



同濟大學
TONGJI UNIVERSITY

漫话自动驾驶

课程学习报告

姓 名：姚天亮

学 号：2150248

联系电话：15000476845

指导教师：王峻、刘成菊

上课时间（课号）：10072701

二〇二三年十月

摘要

本文回顾了自动驾驶技术的历史起源及其核心技术发展情况。介绍了自动驾驶的主要技术，包含环境感知、决策规划、路径控制等方面。文章阐述了自动驾驶各技术模块的最新进展，如多传感器融合感知、端到端深度学习决策等。并指出目前面临的挑战，如复杂场景下的环境识别。此外，文章从驾驶员行为认知研究介绍了人类驾驶模式对决策的影响。并以医疗机器人为例，说明自动驾驶技术对其他领域的启发效应。

通过不断提升自动驾驶的感知分析能力、人机交互性能，并与不同领域交叉学习，有助于自动驾驶水平不断提升，最终实现完全自动驾驶。

关键词：自动驾驶；环境感知；自动控制；模式识别；规划算法；机器人

一、自动驾驶的发展历程

（1）自动驾驶的缘起

自动驾驶汽车的故事始于二十世纪上半叶的美国。此时，致命交通事故的急剧上升正在成为一个社会问题。20 世纪 20 年代，汽车导致的交通事故使得约 20 万美国公民意外死亡。驾驶员的失误和车辆设计被认为是事故的主要原因。因此，用技术取代容易犯错的人类，是一个非常自然的想法，由此开启了自动驾驶的伟大征程[1]。

（2）无人驾驶

1921 年 8 月 5 日，美国工程师研制了一款第一辆无人驾驶汽车，使用无线电作为控制信号。其实从技术上讲，这并不是一辆自动驾驶汽车，而是一辆遥控汽车——只是驾驶员在车外。值得注意的是，无人驾驶汽车的历史与军事有关，因为这辆原型车是在军用卡车的基础上改进的，且是在俄亥俄州的一个军事基地进行了发布与公开测试。

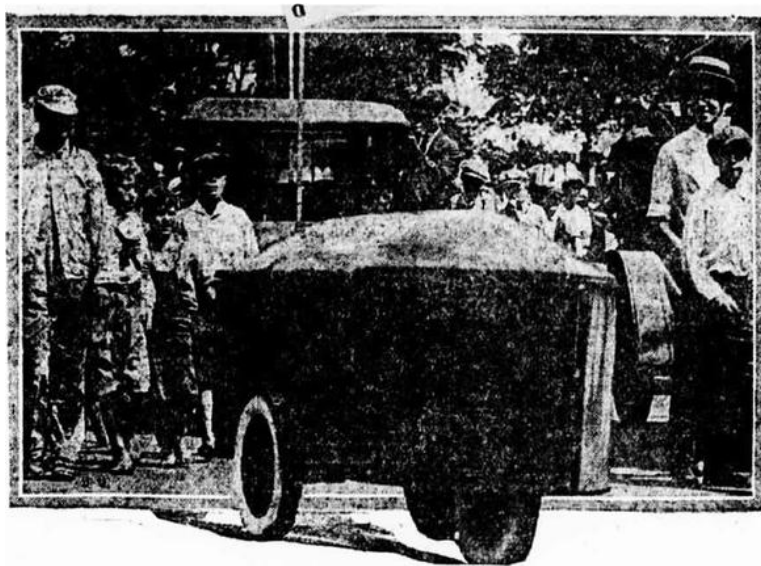


图 1. 世界上第一款无人驾驶汽车

20 世纪 80 年代，自动驾驶汽车的研究成为许多国家学术和工业研究的重要研究课题。对所有这些努力进行代表性描述将超出本文的框架。因此，我们将专注于最重要的开创性工作。德国慕尼黑联邦武装部队大学的恩斯特·迪克曼斯（Ernst Dickmanns）首次开发了基于多个边缘元件感知的带有数字处理器的视觉

引导自动驾驶汽车。1984 年，他的团队概念化了第一辆使用动态模型进行视觉自主引导的车辆：VaMoR（Versuchsfahrzeug für autonome Mobilität und Rechnersehen）是一辆 5 吨货车（Mercedes 508 D），能够携带大型计算机和这个时期的相机。1987 年夏天，VaMoR 可以自动驾驶（仅借助摄像头，无需雷达和 GPS）20 公里，速度高达 96 公里/小时（60 英里/小时）。该技术基于称为 4-D 方法的时空动态模型，该模型在空间的三个维度中添加了时间类别，并集成了预测误差的反馈。直到这次成功之后，汽车行业（戴姆勒-奔驰股份公司）才对迪克曼斯的研究产生了更大的兴趣。

