

Review of the Innovation and Development of Collaborative Sensing Technology for Driverless Vehicles

Zhuoer Tang^{1,a,*}

¹*Arizona College of Technology at Hebei University of Technology, Tianjin, China*

a. tang66@arizona.edu

**corresponding author*

Abstract: Recent advances in autonomous driving technology have marked a significant shift in transportation systems, bringing self-driving capabilities closer to widespread adoption. This technology promises to reduce human intervention in driving tasks while enhancing transportation efficiency and convenience. Collaborative sensing technology is one of the hot research directions of unmanned driving technology. The birth of this technology can help us effectively solve the problems of occlusion and detection distance limitation in the process of vehicle perception. The first part of the paper introduces the current dilemma of unmanned driving technology and proposes collaborative perception as an effective means to solve the problem. Then "The advantages of collaborative perception technology" introduces the specific advantages of this technology. In the section of Methods of collaborative perception, it introduces the specific methods of collaborative perception technology from three perspectives: vehicle communication, data fusion and data sharing. In The challenge and innovation of collaborative perception, the difficulties faced by this technology and the means of innovation are discussed. This comprehensive analysis provides readers with fundamental insights into collaborative perception systems, establishing a foundation for future research and innovation in autonomous driving.

Keywords: Collaborative Perception, driverless technology, multi-agent.

1. Introduction

With the rapid development of artificial intelligence technology, people have started exploring its applications. One significant application is the integration of AI with vehicles, which led to the emergence of driverless technology. Many believe that driverless technology will free people to do other things. Additionally, handing over control to computers may reduce dangerous driving caused by human factors, such as emotional drivers. This will significantly improve the safety of transportation. Currently, autonomous driving technology still falls short of matching experienced drivers' capabilities in terms of quick reactions and risk anticipation. Similarly, there is also the problem of misidentification of objects in autonomous driving technology, which can lead to unexpected traffic accidents. In Collaborative Perception for Connected and Autonomous Driving: Challenges, Possible Solutions and Opportunities, the authors once mentioned that in a crash in California, a Tesla car mistakenly identified a concrete fence as a safe area, resulting in loss of life and property [1]. This incident highlights the critical importance of accurate object recognition in autonomous driving systems. Such incidents are not rare.

As a result, increasing attention has been paid to the safety concerns of autonomous driving technology. The Society of Automotive Engineers (SAE) has proposed six levels of autonomous driving. In short, levels 0 to 2 are still human drivers who need to respond to external situations, such as braking or accelerating to ensure safety. Automatic emergency braking and automatic cruise and other functions play a role in assisting driving. In contrast, levels 3 to 5 do not require the driver to actively drive the vehicle, level 3 requires the driver to take over in special circumstances, and level 5 is the most ideal form that does not require human participation at all [2]. At present, the mainstream level of autonomous driving on the market is still between level 2 and level 3. While this partially reduces the need for human drivers, the technology still requires human supervision due to its imperfections. Building upon the earlier discussion of safety concerns, it's worth noting that current autopilot technology, as demonstrated by the Tesla incident, still poses significant safety risks. In order to work towards Level 5 autonomous driving, this paper needs to further advance the research of autonomous driving technology. This paper examines collaborative sensing technology as a promising approach to advancing autonomous driving systems, with particular focus on its implementation, challenges, and potential benefits. This paper focuses on the reasons and advantages of collaborative perception technology in the first part, the specific methods of collaborative perception technology in the second part, the challenges and future prospects of collaborative perception technology in the third part, and summarizes the full text in the conclusion part.

