



多模态信息质量评价与技术挑战

方向¹, 马楠^{2*}, 韩义恒², 夏芹³, 吴祉璇⁴, 肖荣春², 刘怡光⁵, 熊行创¹, 黄晋⁶

1. 中国计量科学研究院, 北京 100029

2. 北京工业大学信息科学技术学院, 北京 100124

3. 中国汽车工程研究院股份有限公司, 重庆 401122

4. 北京邮电大学计算机学院, 北京 100876

5. 四川大学计算机学院, 成都 610065

6. 清华大学车辆与运载学院, 北京 100084

* E-mail: manan123@bjut.edu.cn

收稿日期: 2025-09-07; 接受日期: 2025-12-11; 网络版发表日期: 2026-01-08

国家重点研发计划“国家质量基础设施体系”重点专项“面向质量信任的多维数据校准关键技术和标准研究”项目(编号: 2023YFF0615800)资助项目

摘要 随着感知智能、多模态信息融合、视觉语言模型等技术的发展, 智能时代获取信息呈现类型多样化、多维度等特征, 给处理、校准和质量评价带来了诸多挑战. 如何高效、准确地处理和评估这些多源异构信息的质量, 已成为智能化信息处理发展中的一个关键问题. 本文提出以据数为核心的多模态信息质量评价, 旨在通过规范信息采集源头和处理流程, 以提升可信性和利用效率. 文章首先探讨了信息采集的标准化流程, 强调信息来源的清晰界定、采集方式的合理选择以及信息格式的统一; 其次, 提出了信息处理流程的规范化, 并建立多维度信息的验证、评价方法和校准技术, 以确保信息的完整性、准确性、一致性、时效性和可访问性; 最后, 以自动驾驶场景下多传感器信息处理为例, 从据数的角度探讨了多模态信息融合和处理的质量评价. 本文的研究将为多模态信息智能化处理的可信性提供系统的质量评价方法, 助力“国家质量基础设施体系”建设, 推动人工智能技术高质量发展.

关键词 据数, 信息可信, 采集源, 质量评价, 自动驾驶

1 引言

信息技术已成为驱动科技创新与社会进步的重要力量. 然而, 由于信息在采集端装置和处理过程等方面的多样性, 其本身的质量、可信度和可用性, 仍然面临诸多挑战. 信息源于物质和能量, 可被感知、存储、加工、传递和再生. 今天, 所有信息记录的符号表示或者

载体, 都被统称为数据. 传统意义上的数据是人类对事物进行测量的结果, 是作为“量”而存在的数据, 可以称为“量数”; 照片、视频、音频不是源于测量而是源于对周围环境的记录, 是作为一种证据、根据而存在的, 可以称为“据数”^[1], “据数”更强调其在多模态场景中的可验证性与可追溯性, 更偏向于在技术层面提供一种“源头之尺”, 用以评价和规范信息的质量, 如图1所示.

引用格式: 方向, 马楠, 韩义恒, 等. 多模态信息质量评价与技术挑战. 中国科学: 技术科学, 2026, 56: 1626–1632

Fang X, Ma N, Han Y H, et al. Multimodal information quality assessment and technical challenges (in Chinese). Sci Sin Tech, 2026, 56: 1626–1632, doi: 10.1360/SST-2025-0251

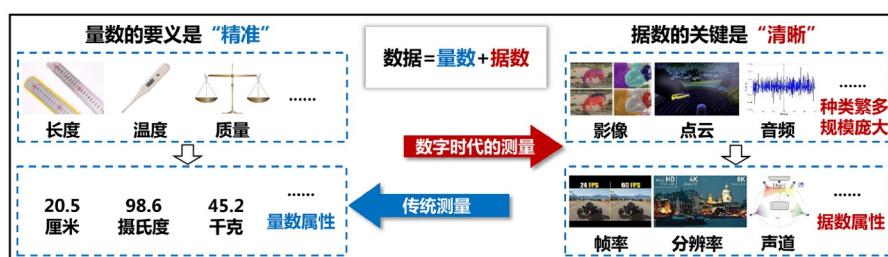


图 1 (网络版彩图)量数与据数

Figure 1 (Color online) Data and evidence-based data.

