

车路群智协同感知方法综述

洪东洋¹, 张慧¹, 张译匀¹, 王雨桐², 韩瑜珊¹, 曹原周汉¹, 李浥东¹

(1. 北京交通大学计算机科学与技术学院, 北京 100044;

2. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室, 北京 100190)

摘要: 单车智能系统在感知层面存在一定的局限性, 如超视距感知能力不足、缺乏全局环境感知视角、在极端场景下的感知稳定性受限等。为提升车辆感知能力, 车路群智协同感知技术应运而生。该技术通过多个智能体(如车辆、基础设施等)共享感知信息, 并对信息进行整合和分析, 获取更全面、准确的环境认知。聚焦于车路群智协同感知方法的最新进展: 首先, 介绍车路群智协同感知方法的基础技术支撑; 其次, 根据数据处理的交互方法, 将其划分为4类协同模式; 然后, 深入探讨了3种主要的协同感知方法以及常用的大规模协同感知数据集; 最后, 讨论了车路群智协同感知技术面临的挑战和未来研究展望。车路群智协同感知方法有助于弥补单一智能体感知的不足, 实现对环境的全面感知。

关键词: 车路协同; 协同感知; 信息融合; 通信技术; 自动驾驶

中图分类号: TP39

文献标志码: A

doi: 10.11959/j.issn.2096-6652.202528

A review of collaborative intelligent perception methods for vehicle-road systems

HONG Dongyang¹, ZHANG Hui¹, ZHANG Yiyun¹, WANG Yutong², HAN Yushan¹,

CAO Yuanzhouhan¹, LI Yidong¹

1. School of Computer Science & Technology, Beijing Jiaotong University, Beijing 100044, China

2. The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

Abstract: The intelligent system of a single vehicle has certain perception limitations, such as the inability to perceive beyond visual range, the lack of global perception capabilities, and restricted perception capabilities in extreme situations. To enhance vehicle perception capabilities, vehicle-road collaborative intelligent perception technology has emerged. This technology enables multiple intelligent agents (such as vehicles, infrastructure, etc.) to share perceived information, integrate and analyze it, and thereby gain a more comprehensive and accurate understanding of the environment. This paper focuses on the latest developments in vehicle-road collaborative intelligent perception methods. Firstly, it introduces the fundamental technological support for vehicle-road collaborative intelligent perception methods. Secondly, it categorizes collaborative modes into four types based on data processing and interaction methods. Then, it delves into three primary collaborative perception methods and commonly used large-scale collaborative perception datasets. Finally, it discusses the challenges faced by vehicle-road collaborative intelligent perception technology and the outlook for future research. Vehicle-road collaborative intelligent perception methods help to compensate for the deficiencies of single-agent perception and achieve comprehensive environmental perception.

收稿日期: 2025-03-07; 修回日期: 2025-06-09

通信作者: 张慧, huizhang1@bjtu.edu.cn

基金项目: 中央高校基本科研业务费项目 (No.2024JBMC045); 国家自然科学基金青年基金项目 (No.62203040)

Foundation Items: Fundamental Research Funds for the Central Universities (No.2024JBMC045), Youth Fund of The National Natural Science Foundation of China (No.62203040)

Key words: vehicle-road collaborative, collaborative perception, information fusion, communication technology, autonomous driving

0 引言

随着智能交通系统的发展,车辆与道路基础设施之间的信息交互日益成为提升交通系统整体效能的关键^[1-6]。其中,感知模块作为关键组成部分,负责实时感知周边环境并提取关键导航信息。近年来,随着大规模训练数据的累积和深度学习算法的进步,感知性能已得到显著提升^[7-8]。然而,要实现更高级别乃至完全自动化的无人驾驶,仅依赖单车感知传感器难以满足需求。单车感知存在的局限性,如视野限制、模态缺失以及传感器数据稀疏等问题,都可能对感知性能产生不利影响。为了克服单车感知的局限性,协同感知技术成为重要的解决方案。

车路群智协同感知技术基于多智能体交互^[9],通过车辆之间以及车辆与基础设施之间的感知信息共享,实现了对环境的全面感知。随着通信基础设施的改善和V2X(vehicle-to-everything)通信技术的发展,车辆能够以可靠的方式交换信息,从而实现协同感知。研究表明,车路群智协同感知技术能够显著提升环境感知的准确性和交通系统的安全性^[10-11]。然而,协同感知领域面临着数据融合的高度复杂性、传感器校准与对准的精确性要求,以及多源信息融合时的不确定性等挑战。为了实现实时且稳健的协同感知,还需应对通信容量限制和噪声干扰等技术挑战,研究人员正积极探索合适的协同模式和方法以解决这些问题。

当前,车路群智协同感知技术正处于蓬勃发展的阶段,众多前沿技术不断涌现,为智能交通系统注入了新的活力。然而,关于这一领域的最新进展与成果尚未形成全面而系统的总结。已有的车路集群协同感知综述文章^[12-20]多从技术架构、应用场景、协同方式等角度论述:一方面,通过系统梳理车路、车车协同感知方法及协同控制手段,详细介绍相关融合感知技术;另一方面,结合常规道路、路口等典型应用场景,阐述协同功能与实际案例,并对现存问题及未来发展方向进行深入探讨。与上述工作相比,本文的贡献点如下。

(1) 本文从协同时机的角度出发,将车路集群协同感知方法划分为早期协同、中期协同、后期协

同和混合协同4种模式,侧重于分析感知信息在协同流程中的交互时间点;进一步从感知数据类型与融合方式的角度出发,将协同感知方法细分为基于图像融合、点云融合和多模态融合3类,重点探讨不同数据类型在车路协同中的融合策略。这两类划分方式相辅相成,协同模式决定了信息何时协同,而融合方式则决定了如何协同。不同协同模式下可采用不同的融合策略,具体实现受限于数据类型、通信条件与处理能力等因素。其中,基于点云融合的方法根据融合阶段,进一步分为数据级、特征级和决策级3类,分别在原始数据、特征提取和决策结果层面进行融合。

(2) 本文整理了近年来车路协同感知领域常用的仿真平台和开源数据集,并将这些数据集根据真实性分为虚拟数据集和真实数据集。本文补充了最新相关的研究数据集,以进一步提升数据集的全面性和完备性。

1 基础技术支撑

车路群智协同感知系统的实现依赖于一系列先进的基础技术支撑,这些技术共同为系统的高效、安全和智能运行提供了坚实保障。车路群智协同感知中涉及的传感器技术(如相机、激光雷达、多传感器融合等)通过精确捕捉和解析环境信息,使车辆能够实时了解周围情况;通信技术,如V2V(vehicle-to-vehicle)、V2I(vehicle-to-infrastructure)和V2X,则通过信息交换,确保车辆、基础设施和其他交通参与者之间的协同合作。本节将详细介绍协同感知技术包含的主要基础支撑技术。

1.1 传感器技术

(1) 相机

相机在协同感知系统中是一种关键的感知设备,通过捕捉道路和周围环境的图像,帮助车辆识别交通标志、车道线、行人等重要信息。相机主要分为单目相机和立体相机两类:单目相机提供高分辨率的图像,但无法直接测量距离;立体相机则通过比较不同视角下的图像计算物体距离,从而提供深度信息。相机的优势在于较低的成本和丰富的视觉信息,但其性能受光线和天气条件影响较大,因此常与其他传感器(如激光雷达)结合使用,以提

供更加全面和可靠的环境感知能力。

(2) 激光雷达

激光雷达感知设备通过发射激光束并测量反射时间来生成高精度的三维点云图，从而精确感知周围环境。机械式激光雷达通过旋转部件进行全方位扫描，提供360°的视野和高分辨率数据；固态激光雷达则采用无机械部件的设计，如光学相控阵或微机电系统（micro-electro-mechanical system, MEMS）技术，具有更高的可靠性和更小的体积。激光雷达能够在各种光照条件下工作，对距离和物体形状的测量极为准确，但其成本较高，且在恶劣天气条件下（如大雨或大雾）性能可能受到一定影响。

(3) 多传感器融合

多传感器融合是指将多种不同类型的传感器（如相机、激光雷达、毫米波雷达、全球定位系统等）获取的数据进行集成和处理，以提高环境感知的准确性、鲁棒性和全面性。通过融合不同传感器的信息，系统能够综合利用各传感器的优势，例如相机提供高分辨率的视觉信息、激光雷达提供精确的距离测量、毫米波雷达提供强大的穿透力，从而在复杂的道路和天气条件下，更可靠地识别和跟踪车辆、行人及其他障碍物，为车辆自主决策和安全驾驶提供关键支撑。

1.2 通信技术

(1) V2V通信技术

V2V通信技术，即车辆对车辆的通信技术，旨在提高驾驶安全性和交通效率。该技术通过专用短程通信（dedicated short-range communications, DSRC）技术，特别是采用5.9 GHz频段进行专项通信，实现车辆之间直接、实时的信息交换。V2V通信的覆盖范围可达300 m，允许车辆自动向附近的其他车辆发送自身的行驶信息，如速度、位置等，同时接收其他车辆的信息。V2V通信技术不需要依赖固定的基站，而是直接在车辆之间传播信息，这使其具有更高的灵活性和可靠性。同时，它使用专用短程通信，确保了通信的安全性和私密性。

(2) V2I通信技术

V2I通信技术，即车辆对基础设施的通信技术，通过无线通信技术，实现了车辆与道路基础设施之间的实时信息交互。V2I通信技术利用无线通信技术为车载智能交通运输系统设立了专门的通信频段，通常使用5.9 GHz的专用短程通信频段。这

使得车辆能够实时接收并解析来自交通信号灯、道路监控系统、路侧单元等基础设施的关键信息。V2I通信可以传递多种类型的信息，包括但不限于交通信号灯的状态、路面结冰警告、限速、交通拥堵、道路施工警示等。V2I通信技术具有实时性、准确性和可靠性的特点，能够确保车辆及时获取最新的道路信息，为驾驶者提供有力的支持。此外，V2I通信使用专用频段进行通信，因此具有较高的安全性和抗干扰能力。

(3) V2X通信技术

V2X通信技术即车辆与万物的通信技术，旨在实现车辆与周围环境、其他车辆、行人以及云端之间的实时通信。V2X通信技术基于无线通信技术，如Wi-Fi、蜂窝网络（如4G和5G）以及专用短程通信技术，实现了车辆与道路基础设施（V2I）、车辆与车辆（V2V）、车辆与行人（V2P）以及车辆与云端（V2N）的信息交换。V2X通信技术允许车辆发送和接收关于行驶速度、位置、意图等的实时信息，同时也包括交通信号灯状态、道路施工警示、行人位置等环境信息。这些信息对于提高交通安全性、缓解交通拥堵、实现自动驾驶等至关重要。

2 协同模式

车路群智协同感知系统通常包含4个主要阶段：数据预处理、基础网络、特征提取、感知输出。在感知过程中，车辆采集的原始数据首先通过编码器进行处理以提取关键特征，随后这些特征由解码器综合以生成最终的感知结果。基于在感知流程中数据共享及协同的不同阶段，协同感知的模式可大致划分为以下4类：早期协同、中期协同、后期协同和混合协同。

在实际部署中，选择哪种协同模式往往需在通信资源、计算资源和任务时效性等因素之间进行权衡。表1列举了4种协同模式在带宽消耗、计算负载、时延敏感性以及对时间同步精度要求等方面的比较。可以看出，早期协同依赖高带宽与高同步精度以实现高效融合，适用于通信条件良好且时延要求高的场景；而后期协同更适用于通信受限但计算资源丰富环境；混合协同则提供灵活的策略选择，适应动态变化的感知需求。

2.1 早期协同

早期协同主要发生在协同感知系统的数据预处理阶段，此模式主要依赖于原始数据融合策略，亦

表1 不同协同模式在关键指标上的对比

协同模式	带宽消耗	计算负载	时延敏感性	感知精度
早期协同	高	低	高	高
中期协同	中等	中等	中等	中等
后期协同	低	高	低	低
混合协同	动态可调	中等~高	可适配	取决于主路径

称数据层融合或低层融合,如图1(a)所示。在此模式下,智能体(即主体车辆)接收并处理来自其他智能体(如其他车辆和基础设施)的原始传感器数据。这些数据在车辆内部进行聚合,形成融合原始数据。早期协同策略能够充分利用原始数据信息,在处理多种传感模态数据时,对计算要求较低,内存预算较小。例如,文献[21]将多个相邻智能体的原始传感器数据共享给边缘服务器,服务器进而将智能体的各个视图合并成更完整、分辨率更高的视图,显著提高了车辆的感知质量。

然而,共享原始数据需占用大量通信带宽,容易造成网络拥堵,因此对数据传输带宽和实时计算能力有较高要求。

2.2 中期协同

中期协同主要在协同感知系统的中间特征层进行。在这一阶段,各智能体首先对采集的原始数据进行语义特征提取,然后将提取后的信息与来自其他智能体的感知数据进行特征融合,最后做出更准确的预测和决策,如图1(b)所示。中期协同使网络能够学习具有不同特征表示和不同深度的跨模态数据,同时通过将信息压缩至特征中,平衡了性能和带宽的需求。许多工作^[22-24]采用了中期协同进行特征共享。

然而,特征提取过程可能会导致原始数据中重

要信息的丢失以及产生不必要的信息冗余,这可能削弱决策的准确性,甚至带来计算资源的浪费。因此,选择合适的特征提取算法和融合策略至关重要,未来仍需要对此进行深入研究。

2.3 后期协同

后期协同主要在网络的输出阶段进行预测融合,如图1(c)所示。在此协同模式下,每个智能体分别训练自己的感知模型,并将输出结果共享给其他智能体^[25-28]。后期协同具有高度的灵活性和模块化,当引入新的传感方式时,只需训练其特定领域的网络,而不影响其他网络。Miller等^[29]采用后期协同策略开发了一套感知和定位系统,有效解决了两车通信链路的时延和丢包问题。文献[30]融合激光雷达和摄像头传感器分别独立生成的行人检测结果,提高了存在盲区、遮挡等情况下行人检测的准确率。文献[31]先用激光雷达检测目标,再将点云目标映射到图像上,实现图像信息与激光雷达信息的融合。