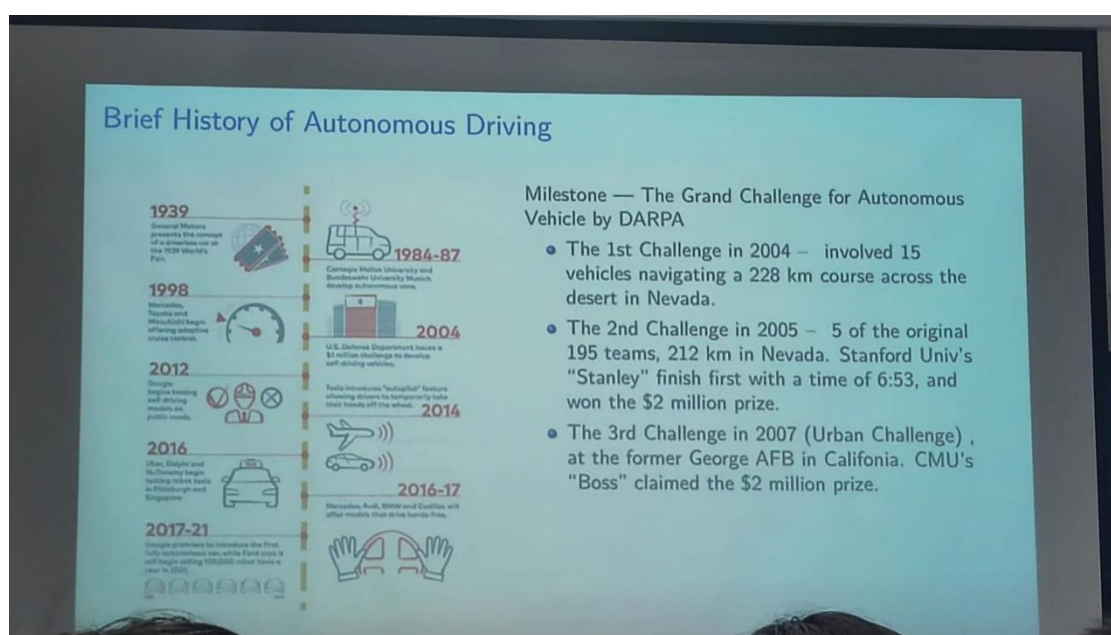


图 2. 自动驾驶发展史

基于视觉的自动驾驶概念随着欧盟的欧洲最高效率和前所未有的安全交通的尤里卡计划（PROMETHEUS）（1987-1994）而获得动力。起初，业界更倾向于利用道路上电缆产生的电磁场对汽车进行横向引导，这一点自 20 世纪 30 年代以来就广为人知。但迪克曼斯团队成功说服业界认可机器视觉概念，该概念将允许检测障碍物并避免基础设施的额外成本。如今，这可以被视为无人驾驶汽车历史上的重大范式转变。

二、 自动驾驶的核心技术

随着随着人工智能和机器视觉、智能感知、控制科学等技术的不断进步，自

自动驾驶车辆正快速发展。主要体现在：

- 1.环境感知能力提升。自动驾驶车辆利用多模态传感器收集周围环境信息，例如道路标记、交通标志、周边车辆等，以实现实时环境识别。
- 2.决策规划和控制执行能力增强。自动驾驶系统根据实时环境识别结果，进行路径规划和速度控制，以确保安全驾驶。
- 3.高精度地图和定位基础设施建设。高精度地图能为自动驾驶提供详细的道路几何信息支持，同时依靠实时定位系统保证自动驾驶车辆的精确定位。
- 4.软件和硬件技术的完善。自动驾驶系统依赖于软硬件技术的深度整合，例如机器视觉算法、深度学习、高性能计算等，以实现稳定可靠的自动驾驶功能。

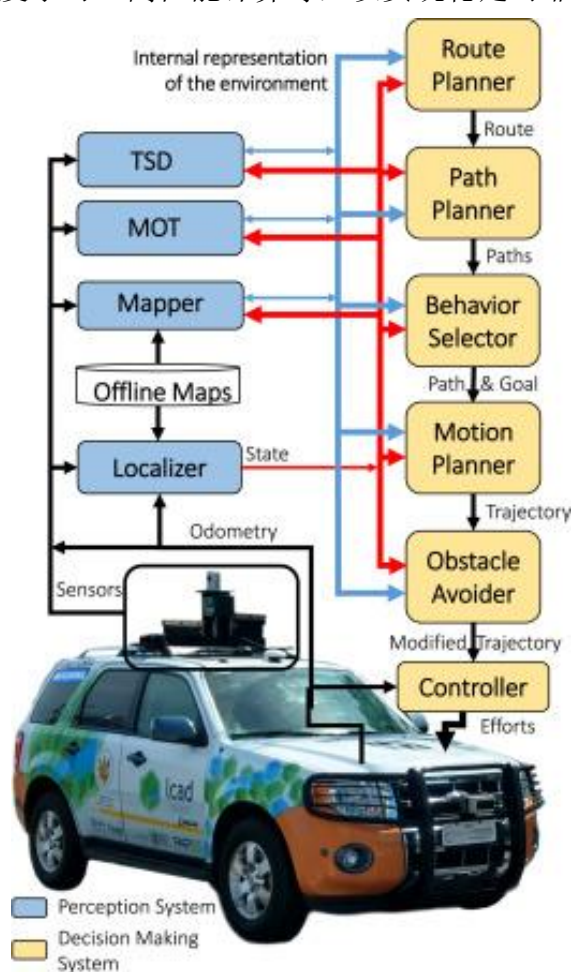


图 3. 自动驾驶系统各模块组成

根据自动驾驶车辆涵盖的功能范围和智能化水平，2014 年美国国际汽车工程师学会（SAE）在 3016 标准，将自动驾驶汽车的智能化程度划分成了 0 至 5 级，共 6 个等级，它们分别是：

