Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Views & Comments

智能驾驶视觉感知——技术挑战与创新解决方案

宫新乐[a,b]，钟志华[c,*]

[a] *School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China*
[b] *School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China*
[c] *School of Automotive Studies & College of Transportation, Tongji University, Shanghai 200092, China*

1. 引言

随着汽车智能技术飞速发展，车载视觉传感器已成为智能驾驶系统的核心部件，应用愈发广泛。它能够采集关键路况数据、辅助驾驶决策，并支撑自动驾驶功能落地。相较于其他车载感知传感器，视觉传感器可采集更细致、更全面的环境信息，为智能驾驶的决策与控制提供重要数据支撑[1]。因此，视觉感知的精度及可靠性直接决定智能驾驶系统的安全性与运行性能，也是实现完全智能化、自动驾驶车辆的关键所在。图1展示了智能驾驶中的视觉感知应用。

工业及消费级相机的技术标准相对宽松，而车规级相机面临着更为复杂、严苛的技术要求，亟须采取针对性的创新工程方案予以解决[2]。首先，从功能与应用层面来看，车规级相机需要适应多变的行车环境，在各类光照条件下保持出色的图像处理能力、硬件鲁棒性及系统集成水平。其次，在设计和制造环节，车规级相机须采用专用材料与先进封装工艺，同时还要经过严苛的环境适应性及抗振动冲击测试，确保在恶劣天气下依旧能输出清晰影像。除此之外，车规级相机在镜头、互补金属氧化物半导体（CMOS）图像传感器以及图像信号处理器三大核心部件上，仍存在诸多技术难题，具体包括高动态范围、强弱光环境适配、振动模糊补偿、低延迟、高分辨率以及运行稳定性等方面的要求[3]。攻克这些技术难点，是智能驾驶场景下实现安全、精准环境感知与目标识别的基础。

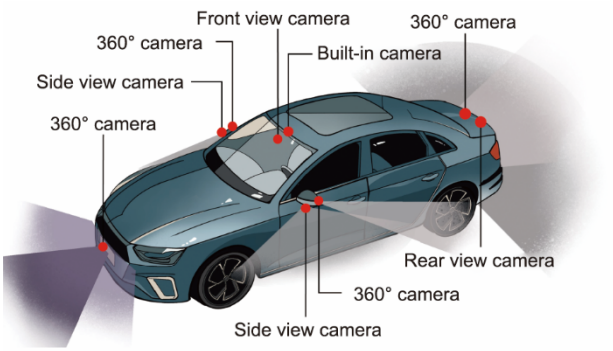


图1. 智能驾驶视觉感知示意图。

2. 核心技术与现存挑战

2.1. 结构与核心技术

车载视觉传感器主要由镜头、CMOS图像传感器、图像信号处理器以及串行器组成[4]。镜头对物体反射的光线进行折射，并将光线汇聚至CMOS图像传感器。CMOS图像传感器利用光电探测器的光电转换效应，把感光面上的光学影像转换为电信号，电信号强度与入射光亮度呈正

* Corresponding author.
E-mail address: zzh@cae.cn (Z. Zhong).

2095-8099/© 2025 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
英文原文：Engineering 2026, 57(2): 18–22
引用本文：Xinle Gong, Zhihua Zhong. Vision Sensing for Intelligent Driving: Technical Challenges and Innovative Solutions. *Engineering*, https://doi.org/10.1016/j.eng.2025.06.038

比。随后，图像信号处理器对原始图像数据进行预处理，完成图像缩放、自动曝光（AE）、自动白平衡（AWB）、自动对焦（AF）、图像降噪等操作，并将数据转换为适配格式。最后，串行器完成处理后图像数据的传输工作。图2为视觉传感器的结构及工作原理示意图。

2.2. 主要挑战

车载视觉传感器是智能驾驶系统实现实时目标检测与场景解析的重要设备，相关技术近年来已取得长足进步。但由于行车环境复杂多变，当前视觉感知技术仍存在诸多短板，难以兼顾高分辨率与高帧率、拍摄高动态范围图像、抵御强光干扰、保障弱光环境灵敏度，同时在快速响应、低延迟以及运行安全可靠性方面也存在不足[5]。此外，车辆行驶过程中产生的振动与颠簸会大幅影响视觉传感器的工作性能。受上述问题制约，在光照突变、极端天气及特殊行车场景中，设备的环境感知精度会下降，目标识别会出现延迟，这不仅增加了行车安全隐患，也阻碍了智能驾驶技术的快速发展[6]。本文围绕视觉传感器的三大核心部件——光学镜头、CMOS图像传感器和图像信号处理器——逐一分析其面临的主要技术挑战。