然而,后期协同高度依赖于各个智能体的输出数据质量,由于个别智能体的输出包含噪声而产生误差累积,整个系统的感知能力受到影响^[32]。此外,定位误差和时间不同步也可能影响合并结果的准确性。

2.4 混合协同

为了进一步提高多智能体协同的效率,研究者们通过结合两种或多种协同模式来优化整体协同感知的效果,即混合协同模式,如图1(d)所示。例如,文献[33]提出了中央融合系统,该系统在传感器具有高可见性的情况下采用高级信息融合(后期协同),而在可见性较差的情况下则采用低级信息融合(早期协同)。Li等^[34]提出了一种蒸馏协同

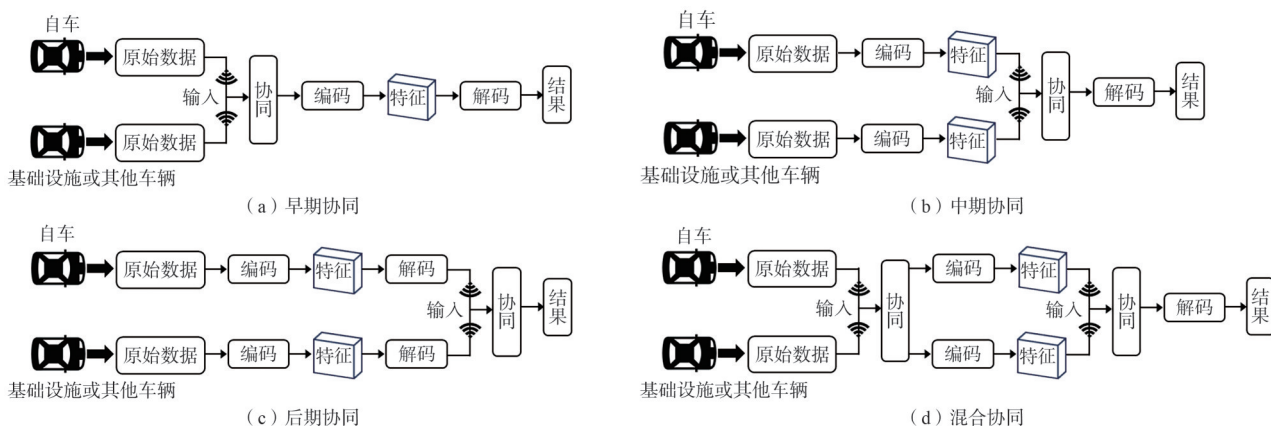


图1 不同协同模式对比

图的方法，该方法引入了基于知识蒸馏的教师-学生框架，教师模型利用多个智能体整体视角的输入进行早期协同，而学生模型则利用单智能体单视角的输入进行中期协同。在推理阶段，教师模型需要消耗大量的通信带宽，因此被舍弃；而学生模型因在训练阶段从教师模型学到了知识，在较低的通信带宽下仍能够保持优异的性能。

3 协同感知方法

协同感知方法通常采用视频摄像机、激光雷达、毫米波雷达等多种传感器设备进行环境感知，不同的传感器数据经过预处理或同一种传感器数据经过不同的预处理，都可视作不同的数据模态，融合这些数据并利用它们的互补特性成为当前的主流研究趋势。本节主要聚焦于协同感知信息融合方法^[35-97]，并从数据源和数据类型的角度划分成3类，聚焦不同传感器数据在协同感知中的作用和特点：基于图像融合的协同感知方法、基于点云融合的协同感知方法、基于多模态融合的协同感知方法。图像融合方法侧重于丰富的纹理与语义信息，点云融合方法强调三维空间的几何精度，而多模态融合方法结合多种数据源优势，在复杂场景下展现更强的鲁棒性。不同协同感知方法比较见表2。本文将方

法整理成时间线，如图2所示。图3展示了3类协同感知方法的比较，以中期协同为例。

3.1 基于图像融合的协同感知方法

图像数据因具有高分辨率、良好的目标语义表达能力与较低的成本，在协同感知中被广泛采用，是自动驾驶车辆与其他感知节点间常用的信息共享类型。然而，在通信带宽受限的场景下，要实现高效、可靠的图像感知信息共享仍面临诸多挑战。

面对通信带宽受限问题，研究者从通信内容压缩和通信策略优化两方面入手，提出了一系列创新方法来提高通信效率^[36,47]。在通信内容压缩方面，Liu等^[35]提出多阶段握手通信机制，其通过神经网络学习并压缩各阶段必需的信息，从而显著优化语义分割等场景理解任务的性能。此外，他们进一步提出了基于学习的通信组构建与通信时机决策框架^[36]，如图4所示。该框架在保持准确度的同时，显著降低了通信带宽的消耗。在通信策略优化方面，Hu等^[47]提出了一种协同式仅相机3D检测方法CoCa3D。该方法允许智能体通过通信彼此共享补充信息，并通过选择最具信息量的线索来优化通信效率。该方法通过来自多个视点的共享消息，消除了单智能体估计深度的歧义，同时补充了单智能体视图中的遮挡和远距离区域，显著提升了感知的准

表2 不同协同感知方法比较

方法	优点	缺点
图像融合 ^[35-47]	分辨率高,感知信息丰富,目标分类效果好,部署成本低	数据量大,受光照与恶劣天气影响较大
点云融合 ^[54-75]	数据级 信息完整,保留原始空间结构,利于高精度检测与融合 特征级 信息压缩后传输,平衡通信效率与融合精度,适用于多车协同 决策级 通信负担小,适用于资源受限场景,易于系统集成	通信带宽要求高,数据冗余度大,实时性差 特征提取依赖模型质量,不同设备间特征兼容性差 信息量少,融合精度有限,对个体感知结果依赖强
多模态融合 ^[83-97]	融合图像与点云优势,感知精度高,抗干扰能力强,适应性好	处理流程复杂,需解决跨模态对齐与融合问题,系统异构性大

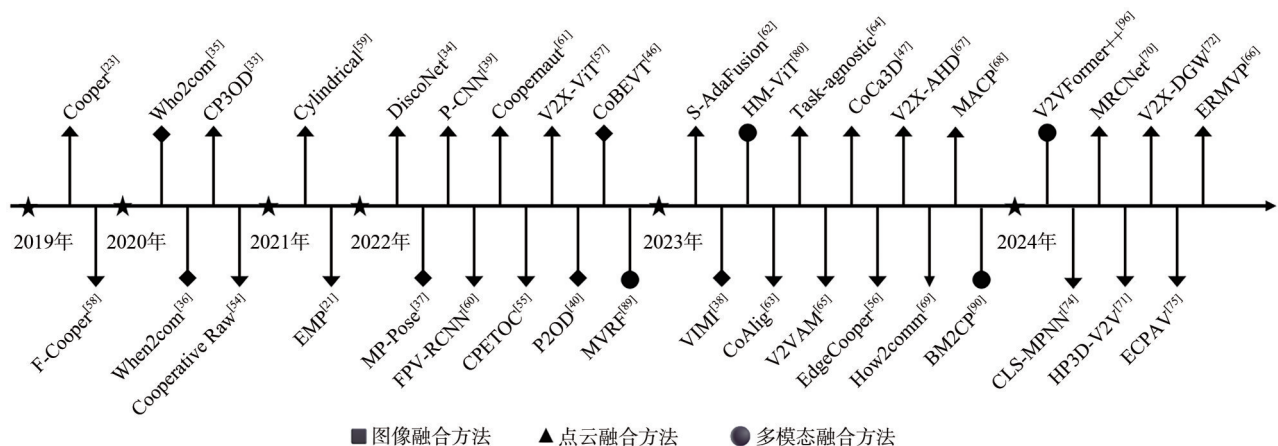


图2 协同感知方法时间线

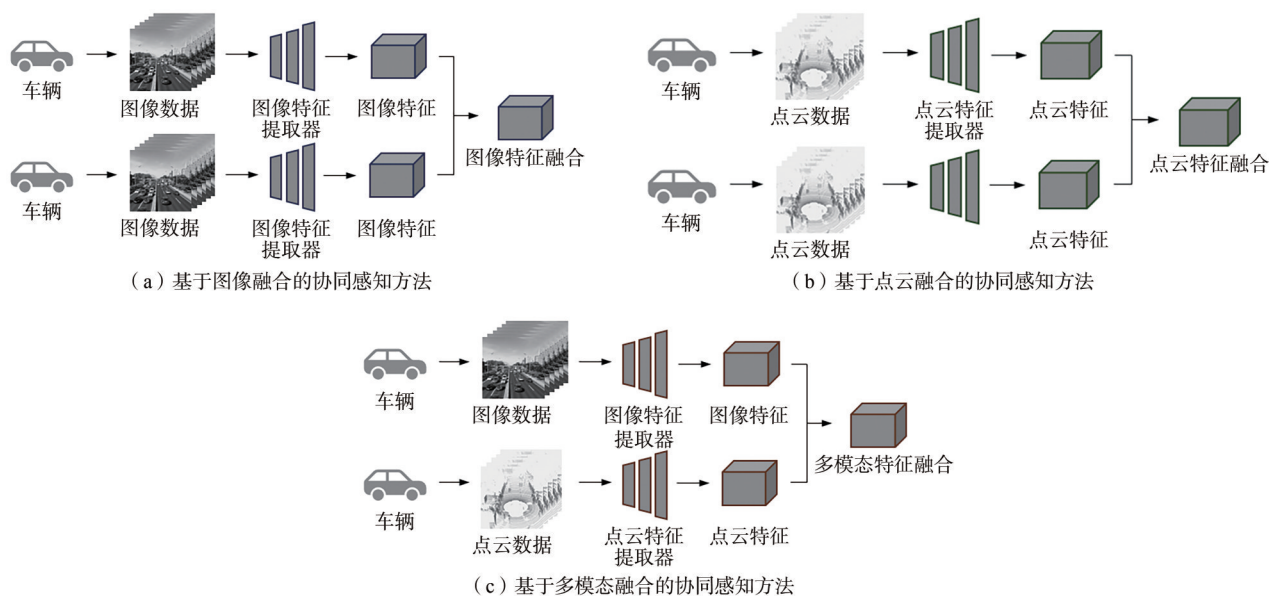
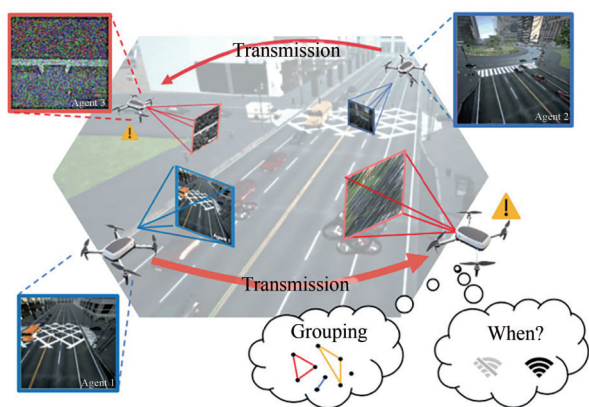


图3 不同协同感知方法对比

图4 多智能体协同感知示意图^[36]

确性和可靠性。Zhou等^[37]提出通用的图神经网络模型，高效解决了多视图视觉感知问题。Wang等^[38]则从视觉特征压缩角度出发，引入了多尺度交叉注意力模块和通道掩模机制，结合相机参数先验，校正多摄像头异步误差，并压缩冗余特征，有效降低了通信负荷。

针对多智能体协同感知中的隐私保护问题，研究者提出了多种方法来保障数据隐私，确保在不泄露敏感信息的情况下完成感知任务。Xiong等^[39]设计了轻量级且注重隐私保护的协作目标分类框架，允许自动驾驶车辆在交换原始传感器数据时保持隐私信息的安全性，确保了数据的隐私性和系统的可靠性。Bi等^[40]则进一步引入多个安全计算协议，构建出具备隐私保障的Faster R-CNN模型，能够在保证边界框与图像特征精度的前提下，防止敏感

数据泄露，为图像类感知任务提供了可落地的隐私保护机制。

针对感知视野和范围的局限性与盲点问题，研究者提出了从图像融合、BEV (bird's-eye view) 增强到深度估计优化的多种方法，力求提升协同感知的精度和可靠性。Sridhar等^[41]基于协同相对定位和地图融合技术^[42]，提出了一种短距离协同感知方法。该方法通过相邻车辆将图像感知信息共享给主车辆，在共享视野下实现了车辆的协同感知，从而有效拓宽了感知范围。Liu等^[43]则通过3D投影模型和特征点匹配技术估计几何变换参数，并设计了一种创新的深度仿射变换方法，实现了车辆间图像的融合。这种方法有效地解决了交通堵塞中的感知盲点问题，但在视角变化时可能产生尺寸偏差。为了解决视角变化带来的问题，Ma等^[44]提出了面向协同自动驾驶的多接入边缘计算辅助5G-V2X原型系统设计。在该系统中，路侧单元摄像机利用分布式数据融合方法，将不同区域检测到的环境信息进行交互感知融合，从而显著提升了车辆的感知范围和精度。Xiao等^[45]则通过融合基于车辆网络的多媒体信息，进一步扩大了车辆的感知范围并消除了盲区。他们采用深度学习方法从视觉系统中提取关键信息，并融合车辆网络中其他车辆的感知信息，生成鸟瞰图，极大地扩展了主车辆的感知范围。为了进一步解决视角变化带来的问题，Xu等^[46]提出了通用的多智能体多相机感知框架CoBEVT，该框架能够协同生成BEV地图预测。虽然基于图像的