0 级：无自动驾驶功能。车辆完全由人驾驶操作。

1 级：辅助驾驶。部分控制功能由电子系统提供辅助，如防滑系统、车距控制等。但人工仍需完全控制车辆。

2 级：部分自动驾驶。电子系统可以在特定条件下执行部分驾驶任务，如车道道线识别系统可实现自动驾驶，但人工必须时刻监控车辆行驶状况可随时接管控制。

3 级：条件自动驾驶。在一定条件下，车辆可以独立行驶、转向、加速和制动。然而，驾驶员必须做好随时接管驾驶的准备。

4 级：高自动驾驶。系统可以在大部分驾驶环境下独立完成驾驶任务，但在某些特殊条件下或特定区域还需人工辅助。

5 级：完全自动驾驶。系统可在所有路面与环境条件下完全独立完成驾驶任务，不需要人工参与。这是真正意义上的无人驾驶。

SAE美国汽车工程师学会对自动驾驶汽车的分级



图 4. SAE 自动驾驶汽车分级

(一) 感知

感知系统在自动驾驶中处于至关重要的地位。它通过各类传感器收集车辆内外的信息，实现对车辆状态和驾驶环境的实时识别与理解，为自动驾驶提供决策依据。自动驾驶感知系统需要像人类驾驶员一样，对所获得的原始数据进行解读识别，对车况环境进行快速准确评估，以确保行驶安全与合理决策。它通过各类

传感器实时感知车辆当前参数，如车速、转向量等，同时也能识别道路线型、交通信号、障碍物等环境信息。然后，感知系统将多源复杂数据进行语义化处理，从表象层面把握车辆与外部环境的语义关系。这为自动驾驶系统提供了深层次的决策依据。作为其他组件的基础输入，感知系统的工作决定了自动驾驶整体能力的上限。

对于自动驾驶车辆，其上搭载的常见传感器有：摄像头、激光雷达、超声波雷达、GPS 系统、多维惯性传感器等[2]。这些传感器的原理、功能、适用场景不同，但具有各自的特点与优势。摄像头能够采集自动驾驶汽车周围的图像与视频信息，信息量大、内容丰富，但采集的图像质量容易受光线变化、天气情况等外部条件影响。而激光雷达具有测量距离长、分辨率高等优点，但成本高且易受天气条件干扰。因此，单一传感器难满足多变路面环境下的复杂感知需求。自动驾驶系统通常采取多传感器融合策略，充分发挥不同传感技术的长处，例如视频图像提供环境特征，毫米波为静态障碍提供轮廓，激光传感进行精细测距等。这可以提升整体感知系统在准确性、鲁棒性和实时性方面的能力，从而更好地满足自动驾驶不同场景下的判断要求。

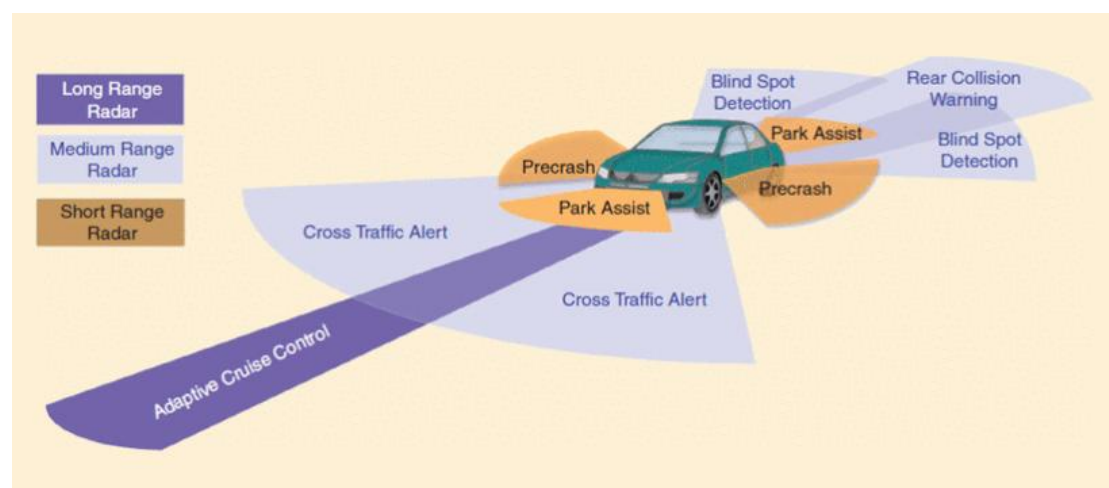


图 5. 不同雷达的有效范围区域

2005 年美国 DARPA (Defense Advanced Research Projects Agency, DARPA) 举办的无人车挑战赛中，斯坦福大学的 Stanley 智能驾驶汽车获得了冠军，其获胜的秘诀就在于具有 SOTA 的智能感知与分析系统[3]。目前，我国同济大学、清华大学等高校和科研院所在智能感知领域也已经取得了不少理论和技术进展。

