combination.

5.4 Transportation Efficiency in Shared Spaces

In a future project we will consider the application of these modeling results to improving the performance of autonomous vehicles in shared spaces. Here we summarize our initial proposed approaches to this problem.

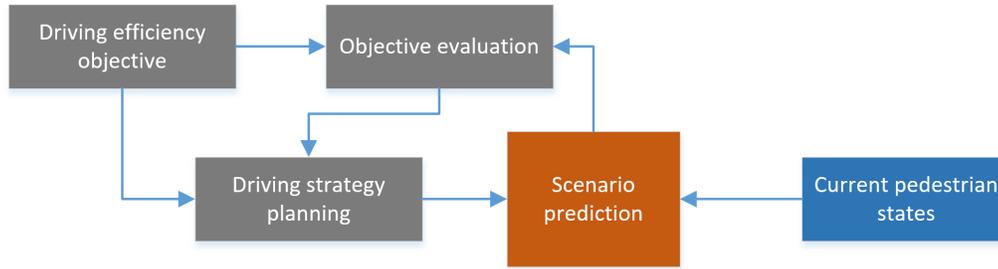


Figure 5.9: Online driving strategy planning process.

5.4.1 Scenario Prediction

The scenario in a short time horizon, for example 10 to 30 seconds, can be predicted in simulation given the current states of the individual pedestrians, models that describe the motion of both the pedestrians and the vehicle, and the driving efficiency objective of the autonomous vehicle.

Once the states of all individual pedestrians are available, the next step would be to estimate the current pedestrian goals. It is common to assume that a pedestrian walks linearly at a constant speed for a short horizon. Under this linear assumption, a near-term goal of the pedestrian can be inferred. The social force based pedestrian interaction model in chapter 2 will be applied to describe the motion of pedestrians under the influence of vehicles in shared spaces.

5.4.2 Approaches

As the pedestrian motion can only be predicted but hardly controlled by intelligent systems, the driving efficiency of the autonomous vehicle plays the main role in affecting the transportation efficiency in shared spaces. Depending on the driving efficiency objective, the driving strategy could be determined by either online or offline approaches.

Figure 5.9 shows the process of online driving strategy planning. Once the driving efficiency objective has been chosen and the current pedestrian states are available, modules of scenario prediction, objective evaluation, and driving strategy planning will be executed sequentially and iteratively until a driving strategy that satisfies the driving efficiency objective is generated. This can be achieved by forming a model predictive control (MPC) problem and solving the objective function. Specific vehicle driving objectives could be, for example, finding the shortest passing time through the crowd while guaranteeing the pedestrian safety.

The driving strategy can also be determined offline as shown in figure 5.10. This approach requires training of a model that represents the relationship between the inputs (the driving efficiency objective and the current pedestrian states), and the output (the driving strategy). Training will be achieved by the scenario prediction and the objective evaluation module given the data of possible initial pedestrian states and different candidates of driving efficiency objectives and driving strategies. One example of a problem that can be attacked by this approach is to estimate the desired time for the autonomous vehicle to pass an area of interest in a crowded shared space [58].

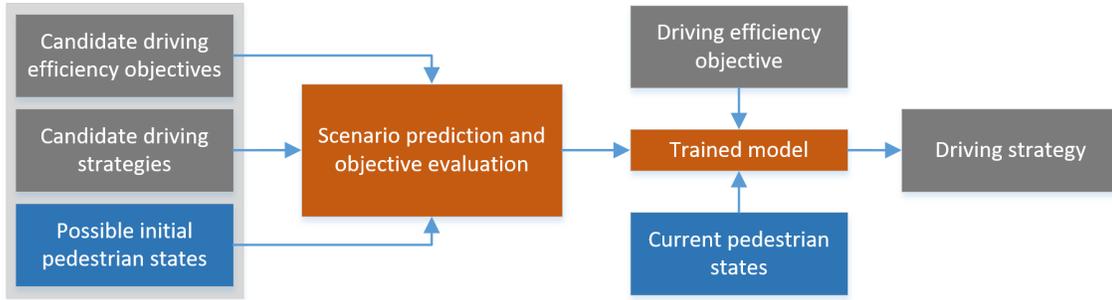


Figure 5.10: Offline driving strategy determination process.

5.5 Conclusion

This chapter presented a framework that combines pedestrian detection via multiple sensors, vehicle-crowd interactive scenario prediction, and approaches to improve the driving efficiency of autonomous vehicles, which finally affects the transportation efficiency in shared spaces. Methods of pedestrian detection on different types of sensors were introduced and the corresponding initial results were presented. Finally, we propose both online and offline approaches for improving driving efficiency. The proposed framework has the potential to solve transportation problems in shared spaces where crowd pedestrians and autonomous vehicles interact with each other.