据数源于多个维度和属性的记录集合, 是对客观存在的多种感知对象的数字化刻画, 如影像、点云、音频等, 主要描述物理对象的形态及其他可测属性^[1]. 据数对智能时代发展至关重要, 不同模态的信息可称其为多维据数. 例如, 一张RGB图像不仅是像素值的集合, 还包含物体类别、几何特征、时空上下文语义关联等信息, 强调了结构化语义与属性关联, 然而, 当前许多信息, 如深度图像、雷达点云, 仅具备粗粒度标签, 缺乏精细化属性描述与多模态语义关联, 这种缺陷导致质量参差不齐, 缺少用于区分和判别的属性标注和自动化的判别手段. 而随着大模型^[2]、多模态感知及交互^[3]等技术的快速发展, 科技社会获取的信息种类愈加复杂, 呈现出高维、多模态、动态变化等特征, 这给信息的处理、校准、融合以及质量可信评价带来了前所未有的挑战. 同时, 信息规模的指数级增长与人工智能技术的深度演进, 使得可信性问题成为制约智能系统纵深发展的核心瓶颈之一^[4]. 从信息的采集、存储、处理、分析到应用, 其可信性挑战贯穿整个生命周期, 仅依靠局部清洗和人工校验等手段, 难以解决全局性、系统性的可信问题, 导致质量管理具有局限性. 究其根本, 在于缺乏一个“可信之尺”. 为了寻找衡量多维据数可信的“尺”, 可从两个维度展开探索.

(1) 规范信息采集的源头, 确保据数的初始质量和采集过程的标准化, 并建立完整的计量体系. 例如, 在智能制造场景中, 建立统一的传感器标定标准, 时空同步精度、元信息标准化尤为重要, 作为据数采集标尺的具体体现, 不同设备采集的信息可以进行一致性度量 and 对比.

(2) 推进处理流程的规范化, 构建多维据数融合质量验证与校准方法, 并建立科学的评价与校准体系. 这部分包括多模态信息处理的标准化、融合准确性评估

以及全生命周期质量溯源等. 例如, 自动驾驶中多模态信息融合的置信度评估方法构建, 是标尺在据数处理层面的延伸.

在众多信息类型中, 已经存在较为成熟的信息组织形式, 文本便是一个典型的例子. 在人类文明的漫长发展过程中, 文字体系凝聚了丰富的“尺”, 例如, 语法规则(句法结构、语义逻辑)、时态变化(过去、现在、将来)、词性划分(名词、动词、形容词等). 正因如此, 文本的处理方式得到极大的发展, 使其在人工智能领域的应用远远领先于其他信息类型. 大语言模型(large language model, LLM)的广泛应用, 得益于长期积累的大规模、高质量的结构化文本, 使得模型可以通过预训练和微调掌握丰富的语言知识. 这一事实充分证明了“尺”对于可信性的关键作用. 文字之所以能够实现高效的计算和应用, 正是因为其自带的一系列“标尺”, 这些规则使得文本具备高度结构化的特点, 从而可以被计算机高效处理. 相比之下, 影像、音频、点云等内容仍缺乏类似的度量体系, 导致难以被高效组织和利用.

本文以据数为研究对象, 提出一套适用于多维据数的可信衡量方法. 通过引入据数视角, 我们建立可信度量标准, 探索规范化组织方式, 推动结构化、属性化, 使其具备更高的利用价值, 为人工智能的发展奠定坚实的信息处理规范化基础, 保障各类智能系统建设的信息可信.

