



南開大學
Nankai University

计算机学院
软件工程报告

无人驾驶网约车多模态交互系统需求分析

姓名：何畅

学号：2210620

专业：计算机科学与技术

2025 年 4 月 6 日

目录

1 引言	3
1.1 编写目的	3
1.2 项目背景	3
2 任务概述	3
2.1 任务目标	3
2.2 用户特点	4
3 假定与约束	4
3.1 假定	4
3.2 约束	5
4 业务概述	5
4.1 系统总业务流程图及其描述	5
4.2 各个子业务流程图及其描述	6
4.2.1 乘客子系统功能	6
4.2.2 系统管理员子系统功能	7
4.2.3 车辆子系统功能	8
4.2.4 车辆维护人员子系统功能	9
5 数据需求	10
5.1 数据需求描述	10
5.1.1 用户行为与偏好数据	10
5.1.2 车辆状态与环境数据	10
5.1.3 多模态交互数据	11
5.1.4 交通与外部环境数据	11
5.1.5 系统管理与安全数据	11
5.1.6 应急响应数据	11
5.2 数据流图	12
5.2.1 顶层数据流图	12
5.2.2 乘客信息管理数据流图	12
5.2.3 多模态交互数据流图	13
5.2.4 应急响应信息数据流图	13
5.2.5 乘客反馈与系统优化数据流图	14
5.3 数据字典	14
5.3.1 多模态交互数据	14
5.3.2 用户行为与偏好数据	16
5.3.3 车辆状态与环境数据	17
6 功能需求	18
6.1 功能划分	18
6.2 功能描述	18

6.2.1	多模态交互整合	18
6.2.2	车内环境与状态监控	18
6.2.3	智能决策与优化	18
6.2.4	系统数据与日志管理	19
6.2.5	系统维护与优化	19
7	创新功能	19
8	性能/非功能需求	20
8.1	准确性	20
8.2	及时性	20
8.3	可扩充性	20
8.4	易用性	20
8.5	易维护性	21
8.6	标准性	21
8.7	先进性	21
9	系统运行要求	21
9.1	硬件配置要求	21
9.2	软件配置要求	22

1 引言

1.1 编写目的

本项目需求分析报告的目的是详细阐述开发一个无人驾驶网约车多模态交互系统的需求和预期功能。该系统旨在为无人驾驶网约车的乘客、车辆控制系统以及车载管理人员提供一个高效、便捷的智能交互解决方案。通过整合语音识别、手势控制、眼动追踪、触控反馈等多种交互方式，平台将提升乘客在不同环境下的操作便捷性和舒适度，增强紧急情况下的应急响应能力。本报告将明确系统的用户需求、功能需求、性能指标及安全要求，作为软件开发团队的指导文件，确保开发过程中的每个环节都能满足最终用户的期望，同时提升无人驾驶网约车的市场竞争力。

1.2 项目背景

随着无人驾驶技术的快速发展，部分城市已经开始试点无人驾驶网约车服务。然而，尽管在自动驾驶和车辆控制方面取得了进展，人机交互，尤其是多模态交互仍然是一个亟待解决的问题。当前市场上的无人驾驶网约车大多依赖于单一的交互方式，如触控屏或语音助手，但这种方式无法适应不同乘客的需求和复杂的使用环境。例如，在嘈杂环境下语音识别效果不佳，触控操作则可能在行驶过程中带来不便，无法提供快速、直观的反馈。

乘客对车内互动方式的需求多样化，传统的交互方式已无法满足所有场景，尤其在突发的紧急情况下，现有交互系统缺乏高效应急响应能力，严重影响了乘客的安全感与使用体验。因此，开发一个智能、灵活的多模态交互系统，能够提升无人驾驶网约车在各种复杂环境下的交互效率和安全性，成为当前行业亟需解决的痛点。

本项目旨在设计并实现一个多模态交互系统，结合语音、手势、眼动追踪等方式，使乘客能够根据具体情境选择最适合的交互模式，从而提供更高效、便捷的乘车体验，满足不同用户需求，提升无人驾驶网约车的市场竞争力。

2 任务概述

2.1 任务目标

本项目致力于构建一个创新的无人驾驶网约车多模态交互系统，旨在通过以下几个关键目标，解决当前无人驾驶网约车在用户交互体验和应急响应中的挑战：

- **高效多模态交互**：开发一个基于语音、手势、触控、眼动追踪等多种交互方式的智能系统。该系统将利用最新的人工智能技术，实时分析乘客的需求和环境变化，自动选择最适合的交互方式。无论是安静的车内环境还是嘈杂的外部环境，乘客都能够轻松、直观地与车辆进行互动，大大提升交互效率，减少因操作不便产生的不良体验。
- **智能应急响应**：为确保乘客在复杂和紧急情况下能够快速得到反馈和指引，系统将基于人工智能和机器学习算法，智能分析交通环境和突发事件，提供实时、精准的应急指引。例如，系统将在交通堵塞或事故发生时，迅速调整交互方式并向乘客提供实时反馈，确保乘客能够高效、安全地做出反应。
- **个性化交互体验**：系统将根据乘客的行为习惯、情绪状态等进行智能学习和优化，提供个性化的交互方式。例如，系统可以根据乘客的历史交互数据，自动调整语音助手的语调、响应方式。此

外，所有无人驾驶网约车的系统将能够共享乘客的数据，使得乘客在不同车辆中的使用体验保持一致。通过这种数据共享，系统能够持续学习和适应乘客的偏好，不断提升交互体验。通过不断学习用户的偏好和习惯，平台能够提高用户的长期满意度和使用粘性。

- **优化用户体验：**系统的设计将遵循人机交互的理念，提供简洁、直观且易于操作的界面。平台将根据不同的用户角色（乘客、系统管理员）定制化交互方式，帮助他们快速访问相关功能。系统还将提供详尽的使用指南，帮助用户快速掌握操作流程，提升系统的易用性和用户满意度。