Chapter 6

Conclusion

6.1 Conclusion

This report summarizes the design of a social force based pedestrian motion interaction model that specifically considers the vehicle effect on the behavior of pedestrians. The parameters of the proposed model were calibrated by the genetic algorithm with the help of visual inspection. A pedestrian trajectory dataset that involves vehicle-pedestrian interaction was established as the ground truth data to support the model parameter calibration. The dataset also has the potential to benefit a variety of pedestrian crowd related studies. Applications of the proposed pedestrian motion model were also presented. We proposed a framework that combines the pedestrian detection, scenario prediction, and vehicle driving efficiency improvement in pedestrian-dense scenarios. In a subsequent project, we will explore and develop these approaches and examine case studies that will demonstrate how to apply the proposed model into driving efficiency application.

6.2 Future Work

To further improve the pedestrian motion model and its calibration, there are several points to consider:

- The assumption of homogeneity could be removed; hence the pedestrian motion model can describe different types of interactions. For example, pedestrians could have different radii, different action capabilities (fast or slow), and different preferences of avoiding the collision. This can be achieved by individually calibrating a parameter set for each pedestrian or clustering a certain number of features based on the individually obtained parameters.
- The utilization of mean square error for calibration may not be a perfect choice. As the simulation time step increases, the error accumulates. The final prediction error (the difference between the last simulated position and the last recorded position) could be considered together with the mean square error.
- Instead of using the whole pedestrian trajectory in the data for calibration, a fixed-time length trajectory can be applied. This will avoid larger cumulative error in longer trajectories.

- Different fundamental vehicle-pedestrian interaction scenarios (front, back, and lateral) may require different designs of vehicle influence, or at least different parameter sets for vehicle influence. That means the calibration of vehicle-pedestrian could be done separately.
- Regarding multi-modality, various types of road users (e.g. cyclists, animals, etc.) should be considered in the future to make the pedestrian motion model more generalizable. But still, among these road users, vehicles should be the primary concern, as they are the most dominant and dangerous participants in the traffic.

The pedestrian trajectory dataset has the following limitations:

- There is no much variation of the vehicle speed in the dataset. Our data has an average vehicle speed of 3m/s. Whether the proposed model generalizes to the influence of the vehicle of higher speed has not been validated by the data. The dataset with different vehicle velocities is desirable.
- Pedestrian participants in our dataset do not represent all kinds of pedestrians, since they are primarily composed of college students. Also, the density of pedestrian crowd does not vary too much.
- It is also expected to build a benchmark that tests a couple of famous pedestrian motion models, which is our major future work.
- Another improvement could be automatically detecting/selecting initial positions of pedestrians when they entered the ROI, hence totally removing manual operation.
- From the aspect of personal characteristics, it would help if the pedestrians in the dataset could be identified according to their age, gender, head direction, and other features, although manual annotation of these features seems to be the only option at current stage.

For the applications presented in this report:

- Pedestrian Detection: The fusion of detection results from different sensors should be further improved. As a fundamental structure for sensor fusion, Dempster-Shafer theory should be adapted to fit the specific situation. The communication between UAV-based infrastructure sensors and on-vehicle sensors also requires further exploration, especially for how to guarantee real-time information exchange.
- Pedestrian Goal Estimation: Although linear assumption is generally acceptable in practice, applying a high-fidelity estimation model or incorporating more environment information can improve the estimation result. However, high-fidelity models and additional information require high computational capability. It is necessary to find an approach that can balance the estimation performance and the computational load.

Appendix A

Social Force Modeling

A.1 Validation of Pedestrian-Pedestrian Interaction

Figure A.1 shows the validation results of all scenarios of pedestrian-pedestrian interaction.

