HAVIT: A VR-Based Platform to Support Human-Autonomous Vehicle Interaction Study

Xiaolei Guo[1], Dayu Wan[1], Dongfang Liu[2], Christos Mousas[1], and Yingjie Chen[1](✉)

[1] Department of Computer Graphics Technology, Purdue University, West Lafayette, IN, USA
{guo579,wand,cmousas,victorchen}@purdue.edu
[2] Department of Computer Engineering, Rochester Institute of Technology, Rochester, NY, USA
dongfang.liu@rit.edu

Abstract. We propose the Human-Autonomous Vehicle Interaction Testbed (HAVIT), a VR-based platform that enables researchers and designers to quickly configure AV-pedestrian interaction scenarios and evaluate their design concepts during the design process in a holistic and consistent manner. The HAVIT presents an efficient workflow that combines the Scenario Configuration, Experimental Setting, and Batch Configuration. Our workflow enables researchers to quickly and flexibly configure motion behaviors of AVs and external human-machine interfaces (eHMIs) through visual panels and direct manipulation; complete experimental setting through Data Collection component and Testing Instruction component; and immediately enact and immersively experience them to reasonable iterate and generate virtual scenarios for testing. We conducted an usability testing with domain experts and designers to test the effectiveness of how HAVIT can support AVs-pedestrian interaction design process.

Keywords: Human-autonomous vehicles interaction · Virtual reality testbed · eHMI

1 Introduction

Autonomous vehicle (AV)-pedestrian interactions directly impact pedestrian safety, etiquette, and overall acceptance of AV technology [1]. It is vital to fully explore this emerging interaction type to address potential ambiguities and conflicts in the future of transportation. However, designing and evaluating communication techniques for AV-pedestrian interactions is a challenging task due to the unavailability of AVs for experiments and the potential harms involved in physical field tests. As such, within these circumstances, virtual reality (VR)-based methods have received considerable attention from the research

H. Krömker (Ed.): HCII 2022, LNCS 13335, pp. 371–390, 2022.
https://doi.org/10.1007/978-3-031-04987-3_25

community and are increasingly being used to investigate pedestrians' behavior in relation to AVs and to understand different interaction solutions [2]. Compared with other methods (e.g., the Wizard-of-Oz [3,4] and video-based [5–7] methods), VR-based approach offers a number of benefits for examination of the simulated scenarios and to conduct AV-pedestrians interaction evaluation. An immersive environment enables improved spatial perception ability that can facilitate realistic judgment of the speed and distance of AVs [8]. Furthermore, greater flexibility in parameter manipulation and environmental control can be easily achieved [9].

While VR-based approaches are becoming increasingly powerful and popular, they can involve difficulties in achieving consistency and reproducibility in AV-pedestrian interaction design and testing [10]. The root cause of these issues is that the design concepts in existing studies are usually evaluated in different virtual environments or applications, leading to varying levels of fidelity in terms of traffic scenarios, communication interface prototypes, system settings, etc. [11]. Therefore, it is hard for designers and researchers to evaluate their design and compare the results across different scenarios and reach a consensus about the knowledge gained.

At the same time, to create virtual scenarios for evaluation, researchers and designers are limited to expert tools for the design of VR environments, which requires considerable effort. Even with expertise on such tools, it is time-consuming and labor-intensive to perform the process of design iteration, and evaluation of the testing scenarios [10], and this issue will become more serious when multiple testing scenarios are required. As a result, new tools and methods need to be developed to overcome the methodological and process issues raised above, which impeding knowledge development in the research community.

To this end, we propose the Human-Autonomous Vehicle Interaction Testbed (HAVIT), a VR-based platform, as a possible solution for enhancing the consistency and efficiency of AV-pedestrian interaction design and study process. To implement our testbed, (1) we were inspired by previous studies to develop the components of the key parameters of the HAVIT in terms of Physical Context, Vehicle Behavior, and External Human-Machine Interface (eHMI) Behavior, which can be adapted to address the factors critical to future AV-pedestrian interaction and interface design; (2) In addition, corresponding structured visual panels in the HAVIT allow users to easily manipulate those parameters for intuitiveness interaction scenarios creation; (3) Finally, the HAVIT provides a coherent and iterative workflow to facilitate the efficiency of formative study towards the eHMI concepts, starting with the Scenario Configuration, moving to Experimental Setting, and ending with Batch Exportation, which enables a rapid examination and generation of virtual testing scenarios.

2 Related Work

We classify the related work into two categories: (1) AV-pedestrian interaction studies and (2) VR simulation for AV-pedestrian interaction research.

2.1 Interaction Between AVs and Pedestrians

Many factors have been explored and proven to influence the decision-making processes of pedestrians. Rasouli and Tsotsos [12] provided a comprehensive summary of the factors influencing pedestrian behavior through a review of the related literature. These factors can be divided into two main categories: environmental factors and pedestrian factors. Environmental factors include traffic characteristics (e.g., vehicle appearance and traffic flow), dynamic factors (e.g., vehicle speed and spacing), and the physical environment (e.g., road structure, traffic signs, and weather), while pedestrian factors include demographics, status (pedestrians' physical status includes attention, walking pattern, speed, and trajectory), ability, characteristics (features that define how pedestrians' think and behave, including culture, past experience, and faith), and social factors. It is worth noting that the above influences are often interrelated in real-life traffic scenarios, and they combine to influence road users' perceptions and understandings of the state and intent of AVs [11]. Studying the interaction between these influences is essential to understand traffic situations' complexity and to facilitate safer AV-pedestrian interactions.

eHMI is an important aspect of AV-pedestrian interactions, which is the form of communication external to an AV that is typically used to communicate the AV's current state and future behavior to pedestrians; it can help in overcoming AV trust issues and improving the effectiveness and experience of AV-pedestrian communication [13]. Various eHMI concepts have been proposed and tested, such as text [8,14], symbols [5,15], street projections [16], light animations [17,18], and information from mobile devices [19]. However, researchers have not yet reached a consensus about how different eHMIs should be used [9]. More and more research is now focusing on details related to the implementation of eHMIs to achieve the best interactions in terms of usability, security, and efficiency. For example, many eHMI studies have started to explore in-depth the dimensions of communication perspectives [20,21], communication subjects [22], and covered states [7]. Other studies have analyzed the design of the interactive elements of a particular type of eHMI, such as color [23,24], placement location [17,21], and display mode [6].

