TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
2024, Vol. 2678(12) 1498–1510

DOI: 10.1177/03611981241252779
journals.sagepub.com/home/trr
Sage

Cooperative Perception System for Aiding Connected and Automated Vehicle Navigation and Improving Safety

Hanlin Chen[1], Vamsi K. Bandaru[2], Yilin Wang[2], Mario A. Romero[2], Andrew Tarko[2], and Yiheng Feng[2]

Abstract

Cooperative perception that integrates sensing capabilities from both infrastructure and vehicle perception sensors can greatly benefit the transportation system with respect to safety and data acquisition. In this study, we conduct a preliminary evaluation of such a system by integrating a portable lidar-based infrastructure detection system (namely, Traffic Scanner [TScan]) with a Society of Automotive Engineers (SAE) Level 4 connected and automated vehicle (CAV). Vehicle-to-everything (V2X) communication devices are installed on both the TScan and the CAV to enable real-time message transmission of detection results in the form of SAE J2735 basic safety messages. We validate the concept using a case study, which aims at improving CAV situation awareness and protecting vulnerable road user (VRU) safety. Field testing results demonstrate the safety benefits of cooperative perception from infrastructure sensors in detecting occluded VRUs and helping CAVs to plan safer (i.e., higher post-encroachment time) and smoother (i.e., lower deceleration rates) trajectories.

Keywords

cooperative perception, connected and automated vehicles, roadside instrumentation, V2X communication, intersection safety, vulnerable road users

Sensing and perception of road traffic and driving environment are critical in autonomous driving and intelligent transportation system (ITS) applications. Camera, radar, and lidar are commonly used types of sensor in transportation systems, for both vehicles and infrastructure. The operational effectiveness of such sensors installed in a vehicle is affected by their limited range and by other vehicles and fixed objects that partially occlude the surroundings of the vehicle. Elevated sensors at the infrastructure side can effectively alleviate the occlusion issue. However, the main downside of roadside sensors is their fixed location, with their detection ability decreasing with distance. This problem might be aggravated with a fixed tilting angle that might be non-optimal for some traffic situations. Nevertheless, adding roadside sensors with real-time data transmission and processing, and fusing complementary information from connected and automated vehicles (CAVs) and roadside sensors in a cooperative sensing and perception system could significantly enhance the vehicles' perception range and navigation quality. This potential improvement is expected to benefit individual CAVs and flow-level traffic control and analysis in many ITS applications.

Although some vehicles are connected and capable of receiving and broadcasting traffic and road information, the initial low percentage of such vehicles does not offer sufficient coverage or rate of information transmission in vehicle-to-vehicle communication. Thus, supplementing the sensing capabilities of CAVs with roadside units is justified. Such a roadside perception system will continue to be beneficial even when the proportion of CAVs is higher, since the information redundancy improves the resilience and reliability of the system. The limitation of most existing infrastructure-based sensing systems is that they are installed at fixed locations (e.g., at intersections)

[1]Applied Research for Mobility Systems, Oak Ridge National Laboratory, Oak Ridge, TN
[2]Lyles School of Civil Engineering, Purdue University, West Lafayette, IN

Corresponding Author:
Yiheng Feng, feng333@purdue.edu

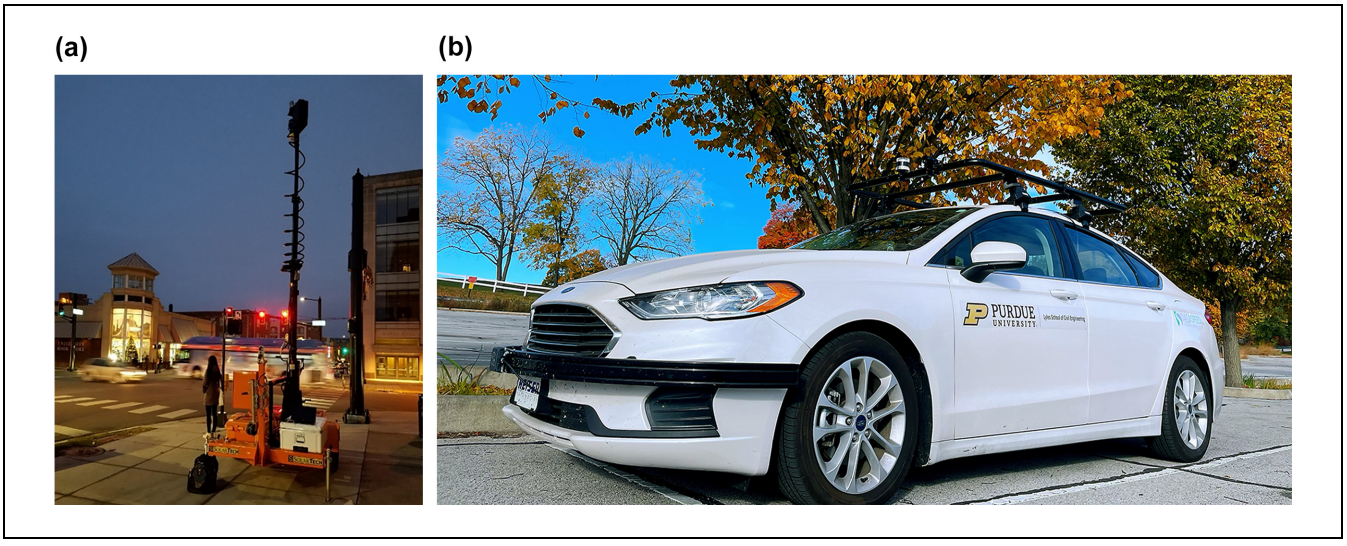

Figure 1. (*a*) Traffic Scanner (TScan) (photo credit: Andrew Tarko); (*b*) Level 4 connected and automated vehicle (photo credit: Yiheng Feng).

and cover a relatively small area (e.g., 100–150 m radius). Because of the high cost of infrastructure sensors, only a limited number of critical locations (e.g., with complex traffic conditions, higher accident rates) are likely to be selected for sensor installation. Furthermore, because of the changing traffic and environmental conditions, the need for such sensing at road network locations might change over time. For instance, although a highway segment is not typically considered a critical location, a temporary work zone or poor pavement conditions during adverse weather might pose a challenge to travelers for safe and efficient navigation in such cases.

To address the aforementioned needs for a cooperative sensing and perception system, we conduct a preliminary evaluation of such a system by integrating a portable lidar-based infrastructure detection system (namely, Traffic Scanner [TScan] [*1*]) with a Society of Automotive Engineers (SAE) Level 4 CAV, as shown in Figure 1. The TScan uses lidar sensors to detect and track various types of road user in real time, including trucks, cars, pedestrians, and bicycles. The CAV is equipped with a wired control system and various onboard sensors, including lidar, camera, and radar. In addition, vehicle-to-everything (V2X) communication devices are installed on both the TScan and the CAV to enable real-time message transmission of perception data, including detection results. We validate the functionality of the cooperative perception system and evaluate its benefit using a case study, which aims at improving CAV situation awareness and protecting vulnerable road user (VRU) safety. Field testing results show that the cooperative perception system provides the vehicle with advanced information, which would otherwise not be available because of occlusions, in a timely manner; this improves both safety and driving comfort.

The main contributions of this paper can be summarized as follows.

1. We conduct a preliminary evaluation of a lidar-based infrastructure–vehicle cooperative perception system between the TScan and CAV.
2. We demonstrate the safety benefits of such cooperative perception systems in real-world applications (i.e., VRU protection) using field tests.

The rest of the paper is organized as follows. The next section first gives an overview of the system and then introduces each system component in detail. Then follow an analysis of the detection accuracy of the TScan and a description of a case study to show the benefit of the cooperative perception system. The conclusion of the paper includes a discussion of potential implementation scenarios.

