

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(XXXX)XX-0001-40

论文引用格式: Liu Xuelin, Yan Jiebin, Fang Yuming, Hou Jingwen, Zhu Hanwei, Sui Xiangjie. A Survey on Quality of Experience Evaluation for Virtual Reality Content[J/OL]. Journal of Image and Graphics, XXXX: 1-40. DOI: 10.11834/jig.260101. (刘学林, 鄢杰斌, 方玉明, 侯敬文, 诸汉炜, 眭相杰. 面向虚拟现实内容的用户体验质量评价综述[J/OL]. 中国图象图形学报, XXXX: 1-40. DOI: 10.11834/jig.260101.) [DOI: 10.11834/jig.260101]

面向虚拟现实内容的用户体验质量评价综述

刘学林¹, 鄢杰斌¹, 方玉明¹, 侯敬文¹, 诸汉炜¹, 眭相杰²

1. 江西财经大学计算机与人工智能学院, 南昌 330032; 2. 澳门城市大学数据科学学院, 澳门 999078

摘要: 随着多媒体技术的飞速发展, 虚拟现实(virtual reality, VR)技术以及沉浸式媒体日益受到重视, 已逐渐成为一个重要的研究和开发领域。由360°相机捕获的VR内容, 如全景图像和全景视频, 相比传统2D视频图像能够提供更为丰富和完整的周围环境细节。凭借其全方位视角与沉浸式的观看体验, 全景图像和全景视频能够增强用户的真实性、互动性和沉浸感, 因此已在众多领域中得到广泛应用。然而, VR内容的固有球形特性对其存储与传输提出了巨大挑战。此外, 在内容采集到终端显示的复杂处理流程中, VR内容都可能引入导致质量下降的潜在因素, 这些因素共同影响用户的最终体验。因此, 准确描述和量化整个处理流程中感知损伤, 并提出有效评价用户体验质量的方法至关重要。早期用户体验质量评价主要依赖吞吐量、带宽、延迟、抖动和丢包率等网络性能参数, 但难以全面反映多媒体服务中网络波动对用户体验的影响。为弥补这一不足, 体验质量作为关键方案被提出, 通常指用户使用应用或服务时的满意程度, 受用户期望、个性特征以及当前状态影响。体验质量强调以用户为中心, 能更准确地反映整体体验。其量化过程将系统性能与人类感知转化为可测指标, 并可通过主观和客观两种方式进行测量。尽管已有不少研究致力于面向VR内容的用户体验质量评价, 但目前仍缺乏对这一领域现有方法的系统综述, 这一定程度上制约了对该领域的深入理解与分析。本文对面向虚拟现实内容的用户体验质量的最新研究进行全面梳理, 旨在推动对该领域的统一认识。首先, 通过识别并归类当前虚拟现实内容体验质量评价方法的关键特征, 建立了一套分类体系与特征描述, 从而构建出用于组织现有方法的结构化框架。进一步地, 本文将虚拟现实内容体验质量评价技术划分为两个主要阶段, 即主观评价与客观评价。一方面, 本文对各类VR内容主观数据集的特点及其构建流程进行了分析, 并探讨了影响因素、测试方法、环境设置及评分步骤等关键要素。另一方面, 本文根据VR内容的不同表达形式, 将面向VR内容的客观体验质量评价方法划分为基于二维平面、基于球面和基于视口的方法三类。其中, 基于2D平面的方法主要补偿投影伪影, 分为启发式方法和数据驱动方法。基于球面的方法直接在球域操作, 采用全局或局部采样策略。基于视口的方法从多视角评价内容质量, 按采样策略分为基于规则、关键点和扫视路径三类。在完成各类模型分类阐述后, 还深入剖析了不同方法的优缺点。此外, 本文对面向VR内容体验质量评价模型性能的评价指标进行了归纳, 并探讨了客观评价模型的实际应用场景。最后, 本文对当前面向VR内容的用户体验质量主客观评价研究进行总结, 剖析该领域尚未解决的研究局限以及当前面临的障碍, 并指出未来可能的发展方向。

关键词: 虚拟现实; 体验质量; 质量评价; 主观评价; 客观评价

A Survey on Quality of Experience Evaluation for Virtual Reality Content

收稿日期: 2026-02-14; 修回日期: 2026-09-25

基金项目: 国家自然科学基金项目(62562034, 62461028, 62501257, 62132006, U24A20220); 江西省自然科学基金青年直接支持项目(20252BAC200010); 江西省职业早期青年科技人才培养项目(20252BEJ730134); 江西省管理科学项目(20262BAA100074); 澳门科学技术发展基金项目(0007/2025/AMJ, 0001/2025/AIJ)

Supported by: National Natural Science Foundation of China(62562034, 62461028, 62501257, 62132006, U24A20220); Natural Science Foundation of Jiangxi Province (20252BAC200010); Early-Career Program for Young Science and Technology Talents of Jiangxi Province (20252BEJ730134); Management Science Foundation of Jiangxi Province (20262BAA100074); Science and Technology Development Fund (0007/2025/AMJ, 0001/2025/AIJ)

Liu Xuelin¹, Yan Jiebin¹, Fang Yuming¹, Hou Jingwen¹, Zhu Hanwei¹, Sui Xiangjie²

1. School of Computing and Artificial Intelligence, Jiangxi University of Finance and Economics, Nanchang 330032, China; 2. Department of Data Science, City University of Macau, Macau 999078

Abstract: With the rapid development of multimedia technologies, virtual reality (VR) and immersive media have attracted increasing attention and have gradually become an important area of research and development. VR content, such as omnidirectional images (OIs) and omnidirectional videos (OVs) captured by 360° cameras, provides richer and more complete surrounding environmental details than conventional perspective images. Benefiting from their panoramic viewing angles and immersive viewing experiences, omnidirectional images and videos enhance users' sense of realism, interactivity, and immersion, and have therefore been widely applied across numerous fields. However, the inherent spherical nature of VR content poses significant challenges to both storage and transmission. In addition, throughout the complex processing pipeline from content acquisition to terminal display, VR content may introduce various potential factors that cause quality degradation, which collectively affect the final user experience. Therefore, it is crucial to accurately characterize and quantify perceptual impairments across the entire processing pipeline and to develop effective methods for evaluating user quality of experience. In previous exploratory research on multimedia quality assessment, early user quality of experience (QoE) evaluation schemes mainly depended on objective network performance parameters, including throughput, bandwidth, transmission delay, jitter, and packet loss rate, to evaluate multimedia service quality. However, these single network-oriented technical indicators cannot fully reflect the complex correlation between dynamic network changes, content distortion characteristics, and human subjective visual perception, which shows obvious deficiencies in describing real user experience for high-immersive VR media. To effectively compensate for such technical limitations, the user-centered QoE evaluation system has been widely proposed and recognized as the core evaluation solution for modern multimedia services. In general, QoE refers to the comprehensive satisfaction level perceived by end users during the use of specific multimedia applications or services, which is comprehensively affected by user psychological expectation, individual personality traits, and real-time viewing physiological and psychological states. Different from traditional objective technical evaluation, QoE adheres to a user-centric evaluation concept and can more comprehensively and accurately reflect the overall service experience of immersive media. Its quantitative evaluation mechanism converts abstract system operating performance and human visual perception rules into standardized measurable indicators and mainly includes two mainstream technical branches, namely subjective manual measurement and objective algorithm measurement. Although existing scholars have carried out abundant research explorations on QoE evaluation for VR content, systematic, comprehensive, and standardized sorting and in-depth summary of existing evaluation theories, technical frameworks, and algorithm methods are still insufficient, which greatly restricts the holistic understanding, systematic analysis, and further innovative development of this research field. This paper conducts a comprehensive and targeted review of the latest research progress and technical achievements on VR content QoE evaluation, aiming to establish a unified, standardized, and systematic theoretical cognition for this research domain. First, by sorting, identifying, and refining the core technical characteristics of current VR content QoE evaluation methods, this paper constructs a complete classification system and detailed feature description, forming a standardized structured framework for combing existing research approaches. Furthermore, this paper divides mainstream VR content QoE evaluation technologies into two core research directions, including subjective evaluation and objective evaluation. In terms of subjective evaluation, we systematically analyze the construction rules, structural characteristics, and building processes of various open-source VR subjective evaluation datasets and comprehensively discuss core research elements including major influencing factors, standardized testing schemes, experimental environment configuration, and unified scoring specifications. In terms of objective evaluation, according to the diverse spatial representation forms and rendering characteristics of VR content, existing objective QoE evaluation methods are classified into three typical categories, namely 2D-plane-based, sphere-based, and viewport-based evaluation methods. Among them, 2D-plane-based methods focus on compensating geometric deformation and projection artifacts caused by spherical plane mapping and can be further divided into traditional empirical heuristic methods and data-driven methods. Sphere-based methods directly implement quality calculation and distortion analysis in the original spherical domain through global full-scene

sampling or local fine sampling strategies. Viewport-based methods fully simulate human eye active viewing habits and evaluate local visual quality from multiple dynamic viewing angles, which can be subdivided into rule-based, keypoint-based, and scanpath-based methods according to different sampling logic. On the basis of categorical elaboration of various evaluation models, this paper conducts an in-depth comparative analysis of the inherent merits, practical limitations, and applicable conditions of different technical approaches. Meanwhile, we summarize mainstream quantitative performance evaluation metrics for verifying the effectiveness of VR QoE evaluation models and explore the practical application boundaries and typical application scenarios of various objective evaluation models. Finally, this paper systematically concludes the overall research status of existing VR content QoE subjective and objective evaluation studies, summarizes unresolved technical bottlenecks and current research obstacles in this field, and effectively prospects several valuable potential research directions for future follow-up exploration and innovation.

Key words: virtual reality; quality of experience; quality assessment; subjective evaluation; objective evaluation

论文引用格式: Liu Xuelin, Yan Jiebin, Fang Yuming, Hou Jingwen, Zhu Hanwei, Xiangjie Sui. A Survey on Quality of Experience Evaluation for Virtual Reality Content [J/OL]. Journal of Image and Graphics. DOI: 10.11834/jig.260101. (刘学林, 鄢杰斌, 方玉明, 侯敬文, 诸汉炜, 睦相杰. 面向虚拟现实内容的用户体验质量评价综述[J/OL]. 中国图象图形学报. DOI: 10.11834/jig.260101.)

0 引言

随着多媒体和信息技术的快速发展以及用户需求的不断提升,沉浸式媒体受到广泛关注,并逐渐成为研究与应用的重要方向。全景图像(omnidirectional images, OIs)和全景视频(omnidirectional videos, OVIs),通常被称为360°内容,能够捕捉观看者周围360°×180°的完整视域范围。这类内容不仅为用户带来了前所未有的交互体验,也提供了更高自由度的场景探索可能。这些技术进步显著改变了用户与媒体内容的互动模式,使用户能够自主选择观看方向,而不再像传统媒体那样受限于固定的预设视角。

在虚拟现实(virtual reality, VR)应用中,OIs和OVIs发挥着关键作用。它们通常采用球形内容表示形式,相比传统二维(two-dimensional, 2D)媒体,在真实性、交互性和沉浸感方面具有明显优势。随着商业头戴式显示器(head-mounted displays, HMDs)的日益普及,用户得以通过自然的头部运动自由探索360°全景环境,获得独特的沉浸式观看体验。然而,高质量全景内容的处理与传输也带来了显著的技术挑战。OIs和OVIs因其固有的球形特性,通常需

要超高的分辨率,这给传输带宽提出了极高要求。此外,在VR内容从采集到最终显示的复杂流程中,多个环节均可能引入质量损失,影响整体观感效果。

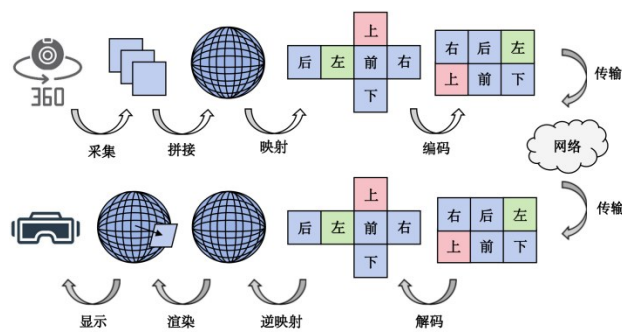


图1 VR内容处理流程

Fig. 1 VR content processing pipeline

如图1所示,VR媒体内容的完整处理流程包含多个关键阶段(Azevedo等,2020b;Lopes等,2018)。 workflow始于采集阶段,通过多个时间同步的相机从不同方向捕获2D图像,共同覆盖完整的球形视域。随后在拼接过程中,这些图像经过配准、对齐与融合,形成无缝的非平面球形表示(Ullah等,2022)。为了便于传输,球形内容会通过投影技术转换为平面格式,常见的投影方式包括等距柱状投影(equirectangular projection, ERP)、立方体贴图投影(cubemap projection, CMP)、截顶方锥投影(truncated square pyramid projection, TSP)和分段球形投影(segmented sphere projection, SSP)等,如图2所示。接着,平面表示被进一步编码,并可结合分块策略将内容划分为独立编码区域,以适配视口自适应流媒体的需求,从而优化带宽使用(Chen等,2018b)。在VR视频系统中,视口是指从球形内容中实时提取并

渲染的、由用户当前注视方向确定的有界视场(field of view, FoV)区域。编码后的内容经标准通信信道传输后,需进行解码与逆映射以恢复球形表示。最后,在渲染阶段根据用户注视方向生成对应的视口2D投影:HMDs通过运动跟踪提供沉浸式观看体验,而传统设备则通过导航交互实现视角控制。在这一复杂处理链中,每个阶段都可能引入不同类型的失真,共同影响最终感知质量(Yan等,2022)。这些挑战主要源自但不限于拼接过程中图像配准与融合算法的技术局限、投影与压缩操作导致的失真,以及传输环节中的带宽限制。这些因素凸显了开发鲁棒的质量评价方法的迫切性,这些方法必需能够准确表征并量化全流程中的感知损伤。目前,研究持续致力于应对上述挑战,旨在兼顾高沉浸感体验的同时,满足现代VR应用对画质的高标准要求。

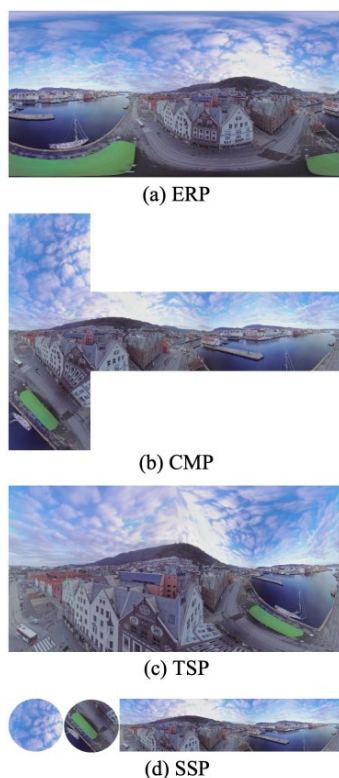


图2 多种投影示意图

Fig. 2 Illustration of various projections

早期研究通过设计服务质量(quality of service, QoS)指标来评价用户体验。服务质量是从技术层面评价网络性能的方法,其通过吞吐量、带宽、延迟、抖动和丢包率等可测量参数来量化网络表现。广义上,QoS可被视为一系列保障网络性能可预测性的技术和测量工具(Möller和Raake,2014),是评价多

媒体应用和网络效率的关键指标(Akhtar和Falk,2017)。其主要特征涵盖系统性能、响应性、服务可用性等参数,以及与应用相关的各类需求。然而,传统QoS指标难以全面反映网络波动对多媒体服务用户体验的影响。因此,体验质量(quality of experience, QoE)日益受到重视。QoE最初定义为用户在与应用或服务交互过程中所感受到的愉悦或烦恼程度,其受用户、个人特征及当前状态等因素的影响(Callet和Moller,2013; (Moller和Raake,2014)。与主要关注技术性能的QoS不同,QoE进一步引入了以用户为中心的维度,如可用性、可接受性和满意度,从而扩展了评价框架。此外,QoE还受到心理与认知因素的影响,包括用户期望、情绪状态、行为模式和即时反应等。将QoE量化为可解释、可测量的指标,对于建立系统性能与人类感知之间的关联至关重要,也为服务优化提供了可操作的依据。QoE的量化可通过“直接QoE测量”(即主观QoE)或“间接QoE测量”(即客观QoE)实现。其中,主观QoE通过在受控实验环境中由人类受试者进行评价,而客观QoE则致力于开发能够自动估计QoE的计算模型。

针对VR内容这一主题,已有多篇相关综述文献。表1系统梳理了近十年来发表的相关综述文献,展现了该领域的研究脉络与发展动态。由于VR内容的处理流程较为复杂,现有综述大多聚焦于流程中的特定环节或具体技术方向。具体而言,Fan等人(2019)和Yaqoob等人(2020)重点探讨了360°视频流传输及自适应传输框架;Yuan和Xie(2021)侧重于VR视频流传输中的用户体验质量建模;Fang等人(2023)则系统回顾了基于分块的视口自适应视频流系统中的视口预测方法;Jeroen等人(2023)从全景视频和全息技术角度对沉浸式视频传输进行了梳理与介绍;Huang等人(2026)则从人工智能赋能的角度对体积视频传输系统进行系统性综述。除网络传输与流媒体技术外,也有部分综述关注VR内容自身特性及其评价方法。Elwardy等人(2019)梳理了公开标注的VR图像与视频数据集;Azevedo等人(2020b)分析了传统视觉通信框架下360°信号常见的视觉失真;Chen等人(2018b)回顾了全景视频投影与评价方面的近期进展;Zeng等人(2022)对沉浸式视频主观质量评价数据库和客观质量评价方法进行了介绍;Zhou等人(2022c)和Sui等人(2024)则

对全景图像客观质量评价方法进行了综述。此外,也有研究工作尝试覆盖处理流程中的多个阶段。Xu 等人(2020)讨论了当前 360° 内容处理的先进技术,重点关注感知建模、质量评价与内容压缩;Chiariotti 等人(2021)对沉浸式视频流系统进行了整体概述,涵盖编码技术、用户体验质量及流传输等;

Biswas 等人(2024)对 VR 环境下的晕动症及其影响因素进行了全面回顾,并进一步总结与分类;Duan 等人(2024)则综述了沉浸式媒体(包括 VR、增强现实、混合现实及三维内容)中的人类感知研究及其应用。

表 1 相关综述汇总
Table 1 Summary of the existing surveys

文献	出版	范围
Chen 等人(2018b)	SP	投影格式、评价方法
Elwardy 等人(2019)	ICSPCS	数据库
Fan 等人(2019)	ACM CSUR	采集、传输、显示
Azevedo 等人(2020b)	IEEE TCSVT	视觉失真
Xu 等人(2020)	IEEE TCSVT	注意力、评价、压缩
Yaqoob 等人(2020)	IEEE COMST	流传输
Chiariotti(2021)	CC	编码、体验质量、流传输
Ruan 和 Xie(2021)	Electronics	流传输、体验质量
Zeng 等人(2022)	信号处理	VR 视频质量评价
Zhou 等人(2022c)	中国图象图形学报	VR 图像客观质量评价
Fang 等人(2023)	/	视口预测
Jeroen 等人(2023)	IEEE COMST	VR 视频传输
Biswas 等人(2024)	ACM CSUR	晕动症
Sui 等人(2024)	APSIPA	全景图像客观质量评价
Duan 等人(2024)	IEEE OJID	人类感知
Huang 等人(2026)	IEEE COMST	体积视频传输

本文与现有综述工作的区别主要体现在以下几个方面。1)视角综合性与全面性:以往综述多聚焦于 VR 内容处理流程中的独立环节,如流媒体传输(Fan 等,2019;Yaqoob 等,2020)、编码技术(Chiariotti,2021)、投影方法(Chen 等,2018b)、数据集(Elwardy 等,2019)、感知研究(Xu 等,2020)及评价手段(Zhou 等,2022c;Sui 等,2024;Duan 等,2024)等。这些研究虽在各领域内提供了有价值的总结,但因视角相对局限,未能系统呈现 VR 内容用户体验质量评价的整体图景。相比之下,本综述并不局限于特定技术环节,而是对影响用户体验质量的各方面因素进行全面审视与综合分析。2)专精深度与广泛覆盖相结合:尽管许多研究方法同时适用于 VR 图像与视频,本综述特别聚焦于 VR 内容的用户体

验质量评价。在此范畴内,本文系统评述了两百余篇相关文献,确保对该领域实现充分覆盖。同时,文中也提供了基础性内容,以增强文章的可理解性与可访问性,尤其便于新进入该领域的研究者参考。与以往仅关注单一阶段或任务的综述不同,本研究提供了贯穿 VR 内容处理流程的综合视角。

综上所述,本文旨在对面向 VR 内容的用户体验质量评价最新进展提供系统而整体的概述,而对尚处于起步和探索阶段的六自由度虚拟现实内容、增强现实或混合现实内容以及多模态体感交互的 QoE 问题进行前瞻性讨论。具体贡献如下:

1)全面性综述:本文从端到端流程的视角出发,对近期面向 VR 内容的用户体验质量相关研究进行了系统分析,致力于推动对该领域形成统一、整合的

理论认知。

2)分类体系与特征归纳:本文归纳了当前VR内容用户体验质量评价方法的定义性特征,构建了一套结构化的分类体系,以系统梳理现有研究脉络。

3)用户体验质量评价方法分类:基于既有研究,本文将VR内容用户体验质量评价技术划分为主观评价与客观评价两大范式。进一步地,根据其内在动机与技术原理,对相应评价方法进行了细致划分对比和深入论述。

4)对开放挑战的深入探讨:在系统梳理文献的基础上,本文还对领域内尚未解决的研究局限、当前面临的挑战以及未来潜在发展方向进行了审视与前瞻性展望。

1 主观体验质量评价

1.1 影响因素

用户体验质量的影响因素(influence factors, IFs)是指用户、系统、服务、应用或环境背景中,任何可能因其实际状态或属性设置而对用户体验质量产生影响的特征。一般而言,这些影响因素可归纳为四个主要类别:用户层面(与人相关)、系统层面(与系统相关)、情境层面(与环境背景相关)以及内容层面(与内容相关)(Bouraqia等,2020),具体分类详见表2。具体而言,用户层面关注用户的心理因素,包括其期望、过往经验以及当前所处的时段。系统层面涵盖与网络性能(如带宽、时延)、终端设备特性(如系统配置、电池续航)以及应用层因素(如编解码策略、性能表现)相关的技术要素。情境层面包括影响用户体验的环境因素,例如所处地理位置和使用场景)。内容层面则指特定服务在构成与功能上的固有属性及区分特征,例如编码码率、分辨率与播放时长等(Juluri等,2016)。

1.2 主观评价方法

主观评价方法主要包括两种范式:单刺激法与双刺激法(Akhtar和Falk,2017)。单刺激法仅向受试者呈现待评价样本,不提供参考信息;而双刺激比较法则同时呈现参考样本(原始内容)与处理样本(测试内容),以支持直接质量对比。评分可采用数值或等级量表,并以连续或离散的形式实施。根据评价时机不同,评价协议也有所差异。在刺激呈现过程中进行的瞬时评分用于捕捉质量随时间的变化,而在刺激呈现后的评价则提供整体质量判断。

国际电信联盟(international telecommunication union, ITU)提出的标准化绝对等级评分(absolute category rating, ACR)法(Akhtar和Falk,2017)是目前应用最广泛的方法,其采用五级量表,1~5分主观评分分别对应坏、差、一般、好和很高五个体验质量等级。通过统计受试者的反馈得到单刺激评价产生平均意见得分(mean opinion score, MOS),即某一刺激获得的所有评分的算术平均值;双刺激范式则生成差分平均意见得分(differential MOS, DMOS),反映参考样本与测试样本间的平均感知差异。ITU对多种评价框架进行了标准化,针对不同的评价场景提出了具体的流程要求。表3对经典主观评价方法进行了比较,涵盖绝对等级评分/单刺激法(single-stimulus, SS)、隐藏参考的绝对等级评分(ACR with hidden reference, ACR-HR)、双刺激损伤等级(double-stimulus impairment scale, DSIS)/退化分类评分(degradation category rating, DCR)、双刺激连续质量尺度法(double-stimulus continuous quality-scale, DSCQS)、成对比较法(pairwise comparison, PC)、视频质量主观评价方法(subjective assessment methodology for video quality, SAMVIQ)、单刺激连续质量评价法(single stimulus continuous quality evaluation, SSCQE)以及同时双刺激连续评价法(simultaneous double stimulus for continuous evaluation,

表2 体验质量影响因素分类

Table 2 Summary and examples of influence factors of QoE

类别	示例
人类因素	年龄、性别、教育水平、情绪、精神状态、偏好、喜好、期望、个性、记忆等
系统因素	网络、设备、时延、丢包、信任程度、可获得性等
上下文因素	目的、背景、地点、时间、环境、代价等
内容因素	视频内容、编码类型、分辨率、时长、帧率、流行度、情感模式等

SDSCE)等。

表3 主观评价方法对比

Table 3 Comparison of typical subjective evaluation methods

方法	ITU 文档	序列类型	参考	评分规则	评分尺度
ACR/SS	ITU-R Rec. BT.500 (ITU, 2012), ITU-T Rec. P910 (ITU, 2008)	单一	/	仅测试序列	ITU-R 质量尺度
ACR-HR	ITU-T Rec. P910	单一	单独显示	仅测试序列	ITU-R 质量尺度、DMOS
DSIS	ITU-R Rec. BT.500, ITU-T Rec. P910	双序列	显式	仅测试序列	ITU-R 质量尺度
DSCQS	ITU-R Rec. BT.500, ITU-T Rec. P910	双序列	隐式	测试序列、参考序列	ITU-R 质量尺度
PC	ITU-T Rec. P910	双序列	/	测试序列	二选一
SAMVIQ	ITU-R Rec. BT. 1788 (ITU, 2007)	多序列	显式	测试序列、参考序列	连续尺度 0~100
SSCQE	ITU-R Rec. BT.500	单一长序列	/	测试序列	连续尺度 0~100
SDSCE	ITU-R Rec. BT.500	双长序列	显式	测试序列、参考序列	连续尺度 0~100

1.3 主观测试环境

有效的主观质量评价需要在方法严谨性与场景真实性之间取得平衡,即既要确保实验过程的科学性,又要尽量贴近真实的观看环境。图3展示了面向VR内容的主观评价实验设置。目前主流的测试方法主要基于以下两种环境范式(Yang等,2018b):

● 实验室环境:作为感知质量评价的“黄金标准”,受控的实验室环境能够对实验参数和观看条件进行精确控制(Moorthy等,2012)。然而,该方法也存在一些局限,如实验搭建与参与者招募成本较高,参与者群体相对单一,通常局限于高校学生或研究人员,人工设置的实验条件可能难以完全还原真实使用场景等。

● 众包环境:作为新兴的替代方案,众包评价平台在通用QoE评价中展现出多方面优势(Hoßfeld等,2014; Ghadiyaram和Bovik,2016),包括效益高、参与者来源多样以及良好的可扩展性。商业平台(如Amazon mechanical turk、CrowdFlower)与专业框架(如crowdMOS、QualityCrowd)为基于网络的主观测试提供了基础设施。这些平台既支持标准化的测试流程,也能灵活适应基于浏览器的实验部署。

1.4 主观评分

当前用于用户体验质量评价的主观评分主要包括两类:平均意见得分和差分平均意见得分(Hoßfeld等,2011)。具体说明如下:

1)平均意见得分(MOS):作为感知质量评价的基础心理测量指标,MOS代表了所有受试者对给定



(a) 主观实验环境



(b) 主观实验流程

图3 主观评价实验设置示意图

Fig. 3 Illustration of subjective experiment setup

测试刺激所给评分的算术平均值。计算时,首先将每位受试者对该刺激的原始评分转换为标准化Z分数,转换过程中需先计算所有受试者评分的均值和标准差。在剔除不符合ITU-R BT500-11标准中可靠性要求的受试者后,再将剩余有效受试者的Z分数线性缩放到指定的百分制范围内。最终,该刺激的MOS值即为所有有效受试者缩放后得分平均值。

2)差分平均意见得分(DMOS):该指标用于量化处理后的媒体相对于其参考版本在感知质量上的退化程度。计算时,先求出每位受试者对参考版本与处理版本的原始评分差值,再对该差值应用与MOS相同的标准化流程(包括Z分数转换及后续缩放),最终得到DMOS值。

1.5 头戴式显示设备

先进头戴式显示设备的出现,显著提升了VR内容所能实现的用户体验质量。这些设备凭借宽广的视野提供深度沉浸的观看环境,使用户能够更自然地与内容互动。随着技术演进,各类头戴式显示器不断推陈出新,共同推动了VR应用的发展。表4汇总了当前广泛用于观看VR内容的头戴式显示设备,图4则以可视化方式展示了其设计与外观特征。

如表4和图4所示,为优化沉浸感、舒适度和可用性而设计的先进技术特征,已广泛集成于现代头戴式显示器之中。当前主流设备呈现出以下关键特点:首先,普遍采用高分辨率OLED或LCD显示屏,具备较高的像素密度(pixels per inch,PPI),既有效缓解了纱窗效应,也显著提升了视觉保真度。其次,提供90°至120°的宽广视场角,接近人类自然双目视觉范围,有力增强了感知沉浸感。第三,为降低晕动症影响,设备通常支持90Hz或更高的刷新率,以实

现流畅的运动渲染效果。第四,集成空间音频系统,无需外接设备即可维持高质量的听觉

沉浸。第五,现代头戴式显示设备在架构上支持独立运行与有线连接高性能渲染两种模式。先进型号还配备眼动追踪以实现注视点渲染,并搭载精准的手部追踪功能,为用户提供直观、自然的交互体验。这些特征共同定义了当代头戴式显示器,在技术性能、用户可访问性与功能适应性之间实现了平衡。

1.6 晕动症

通过头戴式显示器观看VR视频可能引发模拟器眩晕,也称为晕动症,其症状包括恶心、头晕、眩晕与出汗等(Kennedy等,1993;Balk等,2013;Duan等,2017a)。这些症状可能在VR环境暴露期间或之后出现,因此在沉浸式用户体验质量研究中需要对其加以密切监测。目前广泛使用的自评工具是模拟器眩晕问卷(simulator sickness questionnaire,SSQ)(Kennedy等,1993),它通

表4 主流头戴式显示器对比
Table 4 Comparison of the current widely used HMD

索引	设备	分辨率	PPI	视场角	刷新率	显示	年份
a	Oculus Rift DK2	1080×960	386	100°	75Hz	AMOLED	2014
b	Razer OSVR HDK	1920×1080	400	110°	90Hz	OLED	2015
c	Merge VR	1334×750	/	96°	/	/	2015
d	Google CardBoard	1920×1080	/	90°	60Hz	/	2016
e	Samsung Gear VR	2560×1440	466	101°	60Hz	LCD	2016
f	Oculus Rift	2160×1200	461	110°	90Hz	OLED	2016
g	HTC VIVE	2160×1200	448	110°	90Hz	OLED	2016
h	Lenovo Mirage Solo	2560×1440	615	110°	74Hz	LCD	2018
i	Pico G2	2880×1600	615	101°	90Hz	TFT	2018
j	Oculus Rift S	2560×1440	538	110°	80Hz	LCD	2019
k	HTC VIVE Pro Eye	2880×1600	615	110°	90Hz	OLED	2019
l	HTC VIVE Pro 2	4896×2448	950	120°	90/120Hz	LCD	2021
m	FOVE 0	2560×1440	/	100°	70Hz	OLED	2021
n	Apple Vision Pro	3660×3200	/	100°	90/96/100/120Hz	Micro-OLED	2023

过16个项目以四级量表对头痛、疲劳等症状进行量化(见表5)。各症状被归类为三个相互重叠的维度:恶心、眼动不适与定向障碍。每个维度的得分通过对症状分数求和并乘以预设系数得出,总模拟

器眩晕分数则由这三个维度分数加权组合得出。后续改进版本如虚拟现实眩晕问卷(VR sickness questionnaire,VRSQ)(Kim等,2018)根据经验数据删除了与恶心相关的问题,将项目精简至9项。其他工



图4 主流头戴显示器

Fig. 4 Illustrations of mainstream HMD

具如晕动症评价问卷 (motion sickness assessment questionnaire, MSAQ) (Gianaros 等, 2018) 则采用九点量表, 将 16 项症状划分为四个独立类别 (如胃肠道症状、中枢症状等)。为实现快速评价, 研究者引入了简化版工具, 例如痛苦指数量表 (misery scale index, MISC) (Bos 等, 2010)、快速晕动症量表 (fast motion sickness scale, FMS) (Keshavarz 和 Hecht, 2011) 与简化版模拟器眩晕问卷 (Tran 等, 2017), 此外还有用于参与者预筛选的晕动症易感性问卷 (motion sickness susceptibility questionnaire, MSSQ) (Golding, 1998)。Singla 等人 (2019b) 发现, 简化版问卷 (如简化版模拟器眩晕问卷) 虽无法完全替代综合问卷进行细节分析, 但能有效区分高致晕与低致晕内容。此外, 模拟器眩晕症状会随暴露时间延长而加剧, 但可通过间歇性休息进行缓解 (Singla 等, 2019c)。然而在实际场景中, 持续的主观或生理监测, 如脑电图 EEG (Kim 等, 2010), 仍然难以实现。

为突破这些局限, 研究者已转向客观的虚拟现实眩晕评价 (VR sickness assessment, VRSA) 方法。目前的计算模型不完全依赖生理信号 (Dennison 等, 2016) 或主观报告, 而是在观看过程中自动预测眩晕程度。例如, Kim 等人 (2019) 提出了一个深度生成网络, 包含 360° 视频生成器与眩晕分数预测器, 利用配对生理数据与主观反应的数据集进行训练。该模型通过分析运动模式 (运动感知相关眩晕的关键诱因) 对眩晕程度提供客观度量指标, 为传统评价方法提供了一种可扩展的替代方案。

1.7 临场感

VR 内容支持交互式场景探索, 使用户在虚拟环境中产生强烈的身体在场感。这一现象被称为“临场感”, 反映了用户在虚拟现实场景中主观体验到“身临其境”的感受, 其行为反应也高度接近真实

表5 模拟器眩晕问卷
Table 5 Simulator sickness questionnaire

索引	症状	权重		
		恶心	眼动不适	定向障碍
01	全身不适	1	1	
02	疲劳		1	
03	头痛		1	
04	眼睛疲劳		1	
05	难以集中		1	1
06	唾液增加	1		
07	出汗	1		
08	恶心	1		1
09	注意不集中	1	1	
10	头部充盈			1
11	视力模糊		1	1
12	睁眼眩晕			1
13	闭眼眩晕			1
14	晕头转向			1
15	胃部感知	1		
16	打嗝	1		
整体		[a]	[b]	[c]

世界中的互动 (Kalawsky, 2000)。目前, 临场感问卷 (presence questionnaire, PQ) 已成为量化该体验的标准工具, 它通过涵盖真实感、干扰因素、感官保真度和控制感四个维度的 32 个问题进行评估。参与者依据七点量表对每个维度进行评分, 并在部分维度上接受明确的描述性指引。为满足高效评价的需求, 研究者开发了简化版本, 例如简化版临场感问卷 (short PQ) (Tran 等, 2017), 将评价简化为单一点的五点量表, 提升了实际应用的便捷性。已有研究对上述测量工具在全景视频用户体验质量研究中的有效性进行了比较。Singla 等人 (2019b) 指出, 简化版临场感问卷虽能有效捕捉特定视频序列的整体临场感知, 可作为完整问卷的实用替代工具, 但其在详细

技术分析方面缺乏精细度。特别是在探究分辨率、码率和帧率等技术参数如何系统性地影响临场感知时,简化版问卷就稍显不足。因此,完整版临场感问卷在需要深入分析技术因素影响的实验场景中仍具有重要价值(Souza等,2021)。

1.8 数据库

本节综述了当前面向VR图像与视频用户体验质量研究的公开数据集。现有数据集主要分为三大类别:360°视觉注意力数据集、VR图像主观质量评价数据集以及VR视频主观质量评价数据集。下文将对上述基准数据集分别进行详细介绍。

1.8.1 360°视觉注意力

针对传统图像和视频等视觉刺激的人类观看行为研究,通过识别感知显著区域并优化用户体验,已在多媒体编码、传输、质量评价与渲染等多个领域发挥了重要推动作用。类似地,分析用户在沉浸式虚拟环境中的探索行为,对于VR内容设计、压缩算法开发以及显著性建模等应用亦至关重要。尽管近年来已有研究关注桌面观看条件下的用户体验质量,但VR环境因涉及立体视觉、头部朝向、注视动态与运动限制之间的复杂交互而存在本质差异。目前主流的360°内容视觉注意力数据集如表6所示。

1)360°图像视觉注意力数据集:360°图像视觉注意力研究起步较早。如表6所示,相关数据集历经从大规模无标注到小规模多模态标注的发展,并逐步引入语义信息,类型涵盖大规模通用型、注视/扫视行为型和语义分割型,当前仍受限于静态内容与样本规模,未来趋向动态内容、跨设备以及多受试拓展。这些数据集为理解VR视觉行为以及开发全景分析计算模型提供了重要基础。

2)360°视频视觉注意力数据集:如表6所示,近年来涌现多个用于探究VR用户行为与视觉注意的全景视频数据集,研究目标包括视口预测、显著性检测、头动/眼动建模以及体验质量评价。其发展呈现

从单一头动数据向融合眼动、情感与语义等多模态信息演进的趋势,硬件设备同步更新。相较于360°图像数据集,当前视频数据集在样本规模、内容多样性和标准化程度仍有较大提升空间。

小结:综上,现有数据集为沉浸式媒体中的视觉注意、运动预测、情感分析及内容适配等多领域研究提供了重要数据支撑。

1.8.2 VR图像主观质量评价

VR图像质量评价研究最早可追溯至2016年。目前,研究人员已构建了多个重要的全景图像质量评价(OI quality assessment, OIQA)数据集。表7汇总了近十年的主流VR图像主观质量数据库。如表7所示,在规模上,从早期的少量原图的探索性研究,跃升至万级样本的大规模数据库,为数据驱动模型提供了充足训练源。在内容上,多样性增强,分辨率提升,失真类型从传统JPEG、HEVC压缩逐步拓展至多重失真、拼接伪影、亮度不连续、观看条件变化、投影格式以及人工智能(artificial intelligence, AI)生成失真等VR特有失真,评价维度更加贴近真实沉浸式体验。在标注上,从单一

MOS分数发展为包含主观评分、头动轨迹以及眼动数据的多模态标注范式,为探索感知质量与视觉行为的内在关联开辟了新路径。在实验设置上,受试人数从20人扩展至200人,HTC VIVE系列成为设备标准,ACR协议也因其高效性被广泛采用。

小结:上述VR图像质量数据库共同展现了VR图像质量评价领域的发展脉络,然而,尽管当前取得了显著进展,仍然存在着若干亟待解决的问题。如跨设备一致性校准尚未建立、交互式VR内容的主观质量评价近乎空白、MOS评分尺度不统一为跨库模型训练带来挑战、眼动数据标准化问题亟待解决。未来应聚焦动态全景视频质量评价、交互场景下的感知质量建模,以及跨数据库统一评测基准的构建。

1.8.3 VR视频主观质量评价

表8 汇总了主流VR视频主观质量评价数据库及其分类。根据表8内容,可以得出以下结论。

表6 360°图像/视频显著性预测或头部/眼部追踪数据库
Table 6 Summary of existing databases for 360°image/video saliency detection or head/eye tracking

数据库	类型	场景数量	分辨率	时长	受试人数	设备	数据
SUN360(Xiao等,2012)	图像	67,538	9104×4552	/	/	/	/
Salient360!(Rai等,2017)	图像	60	5376×2688 ~18332×9166	/	63	Oculus DK2	眼动/头动数据、 扫视路径、显著图
ODI Saliency(Abreu等,2017)	图像	21	4096×2048	/	32	Oculus DK2	视口中心轨迹、显著图
VR Saliency(Sitzmann等,2018)	图像	22	/	/	213	Oculus DK2	头动方向、注视数据
F-360iSOD(Zhang等,2020)	图像	107	2048×1024	/	3	/	物体级掩膜
Moving Target(Bao等,2016)	视频	16	1080P, 2K, 4K	30	153	Oculus DK2	运动数据
360-Video-Database(Li等,2017)	视频	73	/	29~668	95	Oculus Rift	头动数据、愉悦度、兴奋度
360 Head Movement(Corbillon等,2017)	视频	5	3840×2048	70	59	Razer OSVR	头动数据
VR User Behavior(Wu等,2017)	视频	18	/	164~655	48	HTC VIVE	头动数据
VR-HM48★(Xu等,2017,2019b)	视频	48	2840×1440 ~7680×3840	20~60	40	HTC VIVE	头动数据
360 Dataset(Lo等,2017)	视频	10	4K	60	50	Oculus Rift	显著图、运动图、传感器数据
Sports-360(Hu等,2017)	视频	342	/	/	5	/	显著目标掩膜
PVS-HM(Xu等,2019c)	视频	76	3K~8K	10~80	58	HTC VIVE	头动/眼动数据
VR Eye-Tracking(Xu等,2018b)	视频	208	4K	20~60	45	HTC VIVE	头动/眼动数据
Destop360(Duanmu等,2018)	视频	12	3840×1024, 3840×2048	60~102	50	/	观看轨迹
AVTrack360(Fremerey等,2018)	视频	20	4K	30	48	HTC VIVE	头动数据
UN Salient360(David,2018)	视频	19	3840×1920	20	57	HTC VIVE	头动/眼动数据
PanoSaliency(Nguyen和Yan,2019)	视频	24	/	60~655	48	/	显著图
360 EM Dataset(Agtzidis等,2019)	视频	15	/	38~85	13	FOVE	注视数据

