人形机器人智能驾驶：概念、方法、现状与展望

季旭东1，薛云涵2，余贵珍3，陈鹏3，魏洪兴3，周彬1，张俊杰1

（1．北京航空航天大学合肥创新研究院，合肥，

2．索邦大学，巴黎

3．北京航空航天大学，北京，100191）

摘要：人形机器人智能驾驶融合了机器人学、智能感知、认知决策和自动驾驶等多个前沿领域，旨在使具有人体结构和动作能力的机器人替代或协同人类完成复杂交通驾驶任务。该技术突破了传统自动驾驶中对特定硬件结构和规则模型的依赖，具备更高的人机交互适应性与场景通用性。本文综述了人形机器人智能驾驶的基本概念与发展动因，系统归纳了其核心关键技术，包括多模态感知融合、类人驾驶行为建模、任务规划与协同控制机制，以及仿真与实物验证平台。在分析当前国内外研究现状的基础上，指出该领域在动作协调性、控制精度、实时性与安全保障方面仍面临诸多挑战。同时，本文展望了人形机器人在智能交通、特种作业与人机协同等关键场景中的广阔应用前景，强调了融合人工智能与机器人动力学建模、多任务学习与决策自适应策略等方向的重要发展趋势。本综述旨在为人形机器人智能驾驶系统的研究者和工程实践者提供系统性的参考与思考。

关键词：人形机器人；智能驾驶；协同控制

**Abstract**

.

# 引言

随着人工智能、机器人技术与智能驾驶领域的深度融合，传统以人类为驾驶主体的模式正逐步向自动化、智能化转变。然而，现有的自动驾驶系统仍面临着在非结构化环境中自主决策、复杂人机交互、以及极端工况下应急处置等方面的挑战。在此背景下，人形机器人智能驾驶（Humanoid Robot Autonomous Driving）作为一种具有生物结构适配性与通用执行能力的新型驾驶模式，正受到国内外研究机构与产业界的广泛关注。

人形机器人具有人体相似的骨骼结构、末端执行器和自由度配置，具备通过“仿人操作”执行方向盘控制、油门/制动踏板操作、档位切换等物理驾驶任务的潜力。相较于基于线控底盘的无人车或固定结构的自动驾驶平台，人形机器人智能驾驶系统可以直接“接管”传统车辆，无需改装，具备替代人类完成驾驶任务的优势，具有更强的跨平台适应性、可移植性和对突发情况的人类行为模仿能力。

当前，自动驾驶技术在高速公路、封闭园区等规则明确的环境中取得了显著进展，但在结构复杂、动态变化频繁的城市道路、灾害场景或军用任务中，依赖规则库与深度模型的无人车系统依然面临诸多瓶颈。传统自动驾驶算法在面对突发状况、人车协同互动及非结构化语义理解等方面表现出有限的鲁棒性与灵活性。而人形机器人具备类人认知结构、物理操作能力和交互潜力，以其类似人类的操作方式和环境适应能力，成为智能系统与物理世界之间的重要桥梁，被认为是未来“智能驾驶—智慧交通—泛在机器人”融合发展的关键支撑技术之一。

尽管近年来已有部分研究关注“人形机器人驾驶”的可行性验证和系统集成，如基于运动规划控制的机器人操控方向盘试验、仿人脚步控制算法在踏板操作中的应用等，但整体而言，该领域仍处于早期探索阶段，在控制精度、系统集成性、交互智能与安全冗余等方面尚未形成成熟体系。因此，有必要从人形机器人智能驾驶概念、关键技术、现状与问题、已有研究基础与未来发展方向出发，对人形机器人智能驾驶的研究现状进行全面梳理和深入分析。

# 人形机器人智能驾驶概念

## 人形机器人

人类对机器人的探索与想象早在20世纪就开始萌芽。剧作家卡雷尔·恰培克 [1]在1920年的剧作中第一次提出了“机器人”（Robot）这个词，这一创造也成为该术语的起源。1950年，人工智能之父艾伦·图灵 [2]在其论文《计算机器与智能》中，前瞻性地提出了“具身智能”（embodied intelligence）的概念。他强调，智能行为需要通过相应形态的智能体（即具备身体的机器人）与环境互动习得。这一理论为后来人形机器人的发展奠定了重要的思想基础。与此同时，随着第三次工业革命的浪潮的蓬勃兴起与深入发展，特别是电子计算机处理能力的飞跃和新型工程材料的涌现，克服了早期在计算、控制和物理结构上的诸多限制，为机器人正式从探索性原型走向实际应用提供了关键支撑。理论与技术基础的双重具备，正式开启了人形机器人的发展历程。