2.2 用户特点

- **乘客：**作为主要使用者，乘客的需求多样且个性化。不同乘客可能对交互方式有不同的偏好，例如一些乘客可能倾向于使用语音命令，而另一些则可能偏好触控或手势交互。乘客还期望系统能够理解他们的情绪状态和行为习惯，并根据个人偏好提供个性化的操作体验。此外，乘客对安全性和应急响应能力有较高的期望，希望系统能够在复杂或突发的交通状况下提供准确及时的反馈。
- **系统管理员：**系统管理员负责平台的整体运营，包括用户数据管理与权限控制、系统维护、数据分析和反馈。对于该类用户，系统需要提供强大的数据分析工具，帮助他们深入理解用户行为、交互模式以及系统运行状态。系统管理员还需能够高效地处理用户反馈，优化系统功能并进行定期的维护更新。
- **车辆：**作为系统中的重要角色，车辆不仅执行驾驶任务，还负责与乘客进行多模态交互。车辆具备感知周围环境的能力，并根据实时数据做出决策，确保安全行驶。在行程中，车辆通过语音、显示屏、手势识别、眼动追踪等多种交互方式与乘客互动，提供动态信息反馈。同时，车辆能够实时监控自身状态，如电量、传感器健康、轮胎状态等，并主动向平台反馈异常情况。车辆还需要根据乘客的行为和情绪状态进行智能调整，例如通过语音语调、车内温度等个性化设置提升用户体验。对于突发的交通状况或安全事件，车辆需要快速响应，采取应急措施，确保乘客的安全。
- **车辆维护人员：**车辆维护人员主要负责无人驾驶网约车的状态监控、故障诊断、维护保养和远程支持。为了确保车辆的正常运行，他们需要系统提供实时的设备状态数据，包括传感器、电池、制动系统等关键部件的运行情况。当出现故障时，维护人员依靠系统的智能诊断功能快速定位问题，获取故障代码、异常日志和初步的故障分析结果。此外，系统应具备数据分析和预测能力，通过历史数据进行趋势分析，提前发现潜在故障并进行预防性维护。

3 假定与约束

3.1 假定

假定相关的自动驾驶技术、语音识别、手势识别、触控技术以及眼动追踪技术已经足够成熟，并能够有效集成到系统中。假定每辆车都配备了所需的硬件设备，包括但不限于高质量的语音识别硬件、手势识别摄像头、触控屏幕以及眼动追踪传感器，以支持多模态交互。假定所有车辆与平台之间可以通过统一的数据共享协议交换用户数据，以确保乘客的个性化体验在不同车辆之间无缝延续。假定车辆与云端之间有稳定的网络连接，保证实时数据传输和系统的高效运行，尤其是在紧急情况下需要快速响应。假定乘客的智能手机能够与车载系统良好兼容，以便进行更好的个性化配置和交互。

3.2 约束

系统必须遵守自动驾驶和道路交通等相关法律法规。系统必须在噪音较大的环境下（例如城市街道、交通繁忙的区域等）仍能有效识别和响应用户的语音和手势指令，因此需要开发更加鲁棒的语音识别和手势识别算法。系统必须遵循严格的数据隐私和安全标准，在处理共享的乘客数据时，确保乘客个人信息的保护。由于无人驾驶网约车需要对交通环境和突发事件做出及时反应，系统必须具备高效的实时数据处理能力，以保证应急响应的及时性和准确性。系统的硬件（如传感器、计算单元）可能存在处理能力和带宽的限制，需要优化系统架构以适应这些硬件条件。由于不同乘客的技术适应能力不同，系统的设计必须确保即使是技术上不太熟悉的用户，也能够轻松使用和掌握多模态交互系统，避免过度依赖复杂操作。

4 业务概述

4.1 系统总业务流程图及其描述

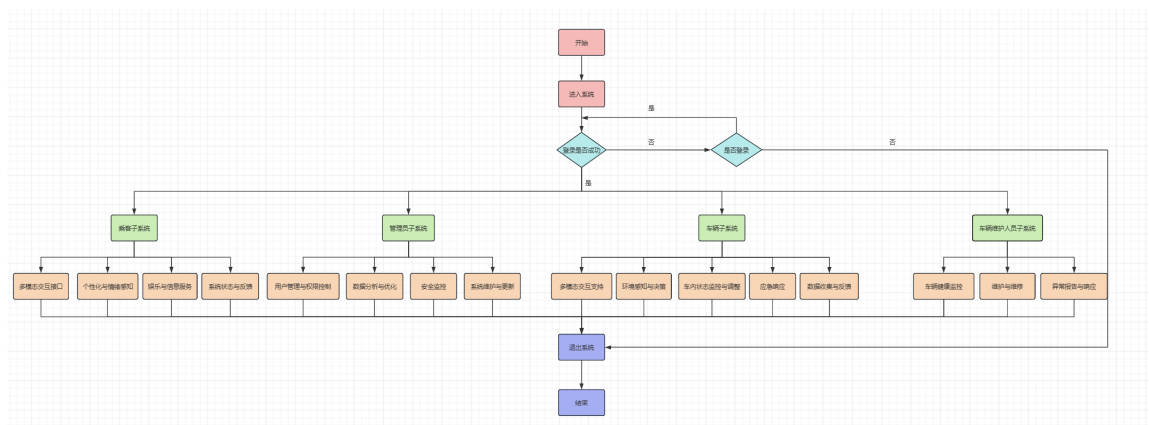


图 4.1: 总业务流程图

用户登录与身份验证首先，乘客、系统管理人员以及车辆维护人员通过智能手机或车载系统进行登录。系统通过身份验证确保用户的权限，乘客登录后，系统会根据其身份提供相应的交互功能和界面，系统管理人员和车辆维护人员则会根据不同角色进入管理或维护界面。

车辆健康监控与诊断在登录成功后，车辆的健康监控与诊断功能会自动启动。车辆系统会实时检测并评估车辆的状态，包含电池电量、传感器功能、轮胎状态等。这些信息会实时传输给车辆维护人员，确保车辆始终处于最佳工作状态，并且如有故障或异常，能够迅速被识别和处理。

多模态交互功能启动当乘客进入车内后，系统会启动多模态交互功能，提供包括语音识别、手势识别、触控屏幕、眼动追踪等多种交互方式。乘客可以通过不同的方式与系统进行互动，设定目的地、调整车内温度、座椅等，或者获取娱乐、信息等服务。此时，系统还能够根据乘客的行为和语音分析，感知其情绪状态并提供个性化的交互体验。

个性化与情绪感知在车辆行驶过程中，系统还能够进行个性化调整，基于乘客的历史行为和实时情绪状态调整车内环境，提升乘客的舒适度。如果乘客处于紧张状态，系统可以通过语音和车内设置帮助其放松，提高乘车体验。

应急响应与安全管理同时，系统具备应急响应功能，能够在遇到交通事故或其他紧急情况时迅速作出反应。车辆会自动采取刹车、避让障碍物等措施，系统则通过语音或触控界面向乘客提供相关安全信息和指导。

数据收集与系统反馈系统的后台会不断收集和分析用户的交互数据、车辆状态数据等，并由系统管理员进行实时监控。管理员可以根据数据分析优化系统，解决潜在问题，保证系统始终稳定运行。

车辆维护与更新对于车辆的日常维护，车辆维护人员会根据系统提供的数据进行预防性维修，确保车辆时刻处于良好的工作状态。当车辆发生故障时，系统会立即通知维护人员并提供故障诊断信息，帮助快速解决问题。