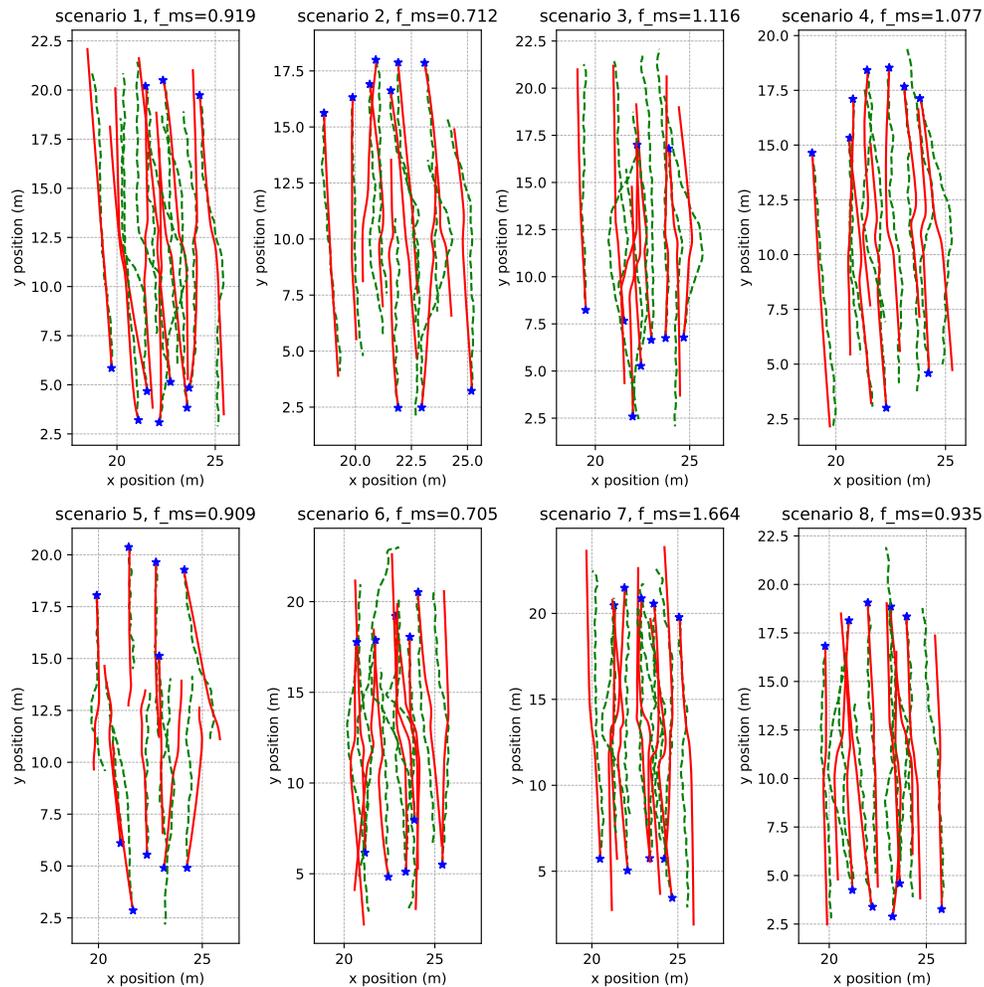


Figure A.1: A comparison between simulated trajectories (red solid lines) and recorded trajectories (green dashed lines) in pedestrian-pedestrian interaction scenarios. f_{ms} shown in the title is the average of the mean square errors (as defined in equation 4.3) of all pedestrians in the scenario. The asterisks indicate the initial positions of each pedestrian.

A.2 Validation of Vehicle-Pedestrian Interaction

Figure A.2 shows the validation results of all scenarios of vehicle-pedestrian interaction.