注: ★表示该数据库同时包含 MOS 和头部运动数据。相关公开数据库信息请见 <https://github.com/LXLHXL123/OpenVRQoE>。

/1)表8展示了从普通压缩视频质量评价、自适应流媒体传输体验到音视频联合感知三个递进层次,系统呈现了该领域从初步探索走向规范化与多元化的演进历程。相较于VR图像主观质量数据库,视频数据库在实验设计维度上呈现出更为复杂的参数空间,这主要源于视频内容所固有的时间维度特征。早期研究如Tran等人仅使用少量样本,侧重于验证编码压缩与分辨率变化对全景视频质量的基本影响,而最新的数据库则包含了502个源自实际应用场景的带有真实失真的测试样本,这种从实验室受控失真向真实世界失真迁移的趋势,显著增强了数据库的生态效度与应用价值。际应用场景的带有真实失真的测试样本,这种从实验室受控失真向真实世界失真迁移的趋势,显著增强了数据库的生态效度与应用价值。

表7 VR图像主观体验质量评价数据库
Table 7 Summary of existing subjective QoE evaluation databases for VR image

数据库	原始/ 测试 数量	分辨率	失真	受试 人数	设备	协议	数据
Upenik 等人(2016)	6/60	3000× 1500, 2250× 1500	JPEG(4), 投影(2)	48	MergeVR, iPhone 6s	ACR- HR	MOS [0, 5], HM

表 8 续表

表 7 VR 图像主观体验质量评价数据库 Table 7 Summary of existing subjective QoE evaluation databases for VR image							
数据库	原始/测试数量	分辨率	失真	受试人数	设备	协议	数据
Upenik 等人 (2017)	4/104	3000×1500	JPEG, JPEG2000, 投影(2)	45	MergeVR, iPhone 6s	ACR-HR	MOS [0, 5], HM
CVIQD(Sun 等, 2017a)	5/170	4096×2048	JPEG(11), H.264(11), H.265(11)	20	HTC VIVE	ACR	MOS [0, 10]
CVIQD2018(Sun 等, 2018)	16/544	4096×2048	JPEG(11), H.264(11), H.265(11)	20	HTC VIVE	ACR	MOS [0, 10]
Huang(Huang 等, 2018)	12/156	9104×4552	JPEG(3), 分辨率(4)	98	HTC VIVE	ACR	MOS [0, 100]
OIQA★(Duan 等, 2018)	16/336	11332×5666, 13320×6660	JPEG(5), JPEG2000(5), 高斯模糊(5), 高斯噪声(5)	20	HTC VIVE	ACR	MOS [0, 10], HM, EM
Upenik 和 Ebrahimi (2019)	4/52	5760×2880	JPEG(4), JPEG2000(4), HEVC(4)	19	MergeVR, Galaxy S7	ACR-HR	MOS [1, 5]
Kim 等人 (2020)	60/780	2048×1024	JPEG(4), JPEG2000(4), HEVC(4)	20	Oculus Rift CV1	ACR	MOS [0, 100]
Jabar 等人 (2020b)	8/200	856×856	观看方向(3), 距离(4), 视口(2)	20	-	SCACJ	CMOS [0,10]
MVAQD(Jiang 等, 2021)	15/315	12K, 10K, 8K, 7K, 6K, 5K, 4K	JPEG(4), JPEG2000(4), HEVC(4), 高斯模糊(4), 高斯噪声(4)	26	HTC VIVE Pro	ACR	MOS [1, 5], HM
IQA-ODI(Yang 等, 2021b)	120/1, 080	7680×3840	JPEG(4), 投影(4)	200	HTC VIVE	ACR	DMOS [0,100]
JUFE-VRIQA★ (Fang 等, 2022)	258/1, 290	8192×4096	高斯模糊(3), 高斯噪声(3), 亮度不连续(3), 拼接(3)	120	HTC VIVE Pro Eye	ACR	MOS [1, 5], HM, EM
ISIQA◆(Madhusudana 和 Soundararajan, 2019)	26/290	8K, 10K	拼接失真	35	Samsung Gear VR	SSCQE	MOS [0, 100]
OSIQA◆(Duan 等, 2023)	12/312	6080×3040	拼接失真	20	HTC VIVE Pro	SSCQE	MOS [0, 100]
Simka 等人 (2024)	5/100	8K, 4K	JPEG(4), JPEG XL(5), HEIC(5), AVIF(5)	66	Oculus Rift S, HTC VIVE Pro	ACR-HR	MOS [1, 5], EM

表 8 续表

表 7 VR 图像主观体验质量评价数据库							
Table 7 Summary of existing subjective QoE evaluation databases for VR image							
数据库	原始/测试数量	分辨率	失真	受试人数	设备	协议	数据
JUFE-10K★(Yan 等, 2025b)	430/10, 750	8192×4096	高斯模糊(3), 高斯噪声(3), 亮度不连续(3), 拼接(3)	172	HTC VIVE Pro Eye	ACR	MOS [1, 5], HM、EM
OIQ-10K★(Yan 等, 2025c)	2, 500/10, 000	7680×4320	均匀, 非均匀	160	HTC VIVE	ACR	MOS [1, 3], HM、EM
SOLID♣(Xu 等, 2018a)	6/162	8192×8192	BPG 压缩(6), 深度(3)	18	Samsung Gear VR	ACR-HR	MOS [1, 5]
LIVE 3D VR IQA♣(Chen 等, 2020a)	15/465	4096×2048	高斯模糊, 高斯噪声, 下采样, VP9 压缩, H.265	42	HTC VIVE	SSCQE	DMOS [0, 100], EM
NBU-SOID♣(Yang 等, 2020)	12/408	2560×2560, 4096×4096	JPEG(5), JPEG2000(5), HEVC(5)	30	HTC VIVE Pro	ACR-HR	MOS [1, 5]
AIGCOIQA2024(Yang 等, 2024)	-/300	/	AI 生成失真, 感知质量, 舒适度, 图文一致性	20	/	ACR	MOS [0, 100]
OHF2024(Yang 等, 2026)	-/600	/	AI 生成失真, 感知质量, 舒适度, 图文一致性	20	HTC VIVE Pro Eye	ACR	MOS [0, 100], 显著图

注: ◆表示拼接全景图像数据库。♣表示立体全景图像数据库。表示 AI 生成数据库。失真列中括号中内容表示失真等级设置。

表 8 VR 视频主观体验质量评价数据库

Table 8 Summary of existing subjective QoE evaluation databases for VR video									
数据库	原始/测试数量	分辨率	时长(s)	帧率	失真	受试人数	设备	协议	数据
Tran 等人(2017a)	3/60	/	30	30	H.264(5), 分辨率(4)	18	Samsung Gear VR	ACR	MOS [0,5]
Tran 等人(2017b, 2021)	3/60	/	30	30	H.264(5), 分辨率(4), 内容(3), 观看模式(2), 设备(2)	36	Samsung Gear VR, Google Cardboard	ACR	MOS [1,5]
Singla 等人(2017a)★	6/60	7680×3840	10	30	H.265(5), 分辨率(2)	30	Oculus Rift	ACR	MOS [0,5]
Singla 等人(2017b)★	6/24	3840×2048, 1920×1080	60~65	25~30	设备(2), 分辨率(2)	28	HTC VIVE, Oculus Rift	ACR	MOS [0,5]

表 8 续表

数据库	原始/ 测试 数量	分辨 率	时长 (s)	帧率	失真	受试 人数	设备	协议	数据
Zhang 等人(2017)	16/ 384	4096 × 2048	10	30	编码(3), 码率(6), 丢包, 打分(3)	/	HTC VIVE	ACR	MOS [0,5]
IVQAD(Duan 等,2017b)	10/ 150	4096 × 2048	15	/	分辨率(3), 帧率(3), 码率(3)	/	HTC VIVE	ACR	MOS [0,5]
VR-VQA48★(Xu 等,2017, 2019b)	12/ 36	4096 × 2048	12	25	H.265(3)	48	HTC VIVE	ACR	DMOS [0,100]
VQA-ODV★(Li 等,2018a)	60/ 540	3840 × 1920 ~ 7680 × 3840	10~ 23	24~30	QP(3), 投影(3)	221	HTC VIVE	ACR	DMOS [0,100]
VRQ-TJU(Yang 等,2018a)	13/ 364	2560 × 2560	16.9	30	H.264(5), JPEG2000(5)	30	/	ACR	MOS [0,5]
Lopes 等人(2018)	6/-	8192 × 4096, 3840 × 1920	10	30,60	分辨率(2), H.265(5)	37	Oculus Rift	ACR	MOS [0,5]
Zhang 等人(2018b)	10/ 50	8192 × 4096, 3840 × 1920	10	30,60	H.265(5), QP(5)	30	HTC VIVE	ACR	DMOS [1,5]
Zou 等人(2018)	8/-	3840 × 1920	20	30	H.264(5)	25	/	ACR	MOS [1,5]
ODV(Croci 等,2019)	6/75	8128 × 4064	10	/	分辨率(3), QP(5)	24	HTC VIVE	M- ACR	DMOS [0,100]
Yao 等人(2019)	4/36	3840 × 1920	30	30	内容(2), QP(3), 投影(3)	60	Oculus Rift	ACR	MOS [0,9]
Zepernick 等人(2019)	4/ 120	7680 × 3840	10	29.97	QP(5), 分辨率(5), 受试人数	30	HTC VIVE	ACR	MOS [0,5]

表 8 续表

数据库	原始/ 测试 数量	分辨率	时长 (s)	帧率	失真	受试 人数	设备	协议	数据
Zhang 等人(2019)	12/ 120	3840 × 1920	10~ 30	/	H.265(5), 丢包(5), 观看 模式(3)	30	HTC VIVE	ACR	/
Orduna 等人(2020b)	6/30	3840 × 2160 ~ 3840 × 1920	30	25~50	H.265(5)	48	Samsung Gear VR, Lenovo Mirage Solo	ACR	MOS [0,5]
Jabar 等人(2020a)	16/ 96	1816 × 1020	10	/	视场角(6)	44	/	PC	/
Anwar 等人(2020a)	3/10	/	60	30	QP(4), 分辨率(3), 5 秒卡顿(3)	34	HTC VIVE	ACR	MOS [0,3]
Anwar 等人(2020b)	6/96	/	60	24~30	H.264(4), 分辨率(4)	29	HTC VIVE, Google Card- board	ACR	MOS [0,5]
Zheleva 等人(2020)	1/4	/	300	/	质量等级(4)	32	HTC VIVE Pro	ACR	MOS [0,5], EEG
Chu 等人(2020)	4/ 120	7680 × 3840	10	29.97	H.265(5), 分辨率(5), 打分(2)	30	HTC VIVE Pro	M- ACR	MOS [0,5]
Zou 等人(2020)	8/-	8192 × 4096	10	30	屏幕分辨率(4)	30	HTC VIVE (Pro), Pico G 2 4K	ACR	MOS [0,5]
Mahmoudpour 和 Schelkens (2021)	15/ 60	3840 × 1920	6~8	30	H.265(3), 速度(3)	20	Oculus Rift	ACR	C/O- MOS [0,100]
SSV360 ★ (Elwardy 等 , 2022)	4/ 120	7680 × 3840	10	29.97	H.265(5), 分辨率(5),观看 方式(2)	5	HTC VIVE Pro	M- ACR	MOS [0,5]
ACR360 ★ (Elwardy 等 , 2023)	4/ 120	7680 × 3840	10	29.97	H.265(5), 分辨率(5), 打分(2)	65	HTC VIVE Pro	M- ACR	MOS [0,5]
Fang 等人(2023)★	0/ 502	1280 × 720~ 5120 × 2560	15	20~60	真实失真	139	HTC VIVE Pro Eye	ACR	MOS [1,5]
Schatz 等人(2017)	3/15	/	60	/	卡顿(5), 观看设备(2)	27	Oculus Rift	ACR	MOS [0,5]

表 8 续表

数据库	原始/测试数量	分辨率	时长 (s)	帧率	失真	受试人数	设备	协议	数据
Curcio 等人(2017)	3/24	3840 × 1920	21	30	H.265(4), 分辨率(2)	12	Samsung Gear VR	ACR	DMOS [0,5]
Li 等人(2018b)	10/30	3840 × 2160	30~40	25.30	分块方案(3), 卡顿, 分辨率切换	40	HTC VIVE	ACR	MOS [0,100]
Singla 等人(2019a)	3/72	7680 × 3840 6144 × 3072 8192 × 4096	30	25,60	分辨率(2), 带宽(4), 延迟(4)	53	Samsung Gear VR	ACR	MOS [0,11]
Schatz 等人(2019)	2/20	8192 × 4096 7680 × 3840	30	/	H.265(3), 头动速度(2), 传输策略(2)	35	/	ACR	MOS [0,7]
Cortes 等人(2020)	3/180	4096 × 1980	20	/	延迟(5), 自适应方案(4)	20	Lenovo Mirage Solo	ACR	DMOS [0,5]
Anwar 等人(2020c)	16/208	3840 × 1920	30~40	24,25,30	H.264(3), 卡顿(3)	31	HTC VIVE	ACR	MOS [0,10]
Fei 等人(2020)	9/468	/	/	/	带宽(5), 丢包(5), 延迟(5)	234	/	ACR	MOS [0,5]
Xie 等人(2020)	30/1, 608	3840 × 1920, 3840 × 2048, 3840 × 2160	10	30	H.264(5), 分辨率(3), 切换时间(10)	340	HTC VIVE	ACR	MOS [0,5]
Zou 等人(2021)	10/354	4K	20	30	H.265(6), 低质量区域比例(5)	30	HTC VIVE Pro	ACR	MOS [0,5]
VOD-VQA (Meng 和 Ma, 2022)	18/774	3840 × 1920	10	30	H.264(4), 分辨率(4), 帧率(3)	160	HTC VIVE	ACR	MOS [0,10]
VRQoE-JUFE★ (Liu 等, 2023a)	30/1, 440	7680 × 3840	15	24,25, 29,97,30, 45,59,94	初始缓冲, 卡顿, 质量切换	180	HTC VIVE Pro Eye	ACR	MOS [1,5]

表8续表

数据库	原始/ 测试 数量	分辨 率	时长 (s)	帧率	失真	受试 人数	设备	协议	数据
JUFE-OVQoE (Liu 等, 2024c)	12/ 378	7680 × 3840	10	29.97, 30	质量自适应	20	HTC VIVE Pro Eye	ACR	MOS [1.5]
Zhang 等人(2018a)†	10/ 810	4096 × 2048	10	30	视频失真, 音频失真, 音画不同步	20	HTC VIVE	ACR	DMOS [0,5]
Fela 等人(2022)†	12/ 192	7680 × 3840	20	30	H.264(4), 分辨率(3), 音频码率(4)	21	Samsung Odyss- sey+	ACR	MOS [0,100]
OAVQAD† (Zhu 等, 2024)	15/ 375	7680 × 3840	6	29.97	H.265(3), 分辨率(3), 噪声, 模 糊, 卡顿, 音频码率(3)	22	HTC VIVE Pro Eye	ACR	MOS [1,10]
OAVQAD+† (Zhu 等, 2025a)	25/ 625	7680 × 3840	6	29.97	H.265(3), 分辨率(3), 噪声, 模 糊, 卡顿, 音频码率(3)	22	HTC VIVE Pro Eye	ACR	MOS [0,100]

注:†表示音视频数据库。

2)在参数规模方面,随着研究问题复杂度的提升,研究者对于覆盖更广泛参数空间和获取更高统计效力数据的需求同步增长。当前VR视频质量数据库的测试样本数量普遍集中在30至400个之间,Xie等人(2020)构建的数据库规模最大,共包括1,608个测试样本。视频时长设计呈现出高度分化,最短的片段仅6至8秒,而Zheleva等人于2020年采用的长达300秒的视频片段则首次将研究视野延伸至长时间观看情境下的疲劳效应与质量感知漂移,这种时间跨度的极端分化反映了研究者对于时域累积效应关注度的日益提升。帧率参数在多数数据库中被设定为30帧每秒以匹配常规视频格式,但Lopes等人(2018)和Zhang等人(2018b)系统性地引入了60帧每秒的高帧率条件,为探索帧率对沉浸式体验的影响提供了数据支撑。

3)在失真类型方面,普通VR视频数据库的失真设计相对集中于编码压缩领域,分辨率缩放和编码量化参数变化构成了主要的参数调节手段。除H.264、H.265及分辨率缩放等传统压缩失真外,VR视频数据库还引入了网络丢包损伤、缓冲延时、播放卡顿、码率自适应、分块方案以及投影格式比较等复杂动态流媒体传输场景中的典型劣化因素。这些设计更贴近实际网络传输下的用户体验。值得注意的是,Jabar等人(2020)单独对视场角变化进行了系统

研究,在六种不同视场角设置下评价其对感知质量的影响。Zou等人(2021)研究基于图块的流媒体传输方案对用户体验的影响,为优化分块传输中的码率分配策略提供了实证依据。这种针对VR特有参数的专项研究在图像数据库中较为少见。

4)在设备使用方面,HTC VIVE系列凭借其稳定的追踪性能和逐渐普及的眼动追踪模块,在所有数据库中占据了超过半数的份额,而Samsung Gear VR和Oculus Rift作为移动端与电脑端的代表性设备亦被广泛采用。其他设备的使用则反映了VR硬件生态的多元化趋势。部分研究(Singla等,2017;Tran等,2017b;Zou等,2020)通过在同一实验框架内并行使用多种头显设备,揭示了硬件差异对主观评分的显著影响,为后续研究建立跨设备质量映射模型提供了宝贵数据。

5)在主观评价协议方面,绝对类别评分占据绝对主导地位,体现出其操作简便性和高效性。其变体形式如M-ACR则通过引入隐藏参考序列和多次重复评分,显著提升了评分结果的可靠性与内部一致性。此外,不同数据库之间的主观评分尺度差异较大,多种评分尺度并存。

6)在标注数据类型方面,尽管多数早期数据库仅输出MOS或DMOS,但部分研究已开始同步采集头部运动轨迹。在此基础上,Elwardy等人(2022,

2023)率先将瞳孔测量、皮电反应、晕动症等心理生理指标与传统质量评分及姿势相关的头部运动学数据进行整合,而Zheleva等人(2020)更是开创性地将脑电图信号与主观评分联合记录,为探索神经生理反应与质量感知之间的内在关联开启了全新范式。这种将行为数据与生理信号融入主观质量评价框架的趋势,预示着VR体验质量研究正从纯粹的心理物理判断向包含认知负荷、沉浸感和疲劳度在内的多维感知体验评价转变。与此同时,音视频联合数据库(Zhang等,2018a; Fela等,2022; Zhu等,2024, 2025a)研究进一步揭示了音频质量在沉浸式体验评价中不可或缺的作用。这些进展表面该领域已从基础质量测量迈向沉浸式环境中复杂感官整合建模的

成熟阶段。

小结:上述数据库为学术界和工业界深入理解VR中技术参数、观看行为与感知体验之间的复杂关系提供了重要支撑。然而,这一领域仍然面临若干挑战,如数据异构阻碍跨数据库模型训练及定量比较、时域维度覆盖不足难以支撑对完整自适应会话中用户质量动态变化的研究、真实失真与合成失真之间存在生态效度差距、多模态标注数据缺乏统一评价范式与跨设备校准方案等。未来研究应聚焦于建立统一的跨数据库评分框架,设计面向长时自适应会话的连续质量评价协议,并发展多模态标注体系,推动从静态片段评价向动态沉浸体验评价的实质性跨越。

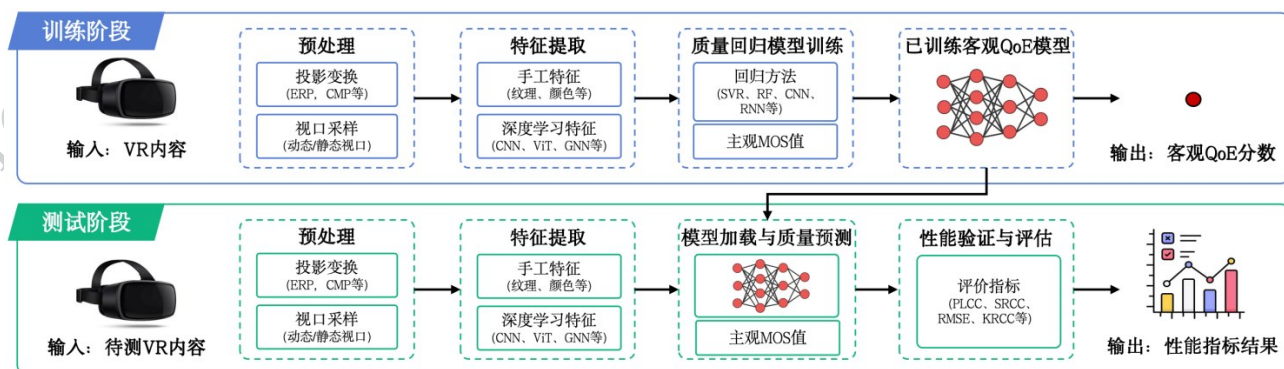


图5 面向VR内容的客观体验质量评价模型训练与测试一般流程

Fig. 5 General workflow for training and testing objective QoE models for VR content