2 规范采集源头, 建立源头之“尺”

信息可信性的根基始于采集源头, 影像、点云、音频等作为据数的基本单元, 其采集质量和规范性直接影响多模态信息的来源质量^[5]. 国民经济和社会发

展十四五规划和2035年远景目标¹⁾中指出“完善国家质量基础设施, 强化标准、计量等体系能力建设”, 但当前智能机器感知系统的多类型据数仍面临采集装置标准不统一、格式多样、环境干扰复杂、质量校验机制缺失等问题, 导致据数质量参差不齐, 难以支撑高可信需求^[6]. 因此, 亟需为据数构建源头标尺, 确保采集过程的规范化与可度量性。

2.1 构建多维据数采集的源头标尺

针对多维据数采集端可能出现采集质量低、抗干扰能力弱, 难以对多模态信息, 如RGB图像、深度图像、视频、激光点云、合成孔径雷达(synthetic aperture radar, SAR)、高光谱图像等进行精确采集的难题, 建立据数采集源头的“标尺”, 需覆盖格式、精度、元信息、编码规则及存储协议等核心维度, 使其高质量共享、交换和集成, 提升自动化处理能力, 为智能分析奠定基础. 具体而言, 可以采取以下措施. (1) 标准化采集装置与技术参数: 针对不同据数类型, 制定统一的采集装置标准和技术参数规范. 例如, 影像采集设备应明确分辨率、帧率、色彩空间等核心参数; 点云采集设备应规定扫描精度、点密度、扫描范围等技术指标; 音频采集设备应明确采样率、信噪比、频率响应等关键参数. (2) 优化采集环境与抗干扰设计: 针对环境干扰问题, 提出优化采集环境的指导原则和抗干扰设计方法. 例如, 影像采集应制定光照条件、背景设置、遮挡物处理等环境规范; 点云采集应提供反射表面处理、噪声点过滤等技术方案; 音频采集应规定降噪算法、回声消除等技术要求. (3) 建立实时质量校验机制: 针对采集过程中可能出现的质量失真等问题, 建立采集过程中及时发现和质量校验的机制. 例如, 影像采集过程可通过实时图像分析技术检测模糊、失真、过曝或欠曝等问题; 点云采集过程中可通过点云滤波算法去除噪声点和异常点; 音频采集过程可通过频谱分析技术识别背景噪声和失真. (4) 计量体系的协同设计: 建立与标准体系相匹配的计量体系^[7], 确保据数采集过程的可度量性. 例如, 影像采集可通过光学计量技术校准设备参数; 点云采集可通过激光计量技术验证扫描精度. 通过多模态计量技术与采集源深度融合, 提升源据数质量, 为后续多模态据数处理、融

合与质量评价提供可靠的源头质量保障。

2.2 源据数的数字化智能校准技术

源据数校准技术聚焦于采集源头的质量评价, 通过量化质量评价中的指标, 从据数智能采集、预处理、探测、感知的通用架构角度研究多维据数数字化智能校准问题, 以RGB图像为例包括单图像质量评价(single-image quality assessment, SIA)和多图像质量评价(multi-image quality assessment, MIA), 对采集到的多源据数进行实时智能预处理和配准, 建立基于源据数感知与评价模型, 对据数可能存在的畸变、晕影、白平衡等需要校准的方面提出峰值信噪比和结构相似度等量化的评价指标, 根据结果反馈, 评估源据数的质量可信。

3 规范多维据数处理, 建立使用之“尺”

规范多维据数处理是在采集源头的质量可信基础上, 研究多维据数感知模型及融合技术、多维据数处理的标准化格式构建、多维据数融合评价体系建立和多维据数端到端校准能力的检测验证方法等. 这些研究将有助于解决多维据数感知模型处理过程中的质量评价与可信溯源验证等核心问题^[8]. 为此, 在规范采集源头的基础上, 建立多维据数处理流程的规范化。

3.1 从据数的角度组织多模态信息

从据数的角度组织多模态信息, 可为其在智能化处理过程中提供标准化和质量溯源的依据. 例如在自动驾驶领域, 通过对车载传感器采集的多源据数进行联合标定, 提高信息融合精度和感知一致性. 其次是信息的多维度分析, 不仅关注信息的表面特征, 还需提取信息的隐含结构和属性, 例如对图像的纹理、几何、深度信息, 点云的密度、法向量、曲率等特征进行可视化的多维度分析^[9].