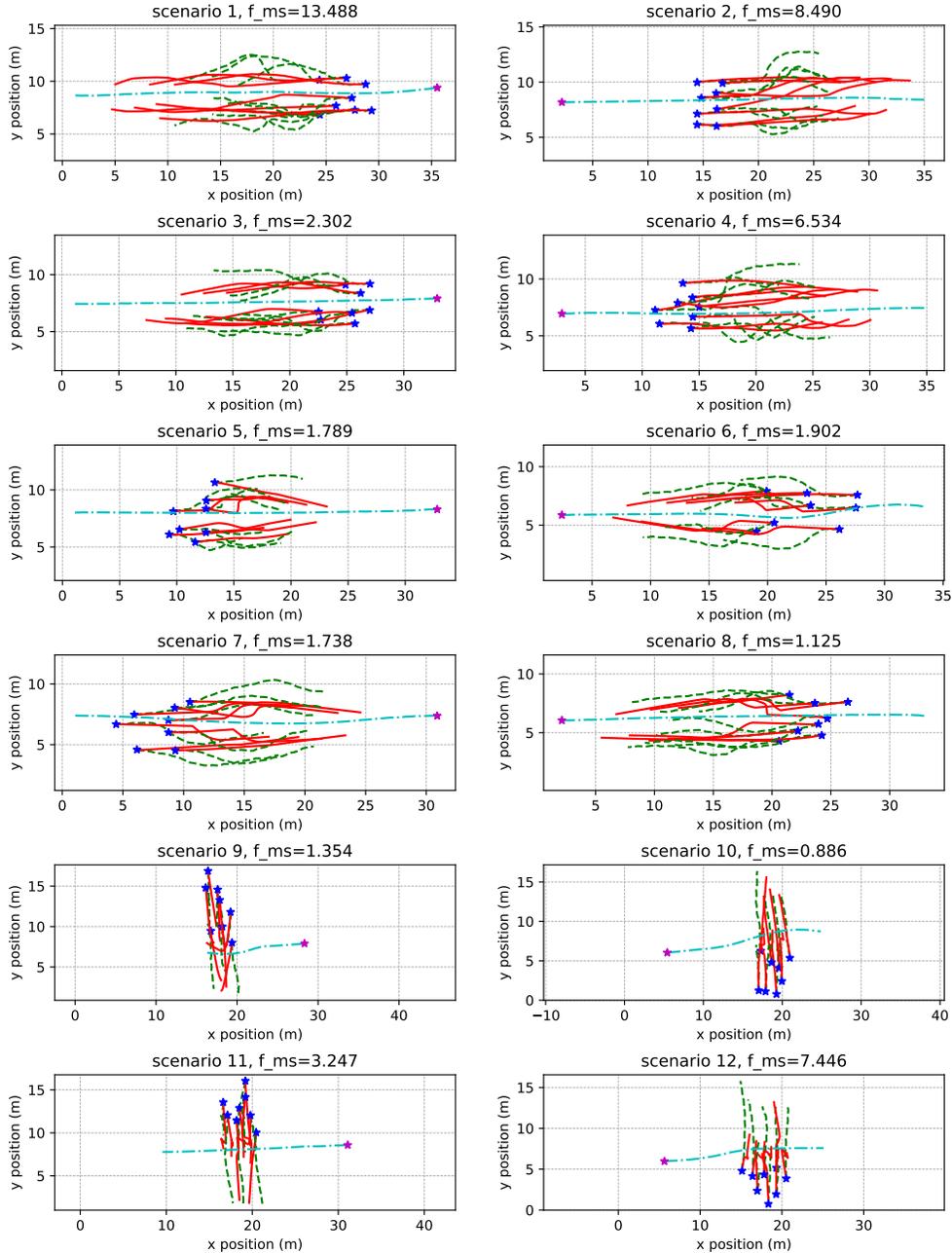


Figure A.2: A comparison between simulated trajectories (red) and recorded trajectories (black) of pedestrians in vehicle-pedestrian interaction scenarios, i.e., scenarios that considers vehicle influence. f_{ms} is the mean square error as defined in equation 4.3. The vehicle motion uses ground truth, of which the trajectories are indicated in dash-dotted cyan. The asterisks indicate initial positions. Rows 1-2, 3-4, and 5-6 show scenarios of back interaction, front interaction, and lateral interaction, respectively.

A.3 Post-Simulation of Vehicle-Pedestrian Interaction

Figures A.3, A.4, and A.5 show the trajectories and velocities back, front, and lateral interactions that correspond to figures 4.9, 4.10, and 4.11, respectively, in the post-simulation.

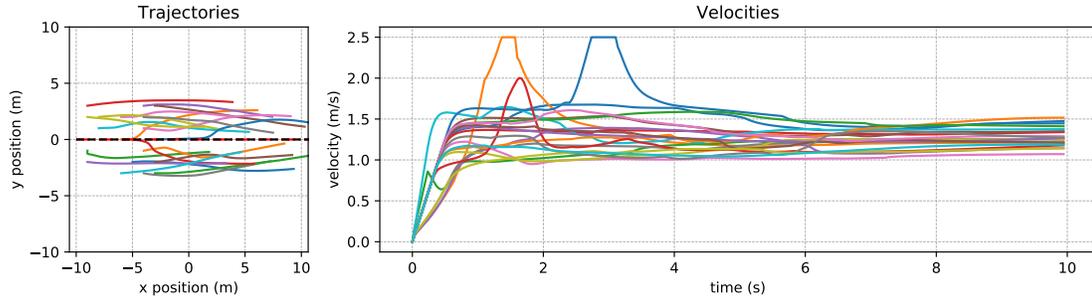


Figure A.3: Trajectories and velocities of back interaction. Left: Trajectories of pedestrians (solid lines) and the vehicle (dashed line) in the simulation. Right: The evolution of velocities of all pedestrians in the simulation