The scalability of eHMIs is another aspect that needs to be explored in the long term [10,25]. Most eHMIs have been tested in relatively simple and unrealistic situations, which has led to many eHMI concepts becoming viable options. The problem, however, is that the results of these studies often only show that the eHMIs improve simple interactions; most studies do not provide insights into using eHMIs in more complex traffic scenarios [1]. Therefore, more evaluations of interactions between pedestrians and eHMIs in diverse traffic scenarios–such as those involving multiple pedestrians [27] or different weather conditions–are needed in the future [11].

Most of the previous research has focused on common traffic scenarios and strategies for communicating the status or intent of AVs to the normal road user. However, research on AV-pedestrian interactions are equally critical in special cases and situations [28], such as sensor failure, a lack of system action, or action errors. As such, future research should include evaluations of (1) how pedestrians should be informed and instructed to act depending on the type of malfunction occurring, (2) how to optimize safe interactions between eHMIs and pedestrians in special scenarios, and (3) how to conduct interactions in a way that ensures the public acceptance and trust of AVs [6]. In addition, while people with disabilities are among the most vulnerable road users in traffic, only a few studies have been done on forms of external communication for people with disabilities (e.g., physical, visual, or hearing impairment) [29–31]. When conducting studies on these specific conditions, a method with high degree of flexibility in manipulating the vehicles behaviors is required.

Many dimensions of AV-pedestrian interaction have not yet been adequately studied, and research on each dimension is indispensable. More importantly, the complexity of the study of AV-pedestrian interaction will increase with the number of studies being conducted and the aspects being studied, making many traditional research methods infeasible. In the face of such challenges, a VR-based method can provide more flexible, scalable approaches that can support more aspects of AV-pedestrian interaction research. This was one of the critical motivations behind the development of HAVIT.

2.2 VR Simulation for AV-Pedestrian Interaction Design and Study

VR simulation has been widely used to study AV-pedestrian interactions. Compared to traditional, non-immersive virtual environments (e.g., paper-based [26], video-based [6,7,32], or real-world-based [3] environments), the VR-based method combines the advantages of the above approaches to allow for flexible environmental control and simulations under highly realistic conditions. For example, Chang et al. [33] designed and evaluated an eHMI concept using the "eyes on the front of the car" eHMI; specifically, they developed two VR scenarios, each including five components: the environment, the user (i.e., a three-dimensional computer-generated pedestrian model), the car, the eHMI, and the car movement route (i.e., a straight line). Doric et al. [34] implemented a VR-based simulation that includes a simple, uncontrolled pedestrian crossing scenario in which virtual vehicles are continuously generated at regular intervals. de Clercq et al. [14] used VR simulation to evaluate the impact of four interface concepts on pedestrian crossing intentions. Specifically, VR was used to simulate the vehicle behavior (e.g., giving way or not giving way), vehicle size, eHMI (four types), and display time of the eHMI (i.e., early, middle, and late). Studies have also begun to evaluate sound interfaces using VR simulations [30]. To evaluate auditory concepts for people with visual impairments, VR-simulated scenarios have included background noise (e.g., a mixture of human voices and engine sounds) and have given participants the ability to control the direction and location of the sounds so that testers can immerse themselves in a realistic sound experience. The closest related work to ours describes the On-Foot [1], a VR-based simulator that aims to simulate mixed traffic scenarios. It provides users with a set of control modules that can be modified by coding. In this case, VR showed a higher degree of control over the simulation of diverse AV-pedestrian interaction conditions.

Furthermore, one of the essential reasons VR-based simulation have produced convincing evidence is that they primarily utilize objective measures [9], such as reaction time, duration, and accuracy. In addition, VR can capture information about the test taker's body movement. For example, Schmidt et al. [35] used an immersive VR environment to explore the intricate social cues that underlie the non-verbal communication involved in pedestrians' crossing decisions. They collected motion trajectories generated by moving the body, legs, arms, and head of each subject in the physical and virtual world.

3 Design Principles

Our primary aim with the HAVIT was to create a VR-based research tool that could improve the consistency and efficiency of AV-pedestrian interaction design studies. To achieve this, we identified three key design principles (DPs) from the literature review and the interview with domain designers and experts.

DP1: Customizable and Flexible. Most VR simulations are unable to support setting and modifying the factors that influence pedestrians' decision to cross the street [10] and are limited in their flexibility to fulfill requirements of diverse AV-pedestrian interaction scenarios creation. The HAVIT builds upon the previous works, sythsized a set of key parameter components to address the factors critical to future AV-pedestrian interaction and interface design.

DP2: Easy to Use. The implementation of virtual AV-pedestrian interaction scenarios often requires researchers to set various factors by coding [15,25]. Our goal was to achieve an intuitive configuration process of virtual testing scenarios that easy-to-use for designers.

DP3: Efficient and Iterative. HAVIT aimed to present an iterative workflow that enables researchers and designers to modify and generate multiple virtual scenarios simultaneously. Besides, HAVIT supports the set up of experiments in the virtual testing scenarios and the collect of data within the developed scenarios; thus, achieving the quick exploration and iteration of design concepts.

4 The HAVIT System

This section presents the HAVIT, providing descriptions of the user interface (UI), the main process it enables, and the relevant components. As Fig. 1 shows, the HAVIT supports three processes at the highest level: Scenario Configuration, Experimental Setting, and Batch Exportation. Each process can be configured via user panels and scripts provided by the HAVIT.

Scenario Configuration. We organized the key parameters into three user panels—Physical Context, Vehicle Behavior, and eHMI Behavior—to guide users to create the scenario. The HAVIT also allows users to quickly add and remove objects from the scenario by interacting directly with them. In addition, users can preview the current scenario at any time during the configuration process.

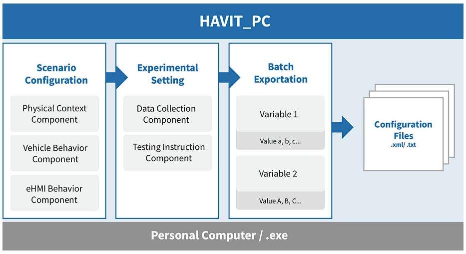

Fig. 1. The main processes of the HAVIT.