1972年世界上第一款全尺寸人形智能机器人WABOT-1 [3]诞生，他能够搬运物体，拥有基础的感知系统（如视觉、听觉），但功能有限。现代人形机器人的发展由此进入起点阶段。从1986起至2010年，以日本公司为主导，人形机器人发展成了配置感知、控制、执行等子系统的集成机械装置。ASIMO [4]（本田）能够跑步，上下楼梯，在运动控制和动作协调方面取得明显进步；NAO [5]（软银）能够通过语音、视觉等方式与人类交流，具备基础的人机交互能力。这一阶段为后续的高动态运动和智能化发展奠定了基础。

近年来，伴随着人工智能、机器学习、大模型技术等领域的发展，通过深度学习、自然语言处理、图像感知识别等技术，人形机器人能够更好地理解和适应复杂环境，执行更高层次的任务。例如，波士顿动力的Atlas [6]能够完成后空翻等高复杂度的动作；特斯拉的Optimus [7]展示了人形机器人在家庭和工业场景中的应用潜力。同时，随着技术的成熟和成本的降低，人形机器人正逐步从实验场景走向现实应用，进入商业化与产业化阶段。它们将会被部署于家庭服务、医疗护理、教育辅助、工业协作等多个领域，展现出广泛的应用前景。

与此同时，智能驾驶作为机器人应用的前沿领域之一，也正在发生深刻的演变。相比于单纯的智能驾驶，人形机器人优势在于其对人类环境的天然适配、具备多任务能力和逐步增强的智能水平，这使得它们在复杂场景（如物流、安保、救援）中比传统自动驾驶系统更具灵活性，同时还具有更强的任务迁移能力和泛化能力。探索人形机器人与智能驾驶技术的融合，有望为复杂环境下的自动驾驶提供全新的解决方案。

## 智能驾驶

为了更全面理解人形机器人在驾驶场景中的潜力，还需回顾当前智能驾驶系统本身的发展脉络与技术体系。早在20世纪初期，智能驾驶（或称自动驾驶）概念就被提出。1925年，美国工程师弗朗西斯·P·霍迪纳在纽约曼哈顿街头完成无人驾驶演示。尽管本质是无线电遥控另一辆汽车，但首次将“无人驾驶”从科幻带入现实。1939年，在纽约世界博览会上，通用汽车首次提出“自动高速公路系统”（Automated Highway System, AHS） 的完整构想。该方案设想通过预埋电子轨道与车辆协同，汽车能在特定高速公路上自动行驶，奠定了车路协同（V2X）的技术雏形。

第一部真正自动化的汽车到1980年代才首次出现。在1984年，卡内基美隆大学推动Navlab与ALV计划，并在1987年实现了昼夜越野驾驶。Navlab 在1995年驾驶车辆穿越美国，其中98%的路程实现了自动驾驶。1998年，王飞跃教授团队研发的VISTA 自动驾驶车辆在凤凰城51号高速公路上进行了公众演示，成为国内较早探索该领域的先行者之一。他们在实现过程中引入了“代理系统”（agent-based systems）理念，进一步演化为“平行驾驶”框架 [8]。该理论首次于 2005年在西安举行的国际汽车电子与安全大会上被正式提出，主张通过网络代理构建虚拟驾驶员，参与车辆控制与决策过程，仅保留核心决策在车内执行。思想随后在 DVPG 设计与 VISTA 系统的演进中得到深化，并作为关键技术之一，纳入国家863计划中相关课题，显著推动了我国无人驾驶技术的快速发展。

21世纪以后，随着人工智能与通信技术的突破性进展，智能驾驶技术迈入快速发展阶段。在2005年第二届DARPA无人驾驶挑战赛中，有五辆车成功完成了212千米的赛道路程，这一突破在全球范围内掀起智能驾驶技术研发的热潮。此后，中国自主研发的智能驾驶技术也加速追赶。自2009年起，国家自然科学基金委员会启动了“中国智能车未来挑战赛”，并将其打造为年度行业盛事。挑战赛围绕开放式环境下的无人驾驶任务，针对性设计多维度测试场景，为测试智能车辆在接近真实交通条件下的能力提供了实验平台。2015年8月，常熟市政府携手西安交通大学、中国科学院自动化研究所、长安大学及青岛智能产业技术研究院，共同在常熟高新技术产业开发区建设了“中国智能车综合技术研发与测试中心”。不仅填补了国内在无人驾驶领域专业测试基地上的空白，也标志着中国在智能车实地验证平台建设方面迈出了关键一步。