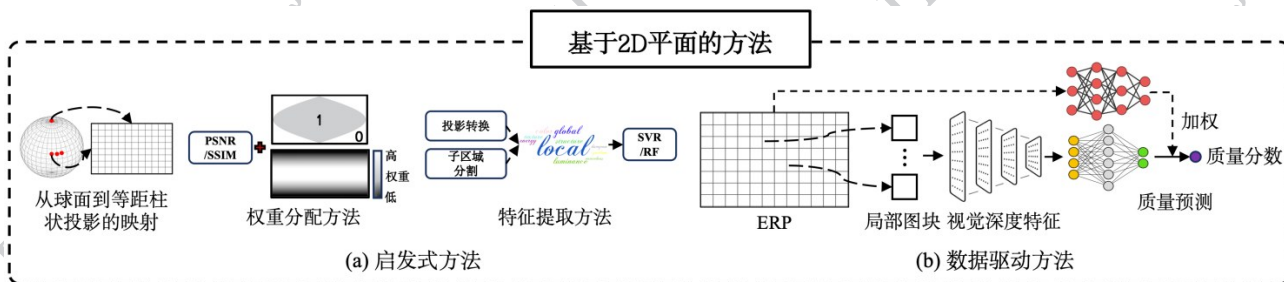


图6 基于2D平面的OIQA/OVQA方法框架图

Fig. 6 Illustrations of 2D-plane-based OIQA/OVQA methods

2 客观体验质量评价

VR图像和视频通常以二维投影平面、球面或视口三种形式进行表示。根据这些表示形式,当前VR图像与视频的客观用户体验质量评价方法可分为基于二维平面、基于球面以及基于视口三类(Fang等,2021;Sui等,2024;Yan等,2025e)。表9和

10分别汇总了现有VR图像与视频的客观质量评价方法。本节将详细介绍这三类方法。图5展示了面向VR内容的客观体验质量评价模型训练与测试的一般流程。

2.1 基于2D平面的方法

VR视频图像通常以等距柱状投影格式存储于二维平面,观看时再重新投影至球面进行渲染显示。这一表示形式使得PSNR和SSIM(Wang等,2004)

等传统图像质量指标可简单应用于VR内容的质量评价。然而,由于360°内容的球面特性,其在平面投影过程中引入了几何畸变等特殊失真。因此,基于二维平面的质量评价方法主要致力于补偿此类由投影引起的失真。可进一步划分为启发式方法与数据驱动方法两类。

2.1.1 启发式方法

启发式方法可大体分为两类。第一类方法通过加权矩阵来建模球面上采样点的密度分布,如图6所示。具体而言,对于给定的VR图像 $I \in \mathbf{R}^{h \times w}$ 或视频 $V \in \mathbf{R}^{h \times w \times t}$,在质量预测过程中使用二维加权矩阵以补偿非均匀采样带来的影响:

$$\hat{Q} = \sum_{(x,y) \in I} \hat{q}(x,y) \cdot W(x,y) \quad (1)$$

式中, $\hat{q}(x,y)$ 和 $W(x,y)$ 分别表示位于坐标 (x,y) 处图像像素的预测局部质量和加权权重, $\hat{Q}$ 为VR图像或视频的整体预测评分。为简化表述,此处省略VR视频中的时间维度 t 。基于此原理,多种加权质量指标被提出。CPP-PSNR(Zakharchenko等,2016)通过将球面坐标 (ϕ, ψ) 映射至平面坐标 (x,y) 以确保均匀采样;WS-PSNR(Sun等,2017b)和WS-SSIM(Zhou等,2018)则球面拉伸比进行加权。此外,视觉注意机制也被引入到加权过程中,例如SAL-PSNR(Luz等,2017)依据区域显著度进行加权,AW-SPSNR(Xiu等,2017)通过面积加权球面PSNR来修正投影格式偏差。其他方法如BP-QAVR(Yang等,2017)采用融合模型结合多层次质量因子与感兴趣区域映射。另外,Lopes等人(2018)提出对SSIM(Wang等,2004)和MS-SSIM(Wang等,2003)的改进方案,基于像素区域覆盖率引入球面校正因子,进而预测VR内容的质量。

另一类启发式方法则考虑人类视觉感知特性,通过提取感知特征以及质量回归映射来评价VR内容质量,如图6所示。该类方法大致分为基于自然场景统计与结构特征的模型(Liu等,2021;Poreddy等,2023)、多感知特征融合方法以及专用失真补偿方法。在多感知特征融合方面,Liu等人(2023b,2023c)分别采样直方图、统计、显著性特征及结构、颜色、熵等指标构建质量评价方法,但未考虑二维投影平面上的非均匀采样问题。针对特定失真,OS-IQA(Li等,2019b)通过结合局部拼接区域与全局沉浸感因子评价含有拼接失真的VR图像质量。为应

对ERP格式中的极区畸变,SSP-BOIQA(Zheng等,2020)通过分段球面投影保留赤道区域,并提取局部与全局特征以预测质量。此外,Orduna等人(2020a)验证了经典视频评价方法VMAF(Netflix,2016)可一定程度用于VR视频质量评价。

2.1.2 数据驱动方法

数据驱动方法采用深度神经网络,通过基于图像块的分析框架来评价VR内容质量。其标准流程通常包含三个核心步骤:1)将等距柱状投影内容划分为 N 个图像/视频块;2)利用预训练网络提取视觉特征 $f_{1:N}$;3)通过特定的融合策略 W 聚合局部质量得到预测结果。该过程可表述为如下优化问题:

$$\arg\max_{\theta} P(Q|W(Q(f_1), Q(f_2), \dots, Q(f_N)); \theta) \quad (2)$$

式中, Q 表示真实质量标签, Q 为质量预测模块, θ 代表可学习的网络参数。该类方法中对于图像块的选择与处理至关重要。基于VR IQA NET(Lim等,2018),Kim等人(2020)提出基于双分支架构的DeepVR-IQA,利用ResNet-50(He等,2016)进行视觉特征提取,并通过对抗学习结合质量预测与人类感知引导模块,分别预测图像块质量与权重,最终聚合得到整体质量。在此基础上,后续涌现出多种创新方法。Truong等人(2019)强调赤道区域的感知重要性,提出基于选择性图像块采样的质量评价方法。Ding等人(2020)采用级联网络实现显著性引导的图像块选择。Yang等人(2021b)结合离散小波变换与空间注意力进行自监督图像增强,并将增强后的图像块作为质量预测的伪参考。Yang等人(2023)提出了一个面向VR视频质量评价的渐进学习框架,利用三个网络实现像素、帧和视频三级联合优化。针对纬度变化带来的混淆效应,Hu等人(2024c)提出了基于因果干预的CIQNet,通过区域特定编码实现纬度不变特征表示。此外,Yan等人(2025f,2026a)创新性地提出了系列视口无关的全景图像质量评价模型,通过自适应采样,在不依赖视口信息的情况下实现有效的质量预测。另外,Yan等人(2026b)还将大模型创新地引入到该任务中。

2.1.3 小结

当前基于2D平面的方法分别从手工设计特征与数据自适应学习的角度出发,共同推动了VR质量评价从平面指标向沉浸式感知模型的演进。其

中,启发式方法一方面通过加权矩阵建模球面采样密度分布,另一方面,该类方法也通过提取自然场景统计特征、多感知特征或专用失真补偿策略来预测质量。数据驱动方法则采用深度神经网络,遵循图像块划分、特征提取、局部质量聚合的通用框架,后

续研究进一步拓展了选择性采样、显著性引导、自监督增强、渐进学习、因果干预及视口无关建模等策略,并初步探索了大模型的应用,逐步提升了对VR内容质量感知的表征能力。

表9 基于2D平面和基于球面的VR图像/视频客观体验质量评价方法

Table 9 Summary of existing 2D plane-based and sphere-based objective QoE evaluation methods for VR content

类型	方法	启发式	数据驱动	扫视路径	观看条件	权重分配	重采样
基于2D平面的方法	CPP-PSNR♠(Zakharchenko等,2016)	✓	-	-			✓
	SAL-PSNR(Luz等,2017)	✓				✓	
	AW-PSNR♠(Xiu等,2017)	✓				✓	
	BP-QAVR♠(Yang等,2017)		✓				
	WS-PSNR♠(Sun等,2017b)	✓				✓	
	WS-SSIM♠(Zhou等,2018)	✓				✓	
	W/WMS-SSIM♠(Lopes等,2018)	✓				✓	
	Huang等人(2018)	✓					
	NR-CSC-SIQA(Ling等,2018)	✓					
	VR IQA NET(Lim等,2018)		✓			✓	
	Omni-IQA(Truong等,2019)		✓			✓	
	ASY-PIQA(Xia等,2019)	✓					
	OS-IQA(Li等,2019b)	✓					
	Ding等人(2020)♠		✓				
	AQA(Li等,2020)	✓				✓	
	Orduna等人(2020a)	✓					
	SSP-BOIQA(Zheng等,2020)	✓					✓
	DeepVR-IQA(Kim等,2020)		✓			✓	
	APC(Ding等,2021)	✓					
	SAP-Net(Yang等,2021b)		✓				
	Liu等人(2021)	✓					
	ProVQA♠(Yang等,2023)		✓			✓	
	SSS(Liu等,2023c)	✓					
	Proeddy等人(2023)	✓					✓
	Liu等人(2023b)	✓					✓
	CIQNet♠(Hu等,2024c)		✓				
	VU-BOIQA(Yan等,2025f)		✓				✓
	IPSS(Yan等,2026a)		✓				✓
	MMBO(Yan等,2026b)		✓				✓
	VI-AGOIQA(Liu等,2025a,2025b)		✓				✓

表9续表

类型	方法	启发式	数据驱动	扫视路径	观看条件	权重分配	重采样
基于球面的方法	BLIP2OIQa (Yang 等, 2026)		✓				✓
	S-PSNR ♠ (Yu 等, 2015)	✓					✓
	S-SSIM ♠ (Chen 等, 2018a)	✓				✓	✓
	OV-PSNR ♠ (Gao 等, 2022)	✓				✓	✓
	IS-PSNR ♠ (Hristova 等, 2021)	✓					✓
	Zhang 等人 (2021)	✓					
	Yang 等人 (2021a) ♠		✓				
	Equi-StairNet (Chen 和 Niu, 2022)		✓				
	Sendjasni 和 Larabi (2023a)		✓			✓	✓

注: ♠表示 VR 视频用户体验质量评价方法。相关开源方法请见 <https://github.com/LXLHXL123/OpenVRQoE>。✓代表该方法具备对应模块与特性。

2.2 基于球面的方法

基于球面的 VR 内容质量评价方法直接在球面域进行操作,主要采用全局均匀采样或局部图块采样两种策略,如图 7 所示。

2.2.1 全局均匀采样方法

全局均匀采样方法以全参考质量评价方法 S-PSNR (Yu 等, 2015) 为代表。该方法通过在球面上设置密集均匀分布的采样点 (通常为 655,362 个点) 提取对应二维像素值进行质量计算。相关扩展方法包括 OV-PSNR (Gao 等, 2022) 和 IS-PSNR (Hristova 等, 2021) 等。

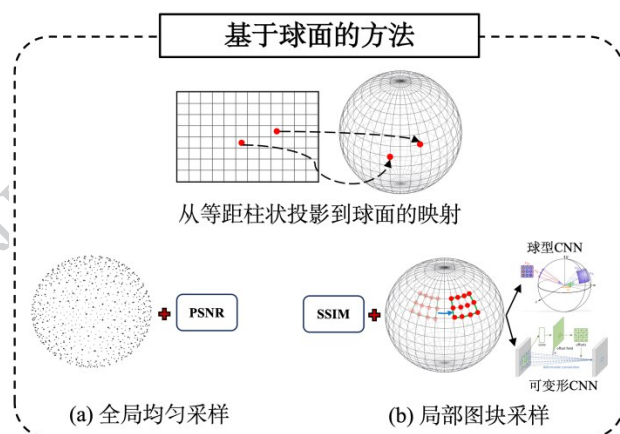


图7 基于球面的 OIQA/OVQA 方法框架图

Fig. 7 Illustrations of sphere-based OIQA/OVQA methods

2.2.2 局部图块采样方法

局部图块采样方法通过分析球面图像块的结构信息来推断整体质量。如 S-SSIM (Chen 等, 2018a)

将循环对称高斯加权应用于基于图像块的 SSIM 计算中,并按坐标拉伸比进行加权。Sendjasni 和 Larabi (2023a) 提出了一种基于注意力的球面图像块加权方法,在质量融合过程中优先处理显著区域。鉴于球形卷积神经网络 (Cohen 等, 2018) 在处理全景内容时的几何不变性, Yang 等人 (2021a) 将其与非局部网络 (Wang 等, 2018) 结合构建时空质量评价模型。Chen 和 Niu (2022) 提出采用等距柱状卷积 (Labrador 等, 2020) 处理 ERP 投影引起的形变。

2.2.3 小结

以上基于球面的方法在处理 VR 内容的几何特性方面展现出显著优势。其中,基于全局均匀采样的方法通过在球面设置大量均匀分布采样点,能够实现全面而统一的质量评价,而基于局部图块采样的方法则通过分析球面图像块的结构信息推断整体质量。两类采样策略分别从全局统计均匀性和局部结构感知的角度出发,为球面域质量评价提供了互补的技术路径。此外,球形卷积技术与注意力机制的持续发展助力该领域的研究。

2.3 基于视口的方法

由于人类视觉系统的视场角有限,对 VR 内容进行整体评价往往无法准确反映用户的真实感知。因此,基于视口的方法通过聚合多个视角下的质量评分来弥补这一局限。根据采样策略的不同,现有方法主要可分为三类:基于规则的方法、基于关键点的方法和基于扫视路径的方法,如图 8 所示。前两类方法主要关注视口图像的空间质量特征建模;而

基于扫视路径的方法则通过模拟用户的动态观看行为,实现对VR内容的时空感知质量的综合评价。

2.3.1 基于规则的方法

基于规则的质量评价方法通过系统化视口采样策略分析选定视口,以评价VR内容整体质量。早期研究致力于跨越人类感知行为与VR质量评价之间的鸿沟,Li等人(2018a)提出了融合头部运动与眼动数据的深度学习模型。Xu等人(2019b)分别提出基于像素与前景区域相似度的失真加权方法和基于预测观看方向的失真权重分配方法。Li等人(2019a)提出了视口卷积神经网络框架(viewport-based convolutional neural network, VCNN),涵盖视口生成与质量评价两阶段,后续提出了相应扩展版本(Xu等,2022)。类似地,Guo等人(2021)提出了基于生成对抗网络的质量评价模型。Sendjasni等人(2021a, 2022a, 2022b)提出了一种显著性加权的CNN框架,采用基于纬度的非重叠球面采样优先处理赤道区域。

另一类基于规则的方法侧重通过优化采样策略应对计算挑战。MC360IQA(Sun等,2020)与Fang等人(2022)采用纬度感知视口采样策略,采样密度从

赤道向两极递减,并利用多分支网络模型共同预测内容质量。Fang等人(2022)还提出了利用可学习的视口坐标辅助质量预测。Chen等人(2020b)和Xu等人(2019a)采用非均匀视口采样策略,在保持质量评价精度的同时优化特征提取效率。此外,Yan等人(2025a, 2025b, 2025c)采用赤道采样法(JVET, 2023)提取覆盖低纬度区域的视口,并融合失真自适应感知模块预测含有非均匀失真的VR图像质量。在其他投影表达空间方面,将VR内容投影映射到其他投影表达空间以克服ERP映射在两极区域引入拉伸形变的问题。多项研究(Chai和Shao, 2021; Fan和Chen, 2024b; Jiang等, 2021; Jiang等, 2022; Zhou等, 2023a; Zhou等, 2022a; Zhou等, 2022b; Hu等, 2025)利用立方体映射投影获取VR视频图像的六个最小重叠视口以预测质量。并且跨投影格式的质量评价研究(Hu等, 2025)也逐步出现。

考虑到人类视觉感知特性,基于手工特征的VR内容质量评价研究也取得了相应进展。Azevedo等人(2020a, 2021)提出了基于视口的多指标融合方法,通过结合多种时空特征评价VR内容质量。

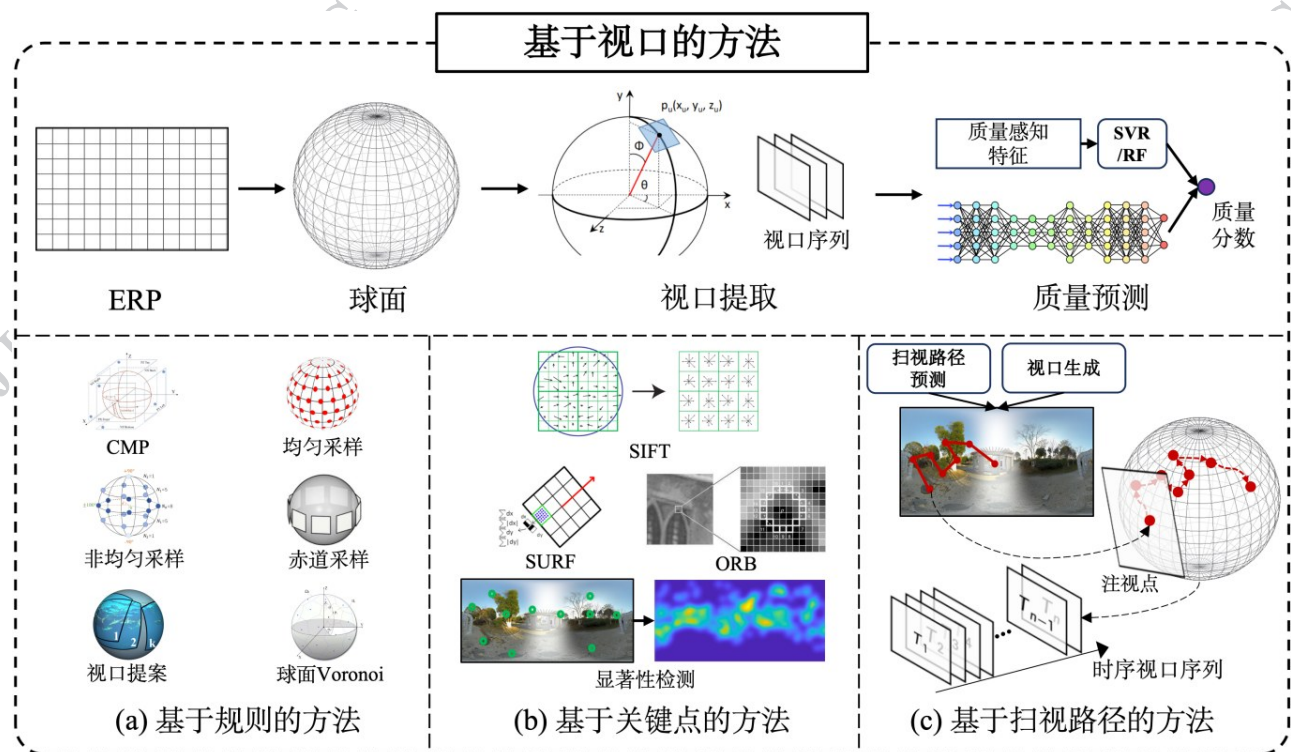


图8 基于视口的OIQA/OVQA方法框架图

Fig. 8 Illustrations of viewport-based OIQA/OVQA methods

MFILGN (Zhou 等, 2022a) 与 MP-BOIQA (Jiang 等, 2022a) 提取视口统计特征, 通过支持向量回归或随机森林映射为质量分数。Tian 等人 (2022) 通过提取三类特征以量化拼接导致的几何畸变与色差, 进而实现质量预测。Liu 等人 (2023d) 提出了一种结合结构、色度及视口统计特征的无参考评价模型。Zhou 和 Wang (2023b) 提出将图像底层统计量与高层语义关联, 同时处理局部视口与全局表征以预测图像质量。

后续研究 (Fan 和 Chen, 2024b; Zhou 等, 2023a) 进一步引入了图卷积网络 (Fu 等, 2022)、视觉 Transformer 网络 (Dosovitskiy 等, 2021; Tofighi 等, 2023)、可变形卷积网络 (Dai 等, 2017; Tian 等, 2023)、VMamba 网络 (Liu 等, 2025)、掩码自编码器 (Hu 等, 2026)、多任务学习 (Zhou 等, 2022b)、知识蒸馏 (Hu 等, 2025) 以及多尺度时空特征聚合 (Hu 等, 2024b; Yan 等, 2025g) 等网络架构, 实现对人类真实观看行为的精准建模和综合质量评价。

此外, 针对 VR 内容流传输中的特定失真, 如投影渲染失真、运动失真与视口质量自适应等, 相关体验质量评价方法也取得了重要的进展。例如, Jabar 等人 (2022b) 提出了一种内容感知质量指标, 通过几何与空间失真特征评价 VR 内容投影渲染的质量。Mahmoudpour 和 Schelkens (2021) 提出了抖动质量指数 (judder quality index, JQI), 通过结合运动抖动检测阈值与视网膜滑动率模型, 实现对运动抖动、视觉掩蔽与 VR 图像质量评价的集成。针对基于分块的全景视频流传输优化, Meng 和 Ma (2022) 建立了连接视口压缩与空间、时间、幅度及分辨率特征的多维质量模型, 通过显著性加权视口聚合实现质量评价。针对人工智能生成内容和特定失真, Liu 等人 (2025a, 2025b) 和 Yang 等人 (2026) 相继提出了面向 AI 生成全景图像的多维质量评价模型。