多维据数要进行有效表征, 其信息处理与验证是质量可信的关键. 在信息处理阶段, 需对多维据数进行时空对齐、动态对齐和语义融合等操作, 以消除时空漂移、修复数据断层、增强跨域一致性. 时空对齐是通过构建时空约束模型消除维度间的不一致性, 重点关注多维据数的时空基准统一及滤波处理的技术;

1) https://www.xinhuanet.com/politics/2021lh/2021-03/05/c_1127172897.htm?baike.

动态对齐运用滑动窗口机制实现多模态信息的对象关联矩阵和动态关联配准; 语义融合可通过时空图神经网络构建跨域特征映射, 在统一时空参照系下实现语义信息的层次化重组。

在规范据数处理过程中, 建立多维度信息验证方法同样至关重要。具体而言, 多模态信息可以从以下几个维度进行验证: (1) 完整性, 检查感知信息是否缺失, 确保其属性的完整性; (2) 准确性, 通过多维据数交叉比对、基准对比等方式, 识别信息偏差、矛盾或错误, 完成准确性的量化过程; (3) 一致性, 确保同一对象在不同信息源中的表示方式相一致, 例如时间戳、坐标系等; (4) 时效性, 确保信息的更新频率符合应用需求, 如实时监测信息的传输延迟; (5) 可访问性, 验证信息的存取权限、存储路径和可读性, 确保信息的高效调用。可结合统计分析、机器学习等方法, 对信息的分布、离散程度、异常值等特征进行智能检测, 并采用信息质量监控工具实现实时评估。

3.2 多维据数的质量评价

仅有信息验证方法还不足以全面提升多维据数质

量, 还要建立完善的评价体系, 其核心目标是对信息质量进行客观、量化的评估, 并基于评价结果实施校准优化, 包括: 信息不确定度计算, 量化信息的误差范围, 提高信息可信度估计; 信息质量评分, 通过加权评分机制衡量信息的完整性、准确性和一致性; 信息可信度评级, 基于历史信息质量和溯源信息, 对信息进行分级管理。从据数的角度规范信息处理流程, 探索多模态信息的组织和使用方法, 更高效、更准确地利用和发挥多模态信息的价值。

4 场景案例: 自动驾驶场景下多维据数质量评估方法

针对自动驾驶场景, 本文提出一种面向多维据数的质量评估方法, 整体设计遵循“据数”框架理念, 主要包括三个核心阶段: 多维据数采集的标准化、多维据数处理阶段质量评估以及动态质量监控。图2展示了自动驾驶场景下的多维据数质量评估方法的整体框架。