- Self-driving vehicle of VECAN @ Tongji University



图 6. 同济大学 VECAN 团队自动驾驶样车

随着人工智能的再度兴起，尤其是深度学习技术的发展，自动驾驶汽车对环境的感知、分析能力得到了大幅提高，用以分析环境的深度神经网络模型被大量提出与应用，在目标识别、语义分割、图像优化、场景理解、轨迹跟踪、图像配准等领域都取得了不俗的表现，使得智能体对于环境的感知能力得到了显著提升[4]。

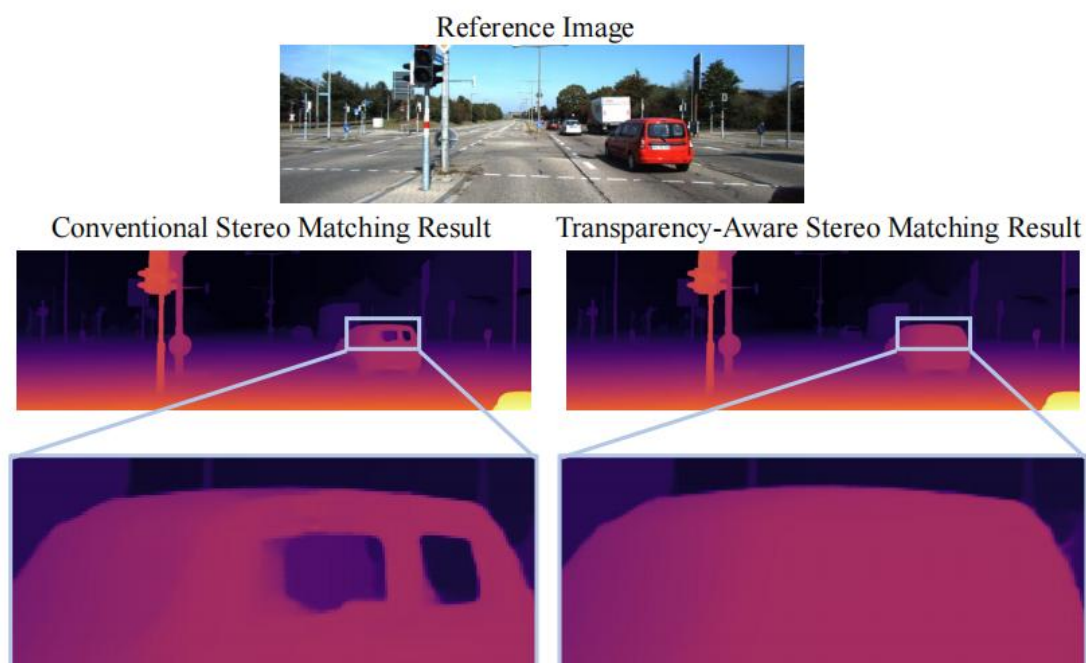


图 7. Zhiyuan Wu 等提出的环境感知算法效果图

尽管在深度学习的加持下，自动驾驶中的智能感知技术取得了跨越式发展，目前看起来效果也不错，但这一领域仍然面临许多挑战。特别是在边缘侧算力资

源不够充足的情况下、在待检测的环境复杂度高的情况下，自动驾驶感知系统要实现快速、准确、可靠的环境感知，仍然十分困难。

(二) 决策

自主决策是实现自动驾驶的核心。由于道路环境复杂多变，各参与主体行为难以预测，找到高效安全的决策规则是一项极具挑战性的工作。主要困难来源于：一是道路结构和交通规则의复杂多样性；二是交通流动态变化的随机性；三是各参与者间难预测的互动关系。自动驾驶车必须在这些不确定因素影响下做出实时决策。

目前，主流的自主决策设计方案大致有两类：

一是分层式设计，即将决策任务分解为感知、决策、执行等多个层级完成。这种方式简化问题，但分类可能过于明显，局限性较大。

二是端到端式设计，直接将大量数据作为输入，通过深度学习找到隐含规则，一步到位实现决策。这种方法拟合性强，但难以解释和优化模型内部逻辑。

分层式设计将决策过程细分为多个独立但互相衔接的子问题，如您分别介绍的情景认知、周车预测、行为选择等模块。这种方法的主要优点在于：

- 1.将复杂问题简化为多个易定解的子任务，提高算法可控性；
- 2.每个子模块独立开发，利于代码模块化和并行开发；
- 3.每个子问题的解可以直观解释，增加系统透明度；
- 4.实施和工程实现难度小。