4.2 各个子业务流程图及其描述

4.2.1 乘客子系统功能

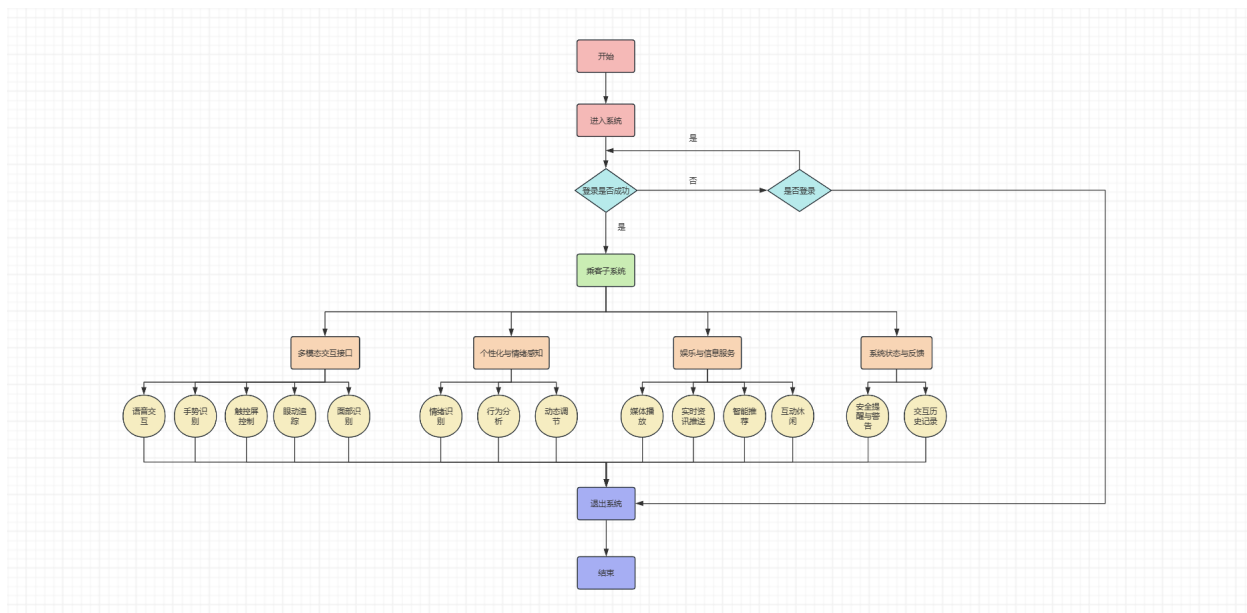


图 4.2: 乘客子系统流程图

多模态交互接口 乘客子系统的多模态交互接口是乘客与系统之间最重要的沟通桥梁，提供语音、手势、触控以及视觉交互等多种方式。乘客可以根据不同环境和个人偏好选择合适的交互方式。具体功能包括：

语音控制：乘客通过语音指令与系统进行交流，语音识别系统能够在嘈杂环境中识别准确的命令，例如调节车内温度、控制音乐播放、查询目的地等。系统还能够根据乘客的语音指令进行反馈，并通过自然语言处理技术提供流畅的对话式交互。

手势交互：车内安装的手势识别传感器使得乘客可以通过简单的手势进行操作。例如，手势滑动可调节座椅位置，指向或挥手可控制车窗的升降或空调的开关。

触控屏幕：车载屏幕提供直观的图形化界面，乘客可通过触摸操作控制车内环境、查看行驶信息或调整其他设置。

眼动追踪与视觉识别：系统通过眼动追踪技术捕捉乘客的视线焦点，并根据乘客视线的变化自动调节显示屏内容或提示乘客。图像识别功能还能够识别乘客的面部表情，进一步增强互动体验。

通过这些多模态交互接口，系统能够在不同的乘车环境下为乘客提供便捷、灵活的操作方式。

个性化情绪感知 个性化情绪感知是乘客子系统的核心功能之一，旨在通过分析乘客的情绪和行为状态，提供更加贴心和舒适的服务。系统通过**面部识别**、**语音语调分析**等技术识别乘客的情绪状态，并

根据其情绪调整车内环境。例如，若系统识别到乘客情绪紧张，车内的音乐和光线可以调节为舒缓的风格；若识别到乘客疲劳，车内温度和座椅可自动调节为更加舒适的状态。此外，系统还可以根据乘客的历史偏好和当前情绪状态推荐个性化的服务，如推荐放松的音乐、休息提示等，进一步提升乘客的体验感。

娱乐与信息服务 为了提升乘客的乘车体验，乘客子系统还包括丰富的娱乐和信息服务功能。系统通过车载娱乐系统为乘客提供多种娱乐内容，包括电影、音乐、广播等，乘客可以通过语音或触控屏幕选择所需的内容。此外，系统会根据乘客的偏好和情绪状态推荐合适的娱乐内容。例如，在长途旅行中，系统可以为乘客推荐轻松的音乐或热门电影，提升其娱乐体验。系统还提供实时的交通信息，包括道路状况、预计到达时间、路况变化等。乘客还可以通过系统获取个性化的新闻推送，包括天气、运动赛事等内容，或与朋友、家人保持联系，进行语音或视频通话，甚至通过社交平台进行互动。

系统状态与反馈 乘客子系统还包含实时的系统状态监控与反馈功能，确保乘客对车辆的当前状态和行驶进度始终保持清晰了解。系统实时监控车辆的各项健康指标，包括电量、油量、轮胎压力、传感器状态等，并通过屏幕或语音向乘客反馈车辆当前的运行状态。若发现异常情况，系统会主动警告乘客并提出建议。系统还提供实时的导航和行驶进度信息，乘客可以随时查看预计到达时间、路况信息及路线规划。如果有意外堵车或事故，系统会主动提醒乘客，并提供替代路线选择。此外，系统能够根据车辆周围的环境进行实时监测，并提供安全提示和预警。例如，若周围有障碍物或前方出现急刹车情况，系统会通过语音或屏幕及时通知乘客。