Experimental Setting. The HAVIT enables the rapid setup of experiments by providing a Data Collection component and a Testing Instruction component. The Data Collection component allows researchers to collect assigned quantitative data (e.g., the start/end time of crossing behavior, time required for decision-making, distance traveled, and average speed) and qualitative data (i.e., the participants' subjective experiences). The Testing Instruction component provides a set of adjustable panels that will be displayed in the VR environment to enable the guidance of participants during the testing.

Batch Exportation. Studies often involve manipulating a group of variables and, thus, generate a set or several sets of trials for an experiment. Repetitive manual configuration reduces the development efficiency and increases the risk of human error when many test scenarios are required. The HAVIT allows users to add variables and values according to the experimental requirements and generate multiple scenarios simultaneously to enhance testing scenarios' development efficiency.

The HAVIT is a Unity-based desktop program that can easily be used on a personal computer at the system level. With the HAVIT, users can configure interaction scenarios and export configuration files according to their needs. These configuration files can then be read easily by a VR device, such as an Oculus, and loaded with the appropriate environment and parameters to generate scenarios for testing. Below, we detail the parameter components, UI, and design of each system component.

4.1 Parameter Components

The design of the key parameters of HAVIT was inspired by previous literature, mainly refers to the important factor aspects of AV-pedestrian interaction that need to be explored in-depth in the future (see Sect. 2.1 of the literature review). Figure 2 shows the key parameter components of the HAVIT, which are (1) Physical context component, (2) Vehicle behavior component, and (3) eHMI behavior component. Each parameter component is consisted of several key classes. Specifically, a `ScenarioController` manages the road structure, and the natural conditions show the corresponding scenario to the user and contain several `RouteControllers`. A `RouteController` controls a specific route in the current scenario, as well as the set of vehicles driving on this route. In addition to a vehicle's appearance and behavior information, a `Vehicle` can display several `eHMIs` when interacting with pedestrians. Also, the `eHMI` class has two child classes: `Visual eHMI` and `Sound eHMI`. Each child class has different properties to control the display behavior of the eHMI. The `PedestrianTaskController` and `DataCollectionController` manage the testing task and the data to collect in the VR experiment. The `ExportationController` contains several variables, each of which points to a specific parameter and has multiple values.

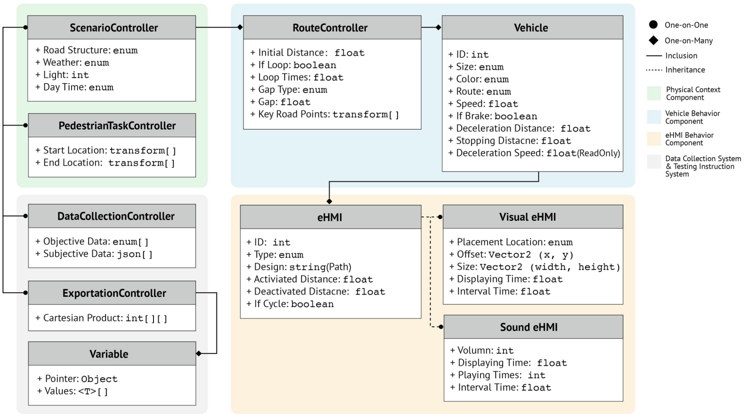

Fig. 2. Summary of the parameter components and key classes of the HAVIT.

4.2 User Interface

As Fig. 3 shows, the HAVIT consists of three UI components: the Main View, user panels, and Mini-Map. The Main View allows users to inspect the entire scenario two modes: (1) Edit Mode, in which users can directly interact with objects within the scenario, such as turning the vehicle to move in a different direction or moving the position, and (2) Preview Mode, in which users can check the effect of the scenario in the current configuration conditions. Users can also freely move the camera to change the viewing angle. Next, the key parameter components and functionalities are implemented and presented to users through the user panels, of which there are five: the Physical Context Panel (Fig. 3a [U1]), Vehicle Behavior Panel (Fig. 3a [U2 and U3]), and eHMI Behavior Panel (Fig. 3a [U4]), which are used to manipulate the parameters for creating scenarios; the Experimental Setting Panel (Fig. 3a [U5]), which is used to set up the collected data and the testing instructions in the VR environment; and the Batch Exportation Panel (Fig. 3b [U6]), which is used to generate multiple scenarios at the same time. Last, the Mini-Map (Fig. 3a [M]) is used to provide an overview of the current scenario. Also, location markers and vehicle routes are displayed on the Mini-Map to allow for quick checks.