2. Innovation and development of collaborative sensing technology for driverless vehicles

2.1. The advantages of collaborative perception technology

Self-driving cars consist of three basic subsystems, they are perception, planning, and control [3]. Among these, perception plays a crucial role. It is an important system for unmanned vehicles to obtain external information and judge their own position. Single-agent perception based on visual tasks have developed rapidly in recent years, such as 2D/3D object detection, semantic segmentation, BEV segmentation, and tracking [1]. In addition, the further combination of monomer perception and deep learning has also made significant progress and breakthroughs [4]. Although the monomer sensing technology has been relatively complete and mature, it has not been able to solve the two main problems, one is the occlusion induction problem in the induction process, and the other is the induction distance problem [1]. These two difficult problems are more because of the inability to make breakthroughs in hardware, which causes the limitations of monomer induction technology. Occlusion is a common phenomenon in the real world, and for a single sensing system, occlusion means that it cannot detect the situation of the occluded object, such as the occluded pedestrian, bicycle, etc. This can lead to a huge security risk. The limitation of sensing distance is the limitation of the detection distance of the vehicle sensor, such as the camera is limited by the field of view, and the LiDAR is limited by the maximum detection distance [1]. Objects beyond the detection range pose significant security risks to autonomous vehicles due to delayed identification. In order to make up for the limitations of individual perception process, a new way, collaborative perception, has been proposed.

This method uses the vehicle to network with other vehicles, infrastructure and other facilities to carry out collaborative perception to expand the field of vision and eliminate the benefits of blind spots. To put it simply, collaborative perception is a multiagent system in which agents communicate with each other to overcome the sensing limitations of individual agents [4]. Compared with single sensing, collaborative sensing has richer information sources, wider viewing angles and stronger adaptability, which can achieve higher sensing accuracy and distance [5]. In addition, the sensors required for collaborative sensing are cheaper than the traditional remote sensors used for monomer

sensing, and have very good economic benefits. With the increasing popularity of unmanned driving technology, collaborative sensing technology has broad application prospects [3].

2.2. Methods of collaborative perception

The initial technology of vehicle communication has two strategies: vehicle-to-vehicle (V2V) communication and vehicle-to-infrastructure (V2I) communication. For the practical application of V2V, this paper can examine the vehicle perception scenario shown in Figure 1. When a pedestrian walks through a zebra crossing to the middle of the road, due to the shelter of car 2, car 1 will not be able to detect the pedestrian, so it cannot brake in time to avoid a collision. This will bring great security risks to single-agent unmanned driving. If the collaborative perception strategy is adopted in this case, the sensor data of the car 3 can be shared with the car 1 to sense the presence of pedestrians in time, thereby avoiding collisions.

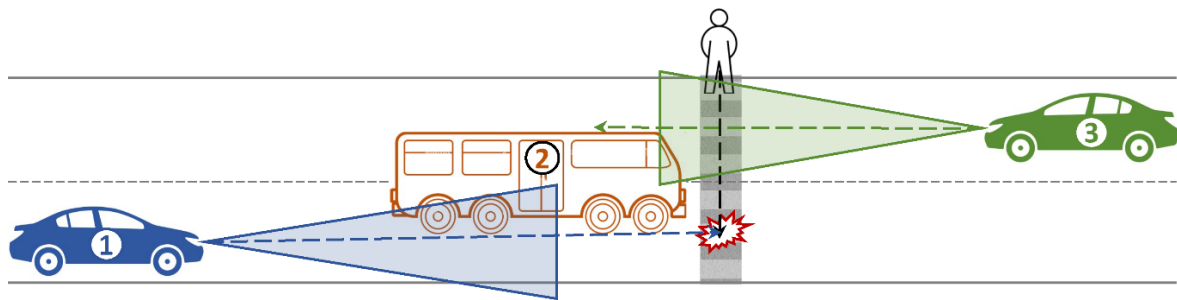


Figure 1: V2V application scenarios [2]

In the study of V2I, Zhengwei et al. put forward a collaborative perception system [6]. Specifically, there are four stages in the collaborative perception process. The first is the information collecting phase, which aims to use traditional sensors to provide the presence of an object in a specific location, uploading the collected data to a server for further processing. Then there is the edge processing stage, which is used to process the raw data, the main steps are preprocessing, object perception and storage. After that comes the cloud processing phase, which is designed to synchronize information collected in different locations and at different times. Finally, there is the message distribution phase, where the processed information can be distributed to roadside infrastructure such as signals for adjustment, or sent to vehicles for further action. Through these steps, this paper can initially complete the collaborative perception of the vehicle and infrastructure.

