Towards a Smart Interaction Strategy in Driverless Minibuses for Optimizing Pedestrian Interaction

Thesis approved by

the Department of Computer Science

University of Kaiserslautern-Landau

for the award of the Doctoral Degree

Doctor of Engineering (Dr.-Ing.)

to

Qazi Hamza Jan

Date of the defense	:	8 August 2024
Dean	:	Prof. Dr. Christoph Garth
Reviewer	:	Prof. Dr. Karsten Berns
Reviewer	:	Prof. Dr. Mian M. Awais

DE-386

Acknowledgement

"To my mentor, Professor Karsten Berns, whose wisdom is like a lighthouse guiding ships through fog—thank you for illuminating my path. To my peers in the hallowed RRLab, who danced with me through storms of experiments and deadlines—your camaraderie has been my anchor. To the heartbeats at home, who bear the names Mom, Dad, and countless beloveds—your love, the wind beneath my wings."

Abstract

Navigating driverless minibuses through pedestrian-dense areas presents a significant challenge, exacerbated by the necessity for these vehicles to transport passengers efficiently to their destinations. For such vehicles to maintain a pace equal to or faster than walking speed, it is important to minimize unnecessary braking and avoid unwarranted stops. Achieving such navigation proficiency in pedestrian zones requires driverless minibuses to possess advanced interactive systems capable of signaling their intentions to pedestrians preemptively, thereby reducing potential misunderstandings.

This thesis focuses on the development and implementation of strategies for the efficient and safe operation of driverless minibuses in pedestrian environments. Although these vehicles are equipped with comprehensive safety systems designed to initiate emergency stops to avert collisions, such interventions can disrupt their smooth flow. It is essential for such vehicles to interpret pedestrian intentions and their awareness levels accurately to prevent the undue activation of emergency stop mechanisms and pauses. Moreover, it is vital to alert pedestrians who may be oblivious to the vehicle's proximity. To this end, the thesis explores various interaction strategies, including auditory signals, visual cues, and car control adjustments, to facilitate clear communication with pedestrians.

Smart interaction strategy is proposed to enhance interactions between driverless minibuses and pedestrians, incorporating four pivotal elements: pedestrian behavior identification, interaction zone establishment, decision-making processes, and the activation of interaction modules. These elements are foundational for achieving a sophisticated interaction level. Identifying pedestrian behavior is crucial for detecting their potential unawareness of the vehicle's proximity. Interaction zones are crafted to calculate a risk value based on pedestrian behavior and proximity, guiding the decision-making process within the navigable space. Following this assessment, interaction modules are deployed to actively engage with pedestrians, embodying the full spectrum of the interaction strategy developed in this study.

The integration of these modules into a standard navigation system enables effective operation in high pedestrian traffic conditions. The smart interaction strategy was evaluated using a practical system referred to as "*Autobus*", which is analogous to a driverless minibus. Initial tests of these components were conducted separately to verify their efficacy, followed by their amalgamation into the system to realize the intended outcomes. Experimental findings confirm that the implemented system enables Autobus to communicate effectively with pedestrians across various scenarios, significantly reducing the frequency of unnecessary stops.

Contents

1. Introduction

The concept of smart cities is evolving as a vision of transformation for the future in dynamic urban development environments. This thesis analyzes the main issues of the development of smart cities: the integration of driverless minibuses into pedestrian areas. The development of smart cities is moving from traditional urban projects, which rely heavily on traditional means, to a future in which technology can coexist with the environment and society. Smart cities are innovative urban models that integrate physical devices with information and communication technologies (ICTs) and the Internet of Things (IoTs), optimizing urban functions, increasing economic growth and improving the quality of life of citizens. The model focuses on efficiency, sustainability and connectivity in order to create a more interactive and responsive urban environment. Smart cities are based on the use of data and technology to simplify and optimize urban services and infrastructure.

In the rapidly developing landscape of intelligent cities, when technology and urban planning are converged to improve quality of life, pedestrian areas are an essential element. These areas are often rich in cultural and social vitality, and not only complement the technological advances of intelligent cities, but also advance the human aspects of urban development. One of these views can be seen in Figure 1.1. These areas have a profound impact on environmental sustainability, social interaction and economic vitality and are an essential component of smart urban planning.

The pedestrian area is designed mainly to promote comfortable and safe walking environments and is a city space reserved exclusively for pedestrians. These areas often range from small public squares to all towns and cities, and are characterized by significant restrictions or total bans on car traffic. The concept of pedestrianization is to transform a street or a district into a pedestrian-only space, improve pedestrian mobility and safety, stimulate local economic activity and improve the aesthetic quality of the urban environment. The main attraction of the pedestrian area is the freedom of walking and cycling. Such areas facilitate safe environments, enabling families to allow children to move freely without concern for safety. The integration of these areas enables all demographic groups from children to the elderly, ensuring safe walking for all pedestrian classes. Recognizing their various benefits, such as environmental, economic and health benefits, urban planners

Figure 1.1: An example representation of the pedestrian zone of smart city that presents a harmonious mixture of advanced urban planning and human-centered design. The visualization shows seamless integration between pedestrian and automated public transport, demonstrating environmental sustainability and encouraging social interaction. The area is characterized by pedestrian spaces. This approach emphasizes the fundamental role of pedestrian areas in improving economic vitality and social well-being, thereby combining their status as an essential part of the Smart cities framework. [AI-generated]

are increasingly focused on the development and expansion of these areas, as highlighted in [Soni 16]. Taking into account these short- and long-term advantages, cities around the world are increasingly adopting pedestrianization strategies, as Yassin et al.Yassin et al. [Yassin 19] has demonstrated. This change leads to an increase in the length and number of pedestrian zones. With the limitation or modification of conventional means of transport, a safe mobility solution for the elderly and people with disabilities is required.

1.1 The Role of Sustainable Transportation in Pedestrian Spaces

Since the population is growing rapidly, the need for green transportation is increasing, and it is more important to find effective ways of moving people to pedestrian areas. Minibuses play a major role in this. As shown in Figure 1.1, minibuses connect different parts of the city and are an essential part of sustainable transportation. Minibuses make the city more accessible, reduce carbon emissions and keep pedestrian areas lively and easy for everyone to get there. They meet the complex needs of urban mobility and contribute to the creation of a more sustainable urban environment.

Enhanced Comfort The comfort of transport, especially for minibuses, has a major impact on passengers' happiness and overall experience. This comfort includes the speed with which the vehicle can reach the destination, even in extreme weather. It ensures smooth travel for passengers without facing excessive overcrowding or difficult conditions. But comfort not only involves physical aspects, it also includes a comfortable and stress-free journey. People want reliable services and safe places during their travels, as well as comfortable seats. When public transport, such as minibuses, is considered comfortable and practical, people are more likely to use them than reliance on private vehicles. This change helps to make urban transportation more sustainable and efficient. Investing in transportation is therefore not only a matter of making passengers happy. It also encourages more people to use public transport on their daily journeys, which is beneficial to the entire urban transport system.

Improved Accessibility for the Elderly and Disabled By integrating minibuses into city transportation networks, the access of elderly passengers and disabled people is greatly improved and the city becomes more inclusive.Minibuses help users stay independent by connecting directly to essential services and improving mobility.They provide direct transportation to medical facilities, shopping centres and recreational areas, reducing the need for long and difficult walks or transfers.This convenience is the key to maintaining active and social activities for older people, which is beneficial to their physical and mental health.In addition, the small size and flexibility of minibuses often allow door-to-door service and make it easier for older people to travel.

First and Last Mile Transit Minibuses play an important role in bridging the first and last mile in pedestrian areas and provide flexible and efficient solutions to one of the most important public transport challenges. The first mile refers to the first part of the traveler's journey from the starting point to the main hub, while the last mile refers to the final segment of the hub's journey to the destination. These segments may discourage people from using public transportation because of inconveniences such as getting to and leaving transit stops. Smaller than traditional buses, smaller and better maneuverable minibuses can navigate more effectively through urban areas, narrow streets, densely populated areas, etc. This will enable them to offer more direct and personalized routes, reduce walking distances to and from the major transit stations and thus greatly improve public transportation convenience.

On-demand Service Moreover, minibuses can operate on dynamic routes and schedules, including on-demand services, further minimizing wait times and making public transportation more appealing. By efficiently connecting passengers with major transit nodes, such as metro stations, bus terminals, and train stations, minibuses ensure a seamless travel experience. This connectivity encourages more people to opt for public transit over personal vehicles, contributing to reduced traffic congestion, lower emissions, and a more sustainable urban environment.

In this context, the focus of this thesis is to examine the integration of driverless minibus within pedestrian zones as a critical component of smart city development. It delves into the challenges and opportunities presented by this integration, exploring how such vehicles can coexist with pedestrians in a shared space. driverless minibuses constitute a major focus of research and development within the automotive sector.

1.2 Advancements in driverless Minibus Technology

In response to growing demand for better mobility in smart city shared spaces, numerous companies are actively involved in the development of specially designed specialized vehicles for pedestrian areas. These initiatives are part of a broader effort to integrate advanced transportation solutions that can seamlessly exist with high density pedestrian areas, thereby contributing to overall efficiency and sustainability of pedestrian areas. EasyMile [1] is the leader in the sector, providing fully autonomous services at Level 4 of the autonomous driving spectrum known as the EZ10 self-driving shuttle shown on the left in Figure 1.2. This level means that vehicles can operate independently without human intervention in most scenarios. EasyMile's applications extend to various environments such as airports, parks and industrial areas. The technology that supports these shuttles is designed to navigate a wide range of scenarios autonomously and with minimum supervision. Furthermore, these vehicles' robust design is able to operate efficiently in mixed traffic environments and different weather conditions, and they demonstrate significant progress in autonomous urban transport.

The Karlsruher Verkehrsverbund (KVV) [2] details their involvement in the autonomous driving project EVA-Shuttle, part of the EU project SHOW. The initiative focuses on integrating autonomous shuttles into public transport. They are using shuttles from EasyMile. These electric, emission-free minibuses operate in the Weiherfeld-Dammerstock area and can be booked via an app for free rides. Operational from February 2023, the shuttles run on weekends, showcasing KVV's commitment to innovative, sustainable urban mobility solutions.

Navya[3], a leader in autonomous mobility solutions, offers self-driving vehicles primarily for passenger transportation. Their product, the Autonom® Shuttle Evo 1.2 in the middle of the figure, is an electric self-driving shuttle designed for urban and private site use, emphasizing efficiency and environmental friendliness. They have sold around 160 units around the world. Navya's technology, Navya Driver®, boasts advanced software and a unique sensor kit, ensuring optimal autonomous navigation. Their solutions aim to address urban transport challenges, optimize travel time, reduce carbon footprint, and enhance mobility in various settings like cities, airports, and campuses.

Zoox[4] has achieved a significant milestone in driverless minibus technology by introducing a purpose-built robotaxi shown in the right of figure 1.2 that operates without manual controls. This vehicle, designed exclusively for riders, has successfully undergone extensive testing and has received the necessary certifications to operate autonomously on public roads. This development represents a major step forward in the driverless minibus industry and is a testament to Zoox's vision and commitment to reinventing transportation.

[1] https://easymile.com/
[2] https://www.kvv.de/index.html
[3] https://www.navya.tech/en/
[4] https://zoox.com/journal/publicroads/

Figure 1.2: The image shows three driversless minibuses designed for pedestrian-dense smart city environments and represents the forefront of autonomous driving technology. On the left, EasyMile's self-driving EZ10 shuttle, which operates at Level 4 automation, shows its ability to operate independently in different environments such as airports and parks. In the middle, Navya's Electric Autonom® Evo electric shuttle demonstrated its application to cities and private sites, highlighting its worldwide deployment, with 160 units in service. On the right, Zoox's revolutionary robot axis, which lacks manual control, highlights an important advance in driverless minibus technology that is certified to operate on public roads. These developments combine the efforts of companies and urban transportation initiatives to integrate sustainable and efficient mobility solutions into the common spaces of growing smart cities.

1.3 Challenges of Driverless Minibuses within Pedestrian Zones

The integration of driverless minibuses into pedestrian zones introduces a spectrum of complex challenges, encompassing aspects of autonomous navigation and pedestrian safety.

Pedestrian Behavior Prediction: Given that humans constitute the primary occupants of pedestrian zones, accurately forecasting pedestrian behavior in shared environments poses a considerable challenge. The unpredictability of pedestrian movements, with individuals having the freedom to walk in any direction across the breadth of pedestrian zones, complicates the task of ensuring accurate, real-time behavioral prediction essential for the safe navigation of driverless minibuses. The randomness of pedestrian behavior in pedestrian zones, presents a complex challenge for the planning and operation of autonomous transportation systems. Pedestrian movements are inherently unpredictable due to a variety of factors such as individual choices, group dynamics, environmental influences, and spontaneous decisions. Unlike vehicles that typically follow predefined paths and rules, pedestrians may change their direction, speed, and patterns of movement without warning. This unpredictability is influenced by social interactions, distractions (such as mobile devices), or obstacles within the environment. Understanding and anticipating this randomness requires sophisticated algorithms and sensors capable of real-time data processing and decision-making. The challenge lies not only in detecting pedestrians and predicting their immediate actions but also in interpreting subtle cues that might indicate a sudden change in behavior. For autonomous vehicles operating in such environments, this unpredictability necessitates the development of advanced machine learning models that can learn from vast amounts of data and adapt to new scenarios they have not been explicitly programmed for.

Pedestrian Risk Assessment: Analyzing the risk posed by pedestrians for autonomous minibuses in pedestrian zones presents multifaceted challenges. At the core, the unpredictability and complexity of human behavior create substantial hurdles, compounded by the dynamic and often unstructured nature of these environments. The interaction between multiple agents—pedestrians, cyclists, and other vehicles—adds layers of complexity to risk assessment. The accuracy of such analyses is further constrained by the limitations of sensors and perception algorithms, which can struggle under adverse conditions or due to inherent technical limitations. Moreover, ensuring that risk assessment models are trained on sufficiently diverse data to accurately reflect the wide range of pedestrian behaviors encountered in different cultural and urban contexts is a daunting task. Legal and ethical considerations also play a critical role, as the algorithms must navigate dilemmas on how to prioritize safety in scenarios where risks cannot be entirely mitigated. Ultimately, building public trust hinges on the transparency and reliability of these risk assessment methods, underscoring the necessity for a holistic, multidisciplinary approach that melds advances in technology with ethical and urban planning insights to ensure the harmonious coexistence of autonomous minibuses and pedestrians in urban spaces.

Interaction with Pedestrians: In traditional vehicular settings, non-verbal cues play a crucial role in facilitating communication between drivers and pedestrians. Eye contact and hand gestures from a driver can significantly reassure pedestrians about their intentions, such as yielding the right of way or signaling that it's safe to cross the road. These gestures help build a rapport and trust between human road users, contributing to a smoother and safer traffic flow. However, the introduction of driverless minibuses disrupts this established mode of communication. With the absence of a human driver, pedestrians might find themselves at a loss, unsure of how to interpret the vehicle's intentions. This uncertainty can lead to hesitation and potentially unsafe situations, as the natural human instinct to seek affirmation from another's gaze or gesture is unmet. The challenge, then, is for autonomous vehicle technology to bridge this communication gap.

Designing of Universal Interaction Signals: In addressing the complexity of pedestrian interaction, a significant challenge emerges: the creation of signals that are universally understandable and intuitive. This challenge is multifaceted, necessitating a consideration of the wide range of cultural backgrounds, languages, and individual capabilities. Within the contexts of safety, urban planning, and the integration of novel technologies into public spaces, the importance of this issue cannot be overstated. Signals designed to direct, inform, or caution pedestrians must effectively bridge cultural and linguistic divides to achieve universal applicability. For example, the color red, often linked with cessation or hazard in numerous Western societies, may symbolize fortune and prosperity within certain Asian cultures. Such cultural variance underscores the need for meticulous selection of symbols, colors, and signals to prevent misinterpretation and guarantee the safety of all individuals involved. Additionally, the endeavor to make signals accessible extends to the accommodation of diverse literacy levels and language proficiencies among the pedestrian populace. Reliance on textual directions could potentially exclude those who are unable to read or do not understand the language employed, thus compromising the efficacy of the intended signals.

Mapping of Pedestrians: Integration of pedestrian dynamics into terrain mapping poses unique challenges compared to traditional mapping of vehicular paths and obstacles. Traditional mapping techniques that are good at cataloging physical terrain characteristics and identifying static and dynamic obstacles are inadequate when it comes to complex behavioral patterns and pedestrian movements in shared spaces. As mentioned above, one of the main challenges is the unpredictable nature of pedestrian behavior. Unlike static obstacles, cars usually follow predictable routes and habits based on traffic rules, and pedestrians have unique perception abilities that can change direction and stop suddenly. They navigate the environment not only on the basis of physical obstacles, but also on the basis of social indications, personal preferences and spontaneous decisions. This uncertainty requires a more sophisticated approach to mapping system representation.

Navigation of driverless minibuses: The navigation of driverless minibuses in pedestrian zones is particularly challenging due to the lack of lane divisions, narrow pathways, the constant presence of pedestrians, and the absence of formal traffic regulations. These driverless minibuses must be adept at interpreting and adhering to social protocols and informal rules to effectively navigate these shared spaces. Moreover, these systems must also contend with the need for real-time decision-making, requiring algorithms capable of processing data and evaluating risks swiftly to respond to rapidly changing scenarios. Given that minibuses are designed to transport passengers, this introduces additional complexities to autonomous navigation. The autonomous driving system must prioritize passenger safety while also ensuring timely arrival at their destinations. It is imperative for the system to anticipate and mitigate the need for emergency braking well in advance, which could cause discomfort to passengers who are not cognizant of the external environment.

Safety of Pedestrians: Even though driverless minibuses in shared spaces are designed to follow social rules, there are times when pedestrian or vehicle safety could be at risk. For example, in areas where it's hard to see everything clearly, the sensors might miss a pedestrian or an obstacle. Kids, in particular, are more vulnerable because they're curious and like to explore, which can put them in danger around these vehicles.

Undefined and Variable Bus Stops: In pedestrian zones, movement and destination patterns differ greatly from regular urban traffic. Here, the idea of fixed bus stops is transformed to fit a more fluid and flexible transit system. This change is vital in areas mainly designed for walking, where people have many different places to go that fixed stops can't always serve. In these zones, minibuses often use a flexible routing system with on-demand stops or service areas instead of fixed points. This setup allows for a more personalized transit experience, meeting the diverse needs of users heading to various spots like shops, cafes, cultural sites, and parks. It also improves accessibility, letting passengers get on and off closer to where they actually need to be, cutting down on extra walking or transfers.

Testing of Autonomous Driving in Pedestrian Zones: Given the challenges above, there are a few major hurdles when it comes to testing new algorithms for autonomous vehicles. One major issue is safety. It is impractical to test out new algorithms in real-world settings without knowing they will work, especially with pedestrians around. When these systems are tested in simulations, they run into another problem: it's hard to accurately mimic how pedestrians behave. Simulations can't capture all the randomness and variety in human behavior. Because of this, there's a lot of uncertainty about whether these systems are truly ready for the real world after just being tested in a simulated environment.

1.4 Contribution of This Thesis

In today's market, there are plenty of products aimed at making transportation in pedestrian areas better, focusing on customer experience and fitting smoothly into shared spaces (Section 1.2). Even with these advancements, there's still a lot not known about how driverless minibuses interact with and ensure the safety of pedestrians. Currently, state-of-the-art systems use a basic, one-way communication setup for bus stop info and audio announcements. This fails to adequately engage with pedestrians, especially in areas without clear lanes or established driving rules. We need more advanced systems that can clearly communicate what the vehicle's intentions are. This would help build trust and ensure safety in shared spaces. Also, the literature review shows that testing these driverless minibuses in simulated environments is common. While this is useful for seeing how people might respond in a controlled setting, it doesn't quite match up to the real-world experience. Physical tests with these minibuses, though riskier, bring out more genuine human reactions. This highlights the gap between simulated tests and real-world interactions. To bridge this gap, there is a need to focus on developing better interaction mechanisms and conduct more physical tests. This will help ensure that driverless minibuses are safe and effective in pedestrian zones.

The main objective of this thesis is to address research gaps by creating a system that improves transport and autonomy for people relying on pedestrian-assisted mobility. These self-contained minibuses are designed to move safely and efficiently through pedestrian areas, which are complicated by their common nature. As has been said in previous posts, the main problems faced by these systems are related to interaction with pedestrians in tight and unstructured environments. To address these challenges, we are developing an innovative *"smart interaction"* strategy. This strategy is aimed at facilitating effective communication between driverless minibuses and pedestrians. The main purpose of intelligent interactions is to improve the safety and efficiency of driverless minibuses in pedestrian areas. Therefore, the research statement of the thesis is as follows:

To facilitate effective and intelligent interaction between driverless minibuses and pedestrians, ensuring efficient and safe navigation.

This emphasizes the development of a sophisticated communication system aimed at enabling driverless minibuses to interact dynamically with pedestrians. This interaction is crucial for the vehicles' ability to navigate pedestrian zones safely and efficiently. The term "effective and intelligent interaction" suggests the use of advanced technologies, including artificial intelligence, sensors, and machine learning algorithms, to interpret pedestrian behavior and predict their movements. By doing so, the system enhances the vehicle's responsiveness and decision-making capabilities in real-time, thereby ensuring both the safety of pedestrians and the smooth operation of the autonomous minibus within shared pedestrian zones.

Smart Interaction encompasses more than merely the emission of arbitrary signals by the vehicle. Rather, it should entail a sophisticated, context-sensitive procedure that incorporates an analysis of diverse factors, including pedestrian behaviors, the ambient environment, and potential navigational routes. This system should engage in a comprehensive process, starting with the identification of pedestrian behaviors and culminating in the vehicle's tailored response to these observed behaviors. Hence, *Smart Interaction* necessitates a multi-layered, technologically advanced approach to ensure seamless and safe integration into human-centric environments. This system should be built upon several key technical pillars:

Sensing and Perception: The foundation of Smart Interaction is a robust sensing and perception system that detect, classify, and track pedestrians in real time, distinguishing between static and dynamic obstacles while accurately predicting pedestrian behaviors. Novel machine learning models, particularly those using deep learning, is trained on datasets to recognize and predict pedestrian behaviors in varying contexts, enhancing recognition.

Context-Aware Decision: Beyond mere detection, the system incorporates context-aware processing capabilities. This involves analyzing the context of the vehicle's environment, including understanding pedestrian density, position, interpreting traversable pathways, recognizing unawareness, and adapting to crowd of people.

Adaptive Control System: Adapting vehicle control, capable of modifying its behavior based on the detection of pedestrian risk and available driving area for the vehicle. Selecting best path based on traversable path availability.

Communication with Visual and Auditory Cues: Effective communication mechanisms which are essential for signaling the vehicle's intentions to pedestrians. This involves visual signals (e.g., LED displays showing crossing symbols) and auditory cues. Controlling the content and activation of such modules based on pedestrian risk and behavior.

Hardwired Safety System: The development of a safety system intricately integrated with the braking mechanism is proposed to guarantee immediate cessation in response to unforeseen circumstances, including erroneous pedestrian predictions, abrupt pedestrian incursions in the vehicle's path, or delays emanating from the navigation architecture.

Testing Environment: The practical implementation of the concepts presented in this thesis is exemplified through the deployment of an driverless minibus, termed as ***"Autobus"***, as depicted in Figure 1.3. Comparable in dimensions to existing state-of-the-art systems, the Autobus is engineered to facilitate the transportation of individuals across the RPTU Kaiserslautern-Landau campus. Distinguished by its capability to steer in both directions, this feature renders the Autobus exceptionally adept at maneuvering within pedestrian-dense areas. Comprehensive details regarding the system's specifications and its operational context are elaborated upon in Section 3. In addition, a simulation environment has been established for initial concept validation tests.

Figure 1.3: The "Autobus", a driverless minibus deployed at RPTU Kaiserslautern-Landau University, is a practical example of the concept of autonomous transportation discussed in this thesis. The image shows the Autobus in its operational position and shows its compact design, which can navigate high pedestrian traffic areas with bidirectional driving capabilities. The deployment of Autobus to transport individuals within the campus environment is a direct application of advanced autonomous systems described in this thesis.

1.5 Thesis Structure:

This thesis explores the interaction requirements between Autobus and pedestrians in pedestrian zones, grounding its investigation in human psychology and the development of an effective interactive system between vehicle and pedestrians. The document is organized as follows:

Chapter 2 delineates the fundamental concept underlying this research, emphasizing the essential components requisite for the implementation of *Smart Interaction* within the Autobus in pedestrian settings. Leveraging insights from the psychology of human drivers, this section elucidates the process of translating principles of human interaction to the context of a driverless minibus. Furthermore, it presents an overview of the critical four components necessary for enabling effective communication between the Autobus and pedestrians. The four components are: Pedestrian activity; Interaction fields; Decision making; and Interaction interface.

Chapter 3 offers an in-depth analysis of the test vehicle's design, i.e., Autobus, employing it as a case study to evaluate the interaction concept proposed in this thesis. This chapter outlines the strategic placement of sensors on the vehicle and discusses the integration of a hardwired safety system, specifically tailored for such driverless minibuses. Additionally, it provides a comprehensive description of the testing environment to elucidate the identical challenges present in the designated pedestrian zones. Moreover, the chapter presents a visualization of the simulation system developed for preliminary testing, further illustrating the research's methodological approach.

Chapter 4 presents the concept of pedestrian activity recognition as a foundational element of *"Smart Interaction"*. After examining contemporary techniques, this chapter delves into the pedestrian detection and activity recognition approaches formulated in this research, concluding with an exposition of the findings. The methodology employed involves analyzing 3D skeleton joint points alongside time series data to interpret pedestrian behaviors.

Chapter 5 explores the creation of interaction fields, a pivotal component of the interaction concept. Drawing inspiration from the varied scenarios encountered in pedestrian zones, this chapter details the methodologies and design principles employed in establishing interaction fields that cater to both vehicles and pedestrians.

Chapter 6 investigates the decision-making mechanisms in vehicles integrated with the interaction system. It scrutinizes the impact of planning parameters on interactive navigation, utilizing interaction modules to enhance communication with pedestrians. The chapter adopts a tentacle navigation strategy to determine the optimal control approach for the vehicle.

Chapter 7 discusses the interaction interface developed for this thesis. It transitions from a general discussion on human-robot interaction to focus on vehicle-pedestrian interaction dynamics, addressing the optimization of these interactions and the selection of content for the interaction modules.

Chapter 8 details the results of the fully realized concept introduced in this thesis. It examines various scenarios encountered while navigating the campus with the Autobus and discusses how the implemented concept contributes to improved driving performance.

Chapter 9 concludes the thesis by summarizing the findings and discussing the implications of the research. It also offers a perspective on potential future directions for further investigation.

Appendix A provides the technical description of the sensors and modules for safety system mounted on the Autobus as explained in Chapter 3. It provides the overall schematics and integration of the safety modules.

Appendix B presents the questionnaire for testing the interaction modules discussed in Chapter 7. It lists the questions designed to evaluate the effectiveness and usability of the content for the interacting modules.

2. Establishing Core Principles of Interaction for driverless Minibus in Pedestrian Zones

The previous chapter (Chapter 1) describes the complexity associated with the navigation of driverless minibuses in pedestrian areas and gives an overview of the current state-of-the-art systems, focusing on their general applicability. Despite these advances, there is still a large gap in research to ensure the reliable operation of such vehicles. In order to fill this gap and address these challenges, this chapter proposes a new methodology to conceptualize autonomous driving from the point of view of human experience. This approach includes a comprehensive investigation and interpretation of human drivers' behaviors in pedestrian areas to apply these insights to the development and optimization of driverless minibuses.

The discussion begins with the introduction of *"**Smart Interaction**"* as the main contribution to this paper. It then studies human driving psychology and explores cognitive processes and decision-making strategies used by human drivers in pedestrian-intensive areas. This analysis not only illuminates human driving behavior but also provides a framework for improving the autonomy of driving in pedestrian areas. Then, we examine the complexity of the interaction between humans and vehicles. This involves a detailed study of the interaction between driverless minibuses and pedestrians, and the breakdown of their key components and dynamics. Factors such as communication cues, safety protocols, and pedestrian perception of the vehicle are considered—all key factors in this interaction.

2.1 Enhancing Vehicle-Pedestrian Communication via Smart Interaction in Driverless Minibuses

In environments where the minibus is not driven near the pedestrian, especially in areas without clear markings, the risk of conflict is high. In order to safely navigate around these vehicles, pedestrians need to understand the intentions of these vehicles. It is not enough to simply show the next stop of these driverless vehicles to ensure safety. On the contrary, these vehicles must actively communicate their intentions and react appropriately to pedestrians' actions. This paper introduced the concept of *"smart interaction"* as a foundational framework for establishing an effective communication strategy between driverless minibuses and pedestrians. Through *"Smart Interaction"*, the vehicle evaluates the current situation of its pedestrians and nearby pedestrians in order to provide the most appropriate action plan. The aim is to reach mutual understanding to ensure safety and efficiency in shared spaces. This approach not only improves pedestrian safety but also facilitates smoother integration of driverless minibuses into pedestrian-dense environments, paving the way for more harmonious interactions between people and vehicles.

Traditionally, vehicle and pedestrian interactions are managed by human drivers through visual signals, body language, and sometimes verbal communication. These include direct contact with the eye, gestures, indicators and brake lights. The introduction of driverless minibuses into these complex interaction networks poses significant challenges. Without a human driver, many forms of intuitive and spontaneous communication are lost, so a new approach is needed for the interaction between vehicles and pedestrians. Here the concept of smart interaction becomes essential. Smart interactions are inspired by the way humans communicate with pedestrians and aim to reproduce this level of understanding and response in minibuses without drivers. Advanced sensors, artificial intelligence, and visual indicators enable driverless minibuses to understand pedestrian actions and intentions and clearly communicate their intentions. This method integrates several elements into the traditional navigation framework and establishes a complete interaction cycle. The existing navigational framework is assumed to be fully autonomous, ensuring that the driverless minibus moves safely and efficiently, while interacting smoothly with pedestrians.

The conclusions drawn are that the basic principles of smart interaction are directly inspired by the dynamic and intuitive way human drivers interact with pedestrians. This thesis recognizes the importance of these interactions in maintaining safety and order in common spaces and proposes a technology-based approach to imitate this human aspect in the field of autonomous vehicle technology. Smart interaction uses a human-based strategy to bridge the gap between the impersonal nature of autonomous technology and pedestrian expectations. This approach not only promises to improve safety and efficiency but also fosters a sense of trust and understanding between humans and autonomous vehicles, thereby ensuring the smooth integration of this technology into our daily lives.

2.2 The Psychological Aspects of Human Drivers in Proximity to Pedestrians

Mimicking human driver interaction for driverless minibus interaction with pedestrians is paramount for enhancing safety, improving traffic efficiency, and fostering societal acceptance of autonomous technology. Human drivers utilize a combination of non-verbal cues, such as eye contact and gestures, along with vehicle dynamics, to communicate intentions to pedestrians, significantly reducing the likelihood of accidents. By adopting similar interaction mechanisms, such driverless minibuses can signal their intentions more clearly, making their behavior more predictable and understandable to pedestrians, thereby enhancing safety. Furthermore, human-like interaction strategies enable driverless minibuses to navigate pedestrian-dense environments more smoothly, optimizing traffic flow and reducing unnecessary stops. This not only contributes to overall traffic efficiency but also aids in the seamless integration of such driverless minibuses into pedestrian rich environments, where complex human-pedestrian interactions are the norm. Importantly, vehicles that interact in a manner akin to human drivers may appear more trustworthy to the public, which is crucial for the widespread acceptance of such vehicles.

Hence, it becomes important to understand the dynamics and often complex interactions between human drivers and pedestrians. A crucial aspect of the interaction between drivers and pedestrians is non-verbal communication [Ren 16]. It examines the impact of pedestrians' eye contact on the comfort level of drivers during crossing situations. The study [Ren 16] highlights the increasing incidence of people, particularly young individuals, using cellphones or earphones while crossing roads, which raises concerns about pedestrian safety and driver comfort. The research aims to understand how pedestrians' eye contact affects drivers' comfort boundaries, which is observable through car deceleration patterns. The study considers drivers' gender and the attempt to make eye contact as variables and measures the resulting changes in car speed. The methodology involved the use of people to simulate pedestrian behavior during the experiment. For this reason researchers in [Onkhar 21] have been addressing the research gap in objectively measuring eye contact between drivers and pedestrians, a non-verbal communication method thought to reduce accident risks. In [Rasouli 17, Sucha 17], the study analyzed pedestrian and driver behaviors focusing on non-verbal communication cues used at crossing points. It found that in over 90% of observed cases, pedestrians looked at oncoming cars before crossing, indicating the critical role of gaze and eye contact in pedestrian safety. One of the key findings identifies various patterns of pedestrian behavior in crossing scenarios, such as the sequence of looking, reacting to the driver's actions, and finally crossing. It categorizes these patterns based on whether attention or crossing action is observed. The findings in [Ghosh 22] reveals that participants' emotional responses, particularly valence, were significantly influenced by the observed pedestrian actions. Positive actions like hand waving and nodding were associated with positive valence, while negative actions such as impolite gestures or inattentiveness led to negative valence. Arousal ratings showed less variation across scenarios. A study by Onkhar et al. [Onkhar 22] shows the impact of driver's eye contact on pedestrian safe feeling. People crossing feel safe when having eye contact with the driver. This emphasizes the importance of non-verbal cues in these interactions. Such findings underscore the challenge for driverless minibuses, which must be programmed to interpret these subtle human behaviors and respond appropriately.

Another dimension is the expectation and anticipation in these interactions. Drivers often anticipate pedestrian actions based on context. For example, a driver might slow down when they see a pedestrian near a crosswalk, predicting the possibility of them deciding to cross. This anticipatory behavior is critical for safe driving but poses a significant challenge for autonomous drive programming, which must transfer this predictive capability.

It's also important to recognize the role of cultural and environmental contexts [Nordfjærn 11]. In dense urban environments, where pedestrian activity is high, drivers might be more attuned to the presence of pedestrians and thus more cautious. In contrast, in rural or suburban areas where pedestrian activity is less common, drivers may be less expectant of pedestrian interactions, leading to a different set of behavioral responses. For instance, a study [Papadimitriou 13] conducted in both urban and rural settings in Sweden indicated

Figure 2.1: An illustration of the dynamics of human-vehicle interaction in urban traffic. The image shows a scenario in which a pedestrian crossing an approaching vehicle and emphasizes the multifaceted nature of communication and negotiation occurring during these encounters. The main factors influencing this interaction include factors affecting pedestrians such as distance from the vehicle, speed of movement, directional intention, body language including eye contact and gestures, as well as personal characteristics such as clothing and style of movement. At the same time, vehicle-related factors include vehicle kinematics, vehicle type, sound signals such as engine noise and horns, visual signals such as headlights and turn signals, and driver behaviour characterized by engine revving and general driving styles. Other environmental variables such as visibility conditions, ambient noise levels and the number of road users also play an important role in determining the result of interactions. This complex interaction is essential to understand and improve safety protocols, traffic flow efficiency and the development of autonomous driving systems that can correctly interpret and react to human behaviour. [Habibovic 18]

that drivers in urban areas were more likely to slow down and yield to pedestrians, partly due to higher pedestrian density and the drivers' habitual exposure to such scenarios. This highlights the need for driverless minibus to adapt to varying environmental contexts such as width of the pathways in their interaction logic.

Hence, for a safe and enjoyable interaction on the road, it's crucial that all road users share a common understanding of the situation. If there's a discrepancy in how the situation is perceived or understood among road users, it's likely that disruptions and conflicts will arise [Endsley 95]. The relationship between drivers and pedestrians remains complicated, influenced by a range of factors, as depicted in Figure 2.1.

The observation in [Schmidt 09] was made that 84% of pedestrians initiate eye contact with drivers. It reported that pedestrians intending to cross the street engage in visual contact with oncoming drivers as a means of seeking "acknowledgment." This acknowledgment, particularly if the driver reciprocates the eye contact, is interpreted by pedestrians as a sign of being noticed and achieving a mutual recognition of their intent to cross.

In such scenarios, the interaction between the pedestrian and the driver serves as a crucial safety mechanism. When a pedestrian makes eye contact with a driver and receives acknowledgment, it creates a shared understanding that can significantly reduce the risk of accidents. This mutual recognition acts as an informal, yet powerful, agreement that the pedestrian has been seen and that the driver is aware of their intention to cross. Moreover, this interaction reflects the broader dynamics of human behavior in traffic systems, where non-verbal cues play a vital role in coordinating actions and intentions. It highlights the need for both drivers and pedestrians to remain alert and engaged with their surroundings, particularly at crosswalks and intersections where the potential for conflict is high.

Building on the insights from the human driving in proximity to pedestrians, it becomes evident that interaction with pedestrians is a fundamental solution for improving driving performance in pedestrian-rich environments. This finding is pivotal, suggesting that driverless minibuses need sophisticated systems to simulate or interpret such non-verbal cues effectively.

Developing driverless systems that can 'mimic' or recognize these non-verbal human interaction is crucial. This includes programming such systems to understand and respond to gestures, body language, and other subtle cues that humans use intuitively. For instance, driverless minibuses (specifically in shared spaces) might need to be equipped with extra modules and algorithms that can detect a pedestrian's gaze direction or body posture, allowing the vehicle to make more informed decisions about when to yield or proceed.

Central to this discussion is the understanding that driverless minibus must do more than merely navigate and avoid obstacles; they must actively communicate with pedestrians. This involves conveying intentions and acknowledging pedestrian presence in a manner similar to human drivers. The development and integration of smart interaction is therefore not just an enhancement but a necessity.

2.3 Realizing Smart Interaction for Driverless Minibuses

In light of the aforementioned motivation and critical elements, this thesis introduces a novel approach for implementing smart interaction in driverless minibus driven in pedestrian zone. Given that such vehicles transport passengers in areas predominantly occupied by pedestrians, it is essential to comprehend the underlying principles of their operation. This understanding is vital for optimizing the effectiveness of the interactive modules.

The smart interaction draws inspiration from our preceding work [Jan 20], wherein a framework for safe and efficient navigation for Autobus, through interactions with pedestrians, is proposed. This prior study, delineated in detail towards the conclusion of Chapter 3, utilized a simulated setting mirroring the RPTU campus environment to conduct experiments with a hypothetical driverless minibus. The investigation targeted specific pedestrian behaviors to evaluate a variety of scenarios. Central to this research is the adoption of a behavior-based architecture for basic vehicle navigation [Proetzsch 10], which has been specifically adapted for managing the shuttle's movement in pedestrian-dense areas. This navigational architecture was further enhanced through the integration of the Pedestrian Interaction System (PIS), a system that amalgamates interaction with pedestrians and evasion tactics, as illustrated in figure 2.2. The efficacy of the PIS was scrutinized by monitoring the shuttle's travel duration from its point of origin to its intended destination. Acknowledging the variable nature of pedestrian movements, the mean travel time was derived from conducting three distinct trials for each scenario under consideration. The research assessed four disparate scenarios, featuring 5, 10, 15, and 20 pedestrians, respectively. Comparative analysis was performed to gauge the performance in scenarios with the PIS enabled against those where it was not activated.

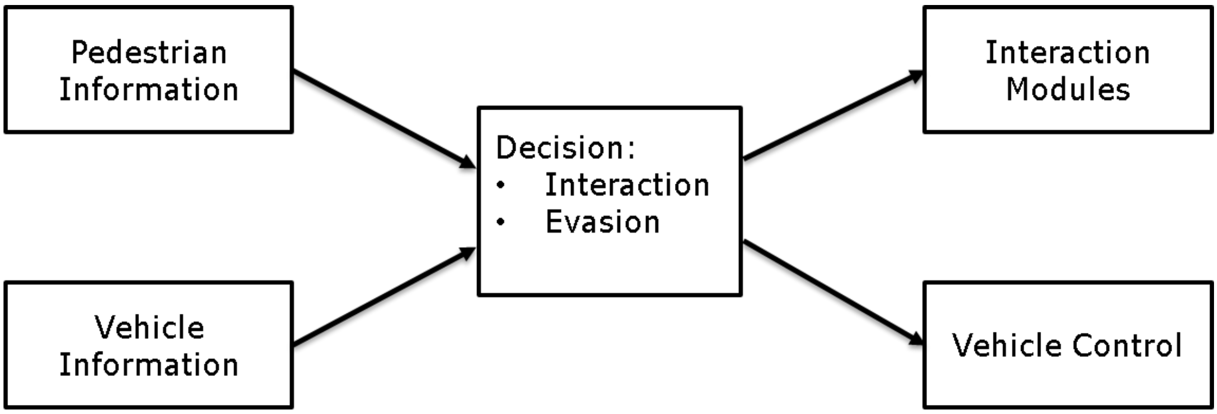


Figure 2.2: Block diagram of Pedestrian Interaction System (PIS) for Autobus in simulated environment. It uses interaction strategy along with normal vehicle control based on pedestrian information. Based on fixed rule, it decides for evasion and/or interaction [Jan 20].

2.3.1 Driving Principle

In contrast to typical driverless minibuses, the successful integration of interaction modules in this context demands consideration of additional factors, particularly due to the unique communication needed with pedestrians. These vehicles operate within a dual-state environment, catering to both pedestrians and passengers, each group characterized by a common goal: reaching their destination efficiently, barring those on leisurely or recreational journeys. These two groups often share a mutual desire for priority, highlighting the complexity of managing the interactions.

To address this, the system does not solely rely on pedestrian-focused attributes. Instead, it incorporates both external and internal factors – 'content' pertaining to the external environment and 'comfort' concerning the internal passenger experience. This comprehensive approach ensures a balanced interaction that caters to the needs and priorities of both pedestrians and passengers. The distinct states and their corresponding factors are delineated in Figure 2.3, illustrating the system harmonizing these dual aspects.

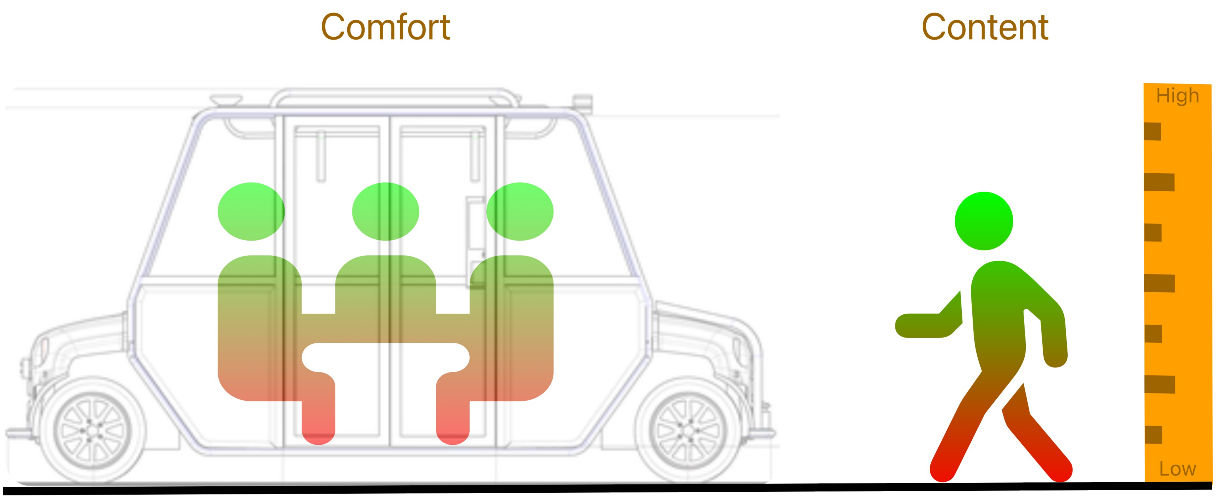


Figure 2.3: Comfort for passengers and contentment for pedestrians are defined as states that exist on the same scale regarding their intensity levels. Green represents higher levels of comfort and contentment, while red denotes lower levels. The color transition between these states indicates that both follow a similar pattern.

Comfort: The presence of passengers in the vehicle significantly influences the operational choices regarding its interaction module. Passengers typically choose driverless minibuses over walking for reasons of comfort, in addition to necessity. Comfort allows a widespread adoption, and one of the influential factor for acceptance [Paddeu 20]. To meet these expectations, the driverless minibus must consider certain factors in its decision-making process. Key factors include the time taken to reach the destination and the duration of stops made to yield the right of way. An increase in either factor can lead to reduced passenger comfort. Therefore, the interaction module must strive to optimize these factors, maintaining them at the highest feasible level, while not compromising pedestrian safety.

Content: In the research documented in [Chaloupka-Risser 07], an exploration was conducted into pedestrian contentment with the design of public spaces, the safety of traffic systems, and overall quality of life, with the aim of deriving principles for creating infrastructure that accommodates the needs of vulnerable road users. This dissertation broadens the concept of pedestrian contentment to include their reactions to the proximity of driverless minibuses. The presence of such vehicles in pedestrian-designated areas frequently incites dissatisfaction, as pedestrians generally do not anticipate encountering vehicles in these spaces. However, as delineated in the introductory chapter, these vehicles fulfill a specific function within such zones. Addressing pedestrian negative reaction necessitates the establishment of a user-centric interface between driverless minibuses and pedestrians. Enhanced pedestrian satisfaction is achievable when the vehicle is capable of precisely recognizing and reacting to pedestrian intentions. Conversely, pedestrian satisfaction diminishes when they feel compelled to yield to the driverless minibus.

The metrics for measuring comfort and contentment are distinct and vary over the duration of the vehicle's journey. Comfort is a continuous measure, extending from the start to

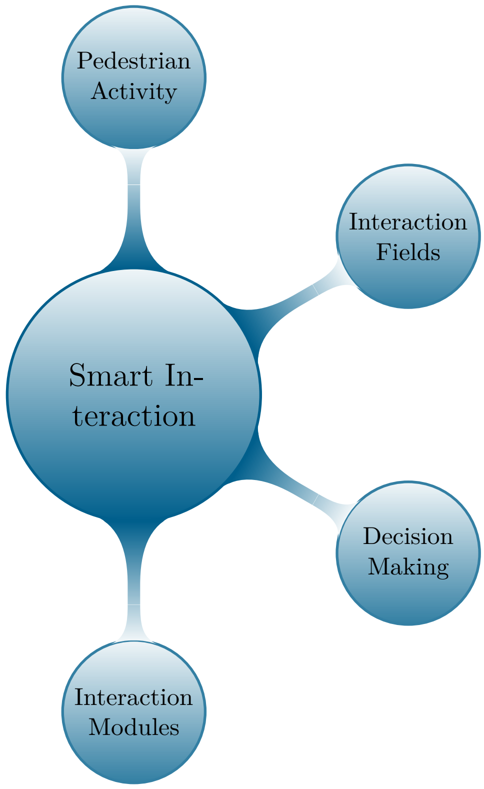


Figure 2.4: Schematic representation of the components involved in conceptual framework of smart interaction stretegy presented in this thesis. The diagram illustrates the central role of smart interaction, delineating its relationship with key components such as pedestrian activity, interaction fields, decision-making processes, and interaction modules.

the end of the trip. In contrast, contentment is more fleeting, resonating primarily during the brief interactions between the pedestrian and the Autobus. To achieve equilibrium among the specified parameters, it is prudent to establish a series of defined interaction fields. These fields will serve as a basis for smart interaction to determine both the nature and content of the interaction, thereby circumventing any superfluous engagement. This strategic approach ensures that interactions are not only contextually relevant but also efficiently executed, enhancing the overall effectiveness of the system.

2.3.2 Smart Interaction Components

Understanding smart interaction in Autobus means exploring its various essential components for smooth and effective communication. This system follows a complete cycle, beginning with recognizing pedestrians and ending with different signaling methods. The main parts of this smart interaction are shown in Figure 2.4 and described below. The highlighted section of Figure 2.4 specifically relates to the content of this chapter.

Pedestrian activity: The introduction of the interaction mechanism within the Autobus is based on a successful pedestrian detection. These systems are designed to detect and identify human activities autonomously in a specific environment, primarily through video and sensor data analysis. The aim is to equip the system with a capacity to understand human activities similarly to human perception, a feature of the interaction processes discussed in this thesis. The component of pedestrian activity comprises several subcomponents, starting with the data acquisition phase where data representing pedestrian activity are captured via a camera. Subsequent preprocessing steps are implemented, including application of filtering technology to refine and prepare data for analysis. Then the function extraction leads to the classification phase, which is performed by machine learning algorithms. In the scope of this study, the most commonly observed activities are sitting, walking, taking phone conversations, texting, and directions. A thorough understanding of pedestrian behaviour is essential to customize interactions in a way that is appropriate to each unique context.

Interaction Fields: The concept of interaction fields is pivotal for ascertaining both the timing and characteristics of interactions between Autobus and pedestrians. Interaction fields are defined for both, vehicle as well as pedestrians explained below :

- **Vehicle Interaction Field:** The term refers to a designated spatial area surrounding vehicles in which interaction with pedestrians is considered necessary. The purpose of this thesis is to determine three different zones of the vehicle interaction field, based on the proximity of the pedestrian to the vehicle: the risky zone indicating the immediate possibility of collision; the direct interaction area indicating the need for active participation or attention by the vehicle; and the ambient zone, which included a wide range of areas in which the vehicle must recognize pedestrian movements but where direct interaction was not immediately necessary.

- **Pedestrian Interaction Field:** The area is determined by the pedestrian's orientation, perceived risk levels and orientation intentions, and integrated into the vehicle interaction area to optimize the interaction strategy. The pedestrian interaction field helps to understand how pedestrians navigate around the vehicles, including how their movement patterns influence the decision-making processes of vehicles in real time.

These interaction fields are essential for the development of algorithms that allow Autobus to interpret and respond precisely to the dynamics in pedestrian behavior. Through this design, the perceived risk of pedestrian movement into a complex cycle of interactions. This cycle facilitates more awareness and decision-making in the Autobus, while ensuring a safer and more efficient positioning system that takes into account the complexity of pedestrian interactions.

Decision-Making: The process of deciding when to activate interaction modules or vehicle control systems is predicated on a holistic assessment of both the vehicle and pedestrian interaction fields. This entails a detailed analysis that integrates spatial zones around the vehicle with the trajectories and behaviors of pedestrians. By synthesizing information from these interaction fields, a comprehensive decision-making module is established. This module is crucial for determining the appropriate moments for engaging interaction modules or implementing vehicle control actions, thereby ensuring that responses are finely tuned to the dynamics of the surrounding environment and pedestrian activities.

Interaction Modules: Within the scope of this research, three principle methods are used to engage pedestrians: LED displays, voice communication systems, and auditory signals (beepers). The use of these interaction modules depends on their knowledge of the location of the pedestrian and the presence of the vehicle in relation to the interaction areas.

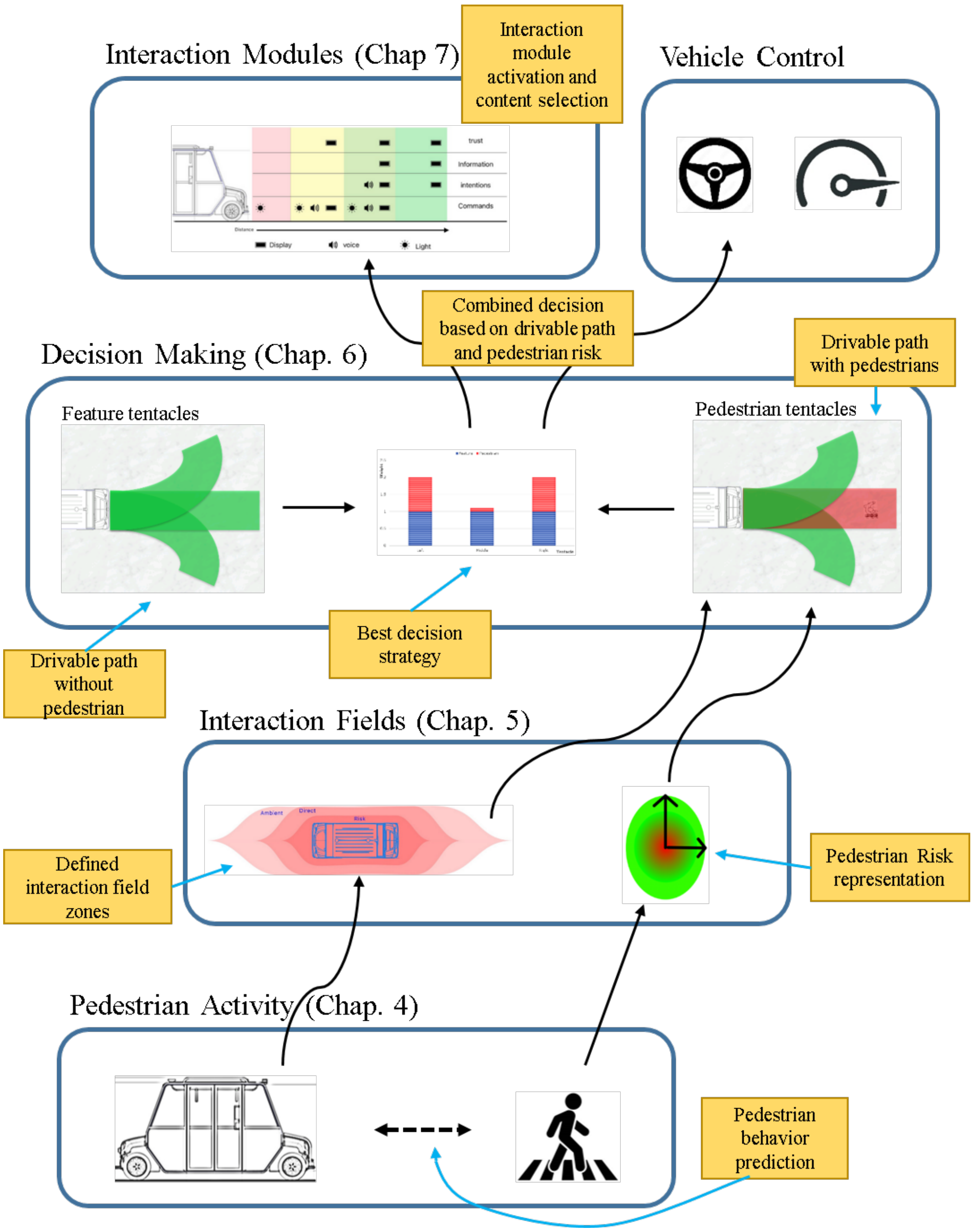


Figure 2.5: Hierarchical representation of smart interaction strategy in the context of this thesis concept along with chapter number for detailed description. This diagram delineates the logical flow of processes, starting from the bottom tier, where pedestrian behavior is predicted, to inform the Autobus's navigation system (details in Chapter 4). Progressing upwards, the next tier showcases the defined interaction field zones, delineating spatial areas for vehicle-pedestrian interaction with an emphasis on safety and communication (Chapter 5). The subsequent level illustrates the optimal decision-making strategy (Chapter 6), derived from the drivable paths analysis with and without pedestrian presence, to ensure secure navigation. The top tier captures the interaction module activation (Chapter 7), highlighting the selection of content for communication through various channels such as display, voice, and light, based on the combined analysis of trust, informational content, intentions, and commands. This tiered approach elicits a comprehensive decision-making framework, essential for the Autobus's operation within pedestrian zones, as discussed in the thesis.

Furthermore, the information presented by these modules has been carefully adapted to match the specific activities observed by pedestrians. This approach ensures that the communication strategy is appropriate and effective, increasing the safety and efficiency of pedestrian-vehicle interactions.

2.3.3 Smart Interaction Framework Overview

Overview of smart interaction framework is presented in Figure 2.5. Starting from the foundation, this diagram presents pedestrian behavior prediction, moving to the definition of interaction zones around the vehicle. It progresses to illustrate decision-making strategies based on pedestrian-inclusive and exclusive drivable paths, factoring in pedestrian risk assessments. The topmost part of the figure encapsulates the interaction module, which selects appropriate communication methods for the autonomous vehicle's operation in pedestrian environments, as outlined in the thesis.The diagram references specific chapters corresponding to each component for a more detailed exposition.

Overall, smart interaction is a pivotal development in the field of driverless minibuses, especially in addressing the challenges of the pedestrian-dominated environment. It gives priority to safety, context awareness and effective communication and facilitates the integration of these driverless minibuses into pedestrian areas, fostering a future in which humans and machines can coexist seamlessly in shared spaces.

In the forthcoming chapter, a detailed exposition of smart interaction is presented, building upon the foundational concepts and formulations introduced in this chapter. This subsequent section delves deeper into the intricacies of the system, elucidating the practical application and operational nuances that underpin effective vehicle-pedestrian interactions.

3. System Fundamentals

Before exploring the specific features of various modules related to the interaction concept in this research (as indicated in the Chapters 4, 5, 6, and 7), it is necessary to provide an overview of system architecture and test environments. This introduction is crucial to understanding the rationale and hypotheses behind the experiments and theories presented in these modules. Advanced systems show unique designs and functions, as discussed in Chapter 1. In the context of human interactions with vehicles, seemingly small characteristics such as the size and shape of the vehicle become important factors. The vehicles in question here are distinguished by the symmetrical front and rear design, lack of driver seats, lack of steering wheel. The absence of this driver leads to a decrease in pedestrian predictability [Guéguen 15] – due to the lack of regular driver gestures.

The Wizard of Oz technology [Detjen 20] is an attempt to simulate the functions of autonomous vehicles, focusing on simulation of driverless vehicles in experimental research. The observation indicated that pedestrians showed greater concern when they did not have a noticeable driver, which affected their confidence and decision-making processes when considering crossing the vehicle's path. This methodology allows a more authentic experimental setup. However, this method is harmful to the feasibility of driverless minibuses because there is no driver seat and thus enables the possibility of realistic experiments.

This section is aimed at investigating the design of the Autobus used in experimental validation of the concepts presented in this thesis. It includes a comprehensive description of the sensors installed on the vehicle and a detailed analysis of the safety systems integrated into the vehicle safety chain. As already stated in the introduction (Chapter 1), prioritization of safety is crucial, especially due to pedestrians' involvement and the need to meet the certification standards for vehicle operations in public spaces. To explain the complexity associated with pedestrian navigation, this chapter provides a detailed study of the environment in which the vehicle has been tested. In addition, it presents a simulated version of this environment that was used for preliminary tests and partially to verify the feasibility and effectiveness of the proposed concept.

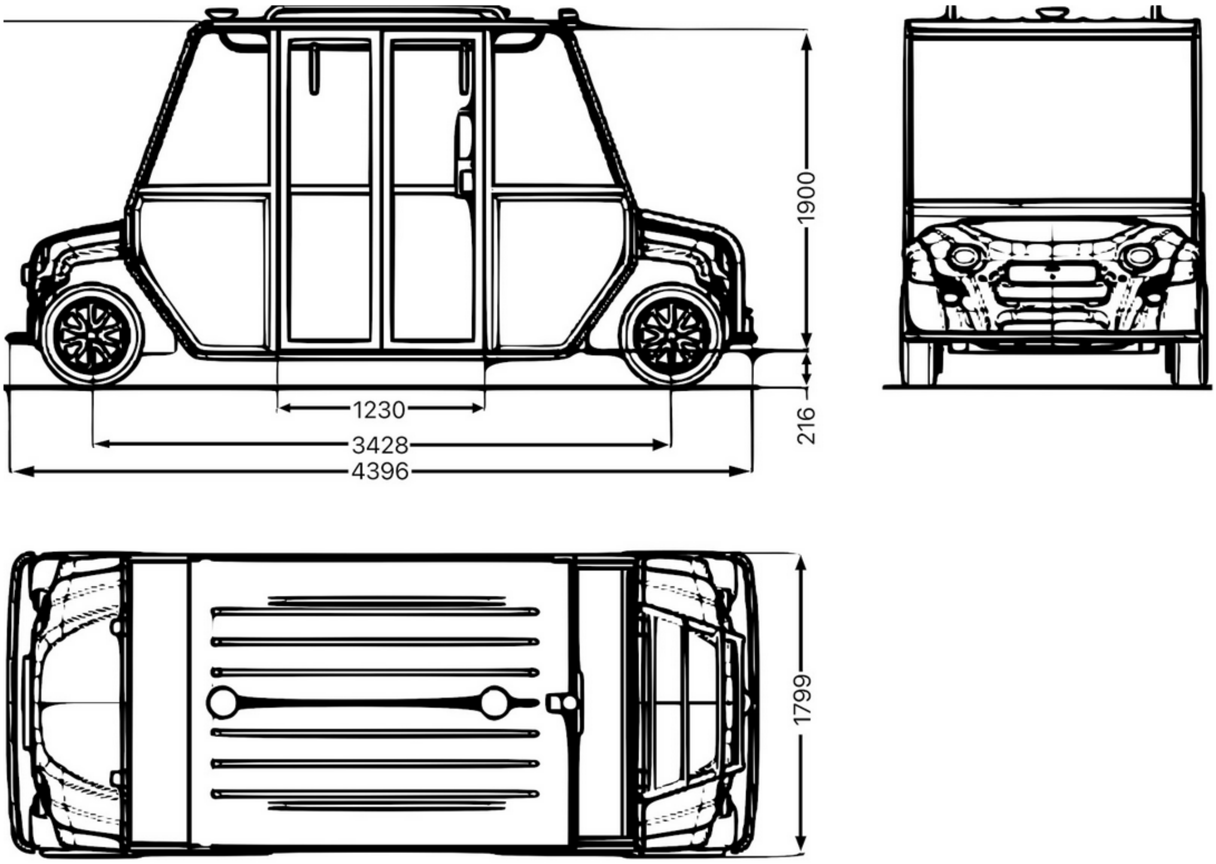


Figure 3.1: Technical schematic of the bidirectional Autobus used for testing the concept of this thesis, indicating key dimensions. The diagram provides a side view (top left), front view (top right), and top-down view (bottom center) with measurements in millimeters. The vehicle is designed to accommodate six seated and three standing passengers, as reflected by its dimensions: a length of 4396 mm, width of 1799 mm, and height of 1900 mm. The compact design is instrumental for maneuvering in pedestrian zones.