（1）光学镜头。镜头的光圈大小、焦距范围等参数会受到景深、光圈衍射、镜头体积等物理条件限制，进而影响不同光照环境下的成像质量[7]。受这类光学特性制约，现有镜头在明暗反差极大的场景中往往无法完整保留关键画面细节，导致路况与交通标识识别出现偏差[8]。车载广角镜头虽拓宽了视野范围，却会产生光学畸变，造成目标识别与距离测算出现较大误差。例如，智能车辆常采用190°鱼眼相机拍摄图像以构建鸟瞰图（BEV），但画面边缘的严重畸变会引发障碍物定位偏移，在窄路行驶场景中该问题尤为突出。同时，传统镜头的疏水涂层防污性能不足，雨天镜头表面残留的水滴会造成画面模糊，致使BEV系统无法识别湿滑反光路面上的车道线。

受衍射极限制约，镜头成像清晰度存在天然上限。现有镜头技术难以在缩小体积的同时进一步提升分辨率，无法清晰捕捉、识别远距离目标或小型物体（如远处的交通标识、路面障碍物等）。而且镜头所用材料对温度变化较为敏感，极端温度环境下会出现焦距偏移、色彩还原失真等问题，恶劣天气下的成像质量大打折扣。长期使用后，镜头性能也会逐步衰减。另外，车辆行驶时的机械振动与颠簸还会造成镜头失焦，降低画面清晰度和稳定性[9]。车载测试数据显示，当车辆以60 km·h^{-1}以上的速度驶过颠簸路面时，相机的振动频率可达30 Hz，进而造成画面模糊，导致自动紧急制动（AEB）系统响应延迟。

（2）CMOS图像传感器。空间分辨率提升意味着像素数量增加，图像数据量也随之变大。但受限于CMOS图像传感器的读出速度与模数转换效率，数据处理耗时会相应增加，最终造成时间分辨率（即帧率）下降[10]。现阶段技术通常只能牺牲帧率来换取更高的画面分辨率，导致车载视觉传感器难以同时实现远距离小目标探测与快速运动物体实时追踪。人眼的动态范围约为160 dB，而目前车载CMOS图像传感器的动态范围普遍仅为120~140 dB，尽管相关性能在持续优化，但其依旧无法全面适配智能驾驶全场景下复杂的光照变化[11]。不仅如此，提升分辨率会压缩单个像素的尺寸，降低像素饱和电荷容量，进一步限制CMOS图像传感器的动态范围。在隧道出入口、夜间强光环境等明暗交错的复杂场景中，画面极易出现局部过曝或欠曝，丢失关键细节，影响高反光物体（如交通标识）与阴影区域的同步识别。例如，日出、日落时分的逆光场景会大幅降低车牌识别准确率。

受动态范围、电子溢出现象、镜头散射、模数转换器位深以及曝光控制等特性制约，现有CMOS图像传感器对强光光源（如直射阳光、车辆大灯）的抑制能力有限，极易产生眩光，干扰光源周边区域的成像效果[12]。这会造成检测盲区，例如，夜间行驶时，对向车辆开启远光灯会短暂造成相机“致盲”。高分辨率视觉传感器会缩小像素尺寸，光电探测器随之变小。受物理条件限制，单个像素的光电转换效率有所下降，读出噪声也会增加，进而弱化弱光成像性能。在夜间、黄昏等低光环境下，画面质量变差，远距离、小型目标的检测准确率随之降低；同时降噪运算量大幅增加，降低图像处理的实时性[13]。例如，在黄昏时