2.3.2 基于关键点的方法

基于关键点的视口质量评价方法通常在二维投影平面上提取具有区分性的关键点, 并以其为中心提取视口进行质量预测。VGCN (Xu 等, 2021) 利用加速稳健特征 (speeded up robust features, SURF) 算法 (Bay 等, 2006) 识别 VR 图像的 20 个关键点, 进而构建图卷积网络以建模视口间的空间关系, 并自适应融合各视口的质量分数。Wang 等人 (2025) 进一步将其扩展至立体全景图像质量评价, 采用 SURF

检测关键点并基于双目视口超图卷积网络进行预测。Zhang 和 Liu (2022) 则使用定向 FAST 和旋转 BRIEF 算法 (Rublee 等, 2011) 提取关键点, 进而提取视口以预测质量。Zhou 等人 (2023c) 用尺度不变特征变换 (scale invariant feature transform, SIFT) 算法 (Lowe, 2004) 提取关键点并进行非重叠时空采样, 提出面向拼接全景图像的质量评价模型。Guo 等人 (2024) 提出结合运动区域分析与显著性视口选择的质量评价方法。为缓解仅依赖局部视口可能引起的预测偏差, 相关研究 (Xu 等, 2021; Zhang 和 Liu, 2022) 分别引入二维投影图像全局质量 (Xu 等, 2021) 或沿假设扫视路径提取的视口集合质量 (Zhang 和 Liu, 2022), 将全景图像中的局部与全局信息进行综合, 最后计算得出全景图像的最终质量评分。

2.3.3 基于扫视路径的方法

基于扫视路径的 VR 内容质量评价方法通过分析基于预测视觉扫视路径生成的视口序列来预测 VR 内容的感知质量。这类方法主要解决的是 VR 环境下的人类多样化的观看行为建模, 并孕育出扫视路径预测相关任务。早期 Sendjasni 等人 (2021b, 2023b) 利用预测用户扫视路径、视觉注视数据及恰可察觉差异信息识别重要视口以及实现质量评价。Sui 等人 (2022) 提出了基于人类扫视路径的全景图像质量计算框架, 采用二维视频质量评价指标或序列模型评价视口序列的综合质量。类似地, Liu 等人 (2024a, 2024b) 构建了包含视口生成、特征提取与质量预测的综合框架, 并使用真实眼动数据训练模型。

由于真实环境下用户行为数据的缺乏, Wang 等人 (2022) 提出了动态注意力视口序列重建技术, 以模拟用户体验过程。Yang 等人 (2022) 提出了结合视觉 Transformer 的扫视路径预测与双分支质量预测的模型 TVFormer。Wu 等人 (2023) 构建了 Assessor360 模型, 通过递归概率采样从给定坐标起点生成多个伪视口序列, 并辅以多尺度特征聚合以预测 VR 图像的质量。Hu 等人 (2024a) 则联合扫视路径预测、主观热力图分析与视口显著性检测等任务实现 VR 内容质量评价。Fan 等人 (2024a) 通过显式建模用户视觉扫视路径来刻画用户观看模式, 并联合优化扫视路径生成与质量预测。Yan 等人 (2025d) 采用多坐标注意力机制与基于默认扫视路径的视口采样策略进行 VR 图像质量评价。Sui 等人 (2025) 提

出了生成式扫视路径表征模型,利用 ScanDMM(Sui 等,2023)预测扫视路径,并结合视频骨干网络(Ni 等,2022)建模用户多样化感知。Wang 等人(2026)首次将强化学习与扫视路径预测相结合,并进一步应用于全景图像质量评价任务中。此外,为了应对 VR 视频中可能存在的多样真实失真,Wen 等人(2024)将伪柱面表征与卷积技术(Li 等,2024)引入到 VR 图像质量评价中,从而实现全景视频质量的有效预测。

2.3.4 小结

基于视口的方法通过聚合多个视角下的质量评分来克服整体评价难以反映真实感知的局限。其中,基于规则的方法采用系统化视口采样(如纬度感知采样、立方体映射投影),基于关键点的方法通过关键点检测方法在二维投影平面提取关键点并截取视口,基于扫视路径的方法则通过预测视觉扫视路径生成视口序列。三类方法分别从静态采样策略、关键点驱动和动态路径模拟的角度,共同构成了面向 VR 内容质量评价的视口分析技术体系。

2.4 VR 内容 QoE vs 2D 视频 QoE

当前面向 VR 内容的 QoE 评价与传统 2D 视频内容的 QoE 评价主要存在以下四个维度的差异。在感知特征差异方面,传统 2D 视频 QoE 评价主要关注平面视觉质量,核心感知维度包括清晰度、压缩失真以及时域质量等;而 VR 内容 QoE 评价在此基础上扩展了沉浸感、临场感、晕动症、空间感知等独特感知维度。在主观数据集差异方面,2D 视频 QoE 领域已建立了较为完善的数据集体系,便于不同模型的公平对比;VR 内容 QoE 数据集则存在规模偏小、失真类型局限、评分协议不统一和场景多样性不足等问题。在客观评价方法差异方面,2D 视频 QoE 客观评价已形成较为成熟的方法体系;VR 内容 QoE 客观评价则在其基础上进行扩展与改造,引入投影加权、视口聚合和多维度融合等策略以适应球面特性与观看行为。在技术挑战差异方面,2D 视频 QoE 评价面临的技术挑战主要集中在编解码优化和网络传输优化等方面;VR 内容 QoE 评价则面临球形投影与映射、视口自适应传输以及人类行为建模等额外的技术难题。

表 10 基于视口的 VR 图像/视频客观体验质量评价方法
Table 10 Summary of existing viewport-based objective QoE evaluation methods for VR content

类型	方法	启发式	数据驱动	扫视路径	观看条件	权重分配	重采样
基于规则的方法	Li 等人(2018a)♠		✓		✓	✓	✓
	P-VQA♠(Xu 等,2019b)	✓			✓	✓	✓
	V-CNN♠(Li 等,2019a;Xu 等,2022)		✓			✓	✓
	Guo 等人(2021)♠		✓				
	MMF♠(Azevedo 等,2020a,2021)	✓					✓
	MC360IQA(Sun 等,2020)		✓				✓
	Fang 等人(2022)		✓		✓		✓
	CPBQA(Jiang 等,2021)	✓					✓
	MFILGN(Zhou 等,2022a)	✓					✓
	MP-BOIQA(Jiang 等,2022a)	✓					✓
	Chai 和 Shao(2021)♠		✓				✓
	Zhou 等人(2022b)		✓				✓
	Jabar 等人(2022b)	✓					✓
	JQI♠♥(Mahmoudpour 和 Schelkens,2021)	✓					
	TB-VMAF♠♥(Jiang 等,2022b)	✓					
	Q_{STAR}^{VP} ♠♥(Meng 和 Ma,2022)	✓					

表 10 续表

类型	方法	启发式	数据驱动	扫视路径	观看条件	权重分配	重采样
	Sendjasni 等人 (2021a)		✓				✓
	SAL-360IQa (Sendjasni 和 Larabi, 2022a)		✓				✓
	Sendjasni 等人 (2022b)		✓				✓
	OmiQnet (Fan 和 Chen, 2024b)		✓				✓
	Zhou 等人 (2023a)		✓				✓
	VSOIQE (Tian 等, 2022)	✓					✓
	AHGCN (Fu 等, 2022)		✓				✓
	Liu 等人 (2023d)	✓					✓
	LSGS (Zhou 和 Wang, 2023b)		✓			✓	✓
	ST360IQ (Tofighi 等, 2023)		✓				✓
	VSBNet (Tian 等, 2023)		✓				✓
	STFINet (Hu 等, 2024)		✓				✓
	TSMT♠ (Guo 等, 2024)		✓				✓
	MTAOIQa (Yan 等, 2025a)		✓				✓
	OIQAND (Yan 等, 2025b)		✓				✓
	IQCaption360 (Yan 等, 2025c)		✓				✓
	Yan 等人 (2025g)♠		✓				✓
	Hu 等人 (2025)		✓				✓
	Liu 等人 (2025)		✓				✓
	OmniVQA (Hu 等, 2026)♠		✓				✓
基于关键点的方法	VGCN (Xu 等, 2021)		✓				✓
	Zhang 和 Liu (2022)		✓				✓
	Zhou 等人 (2023c)		✓				✓
	BVHGCN (Wan 等, 2025)		✓				✓
基于扫视路径的方法	Sendjasni 等人 (2021)		✓	✓			✓
	Sui 等人 (2022)	✓		✓	✓	✓	✓
	Wang 等人 (2022)		✓	✓		✓	✓
	TVFormer (Yang 等, 2022)		✓	✓			✓
	PW-360IQa (Sendjasni 和 Larabi, 2023b)		✓	✓			✓
	Assessor360 (Wu 等, 2023)		✓	✓	✓		✓
	Liu 等人 (2024a, 2024b)		✓	✓	✓		✓
	ESPOIQa (Hu 等, 2024a)		✓	✓			✓
	Wen 等人 (2024)♠		✓	✓	✓		✓
	Fan 等人 (2024a)♠		✓	✓			✓
	Max360IQ (Yan 等, 2025d)		✓	✓	✓		✓
	GSR-X (Sui 等, 2025)		✓	✓	✓		✓

表 10 续表

类型	方法	启发式	数据驱动	扫视路径	观看条件	权重分配	重采样
	SPP-BFR(Wang等,2025)		✓	✓			✓
	RL-ScanIQA(Wang等,2026)		✓	✓			✓

注:♥表示面向360°视频流传输的用户体验质量评价方法。

2.5 评价指标

在用户体验质量评价领域,常用于评价客观质量模型方法性能的指标主要有以下五类(Yang等,2018b),分别从不同角度衡量模型预测质量的表现。

皮尔逊相关系数(pearson's Correlation coefficient, PLCC)用于评价模型客观预测结果与主观评分之间的线性相关程度,反映的是模型“绝对数值预测准不准”。通常适合需要输出精确质量分值以生成质量报表、进行横向产品对比的场景。其计算公式如下:

$$PLCC = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} \quad (3)$$

式中, (x_i, y_i) 表示第 i 组对应的主观分数与客观预测值, $(\bar{x}, \bar{y})$ 分别为两类数据的均值。PLCC 取值范围为 $[-1, 1]$, 数值越大表明线性相关性越强, 1 表示完全正相关, -1 表示完全负相关, 0 表示无线性关系。

斯皮尔曼等级相关系数(spearman's rank order correlation coefficient, SRCC)用于衡量模型预测分数与主观评分之间的单调排序一致性,反映的是“相对优劣排得对不对”。通常适合如码率分配、内容推荐排序等需要做决策排序的场景。相应计算公式如下:

$$SRCC = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)} \quad (4)$$

式中, d_i 表示第 i 个样本的主观排序与客观排序之差, N 为样本总数。SRCC 的取值范围同样为 $[-1, 1]$, 值越接近 1 说明排序一致性越高。

肯德尔等级相关系数(kendall rank correlation coefficient, KRCC)与一致对和不一致对的数量衡量排序相关。其统计上较为保守,对样本量敏感度较低,适合数据集样本量较小或存在大量相同主观评分的平局样本的场景。相应计算公式如下:

$$KRCC = \frac{N_c - N_d}{1/2N(N-1)} \quad (5)$$

式中, N_c 和 N_d 分别表示一致对和不一致对的数量。

KRCC 取值在 $[-1, 1]$ 之间, 1 表示等级完全一致, -1 表示完全相反。

异常值比率(outlier ratio, OR)用于衡量预测分数与主观评分偏差超过一定阈值的样本比例,反映的是“模型会不会犯大错”。其适合可靠性要求极高的场景(如医疗VR、工业远程操控)。相应计算公式如下:

$$OR = \frac{N_o}{N} \quad (6)$$

式中, N_o 表示超出主观评分 ± 2 个标准差范围的预测数量。OR 值越低,表明预测结果与主观评价的一致性越好。

均方根误差(root mean square error, RMSE)用于衡量预测值与真实值之间的绝对偏差程度,反映的是“平均偏差多大”。RMSE 是设定工程容差阈值的直接依据。但受离群点影响较大,建议与 PLCC 或 SRCC 配合使用。相应计算公式如下:

$$RMSE = \sqrt{\frac{1}{N} \sum (x_i - \bar{x})^2} \quad (7)$$

RMSE 值越小,表示预测准确性越高,与主观评分的匹配度越好。

以上指标从线性相关性、单调性、等级一致性、鲁棒性及误差幅度等多个维度提供了全面评价。综合运用这些指标,可以在人类感知基准上对客观用户体验质量评价方法进行系统且稳健的验证。

3 VR QoE 模型性能对比

本文对现有主流 VR 内容体验质量评价方法在各个代表性数据集上的性能进行了系统测试,涵盖的数据集包括当前规模最大的两个全景图像主观体验质量数据库 JUF-10K(Yan等,2025b)和 OIQ-10K(Yan等,2025c),含模拟失真的全景视频主观体验质量数据库 VQA-ODV(Li等,2018a)、Fang等人(2023)提出的含真实失真的数据库,以及面向全景视频自适应传输的主观数据库 VOD-VQA(Meng和

Ma, 2022)。其中, Fang 等人(2023)所构建的数据库包含两种用户体验时间条件设置(7秒与15秒), 在该数据集上用于比较的 ERP-VQA 与 Scanpath-VQA 两种客观方法分别代表以 ERP 格式帧和沿用户扫描路径提取的视口序列为输入的质量评价模型。针对各主观数据库, 本文选取了相应的代表性客观 QoE 评价模型, 并以 SRCC 与 PLCC 为指标进行性能比较。结果如图9所示。根据图中结果, 可以得出以下结论。首先, 各方法在不同数据库上的性能表现存在差异, 而部分方法在同一数据库上的性能差

异相对较小。其次, 相较于传统的全参考质量评价模型, 现有基于深度学习与面向特定内容的无参考质量评价模型表现更佳。同时, 相较于传统模拟失真数据库, 面向真实失真的全景视频质量评价更具挑战性, 整体提升空间较大, 并且实验结果也表明模型性能与用户观看条件与观看行为密切相关。此外, 当前面向全景视频自适应传输场景的客观 QoE 评价方法仍较为有限, 尽管部分最新方法取得了不错的性能, 但在实际应用中仍存在较大的局限性。

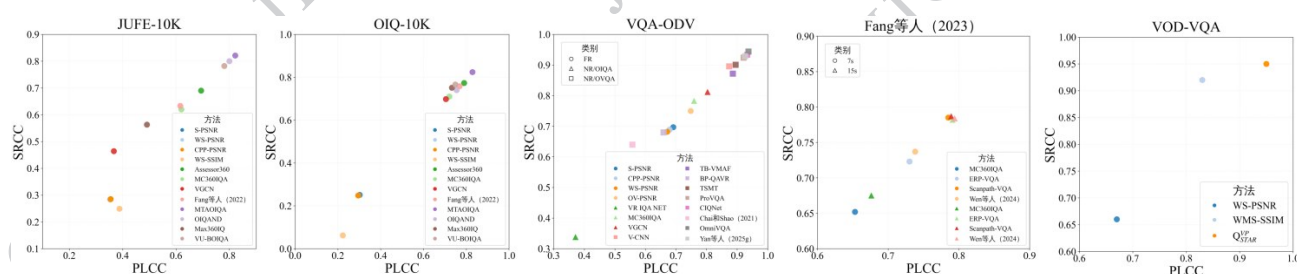


图9 主流VR内容体验质量客观评价模型在各类数据集上的性能对比

Fig. 9 Performance comparison of existing VR QoE objective methods on different subjective datasets

4 VR QoE模型应用

VR QoE客观模型在实际应用中发挥着重要作用, 主要体现为对VR全链条相关技术的参数优化与方法评价与选择, 最终综合提升用户的整体体验。在参数优化方面, VR内容的生成、处理、传输、存储、投影和显示等环节均涉及大量可调软硬件参数。例如在VR内容生成与拼接过程中, 不同参数组合会显著影响最终输出内容的质量。在VR内容传输方面, 可以为自适应传输策略提供感知质量反馈, 从而支撑动态码率调节、分块调度和视口自适应传输等机制的优化。而传统的优化方式主要依赖主观实验, 即由人类观察者逐一评判各参数下的输出效果, 但参数组合空间通常较大, 导致人力成本高、效率低, 难以在工程实践中推广。因此, VR QoE客观模型可作为高效的替代方案, 充当自动化质量评判器, 对不同参数设置下输出的VR内容进行客观评分, 从而实现快速、可复现的最佳参数选择。在技术方法的评价与选择方面, 同类任务往往存在多种实现方案, 各有其适用条件与性能侧重。以拼接技术为例, 不同算法在拼接精度、几何一致性、光影过渡等

方面表现各异, 实际场景中难以先验确定最优方案。借助VR QoE客观模型, 可以对多种候选方法在同一输入上的输出结果进行统一质量评价, 客观衡量其视觉表现, 从而指导最终的技术选型。这种基于用户感知与体验质量评价的决策机制, 兼顾了客观性与执行效率, 能够为VR系统设计与多样化场景下的工程化部署提供了可量化、可比较的依据。

5 结 语

5.1 总结

本文对VR内容在整个处理流程中的用户体验质量评价最新进展进行了系统而全面的综述。本文构建了一个结构化的分类体系, 系统归纳了当前面向VR内容的用户体验质量评价框架中的核心方法, 将其划分为两种基本范式, 并深入剖析了其内在动机与技术原理。为方便本领域研究者, 本文总结了当前可用的开源数据集与模型方法。本综述不仅整合了现有知识体系, 还对当前面临的持续挑战与新兴机遇进行了评述, 促进该领域的未来发展。

尽管面向VR内容的用户体验质量评价研究已取得显著进展, 当前存在的若干关键挑战仍制约着

相应方法的开发与应用。首要局限在于当前研究对于用户观看条件(如分辨率、视场角及头戴式显示器性能限制)的忽视,而这些因素对用户感知具有显著影响。其次,多数客观体验质量评价方法依赖确定性视口序列,未考虑用户真实观看行为的概率特性,从而不可避免的引入了预测偏差。同时,客观评价方法的计算复杂度进一步限制了实时应用,尤其在基于扫视路径的方法中,低效的视口提取过程成为瓶颈。此外,数据稀缺仍是一大障碍,现有数据集在规模与多样性方面仍显不足,制约了模型的整体泛化能力。最后,当前面向VR内容的体验质量评价框架普遍缺乏跨投影格式的兼容性,限制了其在各类VR内容表示形式之间的适用性,因此,亟需开发与投影格式无关的解决方案。

5.2 未来发展方向

未来研究重点可以关注以下几个方向,以推动面向VR内容的用户体验质量评价的进一步发展。

1)随着扩展现实技术向全感官沉浸方向演进,面向多模态体感的VR内容QoE评价除考虑视听双模态外,还需进一步纳入触觉、嗅觉、味觉以及运动感知等多模态体感维度。模态信息的增加带来了跨模态感知一致性与同步约束、多模态质量融合和多模态主观实验复杂性等新挑战(Yang等,2025;Zhu等,2025b)。然而,该方向当前研究成果极为有限,尚缺乏系统性综述基础。因此,亟需深入探索。未来研究可以从多通道标注标准化数据集与评测基准构建、结合生物信号的客观生理指标映射模型建立,以及基于大模型的语义级评价与交互机制探索等方向推进。

2)随着沉浸式媒体技术的不断演进,六自由度虚拟现实内容以及增强/混合现实内容正逐步成为学术界和产业界关注的新焦点。其中,六自由度虚拟现实内容能够支持用户在三维空间中自由移动与转向,赋予了用户动态空间体验,但也引入了空间连续性与视觉一致性以及自由视点伪影等挑战(Chen等,2024)。而增强/混合现实内容将虚拟内容叠加于真实物理世界之上,其QoE评价需额外考虑虚实融合自然度、环境感知自适应性和交互反馈真实感等问题(Wang等,2024;Combe等,2024)。相较于普通VR内容,上述两类内容的QoE评价研究尚处于早期探索阶段,无论是标准化主观实验方法还是成熟客观评价模型亟待完善。

3)基于多模态大模型(large multimodal models, LMMs)的用户体验质量预测潜力初显。相关研究,如Q-Bench(Zhang等,2024)、Q-Align(Wu等,2024)等,已证明其在感知质量预测和推理任务中的能力。相关进展可参加Zhang等人(2025a,2025b)的研究。然而,将其应用于VR内容质量评价尚处于起步阶段。未来工作应探索基于多模态大模型的可扩展、强泛化能力的体验质量评价框架。

4)随着生成式人工智能技术的快速发展,AI生成VR内容正成为高质量虚拟环境构建的重要途径(Liu等,2025b;Yang等,2026)。然而,此类内容在丰富内容创作形式的同时,也带来了质量评价与安全检测层面的新挑战。与自然采集或合成内容不同,AI生成内容可能出现语义不一致、结构失真、区域伪影及逻辑冲突等问题,且不同生成模型引发的失真类型存在显著差异。为此,亟需建立涵盖真实性、结构一致性、语义保真度、视觉自然度及对常见失真鲁棒性等多维度的新型质量评价标准,确保AI生成内容满足VR应用的体验质量阈值,保障用户接纳度与系统可靠性。