3.1 Autobus Design

Creating the illusion of genuine driverless minibuses by removing the driving seat and its associated components leads to confusion among pedestrians, who seek to understand the control mechanism and operator of the vehicle. In such contexts, interactive systems can be leveraged to their fullest potential to maintain authenticity in human responses to driverless minibuses. To embody this authenticity, the Robotic Research Lab at RPTU Kaiserslautern, previously mentioned in Chapter 1 and depicted in Figure 1.3, acquired a specific vehicle from Kompaii Robotics [1], France. This vehicle, designed to carry three additional standing passengers in addition to six seated ones, is bidirectional, enabling it to navigate equally effectively in both directions, as illustrated in Figure 3.1. Consequently, the deployment of the Autobus in a real-world setting serves as a practical validation of the thesis's underlying concept.

Figure 3.1 delineates the dimensions of the Autobus. The vehicle is specifically designed to navigate pedestrian zones; its size is optimized to ensure that it is neither too large to

[1] https://kompai-robotics.odoo.com/

Designation	**Features**
Dimensions Overall	L x W x H : 4.4 x 1.8 x 2.2 m
Estimated weight	1500 Kg
Capacity	10 passengers: 6 people seated + 4 people standing
Motorized	4 wheel drive: 2 x 15 KW driven by CURTIS drivers Mechanical differential front and rear
Steering	Front axle (car-like steering axle) Rear steering axle
Maximum speed	50 km/h on flat surface
Acceleration	nominal acceleration, set by software: 0.46 m/s^2 maximum deceleration, controlled by soft: 1 m/s^2 These values are configurable
Turning radius	7 m single steering axle 3.4 m double steering axle
Largest incline	10%
Suspensions	OLEOPNEUMATIC
Lead batteries	72 V, 200 Ah, 14.40 kW
Autonomy	several hours(depends on the usage scenario)
Charging time	8 hours
Doors	Motorized swing door
Wheels	16 inch
Control system	NVIDIA Jetson TX2 implementing basic low level control with UDP
Operating conditions	-10, +50°C, rain, fog, snow (possibility to put studded tyres)
Interior Equipment	Seat belts for sitting passengers Ventilation opening Handle for standing passengers

Table 3.1: Technical Specifications of the Autobus

Designation	Motor	Values
Maximal torque		75 Nm
Maximal speed motor		8000 tr/min
	Reductor	
Maximal torque		1400 Nm
Gear ratio		1/16
	Emergency Brake	
Maximal brake torque		110 Nm
Tightening force		1850 N
Diameter		200
Power		10 W
Voltage		AC 223V

Table 3.2: Motor, Reductor, and Emergency Brake Specifications

Designation	Values
Steering	
Maximal variation speed	450 °/s
Maximal torque	716 daN
Response time	30-40ms
Gear ratio	55.55 mm / revolution

Table 3.3: Steering Specifications

operate within these areas nor too small to accommodate more than two individuals. The width of the vehicle, at 1.8 meters, is of particular note. This dimension poses operational challenges in pedestrian zones, where space constraints may impede the free movement of pedestrians around the vehicle. During trials conducted on the RPTU campus, the Autobus navigated through the narrowest pathways, which measure at least 3.5 meters in width. To address the potential maneuverability issues arising from its width, the vehicle incorporates a double Ackermann steering system, enhancing its ability to negotiate turns with a reduced turning radius of 3.4 meters, despite its 4.4-meter length. Table 3.1 show the technical specification of the Autobus.

Further technical details of the traction axle, steering axle, and emergency brakes are provided in Tables 3.2, 3.3, and 3.4, respectively. The Autobus features two motors and emergency braking systems located at both the front and rear. Each motor delivers a maximum torque of 75 Nm, which ensures adequate acceleration when the Autobus is fully loaded with passengers. Additionally, a reduction gear with a torque capacity of 1400 Nm effectively brings the vehicle to a stop when it reaches zero velocity, based on deceleration values. However, this system is insufficient on slopes, necessitating the use of hydraulic brakes on each tire. The emergency braking system, with a clamping force of 1850 N, is capable of locking the wheels at varying speeds.

3.2 Sensor System

Sensor systems are a critical component in the architecture of driverless minibuses, providing the necessary data for safe and efficient operation. These systems, comprising a diverse

Designation	Values
Disc diameter	250 mm
Braking torque	140 Nm
Holding torque	1850 Nm

Table 3.4: Brake System Specifications

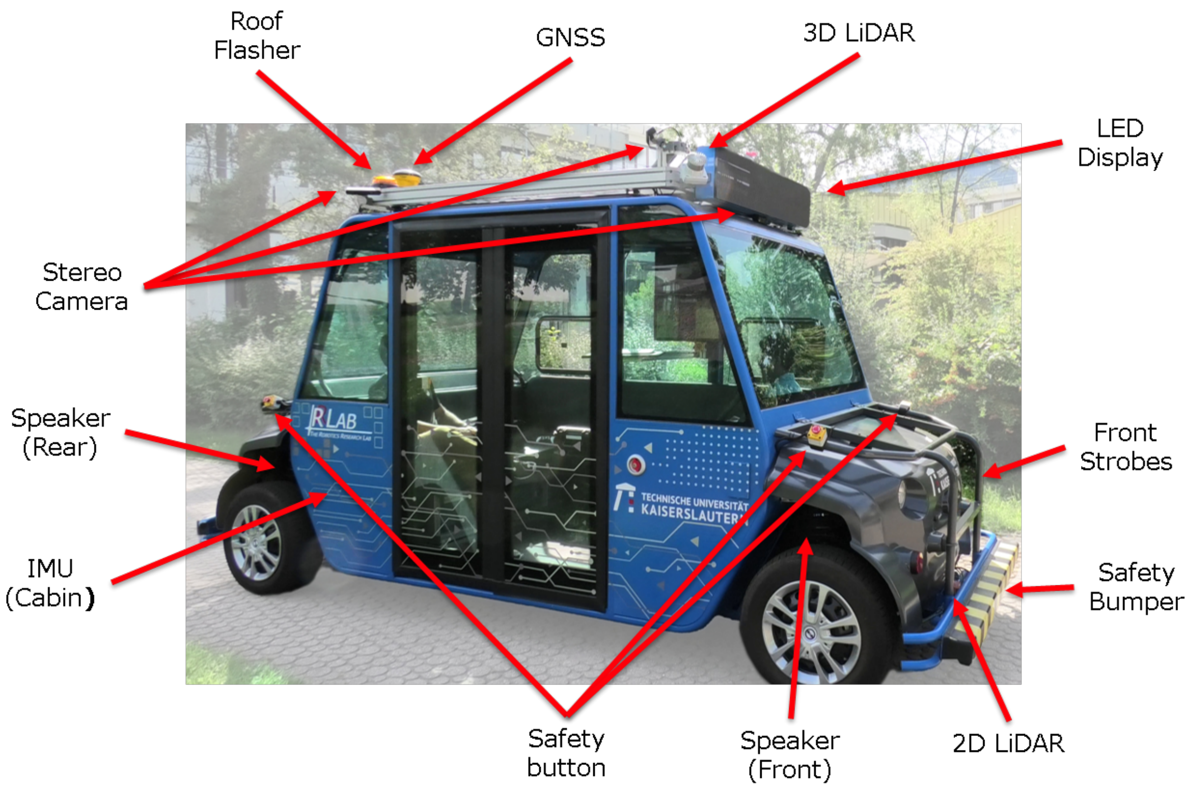


Figure 3.2: This image illustrates the sophisticated array of sensors equipped on the Autobus. Red arrows point to various sensor locations, highlighting the vehicle's comprehensive sensory configuration essential for its navigation and safety systems. The sensors, distributed strategically across the vehicle, include LiDARs and cameras, which together facilitate real-time environmental monitoring, pedestrian detection, and collision avoidance. This configuration enables the Autobus to perceive its surroundings, make informed decisions, and maneuver safely within the complex ecosystem of urban pedestrian zones, thereby exemplifying the integration of advanced autonomous technologies in public transportation.

array of sensors, form the cornerstone of a driverless minibus's perception capabilities, enabling it to interpret and interact with its environment effectively. Among the most critical sensors are 2D and 3D Light Detection and Ranging (LiDAR) systems. These LiDAR sensors provide high-resolution, three-dimensional information about the vehicle's surroundings, essential for obstacle detection and terrain mapping. Complementing LiDAR, stereo cameras offer valuable visual data, capturing detailed images that aid in object recognition and scene understanding. Furthermore, Global Navigation Satellite System (GNSS) units are integral to these sensor arrays, offering precise geolocation data that assists in route planning and navigation. Together, these sensor systems form a comprehensive suite that empowers Autobus to operate reliably in a various environmental conditions, from crowded urban centers to remote rural areas. The following subsections talks about the particular sensors mounted on Autobus for navigation use.

3.2.1 Types and Applications of Sensors in Autobus

Autobus integrates an advanced sensor, including bidimensional and tridimensional laser scanning systems (2D and 3D LIDAR), stereo cameras and global navigation satellite systems (GNSS). Each of these sensory components plays an important role in the Autobus navigational scheme and provides essential data for real-time environmental modeling, obstacle detection, and traffic planning. The application of these sensors, as explained in the theoretical framework presented in Chapter 2, goes beyond basic location and navigation tasks. They are essential to facilitate intelligent interaction. By high-resolution spatial mapping and dynamic object tracking, the Autobus displays accurate detection and responsive behaviour towards pedestrians, thus allowing a symbiotic coexistence within shared spaces. Furthermore, this sensory fusion increases the autonomy of the Autobus and improves decision-making processes under complex urban scenarios. At the same time, it guarantees the interactional harmony between the automated system and human users - the prerequisite for the effective integration of Autobus into the pedestrian zone. Figure 3.2 presents a comprehensive visual overview of the Autobus platform, indicating the strategic positioning of various sensory devices. This figure emphasizes the physical integration of sensors into the vehicle chassis, giving insight into the spatial distribution and field of view of each sensor command. The annotations within the figure specify the locations of the 2D and 3D laser scanners, stereo cameras and GNSS modules, providing a clear reference to understand how these components together contribute to the vehicle's perception system. This visualization is instrumental in understanding the synergistic operation of the sensor array, which supports the autonomy of the navigation and pedestrian interaction of the Autobus. The following subsections are detailed for each sensor implemented within the Autobus system. These descriptions will clarify the operational principles, technical specifications and integration methods for the following sensory devices.

Description 3.1: OutdoorScan3

The outdoorScan3 safety laser scanner ensures the protection of individuals in a variety of outdoor environments, whether for mobile or stationary applications. Leveraging sophisticated algorithms and its exclusive outdoor safeHDDM® scanning technology, it delivers consistent productivity under adverse weather conditions built by SICK [a].

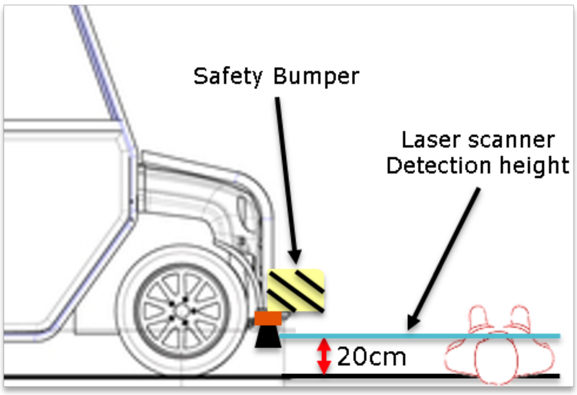
[a] https://www.sick.com/cz/en/

3.2.1.1 2D-LiDAR

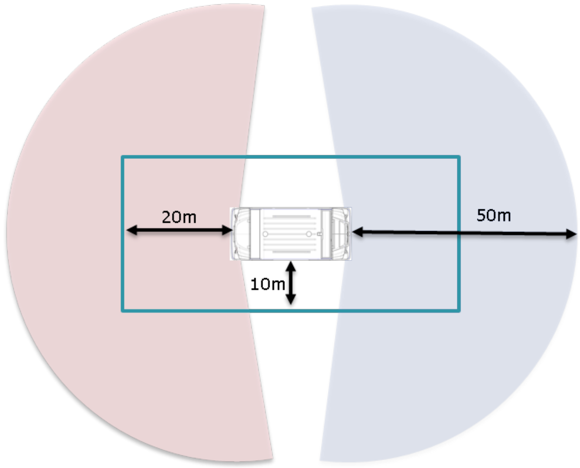
The two-dimensional laser scanner, an essential element in the sensor suite of Autobus, deserves careful attention for its vital function in their system. This unit relies on the LiDAR technology, which involves the emission of laser pulses to gauge distances by timing the reflection of the light from surfaces. This scanner differs from its three-dimensional variant by mapping the environment on a flat plane. Within the Autobus, outdoorScan3 (Description 3.1) is used, which is crucial for recognizing and circumventing obstacles close to the ground, playing a key role in preventing collisions. Its capability to furnish immediate situational data is invaluable for navigating dynamic environments, particularly in densely populated or urban areas where sudden obstructions might occur.

Strategically positioned at the front and rear within the protective bumper and inverted, its scanning beam sits approximately 20cm above the ground, aligning with the lower part of the vehicle's chassis. This placement is significant for prompt detection of obstacles and for basic landscape mapping, both necessary for secure vehicular movement. The scanner's height also allows for the potential identification of a person lying prone on the road surface. In the provided illustration in Figure 3.3a of the Autobus, the side view reveals the 20cm elevation of the scanner from the ground, a height optimized for detecting the width of a human torso in a horizontal position. Average human torso width is estimated to be more than 20cm. Figure 3.3b presents a top view of the scanning field in its original scale. The scanner possesses a maximum operational range of 50 meters; however, such an extensive range is superfluous for the Autobus operating within confined pedestrian areas. Consequently, the depicted field is truncated to the cyan rectangle, which represents a more practical detection zone extending 20 meters horizontally from the vehicle's front and 10 meters laterally from its sides. This positioning also confers the advantage of facilitating detection across an arc exceeding 180°from the aggregate angle of 175°. Extra 10°are added on each side as can be seen in the shaded region of Figure 3.3b. Detection of the vehicle's sides using analogous methodology is impracticable, as any ground elevations, such as humps, would result in constant detection by the scanner, notwithstanding the vehicle's capacity to navigate over them.

Table 3.5 outlines the characteristics of single-line scanners producing 2D polar scan points, also employed in safety systems detailed in Section 3.3.



(a) Side view illustration of the Autobus highlighting the laser scanner's detection height set at 20cm above the ground surface, designed to identify low-lying obstacles, specifically humans, and ensure ground clearance.



(b) Top-down schematic of the Autobus displaying the 2D laser scanner's fields of view. The outer blue sector illustrates the maximum range of 50m for front, while the inner pink sector denotes the operational field for rear. The cyan rectangle demarcates the optimized 20m forward and 10m lateral detection area utilized for navigation in confined environments.

Figure 3.3: 2D-LiDAR configuration for Autobus with (a) side view and (b) top view.

Parameter	**Value**
Protective field range	4 m
Warning field range	40 m
Number of simultaneously monitored fields	$\leq 8^1$
Number of fields	128
Number of monitoring cases	128
Scanning angle	275°
Resolution (can be configured)	50 mm 70 mm
Angular resolution	0.39°
Response time	$\geq$ 115 ms
Protective field supplement	65 mm

Table 3.5: Technical Specifications of OutdoorScan3.

3.2.1.2 3D-LiDAR

In contrast to 2D laser-scanner, 3D LiDAR represents a quantum leap in spatial data acquisition. It extends beyond the limitations of a single plane, capturing comprehensive three-dimensional data. This is achieved through an intricate array of sensors and laser beams, which collectively map environments in three dimensions with a high degree of precision and detail. The key advantage of 3D LiDAR lies in its ability to generate detailed point clouds that represent the environment in volumetric form. This capability is crucial for numerous applications, including autonomous navigation, topographical mapping, and complex environmental modeling, where understanding the vertical structure and complexity of the scene is paramount.

Description 3.2: Ouster-OS0

The OS0 short-range sensor achieves a detection range of up to 35 meters on a 10% reflectivity target, complemented by an expansive 90-degree vertical field of view. This sensor offers a superior balance of cost, efficiency, dependability, dimensions, mass, and energy consumption within the industry. Engineered by Ouster [a] for operation across diverse meteorological conditions, its compact form factor facilitates seamless integration with autonomous vehicles, industrial machinery, robotic systems, unmanned aerial vehicles, and geospatial mapping technologies.

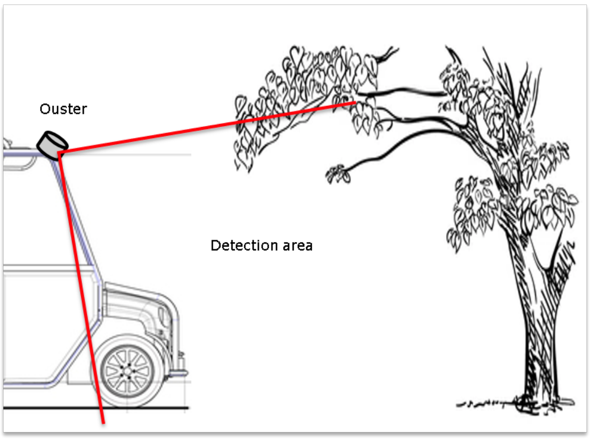
[a] https://ouster.com/

Within the Autobus platform, the employment of the Ouster-OS0 sensor (as described in Description 3.2) is critical for complex navigational functions. This sensor enhances the vehicle's perception capabilities, offering an extensive and detailed awareness of the immediate terrain and built environment. This allows for the real-time acquisition of intricate details within pedestrian areas. The decision to utilize this particular sensor was informed by its high-resolution 128-channel output and its broad 90-degree vertical field of view. The 128 channels provide granular measurements at greater distances for effective object detection. This granularity ensures that even minor miscellaneous items within pedestrian zones are promptly recognized as potential obstacles. Furthermore, the sensor's 90-degree opening angle affords a comprehensive sweep of the area, ranging from the ground level proximate to the vehicle up to higher elevations, thereby maximizing coverage. Two ousters are mounted in diagonal configuration as can be seen in Figure 3.4b. Such configuration allows to cover the blind spots around the vehicle. The total range of the scanner is 50m, but as discussed for 2D laser scanners, the range is minimized to rectangle in cyan shown in the figure. The mounting angles of the scanner in rectangular cordinate system are *x=0°, y=35°, and z=-45°for front and z=135°for rear.* At this position, it is possible to cover range above the ceiling for over hanging object as demonstrated in Figure 3.4a.

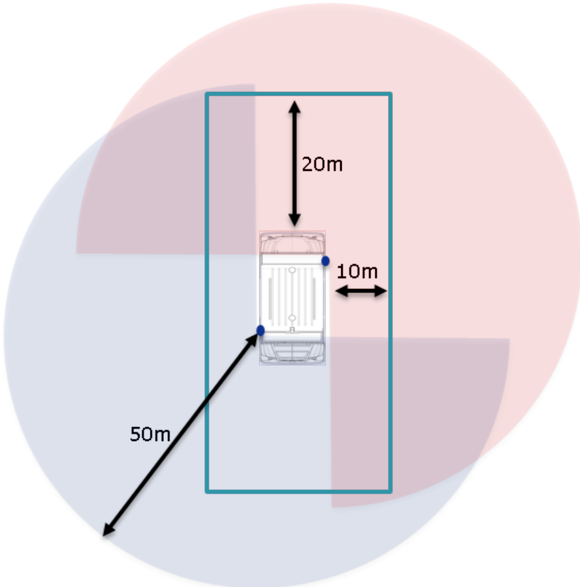
Table 3.6 details the specifications of the 3D laser scanner, featuring a vertical resolution of 128 channels and a horizontal resolution of 2048, enabling it to create a detailed environmental map. Consequently, the Ouster-OS0 is utilized for environmental mapping, enhancing obstacle detection and path segmentation.

3.2.1.3 Stereo Camera

The comparative analysis of stereo camera systems and LiDAR sensors is crucial for elucidating their distinct capabilities and potential applications. LiDAR, celebrated for its precision in measuring distances through the reflection of laser light, is adept at producing



(a) Side view illustration of the Autobus highlighting the laser scanner's detection angle, designed to identify overhanging obstacles, specifically trees and buildings, and ensure drive clearance.



(b) Top-down schematic of the Autobus displaying the 3D laser scanner's fields of view. The outer blue sector illustrates the maximum range of 50m for front, while the inner pink sector denotes the operational field for rear. The cyan rectangle demarcates the optimized 20m forward and 10m lateral detection area utilized for navigation in confined environments.

Figure 3.4: 3D-LiDAR configuration for Autobus with (a) side view and (b) top view.

Specification	**Value**
Range (80% Lambertian reflectivity, 1024 @ 10 Hz mode)	75 m @ 100 klx sunlight, >90% detection probability
Range (10% Lambertian reflectivity, 1024 @ 10 Hz mode)	35 m @ 100 klx sunlight, >90% detection probability
Minimum Range	0.5 m (to be reduced in FW 3.1)
Vertical Resolution	32, 64, or 128 channels
Horizontal Resolution	512, 1024, or 2048 (configurable)
Rotation Rate	10 or 20 Hz (configurable)
Field of View	Vertical: 90° (+45° to -45°) Horizontal: 360°
Angular Sampling Accuracy	Vertical: ±0.01° / Horizontal: ±0.01°
False Positive Rate	1/10,000
Range Resolution	0.1 cm
# of Returns	2 (strongest, second strongest)

Table 3.6: Technical Specifications of Ouster-OS0.

high-resolution three-dimensional point clouds. This attribute renders it highly effective under various lighting conditions, offering robust data essential for autonomous driving applications. In contrast, stereo camera systems adopt an alternative methodology. By employing a duo of cameras to mimic human binocular vision, they determine depth through analyzing the disparity between two concurrently captured images. Although typically more cost-effective and less computationally intensive than LiDAR, the efficacy of stereo camera systems is significantly contingent upon favorable lighting conditions for optimal performance. Additionally, these systems excel in providing detailed texture and color information, rendering them especially advantageous for tasks necessitating visual detail, such as object recognition and scene segmentation.

Description 3.3: ZED-2i

The ZED 2i, distinguished camera for depth perception, motion detection, and artificial intelligence applications, is a robust and adaptable stereo camera suitable for deployment across a broad spectrum of environments. Built by StereoLabs [a], it Features a 120mm stereo baseline, USB 3.1 connectivity, an integrated Inertial Measurement Unit (IMU), barometer, and magnetometer, and the capability to deliver 1080p resolution video at 30 frames per second, this device is also certified with an IP66 rating. It uniquely combines advanced long-range depth perception with artificial intelligence to facilitate three-dimensional environmental recognition within a 120° (Diagonal) wide-angle field of view. Engineered to be resistant against dust, water, and humidity, the ZED 2i caters to a variety of demanding applications, including outdoor deployment and use in challenging medical, industrial, and agricultural settings.

[a] https://www.stereolabs.com/

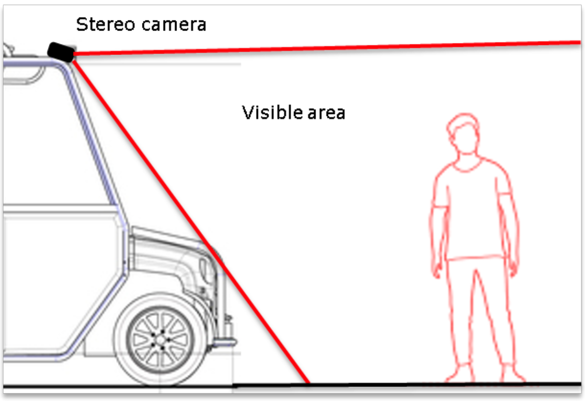
In the context of the Autobus, the ZED-2i cameras (as delineated in Description 3.3) are deployed primarily for the detection of human skeletal structure and the recognition of activities, selected for their proficiency in capturing and interpreting visual signals, hues, and textures. Additionally, the Autobus employs the color imaging capabilities of the ZED cameras for the segmentation of travel paths. The strategic configuration of these cameras is critical, particularly in light of the operational milieu of the Autobus. Within pedestrian domains, the erratic movements of individuals in proximity to the vehicle necessitate a panoramic view to ensure secure maneuvering. This requirement is accentuated at bus stations, where the monitoring of passenger dynamics during boarding and disembarking processes is imperative for safety and service efficacy. To accomplish comprehensive surveillance, six ZED cameras is methodically affixed to the Autobus: trio facing anteriorly and trio posteriorly. This distribution is designed to ensure an expansive visual field, mitigate potential occlusions, and augment situational perception. Figure 3.5a

Feature	**Specification**
Output Resolution	Side by Side 2x (2208x1242) @15fps 2x (1920x1080) @30fps 2x (1280x720) @60fps 2x (662x376) @100fps
Interface	USB Type C - External cable (up to 10m - 32.80ft)
Baseline	12cm (4.73 in)
RGB Sensors	Dual 1/3" 4MP CMOS 2688 x 1520 pixels $2\mu m$ x $2\mu m$ Rolling shutter YUV 4:2:2 - UYV (8bits)
Motion Sensors	Gyroscope, Accelerometer, Magnetometer
Environmental Sensors	Barometer, Temperature

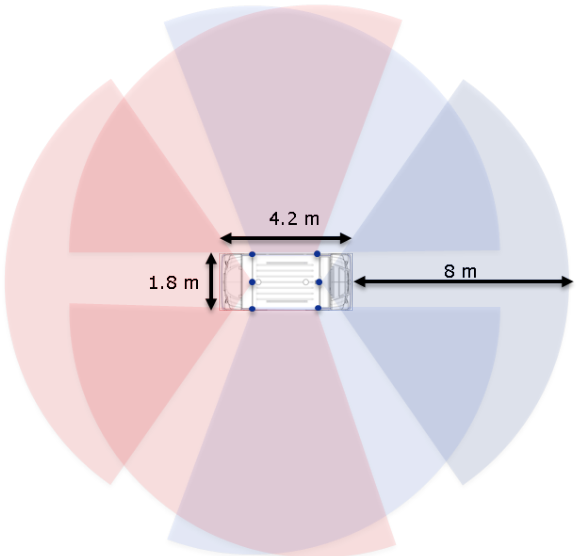
Table 3.7: Technical Specifications of ZED-2i.

shows front stereo camera mounting from the side view of the Autobus. It gives the visible range of the camera where a standing pedestrian of average height is easily visible. -Figure 3.5b offers an explicit representation of the six camera configuration from top, elucidating their locational bearings and observational extent. The shade in blue shows the front view cameras, whereas the pink shade shows the rear viewing cameras. An exhaustive delineation of the ZED camera system's specifications and technical attributes is provided in Table 3.7.

Table 3.7 presents the technical details of ZED stereo camera, which has a 12cm baseline and operates at various resolutions. Given this baseline and the Autobus's slow speed in pedestrian areas, it effectively provides depth estimation up to 8m. Different output resolutions can be adjusted based on the application and processing power.



(a) Schematic representation of a stereo camera's field of view mounted on a vehicle, delineating the visible area with respect to a standing human figure for perspective.



(b) Configuration of only front viewing stereo cameras on Autobus.

Figure 3.5: Stereo camera configuration for Autobus with (a) side view for front camera and (b) all camera top view.

Description 3.4: Trimble-BX992

The BX992, featuring a dual-antenna receiver enclosure, integrates an inertial navigation system powered by the BD992-INS, facilitating the generation of robust, high-accuracy positions and orientations across various environments through the amalgamation of inertial sensors within a singular module. This system provides 336 tracking channels, employs Trimble Everest Plus™ for multipath mitigation, incorporates advanced RF spectrum monitoring/analysis, and utilizes proven technology for low-elevation tracking. It ensures continuous positioning capabilities in areas where GNSS signals are obstructed or unavailable, simultaneously offering solutions with a high-update rate for position and orientation. Moreover, the BX992 supports flexible interfacing options, including RS232, USB, and Ethernet connections, enhancing its adaptability for diverse operational requirements.

Specification Category	**Details**
Product Type	Dual Antenna GNSS Receiver Module
GNSS Signals	GPS: L1 C/A, L2C, L2E, L5 GLONASS: L1 C/A, L1P, L2 C/A, L2P, L3 Galileo: E1, E5 AltBOC, E5a, E5b, E6 BeiDou: B1, B2, B3 QZSS: L1 C/A, L1 SAIF, L2C, L5, LEX SBAS: L1 C/A, L5
Channels	336 Channels
Heading Accuracy	0.1 degrees with 2 m baseline
Position Accuracy	Horizontal: 1 cm + 1 ppm RMS Vertical: 2 cm + 1 ppm RMS
Update Rate	Up to 20 Hz
Communication Ports	Ethernet, Serial, USB, CAN
Operating Temperature	-40°C to +75°C
Power Requirements	3.3V DC to 5.5V DC
Dimensions	60 mm x 67 mm x 15 mm
Weight	Approx. 60 g
Certifications	FCC, CE, RoHS

Table 3.8: Specifications of Trimble BX992

3.2.1.4 GNSS

Global Navigation Satellite System (GNSS), an indispensable component in the technological ecosystem of autonomous driving, plays a fundamental role in their navigational capabilities. Anywhere on or near the Earth, the satellite-based navigation system, or GNSS, gives precise location and timing information in all weather situations. In the domain of autonomous driving, this system is crucial for providing precise geospatial positioning, which is a cornerstone for route planning, navigation, and overall orientation of the vehicle in its environment.

Navigating within pedestrian areas can be difficult due to the changing nature of pedestrian traffic, obstacles and environmental conditions. To rely solely on local navigational functions is often not enough because sensors have limited perception ranges. In crowded places, people, vehicles and urban infrastructure often block these sensors from tracking important indicators. Systems that use tools such as Inert Measurement Units (IMUs) or visual odometry may suffer from cumulative errors, which cause them to drift over time. Without a way to correct these drifts, navigation accuracy is severely affected. Here the Global Satellite Navigation System (GNSS) plays its role. The Trimble-BX992 unit, for example, is a high-precision GNSS that offers up to 10 cm of accuracy, making it vital for the Autobus. It is particularly useful in urban areas with complex environments and tall buildings. Although traditional GNSS can face signal variations and obstructions in such environments – known as urban canyon effect – high-precision GNSS can overcome these problems. The greater accuracy of high-precision GNSS is crucial for various applications. It improves pedestrian tracking, supports autonomous vehicle navigation and improves location-based services. In these scenarios, even small position differences can lead to significant navigation errors. High-precision GNSS ensures accurate location tracking and

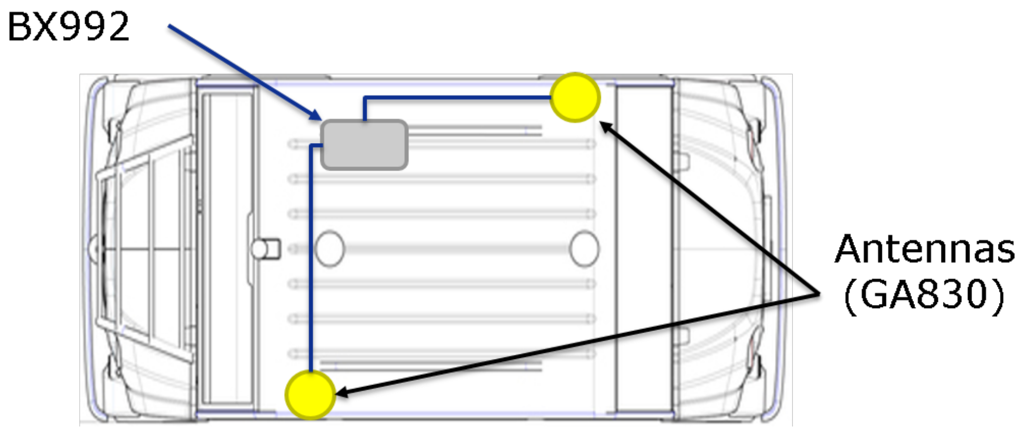


Figure 3.6: The diagram illustrates the top-down view of the Autobus, highlighting the placement of the Trimble-BX992 unit (indicated by the blue arrow) which is integral to its high-precision GNSS. The two yellow circles represent the primary GNSS antennas, crucial for providing locational accuracy up to 10 cm. This system is essential for reliable navigation in complex urban environments, counteracting the limitations posed by local navigation systems and the 'urban canyon' effect caused by tall buildings and dense infrastructure.

enhances navigation capabilities by improving the accuracy of bus stop locations based on pedestrian demands.

Moreover, the two antennas provide two distinct points from which satellite signals can be received. By analyzing the phase difference between the signals received at each antenna, the system can determine the vehicle's heading, which is its directional orientation in relation to the Earth's surface. Figure 3.6 delineates the configuration of the Trimble-BX992 integrated with dual antennas (GA830) affixed to the vehicle's roof, providing better vehicular orientation capabilities for the Autobus. For more technical specifications, refer to Table 3.8.

3.2.2 Interaction Modules

In addition to the standard navigational sensors, the Autobus incorporates specific interaction modules, notably LED displays, speakers, beepers, and flashers, which are fundamental to the overarching concept of this work. These modules play a crucial role in facilitating effective communication between the vehicle and its external environment, particularly with pedestrians and other road users. The integration of these interactive elements is central to the design philosophy of the Autobus, ensuring not just autonomous navigation but also interactive safety and information dissemination. Figure 3.7 shows the different interaction modules mounted on the Autobus. The subsequent subsections provide a detailed exposition of the configuration and operational principles of these interacting modules. This includes an in-depth analysis of how the LED displays, speakers, beepers, and flashers are configured on the Autobus.

3.2.2.1 LED Display

The LED display is a key part of the Autobus user interface and helps to communicate important information to nearby pedestrians. To maximize its efficiency, the display is

Figure 3.7: The image shows various components of the external communication system of the Autobus, which are designed to interact with its environment. The LED display, strategically placed to disseminate essential information to pedestrians shown within the red dotted rectangle, is highlighted for increasing safety and awareness on the road. The blue square surrounds the omnidirectional beacon lights mounted on the vehicle, which provide maximum visibility and are used as indicators of the autonomous operational status of the car. Turn signals are defined by yellow dots to indicate the intent of vehicles to nearby traffic participants and reduce the likelihood of collision. Finally, flashers in front of the pedestrians are activated in front of them shown by yellow dotted line rectangle to warn them of the immediate presence of the Autobus, particularly important cautionary measure in densely populated areas where buses are operating in close connection with people walking.

Feature	**Details**
Display Type	LED
Customization	Custom-made for Autobus
Dimensions	Approx. 1250 x 250 x 100 mm
Brightness	6,000 cd/m^2
Pixel Pitch	2.97 mm (P. 2.97)
Resolution	420 x 84 pixels
Housing Material	Aluminum
Protection Class	IP65, all-around
Operating Voltage	24VDC
Brightness Adjustment	Via software
Control Interfaces	LAN, RS232
Sensors	Ambient light sensor, Temperature sensor
Connectivity	WLAN module

Table 3.9: Specifications of LED Display

carefully mounted on both front and rear sections of the bus roof. The red point limit of Figure 3.7 ensures that pedestrians and other vehicles can easily view the display. By placing the display on this elevated position (especially 2.5 meters) it helps to avoid visual obstruction and make information clear and accessible. This height is perfect because it aligns with the typical sight line of a pedestrian who walks upright. As a result, people can easily see and understand information without having to change their eyes or postures significantly. This thoughtful position not only improves visibility, but also improves the overall communication effectiveness of the screen.

Featuring a resolution of 420 x 84 pixels and a luminance of 6000 cd/m^2, the display ensures legibility from extended distances. Complementary to this, the appendix summarizes the display's specifications in Table 3.9, which encompasses the particulars of the display's interface for communication. This includes an overview of its integration with the central processing unit of the bus and the protocols employed for the transmission of the data to be exhibited.

3.2.2.2 Speaker

Apart from visual cues, audio also plays a huge role in getting the attention of pedestrians who aren't looking at the screen. To make sure these people notice important messages, the bus has a speaker system placed next to both the entrance and exit points, as shown in Figure 3.8. This setup was chosen carefully to ensure messages are heard clearly, especially when passengers are getting on or off the bus. The speakers near the doors mean that the audio is right where passengers need to hear it the most.

Moreover, the speaker system is meticulously designed to send sound to both the front and back of the bus. This way, people nearby—whether they're walking in front of or behind the bus—can hear notifications and warnings. Having a wide-reaching sound system is crucial for keeping everyone safe and making sure messages are heard throughout the bus's surroundings.

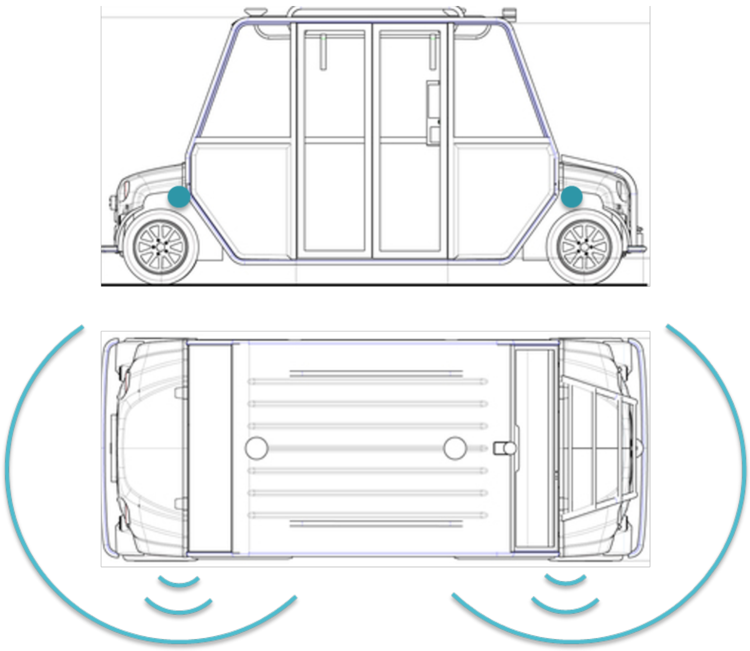
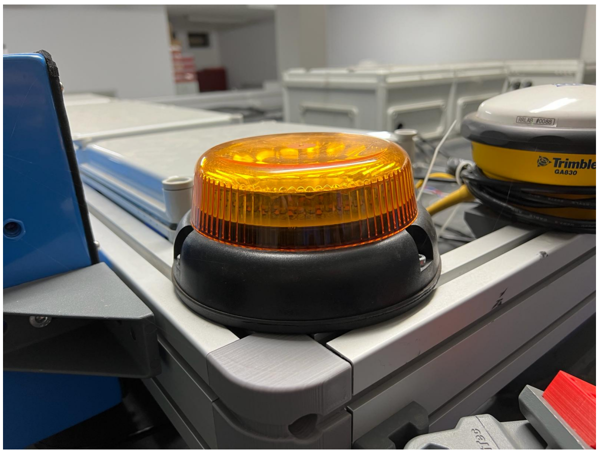


Figure 3.8: The schematic diagram shows the strategic placement of the speaker system in the Autobus. The speakers are located on both sides of the vehicle next to the entrance and exit points to maximize the effectiveness of audio communication, especially during important interactions such as boarding and alighting. The placement of the speakers near the doors is designed to ensure clear audio messages to passengers in these important areas. In addition, the system was designed to extend the acoustic field, transmit sound to the front and back of the Autobus and to make announcements to pedestrians and others nearby. This wide sound projection is essential to maintain safety and enable continuous communication near the Autobus.

(a) The figure illustrates the utilization of front-mounted flashers on the Autobus, which are designed to alert pedestrians to its presence.



(b) The roof-mounted flashing lights serve as indicators of the Autobus's autonomous operational status. Activation of these lights, via flashing, denotes that the vehicle is navigating autonomously.

Figure 3.9: Front and roof-mounted flashers as pedestrian alerts and autonomous operation indicators.

3.2.2.3 Flashers

In vehicle communication, lights and flashes signal different operational states. The autobus uses two types of flashers, each serving a unique purpose. Figure 3.7 shows an overview of these flashers, highlighted with blue/cyan and yellow borders. As depicted in Figure 3.9a, the flashers at the front and back of the vehicle activate only when pedestrians are in a critical position near the vehicle, serving as warning signals. Conversely, the flashers in Figure 3.9b indicate the Autobus is in autonomous navigation mode. These flashers turn on to inform pedestrians that the bus is driving itself. They are placed with GNSS antennae in a diagonal pattern to ensure they are visible from all directions.

Feature	**Detail**
Routing/Switching Capacity	48 Gbit/s
10/100/1000 MBit/s Ports	24 x
PoE-at Ports	12
Features	PoE Function
Weight	3.3 kg
Width	440 mm
Height	44 mm
Depth	238 mm
Type	HPE 1420-24G-PoE+ (124W) Switch - Switch
Interfaces	RJ45
LAN Transmission Rate	1000 MBit/s

Table 3.10: Specifications of the switch

3.2.3 Hardware Configuration

The deployment of a diverse array of sensors in the Autobus facilitates various hardware configurations, addressing the vehicle's substantial dimensions which prevents direct connections of certain rooftop-mounted sensors to the central processing unit. Hence, the communication for some sensors becomes impossible. It is aimed to ensure that the sensor data is delivered to the main CPU without delay. Consequently, the adoption of switch-on mechanisms alongside fiber optic technology is critical for the expedited and efficient handling of data across an extensive network of onboard sensors. The switch has a switching capacity of 48 *Gbit/s* with LAN transmission rate of 1000 *MBit/s*. The technical details are given in Table 3.10. The fiber optic framework from the switch to main PCs provides high-bandwidth and low-latency communication channels, essential for processing the voluminous data produced by the vehicle's assortment of sensors, such as LiDAR, cameras, and GPS. This infrastructure enables the Autobus to swiftly assimilate and analyze sensor data, underpinning real-time decision-making processes vital for autonomous navigation and ensuring safety.

Additionally, the system's architecture incorporates a distributed computing model, with two primary embedded PCs allocated to specific functions and data streams from distinct sensors. This modular data processing strategy permits simultaneous computations, markedly improving the system's capacity for managing and interpreting the copious data emanating from the sensor array. Distributing the computational tasks among multiple PCs allows for the efficient handling of intricate datasets, including high-definition imagery and comprehensive 3D mappings, thereby enhancing processing speed and precision. Such an architecture not only elevates the system's responsiveness but also supports its ability to execute prompt and accurate navigational decisions in the fluid and unpredictable context of urban settings. To accommodate the demands of deep neural network applications—vital for this thesis's concept—without overburdening the primary PCs, three Jetson AGX boards are also installed on the roof. This unit is directly connected to cameras for the preliminary data processing executing neural networks, thereby alleviating computational loads on the main processors. The synergistic application of advanced fiber optic communication and distributed computing frameworks is thus imperative for achieving the high degrees of accuracy and dependability requisite for autonomous urban transit systems. Figure 3.10 delineates the sensor system's hardware configuration and the arrangement of embedded PCs.

3.3 Safety-System

Pedestrian areas are unstructured pathways having direct access to buildings and living premises. Main focus of such areas is to have better accessibility for pedestrians without commercial traffic. Such zones are best suited for the environment in terms of traffic, attractiveness, safety, and noise/air pollution. Having the appreciation for all the given benefits for pedestrians, people often tend to walk carefree along with their adolescent and pets walking unrestrictedly. The presence of such obstacles around the Autobus puts safety a foremost concern in autonomous driving. This necessitates an uninterrupted safety system to be able to react in critical situations. Often perception systems using sensors, such as cameras and LiDARs, are prone to errors. These errors, perchance, could be due to missing measurement or fault in the system itself. Misclassifications may arise from

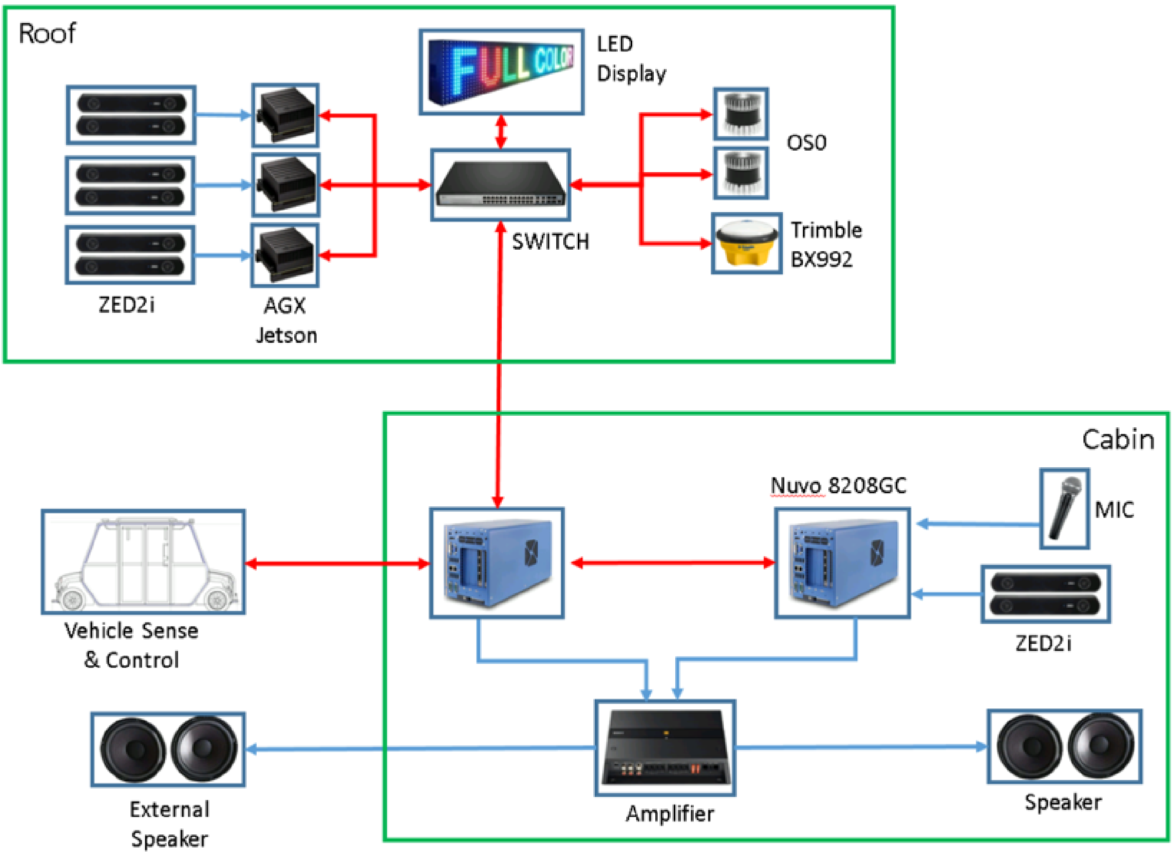


Figure 3.10: The schematic shows the extensive deployment of sensor arrays on Autobus and the related distributed computing architecture. The 48 Gbit/s and 1000MBit/s LAN switches facilitate fiber optic data transmission to main PCs and ensure high bandwidth and low latency communication, which is essential for real-time processing of substantial sensor data. The embedded PC is assigned to special functions and receives data from sensors such as LiDAR, cameras and GPS. The distributed computing model allows parallel data processing and improves the system's ability to process and interpret high-definition images and 3D mappings. The three Jetson AGX panels installed on the roof are specialized in the pre-processing and operation of neural networks and reduce the computational load of the main processors. This strategic setting supports the high accuracy and reliability of the system, which are essential to autonomous navigation in urban environments.

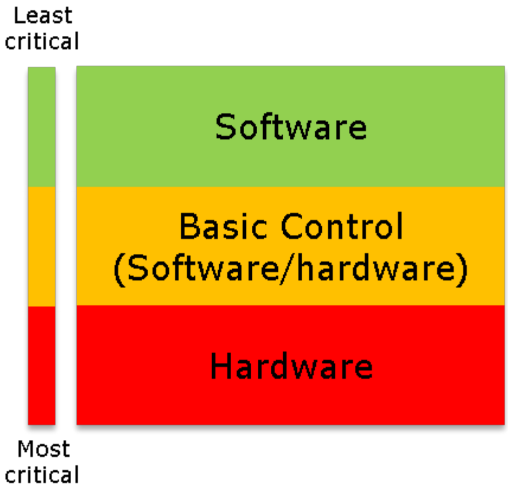


Figure 3.11: The diagram represents the stratification of safety system in Autobus, categorizing them into three primary levels: Software, Basic Control (integrating both software and hardware components), and Hardware. Each level is color-coded, with green indicating the software layer, yellow for the basic control layer, and red for the hardware layer, signifying the increasing level of criticality from top to bottom.

detection algorithms, or instances may occur where no classification is rendered. Hence, a hardwired safety concept is designed in the Autobus.

3.3.1 Safety Concept

The configuration depicted in Figure 3.11 illustrates a tiered safety architecture, a work done on safety configuration for Autobus [Jan 21a], segregated into three distinct levels, each integral to the Autobus's overall safety strategy. The lowest block represents the most critical safety level, denoted by its foundational position in the hierarchy. This critical layer is characterized by its autonomy and resilience, functioning independently of external inputs. Its primary role is to ensure a 'safe stop' in the event of a system failure or when a predefined safety parameter is triggered. This fundamental safety mechanism operates as a fail-safe, designed to cease all vehicle operations securely, thereby mitigating the risk of accidents or damage.

The intermediate layer, as visualized above the critical layer, contains non-critical safety functions. Although this layer is less critical, it is essential for safe operation of the system and provides additional control and balance. The upper limits of this intermediate layer are the minimum safety thresholds of the critical layer below. It is equipped with a monitoring dog timer that continuously monitors inputs at higher levels and starts stopping in the absence of inputs. Therefore, any security measures implemented at this level must not compromise the integrity of the critical layer's essential safety functions.

The topmost block represents the least critical layer for immediate safety response, but is crucial for complex operation of the system and interactions with its environment. This layer includes high-level algorithms and decision-making processes. Although it is the least safe layer, it is still designed to respect and maintain the safety constraints defined by the fundamental layer. It includes a safety module to detect immediate obstacles, using hydraulic brakes instead of emergency brakes. The basic principles of this research apply to this layer.

Each of these layers is designed to work in concert, forming a cohesive safety net. The safety concept ensures that higher-level malfunctions are contained and managed without impeding the operation of more critical safety functions. The subsequent section provides a detailed description of the lower most layer, i.e., Hardware.

3.3.2 Safety Hardware Design

In a vulnerable situation where collision with humans can occur, the aforesaid errors cannot be tolerated. This demands in designing a safety certified system for closeby interaction process which is part of this work for assisting the smart interaction concept. Safety certified systems are designed based on Safety Integrity Level (SIL) [Gulland 04] and Performance Level (PL). Such systems are directly connected to the safety chain of a vehicle. In general, the certification process is typically conducted throughout the manufacturing phase and is completed before the system's finalization. However, for Autobus, a unique prototype, certification was not performed during its manufacturing. To facilitate potential future certification, safety modules adhering to specific SIL and PL were incorporated. These modules are directly integrated into Autobus's safety chain. In Figure 3.12, a schematic delineation of the safety system's interconnections within the Autobus is presented. A safety circuit utilizing a 12V supply line, depicted in brown, traverses the SICK safety monitoring system, proceeding to Relay 1 and Relay 2. These components govern the activation and deactivation of the vehicle's motors and emergency brakes, respectively. The depicted state of the system indicates operational safety, with current coursing through the 12V line ensuring the relays are closed, thereby activating the motors. In the event of a circuit interruption instigated by any element within the SICK system, the relays transition to an open state, resulting in motor deactivation and the concurrent activation of the emergency brakes. For the overview of the coupled components, figure A.2 in Appendix shows a detailed schematics of the components connected to the safety chain of the vehicle.

Table 3.11 provides an overview of specific conditions that trigger the activation of safety protocols in a vehicular safety system. These conditions, categorized under 'Causes,' indicate various failure or unsafe scenarios that can occur during the operation of the vehicle. The 'Status' elucidates the specific state or action that corresponds to each cause, for instance, a vehicle exceeding a speed of 7 Km/h or the failure of a battery. Upon the occurrence of any listed cause, the corresponding response of the vehicle's safety system is twofold, as indicated in the 'E-brake and Motor' column. This column is merged across all conditions to denote a uniform response: the emergency brake (E-Brake) is activated ('enabled'), and the Motor is simultaneously deactivated ('disabled'). This standard response is designed to immediately halt the vehicle and prevent any further motion that could exacerbate the situation or pose additional risks. The conditions outlined include scenarios such as 'Over speeding,' 'Battery Failure,' activation of the 'Emergency button,' 'Safety bumper' being pressed, 'Door Contacts' status during driving, and the interruption of 'Laser scanners' safety field. Each of these scenarios is deemed significant enough to warrant an immediate stop of the vehicle to ensure the safety of the occupants and surrounding environment.

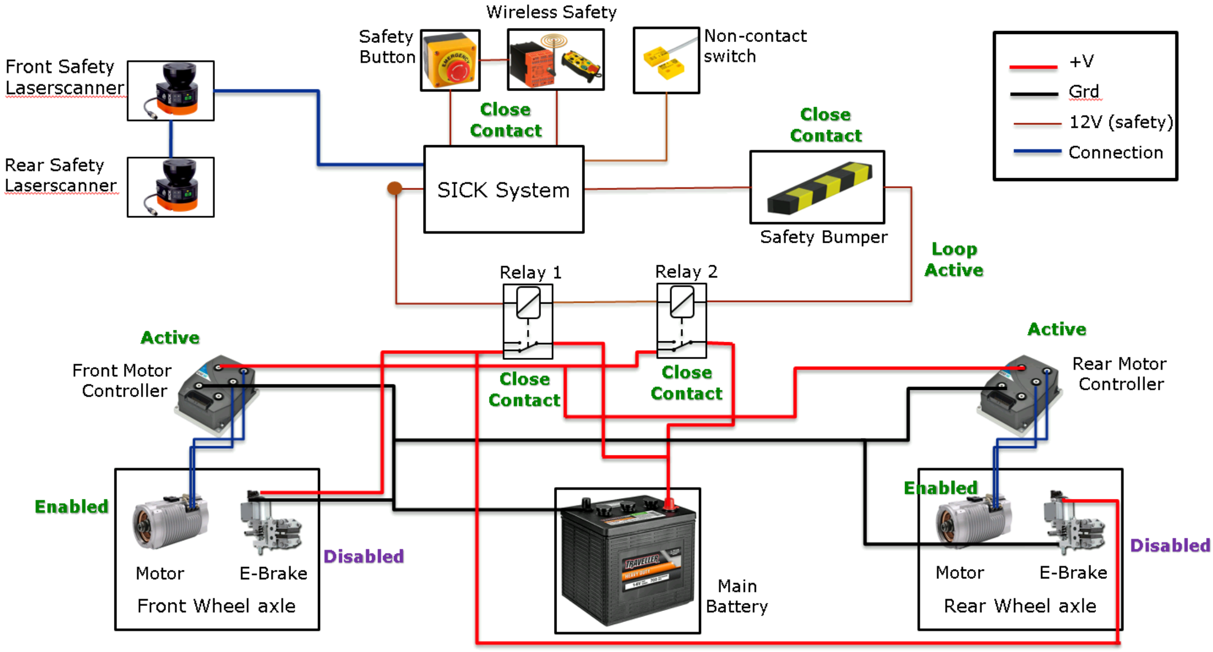


Figure 3.12: The diagram illustrates the interconnection of front and rear safety laserscanners, a wireless safety system with a non-contact switch, SICK system for safety monitoring, and safety bumpers. It includes the main battery supplying power to front and rear motor controllers through relays. The loop is active with close contact indicating that the motors are enabled and the emergency brakes (E-Brakes) are disabled, ensuring operational safety.

Causes	Status	E-brake and Motor
Over speeding	>7 Km/h	E-Brake = enabled
Battery	Failure	
Emergency button	Pressed	
Safety bumper	Pressed	
Door Contacts	Opened (Driving)	Motor = disabled
Laser scanners	Safety field interrupted	

Table 3.11: The table shows failure/unsafe conditions for the safety system. Depending on the cause and the corresponding status, the E-brake and motor are enabled and disabled, respectively.

3.3.3 Components and Working

In the design of the safety system, several key components are required, as indicated in Table 3.11. These include a speed encoder module to monitor the vehicle's speed, a battery monitoring relay to track the battery status, an emergency stop switch to quickly shut down, a magnetic switch to control the door, and a laser scanner to detect obstacles. All these components are sourced from SICK, a specialized safety system company, and interact with specialized programmable CPUs through SICK's safety design software. These relays are particularly crucial because they are activated within the vehicle's safety chain. The detailed technical specifications for each module can be found in Appendix B. As laser scanners play an important role in reducing safety risks and avoiding obstacles, detailed descriptions of these components are given in the following sections.

Assessment of Laser-scanner for Enhanced Outdoor Safety Operations

For the purpose of scanning for obstacle, the system uses 2D-LiDAR mentioned in Subsection 3.2.1.1 with the same configuration. The scanners in addition for higher level navigation algorithm is connected to the processing unit of the SICK for wired safety. OutdoorScan3 used for the purpose has safety related feature pointed in Advantages 3.1. OutdoorScan3 is distinguished by its certification for outdoor use, bolstered by a robust design that incorporates enhanced shock resistance. It is capable of operating under moderately adverse weather conditions, such as rain, snow, and fog. The OutdoorScan3 complies with Safety Integrity Level 2 (SIL2) and Safety Integrity Level for Control Level 2 (SILCL2) standards, ensuring a high level of reliability and safety.

Advantages 3.1:

OutdoorScan3

- Ensures high productivity by enabling safe interactions between humans and machines in outdoor environments.
- Maintains exceptional operational reliability, regardless of adverse weather conditions.
- Designed with user convenience in mind, making it ideal for outdoor applications.
- Features flexible protective fields that can be customized for secure automation tasks.
- Offers straightforward diagnostic access, allowing for quick problem-solving.
- Supports a steady and dependable flow of materials, both inside and outside, adapting to varying weather conditions.

In the event of a system fault, the OutdoorScan3's safety output is designed to switch to a logic 0 state over the network. This state change is detected by the main module of the system, triggering the activation of vehicular safety mechanisms.

A primary function of the OutdoorScan3 involves the generation and regulation of protective and warning zones (enumerated below). These zones are defined as adjustable geometric areas within the scanner's range. Upon detection of an obstacle within these preset areas, the scanner alters its output signals to indicate safe or unsafe conditions. Since the stopping distance of the vehicle is not the same, it is important to have dynamic safety fields. Therefore, outdoorScan3 allows capability for dynamically modifying the safety fields based on vehicle's stopping distances. Examples of the configurations of these monitoring zones are illustrated in Table 3.12.

1. **Warning field:** For functional use only (shown in yellow in Table 3.12), with a range of 40m.
2. **Protective Field:** For detection and protection (shown in red in Table 3.12, with a range of 4m.

As delineated in Table 3.5,the OutdoorScan3 system can monitor up to eight different areas simultaneously, as shown in Table 3.5. This includes the specifications of the protective

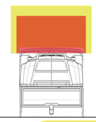
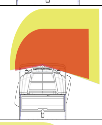
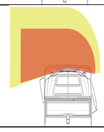
Monitoring Case	Description	Fields
Driving straight	When driving straight the fields are perfect square	
Turning right	At full steering towards right	
Turning left	At full steering towards left	

Table 3.12: This table exemplifies the OutdoorScan3's capability to dynamically adjust safety monitoring zones according to vehicle steering direction. The "Monitoring Case" column identifies three distinct scenarios: driving straight, turning right, and turning left. The "Description" column clarifies the vehicle's steering status. In "Fields," the diagrams represent changes in the warning (yellow) and protective (red) fields relative to the vehicle orientation. When driving straight, both fields form a perfect square, indicating equal safety distance around the vehicle. During turns, the warning field expands opposite to the turn direction, reflecting the need for a greater safety margin due to the vehicle's momentum and altered stopping distance. These adjustments are vital for maintaining vehicle safety and are part of the system's broader capability to monitor multiple zones simultaneously, enhancing responsiveness to operational conditions.

and warning fields managed by these scanners. This safety function is enabled through the "Safety Designer" software, allowing to create up to 4 meters of protective fields in radius. These fields can trigger the braking mechanism, if needed. These fields can be adjusted in terms of the shape, size and conditions for the safety system to encounter various scenarios. This flexibility helps to tailor safety measures to the specific environment and situation of the vehicle. Consequently, the speed of the vehicle is divided into seven different zones, with speeds ranging from 0 to 6 kilometres per hour. Each speed zone corresponds to a specific length of the safety field, as illustrated in Figure 3.13, which shows how speed zones correspond to the length of the safety field.

Dynamic Configuration of Safety Fields in Response to Vehicle Speed and Steering Dynamics

Two encoders are integral to this system. The first encoder is connected to the motor and is responsible for measuring the speed of the vehicle. The second encoder is associated with the steering mechanism, providing data on the steering angle. The interplay between these two encoders is crucial for the dynamic adjustment of the safety fields.