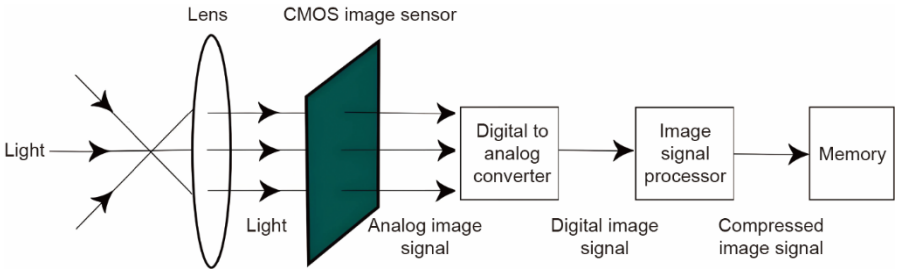


图2. 视觉传感器结构与工作原理。

段，现有视觉传感器的行人检测召回率会明显下降，主要原因就是低光环境下CMOS传感器的信噪比大幅降低。

目前主流CMOS图像传感器采用Bayer滤镜阵列区分色彩，导致色彩分辨率仅为原始分辨率的四分之一[14]。传感器感光度提升后，色彩噪点会变得十分明显，使得目标物体在不同光照下难以保持稳定的视觉特征，影响基于色彩的目标识别（如交通标识与信号灯识别）。此外，CMOS图像传感器的暗电流会随温度升高而增大，进一步拉低成像质量。为此，现有技术需要在不同温度环境下对传感器进行重新校准，但即便如此，酷暑、严寒等极端温度依然会造成传感器性能衰减。设备长时间工作还会出现温升问题，影响传感器运行稳定性。

（3）图像信号处理器。现有图像信号处理器的片上处理与并行运算能力存在局限，图像处理速度受到制约，处理高分辨率、高帧率图像数据时该问题尤为突出[15]。这一短板直接影响智能车辆的实时响应能力。车辆高速行驶或处于复杂路况时，处理延迟会导致系统无法及时识别并应对突发状况，降低智能驾驶系统的安全性与可靠性。低光环境下传感器噪声会增大，而当前图像信号处理器的降噪、信号增强算法仍有较大优化空间[16]。这些问题会造成成像质量下降，致使系统在夜间、隧道等场景中难以精准识别和追踪车辆及行人。

面对强光、眩光、雨雪天气以及路面反光等复杂工况，现有图像信号处理器无法有效去除画面噪点和眩光，会降低图像清晰度与可用性，进而导致障碍物识别出错或识别失败[17]。多数图像信号处理器采用固化算法设计，可编程性差，难以通过软件更新适配全新图像处理算法，也无法满足智能驾驶的各类定制化需求。同时，这类处理器缺少场景理解与环境感知能力，无法针对隧道、夜间等特定行车场景做专项优化，这限制了视觉感知系统的智能化水平。车辆行驶产生的振动、颠簸会造成图像模糊与畸变，而现有图像信号处理器对此类问题的修正能力不足，会进一步劣化画质，影响目标检测与追踪精度，在高速行驶、颠簸路面上该问题更为明显。表1汇总了车载视觉传感器存在的各类技术挑战。

3. 创新解决方案

针对智能车载视觉传感器现存的技术难题，本文结合跨领域技术融合思路，提出针对性的优化方向与创新解决方案。

3.1. 新型材料与技术应用

传统视觉传感器在分辨率、感光度、感光能力等方面遭遇性能瓶颈，想要突破这一困境，须研发量子点、钙钛矿等新型高效感光材料。已有研究证实，纳米结构材料能够提升光子吸收效率，优化传感器工作效能，有望为图像传感技术带来变革[18]。这类新型光电材料可显著提升智能驾驶系统的环境感知、障碍物检测及夜间成像能力。因此，借助纳米结构设计优化光学特性是提升感光效率与整体画质的有效手段。例如，文献[19]提出的全模拟光电芯片专门面向高速视觉场景设计，印证了优化材料属性能够大幅提升成像表现，尤其适用于高速行驶与低光环境。依托光电效应探索全新技术路径、提升光电转换效率也能有效改善传感器综合性能。此外，研究表面等离子体共振等纳米光学技术可增强传感器的光线吸收能力，提升复杂光照条件下的画面清晰度。融合以上各类前沿技术能够全面提升视觉传感器的成像水准，让自动驾驶车辆在真实复杂路况中实现更稳定、更可靠的环境感知。

3.2. 架构优化升级

为解决视觉传感器分辨率与帧率难以兼顾的问题，研究人员开始借鉴人类视觉系统，探索神经形态视觉感知架构。通过研发仿生自适应视觉传感器，可大幅提升自动驾驶车辆在复杂环境下的感知能力，使其拥有更灵敏、更可靠的视觉感知能力。例如，文献[20]提出的视觉芯片采用仿人视觉互补通路结构，兼具高动态范围与快速响应能力，适用于对动态行车环境的实时感知。为进一步提升高速成像性能，相关研究基于连续连拍原理研发出超高速动态成像技术[21]。该技术能够对高速移动目标进行高清成像，强化障碍物检测、动态场景感知、运动轨迹追踪等核