Cooperative Perception System

System Overview

Figure 2 shows the overall structure of the cooperative perception system. In the vehicle platform, a lidar sensor is used for perception, owing to its high-level accuracy in providing depth information and a 360° field of view. The main reason for choosing lidar over other perception sensors is that lidar is usually considered to have high detection accuracy among all perception sensors, especially in providing accurate distance measurements in the

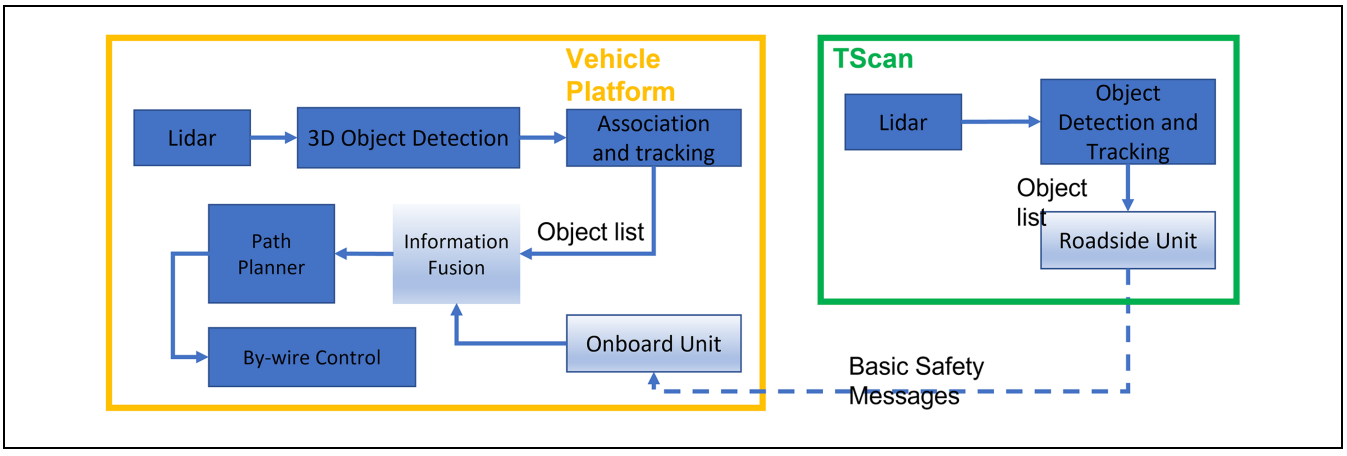

Figure 2. Overall structure of the cooperative perception system.

3D environment; this is highly important in safety-critical applications, such as VRU protection. Furthermore, lidar is not sensitive to changes in ambient light conditions; therefore, it can also work during nighttime. A clustering-based 3D object detection module from the Autoware autonomous driving platform is used to detect the surrounding objects and the Hungarian algorithm (*2*) plus Kalman filter (*3*) association and tracking module is applied to track the objects and generate continuous trajectories. The object list generated from the tracking algorithm is fed into an information fusion module. Meanwhile, the TScan also utilizes lidar and applies a similar object detection and tracking pipeline to generate a detected object list. Coordinate transformation is performed to convert local coordinates to Global Positioning System (GPS) coordinates. Each object in the list is then encoded into an SAE J2735 basic safety message (BSM) and broadcast through a roadside unit (RSU). Note that if there are several objects in the list, then several BSMs are sent at the same time. An onboard unit (OBU) is installed on the vehicle platform to receive the BSMs and to forward the information to the information fusion module. Two Cohda MK5 devices are used to form the V2X communication device. The immediate forwarding of messages (IFM) function is used to send the detection results through a radio signal. Note that although MK5 devices only support dedicated short range communication (DSRC), the same message protocol and forwarding function can be applied in the V2X communication environment. The information fusion unit uses a predefined area (i.e., the roadway area) to geofence the objects and fuses the object lists by removing redundant observations. The fused information is then input to a path planner, in which a customized heuristic algorithm provides longitudinal vehicle control, while the pure pursuit algorithm (*4*) provides lateral vehicle control. Finally, the path planner outputs the vehicle speed and steering angle to a wired control system for execution. In the vehicle platform, all components are synchronized through the robot operating system (ROS).

In the next section, the components of the TScan and the vehicle platform are described in detail.

TScan

The TScan is a portable microscopic traffic data acquisition system that utilizes lidar technology. Two trailer-based prototypes were developed at the Purdue Center for Road Safety with support from the Joint Transportation Research Program (JTRP) of the Indiana Department of Transportation and Purdue University and the NEXTRANS Center at Purdue University. The hardware platform of the TScan is shown in Figure 3.

The TScan has two lidar sensors:

- Velodyne HDL 32E, with a maximum range of 80 m and a vertical field of view from + 10° to −30°
- Ouster OS2 64 BH, with a maximum range of 200 m and a vertical field of view from 0° to −11.25°.

The sensors are mounted on pan-tilt motors. The two sensors' range and field of view complement each other and are combined with the pan-tilt motors. This provides users with the flexibility to tilt the sensors individually, as required, to cover the intersection effectively, minimizing the blind spot caused as a result of the limited vertical field of view of the lidar sensors. Data from both the lidar sensors are fused at the point cloud level. Perception algorithms described in Bandaru et al. (*1*) are applied to the fused point cloud data to detect objects. A fisheye camera is used to record video, which is used for

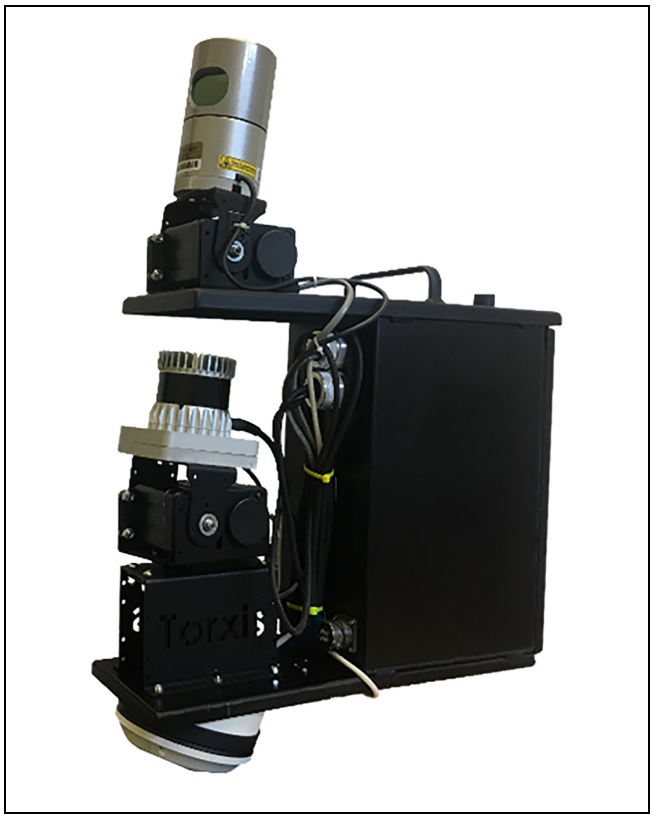

Figure 3. Traffic Scanner (TScan) hardware platform (photo credit: Andrew Tarko).

verification of events by human operators. A Garmin 18x LVC GPS unit is used to obtain GPS coordinates in the field and keep time. All the sensors and motors are mounted on a custom-fabricated aluminum casing. All the electronics needed to control and operate the sensors and motors are present inside the enclosed portion of the casing, along with a fanless industrial computer equipped with an eighth-generation Intel core i7 processor to process the data in real time.

The TScan uses a few minutes of data to identify background. Then, during real-time operation, the background is removed and the remaining points are clustered. Each cluster potentially represents a moving object. These clusters are tracked over time using a combination of a Kalman filter, for state estimation, and a Hungarian assignment algorithm, for association. The corrected Kalman estimates for every object in the field of view are then broadcast every 0.1 s. More details of the TScan can be found in Bandaru et al. (*1*).

Vehicle Platform

The CAV testing platform is equipped with sensors including a 64-layer high-resolution lidar sensor, a long-range radar for vehicle and collision detection, and an RGB camera for front view image perception. For vehicle localization, it features a dual-antenna high-resolution GPS with a real-time kinematics correction function enabled, and an inertial mass unit installed for kinematics measurement.

Perception Based on 3D Point Cloud Data. In this work, we mainly use lidar as the perception sensor, deployed in Autoware AI (*5*). First, we execute a point cloud downsampling process, with the aim of downsampling the raw point cloud data obtained from the lidar sensor on the CAV. The primary goal of downsampling is to eliminate noise from the raw point cloud data, a process influenced by the number of points within a voxel. We have configured the voxel size to be 0.01 m and the measurement range to be 200 m.