5)晕动症仍是制约VR应用推广与用户沉浸式体验持续性的主要障碍(Venkatkrishnan等,2024)。尽管已有研究初步探讨了其引发的恶心、视觉疲劳及用户排斥等主观不适反应,但尚未建立能够系统量化其严重程度的标准质量评价指标和计算模型。现有评价手段多依赖主观问卷,缺乏客观、连续、可泛化的生理信号映射与感知度量体系。弥补这一缺口对于设计内容呈现策略、优化视口切换逻辑及发展个性化防晕动方案具有重要意义,也是提升VR系统用户安全性与可接受度的关键环节。

6)面向个体差异的个性化用户体验质量建模研究仍显薄弱。现有方法通常采用“一刀切”范式,默认所有用户的感知偏好、观看行为与感官敏感度具有一致性,而忽视了个体差异对主观质量评价的显著影响。因此,亟需构建可从用户交互行为与生理反馈中持续学习的自适应评价模型,实现真正以用户为中心的QoE精准预测,支撑个性化内容推荐与体验优化。

7)VR自适应流传输场景对用户体验质量评价提出了新的需求。传统视频流媒体评价模型主要面向平面内容,难以充分应对VR环境的复杂性,例如视口依赖的分块传输策略中动态码率分配与延迟敏

感性、视口切换引入的卡顿与不一致等问题。现有指标方法未能有效关联网络传输参数与用户主观感知之间的耦合关系。因此,亟需发展面向VR实时流传输的定制化QoE评价框架,以指导自适应调度策略设计和系统性能优化,保障在带宽波动等网络条件下依然维持稳定的沉浸式体验。

参考文献(References)

- Abreu A D, Ozcinar C and Smolic A. 2017. Look around you: saliency maps for omnidirectional images in VR applications//Proceedings of the Ninth International Conference on Quality of Multimedia Experience. Erfurt, Germany. 1-6 [DOI: 10.1109/QoMEX.2017.7965634]
- Agtzidis I, Startsev M and Dorr M. 2019. 360-degree video gaze behaviour: A ground-truth data set and a classification algorithm for eye movements//Proceedings of the 27th ACM International Conference on Multimedia. Nice, France: ACM: 1007-1015 [DOI: 10.1145/3343031.3350947]
- Akhtar Z and Falk T H. 2017. Audio-visual multimedia quality assessment: a comprehensive survey. IEEE Access, 5: 21090-21117 [DOI: 10.1109/ACCESS.2017.2750918]
- Anwar M S, Wang J, Ahmad S, Khan W, Ullah A, Shah M, et al. 2020a. Impact of the impairment in 360-degree videos on users' involvement and machine learning-based qoe predictions. IEEE Access, 8: 204585-204596 [DOI: 10.1109/ACCESS.2020.3037253]
- Anwar M S, Wang J, Khan W, Ullah A, Ahmad S and Fei Z S. 2020b. Subjective qoe of 360-degree virtual reality videos and machine learning predictions. IEEE Access, 8: 148084-148099 [DOI: 10.1109/ACCESS.2020.3015556]
- Anwar M S, Wang J, Ullah A, Khan W, Ahmad S and Fei Z S. 2020c. Measuring quality of experience for 360-degree videos in virtual reality. Science China Information Sciences, 63: 202301 [DOI: 10.1007/s11432-019-2734-y]
- Azevedo R G A, Birkbeck N, Janatra I, Adsumilli B and Frossard P. 2020. A viewport-driven multi-metric fusion approach for 360-degree video quality assessment//Proceedings of the IEEE International Conference on Multimedia and Expo. London, UK: IEEE: 1-6 [DOI: 10.1109/ICME46284.2020.9102936]
- Azevedo R G A, Birkbeck N, Janatra I, Adsumilli B and Frossard P. 2021. Multi-feature 360 video quality estimation. IEEE Open Journal of Circuits and Systems, 2: 338-349 [DOI: 10.1109/OJCS.2021.3073891]
- Azevedo R G A, Birkbeck N, Simone F D, Janatra I, Adsumilli B and Frossard P. 2020b. Visual distortions in 360 videos. IEEE Transactions on Circuits and Systems for Video Technology, 30(8): 2524-2537 [DOI: 10.1109/TCSVT.2019.2927344]
- Balk S A, Bertola D B and Inman V W. 2013. Simulator sickness questionnaire: twenty years later//Proceedings of the Driving Assessment Conference, vol. 7. University of Iowa.
- Bao Y N, Wu H S, Zhang T X, Ramli A A and Liu X. 2016. Shooting a moving target: Motion-prediction-based transmission for 360-degree videos//Proceedings of the IEEE International Conference on Big Data. Washington, DC, USA: IEEE: 1161-1170 [DOI:]
- Bay H, Tuytelaars T and Gool L V. 2006. SURF: speeded up robust features//Proceedings of the European Conference on Computer Vision. Berlin, Heidelberg: Springer: 404-417 [DOI: 10.1109/BigData.2016.7840720]
- Biswas N, Mukherjee A and Bhattacharya S. 2024. Are you feeling sick? - A systematic literature review of cybersickness in virtual reality. ACM Computing Surveys, 56(11): 1-38 [DOI: 10.1145/3670008]
- Bos J E, Vries S C, Emmerik M L and Groen E L. 2010. The effect of internal and external fields of view on visually induced motion sickness. Applied Ergonomics, 41(4): 516-521 [DOI: 10.1016/j.apergo.2009.11.007]
- Bouraia K, Sabir E, Sadik M and Ladid L. 2020. Quality of experience for streaming services: Measurements, challenges and insights. IEEE Access, 8: 13341-13361 [DOI: 10.1109/ACCESS.2020.2965099]
- Boyce J M, Alshina E, Abbas A and Ye Y. 2016. JVET common test conditions and evaluation procedures for 360 video. In Joint Video Exploration Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11.
- Callet P L and Möller S. 2013. Qualinet white paper on definitions of quality of experience. Technical Report.
- Chai X L and Shao F. 2021. Blind quality assessment of omnidirectional videos using spatio-temporal convolutional neural networks. Optik, 226: 165887 [DOI: 10.1016/j.ijleo.2020.165887]
- Chen J H and Niu Y Z. 2022. No-reference omnidirectional image quality assessment via equirectangular convolution and hierarchically fusion//Proceedings of the International Conference on Artificial Intelligence and Pattern Recognition. Xiamen, China: ACM: 643-648 [DOI: 10.1145/3573942.3574074]
- Chen M X, Jin Y Z, Goodall T, Yu X X and Bovik A C. 2020a. Study of 3d virtual reality picture quality. IEEE Journal of Selected Topics in Signal Processing, 14(1): 89-102 [DOI: 10.1109/JSTSP.2019.2956408]
- Chen S J, Zhang Y X, Li Y M, Chen Z Z and Wang Z. 2018a. Spherical structural similarity index for objective omnidirectional video quality assessment//Proceedings of the IEEE International Conference on Multimedia and Expo. San Diego, CA, USA: IEEE: 1-6 [DOI: 10.1109/ICME.2018.8486584]
- Chen Y C, Saha A, Chapiro A, Hane C, Bazin J, Qiu B, et al. 2024. Subjective and objective quality assessment of rendered human avatar videos in virtual reality. IEEE Transactions on Image Processing, 33: 5740-5754 [DOI: 10.1109/TIP.2024.3468881]

- Chen Z B, Xu J H, Lin C Y and Zhou W. 2020b. Stereoscopic omnidirectional image quality assessment based on predictive coding theory. *IEEE Journal of Selected Topics in Signal Processing*, 14 (1): 103-117 [DOI: 10.1109/JSTSP.2020.2968182]
- Chen Z Z, Li Y M and Zhang Y X. 2018b. Recent advances in omnidirectional video coding for virtual reality: projection and evaluation. *Signal Processing*, 14 (6): 66-78 [DOI: 10.1016/j.sigpro.2018.01.004]
- Chiariotti F. 2021. A survey on 360-degree video: coding, quality of experience and streaming. *Computer Communications*, 177 (2021): 133-155 [DOI: 10.1016/j.comcom.2021.06.029]
- Chu T M C, Zepernick H J and Elwardy M. 2020. Rating duration analysis for subjective quality assessment of 360 videos//*Proceedings of the International Conference on Virtual Reality and Visualization*. Recife, Brazil: IEEE: 42-46 [DOI: 10.1109/ICVRV51359.2020.00020]
- Cohen T S, Geiger M, Köhler J and Welling M. 2018. Spherical cnns//*Proceedings of the International Conference on Learning Representations*. Vancouver, Canada
- Combe T, Fribourg R, Detto L and Normand J. 2024. Exploring the influence of virtual avatar heads in mixed reality on social presence, performance and user experience in collaborative tasks. *IEEE Transactions on Visualization and Computer Graphics*, 30 (5): 2206-2216 [DOI: 10.1109/TVCG.2024.3372051]
- Corbillion X, Simone F D and Simon G. 2017. 360-degree video head movement dataset//*Proceedings of the 8th ACM Multimedia Systems Conference*. Taipei, Taiwan, China: ACM: 199-204 [DOI: 10.1145/3083187.3083215]
- Cornia M, Baraldi L, Serra G and Cucchiara R. 2016. A deep multi-level network for saliency prediction//*Proceedings of the 23rd International Conference on Pattern Recognition*. Caocun, Mexico: IEEE: 3488-3493 [DOI: 10.1109/ICPR.2016.7900174]
- Cortés C, Pérez P, Gutiérrez J and García N. 2020. Influence of video delay on quality, presence, and sickness in viewport adaptive immersive streaming//*Proceedings of the Twelfth International Conference on Quality of Multimedia Experience*. Athlone, Ireland: 1-4 [DOI: 10.1109/QoMEX48832.2020.9123114]
- Croci S, Ozcinar C, Zerman E, Cabrera J and Smolic A. 2019. Voronoi-based objective quality metrics for omnidirectional video//*Proceedings of the Eleventh International Conference on Quality of Multimedia Experience*. Berlin, Germany: 1-6 [DOI: 10.1109/QoMEX.2019.8743345]
- Curcio I D D, Toukoma H and Naik D. 2017. Bandwidth reduction of omnidirectional viewport-dependent video streaming via subjective quality assessment//*Proceedings of the 2nd International Workshop on Multimedia Alternate Realities*. Mountain View, California, USA: ACM: 9-14 [DOI: 10.1145/3132361.3132364]
- Dahou Y, Tliba M, McGuinness K and Connor N O. 2021. ATSal: an attention based architecture for saliency prediction in 360 videos//*Proceedings of the International Conference on Pattern Recognition*. 305-320
- Dai J F, Qi H Z, Xiong Y W, Li Y, Zhang G D, Hu H, et al. 2017. Deformable convolutional networks//*Proceedings of the IEEE International Conference on Computer Vision*. Venice, Italy: IEEE: 764-773 [DOI: 10.1109/ICCV.2017.89]
- David E J, Gutiérrez J, Coutrot A, Silva M P D and Callet P L. 2018. A dataset of head and eye movements for 360 videos//*Proceedings of the 9th ACM Multimedia Systems Conference*. Amsterdam, Netherlands: ACM: 432-437 [DOI: 10.1145/3204949.3208139]
- Dennison M S, Wisti A Z and Zmura M D. 2016. Use of physiological signals to predict cybersickness. *Displays*, 44: 42-52 [DOI: 10.1016/j.displa.2016.07.002]
- Ding W X, An P, Liu X, Yang C and Huang X P. 2021. No-reference panoramic image quality assessment based on adjacent pixels correlation//*Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*. Chengdu, China: IEEE: 1-5 [DOI: 10.1109/BMSB53066.2021.9547132]
- Ding W X, An P, Yang C and Huang X P. 2020. Panoramic video assessment based on cascaded network using saliency map. In *Optoelectronic Imaging and Multimedia Technology VII*, Vol. 11550. SPIE, 289-295 [DOI: 10.1117/12.2575107]
- Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X H, Unterthiner T, et al. 2021. An image is worth 16x16 words: transformers for image recognition at scale//*Proceedings of the International Conference on Learning Representations*
- Duan H Y, Min X K, Sun W, Zhu Y C, Zhang X P and Zhai G T. 2023. Attentive deep image quality assessment for omnidirectional stitching. *IEEE Journal of Selected Topics in Signal Processing*, 17(6): 1150-1164 [DOI: 10.1109/JSTSP.2023.3250956]
- Duan H Y, Zhai G T, Min X K, Zhu Y C, Fang Y and Yang X K. 2018. Perceptual quality assessment of omnidirectional images//*Proceedings of the IEEE International Symposium on Circuits and Systems*. Florence, Italy: IEEE: 1-5 [DOI: 10.1109/ISCAS.2018.8351786]
- Duan H Y, Zhai G T, Min X K, Zhu Y C, Sun W and Yang X K. 2017a. Assessment of visually induced motion sickness in immersive videos//*Proceedings of the Pacific Rim Conference on Multimedia*. Harbin, China: Springer: 662-672 [DOI: 10.1007/978-3-319-77380-3_63]
- Duan H Y, Zhai G T, Yang X K, Li D and Zhu W H. 2017b. IVQAD 2017: an immersive video quality assessment database//*Proceedings of the International Conference on Systems, Signals and Image Processing*. Poznan, Poland: IEEE: 1-5. [DOI: 10.1109/IWSSIP.2017.7965610]
- Duan H Y, Zhu X L, Zhu Y X, Min X K and Zhai G T. 2024. A quick review of human perception in immersive media. *IEEE Open Journal on Immersive Displays*, 1: 41-50 [DOI: 10.1109/OJID.2024.3351089]
- Duanmu F Y, Mao Y X, Liu S, Srinivasan S and Wang Y. 2018. A sub-

- jective study of viewer navigation behaviors when watching 360-degree videos on computers//Proceedings of the IEEE International Conference on Multimedia and Expo. San Diego, CA, USA: IEEE: 1-6 [DOI: 10.1109/ICME.2018.8486537]
- Elwardy M, Zepernick H J and Hu Y. 2022. SSV360: a dataset on subjective quality assessment of 360 videos for standing and seated viewing on an hmd//Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops. Christchurch, New Zealand: IEEE: 1-6 [DOI: 10.1109/VRW55335.2022.00047]
- Elwardy M, Zepernick H J, Hu Y and Chu T M C. 2023. ACR360: a dataset on subjective 360 video quality assessment using acr methods//Proceedings of the 16th International Conference on Signal Processing and Communication System. Bydgoszcz, Poland: IEEE: 1-8 [DOI: 10.1109/ICSPCS58109.2023.10261151]
- Elwardy M, Zepernick H J and Sundstedt V. 2019. Annotated 360-degree image and video databases: A comprehensive survey//Proceedings of the International Conference on Signal Processing and Communication Systems. Gold Coast, QLD, Australia: IEEE: 1-6. [DOI: 10.1109/ICSPCS47537.2019.9008715]
- Fan C L, Lo W C, Pai Y T and Hsu C H. 2019. A survey on 360 video streaming: acquisition, transmission, and display. *ACM Computing Surveys*, 52(4): 1-36 [DOI: 10.1145/3329119]
- Fan K L, Wen W, Li M, Peng Y F and Ma K D. 2024a. Learned scan-paths aid blind panoramic video quality assessment//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA: IEEE: 2599-2608 [DOI: 10.1109/CVPR52733.2024.00251]
- Fan Y and Chen C Y. 2024b. OmiQnet: multiscale feature aggregation convolutional neural network for omnidirectional image assessment. *Applied Intelligence*, 54: 5711-5727 [DOI: 10.1007/s10489-024-05421-1]
- Fang Y C, Cagri O and Aljosa S. 2023. A qoe study for viewport prediction approaches in tile-based 360 video streaming. [DOI: 10.13140/RC.2.2.15543.80802]
- Fang Y M, Huang L P, Yan J B, Liu X L and Liu Y. 2022. Perceptual quality assessment of omnidirectional images//Proceedings of the AAAI Conference on Artificial Intelligence, 36(1): 580-588 [DOI: 10.1609/aaai.v36i1.19937]
- Fang Y M, Sui X J, Yan J B, Liu X L and Huang L P. 2021. Progress in no-reference image quality assessment. *Journal of Image and Graphics*, 26(2): 265-286 (方玉明, 睦相杰, 鄢杰斌, 刘学林, 黄丽萍. 2021. 无参考图像质量评价研究进展. *中国图象图形学报*, 26(2): 265-286)[DOI: 10.11834/jig.200274]
- Fang Y M, Yao Y R, Sui X J and Ma K D. 2023. Subjective quality assessment of user-generated 360 videos//Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops. Shanghai, China: IEEE: 723-724 [DOI: 10.1109/VRW58643.2023.00204]
- Fei Z S, Wang F, Wang J and Xie X. 2020. Qoe evaluation methods for 360-degree vr video transmission. *IEEE Journal of Selected Topics in Signal Processing*, 14(1): 78-88 [DOI: 10.1109/JSTSP.2019.2956631]
- Fela R F, Pastor A, Callet P L, Zacharov N, Vigier T and Forchhammer S. 2022. Perceptual evaluation on audio-visual dataset of 360 content//Proceedings of the IEEE International Conference on Multimedia and Expo Workshops. Taipei, Taiwan, China: IEEE: 1-6 [DOI: 10.1109/ICMEW56448.2022.9859426]
- Fremeray S, Singla A, Meseberg K and Raake A. 2018. AVtrack360: An open dataset and software recording people's head rotations watching 360 videos on an hmd//Proceedings of the 9th ACM Multimedia Systems Conference. Amsterdam, Netherlands: ACM: 403-408 [DOI: 10.1145/3204949.3208134]
- Fu J, Hou C, Zhou W, Xu J H and Chen Z B. 2022. Adaptive hypergraph convolutional network for no-reference 360-degree image quality assessment//Proceedings of the 30th ACM International Conference on Multimedia. Lisboa, Portugal: ACM: 961-969 [DOI: 10.1145/3503161.3548337]
- Gao P, Zhang P W and Smolic A. 2022. Quality assessment for omnidirectional video: A spatio-temporal distortion modeling approach. *IEEE Transactions on Multimedia*, 24: 1-16 [DOI: 10.1109/TMM.2020.3044458]
- Ghadiyaram D and Bovik A C. 2016. Massive online crowdsourced study of subjective and objective picture quality. *IEEE Transactions on Image Processing*, 25(1): 372-387 [DOI: 10.1109/TIP.2015.2500021]
- Gianaros P J, Muth E R, Mordkoff J T, Levine M E and Stern R M. 2001. A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, Space, and Environmental Medicine*, 72(2): 115-119
- Golding J F. 1998. Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness. *Brain Research Bulletin*, 47(5): 507-516 [DOI: 10.1016/S0361-9230(98)00091-4]
- Guo J F and Luo Y. 2021. No-reference omnidirectional video quality assessment based on generative adversarial networks. *Multimedia Tools and Applications*, 80: 27531-27552 [DOI: 10.1007/s11042-021-10862-8]
- Guo J F and Luo Y. 2024. Two-stream network with viewport selection for blind omnidirectional video quality assessment. *Multimedia Tools and Applications*, 83: 12139-12157 [DOI: 10.1007/s11042-023-15739-6]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hoßfeld T, Keimel C, Hirth M, Gardlo B, Habigt J, Diepold K, et al. 2014. Best practices for qoe crowdtesting: Qoe assessment with crowdsourcing. *IEEE Transactions on Multimedia*, 16(2): 541-