The system architecture includes a safety field that is directly proportional to the stop distance of the vehicle and varies according to the speed. A progressive velocity is needed to accommodate the required extended stop distance by extending the protective area proportionally. This adaptive mechanism ensures that the vehicle is constantly protected from potential obstacles at all speeds and significantly reduces the probability of collisions. Figure 3.13 shows the empirical relationship between the speed of the Autobus and the length of the safety field. These relationships were determined by experimental methodological experiments, which involved driving the automobile at different speeds

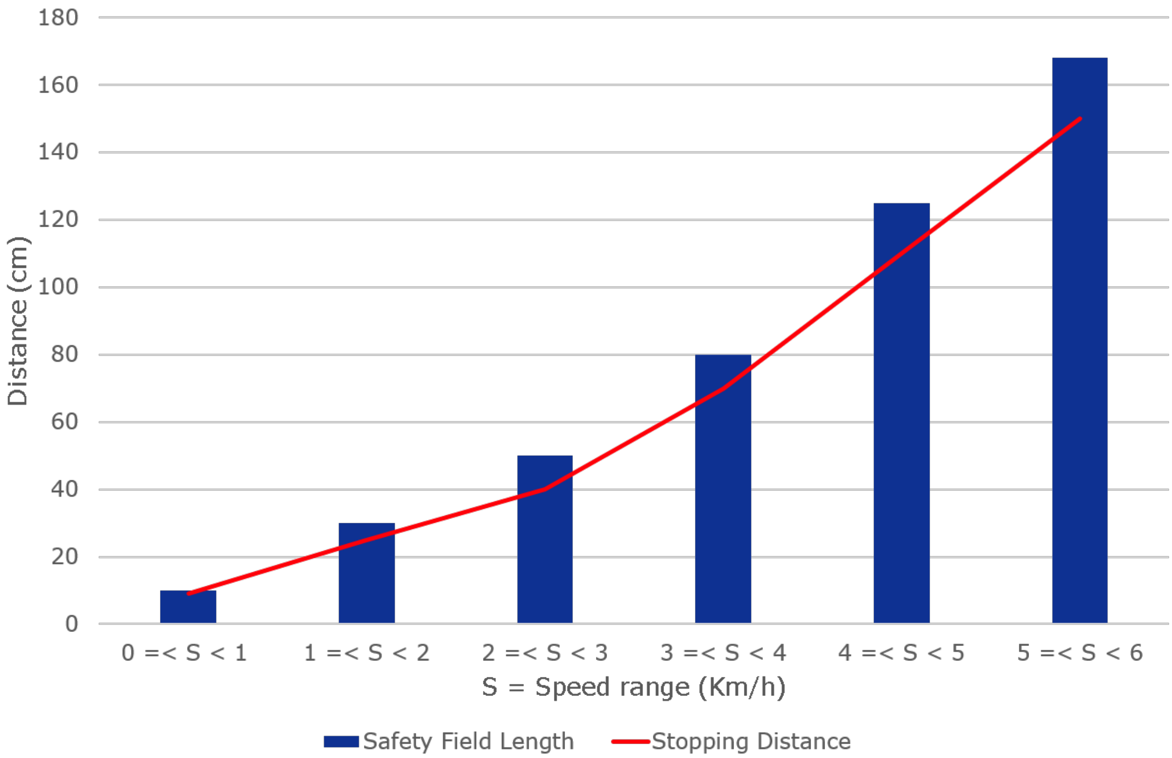


Figure 3.13: The plot illustrates the empirical correlation between the velocity of the Autobus (denoted as 'S' for speed range in km/h) and the corresponding safety field dimensions, which are crucial for vehicular safety. The blue bars represent the length of the safety field at various speed ranges, demonstrating an increasing trend as the velocity escalates. This increment is designed to ensure a proportional buffer zone for stopping the vehicle safely. The red line indicates the actual measured stopping distances, obtained through controlled testing wherein the vehicle was subjected to a series of stops from different speeds. These stopping distances represent the average lengths required for the vehicle to come to a complete halt across multiple trials. The safety field lengths (blue bars) are consistently set larger than the empirical stopping distances (red line) to provide an additional safety margin, which is especially heightened within the highest speed bracket. This safety margin is calibrated to be sufficient yet not excessively large to accommodate the restricted spaces often encountered in pedestrian zones. The graph conveys the strategic balance between safety and operational practicality, ensuring that the Autobus maintains a safe operating distance from obstacles while accounting for dynamic speed variations.

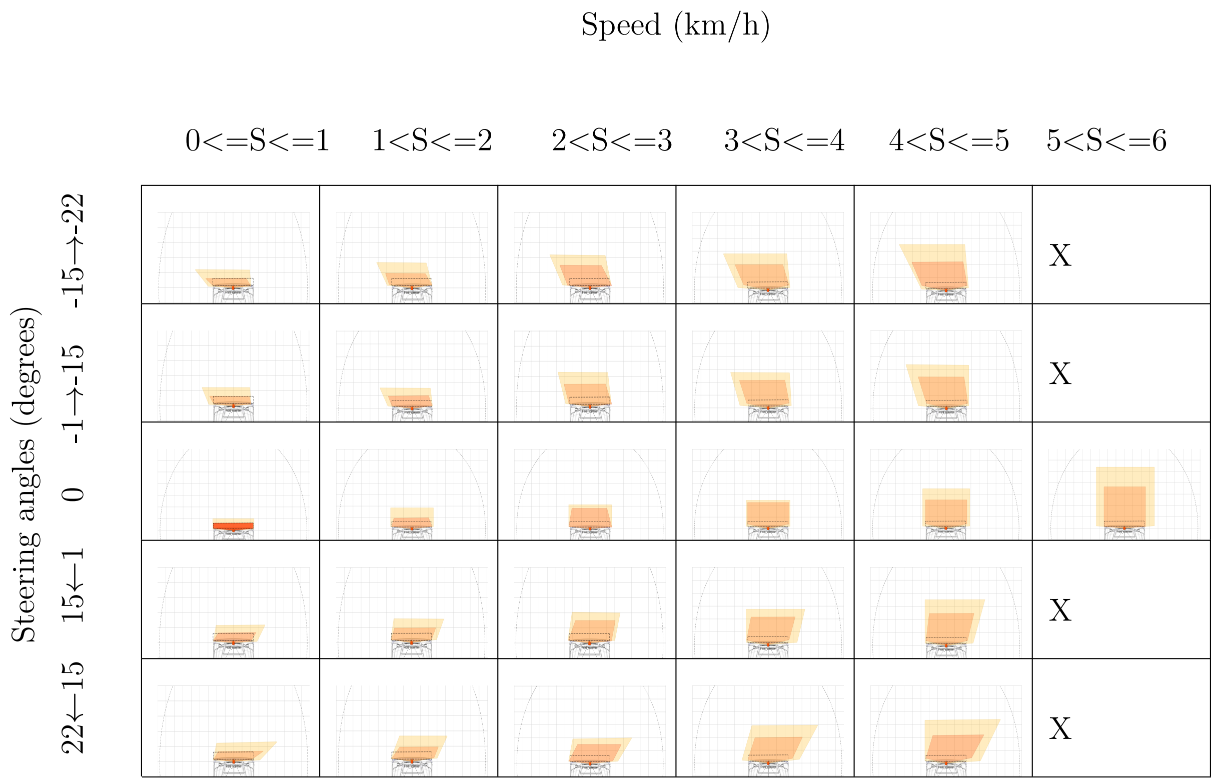


Table 3.13: The table presents an intricate delineation of the velocity and steering angle intervals, alongside the identifiers for the respective monitoring cases. It articulates the distribution of field samples, categorized into protective and warning types, assigned to the outdoorScan3 for the purpose of operational surveillance. Instances that are marked with 'X' signify scenarios that are considered inapplicable within the monitoring framework; At elevated velocities, the act of steering is restricted. Furthermore, the system is designed to reduce its speed in response to steering inputs, ensuring safety and control during operation.

and initiating safety stops on specific markers to measure the necessary stopping distance. Table 3.13 shows the average stopping distance of multiple trials at different speeds, while the blue bar shows the length of the protected field in the direction of driving, which is deliberately set to be larger than the stopping distance in order to include additional safety margins. This surplus margin increases marginally at the higher speeds, but this margin cannot be increased significantly given the constraints of the narrow pedestrian zone. In addition, the responsiveness of the safety mechanism includes changes in vehicle speed and direction, allowing an integrated safety framework that anticipates various operational environments and the dynamic maneuvers of the vehicle.

The laser-scanner, specifically the OutdoorScan3 model, has the capability to be configured with up to 128 distinct monitoring cases. Each of these monitoring cases can be associated with a specific set of field configurations, which are activated based on inputs received by the laser-scanner. The activation of a particular monitoring case is contingent upon the values derived from the speed and steering encoder readings. Table 3.13 shows the different monitoring cases for range of steering angles and speed.

When considering the impact of speed alone, it becomes imperative to adapt the shape of the safety field in accordance with the steering dynamics of the vehicle. This necessity stems from two primary reasons:

1. **Complexity in Steering Dynamics:** The vehicle is equipped with a double-Ackermann steering system, which allows for symmetrical steering of both the front and rear wheels. This configuration typically results in a reduced turning radius. In scenarios where the vehicle is maneuvering a curve, it is crucial to design the protective fields in such a way that they can be timely intercepted by pedestrians walking alongside the vehicle. The aim is to ensure that the protective fields are sensitive and responsive enough to detect pedestrians in close proximity, especially during tighter steering maneuvers.

2. **Considerations for Narrow Turns:** The design of the fields becomes increasingly critical in situations involving narrow turns, as illustrated in Figure 3.14. When approaching a narrow turn, the shape of the safety fields must be adjusted to prevent the inadvertent inclusion of obstacles not present on driving path, such as peripheral bushes or stationary structures. Since the Autobus makes a turn, these obstacles are not existing in the drivable way. If these objects are consistently detected as potential hazards, the vehicle would be prone to unnecessary safety stops. Therefore, it is essential to tailor the configuration of the protective fields to distinguish between actual hazards and benign objects, particularly in constrained turning scenarios.

These considerations underscore the importance of a dynamic and context-sensitive safety system. By adjusting the protective fields based on steering dynamics and the vehicle's immediate environment, the system enhances safety without compromising operational efficiency.

Figure 3.14: Turning example of Autobus in small spaces.

To investigate the proper usage of the aforementioned postulation, a sensible configuration of the system is of significance. This is the first step to start-off before proving the hypothesis. It has to be done meticulously to avoid any knock down effect of error which will further increase the discrepancy of the tests made during the whole process. This

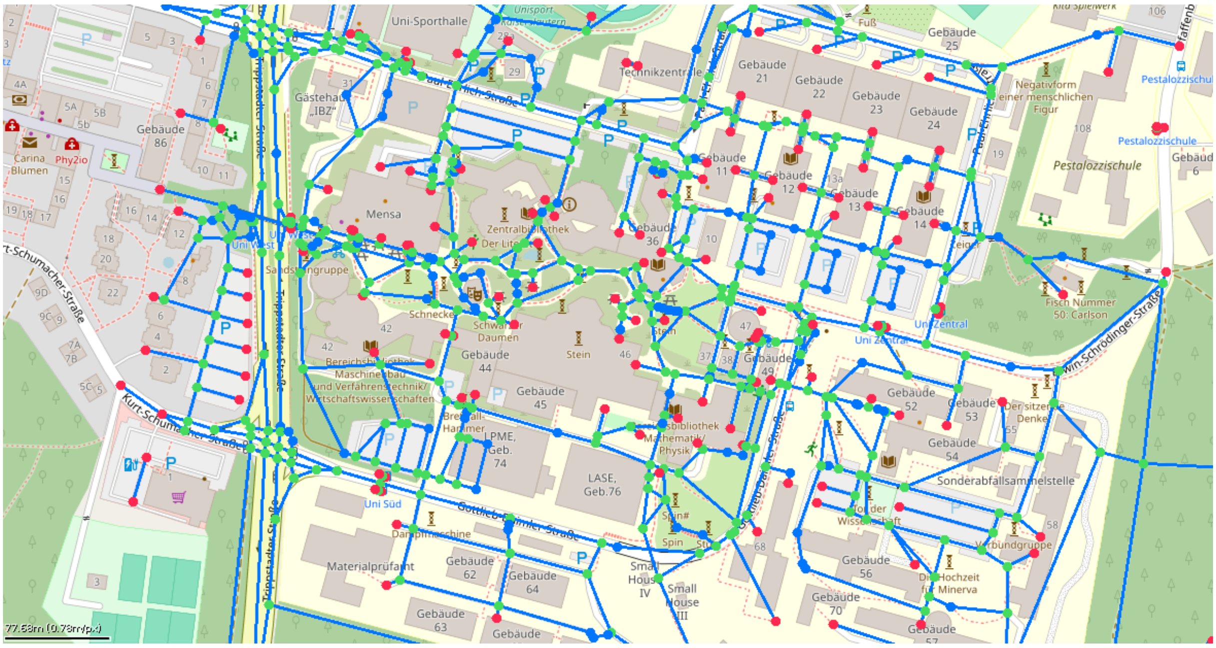


Figure 3.15: OpenStreetMap Visualization of the Autonomous Vehicle Testing Ground. This map represents the designated testing area for the Autobus, highlighting the intricate network of pathways, including pedestrian walkways and vehicular roads within a university campus setting. Key locations and infrastructure are annotated to facilitate navigation and operational testing.

may include ineffectual location of a certain module which the subject might not even consider. To avoid such a scenario, it is validated through different means, depending on the module, the correctness based on its application. Since the vehicle also interacts with the pedestrian, section of interacting modules explains the reasoning of each module.

3.4 Test Environment Description

General description of the pedestrian zone and its advantages are well explained in the introduction chapter. This section elaborates the environment structure and dynamics for better understanding the use case in experiments. The driverless bus (explained in section 3.1) is driven within the campus of RPTU Kaiserslautern. Campus environment offers an unmarked network of pavement for accessibility to every building within and outside the campus. The connectivity of the testing area is shown in Figure 3.15 for RPTU campus.

The roads are designed to accommodate the navigability of vehicles, taking into account the requirements of transport, construction and emergency vehicles. This design allows for relatively easy operation of vehicles which has size up to minibus. The minimum width of these paths is shown in Figure 3.16. The path is not only narrow, but also incorporates curves, which add complexity especially in the presence of pedestrians. In such narrow spaces it is not possible to overtake pedestrians as well.

Figure 3.16 also shows a detailed view of the selected narrow sections in the campus environment. These sections have a width greater than one meter of the width of the vehicle. Two places that are frequently crossed are highlighted to better understand. The areas marked with green are the most narrow sections of driving routes and the navigation

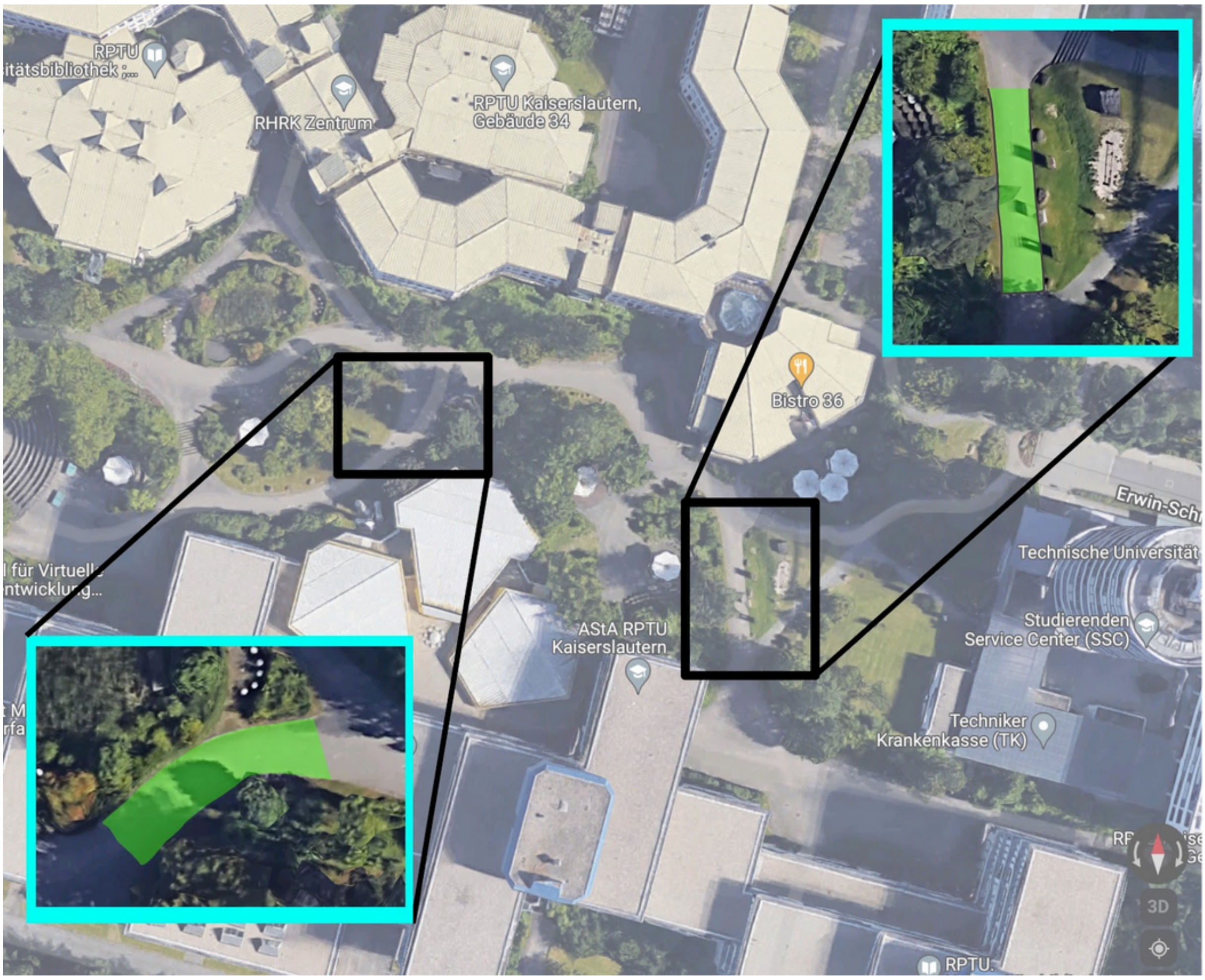


Figure 3.16: Aerial Perspective of RPTU Campus highlighting narrow pathways. This satellite image illustrates the campus layout with specific emphasis on the constricted pathways, delineated by black outlines and augmented insets, that are critical for the navigation of the Autobus within the testing environment. The insets provide an enhanced view of the areas of interest, indicating the potential navigational challenges posed by these narrow routes.

becomes more and more difficult. This complexity is further reinforced in areas where these narrow paths intersect with turns.

To really understand how well a proposed system configuration works, there is a need to plan carefully and thoroughly from the start. This initial setup is crucial for making sure the hypothesis is correct. Getting this right helps avoid a chain reaction of mistakes that could hamper the results of later tests. Errors in the beginning, such as placing parts incorrectly or not integrating modules properly, can lead to false conclusions or even system failures. To circumvent such scenarios, each module's functionality and integration are meticulously verified through various methods appropriate to their specific application and context within the system.

3.5 Simulation Environment

In the field of autonomous driving, especially for pedestrian areas, the importance and complexity of simulation environments are essential. Realistic simulations faithfully reflecting the conditions of the real world are imperative to ensure that the trials carried out have virtually similar implications to those carried out in the physical field. This idea allows for preliminary validation of algorithms and provides an important step before empirical tests.

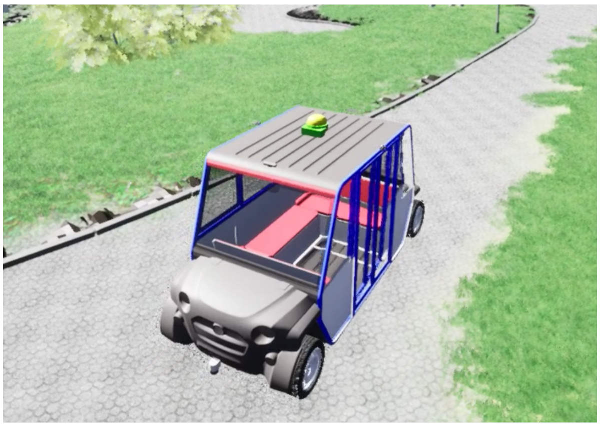
This section explores the complex simulated environment used to test the Autobus in a scene with a dense pedestrian population. Simulation is a controlled, but multifunctional platform that allows researchers and engineers to replicate a wide range of pedestrian behaviours and traffic scenarios that driverless minibuses might encounter in urban environments. The advantages of using simulated environments for autonomous driving are many:

- It provides a safe and risk-free setting where potentially hazardous situations can be recreated and studied without any real-world consequences. This is particularly crucial when examining the interactions between Autobus and pedestrians, where safety is paramount.
- Simulation allows for the testing of extreme or rare scenarios, which although infrequent in reality, are essential for the comprehensive training and evaluation of Autobus algorithms.
- It enables rapid iteration and refinement of algorithms and systems. Changes can be implemented and tested much faster in a virtual environment than in physical testing, accelerating the development process.
- Simulations can generate vast amounts of data, essential for training machine learning models that underpin Autobus decision-making processes.
- It offers scalability, allowing researchers to test a wide range of scenarios and conditions, from different weather and lighting conditions to varying pedestrian densities and behaviors, something that is time-consuming and resource-intensive to replicate in real life.

Based on the fundamental role of simulation in autonomous driving research, our work uses the Unreal Engine (UE) as a central platform for the simulation environment. UE is known for its high-fidelity and photorealistic rendering capabilities and provides an ideal backdrop for sophisticated and complex scenarios in pedestrian areas. UE is widely praised in the game industry for its ability to create an immersive and visually astonishing environment, and it has also proved to be an expert in the field of autonomous vehicle simulation.

The photorealistic nature of UE provides a major advantage to the Autobus perception system. Since Autobus relies heavily on cameras and sensors to interpret the environment, UE's realistic textures, lighting, and dynamics can be replicated. This level of detail ensures that simulations are not only hypothetical scenarios, but are close to real-world approximations. This realism is crucial for training and verifying Autobus' perception algorithms, ensuring that they can accurately recognize and respond to the behaviour of pedestrians and various environmental conditions.

Moreover, the robust UE physics engine and its ability to simulate complex interactions in real time further strengthen its adaptability for Autobus testing. It allows the creation of



(a) This figure illustrates the sophisticated three-dimensional simulation model of the Autobus. The model has been crafted to closely replicate the real vehicle's design specifications, from its structural dimensions to its color scheme and surface textures. It has been developed with an aim to simulate not only the visual aspects but also the operational dynamics of the real-life counterpart. The virtual Autobus is programmed with kinematic equations that mirror the actual vehicle, ensuring it exhibits comparable movement dynamics, acceleration, and braking behaviors. The inclusion of a double Ackermann steering mechanism further refines its movement simulation, providing a realistic representation of the vehicle's maneuvering capabilities within the simulation environment. This visual and operational parity allows for precise analysis and testing of the Autobus's performance in a controlled virtual setting.

(b) Displayed here is the intricate virtual recreation of the RPTU-Kaiserslautern campus as it exists within the UE. This simulation of the campus encompasses an approximately precise and to-scale representation of the campus buildings, roads, pathways, and surrounding landscape, modeled to reflect the actual campus's architecture and spatial layout. The simulation extends to include varying weather conditions, providing a diverse range of scenarios for evaluating the sensory and perception systems of the Autobus in conditions that mimic real-world challenges. The precision in the environmental modeling ensures a high-fidelity platform for the comprehensive testing and optimization of the vehicle's performance metrics in a realistic setting.

Figure 3.17: Autobus and RPTU Campus Model in Simulation

dynamic, unpredictable pedestrian behavior, which is essential to test the limits of AV decision-making algorithms. Using UE, various pedestrian scenarios can be simulated, from typical city crossings to more complex, non-structured environments, providing a comprehensive test platform for advanced autonomous driving technology. This section will explain in greater depth how we used UE's capabilities to create a realistic and effective simulation environment for our research into autonomous driving.

3.5.1 Tailored Design Specifications

Continuing from our look at UE's ability to simulate realistic environments for this thesis, we discuss about how the "Autobus" model and the RPTU campus within the engine are carefully recreated. This step is crucial for making a simulation that truly reflects real-world conditions, giving us a robust platform to test and improve the interaction systems concept for this thesis.

Vehicle Model:

Autobus depicted in Figure 3.17a and detailed in Section 3.1 were carefully reconstructed in simulation to imitate their realistic analogs in aesthetics and operational behavior. This process requires extensive three-dimensional rendering of the Autobus, achieving a high accuracy of dimensions, structure design, color palette, and surface details. The aim was not only to visually resemble a physical vehicle, but also to generate a simulated vehicle that mimics the physical characteristics and performance metrics of a physical vehicle, including movement dynamics, acceleration patterns, dimensions and acceleration/deceleration characteristics. Therefore, the mechanical equations governing the motion of virtual vehicles are consistent with those of the actual system and ensure identical time and space effectiveness in achieving its navigation objectives. In addition, the simulation features a double Ackermann steering mechanism that enhances its functionality.

Campus Model:

Furthermore, the RPTU-Kaiserslautern campus was recreated in UE as shown in Figure 3.17b, the real environment in which the Autobus is designed to work. It involved detailed mapping and modeling, replicating buildings, roads, pathways and other campus environmental features. Replication is not only about visual accuracy, but also about maintaining the correct scale and spatial relationship between different elements, in order to ensure that simulation is consistent with the actual campus architecture and design. It also includes different weather conditions to test the performance of the perception system.

Pedestrian Model:

To test the Autobus interaction system, pedestrians are also incorporated into the simulation. These characters are programmed to imitate the different pedestrian behaviours observed in real life, such as walking, crossing the streets, and reacting to the Autobus. In one of our earlier work [Jan 21b], we designed the movements and behaviours of the characters as realistically possible and provided a realistic context for testing the Autobus interaction system. The characters move around simulated RPTU campus and create dynamic scenarios in which the Autobus must interact with these pedestrians in real time. Figure 3.18 shows the trajectory of two characters when they encounter an upcoming vehicle. The plot reveals the characters' path deviations from their original path, showing their reaction to the approaching vehicle by turning around and changing direction. This is similar to reality, where pedestrians partly take responsibility for opening the way to vehicles.

Sensor Model:

In addition to physical characteristics, the real Autobus sensors are replicated in simulation as well. These include LiDAR, cameras, GPS and other sensors used in car buses to perceive the environment. Replicating these sensors in a virtual environment is crucial to testing the autobus' perception system. The simulated sensors are programmed to mimic the functions and limitations of real-world sensors, and generate data for the use of Autobus algorithms to navigate and interact in the simulation environment.

Using UE to recreate Autobus, RPTU campus, its sensor suite and pedestrian dynamics, required to create a very detailed and authentic simulation framework. This virtual platform was crucial to the iterative development and validation of Autobus' interactive systems. It provides a rigorously controlled but very realistic environment for a comprehensive

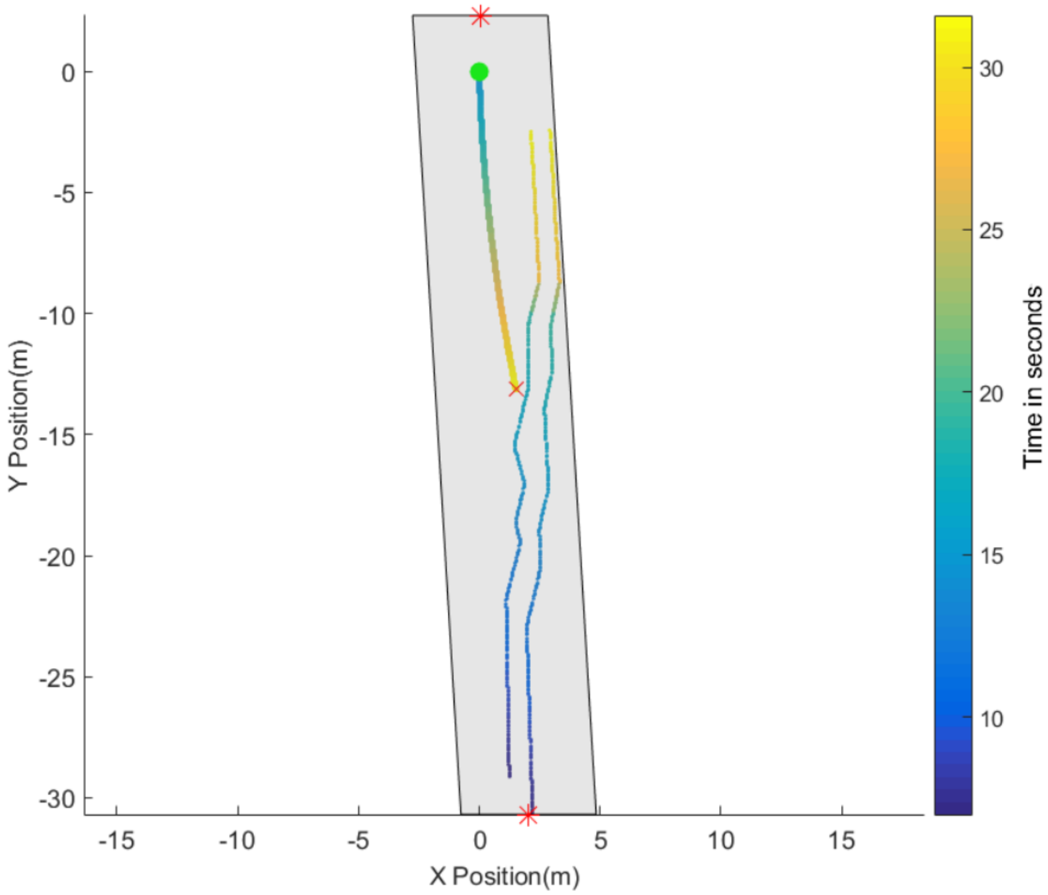


Figure 3.18: The plot captures two pedestrian characters simulated in the campus environment of the RPTU campus, showing their reactions to approaching vehicle. The paths of each character are represented by different colored lines, which change from blue to yellow to red over time, showing their initial paths and their subsequent deviations to the vehicle. These deviations are represented as the deviation of the original path towards a new trajectory, which symbolizes the behavior of pedestrians who yields. This behaviour replicates realistic pedestrian actions documented in our previous work [Jan 21b], showing that pedestrians share the responsibility for avoiding vehicles by actively changing their course. Autobus's ability to interact with these dynamic agents shows the sophisticated levels of the navigation system and interaction system in simulation settings.

evaluation of vehicle performance and boundary conditions. In this virtual environment, a favourable application of simulation methodology was particularly evident when Autobus' end-to-end driving paradigm was assessed, as demonstrated in two of our papers [Jan 23b, Jan 23c]. The end-to-end model was re-adapted to the actual campus parameters to achieve promising results in the real-world environment, and was deployed for empirical tests on physical vehicles. The empirical results show a high correlation with the simulation results, underlining the value of the simulation in the prediction of system performance in the real world.

3.6 Autonomous Navigation Software Architecture

The REACTiON architecture [Wolf 18], as delineated in Figure 3.19, represents an evolving model for off-road autonomous navigation, specifically adapted for the Robotic Research Lab (RRLAB) autonomous vehicles at RPTU Kaiserslautern-Landau. Developed within the modular FINROC environment [Reichardt 12], the architecture accommodates all requisite components for autonomous navigation—including mapping, classification, user interaction, and fail-safe measures. Its hardware interface ensures seamless integration with the vehicle hardware, simulation environments, and data recording functionalities.

Like other RRLab vehicles, the Autobus also uses REACTiON architecture for basic autonomous driving. The need to adapt this framework to pedestrian zones is evident in comparison with unstructured environments, such as pedestrian areas with random crossings and lack of markings. The adaptation consists of innovative integration of the "interaction" component, and this thesis highlights the distinctive contribution shown in Cyan in Figure 3.19. As stated in Chapter 1, this interaction is essential to avoid the potential for the Autobus to enter a "frozen" state when they face pedestrians, who can merely be considered an obstacle in a dense environment and stop its progress.

The REACTiON architecture is structured into discrete groups to facilitate perception and control, thereby allowing for independent and modular modifications to enhance system adaptability. This structure ensures that interaction components, fundamental to the interaction with pedestrians and critical to the framework's operation in pedestrian zones, are methodically placed within the architecture. These components facilitate the complete cycle of interaction—from perceiving pedestrians to the expression of verbal and non-verbal cues.

The technical design considerations for the interaction processes is shown in Figure 3.20, illustrating how data flow between the framework's components underpins the entire cycle with respect to smart interaction concept. This integration of interaction components not only provides an efficient data access path but also ensures a comprehensive understanding of the system's functionalities. Figure 3.20 further elucidates the deployment of these components within the groups as initially presented in Figure 3.19, reinforcing the framework's approach to smart interaction strategy in autonomous driving for Autobus. The different groups used for interaction are described as follows.

Hardware: The hardware interface includes all the sensor systems and driver interfaces. For stereo cameras, which are mounted on the roof are linked to the Jetsons for initial data processing, which involves detecting skeletal structures. This data, along with images and point clouds, is directly imported into the hardware for further analysis. The skeletal data

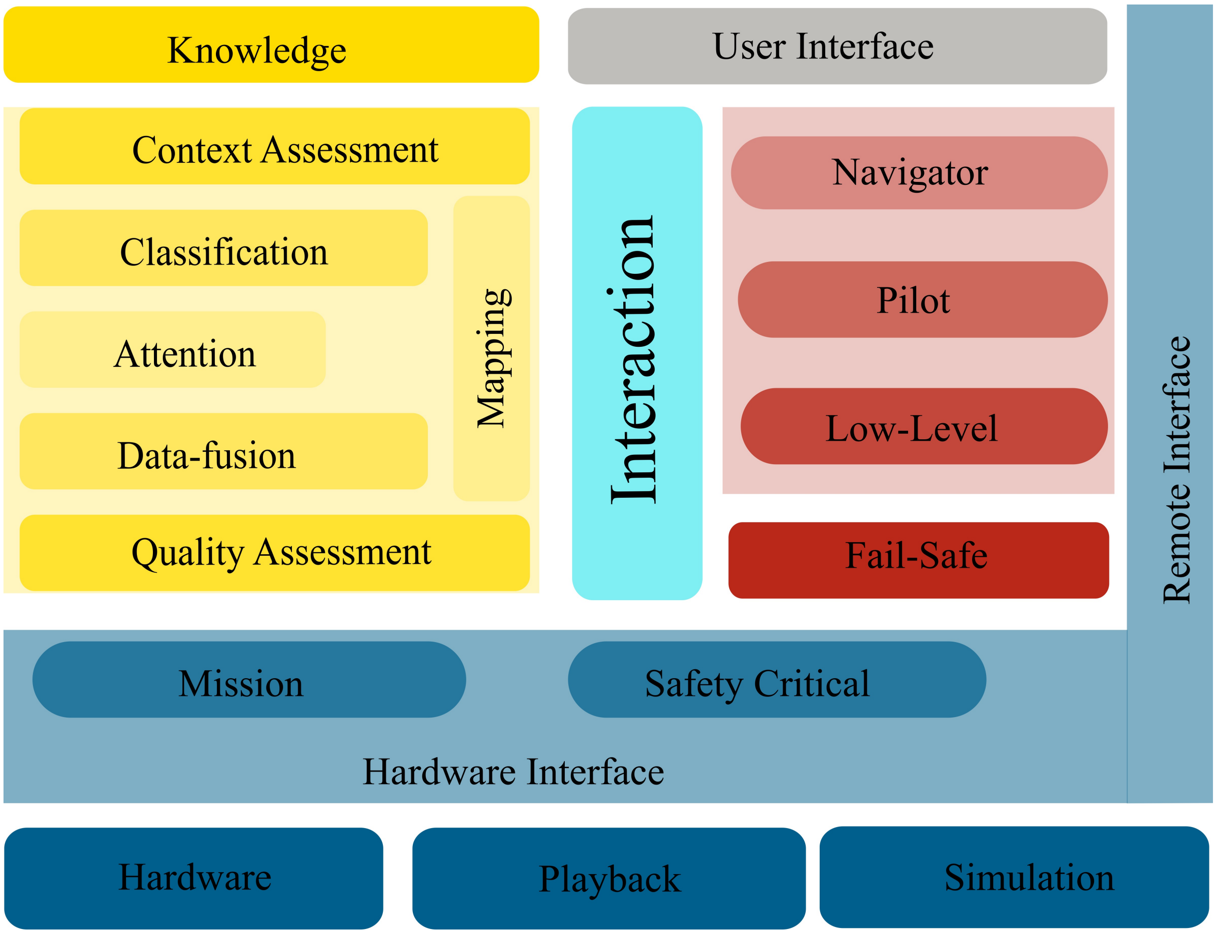


Figure 3.19: This schematic represents the hierarchical organization of the REACTiON architecture used in the autonomous driving system of the Autobus. The architecture is segmented into four main categories: Knowledge, Interaction, Mission, and Safety Critical, with a Remote Interface for external interactions. Each category is subdivided into functional modules: Knowledge encompasses Context Assessment, Classification, Attention, Data-fusion, and Quality Assessment; Interaction is divided into User Interface functions such as Navigator, Pilot, Low-Level control, and Fail-Safe mechanisms; Mission and Safety Critical elements include Hardware Interface operations; and the lowermost tiers consist of Hardware, Playback, and Simulation components for system execution and testing. This structure facilitates a modular and scalable approach to autonomous driving by delineating clear functional groups that can be independently modified and optimized.

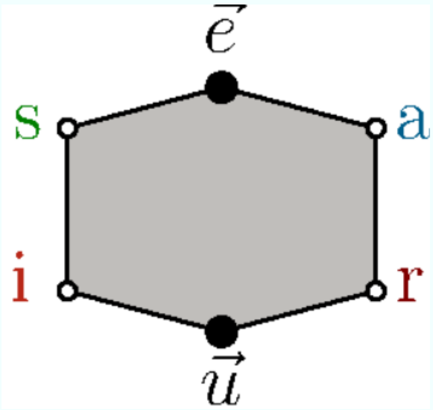
is converted into 3D points and then transmitted along with other sensor data through ports for quality evaluation.

Quality Assessment: The dissertation of Patrick Wolf [Wolf 22] explains the details of quality assessment group. In summary, quality assessment is a multifaceted and integral part of data processing and analysis, ensuring the reliability and accuracy of data used in various applications. This process encompasses several key aspects and features, each crucial for maintaining data integrity and optimizing data fusion results. Firstly, Quality Degradation Assessment plays a pivotal role. It involves a meticulous examination of how data quality diminishes across different stages of a behavior network's processing chain. By pinpointing the stages where significant quality loss occurs, this assessment aids in identifying critical points that require rectification or improvement. This continuous monitoring is essential for maintaining the overall quality of the data as it undergoes various transformations or analyses throughout the processing chain.

By pointing to the camera mounted on the autobus, the skeleton point of the camera's axis is converted from the vehicle's kinetic center, as described in the sensor configuration section. This alignment with the vehicle's kinetic center ensures that the skeleton coordinates match other sensor data and facilitates accurate positioning on the map. The stereo camera is set to a fixed 20° angle and captures data that changes at the tracking points when the angle of the camera is changed from the original position. To address these changes, the data transformation is performed by rotating the camera at 15 and 25° angles and translating it along the x and z axes between -1m and 1m in relation to the data of each joint. Given that the camera is stationary and only rotates on the x-axis, data transformations include the adjustment of the x-axis rotation matrix and translations on the x and z-axis.

Pedestrian Activity: The navigation architecture, as elaborated in Chapter 4, is predicated on the integration of iB2C modules (Description 3.5). It employs a geometrical method to ascertain the postures of pedestrians within the iB2C framework. The efficacy and methodology of this approach have been substantiated in our publication [Jan 22b], wherein the same dataset of 3D skeleton joint points were utilized to calculate various angles between different body parts. Furthering this research, a comprehensive comparative study was conducted by us and detailed in an extended paper [Jan 24]. Moreover, LSTM models are also used for activity recognition. This thesis plays together with the LSTM model and the geometrical method within the iB2C network, offering an in-depth analysis of both approaches. This dual-approach strategy exemplifies the integration of machine learning and geometric analysis in enhancing the robustness and reliability of pedestrian behavior prediction in autonomous navigation architectures. The 3D joint points of the pedestrians, taken from "Quality Assessment," are processed here for activity recognition.

Description 3.5: Integrated Behavior-Based Control (iB2C)



iB2C architecture [Armbrust 11] is a comprehensive system designed for complex behavior-based applications. It leverages the principles of behavior-based systems such as modularity, parallelism, multi-goal orientation, and redundancy, along with intelligent data fusion and behavior interaction, influenced by Brooks Subsumption architecture yet differing in crucial aspects. iB2C avoids strict state-switching and allows multiple system nodes to operate simultaneously, setting it apart from traditional behavior trees and FSMs. It facilitates enhanced system design and better behavior arbitration by incorporating uncertainties into its framework. The iB2C framework provides a common interface with meta signals for uniform behavior interaction, featuring four main types of behavior inputs and outputs: stimulation (s), inhibition (i), activity (a), and target rating (r). Stimulation is used to enhance the significance of a behavior within the network, whereas inhibition serves to constrain it. The values for both stimulation and inhibition are confined within the range of [0, 1]. The internal state of a behavior, defined as the potential ϕ, emerges from the dynamic balance between stimulation and inhibition.

$$\phi = \min(s, (i - 1))$$

Classification: Interaction fields, as discussed in Chapter 5, are developed within the Classification group. This classification leverages data from the "Pedestrian Activity" group to represent various attributes of pedestrians in a comparable fashion.

Mapping: Classification data is integrated into the mapping system to identify navigable zones for the Autobus. For navigation purposes, tentacles are utilized. The weighting of these tentacles is determined by the values assigned to each cell within the map. Considering that the current system employs a multi-feature map, various sets of tentacles are produced. Specifically for this study, a new category of tentacles, named pedestrian tentacles, is introduced to facilitate subsequent decision-making processes.

Pilot: The integration of the pedestrian-aware tentacles into the overall navigation system is a crucial aspect to decide on what action should be taken by the vehicle. It is incorporated into the Pilot group, a decision-making module within the navigation framework. This module is responsible for adjudicating between the standard feature-based tentacles, which primarily focuses on path segmentation, and the newly introduced pedestrian-focused tentacles. The decision-making process within the Pilot group is thus enriched, enabling a more nuanced and context-sensitive navigation strategy. This strategy takes into account not only the physical aspects of the environment but also the dynamic and potentially unpredictable elements introduced by pedestrian movements. Detail of decision making concept is given in Chapter 6. Once the preferred tentacle is chosen, the decision is made either to control the vehicle or/and interact with the pedestrians.

Interaction Modules: A detailed design implementation of interaction modules is presented in Chapter 7, which constitutes the fourth component of the smart interaction strategy for this thesis. The content of the interacting modules, including LED displays, audio, and lights, is controlled within this section.

The next four chapters provide an in-depth analysis of the four essential components integral to the interaction concept of this thesis, as illustrated in Figure 2.4.

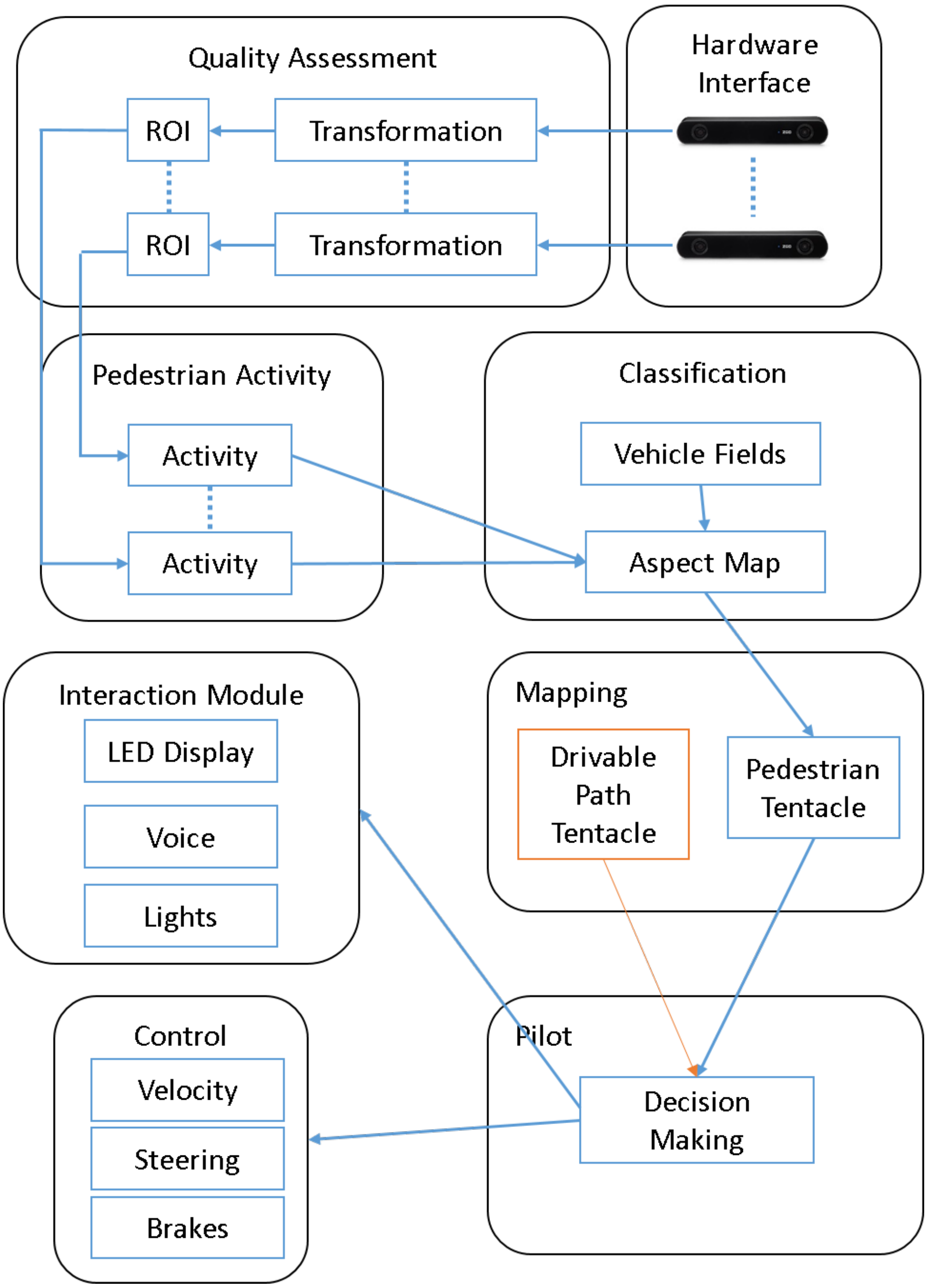
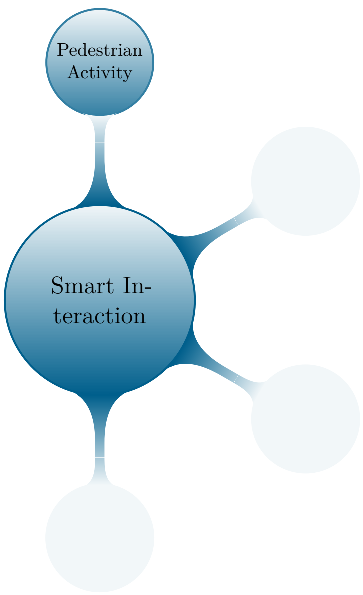


Figure 3.20: This diagram provides a comprehensive view of the interrelated components of the REACTiON framework for autonomous vehicle navigation. Quality Assessment is initiated through Regions of Interest (ROIs) that undergo Transformation processes, interfacing with the Hardware Interface, which receives sensory input data. The Classification module uses this data to generate Vehicle Fields and an Aspect Map. Simultaneously, the Pedestrian Activity modules monitor and evaluate pedestrian movements, labeled as "Activity." The Interaction Module, comprising LED Display, Voice, and Lights, actively communicates with pedestrians. Mapping components synthesize Drivable Path Tentacle and Pedestrian Tentacle data to inform the Pilot module. The Pilot's Decision Making component integrates these inputs to control the vehicle effectively, ensuring adaptive responses to real-time environmental and pedestrian activity. The flow of information between the systems underlines the complexity and integration required for successful autonomous navigation and interaction with pedestrians in dynamic environments.

4. Pedestrian Activity



In the field of driverless minibus technology, it is especially difficult to determine what pedestrian activities are in busy areas. For this reason vehicular intelligence is extremely important in order to ensure that these minibuses work safely and efficiently in places where many people walk around. Thanks to recent advances in sensors, machine learning and data processing, these systems have actually increased their ability to detect, understand and predict what pedestrian intentions.

The recognition of pedestrian activities refers to a systematic process of identifying and classifying the movements and behaviors of persons walking in a specific environment. This interdisciplinary field combines elements of computer vision, artificial intelligence, urban planning and human movement to analyze patterns and characteristics of pedestrian movement. The goal is to better understand how pedestrians navigate and interact with the environment, which can be used to design safer, more efficient and user-friendly urban spaces. The main and fundamental element of the smart interaction posited in this thesis (Chapter 2) is *"Pedestrian Activity"*. To enhance interaction, it is important to know the intentions of pedestrians. This section begins with a discussion of the current state of detection, focusing on the models associated with approach used in this thesis. Subsequently, a critical analysis of pedestrian activity recognition is provided and the methods used to calculate pedestrian activity based on the detection model used is explored in this thesis.

4.1 Pedestrian Detection

The process of pedestrian activity recognition starts with detecting a person. Mostly researchers have been using RGB (Red, Green, Blue) imagery as a cornerstone, primarily due to its capacity to provide a rich, multi-dimensional view of the visual world. RGB images, capturing the full spectrum of colors perceivable by the human eye, offer a comprehensive and nuanced perspective that is crucial for accurate pedestrian detection. This color depth enables algorithms to distinguish pedestrian with greater precision, as subtle variations in hue, saturation, and brightness can be critical differentiators. Furthermore, the widespread availability and compatibility of RGB data make it a universally accessible resource for researchers and developers. This accessibility ensures that advancements in pedestrian detection technology can be rapidly disseminated and implemented across various applications, specifically for driverless minibuses. Additionally, the inherent richness of RGB images facilitates more advanced processing techniques, such as convolutional neural networks (CNNs), which can extract and learn from the complex patterns present in real-world scenarios. Thus, the use of RGB images in object detection not only enhances accuracy but also fosters innovation and applicability in diverse and evolving technological landscapes. For this work, a simple and fast model is used to recognize activities.

The quality of pedestrian recognition achieved by this method is the key to vehicle safety. The model needs to interpret the input data accurately to quickly and accurately identify what pedestrians are doing, which is essential for real-time self-driving. The system detects subtle differences in the movement and appearance of pedestrians and greatly improves the reliability of detection. This reliability is very important because it directly affects the ability of the vehicle to make intelligent decisions and navigate safely through the busy pedestrian area.

4.1.1 State-of-the-art Pedestrian Detection Techniques

Pedestrian detection is a specialized subset of a wider field of object detection. Object detection models are designed to recognize many types of objects, ranging from everyday objects such as chairs and cars to more specific entities such as faces and animals. Pedestrian detection limits the scope of object detection and focuses exclusively on the identification of humans within images and video frames. This concentration is crucial for applications that are particularly interesting in the field of human presence, movement and behavior, such as surveillance systems, autonomous vehicles and pedestrian traffic analysis. Although the task focuses on a single category of objects (i.e. pedestrians), it is a unique challenge, including the interaction between various poses, occlusions, different clothing, and the environment that can have a significant impact on appearance.

State-of-the-art pedestrian detection techniques have come a long way with advancements in deep learning and computer vision. Convolutional Neural Networks (CNNs)[Albawi 17] are now the backbone of many pedestrian detection systems, offering big improvements over older methods that relied on hand-crafted features. Models like YOLO (You Only Look Once)[Redmon 16] and Faster R-CNN [Girshick 15] have been adapted to achieve high accuracy by learning complex features from images of pedestrians. Multi-Scale Detection [Liu 16a] tackles the issue of varying pedestrian sizes and scales in images by applying detectors at different scales or using networks designed to capture features at multiple resolutions. The SSD (Single Shot MultiBox Detector) [Liu 16b] framework is a good example, as it detects objects of different sizes using feature maps from multiple layers. Attention Mechanisms integrate with CNNs to focus on the most relevant features of the input image for pedestrian detection, enhancing performance in crowded or complex scenes. CBAM (Convolutional Block Attention Module) [Woo 18] is a notable example, selectively emphasizing important features in both spatial and channel dimensions. Part-based Models[Girshick 14] break down the human body into parts and detect pedestrians by recognizing and assembling these parts, which is particularly useful when the view is obstructed or the pedestrian is partially visible. The Part-based R-CNN (P-R CNN) [Li 18] extends the R-CNN framework to detect and assemble these parts for full pedestrian detection. Using 3D and Depth Information, such as depth sensors and 3D modeling techniques, improves pedestrian detection when 2D information isn't enough. Methods like 3D Convolutional Neural Networks (3D CNNs) [Qi 17] and LiDAR-based detection systems offer enhanced spatial understanding and robustness to occlusions. Adversarial Training employs Generative Adversarial Networks (GANs) [Creswell 18] to create training datasets or improve the robustness of detection models against tough scenarios, generating realistic examples that cover a wide range of appearances and occlusion patterns. Transfer Learning and Domain Adaptation [Tzeng 14] leverage pre-trained models on large datasets to improve pedestrian detection in specific or challenging domains, like night-time scenes or bad weather. These methods adapt models to new domains with limited labeled data. The models mentioned here form the foundation for advanced methods in pedestrian detection.

4.1.2 Skeleton Detection Techniques for Precise Activity Recognition

In the context of Autobus, particularly within pedestrian-rich environments, the implementation of an optimal detection system for precise activity recognition is paramount to ensure both safety and better interaction. In the array of available techniques, skeleton detection has distinguished itself as a notably effective method for the recognition of pedestrian activities. Advantages 4.1 highlights the advantages of perceiving a skeleton for activity recognition.

Advantages 4.1: Skeleton Points for Activity Recognition

Superior Method for Activity Recognition Skeleton detection is highlighted as a superior technique for accurately recognizing and predicting pedestrian behavior. This is due to its focus on the skeletal structure of individuals, which allows for sophisticated analysis of human postures and movements within complex urban landscapes.

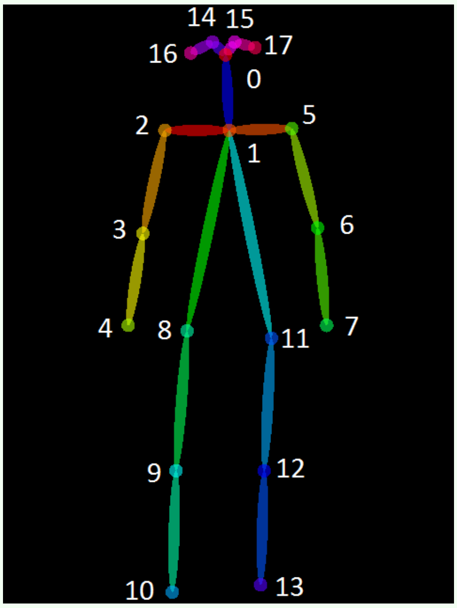
Reduced Computational Load A significant advantage of skeleton detection is its efficiency in processing. By focusing on skeletal joint points, the method requires processing fewer data points compared to other techniques, which reduces the computational load while maintaining high accuracy, especially at peak performance levels.

Granular Examination of Activities Skeleton detection enables a more detailed examination of pedestrian activities by analyzing the angles between various joint points. This allows for a comprehensive understanding of different human postures and movements, providing a nuanced view of pedestrian behavior that is critical for accurate activity recognition.

Enhanced Decision-Making Capabilities The precision with which skeleton detection captures the nuances of pedestrian activities significantly enhances the decision-making capabilities of Autobus. By accurately interpreting pedestrian behavior, Autobus can make more responsive and informed actions, improving its ability to navigate safely in environments populated by pedestrians.

Consequently, skeleton detection stands as a pivotal innovation in the evolution of safe and reliable urban mobility solutions, underscoring its importance in the ongoing development of autonomous transportation technologies. Different body skeleton joints are available. The model used for this thesis is shown in Example 4.1. In Example 4.1, a concise description is provided for each of the numbered points in the 18-point joint skeleton.

Example 4.1: 18-Point Joint Skeleton Model



0. Nose joint.
1. Neck: Serves as a central connection between the head and the torso.
2. Shoulder (R): Right shoulder joint.
3. Elbow (R): Right elbow joint.
4. Wrist (R): Right wrist joint.
5. Shoulder (L): Left shoulder joint.
6. Elbow (L): Left elbow joint.
7. Wrist (L): Left wrist joint.
8. Hip (R): Right hip joint.
9. Knee (R): Right knee joint.
10. Ankle (R): Right ankle joint.
11. Hip (L): Left hip joint.
12. Knee (L): Left knee joint.
13. Ankle (L): Left ankle joint.
14. Eye (R): Right eye.
15. Eye (L): Left eye.
16. Ear (R): Right ear.
17. Ear (L): Left ear.

4.1.3 State-of-the-art Skeleton Detectors

AlphaPose [Fang 22] focuses on improving the accuracy of crowd scenes. It uses a special framework called regional multi-person Pose Estimation (RMPE) along with some other techniques, such as symmetrical spatial transformation networks (SSTNs), parametrical non-maximum compression of partial poses (NMSs), and proposal generators guided by partial poses (PGPGs). It employs a top-down approach, it first detects an individual and then identifies their posture. This two-step process manages complex interactions well, but can be slow less capable devices. DeepCut [Pishchulin 16] uses another strategy. It is a partial-based model that combines CNN and integer linear programming to effectively group the parts of a body. This model does not require separate human detection steps and sets high accuracy standards. However, it requires a lot of computational power because it evaluates many pairwise terms during the pose assembly. DeeperCut [Insafutdinov 16] improves DeepCut by using deeper ResNet architectures to better identify parts and optimize parts group algorithms. This allows even better accuracy, but requires significant computational resources. DensePose [Güler 18] offers a unique approach for mapping all pixels of the human body to a 3D body surface in RGB images. This gives a detailed view of body position and movement, and is ideal for high-fidelity applications. However, it is quite demanding in terms of computation and data. EfficientPose [Eweiwi 15] aims to balance accuracy and efficiency. Using the scalable architecture of EfficientNet, it performs well in real-time applications on various devices. However, this balance sometimes means trading some precision for speed. Finally, SimpleBaseline [Zhu 23] is about keeping detection simple. It uses ResNet and deconvolution layer basic architectures to make it efficient and accurate. Its simplicity improves performance and computational efficiency and makes it a good benchmark even if it is not the most advanced model.

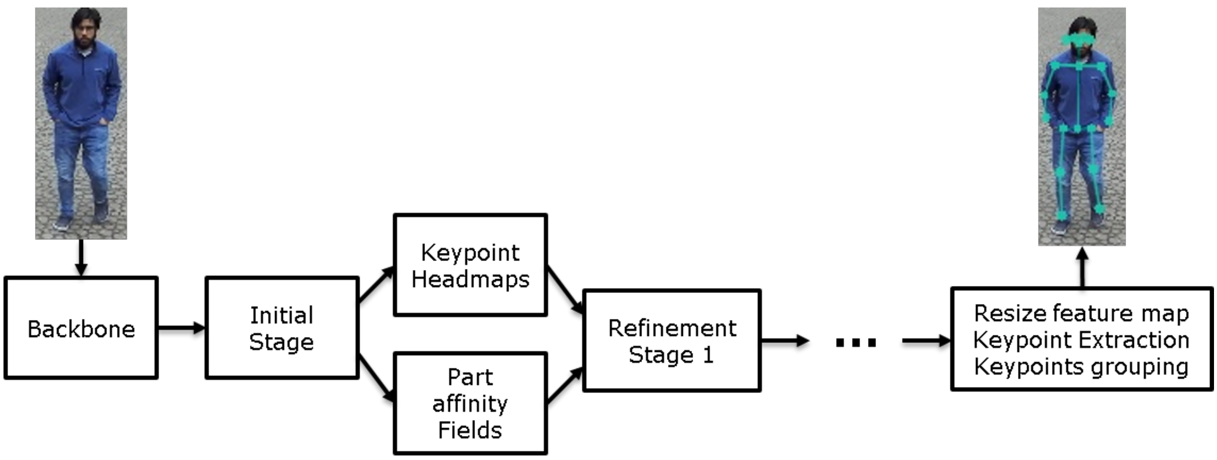


Figure 4.1: OpenPose pipeline.

OpenPose [Osokin 18] provides multi-person 2D real-time positioning detection and uses CNN, which is essential to understanding the positions of the human body. It leverages a non-parametric representation known as the Part Affinity Field (PAF) to associate parts of the body with individuals in the image. The system works from the bottom up, first detecting the parts of the body (e.g., elbows, wrists, etc.) and then combining them into a full pose for each person. Network architecture starts with the extraction of features, followed by the initial estimation of the heat maps of key points and part affinity fields (PAFs). Subsequently, the system undergoes five improvement steps, enabling the detection of 18 different keypoint types. Then, a grouping algorithm determines the best pairing of each keypoint based on affinity, selecting from 19 predefined keypoint pairs (e.g., left elbow-left wrist, right hip-right knee, left eye-left ear). The process is shown in Figure 4.1. Inference is that the input image is resized to fit the input dimensions of the network, the width is adjusted to maintain the aspect ratio, and the padding is multiplied by 8.