After downsampling the point cloud, we execute the first stage of ground removal using a ray ground filter from the Autoware platform. In the downsampled point cloud, we initially separate the ground details radially. Then the ground is identified using geometric information related to the ego vehicle. Within each ray, we determine whether a point belongs to the ground based on the distance and the angle between points. This method can successfully be used to remove the ground that is far away from the ego vehicle, thereby potentially reducing errors in the downstream clustering algorithm. In this study, we set the clipping height at 1 m and the minimum point distance at 1.5 m. We divide the point cloud into different rays at intervals with a radius of 0.08, and points are checked if their corresponding radius is larger than 0.01. For ray ground filtering, we set the local maximum slope at 8, the general maximum slope at 5, and the minimum height threshold at 0.05. Points will be rechecked and reassigned to different classes if their distance to the closest point is greater than 0.2 m.

After the ground removal, the Euclidean-based clustering method is used for clustering-based detection, to identify the locations of vehicles and VRUs. Data points that are less than 60 cm above the ground are ignored. This setting is determined by the height of the lidar sensor mounted on the CAV to avoid impact with stationary road obstacles, such as cement piers, which are near the area of interest of the experiment. In the clustering algorithm, the clustering distance is set to 0.75 m. The minimum number of points within a cluster is set to 20, and the maximum to 100,000. We also perform another voxelization process to downsample the point cloud and ensure the accuracy of the clustering result. At this stage, the leaf size for clustering is set to 0.2 m. A size-based naïve filtering method is applied to filter out stationary and large objects based on the results from the Euclidean point cloud clustering algorithm. Only clusters of lengths, widths, and heights ranging between 0.1 m and 1 m are

recognized as VRUs. The lower bound is incorporated to reduce the effect of noisy predictions.

Association and Tracking Algorithm. After obtaining results from the perception module, the detection results are assigned to tracks and then updated with a Kalman filter. This Kalman filter-based tracking mechanism is inherited from the Simple Online and Real-time Tracking (SORT) algorithm (*6*), which was previously employed for image-based object tracking. The SORT algorithm operates by combining detection information (e.g., from object detectors) with a simple linear assignment problem (LAP) solver for frame-to-frame tracking of object identities.

Detection results, accompanied with corresponding time information, are obtained using the perception module. A constant-speed Kalman filter is then applied to predict the expected object location from the previous detection. Subsequently, a cost matrix is calculated based on the prediction and the current observation result. Using the cost matrix, the Hungarian matching algorithm (*2*) is applied to match the prediction with the current observation. For each initial matching result, the cost is checked to ensure that it does not exceed a predefined threshold. If the cost is larger than the threshold, the prediction for the previous observation is unmatched, and its corresponding track is updated with the prediction only. The unmatched observations are used to initialize new tracks, and each observation is considered as the first frame of each new track. If the prediction from the previous observation matches the current observation, the corresponding track is updated with the result of the Kalman filter, which is the posterior result after correction with the matched observation. In this framework, the quality of association relies on the detector and, thus, is sensitive to noise.

Information Fusion. After obtaining data from the vehicle perception module and BSMs from the TScan, we implement decision-level information fusion to reduce redundancy in the downstream path planning module. As implemented in the CARMA system (*7*), we execute decision-level fusion.

In this work, given the same object, we might have detections from both the CAV and the TScan. Duplication of detections can affect the performance of the downstream module. Therefore, we first perform a redundancy check and then combine duplicated detections, based on location data. If the distance between two detections is less than a certain threshold (e.g., one vehicle length), we consider them to belong to the same vehicle and use the averaged location.

Owing to the complexity of the experiment site, there might be objects with similar locations but completely different kinematic profiles. Therefore, the speed of detection results is also taken into consideration when performing decision-level merging. Based on the location, speed, and dimensional information of the detected object, we construct a complete bipartite graph $G = (S, T; E)$ for decision-level fusion. In this bipartite graph, S represents the set of infrastructure vertices (i.e., bounding boxes detected from the infrastructure side), T represents the set of vehicle vertices (i.e., bounding boxes detected from the vehicle side), and E denotes the set of edges in the graph connecting S and T. Each edge carries a nonnegative cost $c(i,j)$ of objects i and j. This cost function is the weighted sum of various factors, including vehicle dynamic features (e.g., location and speed) and appearance features (e.g., the shape of the bounding box). Any edge that carries a cost higher than a predetermined threshold is removed. For vertices connected to at least one edge, we apply the Kuhn–Munkres algorithm (*8*) to find the maximum matching between two observations representing a single object. Any unmatched vertices are considered to be unique observations. We retain these as the results of the cooperative perception process.

In addition to redundancy removal, we conduct geofencing to eliminate all detected objects outside the anticipated detection area. The results, after decision-level fusion from the CAV perception and the TScan detection, are then fed into the CAV's path planning module, which will be introduced in the next section.

Path Planner. The path planning for the CAV can be divided into two parts: lateral control and longitudinal control. Similar to the work of Chen et al. (*9*), the pure pursuit lateral controller is utilized for trajectory following. For longitudinal control of the CAV, the Gipps car-following model (*10*) is utilized to interact with objects in the same lane. Finally, a post-encroachment time (PET)-based speed planner (*11*) is implemented when the CAV is interacting with objects from the side (e.g., at an intersection).

For the PET-based speed planner, once an object is detected, we first calculate the potential conflict point location, considering the object's current location, its speed in the X and Y directions, and the planned trajectory of the CAV. We then ascertain whether the object is close to the conflict point, and a virtual front vehicle (serving as a red signal) is placed at the conflict point if the object is within a predefined geofencing area, regardless of its speed. In this case, the CAV will yield to the object to avoid a crash. If the object is outside the predefined geofencing area, we calculate the PET for the object and the CAV. If the PET for the object and the CAV is greater than 2 s, then the CAV will maintain its constant speed. However, if the PET for the object and CAV is less than 2 s, we determine whether the object has already crossed the conflict point. If not, the CAV's

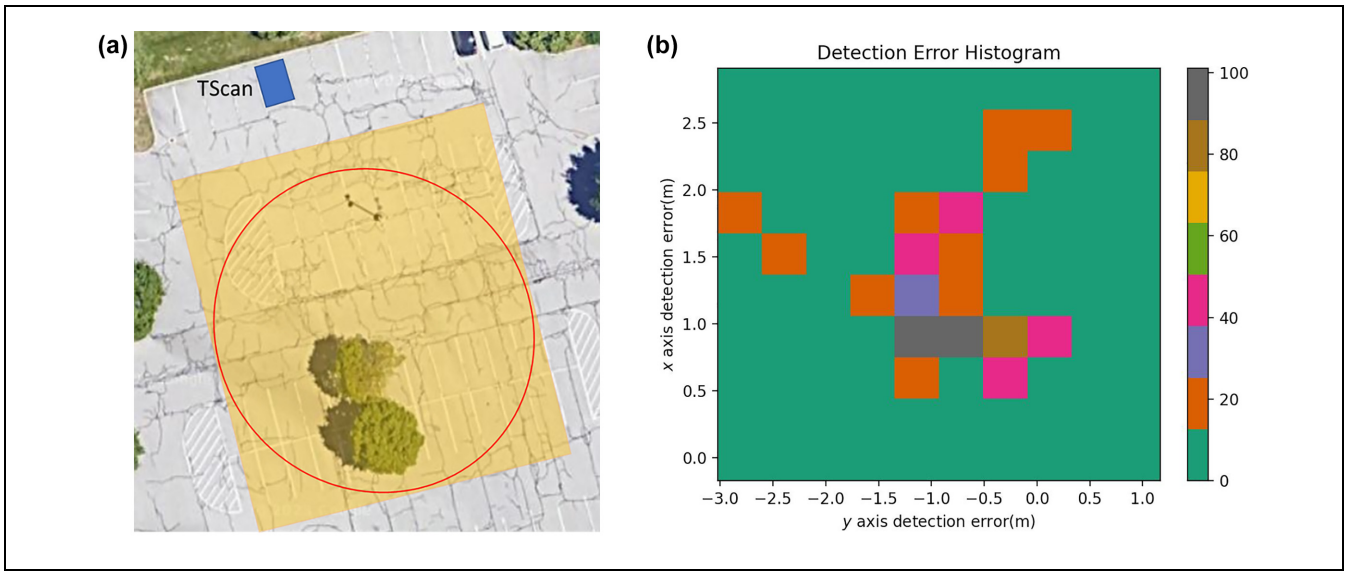

Figure 4. Traffic Scanner (TScan) detection error analysis: (*a*) test setup and (*b*) error distribution.

speed is reduced to obtain the 2 s PET (and also yield to the object). If the object already crossed the conflict point more than 2 s before, the CAV resumes the free-flow speed.