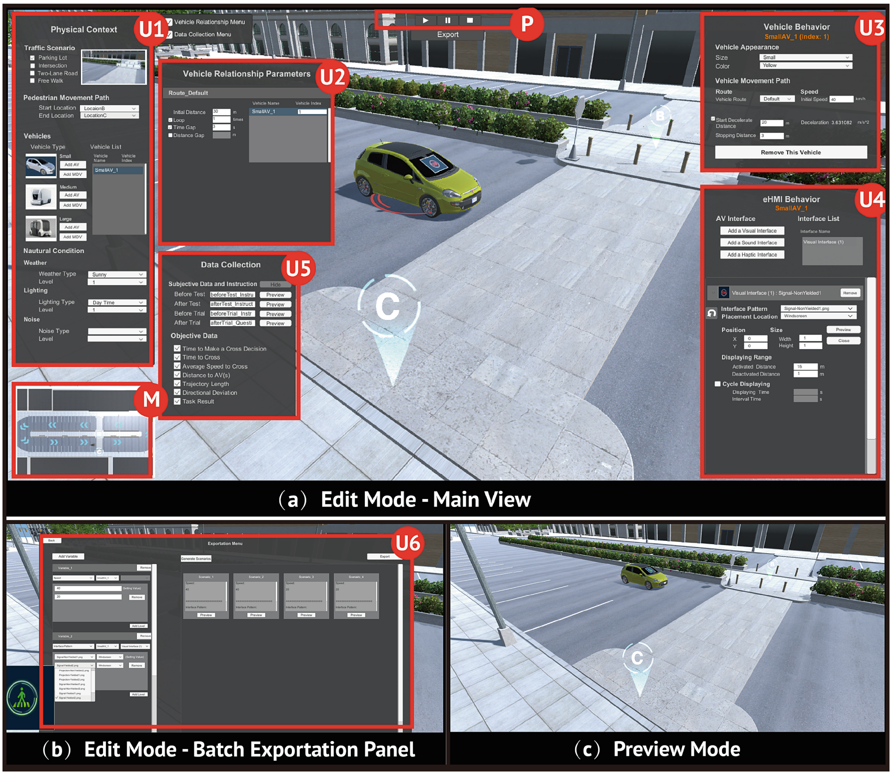

Fig. 3. The UI of the HAVIT. (a) Edit Mode - Main View, (b) Edit Mode - Batch Exportation Panel, and (c) Preview Mode. The panels and interactive components are marked with red borders in the image.

4.3 Scenario Configuration

The first main process enabled by the HAVIT is scenario configuration. Users can quickly and easily configure various AV-pedestrian interaction scenarios using a set of user panels.

Physical Context. The Physical Context parameter allows users to define the road structure, the participant's movement path, the test vehicles, and the natural conditions. These elements are discussed in greater detail in this section.

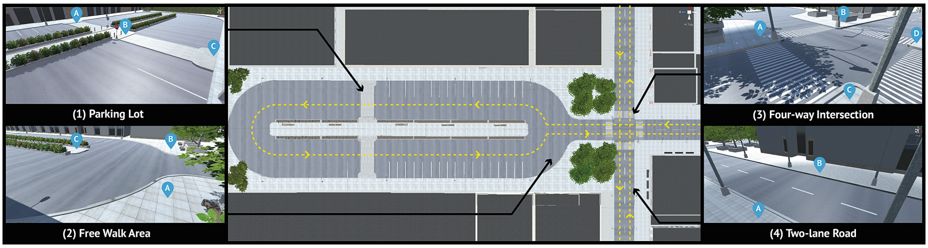

Fig. 4. Road structures in the HAVIT. The aerial-view map of the HAVIT is shown in the middle. The yellow dotted lines represent the routes on which a vehicle can move. The four road-structure scenarios are visualized on either side of the aerial-view map: (1) Parking Lot, (2) Free Walk Area, (3) Four-Way Intersection, and (4) Two-Lane Road. The blue markers with capital letters represent the optional locations that are used to specify the pedestrian movement path in each scenario.

Road Structure. The HAVIT contains a street map showing an area of 396 × 561 square feet (as shown in Fig. 4) that includes four road structure types: (1) Parking Lot, (2) Free Walk Area (i.e., the entrance/exit area of a parking lot), (3) Four-Way Intersection, and (4) Two-Lane Road. All of these are locations where AVs and pedestrians frequently interact in daily traffic. Each road structure has a predetermined travel route for vehicles. The initial settings do not include traffic signals, and users can add traffic signals (e.g., traffic lights, crosswalks, and stop signs) to the scenarios as needed.

Pedestrian (Participant) Movement Path. Once the road structure scenario has been decided on, the user needs to assign the movement path for the participants in the testing scenario. Users can determine the path by specifying the start location and end location, which are dynamically generated according to the road structure selected by the user.

Vehicles. The HAVIT provides three vehicle model sizes: Large (i.e., buses and trucks), Medium (i.e., vans), and Small (i.e., passenger cars). This is because prior works showed that vehicle size leads to differences in the subjective risk perception and objective distance perception of pedestrians [13].

Natural Conditions. The natural condition parameters supported by HAVIT include Weather Condition, Lighting Condition, and Noise. Currently proposed external communication interfaces mainly rely on visual cues and auditory cues. Lighting and weather factors are critical to examining the visibility and interactive performance of visual-based interface concepts. In addition, noise is an important consideration when designing auditory-based interfaces, which are effective solutions for visually impaired people [6].

Vehicle Behavior. One of the critical features of the HAVIT is the ability to provide the flexible control of vehicle behavior. There are two control modes offered for users to achieve this: controlling the behavior of vehicle groups and controlling the behavior of individual vehicles. These act on vehicle groups and on specific vehicles within the scenario, respectively.

Vehicle Relationship Behavior. The HAVIT allows users to control the vehicle behavior for multiple vehicles at the same time based on the vehicles' travel routes. The controllable parameters include the sequence of vehicles, the initial distance (i.e., the distance between pedestrians and the generation point), the number of travel loops, and the generation gap (i.e., the time gap or distance gap) between vehicles.

Individual Vehicle Behavior. The HAVIT also allows users to configure parameters for each vehicle in a scenario. This control mode is useful when the vehicles in the scenario all have different behaviors or when more complex changes in vehicle behavior need to be simulated. Specifically, the HAVIT supports the initial speed (in km/h) of the vehicle, the acceleration/deceleration distance (i.e., the distance at which the vehicle starts to decelerate, in meters) of the vehicle, and the stopping distance of the vehicle (i.e., the distance to pedestrians, in meters). When the user selects the above three parameters, the acceleration/deceleration speed is calculated and shown on the user panel.

The eHMI Behavior. The HAVIT supports the import of user-defined interfaces. Further, the HAVIT allows the simultaneous placement of multiple eHMIs on an AV to evaluate the effects of multiple eHMI combinations. To enable the exploration of the functional details of these interaction concepts, we also provided controllable parameters related to visual and auditory eHMIs and ensured that the parameters of each added eHMI could be adjusted independently. The HAVIT's parameters are as follows:

- **Placement Location (Only for Visual eHMIs):** This refers to the placement area of an eHMI on an AV, which can be the windshield, bumper, roof, side windows, front road, or front cover of the vehicle.
- **Display Position (Only for Visual eHMIs):** This refers to the specific position of an eHMI in the placement area, controlled by the offsets in the horizontal and vertical coordinates based on the center of this area.
- **Display Size (Only for Visual eHMIs):** This refers to the exact size of an eHMI, controlled by width and height.
- **Activate and Deactivate Distance (for Both):** This refers to the distance to pedestrians from where an eHMI starts to appear and disappears on the AV.
- **Cycle Display (for Both):** This refers to whether an eHMI is displayed periodically or not. It can be controlled by setting the displaying time and the interval time.
- **Play Volume (Only for Auditory eHMIs):** This refers to the sound volume of an auditory eHMI.