然而，该设计方式同时也存在一定缺点：

- 1.子模块之间存在信息重合和语义损失；
- 2.人工定义难全面涵盖各种复杂驾驶场景及行为；
- 3.子模块耦合性强，影响系统整体鲁棒性。

与分解的解决方案不同，端到端的决策方法将决策过程视为一个不可分解的集成模块，主要包括两个子类：监督学习型和强化学习型。监督学习决策使用深度神经网络来模仿优秀驾驶员的驾驶行为。网络的输入是来自三个前置摄像头的信息，学习到的驾驶行为（例如方向盘角度）作为深度神经网络的输出。监督学习需要大量包含所有驾驶场景的驾驶员数据来训练深度神经网络，这是很困难的。强化学习是另一种端到端的决策方法。通过与环境交互来控制驾驶决策以获得

奖励（例如更高的安全性、更高的舒适度、更高的能源效率等）。目标是通过让智能汽车完成任务并获得最大的累积奖励来不断提高驾驶技能和决策水平。

当前，分层式框架中的基于规则的行为决策方法虽然简单，但灵活性不足。而端到端深度学习方法可以很好地处理特征表达难的场景，但其很难通过参数调整等方式在逻辑层面改进系统性能。

将来，将分层式框架和端到端深度学习方法进行融合，将自主学习能力与先验知识(如道路结构、车辆动力学模型、驾驶经验、交通规则等)进行有机结合，也将是自动驾驶技术发展的趋势。

此外，由于车载计算资源有限，未来可以利用云计算等资源，获取更多环境信息来分担计算负担，这也是自动驾驶技术发展的一个方向。通过云计算、联邦学习架构甚至卫星联网等高并行计算模式可以获得更广阔的上下文环境知识，有利于提升驾驶决策水平[5]。

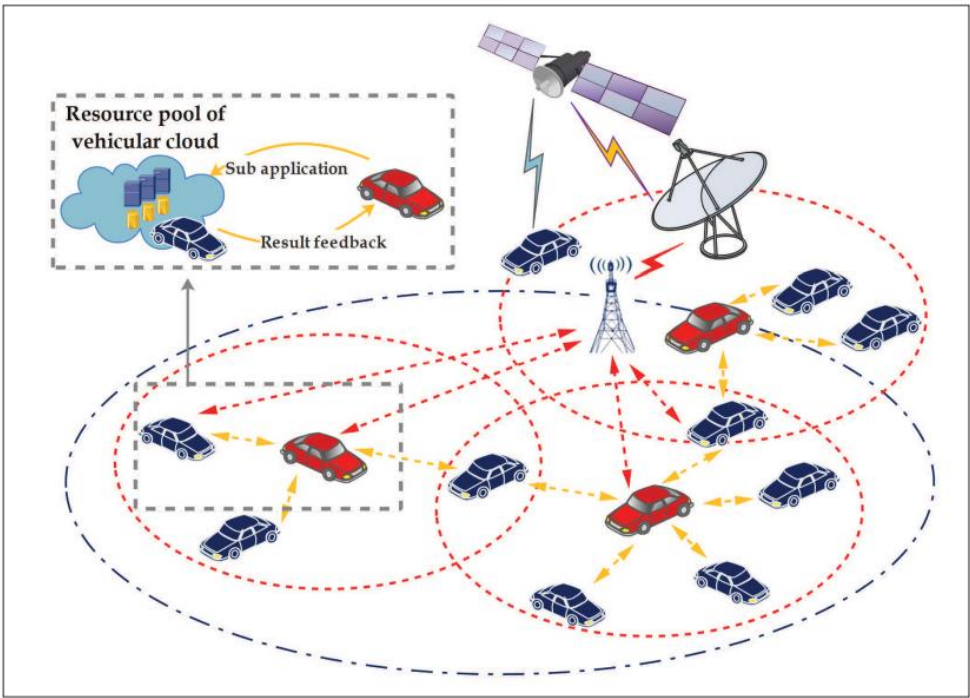


图 8. Minghui Liwang 等提出的基于卫星联网的智能驾驶计算系统架构

(三) 规划与控制

车辆系统的路径规划为自动驾驶车辆指出前往目标点的最优路径。运动规划组件的输入是车辆周围静态和动态障碍物的信息，输出是满足车辆动态和运动学约束的无碰撞轨迹。通常，基于采样的路径规划方法不需要对道路几何形状进

行建模，而是利用碰撞检测来探索自由空间和端点可达性。常用的算法包括随机转向算法和启发式转向算法。目前，应用最广泛的启发式搜索算法是 A* 算法，同时还有 D* 算法及其扩展算法[6]。