Experiments

TScan Detection Accuracy Test

First, we validate the accuracy of the TScan detection results. It is critical that the TScan sends accurate information about object locations. The experiment was conducted in the North Stadium parking lot on Purdue University campus on March 17, 2023. The TScan was staged at the north end of the parking lot, while the CAV circled the parking lot, as shown in Figure 4*a*. The blue block represents the TScan. The yellow shaded area represents the geofencing area of the TScan and the red circle approximates the CAV's driving route. Detection results from the TScan (i.e., the CAV) were sent through BSMs to the CAV. The location data in the BSMs (i.e., GPS coordinates) were compared with the GPS coordinates collected from the CAV's real-time kinematics (RTK) GPS. We consider the coordinates collected from the RTK GPS as ground truth since they are given with centimeter-level accuracy.

The distribution of the detection error is shown in Figure 4*b*. The coordinates are defined to be consistent with the vehicle coordinate system, where the *X*-axis points forward from the vehicle and the *Y*-axis points to the left when the vehicle is facing forward. Most of the errors in the *Y*-axis range from −1.5 m to 0.5 m, with a few outliners between −2.5 m and −3 m. There is a systematic error in the *X*-axis since all errors are positive, ranging from about 0.5 m to 2.5 m. This is because the TScan considers the center of the bounding box as its coordinate while the RTK GPS receiver is not installed in the middle of the vehicle. The distance from the GPS receiver to the center of the vehicle along the *X*-axis is about 1.60 m. As a result, the actual error distribution along the *X*-axis is about −0.9 m to 0.9 m. Since the distance from the GPS receiver to the center of the vehicle along the *Y*-axis is only 0.09 m, we ignore this small difference. In summary, the error distribution validates the detection accuracy of the TScan.

Case Study: Scenario Setup

We designed a case study to demonstrate the benefit of the cooperative perception system in improving VRU safety at intersections. The experiment was staged at the Purdue Research Park parking lot, as shown in Figure 5*a*. In the experiment design, a VRU travels from west to east (red route), while the CAV travels from south to north (green route). These routes have a conflict point at the T-intersection. Because of the building and parked vehicles (blue blocks), the VRU is occluded and cannot be observed by the onboard sensors of the CAV when it approaches the T-intersection. The TScan (yellow circle) is located at the northeast corner of the intersection so that it can observe the VRU and CAV at the same time. The orange-shaded areas represent the TScan's geofencing areas. To ensure safe tests, we constructed a model VRU by attaching two large U-Haul boxes to a Backfire G2 skateboard, as shown in Figure 5*b*. The size of the model VRU is similar to a teenager (about 50 in. tall); we named him Reckless Jack. The skateboard can be

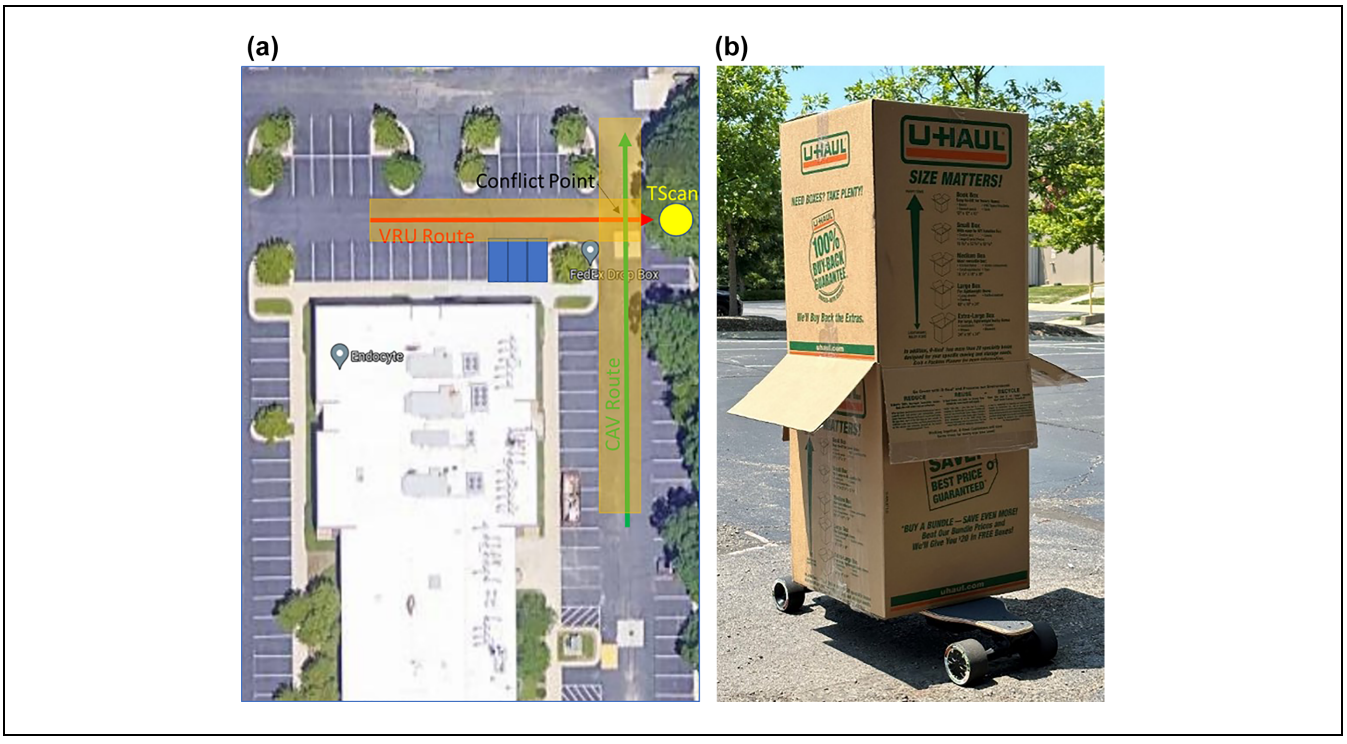

Figure 5. Case study scenario design: (*a*) scenario demonstration; (*b*) vulnerable road user, "Reckless Jack". (Photo credit: Yiheng Feng.)