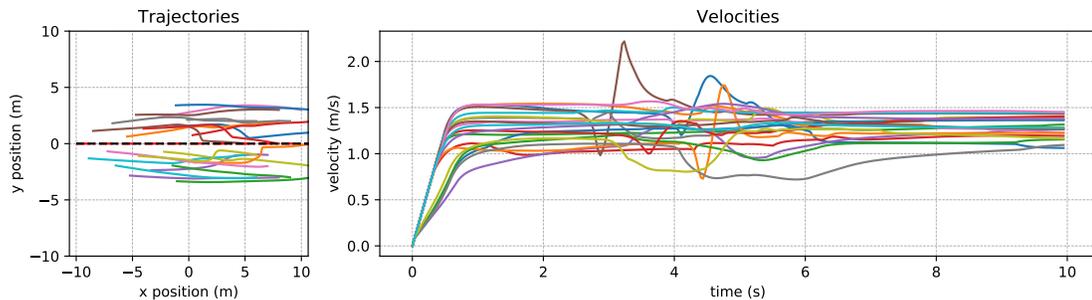


Figure A.4: Trajectories and velocities of front interaction. Left: Trajectories of pedestrians (solid lines) and the vehicle (dashed line) in the simulation. Right: The evolution of velocities of all pedestrians in the simulation

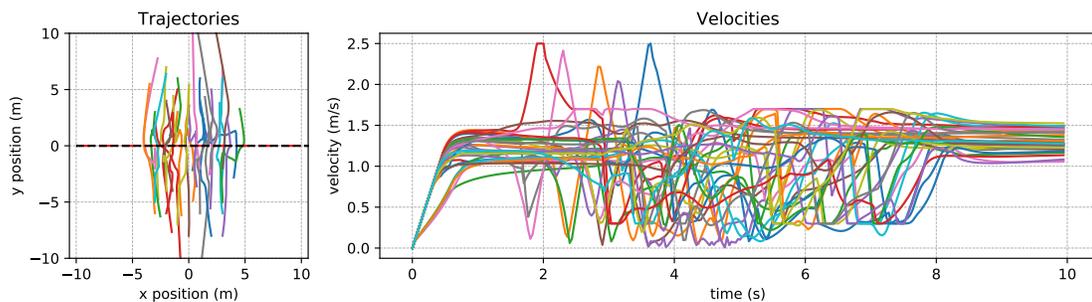


Figure A.5: Trajectories and velocities of lateral interaction. Left: Trajectories of pedestrians (solid lines) and the vehicle (dashed line) in the simulation. Right: The evolution of velocities of all pedestrians in the simulation

Appendix B

Research Products for this Project

B.1 Journal Publications

Dongfang Yang, Ümit Özgüner, and Keith Redmill, “A Social Force Based Pedestrian Motion Model Considering Multi-Pedestrian Interaction with a Vehicle”, *ACM Transactions on Spatial Algorithms and Systems (TSAS)*, 6:2(2020), pp. 1–27.

B.2 Conference Publications

Dongfang Yang, Arda Kurt, Keith Redmill, Ümit Özgüner, “Agent-based microscopic pedestrian interaction with intelligent vehicles in shared space”, *Proceedings of the 2nd International Workshop on Science of Smart City Operations and Platforms Engineering*, pp. 69–74, 2017.

Dongfang Yang, John Maroli, Linhui Li, Menna El-Shaer, Bander Jaber, Keith Redmill, Ümit Özgüner, and Füsün Özgüner, “Crowd motion detection and prediction for transportation efficiency in shared spaces”, *2018 IEEE International Science of Smart City Operations and Platforms Engineering in Partnership with Global City Teams Challenge (SCOPE-GCTC)*, pp. 1–6, 2018.

Dongfang Yang, Ümit Özgüner, and Keith Redmill, “Social force based microscopic modeling of vehicle-crowd interaction”, *2018 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1537–1542, 2018.

Dongfang Yang, Linhui Li, Keith Redmill, and Ümit Özgüner, “Top-view trajectories: A pedestrian dataset of vehicle-crowd interaction from controlled experiments and crowded campus”, *2019 IEEE Intelligent Vehicles Symposium (IV)*, pp. 899–904, 2019.

B.3 Thesis

Rayan El Helou, Agent-based Modelling of Pedestrian Microscopic Interactions, Masters Thesis, The Ohio State University, 2016.

B.4 Datasets

CITR Ohio State University dataset is available at <https://github.com/dongfang-steven-yang/vci-dataset-citr>

Dalian Institute of Technology dataset is available at <https://github.com/dongfang-steven-yang/vci-dataset-dut>