4.2.2 系统管理员子系统功能

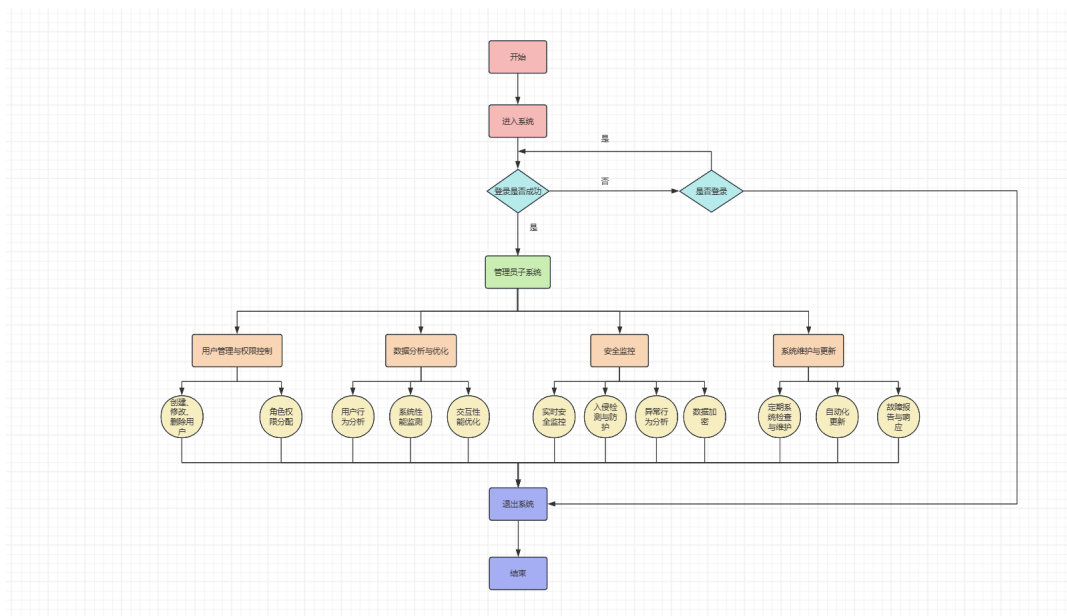


图 4.3: 系统管理员子系统流程图

用户管理与权限控制 管理员子系统提供了强大的用户管理与权限控制功能，确保不同角色的用户能够按照其权限进行相应的操作。该功能模块允许管理员创建、删除和管理平台上的所有用户账户，包括乘客、系统管理人员、车辆维护人员等。每个用户根据其职责拥有不同的权限，系统管理员可以精细

化配置这些权限，例如限制某些用户访问敏感数据或特定功能。管理员可以查看用户的使用日志，分析用户行为并根据需要调整权限设置，确保系统的安全性和合规性。

数据分析与优化 数据分析与优化是管理员子系统的另一个核心功能。系统管理员通过该模块能够实时监控和分析多模态交互系统的各类数据，包括用户行为数据、系统运行数据、车辆状态数据等。基于这些数据，管理员可以识别系统中的潜在问题和瓶颈，提出优化方案。例如，通过分析用户的交互习惯，系统管理员可以调整交互方式的优先级或改进算法，从而提升用户体验。管理员还可以根据大数据分析结果进行系统性能的调优，确保系统在不同负载情况下的稳定性和响应速度。

安全监控 安全监控是管理员子系统的重要组成部分，旨在确保系统的各项操作安全、数据传输安全以及用户隐私保护。该功能模块实时监控系统中的所有安全风险，包括网络攻击、恶意行为、系统漏洞等。一旦检测到安全事件，系统管理员会立刻收到警报并进行处理。管理员可以通过安全监控系统查看各类安全日志、分析潜在的安全隐患并采取相应的应急措施。此外，管理员还负责定期进行系统的安全检查，更新安全策略，确保系统免受外部攻击以及数据泄露的威胁。

系统维护与更新 系统维护与更新功能确保多模态交互系统长期稳定运行。管理员通过该模块执行定期的系统检查、漏洞修复以及性能优化，确保系统不会出现故障或错误。管理员还负责系统的版本更新，包括更新操作系统、驱动程序、算法模块等，确保系统始终使用最新的技术和功能。在进行更新时，管理员需要保证更新过程的平稳进行，避免系统中断或影响用户体验。管理员还需要在更新前进行详细的测试和验证，确保新版本没有引入新问题。在系统运行过程中，管理员也负责对异常状态进行排查和修复，确保系统高效、稳定地运行。

4.2.3 车辆子系统功能

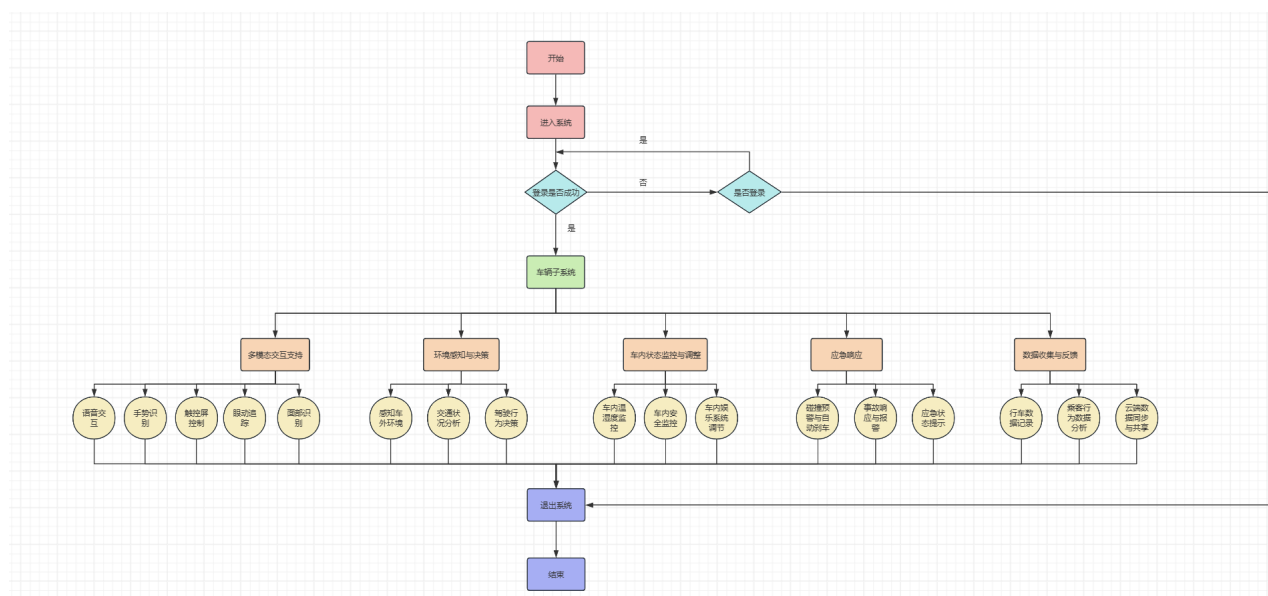


图 4.4: 车辆子系统流程图

多模态交互支持 车辆子系统支持多模态交互功能，使乘客可以通过多种方式与车辆进行互动，包括语音、手势、触控屏以及眼动追踪等。系统能够实时处理乘客的语音命令、手势动作、眼动数据等，智能分析乘客的需求并选择最适合的交互方式。例如，当车内环境较为嘈杂时，车辆可以自动切换为手

势控制或视觉交互，以确保交互的便捷性和准确性。车辆子系统还可以根据乘客的情绪状态进行动态调整，如调整语音助手的语气或语速，以提升乘客的舒适度和互动体验。

环境感知与决策 车辆子系统通过一系列传感器和算法进行环境感知，以实现对外围交通环境的实时监控与分析。包括摄像头、雷达、激光雷达等传感器，车辆能够准确感知周围的障碍物、行人、其他车辆等信息。基于感知数据，车辆能够进行智能决策，例如调整行驶路线、变更车速或发出警告，以确保行车安全。车辆子系统不仅在常规驾驶情况下进行决策，也能应对复杂或突发的交通状况，如交通堵塞、道路施工等，做出最优响应。