多维据数采集标准化: 自动驾驶系统通常集成多视角相机、激光雷达、毫米波雷达等环境感知传感

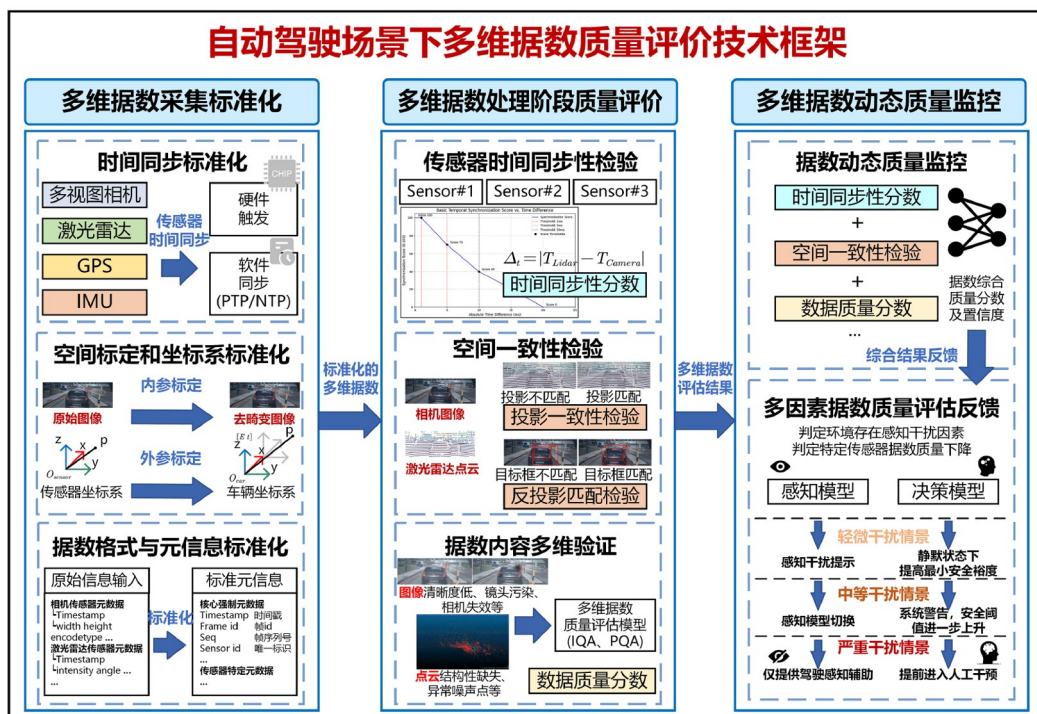


图 2 (网络版彩图)自动驾驶场景下多维据数质量评价技术框架

Figure 2 (Color online) Multi-dimensional evidence-based data quality evaluation technology framework in autonomous driving scenarios.

器,同时辅以全球导航卫星系统(global navigation satellite system, GNSS)与惯性测量单元(inertial measurement unit, IMU)实现高精度定位与运动状态估计.在时间同步标准化阶段,数据通过硬件触发器及软件时间同步方法,实现毫秒级乃至微秒级的多级时间同步,确保不同传感器记录同一环境状态的时间戳严格同步.其次,通过确定传感器的投影、畸变矩阵实现传感器的内参标定,车辆到传感器、传感器间的外参标定,保证相机、激光雷达等传感器的空间关系稳定,防止因车辆颠簸或热漂移导致的外参失效.采集源头质量的规范化不仅减少了后续算法在信息对齐、纠错上的额外开销,还直接决定了后续融合感知的空间准确度与稳定性,为后续的融合处理和评估等奠定坚实基础.

多维数据处理阶段质量评价:采集的多维数据需要在处理阶段进行系统化质量检验与多模态融合一致性分析.首先,针对时间同步性,通过对多传感器时间戳差异进行统计与阈值检验,快速判定是否存在超限时间漂移,防止动态场景下物体在不同传感器视野中出现显著位置差异.其次,针对空间一致性校验,利用投影一致性和反投影验证方法,对激光点云投影至图像平面后,与原图像物体边缘、纹理进行比对,计算投影一致性或平均投影偏差,从而发现因标定退化或传感器失效引起的几何误差.此外,还需针对具体模态数据内容进行质量评估,如图像质量评估(image quality assessment, IQA)和点云质量评估(point cloud quality assessment, PQA)等模型,可分别检测图像异常模糊、过曝及点云稀疏等问题.此阶段不仅可剔除显著异常帧,还为下游任务提供逐帧的质量标签与可解释的质量依据,从而构建系统、全面的多维数据质量评估体系.

多维数据动态质量监控:在自动驾驶场景中,不仅需要传统静态感知信息的质量评价,更关键的是在动态行驶过程中还持续监测多传感器间的时间同步性、空间一致性及动态信息质量的实时变化.此类动态指标具有显著的时序依赖与跨模态关联特性,往往难以通过非时序或线性模型准确捕捉.数据动态质量监控的核心在于,基于数据自身特征、算法反馈与传感器间物理约束,构建覆盖多维度与多尺度的质量评估体系,持续追踪质量演变过程,输出实时的综合质量分数与置信度,为系统提供直接的质量支撑.