4.4 Experimental Setting

Data Collection Component. Considering data collection is an indispensable part of the experimental process, we implemented a Data Collection component in the HAVIT to allow users to collect the data generated by tests. We classified the types of data collected by the HAVIT, based on previous research, as subjective data and objective data.

Objective Data. One of the key reasons that VR-based approaches can produce convincing results is the use of objective measurements. The HAVIT provides a variety of pedestrian task-related metrics that can be automatically activated through the provided scripts. When the task is completed for each scenario, the tracking component automatically reports data information for the corresponding metric. The metrics currently covered by the HAVIT are the time to make a cross decision, time to cross, trajectory length, average speed to cross, distance to AV(s), directional deviation, and task result.

Subjective Data. To improve the validity of subjective data collection, we utilized InVRQs [37], an existing VR questionnaire toolkit, as a complement to the HAVIT. This toolkit was useful, as it provided the questionnaire structure and question types. The HAVIT allows the user to determine where the questionnaire panel appears in the VR scenario.

Testing Instruction Component. The Testing Instruction component is provided to display experimental instructions for participants in VR testing scenarios, which rely on a set of panels inside the VR that are shown to the participants. Four display timings are provided: before the test, after the test, and before and after each trial—all of which support the customization of the questionnaire or text presentation.

4.5 Batch Exportation

To reduce the repetitive manual configuration process, the HAVIT provides a Batch Exportation component. Specifically, after configuring one testing scenario, the user can specify one or more variables (i.e., parameters in the HAVIT) required for batch configuration and assign specific values accordingly through the user panel. After the user specifies all the variables and their values, the system will create a set for each of them and perform the Cartesian product operation on these sets. The HAVIT also provides a scenario list to allow users to preview the generated scenarios. When a user previews a specific scenario, the system will modify the values of each involved parameter according to the corresponding combination. Also, when exporting batch scenarios, the system will iterate through all combinations and export the corresponding configuration file for each combination.

4.6 Preview Mode

To allow users to check the effect of the configured scenario, the HAVIT supports Preview Mode. Specifically, the system refers to the Unity game engine's Play Mode and provides three buttons in game Game window: Play, Pause, and Stop. The user can click the Play button to preview, click again or click the Stop button to exit.

5 Implementation

The HAVIT was developed in the Unity game engine (v. 2020.1.9f1), with all related scripts written in `C#`. The project has been packaged to the Windows platform to work independently from the engine. After testing on an HP OMEN Gaming Laptop with a GTX 1650 graphics processing unit (GPU), the HAVIT was found to have a guaranteed a framerate 60 Hz (default setting) when 20 vehicles are running simultaneously.

6 Evaluation

This work presents the first version of HAVIT. Therefore, we primarily focused on its overall usability in this evaluation, i.e., whether the designed system and features of HAVIT are understandable and easy to use for our intended users. Specifically, we conducted a formative user study with professionals. We were interested in the participants' performance when using the testbed and the qualitative impressions of their experience.

6.1 Participants

We relied on the intended users of the HAVIT to gain insights from their workflows, and we expect that this initial feedback will help distill the strengths and areas for improvement of the HAVIT for the future. As such, we recruited professionals in fields related to human-computer interaction (HCI; $n = 8$; 3 females), including VR experience researchers (P2, P3, P4, and P5), intelligent systems researchers (P6 and P8), and user experience designers (P1 and P7). None of the users had prior experiences with our testbed.

6.2 Procedure

The participants were first introduced to the HAVIT, and they then were instructed to configure a set of testing scenarios for an AV-pedestrian study that featured a within-subject study design, 2 independent variables, and a total of 9 (3×3) testing scenarios (trials). We chose this study topic because its complexity allowed us to demonstrate and test many of the HAVIT's features. The participants then completed questionnaires evaluating the HAVIT's main features and answered interview questions from the researchers. The testing process lasted about 50 min. One researcher took observational field notes, which were analyzed and used to help interpret the results from our survey data.

*Introduction and Training (*15 min*).* Following the signing of informed consent forms and obtaining recording permission, the participants were provided with some background knowledge about the AV-pedestrian interaction study and all the features of the HAVIT system. They were then guided through configuring a simple scenario and allowed to explore freely.

*Tasks (*25 min*).* In this part, the participants configured a set of testing scenarios using the HAVIT, following specific instructions. They were provided with a Study Method document that included the study goal and the experimental design (i.e., independent variables, dependent variables, pedestrian tasks, and experimental setup). We took care to ensure that the content was as short and concise as possible. The instructions were as follows: (1) Manual Configuration (Task 1): Configure 1 of the 9 testing scenarios (a total of 15 parameters need to be set); and (2) Batch Configuration (Task 2): Generate the 9 required testing scenarios at the same time (a total of 6 parameters need to be set). We emphasized that there was no correct order for the configuration of the scenarios and that they could complete the task according to their understanding of it. The participants were asked to verbally report, "I'm done" after completing the first task. The researcher checked the configuration results and informed the participants to make adjustments if necessary, after which they continued with the second task. The participants were also told to complete the tasks as quickly and accurately as possible. The whole process was screen recorded.

*Questionnaire (*5 min*).* After the two configuration tasks were completed, each participant was then asked to answer Likert-type questions related to the system features. Each Likert-type item was graded by users from 1 to 5 in relation to the usefulness of the feature and their level of agreement with the item. Our questions were inspired by the "first-use study" in Exemplar [38].

*Semi-Structured Interview (*5 min*).* Finally, we conducted a semi-structured interview with each participant, which addressed the following: (1) the ease of configuration with the HAVIT, (2) its usefulness, (3) the scenario results achieved and the participant's satisfaction with those, and (4) the potential for the future use of the add-on. The interviews were all audio recorded.