The advantages of V2I are illustrated in Figure 2. The smart tower located on the road can detect the green oval area in the diagram. For the first car, its own detection range is limited, and the intelligent tower can expand the detection range, so as to collect more road information and make timely deceleration actions. From the driving condition of the third car, this paper can find that the intelligent tower can identify the dangerous situation on the road in advance than the third car itself, and through interactive data, the third car can avoid risks in a timelier manner. For Car 5, the presence of the smart tower allows it to know the road construction situation in advance, so that it can plan the best route in advance. In short, V2I can reduce the safety risks of driverless vehicles and improve the efficiency of road traffic.

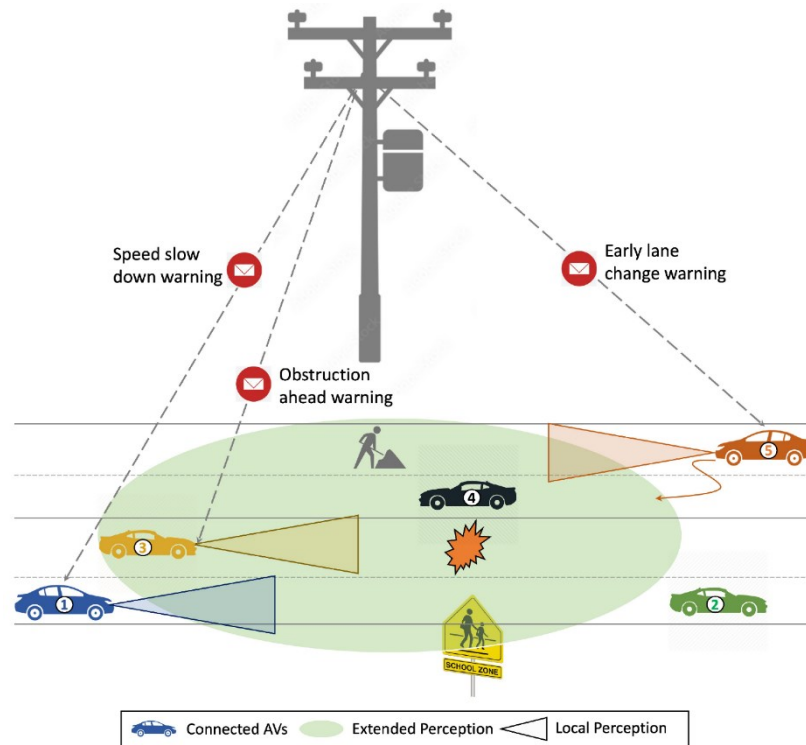


Figure 2: V2I application scenarios [3]

With the deepening of the research in the field of Internet of Things, the concept of car networking has also emerged. In other words, this is the concept of vehicle-to-everything (V2X) [1]. One of the most effective ways to implement V2X is dedicated short-range communications (DSRC) [7]. This is an advanced wireless communication technology, through which vehicles can quickly exchange position information and sense their own situation. However, the technology is still not complete. Despite DSRC's ability to exchange information over a range of 1000 meters, its data exchange rate remains low, typically only 2-6 Mbps [7]. In the future, the emergence of 6G will bring ultra-low latency, high data transfer rates and reliable connectivity to V2X [1]. This can effectively solve the problem of communication delay.

Collaborative driving technology is a technology that coordinates multiple sensors for analysis, and certain breakthroughs have been made in the research of data fusion of each sensor. At this stage, there are three main types of sensor data fusion: image fusion, point cloud fusion, and image–point cloud fusion [5]. For image fusion, it started early and the technology is relatively perfect. It has the advantages of high resolution and good classification effect. On top of that, vision sensors for image fusion are cheap and it is more economical to use such sensors. However, the information collected through the camera will have a large capacity, and the data transmission process depends on high network speed, which will cause delays. In order to solve this problem, the traditional image compression technology and deep learning can effectively reduce the size of image information [5]. These methods can ensure that the timeliness of perceived information is improved without losing important information. As for point cloud fusion, it uses point cloud data from different angles to sense the environment. Compared with image fusion, point cloud fusion has the advantages of a wider detection range and less susceptibility to light conditions [5]. This plays an important role in the process of vehicle positioning and ranging. Image–point cloud fusion is actually the fusion of the above two strategies, and it is the mainstream data fusion technology in the current collaborative

perception technology [5]. This method improves the accuracy of vehicle induction, but the lack of timeliness is still an important problem it will face.