Model	Description	Approach	Computational Demand	Accuracy
AlphaPose	Utilizes RMPE with SSTN, Parametric Pose NMS, and PGPG. Targets crowded scenes.	Top-down	High for real-time on less capable devices	74.2% mAP
DeepCut	Combines CNNs with integer linear programming for body part grouping without separate human detection.	Part-based	Substantial	59.8% mAP
DeeperCut	Builds on DeepCut with deeper ResNet architectures and optimized part grouping algorithms.	Part-based	Very substantial	61.3% mAP
DensePose	Maps every pixel to a 3D surface, providing detailed body position and movement analysis.	Dense mapping	High and data-intensive	57.0% mAP
EfficientPose	Uses EfficientNet architecture to balance accuracy and efficiency, favoring real-time applications.	Scalable architecture	Moderate, efficient for diverse environments	73.8% mAP
SimpleBaseline	Employs a simple architecture with ResNet and deconvolutional layers. Ideal for benchmarking due to its efficiency and accuracy.	Straight forward architecture	Lower, efficient	71.4% mAP
OpenPose	Real-time multi-person system to jointly detect human body, hand, facial, and foot keypoints.	Multi-person	High, real-time capable	65.3% mAP

Table 4.1: Comparison of Pose Estimation Models

Implementing OpenPose for Enhanced Real-Time Multi-Person Pose Estimation

Table 4.1 summarizes different pose estimation models using skeletons. We discovered that, from our experiments with various modern skeleton detectors and from the observation of computational requirements, OpenPose is an unparalleled tool, especially when it comes to 2D multi-person real-time pose detection in Autobus. What distinguishes OpenPose from us is its bottom-up approach. This method contrasts with the top-down method of first detecting individuals and then estimating their pose. The OpenPose architecture is modular and can be adjusted to various parts of the body including the face and hands. Since detection is performed in small integrated PCs (Jetson AGX), it is important to have a model with real-time capabilities. OpenPose is optimized for speed and accuracy, making it suitable for real-time applications, and has set the standard for the estimation of poses. This ability is crucial to this thesis because it allows to deal with scenes with various numbers of individuals without having to obtain previous knowledge of them. This method is a significant step forward from the top-down methods, especially in busy scenes, and is often slow and scalable.

Moreover, OpenPose's ability to detect up to 135 keypoints, covering body, foot, hand, and facial landmarks, gives a level of detail in human posture and movement analysis that is unmatched by most other models, which typically focus on fewer body keypoints. This comprehensive detection is vital in the field of this thesis's application, such as vehicle-pedestrian interaction, where understanding human movement is essential. Another aspect where OpenPose shines, and which is particularly beneficial for this thesis, is its real-time performance. Even in scenarios involving multiple individuals, it maintains its efficiency, making it ideal for interactive systems and live monitoring applications. While the challenges are acknowledged, including the need for high-powered hardware for optimal performance and some limitations in handling occlusions, the balance of speed, accuracy, and the level of detail that OpenPose offers makes it a go-to tool for this thesis in scenarios where rapid and comprehensive pose estimation is key to success. Hence for this thesis, OpenPose's 18-point joint skeleton model is used for pedestrian activity recognition as shown in Example 4.1.

4.2 Pedestrian Activity Recognition

Following the critical phase of skeleton detection, we now move to the complex domain of activity recognition, an essential component to fully understand the dynamics of a scene. Based on advanced object detection techniques, such as OpenPose's human pose estimation technology, activity recognition aims to interpret and classify actions and interactions in the captured data. This section gives detailed information about detecting and analyzing patterns of movement and behaviour and converting raw data into meaningful insights about the activities carried out. Through advanced algorithms and machine learning techniques, activity recognition extends beyond mere detection, providing a deeper layer of interpretation, which is essential for applications ranging from monitoring and interaction between humans and computers to health monitoring - particularly driverless minibus navigation. As mentioned earlier, the human skeleton joint point is used to predict different activities around the Autobus in shared spaces to recognize activities. In this thesis, a personalized data set is created to accommodate specific inputs and activities. This section explores how the subtleties of human gestures and movements are decoded and lay the foundations for intelligent interactions that can not only see, but also understand and respond to the complexity of human activity.

There exist multiple state-of-the-art works for activity recognition using skeletons from images. Various deep neural networks are used, such as High Resolution-Net [Sun 19], through heat maps, and ClueNet [Kishore 19], which is an unsupervised method for estimating occluded pedestrian pose. Li et al. [Li 20] use three main components: AlphaPose, a sequence-to-sequence network, and a softmax classifier, with an improved accuracy of 11.6% on the JAAD dataset compared to prior methods. Finally, tracking using a Kalman filter is used in [Fang 19], followed by CNNs. This methodology proficiently forecasts pedestrian intentions through 2D poses and images. However, it's crucial to emphasize that the focal point of this thesis is on detecting 3D poses. To achieve this, 3D pose data needs to be reconstructed from 2D sources, necessitating the integration of perspectives from various viewpoints.

Enhancing Activity Recognition with 3D Skeleton Joint Points

The reason 3D skeleton points are so much better for active detection is because they capture deep spatial information and precise information. Unlike 2D joint points that track movement only in two dimensions (x and y axes), 3D points add the important z axis. This offers a comprehensive representation of exactly where things are in the real world space. This understanding—when analyzing movements and postures, having this third dimension- allows for more information. For activities such as stepping forward or reaching out, which involve a lot of depth, 3D data gives a much clearer picture. Assumptions or use complex calculations is not needed to know how far away pedestrian is, as with 2D data processing. This makes the whole process easier and more precise. Additionally, 3D data can distinguish between activities that may appear the same in 2D but are quite different in reality. For example, from a 2D perspective, one may see the same hand movement forward vs backward, but in 3D, the difference can be seen clearly. Thus, in the short term, 3D skeleton points will give richer and more detailed information to recognize activities. They can analyze human movements more accurately by capturing the whole picture of the real world space.

Utilizing 3D skeleton joint points in machine learning models for activity recognition presents significant advantages over using images, primarily due to the substantial reduction in input data size and complexity. Images, especially high-resolution ones, consist of a vast array of pixel data, encompassing color, brightness, and texture information, which results in a high-dimensional input space for the model. Processing such extensive data requires considerable computational resources and can lead to longer training times and increased complexity in model architecture. In contrast, 3D skeleton joint points represent movement and posture through a relatively small set of coordinates in three-dimensional space. Each joint is defined by just three values (x, y, and z coordinates), drastically reducing the amount of data fed into the model. This reduction simplifies the data preprocessing steps, decreases the computational load, and often leads to more efficient training of the machine learning models. Moreover, using 3D joint points focuses the model's attention directly on the human body's movement dynamics, eliminating potential distractions and noise present in full images, such as background details or varying lighting conditions. This focused approach not only improves the efficiency and speed of the learning process but also enhances the model's ability to generalize across different environments, making it more robust and accurate in real-world applications of activity recognition. Hence, this capability to accurately capture the full spectrum of human motion, including the intricacies of body language and interaction with surroundings, makes 3D joint points particularly effective for activity recognition systems that require a deep understanding of pedestrian behaviors.

4.2.1 State-of-the-art Skeleton Activity Recognition

The methodology introduced in [Sanchez-Caballero 22] is a completely convolutionary 3D neuronal network that uses 3DFCNN to automatically encode spatial-temporal patterns derived from depth sequences, eliminating the need for pre-processing. This study outlines a technique for real-time human observation using an unprocessed depth image sequence captured by RGB-D cameras. The model architecture integrates both 2D and 3D convolutionary layers, complements maximum and average pool strategies. The network is trained and evaluated on the NTU RGB+D dataset [Shahroudy 16], which includes 60 different human actions. A graph-convolution network (GCN) alternative approach named pose3D was presented in the study [Duan 22]. This skeleton-based activity recognition framework is designed based on the NTU RGB data set, and uses 2D skeletons to generate 2D heat maps. These heat maps are layered in the temporal dimension to form a volume of a 3D heat map. Next, the 3D CNN processes these 3D heat maps to recognize the actions. This method demonstrated superior performance than all of the latest approaches, proving that it is more robust, interoperable, and scalable than the GCN as shown in Table 4.2. Despite its success, the initial 2D to 3D conversion and its training on NTU RGB data, collected via Kinect, are noted limitations.

Dataset	**Previous state-of-the-art**	**Pose3CD**
FineGYM-99	25.2 (2D-Ske)	94.3
NTU60 (X-Sub)	91.5 (3D-Ske)	94.1
NTU60 (X-View)	96.6 (3D-Ske)	97.1
NTU120 (X-Sub)	86.9 (3D-Ske)	86.9
NTU120 (X-Set)	88.4 (3D-Ske)	90.3
Kinetics (OpenPose)	38.0 (2D-Ske)	38.0
Kinetics (Ours)	44.9 (2D-Ske)	47.7

Table 4.2: Pose3CD accuracy performance compared to other approaches [Duan 22].

Since this thesis uses pure 3D-skeleton joint points without the use of depth images, the methodology of this thesis is inspired from [Botache 19]. Botache et al. employed a smart device-based approach to detect early pedestrian movement transitions, delineated into four states: waiting, starting, moving, and stopping. Following the Human Activity Recognition (HAR) baseline methodology, the process encompasses: preprocessing of acceleration and gyroscope data, data segmentation and feature extraction, feature selection alongside dimensionality reduction, and classification via machine learning algorithms, specifically linear SVM and XGBoost. Testing with 11 individuals across 79 scenarios demonstrated the model's capability to preemptively predict movement transitions, achieving an F1-score of 85%. The HAR pipeline central to our methodology for detecting pedestrian movement is illustrated in Figure 4.2.

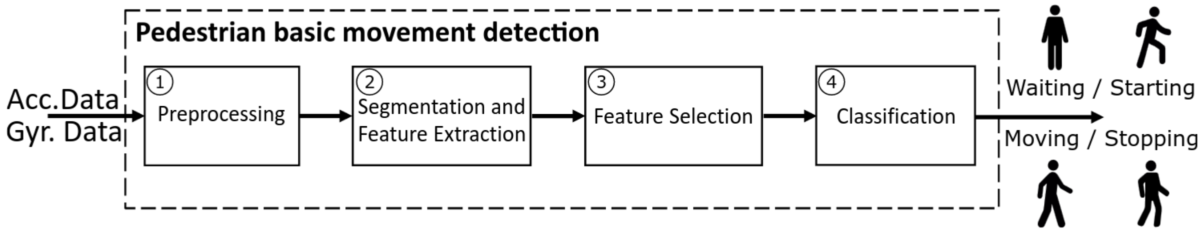


Figure 4.2: Smart device-based method for detecting fundamental pedestrian movements [Botache 19].

Replacing inertial data with 3D skeleton joint points as input for an LSTM model in activity recognition tasks presents a viable and potentially advantageous alternative. Both inertial data and 3D-joint points inherently share similar dimensional characteristics, typically operating in three-dimensional space (x, y, z coordinates), which ensures compatibility with LSTM models designed to process this kind of multidimensional data. Inertial sensors capture movements by recording accelerations and angular velocities, reflecting how different parts of a device (or, by extension, a body part) move in space. Similarly, 3D skeleton joint points track the spatial positions of various body joints over time, offering a direct visualization of bodily movements. The transition from inertial data to 3D joint points in an LSTM model is relatively seamless due to these shared characteristics. Both types of data provide a temporal sequence of movements, which is the kind of data LSTM models excel in processing. LSTMs are designed to understand and predict patterns in sequential data, making them equally adept at interpreting the sequences of movement data, whether they originate from inertial sensors or skeletal tracking.

Moreover, 3D joint points can potentially offer more precise and explicit information about body posture and movement compared to inertial data. While inertial sensors provide a general sense of movement and orientation, 3D skeletal data precisely locates each joint in space, allowing for a more detailed and nuanced analysis of human activities. This leads to more accurate activity recognition, as the model can directly learn from the specific positions and movements of body parts, rather than inferring them from acceleration or gyroscopic data. Hence, utilizing 3D skeleton joint points instead of inertial data in an LSTM model for activity recognition is not only feasible due to their similar dimensional properties but also enhances the model's accuracy and effectiveness. Both data types record movement sequences in three-dimensional space, aligning well with the LSTM's capabilities, and the transition from one data type to the other should be straightforward while potentially offering richer and more detailed insights into human motion. The LSTM model contains four gates shown in Figure 4.3:

- **Forget Gate (f)**: The initial phase in an LSTM network involves determining the portions of the cell state to be omitted from further calculations. This decision is facilitated by a sigmoid activation layer known as the forget gate layer. It evaluates the previous hidden state h_{t-1} and the current input x_t, producing a value within the range [0, 1] for each component of the cell state C_{t-1}. Here, a value close to 1 suggests retaining the information, whereas a value near 0 implies its exclusion.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{4.1}$$

- **Input Gate (i)**: Following the forget gate, the LSTM decides on the new information to be stored in the cell state. This process is two-fold: initially, a sigmoid layer—the input gate—identifies the values in the cell state to update. Subsequently, a tanh layer produces a vector $\tilde{C}_t$ of candidate values for addition to the cell state. These components are combined in the subsequent update phase.

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{4.2}$$

$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{4.3}$$

- **Input Modulation Gate (g)**: The transition to the updated cell state C_t from the previous state C_{t-1} is executed at this stage. Following the decisions made by the forget and input gates, the cell state is updated by multiplying the old state by f_t, thereby discarding the deemed unnecessary information. Afterwards, the product of i_t and $\tilde{C}_t$—the new candidate values adjusted by their respective update rates—is added.

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \tag{4.4}$$

- **Output Gate (o)**: The final step involves determining the output based on the current cell state. A sigmoid layer identifies the components of the cell state to be transmitted as output. The cell state is then processed through a tanh function, and the result is element-wise multiplied by the output of the sigmoid gate, ensuring that only the selected parts are outputted.

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \quad h_t = o_t \cdot \tanh(C_t) \tag{4.5}$$

4.3 Methodological Framework for 3D Skeleton-Based Activity Recognition

The process of activity recognition for this thesis is based on our previous work [Jan 22a, Jan 24]. The recognition process includes a stereo camera mounted on the roof of the vehicle to get images of the pedestrians. Initially, OpenPose is employed to detect and map the 2D skeleton points of a subject in a scene, effectively capturing their pose in two dimensions. Following this, these 2D points are meticulously overlaid onto the point cloud generated by the stereo camera, which provides the crucial third dimension: depth. This stereo camera, by capturing two slightly different views of the same scene, is able to calculate the distance of objects from the camera (depth information) by comparing the two images. When the 2D skeleton points are aligned with this depth data, each point is enriched with a depth value, effectively transforming it from a flat, two-dimensional coordinate to a three-dimensional one. This process involves careful calibration and alignment of the 2D pose data with the depth field to ensure accuracy. The resulting 3D skeleton points offer a more comprehensive understanding of the subject's spatial orientation and movement in

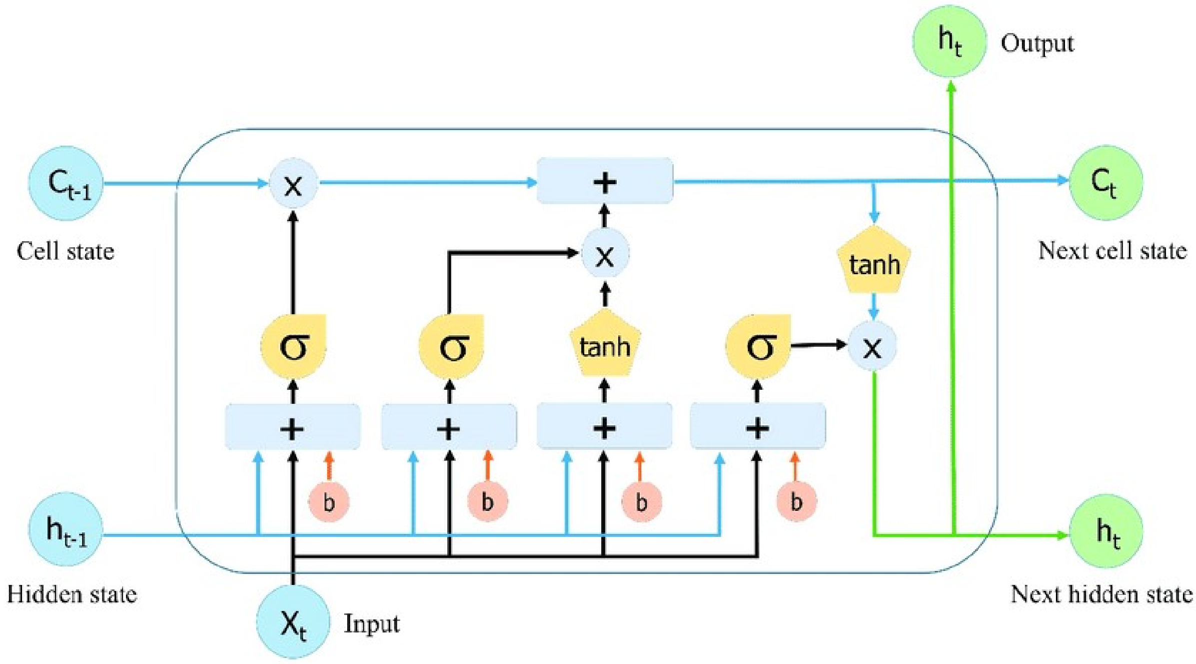


Figure 4.3: Diagram of a Long Short-Term Memory (LSTM) Unit. This figure illustrates the components and operations within an LSTM cell at a given time step t. It demonstrates the interplay between the cell's input X_t, the previous hidden state h_{t-1}, and the previous cell state C_{t-1}. Gates within the LSTM, namely the forget gate (in yellow, with a sigmoid activation σ), the input gate (in blue), and the output gate (in green), are crucial for information regulation within the cell. These gates control the cell state's update mechanism, represented by pointwise multiplication (denoted by $\times$) and addition (denoted by $+$). The activation function tanh is applied to the cell state and to the combined input and cell state to produce the current hidden state h_t. The updated cell state C_t, along with h_t, is then forwarded to the subsequent time step. Weights and biases associated with these gates, symbolized by 'b', are adaptively modified during the training process to refine the cell's information processing capabilities.

Label	Activity	Augmentation	
		Without	With
Secondary			
1	Calling	331	1324
2	Texting	310	1240
3	None	382	1528
4	Waving	94	376
Primary			
5	Parallel Crossing Towards	260	1040
6	parallel Crossing Away	208	832
7	Left Perpendicular Crossing	220	880
8	Right Perpendicular Crossing	219	876
9	Standing	210	840

Table 4.3: Labels for pedestrian activities with number of sequences for 18 keypoints data. The augmentation column shows the number of recorded sequences for every activity with and without augmentation.

the real world, which is pivotal for our dataset of understanding behaviors and directions. For this thesis LSTM models is used to predict pedestrian activities. To predict specific activities existing around the Autobus, a custom dataset is created and the occurring activities are explained within.

4.3.1 Custom dataset creation

In the development of training processes, custom dataset was constructed because the existing datasets did not allow to capture activities specific to pedestrian areas around the Autobus. Common datasets often differentiate between actions such as crossing and non-crossing, mainly in the context of zebra crossings. This limitation required the creation of a dataset containing scenarios involving Autobus in university campus environments, thereby reflecting more realistic pedestrian behaviour.

The behaviour of pedestrians is heavily influenced by the environment, especially in traffic. Observations show that pedestrians generally show a sense of security and therefore potentially dangerous behaviour at zebra crossings and intersections. This behaviour is remarkable in areas where there are no clear markings and traditional transport infrastructure.

To investigate these dynamics, the minibus was navigated through the campus of RPTU Kaiserslautern to observe and record pedestrian responses in proximity to the vehicle. During the observational study conducted around the campus, a variety of frequent and critical pedestrian activities were identified, including calling and texting. These activities were particularly noted while pedestrians crossed in various directions relative to the vehicle. The diverse nature of these movements is attributed to the campus environment, where pedestrians have the freedom to traverse the entire path area, as opposed to structured urban streets with defined sidewalks. Consequently, the activities identified in this context are detailed in Table 4.3.

Upon analysis, these activities were classified into two distinct categories: *"Primary"* and *"Secondary"* activities. This bifurcation was based on the nature of the activities and the specific body joints involved. Primary activities are dynamic and predominantly associated with the lower body joint points, specifically from the hip to the toes (joints 8-13). These activities typically involve motion and are essential for pedestrian mobility. In contrast, secondary activities are primarily related to the upper body and involve joint points above the hip, specifically joints 0-7 and 14-17. These activities often encompass actions such as holding objects or using a mobile device and are less about locomotion and more about interaction or engagement with objects or devices. Figure 4.4 displays a sequence of frames showing pedestrians engaged in various activities, which serve as training samples for the LSTM model. Each frame shows the 2D skeletal model that maps the posture and movement of the individual. The primary and secondary activities are concatenated to show full person activity. The selection of 46 frames per activity is predicated on identifying the minimum number of frames required to fully encapsulate the essence of each activity.

The dataset includes 240,000 records, each describing 18 body joint coordinates. These records were used both for the training and testing of the neural networks. Each instance is a time series sequence encoded in 46 frames and selected as the standard sequence length in the dataset. This length was determined to effectively capture the average walk, covering about 6-8 steps, representing a full walk cycle for an average stride. The dataset was carefully chosen to include participants showing a variety of postures and speeds during walking, with 23 frames that sufficiently cover the spread of almost all actions. The classes were designated by activity and focused on the posture and movements of the upper and lower body, respectively.

Calling and left perpendicular crossing | Texting and parallel crossing towards | Calling and parallel crossing towards | None and right perpendicular crossing | waving and standing

Figure 4.4: The figure displays a collection of representative images demonstrating various pedestrian activities, each tagged with corresponding actions for use in LSTM model training. From left to right: (1) An individual engaged in a phone call while crossing the street perpendicularly, which may indicate a potential distraction. (2) A pedestrian texting on a mobile device while crossing the street in a path parallel to the camera's viewpoint (3) A person making a phone call and crossing parallel to the camera, which is used to predict pedestrian intent and trajectory. (4) A person crossing the street perpendicularly without any visible engagement with a device, representing undistracted walking behavior. (5) An individual exhibiting a waving gesture while standing, which signify attempting to stop a vehicle or interacting with other pedestrians. These images are processed to include skeletal tracking overlays in fluorescent colors to facilitate the identification of limb positions and body movement patterns essential for the algorithm's learning process.

The rationale behind this categorization is to facilitate targeted analysis and classification. By segregating activities based on the joint points involved, it becomes possible to focus on specific groups of joints independently. This distinction is crucial in understanding and analyzing pedestrian behavior in a campus setting, where the range of activities and movements can be more varied and less predictable than in structured urban environments. This approach allows for a more nuanced and accurate classification of pedestrian activities, enhancing the understanding of pedestrian behavior in different contexts.

4.3.2 Segmented Upper and Lower Body Model Approach

The methodology presented in this study is structured into three sequential phases, as illustrated in Figure 4.5: Data Preprocessing, Modeling, and Post-Processing. Initially, raw 3D skeleton joint data are acquired and subsequently segregated into upper and lower body points during the data preprocessing phase. This segregation is imperative given the objective to discern the primary and secondary activities. Additionally, different mathematical operations are performed for better accuracy based on the lower and upper body. Consequently, two separate models are trained to identify primary and secondary activities, respectively. In the final phase, i.e., post-processing, the activities predicted

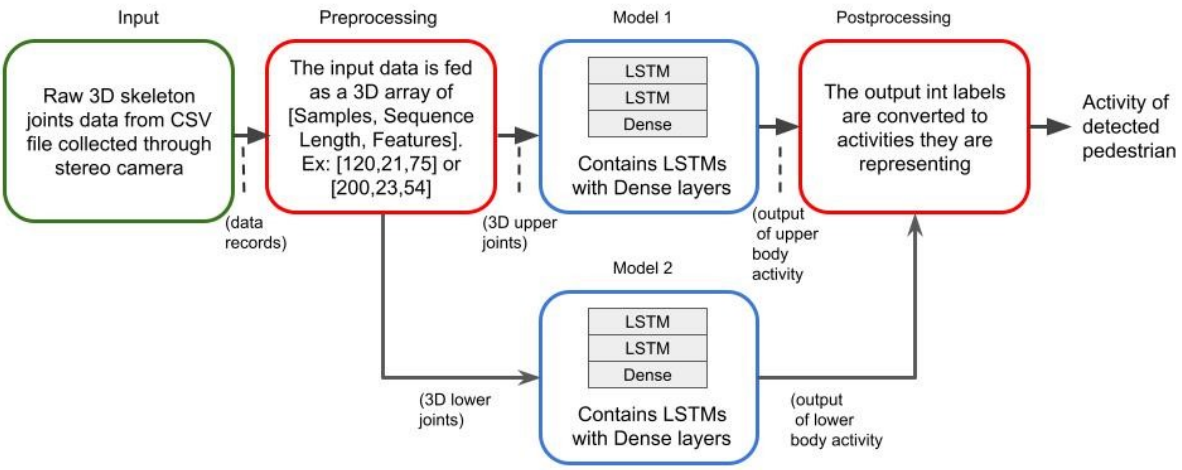


Figure 4.5: This schematic represents the methodology utilized in this thesis for the recognition of complex activities using 3D skeletal data. The initial stage involves the collection of raw 3D skeletal joint data from CSV files, obtained via stereo camera systems. During preprocessing, this data is transformed into a structured 3D array consisting of [Samples, Sequence Length, Features], which allows for the handling of both the spatial and temporal dimensions of the skeletal sequences. The data is bifurcated into upper and lower body joint indices to facilitate specialized processing by two separate LSTM-based neural network models—Model 1 for the upper body and Model 2 for the lower body. Each model employs LSTM layers followed by Dense layers to effectively capture and learn from the temporal dependencies inherent in the sequential data. The post-processing stage involves the conversion of the output labels from each model into descriptive activity labels, which are then synthesized to form a complete representation of the subject's activity. This approach allows for an understanding of primary and secondary activities, necessary for finalizing the results for activity recognition component of the concept.

by both models are combined to yield a comprehensive representation of the combined activity.

To feed the sequences to the model, the raw data is reshaped into a 3-dimensional array where it considers the number of samples, Sequence length of each skeleton, and number of features i.e., total joints (18x3=54). The lower joints are defined in indices 8, 9, 10, 11, 12, and 13 from the entire skeleton, while the rest represents upper body data in the framework. The frames within the window are fed into the LSTM sequentially, respecting the temporal order. The LSTM processes these sequences, extracting temporal features and patterns critical for understanding the activity. After processing one window, the window is "slid" forward by a certain number of frames (determined by the degree of overlap), and the process repeats. As the LSTM processes each window, it builds an internal state that captures the temporal context. This state is updated as the window slides across the sequence, allowing the LSTM to maintain a continuous understanding of the temporal dynamics.

This thesis explores the use of various optimization algorithms, including Stochastic Gradient Descent (SGD), Adaptive Moment Estimation (Adam), and Root Mean Square Propagation (RMSprop). The data labels are converted to one-hot encoded vectors prior to being inputted into the model. To mitigate the risk of overfitting during the training of expansive networks, regularization strategies such as Early Stopping and Dropout are employed. Overfitting occurs when a model, instead of generalizing, starts to memorize the noise in the training dataset, leading to a rise in generalization error and diminished predictive performance on unseen data. The objective is to calibrate the training duration to ensure the network comprehends the mapping from inputs to outputs without memorizing the training data. The model's performance is assessed on a validation set post each training epoch, and training ceases under the Early Stopping criterion if the model's validation performance degrades, such as an increase in loss or a decrease in accuracy.

In multi-class classification tasks, categorical cross-entropy is used as a loss function. For example, a sample can only fit into one of many possible categories, and the model must choose one. This approach uses a cross-entropy loss function, as the task is to classify different pedestrian activities. The categorical cross-entropy loss function computes a sample's loss by adding the following values:

$$Loss = -\sum_{i=1}^{\text{output}} y_i \cdot \log \hat{y}_i \tag{4.6}$$

where $\hat{y}_i$ is the i-th scalar value in the model output, y_i is the corresponding target value, and output size is the number of scalar values in the model output. It uses softmax as the activation function. After the model outputs the array of softmax probabilities, highest probability index is chosen as the predicted label.

4.4 Experiments & Results

The experimental methods of this thesis were adopted step by step, tailored to the specific nature of the methodology. The incremental approach is the key to tracking the progress of improving prediction accuracy. The thesis methodology here consists of two different models, aimed at the lower and upper joints, respectively. As a result, each model's evaluation is performed separately and detailed analysis of results for each body region is given.

Following this approach, upper body joint data are standardized at a uniform scale to ensure invariability for position and location. This standardization is particularly relevant for actions such as calls or text, which involve minimal movements. Thus, the upper joints are aligned within similar ranges, a process that is feasible because the secondary actions under consideration do not necessitate temporal sequencing. For standardization purposes, each data sequence of 23 units is subject to standard scale. This scale is calculated on the basis of the lowest and highest values of the entire sequence, meaning that each sequence with 23x54 data points is scaled on the basis of its minimum and maximum values. Equation 4.7 shows the normalization process. Here, x_i represents the individual data points of the sequence, min(x) represents the minimum value of 23 units in the sequence (assuming the total number of data points is 54), and max(x) represents the maximum value of the sequence. This formula effectively scales each data point from 0 to 1, compared to the minimum and maximum values found in each sequence, thereby standardizing the scale of the upper body data.

$$x_{i,\,\text{normalized}} = \frac{x_i - \min(x)}{\max(x) - \min(x)} \tag{4.7}$$

Contrastingly, lower-body joint data are not normalized. This decision stems from the consideration that normalization could potentially impact the performance of the network. The rationale is that the lower-body joint data need to accurately reflect changes in movement, which is crucial for identifying activities like parallel or perpendicular crossing. Hence, maintaining the original scale of the lower-body joint data is essential for capturing these dynamic movements effectively.

4.4.1 Upper Body

The upper body activity classification model undergoes training over 50 epochs (plateau reached) utilizing SGD, Adam, and RMSprop as optimizers, with corresponding learning rates set at 0.1, 0.01, and 0.001. Training is conducted on both normalized and non-normalized data. In the preprocessing phase, the data undergoes uniform scaling facilitated by the application of the `StandardScaler` transformation, a component of the Scikit-learn library.

Results before data augmentation: Table 4.5 presents the model's accuracy on test data when trained with normalization, while Table 4.4 shows the outcomes for the model trained without data normalization prior to augmentation.

Optimizers	**Learning Rates**		
	0.1	0.01	0.001
SGD	.45	.41	.29
Adam	.2804	.2805	**.46**
RMSprop	.3018	.32	.45
Adagrad	.416	.408	.422

Table 4.4: Accuracies of LSTM model for upper body data classification *without normalization*

Optimizers	Learning Rates		
	0.1	0.01	0.001
SGD	**.82**	.69	.309
Adam	.28	.57	**.82**
RMSprop	.80	.78	.35
Adagrad	.60	.672	.719

Table 4.5: Accuracies of LSTM model for upper body data classification *after normalization*

The data presented in the aforementioned tables unequivocally indicates an enhancement in the model's performance following normalization, as evidenced by the increase in accuracy from 40% to between 70-80%. Notably, at a learning rate (η) of 0.1 with the SGD optimizer, the model attained its peak accuracy at 82%. Conversely, employing Adam with the same learning rate resulted in the lowest performance relative to the other configurations.

Results after data augmentation: In the subsequent experiment, the previously described models were replicated utilizing analogous hyperparameters. However, this iteration incorporated a blend of augmented and original data, with the training extending over 40 to 50 epochs. Additionally, an alternative optimization algorithm, Adagrad [Défossez 20], was employed. The outcomes of this experiment are quantitatively detailed in the accompanying tables 4.6, specifically focusing on the models' accuracy when evaluated against test data.

Optimizers	Learning Rates			
	0.1	0.01	0.001	0.0001
SGD	.493	.483	.591	.613
Adam	.40	.34	.34	.27
RMSprop	.79	.809	.751	.78
Adagrad	.817	.80	.808	**.833**

Table 4.6: Accuracies of LSTM model for upper body data classification *without normalization after augmentation*

Optimizers	Learning Rates			
	0.1	0.01	0.001	0.0001
SGD	.4685	.57	.85	.85
Adam	.89	.33	.68	.43
RMSprop	**.9485**	.925	.88	.52
Adagrad	.9138	.89	.86	.89

Table 4.7: Accuracies of LSTM model for upper body data classification *after normalization after augmentation*

The model using normalized data for upper body joints consistently exceeds its counterpart trained in non-normalized data. As shown in Table 4.6, Adagrad optimizers with a learning rate of 0.0001 yield the highest accuracy among tested configurations. In contrast, other optimizers show average performance at different learning rates. Especially the generalization of the network in test data shows a significant improvement with normalized data, as shown in Table 4.7. The best results are obtained using RMSprop Optimizer with learning rates of 0.1 and 0.01, while SGD Optimizer with learning rates of 0.1 shows low performance. Remarkably, some optimizers at certain learning rates have achieved more than 80% accuracy, a significant advancement over the results observed in previous analyses.

4.4.2 Lower Body

In the final phase of the study, the LSTM model, designed for the classification of lower body activity in pedestrians, was trained under two conditions: with and without data augmentation. The plateau was reached at 75 epochs for the training. Impressively, in at least one of the tested configurations, the model's accuracy on the test data surpassed the 90% threshold. The model underwent experimental training using a variety of optimizers, including SGD, Adam, RMSprop, and Adagrad, each tested with different learning rates of 0.1, 0.01, 0.001, and 0.0001.

Results before data augmentation: Table 4.8 presents results indicating that the Adagrad and Adam optimizers, when set to a learning rate of 0.1, achieve the highest accuracies, registering at 98% and 96% respectively. The performance of other optimizers varies, ranging from average to poor across the spectrum of tested learning rates.

Optimizers	**Learning Rates**			
	0.1	0.01	0.001	0.0001
SGD	.58	.65	.44	.76
Adam	.96	.28	.26	.20
RMSprop	.89	.23	.28	.20
Adagrad	**.98**	.47	.20	.93

Table 4.8: Accuracies of LSTM model for lower body data classification *before augmentation*

Results after data augmentation: Notably, as shown in Table 4.9, the network's performance has experienced significant enhancement following data augmentation. In terms of optimization, RMSprop demonstrates superior results across all η, while Adam shows the least favorable outcomes, particularly at learning rates of 0.001 and 0.0001. It is noteworthy that several optimizers have managed to surpass the 90% accuracy mark at specific learning rates, marking a substantial improvement.

Optimizers	**Learning Rates**			
	0.1	0.01	0.001	0.0001
SGD	.974	.97	.965	.964
Adam	.94	.96	.19	.21
RMSprop	.98	.986	.984	**.99**
Adagrad	.91	.727	.986	.98

Table 4.9: Accuracies of LSTM model for lower body data classification *after augmentation*

4.4.3 Real-world Experiments

The practical applicability of the methodologies introduced in this chapter has been evaluated through a series of experiments. These experiments were designed to test the methods under various scenarios and with real-time, previously unseen data. For this thesis, the test videos were captured using a stereo camera mounted on the autobus operating at the RPTU Kaiserslautern campus. Autobus served as a mobile recording unit, capturing videos across various locations and at different times.

In the context of this study, LSTM networks exhibited marginally superior performance compared to Gated Recurrent Units (GRUs) across a range of hyperparameters. Therefore, LSTMs were the chosen architecture for classifying pedestrian activities. The focus of this analysis was on the separate joints of the detected pedestrian skeletons, which are divided into upper and lower body parts. The activity sequence for each pedestrian comprised 23 frames, aligning with the training sequence length of the model. During preprocessing, data passed through the SequenceConversion module, and thereafter, it was fed into the LSTM model via the TensorflowInterface module for the prediction of pedestrian activities.

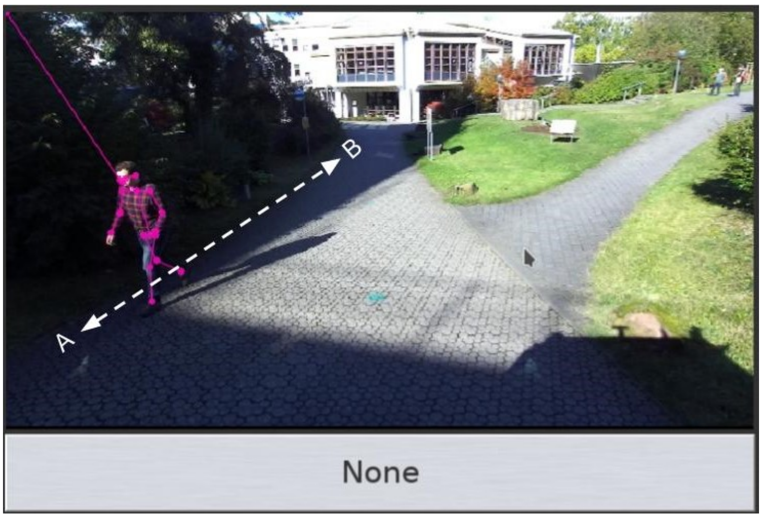
4.4.3.1 Upper Body Pedestrian Activity Classification

The frozen model, designed for recognizing upper body activities, employs LSTM layers and utilizes an RMSprop optimizer with a learning rate of 0.1. This model effectively classifies upper body activities, notably identifying actions such as "None" and "Calling," as demonstrated in Figure 4.6 (a) and (b), respectively. These classifications were observed while monitoring a pedestrian moving between point A and B on the campus. The corresponding videos were recorded at two distinct campus locations, capturing random pedestrians near the vehicle.

However, the model also produced some false predictions, which might have been accurate based on the joint data. A notable source of error is the misclassification of actions involving arm movements, such as holding a cup or bag, as "Calling" or "Texting." This misclassification arises because these activities share similar shoulder, arm, and head pose data, with joint coordinates closely aligned. Additionally, instances where phones were used in speaker mode by other pedestrians were incorrectly categorized as "Texting." Furthermore, discrepancies in the labeling of the custom data, where holding a cup or bag was sometimes tagged as "None" or "Texting," likely contributed to the network's confusion.

It's important to note that the model bases its classifications solely on the 3D joint data of the upper body, without considering the objects held by individuals. This limitation leads to scenarios where pedestrians holding objects, yet facing straight ahead, are classified as "None," indicating no upper body activity. In such instances, even if the real-time activity class is incorrect, the model's prediction, based on joint pose data, could be considered accurate.

These factors, alongside general inaccuracies, significantly influence the model's performance. Two examples of such classifications are illustrated in Figures a and b of Figure 4.7.



(a) A pedestrian detected with no upper body activity walking from B to A.

(b) A pedestrian detected while calling, walking from A to B.

Figure 4.6: Correct Classification of upper body activity of a pedestrian and walking.

(a) A pedestrian detected as Calling while holding a paper.

(b) A pedestrian detected as Texting who actually had no upper body activity.

Figure 4.7: False positive classification of upper body activity of a pedestrian moving from distance A to B.

4.4.3.2 Lower Body Pedestrian Activity Classification

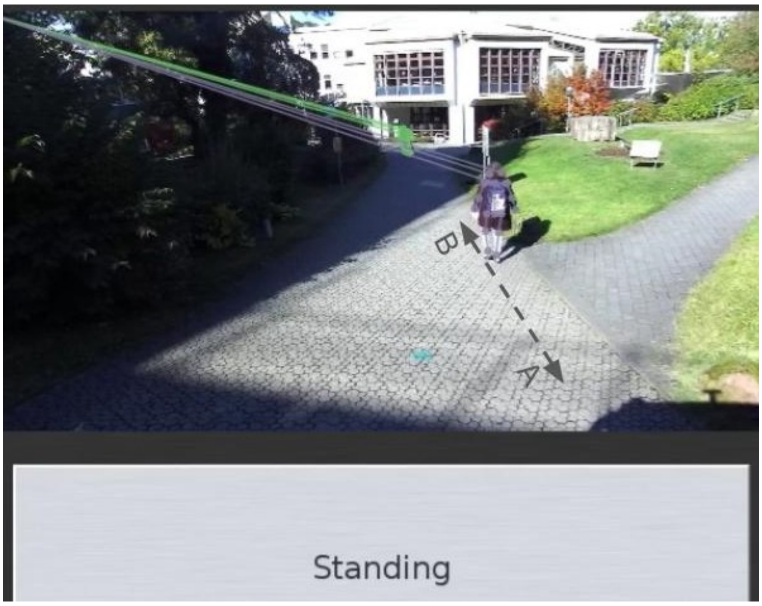
A modified version of the previously described model, employing the same hyperparameters but trained for fewer epochs, was applied to lower joint data without normalization. This model achieved a successful classification rate of upper body actions in 90% of cases. The approach of focusing exclusively on the Hip, Knee, and Ankle joints for identifying primary activities such as Parallel Crossing (Away and Towards), Perpendicular Crossing (Left and Right), and Standing has proven to be effective. This was corroborated by the accurate results obtained on test videos, as depicted in the visualization and illustrated in the corresponding figure.

However, a few instances of misclassification were observed, as shown in Figures 4.8 and 4.9. These were less frequent compared to those in the upper body model. Potential sources of error include the absence of certain frames or missing joint data. Additionally, a lag in the input data coupled with minimal movement could erroneously result in a classification of "standing."

(a) A pedestrian detected with activity Right Perpendicular Crossing.

(b) A pedestrian detected with Parallel Crossing Away from the vehicle activity.

Figure 4.8: True positives classification of lower body activity of a pedestrian moving from distance A to B



(a) A pedestrian detected as Standing instead of Parallel Crossing Away from the vehicle.

(b) A pedestrian detected as Parallel Crossing Towards who actually is Perpendicular Crossing.

Figure 4.9: False negative classification of lower body activity of a pedestrian moving from distance A to B.

4.4.3.3 Upper and Lower body classification simultaneously

In a subsequent experiment, a different approach was employed where models for upper and lower body activity classification were operated concurrently, as opposed to the separate execution observed in the previous scenarios. This setup involved the simultaneous implementation of two distinct frozen models, each tailored to classify activities related to either the upper or lower joints, on a single detected pedestrian. The test videos were identical to those used in the earlier scenario, with the addition of a new video captured while the bus was slowly traversing the campus.

For this experiment, the complete skeleton data of each pedestrian was bifurcated into upper and lower joint datasets. These datasets were then fed into their respective classification models. The predicted outcomes were visualized, where each identified activity was represented by changing labels. These labels shifted following each activity sequence or in response to variations in pedestrian activity. It's important to note that the frozen models utilized in this experiment are identical to those described in the previously mentioned scenario.

The models demonstrated a high degree of accuracy in recognizing various pedestrian activities on unseen data. As evidenced in Figure 4.10, activities of pedestrians moving from point A to B were successfully identified in terms of both upper and lower body classifications, including scenarios (a) and (b), and even in data captured while the vehicle was in motion (c). However, despite their enhanced performance on new skeleton data, the models also exhibited misclassifications for both upper and lower body activities, as depicted in Figure 4.11.

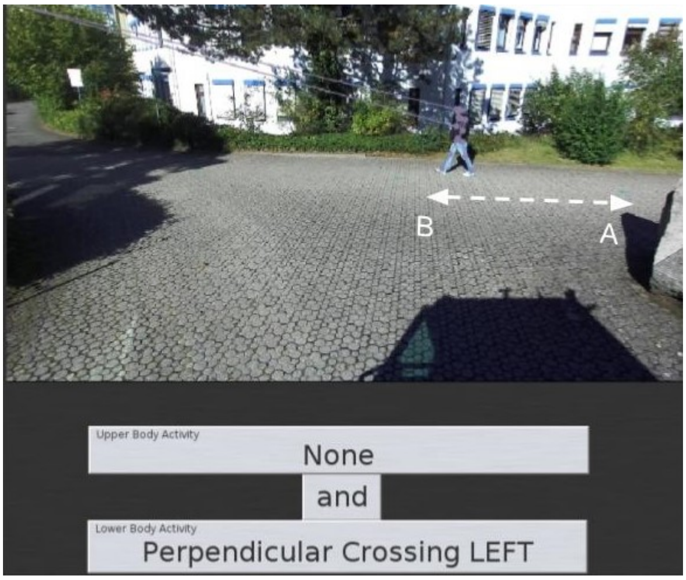
A common mistake in the classification of upper body activities is that pedestrians activity is misinterpreted for movement of their arms holding bags, cups or other objects as "Calling" or "Text". This is due to the similarity in shoulder and hand joint positions when holding objects and participating in these activities. Another factor contributing to the wrong classification of upper and lower body activities is the relative directions of vehicles and pedestrians. In particular, as shown in Figure4.11 (a), the movement or position of the car in combination with the movement of pedestrians on an undefined path has a significant impact on the classification accuracy. The complexity of the scenario is that the car is not on a normal straight road or open terrain, which is an additional challenge. For example, pedestrians moving towards vehicles may normally indicate "Parallel Crossing Towards", but given the distance and orientation of the path in relation to the stereo camera, it may also be interpreted as "Perpendicular Crossing Left", complicating the model predictions.

In real-world environments, such challenges are often encountered, with factors such as unclear paths or the position of the pedestrian in relation to the camera angle affecting the model's results considerably. In addition, missing data and incorrect joint position estimates lead to inaccurate classification. This problem is particularly evident in Figure 4.11 (a) and 4.11 (b), where the opposite joints (e.g., shoulders or hand) on the pedestrian are missing or incorrectly estimated. Although the lower body (heel, knee, ankle) has a different leg movement and visibility, the upper body, including the shoulder and hand, has limited movements and a limited range of joints values. This may confuse the distinction between "Text" and "None" activities, especially when objects are held and transported.

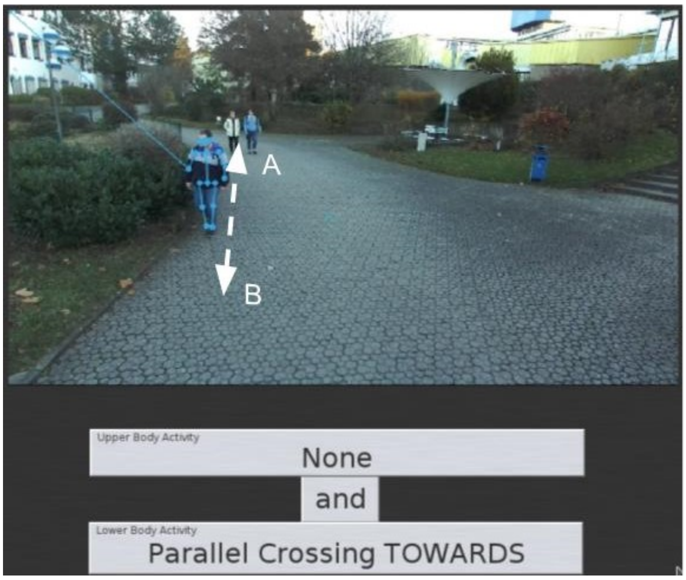
In Table 4.10, activities are classified into safe (S) and unsafe (U) categories. For example, pedestrians who walk away from a vehicle are always labeled as unsafe, regardless of what they are doing. However, when a person goes to the vehicle, the safety depends on what the person is doing — for example, when the pedestrian is busy with texting or calling, it is considered unsafe. Similarly, it is absolutely unsafe if a pedestrian passes perpendicularly while texting or calling. This classification facilitates smart interaction, enabling the conveyance of messages in a manner tailored to the pedestrian's current state and actions.

4.5 Geometric Analysis Approach

The efficacy of machine learning models in learning pedestrian behaviors is fundamentally dependent on the diversity and authenticity of the training data employed. Analysis in the



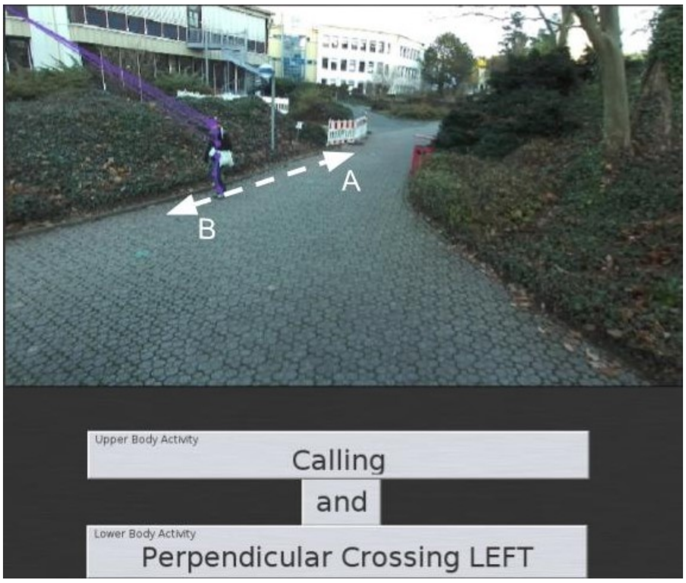
(a) A pedestrian detected while Perpendicularly crossing to the vehicle in left direction with no upper body activity.



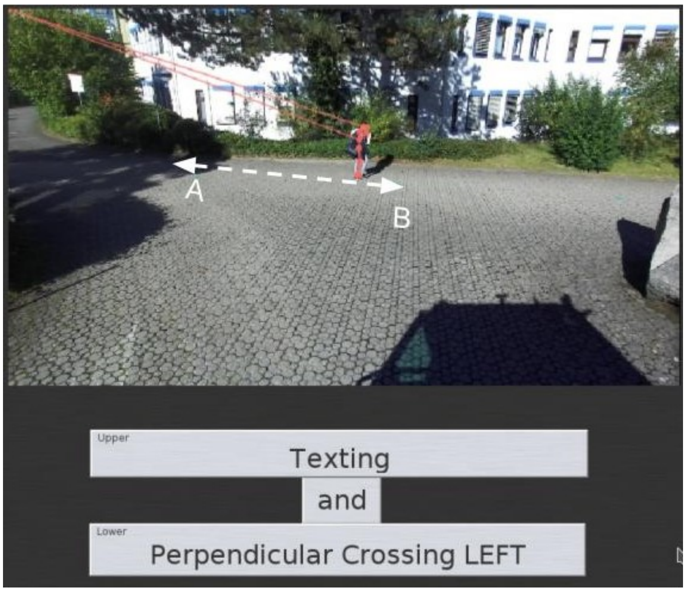
(b) A pedestrian detected as Parallel Crossing Towards the autobus.

(c) A pedestrian detected as Parallel Crossing Towards while calling or talking on phone.

Figure 4.10: Correct classification of upper & lower body activity of a pedestrian moving from distance A to B.



(a) Activity of pedestrian detected as Calling and Perpendicular Crossing instead of Parallel Crossing Towards and None.



(b) Pedestrian activity detected as Perpendicular Crossing but in wrong direction and Texting instead of None.

Figure 4.11: Misclassification of upper & lower body activity of a pedestrian moving from distance A to B.

Activities		Secondary			
		Calling	**Texting**	**None**	**Waving**
Primary	**Parallel Crossing Towards**	U	U	S	S
	Parallel Crossing Away	U	U	U	U
	Left Perpendicular Crossing	U	U	S	S
	Right Perpendicular Crossing	U	U	S	S
	Standing	U	U	S	S

Table 4.10: The table presents an integration of primary and secondary activities to evaluate conditions as Safe (S) or Unsafe (U).

use case, accuracy of 94% for upper body (Table 4.7) and 99% for lower body (Table 4.9) is adequate. However, to the inherent unpredictability of pedestrian behavior, curating a dataset that is both exhaustive and reflective of real-world conditions presents a significant challenge. Therefore, to ensure robustness, this thesis employs a geometric approach to posture analysis serving as a supplementary method to address potential deficiencies in case of wrong predictions. Moreover, through this approach other attributes can be reckoned such as distance, orientation which is further used for interaction fields (Chapter 5). The efficacy and methodology of this approach have been substantiated in [Jan 22b], wherein the same data type of 3D skeleton joint points were utilized to calculate various angles between different body parts. Furthering this research, a comprehensive comparative study was conducted by us and detailed in an extended paper [Jan 24]. This study plays together with the LSTM model and the geometrical method within the iB2C network, offering an in-depth analysis of both approaches. This dual-approach strategy exemplifies the integration of machine learning and geometric analysis in enhancing the robustness and reliability of pedestrian behavior prediction in autonomous navigation architectures.

Upon detecting the skeleton, the 18 joint points of the skeleton are represented as vectors within a rectangular coordinate system. These vectors are instrumental in determining the angles between various body segments. For instance, the vector representing the spine which is lying between neck and right hip joint (Example 4.1) is denoted by $\overrightarrow{Spine_{1,8}}$ and defined as follows:

$$\overrightarrow{Spine_{1,8}} = \langle x_1 - x_8, y_1 - y_8, z_1 - z_8 \rangle \tag{4.8}$$

Angles between each pair of vectors are computed to ascertain the posture of the skeleton. The orientation of the person is calculated by finding the angle between the shoulder vector $\overrightarrow{Shoulder_{2,5}}$ and y-normal vector. The posture is determined by analyzing a series of angles formed by different parts of the skeleton. These angles are calculated using the dot product formula between adjacent vectors, as shown in the equation:

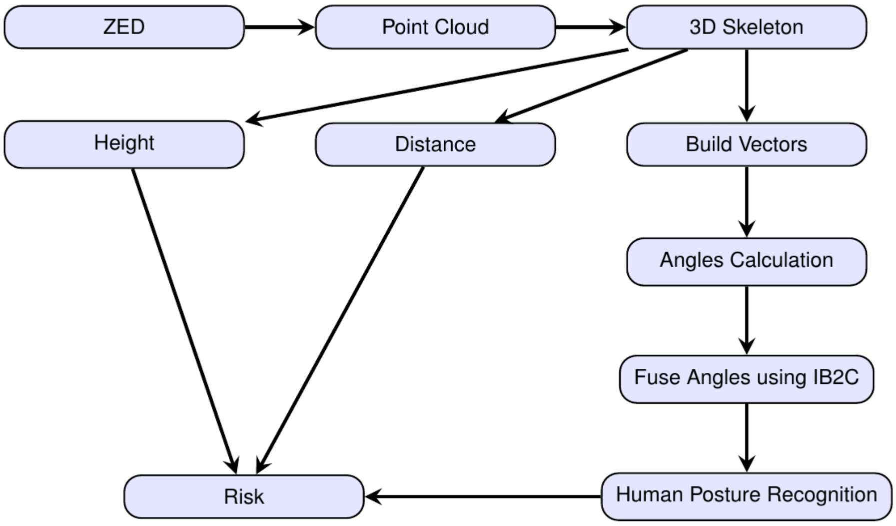


Figure 4.12: The image depicts a flowchart for a computational methodology designed to assess pedestrian postures through 3D skeletal joint point analysis. The process begins with a ZED camera capturing raw data, which is then converted into a point cloud representation. This point cloud serves as the foundation for constructing a 3D skeleton of the pedestrian. From the 3D skeleton, vectors are established to represent the orientation and position of various joints. Subsequently, angles between these vectors are calculated to provide detailed information on joint articulation. These angles are fused using iB2C. The resulting refined data feeds into a risk assessment algorithm that evaluates the potential hazards related to the pedestrian's posture, as well as into a human posture recognition system that categorizes the posture according to predefined classes for further analysis or action.

$$\cos\theta = \frac{\vec{p}\cdot\vec{q}}{|\vec{p}|\,|\vec{q}|} \tag{4.9}$$

In conjunction with the classes delineated for LSTM model, the geometric approach incorporates additional postures suggestive of an individual's lack of awareness:

- Texting on a phone
- Sitting
- Bending
- Looking away from the vehicle
- Having their back turned

These specific postures were identified and classified as activities within a behavior node, following the framework proposed by [Ropertz 17]. The activity level within this behavior node is quantified, reaching a maximum value (α=1) when a complete posture is exhibited, and reducing to zero in the absence of any posture. The risk associated with each behavior is correlated with the skeleton's proximity to the vehicle. A state of awareness in the person negates the risk, whereas unawareness triggers a risk activity proportional to the stimulation of the risk behavior.

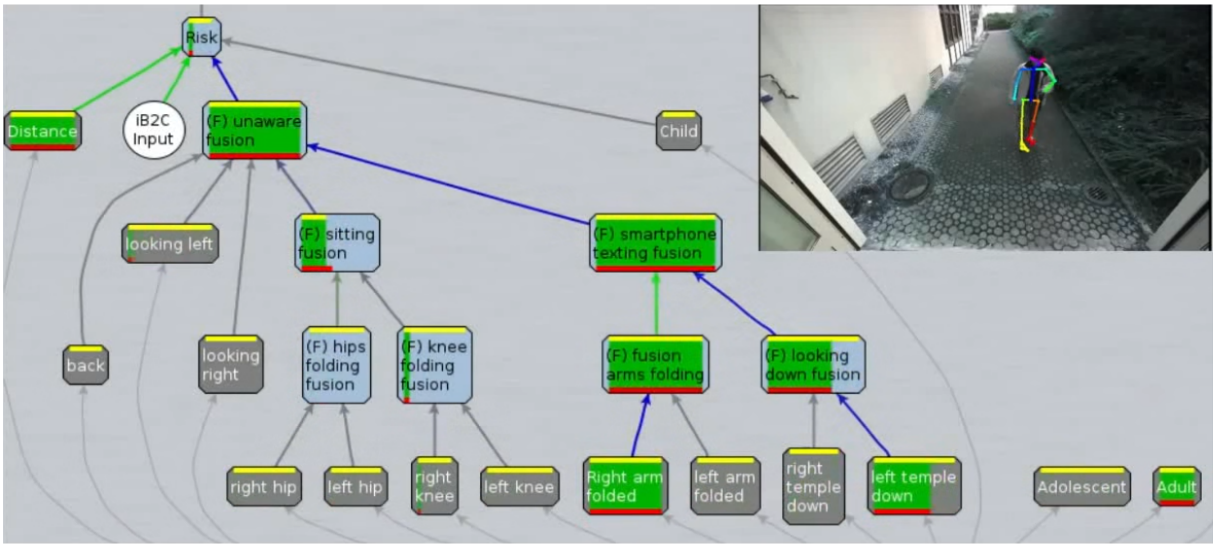


Figure 4.13: The figure presents a case study within the framework of posture recognition, which explains the detection of person with texting postures. This methodology involves a fusion process for the positions of the elbow, assessing whether the right elbow and the left elbow are bent upwards in a way consistent with the text on the smartphone. At the same time, it evaluates the angle of the temple, integrates the position of the visible temple, and further improves the posture analysis. This double evaluation activates the smartphone text fusion module, which requires a typical downward view of text and a typical elbow flexion. The system is specific; the smartphone text recognition module remains inactive unless the position of the individual complies with the text recognition parameters (the downward eye and curved elbow). Once these conditions are met, the system stimulates the unconscious fusion module, which is suggested to play a broader role in the context of behavior analysis, possibly assessing the individual's lack of understanding of their environment through their interaction with smartphones. As a result, the unaware fusion module becomes correspondingly active.

The methodology for computing human postures is illustrated in Figure 4.12. This structured approach facilitates the identification of potential risks based on the posture and proximity of pedestrians relative to the vehicle.

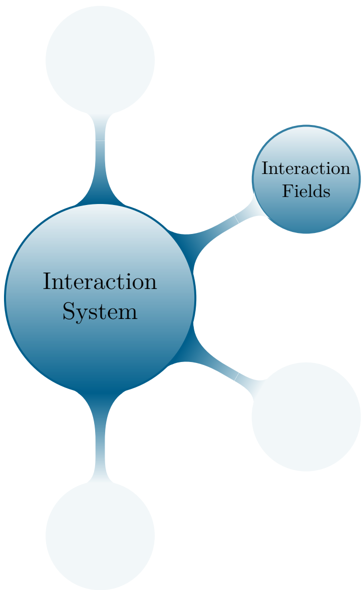
In examining the texting posture as illustrated in Figure 4.13, it is commonly observed that individuals often exhibit one or both elbows bent, with their heads tilted downward to focus on the smartphone. This behavioral pattern facilitates the establishment of anticipated angles for elbows and the temporal region of the head. To accommodate variations in body side involvement, a 'maximum fusion behavior' approach is implemented. This approach ensures that if either elbow or temple exhibits increased activity, the overall activity level of the maximum fusion behavior matches that of the most active module.

The detection of the texting posture is contingent upon the activation of two key fusion modules: the 'folding arms fusion' and the 'looking down fusion'. The activity level of the texting posture module is directly influenced by the 'folding arms fusion' activity, while also being equivalent to the activity of the 'looking down fusion'. The rationale behind this approach is that the texting posture should primarily be identified by the characteristic arm position associated with texting. While looking down can occur in various contexts and may not necessarily indicate texting, the combination of this behavior with the specific arm posture significantly increases the likelihood of identifying a texting posture.

Moreover, the 'unaware fusion behavior' amalgamates various indicators of a pedestrian's lack of awareness, such as looking away, sitting, or the presence of a child. This fusion is crucial in assessing the risk posed by a pedestrian.

In geometric approach, the assessment of pedestrian risk is based on two principal factors: proximity to the vehicle and the level of awareness. The risk module's response is proportionate to the pedestrian's distance from the vehicle, with closer proximity resulting in higher stimulation. However, the actual activation of the risk module is contingent upon the pedestrian's awareness. This implies that a pedestrian in close proximity to the vehicle does not inherently constitute a significant risk if they are aware of their surroundings. Conversely, a lack of awareness in such situations directly elevates the risk level, equating it with the proximity factor. This dual-criteria mechanism ensures a more nuanced and accurate evaluation of pedestrian risk in varying contexts.