Bibliography

- [1] Alexandre Alahi, Kratarth Goel, Vignesh Ramanathan, Alexandre Robicquet, Li Fei-Fei, and Silvio Savarese. Social lstm: Human trajectory prediction in crowded spaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 961–971, 2016.
- [2] Gianluca Antonini, Michel Bierlaire, and Mats Weber. Discrete choice models of pedestrian walking behavior. *Transportation Research Part B: Methodological*, 40(8):667–687, 2006.
- [3] Bani Anvari, Michael GH Bell, Aruna Sivakumar, and Washington Y Ochieng. Modelling shared space users via rule-based social force model. *Transportation Research Part C: Emerging Technologies*, 51:83–103, 2015.
- [4] Stefan Becker, Ronny Hug, Wolfgang Hübner, and Michael Arens. An evaluation of trajectory prediction approaches and notes on the trajnet benchmark. *arXiv preprint arXiv:1805.07663*, 2018.
- [5] Ben Benfold and Ian Reid. Stable multi-target tracking in real-time surveillance video. In *CVPR 2011*, pages 3457–3464. IEEE, 2011.
- [6] Gabriele Bernardini, Enrico Quagliarini, and Marco D’Orazio. Towards creating a combined database for earthquake pedestrians’ evacuation models. *Safety science*, 82:77–94, 2016.
- [7] Hao Cheng and Monika Sester. Mixed traffic trajectory prediction using lstm-based models in shared space. In *The Annual International Conference on Geographic Information Science*, pages 309–325. Springer, 2018.
- [8] CityMobil2. Cities demonstrating automated road passenger transport. <http://www.citymobil2.eu/en/> [Online; accessed 2017-02-25], 2016.
- [9] Winnie Daamen and Serge Hoogendoorn. Calibration of pedestrian simulation model for emergency doors by pedestrian type. *Transportation Research Record*, 2316(1):69–75, 2012.
- [10] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.
- [11] Piotr Dollar, Christian Wojek, Bernt Schiele, and Pietro Perona. Pedestrian detection: An evaluation of the state of the art. *IEEE transactions on pattern analysis and machine intelligence*, 34(4):743–761, 2012.

- [12] Rayan El Helou. Agent-based modelling of pedestrian microscopic interactions. Master's thesis, The Ohio State University, 2016.
- [13] Markus Enzweiler and Dariu M Gavrilă. Monocular pedestrian detection: Survey and experiments. *IEEE transactions on pattern analysis and machine intelligence*, 31(12):2179–2195, 2009.
- [14] Ramsey Faragher et al. Understanding the basis of the kalman filter via a simple and intuitive derivation. *IEEE Signal processing magazine*, 29(5):128–132, 2012.
- [15] Claudio Feliciani, Luca Crociani, Andrea Gorrini, Giuseppe Vizzari, Stefania Bandini, and Katsuhiko Nishinari. A simulation model for non-signalized pedestrian crosswalks based on evidence from on field observation. *Intelligenza Artificiale*, 11(2):117–138, 2017.
- [16] Xiangjun Fu, Michael Vernier, Arda Kurt, Keith Redmill, and Umit Ozguner. Smooth: improved short-distance mobility for a smarter city. In *Proceedings of the 2nd International Workshop on Science of Smart City Operations and Platforms Engineering*, pages 46–51, 2017.
- [17] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
- [18] David Geronimo, Antonio M Lopez, Angel D Sappa, and Thorsten Graf. Survey of pedestrian detection for advanced driver assistance systems. *IEEE transactions on pattern analysis and machine intelligence*, 32(7):1239–1258, 2010.
- [19] Andrea Gorrini, Luca Crociani, Giuseppe Vizzari, and Stefania Bandini. Observation results on pedestrian-vehicle interactions at non-signalized intersections towards simulation. *Transportation research part F: traffic psychology and behaviour*, 59:269–285, 2018.
- [20] Agrim Gupta, Justin Johnson, Li Fei-Fei, Silvio Savarese, and Alexandre Alahi. Social gan: Socially acceptable trajectories with generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2255–2264, 2018.
- [21] Dirk Helbing, Illés Farkas, and Tamas Vicsek. Simulating dynamical features of escape panic. *Nature*, 407(6803):487, 2000.
- [22] Dirk Helbing and Peter Molnar. Social force model for pedestrian dynamics. *Physical review E*, 51(5):4282, 1995.
- [23] M. Himmelsbach, F. v. Hundelshausen, and H. J. Wuensche. Fast segmentation of 3d point clouds for ground vehicles. In *2010 IEEE Intelligent Vehicles Symposium*, pages 560–565, June 2010.
- [24] Serge P Hoogendoorn and Piet HL Bovy. Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological*, 38(2):169–190, 2004.

