

地理信息数据采集与合规处理技术在智驾数据闭环建设中的应用

黄浩 高扬 耿东伶

湖北亿咖通科技有限公司

DOI:10.12238/gmsm.v8i2.2168

[摘要] 地理信息数据在驱动智能驾驶技术的发展中起着基础性作用,其采集与合规处理技术对于构建高效、安全的智驾数据闭环至关重要。基于此,本文深入探讨了地理信息数据采集和合规处理技术在智驾数据闭环中的应用,阐述了其在智驾数据闭环各环节的作用和实施方法,旨在为智能驾驶行业在地理信息数据管理方面提供技术参考和实施方案,通过技术创新与合规管理的协同发展,推动智能驾驶技术的安全、可持续发展。

[关键词] 地理信息数据; 数据采集; 合规处理; 智能驾驶; 数据闭环

中图分类号: G623.45 文献标识码: A

Application of Geospatial Data Acquisition and Compliance Processing Technologies in Intelligent Driving Data Closed-Loop Construction

Hao Huang Yang Gao Dongling Geng

HuBei Yikatong technology Co.,Ltd

[Abstract] Geospatial data plays a foundational role in advancing intelligent driving technologies, where its acquisition and compliance processing are pivotal for establishing efficient and secure intelligent driving data closed loops. This paper conducts an in-depth exploration of the application of geospatial data acquisition and compliance processing technologies within intelligent driving data closed loops, with detailed elaboration on their functions and implementation methodologies across all operational phases. The study aims to provide technical references and implementation frameworks for geospatial data management in the intelligent driving industry. Through synergistic integration of technological innovation and compliance governance, this work ultimately advances the safe and sustainable development of intelligent driving systems.

[Key words] Geospatial data; Data acquisition; Compliance processing; Intelligent driving; Data closed loop

引言

智能驾驶作为未来交通的重要发展方向,正受到全球广泛关注。其发展依赖于数据、算法和算力的协同作用,其中数据方向与地理信息技术密切相关。一方面,地图与地理信息数据包含丰富的道路、地形、环境等信息,为智能驾驶提供了精确的位置参考和环境感知基础,是智能驾驶数据闭环建设中不可或缺的部分。另一方面,装配有GPS、摄像头、激光雷达(部分)的智能汽车,也在深刻改变传统道路的测绘模式,不管是智驾地图的众包更新,还是自动驾驶的数据闭环,都需要采集和处理大量地理信息相关的汽车传感器数据。

然而,地理信息数据的采集涉及复杂的技术和严格的法规

要求,研究地理信息数据采集与合规处理技术在智驾数据闭环建设中的应用,对于突破智能驾驶技术瓶颈、促进产业规模化健康发展具有重要意义。

1 地理信息数据采集技术在智驾数据闭环中的应用

1.1 传统地理信息数据采集方式

1.1.1 基于专业采集车的测绘

专业采集车配备了高精度的GPS、惯性导航系统(INS)、激光雷达、摄像头等多种传感器。激光雷达通过发射激光束并测量反射光的时间来获取周围环境的三维点云数据,能够精确描绘道路的形状、路沿石位置、交通标志和标线等信息^[1]。摄像头则采集道路场景的图像数据,用于识别交通信号灯状态、车道

线颜色和类型等。

国内头部图商,通常有几十辆至上百辆专业采集车,利用专业采集车获取了大量精确的地理信息数据,为后续地图制作和智能驾驶应用提供了基础的地图/真值数据。而对于汽车主机厂、智驾公司,他们把这类车一般定义为真值车。他们与有资质的图商合作,驾驶这些真值车在行驶过程中,通过各传感器协同工作,每分钟可采集数GB的道路真实数据,对数据存储和传输能力提出了很高要求。

1.1.2 基于卫星遥感的采集

卫星遥感技术通过搭载在卫星上的各种传感器,获取地球表面的图像和数据。高分辨率卫星影像能够清晰显示道路、建筑物等地理特征,结合卫星定位系统,可精确获取地理坐标信息。在智能驾驶领域,卫星遥感数据可以提供面向智能驾驶发展初期“重地图”技术路线下的道路高精度地图底图数据^[2],辅助智能驾驶系统进行路径规划和环境感知。然而,卫星遥感数据更新周期较长,难以满足智能驾驶对实时性地理信息的需求。

1.2 新兴地理信息数据采集技术

1.2.1 众源数据采集

众源数据采集借助大量普通车辆(如量产车)上的传感器,实时采集车辆行驶过程中的位置、速度、行驶轨迹、周围环境感知等数据,并上传至云端服务器。例如,特斯拉通过影子模式,让车辆在行驶过程中传感器持续对决策算法进行验证,当系统算法决策与驾驶员行为不一致时,判定为“极端工况”,触发数据回传。这种方式可以收集到海量的真实驾驶场景数据,尤其是各种Corner case(极端情况)数据,极大地丰富了地理信息数据库。

基于量产车的众源测绘,实现了地理信息与车载感知的良性互动,采集成本相对较低,且鲜度高、场景覆盖广,但数据质量参差不齐,需要进行严格的数据筛选和清洗,同时满足数据隐私要求^[3-4]。