表1 车载视觉传感器技术挑战汇总

Component	Key challenges	Impact on performance
Optical lens	Diffraction limits, distortion, temperature sensitivity, and vibration-induced defocus	Loss of detail; inaccurate object recognition and distance estimation
CMOS image sensor	Trade-off between resolution and frame rate, limited dynamic range, glare effects, increased noise in low light, and reduced color resolution	Delayed target detection; degraded performance in extreme lighting conditions
Image signal processor	Limited processing speed, inadequate noise reduction, and lack of algorithmic flexibility and anti-shake capability	Reduced real-time performance; misjudgment in dynamic scenarios

心功能。可精准捕捉瞬时运动状态对于要求快速响应、高可靠性的自动驾驶场景至关重要。研发专用集成电路是提升图像处理效率与灵活性的关键。文献[22]介绍了一款用于畸变校正三维成像的集成视觉传感器，证明优化传感器架构能够在减少画面畸变的同时，提升图像处理效率。设计可重构图像信号处理器架构、优化数据传输性能，能够有效提高视觉传感器的运算效率，使其适配多样化的处理需求。此外，车辆振动和突发位移会严重影响成像质量，因此须融合先进的机械防抖与电子防抖技术。结合光学防抖与电子防抖的混合方案可动态补偿运动带来的画面畸变，保障自动驾驶全程成像清晰、画面稳定。

3.3. 智能图像处理算法

传统计算架构在实时图像处理中面临较大的计算负载、功耗受限等问题，因此须探索神经形态计算、量子计算等新型计算模式。这类前沿技术为研发轻量化、低功耗的人工智能算法提供了新思路，可满足智能驾驶场景下高效、实时的图像处理需求。例如，文献[23]基于事件相机构建高频周期信号检测概率模型，利用事件相机异步工作的特性有效区分目标信号与随机噪声，突破了传统帧式相机的采样率瓶颈。面对复杂多变的行车环境，须研发自适应曝光与色彩校正算法，削弱光照波动带来的干扰。同时，结合物理原理研发去雾、去雨算法，可改善恶劣天气下的成像质量。文献[24]提出一种多尺度频率分离网络，大幅提升图像去模糊效率，并在此基础上融入雨滴折射模型，进一步优化了车道识别效果。在夜间、低能见度场景中，融合红外热成像技术可补充可见光之外的环境信息，提升自动驾驶感知系统的稳定性。借助深度学习技术优化图像处理流程能够提升目标检测、分类任务的识别准确率，尤其适用于高速移动、目标遮挡等复杂工况。此外，还须研发基于人工智能的实时防抖算法[25]，通过分析连续帧之间的运动规律，动态补偿车辆颠簸、振动引发的画面畸变，保障自动驾驶系统实现高质量视觉感知。

4. 结语

本文系统分析了用于智能驾驶的车载视觉传感器的核心结构、关键技术及现存技术难题。研究发现，当前车载视觉传感器在光学镜头、CMOS图像传感器及图像信号处理器这三大部件上均存在明显短板，包括衍射限制、光学畸变、动态范围不足、易受眩光影响、色彩分辨率偏低、处理速度有限等问题。这些缺陷共同制约着智能驾驶系统环境感知的精度、稳定性与可靠性，在极端光照、恶劣天气及复杂行车场景中表现尤为凸出。针对上述问题，本文提出多项创新解决思路：研发新型材料与高效光电转换技术，提升传感器基础性能；借鉴人类视觉系统完成架构升级，实现高动态范围成像与超高速成像；基于神经形态计算、量子计算开发智能图像处理算法，保障复杂环境下感知系统实时运行、低功耗且抗干扰能力强。随着相关技术不断迭代，车载视觉感知系统将朝着高度智能化、高可靠性、高度集成化的方向持续发展，为未来实现更加安全、高效的完全自动驾驶提供助力。

CRediT authorship contribution statement

Xinle Gong: Writing-original draft, Formal analysis, Methodology. **Zhihua Zhong:** Supervision, Writing-review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

致谢

本研究得到国家自然科学基金（52102438）、中国博士后科学基金（2022M711803 和 2024T170482）以及智能绿色车辆与交通全国重点实验室的支持。