车内状态监控与调整 车辆子系统通过车载传感器和监控系统实时监测车内的环境与乘客的状态。例如，车内温度、湿度、空气质量、座椅位置等都可以被监控和调整。当车辆检测到车内温度过高或湿度过大时，系统会自动调整空调温度或湿度，提升乘客的舒适度。此外，车辆还能够检测乘客的行为，例如是否系好安全带，是否出现异常行为，并在必要时发出提醒，确保乘客的安全。

应急响应 应急响应功能是车辆子系统中至关重要的一部分。在突发情况下，如交通事故、车辆故障或乘客出现紧急情况时，车辆子系统需要快速响应。系统能够根据不同情境做出应急决策，如自动停车、通知系统管理人员、启动紧急制动系统等。此外，车辆还可以通过多模态交互方式向乘客提供紧急指导，例如通过语音或显示屏提醒乘客如何处理当前情况，确保乘客的安全和冷静应对。

数据收集与反馈 车辆子系统需要不断收集和反馈关于系统运行、乘客行为和车辆状态的数据。这些数据可以包括车辆的行驶数据、乘客的交互数据、环境感知数据等，所有数据都会上传到平台管理系统。通过数据收集，车辆能够实时向平台反馈当前状态、异常情况以及系统的运行效率，帮助系统管理人员进行优化和决策。此外，收集的数据还可以用于后续的分析与优化，提升车辆子系统的性能和交互体验。

4.2.4 车辆维护人员子系统功能

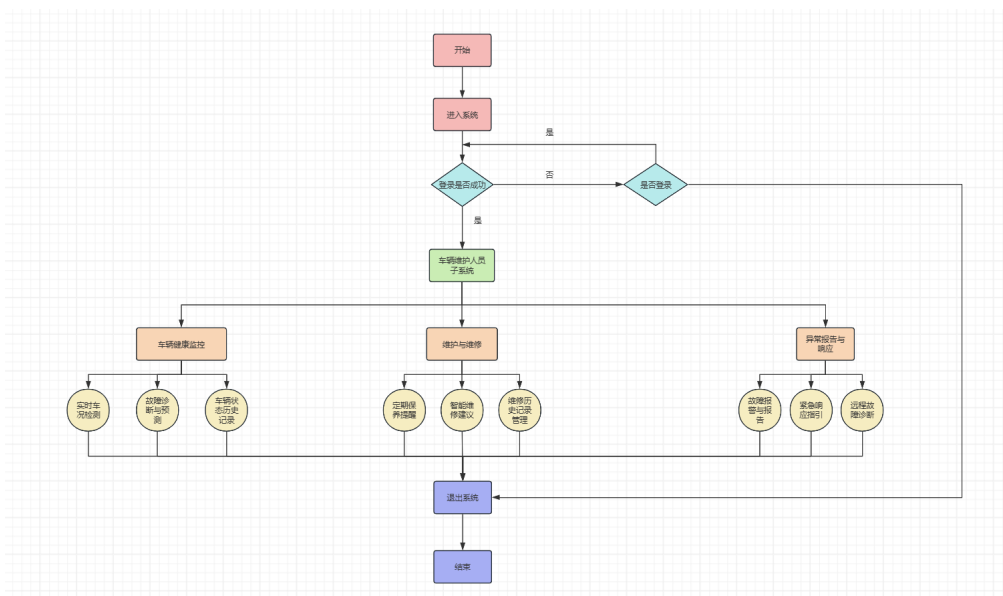


图 4.5: 车辆维护人员子系统流程图

车辆健康监控 车辆健康监控功能通过传感器和诊断系统实时检测车辆的各项性能指标，包括引擎状态、电池电量、轮胎压力、传感器健康、刹车系统等。该功能能够提供车辆各个组件的详细运行数据和状态，及时发现潜在问题并生成健康报告。通过健康监控系统，维护人员能够准确地了解车辆的运行状况，识别需要检修或更换的部件。此外，车辆健康监控功能还可以根据历史数据分析，预测车辆未来可能出现的故障，提供预防性维护建议，减少突发故障的发生。

维护与维修 维护与维修模块是车辆维护人员子系统中的核心功能，旨在确保车辆在出现故障或损坏时能够及时得到修复。该功能支持对车辆的定期维护、故障诊断和维修操作。维护人员可以通过系统查询到车辆的详细维护记录、保养周期和维修历史，并根据系统提示进行定期的维护和检修。系统还支持对故障问题进行精准诊断，帮助维护人员快速定位故障原因，并根据故障类型进行维修操作。在维修过程中，维护人员可以记录维修日志，确保每次维修都有详细的记录，以便后续追踪和审计。

异常报告与响应 异常报告与响应功能旨在实时检测车辆运行过程中的异常情况，并及时采取相应措施。车辆在运行过程中如果遇到故障或异常情况（如传感器故障、动力系统异常等），会自动生成异常报告并通知车辆维护人员。系统会通过多个渠道（如短信、邮件、系统通知等）将异常报告及时传送给维护人员，确保他们能够在第一时间内得知问题。车辆维护人员收到报告后，可以立即采取应急措施，如远程重启系统、调整控制策略或安排现场维修。此外，维护人员还可以通过异常响应模块与系统管理人员进行协作，确保异常问题能够尽快解决，避免影响车辆的正常运营。

5 数据需求

5.1 数据需求描述

5.1.1 用户行为与偏好数据

- **交互方式偏好**: 记录乘客与车辆的交互方式偏好（语音、手势、触控等），以便为每个用户提供个性化的交互体验。
- **情绪状态数据**: 通过眼动追踪、面部识别或语音分析等手段采集乘客的情绪状态数据，帮助系统调整交互方式或车内环境，以增强乘客的舒适度。
- **历史交互数据**: 记录用户的历史交互行为，便于系统优化个性化推荐，例如根据乘客的过去命令来调整语音助手的语调或响应方式。

5.1.2 车辆状态与环境数据

- **车辆传感器数据**: 来自车辆各类传感器的数据，如摄像头、雷达、激光雷达（LiDAR）等，提供车辆周围环境的信息，用于辅助决策（例如避障、路线规划、障碍物检测等）。
- **车内监测数据**: 包括温度、湿度、座椅加热/通风状态、车内空气质量等数据，用于优化乘客体验。
- **车辆健康数据**: 车辆的实时健康监控数据，涉及到发动机、刹车系统、传动系统等状态，用于维护与修复管理。