3D检测相较于基于LiDAR的检测系统降低了成本,但在精确的深度估计方面仍有所不足。

然而,使用摄像机采集视频数据易受网络过载、传输时延等因素的影响,进而对协同感知的实时性产生负面影响。为了有效应对这一挑战,研究者们积极探索了多种策略,旨在减少传输数据的大小,确保协同感知的及时性和有效性。传统图像压缩技术^[48]通过高效的编码算法对视频数据进行压缩,以减少传输所需的带宽。而深度学习技术^[49-50]则通过构建复杂的神经网络模型,学习视频数据的内在特征,实现更高效的数据压缩。

针对协同感知系统中的实时性难题,尤其是在视频数据传输过程中,容易受到网络过载和传输时延的影响,研究者提出了多种压缩策略来减少传输负载,保证系统的实时响应能力。Löhdefink等^[51]利用具有对抗性损失函数的自动编码器进行有损图像压缩。他们使用二进制掩码控制感兴趣区域(region of interest, ROI),提高了压缩效率并保留了关键区域的图像质量;在传输过程中,仅传输经过量化处理的瓶颈数据,进一步减少了数据量。与传统方法不同,Lyu等^[52]提出了一种视频感知数据的传输负载优化方法。他们将视频帧中的静态背景与动态前景分离,仅在初始时传输静态背景,之后仅传输动态前景信息,从而显著降低传输负载,通过像素计算技术进行视频帧背景去除和降噪,快速提取视频中的动态信息;然后,利用生成对抗网络技术,快速地将静态背景和动态前景重新融合成视频帧,并基于视频帧反映的行车环境做出正确的驾驶决策。此外,Glaser等^[53]在多智能体语义分割的背景下,提出了端到端可学习的多智能体空间握手网络。该网络通过处理和压缩视觉信息,实现了多智能体之间的高效协同感知。MASH网络利用深度学习技术,学习视频数据的空间特征和语义信息,从而实现了对视频数据的高效压缩和传输。这种方法不仅减少了传输所需的数据量,还提高了协同感知的准确性和实时性。

3.2 基于点云融合的协同感知方法

激光雷达技术由于其高空间分辨率、广泛的探测范围以及对光照影响的低敏感性,在协同感知技术中扮演着关键角色。根据点云数据处理的不同阶段,点云融合技术可以细分为数据级融合、特征级融合以及决策级融合。该划分能够明确各类方法在处理流程中的信息交互粒度与计算复杂度差异。数

据级融合直接共享原始点云信息,信息最完整但带宽开销较大;特征级融合在压缩后的语义层面进行交换,兼顾效率与性能;决策级融合仅共享最终检测结果,通信负担最小但融合精度受限。每种融合都有其独特的应用场景和优势,共同构成了协同感知方法的重要组成部分。接下来,本节将逐一介绍数据级融合、特征级融合以及决策级融合在协同感知方法中的研究进展。

(1) 数据级融合

数据级融合方法的研究起步较早,因其能够直接处理原始传感器数据,避免了预处理阶段的复杂性,且融合精度较高。然而,这种方法需要处理大量的传感器数据,导致计算负担较重。数据级融合的关键优势在于,不同数据处理算法的差异不会影响车辆间共享数据的准确性,因此它适用于多车协同感知场景。

早期研究多聚焦于雷达点云数据的几何对齐与融合,着重提升检测精度和视野扩展。Chen等^[23]提出了一种在协作稀疏雷达点云数据上检测物体的方法。不同车辆的雷达数据经过坐标变换后,被映射到球面上以生成更密集表示。他们通过体素特征提取器和稀疏卷积中间层处理点云,提取特征;最终,利用区域提案网络进行对象检测,从而实现不同分辨率雷达数据的融合,改善检测精度。Ye等^[54]将多辆车的雷达数据进行融合,提出了一种状态估计框架,用于本地化和传感器数据融合,以扩展单个车辆的视野;通过融合差分全球定位系统数据和相互检测的测量,以及共同观测的静态环境注册,提高定位精度和对齐传感器数据的效果。这两项工作均依赖于高精度位姿信息,其优点是保留了点云的完整几何结构,但也因此带来了通信带宽压力大、实时性较差的问题。

为应对上述挑战,研究者尝试将数据级融合从纯车端转移至边缘节点。Zhang等^[21]提出了一种名为边缘辅助多车辆感知(edge-assisted multi-vehicle perception, EMP)的系统,通过边缘计算实现多车辆传感器数据共享。车辆将传感器数据上传至边缘计算节点,节点将各车辆的数据合并以创建全局视图。边缘节点运行定制的网联自动驾驶汽车算法,并将结果返回给车辆。通过边缘支持,车辆的工作负荷和网络带宽使用量大幅减少。EMP在进行数据级融合时,将各车辆上传的数据进行合并和处理,以提高检测覆盖率和准确性,从而实现协同

感知。相较于纯车端融合,边缘计算方案显著减少了每辆车的计算压力与网络流量,但也带来了边缘节点可用性和网络稳定性的新依赖。

随着研究深入,数据级融合逐步与特征抽象、图神经机制结合,形成跨层级融合范式。Li等^[34]提出了DiscoGraph的协作图,通过知识蒸馏框架,将早期协作中的整体视角输入作为指导,用于中期协作的单一视角输入优化。在早期协作中,各个智能体共享同一个教师模型,并将自己的视角数据输入该模型进行特征编码、解码和输出头的处理。通过知识蒸馏,教师模型的输出用于指导学生模型的特征抽象和聚合,从而提高数据级融合的效果。智能体间的协作通过实时可训练的、姿态感知的DiscoGraph动态实现。此方法兼顾了稠密数据级融合的精度与稀疏特征级融合的效率,提出了“层次融合”的思路。为进一步提升实时性与系统响应能力,后续研究将融合过程中的坐标变换、结构压缩与建模方式进行了系统优化。Chen等^[55]提出EMP算法,首先获取各个雷达设备的位置和旋转信息,并选择其中一个作为主设备,将其他设备的数据转化为主设备的坐标系。通过计算旋转矩阵和平移矩阵,将次级设备的点云数据转换并合并到主设备的点云数据中,从而生成综合的点云数据,用于进一步的3D目标检测。Luo等^[56]将点云数据转化为伪图像,并利用2D卷积架构来提高计算效率。参与者首先根据统一的柱状划分处理点云数据,并进行坐标转换以统一坐标系;然后,通过最大流理论的网络感知传输策略,将处理后的点云数据上传到边缘服务器,在服务器端进行数据聚合;最终将点云数据转换为包含9维张量的柱状数据,用于后续检测和增强感知范围。该方法以2D卷积网络替代计算量大的稀疏卷积,既提高了网络带宽利用率,又兼顾了推理速度,适用于对实时性要求更高的场景。

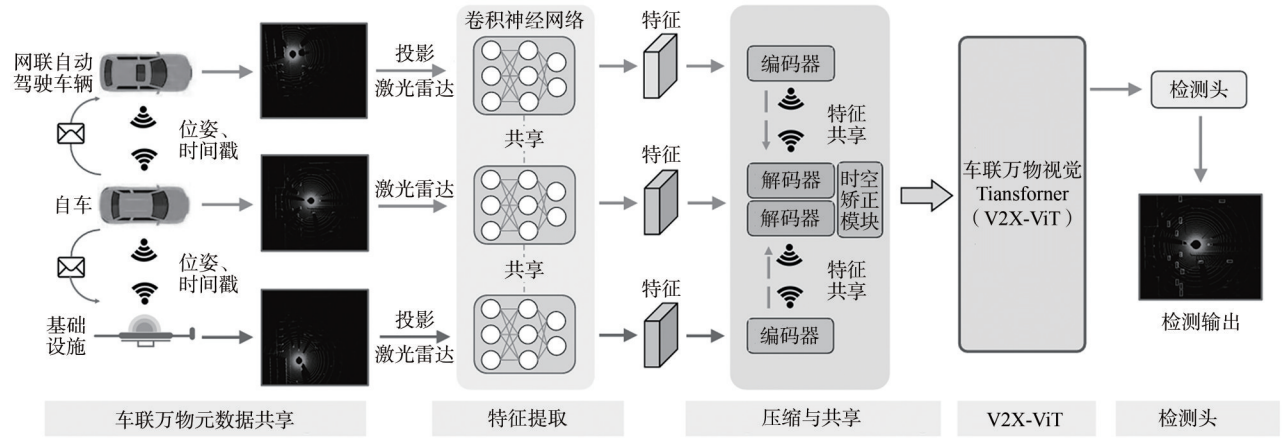
(2) 特征层级融合

特征级融合是利用多模态传感器从原始数据中提取特征信息,并通过各种算法和模型将这些特征进行有效整合的过程。该方法能够充分利用不同传感器的优势,如摄像头的视觉信息和雷达的空间数据,通过精确的特征融合提高目标检测、跟踪和定位的精度和鲁棒性。

早期特征级融合的研究主要集中于如何在统一空间中高效融合多源特征。为解决多车辆生成特征

表达不一致的问题,Chen等^[58]引入了F-Cooper框架,并提出了两种新的特征融合策略:体素特征融合和空间特征融合。体素特征融合策略直接融合体素特征编码层生成的特征图,而空间特征融合策略则在特征学习网络生成空间特征图后进行融合。这种方法不仅使不同雷达设备生成的特征更加有效地融合,还增强了特征的局部空间信息表达,显著提升了协同感知的检测精度和效率。与此同时,为进一步缓解点云稀疏性带来的特征质量问题,Zhu等^[59]提出了另一种解决方案。他们利用圆柱分区来生成更平衡的点云分布,并通过非对称3D卷积网络来处理水平和垂直方向的权重,以匹配驾驶场景中的目标点分布。这种方法增强了系统对稀疏性的鲁棒性,并在分割任务中引入了点级模块来优化输出。Yuan等^[60]则进一步提出通过关键点选择与多尺度特征聚合机制减少通信开销,同时实现精度提升。随后,Xu等^[57]针对车-路协同场景中异构感知节点的特征对齐问题,提出了V2X-ViT方法,如图5所示。该框架在编码阶段引入多尺度窗口注意力机制以提升特征表达能力,并设计了异构自注意力模块与位置时延感知机制,有效缓解了因类型差异与同步误差引发的特征错位问题。

随着协同感知应用向实时性场景拓展,如何降低通信负载、提高融合效率成为研究热点。Cui等^[61]提出的Coopernaut模型则是一种基于V2V通信的协同感知方法。每个邻近的V2V车辆使用点编码器将其传感数据编码成紧凑的消息,并通过空间变换将消息转换到自车坐标系中。自车接收这些消息后,通过体素最大池化来合并多车消息,生成聚合表示。最终,表示聚合器将自车与邻车的联合表示传递给控制模块,生成驾驶决策。这种方法显著增强了自车在复杂情境下的感知能力。与F-Cooper框架不同,Coopernaut更侧重于基于V2V通信的消息传递和特征聚合,增强了自车在复杂情境下的感知能力。针对视频连续帧冗余信息问题,Li等^[64]利用时空自动编码器,对当前帧中的随机子采样补丁和缓存的过去帧补丁进行编码,生成时空混合的特征图,只需传输当前帧的子采样特征,从而在减少通信带宽的同时,实现高效的场景重建和特征级融合。Ma等^[68]提出的网络则关注通信效率和特征融合的问题。该网络通过下采样卷积和激活层压缩特征编码器的输出,与周围车辆进行高效通信。每辆车接收来自周围车辆的特征,并通过一系列的上

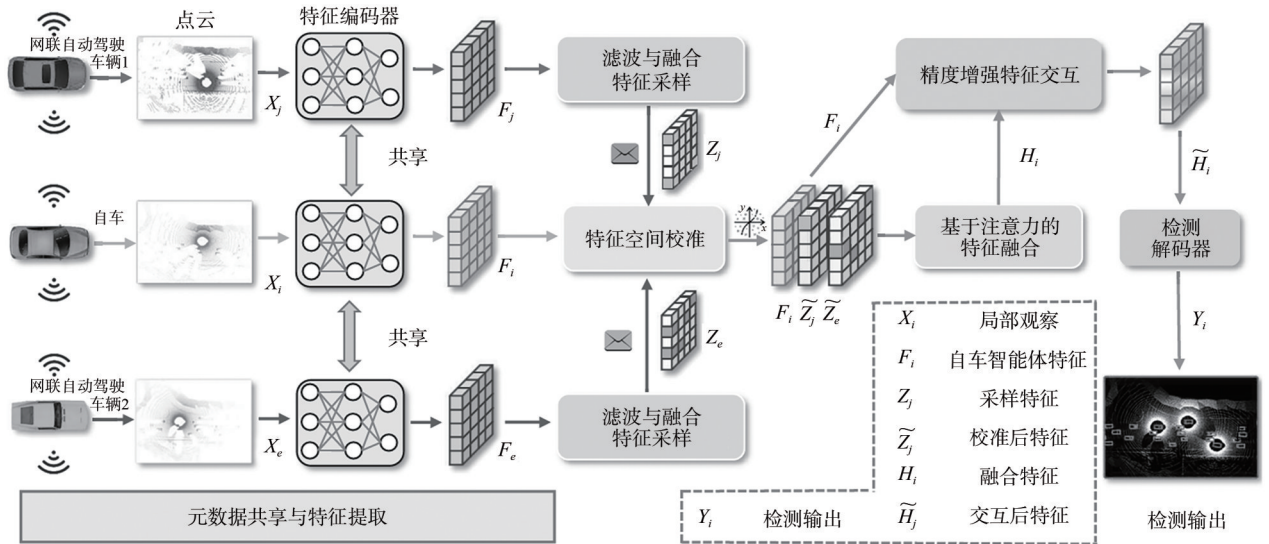
图5 V2X-ViT方法框架^[57]