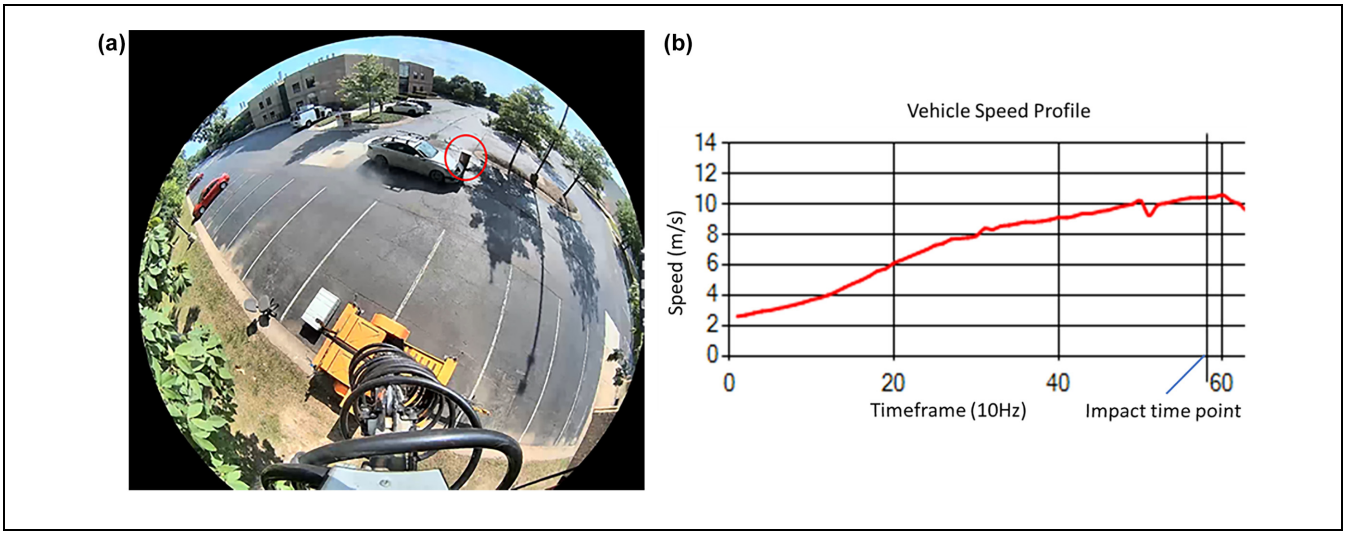

Figure 6. Results of Experiment 1: (*a*) impact of vehicle and vulnerable road user (VRU); (*b*) vehicle speed profile. (Photo credit: Vamsi Bandaru.)

controlled remotely to move at different speeds. We conducted three different experiments, as follows.

- *Experiment 1*. The VRU and the vehicle approach the intersection at the same time without the TScan; the vehicle is driven by a human driver.
- *Experiment 2*. The VRU and the vehicle approach the intersection at the same time without the TScan; the vehicle is driven autonomously.
- *Experiment 3*. The VRU and the vehicle approach the intersection at the same time with the TScan; the vehicle is driven autonomously.

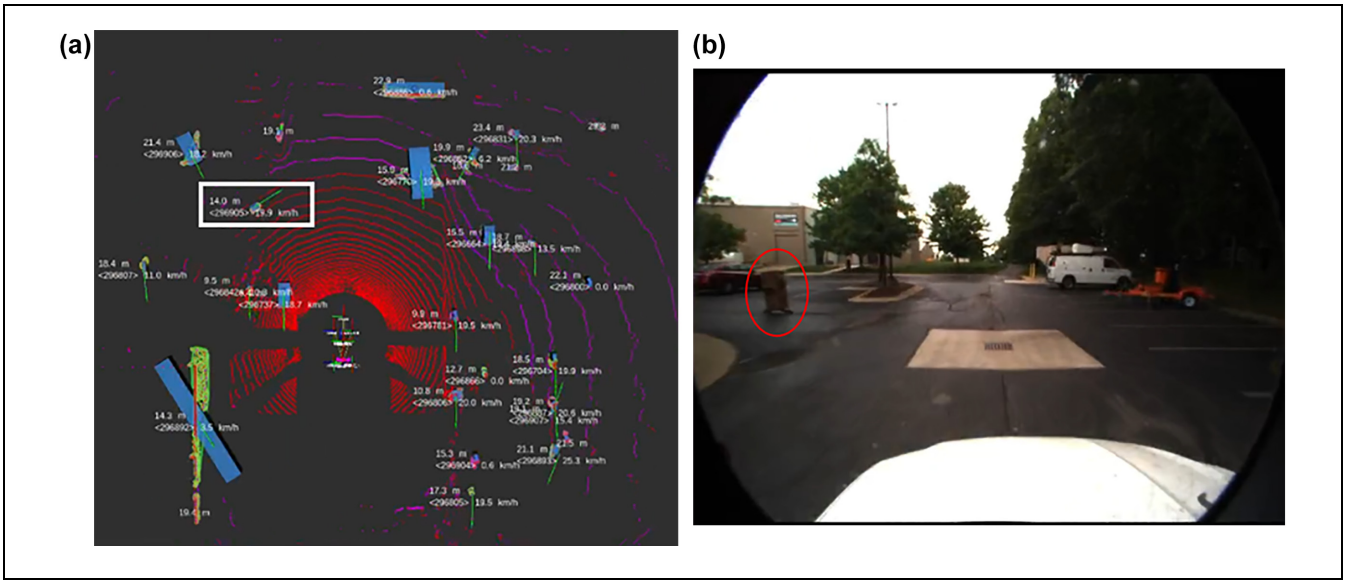

Figure 7. Experiment 2: lidar and camera view: (*a*) lidar detection results; (*b*) camera view. (Photo credit: Hanlin Chen).

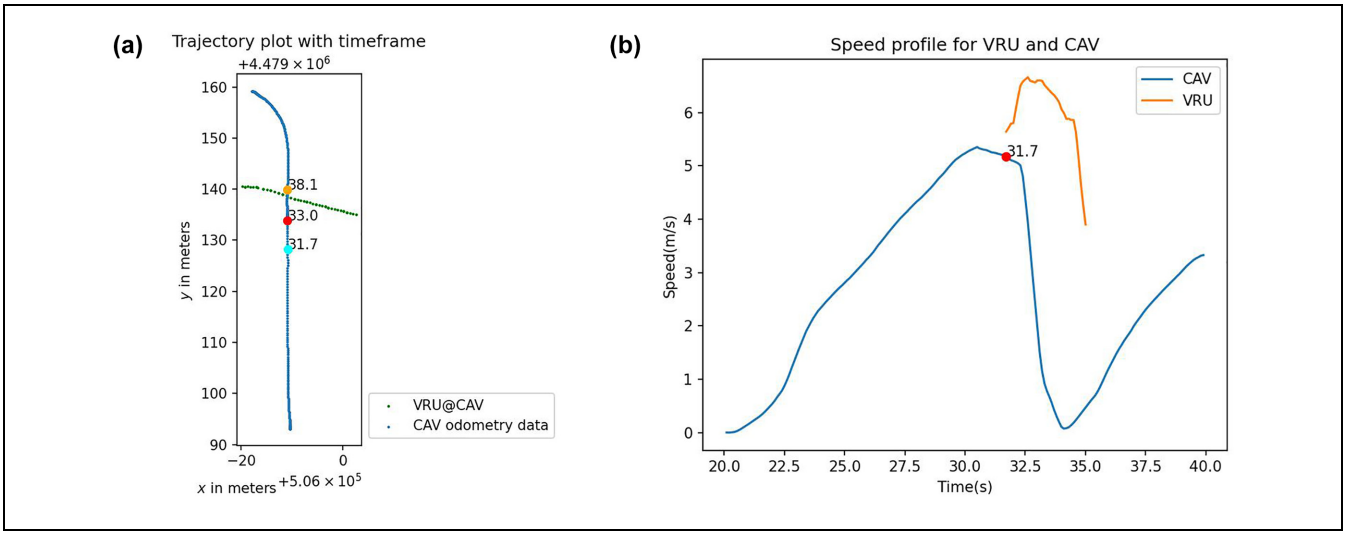

Figure 8. Connected and automated vehicle (CAV) and vulnerable road user (VRU) trajectory and speed profiles of Experiment 2: (*a*) trajectory profiles; (*b*) speed profiles.

In Experiments 2 and 3, the free-flow speed of the CAV was set to 5 m/s. In Experiment 1, the driver was asked to drive at a similar speed. Our central hypothesis is that the TScan is able to detect the VRU before he arrives at the intersection, leading to a safer and smoother vehicle reaction.

Test results are presented next.

In the first experiment, a human driver (from our research team) was asked to drive the vehicle to pass the intersection without knowing about the VRU's existence. The driver did not even notice the appearance of the VRU from the left and crashed into him, as shown in Figure 6*a*. The speed profile of the vehicle observed from the TScan is shown in Figure 6*b*. It can be seen that, at the time of impact, the driver had not taken any actions (e.g., braking). Note that the vehicle's speed exceeded the suggested speed (i.e., 5 m/s) at the time of impact; this might also be the cause of the crash.