- [25] Anders Johansson, Dirk Helbing, and Pradyumn K Shukla. Specification of the social force pedestrian model by evolutionary adjustment to video tracking data. *Advances in complex systems*, 10(supp02):271–288, 2007.
- [26] Fatema T Johora and Jörg P Müller. Modeling interactions of multimodal road users in shared spaces. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 3568–3574. IEEE, 2018.
- [27] Robert Krajewski, Julian Bock, Laurent Kloeker, and Lutz Eckstein. The highd dataset: A drone dataset of naturalistic vehicle trajectories on german highways for validation of highly automated driving systems. In *2018 IEEE 21st International Conference on Intelligent Transportation Systems (ITSC)*, 2018.
- [28] Yoshiaki Kuwata, Justin Teo, Sertac Karaman, Gaston Fiore, Emilio Frazzoli, and Jonathan P How. Motion planning in complex environments using closed-loop prediction. In *Proc. AIAA Guidance, Navigation, and Control Conf. and Exhibit*, 2008.
- [29] Alon Lerner, Yiorgos Chrysanthou, and Dani Lischinski. Crowds by example. In *Computer graphics forum*, volume 26:3, pages 655–664. Wiley Online Library, 2007.
- [30] Robert V Levine and Ara Norenzayan. The pace of life in 31 countries. *Journal of cross-cultural psychology*, 30(2):178–205, 1999.
- [31] Peng Lin, Jian Ma, and Siuming Lo. Discrete element crowd model for pedestrian evacuation through an exit. *Chinese Physics B*, 25(3):034501, 2016.
- [32] Alan Lukežič, Tomáš Vojtř, Luka Čehovin Zajc, Jiří Matas, and Matej Kristan. Discriminative correlation filter tracker with channel and spatial reliability. *International Journal of Computer Vision*, 126(7):671–688, 2018.
- [33] J. M. Maroli, Ü. Özgüner, K. Redmill, and A. Kurt. Automated rotational calibration of multiple 3d lidar units for intelligent vehicles. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–6, Oct 2017.
- [34] D. Matti, H. K. Ekenel, and J. P. Thiran. Combining lidar space clustering and convolutional neural networks for pedestrian detection. In *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 1–6, Aug 2017.
- [35] Melanie Mitchell. *An introduction to genetic algorithms*. MIT press, 1998.
- [36] Betty J Mohler, William B Thompson, Sarah H Creem-Regehr, Herbert L Pick, and William H Warren. Visual flow influences gait transition speed and preferred walking speed. *Experimental brain research*, 181(2):221–228, 2007.
- [37] Mehdi Moussaïd, Dirk Helbing, and Guy Theraulaz. How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences*, 108(17):6884–6888, 2011.
- [38] Sergiu Nedeveschi, Silviu Bota, and Corneliu Tomiuc. Stereo-based pedestrian detection for collision-avoidance applications. *IEEE Transactions on Intelligent Transportation Systems*, 10(3):380–391, 2009.

- [39] U.S. National Highway Traffic Safety Administration (NHTSA). Traffic safety facts (2017 data) - pedestrians. <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812681> [Online; accessed 2019-09-30], 2019.
- [40] Umit Ozguner, Tankut Acarman, and Keith Alan Redmill. *Autonomous ground vehicles*. Artech House, 2011.
- [41] F Pascucci, N Rinke, C Schiermeyer, B Friedrich, and V Berkhahn. Modeling of shared space with multi-modal traffic using a multi-layer social force approach. *Transportation Research Procedia*, 10:316–326, 2015.
- [42] Stefano Pellegrini, Andreas Ess, Konrad Schindler, and Luc Van Gool. You’ll never walk alone: Modeling social behavior for multi-target tracking. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 261–268. IEEE, 2009.
- [43] Mark Pfeiffer, Giuseppe Paolo, Hannes Sommer, Juan Nieto, Rol Siegwart, and Cesar Cadena. A data-driven model for interaction-aware pedestrian motion prediction in object cluttered environments. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–8. IEEE, 2018.
- [44] N Rinke, C Schiermeyer, F Pascucci, V Berkhahn, and B Friedrich. A multi-layer social force approach to model interactions in shared spaces using collision prediction. *Transportation research procedia*, 25:1249–1267, 2017.
- [45] Alexandre Robicquet, Amir Sadeghian, Alexandre Alahi, and Silvio Savarese. Learning social etiquette: Human trajectory understanding in crowded scenes. In *European conference on computer vision*, pages 549–565. Springer, 2016.
- [46] Andreas Schadschneider, Wolfram Klingsch, Hubert Klüpfel, Tobias Kretz, Christian Rogsch, and Armin Seyfried. Evacuation dynamics: Empirical results, modeling and applications. *Encyclopedia of complexity and systems science*, pages 3142–3176, 2009.
- [47] Nicolas Schneider and Dariu M. Gavrila. Pedestrian path prediction with recursive bayesian filters: A comparative study. *german conference on pattern recognition*, 8142:174–183, 2013.
- [48] Robert Schönauer, Martin Stubenschrott, Weinan Huang, Christian Rudloff, and Martin Fellendorf. Modeling concepts for mixed traffic: Steps toward a microscopic simulation tool for shared space zones. *Transportation research record*, 2316(1):114–121, 2012.
- [49] Glenn Shafer. *A mathematical theory of evidence*, volume 42. Princeton university press, 1976.
- [50] Catapult Transport Systems. Self-driving pods. <https://ts.catapult.org.uk/current-projects/self-driving-pods/> [Online; accessed 2017-02-25], 2016.
- [51] Siyu Tang, Mykhaylo Andriluka, and Bernt Schiele. Detection and tracking of occluded people. *International Journal of Computer Vision*, 110(1):58–69, 2014.
- [52] Michael Vernier, Keith Redmill, Umit Ozguner, Arda Kurt, and Bilin Aksun Guvenc. Osu smooth in a smart city. In *2016 1st International Workshop on Science of Smart City Operations and Platforms Engineering (SCOPE) in partnership with Global City Teams Challenge (GCTC)(SCOPE-GCTC)*, pages 1–6. IEEE, 2016.