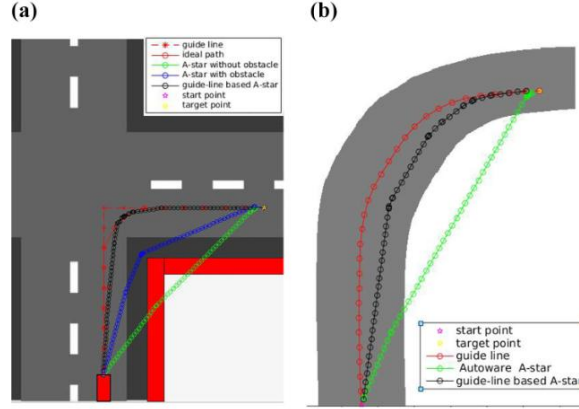


图 9. 路径规划算法样例

自动驾驶车辆的控制方法包括两方面：一方面是运动模型的控制，保持车辆的稳定运行，另一方面是控制车辆的循迹。运动模型的控制通常有纯跟踪控制、输出反馈线性化控制、线性参数变化控制等方法[7]。输出反馈线性化控制可以采用基于车轮/车辆位置的反馈方法[8]。最后线性参数变化控制通常用于横向控制的参数线性变化控制与路径轨迹的稳定性预测控制。在这些系统中，通常利用多个子系统的组合实现改进的处理性能[9]。

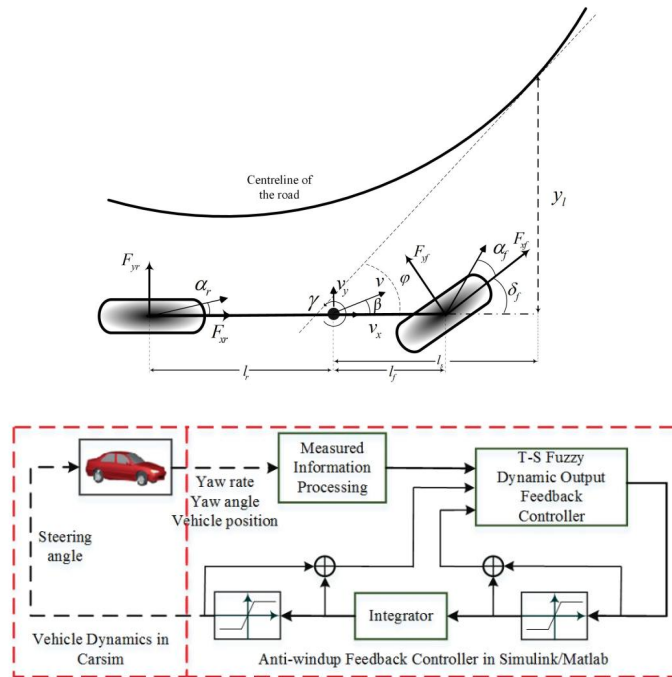


图 10. Changzhu Zhang 等提出的基于模糊推理的车轮反馈控制模型

(四) 驾驶员行为与认知

近年来，随着自动驾驶技术的不断发展，研究人员逐渐意识到仅通过不断升级车辆硬件配置和算法性能，无法从根本上解决自动驾驶车辆广泛应用的问题。自动驾驶汽车在面对更加复杂的道路环境和突发情况下，其决策能力和路径规划能力还无法匹配人类驾驶员的水平。因此，研究人员开始将人类驾驶行为和认知研究纳入自动驾驶技术研发的范畴，希望通过学习和提取人类驾驶模式中的有益经验，来改进自动驾驶车辆的控制算法，促进自动驾驶技术的更大进步。

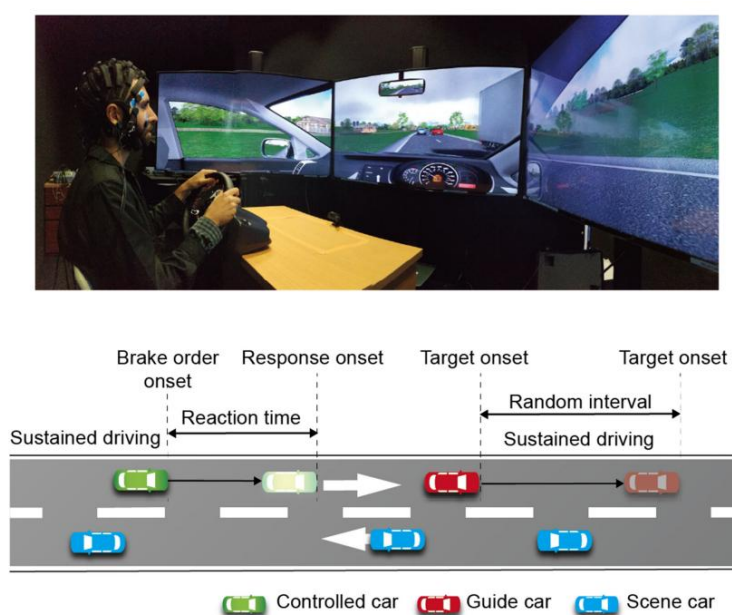


图 11.驾驶员行为与认知测试示意图

驾驶员行为和认知研究对自动驾驶技术的发展有着重要的参考价值和意义。对于 L3.5 级别需要驾驶员在紧急情况下接管车辆的自动驾驶车辆，智能车辆需要实时监测驾驶员的状态，以避免在驾驶员分心或疲劳时转移车辆控制权，从而降低发生事故的风险[10]。相比驾驶员行为研究,驾驶员认知研究具有更高的复杂性和全面性。由于手动驾驶相比自动驾驶可以带来更多的驾驶乐趣和吸引力，因此至今仍然是主流选择。这在一定程度上表明，驾驶员较希望车辆能够符合自己的驾驶意愿和风格。同样，通过研究驾驶员的意图推理和情绪表达，可以使自动驾驶汽车的决策和行为更加贴近驾驶员本身的决策，实现高效和稳定的路况交互与路径规划。因此，对 L4 或 L5 级别完全自动驾驶车来说，驾驶员认知研究的影响力更深远，其成果将决定自动驾驶技术是否能够顺利自然地融入人类社会。