谷歌在2009年1月开始开发自动驾驶技术，经过近两年的道路测试，于2010年10 月揭晓了首个车规级自动驾驶系统Project Chauffeur。它为车辆集成激光雷达等设备，旨在实现全自动驾驶能力的技术验证。2015年，特斯拉以差异化技术路线切入市场，推出特定环境下的高级驾驶辅助系统Autopilot，基于视觉、雷达和超声波传感器融合实现自动驾驶、自动转向等功能，推动了辅助驾驶功能的普及。2017年，百度将技术应用场景拓展至公共交通领域，成功推出无人驾驶公交Apolong，在深圳两条公交线路上进行试运行后于2018年量产，在上海世博会上向国际观众开放。在2021年，清华大学智能产业研究院联合百度Apollo，共同推出全球首份车路协同技术创新白皮书《面向自动驾驶的车路协同关键技术与展望》 [9]，不仅为车路协同技术提供理论参考，也为自动驾驶规模商业化提出了建议。2023年，华为发布了新一代智能驾驶解决方案ADS2.0，以超大规模云端仿真和数据挖掘训练为基础，实现从城市街道到高速公路等多元化的场景的覆盖。随着大模型时代的到来，不断完善的技术水平标志着高阶自动驾驶正加速向实际应用场景渗透。

随着全球范围内各类自动驾驶技术方案不断涌现，行业亟需一套统一的分级标准对不同程度的自动化水平进行明确划分。为此，国际汽车工程师学会（SAE） [10]于2014年发布了自动驾驶六级分类系统作为目前广泛采用的技术分级框架。根据该分级标准，L0表示完全由人工驾驶，L1–L2属于驾驶辅助阶段，车辆可在部分功能上实现自动控制，如巡航控制或车道保持，但驾驶员必须保持注意力并随时准备接管；L3–L4则具备一定自动化能力，属于有条件和高度自动驾驶，系统能够在限定条件（如特定路段或环境）下实现自动化操作；而 L5则代表真正意义上的全自动驾驶，无需人工干预，系统可在任何时间、任何环境下独立完成所有驾驶任务。

在实际应用过程中，辅助系统不同等级间的边界仍存在模糊地带，特别是在城市交通、恶劣天气以及非结构化环境等复杂场景中，系统性能的可靠性与安全性仍未完全达到预期水平。随着感知、决策、控制等技术的不断进步，智能驾驶技术已广泛应用于高速公路、园区物流和城市道路等多种场景，并正在逐步向更高级别的自动化演进。与此同时，现有的智能驾驶系统在面对复杂、多变和高度非结构化的交通环境时仍面临感知不完备、交互不确定、决策泛化能力不足等挑战。在这一背景下，具身智能体的研究逐渐受到关注。尤其是人形机器人，作为兼具类人感知结构、多自由度动作能力和语义交互的集成体代表，被视为未来智能交通系统中有望实现“统一感知—理解—执行”闭环的关键节点。

## 人形机器人智能驾驶

人形机器人智能驾驶，指的是具备类人形态与运动能力的机器人，模拟人类驾驶行为，通过物理交互方式操控现有车辆，实现自主或半自主驾驶任务的过程。与传统自动驾驶系统依赖高精地图、端到端建模不同，人形机器人智能驾驶强调“具身执行”与“结构兼容”，无需对车辆进行专门改造，便可在原生环境中完成驾驶任务。

2012 年，由美国国防高级研究计划局（DARPA）发起的机器人挑战赛（DRC） [11]成为人形机器人研究的重要转折点。这项挑战赛的背景源自2011年日本福岛核事故，此次事故暴露显著出机器人在核泄漏应急控制中的功能性短板。在高温高辐射等灾难场景下机器人作用非常有限，促使DARPA希望通过赛事推动机器人在复杂事故环境中的实用能力。挑战赛要求仿人机器人能够代替人类在灾难环境中执行任务，如利用工具打破墙壁、爬上梯子、在复杂的地形中穿行等紧急响应操作，其中，人形机器人的驾驶能力成为DRC的关键考核方向之一。

正如Paolillo等人（2017） [12]在其研究中指出，灾害救援现场常常涉及对传统人类设备（如挖掘机、钻机、运输车等）的直接操控，这些设备由于设计复杂、结构多样，难以实现统一的自动化改造。但人形机器人在驾驶类任务中展现出显著优势。相较于传统自动驾驶系统，人形机器人可以直接适配现有车辆控制接口，规避了高成本的车辆重构需求，具备高度的场景泛化能力和部署灵活性。