采样卷积层解压信息，然后与上游特征编码器的编码特征进行融合，传递给下游的预测网络。这种方法在保持通信效率的同时，保证了特征融合的有效性。Zhang等^[66]进一步提出了通信高效且协作鲁棒的多车辆感知方法ERMVP，关注于降低通信开销和提高特征融合的效率，如图6所示。通过层次特征抽样策略提取代表性特征向量，ERMVP降低了通信开销，并通过稀疏一致性特征执行精确的空间位置校准，有效减轻了车辆定位误差的影响。此外，引入先进的特征融合与交互范式也整合了不同车辆和数据源之间的整体空间语义信息，促进了特征级的协同感知融合。

在实际应用中，传感器噪声与多车间非同步信息会严重影响协同感知性能。Qiao等^[62]通过雷达网络将点云数据转换为伪图像，并利用2D金字塔

网络提取多尺度特征，这种方法有效地利用了图像处理成熟技术来处理点云数据，提高了特征提取的效率；同时，通过特征融合网络生成组合特征图，进一步增强了特征的表示能力。Lu等^[63]通过智能体-物体姿态图优化，实现多个智能体之间及其检测到的物体之间相对姿态关系的对齐，促进姿态一致性。这种方法不需要准确的姿态监督，具有很强的泛化能力以适应各种姿态误差水平。此外，其采用多尺度中间融合策略，有效地聚合多个空间尺度上的协作信息，以提高特征级融合的效果。

为应对通信不稳定造成的特征缺失问题，Li等^[65]提出损失通信感知恢复网络用于恢复由丢失通信引起的不完整共享特征，而车对车注意力模块则增强了自车与其他车辆之间的交互作用，包括自车的内部注意力和基于不确定性的车辆间注意力。

图6 ERMVP模型架构^[66]

该方法有效地缓解了丢失通信对协同感知性能的负面影响。此外,为应对移动目标运动模糊与序列不一致问题,Hong等^[70]提出的MRCNet利用多尺度特征选择、有效信息融合和运动感知特征提取等技术,实现了智能体间在复杂协同场景下的语义特征交换。通过多轮通信和历史序列中提取的运动上下文信息,MRCNet能够减少噪声影响,并降低移动物体运动模糊的影响,从而提升协同感知的性能。此外,Chen等^[71]通过基于体素网格的统计滤波器有效去噪点云数据,并结合体素和PointPillars方法进行特征融合和BEV特征编码。这种方法解决了PointPillars方法中的空间特征交互问题,并增强了提取特征的语义信息;同时,通过最大池化技术降低维度并生成伪图像,减少了复杂的3D卷积计算。此外,特征级跨车辆特征融合模块也促进了有效的特征融合。

为提升系统在不同场景下的适应性,研究者也探索了特征级知识迁移与未来状态预测机制。He等^[67]提出了一种知识蒸馏网络。在训练阶段,教师模型使用多视角点云数据进行融合特征提取,而学生模型则学习教师模型的特征提取范式,仅使用单视角数据。这种方法使得在测试阶段,学生模型即使只接受单视角数据,也能获得接近教师模型多视角数据的表现,有效解决了单视角数据特征提取的问题。针对如何有效地利用不同车辆之间的信息问题,Yang等^[69]提出的How2comm框架通过共享编码器提取所有智能体的特征,并通过空间通道消息过滤机制最大限度地保留有益的语义信息,以减少通信带宽占用。同时,通过引入流引导的延迟补偿策略和时空协作Transformer模块,How2comm框架能够预测协作者的未來特征并进行时间对齐,实现更稳健的协同感知。此外,Li等^[72]针对恶劣天气下传感器性能下降问题,构建了自适应天气增强模型与柱状特征融合机制,在训练阶段模拟环境退化特性并优化特征表达,增强系统在复杂环境下的鲁棒性,进一步提高了在恶劣天气下的感知性能。

(3) 决策层级融合

决策级融合方法是指各智能体或传感器节点在本地完成目标检测等感知任务后,仅共享决策结果(如3D边界框、目标标签、置信度等),并在后端或中心节点通过一定策略进行融合。相比数据级和特征级融合,该方法通信开销低、处理速度快、系统部署灵活,特别适用于带宽受限或需保护原始数

据隐私的应用场景。然而,决策级融合无法从源头进行深层语义补偿,当前研究相对有限,且仍面临多源信息异步、目标冗余或遗漏等问题。

为减少数据传输量并简化协同结构,Arnold等^[33]利用每个传感器节点预处理点云,并运行检测模型生成3D边界框列表,将这些列表传输到中央融合系统。中央系统使用非极大值抑制算法消除重叠检测框,基于交并比度量重叠,去除置信度最低的检测框,最终将融合后的对象列表广播给附近的所有车辆,以提高检测精度和覆盖范围。考虑到多车间检测结果通常具有时间延迟,Dao等^[73]提出一种被称为“晚-早期协同”的协同感知框架,其中各连接智能体通过广播最接近当前时间的过去时间内检测到的对象(包含三维坐标、类别、置信度及预测速度),并基于点对点场景流计算的速度将这些对象从原始检测时刻的坐标系动态传播到当前时刻的全局坐标系,最终与自车当前收集的点云融合,以增强多车协同感知能力。这种方法通过将传感器检测结果与自车点云融合,解决了后期融合的复杂性,以提高检测精度和实用性。除了提高检测性能外,决策级融合还被用于辅助车辆自身定位。Barbieri等^[74]提出了一种名为CLS-MPNN的新型数据驱动协同定位和感知框架,用于复杂城市环境中的全球导航卫星系统(global navigation satellite system, GNSS)增强。该框架利用车辆上的3D物体检测器处理其本地雷达点云,识别定位周围静态物体。检测结果在集中单元移动边缘计算(mobile edge computing, MEC)上聚合,通过消息传递神经网络(message passing neural network, MPNN)进行数据关联,将多车提供的同一对象的测量数据协同组合。数据关联后,MEC使用协同贝叶斯跟踪算法提高对象的定位精度,并隐性地优化车辆定位。在多车辆冗余感知场景下,如何选择有价值的信息源也成为关键研究问题。Sarlak等^[75]提出自车根据前车的响应选择最佳的 M 辆协助车辆,以增强视觉范围并提高物体检测精度;在后期融合阶段,通过计算自车与协助车辆之间的最佳交并比(intersection-over-union, IoU),选择最匹配的检测结果进行融合,从而提高目标检测的准确性。

3.3 基于多模态融合的协同感知方法

在车路群智协同感知中,仅依赖单一模态的数据进行感知往往存在一定的局限性^[76-77]。这种单模态感知的缺陷主要源于数据获取和处理的片面性。

具体而言,摄像头作为视觉感知的核心组件,其捕获的数据主要聚焦于车辆前方视野的较低位置^[78]。在复杂多变的驾驶场景中,目标物体可能因遮挡而丢失关键信息,特别是在多层遮挡或物体位置较低的情况下,摄像头难以捕捉到完整的物体轮廓和细节。另外,激光雷达作为另一种重要的测距传感器,虽然具备独特的优势,但在检测不同距离的物体时,其分辨率会随距离的增加而逐渐降低^[79]。此外,激光雷达在极端天气条件下,如大雾和暴雨中,性能会受到严重影响,导致检测精度和稳定性显著下降^[76]。

为了克服单模态感知的局限性,研究者们提出了基于多模态融合的协同感知方法。这种方法通过整合不同模态的感知数据,如摄像头捕获的图像数据和激光雷达获取的测距数据,为自动驾驶系统提供更丰富和全面的环境信息。多模态融合不仅能够弥补单一模态在感知范围和精度上的不足,还能显著提升系统在复杂场景和极端天气条件下的鲁棒性和可靠性^[76,81-82]。为解决雾天等低能见度场景下摄像头感知退化问题,Jiang等^[83]提出了一种多模态融合方法,首先对摄像头在雾天拍摄的低能见度图像进行去雾处理,然后利用毫米波(MMW)雷达过滤有效目标,并将这些目标映射到摄像头图像上以获取相应的ROI。最后,结合摄像头视觉网络识别结果和雷达目标估计结果,通过加权方法得到最终的ROI检测结果。仿真结果表明,这种方法显著提高了检测精度,超过了单一传感器的检测效果。类似地,Ji等^[84]的目标分类系统也采用了多模态融合策略。该策略首先通过雷达引导选择少量候选图像供摄像头和学习方法分析;然后通过方向选择性特征检测器处理归一化的注意力窗口,为每个窗口生成稀疏表示;最后,使用多层原位学习网络(multilayer in-place learning network, MILN)区分不同对象的稀疏表示。该系统在多种分类任务上表现出更高的灵活性和准确性。类似的融合方法在其他研究中也得到了广泛探讨。Jha等^[85]通过将77 GHz毫米波雷达数据与摄像头数据相结合,检测和识别车辆视野中的障碍物,并利用雷达传感器提供相对距离信息。该系统在能见度低的环境中也能保持较高的可靠性和准确性。

此外,利用生成对抗网络的深度学习方法在摄像头和雷达数据融合方面也取得了显著进展^[86]。Wang等^[87]利用MMW雷达检测潜在车辆并提供

ROI,然后视觉处理模块在MMW雷达提供的ROI内通过对称性检测和主动轮廓检测来识别车辆;通过车辆跟踪技术,并结合MMW雷达和单目视觉生成的轨迹,进一步验证了检测和跟踪结果的准确性。然而,由于激光雷达的线束数量和探测距离存在差异,采集到的环境信息数据量也会有所不同。为了应对远距离点云数据稀疏的问题,Satio等^[88]提出了一种点云插值方法。该方法首先将LiDAR获取的点云数据校准并转换为像素坐标,然后将点云数据投影到下一帧的像素坐标上,并执行3D重建;通过组合上一帧的点云,实现点云的插值。此外,该方法还利用摄像头获取的RGB或HSV颜色信息来跟踪点云数据,并限制搜索区域以减少计算量。实验结果表明,该方法提高了形状恢复率和可辨识的距离,并且计算成本比基于模板匹配的传统跟踪方法要低。

大多数现有方法主要关注单一模态(特别是激光雷达),并没有充分利用多模态协同感知的优势。为了克服这一局限,Zhang等^[89]提出了一种用于协同感知的多模态虚拟-现实融合的Transformer方法。具体来说,为了从RGB图像和激光雷达点云中获取互补信息,他们提出了多模态虚拟-现实融合(multimodal virtual reality fusion, MVRF)方法,该方法生成虚拟点,并补偿稀疏位置的点信息不足。此外,他们构建了异构图注意力网络(heterogeneous graph attention network, HGAN),用于捕获智能体间的交互,并适应性地整合多个智能体的特征。HGAN包含一系列编码层,每层都包含异构智能体间注意力模块和多尺度自注意力模块,以学习基于各种智能体类型的不同关系,并同时捕获全局和局部空间注意力。Zhao等^[90]提出的多模态协同感知框架BM2CP,采用激光雷达引导的模态融合、协同深度生成和模态引导的中间融合,在不同智能体的模态之间获得深度交互。

Duan等^[91]提出了RGB-PVRCNN环境感知框架,该框架通过V2I(车对基础设施)通信技术显著增强了自动驾驶车辆在交叉路口的环境感知能力。其中,YOLOv3^[92]被用于高效分析摄像头捕获的图像,实现快速且准确的目标位置、速度和类别等关键信息的识别。同时,DBSCAN^[93]聚类算法则对雷达数据进行处理,通过聚类分析精确地确定目标的空间位置和运动特征。Fu等^[94]进一步提出了一种基于边缘计算技术的路边摄像头和毫米波雷

达融合方法。该方法首先利用YOLOv3对摄像头捕获的图像数据进行处理,同时用DBSCAN聚类算法对雷达数据进行处理;经过初步处理后,通过联合校准和直接更新方法,这两种传感器数据在空间和时间上实现了精准同步;随后,利用Munkres算法和卡尔曼滤波器对多个目标进行关联和跟踪,从而显著提升了传感系统的检测性能。尽管该方法在理论分析和模拟环境中展现出了良好的性能,但其在真实道路环境中的实际应用和验证仍需进一步深入。Wang等^[95]提出了一种新型的路边感知融合系统,该系统能够在真实道路环境中实现雷达和摄像机传感器的融合感知。该系统通过背景对象过滤算法有效过滤雷达检测结果中的背景对象,并使用一种自动校准方法,通过自动生成样本集来校准摄像头和雷达的检测结果,从而显著减少了手动校准的时间成本,提高了系统的准确性和实施效率。此外,该系统还显著降低了主动感知目标的时延,加速了协同感知技术的实际部署。

HM-ViT^[80]作为多车协同感知在多模态融合领域的初步探索,是首个统一的多智能体异构模态协同感知框架,如图7所示。该框架能够支持不同数量和类型的智能体,并在动态变化的车对车(V2V)环境中协同预测3D对象。它假设每辆车仅能获取到任意一种模态的数据,并基于此设定来实施不同车辆间的异构模态融合。为应对大规模车队长距离场景下的信息传递需求,Yin等^[96]进一步提出了V2VFormer++框架,该框架通过引入全局-局部变换器(global-local transformer)模块,有效实现了不同车辆间的信息聚合和特征增强,同时显著提升了网联自动驾驶车辆(connected and automated vehicle, CAV)在长距离和大规模场景下的感知能

力。Wang等^[97]对多模态协同感知任务中的特征融合策略进行了深入研究。他们探索了包括通道级拼接、元素级求和以及基于Transformer的融合方法在内的多种策略,并引入了基于注意力机制的融合方法。这种方法能够自适应地调整不同传感器数据的权重,并在不同传感器特征之间建立复杂的关联。在协同感知任务中,该方法展现出卓越的性能和强大的鲁棒性。

4 仿真平台与数据集

近年来,尽管协同驾驶技术发展迅速,但作为一个新兴领域,其研究仍面临诸多挑战。尤其是算法验证所需的现场实验,不仅成本高昂,还存在较大的安全风险。这类实验通常需要部署大量协同自动驾驶车辆和覆盖广泛区域的测试环境,同时还依赖于高精度数据标注,消耗大量人力与时间。为应对这些挑战,在仿真环境中开发和评估协同感知算法成为一种成本可控且安全可靠的替代方案。然而,由于当前缺乏统一和成熟的公共基准,在真实场景中验证算法效果仍比较困难。因此,本节将系统介绍当前在协同感知系统开发中广泛使用的主要仿真平台和开源数据集,为后续研究与评估提供支撑。