References

[1] Gouveia LCP, Choubey B. Advances on CMOS image sensors. Sens Rev 2016; 36(3):231–9.
[2] Innocent M, Velichko S, Lloyd D, Beck J, Hernandez A, Vanhoff B, et al. Automotive 8.3 MP CMOS image sensor with 150 dB dynamic range and light flicker mitigation. In: Proceedings of the 2021 IEEE International Electron Devices Meeting (IEDM); 2021 Dec 11 – 16; San Francisco, CA, USA. Piscataway: IEEE; 2021. p. 30.2.1–4.
[3] Gove RJ. CMOS image sensor technology advances for mobile devices. In: High Performance Silicon Imaging. Amsterdam: Elsevier; 2020. p. 185–240.
[4] Fossum ER. CMOS image sensors: electronic camera-on-a-chip. IEEE Trans Electron Devices 1997;44(10):1689–98.
[5] El-Desouki M, Deen MJ, Fang Q, Liu L, Tse F, Armstrong D. CMOS image sensors for high speed applications. Sensors 2009;9(1):430–44.
[6] Gehrig D, Scaramuzza D. Low-latency automotive vision with event cameras. Nature 2024;629(8014):1034–40.
[7] Liu Z, Hu G, Ye H, Wei M, Guo Z, Chen K, et al. Mold-free self-assembled scalable microlens arrays with ultrasmooth surface and record-high resolution. Light Sci Appl 2023;12(1):143.
[8] Gamal AE, Eltoukhy H. CMOS image sensors. IEEE Circuits Devices Mag 2005;21(3):6–20.

[9] Xu JT, Wang XY, Wang TD, Chen X, Song Z, Lei H, et al. Review on optical visual sensor technology. J Image Graph 2023;28(6):1630–61.
[10] Sukhavasi SB, Sukhavasi SB, Elleithy K, Abuzneid S, Elleithy A. CMOS image sensors in surveillance system applications. Sensors 2021;21(2):488.
[11] Takayanagi I, Kuroda R. HDR CMOS image sensors for automotive applications. IEEE Trans Electron Devices 2022;69(6):2815–23.
[12] Scott-Thomas J. Trends and developments in state-of-the-art CMOS image sensors. Piscataway: IEEE; 2023.
[13] Theuwissen A. CMOS image sensors: state-of-the-art and future perspectives. In: Proceedings of the ESSCIRC 2007—33rd European Solid-State Circuits Conference; 2007 Sep 11–13; Munich, Germany. Piscataway: IEEE; 2007. p. 21–7.
[14] Ohta J. Smart CMOS image sensors and applications. 2nd ed. Boca Raton, FL: CRC Press; 2020.
[15] Maheepala M, Joordens MA, Kouzani AZ. Low power processors and image sensors for vision-based IoT devices: a review. IEEE Sens J 2021;21(2):1172–86.
[16] Wang X, Wong W, Hornsey R. A high dynamic range CMOS image sensor with inpixel light-to-frequency conversion. IEEE Trans Electron Devices 2006; 53(12):2988–92.
[17] Bandoh Y, Qiu G, Okuda M, Daly S, Aach T, Au OC. Recent advances in high dynamic range imaging technology. In: Proceedings of the 2010 IEEE International Conference on Image Processing; 2010 Jun 26–29; Hong Kong, China. Piscataway: IEEE; 2010. p. 3125–8.
[18] Iqbal MA, Malik M, Le TK, Anwar N, Bakhsh S, Shahid W, et al. Technological evolution of image sensing designed by nanostructured materials. ACS Mater Lett 2023;5(4):1027–60.
[19] Chen Y, Nazhamaiti M, Xu H, Meng Y, Zhou T, Li G, et al. All-analog photoelectronic chip for high-speed vision tasks. Nature 2023;623(7985):48–57.
[20] Yang Z, Wang T, Lin Y, Chen Y, Zeng H, Pei J, et al. A vision chip with complementary pathways for open-world sensing. Nature 2024;629(8014):1027–33.
[21] Huang TJ. Spiking continuous photographing principle and demonstration on ultrahigh speed and high dynamic imaging. Acta Electron Sinic a 2022;50(12): 2919–27.
[22] Wu J, Guo Y, Deng C, Zhang A, Qiao H, Lu Z, et al. An integrated imaging sensor for aberration-corrected 3D photography. Nature 2022;612(7938):62–71.
[23] Ben-Ezra DEC, Arad R, Padowicz A, Tugendhaft I. Probabilistic approach for detection of high-frequency periodic signals using an event camera. Nat Comput 2026;22:99–110.
[24] Zhang Y, Li Q, Qi M, Liu D, Kong J, Wang J. Multi-scale frequency separation network for image deblurring. IEEE Trans Circuits Syst Video Tech 2023; 33(10):5525–37.
[25] Li H, Hua Q, Shen G. Tianmouc vision chip designed for open-world sensing. Sci China Mater 2024;67(9):3046–8.