人形机器人智能驾驶系统的实现需要构建一个高度集成、多层次协同的技术架构。其架构可分为四个层级，逐层递进实现从环境感知到自主决策的完整链路。

1. 环境感知与理解层​。相较于传统智能驾驶系统侧重车载外部感知，人形机器人需同时关注车辆内部与外部环境。其中内部包括驾驶舱内仪表盘和座椅等位置的识别、机器人如何正确定位座椅并维持坐姿、加速踏板踩空后的复原；外部则涉及到构建周围道路环境、识别和跟踪道路上的各类目标等。机器人通过多模态传感器（集成单目/双目摄像头、激光雷达、毫米波雷达等）融合构建动态环境的全局认知，为上层决策提供实时、可靠的环境模型。
2. 任务决策层​。其核心任务是基于多源感知信息和环境状态，结合预设任务目标，为机器人生成具备可执行性的行为策略。该层需要以模块化架构整合语义理解、意图推理与任务管理逻辑，综合考虑视觉、触觉、雷达信号等多种信息，协调各子系统的工作优先级，在安全性与效率之间进行动态权衡，确保安全、高效的行驶路径。
3. 运动规划与控制层。该层是连接任务决策层与机器人本体执行机构的桥梁。在这层中，核心任务是将高层驾驶决策转化为对机器人本体（包括机械臂、躯干、头部等）的低层控制指令，具体包括路径生成、姿态调节、抓握与操控动作的协同控制等。在驾驶任务中，机器人则需完成机械臂对方向盘的精准操控、脚部对加减速踏板的协同控制，以及躯干姿态的动态维持等指令。此类多维度操作指令的协同执行，对机器人身体结构的稳定性与自由度提出了要求和挑战。此外，还需兼顾车体震动、地形坡度等扰动因素，确保控制的鲁棒性和驾驶的安全性。
4. 自主学习与适应层​。为了面对真实环境中多样化的车辆类型与非结构化道路条件，人形机器人需要具备长期自主学习和快速适应新环境的能力，确保在突发复杂交通场景中仍能完成稳定操作。在与车辆和环境交互的过程中，通过模仿学习、强化学习或少样本学习等机制，在线或离线地优化感知—决策—控制链路中的关键参数，提升对未知环境的泛化能力。当机器人遇到未知车型或界面布局变化时，系统应能基于历史经验快速调整操作策略，保证驾驶行为的流畅性和安全性。

# 人形机器人智能驾驶关键技术

人形机器人智能驾驶作为一个复杂的多学科交叉领域，涉及感知、决策、执行和系统集成等多个技术层面。基于现有研究成果，本节将从感知技术、执行技术和系统集成平台三个维度，深入分析人形机器人智能驾驶的关键技术体系。



## 人形机器人驾驶感知

感知是人形机器人智能驾驶系统的核心支撑模块，其感知模块的设计逻辑与传统自动驾驶系统存在本质差异——后者以"车辆为中心"构建环境感知框架，而前者需以"类人驾驶员"为参照，满足环境理解深度与交互功能拓展的双重要求。一方面，需要实现对车辆状态、道路环境、交通参与者等多维度信息的实时获取和理解；另一方面，需额外构建车内空间感知子系统，通过视觉、触觉等多传感器协同，实现对仪表板、控制装置的精确定位和交互。

### 多模态信息融合

人形机器人在执行驾驶任务时，需要整合来自视觉、力觉、惯性和触觉等多模态传感器的信息，实现对环境与车辆状态的统一认知。Paolillo等人（2017） [12]研发的HRP-2机器人通过单目摄像头实现对道路边界和车道线的识别，并采用光流法结合IMU数据估算自身与车辆的运动状态，为路径规划和决策提供基础数据。Rasmussen等人（2014） [13]则配置了前向摄像头和额外的车内摄像头，用于识别车辆内部控制装置的位置和状态（如车内方向盘、踏板等），并集成多点力传感器感知接触强度，实现触觉交互。更先进的系统已突破单一模态限制，如东京大学研发的Musashi [14]，搭载了多样的触觉与压力传感器网络，融合力觉、位置、视觉与内部状态感知能力，实现从环境接触（如抓取物体时的触觉反馈）到动作调整（如根据压力分布改变握力）的自然运动闭环。这种多模态融合的感知层级设计，为机器人提供了“接触-感知-调整”的类人交互决策路径。

### 动态三维感知

动态三维感知技术旨在实时获取车辆内部及道路环境的三维结构信息，为人形机器人在驾驶过程中的空间定位与避障提供关键支持。与固定位置的传感器不同，人形机器人作为具身智能体，在车辆行驶过程中需不断感知并适应环境变化。为了实现精准感知，机器人需在车辆坐标系中定位自身，同时解析周围三维环境的几何结构，例如踏板、方向盘等关键物体，进而构建点云模型以实现精确对齐。在Rasmussen等人（2014） [13]提出的系统中，通过两个相机提供的数据信息，机器人可以简历车辆内部布局的空间模型并计算机器人相对位置，从而动态调整坐姿与操作策略。然而，在具身智能驾驶场景下，机器人还需面对额外的动态感知挑战。由于道路颠簸等路况影响，人形机器人的身体运动会引入额外的动态因素带来感知偏移，需要通过姿态估计、坐标变化和运动补偿等机制保证机器人的感知精度。例如在光照不足或强烈抖动的情况下，通过滤波、图像增强等方法提升视觉感知稳定性，或优先采信稳定性更高的传感器通道，以保障整体感知的可靠性。