4.1 仿真工具

为了在真实世界数据集广泛可用之前推进多智能体协同感知的发展,研究者开发了许多仿真平台。以下将概述几种用于开发和分析协同驾驶系统的主要仿真工具的基本特征。

(1) SUMO

SUMO^[98]是传统的交通和驾驶行为仿真器,能处理大规模真实交通流。它结合每辆车的动态建

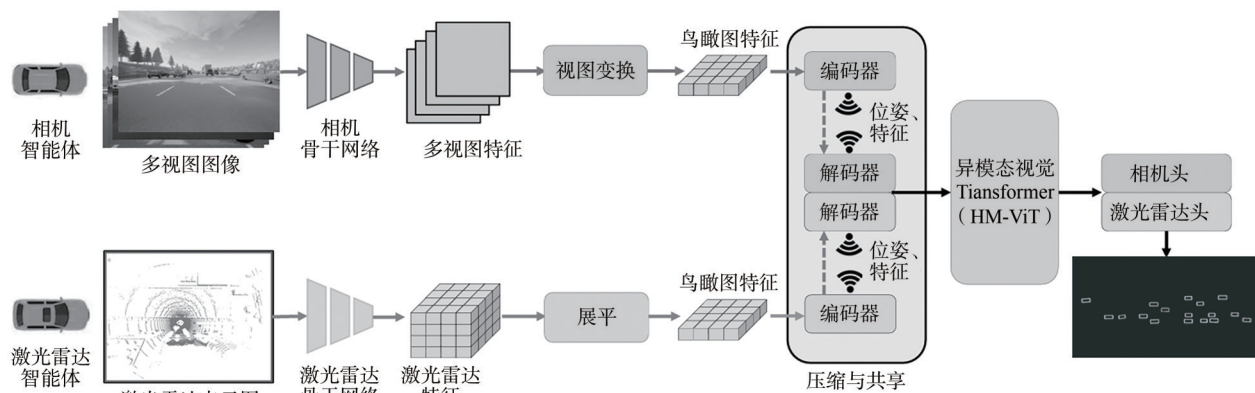


图7 HM-ViT方法框架^[80]

模,使用户能够通过交通控制接口 (TraCI) API 快速创建定制的交通场景。SUMO 非常适合集成自然驾驶轨迹数据,这些数据可以直接应用于周围环境来测试自动驾驶车辆。然而,它在提供具有极高保真度的高分辨率传感器数据方面存在限制。

(2) CARLA

CARLA^[99]是开创性的自动驾驶仿真器,旨在促进新型自动驾驶系统的开发、培训和验证。CARLA 的架构具有可扩展性,采用服务器-多客户端模式将计算分布到多个节点上。服务器不断更新环境物理特性,而用户通过 CARLA API 在客户端上进行控制。尽管 CARLA 允许扩展环境,但快速创建场景和管理大量交通变得复杂。CARLA 提供了交通管理模块来生成背景交通,但这些模块依赖于简化的行为规则,可能无法真实地再现实际驾驶行为。

(3) OpenCDA

OpenCDA^[100-101]是开放式协同仿真框架。它无缝集成了 CARLA 和 SUMO,用于逼真场景渲染、车辆建模和交通仿真。该框架支持在 CARLA+SUMO 协同仿真环境中评估协同驾驶自动化 (cooperative driving automation, CDA) 程度,并支持在仿真过程中不同级别和类别的 CAV 之间的协同。OpenCDA 还提供了全面的 CDA 研究管道、源代码和标准自动驾驶组件库,包括感知、定位、规划、控制和 V2X 通信模块。其主要优势是灵活可扩展性,OpenCDA 允许用户用定制设计的模块替换默认模块,以评估新功能对整体 CDA 性能的影响。

(4) V2XP-ASG

V2XP-ASG^[102]旨在通过生成复杂的现实场景来评估基于 LiDAR 的 V2X 感知模型。该仿真框架包括两个关键组件:对抗协同搜索 (adversarial collaborator search, ACS) 和对抗扰动搜索 (adversarial perturbation search, APS)。ACS 基于现有场景创建对抗协同图,识别并选择可能导致感知任务中出现次优结果的协同者。APS 使用针对感知任务的对抗目标量身定制的黑盒优化算法来扰动场景中车辆的位置。将 V2XP-ASG 中的两个阶段结合起来可以提高仿真场景的复杂性和难度。

(5) AUTOCASTSIM

AUTOCASTSIM^[61]是基于 CARLA 平台的自动驾驶仿真框架。AUTOCASTSIM 旨在评估协同驾驶模型,该平台提供了 3 个基准交通场景 (超车、

左转和闯红灯),这些场景由于视线受限对决策构成了挑战。AUTOCASTSIM 包含集成的网络仿真功能,便于定制 V2V 通信动态。此外,它还集成了高级驾驶模型,能够创建定制且逼真的交通场景。

(6) CMM

文献[103]强调了当前大多数交通仿真器的一个共同限制,即这些模拟器通常使用目标对象的固有属性,而不考虑 CDA 模型中的不完美感知。为解决这一问题,作者开发了基于开源自动驾驶仿真平台 CARLA (car learning to act) 的协同仿真平台——网络移动镜像 (cyber mobility mirror, CMM)。CMM 平台通过生成真实的感知数据来增强 CDA。将基于深度学习的感知管道集成到仿真系统中,以建立能够感知和重建 3D 仿真交通对象的 CMM 框架。这一改进使 CDA 模型能够利用 CMM 系统的感知信息,提高模型保真度,并支持更强大的系统验证。

(7) OpenCDA-ROS

现有的 CDA 仿真平台往往在仿真和真实世界场景之间表现出差异。为了全面解决这个问题,需要统一的工具和框架来支持在仿真真实世界环境中开发 CDA。OpenCDA-ROS^[104]是作为 OpenCDA 的扩展而开发的,专为机器人操作系统 (ROS) 量身定制。这种调整在仿真和真实世界的 CDA 之间架起了一座桥梁,确保算法和车辆在两种环境中的表现相似。通过将协同功能重新组织成更易于共享的结构,OpenCDA-ROS 有利于仿真和现实世界的原型设计与部署,促进了这两个领域在 CDA 开发中的无缝集成。

4.2 数据集

在协同感知领域,大型开放数据集对于推动技术进步和模型评估至关重要。尽管已经存在 KITTI^[77]、nuScenes^[105]和 Waymo^[106]等与自动驾驶感知相关的数据集,但这些数据集主要侧重于个体感知,对协同感知的需求尚显不足。鉴于此,本文整理并总结了目前主流的车路群智协同感知数据集。表 3 对近年来主流的协同感知数据集进行了系统性梳理与对比,涵盖了来自仿真环境 (如 CARLA、OpenCDA、SUMO) 与真实世界的多种协同感知场景,如 V2V、V2I、V2X、车路协同 (vehicle-infrastructure cooperation, VIC) 等,并详细列出了每个数据集的帧数、所用传感器类型 (如相机、激光雷达、毫米波雷达、IMU/GPS) 以及支持的感知任务 (包括物体检测、物体跟踪和运动预

测)。表4则展示了协同感知方法在公共数据集上的性能比较,对于同一方法的多个数据集结果,仅在首行标注其基本信息。性能评估采用了平均精度(AP)指标,AP@0.5和AP@0.7分别表示在IoU阈值为0.5和0.7时的平均精度。IoU阈值是用于判断预测边界框与真实标注框是否重合的标准,重合度越高,AP值越大,表示方法在目标检测上的表现越好。

4.2.1 虚拟数据集

(1) OPV2V

OPV2V^[107]是专为V2V(车对车)协同感知设计的模拟数据集,通过OpenCDA框架和CARLA模拟器收集。该数据集提供了完整的配置文件以重现实验环境,包含11 464帧的激光雷达点和RGB摄像头数据。OPV2V还提供了名为Culver City的真实模拟测试集,对于评估模型的泛化能力极具价值。该数据集支持三维物体检测和鸟瞰视图语义分割的基准测试,目前主要关注车辆这一类型的物体。

(2) V2X-Sim

V2X-Sim^[108]是模拟多智能体感知的数据集,由交通模拟软件SUMO和CARLA模拟器共同生成,数据格式遵循nuScenes^[105]标准。V2X-Sim配备了RGB摄像头、LiDAR(激光雷达)、GPS(全球定位系统)和IMU(惯性测量单元),共收集了100个场景,每个场景包含2~5辆车,总计10 000帧数据。数据集按8 000/1 000/1 000的比例划分为训练集/验证集/测试集。V2X-Sim的基准测试支持检测、追踪和分割3个关键感知任务,所有任务均采用BEV表示法,并在2D BEV中生成结果。

(3) V2XSet

V2XSet^[57]是大规模的开放模拟数据集,旨在用于V2X(车对外界)感知研究。其数据集格式与OPV2V^[107]类似,包含11 447帧数据。与V2X协同数据集V2X-Sim^[108]和V2I协同数据集DAIR-V2X^[109]相比,V2XSet涵盖了更多场景,并在基准测试中考虑了不完美的现实世界条件。该基准支持3D物体检测和BEV分割任务,并提供完美和噪声两种测试设置以供评估。

(4) CODD

CODD^[110]是基于CARLA模拟器生成的协同驾驶数据集,由108个片段组成。每个片段包含125个时间帧的LiDAR点云数据。每个时间帧对应一个数据切片,包括场景中所有车辆的LiDAR数据、传感器的绝对姿态信息,以及汽车和行人的3D边界框的地面实况注释。与OPV2V数据集^[107]相比,CODD专注于单一驾驶场景,而OPV2V则包含来自不同场景的联网车辆的相机图像和LiDAR点云数据。

4.2.2 真实数据集

(1) DAIR-V2X

DAIR-V2X^[109]是首个来自真实场景的大规模V2I协同感知数据集。其中,DAIR-V2X-C数据集特别适用于V2I协同研究,提供了VIC3D基准,以探索V2I物体检测任务。与主要基于激光雷达点的V2X-Sim^[108]和V2XSet^[57]不同,VIC3D物体检测基准的同时支持基于图像和基于激光雷达点的协同方法。

(2) V2X-Seq

V2X-Seq^[111]提供了真实世界场景中的顺序V2I

表3 协同感知相关数据集对比

数据集	年份	来源	场景	帧数	传感器			任务		
					相机	激光雷达	IMU/GPS	物体检测	物体跟踪	运动预测
OPV2V ^[107]	2022年	CARLA&OpenCDA	V2V	11 000	√	√	√	√	√	√
V2V-Sim ^[108]	2022年	CARLA&SUMO	V2X	51 000	√	√	×	√	√	×
V2XSet ^[57]	2022年	CARLA&OpenCDA	V2X	11 000	√	√	√	√	×	×
DAIR-V2X-C ^[109]	2022年	真实世界	V2I	39 000	√	√	×	√	×	×
CODD ^[110]	2022年	CARLA	V2V	135 000	×	√	×	×	×	×
V2X-Seq ^[111]	2023年	真实世界	V2I	15 000	√	√	√	√	×	√
V2V4Real ^[112]	2023年	真实世界	V2V	20 000	√	√	√	√	√	×
LUCOOP ^[113]	2023年	真实世界	V2V	54 000	×	√	√	√	√	×
HoloVIC ^[114]	2024年	真实世界	VIC	100 000	√	√	×	√	√	×
TUMTraf-V2X ^[115]	2024年	真实世界	V2X	2 000	√	√	√	√	√	×
R-Cooper ^[116]	2024年	真实世界	I2I	30 000	√	√	×	√	√	√

表4 协同感知方法在公共数据集上的性能比较

方法	年份	模态	阶段	数据集	AP@.5	AP@.7	代码开源
早期协同	2019年	激光雷达	早期	V2XSim	64.2	60.3	—
晚期协同	2020年	激光雷达	晚期	V2XSim	55.4	41.8	—
Who2com ^[35]	2020年	相机	中期	V2X-SIM	47.2	42.2	×
				V2X-SIM	47.9	42.9	
When2com ^[36]	2020年	相机	中期	OPV2V	69.12	53.76	√
				DAIR-V2X	51.88	37.05	
P-CNN ^[39]	2022年	相机	早期	—	—	—	×
MP-Pose ^[37]	2022年	相机	中期	—	—	—	×
CoBEVT ^[46]	2022年	相机	中期	OPV2V	88.3	79.7	√
MVRF ^[89]	2022年	相机, 激光雷达	中期	V2XSet	88.9	71.5	×
P2OD ^[40]	2022年	相机	早期	—	—	—	×
FPV-RCNN ^[60]	2022年	激光雷达	中期	OPV2V	—	82	√
V2X-ViT ^[57]	2022年	激光雷达	中期	V2XSET	—	71.2	√
CoCa3D ^[47]	2023年	相机	中期	DAIR-V2X	0.226	0.0985	√
VIM ^[38]	2023年	相机	中期	DAIRV2X	15.61	—	×
HM-ViT ^[80]	2023年	相机, 激光雷达	中期	OPV2V	0.695	0.515	×
BM2CP ^[90]	2023年	相机, 激光雷达	中期	OPV2V	83.72	63.17	√
				DAIRV2X	64.03	48.99	
S-AdaFusion ^[62]	2023年	激光雷达	中期	OPV2V	91.6	85.6	×
				CODD	86.2	83.9	
				V2X-Sim	0.832	0.726	
CoAlig ^[63]	2023年	激光雷达	中期	OPV2V	0.958	0.892	√
				DAIR-V2X	0.726	0.585	
Task-agnostic ^[64]	2023年	激光雷达	中期	V2X-Sim	65.09	60.26	√
V2VAM ^[65]	2023年	激光雷达	中期	OPV2V	0.926	0.861	√
				V2XSET	0.855	0.724	
V2X-AHD ^[67]	2023年	激光雷达	中期	DAIR-V2X	64.03	48.99	√
				V2V4Real	67.6	47.9	
MACP ^[68]	2023年	激光雷达	中期	OPV2V	93.7	90.3	√
				DAIR-V2X	62.36	47.18	
How2comm ^[69]	2023年	激光雷达	中期	V2XSET	84.05	67.01	√
				OPV2V	85.42	72.24	
V2VFormer++ ^[96]	2024年	相机, 激光雷达	中期	OPV2V	93.5	89.5	×
				V2XSim	85.33	69.82	
MRCNet ^[70]	2024年	激光雷达	中期	OPV2V	89.77	76.12	√
				V2XSET	85	66.31	
HP3D-V2V ^[71]	2024年	激光雷达	中期	OPV2V	67.4	56.5	×
				V2V4Real	65.01	40.45	
ERMVP ^[66]	2024年	激光雷达	中期	OPV2V	85	66.32	√
V2X-DGW ^[72]	2024年	激光雷达	中期	OPV2V	89.23	81.52	√