5.1.3 多模态交互数据

- **语音数据**：乘客的语音指令，包括语音识别数据和相应的语音响应数据，用于训练和优化语音助手系统。
- **手势数据**：乘客的手势动作数据（通过摄像头或手势识别设备），用于解读乘客的指令和行为。
- **眼动数据**：乘客的眼动数据，用于分析乘客的注意力和情绪，优化交互界面和反馈机制。
- **触控数据**：触控屏上的操作数据，用于支持触控交互方式。
- **面部数据**：乘客的面部表情数据，用于分析乘客的情绪状态，优化交互方式和提供更人性化的反馈。

5.1.4 交通与外部环境数据

- **实时交通数据**：交通流量、拥堵情况、交通事故等外部数据，帮助优化路径规划、避免交通拥堵。
- **天气数据**：实时天气信息（如温度、降水量、风速等），影响车辆行驶的决策和乘客的舒适度。
- **道路状态数据**：路面状况（如施工、道路坑洞、积水等）数据，用于提供安全的行车路线。

5.1.5 系统管理与安全数据

- **用户数据管理**：用户的个人信息、账户信息、权限配置、历史交互数据等，保障系统的个性化体验和安全性。
- **数据安全与隐私保护**：所有涉及到用户、车辆、环境的敏感数据需要加密、脱敏处理，确保符合隐私保护法规和标准。
- **故障诊断数据**：车辆或系统出现故障时的日志数据，用于故障分析和改进。
- **系统运行数据**：系统状态、性能指标、负载信息、网络连接状态等，用于保障系统稳定运行和性能优化。

5.1.6 应急响应数据

- **异常状态数据**：车辆或系统检测到的异常状态（如紧急刹车、事故检测等），需要被实时传输到后台，并触发应急响应机制。
- **应急处理方案数据**：基于实时数据生成应急响应方案（如避障、路线调整、车辆停驶等），并反馈给驾驶员、乘客或平台管理人员。

5.2.5 乘客反馈与系统优化数据流图

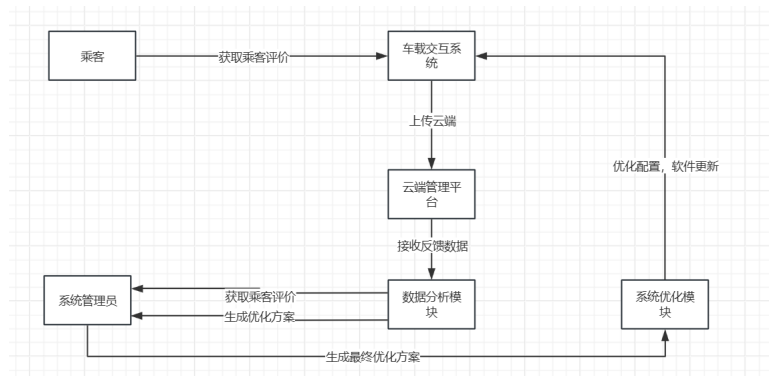


图 5.10: 乘客反馈与系统优化数据流图

5.3 数据字典

5.3.1 多模态交互数据

语音数据

- **数据项名称:** 语音数据
- **数据项定义:** 乘客的语音指令，包括语音识别数据和相应的语音响应数据，用于训练和优化语音助手系统。
- **数据类型:** 字符串/音频
- **数据格式:** 音频文件（如 WAV 格式）或字符串
- **数据来源:** 车载麦克风
- **约束条件:** 语音识别准确度达到 95% 以上
- **用途:** 用于语音助手交互、语音识别和响应

手势数据

- **数据项名称:** 手势数据
- **数据项定义:** 乘客的手势动作数据（通过摄像头或手势识别设备），用于解读乘客的指令和行为。
- **数据类型:** 字符串/二维坐标
- **数据格式:** 手势动作的坐标序列或预定义手势标签
- **数据来源:** 车载摄像头或手势识别设备
- **约束条件:** 动作识别准确率需达到 95% 以上
- **用途:** 用于识别手势命令并执行对应的系统操作

眼动数据

- **数据项名称：**眼动数据
- **数据项定义：**乘客的眼动数据，用于分析乘客的注意力和情绪，优化交互界面和反馈机制。
- **数据类型：**结构体
- **数据格式：**包含注视点坐标、眼睛状态等数据
- **数据来源：**眼动追踪设备
- **约束条件：**追踪精度要求达到 1 度以内
- **用途：**用于判断乘客注意力，调整交互方式和反馈

触控数据

- **数据项名称：**触控数据
- **数据项定义：**触控屏上的操作数据，用于支持触控交互方式。
- **数据类型：**二维坐标
- **数据格式：**触控坐标、触控类型（单击、滑动等）
- **数据来源：**触控屏
- **约束条件：**坐标精度要求为 1mm 以内
- **用途：**用于支持触控交互方式

面部数据

- **数据项名称：**面部数据
- **数据项定义：**乘客的面部表情数据，用于分析乘客的情绪状态，优化交互方式和提供更人性化的反馈。
- **数据类型：**结构体
- **数据格式：**面部表情特征值或表情标签
- **数据来源：**车载摄像头
- **约束条件：**表情识别准确率需达到 90% 以上
- **用途：**用于情绪分析与反馈优化

5.3.2 用户行为与偏好数据

交互方式偏好

- **数据项名称：**交互方式偏好
- **数据项定义：**记录乘客与车辆的交互方式偏好（语音、手势、触控等），以便为每个用户提供个性化的交互体验。
- **数据类型：**字符串
- **数据格式：**可选值：“语音”，“手势”，“触控”
- **数据来源：**乘客设置或系统推断
- **约束条件：**必填项
- **用途：**用于个性化交互方式调整

情绪状态数据

- **数据项名称：**情绪状态数据
- **数据项定义：**通过眼动追踪、面部识别或语音分析等手段采集乘客的情绪状态数据，帮助系统调整交互方式或车内环境，以增强乘客的舒适度。
- **数据类型：**字符串/数字
- **数据格式：**例如：“愉快”，“焦虑”，“疲劳”，或基于数值评分的情绪状态（如 1 到 5 的评分）
- **数据来源：**眼动追踪，面部识别，语音分析
- **约束条件：**值在预定义的情绪范围内
- **用途：**用于情绪感知，调整车内环境和交互策略

历史交互数据

- **数据项名称：**历史交互数据
- **数据项定义：**记录用户的历史交互行为，便于系统优化个性化推荐，例如根据乘客的过去命令来调整语音助手的语调或响应方式。
- **数据类型：**数组/结构体
- **数据格式：**JSON 格式，包含用户 ID、交互时间、交互方式、交互内容
- **数据来源：**用户交互历史记录
- **约束条件：**数据存储周期为 1 年，超过时间限制的数据将被自动删除
- **用途：**用于个性化推荐和交互优化