三、自动驾驶对其他学科的启发——以医疗机器人为例

医疗机器人技术是现代医学与工程技术结合的创新成果，面向人民生命健康，具有重要社会、科学、技术与经济意义。在医疗机器人领域，路径规划、人体环境感知、手术过程人机交互控制等方面存在着不少共性的科学与工程问题。具有相互借鉴和启发的空间。值得一提的是，在手术机器人领域，前帝国理工学院的杨广中院士（Hamlyn 中心创始人，现上海交通大学医疗机器人研究院院长）类比自动驾驶的智能化层级划分，提出了医疗机器人的智能化层级划分[11]。



图 12. 医疗机器人智能化分级

这些智能层级包括：

0 级：无智能能力。此级别的机器人仅能执行遥控操作，如缩放运动等。

1 级：机器辅助。机器人提供一定的机械保护或辅助，人类保持对系统的持续控制。

2 级：任务智能。机器人能独立完成人类发起的某些具体任务，人类实行离散而非连续控制。

3 级：条件智能。系统能生成任务策略但依赖人类选择或批准策略。机器人无需严密监督即可完成任务。

4 级：高智能。机器人能进行医疗决策但须在合格医师监督下进行。类似医生助手级别。

5 级：全智能(无需人类参与)。机器人能独立完成整个手术流程，等同普通外科医生能力。

随着智能程度的增加，机器视觉、决策和执行能力需更强大，机器人的能力也越发接近于人类的高水平医生。同时手术风险和患者的隐私保护也成重要课题。总的来说这为机器手术不同智能级别提供了一个具有参考价值的智能化层级分类标准。这与自动驾驶汽车随着感知、决策、自主化能力的提升，驾驶员在驾驶过程中的参与程度以及自动驾驶任务所面临的场景等分级条件是具有高度相似性的，毕竟二者都是需要保证人（驾驶员/患者）的绝对安全的。

在技术层面上，自动驾驶汽车需要对周边环境进行实时识别与定位，以便选择安全高效的运行路径。这就需要强大的多模态感知能力，而手术机器人需要通过 DSA、CTA、IVUS 等多种模态的传感技术，实现对人体内部复杂环境的精准感知与分析，在环境感知能力方面与自动驾驶技术存在共同之处。此外，手术机器人还需要根据医生指令和人体结构识别，进行智能决策规划选择最佳操作方式，这与自动驾驶进行路径选择和动态规划有着类似的方式。在手术机器人感知领域，Prendergast 等[12]也提出了一种基于多模态图像状态相关的区域估计的内窥镜手术机器人导航策略，这与自动驾驶汽车中的视觉里程计在原理上是一致的。

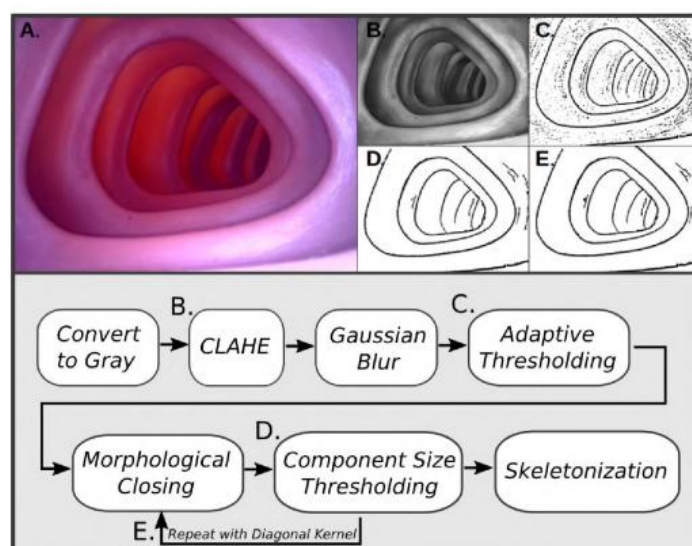


图 13..内窥镜手术机器人“视觉里程计”示意图

在精密机电一体化控制方面，手术机器人实现微米级操作也为自动驾驶实现厘米级精准控制奠定基础，同济大学的 Peng Qi 团队开发的 Venibot 利用视觉指导实时伺服，将静脉穿刺失败率降低到 2.00% [13]。