数据集, 用于开发和验证协同感知和运动预测算法。该数据集包含两个子数据集: 顺序感知数据集和轨迹预测数据集。顺序感知数据集涵盖了来自95个场景的超过15 000帧数据, 包括基础设施图像、点云、车辆端图像、3D检测/跟踪注释和矢量

地图等多种数据类型。轨迹预测数据集则包含21万个场景, 涉及28个路口的50 000个合作案例。

(3) V2V4Real

V2V4Real^[112]是首个大规模真实世界多模态V2V感知数据集。该数据集在俄亥俄州哥伦布市

使用特斯拉和福特车型收集,覆盖了410 km的道路。数据集包含20 000帧激光雷达数据以及超过240 000个三维边界框注释。此外,V2V4Real还为三维物体检测、物体追踪以及领域适应这3项协同感知任务提供了基准测试。

(4) LUCOOP

V2V4Real数据集只包含两种车型,在推进协同应用研究方面缺乏全面性。为了解决V2V4Real数据集在车型多样性方面的限制,LUCOOP^[113]提供了由3辆配备广泛传感器套件的CAV收集的时间同步多模态数据集。每辆车的传感器套件至少包括1个LiDAR传感器、1个GNSS天线和多达3个IMU。此外,每辆车上还安装了超宽带(ultra-wide band, UWB)传感器,并沿轨迹静态放置,以实现V2V和V2X的范围测量。该数据集涵盖了多车辆设置中的传感器测量数据,轨迹超过4 km,时间间隔超过26 min。它包含超过54 000帧LiDAR数据、大约700 000个IMU测量值、超过2.5 h的10 Hz GNSS原始测量值以及来自参考站的1 Hz数据。此外,数据集还提供了超过1 km的轨迹上6 000多个总站测量数据,以及1 874个V2V和267个V2X UWB测量结果。此外,它还提供了用于评估物体检测方法的3D包围盒标注和每辆车在测量活动期间的高精度实际位姿。

(5) HoloVIC

HoloVIC^[114]构建了首个大规模多视点多传感器全息交叉路口和车路协同的VIC数据集。该数据集共包含100 000+帧同步数据,构建了多个全息交叉路口,采用4种不同的传感器布局,包括摄像头、鱼眼和激光雷达,用于从交叉路口的所有传感器收集同步数据。此外,该数据集还制定了多个任务和基准,如单目3D检测、激光雷达3D检测、多对象跟踪、多传感器多对象跟踪、车路协同感知等。

(6) TUMTraf-V2X

TUMTraf-V2X^[115]提供了用于协作3D对象检测和跟踪的多模态多视图V2X数据集,其中包含来自5个路边传感器和4个车载传感器的2 000个标记点云和5 000个标记图像,以及带有轨迹ID的30 000个3D框以及精确的GPS和IMU数据。该数据集标记了8个类别,并涵盖了具有挑战性的驾驶操作的遮挡场景,例如交通违规、险些相撞事件、超车和掉头。

(7) R-Cooper

R-Cooper^[116]是第一个真实世界的大规模数据集,由30 000个手动注释的传感器数据组(图像和点云)组成,涵盖路口和走廊两个典型的交通场景。该数据集可以进一步支持多种协作感知任务,包括检测、跟踪、预测、定位、计数、监控等。

5 挑战及未来展望

车路群智协同感知技术在多智能体感知任务中扮演着重要角色,尤其在自动驾驶车辆领域具有显著应用价值^[117-119]。本节将详细探讨在实现车路群智协同感知过程中面临的通信宽带限制、姿态误差、异步信息共享等多方面的挑战,并对车路群智协同感知技术的未来发展趋势进行展望。

5.1 通信宽带限制

多智能体系统配备了多样化的传感器,如相机和激光雷达,这些传感器能够生成庞大的感测数据,包括点云数据和连续的图像序列。以KITTI数据集为例,其中的3D激光扫描仪每帧能生成约100 000个数据点。在数据集记录的最小场景中,累计帧数达到114帧,总点数超过1 000万。为了增强整体的感知能力,多智能体之间需通过通信共享互补信息,实现高效的协作感知。然而,这一过程不可避免地引入了感知性能与通信带宽之间的关键权衡问题。在实际应用场景中,通信系统往往受到资源限制,难以承受实时传输完整原始观测数据或海量特征信息带来的巨大通信负担。

已有的研究方案通过信息压缩等方法在感知性能和通信带宽之间取得了一定的平衡^[120-122]。Hu等^[120]提出了实用的压缩策略,通过聚焦于感知任务中的关键区域,在减少通信需求的同时提升感知性能。他们进一步设计了动态调整机制,根据通信带宽的变化实时调整通信中涉及的空间区域,从而有效应对实际通信环境中的带宽波动。但该方法存在局限性,信息压缩的方法在一定程度上会加剧高压比下的性能下降。

此外,事件触发控制逐渐成为缓解通信压力的重要研究方向之一。与传统的周期性通信机制相比,事件触发机制仅在系统状态或观测误差超过一定阈值时才进行通信,从而显著降低了通信频次,尤其适用于带宽受限的V2X环境。进一步地,已有研究将事件触发控制与深度强化学习方法结合,提升了系统在复杂交通环境中的自适应能力。例

如,文献[123]提出了一种基于事件触发的强化学习框架,用于在通信资源有限的场景中实现可靠、低时延的数据传输。文献[124]提出了在自动驾驶控制中结合平行控制与深度强化学习进行事件触发策略优化的应用案例。这些研究表明,事件触发机制不仅在降低通信负担方面具有显著优势,而且具备与智能决策算法协同设计的良好兼容性,为实现高效协同感知与控制提供了新的研究方向。

5.2 位姿误差补偿

车路群智协同感知需要精确共享它们的位置和姿态信息,从而促进更全面的感受野,利用多个智能体之间的信息交换来提高在存在遮挡等传感器缺陷的情况下物体检测的准确性,这要求各智能体的数据在统一的空间坐标系统中保持同步,为协同作业提供基础支撑。然而,在实际应用中,每个智能体的定位系统估算的六自由度位置和姿态往往存在一定的误差。这种误差可能导致在计算相对位置和姿态时出现偏差,进而对智能体间的协同效果产生显著影响。

为了减轻位姿误差补偿的负面影响,已有研究考虑了各种方法来提高位姿误差补偿鲁棒性^[60,125-126]。为了增强在复杂多变的位姿环境下的协作感知稳定性,Vadivelu等^[125]设计了位姿回归模块来学习位姿误差的校正。Yuan等^[60]使用3D点的语义标签来查找逐点对应关系。Gu等^[126]设计了位姿误差校正模块,以实现对特征层面未对齐部分的精确对齐,从而确保了感知的稳健性。然而,上述方法的一个关键限制在于它依赖于训练数据中包含的真实姿势信息。尽管训练数据中的位姿误差可以在离线阶段进行纠正,但这种标记过程不仅成本高昂,还可能因为各种原因(如人为错误或数据本身的不完善)而产生误差。

5.3 异步信息共享

在协同感知系统中,智能体之间需要进行信息共享。然而,由于数据量大和传输带宽的限制,信息从一个智能体传输到另一个智能体时会产生不可避免的时间差。这种时间差导致了信息时延,要求系统在接收并融合信息时需要考虑不同时间点的数据。换言之,在融合数据时,系统需整合来自不同时间点的信息,而非完全同步的数据。信息共享的延迟对协同感知系统的效能会产生显著影响,甚至在某些情况下,导致系统的性能低于仅依赖单个智能体的感知系统。

Xu等^[57]提出了一种整体注意力模型,旨在高效整合道路智能体(包括车辆及基础设施)之间的信息。该模型通过结合交替的异构多智能体自注意力机制与多尺度窗口自注意力层,不仅能够捕捉智能体间的动态交互,还能细致分析每个智能体的空间特征关系,从而有效缓解异步信息导致的不对称性问题。然而,该方法在面临大规模应用时可能面临可扩展性的挑战,同时,在数据的收集与共享过程中,也可能引发新的隐私保护问题和对抗性攻击的鲁棒性考验。

5.4 未来展望

针对车路群智协同感知在通信、定位与时序等方面的关键挑战,未来的研究可从以下4个方向进行更深入的探索与优化。

(1) 面向带宽限制的通信感知协同设计

随着感知系统的传感精度不断提升,感知数据的规模也迅速增长,进一步加剧了通信负载。未来可重点发展通信与感知协同优化机制,结合事件触发机制、压缩感知、注意力感知特征选取等策略,动态决定何时发送何种信息,以更好地权衡感知精度与通信成本。同时,深度强化学习与事件触发机制的结合也展现出良好的自适应潜力,应进一步研究其在车路感知资源分配与时延控制中的可拓展性与泛化能力。

(2) 面向位姿误差的自适应鲁棒对齐机制

现有姿态补偿方法大多依赖于离线标注数据,且对动态环境下的鲁棒性仍有限。未来应发展弱监督或无监督的位姿误差估计方法,结合多源冗余信息(如地图、语义标签、车辆轨迹)进行误差校正。此外,通过引入不确定性建模与贝叶斯推理,可在估计误差的同时评估感知结果的置信度,为协同决策提供更可靠的依据。

(3) 面向异步信息的时序一致性建模

考虑到通信时延与感知数据生成频率不一致,协同系统应具备时序容错能力与异步对齐机制。未来可引入时序对齐Transformer、空间-时间融合网络等模型,对不同时间点的数据进行统一建模与时延补偿。此外,异步融合过程中的鲁棒性与隐私保护问题也不容忽视,研究基于联邦学习或差分隐私的异步协同方法将是重要的发展方向。

(4) 从协同感知向智能决策闭环拓展

协同感知的最终目标是提升决策与控制的智能化水平。因此,未来应注重从感知-决策-控制的整

体闭环出发, 构建更强耦合的系统级优化机制。例如, 基于端到端协同学习架构, 同时考虑感知精度、通信成本和决策收益, 从整体上实现最优感知与智能响应。

6 结束语

本文深入讨论了车路群智协同感知的重要性及其实现机制。通过对早期协同、中期协同、后期协同及混合协同等多种协同感知模式的详尽分析, 本文阐明了这些模式如何有效突破单一感知系统的局限性, 有效提升感知系统的准确性。本文聚焦于基于图像融合、点云融合以及多模态融合技术的协同感知方法, 深入探讨了其在构建高效协同感知系统中的独特优势与潜在不足。同时, 本文关注了车路群智协同感知技术面临的挑战, 包括通信带宽限制、姿态校准误差、异步信息共享等问题。这些挑战对车路群智协同感知技术的进一步发展提出了更高的要求, 也为未来的研究指明了方向。最后, 本文对未来的车路群智协同感知技术的发展趋势进行了展望。

参考文献:

- [1] 刘泽禹, 张慧, 李滢东. 面向自动驾驶的感知决策一体化综述[J]. 智能科学与技术学报, 2025, 7(1): 4-20.
LIU Z Y, ZHANG H, LI Y D. A review of integrated perception and decision-making for autonomous driving[J]. Chinese Journal of Intelligent Science and Technology, 2025, 7(1): 4-20.
- [2] 娄云洁, 艾明飞, 庄术洁, 等. DK-Port: 基于大语言模型和强化学习的港口自动驾驶仿真环境构建与验证[J]. 智能科学与技术学报, 2025, 7(1): 98-113.
LOU Y J, AI M F, ZHUANG S J, et al. DK-Port: construction and validation of port autonomous driving simulation environment based on large language models and reinforcement learning[J]. Chinese Journal of Intelligent Science and Technology, 2025, 7(1): 98-113.
- [3] 沈甜雨, 李志伟, 范丽丽, 等. 具身智能驾驶: 概念、方法、现状与展望[J]. 智能科学与技术学报, 2024, 6(1): 17-32.
SHEN T Y, LI Z W, FAN L L, et al. Embodied intelligent driving: concept, methods, the state of the art and beyond[J]. Chinese Journal of Intelligent Science and Technology, 2024, 6(1): 17-32.
- [4] 黄峻, 田永林, 戴星原, 等. 基于深度学习的自动驾驶多模态轨迹预测方法: 现状及展望[J]. 智能科学与技术学报, 2023, 5(2): 180-199.
HUANG J, TIAN Y L, DAI X Y, et al. Deep learning-based multi-modal trajectory prediction methods for autonomous driving: state of the art and perspectives[J]. Chinese Journal of Intelligent Science and Technology, 2023, 5(2): 180-199.
- [5] LUO G Y, ZHANG H, YUAN Q, et al. Complementarity-enhanced and redundancy-minimized collaboration network for multi-agent perception[C]//Proceedings of the 30th ACM International Conference on Multimedia. New York: ACM, 2022: 3578-3586.
- [6] CHEN X H, DENG Y Q, DING H C, et al. Vehicle as a service (VaaS): leverage vehicles to build service networks and capabilities for smart cities[J]. IEEE Communications Surveys & Tutorials, 2024, 26(3): 2048-2081.
- [7] LI X, TIAN Y L, YE P J, et al. A novel scenarios engineering methodology for foundation models in metaverse[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2023, 53(4): 2148-2159.
- [8] ZHANG H, LUO G Y, LI J L, et al. C2FDA: coarse-to-fine domain adaptation for traffic object detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(8): 12633-12647.
- [9] DORRI A, KANHERE S S, JURDAK R. Multi-agent systems: a survey[J]. IEEE Access, 2018, 6: 28573-28593.
- [10] SCHIEGG F A, LLATSER I, BISCHOFF D, et al. Collective perception: a safety perspective[J]. Sensors, 2021, 21(1): 159.
- [11] SHAN M, NARULA K, WONG Y F, et al. Demonstrations of cooperative perception: safety and robustness in connected and automated vehicle operations[J]. Sensors, 2021, 21(1): 200.
- [12] 丁飞, 张楠, 李升波, 等. 智能网联车路云协同系统架构与关键技术研究综述[J]. 自动化学报, 2022, 48(12): 2863-2885.
DING F, ZHANG N, LI S B, et al. A survey of architecture and key technologies of intelligent connected vehicle-road-cloud cooperation system[J]. Acta Automatica Sinica, 2022, 48(12): 2863-2885.
- [13] 张毅, 裴华鑫, 姚丹亚. 车路协同环境下车辆群体协同决策研究综述[J]. 交通运输工程学报, 2022, 22(3): 1-18.
ZHANG Y, PEI H X, YAO D Y. Research review on cooperative decision-making for vehicle swarms in vehicle-infrastructure cooperative environment[J]. Journal of Traffic and Transportation Engineering, 2022, 22(3): 1-18.
- [14] 张新钰, 卢毅果, 高鑫, 等. 面向智能网联汽车的车路协同感知技术及发展趋势[J]. 自动化学报, 2025, 51(2): 233-248.
ZHANG X Y, LU Y G, GAO X, et al. Vehicle-road collaborative perception technology and development trend for intelligent connected vehicles[J]. Acta Automatica Sinica, 2025, 51(2): 233-248.
- [15] 杜泉成, 王晓, 李灵犀, 等. 行人轨迹预测方法关键问题研究: 现状及展望[J]. 智能科学与技术学报, 2023, 5(2): 143-162.
DU Q C, WANG X, LI L X, et al. Key problems and progress of pedestrian trajectory prediction methods: the state of the art and prospects[J]. Chinese Journal of Intelligent Science and Technology, 2023, 5(2): 143-162.
- [16] 刘腾, 王晓, 邢阳, 等. 基于数字四胞胎的平行驾驶系统及应用[J]. 智能科学与技术学报, 2019, 1(1): 40-51.
LIU T, WANG X, XING Y, et al. Research on digital quadruplets in cyber-physical-social space-based parallel driving[J]. Chinese Journal of Intelligent Science and Technology, 2019, 1(1): 40-51.
- [17] 黄哲, 王永才, 李德英. 3D目标检测方法研究综述[J]. 智能科学与技术学报, 2023, 5(1): 7-31.
HUANG Z, WANG Y C, LI D Y. A survey of 3D object detection algorithms[J]. Chinese Journal of Intelligent Science and Technology, 2023, 5(1): 7-31.
- [18] 王雨倩, 丁嵘. 基于群体智能成果的路径规划程序自动生成系统[J]. 智能科学与技术学报, 2022, 4(2): 255-263.
WANG Y Q, DING R. Automatic path planning program generation system based on swarm intelligence results[J]. Chinese Journal of Intelligent Science and Technology, 2022, 4(2): 255-263.
- [19] 鲁思迪, 何元恺, 施巍松. 车计算: 自动驾驶时代的新型计算范式[J]. 计算机研究与发展, 2025, 62(1): 2-21.
LU S D, HE Y K, SHI W S. Vehicle computing: an emerging computing paradigm for the autonomous driving era[J]. Journal of Computer Re-

- search and Development, 2025, 62(1): 2-21.
- [20] CUI G Z, ZHANG W L, XIAO Y Q, et al. Cooperative perception technology of autonomous driving in the Internet of vehicles environment: a review[J]. *Sensors*, 2022, 22(15): 5535.
 - [21] ZHANG X M, ZHANG A L, SUN J C, et al. EMP: edge-assisted multi-vehicle perception[C]//Proceedings of the 27th Annual International Conference on Mobile Computing and Networking. New York: ACM, 2021: 545-558.
 - [22] WANG T H, MANIVASAGAM S, LIANG M, et al. V₂VNet: vehicle-to-vehicle communication for joint perception and prediction[M]//Computer Vision-ECCV 2020. Cham: Springer, 2020: 605-621.
 - [23] CHEN Q, TANG S H, YANG Q, et al. Cooper: cooperative perception for connected autonomous vehicles based on 3D point clouds[C]//Proceedings of the 2019 IEEE 39th International Conference on Distributed Computing Systems. Piscataway: IEEE Press, 2019: 514-524.
 - [24] MARVASTI E E, RAFTARI A, MARVASTI A E, et al. Cooperative LIDAR object detection *via* feature sharing in deep networks[C]//Proceedings of the 2020 IEEE 92nd Vehicular Technology Conference. Piscataway: IEEE Press, 2020: 1-7.
 - [25] LIU H S, REN P F, JAIN S, et al. FusionEye: perception sharing for connected vehicles and its bandwidth-accuracy trade-offs[C]//Proceedings of the 2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking. Piscataway: IEEE Press, 2019: 1-9.
 - [26] HURL B, COHEN R, CZARNECKI K, et al. TruPercept: trust modelling for autonomous vehicle cooperative perception from synthetic data[C]//Proceedings of the 2020 IEEE Intelligent Vehicles Symposium. Piscataway: IEEE Press, 2020: 341-347.
 - [27] RAWASHDEH Z Y, WANG Z. Collaborative automated driving: a machine learning-based method to enhance the accuracy of shared information[C]//Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems. Piscataway: IEEE Press, 2018: 3961-3966.
 - [28] ZHANG Z J, WANG S, HONG Y C, et al. Distributed dynamic map fusion *via* federated learning for intelligent networked vehicles[C]//Proceedings of the 2021 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2021: 953-959.
 - [29] MILLER A, RIM K, CHOPRA P, et al. Cooperative perception and localization for cooperative driving[C]//Proceedings of the 2020 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2020: 1256-1262.
 - [30] WU T E, TSAI C C, GUO J N. LiDAR/camera sensor fusion technology for pedestrian detection[C]//Proceedings of the 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. Piscataway: IEEE Press, 2017: 1675-1678.
 - [31] ZHANG F H, CLARKE D, KNOLL A. Vehicle detection based on LiDAR and camera fusion[C]//Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems. Piscataway: IEEE Press, 2014: 1620-1625.
 - [32] AMBROSIN M, ALVAREZ I J, BUERKLE C, et al. Object-level perception sharing among connected vehicles[C]//Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference. Piscataway: IEEE Press, 2019: 1566-1573.
 - [33] ARNOLD E, DIANATI M, DE TEMPLE R, et al. Cooperative perception for 3D object detection in driving scenarios using infrastructure sensors[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(3): 1852-1864.
 - [34] LI Y M, REN S L, WU P X, et al. Learning distilled collaboration graph for multi-agent perception[EB]. arXiv preprint, 2021: arXiv: 2111.00643.
 - [35] LIU Y C, TIAN J J, MA C Y, et al. Who2com: collaborative perception *via* learnable handshake communication[C]//Proceedings of the 2020 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2020: 6876-6883.
 - [36] LIU Y C, TIAN J J, GLASER N, et al. When2com: multi-agent perception *via* communication graph grouping[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 4105-4114.
 - [37] ZHOU Y, XIAO J H, ZHOU Y, et al. Multi-robot collaborative perception with graph neural networks[J]. *IEEE Robotics and Automation Letters*, 2022, 7(2): 2289-2296.
 - [38] WANG Z, FAN S Q, HUO X L, et al. VIMI: vehicle-infrastructure multi-view intermediate fusion for camera-based 3D object detection[EB]. arXiv preprint, 2023: arXiv: 2303.10975.
 - [39] XIONG J B, BI R W, TIAN Y L, et al. Toward lightweight, privacy-preserving cooperative object classification for connected autonomous vehicles[J]. *IEEE Internet of Things Journal*, 2022, 9(4): 2787-2801.
 - [40] BI R W, XIONG J B, TIAN Y L, et al. Achieving lightweight and privacy-preserving object detection for connected autonomous vehicles[J]. *IEEE Internet of Things Journal*, 2023, 10(3): 2314-2329.
 - [41] SRIDHAR S, ESKANDARIAN A. Cooperative perception in autonomous ground vehicles using a mobile-robot testbed[J]. *IET Intelligent Transport Systems*, 2019, 13(10): 1545-1556.
 - [42] KIM S W, QIN B X, CHONG Z J, et al. Multivehicle cooperative driving using cooperative perception: design and experimental validation[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(2): 663-680.
 - [43] LIU W, MA Y, GAO M Q, et al. Cooperative visual augmentation algorithm of intelligent vehicle based on inter-vehicle image fusion[J]. *Applied Sciences*, 2021, 11(24): 11917.
 - [44] MA H S, LI S F, ZHANG E Q, et al. Cooperative autonomous driving oriented MEC-aided 5G-V2X: prototype system design, field tests and AI-based optimization tools[J]. *IEEE Access*, 2020, 8: 54288-54302.
 - [45] XIAO Z Y, MO Z B, JIANG K, et al. Multimedia fusion at semantic level in vehicle cooperative perception[C]//Proceedings of the 2018 IEEE International Conference on Multimedia & Expo Workshops. Piscataway: IEEE Press, 2018: 1-6.
 - [46] XU R S, TU Z Z, XIANG H, et al. CoBEVT: cooperative bird's eye view semantic segmentation with sparse transformers[EB]. arXiv preprint, 2022: arXiv: 2207.02202.
 - [47] HU Y, LU Y F, XU R S, et al. Collaboration helps camera overtake LiDAR in 3D detection[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 9243-9252.
 - [48] THEIS L, SHI W Z, CUNNINGHAM A, et al. Lossy image compression with compressive autoencoders[EB]. arXiv preprint, 2017: arXiv: 1703.00395.
 - [49] XU D, LU G, YANG R, et al. Learned image and video compression with deep neural networks[C]//Proceedings of the 2020 IEEE International Conference on Visual Communications and Image Processing. Piscataway: IEEE Press, 2020: 1-3.
 - [50] ROKACH L, MAIMON O, SHMUELI E. Machine Learning for Data Science Handbook[M]. Cham: Springer, 2023.
 - [51] LÖHDEFINK J, BÄR A, SCHMIDT N M, et al. Focussing learned image compression to semantic classes for V2X applications[C]//Proceedings of the 2020 IEEE Intelligent Vehicles Symposium. Piscataway: IEEE Press, 2020: 1641-1648.
 - [52] LYU P, LI K, XU J, et al. Cooperative sensing information transmis-