Based on the stage at which data sharing and collaboration occur, collaborative awareness schemes can be categorized into early, intermediate, and late collaboration. Early collaboration is also called as data-level fusion. It fuses data at the network input [4]. The advantage of early collaboration is that the raw data shared by sensors is not processed in any way, which is more reliable, but this places high demands on the network [3]. In order to reduce the high bandwidth requirements of collaborative perception, vehicles are allowed to perform an analysis and prediction of the original data before transmission. From this was born the intermediate perception technology. It has been the most popular perceptual selection, but it has the problem of redundancy and loss of information, which has a high requirement for feature extraction strategy [4]. Unlike the above two methods, late collaboration performs predictive fusion at the network output stage, after individual vehicles have processed their own data [4]. In other words, late collaboration encourages information fusion after each vehicle produces a perceived result. This method can greatly save bandwidth, but it can cause serious errors and noise to the results [3]. In short, there is a trade-off between the timeliness and accuracy of information in these ways.

2.3. The challenge and innovation of collaborative perception

In the process of collaborative perception, each vehicle or facility may have different sensing moments. As a result, shared information might reach the autonomous vehicle asynchronously. The main idea to solve this kind of problem is to compensate the time delay of sensing or communication [1]. To address this issue, one key approach is to ensure temporal synchronization. That is, the timestamps should be consistent across sensors. In addition, there are model differences in collaborative perception. These differences stem from variations in sensor training, algorithms, or models [3]. This requires the design of uniform parameters and framework to ensure the normal operation of multi-agent communication. In the process of collaborative perception, the sharing of data between sensors will also lead to serious privacy leakage problems, and the protection of privacy needs to be deeply studied and discussed. The current solution to this problem is to encrypt the shared raw data using security maintenance techniques, such as homomorphic encryption or secure multi-party computation [4]. However, the proposal still has limitations and needs further discussion. In addition to these privacy concerns, sharing information can also be malicious, which comes from attacks on the network. Tu et al. evaluated the offensive and defensive performance of V2VNet [8]. If the attack model is known, implementing adversarial training can effectively enhance the system's defense against such attacks. The more intelligent there are, the more defensive there will be.

In terms of further innovation in Collaborative Perception, Zhu et al. propose the innovative Collaborative Occlusion-free Perception (COFP) framework, which allows autonomous vehicles to solve occlusion problems without relying on infrastructure to obtain information about objects being obscured [9]. Thus driving decisions can be made correctly. In addition, the Vehicle-Road-Cloud Collaborative Perception Framework has also attracted much attention. Vehicle-Road-Cloud Integration System (VRCIS) is divided into three layers, including base layer, platform layer and application layer. In simple terms, starting from the collection of multi-source sensor data at the physical layer, including vehicle driving status and traffic signal data is uploaded to the edge cloud, and the physical layer is controlled through calculation. The entire process is overseen by central cloud [10]. In conclusion, collaborative perception is a rapidly evolving field in current research. The innovations developed on this basis have the potential to revolutionize autonomous driving, enhancing safety and efficiency in ways this paper is only beginning to explore. As time goes by, the technology of collaborative perception will become more and more perfect and mature.