5. Interaction Fields



The pedestrian zones present a spectrum of structural variations that have a significant impact on the navigation of driverless minibuses such as the Autobus. These areas can range from highly organized environments with narrow paths and crossings to complex, unstructured spaces where traditional navigational clues are minimal or absent. The structures of pedestrian areas are influenced by their design, their intended use and the level of pedestrian and vehicle integration. Well-defined areas promotes predictable pedestrian behaviour, while more open shared spaces necessitate complex navigational interpretation algorithms. The creation of a interaction field around the Autobus and pedestrians is essential for smart interaction in pedestrian areas. These interaction fields are virtual dynamic zones that include the vehicle, modulating size and shape in response to the movements of the vehicle. These fields serve a dual purpose, namely to improve safety and enable useful interaction between vehicles and pedestrians. Therefore, after the recognition of the activity of the pedestrian (Chapter 4), the risk of the pedestrian can be translated into such interaction fields. Based on this assumption, the following section briefly describes the challenges and implications of different environmental structures. It examines the different levels of risk associated with pedestrian encounters and indicates the likelihood of pedestrian and vehicle collisions. The upcoming section reviews research pertaining to interaction fields, drawing inspiration for development of such fields for Autobus. Finally, the practical implementation explains the design concepts that support this research.

5.1 Impact of Environmental Structure on Interaction Fields

Figure 5.1 illustrates the relationship between varying degree of environmental structure to path clarity and pedestrian risk. An approximation graph is depicted that shows a decrease in pathway clarity when the geometrical structure of the environment becomes less defined. The approximation is derived from a generalized understanding of the environment, excluding anomalies. For instance, the presence of a man-made pathway within a forested area is not accounted for in this model. Additionally, factors such as lane markings and fixed road colors further contribute to this approximation, enhancing the accuracy of detection algorithms in structured environments by aiding in the identification of drivable paths.

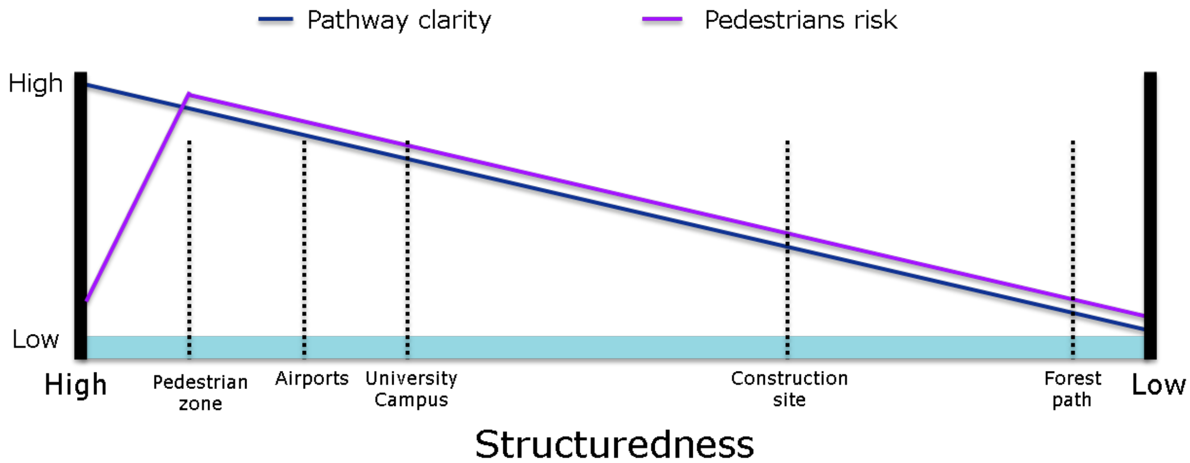


Figure 5.1: The graph illustrates the relationship between "Pathway Clarity" and "Pedestrian Risk" for variability of "Structuredness" in urban environment. It shows how the risk to pedestrians varies with the clarity of pathways in different environments. The x-axis denotes the degree of geometric structuring in pedestrian zones, ranging from low to high, signifying the extent to which pedestrian pathways are distinctly demarcated. The y-axis quantifies pedestrian risk, with higher values denoting increased danger, and pathway clarity with higher value denoting segregation of pathway for pedestrians.

In pedestrian zones, the complexity of autonomous navigation varies widely, influenced by the degree of structure within the environment. Example 5.1 compares different pedestrian zones, highlighting this variance. The left image in the example presents a pedestrian zone with a clear drivable path, indicative of a structured environment. The right image shows trails in a busy pedestrian area with no segregation which is further depicted in Figure 5.1 as "pedestrian zone". It can be seen that the pedestrian risk tremendously increases Figure 5.1 demonstrates a lower dip in pedestrian risk in highly structured environments, such as street settings with separate walkways for pedestrians and designated roads for vehicles, indicating a safer interaction dynamic.

Example 5.1: Comparison of Various Pedestrian Zones

The image showcases a pedestrian zone with a more structured urban design. It features distinct walkways for pedestrians, separated from the areas where vehicles are permitted. This setting aids driver-less minibuses in path planning, providing more predictable pedestrian behavior patterns. However, challenges persist with the occasional pedestrian deviating from the designated paths, necessitating that vehicles maintain an alert monitoring system. Transitioning between different types of zones, like from a clear path to a shared zone, also requires careful negotiation by the vehicle. Interaction fields in such structured zones, while possibly less variable than in open plazas, must still be capable of adapting quickly to unforeseen pedestrian movements to maintain safety.

Contrastingly, the image shows an open and broad space with very few physical guides like marked paths or barriers, challenging autonomous vehicle navigation. In such an environment, the vehicle must rely on complex algorithms to predict pedestrian movements that can be highly random and multidirectional. The absence of structural cues increases the need for vigilant risk assessment, as pedestrians could potentially traverse any part of the plaza at any instant. Driver-less minibuses here must employ dynamic and highly responsive interaction fields to safely navigate through the varying pedestrian densities and activities.

In the context of Autobus, it is imperative to establish clearly defined spatial zones around the vehicle, also referred as "interaction field". These fields are essential for facilitating effective communication between the Autobus and pedestrians. The complexity and inherent risks associated with various pedestrian zones necessitate this stratification. By delineating specific regions around the vehicle, these interaction fields allow the Autobus to gauge the position and potential risk associated with individual pedestrians. Consequently, the Autobus can modulate its behavior and communication strategies accordingly. The idea of interaction fields focuses primarily on pedestrians who are in immediate need of interaction, based on their proximity and the level of risk they pose. This approach ensures that the Autobus can operate safely and efficiently in environments with varying pedestrian densities and behaviors.

Based on the preceding motivation, this chapter delves into the design of interaction fields between Autobus and pedestrians. The chapter commences with literature review, aiming to elucidate and distill the concept of interaction fields. This is followed by an in-depth exploration of the methodologies employed in this study. The chapter concludes with a detailed discussion of the findings in the discussion section.

5.2 Enhancing Pedestrian Engagement through Phased Interaction Zones

This idea of introducing interaction phases is given in [Vogel 04]. The authors have developed an interaction framework that progresses through four phases, moving from public to personal interaction with fluid transitions between each stage. These phases are Ambient Display, Implicit Interaction, Subtle Interaction, and Personal Interaction. Unlike previous models, their framework doesn't rely solely on physical proximity to define phases, nor does it require handheld devices for personal interaction. It emphasizes fluidity and supports multiple users, each in their own phase. The ambient display shows general information as a neutral state, subtly shifting to implicit interaction when a user approaches, inferring their receptiveness to information. During subtle interaction, the display provides more detailed notifications and personalizes public information for the user. Simple explicit actions, like hand gestures, are used for interaction in this phase, allowing shared use and maintaining an overview of the display. Finally, in the personal interaction phase, users can engage more directly with the display, such as through touch, for detailed and personal information. This phase allows for longer interaction while minimizing disruption to others, using the user's body to occlude sensitive information from onlookers. An example is shown in Figure 5.2.

Based on the interaction framework proposed by D. Vogel et. al. [Vogel 04], this thesis proposes the implementation of Autobus interaction zones in order to significantly improve their engagement with pedestrians. This model adopts a phased interaction approach, which seamlessly moves from the environment to direct engagement, which promotes a more natural and effective communication channel between the Autobus and pedestrians. The basic area, similar to the "Ambient Display", serves as a stage of general awareness, where the Autobus communicates its presence to pedestrians nearby through passive interaction, ensuring that pedestrians recognize the intentions of the vehicle without needing direct interaction. As the pedestrians come closer and reflect the "implicit interaction" phase, the Autobus dynamically adjusts its behaviour in response to the pedestrian's path and speed, revealing subtle intentions such as speed adjustments.

Progressing to a more "direct" interaction phase, the Autobus begins a more explicit communication with pedestrians in the surrounding area, using visual or auditory signals to define its forthcoming actions or to recognize pedestrian priority. This level of engagement improves security and clarity in densely populated or complex urban environments. In the most interactive environments, similar to the "personal interaction" phase, the Autobus transitions to the "risk" interaction mode, including emergency braking and other direct communication methods to ensure the highest safety.

By articulating these interaction zones, the Autobus is placed to significantly improve its communication effectiveness with pedestrians, safety, predictability and trust in similar driverless minibuses. This strategic approach is aligned not only with the paradigms of human interaction, but also adapts to the different degrees of involvement required in different pedestrian environments.

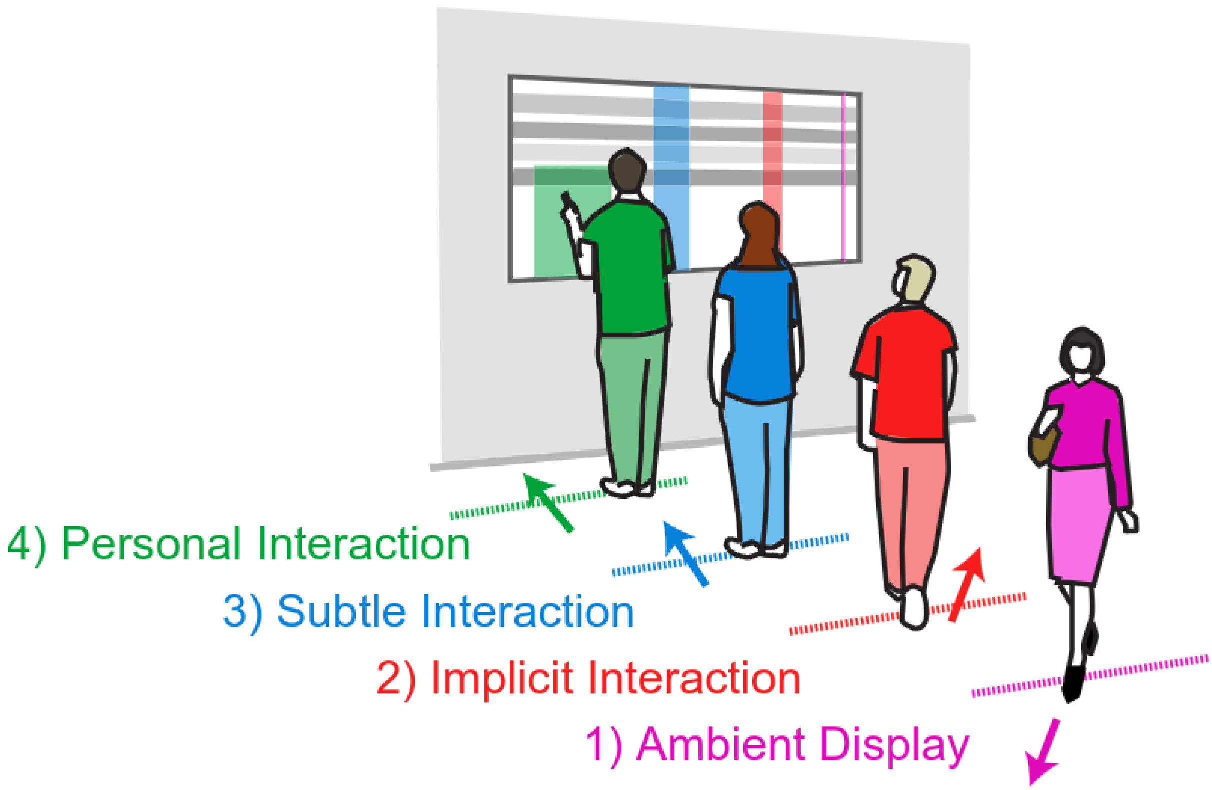


Figure 5.2: The image depicts a sequence of interactions between people and a display, categorized into four levels: Ambient Display - A person walking by without interacting. Implicit Interaction - A person standing and looking at the display with interest in Red. Subtle Interaction - A person (in blue) closer to the display, looking to the screen. Personal Interaction - A person (in green) directly touching and interacting with the display [Vogel 04].

Establishing interaction fields for Autobus offers several advantages. Three primary benefits, as highlighted in this study with respect to the interaction concept defined in Chapter 2, are outlined in Advantages 5.1

Advantages 5.1: Interaction Fields

Understanding the Risk Posed by Pedestrians Understanding how close pedestrians are and how they move is very important for Autobus in pedestrian zone. The interaction fields helps Autobus to identify risks of pedestrians. These interaction fields help the Autobus navigation system not only to point out the presence pedestrians, but also to understand how dangerous a situation could be based on how close they are, their movement direction, and how fast they are moving. For example, a pedestrian moving rapidly towards the vehicle's trajectory represents a higher risk than someone moving away or standing still. By looking at these dynamic fields, the Autobus can identify risks in real-time and distinguish between safe and potentially hazardous situations. In this way, the Autobus can make informed decisions such as slow down, stop or reroute, thereby significantly enhancing safety for both pedestrians and vehicle occupants.

Localizing Pedestrian Risk Around the Vehicle Another key aspect of these interaction fields is their role to map pedestrian risks around vehicles. By dividing the area around a vehicle into different areas with varying levels of risk based on what pedestrians activity, the vehicle can maintain spatial awareness of its immediate surroundings. This is particularly important in urban areas where pedestrians can approach from all directions. These interaction fields are dynamic buffer zones that are updated to respond to how pedestrians move. This real-time information allows Autobus to adjust its path and speeds as needed, ensuring the safety of pedestrians and reduce the risk of accidents.

Determining Required Interaction Types Furthermore, the interaction field is the key to understanding how the Autobus interacts with pedestrians. Different behaviours require different responses. For example, a pedestrian making eye contact and indicating intention to cross may require the vehicle to yield, while walking pedestrian groups on the marked streets with separate sidewalk do not need immediate action. By analyzing these fields, the autonomous system can detect these subtle indications and react accordingly. This capability not only contributes to safety, but also helps to ensure that traffic flows smoothly. This avoids unnecessary stops and maneuvers that may disturb traffic flow.

5.3 Establishing Interaction Fields Design

Understanding and defining these various interaction fields is crucial to the development of the Autobus. It allows customized navigation systems and safety protocols to meet the specific challenges of each environment. In highly structured environments, the autobus can be programmed to expect more predictable pedestrian behaviour, whereas in unstructured zones, it can give priority to the navigation of the ground and the sudden detection of pedestrians. This tailored approach not only improves the safety of pedestrians and passengers, but also contributes to the wider acceptance and trust of driverless minibus technology by showing adaptability and understanding of various urban changes by the degree of access of the population.

In developing vehicle interaction field for Autobus, it is essential to incorporate a comprehensive understanding of pedestrian behavior, particularly the risk of pedestrians. This risk stems mainly from a lack of awareness among pedestrians about the presence of Autobus, which may occur in scenarios where pedestrians are moving in the same direction as the Autobus in the front or where pedestrians are busy with electronic devices. Such behaviour significantly increases the risk of collision or delays in the transportation of the Autobus. This thesis therefore proposes a customized model of pedestrian and vehicle interaction fields to improve the decision-making capabilities of Autobus navigation and interaction systems. The aim of this model is to mitigate risk by facilitating more informed and responsive adjustment of driving and interaction strategies of vehicles.

5.3.1 Vehicle Interaction Field

Not all the visible surrounding regions around the vehicle are of interest for interacting with pedestrians when driving. Due to open space and presence of many people in the surrounding, interaction should, certainly, be more dedicated to the interactee. Interactee may not be reacting to the vehicle, but is anyone in the drivable area of the vehicle. For such purposes, there is a need to segregate the region of interest and further divide them according to the types of interactions. To embrace this strategy, interaction fields are created around the vehicle which assists in encumbering the usage of interaction modules; hence, narrowing the wide range of activated combinations of interacting modules in a sensible way.

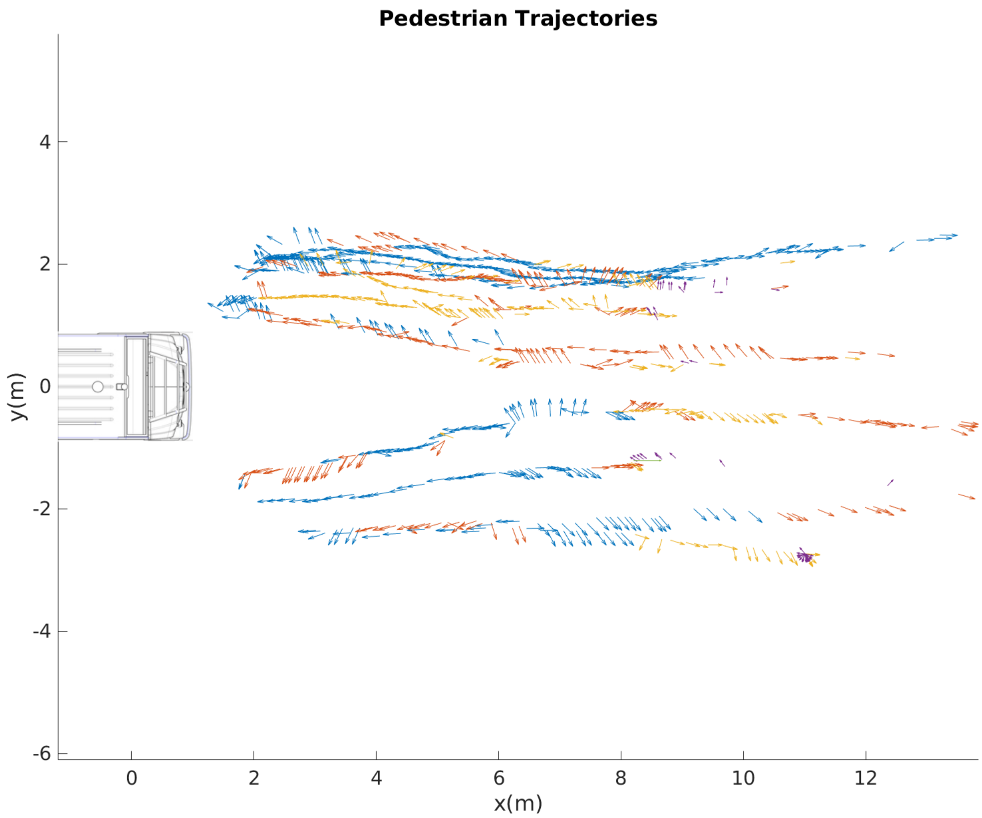


Figure 5.3: The depicted plot illustrates recorded pedestrian trajectories during their encounters with the Autobus. A discernible pattern emerges from the data, suggesting that pedestrians tend to initiate avoidance maneuvers in response to the vehicle's presence in a consistent manner. This observed regularity indicates the potential for predictable pedestrian behavior, which could be leveraged to design the a pattern for interaction field where it is not interrupt in normal circumstances.

The configuration of the interaction field is informed by pedestrian trajectories relative to the vehicle. As depicted in Figure 5.3, there is a clear pattern in the movement of pedestrians who are moving towards the Autobus. Under typical conditions, where pedestrians are cognizant of the approaching Autobus, they tend to alter their path to yield to the vehicle, especially in constrained spaces. This behavior, as observed and analyzed by F. Schneemann and I. Gohl [Schneemann 16], indicates that pedestrians often take the responsibility of avoiding the vehicle, demonstrating a collective pattern of spatial negotiation and collision avoidance.

The interaction fields around the Autobus, as illustrated in Figure 5.4, are designed to facilitate graduated levels of communication and risk assessment based on the proximity of pedestrians. This design is motivated by the observed pedestrian behavior from the plot shown in Figure 5.3, which highlighted the varying responses of pedestrians to the presence of the Autobus in their vicinity. The design of the interaction fields around the Autobus, extending longer in the direction of travel and shorter in the opposite direction, is informed by the dynamics of pedestrian and vehicle interaction as well as the principles of risk assessment.

- The forward elongation of the fields accounts for the vehicle's stopping distance, which is proportional to its speed.
- A longer interaction field in the direction of travel ensures that there is ample space for the Autobus to come to a safe stop if necessary.
- In contrast, the interaction fields are smaller on the sides of the Autobus because pedestrians moving in this direction are more likely to have already observed the vehicle passing by, reducing the level of risk.

Additionally, the Autobus's sensory and navigation systems should be more readily detect and respond to pedestrians who are entering these fields from the front or sides rather than only from sides. Thus, the risk of a collision is naturally lower from the sides, and the fields can be reduced in size without compromising safety. The other reason of narrow size on the sides is because of narrow pathways, as discussed in Chapter 3 Overall, the asymmetric design of the interaction fields around the Autobus optimizes the vehicle's response time and decision-making capabilities, enhancing pedestrian safety in varying traffic conditions and pedestrian behaviors. The three categories of interaction field are:

- **Ambient interaction:** The outermost 'Ambient' field represents a zone of general awareness, where the vehicle signals its presence passively to pedestrians. This field aligns with the observed tendency of pedestrians to acknowledge the vehicle from a distance without requiring direct engagement. The extended Ambient field in the driving direction reflect the focus on pedestrian crossing from the front. This is the furthest distance in interaction. It is meant to show general information of the vehicle's intention. There is no risk of collision in ambient field, hence it has the same idea of "Ambient Display" shown in Figure 5.2.
- **Direct interaction:** The 'Direct' interaction field, as demarcated in Figure 5.2, facilitates overt communication, in contrast to the more nuanced 'Implicit' and 'Subtle' interactions shown in Figure 5.2. This field is notably extended in the direction of the Autobus's travel to preemptively address the safety of pedestrians who have not yet engaged with the vehicle. The lengthened 'Direct' field enables the early detection of pedestrians, particularly those who, due to distraction or orientation away from the traffic, may not be cognizant of the vehicle's proximity. Such anticipatory design affords the Autobus the necessary freedom to modify its trajectory or velocity in response to pedestrian movements, thereby enhancing safety. Upon entry into the 'Direct' field, the Autobus escalates its interaction

with pedestrians to more active forms of communication. This shift is grounded in empirical pedestrian behavior, which shows that individuals tend to realign their walking patterns acknowledging the vehicle's path, thus mitigating the risk of collision. Ideally, the 'Direct' field is to be circumvented by pedestrians to avoid unnecessary alerts. However, if a pedestrian does penetrate this zone—indicating a lack of spatial awareness—the system is designed to issue an immediate auditory signal or verbal warning through speech synthesis to promptly capture the pedestrian's attention and convey critical safety information.

- **Risk interaction:** The 'Risk' interaction field, situated immediately around the Autobus, is designed to initiate a critical response measures when pedestrians enter this area. The establishment of this zone is informed by the risk data of pedestrians, which underscores the necessity for swift action in instances of pedestrian unawareness or erratic movement. Upon a pedestrian's entry into this zone, the Autobus is programmed to elevate its alert status to the maximum, potentially triggering emergency protocols to avert a collision. The zone being closest to the vehicle, the 'Risk' field is inherently associated with a high collision threat. In such events, the Autobus is required to engage its emergency braking systems promptly and employ clearly visible visual signals, such as flashing lights, to alert the pedestrian. This immediate and forceful response is a vital aspect of the vehicle's interaction strategy, ensuring the highest level of safety during close encounters.

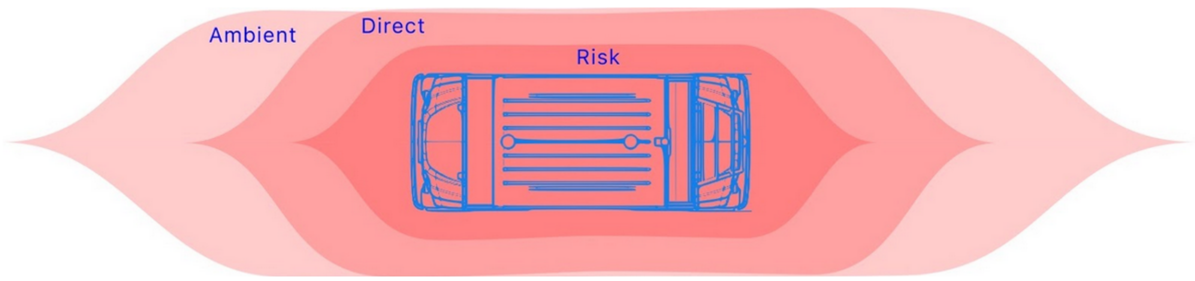


Figure 5.4: Schematic representation of the Autobus with delineated interaction fields. The 'Ambient' field, shown in light red, signifies an initial zone of awareness, extending predominantly in the direction of the vehicle's travel to provide early detection of pedestrians. The 'Direct' interaction field, depicted in a darker shade of red, indicates an intermediate zone where more active communication with pedestrians occurs, allowing for dynamic response as pedestrians approach closer. The 'Risk' field, marked in the darkest red, encompasses the immediate vicinity of the Autobus, where the potential for collision is highest, necessitating prompt response measures. This stratification facilitates graded interaction with pedestrians, enhancing safety protocols based on proximity and relative motion for Autobus.

Overall, the vehicle interaction field is strategically layered to allow the Autobus to modulate its communication and safety measures dynamically. This structure ensures that the vehicle remains aware of its surroundings and can react appropriately to the fluid and sometimes unpredictable nature of pedestrian movement, thereby enhancing safety for all parties involved.

5.3.2 Pedestrian Interaction Fields

For the integration of pedestrian models into vehicular interaction field mappings, it is imperative that pedestrians are represented in a manner analogous to interaction fields. Ellipses constitutes an appropriate geometric representation for this purpose, as suggested by W. Limprasert et al. [Limprasert 13]. Using ellipses to represent pedestrians in vehicular interaction fields provides significant benefits for decision making. Elliptical models effectively capture the spatial extent and movement direction of pedestrians with their major and minor axes, aligning well with natural human motion paths. This simplifies computational processes, enabling quicker decision-making in Autobus due to the geometric simplicity and reduced processing demands of ellipses. Furthermore, ellipses allow for dynamic adjustments in scale and orientation, accommodating variations in pedestrian speeds and directions. This adaptability, combined with the compatibility of ellipses with Pedestrian detection (Chapter 4), facilitates efficient prediction of pedestrian movements and seamless vehicular adjustments. Overall, the use of ellipses strikes a practical balance between representing human movement accurately and ensuring computational efficiency, enhancing the safety and reliability of pedestrian-vehicle interactions. Figure 5.5 shows the representation of pedestrian as ellipses.

An ellipse is mathematically characterized by its geometric parameters, which can be correlated with pedestrian movement aspects relevant to interaction fields for Autobus. The standard equation of an ellipse, centered at the origin and aligned with the coordinate axes, is:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \tag{5.1}$$

where a and b are the lengths of the semi-major and semi-minor axes, respectively. For an ellipse that is centered at a point (x_0, y_0) and rotated by an angle θ, we adapt this equation to encompass pedestrian positioning and orientation.

- **Center (x_0, y_0):** Represents the pedestrian's current location or pose in the plane, acting as the geometric center of the ellipse.

- **Major axis (2a):** The longest diameter of the ellipse, corresponding to the primary direction of pedestrian movement. The semi-major axis a signifies the forward movement direction and intended path length of a pedestrian.

- **Minor axis (2b):** Denotes the shortest diameter of the ellipse, orthogonal to the major axis. The semi-minor axis b reflects the lateral movement or space occupied by the pedestrian, essentially representing the pedestrian's width in the movement space.

- **Angle of rotation (θ):** Indicates the orientation of the ellipse in relation to the coordinate frame, aligning the major axis with the pedestrian's movement direction. This angle provides an alignment of the pedestrian's projected path with the environment for accurate interaction simulation.

- **Color gradient:** The color gradient within the ellipse symbolizes the pedestrian's level of awareness, originating at the center and extending outwards to the ellipse's periphery. A progression from the center towards the outer edge indicates an increase in the pedestrian's unawareness, with values escalating from the core to the boundary.

These parameters collectively define the spatial and directional attributes of a pedestrian within an interaction field, aiding autonomous vehicles in dynamically adjusting their navigation for pedestrian safety.

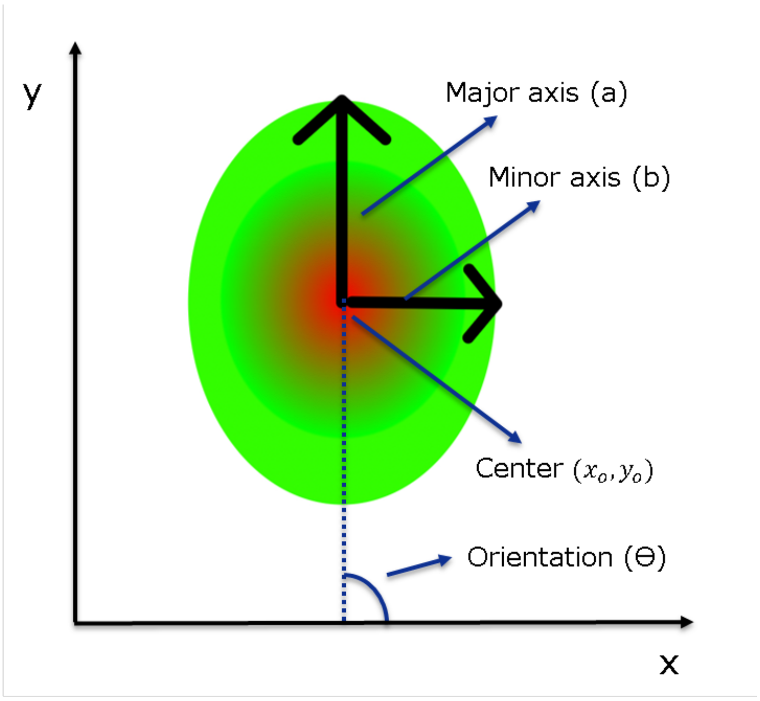


Figure 5.5: A graphical representation of a pedestrian interaction field in a coordinate system, depicted as a gradient-filled ellipse. The axes labeled 'X' and 'Y' represent the coordinate frame of reference on the horizontal and vertical planes, respectively. The gradient intensity within the ellipse signifies the awareness of the pedestrian, with the highest value at the center transitioning to lower values towards the periphery. The values increases with increase in unawareness. The two arrows denote the major (vertical arrow) and minor (horizontal arrow) axes of the ellipse, illustrating the pedestrian's primary forward movement direction and the lateral dimension of space occupation, respectively.

5.4 Integrating Interaction Field Representation as an Entity of Aspect Maps For Existing Navigation Framework

The design of interaction fields is performed using aspect maps. This is a novel framework for applying cognitive processes in robotics, which was developed by GregorGregor [Zolynski 18]. This framework clearly differs from traditional object-oriented approaches and instead chooses parallel data flow networks. These networks are inspired by the first visual perception systems observed in animals and humans. Aspect maps, which are integrated spatial and symbolic structures that represent limited aspects of the environment and capture only one aspect. These maps are constructed using aspects from different sources, including other aspect maps, general storage, expert knowledge and sensor data. Aspect maps are used not only to unite data, but also to combine aspects to answer specific questions, extrapolate new information, and identify contradictions. The methodology focuses on modular and reusable solutions and applies a strategy of division and conquer to address cognitive challenges. This approach uses multilevel abstractions of sub-symbolic information. By focusing on environmental aspects rather than objects, this architecture opens the door to a number of potential applications. Its strengths include early data abstraction and unification, facilitating general processing algorithms and multimodal data integration. In addition, data flow network approaches improve transparency and simplify the control and verification of cognitive processes, thus accelerating development and improving the reliability and effectiveness of architectures.

Since, aspect maps are elements of a representation architecture for processing spatial knowledge, the interaction fields for vehicle and pedestrian are created as one of its representation in its spatial structure so that it can be easily correlated and combined.

5.4.1 Vehicle Interaction Field Aspect

The delineation of the interaction field surrounding a vehicle is an essential process that requires establishing its dimensions and form. As delineated in Figure 5.4, the vehicle's interaction field possesses a distinct configuration, for which Bézier curves are employed in the design process. The application of Bézier curves for crafting the interaction field contours around the bus, as illustrated in Figure 5.4, has proven to be highly efficient, with its benefits detailed in Advantages 5.2.

Four control point bezier curve is created. The cubic Bézier curve equation with four control points is represented in Equation 5.2

$$B(t) = (1-t)^3 P_0 + 3(1-t)^2 t P_1 + 3(1-t) t^2 P_2 + t^3 P_3 \tag{5.2}$$

where:

- $B(t)$ represents the point on the Bézier curve.
- P_0, P_1, P_2, and P_3 are the control points.
- t is the parameter that varies from 0 to 1.

At $t = 0$, the curve starts at the first control point P_0, and at $t = 1$, the curve ends at the last control point P_3. The intermediate control points P_1 and P_2 define the tangents to the curve at the endpoints and govern the curve's shape. As the parameter t progresses from 0 to 1, the curve passes through the intermediate points, generating a smooth path within the convex hull of these control points.

Advantages 5.2: Bézier Curves For Interaction Field Design

Precision of Shape Definition: Bézier curves offer a high level of control over the shape through the manipulation of control points. By adjusting these points, the designer can precisely define the curvature of the interaction field to accurately represent the physical characteristics of the phenomena being modeled.

Smoothness and Continuity: Bézier curves are mathematically smooth and continuous. This characteristic is critical when representing fields that naturally exhibit smooth gradations. The smoothness ensures that there are no sharp edges or abrupt changes in the interaction field that would be physically unrealistic.

Scalability and Flexibility: The interaction fields can be scaled up or down or stretched to simulate different scenarios or to fit into different scene contexts. Bézier curves are parametric, which means they can be easily scaled and transformed without loss of detail or shape integrity.

Visual Clarity: When illustrating concepts, clarity is paramount. Bézier curves provide a clear and aesthetically pleasing way to represent interaction fields.

Animation and Interpolation: Since grid maps are used, the interaction fields need to be animated or interpolated between states (for example, showing how the risk area expands as a Autobus accelerates), Bézier curves facilitate smooth transitions. The intermediate shapes during animation can be easily calculated, providing a realistic and continuous motion.

Mathematical Operations and Collision Detection: When simulating interaction fields, it may be necessary to perform operations such as finding intersections or calculating the area within a curve. Bézier curves have well-defined mathematical properties that facilitate these operations, which can be critical for simulations involving collision detection or overlap analysis.

Figure 5.6 illustrates a top-down schematic view of Autobus with two Bézier curves extending from its front to the right and left, each depicting the interaction field shape for both sides. In aspect maps, they are used to create the boundary of the vehicle interaction field. Algorithm 5.1 shows the creation of Bézier curve on a grid map. Figure 5.7 shows the corresponding implementation of Bézier curves on the grid for aspect map.

In the associated grid map, the contours of the vehicle's interaction field are defined by the placement of the interpolated points. In close proximity of the vehicle, grid values are purposefully designated as null, a detail vividly captured in Figure 5.8. In this depiction, regions with null values are indicated in red, and those assigned a value of one are shown in green. This chromatic demarcation ensures a stark visual contrast between zones in the grid, and later, for establishing a gradient encircling the vehicle.

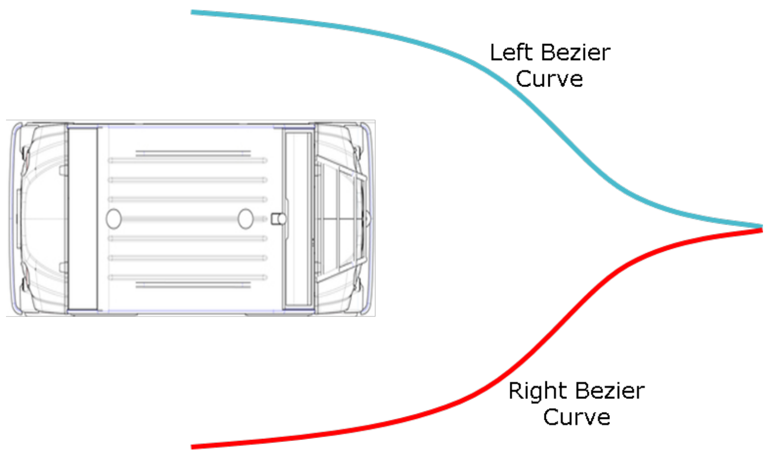
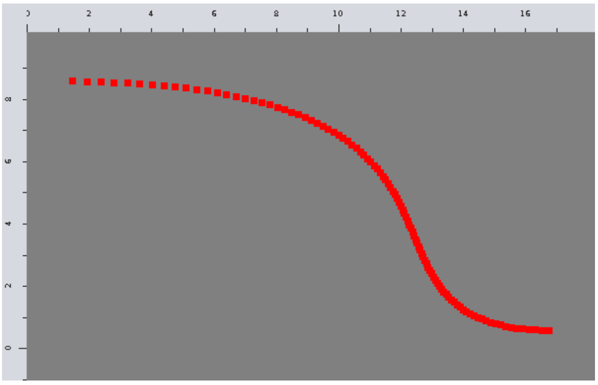


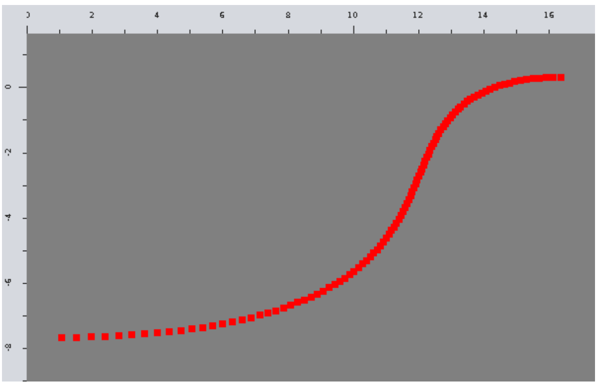
Figure 5.6: Top-down view of Autobus with associated interaction field outline: the cyan curve represents the left side of the interaction field, while the red curve indicates an alternative side, both modeled using cubic Bézier curves.

Algorithm 5.1: Updating Aspect Grid with Spline Function Values

```
in_gradient, in_spline_pts, par_cell_size ;
Initialize aspect grid with gradient values;
pts ← GetPointer(in_spline_pts);
if pts.size() > 0 then
    forall points in pts do
        t ← aspect grid.GetCellIndex(point.X() / GetValue(par_cell_size), point.Y() /
          GetValue(par_cell_size));
        aspect grid.SetCellValue(t, 1.0);
    end
end
return aspect grid
```

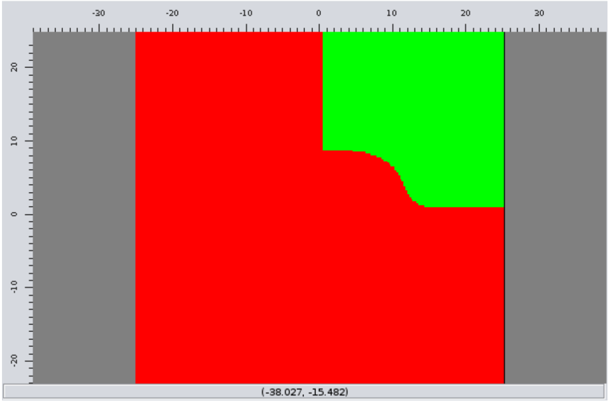


(a) Bézier curve on the left of the vehicle as shown in Figure 5.6.

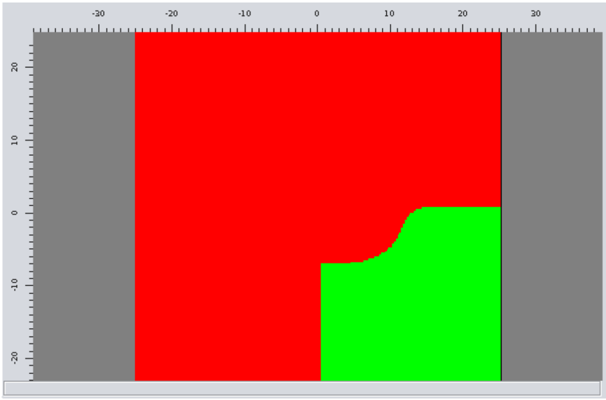


(b) Bézier curve right of the vehicle as shown in Figure 5.6.

Figure 5.7: Implementation of a cubic Bézier curve applied in aspect map modeling.



(a) Left side segmentation curve



(b) right side segmentation curve

Figure 5.8: Visualization of spatial partitioning using a cubic Bézier curve to delineate boundaries: The red region is assigned a value of zero, representing one classification, while the green region is assigned a value of one, indicating a separate classification. The curve provides a smooth transition between two distinct value domains on a two-dimensional plot, with coordinates measured on an orthogonal axis system.

Upon the successful delineation of the boundary maps as depicted in Figure 5.9, an additional step involves the incorporation of gradient data. Initially, the 'right filtered boundary' undergoes a thresholding process using the gradient map as a reference. During this process, the zero values, indicated in red, are assigned values from the gradient, while the green cells are designated as NaN (Not a Number), indicating the absence of a value. A similar procedure is then applied to the composite map resulting from the previous operation, this time employing the 'left filtered boundary' for thresholding. The culmination of these steps is illustrated in the uppermost map of the figure, which presents the finalized vehicle field map. A similar process is done for the rear of the Autobus to have a similar shape shown in Figure 5.4. Depending on the driving direction, half the vehicle fields are active.

The gradient is structured to symbolize three distinct interaction zones within the vehicular field, namely, the ambient zone, the direct interaction zone, and the risk zone. It is parametrically defined such that the values are at their highest proximate to the vehicle's centroid and decrement progressively towards the periphery of the field. Consequently, this gradient serves as a quantifier of spatial interaction potential, delineating regions of varying

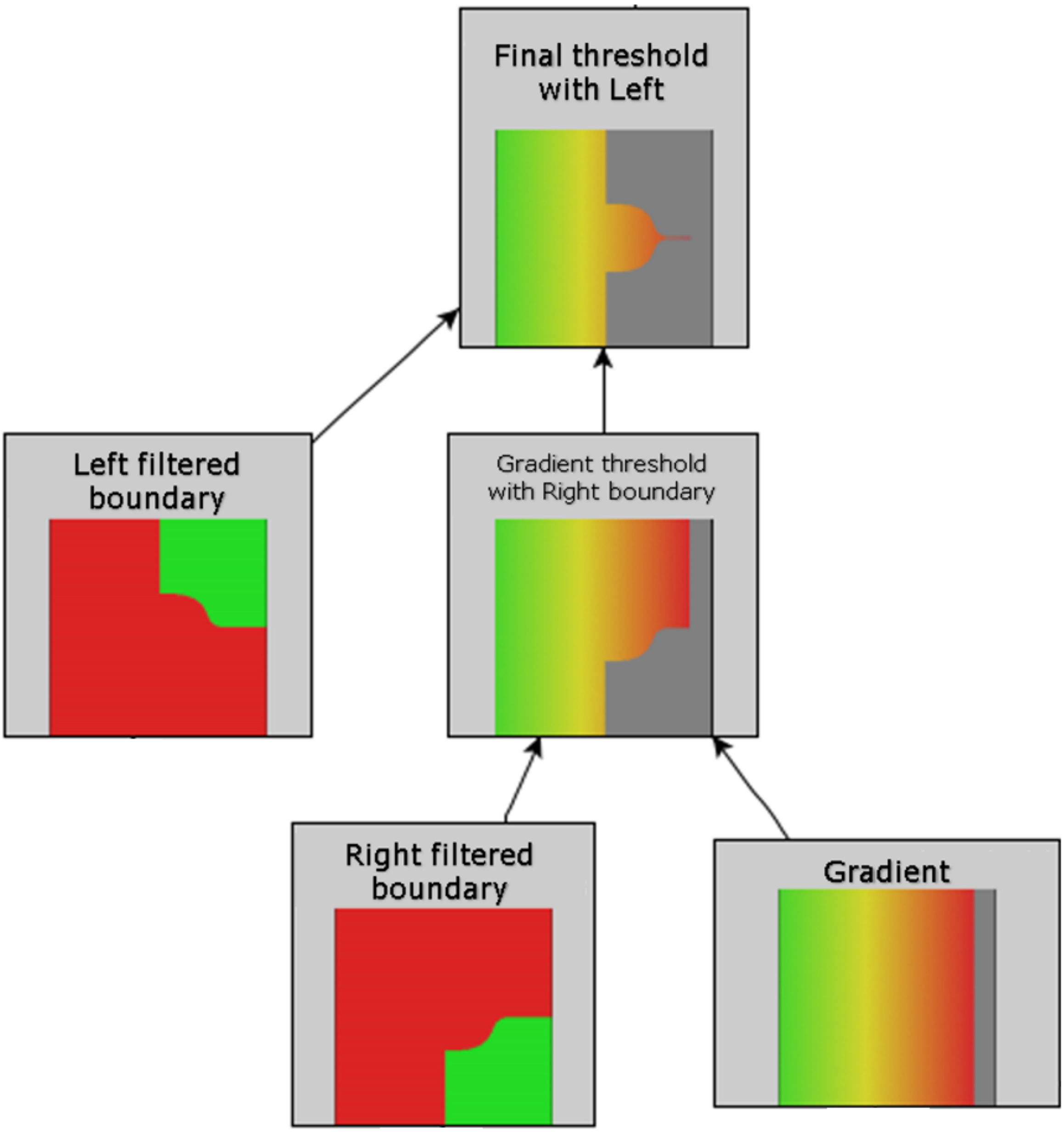


Figure 5.9: Process flow diagram for boundary integration with gradient thresholding. The lower left and adjacent right panels illustrate the 'Left filtered boundary' and 'Right filtered boundary' maps, respectively, with green indicating defined ones and red indicating null values. The right middle panel shows the 'Gradient' map with a color gradient applied. This gradient map is utilized to threshold the 'Right filtered boundary' map, as depicted in the middle left panel, titled 'Gradient threshold with Right boundary', where red areas are replaced by corresponding gradient values and green areas become non-numeric (NaN). The 'Left filtered boundary' is then similarly applied to the outcome of this operation, resulting in the 'Final threshold with Left' map shown in the top panel, which represents the completed vehicle field with integrated gradient thresholds. This sequence of operations demonstrates the stepwise refinement of the boundary maps through gradient integration, ultimately yielding a comprehensive field suitable for further analysis.

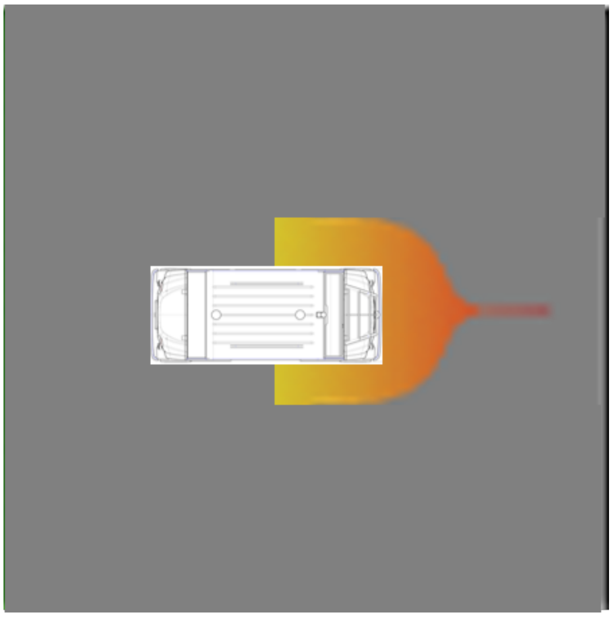


Figure 5.10: The image presents a schematic representation of a vehicle's forward interaction field, visualized from a top-down perspective. The graphic showcases an overlay of the vehicle's outline on a contrasting background, with the interaction field delineated in a distinct orange hue emanating from the front of the vehicle.

engagement intensity with the vehicle. By accumulating this gradient with data indicative of pedestrian unawareness, one can find out the spatial distribution of interaction values. This fusion of datasets allows for an analysis of the interaction intensities, providing a framework to predict potential risk areas based on the pedestrian's level of unawareness in relation to the vehicle's operational domain. Figure 5.10 illustrates the interaction field at the front of the vehicle, with a top-down view of the vehicle provided to visualize its appearance.

5.4.2 Pedestrian Interaction Field Aspect

The vehicle interaction field, a concept crucial for comprehending pedestrian unawareness and associated risk levels, necessitates an analogous representation of pedestrians. While the implementation methodology mirrors the methodology used in the pedestrian interaction field, it diverges in terms of the geometric shape utilized for representation.

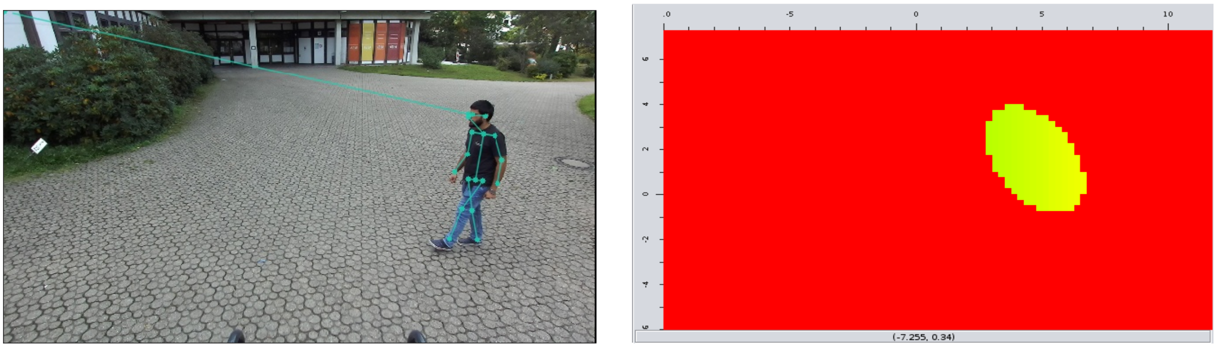
As shown in Figure 5.5 that the represented as ellipse, this specific geometric choice is instrumental for accurately modeling the spatial dimensions and movement dynamics associated with pedestrians. As demonstrated in Figure 5.11, pedestrians are represented by an ellipse, shaded in a yellow-greenish hue. This visual representation is significant for several reasons.

Firstly, the origin point (0, 0) of the coordinate system is aligned with the center of the vehicle. Consequently, the center of the ellipsoid, representing the pedestrian, is placed in accordance with the pedestrian's actual position relative to the vehicle. This alignment is critical for a precise spatial representation of the pedestrian in relation to the vehicle, enhancing the model's accuracy in assessing interactions and potential collision risks. Additionally, the orientation of the ellipsoid's elongated axis is carefully chosen to indicate the pedestrian's direction of movement. This directional aspect is not merely a visual attribute but also plays a pivotal role in the analytical model. It further enables the prediction of the pedestrian's trajectory and enhances the assessment of potential risks arising from their movement patterns. Algorithm 5.2 gives the process of creating ellipse on a 2D grid map.

Algorithm 5.2: Creating Ellipse function on a Grid

in_yaw, *in_center*, *par_cell_size* *in_radius_x*, *in_radius_y*, *in_off_set*, *in_gradient*;
Aspect grid with ellipsoid influence values;
$\theta \leftarrow$ GetYawValue(*in_yaw*);
$\cos_\theta \leftarrow \cos(\theta)$;
$\sin_\theta \leftarrow \sin(\theta)$;
$rot\_matrix \leftarrow \begin{bmatrix} \cos_\theta & -\sin_\theta \\ \sin_\theta & \cos_\theta \end{bmatrix}$;
$origin \leftarrow$ GetPointer (*in_center*);
$c_x \leftarrow origin.X()/$ GetValue (*par_cell_size*);
$c_y \leftarrow origin.Y()/$ GetValue(*par_cell_size*);
for $x \leftarrow -radius\_$ ***to*** $radius\_$ **do**
  for $y \leftarrow -radius\_$ ***to*** $radius\_$ **do**
    $rel_x \leftarrow x - c_x$;
    $rel_y \leftarrow y - c_y$;
    $new_x \leftarrow rot\_matrix[0][0] \cdot rel_x + rot\_matrix[0][1] \cdot rel_y$;
    $new_y \leftarrow rot\_matrix[1][0] \cdot rel_x + rot\_matrix[1][1] \cdot rel_y$;
    $new_x \leftarrow new_x + c_x$;
    $new_y \leftarrow new_y + c_y$;
    $ellipse\_value \leftarrow 1 + \left(\frac{new_x - c_x}{GetValue(in\_radius\_x)}\right)^2 + \left(\frac{new_y - c_y}{GetValue(in\_radius\_y)}\right)^2$;
    SetCellValue(*aspect_*, x, y, *in_off_set* + *ellipse_value* . *in_gradient*) ;
  end
end
return aspect_grid

Figure 5.11 shows the ellipse representing a pedestrian visualization for this thesis. The model enhances the understanding of pedestrian awareness by integrating a dynamic intensity scale into the ellipse on the network map, which varies directly with the level of pedestrian awareness. Decreased awareness leads to an increase in the intensity of grid values within the ellipse. The methodology for calculating pedestrian unawareness based on their activity is described in Chapter 4. This method is crucial to accurately represent the awareness of pedestrians in the context of the interactions with the Autobus. Figure 5.12 provides a zoomed-in view of the pedestrian's representation on the grid map and compares the state of unawareness in the range of zero and one. A value of zero does not completely eliminate the presence of pedestrians in the visualization. The grid cells change color depending on the pedestrian's level of awareness and change from red to green in a spectrum corresponding to the grid value from 0 to 1. In this case, the red represents a lower value of the grid and a higher pedestrian awareness, while the green represents a higher value of the grid and thus a lower awareness. This color scheme is not only a visual aid, but it is also crucial for the model's ability to quantitatively evaluate and visually represent variations in pedestrian awareness. This strategy effectively identified

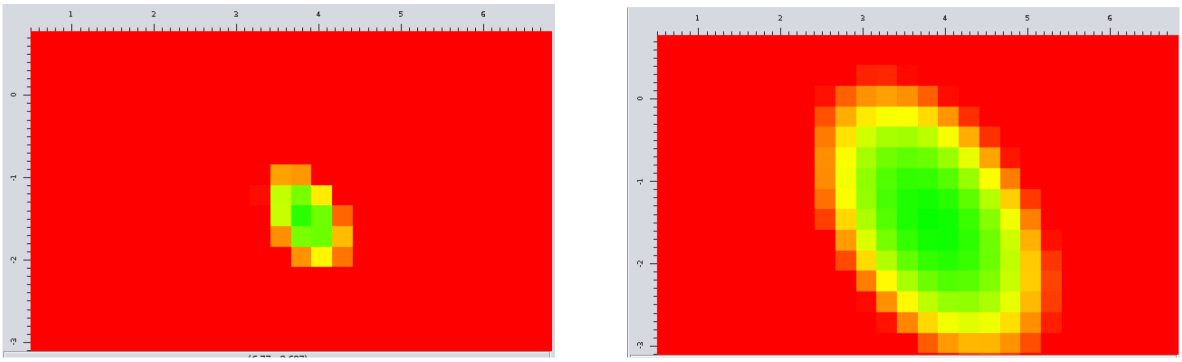


(a) Image of a pedestrian captured from front camera, walking in front the Autobus.

(b) The spacial occupation of the pedestrian in the grid map from top view.

Figure 5.11: Representation of pedestrian space on a grid map. The grid map depicts the spatial occupancy of a pedestrian within a defined area, with the pedestrian's presence illustrated by an ellipsoid in a yellow-green hue. The red background signifies unoccupied space on the grid. This graphical abstraction serves as a quantification tool for analyzing pedestrian dynamics and spatial interaction on the grid, providing a visual approximation of the physical space influenced by a pedestrian's presence.

areas where pedestrians may be more vulnerable or at higher risk, thereby improving understanding of pedestrian behaviour and its impact on safety in vehicle-pedestrian interaction.



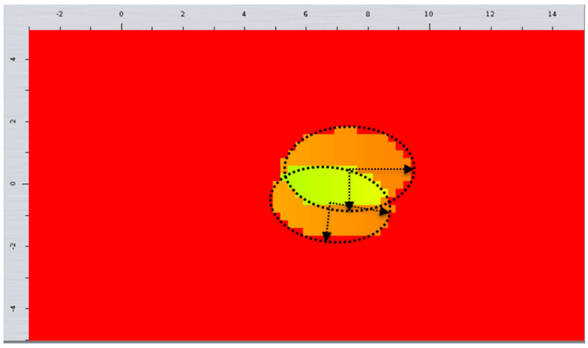
(a) Visualization of pedestrian awareness with a calibrated unawareness value of 0, depicted as an isolated area of heightened intensity amidst a red background indicating maximum awareness.

(b) Visualization of pedestrian awareness with a calibrated unawareness value of 1, demonstrating a gradient of decreasing awareness from green to red within the ellipsoid representation on the grid map.

Figure 5.12: The provided figures present magnified interpretation of Figure 5.5, each corresponding to distinct values of pedestrian unawareness on the same scale.

In order to depict the collective awareness of multiple pedestrians, the respective values assigned to each pedestrian are aggregated. This process is exemplified in Figure 5.13, which illustrates two pedestrians with reduced awareness in the vicinity of the Autobus. The lack of awareness among these pedestrians is indicated by elevated intensity values, which, when coinciding in space, intensifies cumulatively. Consequently, the areas of overlap are visually represented by a more saturated color, denoting increased values. In

(a) Image of two pedestrians captured from front camera, walking in front the Autobus.



(b) The spacial occupation of both pedestrians in the grid map from top view.

Figure 5.13: Composite Grid Map Illustrating Multiple Pedestrians. The grid map visualizes two overlapping pedestrian spaces, each modeled as an ellipse with additive value layers, against an unoccupied red background. The overlay of the ellipses generates a gradient of values where the spatial extents intersect. The central yellow region indicates the highest value concentration, signifying the greatest degree of pedestrian overlap. This composite representation is pivotal for understanding the collective spatial impact of pedestrian movements and for applications requiring the analysis of pedestrian density and interaction.

instances of such overlap, the awareness values are summed, adopting the maximum value from the individual contributions as the representative metric. This methodology ensures that the heightened risk associated with reduced awareness is accurately captured and emphasized in the model's output for multiple pedestrian groups.

5.4.3 Merged Vehicle and Pedestrian Interaction Field Aspect

The analysis of merged values within the interaction fields is crucial for understanding how pedestrians interact with vehicles, such as an Autobus. Figure 5.14 provides a clear depiction of this interaction, illustrating a pedestrian entering the Autobus' interaction field. This scenario is particularly important for evaluating the dynamic relationship between the pedestrian and Autobus.

Upon the pedestrian's entry into the interaction field, the model immediately begins to detect and analyze their awareness levels. This process involves the assessment of the pedestrian's specific location within the field and the corresponding awareness values at that point. The integration of these values is essential for calculating the overall risk associated with the pedestrian's presence in the vicinity of the vehicle.

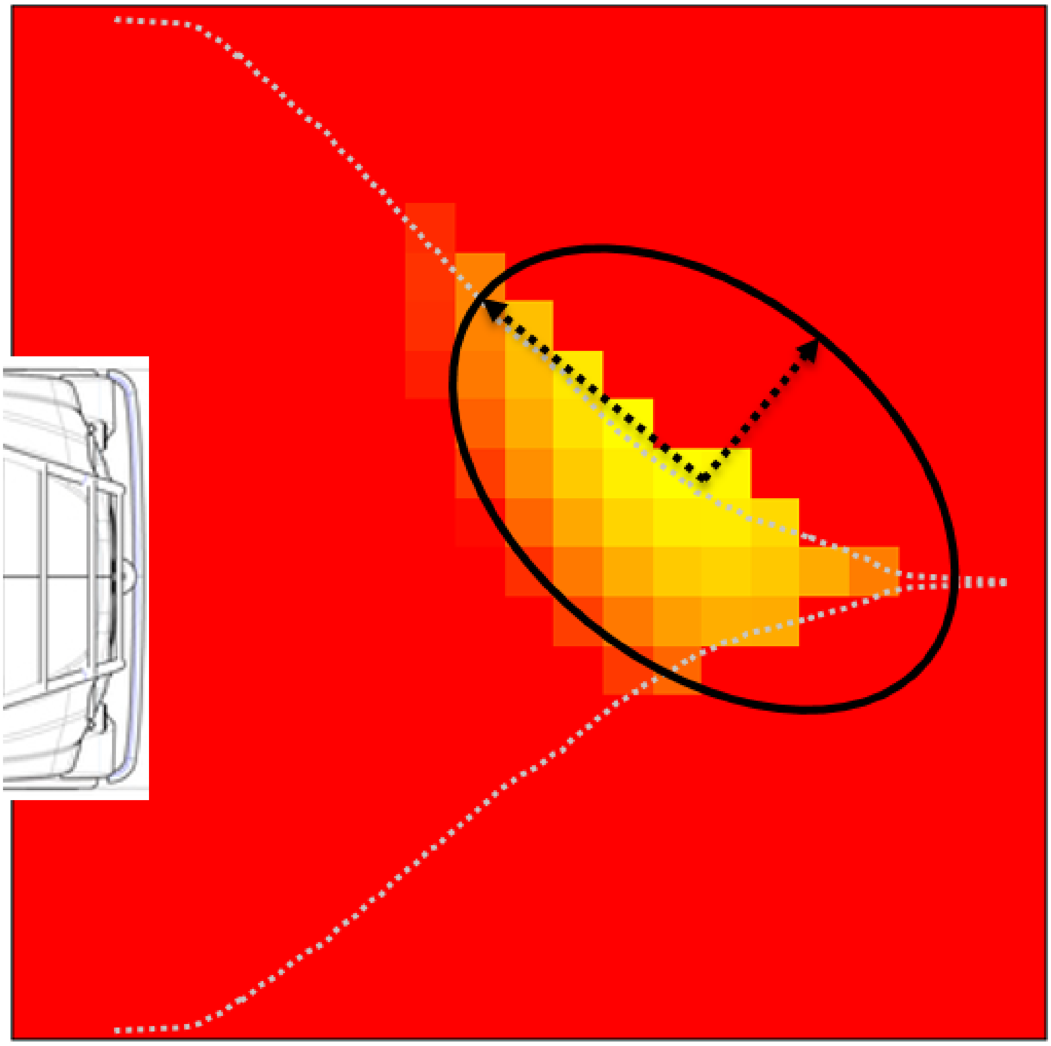


Figure 5.14: Spatial distribution of risk assessment in the interaction field between a pedestrian and the Autobus. The pedestrian's location is indicated by a black elliptical outline on the grid map, and the vehicle interaction field shape is shown with gray dotted line. To the left, a schematic representation of the Autobus is superimposed to orient the viewer to the vehicle's position and direction. The gradient from red to yellow on the heat map signifies a graded increase in perceived risk, with yellow indicating higher concentrations of merged awareness values. These values are computed to evaluate the likelihood and potential severity of an encounter between the pedestrian and the vehicle. This visualization facilitates an understanding of how a pedestrian's position within the interaction field affects the Autobus' risk assessment.