5.3.3 车辆状态与环境数据

车辆传感器数据

- **数据项名称：**车辆传感器数据
- **数据项定义：**来自车辆各类传感器的数据，如摄像头、雷达、激光雷达（LiDAR）等，提供车辆周围环境的信息，用于辅助决策（例如避障、路线规划、障碍物检测等）。
- **数据类型：**结构体
- **数据格式：**JSON 格式，包含时间戳、传感器类型、测量值等
- **数据来源：**车载传感器（摄像头、雷达、LiDAR）
- **约束条件：**数据更新频率不低于 10Hz
- **用途：**用于环境感知、障碍物检测与路径规划

车内监测数据

- **数据项名称：**车内监测数据
- **数据项定义：**包括温度、湿度、座椅加热/通风状态、车内空气质量等数据，用于优化乘客体验。
- **数据类型：**结构体
- **数据格式：**JSON 格式，包含各项传感器数据（如温度、湿度等）
- **数据来源：**车内环境传感器
- **约束条件：**数据精度为小数点后两位
- **用途：**优化车内环境（如空调调节）

车辆健康数据

- **数据项名称：**车辆健康数据
- **数据项定义：**车辆的实时健康监控数据，涉及到发动机、刹车系统、传动系统等状态，用于维护与修复管理。
- **数据类型：**结构体
- **数据格式：**JSON 格式，包含各项车辆系统的健康指标（如油压、电池电量、刹车状态等）
- **数据来源：**车辆诊断系统
- **约束条件：**实时监控，数据更新频率不低于 1 分钟
- **用途：**用于车辆健康评估，及时进行维护和修复

6 功能需求

6.1 功能划分

本节详细描述了无人驾驶网约车多模态交互系统的功能需求，包括：

1. 多模态交互整合
2. 车内环境与状态监控
3. 智能决策与优化
4. 系统数据与日志管理
5. 系统维护与优化

6.2 功能描述

6.2.1 多模态交互整合

智能交互接口：

通过语音、手势、触控、眼动追踪等多模态输入方式，提供灵活、直观的用户交互体验。系统能够根据环境噪音、乘客的需求和上下文场景智能切换交互模式。例如，在噪声较大的环境中，系统会自动提升语音识别灵敏度，或者提供手势和触控的交互选项，确保乘客能够流畅地与系统进行互动。

个性化反馈与情绪感知：

利用面部识别、眼动追踪和语音语调分析等技术，实时感知乘客的情绪状态并动态调整交互策略。如果检测到乘客表现出焦虑、疲惫或不适，系统会自动调整车内环境（如温度、座椅角度）或通过温和的语音风格与乘客互动，以提升乘客的舒适度。同时，系统能够根据历史交互数据提供个性化的服务推荐，例如娱乐内容、导航建议等，进一步优化乘客体验。

6.2.2 车内环境与状态监控

车辆健康与状态监控：

系统持续监控车辆各项健康参数，如传感器状态、车内设备、电池电量、车内温度等，确保车辆运行平稳。通过实时数据反馈，车辆管理人员可即时获知车辆的健康状态，及时发现和处理潜在的故障问题。当系统检测到异常情况时，会自动启动故障诊断流程，并通知相关维护人员进行修复。

应急响应机制：

根据实时交通和环境数据，系统能够预判可能的突发事件（如障碍物、交通堵塞等），并及时调整车辆的行为（如减速、避让）。在突发紧急情况下，系统会立即向乘客发出安全警告，并提供应急反馈，帮助乘客做出正确反应，保障乘客的安全。

6.2.3 智能决策与优化

多模态指令与优先级管理：

系统根据不同输入方式（语音、手势、触控等）智能地设置指令的优先级。在多个指令同时发出时，系统会根据当前驾驶场景、交互模式和安全性进行优先级调整，确保驾驶安全和乘客需求的高效响应。例如，系统会优先处理与安全相关的指令，如停止、刹车等，而对一些低优先级的操作（如调整座椅）则延后处理。

驾驶场景自适应与反馈：

系统能够根据不同驾驶场景（如城市街道、高速公路、复杂交通等）自动调整驾驶决策。根据车辆传感器收集的环境数据（如交通信号、车流量等），系统优化行驶策略并及时向乘客反馈。例如，在高速公路上行驶时，系统会自动增强语音提示音量和清晰度，以确保乘客在较大车速下也能清晰听到指令和反馈。

6.2.4 系统数据与日志管理

数据收集与共享：

系统通过各类传感器收集车辆状态数据、乘客行为数据和环境信息，并与平台实时共享，确保平台能够进行实时分析和优化。乘客的个人数据（如语音指令、手势习惯等）将经过隐私保护处理，用于系统优化和个性化服务。例如，系统能够根据乘客的习惯自动调节车内环境，或推荐乘客可能喜欢的娱乐内容。

日志记录与分析：

系统详细记录所有用户交互行为、车辆状态以及异常事件，并生成日志以供后期分析。管理员可通过日志发现潜在的系统优化点，进行故障排查或对系统性能进行评估。日志数据还可为未来的系统更新和改进提供依据，帮助提升整体服务质量。

6.2.5 系统维护与优化

自动更新与维护：

系统定期进行自动更新和维护，确保乘客和管理员能够体验到最新功能与安全修复。更新不仅包括系统性能的提升，还会修复已知的安全漏洞和故障。当系统检测到故障或异常情况时，会推送紧急修复程序，保障系统的持续稳定运行。

平台监控与分析：

管理员通过平台进行集中监控，可以实时查看每辆车的运行状况，包括车辆的状态、故障报告和异常事件。平台还提供远程诊断功能，帮助管理员及时发现和处理车辆问题，确保每辆车都能在最佳状态下运营。同时，系统会根据收集到的数据持续优化车辆调度和管理策略，提升运营效率。

7 创新功能

情绪驱动的智能环境调节功能利用多模态技术分析乘客的情绪状态，并据此动态调整车内环境，以增强乘客的舒适度和个性化体验。系统通过结合**面部识别**、**眼动追踪**、**语音分析**等多种数据源，实时感知乘客的情绪变化。这些数据不仅有助于精准判断乘客当前的情绪状态，还能有效预测他们的需求。

当系统通过语音分析或面部识别检测到乘客情绪处于焦虑或紧张状态时，车内环境会自动调整。例如，车内温度会适时降低，座椅角度会根据乘客的舒适度需求进行调整，车内音乐则会切换为更加舒缓的旋律。此外，系统还会监测车内空气质量，如若发现空气质量较差，便会自动增加车内的空气清新度或者启动车载香氛装置，以帮助乘客放松身心。

当乘客表现出疲劳的迹象时，系统会通过眼动追踪或面部识别检测到这一点，并自动调整座椅的角度，提供更为舒适的休息状态，同时播放轻柔的背景音乐来缓解乘客的疲劳感。对于情绪愉悦的乘客，系统则会提升车内的舒适感和活力，例如播放动感的音乐，提升空气清新度，让乘客的体验更加愉悦。

通过实时感知乘客的情绪并根据其变化对环境进行自适应调节，情绪驱动的智能环境调节功能能够在不需要乘客主动操作的情况下，提供一种无缝的个性化体验。这种情绪与环境的自动匹配，使得