### 领域自适应

不同于自动驾驶本身，传统智能驾驶系统通常聚焦于固定车型的适配，而人形机器人智能驾驶系统需要面对跨车型，跨环境的实验场景——这一特性既是其区别于传统驾驶的核心优势，也构成了任务面临的主要挑战：泛化能力。因此，具备领域自适应能力是实现通用驾驶的关键。在面对不同车辆品牌、操作布局、部件尺寸的情况下，感知系统需要快速学习和适应新的车型配置，建立和具体操作的映射关系。Rasmussen团队 [13]提出的机器人在两种不同的车型上进行了实验，它能识别出是哪一类型的车辆，并依靠仪表盘的位置对自己定位，以便计算机器人动作与方向盘等部件交互。Musashi系统则得益于其肌肉驱动系统的类人结构与仿人体关节的骨骼框架，借助分布式传感器网络的配置，自身形态就具备较强的环境适应弹性，体现了一种“结构即能力”的自适应设计思路。

## 人形机器人驾驶执行

人形机器人驾驶执行模块负责将感知结果转化为具体的操控动作，包括对方向盘、油门踏板等的物理操纵与动作控制。

### 类人行为驾驶建模

在传统的智能驾驶系统中，驾驶任务常常被建模为轨迹规划与路径跟踪问题，包括路面情况、其他交通参与者等。而在人形机器人系统中，驾驶行为则需要被映射为符合类人身体逻辑结构的动作序列，不仅包括基本的双臂抓握、转动方向盘，还有油门加速踏板的踩踏、驾驶室位置的姿态调整等序列动作。HRP-2机器人目前的驾驶行为就可以被分解为对方向盘操作的建模、对踏板操作的建模和驾驶姿态的建模。Paolillo等 [12]构建了一个传感器驱动的控制框架：通过视觉画面计算出的结果生成方向盘参考角度，卡尔曼滤波估计车速并计算油门指令，使机器人执行转向和加速操作。机器人按照规划好的动作序列调整姿态，其手脚能够精确接触踏板和方向盘。类似地，东京大学研发的 Musashi机器人 [14]被设计为支持“用两臂同时操作方向盘、油门踏板控制、维持驾驶坐姿等多动作协同任务。其肌肉骨骼系统采用仿生驱动设计，通过模仿人体肌肉的收缩特性，能够根据任务需求调节刚度。其类肩关节部位具有在肋骨笼上滑动的肩胛骨结构，能够实现大范围的运动，提高控制的连续性和稳定性，初步实现了结构上的类人行为驾驶建模。

### 任务规划与协同控制

人形机器人在驾驶任务中，需要对整车的任务进行规划，并实现机器人各部分、机器人和人之间的协同控制。系统不仅要完成单一动作，还需具备多通道并发任务调度与控制能力，例如在转向过程中同步调整踏板输入、维持身体稳定姿态等。

Paolillo等人 [12]的框架提供了三种驾驶策略：远程操作、机器人全自主驾驶和人机混合控制。人类监督员根据需要，可以选择让机器人承担全部或部分驾驶任务。这种设计本质上是一种任务级的协同控制策略，保证人类可以随时接管驾驶，从而提高任务的安全性与灵活性。Sohn团队 [15]开发的辅助系统使操作员能够根据当前车辆状态预测出车辆路线，基于避障和行驶距离计算出最优路径，并在用户控制界面上显示。他们还对比分析了不同操作模式下的驾驶模式（如完全人类驾驶、完全远程操控），以评估该控制策略的表现，并在真实道路环境下驾驶车辆验证效果。

Rasmussen 等 [13]构建了一个任务分级控制架构，分为低层动作控制（如关节位置与速度）、中层运动模块（如转方向盘、踩踏板）、高层任务调度（如启动、停车），各级模块通过共享状态信息实现协调。实验验证表明，多模态感知系统在复杂环境中的鲁棒性比单一传感器系统提高了约35%。东京大学的团队 [14]则提出一个集成多个模块的学习系统，包括感知、规划、控制等方面，可以根据Musashi机器人的柔性硬件特征执行连续驾驶动作。目前的研究来说，人形驾驶系统已初步具备人机切换和闭环控制能力，但任务规划仅限于简单路径与避障，尚未形成机器人多智能体协作框架。