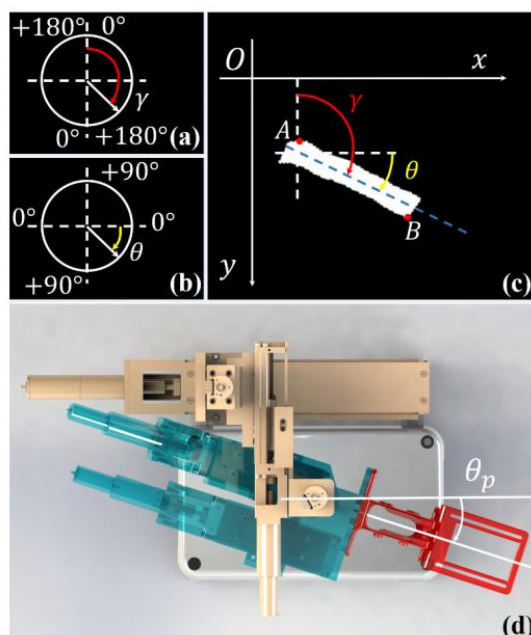


图 14. 扎针机器人水平方向位姿控制示意图

总之，在环境感知、智能决策以及精密控制等方面，自动驾驶系统的发展成果，对提升手术机器人技术的环境识别能力、决策水平以及运行精准度，都能起到一定的借鉴和启发作用。这将有利于推动手术机器人向智能化和安全性的更高程度过渡，并在不久的将来逐步落地。

四、总结

自动驾驶技术的发展历程漫长，从 20 世纪初期的无人驾驶汽车雏形，到 80 年代视觉引导的自动驾驶原型车问世，再到 21 世纪深度学习等前沿技术的引入，自动驾驶车辆的感知、决策、控制和规划能力获得了飞跃式的提升。当前，自动驾驶技术正处于从有人监管向完全自动驾驶过渡的关键时期。

为进一步增强自动驾驶汽车的智能化水平，研究人员开始关注人类驾驶行为认知模式在决策中的应用。与此同时，自动驾驶技术的发展也对其他领域产生了启发，例如在医疗机器人技术中，路径规划、环境感知和精密控制等方面均与自动驾驶技术存在着共通的科学问题，在技术实现上具有相互可借鉴之处。

综上所述，自动驾驶汽车的发展任重而道远，还有许多科学难题有待攻克，这些问题的解决都将推动自动驾驶技术向更高的安全性、智能化和可用性目标迈进。展望未来，自动驾驶技术必将以其巨大的经济效益和社会效益，深刻改变人类的

出行方式和生活模式。

最后，非常感谢电信学院控制系刘成菊老师组织的这一门非常精彩的课程，非常感谢王峻教授带来的精彩报告与分享，让我领略了自动驾驶之美！

参考文献

- [1] Badue C, Guidolini R, Carneiro R V, et al. Self-driving cars: A survey[J]. Expert Systems with Applications, 2021, 165: 113816.
- [2] Patole S M, Torlak M, Wang D, et al. Automotive radars: A review of signal processing techniques[J]. IEEE Signal Processing Magazine, 2017, 34(2): 22-35.
- [3] The 2005 DARPA grand challenge: the great robot race[M]. springer, 2007.
- [4] Wu Z, Su S, Chen Q, et al. Transparent Objects: A Corner Case in Stereo Matching[C]//2023 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2023: 12353-12359.
- [5] Liwang M, Dai S, Gao Z, et al. A truthful reverse-auction mechanism for computation offloading in cloud-enabled vehicular network[J]. IEEE Internet of Things Journal, 2018, 6(3): 4214-4227.
- [6] Arnay R, Morales N, Morell A, et al. Safe and reliable path planning for the autonomous vehicle verdino[J]. IEEE Intelligent Transportation Systems Magazine, 2016, 8(2): 22-32.
- [7] Åström K J, Murray R M. Feedback systems: an introduction for scientists and engineers[M]. Princeton university press, 2021.
- [8] Zhang C, Lam H K, Qiu J, et al. Fuzzy-model-based output feedback steering control in autonomous driving subject to actuator constraints[J]. IEEE Transactions on Fuzzy Systems, 2019, 29(3): 457-470.
- [9] Cardoso V, Oliveira J, Teixeira T, et al. A model-predictive motion planner for the iara autonomous car[C]//2017 IEEE international conference on robotics and automation (ICRA). IEEE, 2017: 225-230.
- [10] Xu T, Wang H, Lu G, et al. E-key: An EEG-based biometric authentication and driving fatigue detection system[J]. IEEE Transactions on Affective Computing, 2021.
- [11] Yang G Z, Cambias J, Cleary K, et al. Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy[J]. Science Robotics, 2017, 2(4): eaam8638.
- [12] Prendergast J M, Formosa G A, Fulton M J, et al. A real-time state dependent region estimator for autonomous endoscope navigation[J]. IEEE Transactions on Robotics, 2020, 37(3): 918-934.
- [13] Chen Y, Wang Y, Lai B, et al. Semi-supervised vein segmentation of ultrasound images for autonomous venipuncture[C]//2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2021: 9475-9481.