乘客不仅享受更加贴心的服务，还能体验到一种前所未有的智能交互。

这种技术的创新之处在于，不仅根据车内外环境变化来调节设置，更重要的是结合乘客的情绪波动进行个性化调整。这一功能突破了传统车载环境调节系统的局限，让系统能够动态地响应每个乘客的情绪需求，创造一个更加舒适、自然且智能的乘车体验。

8 性能/非功能需求

本节描述了无人驾驶网约车多模态交互系统在性能和非功能方面的需求，包括系统的准确性、及时性、可扩充性、易用性、易维护性、标准性和先进性。

8.1 准确性

系统需要具备高度的准确性，特别是在语音识别、手势识别、眼动追踪和情绪感知等多模态交互环节中。以语音识别为例，系统在安静环境下的识别准确率应达到 98%，在典型车载噪声背景下仍应保持在 92% 以上。通过引入预训练大语言模型，语音意图识别准确率应在 95% 以上。手势识别和眼动追踪模块在日间自然光环境下的识别准确率为 95%，在夜间或光照复杂情况下保持在 90%。在涉及安全的决策中，例如紧急制动与避障操作，系统应将误判率控制在 0.1% 以下，确保驾驶安全。

8.2 及时性

为了保证良好的用户体验和行车安全，系统需具备毫秒级的实时响应能力。语音指令的平均响应延迟应小于 150 ms，手势识别响应时间小于 200 ms，系统对突发状况（如行人横穿）做出预警的最大响应时间不超过 300 ms。系统采用边缘计算和异构计算资源调度机制，在高负载场景下仍能保证 90% 以上的交互请求延迟低于 250 ms。此外，决策模块采用并行推理框架，处理高并发事件时每秒可处理指令超过 1000 条，确保系统在实际道路环境中的高效运行。

8.3 可扩充性

系统的设计应支持功能和硬件的灵活扩展。随着技术的发展和用户需求的变化，系统需具备快速集成新交互方式和硬件设备的能力。例如，系统架构采用模块化设计，支持多达 12 种主流传感器接口（如 LiDAR、RGB-D、红外摄像头）及 10 种以上主流交互模块的即插即用。系统支持 OTA（Over-the-Air）远程更新，每次平均更新包体积小于 150 MB，更新成功率达 99.3%。用户可在不更换核心硬件的前提下，部署最新的多模态融合模型，节省 80% 的升级成本。

8.4 易用性

系统界面和交互方式需直观友好，确保不同年龄和技术背景的用户都能轻松上手。多模态交互应提供自然、流畅的体验，系统根据用户的使用习惯和环境自动推荐最合适的交互方式。对于突发事件或紧急状况，系统应提供简化的交互路径，减少用户在紧急情况下的操作负担。针对视障或听障用户，系统还需提供辅助功能，如语音转文字、文字转语音等，确保无障碍交互。系统交互设计经用户测试后，整体易用性评分应达到 4 分（满分 5 分）。

8.5 易维护性

系统需提供便捷的维护和管理功能，包括远程监控、故障诊断和自动更新。通过实时的系统健康监测，运维人员可以快速定位并解决 90% 以上潜在故障。此外，系统应具备自动化运维工具，支持批量更新和维护，支持同时管理超过 10,000 台设备的状态信息，减少人工干预。在异常情况下，系统需具备自我修复能力，尽可能降低对人工介入的依赖。

8.6 标准性

系统的开发和设计应遵循行业标准和相关法规，确保与不同硬件和平台的兼容性。系统接口需采用标准化的通信协议，便于第三方服务和设备的集成。特别是在数据传输和存储方面，系统需符合隐私保护和数据安全的相关标准，确保用户数据的安全性。例如，系统设计全面遵循 ISO 26262、ASIL-B 安全等级标准，数据接口符合 SAE J1939 和 AUTOSAR 通信协议。

8.7 先进性

为保持行业竞争力，系统应采用最新的人工智能和多模态交互技术。通过持续的算法升级和模型训练，提升交互的自然度和智能化水平。系统还应具备自学习能力，通过分析用户的行为和反馈不断优化自身性能。此外，系统应预留对新兴技术的支持能力，确保未来能够快速引入前沿技术，实现技术领先。

9 系统运行要求

本节描述了无人驾驶网约车多模态交互系统在运行过程中所需的硬件和软件配置要求，以确保系统的稳定性和高效性。

9.1 硬件配置要求

1. 计算单元

- **CPU**: 至少 8 核心，主频 2.5 GHz 以上。
- **GPU**: NVIDIA RTX 4090 级别以上或等效 GPU，以支持 AI 推理任务。
- **RAM**: 16 GB DDR4 或以上，以应对多模态数据处理和深度学习推理。
- **存储设备**: 512 GB NVMe SSD 或以上，用于快速数据读写和存储交互日志。

2. 感知设备

- **摄像头**: 高清 RGB 摄像头，分辨率至少为 1080p，帧率不低于 30 FPS。
- **红外摄像头**: 用于夜间或低光环境中的手势识别和眼动追踪。
- **麦克风阵列**: 至少 4 个定向麦克风，用于实现远场语音识别和噪声抑制。
- **触控屏幕**: 12 英寸或以上，全高清分辨率，支持多点触控。

3. 通信与网络

- **网络模块**: 支持 5G、4G LTE、Wi-Fi 6 和蓝牙 5.0。
- **CAN 总线**: 用于与车辆控制系统通信。
- **GNSS 模块**: 支持 GPS、北斗等卫星导航系统。

4. 电源与散热

- **供电系统**: 配备备用电池组，支持至少 30 分钟的紧急供电。

- **散热系统**：采用高效散热模块，确保在高负载情况下维持系统稳定运行。

9.2 软件配置要求

1. 操作系统

- Ubuntu 20.04 LTS 或以上版本，支持 GPU 驱动和 CUDA 环境。
- 兼容 ROS (Robot Operating System) 2 进行传感器数据管理和通信。

2. 深度学习框架

- PyTorch 2.0 或 TensorFlow 2.12，用于模型训练与推理。
- 支持 TensorRT 进行模型优化与加速。

3. 语音识别与处理

- Whisper 或 Kaldi，用于多语种语音识别。
- 使用 Voice Activity Detection (VAD) 进行噪声抑制。

4. 图像与视频处理

- OpenCV 4.8 或以上版本，用于图像预处理与特征提取。
- 使用 NVIDIA CUDA 和 cuDNN 进行加速。

5. 数据管理与存储

- 使用 PostgreSQL 或 MySQL 进行结构化数据存储。
- Prometheus 或 ELK (Elasticsearch、Logstash、Kibana) 用于数据监控与日志分析。

6. 通信与网络协议

- 支持 MQTT、HTTP/HTTPS、WebSocket 等协议。
- 使用 gRPC 或 RESTful API 进行跨模块通信。