多维数据质量评价方法对自动驾驶系统有重要影

响.在模块化架构中,各感知、决策与控制模块相互解耦,质量评估可为每个模块提供实时的传感器质量分数或置信度,使得下游算法能根据不同场景动态调整决策策略.例如,当多视角相机因高曝光或动态场景导致图像质量下降,目标检测模块可降低其权重;若激光雷达在恶劣天气下表现不佳,系统可增强视觉或毫米波雷达的依赖,通过冗余与鲁棒性设计,即便部分传感器退化,系统仍能稳定运行,并有助于问题溯源.对于端到端架构,随着多传感器数据直接输入到大规模多模态模型,影响变得更加直接且难以隔离.此时,通过动态质量分数结合时间同步、空间一致性及标定状态,可引导模型在整体置信度下降时输出更保守的决策,如减速或提前触发人工接管.当个别传感器质量下降时,注意力机制或门控单元可有效降低其负面影响,从而在复杂环境或局部传感器失效的情况下保持系统的安全性与可预测性,增强与驾驶员的交互信任.

5 总结与展望

综上所述,本文以多维数据作为研究对象,实现智能时代多模态信息的质量评价,在自动驾驶、智能机器人领域发挥重要作用.本文探讨了如何建立源头之“尺”以规范采集源,通过构建标准化采集装置、优化抗干扰设计及实时质量校验机制,为多模态信息提供源头质量保障;与此同时,相关领域法律法规和治理框架的逐步完善,也为多维数据理念在实际信息处理系统中的应用提供了制度支撑.在处理部分从数据角度组织多模态信息,构建多维数据的信息处理、验证及评价体系,引入时空对齐、动态对齐及语义融合等规范化流程,并结合现有信息质量评价与互操作标准,为多模态处理的标准化与推广奠定基础.对于通用大模型,本文分析了从数据角度增加纹理、几何等结构化属性标签,推动标准化三维感知数据的有效积累,增强其多模态泛化性的潜力.未来我们还将开展多维数据质量评价和决策的耦合机制、质量评价和端到端系统融合方案等方面的研究,进一步扩展多模态信息质量评价与技术的应用场景,并通过数字化治理平台、可信计算与可解释人工智能的引入,可以确保“数据”在系统实现中具备可操作性与现实意义.同时,随着自动驾驶、智能机器人等应用场景的不断拓展,多

模态信息质量评价将在跨领域协同、实时动态适配以及面向大规模知识驱动系统的可信决策中发挥更加关键的作用。最终, 提升多模态信息的处理质量和可信

性, 为“国家质量基础设施体系”填补此领域空白, 加速数据驱动智能系统的发展, 从而推动智能化社会的可持续发展。

致谢 本文撰写过程中, 得到中国人工智能学会名誉理事长李德毅院士, 中国人工智能学会理事长戴琼海院士, 北京工业大学党委副书记乔俊飞教授, 中国标准化研究院质量研究分院原院长刘碧松研究员, 浙江万丰科技开发股份有限公司副总裁、研究院院长郑建明教授级高级工程师的支持和指导, 在此表示最诚挚的感谢。