## 人形机器人智能驾驶平台

人形机器人智能驾驶平台是整个人机驾驶系统的集成载体，承载了感知、执行、控制、通信和电源等多模块。当前研究主要基于仿真平台与实物验证平台双轨推进。

### 仿真平台

现有研究在实现人形机器人驾驶前，多采用仿真平台进行算法验证与调试。目前在人形机器人驾驶研究中，科研团队普遍借助如 Gazebo、MuJoCo等通用机器人仿真环境，开展车辆动力学模型构建、机器人动力学特性分析与传感器数据的模拟研究。在仿真环境中，研究人员可以高效验证驾驶算法的控制策略，无需依赖实体硬件反复调试。仿真平台通过虚拟环境复现真实驾驶场景，降低了物理测试的硬件损耗、场地占用等成本，并减少了实际操作会产生的设备碰撞等风险，为后续的实物验证提供了可靠的性能预测依据。在Rasmussen [13]的实验中，在实物测试转向和速度控制的方法之后，团队收集了五公里以上的校园环境和高尔夫球场数据在Gazebo仿真平台中进行算法验证，模拟测试了障碍检测、环境建模、路径规划和轨迹跟踪等方面的准确性。

在实际研究中，也有许多团队采用多平台集成的方式，在不同阶段使用不同的仿真平台。例如，在进行机器人运动和控制算法的调整阶段，MuJoCo仿真平台能够精确模拟机器人手臂与方向盘的接触，足部与踏板的交互等复杂的接触动力学过程。在测试视觉感知算法和驾驶决策系统时，CARLA内置的多样化的车辆模型、行人、交通信号灯等元素可以进行复杂的交通场景模拟。这种多平台协同的方式能够充分发挥各平台的优势，提高研究效率和结果的可靠性。

### 实物验证平台

实物验证平台是推动人形机器人驾驶系统落地的关键，在真实环境中测试可以检验系统的感知鲁棒性、执行精度与平台可靠性。目前已有多种人形机器人平台用于驾驶任务验证。Paolillo等人 [12]在实车驾驶实验中依托日本AIST开发的HRP-2Kai人形机器人，在园区内真实车辆上完成了自主驾驶演示。值得关注的是，该团队基于在DRC决赛中驾驶任务的实验经验，进一步证明了该系统共享自主的驾驶模式比完全远程操作更有效。

Sohn等团队 [15]在DRC机器人挑战赛中使用了升级版的DRC-Hubo+机器人，对其本体和感知头进行了改进。引入驾驶辅助系统后​​，经过系统训练的操作员配合完整辅助功能，车辆在躲避障碍时决策速度大幅缩短。

Kawaharazuka团队 [14]则基于日本WARI实验室的Musashi肌肉驱动机器人展开研究，实验选用了超小型电动汽车COMS（Chotto Odekake Machimade Suisui，意为“短途便捷代步车”）系列，实验场地位于东京大学柏校区。在该小车的后备箱内集成了实验所需的元件，包括电源、Wi-Fi路由器等。作为电动车，未来人形机器人的所有电力均可通过COMS车型直接供给。

相比于对车辆硬件进行深度改造，上述平台的共同特点是无需为车辆加装额外元件即可执行驾驶操作，这一特性直接提升了系统的跨平台通用性。未来的研究可进一步改进人形机器人用于驾驶的硬件设计（如轻量化手腕、可自适应踏板接触等）和人机协同策略，以提升安全性和操作效率，实现更加完善的智能驾驶功能。

# 现状与问题

当前，人形机器人驾驶作为无人驾驶与人形机器人技术交叉的研究方向，已在可控环境下实现了基本驾驶任务的执行，显示出初步的可行性。比如，机器人可基于摄像头提供的数据使车辆沿道路中心行驶，完成加速和转向的操作；拥有先进结构的肌肉骨骼机器人能利用自身结构的优越性，在双臂操作方向盘和维持驾驶姿势上具有灵活的操作性能。然而，这些研究成果多在理想化或简化场景（如封闭园区、固定路线）中验证，与实际道路驾驶场景的复杂需求仍存在显著的环境适应性差距。

进一步来看，这种环境适应性的不足，本质上源于人形机器人在核心能力上的短板。首先，人形机器人在机动性上与人类驾驶员仍具有较大差距。尽管人形机器人在外形结构和动作模式上模拟人类，但由于机器人本身结构复杂，受限于当前硬件技术的发展，多数人形机器人仍采用刚性驱动系统。这种结构的机器人关节运动范围有限，在面对复杂驾驶操作（如方向盘大角度快速回转、微调刹车深度等）时，稳定性和灵活性不足。即使是具有优越结构的肌肉骨骼机器人Musashi，团队在方向盘操作中仍面临诸多挑战。与人类实际驾驶时常采用的交叉臂转向方式不同，该系统目前仅能执行推拉式转向，难以完成大角度动作，且在操作过程中易发生机械结构之间的干涉与碰撞。此外，非抓握式的接触姿势导致其手指结构频繁被方向盘卡滞，必须进行人工复位。在下肢控制方面，目前几乎所有的机器人都只实现了单一踏板（油门）的控制，腿部结构对不同车型踏板高度的适应能力较弱。同时，在颠簸路况下，脚部的自主复位功能还需要进一步测试来保证稳定性。