In the second experiment, the autonomous driving system took control of the vehicle, followed the same route, and interacted with the VRU. Figure 7*a* is a snapshot of the first time point, in which the VRU was detected by the lidar sensor (white box). The distance to the CAV was only 14.0 m. Figure 7*b* shows the front camera view

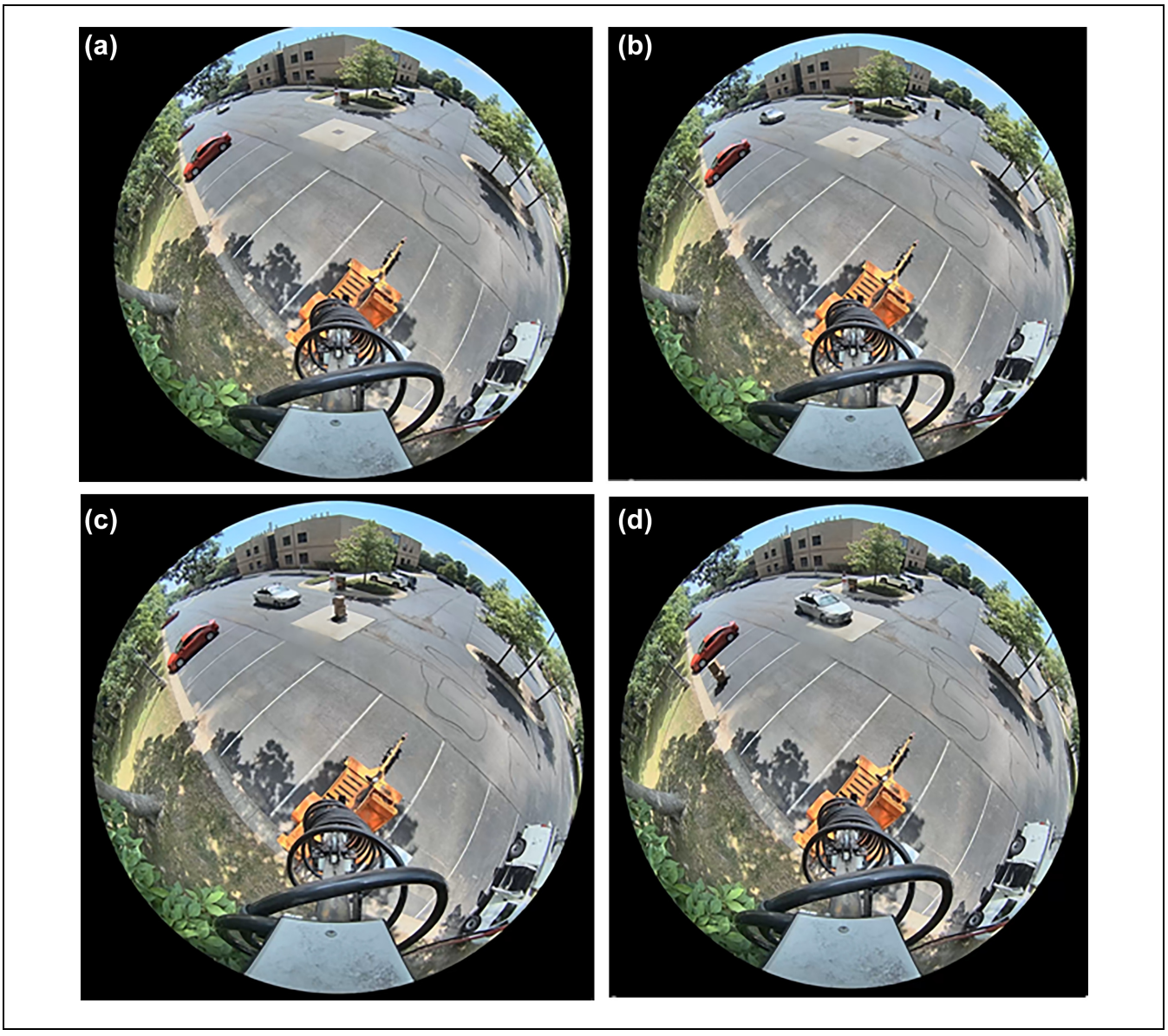

Figure 9. Experiment 3 scenario demonstration: (*a*) objects near the intersection at 5:20 s, (*b*) objects started interaction at 5:23 s, (*c*) connected and automated vehicle (CAV) stopped for vulnerable road user at 5:26 s; (*d*) CAV passed conflict point at 5:30 s. (Photo credit: Vamsi Bandaru.)

at the same moment. It can be seen that, when the lidar sensor captured the VRU, he was already very close to the intersection. The CAV detected a potential conflict and applied hard braking to avoid the crash. The trajectory and speed profiles of the CAV and the VRU (after detection) are shown in Figure 8. The three time points marked in the trajectory plot are the first time that the VRU was detected (31.7 s), the time when the VRU passed the conflict point (33.0 s), and the time when the CAV passed the conflict point (38.1 s). From the CAV's speed profile, it can be seen that, after the detection of the VRU at 31.7 s, it applied hard braking with a maximum deceleration of $-5.12\,\mathrm{m/s^2}$ to avoid the crash. Compared with Experiment 1, with a human driver, the vehicle was able to make a complete stop and avoid a crash, owing to the detection of the VRU and the shorter reaction time.

In the third experiment, the CAV and VRU approached the intersection in the same way as in Experiment 2. In this experiment, the TScan detected the objects in the geofenced area and broadcast details of detected objects in real time to the vehicle through BSMs. Figure 9 shows a few snapshots from the TScan fisheye camera recording of the whole scenario.

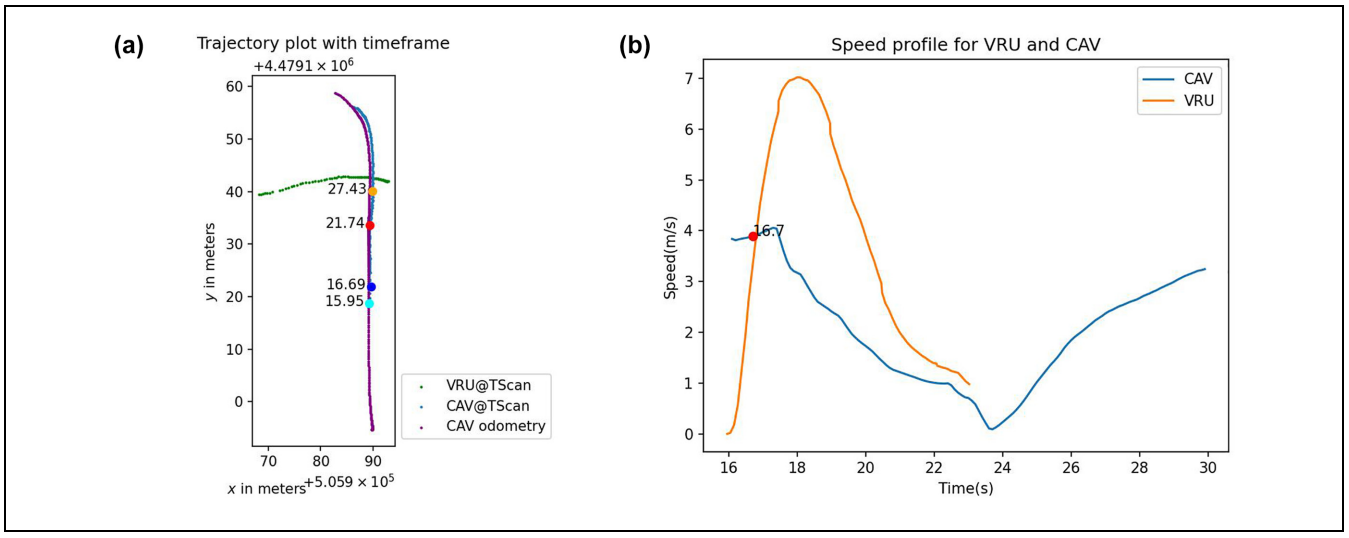

Figure 10. Connected and automated vehicle (CAV) and vulnerable road user (VRU) trajectory and speed profiles of Experiment 3 (from CAV): (*a*) trajectory profiles; (*b*) speed profiles.