参考文献

- 1 Tu Z P. Data Civilization: How Big Data Reshapes Human Civilization, Commercial Patterns, and the Personal World (in Chinese). Beijing: CITIC Publishing Group Co., Ltd., 2018. 48: 44 [涂子沛. 数文明: 大数据如何重塑人类文明、商业形态和个人世界. 北京: 中信出版集团股份有限公司, 2018. 48: 44]
- 2 Che W X, Dou Z C, Feng Y S, et al. Towards a comprehensive understanding of the impact of large language models on natural language processing: Challenges, opportunities and future directions (in Chinese). *Sci Sin Inf*, 2023, 53: 1645–1687 [车万翔, 窦志成, 冯岩松, 等. 大模型时代的自然语言处理: 挑战、机遇与发展. 中国科学: 信息科学, 2023, 53: 1645–1687]
- 3 Zhang H, Li K, Chen B, et al. A task and motion planning method for dual-arm robots in biochemical experiments: Large multimodal model-driven intelligent decision making (in Chinese). *Sci Sin Tech*, 2025, 55: 1689–1704 [张辉, 李康, 陈波, 等. 面向生化实验的双臂机器人任务与运动规划方法: 多模态大模型驱动的智能决策. 中国科学: 技术科学, 2025, 55: 1689–1704]
- 4 Shen C X, Zhang H G, Feng D G, et al. An overview of information security (in Chinese). *Sci China Ser F: Inf Sci*, 2007, 2: 129–150 [沈昌祥, 张焕国, 冯登国, 等. 信息安全综述. 中国科学F辑: 信息科学, 2007, 2: 129–150]
- 5 Tao X M, Peng J Y, Wang Y, et al. Information metrics for data components based on information entropy: Data pricing and its application analysis for electric energy statistics (in Chinese). *Sci Sin Inf*, 2025, 55: 654–680 [陶晓明, 彭劼扬, 王钺, 等. 基于信息熵的数据元件信息计量——以数据定价及其电能统计应用分析为例. 中国科学: 信息科学, 2025, 55: 654–680]
- 6 Schwabe D, Becker K, Seyferth M, et al. The METRIC-framework for assessing data quality for trustworthy AI in medicine: A systematic review. *npj Digit Med*, 2024, 7: 203
- 7 Qin Y Z. Embracing the digital era and accelerating the development of metrology (in Chinese). *China Metrol*, 2022, 6: 15–17 [秦宜智. 拥抱数字时代, 加快计量发展. 中国计量, 2022, 6: 15–17]
- 8 Briney K. Data Management for Researchers: Organize, Maintain and Share Your Data for Research Success. UK: Pelagic Publishing Ltd., 2015
- 9 Xia J Z, Li J, Chen S M, et al. A survey on interdisciplinary research of visualization and artificial intelligence (in Chinese). *Sci Sin Inf*, 2021, 51: 1777 [夏佳志, 李杰, 陈思明, 等. 可视化与人工智能交叉研究综述. 中国科学: 信息科学, 2021, 51: 1777]

Multimodal information quality assessment and technical challenges

FANG Xiang¹, MA Nan², HAN YiHeng², XIA Qin³, WU ZhiXuan⁴, XIAO RongChun²,
LIU YiGuang⁵, XIONG XingChuang¹ & HUANG Jin⁶

¹ National Institute of Metrology, China, Beijing 100029, China

² School of Information Science and Technology, Beijing University of Technology, Beijing 100124, China

³ China Automotive Engineering Research Institute Co., Ltd., Chongqing 401122, China

⁴ College of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

⁵ College of Computer Science, Sichuan University, Chengdu 610065, China

⁶ School of Vehicle and Mobility, Tsinghua University, Beijing 100084, China

The rise of perceptual intelligence, multimodal information fusion, and vision-language models has led to increasingly diverse and complex information acquisition in the intelligent era. This diversity introduces significant challenges for the processing, calibration, and quality assessment of heterogeneous, multi-source information. Addressing these challenges, this paper proposes an evidence-based data framework for multimodal information quality evaluation, aiming to enhance reliability and utilization efficiency. We first establish a standardized information acquisition process that defines clear information source boundaries, rational selection of acquisition methods, and unified information formats. Next, we design a normative processing workflow that incorporates multidimensional validation, evaluation metrics, and calibration techniques to ensure information completeness, accuracy, consistency, timeliness, and accessibility. To demonstrate the practical applicability of our approach, we present a case study on multi-sensor information processing in autonomous driving, evaluating the quality of multimodal information processing and fusion from an evidence-based data perspective. The proposed framework offers a systematic methodology for credible quality assessment of multimodal information in intelligent systems. It contributes to the development of national quality infrastructure and supports the high-quality advancement of artificial intelligence technologies by ensuring that diverse and complex information can be processed and evaluated with consistency and rigor.

evidence-based data, information credibility, information acquisition sources, quality assessment, autonomous driving

doi: [10.1360/SST-2025-0251](https://doi.org/10.1360/SST-2025-0251)