感知系统作为人形机器人实现自主驾驶的前提，其能力直接决定了整个系统对外界环境的认知精度与反应时效。大部分系统在对外部环境的观察上仅使用单目摄像头（或简单传感器组合）进行感知任务，这种配置不仅缺少在夜景状况下的数据支持，在遮挡频繁或信息密度高的场景中还会出现噪声大、鲁棒性低的问题。在空旷、缺少其他交通参与者的测试环境中，简单的感知机制已基本满足需求；但在实际道路应用中，单一视觉手段难以稳定识别车道线、信号灯、行人及其他交通参与者，难以胜任多车道、交叉路口等复杂场景。更为关键的是，当前多数系统尚未建立完善的动态目标预测机制，对道路环境中可能出现的突发状况反应能力有限，难以实现真正的安全与高效驾驶。

尽管已有多项研究探索了人形机器人在驾驶任务中的初步能力，但多数工作仍处于技术验证阶段，缺乏对系统性能的普遍性评估和规范性测试。在许多研究任意的论文成果中，实验结果的展示多为定性描述或少量指标，尤其在感知鲁棒性、多目标协同控制等方面，尚未建立统一的评估指标体系来量化。这种评估方式削弱了结果的说服力，同时也制约了跨平台、跨系统的性能比较与迭代优化。Kawaharazuk团队 [14]提出可以考虑“完成特定路线所需时间”作为性能的衡量指标，然而，测试路线也需要考虑多种指标来进行标准化，若缺乏结构化设计与指标体系支撑，也无法有效支撑横向评估。例如，路径中是否包含变道、掉头、行人避让等子任务；评估是否考虑在坡道或非水平环境下的操作稳定性等。另一方面，Musashi作为首例仅凭双臂在车内完成驾驶操作，不需要夹具辅助的人形机器人，展示了极高的仿生类人结构适配能力。但如何定义机器人智能驾驶的规范性，也是一项有待探索的挑战。建立标准化的测试流程、设计多场景覆盖的验证平台、推动开源评估工具和数据集的建设，是未来需要考虑到的一个重要方向。

# 思考与展望

尽管人形机器人在驾驶任务上的应用前景广阔，但要达到与人类驾驶相媲美的水平，仍需要多项技术突破。综合几位DRC机器人挑战赛参与者的研究成果可以看出，人形机器人智能驾驶目前仍处于开发与验证阶段，尚未形成通用的解决方案，在不同技术方面还存在着诸多挑战。

现有研究大多聚焦于将驾驶任务拆分为某一关键动作的实现，对完整驾驶流程中的具体环节缺乏任务规划。未来亟需发展包括SLAM（同步定位与建图）、目标识别与动态任务规划在内的全流程感知决策系统，由“动作级操控”迈向“任务级执行”，完成从上车、启动车辆，到路径规划、环境感知、指令执行的完整闭环。为实现这一目标，系统需将复杂操作流程进行模块化与系统化整合。例如，Kawaharazuka等人 [14]提出，钥匙旋转、手刹拉拽、油门踏板操作与方向盘的转动等子功能应被整合为统一系统。这一完整流程的实现不仅要求各子模块间的深度融合，还需从控制架构层面建立统一调度机制，以保障多模态传感信息、多关节动作协调与异步执行的稳定运行。

Rasmussen提到 [13]，人形机器人的操作技能需要进一步拓宽，需要涵盖力反馈系统与手眼协调视觉分析等。尽管相关功能在肌肉骨骼结构中已有初步实现，大幅提升了动作的柔性与仿生运动形态上的相似度，机器人在操作策略与执行方式上与驾驶员仍存在较大差距，难以实现真正意义上的人类操作行为还原。例如，手指结构无法完成抓握姿态、转向模式在大范围运动时存在冲突、不平坦地形或急弯路况下维持坐姿等。未来需要基于模仿学习和强化学习等方法构建类人行为建模系统，设计更适用于柔性躯体的控制方法，比如优化肩胛骨大范围运动硬件，采用平滑弹性材料等设计建议。

同时我们要考虑到机器人对不同汽车的泛化能力。大多数平台都基于特定车型优化，这种专车定制带来系统精度提升的同时，也限制了其泛化能力。除了在DRC挑战赛中应用到的标准车辆，在实际应用场景中，可能存在工程用车、左/右舵乘用车、电车等类型迥异，控制装置布局差异巨大的车辆类型。面向实用化部署的系统必须实现对不同车辆接口的自适应控制。一方面，优化多模态感知，构建车辆内部三维结构图谱，提升对方向盘、踏板等关键部件以外其他部件的快速识别与空间标定能力；另一方面，需要开发具备高自由度的上肢、躯干及腿部执行系统，以适应不同车型座舱的结构布局。此外，还可结合强化学习与迁移学习方法，加速系统在新车型上的适配过程，实现动态参数调整与控制策略迁移，从而赋予机器人“感知即适应”的能力。