An illustrative example of how grid values increase in response to the pedestrian's movement within the interaction field is presented in Figure 5.15. This figure demonstrates the gradual escalation of risk values as the pedestrian moves deeper into the interaction field. The increasing grid values are indicative of the rising level of potential risk or conflict between the pedestrian and the Autobus, making it a critical aspect of the safety analysis.

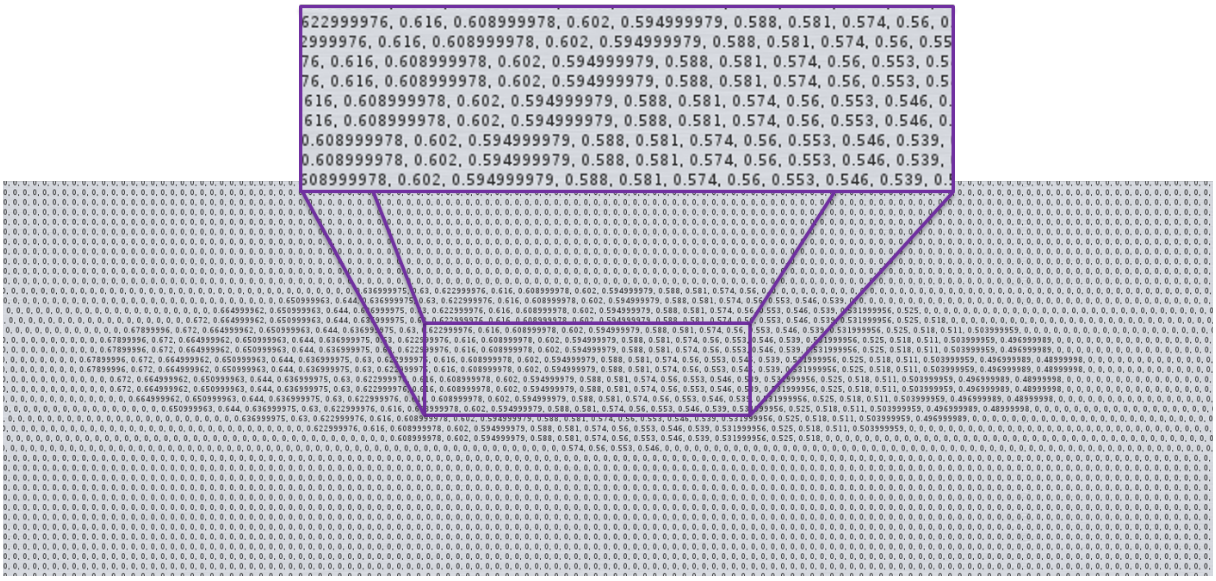


Figure 5.15: This figure presents a quantitative representation of risk values associated with a pedestrian's movement within the interaction field of an Autobus. The data matrix indicates the progressive increase in grid values corresponding to the pedestrian's deeper penetration into the field, which translates to a higher potential for risk or conflict. The gradation from lower to higher values, particularly noticeable in the central columns (from right to left), reflects the criticality of location-based risk assessment for the Autobus stationed to be on the left side. A zoomed version is extracted in purple box.

Associated pedestrian risk: The "risk" parameter from the overall map is calculated by summing up the ellipsoid grid values for a pedestrian and the vehicle interaction field grid values. As the grid is represented as a two-dimensional matrix where each entry corresponds to a grid value at a specific position, the formula for the sum within the overlapped area of the pedestrian's ellipsoid and the vehicle's interaction field can be expressed as follows:

Let E be the set of grid positions occupied by the pedestrian's ellipsoid and V be the set of grid positions occupied by the vehicle's interaction field.

$$S = \sum_{(i,j)\in E\cap V} (E_{ij} + V_{ij})$$

Where:

- S is the total sum of the overlapped grid values.
- E_{ij} is the grid value at position (i, j) due to the pedestrian's ellipsoid.
- V_{ij} is the grid value at position (i, j) due to the vehicle's interaction field.
- $E \cap V$ represents the set of grid positions where the ellipsoid and interaction field overlap.

Since the ellipsoid grid values are designed to represent the level of unawareness of the pedestrian and the vehicle interaction field grid values represent potential risk, then the sum S provides a composite value representing the combined effect of pedestrian awareness and vehicle interaction at each overlapped position. This sum can be used to assess the risk level in the interaction zone between the pedestrian and the vehicle based on the threshold value set for the different interaction fields.

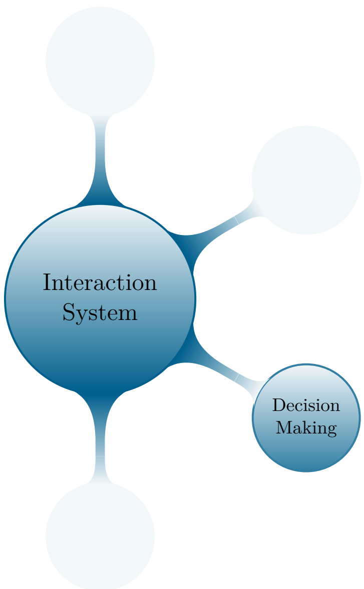
In addition, the aggregated grid values are transmitted to the "Decision Making" process described in Chapter 6. This phase analyzes the integrated data to determine whether escape maneuvers are feasible or whether the "interaction module" activation is necessary. Decision-making protocols integrate these assessments into comprehensive responses, strengthening safety protocols in scenarios where active interactions or prevention measures are required.

Furthermore, the calculated risk value is fed into the "interaction module", as delineated in the Chapter 7. This module uses a threshold mechanism to classify the risk value into three different types of interaction: ambient, direct or risk-based. The interaction classification determines the subsequent activation of the appropriate modules within the interaction module. This stratification is crucial because it dictates the module's reaction to different levels of pedestrian-vehicle interaction and ensures an adaptive and contextual approach to managing potential hazards.

In conclusion, design perspective gives priority to an effective risk assessment using interaction fields adapted to the vehicle's environment. The shape and direction of these fields are strategically designed to be aligned significantly with the travel direction of the Autobus. This alignment reduces the likelihood of engaging in inappropriate interactions, especially in the peripheral parts of the vehicle where these interactions are less critical. On the other hand, the pedestrian interaction area is oriented in accordance with the pedestrian's walking path. This orientation is crucial to the creation of a safety buffer that allows the Autobus to have sufficient time and space to begin communication or to take the necessary steps to react to the movements of the pedestrians.

Furthermore, the adaptability of the design is highlighted by its customizable features. The size and intensity of the interaction fields can be modified to meet specific environmental conditions, increasing the versatility of the system. This customization is essential to ensure that the interaction fields are optimally configured for different scenarios, be it a crowded urban street or a more isolated suburban area.

6. Decision Making



Decision making is a crucial aspect of navigation of driverless minibuses. This is the key to ensure that these vehicles are safe, efficient and reliable. On highways, everything is structured and predictable, and decision-making algorithms can be very confident in predictability. However, as mentioned in the previous chapter (Chapter 5), the driving in pedestrian areas becomes more difficult with these driverless minibuses; such areas are characterized by their unpredictability and the diverse range of variables that the vehicle must consider in real-time. This uncertainty means that the decision-making framework for driverless minibuses must be adaptable. Unlike the relatively predictable pattern of vehicular traffic in marked streets, pedestrian movements are irregular and spontaneous in pedestrian zones. Therefore, the driverless minibuses must not only react to existing behavior, but also anticipate what pedestrians might do next. In order to do so, a driverless minibus must have a high-level perception system to detect and predict the behavior of pedestrians accurately. They also need an interface to interact with pedestrians and adjust their navigation strategies to ensure smooth and safe travel.

This chapter examines the decision-making processes in the context of pedestrian transportation in shared spaces and is an important element of the smart interaction strategy used by Autobus. The Autobus decision-making mechanism is a complex combination of navigation and interaction strategies designed to emulate and strengthen human decision-making capabilities in dynamic and often unpredictable environments. These features are crucial to the core functionality of the Autobus, which allows the vehicle to interpret and react appropriately to the environment. Given the thesis's emphasis on pedestrian interaction, the decision-making module combines both navigation and interactive elements by using audio and visual signals to communicate with pedestrians. This integrated approach is essential to manage the interaction of vehicles in scenarios where traditional evasion tactics are not practical due to space constraints. The process begins with analyzing pedestrian activity (see Chapter 4) and representing pedestrian awareness on the map (Chapter 5). These insights are crucial to the formulation of effective strategies, to guide the Autobus in determining whether to navigate around or interact with pedestrians blocking its paths. The risk assessment of pedestrian proximity is based on the interactive field map defined in Chapter 5 which helps to determine the most suitable navigation and interaction strategy. This decision-making framework emphasizes the main objective of the thesis: improving the interaction between the Autobus and pedestrians, and ensuring safe and efficient navigation within pedestrian areas. This chapter begins with an examination of the design principles underlying decision-making processes. It provides a detailed examination of various factors influencing pedestrian behaviour. Then three planning parameters are explained to develop effective strategies. The discussion then turns to the formulation of an interactive navigation strategy for Autobus. The implementation of these strategies is detailed at the end of this chapter.

6.1 Design Principle

The navigation of pedestrian areas is a unique challenge for the Autobus, and pedestrians' very dynamic behaviour introduce a significant level of uncertainty in predicting their intentions. Unlike other environments, pedestrian areas are characterized by spontaneous and unpredictable movements, and can vary greatly at frequent intervals. This variability requires a sophisticated approach to intention prediction and behavior analysis to ensure safe and effective navigation. In order to understand and effectively manage this complexity, it is essential to identify and analyse factors influencing the behaviour and interactions of pedestrians and the Autobus. These factors are classified and illustrated in a series of graphs, ranging from Figures 6.1 to 6.4. In these graphs, the largest circle represents the primary factor affecting pedestrian behavior in pedestrian areas. The smaller and interconnected circles around this main factor represent various parameters that are influenced by this main factor.

The considerations in Figures 6.1 to 6.4 must be envisaged during the design process. For simplicity, the sub-categories of the different factors have been discretized. There must exist approaches which should consider all these parameters and automate the entire driving process and interaction, accordingly. Due to such complexity, driving algorithms for pedestrian zone is not only enhanced by a single module but consists of different components in a navigation framework to achieve the desired goal.

6.1.1 Planning Parameters

Generally, interaction is intertwined to so many aspects between different agents that there is a need to specify its application and purpose in order to implement a strategy. Looking at the overall application of this work and the relevant literature, three determinants are chosen:

- **Safety:** Safety is essential and is an important factor for the community to accept driverless minibuses, as highlighted in [Rahman 21]. In the autonomous navigation of pedestrian areas, safety focuses primarily on alleviating the risk of collisions with pedestrians. This requires a proactive approach in which the vehicle's decision-making

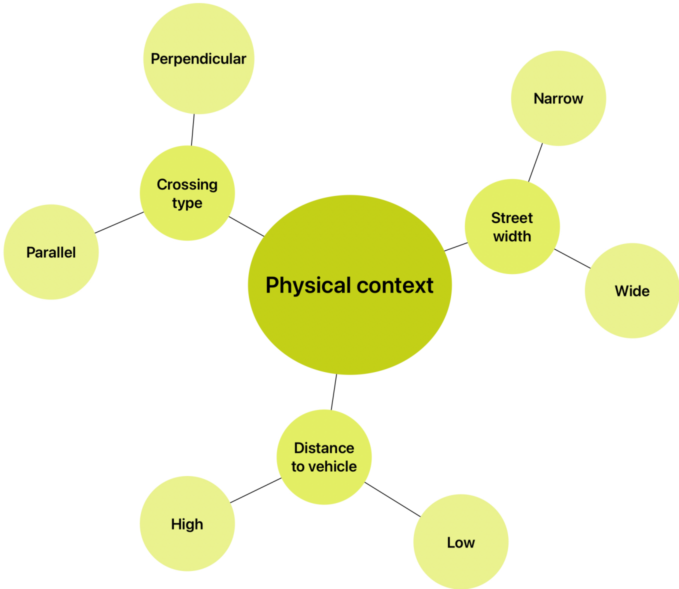


Figure 6.1: Hierarchical influence diagram of pedestrian physical context factors for Autobus Navigation. The central node represents the overarching 'Physical context' of pedestrian environments. The secondary nodes illustrate the critical factors affecting this context, including 'Crossing type' (with 'Perpendicular' and 'Parallel' as sub-factors), 'Street width' (with 'Narrow' and 'Wide' as sub-factors), and 'Distance to vehicle' (with 'High' and 'Low' as sub-factors). This diagram serves as a conceptual map to understand how different physical attributes of a pedestrian zone can influence the decision-making processes of the Autobus. Physical context defines the factors related to the environment such as crossing types, street width, and distance to vehicle. These are sub-categorized based on the type of environment the vehicle is driving through.

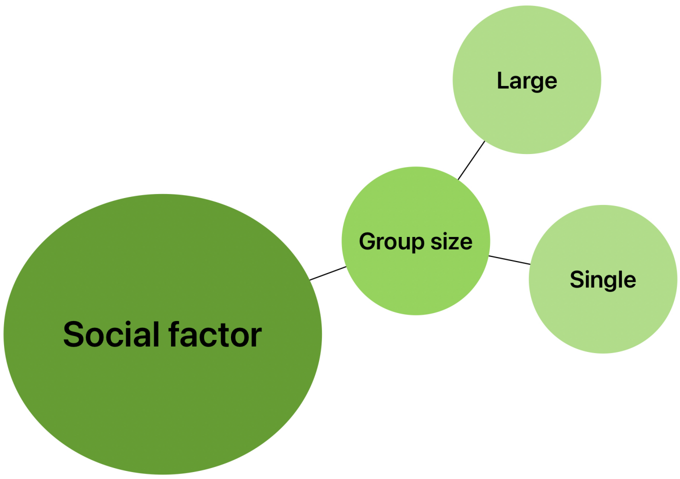


Figure 6.2: Schematic representation of the impact of group size as a social factor on pedestrian dynamics. The central node signifies the 'Social factor' affecting pedestrian behavior, with 'Group size' as a key component. The secondary nodes indicate the two sub-categories of group size: 'Large' groups and individuals ('Single'), each exerting a distinct influence on pedestrian movement patterns and interactions in the context of Autobus navigation. Social factor are dependent on the crowd size. It mainly refers to group size. It should be taken into consideration for some of the driving behaviors such as evasion.

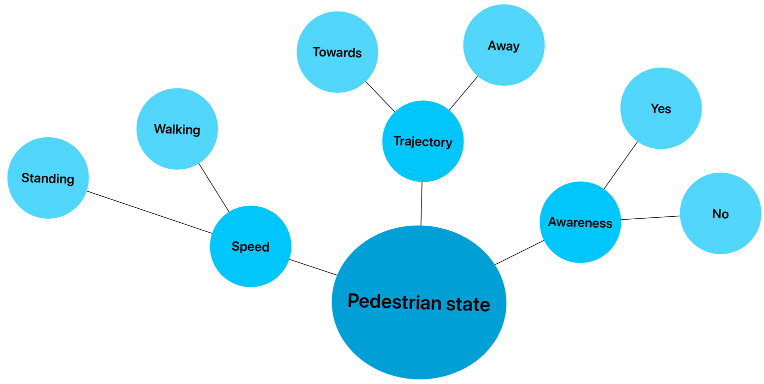


Figure 6.3: Conceptual framework of pedestrian state variables affecting the decision-making of Autobus. The core node labeled 'Pedestrian state' connects to critical state descriptors such as 'Speed,' 'Trajectory,' and 'Awareness.' The 'Speed' category distinguishes between 'Walking' and 'Standing,' while 'Trajectory' divides into 'Towards' and 'Away,' indicating the pedestrian's movement relative to the vehicle. The 'Awareness' node bifurcates into 'Yes' and 'No,' reflecting the pedestrian's consciousness of the vehicle's presence. This framework outlines the fundamental aspects of pedestrian behaviors that Autobus must interpret to navigate safely and interact effectively with humans. Pedestrian state is relevant to pedestrian behavior. For this thesis speed, trajectory, and awareness based on activity (Chapter 4) are considered. The speed is further divided into standing and walking. Due to short interaction intervals, the speed of walking is not considered. Trajectory considers the direction of the pedestrian and attention is the awareness of the pedestrian towards the AV.

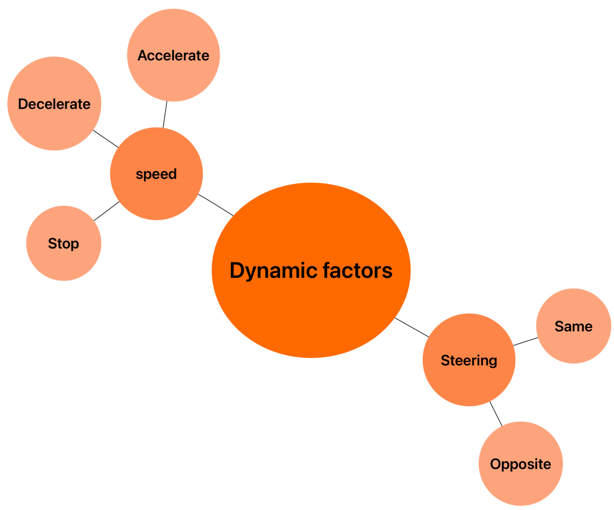


Figure 6.4: Diagram of dynamic factors in Autobus Control. The central, large node represents 'Dynamic factors' that are critical to the operation of Autobus. This is linked to two key operational elements: 'Speed' and 'Steering.' The 'Speed' node further connects to its operational states: 'Accelerate,' 'Decelerate,' and 'Stop.' Similarly, 'Steering' is associated with directional choices of 'Same' and 'Opposite,' denoting the vehicle's steering decisions in response to obstacles or the driving environment. This schematic serves to illustrate the variable elements of vehicle dynamics that are continuously adjusted during autonomous navigation.

algorithms prioritize pedestrian safety. To this end, vehicles must engage in clear and comprehensible interaction with pedestrians. This means effectively signaling intentions, such as stopping or delivering its intentions to avoid potential risks. Safety is an important planning parameter, and it is essential that the Autobus not only detects and responds to pedestrians' immediate movements, but also anticipates and adjusts its course accordingly. This double approach ensures higher safety by reducing the likelihood of accidents and improving vehicle predictability, promoting pedestrian trust and facilitating more smooth integration of Autobus into pedestrian-intensive environments. Based on the risk assessment provided by the interaction field analysis (referenced in Chapter 5, it is preferred to ensure safe navigation.

- **Efficiency:** The main purpose of driverless minibuses such as the Autobus is to transport people through densely populated areas with pedestrians. This scenario presents a unique challenge, as the passengers on the vehicles and the pedestrians who pass share a common goal of arriving at their destination quickly. Because Autobus is designed to navigate a normally wakable path, their interest is based on their ability to provide a cost-effective alternative to walking. As a result, efficiency becomes an essential planning parameter in the decision-making process of the operation of Autobus in pedestrian zones.

  In this respect, efficiency is twofold: not only to minimize travel time, but also to optimize routes and speeds to ensure timely arrival without jeopardising safety. Autobus must intelligently calculate the most direct and less crowded routes taking into account pedestrian traffic patterns and possible delays. This approach requires a sophisticated navigation system capable of processing real-time data and making adaptive decisions to improve overall trip efficiency. In addition to focusing on efficiency, the goal of the Autobus is to provide a convincing and competitive means of transport that meets both the needs of passengers and pedestrians and ensures seamless integration into urban mobility ecosystems. The navigation system evaluates delays to determine whether escape maneuvers or interactions are prioritizing.

- **Trust:** Empirical research has shown that people are afraid and anxious about self-driving, and concerns are often centered on the possibility of system failure [Cugurullo 21]. Trust plays an important role in reducing these fears and bridges the inherent uncertainty of autonomous driving technology with the acceptance by the general public [Ribeiro 22, Lee 04, Ghazizadeh 12]. In the field of driverless minibuses that navigate pedestrian zones, trust is especially important to ensure that pedestrians are aware of the existence of vehicles and consider their existence. Trust transcends emotional meaning in the context of AV-pedestrian interaction and adopts a more symbolic form of communication. This symbolic trust is supported by clear and undoubtedly reliable signs that convey the intention of the minibus and the recognition of pedestrians. For example, by implementing a display that displays animations or uses verbal signals, an effective communication of vehicle's actions can be done in response to real-time situations, increasing the trust of pedestrians. Such measures ensure that trust is not only an abstract concept, but also a tangible element of vehicle operation protocols and an important component of the smooth integration of Autobus into pedestrian environments.

By integrating trust as the basis for planning parameters, Autobus' decision-making process becomes more sophisticated. This requires not only the recognition and reaction of the immediate environment, but also the psychological and emotional comfort of the people nearby. This can be done by displaying an animation or providing verbal indications adapted to a specific situation (Chapter 7). These factors improve the usability of the Autobus in crowded environments. The future acceptance of such driverless minibuses is expected to lead to the deployment of a fleet and thus improve connectivity. In the planning phase, these parameters will be considered depending on the interaction phase. This change is explained in the following subsections and explains the phase distribution of the interaction system in detail.

6.2 Decision-Making with Clothoid-Inspired Tentacle and Multi-Attribute Environmental Analysis

The tentacle-based approach relies on a map to derive weights essential for formulating navigation strategies. Consequently, understanding the operational dynamics of tentacle-based navigation necessitates a preliminary elucidation of the map's structural configuration and its constituent features.

Strategic Exclusion and Specialized Mapping in the Multi Feature Map Framework

Multifunctional mapping has changed the way autonomous system navigates unstructured environments such as pedestrian zones. This method combines various types of data to improve navigation, such as simple obstacle data, elevation data and segmented paths across different terrains. At the heart of this system are the ground elevation data that form the map's "base surface". This foundational layer is necessary to build maps and add additional data layers. These additional layers make the map richer and provide a more detailed and accurate view of the terrain. This technology provides a complete multi-dimensional landscape picture that is crucial for driving complex off-road areas. The data within these maps are categorized into four distinct types: ground elevation data, ground classification data, obstacle elevation data, and obstacle classification data. This detailed breakdown helps in understanding the terrain and obstacles within. Such detailed classification is crucial in an off-road environment where navigation is more difficult than in structured areas such as roads.

The main aspects of the multi-functional map system are the data interface and storage mechanism. The system uses independent data input and retrieval interfaces and adheres to the design of a perception-based design proposed by Schäfer et. al. [Schafer 08]. In this design, the data is stored in the map at the level of the individual cell. Each cell contains a list of data that contains all known information about the specific cell, thereby creating a complete data archive for each segment of the map. An example can be seen in Figure 6.5. One of the most important results of this data integration is the calculation of navigation values for each cell of the map. This value lies in the spectrum from 0 to 1, which indicates the cell's transferability for the vehicle.

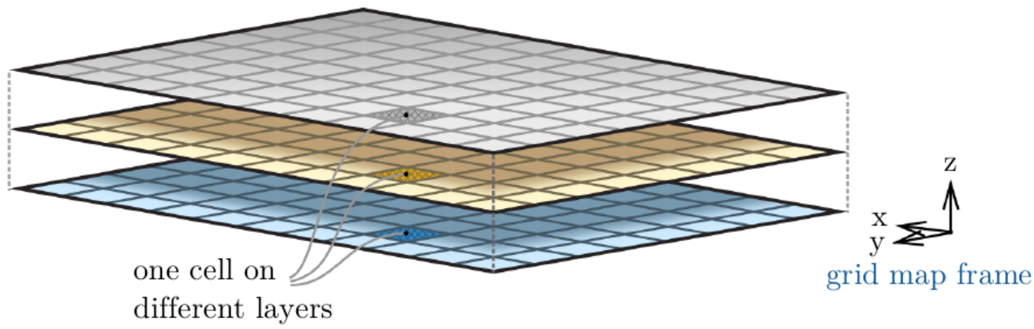


Figure 6.5: The figure depicts a composite structure of a multi-layered feature map, essential for advanced robotic navigation in unstructured environments. The highlighted cell spans across multiple informational layers, each representing a different type of data—ground elevation, such as ground classification, obstacle elevation, and obstacle classification. This multi-dimensional approach enriches the map's texture and provides a nuanced landscape representation, offering a detailed understanding of the terrain and potential obstacles for optimized path finding

In the methods discussed in this thesis, pedestrians are deliberately left out of the primary obstacle map, which is part of the framework of multifunctional maps. This is a deliberate design choice that is performed during the characteristic classification stage, as described in the section that explains interaction fields (Chapter 5). In order to address this exclusion, the different categories of the multi-functional map are specifically focused on pedestrian awareness, as shown in Figure 6.6. This special category is essential to improve the ability of the navigation system to recognize and react to pedestrians in the environment without observing them as obstacles.

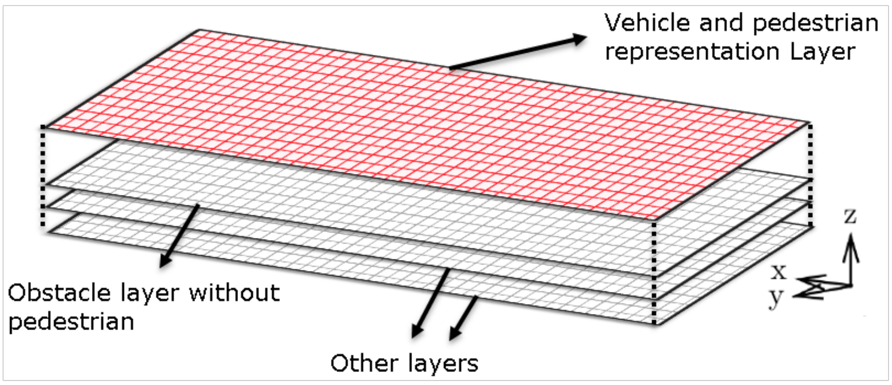
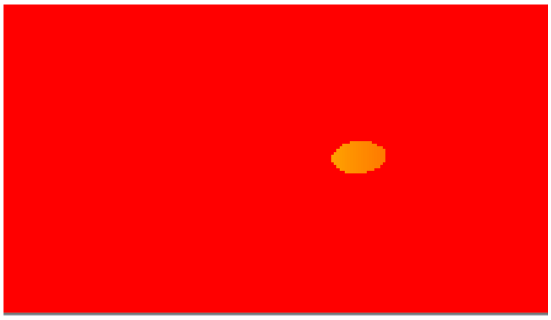


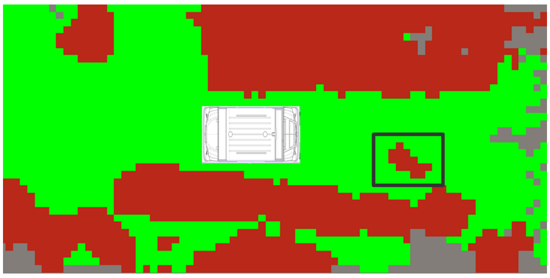
Figure 6.6: The figure depicts an additional layer (in red) for pedestrian awareness which are removed from normal obstacle layer in multi feature map framework.

Figure 6.7 illustrates an instance where a pedestrian engaged in a phone conversation is detected in the vicinity of an Autobus. The first part of the figure, denoted as (a), captures the visual from a camera where the pedestrian's skeletal structure is pinpointed. Subsequently, part (b) of the figure displays the map that exclusively represents the pedestrian. In order to evaluate the accuracy, two variations of obstacle maps are demonstrated. The first variant, marked with a black square in part (c), incorporates the pedestrian as an obstacle. In contrast, the second variant, shown in part (d), depicts the obstacle map after the pedestrian has been removed through aspect map, also highlighted with a black square for reference.

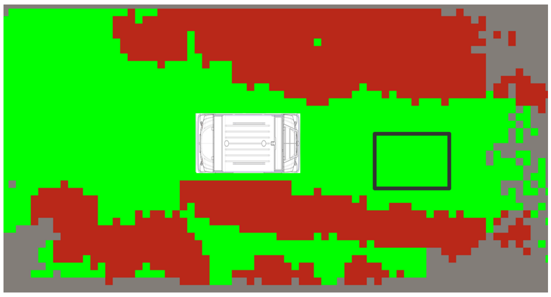
(a) Camera image with pedestrian skeleton detection: The visual input captures a pedestrian, and the processing algorithm outlines the human figure's skeletal structure for identification and tracking.



(b) Pedestrian-Only Map Representation: This high-contrast map isolates the pedestrian from the background, presenting a focused visual representation used for detailed analysis.



(c) Obstacle Map Including the Pedestrian: The comprehensive obstacle map shows various elements with the pedestrian marked within a black box, indicating the presence of a pedestrian as a navigational consideration.



(d) Obstacle Map with Pedestrian Excluded: Here, the obstacle map is displayed after the pedestrian has been removed via aspect map, demonstrating the map's adaptive feature for scenarios excluding pedestrians.

Figure 6.7: Figure illustrates the process and results of a pedestrian detection and exclusion algorithm in Autobus navigation. The system uses a Multi Feature Map approach to separate pedestrian features from other obstacles (as demonstrated in Figure 6.6) during the classification phase for enhanced navigation responsiveness.

By creating separate pedestrian categories, navigation values can be specifically designed to take pedestrian awareness into account. These values take into account factors such as location, how people move and the density of crowded areas. This novel approach ensures that the navigation system responds to the dynamic and often unpredictable movement of pedestrians in a practical way and not just avoiding them as regular obstacles.

Integrating Dynamic Clothoid Tentacles in Multi Feature Map

Advanced technical navigation frameworks based on the principle of *Multi Feature Map* are significant advances in robotic navigation, especially in complex and unstructured environments. This framework is distinguished by the integration of advanced spatial awareness and reactive decision-making mechanisms, which are essential for safe and efficient navigation.

This research centers on the innovative use of clothoid-based tentacles, which are integral to robotic navigation. A clothoid is a curve that changes linearly with its arc length, so that it is perfect for smooth transitions between different curvature values for kinematics of a vehicle. These tentacles are not physical structures, but abstract structures, used for road planning and environmental analysis. Each tentacle represents a possible navigational path and projected on the basis of the current position and direction of the robot. This method allows dynamic path planning, which allows robots to adapt to their surroundings and avoid obstacles. The path of these tentacles is evaluated in real time, taking into account factors such as proximity to objects, terrain slope and overall traversability of the path.

Algorithm 6.1: Clothoid Path Computation for Navigation

in_start_curvature, in_end_curvature, in_path_length, aspect_grid
$\Delta\kappa \leftarrow in\_end\_curvature - in\_start\_curvature$
$a \leftarrow \sqrt{\frac{2 \times in\_path\_length}{|\Delta\kappa|}}$
$s_{start} \leftarrow \frac{in\_start\_curvature \times a^2}{2}$
$s_{end} \leftarrow \frac{in\_end\_curvature \times a^2}{2}$
$\theta_{rotation} \leftarrow \text{GetCurrentRotationAngle}()$
for $s \leftarrow s_{start}$ ***to*** s_{end} **do**
    $x \leftarrow a \cdot \text{FresnelCos}\left(\frac{s}{a}\right)$
    $y \leftarrow a \cdot \text{FresnelSin}\left(\frac{s}{a}\right)$
    $\theta \leftarrow \left(\frac{s}{a}\right)^2 - \theta_{initial}$
    $direction\_vector \leftarrow \text{ComputeDirectionVector}(x, y)$
    $\text{Normalize}(direction\_vector)$
    $\text{Rotate}(direction\_vector, \theta_{rotation})$
    $sample\_point.position \leftarrow (x, y)$
    $sample\_point.direction \leftarrow direction\_vector$
    $\text{StoreSamplePoint}(sample\_point)$
end
return *aspect_grid*

The core of the algorithm described in Algorithm 6.1 hinges on calculating the parameters that define the shape of a clothoid. A clothoid, or Cornu spiral, is characterized by a curvature that changes linearly with the arc length. This property makes clothoids particularly useful in applications like road design, where smooth transitions between different curvatures are required.

The primary equations used to define the Cartesian coordinates (x, y) of a clothoid in terms of its arc length s are given by:

$$x(s) = a \cdot \text{FresnelC}\left(\frac{s}{\sqrt{\pi a}}\right),$$
$$y(s) = a \cdot \text{FresnelS}\left(\frac{s}{\sqrt{\pi a}}\right)$$

Here, s is the arc length along the clothoid, and a is a scaling factor related to the rate of curvature change. The functions FresnelC and FresnelS are the Fresnel cosine and sine integrals, respectively, defined as follows:

$$\text{FresnelC}(u) = \int_0^u \cos\left(\frac{\pi t^2}{2}\right) dt,$$
$$\text{FresnelS}(u) = \int_0^u \sin\left(\frac{\pi t^2}{2}\right) dt$$

These integrals are not the usual cosine and sine functions but special functions used in wave optics and related fields. They provide the Cartesian coordinates $x(s)$ and $y(s)$ of a point at arc length s along the clothoid.

In the implementation of the algorithm, the change in curvature ($\Delta\kappa$) is initially computed. Subsequently, the scaling factor a is determined using the path length and the absolute value of $\Delta\kappa$. The normalized distances along the clothoid from the start to the end (s_{start} and s_{end}) are then calculated using a and the respective curvatures at these points.

The algorithm constructs the clothoid by incrementally calculating the position of points along the path. For each point, the primary clothoid formula is used to compute the Cartesian coordinates based on the scaled arc length s and scaling factor a. It also computes the orientation for each segment and normalizes the direction vector to ensure a consistent tangent along the path. The computed points and directions are stored as pairs to represent the potential navigational paths, which are often referred to as "tentacles."

This dynamic construction allows for real-time adjustments to the path as the vehicle navigates, providing a mechanism that is responsive to changes in the environment. Such adaptability is crucial in complex and unpredictable settings, where continuous reassessment of trajectory is necessary for safe and efficient navigation.

The selection of the optimal tentacle, or navigational path, is based on a rigorous evaluation process. Each tentacle is scored using a set of predefined metrics, with the highest-scoring tentacle being chosen for navigation. This process involves calculating the mean navigation value across all the cells that a tentacle covers, effectively weighing the safety and feasibility of traversing that path. The chosen tentacle's curvature parameter is then fed into the robot's control system, guiding its movement along the selected path.

Let $T = \{t_1, t_2, \ldots, t_n\}$ be the set of all possible tentacles. For each tentacle t_i, calculate its score $S(t_i)$ based on predefined metrics:

$$S(t_i) = \frac{1}{N_{t_i}} \sum_{c \in C_{t_i}} M(c)$$

where:

- $S(t_i)$ is the score of tentacle t_i.
- N_{t_i} is the number of cells covered by tentacle t_i.
- C_{t_i} is the set of cells covered by tentacle t_i.
- $M(c)$ is the metric value for cell c, which evaluates the safety and feasibility of the cell for navigation.

The optimal tentacle t_{opt} is then chosen by:

$$t_{opt} = \arg\max_{t_i \in T} S(t_i)$$

Finally, the curvature parameter $\kappa_{t_{opt}}$ of the chosen tentacle t_{opt} is fed into the vehicle's control system:

$$\text{ControlSystem}(\kappa_{t_{opt}})$$

where ControlSystem represents the function of the robot's control system that processes the curvature parameter and guides the robot's movement along the path determined by t_{opt}.

A crucial aspect of this framework is the way in which it reacts and adapts. Instead of sticking to a fixed path, the system is always checking and adjusting its path by continually recalculating and reevaluating the tentacles. It can be reconfigured at a regular interval, i.e. respond quickly to changes around it. This is especially useful when the environment is changing rapidly, such as obstacles that appear suddenly or terrain conditions may rapidly change. As a result, it constantly adjusts its course to stay on track, making it super efficient in unpredictable situations.

The framework also extends to sophisticated assessment of tentacle pathways incorporating multifunctional content. Each cell of the map is assessed on the basis of several attributes such as the type of ground segment, obstacles, and sensory data quality. This comprehensive assessment ensures that the chosen path is not only free of direct physical obstructions, but is also optimally drivable depending on the robot's capabilities and environmental conditions.

6.3 Pedestrian-Aware Tentacles for Enhanced Decision-Making

Generally, a driverless minibus drives over a traversible path and avoids obstacle present on the path. As illustrated in chapter 3, pedestrian pathways can be considerably narrow. Driving Autobus with a single person on the pathway will impel the vehicle to slow down due to unavailability of evasion. In such case, a traditional navigation will make a safety slow down or completely stop. As stated earlier, humans have the aptitude to take responsibility and give way; therefore, the best strategy in this capacity is to inform them when needed. By doing so will reduce unnecessary breaking and slowdown behavior.

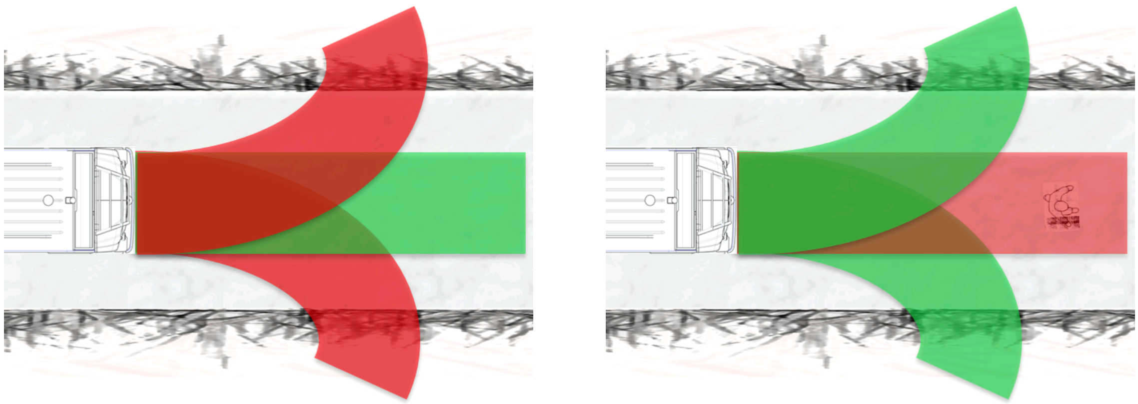
When removing pedestrians from obstacle maps, as shown in Figure 6.7(d), the detection of pedestrians in the environment becomes mandatory. As a result, the development of a set of tentacles specifically designed to assess the risks associated with pedestrians is necessary. To this end, additional pedestrian-oriented tentacles are generated using pedestrian maps to facilitate risk assessments. Pre-existing tentacles use the characteristic map to define driving areas and continue to calculate the best navigation path. These are now called feature tentacles throughout the thesis. The selection of pedestrian tentacles depends on the index of the highest value on the pedestrian map and guarantees the risk in the driving area. The choice between evasion maneuvers and interaction modules is influenced by a global risk assessment of both feature tentacles and pedestrian tentacles, as indicated by the feature tentacle index.

Let $W_{f,i}$ represent the weight assigned by the feature tentacle, and $W_{p,i}$ represent the weight assigned by the pedestrian tentacle. If i denotes the index of the tentacle (with $i =$ being the indexes of the feature and pedestrian tentacle), the aggregation of each tentacle index is given by the equation:

$$W_{\text{total},i} = W_{f,i} + W_{p,i} \tag{6.1}$$

Subsequent sections delineate three scenarios to elucidate the application of both tentacle types and the resultant decision-making processes of Autobus. While myriad situations may arise in pedestrian environments, varying according to factors discussed in subsection 6.1, it is beyond the scope of this thesis to catalogue every potential scenario. However, to validate the proposed methodology, an exemplar scenario involving a pedestrian under varied environmental conditions is presented.

- **Narrow pathway:** Since pedestrian zones include narrow pathways, where the evasion is not possible, an example is shown in figure 6.8. The scene is explained by using three tentacles. Figure 6.8a shows the feature tentacle for traversible path, where as figure 6.8b shows the pedestrian tentacle which avoids other obstacle and calculates the weights based on pedestrian awareness.



(a) Feature Tentacle **(b)** Pedestrian Tentacle

Figure 6.8: Narrow pathway tentacle visualization: Green color show high traversiblity and red color shows low traversibility.

When summing up the weights from both sets of corresponding tentacles as shown in figure 6.9, It is seen that all the corresponding tentacles have equal weights. The best index from feature tentacle is chosen (middle). Since it is not possible to drive due to pedestrian, interaction modules generates commands to make the pedestrian aware of the vehicle.

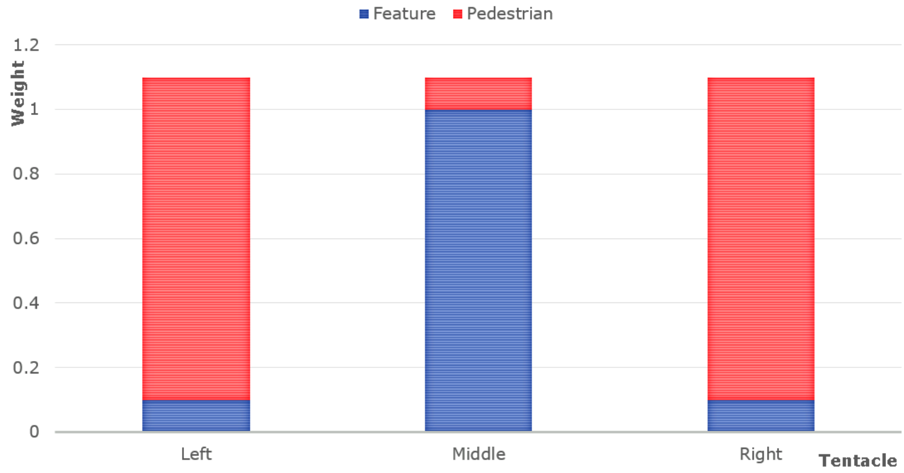
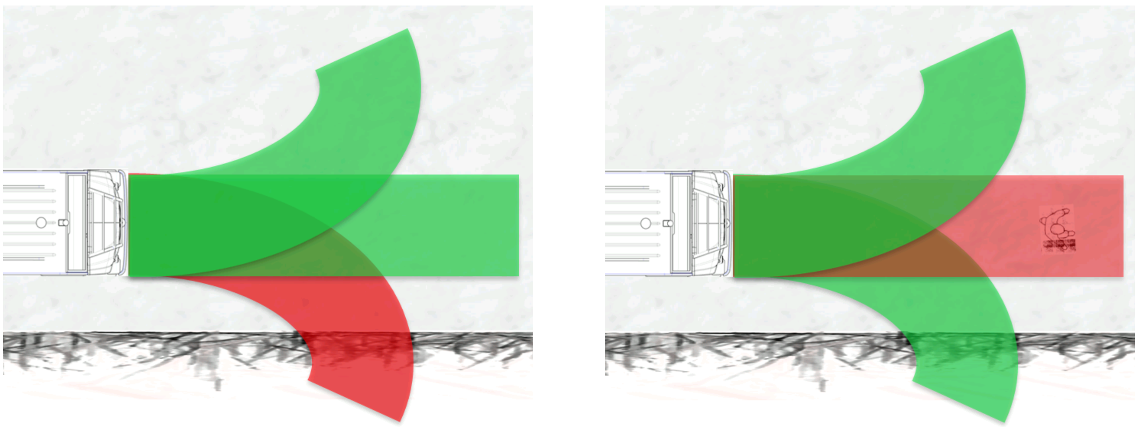


Figure 6.9: Narrow pathway: Sum of tentacle weights for tentacle shown in Figure 6.8.

- **Semi-narrow pathway:** Such situations allow for the possibility of evasion if not too crowded, meaning the AV can maneuver to one side of the pedestrian. In the specific case shown in figure 6.10, an unaware pedestrian is walking in front of the vehicle. The feature tentacle can drive straight and to the left, whereas the pedestrian tentacle has the lowest value in the forward direction due to the unaware pedestrian.



(a) Feature Tentacle (b) Pedestrian Tentacle

Figure 6.10: Semi-narrow pathway tentacle visualization: Green color show high traversiblity and red color shows low traversibility.

Observing the sum of weights in figure 6.11 for the semi-narrow pathway, the highest value is taken by the left most tentacle. For this case the left feature tentacle can be selected, and the vehicle can simply steer in the left direction. For this case the interaction is not necessary.

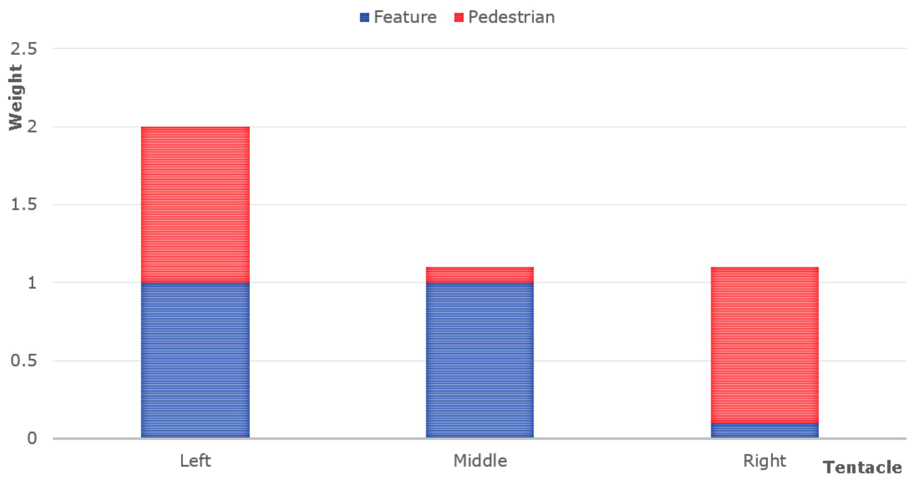
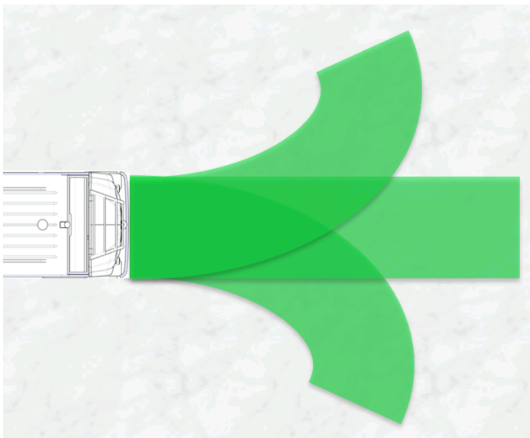
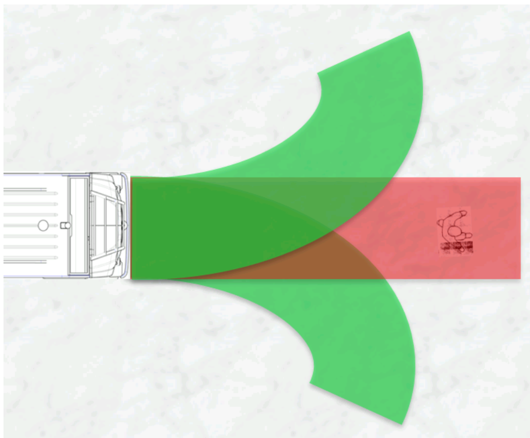


Figure 6.11: Semi-narrow pathway: Sum of tentacle weights for tentacle shown in Figure 6.10.

- **Open area:** Open areas, as the name suggests, are areas with no driving restrictions. AV is able to drive along the width of the pathway. For such scenario, there are more evasion possibilities. Also in this case, less interaction is needed since pedestrians have the possibility to walk in different directions other than the vehicle direction. But for the purpose of understanding, an unaware pedestrian is located in the same location similar to previous scenarios. Figure 6.12b shows the forward tentacle with lowest weight.



(a) Feature Tentacle



(b) Pedestrian Tentacle

Figure 6.12: Open area tentacle visualization: Green color show high traversiblity and red color shows low traversibility.

Examining the weight for this scenario in figure 6.13, it is shown that the left and right tentacle has the highest value; hence, it is possible to drive on either of the tentacles. Generally the tentacle is chosen which is in the direction of global trajectory.

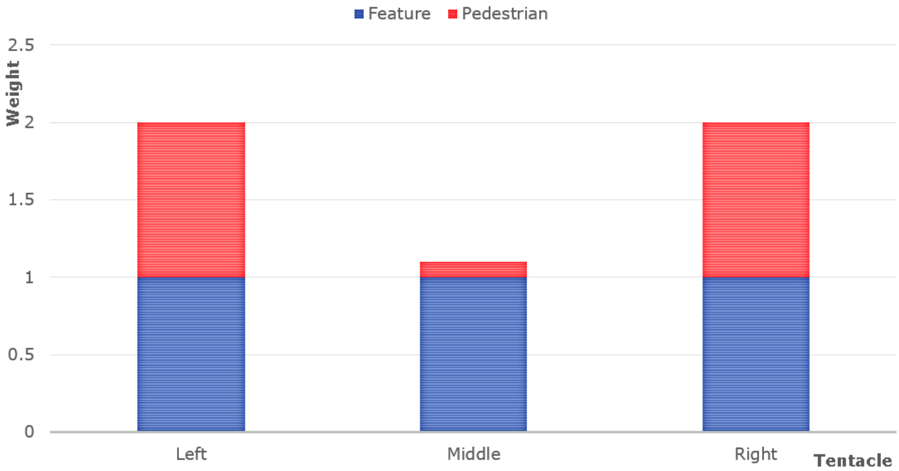
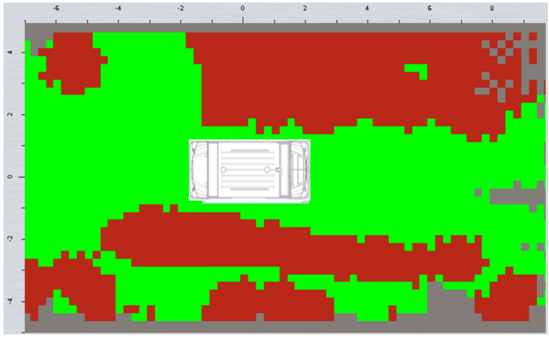


Figure 6.13: Open area: Sum of tentacle weights for tentacle shown in Figure 6.12.

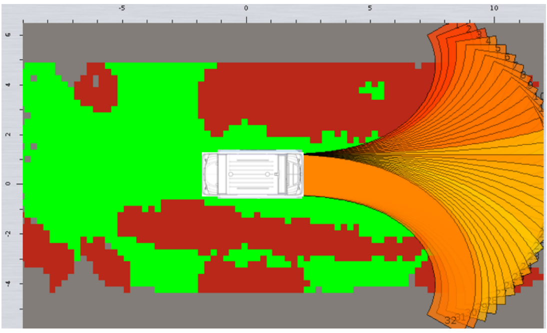
Figure 6.14 illustrates the results from an experimental visualization of the implemented 'tentacle' approach to pedestrian navigation mapping. The depicted 'pedestrian tentacles' are overlaid on a pedestrian-density map, serving as a heuristic representation of potential pathways for Autobus navigation. Each tentacle index is assigned a color-coded weight, which transitions from green to red to represent a spectrum of values ranging from lower to higher. This gradation is directly influenced by the pedestrian unawareness beneath each tentacle path: areas with higher pedestrian risk yield increased risk values, which in turn amplify the corresponding tentacle index's value. The methodology employed for computing these values integrates both the spatial distribution of pedestrians and the projected trajectories of the Autobus, with the aim of optimizing route selection for safety and efficiency.

Overall, the implementation of a dual-tentacle approach for Autobus navigation has achieved significant progress in the area of navigation for crowded environment. Given the sensitivity of pedestrians and the principles of traditional navigation, the system is well equipped to deal with the complexity of pedestrian environments in the real world. This component is the basis for future improvements in the driverless minibus navigation system and emphasizes the critical role of pedestrian awareness in safe and effective autonomous driving.

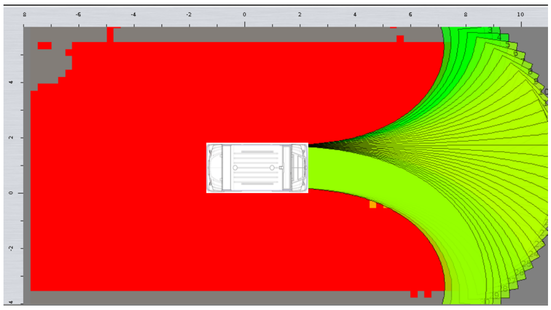
(a) Camera Image with Pedestrian Skeleton Detection: The visual input captures a pedestrian talking on the phone, and the skeleton detection outlines the human figure's skeletal structure for identification and tracking.



(b) Obstacle Map with pedestrian excluded: Here, the obstacle map is displayed after the pedestrian has been removed via aspect map, demonstrating the map's adaptive feature for scenarios excluding pedestrians.



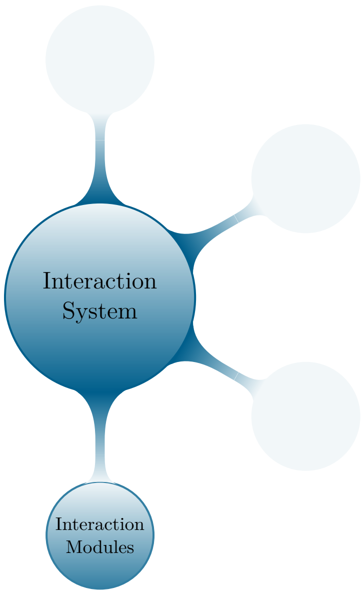
(c) Feature Tentacle: The traversable path for the Autobus is delineated by segmented paths. The regions depicted in red signify areas with higher weights, indicating the presence of obstacles.



(d) Pedestrian Tentacle: The depicted tentacle visualizes pedestrian detection. Areas marked in green indicate low weights, signifying an absence of pedestrians. As the color transitions to yellow, it denotes increasing weights, reflecting the presence of pedestrians within the vicinity.

Figure 6.14: Figure illustrates a scenario of pedestrian in narrow space. (a) shows the scene, (b) shows the obstacle map. Two different tentacles are shown, (c) feature for path detection and (d) pedestrian tentacle for pedestrian awareness.

7. Interaction modules



The interaction module is a key element of smart interaction strategies, as described in Chapter 6, especially when traveling through various pedestrian environments. Interaction modules bridge the gap between the complex driving environment of the Autobus and the human-centered environment they operate in. It includes tools and technologies that allow effective communication and interaction between vehicles, their occupants and external entities such as pedestrians. This chapter focuses on the importance of each interaction module used in Autobus and provides detailed analysis to emphasize its essential role as a major communication channel. These modules transcend basic operational elements and serve as the main interface for Autobus to interact and transmit information to external environments.

Based on the outputs of the decision-making module discussed in Chapter 6, there are specific scenarios where the targeted interaction of the vehicle with pedestrians is mandatory. To solve this problem, the interaction module uses various Human-Machine Interface (HMI) technologies, including LED displays, auditory communication, and light signals. This chapter starts with a detailed analysis of the HMI system that is essential to the dynamic interface between the user and the technology. In particular, the integration of HMI into the Autobus is highlighted, which means a significant advancement in human interaction with technological collection. HMI deployment is essential to ensuring clear and natural communication between individuals and complex automated systems and a key consideration in the era of rapid technological development.

Integration of the HMI system into Autobus has many advantages. These include improving accessibility and interaction fluidity of the system, which is necessary to ensure that pedestrians can effectively interact with these advanced vehicles regardless of their technical knowledge. This improvement in system intuitiveness transcends simple user convenience

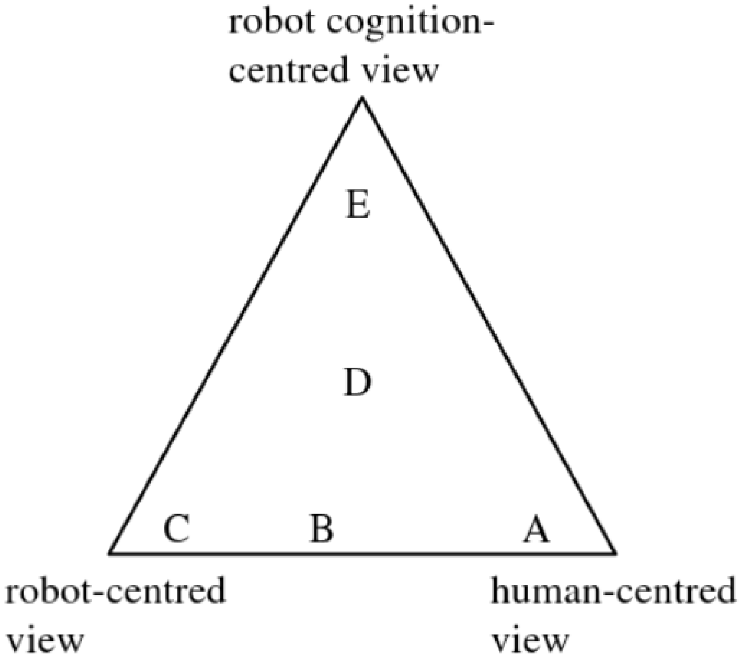


Figure 7.1: This triangular model maps the spectrum of HRI methodologies, ranging from a robot-centered view to a robot cognition-centered view and a human-centered view. At the vertices, we have 'A' denoting socially evocative approaches, 'B' indicating socially situated approaches, 'C' representing sociable robots, 'D' pertaining to socially intelligent robots, and 'E' at the apex as socially interactive robots. Each point within this triangle implies a varying degree of robotic complexity, from sophisticated service robots requiring advanced skills at one extreme to simpler, toy-like robots used for basic HRI research at the other. The underlying implication is that the robotic component's complexity scales with the sophistication of the HRI approach."The conceptual space of HRI approaches. A - Socially evocative, B - Socially situated, C - sociable, D - Socially intelligent, E - Socially Interactive. [Dautenhahn 07]

and plays an important role in safety assurance. It facilitates the clear and transparent transmission to pedestrians of the operational intentions of the Autobus. Additionally, this encompasses the conveyance of the vehicle's navigational intentions or yielding instructions to pedestrians, contingent upon the interaction field delineated in Chapter 5. Such transparency is essential to promote trust and predictability in vehicle behaviour and thus contribute to the overall safety of the ecosystems of the environment. This chapter begins with various communication categories to clarify the importance of the content used in Autobus communications. Afterwards, the existing literature to examined to explore the nature and application of interaction modules in research. The final section focuses on specific interaction modules implemented in the thesis, beginning with a description of the communication process and then a detailed review of the functionality of each interaction module is given.

7.1 Human Robot Interaction

Human-Robot Interaction (HRI) is a diverse topic which is still being extensively researched. Since the invention of technology, one aspect which is always kept in mind is the human interaction with such technologies. Human factors are considered throughout the period of designing, researching and evaluating new technologies. HRI is important where social interaction with people is the central part of the application such as driverless minibuses in shared spaces. The vehicle should exhibit intelligent cognitive skills in order to be accepted and appealing; this results from findings of exemplary social intelligence in humans.

The spectrum of social skills in robots are application dependent [Yanco 02, Scholtz 03]. Robots performing independent tasks to robots used as human companions have low to high levels of social skill requirements, respectively. In the context of vehicle-pedestrian interaction, the inquiry pertains to the requisite social capabilities a vehicle must possess. The responses fall within the scope of understanding human social cognition and its principles. A relationship between different approaches of HRI is developed in [Dautenhahn 07] shown in figure 7.1.

Figure 7.2 is a triangular model that provides a comprehensive framework for categorizing Autobus HRI functions which contains pedestrian interaction modules. Autobus uses these modules to interact with pedestrians on the basis of its activities and communicate its driving state, including aspects of the model's defined HRI approach. It represents "D" for socially intelligent robots through the advanced processing of social indicators and "E" for socially interactive robots through its proactive engagement with nearby pedestrians. As a result, the Autobus is located in a sophisticated link between "D" and "E", showing not only a higher level of robot complexity, but also a higher level of HRI complexity. This position in the triangular model highlights the ability of the Autobus to navigate complex social environments safely and effectively, using nuanced social behaviors that are aware of human social standards and expectations.

Classification and Strategies of Vehicle-to-Pedestrian Communication Modalities

In the specific area of this research, vehicle interaction modes are classified into two broad categories: formal and informal, as previously established in [Färber 15]. Formal interactions include universally recognized signals, and are included in the design and function of vehicles, including turn signals, horn and headlight activation. However, informal interactions are those that arise from the human driver itself and encompass a number of non-verbal signs that are subjective and dependent on individual behavior.