The trajectory and speed profiles of the CAV and the VRU collected from the vehicle platform are shown in Figure 10. The four time points marked in the trajectory plot are the first time that the VRU's information was received by the CAV from the TScan (15.95 s), the time when the CAV started to react to the VRU, owing to the planned PET (16.69 s), the time when the VRU passed the conflict point (21.74 s), and the time when the CAV passed the conflict point (27.43 s). The speed profile of the CAV shows a much smoother trend, with an average deceleration of $-0.49\,m/s^2$ and a maximum deceleration of $-2.21\,m/s^2$, which are much lower than those in Experiment 2.

Figure 11 shows the corresponding trajectory and speed profiles recorded from the TScan for the same experiment. The speed profile of the CAV also shows that it has slowed down to allow the VRU to pass through. The thin yellow line shows the predicted future trajectory of the detected objects. The dotted yellow shows the past trajectory.

This experiment demonstrates the benefit of cooperative perception from infrastructure sensors in improving CAV situation awareness, safety, and driving comfort, as well as VRU safety. The safety effect can be further demonstrated by the PET plots between the CAV and VRU given in Figure 12. Figure 12*a* shows the PET profile in Experiment 2 from the first detection of the VRU until the VRU had passed the conflict point. When the VRU was detected, the PET was already very low. Although the vehicle braked hard, the final PET when the VRU passed the conflict point was still less than 1 s. Figure 12*b* shows the PET profile in Experiment 3 from the time that the TScan detected the VRU until the VRU passed the conflict point. It can be seen that the PET dropped from 5 s to around 1 s when the CAV started to react to the VRU (at 16.69 s). Afterward, the PET started to increase and maintained within a safe boundary.

Conclusions and Discussion

In this study, we conducted a preliminary evaluation of a cooperative perception system by integrating a portable infrastructure detection system (TScan) with a CAV. Both the TScan and the vehicle platform used 3D lidar-based detection to perceive the driving and traffic environment. The detection results were transmitted and shared through V2X communications in real time. A field experiment was designed to determine the benefit of cooperative perception in improving CAV situation awareness and VRU safety. The test results demonstrate that, with perception data from the infrastructure, the CAV was able to detect the VRU earlier to avoid crashes and plan a smoother vehicle trajectory. Although the evaluated system shows benefits from cooperative perception, there are still limitations, including the following. (1) The system has a blind spot, which is around the deployment location of the lidar sensor. To solve this problem, several sensors are needed to provide full coverage. (2) Because of the mechanism of lidar detection, it is difficult to identify the class of the object (e.g., a bicycle or a scooter). Cameras should be added to conduct object classification and the results should be fused with the lidar detection results. (3) Lidar sensors may have reduced performance in adverse weather conditions, such

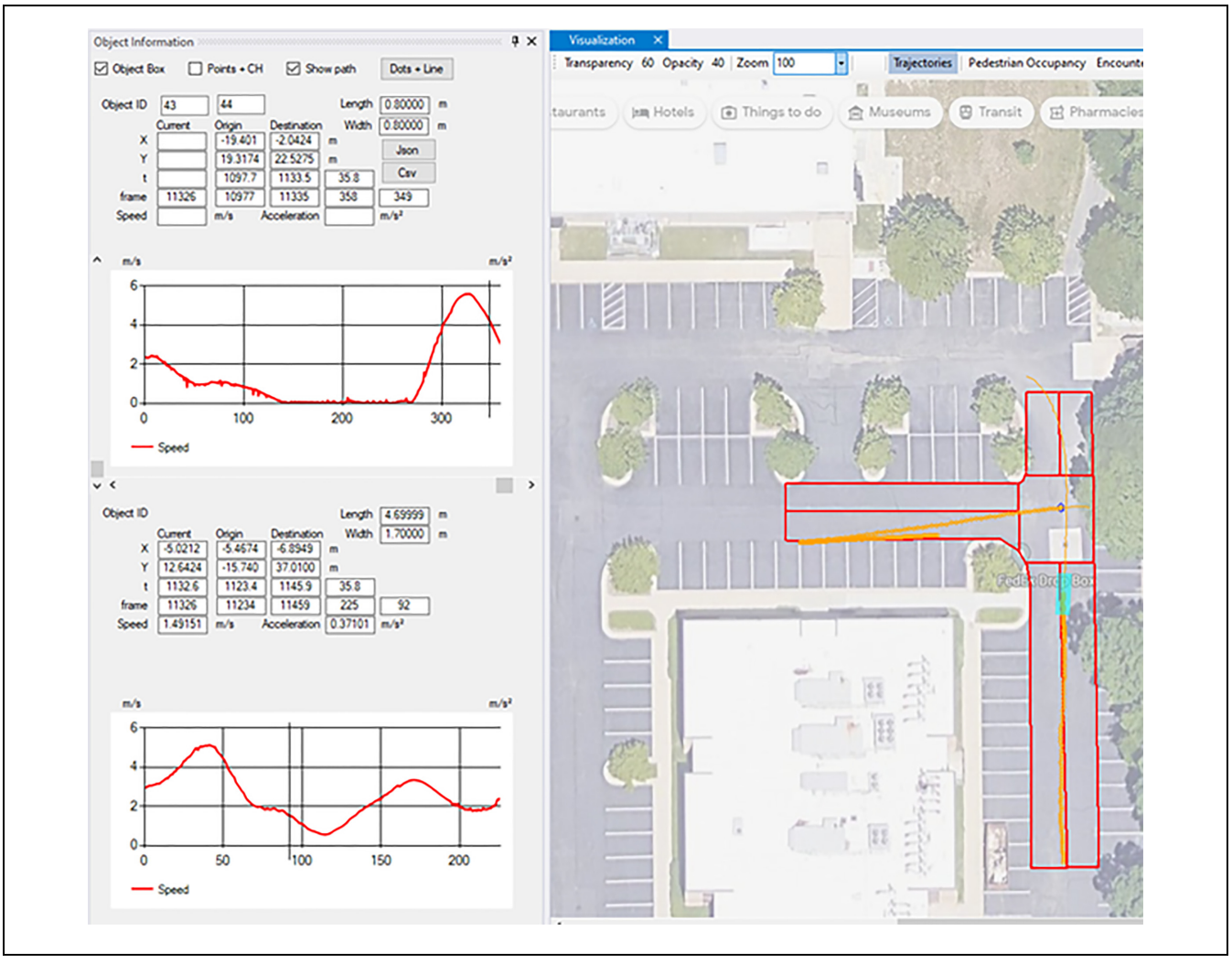

Figure 11. Connected and automated vehicle (CAV) and vulnerable road user (VRU) trajectory and speed profiles of Experiment 3 (from the TScan).

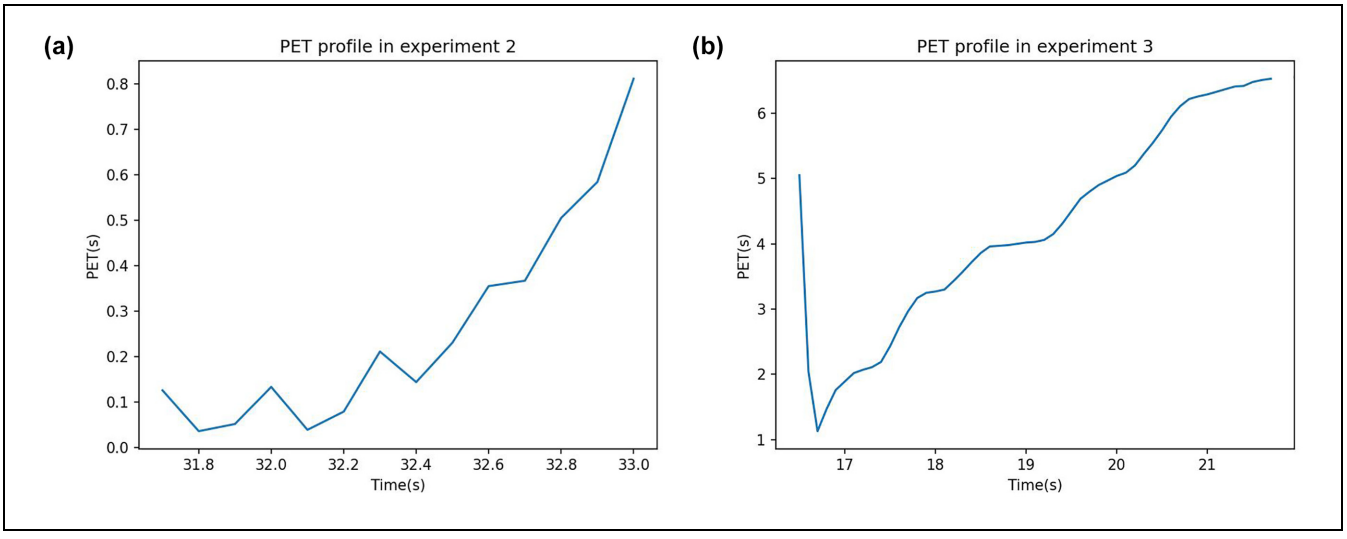

Figure 12. Post-encroachment time (PET) profile comparison: (*a*) PET profile in Experiment 2; (*b*) PET profile in Experiment 3.

as heavy rain or snow. (4) Owing to the changed location and field of view of each deployment, the lidar sensors need to be calibrated every time for reference point calculation, coordinates transformation, and geofencing. An automated calibration process is needed for fast deployment in the future.