1.2.2 数字孪生与虚拟测绘

自动驾驶数据仿真与数字孪生打造了具备真实物理特性的虚拟世界,可以极大地丰富数据集的多样性与覆盖范围。目前比较主流的做法是,通过激光雷达的精准测绘确保厘米级精度的真实数据,借助3DGS/NeRF等神经渲染技术,实现道路环境三维场景的智能重建;同时,在生成的三维场景中,可灵活添加动态对象(车辆、行人、信号灯等),并全面兼容OpenSCENARIO标准,适配多种自动驾驶测试平台。相比传统数月的手动建模,神经网络重建只需几天时间,就能将采集到的图像和点云数据自动生成高保真的三维静态场景仿真数据。

2 地理信息数据合规处理技术的重要性及其应用

2.1 合规处理技术的重要性

随着智能驾驶产业的发展,地理信息数据安全成为国家安全的重要组成部分。自然资源部2022年8月发布的一号文^[5]明确指出,智驾数采车、量产车上的数据采集、回传被归类为测绘数据,必须严格遵守相应的测绘合规管理要求。不合规的数据采集和处理可能导致敏感地理信息泄露,威胁国家主权、安全和发展利益。同时,对于企业而言,合规问题直接关系到企业的运营合法性和商业信誉。若因数据合规问题被处罚,企业可能面临巨额罚款、业务受限等风险。因此,实施合规处理技术是智能驾驶企业和测绘资质图商必须重视的关键环节。

2.2 具体合规处理技术的应用

2.2.1 采集区域控制

在智驾公司委托图商用采集车实施数据采集前,要进行采集区域控制,在一些敏感区域,如军事管理区、重要基础设施周边等,限制或禁止数据采集,确保数据采集活动符合法律法规和国家安全要求。

2.2.2 数据加密与脱敏

数据加密是保障数据安全传输和存储的重要手段。在地理信息数据传输过程中,采用加密算法(如AES对称加密算法)将原始数据转换为密文,只有拥有解密密钥的接收方才能还原数据,防止数据在传输过程中被窃取或篡改。

数据脱敏包含地理信息脱敏(“地信脱敏”)与个人信息脱敏(“个性脱敏”)。地信脱敏处理则是由具备相应测绘资质的单位对地理信息数据中的敏感信息,如坐标、重要桥梁的三限(限高、限重、限宽)信息等,进行变形或删除。而个信脱敏是指在处理车内外视频数据时,对人脸、车牌等涉及个人隐私的信息进行模糊化处理,防止个人隐私泄露^[6]。

2.2.3 数据分类分级管理

对地理信息数据进行分类管理,根据数据的敏感程度、应用场景等因素,划分不同的安全级别,采取不同的存储、访问和处理策略。对于脱敏前的地理信息数据,严格限制访问权限,仅允许特定人员和系统在安全环境下进行操作。

2.2.4 与主管部门沟通协作

智能驾驶企业应与自然资源、网信等主管部门保持密切沟通,及时了解政策法规变化,确保技术实践符合最新的合规要求。在项目尤其是探索性项目的实施过程中,主动邀请主管部门进行合规审查,定期汇报数据采集与处理方案的进展情况,为企业合规运营提供保障。

3 地理信息数据采集与合规处理技术在智驾数据闭环各环节的作用

3.1 数据采集环节

合规的数据采集确保了数据来源的合法性,避免企业因违规采集面临法律风险。国内的各家车厂,均与具备甲级电子导航地图资质的企业建立深度战略合作,研发智能驾驶数据闭环架

构,在合法合规的前提下,为智驾数据闭环提供高效的数据采集方案的同时,保障了数据来源合法。

3.2 数据传输环节

在数据传输过程中,合规处理技术中的数据加密确保了地理信息数据在传输过程中的安全性,防止数据被窃取或篡改。同时,采用高效的数据压缩和网络优化技术,在保障数据安全的前提下,提高了数据传输效率,满足智能驾驶对实时性数据传输的需求^[7]。

3.3 数据存储环节

合规的数据存储要求对不同安全级别的地理信息数据进行分类存储,并采取相应的安全防护措施。云存储的分布式存储架构和对象存储技术,具备大容量、高可靠性、高性能、高伸缩性、低成本等优势,满足了智能驾驶海量地理信息数据存储的需求。同时,通过严格的访问权限控制,确保只有授权的系统和人员能够访问相应的数据,保障数据存储的安全性和合规性。国内多家云商已经陆续发布了具备国密等保认证的智驾云专区,为智能驾驶数据构筑安全“护城河”,已有大量车企在投入使用^[8]。

3.4 数据预处理与清洗环节

地理信息数据在采集过程中可能存在噪声、缺失值、错误值等问题,需要进行预处理和清洗。云计算和边缘计算技术在数据预处理与清洗中发挥着重要作用。云计算强大的计算能力可对大量数据进行高效处理,边缘计算则在数据源附近进行初步处理,降低数据传输延迟,同时保护数据隐私。经过预处理和清洗的数据,质量得到提升,为后续的数据标注和模型训练提供了可靠基础。