In the field of vehicle-to-pedestrian communication, [Mirnig 17] identifies three major communication strategies. These are the practical communication of standard vehicles, the use of vehicles with human-like characteristics, and the concept of placing a social robot within the driver's area. This scope of research focuses on the utility approach. This choice is informed by its alignment with the inherent attributes of the system and the requirements for the interactions it must manage. The aim is to communicate the decision-making process of the vehicle directly to pedestrians without the need for human characteristics on the vehicle. In addition, the idea of placing a social robot in the driver's seat is abandoned because it would have a negative impact on the space otherwise available to passengers.

Moving forward„ this chapter will conduct a comprehensive analysis of the dynamics of interaction between Autobus and pedestrians. The discussion begins with a systematic explanation of the different communication channels that facilitate interaction between these agents. This examination will include an in-depth examination of the important factors that affect the nature and effectiveness of these communications.

7.2 Spectrum of Communication Types in driverless vehicle-Pedestrian Interactions

Building upon the foundational understanding of the Autobus interaction module, it is essential to examine the various types of communication promoted by these systems. In general, these communications can be classified into verbal and non-verbal forms, implicit and explicit, as well as intentional and unintentional. Each category plays a unique role in the transmission of information and intentions, especially in the case of driverless minibuses that interact with pedestrians.

7.2.1 Verbal and Non-verbal Communication

In the field of driverless minibus technology, especially the interaction between driverless minibuses and pedestrians, it is important to classify communication into verbal and non-verbal forms. Verbal communication, which includes oral and written expressions, is the basis for articulating vehicle directives. Its effectiveness depends on the fusion of factors such as volume, intonation, articulation and content selection. Furthermore, the way the receiver absorbs and interprets feedback is equally important to determine the success of verbal interaction.

On the contrary, non-verbal communication refers to communication of messages without the use of spoken or written languages employing medium such as such as speech patterns and animation. Non-verbal signals, such as voice tone, body language, and visual contact, could either strengthen or contradict the messages that are communicated verbally, and provide multifaceted communication layers that cannot be communicated by verbal means alone.

In the context of Autobus, such communication capabilities are instrumental. Integration of verbal communication methods may include sound signals or verbal warnings to inform pedestrians of the intention of the vehicle. Non-verbal communication includes visual signals such as lighting patterns or screen displays that imitate human movements and recognize the presence of pedestrians. With these means of communication, the Autobus can effectively signal their actions and intentions to pedestrians, improve safety and facilitate harmonious coexistence with pedestrians. The subtle application of these types of communication is not only fundamental to interpersonal communication, but also adapted and refined to interaction between pedestrians and Autobus.

7.2.2 Implicit and Explicit Communication

In scientific debate, it is necessary to distinguish between implicit and explicit communication in order to understand the complex relationship between driverless minibuses and pedestrians. Implicit communication involves subtler, often non-verbal, signals that indirectly convey intentions or information. In the case of Autobus, this can be demonstrated by applying specific lighting colours or sequences on the display, indicating caution, or by decelerating as the vehicle approaches a pedestrian, thus indicating, non-verbally, its intention to stop.

In contrast, explicit communication is characterized by clear communication of messages. For the Autobus, this may take the form of a audible announcement like "crossing forward" or a visual message on an LED display that clearly states messages such as "walking for pedestrians". Such straightforward communication is vital to eliminating ambiguities and minimizing the risk of misinterpretation of the dynamic interactions between the Autobus and the pedestrians.

The employment of both implicit and explicit communication methods is vital in the autonomous operation of Autobus. Implicit cues are often used as preemptive signals, subtly alerting pedestrians to the vehicle's immediate intentions and enabling them to anticipate its subsequent actions. Explicit communication serves as a clear confirmation of the vehicle's imminent behavior, providing pedestrians with the certainty needed to navigate their interactions with the Autobus safely. The strategic use of these communication modes enhances the overall efficacy of the interaction, ensuring that the shared space between Autobus and pedestrians is navigated with mutual understanding and safety.

7.2.3 Intentional and Unintentional Communication

Within the scientific framework of driverless minibus design, particularly in relation to pedestrian interaction, communication is bifurcated into intentional and unintentional categories. Intentional communication is characterized by deliberate actions taken by the vehicle to transmit information or warnings to pedestrians and other road users. For Autobus, this encompasses the strategic utilization of auditory signals or the exhibition of specific messages or icons on LED screens, all designed to articulate the vehicle's imminent movements or its operational status.

Unintentional communication, although not explicitly designed in the vehicle's communication suite, nevertheless provides information. This type of communication can be generated by sensory signals such as vehicle electric motor noise, visual indicators such as Autobus speed, or kinetic signals such as tyre orientation. Although these elements have not been deliberately used as communication tools, they can provide pedestrians with information about the current situation of the vehicle or the planned maneuver.

For Autobus operating in shared areas, intentional communication systems are crucial to the role of a link between vehicles and pedestrians and to ensure that pedestrians are aware of mini-bus behaviour. Unintentional communication also plays a supportive role; pedestrians can identify certain behaviours from these indications and help to develop a comprehensive understanding of vehicle behaviour. Together, these forms of communication enhance the symbiotic interaction between the Autobus and pedestrians, ensuring a safer and more predictable shared road environment.

7.3 Key Parameters for Effective Communication

Driving in pedestrian zones differs from driving in a normal street environment. As the name suggests, these zones are meant to be for pedestrians to walk over the entire width of the area. Conventional traffic is diverted from this area. Due to the presence of people, researchers working in autonomous driving in such zones use the jargon of social interaction for many different reasons: focusing on trust factor; clarifying incertitude; cutout indecisiveness and increase overall safety. Interaction modules play a vital part in explicit communication with pedestrians. Pedestrians are comfortable with knowing about the intentions of the vehicle. For a human driver, there is a resonant eye contact between the driver and the pedestrian (explained in psychological aspects of human drivers section of Chapter 2). By means of head and hand gestures – in advance – the driver avoids decelerating the vehicle and politely solicits the vehicle's right to pass. On the other side, knowingly the priority of a pedestrian, the pedestrian volunteers to prorogate its rights to pass. The said communications are likely to happen when an unexpected vehicle drives through tight spaces as pedestrian zones. Unlike urban-streets, a vehicle would never be expected to drive on footpaths and make sure to stop before zebra crossing to give way to the pedestrian. These interactions must exist in this interaction concept of this thesis in order to maintain the same receptivity from pedestrians. In the development of communication systems for interactions between Autobus and pedestrians, several key factors are considered to ensure effective and safe exchanges. These factors are derived from an array of research studies focusing on different aspects of Autobus-pedestrian communication:

- **Trust** [Schmidt-Wolf 22, Matthews 17]: One of the popular problems in the umbrella of relationship is the trusting factor [Ayupova 19]. Risk and uncertainty are reduced with higher trust. Hence, trust becomes a ground for decision making. Trust is a fluctuating belief having multiple stages, from establishment to restoration when undermined. Ontology based recognition and understanding natural language in smart devices is a breakthrough of interaction between users and technology. There are five components extracted which affect trust: social intelligence, voice characteristics and communication style, look of the agent, non-verbal communication, and performance quality [Rheu 21]. Interpersonal relationship trust definition is defined in [Rempel 85]. Faith, predictability and dependability are the dimensions examined for trust. Focusing on Autobus and pedestrian dynamics, trust is a critical factor that underpins the safety and harmony of shared spaces. This trust encompasses several key dimensions, including faith in the autobus's capabilities, the predictability of its actions. Similar work with its dependability across various situations—elements that has been thoroughly explored by researchers in [Rempel 85]. In this context, faith means that pedestrians have confidence in the Autobus that allows them to navigate safely and make decisions without human intervention and to ensure their safety in all interactions. Predictability is also important, so that pedestrians can anticipate the vehicle's movements, such as stopping at the intersection or yielding, which is essential to safe coexistence in shared spaces. Dependability refers to the continuous performance and reliability of cars to detect pedestrians and comply with traffic regulations, strengthening pedestrian confidence in these systems. The implementation of the intent display system in Autobus is an important factor in this process of building trust. By clearly communicating the intention of the vehicle, whether stopping, turning or driving, these systems help to make Autobus movements more predictable and understandable to pedestrians and thus to directly increase trust.

- **Information:** In the interaction between Autobus and pedestrians, the transmission of information is fundamental in fostering trust. Trust is based essentially on knowledge, which, in the case of the Autobus, means that pedestrians have a clear understanding of the operation of the vehicle. Essential details such as the Autobus's speed or its current mode of operation needs to be communicated transparently to pedestrians to demystify the vehicle's actions. For effective pedestrian interaction, this information must be shared in a direct and unambiguous manner, often through verbal announcements. For instance, Autobus might audibly inform nearby pedestrians of its intention to halt or accelerate, thereby eliminating confusion. Such straightforward communication of the Autobus's status empowers pedestrians with the necessary information to safely navigate their interactions with the vehicle. The clarity and precision of this information exchange are pivotal in mitigating misunderstandings and enhancing pedestrian confidence in the safety protocols of the vehicle.

- **Intention:** The articulation of intentions by an Autobus is a pivotal element in its interactions with pedestrians. The explicit clarification of the Autobus's prospective actions plays a significant role in clarifying any doubts or uncertainties that pedestrians may harbor. Communication of these intentions should be direct and clear, often employing verbal cues or clear text displays, typically limited to three or four words for clarity. For example, an Autobus may utilize displays or audio messages to signal actions such as "Turning left" or "Slowing down". This precise expression of intent is crucial for support transparent communication, enabling pedestrians to anticipate and appropriately respond to the Autobus's movements.

- **Commands:** The Autobus's controls are essential to ensure pedestrian safety and awareness. Commands such as information and intentions should be intentional, clear and verbal. The purpose of these commands is to get pedestrians to understand and react immediately. For safety reasons, it is essential to formulate these commands so that they can be deduced quickly and easily. Examples of such commands include direct and friendly instructions such as "Please give the way" or "Please leave." The directness of these commands prevents misunderstandings and ensures that pedestrians can quickly understand and act on the requests of the Autobus. This direct communication helps to maintain the coexistence of Autobus and pedestrians, particularly in situations where immediate pedestrian actions are required to avoid potential safety risks.

Incorporating these factors into the design and operation of Autobus can significantly improve the safety and efficiency of their interactions with pedestrians.

7.4 Multimodal Interaction Process for Enhanced Vehicle-Pedestrian Interaction

In previous discussion, the interaction mechanisms between the Autobus and pedestrians were studied. These interactions are designed to be fluid and intuitive, and the diversity of communication exchanges is proportional to the number of interaction channels used. For example, multimodal communication combines visual and auditory information and usually provides a clearer understanding than a single communication channel. Effective communication is characterized by facilitating understanding and promoting reciprocal exchange. These qualities, fluidity, clarity and bidirectional communication, are especially important in the environment with dense pedestrians.

The configuration of this interaction is depicted in Figure 7.2, elucidating the interactive exchange as established in this research. For illustrative purposes, the diagrammatic representation in Figure 7.2 is bifurcated into two distinct sections: the upper section illustrates the Autobus-to-pedestrian communication vectors, whereas the lower section visualizes the pedestrian-to-Autobus information flow.

The upper part of the figure shows the potential inputs observed by pedestrians from the Autobus. Communication from the vehicle can be performed either by a single channel or by a composite of multiple channels simultaneously. Each of these interaction modules has its own characteristics, display, lighting and sound, and has a specific communication purpose that will be discussed in the following section.

Conversely, the lower segment of the image details the vehicle's perception mechanisms of pedestrian information. This includes the acquisition of 3D human skeletal joint points via stereo cameras strategically mounted atop the Autobus (Chapter 4). These joint points are captured from all cameras in the system (the technical configuration of which is expounded in Chapter 3); however, the figure exemplifies only the lateral perspective from the foremost camera. Such perception capabilities enable the Autobus to discern pedestrian posture and movement, forming a foundational component for responsive vehicle-pedestrian interaction.

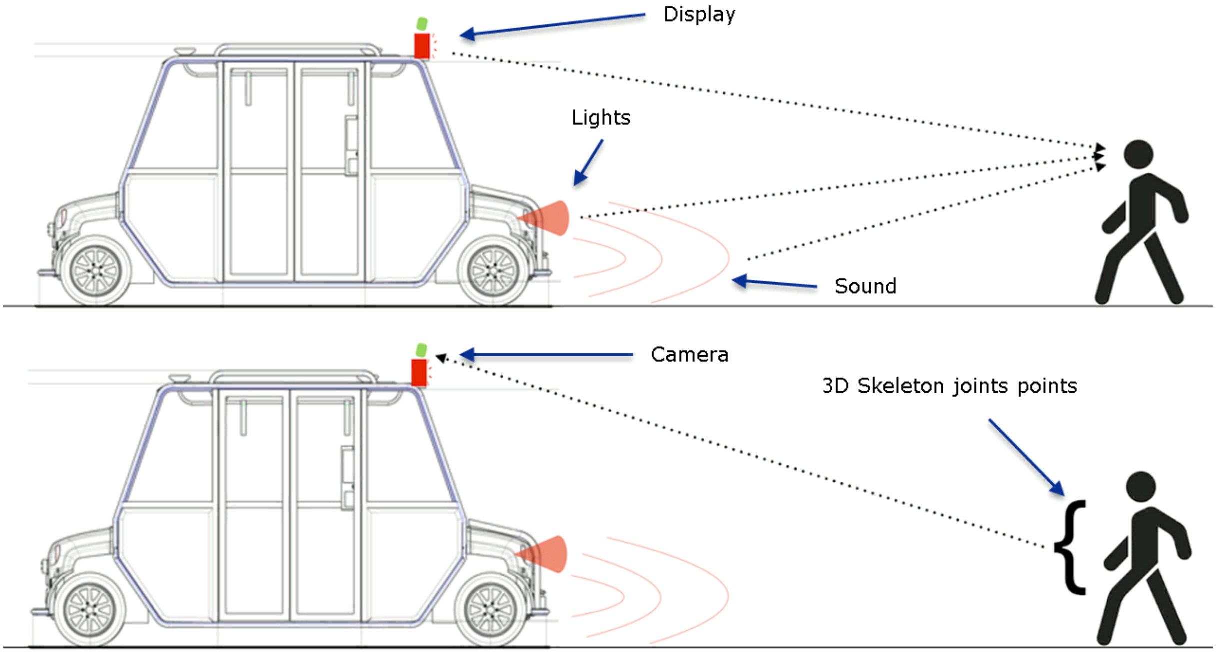


Figure 7.2: The upper section of the illustration presents the Autobus's communication modules, including visual displays, lighting signals, and auditory notifications, which provide pedestrians with multimodal cues about the vehicle's intentions. The lower section depicts the pedestrian information acquisition system utilized by the Autobus, capturing 3D human skeletal joint points through an array of stereo cameras mounted on the vehicle's roof, showcasing the side view from the front camera only. This figure encapsulates the dual aspects of interaction—outgoing communication from the Autobus to the pedestrian and the incoming pedestrian data to the Autobus, underlining the system's comprehensive approach to interaction in pedestrian zones.

7.5 Optimizing Pedestrian-Aware Navigation Through Adaptive Interaction Module Activation

The efficacy of pedestrian-influenced vehicle navigation can be enhanced through judicious activation of interaction modules. These modules offer a range of application possibilities, which are streamlined by activating them in response to specific interaction fields. The conditions for this activation are depicted in Figure 7.3, with the interaction fields, defined in Chapter 5, serving as regulatory mechanisms. As discussed in the interaction process section of this chapter, there are three primary types of interaction: audio, video, and beepers. A key objective of this thesis is to employ these interaction modules in a manner that avoids startling pedestrians. Consequently, the potential activation of each interaction module is delineated for different zones, as illustrated in Figure 7.3.

In the ambient interaction zone, interaction is limited to LED displays, suitable for a safe environment where pedestrian attention is not imperative. This allows for the display of intentions or information for pedestrians who may be interested. The direct interaction zone offers a broader scope for interaction, facilitating most priority negotiations. Here, urgent commands and awareness signals are communicated, and the use of voice and beepers is appropriate given the proximity to the risky interaction zone. Progressing to the risky interaction zone, light flickering is employed to convey a sense of urgency, coinciding with the activation of brakes. This zone is critical for signaling imminent hazards.

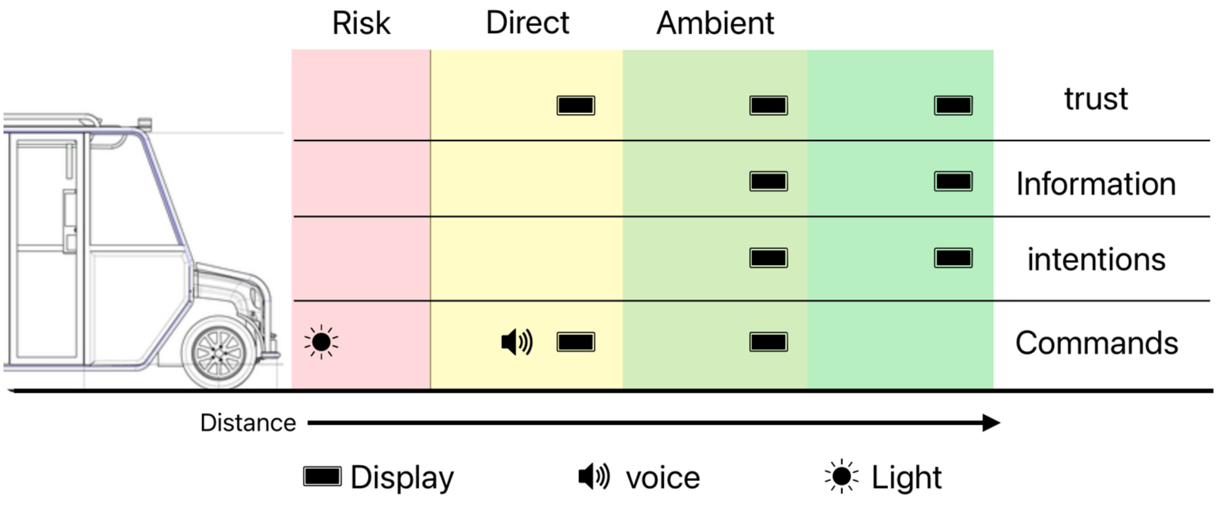


Figure 7.3: This figure illustrates the tailored application of various interaction modules across different pedestrian proximity zones, as established in Chapter 5. Each zone triggers specific modules to optimize the communication of risk, intentions, information, and commands to pedestrians. In the ambient zone, LED displays are used for non-critical information sharing where pedestrian alertness is not immediately necessary. Moving closer, in the direct interaction zone, a combination of displays, voice, and sound signals facilitates essential communication for priority negotiation. The most critical, the risky interaction zone, employs light signals to indicate urgent situations, such as the vehicle's imminent stopping, aligning with the deployment of the braking system. This systematic approach ensures effective, non-startling interaction tailored to the context of pedestrian proximity and the associated level of navigational risk.

7.6 Exploring Interaction Content For Enhanced Communication

This section focuses on the interaction content used by the Autobus for communications with pedestrians. The design and implementation of effective interaction content is the most important to ensure safety and promote trust between the Autobus and pedestrians. This section examines various types of content that can be used by the Autobus to communicate intentions and actions to pedestrians, including visual, auditory and combined modes. The aim of analysing current research and practice in this area is to understand the impact of these modes of interaction on pedestrian behaviour and perception, thereby contributing to the development of more intuitive and safer autonomous transportation systems.

To evaluate the effectiveness of the interaction system, a comprehensive experiment with Autobus [Jan 23a] was conducted, using the two basic components of the interaction system: visual displays and audio signals. An extensive analysis was conducted to assess the pedestrian understanding of the interaction system by displaying various signals on the display. These modules can transmit various content, including text, images and animations, with adjustable attributes such as size, blinking rate, colors etc. One of the distinctive characteristics of this research is the iterative improvements in display content which are made based on feedback from pedestrians to ensure continuous improvement of the system.

Figure 7.4: Initial distribution of content.

The interaction content was initially categorized based on typical pedestrian activities around a moving Autobus, such as yielding, and parallel or perpendicular crossing. These activities were then broken down into various representational forms, including texts, signs, and animations. The representations were further diversified by displaying them in different color combinations and animation styles, resulting in over 200 potential variations. An example of some of these primitive designs is illustrated in Figure 7.4, providing a visual reference for the diversity and scope of the interaction content tested.

The primary objective of this thesis is to improve and optimize various visual elements (such as animation, text, and signs) in order to enable effective and clear interactions, especially in scenarios where pedestrians have limited time to interpret these signals while crossing the Autobus. The research focuses on the clarity and efficiency of these visual signals, establishing a solid foundation for the interaction system. The goal is to significantly reduce the likelihood of misinterpretations and accidents in the real world, thus enhancing the safety and reliability of the interaction between vehicles and pedestrians. The testing process and results are described as follows.

7.6.1 Experimental Configuration and Technique

Static tests were conducted on controlled-environment with the Autobus, in which participants were instructed to position themselves in front of the stationary vehicle. This static test method minimizes uncertainty in pedestrian reactions and provides valuable insight into pedestrian perceptions and responses to specific characteristics of interaction content. The static test consists of two separate phases:

- **Phase 1:** This initial stage involved the system displaying a complete set of primitive variations of the interaction content. The aim was to determine which symbols are most relevant and easily understood by participants. Feedback was actively sought during this phase in order to incorporate these suggestions into the next phase of the testing.
- **Phase 2:** Building on the findings from Phase 1, refined content was sequentially displayed. Participants were asked to cross the front of the vehicle while these sequences were active. This phase aimed to evaluate the ease with which participants could understand and respond to the sequence of interactions, a critical factor given the typically brief duration of real-world Autobus-pedestrian interactions. Participants were encouraged to vary their positions relative to the vehicle, such as standing parallel or perpendicular to it, to simulate different real-world scenarios. After each test scenario, participants were queried about their recognition, visibility, and comprehension of the signs, with their responses collected via a questionnaire.

Participants: During the first phase of static testing, the participants consisted of 14 individuals, ranging in age from 23 to 58 years. This group comprised 1 female and 13 male participants, all of whom were engaging with the interaction system for the first time.

The second stage of static testing collected data from 20 participants. The group had a similar age distribution, with 14 men and 6 women. Participants come from a variety of backgrounds within the university environment, including students, researchers, and faculty members. In particular, six of these participants participated in the first phase of the study, providing continuity of experience to enrich the data collected. Other participants, some had previously experienced similar experimental settings, but were entirely new to this type of experimental interaction. The combination of the experiment and the novice sample was designed to simulate a realistic representation of possible Autobus-Pedestrian interactions that includes familiarity and expertise with Autobus technology.

Questionnaire: The questionnaire designed for this study contained queries specifically aimed at elucidating the participants' comprehension of the content displayed by the Autobus and the logic underpinning the sequence in which this content was presented. Detail of questionnaire can be found in Appendix B. The testing was bifurcated into two distinct phases, each with its own targeted objective.

- **Phase 1:** The primary aim during this initial phase was to gather pedestrian feedback on the various types of displayed content. The goal was to filter and identify the most effective and comprehensible content for pedestrian interaction. The insights gathered in this phase were crucial for refining the interaction modules, as depicted in Figures 7.5, 7.6, and 7.7.

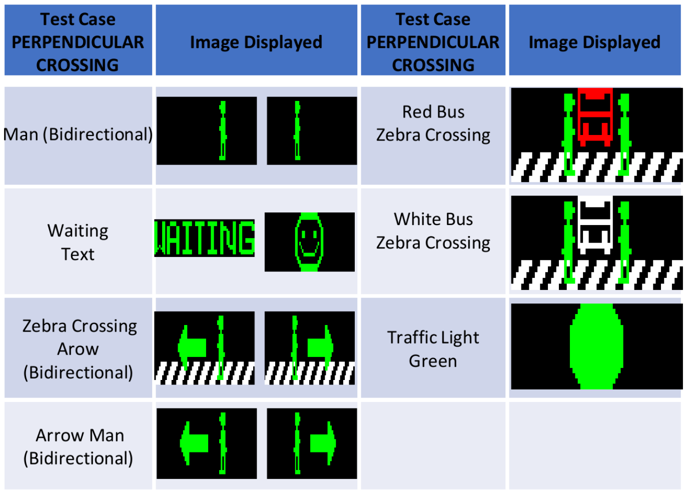


Figure 7.5: The matrix represents Phase 1 of static testing for pedestrian comprehension of Autobus interaction content. The left column categorizes the visual communication strategies: Bidirectional Man, Waiting Text, Zebra Crossing Arrow (Bidirectional), and Arrow Man (Bidirectional), which are designed to direct or instruct pedestrian movement. The right column illustrates practical applications of these strategies: Red and White Bus Zebra Crossings indicate vehicular presence at pedestrian crossings, while Traffic Light Green signals safe crossing conditions. This phase aims to identify the most effective communication content for refining pedestrian interaction modules with Autobus.

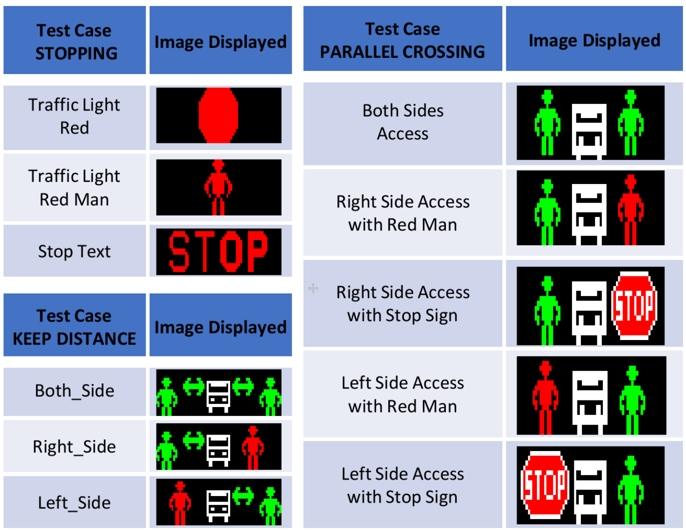


Figure 7.6: Displayed is a compilation of interaction content continued from Figure 7.5 used in static testing for pedestrian communication, categorized under three test cases: STOPPING, PARALLEL CROSSING, and KEEP DISTANCE. In the STOPPING case, the imagery includes a red traffic light, a red figure, and a 'STOP' text, all indicating a clear command for pedestrians to halt. The PARALLEL CROSSING case presents scenarios with both sides of a crosswalk accessible, restricted access on one side indicated by a red figure, and a stop sign commanding no crossing. The KEEP DISTANCE case illustrates variations of vehicular proximity with corresponding pedestrian actions: green figures denote safe passage, while red figures and stop signs advise caution and stopping, tailored for both right and left sides. This figure underscores the importance of clear visual cues for safe pedestrian behaviors around Autobus.

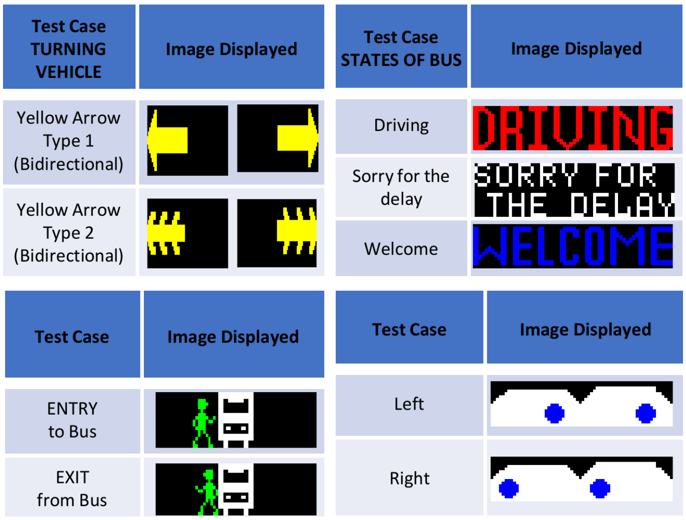


Figure 7.7: This matrix illustrates interaction content for Autobus communication, continued from Figure 7.6, organized into three test cases: TURNING VEHICLE, STATES OF BUS, and BUS ENTRY/EXIT. The TURNING VEHICLE case shows yellow arrows in two styles indicating bidirectional turning options. The STATES OF BUS case presents textual information for the operational status of the bus: 'DRIVING' in red indicates motion, 'SORRY FOR THE DELAY' in mixed colors communicates a halt, and 'WELCOME' in blue signifies readiness for boarding. The BUS ENTRY/EXIT case depicts green figures entering and exiting a bus, coupled with directional eye patterns suggesting the safe direction for these actions. These visual cues aim to provide clear instructions and status updates to pedestrians and passengers interacting with Autobus.

- **Phase 2:** The focus in this subsequent phase shifted towards understanding how pedestrians processed and reacted to the sequence of the displayed content from Phase 1. This phase was critical in evaluating the effectiveness of the content flow and its impact on pedestrian behavior and perception. Figures 7.8 and 7.9 illustrate specific test scenes and the sequences utilized in this phase.

An example sequence for the evaluation of a perpendicular crossing scenario during Phase 2, the test sequence was methodically structured as follows: initiation with 'DRIVING,' transitioning to 'EYES' (denoting the Autobus's environmental scanning), progressing to 'STOPPING' (paired with a descending numerical indicator), proceeding to the perpendicular crossing test case (inclusive of a Red Bus icon, Zebra Crossing representation, and a pedestrian motion animation), then advancing to 'RESUMING' (again, with a numerical countdown), and culminating with the resumption of 'DRIVING.' This sequence is representative of the intricate communication paradigms explored for this thesis.

Participants were engaged in a binary query sequence built to explain their understanding of the content sequence, their tendency to follow the instructions described, and their subjective assessment of safety in realistic conditions. The grouping of these responses gave a decisive idea of the communication effectiveness of the sequences presented, thus forming an integrated assessment of the applicability of the interactive system in a real traffic environment.

Test Scenes	Sequence
Yielding	DRIVING
	EYES
	HAPPY EMOTICON
	STOPPING 5...1
	RED BUS
	RESUMING 5...1
	DRIVING

Test Scenes	Sequence
Non-Yielding	DRIVING
	EYES
	STOP TEXT
	STOP MAN
	HAPPY EMOTICON
	DRIVING

Test Scenes	Sequence
Parallel Crossing	DRIVING
	EYES
	Happy Emoticon
	Parallel Crossing Both
	Driving

Test Scenes	Sequence
Vehicle Turning	DRIVING
	EYES
	VEHICLE TURNING LEFT and Right
	DRIVING

Figure 7.8: Depiction of the sequence of interaction content tested in Phase 2, analyzing pedestrian processing and reaction. Test scenes for 'Yielding,' 'Non-Yielding,' 'Parallel Crossing,' and 'Vehicle Turning' are presented, each with a sequential flow of visual cues. For 'Yielding,' the sequence starts with 'DRIVING,' followed by 'EYES' to capture attention, 'HAPPY EMOTICON' to signal positive intent, a countdown 'STOPPING 5...1,' a 'RED BUS' indicating a full stop, then 'RESUMING 5...1,' and concludes with 'DRIVING.' 'Non-Yielding' replaces the positive cues with direct 'STOP TEXT' and 'STOP MAN' commands, indicating no intention to yield. 'Parallel Crossing' reinforces safe crossing with 'Happy Emoticon' and 'Parallel Crossing Both' directions, and 'Vehicle Turning' alerts to the 'VEHICLE TURNING LEFT and Right' before resuming 'DRIVING.' This visual progression assesses the clarity of intent and ensures pedestrians can anticipate the vehicle's actions, promoting safety and trust in Autobus operation.

Test Scenes	Sequence
KEEP SAFE DISTANCE BOTH, RIGHT, LEFT	DRIVING
	EYES
	KEEP SAFE DISTANCE
	DRIVING

Test Scenes	Sequence
ENTRY_EXIT	DRIVING
	STOPPING 5...
	EXIT FROM BUS
	WAITING
	ENTRY TO BUS
	RESUMING 5...
	DRIVING

Test Scenes	Sequence
SORRY_FOR_DELAY	Driving
	Stopping 5....
	Sorry for delay
	Waiting
	Resuming 5...
	Driving

Figure 7.9: This figure, continued from Figure 7.8, outlines the sequence of interaction content for the second phase of testing, with a focus on 'KEEP SAFE DISTANCE,' 'SORRY_FOR_DELAY,' and 'ENTRY_EXIT' test scenes. 'KEEP SAFE DISTANCE' involves a sequence beginning with 'DRIVING,' attention-grabbing 'EYES,' a directive to 'KEEP SAFE DISTANCE' from the vehicle, and then resuming 'DRIVING.' The 'SORRY_FOR_DELAY' scene communicates a stop in vehicle movement with 'Stopping 5...,' an apologetic 'Sorry for delay,' a holding pattern 'Waiting,' followed by 'Resuming 5...' and then the continuation of 'Driving.' The 'ENTRY_EXIT' scene guides the passenger flow with 'DRIVING,' a countdown 'STOPPING 5...,' 'EXIT FROM BUS,' a pause 'WAITING,' invitation 'ENTRY TO BUS,' countdown 'RESUMING 5...,' and finally, 'DRIVING' resumes. These sequences are crucial in studying pedestrian and passenger responses to dynamic visual communications from Autobus.

7.6.2 Assessing Interaction Content Characteristics for Optimized Communication

In the first phase of the testing, the study focused on establishing the basic requirements for effective interaction between the Autobus and pedestrians. A major finding is that the size of the display significantly influences the readability of various elements such as symbols, texts or figures. Furthermore, the context and interpretation of the message depend on the relative position of the Autobus and pedestrian. This emphasized the need to adapt the position of the display to target different audiences in the vicinity. Thus, the ability to communicate multiple messages simultaneously became a critical factor in the success of interaction.

Testing Phase 1: Two subjective opinions, *Recognition* and *Preferences*, were reckoned for the displayed content. *Recognition* is how easy it is to recognize the content, and *preference* is which content is liked independent of recognition.

Utilizing the responses compiled via the questionnaire delineated in Appendix B, the subsequent graphical representations, specifically bar charts, are presented. These charts are structured with the test case categories positioned along the X-axis and the corresponding percentages, indicative of recognition or preference rates, plotted along the Y-axis. In the context of the perpendicular crossing scenario, a range of combinations derived from Figures 7.5, 7.6, and 7.7 were subjected to evaluation. Their respective recognizability and preference levels are depicted in the upper and lower portions of Figure 7.10, accordingly. Interpretation of these results suggests that the 'Man-Zebra Crossing-Arrow' symbology

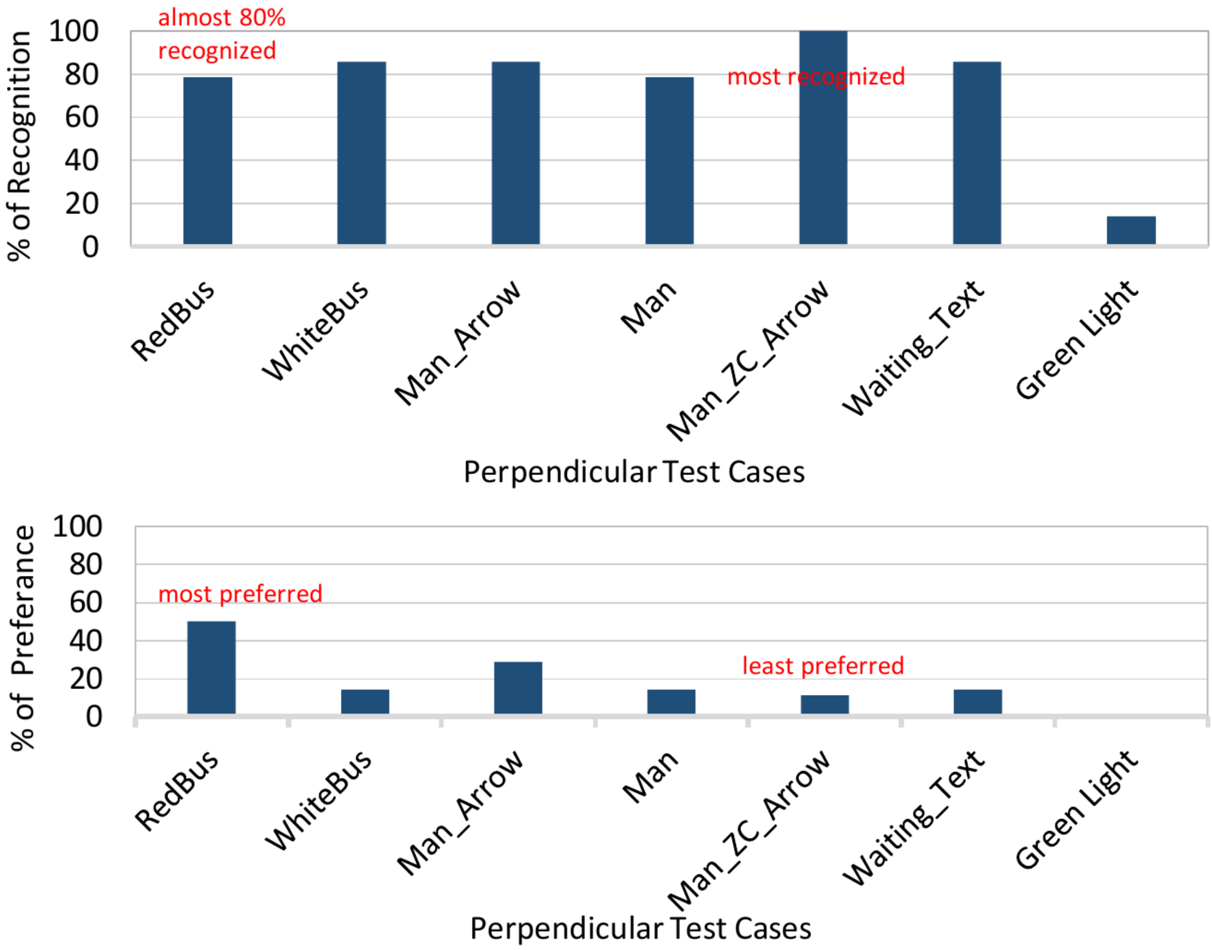


Figure 7.10: The upper chart quantitatively represents participant recognition rates for various signal types, with the 'Red Bus' and 'Man-ZC_Arrow' configurations achieving nearly 80% and the highest recognition, respectively. The lower chart displays participant preference percentages, highlighting a strong predilection for the 'Red Bus' signal, while the 'Man-ZC_Arrow' is identified as the least favored. These metrics reflect the efficacy and desirability of specific visual cues used in Autobus-to-pedestrian communication within a controlled testing environment.

garners the highest recognition yet the lowest preference among the study's participants. Conversely, the 'Red Bus-Zebra Crossing-Man' iconography achieves recognition by 80% of the cohort, simultaneously attaining the status of most favored. In a similar vein, the bar charts within Figure 7.11 (upper chart) corresponding to the non-yielding perpendicular crossing delineate that the 'red stop man' icon registers as the most identifiable sign for the aforementioned scenario. Additionally, Figure 7.11 (lower chart) encapsulates data for various other scenarios, including 'keep safe distance,' 'entry case,' and 'exit' case, indicating that these were perceived as intended by the participants. Notwithstanding that certain test cases did not achieve universal comprehension, a substantial proportion of the participant base exhibited an understanding of the conveyed messages.

Testing Phase 2: Figure 7.12 illustrates bar charts for Phase 2, encompassing test sequences for 'Yielding,' 'Entry-Exit,' and 'Sorry for Delay' from Figure 7.8 and 7.9. These sequences achieved a complete (100%) recognition and comprehension rate among participants, who reported full adherence to and a sense of safety following these directives.

Figure 7.11: The upper graph presents recognition rates for the 'Red Traffic Light,' 'Red Stop Man,' and 'Stop Text' test cases within Phase 2, illustrating a high level of participant recognition for 'Red Stop Man.' The lower graph details recognition percentages across a broader spectrum of test cases: 'Keep Distance,' 'Parallel Yield Man,' 'Parallel Yield...', 'Entry,' 'Exit,' 'Driving Text,' 'Turning Blinking,' and 'Turning Progress...,' with 'Keep Distance' and 'Red Stop Man' achieving near-universal recognition. This graphical representation indicates varied levels of perceptual accuracy among participants for different visual cues in Autobus communication protocols, emphasizing areas for potential enhancement in the 'Turning' related cues.

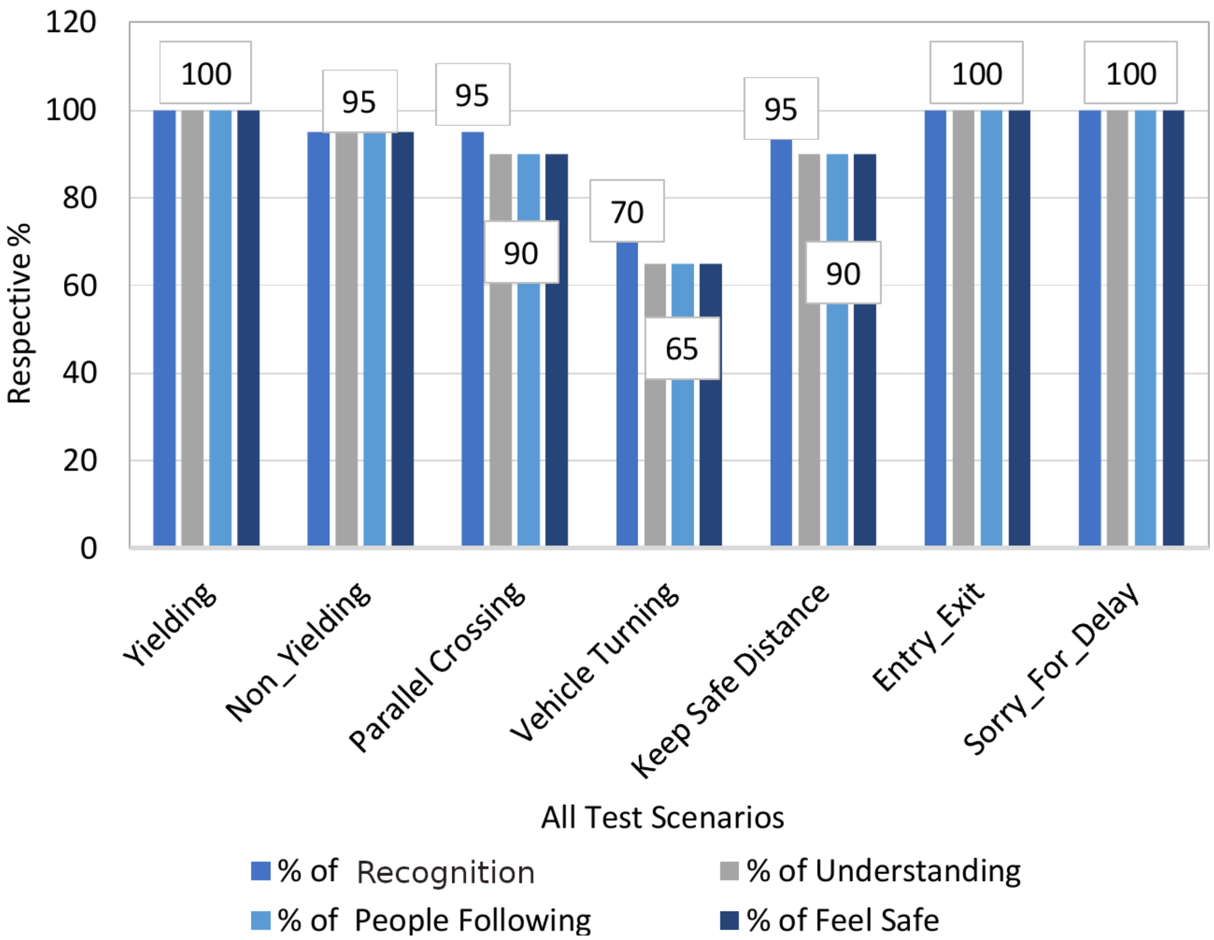


Figure 7.12: Results for Phase II testing.

Conversely, the test scenes 'Parallel Crossing' and 'Keep Safe Distance' were subject to analogous perception by the participants. The 'Non-yielding' scenario was substantially understood by 95% of the participant cohort. However, the 'Vehicle Turning' scenario necessitates enhancements, as its current perceptual recognition lingers between 60 and 70%, which is below the targeted benchmark of 90%.

In conclusion, the testing phases have provided valuable insights into the effectiveness of the interaction between the Autobus and pedestrians. In the first phase, the importance of the readability of symbols, texts and figures is strongly influenced by display size and positioning. The need to adapt displays to different audience groups and the ability to transmit multiple messages simultaneously were highlighted.

Subjective assessment of recognition and preference revealed specific preferences and understandings. The combination of the "Man-Zebra Crossing Arrow" was the most recognized but the least preferred. On the contrary, the combination "Red Bus-Zebra Crossing-Man" achieved a high score both in recognition and in preference. These results show a complex relationship between recognition and preference, and show that the most recognized content is not always the most preferred one.

Phase 2 testing further demonstrated that certain messages, such as 'Yielding', 'Entry-Exit', and 'Sorry for Delay', achieved 100% recognition and were understood and followed by all participants, indicating a high level of clarity and effectiveness. However, scenarios like 'Vehicle Turning' need improvement, as their perception levels were notably lower.

Overall, this thesis highlights the critical importance of clear, adaptable and public-focused communication in driverless minibus-pedestrian interactions. It suggests that recognition of signs and signals is vital, but that the preference and context-specific understanding of preferences are equally important to ensure safe and effective interactions. In this thesis, text is assigned priority to facilitate a clear understanding of the information transmitted. Generally speaking and readable texts are easy to understand and effectively communicates knowledge without confusion.

7.7 Implementation and Integration of the Interaction Module

The final part of this chapter explains implementation and explains the practical applications associated with integration with other modules. It delves into the details of content that facilitates interaction by using a series of experiments to refine and distill this content and improve the effectiveness and utility of interactions. This is the focus of a careful implementation process to ensure the seamless integration of interaction modules into a broader navigation architecture of the Autobus. The deployment of the module is crucial to the effective and safe autonomous navigation of the vehicle, which represents the tangible interface between the vehicle and the pedestrian environment through which it communicates and interacts.

The block diagram depicted in Figure 7.13 defines the structure of the interaction modules in the group, which are mainly designed to communicate with pedestrians nearby. The central component of the system is the activation module, which acts as a switch and triggers interaction protocols based on navigation results obtained from the "decision-making" component (Chapter 6). At the time of activation, this module orchestrates a coordinated response to three output channels: display, voice and blinker. These channels are supplemented by a content database, a repository that provides the necessary content for each mode of interaction. This database is essential for providing visual data for displays, text for voice channels and blinker with a specific signal speed. The interactive content is selected from the content database module and transmitted from the respective mode to ensure that communication is not only appropriate to the context, but also compatible with the predefined protocol.

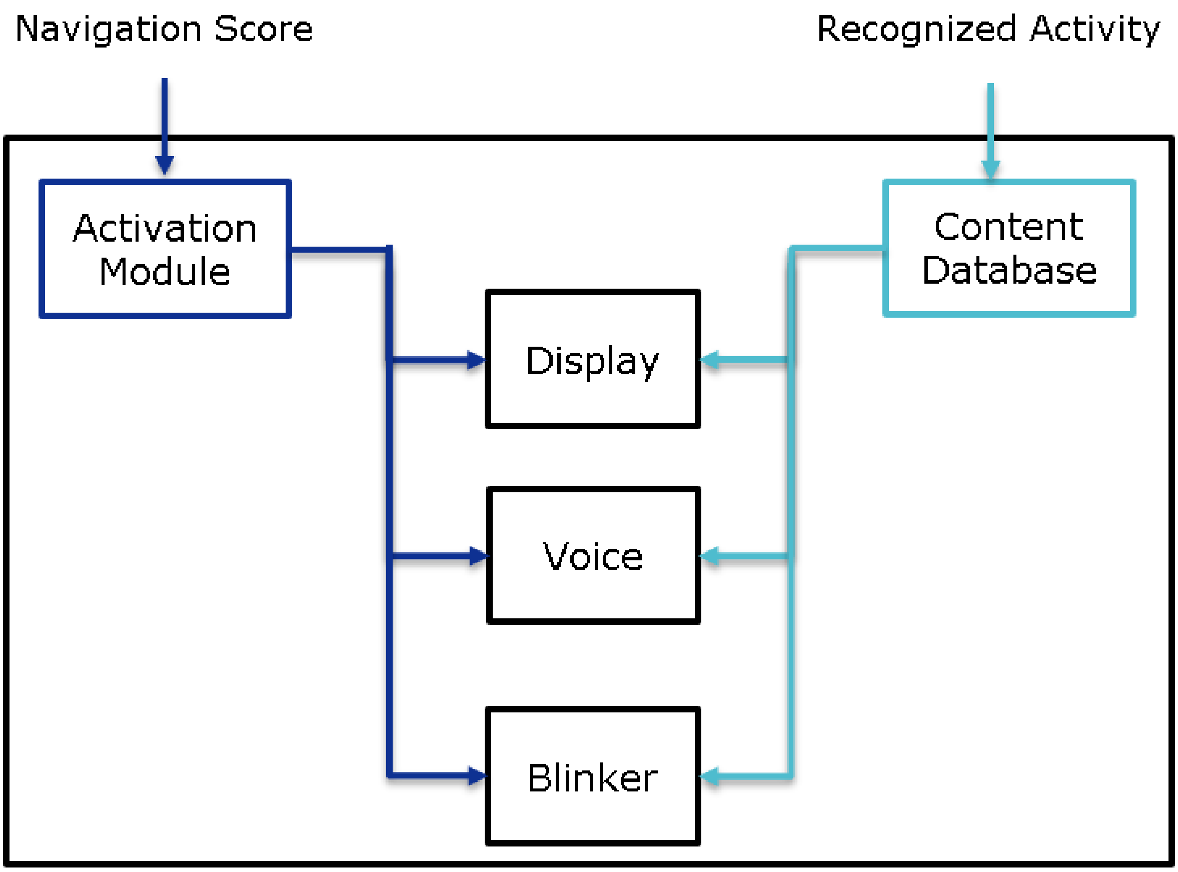


Figure 7.13: At the core of the module lies the Activation Module, tasked with initiating interaction procedures based on the Navigation score derived from the Decision-Making component (detailed in Chapter 6). The Activation Module triggers a synchronized operation across three communication channels: Display, Voice, and Blinker. The integration with the Content Database is pivotal, serving as abundant source of tailored content for each communicative avenue—supplying visual information to the Display, textual data to the Voice channel, and signal timing to the Blinker. The strategic distribution of interaction content from the Content Database to the output channels ensures the conveyance of context-specific and protocol-compliant communications to pedestrians.

The activation and participation of these interaction elements are greatly influenced by the associated risk assessments derived from the interaction fields (Chapter 5). An increase in the risk level may require a system to respond more immediately or more strongly to mitigate potential hazards. On the other hand, the reduction in risk could lead to a subdued or passive response from the system. This layered response strategy ensures that the system's response is in line with the risk level detected, thus improving the efficiency and effectiveness of pedestrian and Autobus interaction. The diagram in Figure 7.14 shows the activation of modules according to pedestrian risk assessments. In the field of ambient interaction, at the lowest risk level, the system uses LEDs to communicate the intentions of the vehicle. At moderate risk levels, within the direct and risk interaction domains, the auditory signals are activated together with the LED displays to communicate with pedestrians. When entering a high-risk zone in the risk area, the system activates a blinker to indicate the imminent danger.

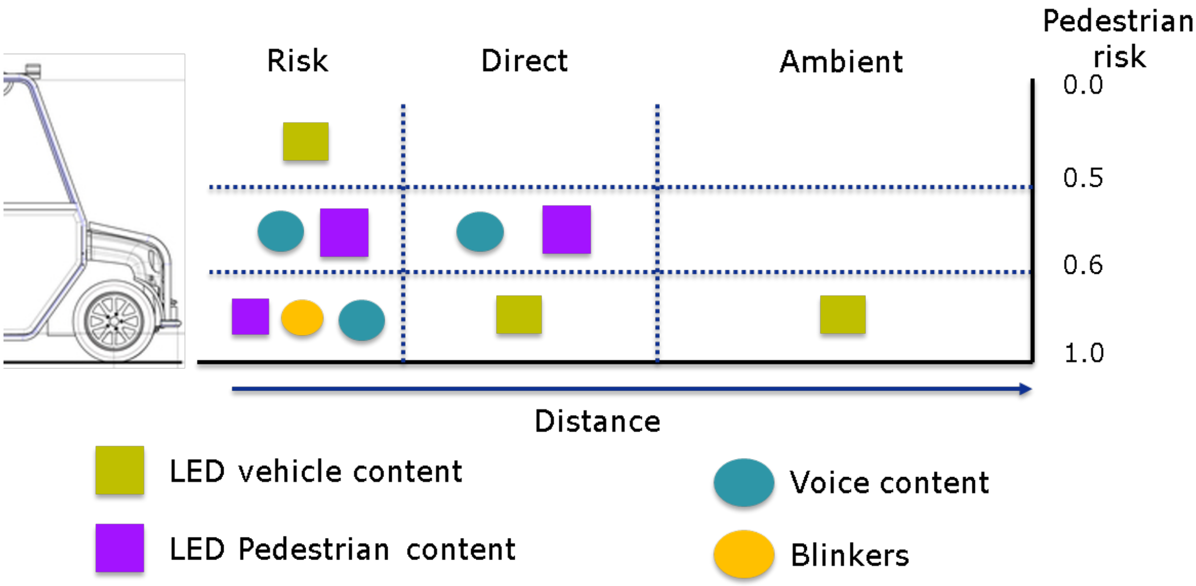


Figure 7.14: The figure illustration delineates the correlation between the risk levels ascertained from the interaction fields (discussed in Chapter 5) and the corresponding activation of the Autobus' communication modules. For minimal risk situations within the ambient interaction domain, the vehicle deploys LED signals. As the risk escalates to a moderate level within the direct interaction and risk domains, the system concurrently employs auditory cues alongside LEDs to enhance communication clarity with pedestrians. In high-risk zones, blinkers are engaged to provide an immediate and conspicuous warning. The strategic gradation of communication methods depicted here demonstrates the system's adaptive response mechanism, ensuring proportional reactivity to the dynamic pedestrian risk assessments.

This chapter concludes the examination of the fourth and final component of the smart interaction concept. The subsequent chapter will synthesize the integrated findings pertaining to the overall performance of the Autobus when operating within a campus setting.

8. Towards an Integrated Interaction Framework for Autobus in Pedestrian Zones

Transitioning from the fundamental four components of smart interaction strategy (Pedestrian Activity - Chapter 4, Interaction Fields - 5, Decision Making - 6, and Interaction Modules - 7), which meticulously detailed the technical development and theoretical underpinnings of individual components in Autobus navigation architecture, we now advance to the experiments of the whole architecture. This subsequent chapter serves as a critical juncture, shifting focus from theoretical concepts and system design to the practical application and real-world testing of the Autobus in pedestrian zones. Here, the implementation strategies elucidated in the previous chapter are put to the test, providing empirical evidence of their efficacy. This phase is pivotal as it bridges the gap between the conceptual framework and its tangible outcomes, offering insights into the functionality, efficiency, and reliability of smart interaction under actual operating conditions. The experiments conducted are designed to rigorously evaluate the various components of the system when integrated in as a whole.

The experiments conducted aim to validate the efficacy of these modules in real-world scenarios, demonstrating the vehicle's enhanced ability to communicate with pedestrians and thereby reducing unnecessary stops and interruptions in its journey. The outcomes of these experiments are critical in showcasing the practical applicability and effectiveness of the developed interaction system in managing the intricacies of autonomous driving in pedestrian zones.

8.1 Experiment and Evaluation

This section of this thesis delves into developing and testing various interaction strategies and facilitate clear communication between the Autobus and pedestrians. These strategies aim to alert pedestrians, especially those not initially aware of the vehicle's presence, and to comprehend pedestrian intentions effectively. The experimental framework is structured around four pivotal elements: identifying pedestrian behavior (Chapter 4), establishing interaction zones (Chapter 5), decision-making (Chapter 6), and the activation of interaction modules (Chapter 7). Each component of the interaction system was initially tested in isolation to assess its effectiveness, results given in their respective chapters. Following this, they were integrated into the standard navigation system, tailored for off-road environments.

8.1.1 Addressing the Complexity of Pedestrian Behavior in Autobus Testing

Before examining the test scenario presented in this thesis, it is important to recognize the numerous challenges associated with the test system that interacts with pedestrians. Pedestrian behaviour is chaotic and poses various challenges in assessing specific interactions between Autobus and pedestrians. This uncertainty introduces complexity in modeling, predicting, and testing to ensure the safety and reliability of these technologies in real environments. Due to the random and varied behaviour of pedestrians, the test challenge in certain scenarios can be examined from multiple perspectives.

Behavioral Unpredictability Unlike programmed machines, pedestrians do not follow predictable rules or behaviours. Factors such as individual differences, social influences, environmental context, mental state, even temporary distraction can dramatically alter the actions of pedestrians. For example, people who usually wait for the pedestrian signal to change before crossing can decide to take the jaywalk if they are running late. This variability means that even the same person may react differently to the same situation at different times, making it more difficult for the Autobus to predict and respond to pedestrian movements.

Scenario Complexity The examination of Autobus interactions with pedestrians must take into account almost infinite scenarios influenced by variables such as weather conditions, time of day and number of people. The complexity increases when dealing with interactions in specific contexts, such as the presence of children, where children's behaviour can be even less predictable than that of adults. Designing tests that simulate these different conditions accurately and comprehensively to evaluate the responses of the driverless minibus is a significant challenge.

Safety and Ethical Constraints The uncertainty of pedestrian behaviour means that ensuring the safety of all participants during the test is important but also complex. It is ethically unacceptable to put pedestrians in potentially dangerous situations to observe how Autobus or pedestrians react. As a result, most of the preliminary tests were conducted in simulated environments or with staged scenarios that do not capture the full range of realistic human behaviour. This limitation raises questions about the completeness and applicableness of the test results in the real world.

Data Collection and Analysis Challenges The collection of data on pedestrian behaviours in a meaningful way for Autobus testing involves significant challenges. A large number of possible behaviours and scenarios require collection and analysis of huge amounts of data to identify patterns, anomalies, and potential responses. This process is complicated by the need to respect privacy and obtain consent, especially in public places where pedestrian behaviour is more diverse and unpredictable.

Adaptation and Evolution Pedestrian behaviour is not static; with the passage of time, social norms, urban environments change, and new technologies are introduced, such as new modalities on the Autobus itself. This evolution means that even an autonomous driving system that has been thoroughly tested and initially effective may need constant updates and re-evaluations to keep pace with changes in human behavior.

Despite these challenges, we strive to capture some of the behaviors present during testing and observe how the Autobus responds.

8.1.2 Behavioral Response of Pedestrians to Visual and Auditory Signals

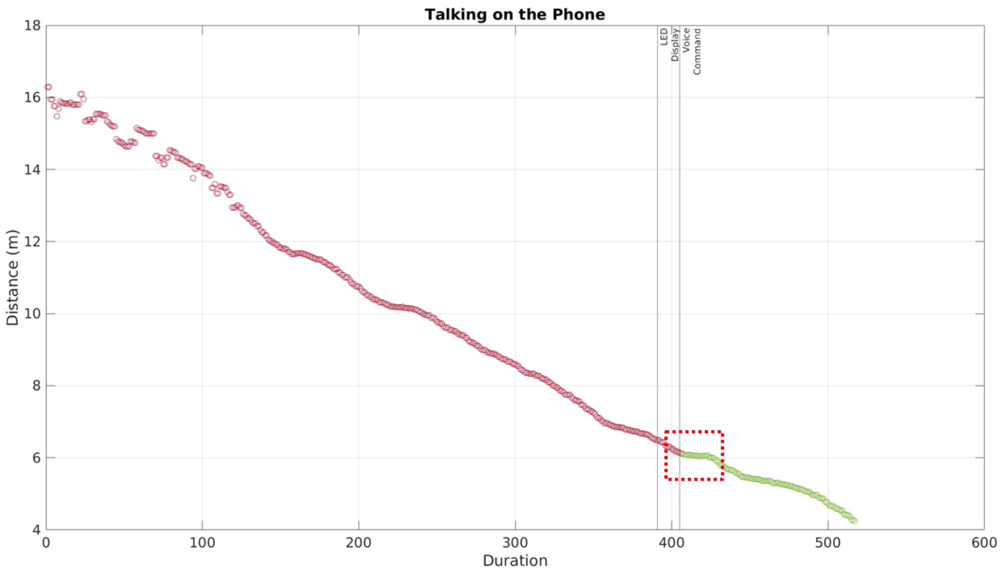
Considering the difficulties outlined in the prior subsection, this dissertation delineates and details the most common and unique scenarios that the Autobus encountered within the campus milieu. The investigation aims to address the concerns of safety, efficiency, and passenger experience through the judicious selection of test scenarios for evaluation. Verification of outcomes is achieved by monitoring alterations in pedestrian awareness and movement, as represented in the figures and plots. The plots have been presented with corresponding images to see how the environment looks like where the situation was encountered. The ensuing sections provide an in-depth examination of the selected test scenarios.