- sion load optimization for automated vehicles[J]. Chinese Journal of Computers, 2021, 44(10): 1984-1997.
- [53] GLASER N, LIU Y C, TIAN J J, et al. Overcoming obstructions via bandwidth-limited multi-agent spatial handshaking[C]//Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2021: 2406-2413.
- [54] YE E, SPIEGEL P, ALTHOFF M. Cooperative raw sensor data fusion for ground truth generation in autonomous driving[C]//Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems. Piscataway: IEEE Press, 2020: 1-7.
- [55] CHEN H L, LIU B, ZHANG X M, et al. A cooperative perception environment for traffic operations and control[EB]. arXiv preprint, 2022: arXiv: 2208.02792.
- [56] LUO G Y, SHAO C Z, CHENG N, et al. EdgeCooper: network-aware cooperative LiDAR perception for enhanced vehicular awareness[J]. IEEE Journal on Selected Areas in Communications, 2024, 42(1): 207-222.
- [57] XU R S, XIANG H, TU Z Z, et al. V2X-ViT: vehicle-to-everything cooperative perception with Vision transformer[C]//Computer Vision-ECCV 2022. Cham: Springer, 2022: 107-124.
- [58] CHEN Q, MA X, TANG S H, et al. F-cooper: feature based cooperative perception for autonomous vehicle edge computing system using 3D point clouds[C]//Proceedings of the 4th ACM/IEEE Symposium on Edge Computing. New York: ACM, 2019: 88-100.
- [59] ZHU X G, ZHOU H, WANG T, et al. Cylindrical and asymmetrical 3D convolution networks for LiDAR-based perception[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(10): 6807-6822.
- [60] YUAN Y S, CHENG H, SESTER M. Keypoints-based deep feature fusion for cooperative vehicle detection of autonomous driving[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 3054-3061.
- [61] CUI J X, QIU H, CHEN D, et al. Coopernaut: end-to-end driving with cooperative perception for networked vehicles[C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 17231-17241.
- [62] QIAO D H, ZULKERNINE F. Adaptive feature fusion for cooperative perception using LiDAR point clouds[C]//Proceedings of the 2023 IEEE/CVF Winter Conference on Applications of Computer Vision. Piscataway: IEEE Press, 2023: 1186-1195.
- [63] LU Y F, LI Q H, LIU B A, et al. Robust collaborative 3D object detection in presence of pose errors[C]//Proceedings of the 2023 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2023: 4812-4818.
- [64] LI Y, ZHANG J, MA D, et al. Multi-robot scene completion: Towards task-agnostic collaborative perception[C]//Proceedings of the Conference on Robot Learning. [S.l.:s.n.], 2023: 2062-2072.
- [65] LI J L, XU R S, LIU X Y, et al. Learning for vehicle-to-vehicle cooperative perception under lossy communication[J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(4): 2650-2660.
- [66] ZHANG J Y, YANG K, WANG Y L, et al. ERMVP: communication-efficient and collaboration-robust multi-vehicle perception in challenging environments[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 12575-12584.
- [67] HE C, WANG H, CHEN L, et al. V2x-ahd: vehicle-to-everything cooperation perception via asymmetric heterogeneous distillation network[Z]. 2023.
- [68] MA Y S, LU J W, CUI C, et al. MACP: efficient model adaptation for cooperative perception[C]//Proceedings of the 2024 IEEE/CVF Winter Conference on Applications of Computer Vision. Piscataway: IEEE Press, 2024: 3361-3370.
- [69] YANG D K, YANG K, WANG Y Z, et al. How2comm: communication-efficient and collaboration-pragmatic multi-agent perception[J]. Advances in Neural Information Processing Systems, 2024, 36: 25151-25164.
- [70] HONG S X, LIU Y, LI Z, et al. Multi-agent collaborative perception via motion-aware robust communication network[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 15301-15310.
- [71] CHEN H M, WANG H F, LIU Z L, et al. HP3D-V₂V: high-precision 3D object detection vehicle-to-vehicle cooperative perception algorithm[J]. Sensors, 2024, 24(7): 2170.
- [72] LI B L, LI J L, LIU X Y, et al. V2X-DGW: domain generalization for multi-agent perception under adverse weather conditions[EB]. arXiv preprint, 2024: arXiv: 2403.11371.
- [73] DAO M Q, BERRIO J S, FRÉMONT V, et al. Practical collaborative perception: a framework for asynchronous and multi-agent 3D object detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(9): 12163-12175.
- [74] BARBIERI L, TEDESCHINI B C, BRAMBILLA M, et al. Deep learning-based cooperative LiDAR sensing for improved vehicle positioning[J]. IEEE Transactions on Signal Processing, 2024, 72: 1666-1682.
- [75] SARLAK A, ALZORGAN H, BOROUJENI S P H, et al. Enhanced cooperative perception for autonomous vehicles using imperfect communication[C]//Proceedings of the 2024 20th International Conference on Distributed Computing in Smart Systems and the Internet of Things. Piscataway: IEEE Press, 2024: 700-707.
- [76] BIJELIC M, GRUBER T, MANNAN F, et al. Seeing through fog without seeing fog: deep multimodal sensor fusion in unseen adverse weather[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 11679-11689.
- [77] GEIGER A, LENZ P, URTASUN R. Are we ready for autonomous driving? The KITTI vision benchmark suite[C]//Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2012: 3354-3361.
- [78] YOO J H, KIM Y, KIM J, et al. 3D-CVF: generating joint camera and LiDAR features using cross-view spatial feature fusion for 3D object detection[C]//Computer Vision-ECCV 2020. Cham: Springer, 2020: 720-736.
- [79] XIE L, XU G D, CAI D, et al. X-view: non-egocentric multi-view 3D object detector[J]. IEEE Transactions on Image Processing, 2023, 32: 1488-1497.
- [80] XIANG H, XU R S, MA J Q. HM-ViT: hetero-modal vehicle-to-vehicle cooperative perception with vision transformer[C]//Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2023: 284-295.
- [81] VORA S, LANG A H, HELOU B, et al. PointPainting: sequential fusion for 3D object detection[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 4603-4611.
- [82] XIE L, XIANG C, YU Z X, et al. PI-RCNN: an efficient multi-sensor 3D object detector with point-based attentive cont-conv fusion module[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12460-12467.
- [83] JIANG Q Y, ZHANG L J, MENG D J. Target detection algorithm based on MMW radar and camera fusion[C]//Proceedings of the 2019 IEEE Intelligent Transportation Systems Conference. Piscataway: IEEE Press, 2019: 1-6.
- [84] JI Z P, PROKHOROV D. Radar-vision fusion for object classification[C]//

- Proceedings of the 2008 11th International Conference on Information Fusion. Piscataway: IEEE Press, 2008: 1-7.
- [85] JHA H, LODHI V, CHAKRAVARTY D. Object detection and identification using vision and radar data fusion system for ground-based navigation[C]//Proceedings of the 2019 6th International Conference on Signal Processing and Integrated Networks. Piscataway: IEEE Press, 2019: 590-593.
- [86] LEKIC V, BABIC Z. Automotive radar and camera fusion using Generative Adversarial Networks[J]. Computer Vision and Image Understanding, 2019, 184: 1-8.
- [87] WANG X, XU L H, SUN H B, et al. On-road vehicle detection and tracking using MMW radar and monovision fusion[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(7): 2075-2084.
- [88] SAITO M, SHEN S C, ITO T. Interpolation method for sparse point cloud at long distance using sensor fusion with LiDAR and camera[C]//Proceedings of the 2021 IEEE CPMT Symposium Japan. Piscataway: IEEE Press, 2021: 116-117.
- [89] ZHANG H, LUO G Y, CAO Y, et al. Multi-modal virtual-real fusion based transformer for collaborative perception[C]//Proceedings of the 2022 IEEE 13th International Symposium on Parallel Architectures, Algorithms and Programming. Piscataway: IEEE Press, 2022: 1-6.
- [90] ZHAO B Y, ZHANG W, ZOU Z N. BM2CP: efficient collaborative perception with LiDAR-camera modalities[EB]. arXiv preprint, 2023: arXiv: 2310.14702.
- [91] DUAN X T, JIANG H, TIAN D X, et al. V2I based environment perception for autonomous vehicles at intersections[J]. China Communications, 2021, 18(7): 1-12.
- [92] FAN J Q, HUO T J, LI X. A review of one-stage detection algorithms in autonomous driving[C]//Proceedings of the 2020 4th CAA International Conference on Vehicular Control and Intelligence. Piscataway: IEEE Press, 2020: 210-214.
- [93] ŞAHİN M Ş, ACARMAN T. An object segmentation approach for mobile lidar point clouds[C]//Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems. Piscataway: IEEE Press, 2021: 554-560.
- [94] FU Y J, TIAN D X, DUAN X T, et al. A camera-radar fusion method based on edge computing[C]//Proceedings of the 2020 IEEE International Conference on Edge Computing. Piscataway: IEEE Press, 2020: 9-14.
- [95] WANG L F, ZHANG Z Y, DI X, et al. A roadside camera-radar sensing fusion system for intelligent transportation[C]//Proceedings of the 2020 17th European Radar Conference. Piscataway: IEEE Press, 2021: 282-285.
- [96] YIN H B, TIAN D X, LIN C M, et al. V₂VFormer: multi-modal vehicle-to-vehicle cooperative perception via global-local transformer[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(2): 2153-2166.
- [97] 王秉路, 靳杨, 张磊, 等. 基于多传感器融合的协同感知方法[J]. 雷达学报, 2024, 13(1): 87-96.
- WANG B L, JIN Y, ZHANG L, et al. Collaborative perception method based on multisensor fusion[J]. Journal of Radars, 2024, 13(1): 87-96.
- [98] LOPEZ P A, BEHRISCH M, BIEKER-WALZ L, et al. Microscopic traffic simulation using SUMO[C]//Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems. Piscataway: IEEE Press, 2018: 2575-2582.
- [99] DOSOVITSKIY A, ROS G, CODEVILLA F, et al. Carla: an open urban driving simulator[C]//Proceedings of the Conference on Robot Learning. [S.l.:s.n.], 2017: 1-16.
- [100] XU R S, GUO Y, HAN X, et al. OpenCDA: an open cooperative driving automation framework integrated with co-simulation[C]//Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference. Piscataway: IEEE Press, 2021: 1155-1162.
- [101] XU R S, XIANG H, HAN X, et al. The OpenCDA open-source ecosystem for cooperative driving automation research[J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(4): 2698-2711.
- [102] XIANG H, XU R S, XIA X, et al. V2XP-ASG: generating adversarial scenes for vehicle-to-everything perception[C]//Proceedings of the 2023 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2023: 3584-3591.
- [103] BAI Z W, WU G Y, QI X W, et al. Cyber mobility mirror for enabling cooperative driving automation in mixed traffic: a co-simulation platform[J]. IEEE Intelligent Transportation Systems Magazine, 2023, 15(2): 251-265.
- [104] ZHENG Z L, HAN X, XIA X, et al. OpenCDA-ROS: enabling seamless integration of simulation and real-world cooperative driving automation[J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(7): 3775-3780.
- [105] CAESAR H, BANKITI V, LANG A H, et al. nuScenes: a multimodal dataset for autonomous driving[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 11618-11628.
- [106] SUN P, KRETZSCHMAR H, DOTIWALLA X, et al. Scalability in perception for autonomous driving: waymo open dataset[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 2443-2451.
- [107] XU R S, XIANG H, XIA X, et al. OPV₂: an open benchmark dataset and fusion pipeline for perception with vehicle-to-vehicle communication[C]//Proceedings of the 2022 International Conference on Robotics and Automation. Piscataway: IEEE Press, 2022: 2583-2589.
- [108] LI Y M, MA D K, AN Z Y, et al. V2X-sim: multi-agent collaborative perception dataset and benchmark for autonomous driving[J]. IEEE Robotics and Automation Letters, 2022, 7(4): 10914-10921.
- [109] YU H B, LUO Y Z, SHU M, et al. DAIR-V2X: a large-scale dataset for vehicle-infrastructure cooperative 3D object detection[C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 21329-21338.
- [110] ARNOLD E, MOZAFFARI S, DIANATI M. Fast and robust registration of partially overlapping point clouds[J]. IEEE Robotics and Automation Letters, 2022, 7(2): 1502-1509.
- [111] YU H B, YANG W X, RUAN H Z, et al. V2X-seq: a large-scale sequential dataset for vehicle-infrastructure cooperative perception and forecasting[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 5486-5495.
- [112] XU R S, XIA X, LI J L, et al. V₂V₄Real: a real-world large-scale dataset for vehicle-to-vehicle cooperative perception[C]//Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 13712-13722.
- [113] AXMANN J, MOFTIZADEH R, SU J Y, et al. LUCOOP: Leibniz University cooperative perception and urban navigation dataset[C]//Proceedings of the 2023 IEEE Intelligent Vehicles Symposium. Piscataway: IEEE Press, 2023: 1-8.
- [114] MA C, QIAO L, ZHU C K, et al. HoloVic: large-scale dataset and benchmark for multi-sensor holographic intersection and vehicle-infrastructure cooperative[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 22129-22138.
- [115] ZIMMER W, WARDANA G A, SRITHARAN S, et al. TUMTra

V2X cooperative perception dataset[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 22668-22677.

- [116] HAO R Y, FAN S Q, DAI Y R, et al. RCooper: a real-world large-scale dataset for roadside cooperative perception[C]//Proceedings of the 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 22347-22357.
- [117] LUO G Y, ZHANG H, YUAN Q, et al. Complementarity-enhanced and redundancy-minimized collaboration network for multi-agent perception[C]//Proceedings of the 30th ACM International Conference on Multimedia. New York: ACM, 2022: 3578-3586.
- [118] LUO G Y, YUAN Q, LI J L, et al. Artificial intelligence powered mobile networks: from cognition to decision[J]. IEEE Network, 2022, 36(3): 136-144.
- [119] ZHANG H, TIAN Y L, WANG K F, et al. Mask SSD: an effective single-stage approach to object instance segmentation[J]. IEEE Transactions on Image Processing, 2019, 29: 2078-2093.
- [120] HU Y, FANG S H, LEI Z X, et al. Where2comm: communication-efficient collaborative perception via spatial confidence maps[J]. Advances in Neural Information Processing Systems, 2022, 35: 4874-4886.
- [121] WANG B L, ZHANG L, WANG Z Z, et al. Core: cooperative reconstruction for multi-agent perception[C]//Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2023: 8676-8686.
- [122] YANG K, YANG D K, ZHANG J Y, et al. What2comm: towards communication-efficient collaborative perception via feature decoupling[C]//Proceedings of the 31st ACM International Conference on Multimedia. New York: ACM, 2023: 7686-7695.
- [123] KHAN N, COLERI S. Event-triggered reinforcement learning based joint resource allocation for ultra-reliable low-latency V2X communications[J]. IEEE Transactions on Vehicular Technology, 2024, 73(11): 16991-17006.
- [124] LU J W, HAN L Y, WEI Q L, et al. Event-triggered deep reinforcement learning using parallel control: a case study in autonomous driving[J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(4): 2821-2831.
- [125] VADIVELU N, REN M Y, TU J, et al. Learning to communicate and correct pose errors[C]//Proceedings of the Conference on Robot Learning. [S.l.:s.n.], 2021: 1195-1210.
- [126] GU J M, ZHANG J Y, ZHANG M Y, et al. FeaCo: reaching robust feature-level consensus in noisy pose conditions[C]//Proceedings of the 31st ACM International Conference on Multimedia. New York: ACM, 2023: 3628-3636.

[作者简介]



洪东洋（2005-），女，北京交通大学计算机科学与技术学院硕士生，主要研究方向为计算机视觉、深度学习。



张慧（1993-），女，北京交通大学计算机科学与技术学院副教授，主要研究方向为复杂环境下的车辆感知、多传感器融合的协同检测、多模态目标检测、群体智能、平行视觉。



张译匀（2002-），女，北京交通大学计算机科学与技术学院本科生，主要研究方向为计算机视觉、深度学习。



王雨桐（1994-），女，博士，中国科学院自动化研究所多模态人工智能系统国家重点实验室副研究员，主要研究方向为计算机视觉、图像异常检测。



韩瑜珊（1996-），女，北京交通大学计算机科学与技术学院博士生，主要研究方向为计算机视觉。



曹原周汉（1988-），男，北京交通大学计算机科学与技术学院副教授，主要研究方向为计算机视觉、深度学习、自动驾驶。



李滢东（1982-），男，北京交通大学计算机科学与技术学院副院长（主持工作）、教授、博士生导师，交通大数据与人工智能教育部重点实验室主任，主要研究方向为大数据智能、数据隐私保护、先进计算、智能交通等。