3.5 数据标注环节

AI自动标注技术利用机器学习和深度学习算法,对地理信息数据中的目标物体(如道路、交通标志、行人等)进行自动识别和标注。合规的数据标注需要确保标注结果的准确性和一致性,同时保护数据中的敏感信息。例如,特斯拉的自动标注系统在车辆行驶过程中,将摄像头收集的路面信息上传到服务器的离线神经网络大模型进行预测性标注,再反馈给车端传感器进行验证,确保标注结果的准确性。标注后的地理信息数据为智能驾驶模型训练提供了丰富的样本,是提升模型性能的关键环节。

3.6 模型训练与优化环节

地理信息数据作为智能驾驶模型训练的重要输入,对模型的准确性和泛化能力有着重要影响。合规处理后的高质量地理信息数据,能够使模型更好地学习道路规则、环境特征等信息,提高模型在复杂场景下的决策能力。同时,随着数据闭环的运行,不断有新的地理信息数据被采集,经过合规处理后回灌到模型训练中,用于模型的持续优化。

3.7 仿真测试环节

在仿真测试中,利用地理信息数据构建逼真的虚拟驾驶环境,对智能驾驶系统进行全面测试。合规的地理信息数据使用确保了仿真测试的合法性和安全性。通过模拟不同的地理环境、天气条件、交通状况等,检验智能驾驶系统的性能和可靠性,发现潜在问题并进行优化。

3.8 车端验证环节

车端验证是将经过模型训练和仿真测试优化后的智能驾驶系统部署到实际车辆上进行实地验证。地理信息数据在车端验证中用于辅助车辆定位、路径规划和环境感知。合规处理后的地理信息数据能够确保车端验证过程中数据使用的合法性,同时为智能驾驶系统在真实道路环境下的运行提供准确的地理信息支持,保障车端验证的顺利进行,推动智能驾驶系统从研发到实际应用的转化。

4 结论与展望

地理信息数据采集与合规处理技术在智驾数据闭环建设中具有不可替代的作用。准确、高效的数据采集技术为数据闭环提供了丰富的数据源,合规处理技术则确保了数据在整个生命周期内的合法、安全使用^[9]。通过在智驾数据闭环各环节的协同应用,这些技术有力地推动了智能驾驶技术的发展和优化。

然而,随着智能驾驶技术的不断演进,自动驾驶数据服务的底层逻辑,正在从静态数据采集转向动态智能决策转化。未来,测绘地理信息将深度融入智能驾驶的“感知-决策-执行”的链条。我们需要进一步研发更先进的数据采集技术,提高数据采集的精度、效率和覆盖范围;持续优化合规处理技术,在满足严格法规要求的同时,降低企业合规成本;加强真实测绘与虚拟测绘融合的技术研究,促进地理信息数据与其他类型数据的深度融合,为智能驾驶系统提供更全面、准确的决策支持。此外,随着智能驾驶全球化的推进,也将推动全球测绘治理体系的重构,测绘标准、数据跨境流动等问题也在上升成为国际博弈的焦点之一。

相信在地理信息数据采集与合规处理技术的不断创新和完善下,通过技术创新与合规管理的协同发展,智能驾驶将迎来更加广阔的发展前景,为人们的出行带来更多便利和安全保障。

[参考文献]

[1]胡文雄.车载移动测量技术在道路测量中的应用[J].工程勘察,2019,47(7):62-65.

[2]杨振凯,华一新,訾璐等.浅析高精度地图发展现状及关键技术[J].测绘通报,2021(6):54-60.

[3]陈会仙,章炜,杨蒙蒙,等.面向无人驾驶的高精地图公开应用趋势研究[J].武汉大学学报(信息科学版),2024,49(4):537-545.

[4] 谢富鑫, 郭圆, 周剑. 智能汽车高精地图众源更新研究现状[J]. 人工智能, 2023(5):48–54.

[5] 自然资源部. 关于加强智能网联汽车有关测绘地理信息安全管理的通知[EB/OL].[2024-07-26].https://www.gov.cn/zhengce/zhengceku/202407/content_6965227.htm.

[6] 朱长青, 任娜, 徐鼎捷. 地理信息安全技术研究进展与展望[J]. 测绘学报, 2022, 51(6):1017–1028.

[7] Guan T, Han Y, Kang N, et al. An overview of vehicular cybersecurity for intelligent connected vehicles[J]. Sustainability, 2022, 14(9):5211.

[8] 杨殿阁, 李庆建, 王艳. 高精动态地图基础平台参考架构和技术路线[J]. 智能网联汽车, 2021(1):74–83.

[9] Karkošková S. Data governance model to enhance data quality in financial institutions[J]. Information Systems Management, 2023, 40(1):90–110.

作者简介:

黄浩(1979--), 男, 汉族, 湖南汨罗人, 工程师, 职务: 总经理, 硕士研究生, 研究方向: 地图导航、时空数据智能。