Pedestrian Engaged in Telephonic Conversation Unaware of Autobus

An experiment was performed to assess the effectiveness of various signaling modalities in attracting the attention of pedestrians engaged in telephone conversations while walking along the path of an approaching Autobus. Given that the pedestrian was distracted and the Autobus was navigating a narrow pathway (as depicted in Figure 8.1a), avoiding the pedestrian was not feasible, necessitating prioritized interaction by the decision-making module. Figure 8.1b displays the pedestrian's temporal response to auditory and visual signals from the Autobus. The graph, marked with circles on the line, indicates the separation distance between the pedestrian and the vehicle, with red highlighting the pedestrian's initial lack of awareness, as indicated by their orientation away from the Autobus. Although a visual cue was provided through LED signaling, the pedestrian remained unaware, due to looking at the opposite side. In contrast, the introduction of a voice command led to a notable shift in the pedestrian's orientation towards the Autobus, indicating a recovery of situational awareness, as marked by a change to green on the graph. This regained alertness is corroborated by the subsequent horizontal line (enclosed in a red dotted box), suggesting a maintained distance between the pedestrian and the Autobus, indicative of the pedestrian making space for the vehicle.

In summary, the graphical representation clearly delineates the crucial shift from distraction to awareness facilitated by an auditory stimulus, highlighting the essential role of effective

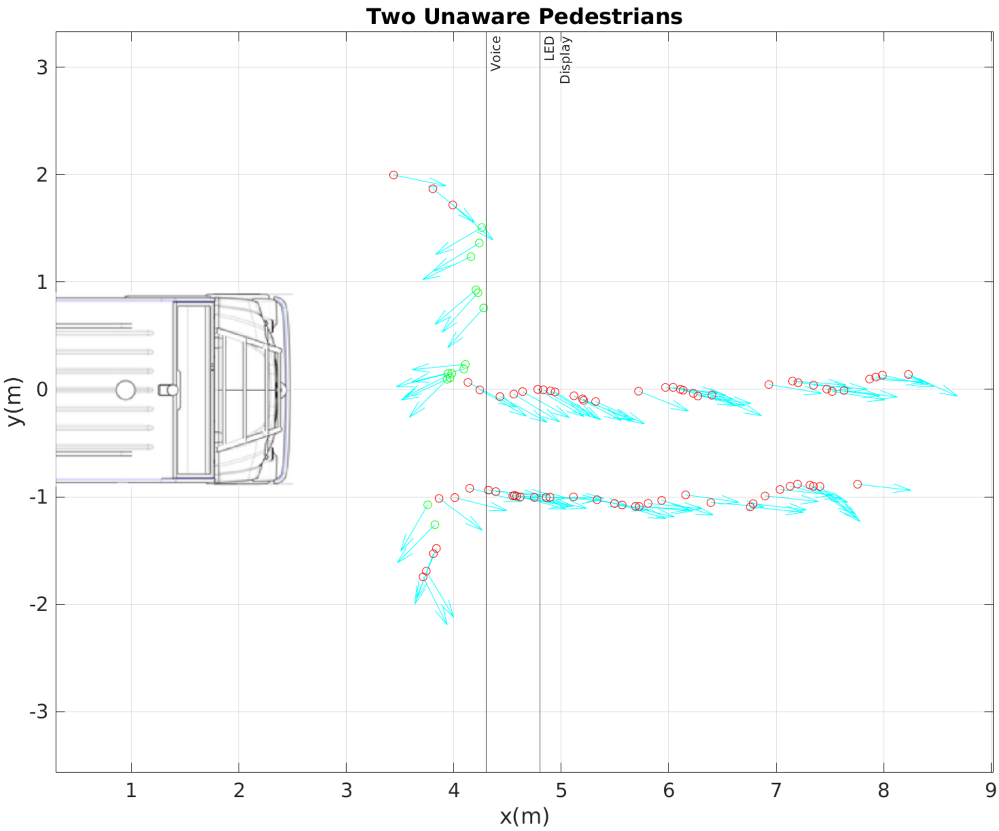
(a) The figure shows a scenario where a pedestrian walking infront of the Autobus in narrow space talking on the phone. The pedestrian is unaware of the approaching vehicle behind.



(b) A pedestrian's response to voice signaling while on the phone: red circles indicate unawareness; awareness is signified by green circles upon receipt of a voice command.

Figure 8.1: Unaware pedestrian talking on the phone reacting to voice command from the Autobus.

(a) The illustration depicts a scenario where two pedestrians are walking ahead of a vehicle, engrossed in conversation and unaware of its presence.



(b) The figure presents a visual analysis of pedestrian response to auditory signals from Autobus in a simulated environment. The vehicle, depicted at the coordinate origin, serves as a reference point for assessing pedestrian awareness and orientation. Initially, two pedestrians, represented by red dots and indicated by their respective trajectories and orientation arrows, are observed to be unaware of the vehicle's presence. Upon activation of a voice signal by the vehicle, a marked change in pedestrian behavior is noted. The red dots transition to green, suggesting a heightened awareness of the vehicle. Concurrently, there is a noticeable shift in the direction of movement, as illustrated by the orientation of the arrows. This reaction demonstrates the efficacy of auditory signaling in alerting pedestrians, thereby potentially enhancing safety measures in autonomous vehicle operations.

Figure 8.2: Unaware pedestrian crossing in front of the Autobus yields as soon as the pedestrian sees the yielding signal on the LED.

audio signal transmission in ensuring pedestrian safety in scenarios where visual signals, such as LEDs, go unnoticed.

Two Unaware Pedestrians Walking and Engaged in Conversation

Figure 8.2 depicts a scatter plot superimposed over a grid, representing the interaction between Autobus and two pedestrians in narrow path. The Autobus is illustrated at the origin (0,0) of the plot, detailed with a top-view image. Two distinct pedestrian trajectories are plotted on the graph. Both pedestrians are captured from 8 meters. As the vehicle approaches them, the pedestrian can be seen getting close to the vehicle.

The pedestrians initially proceeding along their path, oblivious to the approaching vehicle from behind as can be seen in the camera view (Figure 8.2a). Upon approach, the Autobus activates its visual signaling system. However, due to the orientation of the pedestrians—away from the vehicle—the visual cues fail to capture their attention, and they continue along their trajectory undeterred. Consequently, as the Autobus draws nearer, it resorts to an auditory signaling. This auditory prompt evokes a response from the pedestrians, leading them to diverge to opposite sides of the path, thereby allowing the Autobus to pass unimpeded. Under typical conditions, one might expect that two pedestrians entering as a pair would yield to the right, since both tend to be slightly right of the Autobus. However, the pedestrian on the left chooses to separate from the company and move aside in the opposite direction as can be seen from the plot in Figure 8.2b. This interaction underscores the importance of multi-modal signaling in Autobus to ensure effective communication with pedestrians and the safe navigation of around people.

Pedestrian Crossing Yielded

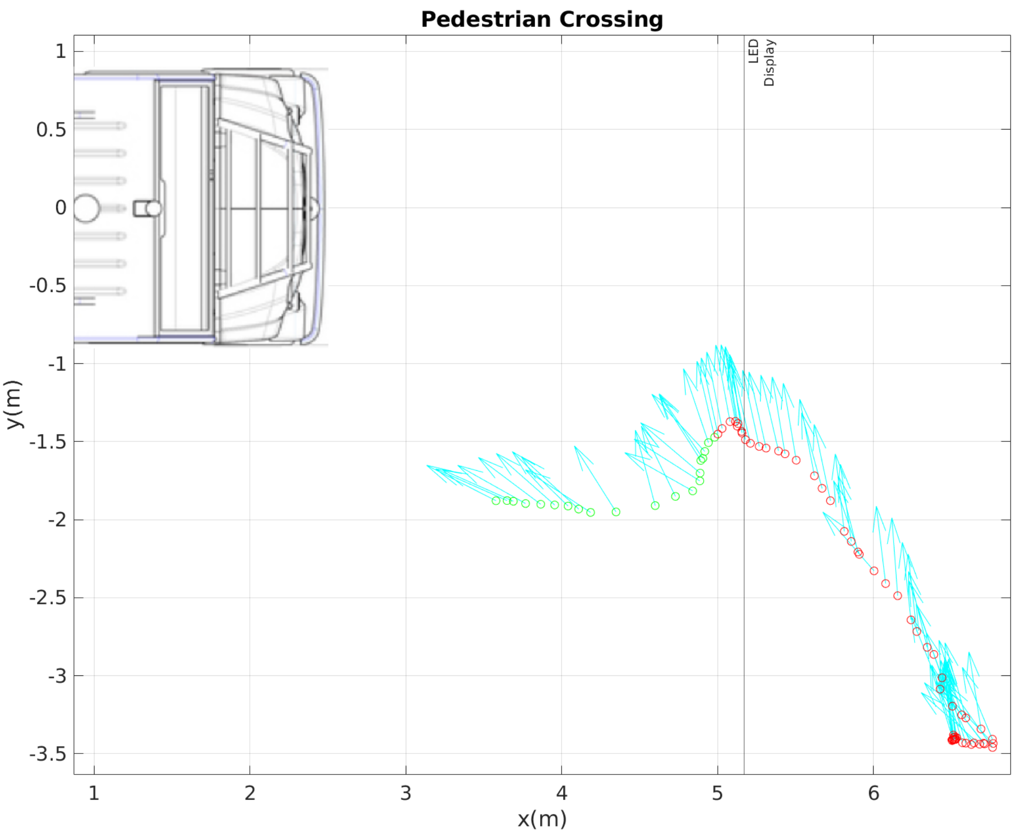
Figure 8.3b shows the interaction between pedestrians and buses in the cross scenario. The Autobus is shown on the left side of the map, and a side view graphic is placed at the source of the coordinate system. The pedestrian path is mapped using a series of points connected by line segments to follow the trail. The dots change in color from red to green, where red indicates the pedestrian's initial unawareness and green marks the transition to awareness of the Autobus. This change coincides with the activation of the display signal from the vehicle, which warns pedestrians not to cross.

The arrows originating from the dots indicate the direction in which the pedestrian is moving. Initially, these arrows move towards the Autobus, but after the display signal is activated, the pedestrian yields and complies with the vehicle instructions. Figure 8.3a provides a complementary view, showing a pedestrian who has a smartphone in texting position and who originally planned to cross before the Autobus. The pedestrian stops after receiving the signal and the vehicle passes safely.

An Incident of Non-compliance: Aware Pedestrian Standing in Front

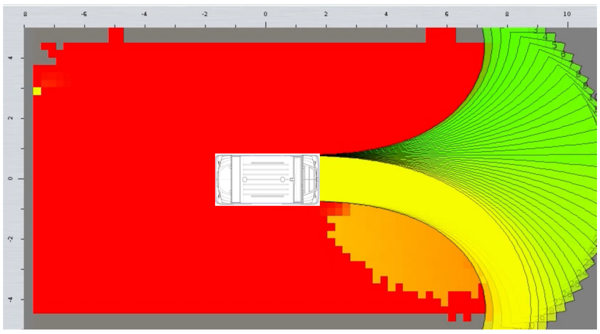
Figure 8.5a illustrates a scenario where two pedestrians, aware of the Autobus, approach the vehicle. As they near the vehicle, one individual steps directly into its path. According to their awareness of the Autobus, the interaction fields register low awareness values. Upon the pedestrian's close approach, both visual and auditory signals are activated. However, the pedestrian ignores the visual cues displayed on the vehicle's LED panel and positions themselves in front of the bus before returning to their original path, as shown by the cyan arrows in Figure 8.5b. This incident exemplifies a distinctive category of pedestrian behavior that appears to be influenced by factors such as curiosity or amusement

(a) The figure illustrates a scenario in which a pedestrian is cross in front of the Autobus. The pedestrian is preoccupied with their mobile device, using it in a texting position. The pedestrian becomes aware immediately after receiving a signal from the Autobus.

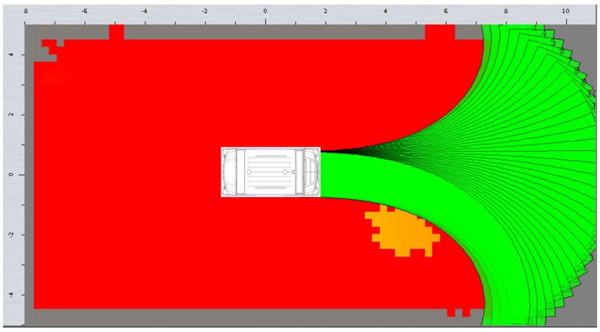


(b) Scatter plot demonstrating a pedestrian's trajectory alteration in response to a visual 'Do Not Cross' signal from an Autobus. The initial pedestrian pathway, indicated by green dots and arrows, shows movement towards the vehicle's trajectory. Upon activation of the visual display, the pedestrian ceases forward motion, and stops, allowing the Autobus to pass safely.

Figure 8.3: Unaware pedestrian crossing in front of the Autobus yields as soon as the pedestrian sees the yielding signal on the LED.



(a) The image illustrates the pedestrian tentacle, when the pedestrian is unaware. Due to this unawareness, the ellipse has higher values and a larger radius because the pedestrian is walking towards the Autobus. The corresponding tentacle is shown in yellow, indicating higher values.



(b) The image shows when the pedestrian becomes aware of the Autobus after the signal is given. The green values of the pedestrian tentacle shows the it is safe to drive.

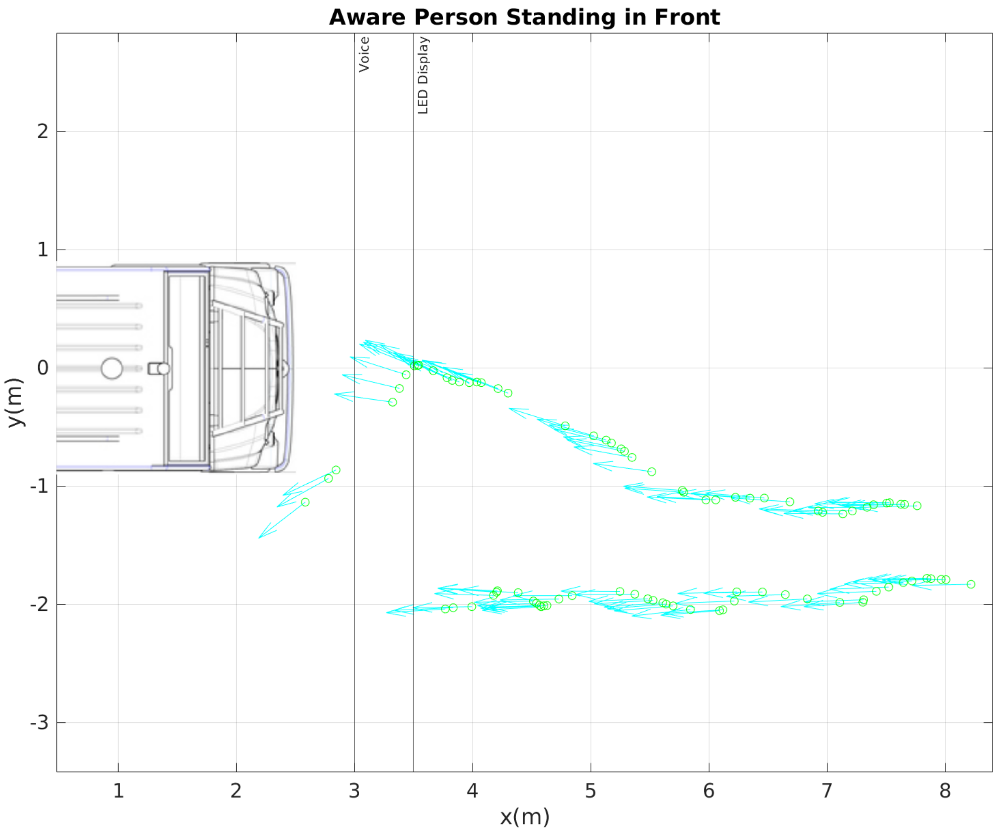
Figure 8.4: Illustration of the pedestrian tentacle in scenarios where the pedestrian is unaware (a) and becomes aware (b).

rather than the vehicle's signals or speed. The dataset contains multiple instances of such unconventional pedestrian behaviors that are unresponsive to the vehicle's communicative signals, ultimately necessitating the Autobus to initiate a full stop for safety considerations.

The above-mentioned unique scenarios show how the Autobus navigates in unexpected situations using smart interaction strategies. This approach improves vehicle efficiency by warning pedestrians not aware of the presence of the vehicle. Through its interactive modules, the vehicle often encounters pedestrians engaged in conversation, texting or calling. The interaction system communicates effectively the presence of the vehicle to them. Most pedestrians courteously make way for the Autobus when they receive the signal. However, the combination of Autobus slowdown and signaling confuses pedestrians, and there is uncertainty about what steps should be taken. Some people leave the path, while others, especially those in hurry, ignore the signals.

It is impractical to measure the time saved using the interaction system accurately compared to normal driving without such interaction. This is due to various influencing factors such as daytime, lecture program, type of area where interactions take place, and the size of pedestrian groups. These variables significantly skew the average values, making it difficult to draw a definitive conclusion about the efficiency of the system in reducing travel time. On the basis of the results presented in this chapter, there are obvious qualitative improvements. However, in order to quantify the extent of improvement, extensive tests are required over several years, taking into account variable such as time, school timetables, lunch breaks and weather conditions.

(a) The figure depicts a scenario in which two pedestrians are walking in front of the Autobus in a semi-narrow space, crossing its path while conversing. One of the pedestrians intentionally stands directly in front of the vehicle, causing it to halt despite the signal displayed.



(b) The plot shows two pedestrians approaching the vehicle from the front. The pedestrian on the left disregards the Autobus signals and stands directly in front of it, causing the vehicle to stop.

Figure 8.5: A situation where two aware pedestrian approaching the Autobus. One pedestrian ignores the signal from the Autobus and stands in front of the vehicle.

9. Conclusion

This thesis successfully developed a smart interaction strategy (Figure 2.5) suitable for driverless minibuses operating in pedestrian areas. An important element of this approach is the ability to adapt to different behaviors of pedestrians, which requires adapting different parameters. These adjustments are based on the size of a driverless minibus and specific characteristics of the pedestrian environment. Therefore, the autonomy of pedestrian areas requires special forms of interaction to ensure smooth operation. The concept of intelligent interaction is inspired by the way humans use to communicate with pedestrians. Implementing this tailored interaction strategy in driverless minibuses has demonstrated promising performance in real-world experiments.

The deficiencies identified in the modern driverless minibus systems highlight the crucial need for interaction mechanisms adapted to pedestrian dynamics and emphasize its crucial role in the effectiveness and safety of such vehicles. This thesis introduces an integrated framework based on four fundamental components to improve the communication between vehicle and pedestrians. The process begins with a detailed analysis of pedestrian behaviour to identify and understand their core activities and movements. The framework then establishes a pedestrian risk assessment zone around the vehicle and uses its location relative to the vehicle to inform safety measures. The decision-making aspects of the framework are then at the forefront, weighing the options between navigating around pedestrians, initiating direct interaction with pedestrians, identifying safe driving regions and responding to pedestrians' unawareness. The last step of this process is to deploy the interaction module. This module is responsible for managing a series of communication tools ranging from visual and auditory signals to more nuanced message transmission. It controls both the dissemination of information to pedestrians and the activation or deactivation of these communication mechanisms, ensuring the effective exchange of essential information. This comprehensive approach not only addresses the gap in pedestrian interaction, but also sets new standards for the development of driverless minibuses.

The results of this research are included in the existing navigation system REACTiON [Wolf 18] (Figure 3.20). This strategic integration greatly improves REACTiON's ability to navigate in pedestrian-dominated areas and ensures an optimal and effective driving experience. To rigorously evaluate the effectiveness of smart interaction strategies to address current limitations, a comprehensive test protocol using Autobus has been implemented (Figure 1.3). Autobus is a vehicle engineered with kinetic and dynamic characteristics similar to other driverless minibuses and has been an ideal platform for empirical research. The validation process is rigorously carried out in the campus environment of the RPTU Kaiserslautern (Figure 3.16), offering a realistic environment for measuring performance improvements attributed to smart interaction strategies. An important component of this assessment focused on Autobus interaction modules, especially in scenarios where pedestrians were previously unaware of the vehicle's presence. In such cases, activating these modules is an important aspect of this thesis, as it directly contributes to improving the safe navigation in the campus environment.

By implementing these interaction modules, the Autobus was able to effectively communicate with pedestrians, thereby significantly mitigating the risks associated with pedestrian unawareness. This enhancement not only underscores the potential of incorporating smart interaction strategies into autonomous navigation systems but also demonstrates the tangible benefits of such innovations in improving safety and efficiency in pedestrian-rich zones. The successful integration and validation of the smart interaction strategy within the navigation architecture mark a significant step forward in the development of autonomous driving technologies, offering promising avenues for future research and application in shared spaces of smart cities.

9.1 Evaluation

This research has made significant contributions to autonomous driving in pedestrian areas, especially by improving efficiency and safety through innovative pedestrian interaction strategies. It clearly shows that pedestrian signals, who are unaware of the existence of the Autobus in their surroundings, can significantly improve navigation results. The research promotes a deeper sense of trust and understanding between the vehicle and pedestrians by adapting traditional human driver signals such as visual and auditory warnings in the Autobus navigation system. This approach not only simulates the typical intuitive interaction of human drivers, but also integrates these indications into vehicle autonomous operation to ensure clearer intentions.

Each of the four interaction components of the smart interaction framework (Figure 2.4) plays an important role in the realization of this comprehensive process. From detection and assessment of pedestrian behaviour to decision-making and real-time navigation and engagement and finally, the management of visual and auditory signals, all elements contribute to a holistic interaction strategy. This multifaceted approach demonstrates the important role, mentioned in this thesis, in the development of autonomous driving technologies for pedestrian areas. It highlights how a precise and pedestrian-focused communication can significantly increase the efficiency and safety of driverless minibuses, and shows the practical application and relevance of these findings in the real world.

The pedestrian activity module provides a sophisticated approach to the recognition of pedestrian activity by using 3D skeleton joints (Example 4.1) through the LSTM model (Figure 4.5) and the geometric approach (Figure 4.12). This method significantly increases the speed of detection activities by focusing on a less but highly informative set of 3D point data. Activities classified in this thesis (Table 4.3) are the most commonly occurring in the close vicinity of the vehicle, including different pedestrian crossing patterns. These patterns not only help identify typical movements, but also detect signs of pedestrian unawareness, such as walking orientation. The strategic categorization of pedestrian activities, combined with the increase in data points, plays a key role in increasing the system's overall accuracy in recognizing and interpreting pedestrian behaviours. This thesis proved that the accuracy of the upper body (Table 4.7) is 95%, and the lower body (Table 4.9) is 99%. The inclusion of specific behaviors that indicate distraction, such as texting or using a phone during a crossing, adds a crucial layer to the system for understanding pedestrian risk (Table 4.10). With the geometric approach using iB2C (Description 3.5), additional/redundant attributes such as position, distance and movement can be obtained (Figure 4.13).

The development of an interaction field around the vehicle (Figure 5.4) and pedestrians (Figure 5.5) has introduced a novel method of assessing pedestrian risk. These fields are seamlessly integrated into existing navigation maps (Figure 5.14) and effectively distinguish pedestrians from other obstacles. This differentiation allows the Autobus to adapt its interaction more appropriately to pedestrians. The characteristics of pedestrians, including direction, awareness level, and movement patterns, are packaged into elliptical shapes and provides a brief, clear and complete representation within interaction field of the vehicle. This representation is crucial for understanding the behaviour of pedestrians and for translating complex human activities into a form that can be easily understood by autonomous driving systems. Dividing the vehicle's interaction fields into ambient, direct and risky zones further improves the system's reaction to pedestrians. This structured representation greatly improves the decision-making process by identifying the type of interaction required, be it simple recognition or more urgent avoidance operations. This representation helps to generalize the number of different pedestrians (e.g. groups) in different areas depending on their level of awareness (Figure 5.13).

The decision-making module within the framework played a pivotal role in navigating pedestrian encounters by evaluating the assessed risk levels. By using separate set of tentacles for pedestrian risk named as "pedestrian tentacle" and drivable pathway as "feature tentacle", the decision focuses on the awareness level of the pedestrian instead of accounting it for normal obstacle (Figure 6.14). When it detected a pedestrian who appears to be unaware of the vehicle's presence, the system initially looked for an alternative path that avoided the need for direct interaction. Based on sum of the weight from feature and pedestrian tentacle (Equation 6.1), it checks for best strategy. If no such path was available (Figure 6.8) or if the pedestrian's actions necessitated closer communication, the system then engaged the interaction module.

Interaction module was responsible for activating specific channels of communication (Figure 7.3), such as LED displays, voice messages, and flashers, tailored to address the situation effectively. The selection of content (Figure 7.4) for these communication tools was meticulously refined through extensive testing with people by the use of questionnaire (Appendix B) to ensure clarity and comprehensibility. Pedestrians demonstrated a high level of responsiveness to these cues, as evidenced by unaware individuals making way for the vehicle following voice prompts. Similarly, pedestrians adjusted their walking paths upon reading messages displayed on the LED screen, facilitating smoother navigation for the Autobus. It can be seen from results of phase-I that "Man with Arrows" for most recognized, but "Red bus" is preferred (Figure 7.10). This shows that, recognition and preference of shown content can vary. For phase-II, 60% of the recognition and people following the content was above 95% (Figure 7.12). These interactions underscore the effectiveness of the chosen communication strategies in enhancing pedestrian safety and vehicle navigability.

Using the smart interaction strategy enabled in the Autobus, it has demonstrated remarkable flexibility in driving strategies, adjusting to different situations in a way that greatly improves the overall passenger experience. With a smart interaction strategy, the Autobus could drive smoothly wherever pedestrians were unaware of the vehicle. The brakes were significantly reduced by giving such signals to pedestrians. People who recognized the intentions of the Autobus yielded, resulting in a reduction in braking when the vehicle was driving around the campus. This dynamic adaptation capacity to the crowded environment not only improves the safety and efficiency, but also ensures the safety and seamless transport of the passengers in the vehicle. The inherent modularity of the design is especially advantageous and allows additional specialized controls to be integrated, when needed. This modular architecture also allows individual components to be updated without affecting the functionality of other components, so that the system can evolve over time and have minimal disruption. Such a framework supports continuous improvements and adaptations to emerging technologies and changing conditions and highlights the potential of the Autobus for the future development of autonomous transportation. This thoughtful approach to system design highlights the readiness of vehicles to meet current and future challenges and is a pioneering solution in the field of driverless minibuses for pedestrian zones.

9.2 Outlook and Future Work

Future work, informed by the discussions and findings within this thesis, will pivot around four core areas: enhancing pedestrian interaction models, refining autonomous decision-making capabilities, and the usage of interaction modules. Moreover, it also discusses the modular system integration and updates. The thesis has demonstrated that the smart interaction strategy within Autobus significantly elevates the driving strategy, particularly in pedestrian-rich environments, suggesting that further advancements in cognition and perception algorithms could yield even greater improvements in driverless minibus driving performance.

Firstly, enhancing pedestrian interaction models involves developing more sophisticated algorithms for interpreting pedestrian behaviors and intentions. This could lead to a deeper understanding of complex semantic scenes, enabling driverless minibuses to navigate more effectively in dynamic urban environments. An emphasis on novel classification algorithms could allow for the recognition of a broader array of objects and pedestrian activities, thus enhancing the vehicle's semantic evaluation capabilities. Such algorithms are expected to broaden the scope of detectable objects and pedestrian actions, thereby elevating the vehicle's capacity for semantic analysis. This advancement will extend to the nuanced detection of pedestrian features, including facial expressions and the presence of handheld objects, like mobile phones or beverages. Recognizing facial expressions could unveil new layers of insight into pedestrian emotional states, potentially informing the vehicle's interaction strategy in a more nuanced manner. For example, detecting a pedestrian's distracted demeanor could trigger specific cautionary measures by the autonomous system. Expanding the scope of pedestrian interaction models to include the detection of animals and bicycles represents a significant advancement in autonomous vehicle technologies. These entities, which frequently share urban spaces with pedestrians, pose unique challenges and opportunities for driverless minibuses navigating dynamic environments. By developing algorithms capable of identifying and interpreting the behaviors and trajectories of both animals and bicycles, autonomous systems can achieve a more comprehensive situational awareness. Additionally, increasing redundancy in the system could help mitigate semantic and measurement errors, contributing to a more robust and reliable autonomous driving system.

Secondly, advancing autonomous decision-making capabilities is crucial. A key aspect of this advancement is the development of sophisticated decision-making frameworks capable of navigating the complex terrain of ethical considerations. The decision-making can be improved by incorporating additional features of both the pedestrian and the environment. Integrating ethical decision-making processes into autonomous systems requires a multidisciplinary approach, drawing on insights from ethics, law, psychology, and robotics to establish guidelines that ensure decisions are made in a manner that is both morally responsible and aligned with societal values. Future research should focus on creating systems that are not only technically proficient but also possess an embedded ethical reasoning capability. This would allow vehicles to make informed decisions in critical situations, balancing the need for safety, efficiency, and ethical compliance. Later, on-demand mobility [Husemann 22], which will utilize a fleet of driverless minibuses, should also include scheduling in its decision-making processes. Such advancements would significantly enhance the trustworthiness and reliability of driverless minibuses, paving the way for their broader acceptance and integration into daily life.

In addition, the potential integration of more sophisticated interaction modules, such as the direct projection of visual signals onto the path of unaware pedestrians, improves safety and clarity in complex or unclear situations. The development of a deep learning model using content from more casual conversations, rather than relying exclusively on fixed databases, will facilitate a more relaxed interaction between pedestrians and vehicles. These models should enable dynamic adaptation of communication strategies based on specific contexts, and ensure that the most effective methods of interaction are always used. This capability is essential for the scalable and flexible deployment of driverless minibuses that adapt to a wide range of environments and scenarios.

By addressing these areas, future research aims to not only refine the capabilities of such interaction strategy in Autobus and similar driverless minibuses but also to contribute to the broader field of autonomous driving and human-robot interaction, ensuring these systems can safely and effectively integrate into human-centric environments.

A. Appendix A

A.1 Safety System

The appendix presented herein details the components of safety system utilized in the Autobus, including technical specifications and parameters necessary for comprehending the system's configuration as outlined in Chapter 3. The components used in the safety system schematics (Figure A.2) are shown in Figure A.1 and explained below:

- **CPU (FX3-CPU000000):** The CPU serves as the primary module, processing all incoming signals from connected modules. It executes actions based on sensor inputs according to the valid configurations stored within its system plug. The included logical editor in the SD software facilitates the implementation of control logic tailored to specific requirements.
- **Gateway (FX3-GEPR00000):** This gateway module facilitates the connection of laser scanners and the external system to the CPU via Ethernet, using the Safety Designer software. It supports EFI-pro and EtherNet/IP™ CIP Safety™ protocols, allowing data transmission rates ranging from 10 Mbit/s on a 10Base-T network to 100 Mbit/s on a 100Base-TX network. The module features autosensing capabilities and a configurable Requested Packet Interval (RPI) from 4 ms to 500 ms for packet production and 1 ms intervals for packet consumption.
- **Input/Output (FX3-XTIO84002):** Positioned adjacent to the gateway, this module receives various signal inputs from external sources and provides corresponding outputs. It features a rapid shutdown response time of 8 ms and includes four safety outputs, which are primarily used for controlling relay activation.
- **Digital input (FX3-XTDI80002):** This module captures digital signals from external safety switches. Upon signal deactivation, it communicates with the CPU to trigger appropriate responses. It is equipped with eight safety inputs.
- **Motion controller (FX3-MOC100000):** Employed to capture encoder readings from steering and motor control units, this module translates these readings for

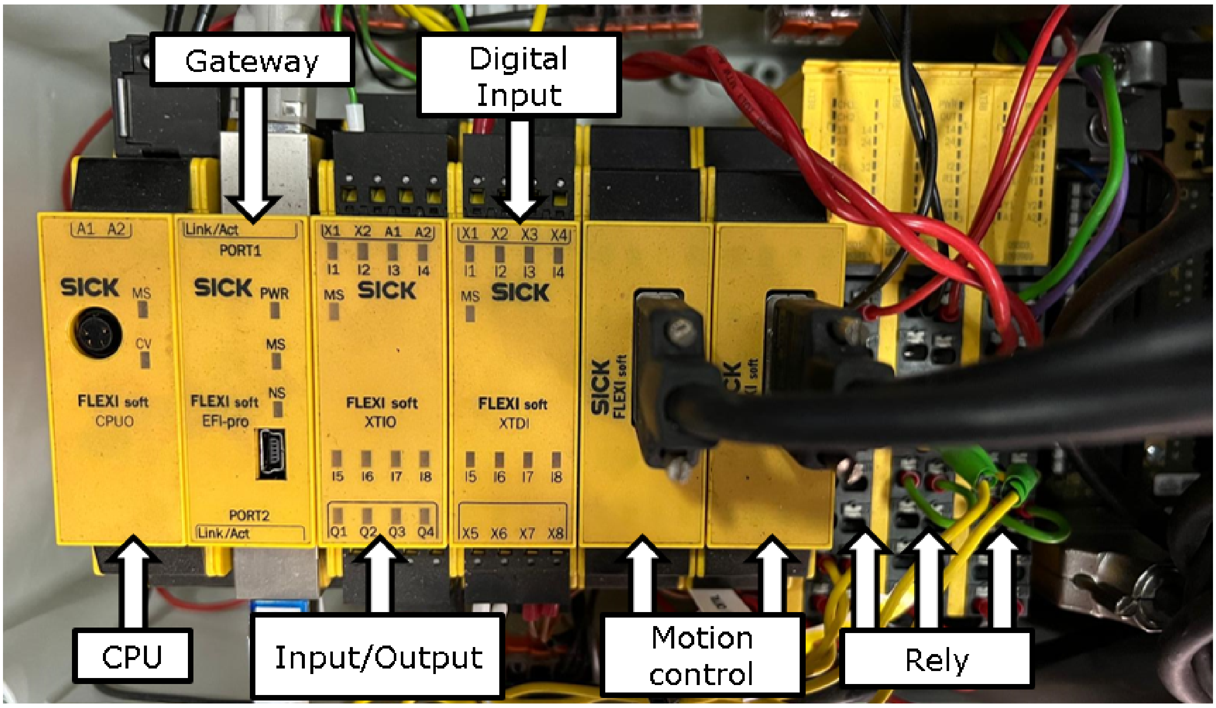


Figure A.1: An integrated safety system module featuring various components of a SICK Flexi Soft safety controller. The Central Processing Unit (CPU) (FX3-CPU000000) on the left, responsible for processing signals from the system's modules based on preconfigured logic. Adjacent to the CPU is the Input/Output Module (FX3-XTIO84002), which manages external signal inputs and actuates safety outputs. The Gateway Module (FX3-GEPR00000) on the right of CPU facilitates Ethernet connectivity, supporting EFI-pro and EtherNet/IP™ CIP Safety™ protocols for network communication. The Digital Input Module (FX3-XTDI80002) is placed next to the Gateway, handling digital inputs from safety switches and signaling the CPU for appropriate action upon deactivation. On the right, the Motion Control Module (FX3-MOC100000) is dedicated to interpreting encoder signals from motion-related components. Finally, multiple Safety Relays (RLY3-OSSD100/200/300) are seen on the far right, which ensure the rapid disengagement of the safety circuit to enable emergency stops, with individual relays dedicated to managing different safety-critical outputs, such as flashers and braking systems.

subsequent processing by the CPU. Two motion controllers are utilized, one for steering and the other for speed regulation, supporting various drive safety functions including safe stop, safely-limited speed, safe direction, and safe limited position.

- **Rely (RLY3-OSSD100/200/300):** This relay is crucial for interrupting the main safety circuit to facilitate emergency stops. It receives inputs from the Input/Output module. The deployment of three relays is strategic; two relays manage different flashers, while the third controls braking mechanisms.

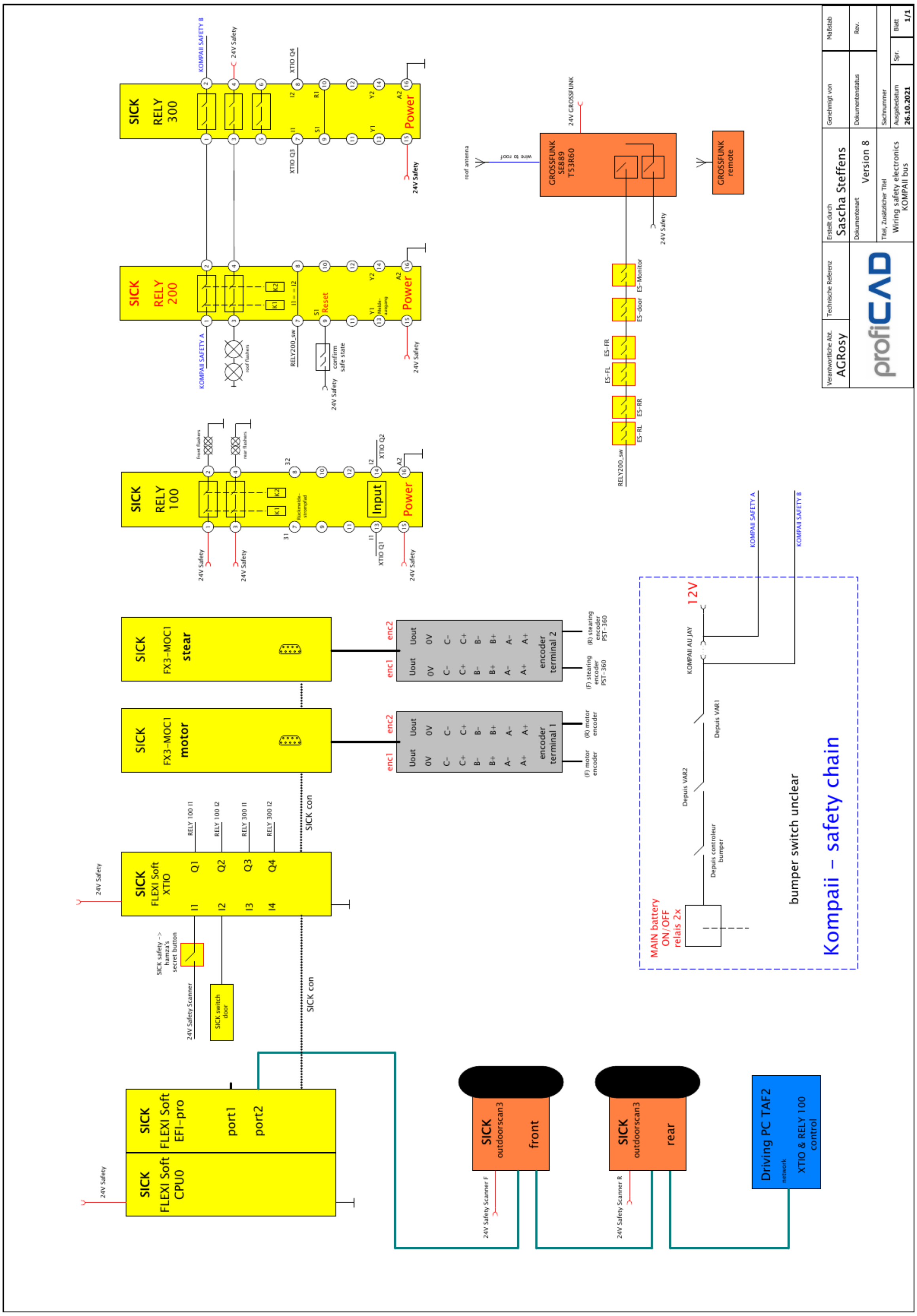


Figure A.2: Safety system schematics.

B. Appendix B

B.1 Questionnaire

The following section introduces a detailed questionnaire tailored to support the empirical data presented in Chapter 7. This questionnaire is the result of a comprehensive effort to identify the most effective methods of content delivery for the Autobus's LED displays. It is designed to explore participants' preferences and recognition experiences, transforming a variety of factors into quantifiable data that define 'optimal content'.

Structured to elicit a wide range of information, the questionnaire starts with basic demographic queries to frame the data within a context of varied participant backgrounds. It then advances to explore a mix of qualitative and quantitative metrics that assess aspects such as the lucidity, engagement, and overall effectiveness of the content displayed. Likert-scale items measure responses to the content's visual appeal and memorability, while open-ended questions encourage participants to share unstructured, in-depth feedback.

Each question comes with clear instructions to ensure comprehensive and accurate participant responses. The responses obtained are critical for the statistical analysis that follows, with outcomes which significantly contribute to the iterative enhancement of the LED interaction modules discussed in the chapter.

Personal Information

Name:	____________________
Age:	____________________
Sex:	____________________
Occupation:	____________________
Date:	____________________

Test Phase I

1. Have you been part of this kind of test before?
   Yes No

2. Is the sign/text/animation understandable?
   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree
3. The color used in the sign/text/animation are easy visible?
   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree
4. Is this sign/text/animation Recognizable?
   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree
5. Is this sign/text/animation preferred?
   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree
6. Do you feel safe with this type of sign/text/animation?
   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree
7. Do you have any extra comment regarding the displayed content?

   ____
8. Do you have any other preferences regarding the shown content?

   ____

Test Phase II

1. Were you part of "Phase I" testing?
   Yes No
2. Is the sequence understandable?
   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree
3. Is this sequence Recognizable?
   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree
4. Is this sequence preferred?
   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree
5. Do you feel safe with this type of sequence?
   ☐ Strongly Agree ☐ Agree ☐ Neutral ☐ Disagree ☐ Strongly Disagree
6. Do you have any extra comment regarding the displayed content?

   ____
7. Do you have any other preferences regarding the shown content?

   ____

I acknowledge that my participation in this questionnaire is entirely voluntary. I have been informed that I am free to discontinue my involvement at any stage without providing a reason and without any negative consequences.

____

Bibliography

[Albawi 17] S. Albawi, T. A. Mohammed, S. Al-Zawi, "Understanding of a convolutional neural network", in *2017 international conference on engineering and technology (ICET)*, Ieee. 2017, pp. 1–6.

[Armbrust 11] C. Armbrust, T. Braun, T. Föhst, M. Proetzsch, A. Renner, B.-H. Schäfer, K. Berns, "RAVON: The robust autonomous vehicle for off-road navigation", in *Using Robots in Hazardous Environments*, Elsevier, 2011, pp. 353–396.

[Ayupova 19] S. Ayupova, D. Bents, E. Kozlova, "Trust as a factor of interaction efficiency", in *Smart Technologies and Innovations in Design for Control of Technological Processes and Objects: Economy and Production: Proceeding of the International Science and Technology Conference" FarEaston-2018" Volume 2*, Springer. 2019, pp. 58–64.

[Botache 19] D. Botache, L. Dandan, M. Bieshaar, B. Sick, "Early pedestrian movement detection using smart devices based on human activity recognition", 2019.

[Chaloupka-Risser 07] C. Chaloupka-Risser, G. Haindl, "Safety and contentment of pedestrians in connection with various mobility conditions", in *ICTCT EXTRA-WORKSHOP. Road user behaviour with a special focus on vulnerable road users: technical, social and psychological aspects*. 2007, pp. 32–41.

[Creswell 18] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, A. A. Bharath, "Generative adversarial networks: An overview", *IEEE signal processing magazine*, vol. 35, no. 1, pp. 53–65, 2018.

[Cugurullo 21] F. Cugurullo, R. A. Acheampong, M. Gueriau, I. Dusparic, "The transition to autonomous cars, the redesign of cities and the future of urban sustainability", *Urban Geography*, vol. 42, no. 6, pp. 833–859, 2021.

[Dautenhahn 07] K. Dautenhahn, "Socially intelligent robots: dimensions of human–robot interaction", *Philosophical transactions of the royal society B: Biological sciences*, vol. 362, no. 1480, pp. 679–704, 2007.

[Défossez 20] A. Défossez, L. Bottou, F. Bach, N. Usunier, "A simple convergence proof of adam and adagrad", *arXiv preprint arXiv:2003.02395*, 2020.

[Detjen 20] H. Detjen, B. Pfleging, S. Schneegass, "A wizard of oz field study to understand non-driving-related activities, trust, and acceptance of automated vehicles", in *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 2020, pp. 19–29.

[Duan 22] H. Duan, Y. Zhao, K. Chen, D. Lin, B. Dai, "Revisiting skeleton-based action recognition", in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2022, pp. 2969–2978.

[Endsley 95] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems", *Human factors*, vol. 37, no. 1, pp. 32–64, 1995.

[Eweiwi 15] A. Eweiwi, M. S. Cheema, C. Bauckhage, J. Gall, "Efficient pose-based action recognition", in *Computer Vision–ACCV 2014: 12th Asian Conference on Computer Vision, Singapore, Singapore, November 1-5, 2014, Revised Selected Papers, Part V 12*, Springer. 2015, pp. 428–443.

[Fang 19] Z. Fang, A. M. López, "Intention recognition of pedestrians and cyclists by 2d pose estimation", *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 11, pp. 4773–4783, 2019.

[Fang 22] H.-S. Fang, J. Li, H. Tang, C. Xu, H. Zhu, Y. Xiu, Y.-L. Li, C. Lu, "Alphapose: Whole-body regional multi-person pose estimation and tracking in real-time", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[Färber 15] B. Färber, "Kommunikationsprobleme zwischen autonomen Fahrzeugen und menschlichen Fahrern", *Autonomes Fahren: Technische, rechtliche und gesellschaftliche Aspekte*, pp. 127–146, 2015.

[Ghazizadeh 12] M. Ghazizadeh, J. D. Lee, L. N. Boyle, "Extending the Technology Acceptance Model to assess automation", *Cognition, Technology & Work*, vol. 14, pp. 39–49, 2012.

[Ghosh 22] S. Ghosh, G. Pons Rodriguez, S. Rao, A. El Ali, P. Cesar, "Exploring Emotion Responses toward Pedestrian Crossing Actions for Designing In-vehicle Empathic Interfaces", in *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 2022, pp. 1–6.

[Girshick 14] R. Girshick, J. Donahue, T. Darrell, J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014, pp. 580–587.

[Girshick 15] R. Girshick, "Fast r-cnn", in *Proceedings of the IEEE international conference on computer vision*. 2015, pp. 1440–1448.

[Guéguen 15] N. Guéguen, S. Meineri, C. Eyssartier, "A pedestrian's stare and drivers' stopping behavior: A field experiment at the pedestrian crossing", *Safety science*, vol. 75, pp. 87–89, 2015.

[Güler 18] R. A. Güler, N. Neverova, I. Kokkinos, "Densepose: Dense human pose estimation in the wild", in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 7297–7306.

[Gulland 04] W. Gulland, "Methods of determining safety integrity level (SIL) requirements-Pros and Cons", in *Practical Elements of Safety: Proceedings of the Twelfth Safety-critical Systems Symposium, Birmingham, UK, 17–19 February 2004*, Springer. 2004, pp. 105–122.

[Habibovic 18] A. Habibovic, V. M. Lundgren, J. Andersson, M. Klingegård, T. Lagström, A. Sirkka, J. Fagerlönn, C. Edgren, R. Fredriksson, S. Krupenia et al, "Communicating intent of automated vehicles to pedestrians", *Frontiers in psychology*, vol. 9, p. 284756, 2018.

[Husemann 22] J. Husemann, S. Kunz, K. Berns, "On demand ride sharing: scheduling of an autonoumous bus fleet for last mile travel", in *International Conference on Intelligent Autonomous Systems*, Springer. 2022, pp. 765–777.

[Insafutdinov 16] E. Insafutdinov, L. Pishchulin, B. Andres, M. Andriluka, B. Schiele, "Deepercut: A deeper, stronger, and faster multi-person pose estimation model", in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part VI 14*, Springer. 2016, pp. 34–50.

[Jan 20] Q. H. Jan, S. Klein, K. Berns, "Safe and efficient navigation of an autonomous shuttle in a pedestrian zone", in *Advances in Service and Industrial Robotics: Proceedings of the 28th International Conference on Robotics in Alpe-Adria-Danube Region (RAAD 2019) 28*, Springer. 2020, pp. 267–274.

[Jan 21a] Q. H. Jan, K. Berns, "Safety-configuration of Autonomous Bus in Pedestrian Zone.", in *VEHITS*. 2021, pp. 698–705.

[Jan 21b] Q. H. Jan, J. M. A. Kleen, K. Berns, "Simulated Pedestrian Modelling for Reliable Testing of Autonomous Vehicle in Pedestrian Zones", in *Smart Cities, Green Technologies, and Intelligent Transport Systems: 9th International Conference, SMARTGREENS 2020, and 6th International Conference, VEHITS 2020, Prague, Czech Republic, May 2-4, 2020, Revised Selected Papers 9*, Springer. 2021, pp. 290–307.

[Jan 22a] Q. H. Jan, Y. S. Baddela, K. Berns, "Pedestrian Activity Recognition from 3D Skeleton Data using Long Short Term Memory Units.", in *VEHITS*. 2022, pp. 368–375.

[Jan 22b] Q. H. Jan, P. Wolf, K. Berns, J. Reich, M. Wellstein, "Integration of human skeleton posture models into REACTiON for realizing dynamic risk management", in *International Commercial Vehicle Technology Symposium*, Springer. 2022, pp. 109–120.

[Jan 23a] Q. H. Jan, P. Dahiwal, K. Berns, "Effective Visual Content of eHMI for Autonomous Vehicles in Pedestrian Zones", in *International Conference on Robotics in Alpe-Adria Danube Region*, Springer. 2023, pp. 203–210.

[Jan 23b] Q. H. Jan, A. A. Khan, K. Berns, "Thorough Analysis and Reasoning of Environmental Factors on End-to-End Driving in Pedestrian Zones", in *Proceedings of the 20th International Conference on Informatics in Control, Automation and Robotics, ICINCO 2023, Rome, Italy, November 13-15, 2023, Volume 1*, G. Gini, H. Nijmeijer, D. P. Filev, Eds. SCITEPRESS, 2023, pp. 495–502.

[Jan 23c] Q. H. Jan, J. M. A. Kleen, K. Berns, "Deep Driving with Additional Guided Inputs for Crossings in Pedestrian Zones", in *Proceedings of the 9th International Conference on Vehicle Technology and Intelligent Transport Systems, VEHITS 2023, Prague, Czech Republic, April 26-28, 2023*, A. V. Vinel, J. Ploeg, K. Berns, O. Gusikhin, Eds. SCITEPRESS, 2023, pp. 402–409.

[Jan 24] Q. H. Jan, Y. S. Badella, K. Berns, “Detailed Analysis of Pedestrian Activity Recognition in Pedestrian Zones Using 3D Skeleton Joints Using LSTM”, *SN Computer Science*, vol. 5, no. 2, pp. 1–15, 2024.

[Kishore 19] P. S. R. Kishore, S. Das, P. S. Mukherjee, U. Bhattacharya, “ClueNet: A Deep Framework for Occluded Pedestrian Pose Estimation.”, in *BMVC*. 2019, p. 245.

[Lee 04] J. D. Lee, K. A. See, “Trust in automation: Designing for appropriate reliance”, *Human factors*, vol. 46, no. 1, pp. 50–80, 2004.

[Li 18] J. Li, H.-C. Wong, S.-L. Lo, Y. Xin, “Multiple object detection by a deformable part-based model and an R-CNN”, *IEEE Signal Processing Letters*, vol. 25, no. 2, pp. 288–292, 2018.

[Li 20] F. Li, S. Fan, P. Chen, X. Li, “Pedestrian motion state estimation from 2D pose”, in *2020 IEEE Intelligent Vehicles Symposium (IV)*, IEEE. 2020, pp. 1682–1687.

[Limprasert 13] W. Limprasert, A. Wallace, G. Michaelson, “Real-time people tracking in a camera network”, *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 3, no. 2, pp. 263–271, 2013.

[Liu 16a] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C. Berg, “Ssd: Single shot multibox detector”, in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, Springer. 2016, pp. 21–37.

[Liu 16b] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C. Berg, “Ssd: Single shot multibox detector”, in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, Springer. 2016, pp. 21–37.

[Matthews 17] M. Matthews, G. Chowdhary, E. Kieson, “Intent communication between autonomous vehicles and pedestrians”, *arXiv preprint arXiv:1708.07123*, 2017.

[Mirnig 17] N. Mirnig, N. Perterer, G. Stollnberger, M. Tscheligi, “Three strategies for autonomous car-to-pedestrian communication: A survival guide”, in *Proceedings of the companion of the 2017 ACM/IEEE international conference on human-robot interaction*. 2017, pp. 209–210.

[Nordfjærn 11] T. Nordfjærn, S. Jørgensen, T. Rundmo, “A cross-cultural comparison of road traffic risk perceptions, attitudes towards traffic safety and driver behaviour”, *Journal of Risk Research*, vol. 14, no. 6, pp. 657–684, 2011.

[Onkhar 21] V. Onkhar, P. Bazilinskyy, J. C. Stapel, D. Dodou, D. Gavrila, J. C. de Winter, “Towards the detection of driver–pedestrian eye contact”, *Pervasive and Mobile Computing*, vol. 76, p. 101455, 2021.

[Onkhar 22] V. Onkhar, P. Bazilinskyy, D. Dodou, J. De Winter, “The effect of drivers’ eye contact on pedestrians’ perceived safety”, *Transportation research part F: traffic psychology and behaviour*, vol. 84, pp. 194–210, 2022.

[Osokin 18] D. Osokin, “Real-time 2d multi-person pose estimation on cpu: Lightweight openpose”, *arXiv preprint arXiv:1811.12004*, 2018.

[Paddeu 20] D. Paddeu, G. Parkhurst, I. Shergold, “Passenger comfort and trust on first-time use of a shared autonomous shuttle vehicle”, *Transportation Research Part C: Emerging Technologies*, vol. 115, p. 102604, 2020.

[Papadimitriou 13] E. Papadimitriou, A. Theofilatos, G. Yannis, “Patterns of pedestrian attitudes, perceptions and behaviour in Europe”, *Safety science*, vol. 53, pp. 114–122, 2013.

[Pishchulin 16] L. Pishchulin, E. Insafutdinov, S. Tang, B. Andres, M. Andriluka, P. V. Gehler, B. Schiele, “Deepcut: Joint subset partition and labeling for multi person pose estimation”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 4929–4937.

[Proetzsch 10] M. Proetzsch, T. Luksch, K. Berns, “Development of complex robotic systems using the behavior-based control architecture iB2C”, *Robotics and Autonomous Systems*, vol. 58, no. 1, pp. 46–67, 2010.

[Qi 17] C. R. Qi, H. Su, K. Mo, L. J. Guibas, “Pointnet: Deep learning on point sets for 3d classification and segmentation”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 652–660.

[Rahman 21] M. T. Rahman, K. Dey, S. Das, M. Sherfinski, “Sharing the road with autonomous vehicles: A qualitative analysis of the perceptions of pedestrians and bicyclists”, *Transportation research part F: traffic psychology and behaviour*, vol. 78, pp. 433–445, 2021.

[Rasouli 17] A. Rasouli, I. Kotseruba, J. K. Tsotsos, “Agreeing to cross: How drivers and pedestrians communicate”, in *2017 IEEE Intelligent Vehicles Symposium (IV)*, IEEE. 2017, pp. 264–269.

[Redmon 16] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, “You only look once: Unified, real-time object detection”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 779–788.

[Reichardt 12] M. Reichardt, T. Föhst, K. Berns, “Introducing finroc: a convenient real-time framework for robotics based on a systematic design approach. Robotics Research Lab, Department of Computer Science, University of Kaiserslautern, Kaiserslautern”, Germany, Technical Report, Tech. Rep., 2012.

[Rempel 85] J. K. Rempel, J. G. Holmes, M. P. Zanna, “Trust in close relationships.”, *Journal of personality and social psychology*, vol. 49, no. 1, p.$\tilde{9}$5, 1985.

[Ren 16] Z. Ren, X. Jiang, W. Wang, “Analysis of the influence of pedestrians’ eye contact on drivers’ comfort boundary during the crossing conflict”, *Procedia engineering*, vol. 137, pp. 399–406, 2016.

[Rheu 21] M. Rheu, J. Y. Shin, W. Peng, J. Huh-Yoo, “Systematic review: Trust-building factors and implications for conversational agent design”, *International Journal of Human–Computer Interaction*, vol. 37, no. 1, pp. 81–96, 2021.

[Ribeiro 22] M. A. Ribeiro, D. Gursoy, O. H. Chi, “Customer acceptance of autonomous vehicles in travel and tourism”, *Journal of Travel Research*, vol. 61, no. 3, pp. 620–636, 2022.

[Ropertz 17] T. Ropertz, P. Wolf, K. Berns, “Quality-Based Behavior-Based Control for Autonomous Robots in Rough Environments.”, in *ICINCO (1)*. 2017, pp. 513–524.

[Sanchez-Caballero 22] A. Sanchez-Caballero, S. de López-Diz, D. Fuentes-Jimenez, C. Losada-Gutiérrez, M. Marrón-Romera, D. Casillas-Perez, M. I. Sarker, “3dfcnn: Real-time action recognition using 3d deep neural networks with raw depth information”, *Multimedia Tools and Applications*, vol. 81, no. 17, pp. 24119–24143, 2022.

[Schafer 08] H. Schafer, M. Proetzsch, K. Berns, “Action/perception-oriented robot software design: An application in off-road terrain”, in *2008 10th International Conference on Control, Automation, Robotics and Vision*, IEEE. 2008, pp. 223–228.

[Schmidt-Wolf 22] M. Schmidt-Wolf, D. Feil-Seifer, “Vehicle-to-pedestrian communication feedback module: A study on increasing legibility, public acceptance and trust”, in *International Conference on Social Robotics*, Springer. 2022, pp. 14–23.

[Schmidt 09] S. Schmidt, B. Faerber, “Pedestrians at the kerb–Recognising the action intentions of humans”, *Transportation research part F: traffic psychology and behaviour*, vol. 12, no. 4, pp. 300–310, 2009.

[Schneemann 16] F. Schneemann, I. Gohl, “Analyzing driver-pedestrian interaction at crosswalks: A contribution to autonomous driving in urban environments”, in *2016 IEEE intelligent vehicles symposium (IV)*, IEEE. 2016, pp. 38–43.

[Scholtz 03] J. Scholtz, “Theory and evaluation of human robot interactions”, in *36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the*, IEEE. 2003, pp. 10–pp.

[Shahroudy 16] A. Shahroudy, J. Liu, T.-T. Ng, G. Wang, “Ntu rgb+ d: A large scale dataset for 3d human activity analysis”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 1010–1019.

[Soni 16] N. Soni, N. Soni, “Benefits of pedestrianization and warrants to pedestrianize an area”, *Land use policy*, vol. 57, pp. 139–150, 2016.

[Sucha 17] M. Sucha, D. Dostal, R. Risser, “Pedestrian-driver communication and decision strategies at marked crossings”, *Accident Analysis & Prevention*, vol. 102, pp. 41–50, 2017.

[Sun 19] K. Sun, B. Xiao, D. Liu, J. Wang, “Deep high-resolution representation learning for human pose estimation”, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019, pp. 5693–5703.

[Tzeng 14] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, T. Darrell, "Deep domain confusion: Maximizing for domain invariance", *arXiv preprint arXiv:1412.3474*, 2014.

[Vogel 04] D. Vogel, R. Balakrishnan, "Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users", in *Proceedings of the 17th annual ACM symposium on User interface software and technology.* 2004, pp. 137–146.

[Wolf 18] P. Wolf, T. Ropertz, K. Berns, M. Thul, P. Wetzel, A. Vogt, "Behavior-based control for safe and robust navigation of an unimog in off-road environments", in *Commercial Vehicle Technology 2018: Proceedings of the 5th Commercial Vehicle Technology Symposium-CVT 2018*, Springer. 2018, pp. 63–76.

[Wolf 22] P. Wolf, *Cognitive Processing in Behavior-Based Perception of Autonomous Off-Road Vehicles*, Dr. Hut, 2022.

[Woo 18] S. Woo, J. Park, J.-Y. Lee, I. S. Kweon, "Cbam: Convolutional block attention module", in *Proceedings of the European conference on computer vision (ECCV).* 2018, pp. 3–19.

[Yanco 02] H. A. Yanco, J. L. Drury, "A taxonomy for human-robot interaction", in *Proceedings of the AAAI fall symposium on human-robot interaction.* 2002, pp. 111–119.

[Yassin 19] H. H. Yassin, "Livable city: An approach to pedestrianization through tactical urbanism", *Alexandria Engineering Journal*, vol. 58, no. 1, pp. 251–259, 2019.

[Zhu 23] Y. Zhu, N. Samet, D. Picard, "H3WB: Human3. 6M 3D WholeBody Dataset and Benchmark", in *Proceedings of the IEEE/CVF International Conference on Computer Vision.* 2023, pp. 20166–20177.

[Zolynski 18] G. Zolynski, *Cognitive maps for autonomous machines in construction and civil engineering*, Verlag Dr. Hut, 2018.

C. Curriculum Vitae

Name: Qazi Hamza Jan
Address: Gottlieb-Daimler-Str. Gebäude 48, Raum 358, D-67663, Kaiserslautern
Email: `hamza.jan@rptu.de`

Qazi Hamza Jan is an academic and professional currently pursuing a PhD in Computer Science at the Rheinland-Pfälzische Technische Universität Kaiserslautern, with a focus since 2019. His educational background also includes a Master of Science in Electrical Engineering from Lahore University of Management Sciences (LUMS), specializing in Control Systems, Computer Vision, and Embedded Systems, achieved between 2013 and 2015. Additionally, he holds a Bachelor of Science in Electrical Engineering from Bahria University, Islamabad Campus, obtained in 2013.

Professionally, Hamza has been extensively involved in robotics research. Since September 2022, he has been a Full-time Researcher at the Robotic Research Lab, RPTU Kaiserslautern-Landau, working on innovative navigation strategies for driverless minibuses, including aspects of perception and control. Prior to this role, from August 2016 to January 2018, Hamza was a lecturer at CECOS University, Peshawar. His academic and professional journey highlights a robust commitment in advancing field of electrical engineering and robotics.