- [53] Valeria Vignali, Federico Cuppi, Ennia Acerra, Arianna Bichicchi, Claudio Lantieri, Andrea Simone, and Marco Costa. Effects of median refuge island and flashing vertical sign on conspicuity and safety of unsignalized crosswalks. *Transportation research part F: traffic psychology and behaviour*, 60:427–439, 2019.
- [54] Yao Xiao, Ziyou Gao, Yunchao Qu, and Xingang Li. A pedestrian flow model considering the impact of local density: Voronoi diagram based heuristics approach. *Transportation research part C: emerging technologies*, 68:566–580, 2016.
- [55] Hao Xue, Du Q Huynh, and Mark Reynolds. Ss-lstm: A hierarchical lstm model for pedestrian trajectory prediction. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1186–1194. IEEE, 2018.
- [56] Dongfang Yang, Arda Kurt, Keith Redmill, and Ümit Özgüner. Agent-based microscopic pedestrian interaction with intelligent vehicles in shared space. In *Proceedings of the 2nd International Workshop on Science of Smart City Operations and Platforms Engineering*, pages 69–74, 2017.
- [57] Dongfang Yang, Linhui Li, Keith Redmill, and Ümit Özgüner. Top-view trajectories: A pedestrian dataset of vehicle-crowd interaction from controlled experiments and crowded campus. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, pages 899–904. IEEE, 2019.
- [58] Dongfang Yang, John M Maroli, Linhui Li, Menna El-Shaer, Bander A Jabr, Keith Redmill, Füsün Özgüner, and Ümit Özgüner. Crowd motion detection and prediction for transportation efficiency in shared spaces. In *2018 IEEE International Science of Smart City Operations and Platforms Engineering in Partnership with Global City Teams Challenge (SCOPE-GCTC)*, pages 1–6. IEEE, 2018.
- [59] Dongfang Yang, Ümit Özgüner, and Keith Redmill. Social force based microscopic modeling of vehicle-crowd interaction. In *2018 IEEE Intelligent Vehicles Symposium (IV)*, pages 1537–1542. IEEE, 2018.
- [60] Dongfang Yang, Ümit Özgüner, and Keith Redmill. A social force based pedestrian motion model considering multi-pedestrian interaction with a vehicle. *ACM Transactions on Spatial Algorithms and Systems (TSAS)*, 6(2):1–27, 2020.
- [61] Francesco Zanlungo, Tetsushi Ikeda, and Takayuki Kanda. Social force model with explicit collision prediction. *EPL (Europhysics Letters)*, 93(6):68005, 2011.
- [62] Weiliang Zeng, Peng Chen, Guizhen Yu, and Yunpeng Wang. Specification and calibration of a microscopic model for pedestrian dynamic simulation at signalized intersections: A hybrid approach. *Transportation Research Part C: Emerging Technologies*, 80:37–70, 2017.
- [63] Weiliang Zeng, Hideki Nakamura, and Peng Chen. A modified social force model for pedestrian behavior simulation at signalized crosswalks. *Procedia-Social and Behavioral Sciences*, 138:521–530, 2014.
- [64] Hengshuang Zhao, Xiaojuan Qi, Xiaoyong Shen, Jianping Shi, and Jiaya Jia. Icn-net for real-time semantic segmentation on high-resolution images. *arXiv preprint arXiv:1704.08545*, 2017.

- [65] Liang Zhao and Charles E Thorpe. Stereo-and neural network-based pedestrian detection. *IEEE Transactions on intelligent transportation systems*, 1(3):148–154, 2000.
- [66] Bolei Zhou, Xiaogang Wang, and Xiaoou Tang. Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2871–2878. IEEE, 2012.