3. Conclusion

In this paper, this paper introduces a key technology of driverless technology - collaborative perception technology. This technology effectively addresses two major challenges in vehicle perception: occlusion problems and detection distance limitations, thereby enhancing the overall safety and efficiency of autonomous driving systems. Vehicle communication primarily relies on V2V and V2I technologies, which serve as the fundamental and well-established methods in collaborative perception. V2X, which represents the research achievements of the Internet of Things, is a deeper and more advanced research on vehicle perception, with greater technical difficulty and improved information processing efficiency. The main methods for sensor data fusion include image fusion, point cloud fusion, and image-Point cloud Fusion. Image-point cloud fusion is the fusion of image fusion and point cloud fusion. It absorbs the advantages of both and is the mainstream method for sensor information fusion. The stages of data sharing and collaboration are mainly divided into early collaboration, intermediate collaboration, and late collaboration. These collaboration stages represent a trade-off between data accuracy and transmission efficiency: earlier stages require higher network speeds but provide more accurate information, while later stages sacrifice some accuracy for reduced transmission loads. The key to choosing a certain approach is to balance accuracy against timeliness of the data, and intermediate collaboration is currently the most popular approach. The main technical problems of collaborative perception are the timely data synchronization among multiple agents and possible network attacks. As research continues to advance new collaborative perception frameworks, particularly in areas of data synchronization and security, this technology will play an increasingly crucial role in achieving level 5 autonomous driving, which can provide a strong help to achieve level 5 autonomous driving.

References

- [1] Hu, S., Fang, Z., Deng, Y., Chen, X., & Fang, Y. (2024). Collaborative perception for connected and autonomous driving: Challenges, possible solutions and opportunities. *arXiv preprint arXiv:2401.01544*. <https://doi.org/10.48550/arXiv.2401.01544>
- [2] SAE International. (2021). SAE levels of driving automation refined for clarity and international audience. *SAE News*. <https://www.sae.org/blog/sae-j3016-update>
- [3] Malik, S., Khan, M. J., Khan, M. A., & Nasir, H. (2023). Collaborative perception-The missing piece in realizing fully autonomous driving. *Sensors*, 23(18), 7854. <https://doi.org/10.3390/s23187854>
- [4] Han, Y., Zhang, H., Li, H., & Wang, K. (2023). Collaborative perception in autonomous driving: Methods, datasets, and challenges. *IEEE Intelligent Transportation Systems Magazine*, 15(5), 131-151. <https://doi.org/10.1109/MITS.2023.3298534>
- [5] Cui, G., Zhang, W., Xiao, Y., & Zhang, X. (2022). Cooperative perception technology of autonomous driving in the internet of vehicles environment: A review. *Sensors*, 22(15), 5535. <https://doi.org/10.3390/s22155535>
- [6] Bai, Z., Wu, G., Qi, X., & Liu, J. (2022). Infrastructure-based object detection and tracking for cooperative driving automation: A survey. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)* (pp. 1-6). <https://doi.org/10.1109/IV51971.2022.9827461>
- [7] Yee, R., Chan, E., Cheng, B., & Gerdes, J. C. (2018). Collaborative perception for automated vehicles leveraging vehicle-to-vehicle communications. In *Proceedings of the IEEE Intelligent Vehicles Symposium (IV)* (pp. 1099-1106). <https://doi.org/10.1109/IVS.2018.8500388>
- [8] Wang, T. H., Manivasagam, S., Liang, M., Yang, B., Zeng, W., & Urtasun, R. (2020). V2VNet: Vehicle-to-vehicle communication for joint perception and prediction. In *Proceedings of the European Conference on Computer Vision* (pp. 605-621). https://doi.org/10.1007/978-3-030-58536-5_36
- [9] Xiao, Z., Shu, J., Jiang, H., & Liu, Y. (2024). Toward collaborative occlusion-free perception in connected autonomous vehicles. *IEEE Transactions on Mobile Computing*, 23(4), 4918-4929. <https://doi.org/10.1109/TMC.2023.3298643>
- [10] Gao, B., Liu, J., Zou, H., & Zhang, X. (2024). Vehicle-road-cloud collaborative perception framework and key technologies: A review. *IEEE Transactions on Intelligent Transportation Systems*, 25(1), 1-24. <https://doi.org/10.1109/TITS.2024.3459799>