- 558 [DOI: 10.1109/TMM.2013.2291663]
- Hoßfeld T, Schatz R and Egger S. 2011. SOS: the mos is not enough! // Proceedings of the Third International Workshop on Quality of Multimedia Experience. Mechelen, Belgium: IEEE: 131-136 [DOI: 10.1109/QoMEX.2011.6065690]
- Hristova H, Simon G, Corbillion X, Devlic A and Swaminathan V. 2021. Heterogeneous spatial quality for omnidirectional video. IEEE Transactions on Circuits and Systems for Video Technology, 31(4): 1411-1424 [DOI: 10.1109/TCSVT.2020.3007620]
- Hu H N, Lin Y C, Liu M Y, Cheng H T, Chang Y J and Sun M. 2017. Deep 360 pilot: learning a deep agent for piloting through 360 sports videos//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE: 1396-1405 [DOI: 10.1109/CVPR.2017.153]
- Hu H X, Shao F and Chen H W. 2024a. Eye scanpath prediction-based no-reference quality assessment of omnidirectional images. IEEE Transactions on Instrumentation and Measurement, 73: 1-15 [DOI: 10.1109/TIM.2024.3470239]
- Hu H X, Shao F, Chen H W, Chai X L and Jiang Q P. 2025. Cross-projection distilling knowledge for omnidirectional image quality assessment. IEEE Transactions on Multimedia, 27: 6752-6765 [DOI: 10.1109/TMM.2025.3590920]
- Hu H X, Shao F, Wang H Z, Mu B Y, Chen H W, Jiang Q P, et al. 2024b. Spatio-temporal feature integration for quality assessment of stitched omnidirectional images. IEEE Transactions on Emerging Topics in Computational Intelligence, 8(2): 1484-1499 [DOI: 10.1109/TETCI.2023.3341333]
- Hu Z Y, Liu L X, Gu K, Li L D and Bovik A C. 2026. Distortion-sensitive masked autoencoder for omnidirectional video quality assessment. IEEE Transactions on Multimedia, 28: 3353-3363 [DOI: 10.1109/TMM.2026.3651128]
- Hu Z Y, Liu L X and Sang Q B. 2024c. Omnidirectional video quality assessment with causal intervention. IEEE Transactions on Broadcasting, 70, 1: 238-250 [DOI: 10.1109/TBC.2023.3342707]
- Huang M K, Shen Q, Ma Z, Bovik A C, Gupta P, Zhou R B, et al. 2018. Modeling the perceptual quality of immersive images rendered on head mounted displays: Resolution and compression. IEEE Transactions on Image Processing, 27 (12): 6039-6050 [DOI: 10.1109/TIP.2018.2865089]
- Huang Y K, Gao Y C, Zhu Y W, Chen Z G, Sidhu R, Qiao X Q, et al. 2026. A review of 6-DoF volumetric video communication systems: AI empowered perspectives. IEEE Communications Surveys & Tutorials [DOI: 10.1109/COMST.2026.3726902]
- ITU. 2002. Methodology for the subjective assessment of the quality of television pictures [EB/OL]. ITU-R Rec. BT. 500-11. [2002-06-28].
- ITU. 2007. Methodology for the subjective assessment of video quality in multimedia applications [EB/OL]. ITU-R Rec. BT. 1788. [2007-01-26]
- ITU. 2008. Subjective video quality assessment methods for multimedia applications [EB/OL]. ITU-T Rec. P.910. [2008-04-06]
- ITU. 2012. Methodology for the subjective assessment of the quality of television pictures [EB/OL]. ITU-R Rec. BT. 500-13. [2012-01-27].
- Jabar F, Ascenso J and Queluz M P. 2020a. Field-of-view effect on the perceived quality of omnidirectional images//Proceedings of the IEEE International Conference on Multimedia & Expo Workshops. London, UK: IEEE: 1-6 [DOI: ICMEW46912.2020.9106059]
- Jabar F, Ascenso J and Queluz M P. 2020b. Objective assessment of perceived geometric distortions in viewport rendering of 360 images. IEEE Journal of Selected Topics in Signal Processing, 14(1): 49-63 [DOI: 10.1109/JSTSP.2019.2962970]
- Jeroen V D H, Hadi A, Maria T V, Yago S, Raimund S and Thomas S. 2023. A tutorial on immersive video delivery: From omnidirectional video to holography. IEEE Communications Surveys & Tutorials, 25 (2): 1336-1375 [DOI: 10.1109/COMST.2023.3263252]
- Jiang H, Jiang G Y, Yu M, Luo T and Xu H Y. 2022a. Multi-angle projection based blind omnidirectional image quality assessment. IEEE Transactions on Circuits and Systems for Video Technology, 32(7): 4211-4223 [DOI: 10.1109/TCSVT.2021.3128014]
- Jiang H, Jiang G Y, Yu M, Zhang Y, Yang Y, Peng Z J, et al. 2021. Cubemap-based perception-driven blind quality assessment for 360-degree images. IEEE Transactions on Image Processing, 30: 2364-2377 [DOI: 10.1109/TIP.2021.3052073]
- Jiang Z Q, Xu Y L, Sun J, Hwang J N, Zhang Y F and Appleby S C. 2022b. Tile-based panoramic video quality assessment. IEEE Transactions on Broadcasting, 68 (2): 530-544 [DOI: 10.1109/TBC.2021.3136748]
- Juluri P, Tamarapalli V and Medhi D. 2016. Measurement of quality of experience of video-on-demand services: A survey. IEEE Communications Surveys & Tutorials, 18(1): 401-418 [DOI: 10.1109/COMST.2015.2401424]
- JVET. 2023. Algorithm description of projection format conversion and video quality metrics in 360lib version. Standard JVET-G1003. Joint Video Experts Team (JVET). <https://jvet-experts.org/>.
- Kalawsky R S. 2000. The validity of presence as a reliable human performance metric in immersive environments//Proceedings of the International Workshop on Presence. Delft, Netherlands: 1-16
- Kennedy R S, Lane N E, Berbaum K S and Lilienthal M G. 1993. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. The International Journal of Aviation Psychology, 3(3): 203-220 [DOI: 10.1207/s15327108ijap0303_3]
- Keshavarz B and Hecht H. 2011. Validating an efficient method to quantify motion sickness. Human Factors, 53(4): 415-426 [DOI: 10.1177/0018720811403736]
- Kim H G, Lim H T, Lee S and Ro Y M. 2019. VRSA Net: VR sickness assessment considering exceptional motion for 360° vr video. IEEE Transactions on Image Processing, 28(4): 1646-1660 [DOI: 10.

- 1109/TIP.2018.2880509]
- Kim H G, Lim H T and Ro Y M. 2020. Deep virtual reality image quality assessment with human perception guider for omnidirectional image. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(4): 917-928 [DOI: 10.1109/TCSVT.2019.2898732]
- Kim H K, Park J, Choi Y and Choe M. 2018. Virtual reality sickness questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. *Applied Ergonomics*, 69: 66-73 [DOI: 10.1016/j.apergo.2017.12.016]
- Kim Y Y, Kim H J, Kim E N, Ko H D and Kim H T. 2010. Characteristic changes in the physiological components of cybersickness. *Psychophysiology*, 42(5): 616-625 [DOI: 10.1111/j.1469-8986.2005.00349.x]
- Labrador C F, Facil J M, Yus A P, Demonceaux C, Civera J and Guerrero J J. 2020. Corners for layout: end-to-end layout recovery from 360 images. *IEEE Robotics and Automation Letters*, 5(2): 1255-1262 [DOI: 10.1109/LRA.2020.2967274]
- Li C, Xu M, Du X Z and Wang Z L. 2018a. Bridge the gap between vqa and human behavior on omnidirectional video: a large-scale dataset and a deep learning model//*Proceedings of the 26th ACM International Conference on Multimedia*. Seoul, Republic of Korea: ACM: 932-940 [DOI: 10.1145/3240508.3240581]
- Li C, Xu M, Jiang L, Zhang S Y and Tao X M. 2019a. Viewport proposal cnn for 360 video quality assessment//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, CA, USA: IEEE: 10169-10178 [DOI: 10.1109/CVPR.2019.01042]
- Li J, Bailenson J N, Pines A, Greenleaf W J and Williams L M. 2017. A public database of immersive vr videos with corresponding ratings of arousal, valence, and correlations between head movements and self report measures. *Frontiers in Psychology*, 8 [DOI: 10.3389/fpsyg.2017.02116]
- Li J, Feng R S, Liu Z, Sun W and Li Q Y. 2018b. Modeling qoe of virtual reality video transmission over wireless networks//*Proceedings of the IEEE Global Communications Conference*. Abu Dhabi, United Arab Emirates: IEEE: 1-7 [DOI: 10.1109/GLOCOM.2018.8647729]
- Li J, Yu K W, Zhao Y F, Zhang Y and Xu L. 2019b. Cross-reference stitching quality assessment for 360 omnidirectional images//*Proceedings of the 27th ACM International Conference on Multimedia*. Nice, France: ACM: 2360-2368 [DOI: 10.1145/3343031.3350973]
- Li J, Zhao Y F, Ye W H, Yu K W and Ge S M. 2020. Attentive deep stitching and quality assessment for 360 omnidirectional images. *IEEE Journal of Selected Topics in Signal Processing*, 14(1): 209-221 [DOI: 10.1109/JSTSP.2019.2953950]
- Li M, Bao Y N, Sui X H, Li J X, Lu G M and Xu Y. 2024. Learning content-weighted pseudocylindrical representation for 360 image compression. *IEEE Transactions on Image Processing*, 33: 5975-5988 [DOI: 10.1109/TIP.2024.3477356]
- Lim H T, Kim H G and Ra Y M. 2018. VR IQA NET: deep virtual reality image quality assessment using adversarial learning//*Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*. Calgary, AB, Canada: IEEE: 6737-6741 [DOI: 10.1109/ICASSP.2018.8461317]
- Ling S Y, Cheung G and Callet P L. 2018. No-reference quality assessment for stitched panoramic images using convolutional sparse coding and compound feature selection//*Proceedings of the IEEE International Conference on Multimedia and Expo*. San Diego, CA, USA: IEEE: 1-6 [DOI: 10.1109/ICME.2018.8486545]
- Liu X L, Yan J B, Fang Y M, Hou J W. 2025a. Opinion-unaware blind quality assessment of AI-generated omnidirectional images based on deep feature statistics. *Journal of Visual Communication and Image Representation*, 110: 104461 [DOI: 10.1016/j.jvcir.2025.104461]
- Liu X L, Yan J B, Huang L P, Fang Y M, Wan Z and Liu Y. 2024a. Perceptual quality assessment of omnidirectional images: a benchmark and computational model. *ACM Transactions on Multimedia Computing, Communications and Applications*, 20(6): [DOI: 10.1145/3640344]
- Liu X L, Yan J B, Lai C Y, Li Y and Fang Y M. 2025b. Viewport-independent blind quality assessment of ai-generated omnidirectional images via vision-language correspondence. *IEEE Signal Processing Letters*, 32: 1630-1634 [DOI: 10.1109/LSP. 2025.3556791]
- Liu X L, Yan J B, Wan Z, Fang Y M and Liu H T. 2024b. Blind quality assessment of panoramic images based on multiple viewport sequences//*Proceedings of the IEEE International Symposium on Circuits and Systems*. Singapore: IEEE: 1-5 [DOI: 10.1109/ISCAS58744.2024.10558321]
- Liu X L, Yan J B, Wan Z, Fang Y M and Wang Z. 2023a. A quality-of-experience database for adaptive omnidirectional video streaming. *IEEE Journal of Selected Topics in Signal Processing*, 17(5): 949-963 [DOI: 10.1109/JSTSP.2023.3300529]
- Liu X L, Zhang H Y, Yan J B, Zhang H, Fang Y M and Wang S Q. 2024c. Quality of experience of viewport adaptive omnidirectional video streaming//*Proceedings of the IEEE International Conference on Image Processing*. Abu Dhabi: IEEE: 3667-3673 [DOI: 10.1109/ICIP51287.2024.10647716]
- Liu Y, Li S F, Duan H Y, Zhou Y, Fan D X and Zhai G T. 2025. Multi-task guided no-reference omnidirectional image quality assessment with feature interaction. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(9): 8794-8806 [DOI: 10.1109/TCSVT. 2025.3551723]
- Liu Y, Yin X H, Wan Z L, Yue G H and Zheng Z. 2023b. Toward a no-reference omnidirectional image quality evaluation by using multi-perceptual features. *ACM Transactions on Multimedia Computing, Communications and Applications*, 19(2): 1-19 [DOI: 10.1145/

- 3549544]
- Liu Y, Yin X H, Wang Y, Yin Z X and Zheng Z. 2023c. HVS-based perception-driven no-reference omnidirectional image quality assessment. *IEEE Transactions on Instrumentation and Measurement*, 72: 1-11 [DOI: 10.1109/TIM.2022.3232792]
- Liu Y, Yin X H, Yue G H, Zheng Z, Jiang J H, He Q G, et al. 2023d. Blind omnidirectional image quality assessment with representative features and viewport oriented statistical features. *Journal of Visual Communication and Image Representation*, 91: 103770 [DOI: 10.1016/j.jvcir.2023.103770]
- Liu Y, Yu H W, Huang B Q, Yue G H and Song B Y. 2021. Blind omnidirectional image quality assessment based on structure and natural features. *IEEE Transactions on Instrumentation and Measurement*, 70: 1-11 [DOI: 10.1109/TIM.2021.3102691]
- Lo W C, Fan C L, Lee J, Huang C Y, Chen K T and Hsu C H. 2017. 360 video viewing dataset in head-mounted virtual reality//*Proceedings of the 8th ACM on Multimedia Systems Conference*. Taipei, Taiwan, China: ACM: 211-216 [DOI: 10.1145/3083187.3083219]
- Lopes F, Ascenso J, Rodrigues A and Queluz M P. 2018. Subjective and objective quality assessment of omnidirectional video//*Proceedings of the Applications of Digital Image Processing XLI*. San Diego, California, USA: SPIE: 107520P [DOI: 10.1117/12.2321679]
- Lowe D G. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2): 91-110 [DOI: 10.1023/B:VISI.0000029664.99615.94]
- Luz G, Ascenso J, Brites C and Pereira F. 2017. Saliency-driven omnidirectional imaging adaptive coding: Modeling and assessment//*Proceedings of the IEEE International Workshop on Multimedia Signal Processing*. Luton, UK: IEEE: 1-6 [DOI: 10.1109/MMSP.2017.8122228]
- Madhusudana P C and Soundararajan R. 2019. Subjective and objective quality assessment of stitched images for virtual reality. *IEEE Transactions on Image Processing*, 28(11): 5620-5635 [DOI: 10.1109/TIP.2019.2921858]
- Mahmoudpour S and Schelkens P. 2021. Omnidirectional video quality index accounting for judder. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(1): 61-75 [DOI: 10.1109/TCSVT.2020.2971357]
- Meng Y and Ma Z. 2022. Viewport-based omnidirectional video quality assessment: database, modeling and inference. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(1): 120-134 [DOI: 10.1109/TCSVT.2021.3057368]
- Möller S and Raake A. 2014. *Quality of experience: advanced concepts, applications and methods*. Springer [DOI: 10.1007/978-3-319-02681-7]
- Moorthy A K, Choi L K, Bovik A C and Veciana G D. 2012. Video quality assessment on mobile devices: subjective, behavioral and objective studies. *IEEE Journal of Selected Topics in Signal Processing*, 6(6): 652-671 [DOI: 10.1109/JSTSP.2012.2212417]
- Netflix. 2016. Video multi-method assessment fusion (vmaf). <https://github.com/Netflix/vmaf>.
- Nguyen A and Yan Z S. 2019. A saliency dataset for 360-degree videos//*Proceedings of the 10th ACM Multimedia Systems Conference*. Amherst, Massachusetts: ACM: 279-284 [DOI: 10.1145/3304109.3325820]
- Ni B L, Peng H W, Chen M H, Zhang S Y, Meng G F, Fu J L, et al. 2022. Expanding language-image pretrained models for general video recognition//*Proceedings of the European Conference on Computer Vision*. Telaviv, Israel: Springer: 1-18 [DOI: 10.1007/978-3-031-19772-7_1]
- Orduna M, Díaz C, Muñoz L, Pérez P, Benito I and García N. 2020a. Video multimethod assessment fusion (vmaf) on 360vr contents. *IEEE Transactions on Consumer Electronics*, 66(1): 22-31 [DOI: 10.1109/TCE.2019.2957987]
- Orduna M, Pérez P, Díaz C and García N. 2020b. Evaluating the influence of the hmd, usability, and fatigue in 360vr video quality assessments//*Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*. Atlanta, GA, USA: IEEE: 682-683 [DOI: 10.1109/VRW50115.2020.00192]
- Poreddy A K R, Ganeswaram R B C, Appina B, Kokil P and Pachori R B. 2023. No-reference virtual reality image quality evaluator using global and local natural scene statistics. *IEEE Transactions on Instrumentation and Measurement*, 72: 1-16 [DOI: 10.1109/TIM.2023.3322995]
- Qi Y B, Jiang G Y, Yu M, Zhang Y and Ho Y S. 2021. Viewport perception based blind stereoscopic omnidirectional image quality assessment. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(10): 3926-3941 [DOI: 10.1109/TCSVT.2020.3043349]
- Rai Y, Gutiérrez J and Callet P L. 2017. A dataset of head and eye movements for 360 degree images//*Proceedings of the 8th ACM on Multimedia Systems Conference*. Taipei, Taiwan, China: ACM: 205-210. [DOI: 10.1145/3083187.3083218]
- Ruan J J and Xie D L. 2021. A survey on qoe-oriented vr video streaming: some research issues and challenges. *Electronics*, 10(17) [DOI: 10.3390/electronics10172155]
- Rublee E, Rabaud V, Konolige K and Bradski G. 2011. ORB: an efficient alternative to sift or surf//*Proceedings of the International Conference on Computer Vision*. Barcelona, Spain: IEEE: 2564-2571 [DOI: 10.1109/ICCV.2011.6126544]
- Schatz R, Sackl A, Timmerer C and Gardlo B. 2017. Towards subjective quality of experience assessment for omnidirectional video streaming//*Proceedings of the Ninth International Conference on Quality of Multimedia Experience*. Erfurt, Germany: IEEE: 1-6 [DOI: 10.1109/QoMEX.2017.7965657]
- Schatz R, Zabrovskiy A and Timmerer C. 2019. Tile-based streaming of

- 8k omnidirectional video: subjective and objective qoe evaluation//
Proceedings of the Eleventh International Conference on Quality of
Multimedia Experience. Berlin, Germany: IEEE: 1-6 [DOI: 10.
1109/QoMEX.2019.8743230]
- Sendjasni A and Larabi M C. 2022a. SAL-360IQA: a saliency weighted
patch-based cnn model for 360-degree images quality assessment//
Proceedings of the IEEE International Conference on Multimedia
and Expo Workshops. Taipei, Taiwan, China: IEEE: 1-6 [DOI:
10.1109/ICMEW56448.2022.9859468]
- Sendjasni A and Larabi M C. 2023a. Attention-aware patch-based cnn
for blind 360-degree image quality assessment. *Sensors*, 23 (21) :
8676 [DOI: 8676.10.3390/s23218676]
- Sendjasni A and Larabi M C. 2023b. PW-360IQA: perceptually-
weighted multichannel cnn for blind 360-degree image quality
assessment. *Sensors*, 23 (9) [DOI: 10.3390/s23094242]
- Sendjasni A, Larabi M C and Cheikh F A. 2021a. Convolutional neural
networks for omnidirectional image quality assessment: pre-trained
or re-trained? //Proceedings of the IEEE International Conference
on Image Processing. Anchorage, AK, USA: IEEE: 3413-3417
[DOI: 10.1109/ICIP42928.2021.9506192]
- Sendjasni A, Larabi M C and Cheikh F A. 2021b. Perceptually-weighted
cnn for 360-degree image quality assessment using visual scan-path
and jnd//Proceedings of the IEEE International Conference on
Image Processing. Anchorage, AK, USA: IEEE: 1439-1443
[DOI: 10.1109/ICIP42928.2021.9506044]
- Sendjasni A, Larabi M C and Cheikh F A. 2022b. Convolutional neural
networks for omnidirectional image quality assessment: a bench-
mark. *IEEE Transactions on Circuits and Systems for Video Tech-
nology*, 32 (11) : 7301-7316 [DOI: 10.1109/TCSVT. 2022.
3181235]
- Simka M, Polak L, Novotny M, Kufa J and Fliegel K. 2024. Perfor-
mance evaluation of objective quality assessment methods for omni-
directional images under emerging compressions. *IEEE Access*,
12: 150419-150429 [DOI: 10.1109/ACCESS.2024.3478793]
- Singla A, Fremerey S, Robitza W, Lebreton P and Raake A. 2017a.
Comparison of subjective quality evaluation for hevc encoded omni-
directional videos at different bit-rates for uhd and fhd resolution//
Proceedings of the on Thematic Workshops of ACM Multimedia.
Mountain View, California, USA: ACM: 511-519 [DOI: 10.1145/
3126686.3126768]
- Singla A, Fremerey S, Robitza W and Raake A. 2017b. Measuring and
comparing qoe and simulator sickness of omnidirectional videos in
different head mounted displays//Proceedings of the Ninth Interna-
tional Conference on Quality of Multimedia Experience. Erfurt,
Germany: IEEE: 1-6 [DOI: 10.1109/QoMEX.2017.7965658]
- Singla A, Göring S, Raake A, Meixner B, Koenen R and Buchholz T.
2019a. Subjective quality evaluation of tile-based streaming for
omnidirectional videos//Proceedings of the 10th ACM Multimedia
Systems Conference. Amherst, Massachusetts: ACM: 232-242
[DOI: 10.1145/3304109.3306218]
- Singla A, Rao R R R, Göring S and Raake A. 2019b. Assessing media
qoe, simulator sickness and presence for omnidirectional videos
with different test protocols//Proceedings of the IEEE Conference
on Virtual Reality and 3D User Interfaces. Osaka, Japan: IEEE:
1163-1164 [DOI: 10.1109/VR.2019.8798291]
- Singla A, Robitza W and Raake A. 2019c. Comparison of subjective
quality test methods for omnidirectional video quality evaluation//
Proceedings of the IEEE International Workshop on Multimedia Sig-
nal Processing. Kuala Lumpur, Malaysia: IEEE: 1-6. [DOI: 10.
1109/MMSP.2019.8901719]
- Sitzmann V, Serrano A, Pavel A, Agrawala M, Gutierrez D, Masia B,
et al. 2018. Saliency in vr: how do people explore virtual environ-
ments? *IEEE Transactions on Visualization and Computer Graph-
ics*, 24 (4) : 1633-1642 [DOI: 10.1109/TVCG.2018.2793599]
- Souza V, Maciel A, Nedel L and Kopper R. 2021. Measuring presence
in virtual environments: a survey. *ACM Computing Surveys*, 54
(8) : 1-37 [DOI: 10.1145/3466817]
- Sui X J, Fang Y M, Zhu H W, Wang S Q and Wang Z. 2023. Scan-
DMM: a deep markov model of scanpath prediction for 360
images//Proceedings of the IEEE/CVF Conference on Computer
Vision and Pattern Recognition. 6989-6999 [DOI: 10.1109/
CVPR52729.2023.00675]
- Sui X J, Ma K D, Yao Y R and Fang Y M. 2022. Perceptual quality
assessment of omnidirectional images as moving camera videos.
IEEE Transactions on Visualization and Computer Graphics, 28
(8) : 3022-3034 [DOI: 10.1109/TVCG.2021.3050888]
- Sui X J, Wang S Q and Fang Y M. 2024. A survey on objective quality
assessment of omnidirectional images//Proceedings of the Asia
Pacific Signal and Information Processing Association Annual Sum-
mit and Conference. Macau, Macao. 1-6 [DOI: 10.1109/
APSIPAASC63619.2025.10849147]
- Sui X J, Zhu H W, Liu X L, Fang Y, Wang S and Wang Z. 2025. Per-
ceptual quality assessment of 360° images based on generative
scanpath representation. *IEEE Transactions on Image Processing*,
34: 4485-4499 [DOI: 10.1109/TIP.2025.3583181]
- Sun W, Gu K, Ma S W, Zhu W H, Liu N and Zhai G T. 2018. A large-
scale compressed 360-degree spherical image database: from sub-
jective quality evaluation to objective model comparison//Proceed-
ings of the IEEE 20th International Workshop on Multimedia Signal
Processing. Vancouver, BC, Canada. 1-6 [DOI: 10.1109/MMSP.
2018.8547102]
- Sun W, Gu K, Zhai G T, Ma S W, Lin W S and Calle P L. 2017a.
CVIQD: Subjective quality evaluation of compressed virtual reality
images//Proceedings of the IEEE International Conference on Image
Processing. Beijing, China: IEEE: 3450-3454 [DOI: 10.1109/
ICIP.2017.8296923]
- Sun W, Min X K, Zhai G T, Gu K, Duan H Y and Ma S W. 2020.
MC360IQA: a multi-channel cnn for blind 360-degree image qual-