6.3 Measurements

To test this first version of the HAVIT, we defined two basic metrics for analysis: (1) completion time, which refers to how much time the participants required to complete each task (the timing started when the participants verbally reported, "I'm ready" and ended when they stated, "I'm done"); and (2) task success result, which refers to whether the task was completed successfully (i.e., if all parameters were set up correctly) or was failed. A thematic analysis of the participants' opinions was conducted; these opinions were collected during the semi-structured interviews. The themes also stemmed from the observations of the participants' behavior during the tasks and the observer's debriefing after the VR testing scenario configuration session.

6.4 Results

Objective Data. Table 1 shows a summary of the participants' completion times and task success. In general, the participants were able to understand the features provided by the HAVIT. All participants completed both tasks. In Task 1 (T1, Manual Configuration), all participants except P1 and P7 finished in approximately 10 min (mean [M] = 9.02). P1 and P7 had less quantitative experimental experience and spent extra time on reading the study method document. Four participants (P1, P3, P7, and P8) were unsuccessful in completing T1, and the errors they made are shown in Table 1. In Task 2 (T2, Batch Configuration), 7 out of the 8 participants completed the task successfully, and P4 was unsuccessful because of one omission error. The average completion time of T2 was 4.22 min.

Table 1. Summary of the participants' task-completion times (T1 and T2) and success results.

Participants	Task 1 (Minutes)	Task 1 (Success & Accuracy Score)	Task 1 (Error Type & Number)	Task 2 (Minutes)	Task 2 (Success & Accuracy Score)
P1	12:05	Fail (14/15)	Input error: 1	4:41	Success (6/6)
P2	7:56	Success (15/15)	None	3:28	Success (6/6)
P3	8:32	Fail (13/15)	Input error: 1 Omission error: 1	4:24	Success (6/6)
P4	5:24	Success (15/15)	None	5:37	Fail (5/6)
P5	6:18	Success (15/15)	None	3:43	Success (6/6)
P6	6:35	Success (15/15)	None	3:35	Success (6/6)
P7	17:46	Fail (14/15)	Omission error: 1	5:17	Success (6/6)
P8	9:43	Fail (14/15)	Omission error: 1	4:58	Success (6/6)

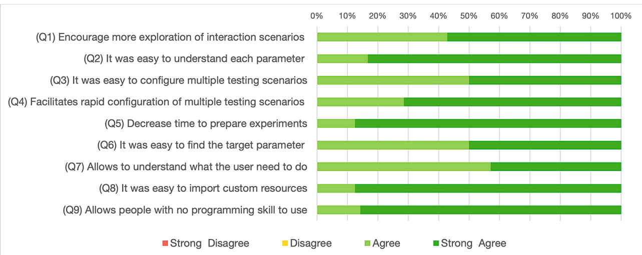

Fig. 5. System feature-related Likert-type question results ($n = 8$).

Subjective Data. Here, we report the results from the Likert-scale questions, in terms of mean (M), median (m), and standard deviation (SD). Most participants reported that the HAVIT enabled them to quickly understand a wide range of impact factors related to AV-pedestrian interaction and that it encouraged exploration (Q1: $M = 4.38$, $m = 4.50$, $SD = .74$). Likewise, 7 participants held positive views about the HAVIT's ability to reduce the time needed to prepare for an experiment (Q5: $M = 4.75$, $m = 5.00$, $SD = .71$). For the Batch Exportation process, most participants reported that the HAVIT could help decrease the time needed to configure multiple scenarios (Q4: $M = 4.50$, $m = 5.00$, $SD = .76$), and the Batch Configuration method is easy to understand (Q3: $M = 4.00$, $m = 4.00$, $SD = .93$). Further, the participants were generally confident about uploading their self-defined interface to the HAVIT (Q8: $M = 4.88$, $m = 5.00$, $SD = .35$) and ranked it highly in relation to the statement that the HAVIT "allows people with no programming skill to use it" (Q9: $M = 4.63$, $m = 5.00$, $SD = .74$). Several participants also agreed that the parameters are intuitive and easy to understand (Q2: $M = 4.38$, $m = 5.00$, $SD = .91$), but they would like some video illustrations and more detailed information. Several participants suggested that better user panels for managing related parameters could help improve the "time to configure the scenario" and "rapid modification" (Q6: $M = 3.75$, $m = 3.50$, $SD = .89$). Finally, the HAVIT's workflow received positive feedback from the participants (Q7: $M = 4.25$, $m = 4.00$, $SD = .71$). Across all questions, the median ratings were at or above 4 on a 5-point Likert scale (5 = best; see Fig. 5).

7 Discussion

7.1 Effectiveness at Facilitating Efficient Exploration Process

This work presented the HAVIT as a promising solution for facilitating the efficiency of the AV-pedestrian interaction design and study process. From the feedback from the questionnaire and interviews, we found that the promotion of efficiency mainly comes from three key features of the HAVIT.

The category-based user panel design was found to guide researchers and designers to explore more AV-pedestrian interaction scenarios, even if they have limited knowledge about this field. P2 and P6 mentioned that the HAVIT provided a framework to improve the efficiency of gaining an understanding of this research topic. For example, P2 explained, "The panels are organized logically, with the relevant parameters all together, which gives me a quick idea of which types of the factors to focus on".

The flexible workflow—such as being able to preview the scenarios at any time during the configuration process—not only helped users explore ideas directly but also helped them focus more on the pedestrians and the potential interactions. P5 explained the main benefit of the HAVIT as being that it "Immediately gets you into a headspace for thinking of spatially instead of having to extrapolate in a text document". P4 also explained that previewing the generated scenarios enabled a quick assessment of the reasonableness of the parameter settings by comparing the effects of multiple scenarios.

The HAVIT combines the authoring phase and evaluating phase in a coherent workflow by implementing the Batch Configuration, Data Collection, and Testing Instruction components. P4 mentioned that quickly generating multiple testing scenarios was helpful for avoiding the repetitive configuration process, which might have discouraged exploration and led to thought fixation. P3 and P6 said that being able to collect data and provide instruction tools in the VR environment was a significant advantage, explaining, “the design of these features is reasonable; they fit the needs of the VR-based experiment process and are very convenient for researchers”. P5 added, “I like how fast it is from planning the task to acting it out; it encourages me to try more”.