现阶段人形机器人智能驾驶感知系统以图像分割、车道检测等基础任务为主。这些任务多依赖于单目视觉或立体视觉系统，主要用于辅助基本路径保持和简单的避障控制。然而，由于城市道路环境中存在多交通参与者交互，包含高动态变化场景情况，摄像头和基础传感器数据表现不稳定，这种感知方法在复杂城市场景中很难维持其鲁棒性，进一步导致智能驾驶表现不佳。这表明，仅依赖低层感知数据已难以满足对复杂环境的认知需求，而语义层理解能力的缺失正是当前感知系统的一大短板。机器人不仅需要能感知环境中各类物体的存在和位置，更需要能够识别这些物体的状态及其与机器人及周围环境的关系，进而推断潜在的行为意图和交互规则。例如，机器人需要理解交通标志的含义，识别行人是否准备穿越马路，判断其他车辆的意图（如变道、转弯），以及对交通信号灯的状态进行正确解读。这种多维度的语义信息是实现智能决策的关键前提。

人形机器人智能驾驶正处于从“实验验证”走向“系统集成”的关键转折点，需从硬件结构、规划算法、泛化能力、感知能力到评估体系等多个层面协同突破。随着AI算法、智能材料与嵌入式控制的快速演进，该领域有望成为推动机器人“进化”至现实交通系统中的核心支点。

# 文献

|  |  |
| --- | --- |
| [1] | B. M. Stableford, Science fact and science fiction: an encyclopedia, CRC Press. pp. 22–23. ISBN 978-0-415-97460-8, 2006. |
| [2] | T. A. M, Computing machinery and intelligence[M]Parsing the Turing Test. Dordrecht: Springer Netherlands 2007 23-65.. |
| [3] | “Humanoid History -WABOT-,”Available: www.humanoid.waseda.ac.jp.. |
| [4] | M. Eaton, Evolutionary humanoid robotics. Heidelberg [Germany]: Springer. p. 40. ISBN 9783662445990. OCLC 902724634., (2 February 2015). |
| [5] | A. Robotics, “A-Robots NAO,” [Archived page, August 20, 2014].. Available: https://web.archive.org/web/20140820134802/http://www.aldebaran.com/en/humanoid-robot/nao-robot. |
| [6] | J. Markoff, “Modest Debut of Atlas May Foreshadow Age of 'Robo Sapiens',” The New York Times., (July 11, 2013). Available: https://www.nytimes.com/2013/07/12/science/modest-debut-of-atlas-may-foreshadow-age-of-robo-sapiens.html. |
| [7] | “Tesla says it is building a 'friendly' robot that will perform menial tasks, won't fight back. ISSN 0190-8286.,” Washington Post, Retrieved August 20, 2021.. [联机]. Available: https://www.washingtonpost.com/technology/2021/08/19/tesla-ai-day-robot/. |
| [8] | W. X. H. S. S. e. a. SHEN Y, “ Agent-based technology in intelligent vehicles and driving: state-of-the-art and prospect[J].,” *Journal of Command and Control, 2019, 5(2): 87-98.,* pp. 87-98, 5(2) 2019. |
| [9] | “全球首份！清华大学智能产业研究院联合百度Apollo发布车路协同技术创新白皮书,”16.06.2021. Available: https://air.tsinghua.edu.cn/info/1056/1920.htm. |
| [10] | O.-R. A. D. (. Committee, Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles[M].[S.l.], SAE International, 2021. |
| [11] | T. H. Chung, V. Orekhov 和 A. Maio, “Into the Robotic Depths: Analysis and Insights from the DARPA Subterranean Challenge. 10.1146/annurev-control-062722-100728,” *Annual Review of Control, Robotics, and Autonomous Systems ,* pp. 477-502, 2023. |
| [12] | A. G. P. C. A. V. M. &. K. A. Paolillo, “Autonomous car driving by a humanoid robot,” *Journal of Field Robotics,* p. DOI: 10.1002/rob.21731. |
| [13] | C. S. K. W. Q. &. O. P. Rasmussen, “Perception and control strategies for driving utility vehicles with a humanoid robot,” 出处 *International Conference on Intelligent Robots and Systems (IROS 2014)*, Chicago, IL, USA, 2014. |
| [14] | K. T. K. K. Y. O. Y. M. T. S. K. O. M. N. Y. A. Y. O. K. K. K. &. I. M. Kawaharazuka, “Toward autonomous driving by musculoskeletal humanoids: A study of developed hardware and learning-based software,” 出处 *IEEE Robotics & Automation Magazine*, 2020. |
| [15] | K. &. J. G. Sohn, “Ground vehicle driving by full sized humanoid,” *Journal of Intelligent & Robotic Systems,* p. 99:407–425. |