We mainly utilized the cooperative perception system for CAV path planning and VRU protection to improve safety. Such systems can be implemented in many other transportation and vehicle applications. First, different sensor fusion models can be tested based on transmitted data. In our implementation, decision-level fusion is applied where the information is transmitted as BSMs. If the raw sensor data are transmitted, data level fusion models, such as those of Zhang et al. (*12*), can be tested. Another alternative is the use of feature level fusion models, such as those of Bai et al. (*13*), where only extracted features (e.g., using convolutional neural networks [CNNs]) are transmitted, instead of raw sensor data. The performance of different fusion technologies can be evaluated from several perspectives, such as accuracy and latency. The cooperative perception system can also serve as a new data source for traffic control applications, such as intersection management (*14*, *15*) and eco-routing (*16*, *17*). However, only a limited number of studies have explored the differences in quality, quantity, and distribution between the data from cooperative perception and existing sources, such as loop detectors and connected vehicle trajectories. For example, traffic data collected from cooperative perception are highly clustered around sensors (e.g., a CAV or an infrastructure location), rather than randomly distributed on a roadway. Based on the new data pattern, existing traffic state estimation or traffic control models may need to be modified. The portability of the infrastructure perception system facilitates its deployment in temporary sites, such as work zones. Through real-time detection and prediction, potential work zone intrusion can be detected in advance and warning messages can be sent promptly to workers' smart vests, through V2X communication, to improve safety. Finally, offline processed trajectories collected from a cooperative perception system can be further utilized to investigate long-term safety performance, such as traffic conflicts or crash estimation (*18*, *19*).

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: H. Chen, Y. Feng, A. Tarko; data collection: H. Chen, V. K. Bandaru, Y. Wang, M. A. Romero; analysis and interpretation of results: H. Chen, V. K. Bandaru, Y. Feng, A. Tarko; draft manuscript preparation: H. Chen, Y. Feng, A. Tarko, V. K. Bandaru, Y. Wang, M. A. Romero. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research is supported in part by the U.S. National Science Foundation (NSF) through Grant CPS #2038215, and U.S. DOT Region 5 University Transportation Center (Center for Connected and Automated Transportation).

ORCID iDs

Hanlin Chen https://orcid.org/0000-0001-6508-7715
Vamsi K. Bandaru https://orcid.org/0000-0002-9222-257X
Yilin Wang https://orcid.org/0000-0002-6002-0535
Andrew Tarko https://orcid.org/0000-0001-9301-3270
Yiheng Feng https://orcid.org/0000-0001-5656-3222

References

1. Bandaru, V. K., M. A. Romero, C. Lizarazo, and A. P. Tarko. *TScan – Stationary Lidar for Traffic and Safety Applications: Vehicle Interpretation and Tracking*. Joint Transportation Research Program Publication No. FHWA/IN/JTRP-2021/31. Purdue University, West Lafayette, IN, 2021.
2. Kuhn, H. W. The Hungarian Method for the Assignment Problem. *Naval Research Logistics Quarterly*, Vol. 2, No. 1–2, 1955, pp. 83–97.
3. Welch, G., and G. Bishop. *An Introduction to the Kalman Filter*. No. 95-041. University of North Carolina at Chapel Hill, 1995.
4. Coulter, R. C. *Implementation of the Pure Pursuit Path Tracking Algorithm*. CMU-RI-TR-92-01. The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 1992.
5. Autoware Foundation. Lidar Euclidean Cluster Detect. Autoware AI Perception. *GitHub*, 2024. https://github.com/autowarefoundation/autoware_ai_perception/tree/master/lidar_euclidean_cluster_detect. Accessed February 23, 2024.
6. Bewley, A., Z. Ge, L. Ott, F. Ramos, and B. Upcroft. Simple Online and Realtime Tracking. *Proc., 2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, IEEE, New York, September 2016, pp. 3464–3468.
7. Chen, H., B. Liu, X. Zhang, F. Qian, Z. M. Mao, and Y. Feng. A Cooperative Perception Environment for Traffic Operations and Control. *arXiv Preprint arXiv:2208.02792*, 2022.
8. Zhu, H., D. Liu, S. Zhang, Y. Zhu, L. Teng, and S. Teng. Solving the Many to Many Assignment Problem by Improving the Kuhn–Munkres Algorithm with Backtracking. *Theoretical Computer Science*, Vol. 618, 2016, pp. 30–41.

9. Chen, H., R. Luo, and Y. Feng. Improving Autonomous Vehicle Mapping and Navigation in Work Zones Using Crowdsourcing Vehicle Trajectories. *arXiv Preprint arXiv:2301.09194*, 2023.
10. Gipps, P. G. A Behavioural Car-Following Model for Computer Simulation. *Transportation Research Part B: Methodological*, Vol. 15, No. 2, 1981, pp. 105–111.
11. Gettman, D., and L. Head. Surrogate Safety Measures from Traffic Simulation Models. *Transportation Research Record: Journal of the Transportation Research Board*, 2003. 1840: 104–115.
12. Zhang, X., A. Zhang, J. Sun, X. Zhu, Y. E. Guo, F. Qian, and Z. M. Mao. Emp: Edge-Assisted Multi-Vehicle Perception. *Proc., 27th Annual International Conference on Mobile Computing and Networking*, New Orleans, LA, Association for Computing Machinery, New York, October 2021, pp. 545–558.
13. Bai, Z., G. Wu, M. J. Barth, Y. Liu, E. A. Sisbot, and K. Oguchi. Vinet: Lightweight, Scalable, and Heterogeneous Cooperative Perception for 3D Object Detection. *Mechanical Systems and Signal Processing*, Vol. 204, 2023, p. 110723.
14. Feng, Y., K. L. Head, S. Khoshmagham, and M. Zamanipour. A Real-Time Adaptive Signal Control in a Connected Vehicle Environment. *Transportation Research Part C: Emerging Technologies*, Vol. 55, 2015, pp. 460–473.
15. Yang, Z., Y. Feng, and H. X. Liu. A Cooperative Driving Framework for Urban Arterials in Mixed Traffic Conditions. *Transportation Research Part C: Emerging Technologies*, Vol. 124, 2021, p. 102918.
16. Djavadian, S., R. Tu, B. Farooq, and M. Hatzopoulou. Multi-Objective Eco-Routing for Dynamic Control of Connected & Automated Vehicles. *Transportation Research Part D: Transport and Environment*, Vol. 87, 2020, p. 102513.https://doi.org/10.1016/j.trd.2020.102513.
17. Zeng, W., T. Miwa, and T. Morikawa. Eco-Routing Problem Considering Fuel Consumption and Probabilistic Travel Time Budget. *Transportation Research Part D: Transport and Environment*, Vol. 78, 2020, p. 102219. https://doi.org/10.1016/j.trd.2019.102219.
18. Tarko, A. P. A Unifying View on Traffic Conflicts and Their Connection with Crashes. *Accident Analysis & Prevention*, Vol. 158, 2021, p. 106187.
19. Tarko, A. *Measuring Road Safety with Surrogate Events*. Elsevier, Amsterdam, 2019.

The views presented in this paper are those of the authors alone.