7.2 The Ease of Use of the Scenario Configuration

The participants’ questionnaire responses were mainly positive and encouraging. Still, from the performance data and researchers’ observations, we identified some key usability aspects that needed to be improved during the configuration process. First, all our participants were able to complete the configuration task in a relatively short period. However, 4 of them made a few errors (1–2 omission errors or input errors) in the first task. Two participants reported that “it would not be easy to find a specific parameter and adjust it when many parameters are involved.” P2 added, “Sometimes I don’t realize that I have adjusted this parameter, so I don’t check if its value is correct”. Although the participants who made these errors reported that they were confident in the configuration process and believed that they would not make similar errors if they used the HAVIT one or two more times. This issue could be circumvented by either providing highlight cues or by implementing a panel to show the parameters that have been set by the user.

In addition, the current user panel features a hierarchy that shows less information, which aimed to improve the efficiency of information access for users. While most users appreciated the usefulness of the user panels in terms of gaining a quick understanding and overview of the information, five of the eight participants mentioned in one way or another that the user panels occasionally became obtrusive and distracting: “There are too many user panels in front of me when I am trying to see and set up parameters” (P5). This feedback emerged after the users became familiar with the system when they started to feel as though they did not need the user panels to be displayed all the time. This finding raises an important question when designing such systems: how can we strike a balance between an intuitive parameter structure and a clear user view while providing both to the user? We believe a further comparative evaluation study with two groups of participants who are given different experiences might help in understanding this phenomenon and identifying a well-balanced solution.

7.3 Prospective Applications of the HAVIT

Based on our investigation of the HAVIT, we see a flexible platform and relatively strong generative power for AV-pedestrian interaction studies. Here, we discuss the potential applications made possible by the HAVIT: (1) Scalability studies of interface concepts related to AVs, which need to evaluate the ability of eHMI concepts to be used in various scenarios with different numbers and types of vehicles, different road structures, etc. (2) The HAVIT can be used to investigate the finer details of the implementation of eHMI concepts, which need to explore how to use and organize interactive elements (e.g., display location, time pattern, and color) in design to communicate crucial messages. (3) Finally, HAVIT can be used to explore pedestrian responses under extreme situations, such as when sensor failures and interface display errors. The above applications are critical to the universality and standardization of AV interaction technology, and they are also significant challenges facing AV-pedestrian interaction research at present.

8 Limitation and Future Work

There are some limitations related to the parameter components of the HAVIT. First, although our testbed is based on generic parameters applicable to AV-pedestrian interaction studies, it provides a limited choice of some parameters. This is because we recognized that some aspects of AV-pedestrian interaction scenarios and influencing factors cannot be fully predicted in advance. We envisioned the HAVIT to be based on core processes and critical functions rather than an all-encompassing solution. To improve the universality of the HAVIT, future versions could include additional road scene types, pedestrian interaction methods, and eHMI interaction prototypes.

Second, the current version of the HAVIT does not support the exploration of interactions among pedestrians. To address this limitation in the future, an initial step could be utilizing characteristically controllable (e.g., in terms of gender, age, moving speed, and group size) virtual pedestrians. Furthermore, we would like to allow multiple participants to be present and tested simultaneously in a scenario. By embedding additional sensory input and body tracking to capture critical features in a user's motion, the HAVIT can support a more realistic, accurate investigation of the effects of interactions between pedestrians. In addition, the HAVIT provides a Preview Mode designed to allow users to preview the final effect of a scenario; however, this preview is based on a two-dimensional (2D) display of the 3D scenario, so there are still some differences in immersion and fidelity between the preview and the final VR scenario.

For the evaluation method, considering the different levels of familiarity of the recruited participants with this research topic, giving them the freedom to construct simulation scenarios may have led to significant differences in the difficulty of the final scenario configuration. Therefore, we assigned them configuration tasks that needed to be completed. However, this causes the experimental results to be in a constrained situation. Further validation is essential to establishing the HAVIT as a research tool. We plan to conduct additional evaluations to benchmark the HAVIT in relation to existing simulation tools. Besides, an end-users (experts in AV-pedestrian interaction domain) goal-driven use case study will be considered in the near future to examine if HAVIT can facilitate their requirements.

9 Conclusion

In this work, we introduced HAVIT, a VR-based platform aimed to facilitate the design process of AV-pedestrian interactions. In this line of work, we structured the key parameter components in a set of panels that users can easily select and modify to create virtual interaction scenarios, with a focus on rich, spatiotemporal interaction between the AVs and pedestrians. We also proposed an efficient workflow with three phases to further ease the creation, iteration, and generation cycles of virtual scenarios. To evaluate the usefulness of HAVIT, we conducted a formative user study with intended professional users. Our evaluation received positive results, indicating that the workflow of the HAVIT is usable and easy to understand. We believe that our workflow can increase the flexibility and efficiency of concept exploration of possible interactions in the early design stage.

References

1. Mahadevan, K., Sanoubari, E., Somanath, S., Young, J.E., Sharlin, E.: AV-Pedestrian interaction design using a pedestrian mixed traffic simulator. In: Proceedings of the 2019 on Designing Interactive Systems Conference, pp. 475–486 (2019)
2. Deb, S., Carruth, D.W., Sween, R., Strawderman, L., Garrison, T.M.: Efficacy of virtual reality in pedestrian safety research. Appl. Ergon. **65**, 449–460 (2017)
3. Habibovic, A., et al.: Communicating intent of automated vehicles to pedestrians. Front. Psychol. 1336 (2018)
4. Rothenbücher, D., Li, J., Sirkin, D., Mok, B., Ju, W.: Ghost driver: a field study investigating the interaction between pedestrians and driverless vehicles. In: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 795–802. IEEE (2016)
5. Ackermann, C., Beggiato, M., Schubert, S., Krems, J.F.: An experimental study to investigate design and assessment criteria: what is important for communication between pedestrians and automated vehicles? Appl. Ergon. **75**, 272–282 (2019)
6. Dey, D., et al.: Distance-dependent eHMIs for the interaction between automated vehicles and pedestrians. In 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 192–204, September 2020
7. Fridman, L., Mehler, B., Xia, L., Yang, Y., Facusse, L.Y., Reimer, B.: To walk or not to walk: crowdsourced assessment of external vehicle-to-pedestrian displays. arXiv preprint arXiv:1707.02698 (2017)
8. Kemeny, A., Panerai, F.: Evaluating perception in driving simulation experiments. Trends Cogn. Sci. **7**(1), 31–37 (2003)
9. Rouchitsas, A., Alm, H.: External human-machine interfaces for autonomous vehicle-to-pedestrian communication: a review of empirical work. Front. Psychol. 2757 (2019)

10. Colley, M., Walch, M., Rukzio, E.: Unveiling the lack of scalability in research on external communication of autonomous vehicles. In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–9 (2020)
11. Dey, D., et al.: Taming the eHMI jungle: a classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. Transp. Res. Interdisc. Perspect. **7**, 100174 (2020)
12. Rasouli, A., Tsotsos, J.K.: Autonomous vehicles that interact with pedestrians: a survey of theory and practice. IEEE Trans. Intell. Transp. Syst. **21**(3), 900–918 (2019)
13. Faas, S. M.: What pedestrians want from autonomous vehicles-Mercedes cooperative car (2018)
14. De Clercq, K., Dietrich, A., Núñez Velasco, J.P., De Winter, J., Happee, R.: External human-machine interfaces on automated vehicles: effects on pedestrian crossing decisions. Hum. Factors **61**(8), 1353–1370 (2019)
15. Holländer, K., Wintersberger, P., Butz, A.: Overtrust in external cues of automated vehicles: an experimental investigation. In: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 211–221 (2019)
16. Nguyen, T. T., Holländer, K., Hoggenmueller, M., Parker, C., Tomitsch, M.: Designing for projection-based communication between autonomous vehicles and pedestrians. In: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 284–294 (2019)
17. Bazilinskyy, P., Kooijman, L., Dodou, D., De Winter, J.C.F.: How should external human-machine interfaces behave? Examining the effects of colour, position, message, activation distance, vehicle yielding, and visual distraction among 1,434 participants. Appl. Ergon. **95**, 103450 (2021)
18. Deb, S., Strawderman, L.J., Carruth, D.W.: Investigating pedestrian suggestions for external features on fully autonomous vehicles: a virtual reality experiment. Transport. Res. F: Traffic Psychol. Behav. **59**, 135–149 (2018)
19. Hussein, A., Garcia, F., Armingol, J.M., Olaverri-Monreal, C.: P2V and V2P communication for pedestrian warning on the basis of autonomous vehicles. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), pp. 2034–2039. IEEE (2016)
20. Dey, D., Martens, M., Wang, C., Ros, F., Terken, J. : Interface concepts for intent communication from autonomous vehicles to vulnerable road users. In: Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 82–86 (2018)
21. Eisma, Y.B., Reiff, A., Kooijman, L., Dodou, D., De Winter, J.C.F.: External human-machine interfaces: effects of message perspective. Transport. Res. F: Traffic Psychol. Behav. **78**, 30–41 (2021)
22. Dey, D., Matviienko, A., Berger, M., Pfleging, B., Martens, M., Terken, J.: Communicating the intention of an automated vehicle to pedestrians: the contributions of eHMI and vehicle behavior. IT-Inf. Technol. **63**(2), 123–141 (2021)
23. Rettenmaier, M., Schulze, J., Bengler, K.: How much space is required? Effect of distance, content, and color on external human-machine interface size. Information **11**(7), 346 (2020)
24. Werner, A.: New colours for autonomous driving: An evaluation of chromaticities for the external lighting equipment of autonomous vehicles. Colour Turn 1 (2018)
25. Löcken, A., Golling, C., Riener, A.: How should automated vehicles interact with pedestrians? A comparative analysis of interaction concepts in virtual reality. In

Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 262–274 (2019)
26. Matthews, M., Chowdhary, G., Kieson, E.: Intent communication between autonomous vehicles and pedestrians. arXiv preprint arXiv:1708.07123 (2017)
27. van Vastenhoven, A.: The effect of distance-dependent eHMIs on the interaction between automated vehicles and two pedestrians. Doctoral dissertation, Master' thesis. Eindhoven University of Technology (2020)
28. Kyriakidis, M., et al.: A human factors perspective on automated driving. Theor. Issues Ergon. Sci. **20**(3), 223–249 (2019)
29. Asha, A.Z., Smith, C., Oehlberg, L., Somanath, S., Sharlin, E.: Views from the wheelchair: understanding interaction between autonomous vehicle and pedestrians with reduced mobility. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–8 (2020)
30. Colley, M., Walch, M., Gugenheimer, J., Askari, A., Rukzio, E.: Towards inclusive external communication of autonomous vehicles for pedestrians with vision impairments. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, pp. 1–14 (2020)
31. Colley, M., Walch, M., Gugenheimer, J., Rukzio, E.: Including people with impairments from the start: external communication of autonomous vehicles. In: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings, pp. 307–314 (2019)
32. Bhagavathula, R., Williams, B., Owens, J., Gibbons, R.: The reality of virtual reality: a comparison of pedestrian behavior in real and virtual environments. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 62, no. 1, pp. 2056–2060. SAGE Publications, Los Angeles (2018)
33. Chang, C.M., Toda, K., Sakamoto, D., Igarashi, T.: Eyes on a car: an interface design for communication between an autonomous car and a pedestrian. In: Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 65–73 (2017)
34. Doric, I., et al.: A novel approach for researching crossing behavior and risk acceptance: the pedestrian simulator. In: Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 39–44 (2016)
35. Schmidt, H., Terwilliger, J., AlAdawy, D., Fridman, L.: Hacking nonverbal communication between pedestrians and vehicles in virtual reality. arXiv preprint arXiv:1904.01931 (2019)
36. Kooijman, L., Happee, R., de Winter, J.C.: How do eHMIs affect pedestrians' crossing behavior? A study using a head-mounted display combined with a motion suit. Information **10**(12), 386 (2019)
37. Feick, M., Kleer, N., Tang, A., Krüger, A: The virtual reality questionnaire toolkit. In Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology, pp. 68–69 (2020)
38. Hartmann, B., Abdulla, L., Mittal, M., Klemmer, S.R.: Authoring sensor-based interactions by demonstration with direct manipulation and pattern recognition. In: Proceedings of CHI 2007, pp. 145–154. ACM (2007)