- ity assessment. *IEEE Journal of Selected Topics in Signal Processing*, 14(1): 64-77 [DOI: 10.1109/JSTSP.2019.2955024] (Sun Y L, Lu A and Yu L. 2017b. Weighted-to-spherically-uniform quality evaluation for omnidirectional video. *IEEE Signal Processing Letters*, 24(9): 1408-1412 [DOI: 10.1109/LSP.2017.2720693] (Tian C Z, Chai X L, Chen G, Shao S, Jiang Q P, Meng X C, et al. 2022. VSOIQE: a novel viewport-based stitched 360 omnidirectional image quality evaluator. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(10): 6557-6572 [DOI: 10.1109/TCSVT.2022.3172135] (Tian C Z, Shao F, Chai X L, Jiang Q P, Xu L and Ho Y S. 2023. Viewport-sphere-branch network for blind quality assessment of stitched 360 omnidirectional images. *IEEE Transactions on Circuits and Systems for Video Technology*, 33(6): 2546-2560 [DOI: 10.1109/TCSVT.2022.3225172] (Tofighi N J, Elfkir M H, Imamoglu N, Ozcinar C, Erdem E and Erdem A. 2023. ST360IQ: no-reference omnidirectional image quality assessment with spherical vision transformers//*Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*. Rhodes Island, Greece: IEEE: 1-5 [DOI: 10.1109/ICASSP49357.2023.10096750] (Tran H T T, Ngoc N P, Bui C M, Pham M H and Thang T C. 2017a. An evaluation of quality metrics for 360 videos//*Proceedings of the Ninth International Conference on Ubiquitous and Future Networks*. Milan, Italy: IEEE: 7-11 [DOI: 10.1109/ICUFN.2017.7993736] (Tran H T T, Ngoc N P, Pham C T, Jung Y J and Thang T C. 2017b. A subjective study on qoe of 360 video for vr communication//*Proceedings of the IEEE International Workshop on Multimedia Signal Processing*. Luton, UK: IEEE: 1-6 [DOI: 10.1109/MMSP.2017.8122249] (Tran H T T, Ngoc N P and Thang T C. 2021. Towards an overall qoe model for 360-degree video//*Proceedings of the IEEE International Conference on Communications and Electronics*. Phu Quoc Island, Vietnam: IEEE: 327-331 [DOI: 10.1109/ICCE48956.2021.9352076] (Truong T Q, Tran H T T and Thang T C. 2019. Non-reference quality assessment model using deep learning for omnidirectional images//*Proceedings of the IEEE International Conference on Awareness Science and Technology*. Morioka, Japan: IEEE: 1-5 [DOI: 10.1109/ICAwST.2019.8923442] (Ullah H, Afzal S and Khan I M. 2022. Perceptual quality assessment of panoramic stitched contents for immersive applications: a prospective survey. *Virtual Reality & Intelligent Hardware*, 4(3): 223-246 [DOI: 10.1016/j.vrih.2022.03.004] (Upenik E and Ebrahimi T. 2019. Saliency driven perceptual quality metric for omnidirectional visual content//*Proceedings of the IEEE International Conference on Image Processing*. Taipei, Taiwan, China: IEEE: 4335-4339 [DOI: 10.1109/QoMEX.2017.7965660] (Upenik E, Rerabek M and Ebrahimi T. 2016. Testbed for subjective evaluation of omnidirectional visual content//*Proceedings of the Picture Coding Symposium*. Nuremberg, Germany: IEEE: 1-5 [DOI: 10.1109/PCS.2016.7906378] (Upenik E, Rerabek M and Ebrahimi T. 2017. On the performance of objective metrics for omnidirectional visual content//*Proceedings of the International Conference on Quality of Multimedia Experience*. Erfurt, Germany: IEEE: 1-6 [DOI: 10.1109/QoMEX.2017.7965660] (Venkatakrisnan R, Venkatakrisnan R, Raveendranath B, Sarno D M, Robb A C, Lin W C, et al. 2024. The effects of auditory, visual, and cognitive distractions on cybersickness in virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 30(8): 5350-5369 [DOI: 10.1109/TVCG.2023.3293405] (Wan Z L, Yan X, Li Z Y, Fan X P, Zuo W M and Zhao D B. 2025. No-reference stereoscopic omnidirectional image quality assessment via a binocular viewport hypergraph convolutional network. *IEEE Transactions on Circuits and Systems for Video Technology*, 35(7): 7196-7209 [DOI: 10.1109/TCSVT.2025.3538511] (Wang P F, Duan H Y, Xie Z Y, Min X K and Zhai G T. 2024. Subjective and objective quality assessment for augmented reality images. *IEEE Open Journal of Immersive Displays*, 1: 135-145 [DOI: 10.1109/OJID.2024.3441511] (Wang X L, Girshick R, Gupta A and He K M. 2018. Non-local neural networks//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT, USA: IEEE: 7794-7803 [DOI: 10.1109/CVPR.2018.00813] (Wang Y, Zhou Y, Li M Y, Li L D, Wu J J, Yue G H, et al. 2025. Panoramic image quality assessment based on scanpath prediction and bidimensional feature representation. *IEEE Transactions on Instrumentation and Measurement*, 74: 1-11 [DOI: 10.1109/TIM.2025.3571161] (Wang Y H, Li H and Jiang Q P. 2022. Dynamically attentive viewport sequence for no-reference quality assessment of omnidirectional images. *Frontiers in Neuroscience*, 16: 1022041. [DOI: 10.3389/fnins.2022.1022041] (Wang Y J, L Y Y, Liu J M, Zhang F L, Zheng X H and Doodgson N A. 2026. RL-ScanIQ: Reinforcement-learned scanpaths for blind 360deg image quality assessment//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Denver, USA: IEEE: 37401-37412 (Wang Z, Bovik A C, Sheikh H R and Simoncelli E P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600-612 [DOI: 10.1109/TIP.2003.819861] (Wang Z, Simoncelli E P and Bovik A C. 2003. Multiscale structural similarity for image quality assessment//*Proceedings of the Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, Pacific Grove, CA, USA: IEEE: 1398-1402 [DOI: 10.1109/ACSSC.2003.1292216] (Wen W, Li M, Yao Y R, Sui X J, Zhang Y B, Lan L, et al. 2024. Perceptual quality assessment of virtual reality videos in the wild. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(9): 8368-8381 [DOI: 10.1109/TCSVT.2024.3378352] (Wu C L, Tan Z H, Wang Z and Yang S Q. 2017. A dataset for exploring user behaviors in vr spherical video streaming//*Proceedings of the 8th ACM on Multimedia Systems Conference*. Taipei, Taiwan, China: ACM: 193-198 [DOI: 10.1145/

- 3083187.3083210] (Wu H N, Zhang Z C, Zhang W X, Chen C F, Liao L, Li C Y, et al. 2024. Q-ALIGN: teaching lms for visual scoring via discrete text-defined levels//Proceedings of the 41st International Conference on Machine Learning. Vienna, Austria: JMLR: 54015-54029 [DOI: 10.5555/3692070.3694286] (Wu T H, Shi S W, Cai H M, Cao M D, Xiao J, Zheng Y Q, et al. 2023. Assessor360: multi-sequence network for blind omnidirectional image quality assessment//Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, LA, USA: Curran Associates Inc: 64957-64970 [DOI: 10.5555/3666122.3668956] (Xia Y M, Wang Y F and Peng Y. 2019. Blind panoramic image quality assessment via the asymmetric mechanism of human brain//Proceedings of the IEEE Visual Communications and Image Processing. Sydney, NSW, Australia: IEEE: 1-4 [DOI: 10.1109/VCIP47243.2019.8965887] (Xiao J X, Ehinger K A, Oliva A and Torralba A. 2012. Recognizing scene viewpoint using panoramic place representation//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA: IEEE: 2695-2702 [DOI: 10.1109/CVPR.2012.6247991] (Xie S W, Xu Y L, Shen Q, Ma Z and Zhang W J. 2020. Modeling the perceptual quality of viewport adaptive omnidirectional video streaming. IEEE Transactions on Circuits and Systems for Video Technology, 30(9): 3029-3042 [DOI: 10.1109/TCSVT.2019.2934136] (Xiu X Y, He Y W, Ye Y and Vishwanath B. 2017. An evaluation framework for 360-degree video compression//Proceedings of the IEEE Visual Communications and Image Processing. St. Petersburg, FL, USA: IEEE: 1-4 [DOI: 10.1109/VCIP.2017.8305084] (Xu J H, Lin C Y, Zhou W and Chen Z B. 2018a. Subjective quality assessment of stereoscopic omnidirectional image//Proceedings of the Pacific-Rim Conference on Multimedia. Hefei, China: Springer: 589-599 [DOI: 10.1007/978-3-030-00776-8_54] (Xu J H, Luo Z Y, Zhou W, Zhang W Y and Chen Z B. 2019a. Quality assessment of stereoscopic 360-degree images from multi-viewports//Proceedings of the Picture Coding Symposium. Ningbo, China: IEEE: 1-5 [DOI: 10.1109/PCS48520.2019.8954555] (Xu J H, Zhou W and Chen Z B. 2021. Blind omnidirectional image quality assessment with viewport oriented graph convolutional networks. IEEE Transactions on Circuits and Systems for Video Technology, 31(5): 1724-1737 [DOI: 10.1109/TCSVT.2020.3015186] (Xu M, Jiang L, Li C, Wang Z L and Tao X M. 2022. Viewport-based cnn: a multi-task approach for assessing 360 video quality. IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(4): 2198-2215 [DOI: 10.1109/TPAMI.2020.3028509] (Xu M, Li C, Liu Y F, Deng X and Lu J X. 2017. A subjective visual quality assessment method of panoramic videos//Proceedings of the IEEE International Conference on Multimedia and Expo. Hong Kong, China: IEEE: 517-522 [DOI: 10.1109/ICME.2017.8019351] (Xu M, Li C, Chen Z Z, Wang Z L and Guan Z Y. 2019b. Assessing visual quality of omnidirectional videos. IEEE Transactions on Circuits and Systems for Video Technology, 29(12): 3516-3530 [DOI: 10.1109/TCSVT.2018.2886277] (Xu M, Song Y H, Wang J Y, Qiao M L, Huo L Y, Wang Z L. 2019c. Predicting head movement in panoramic video: a deep reinforcement learning approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(11): 2693-2708 [DOI: 10.1109/TPAMI.2018.2858783] (Xu M, Li C, Zhang S and Callet P L. 2020. State-of-the-art in 360 video/image processing: perception, assessment and compression. IEEE Journal of Selected Topics in Signal Processing, 14(1): 5-26 [DOI: 10.1109/JSTSP.2020.2966864] (Xu Y Y, Dong Y B, Wu J R, Sun Z Z, Shi Z R, Yu J Y, et al. 2018b. Gaze prediction in dynamic 360 immersive videos//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE: 5333-5342. [DOI: 10.1109/CVPR.2018.00559]
- Yan J B, Fang Y M and Liu X L. 2022. The review of distortion-related image quality assessment. Journal of Image and Graphics, 27(5): 143-1466 (鄢杰斌, 方玉明, 刘学林. 2022. 图像质量评价研究综述——从失真的角度. 中国图象图形学报, 27(5): 1430-1466) [DOI: 10.11834/jig.210790]
- Yan J B, Rao J L, Chen J J, Tan Z W, Liu W D and Fang Y M. 2025a. Multitask auxiliary network for perceptual quality assessment of non-uniformly distorted omnidirectional images. IEEE Transactions on Circuits and Systems for Video Technology, 35(3): 2782-2793 [DOI: 10.1109/TCSVT.2024.3497994]
- Yan J B, Rao J L, Liu X L, Fang Y M, Zuo Y F and Liu W D. 2025b. Subjective and objective quality assessment of non-uniformly distorted omnidirectional images. IEEE Transactions on Multimedia, 1-13 [DOI: 10.1109/TMM.2025.3535372]
- Yan J B, Tan Z W, Fang Y M, Chen J J, Jiang W H and Wang Z. 2025c. Omnidirectional image quality captioning: a large-scale database and a new model. IEEE Transactions on Image Processing, 34: 1326-1339 [DOI: 10.1109/TIP.2025.3539468]
- Yan J B, Tan Z W, Fang Y M, Rao J L and Zuo Y F. 2025d. Max360IQ: blind omnidirectional image quality assessment with multi-axis attention. Pattern Recognition, 162: 111429 [DOI: 10.1016/j.patcog.2025.111429]
- Yan J B, Tan Z W, Rao J L, Wu L, Zuo Y F and Fang Y M. 2025e. Computational analysis of degradation modeling in blind panoramic image quality assessment. ACM Transactions on Multimedia Computing, Communications, and Applications, 1-23 [DOI: 10.1145/3720547]
- Yan J B, Wu K C, Chen J J, Tan Z W, Fang Y M and Liu W D. 2025f. Viewport-unaware blind omnidirectional image quality assessment: a flexible and effective paradigm. ACM Transactions on Multimedia Computing, Communications, and Applications, 1-19 [DOI: 10.1145/3723165]
- Yan X Y, Yang F X, Yang C, An P, Huang X P and Shen L Q. 2025g. No-reference 360 video quality measurement with local and

- global features fusion. *IEEE Transactions on Instrumentation and Measurement*, 74: 1-12 [DOI: 10.1109/TIM.2025.3548176]
- Yan J B, Liu Z Y, Chen J J, Xu X Y, Chen P F and Fang Y M. 2026a. Viewport-unaware full-reference omnidirectional image quality assessment with inter-patch and sequence similarity. *IEEE Transactions on Circuits and Systems for Video Technology*, 36(5): 6732-6743 [DOI: 10.1109/TCSVT.2026.3651643]
- Yan J B, Zhang J Y, Chen J J, Chen P F, Liu X L, Tan Z W, et al. 2026b. Blind omnidirectional image quality assessment: Embracing the magic power of multimodal large language models. *International Journal of Computer Vision*, 134 (78) [DOI: 10.1007/s11263-025-02626-w]
- Yang J C, Liu T L, Jiang B, Lu W and Meng Q G. 2021a. Panoramic video quality assessment based on non-local spherical cnn. *IEEE Transactions on Multimedia*, 23: 797-809 [DOI: 10.1109/TMM.2020.2990075]
- Yang J C, Liu T L, Jiang B, Song H and Lu W. 2018a. 3D panoramic virtual reality video quality assessment based on 3d convolutional neural networks. *IEEE Access*, 6: 38669-38682 [DOI: 10.1109/ACCESS.2018.2854922]
- Yang L, Duan H Y, Teng L, Zhu Y C, Liu L X, Hu M H, et al. 2024. AIGCOIQA2024: Perceptual quality assessment of AI generated omnidirectional images// *Proceedings of the IEEE International Conference on Image Processing*. Abu Dhabi, United Arab Emirates: IEEE: 1239-1245 [DOI: 10.1109/ICIP51287.2024.10647885]
- Yang L, Duan H Y, Wang J R, Liu J, Hu M H, Min X K, et al. 2026. Quality assessment and distortion-aware saliency prediction for AI-generated omnidirectional images. *IEEE Transactions on Circuits and Systems for Video Technology*, 36(3): 3351-3366 [DOI: 10.1109/TCSVT.2025.3616234]
- Yang L, Xu M, Deng X and Feng B. 2021b. Spatial attention-based non-reference perceptual quality prediction network for omnidirectional images//*Proceedings of the IEEE International Conference on Multimedia and Expo*. Shenzhen, China: IEEE: 1-6 [DOI: 10.1109/ICME51207.2021.9428390]
- Yang L, Xu M, Li S X, Guo Y C and Wang Z L. 2023. Blind vqa on 360 video via progressively learning from pixels, frames, and video. *IEEE Transactions on Image Processing*, 32: 128-143 [DOI: 10.1109/TIP.2022.3226417]
- Yang L, Xu M, Liu L, Huo L Y and Gao X B. 2022. TVFormer: trajectory-guided visual quality assessment on 360 images with transformers//*Proceedings of the 30th ACM International Conference on Multimedia*. Lisboa, Portugal: ACM: 799-808 [DOI: 10.1145/3503161.3547748]
- Yang M Z, Wang S G, Calheiros R N and Yang F C. 2018b. Survey on qoe assessment approach for network service. *IEEE Access*, 6: 48374-48390 [DOI: 10.1109/ACCESS.2018.2867253]
- Yang S, Zhao J Z, Jiang T T, Wang J, Rahim T, Zhang B, Xu Z J and Fei Z S. 2017. An objective assessment method based on multi-level factors for panoramic videos//*Proceedings of the IEEE Visual Communications and Image Processing*. St. Petersburg, FL, USA: IEEE: 1-4 [DOI: 10.1109/VCIP.2017.8305133]
- Yang W Y, Zhao Z H and Gao S. 2025. From touch to immersion: A systematic review of haptic and multimodal perception in the metaverse. *IEEE Open Journal on Immersive Displays*, 2: 122-138 [DOI: 10.1109/OJID.2025.3622139]
- Yang Y Y, Jiang G Y, Yu M and Qi Y B. 2020. Latitude and binocular perception based blind stereoscopic omnidirectional image quality assessment for vr system. *Signal Processing*, 173: 107586 [DOI: 10.1016/j.sigpro.2020.107586]
- Yao S H, Fan C L and Hsu C H. 2019. Towards quality-of-experience models for watching 360 videos in head-mounted virtual reality// *Proceedings of the Eleventh International Conference on Quality of Multimedia Experience*. Berlin, Germany: IEEE: 1-3 [DOI: 10.1109/QoMEX.2019.8743198]
- Yaqoob A, Bi T and Muntean G M. 2020. A survey on adaptive 360 video streaming: solutions, challenges and opportunities. *IEEE Communications Surveys & Tutorials*, 22(4): 2801-2838 [DOI: 10.1109/COMST.2020.3006999]
- Yu M, Lakshman H and Girod B. 2015. A framework to evaluate omnidirectional video coding schemes//*Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. Fukuoka, Japan: IEEE: 31-36 [DOI: 10.1109/ISMAR.2015.12]
- Zakharchenko V, Choi K P and Park J H. 2016. Quality metric for spherical panoramic video//*Proceedings of the Optics and Photonics for Information Processing X*. San Diego, California, USA: SPIE: 99700C [DOI: 10.1117/12.2235885]
- Zeng H Q, Ding R, Huang H L, Chen J and Zhu J Q. 2022. Immersive video quality assessment: an overview. *Journal of Signal Processing*, 38(9): 1853-1862 (曾焕强, 丁瑞, 黄海靓, 陈婧, 朱建清. 2022. 沉浸式视频质量评价综述. *信号处理*, 38(9): 1853-1862) [DOI: 10.16798/j.issn.1003-0530.2022.09.008]
- Zepernick H J, Elwardy M, Hu Y and Sundstedt V. 2019. On the number of participants needed for subjective quality assessment of 360 videos//*Proceedings of the 13th International Conference on Signal Processing and Communication Systems*. Gold Coast, QLD, Australia: IEEE: 1-10 [DOI: 10.1109/ICSPCS47537.2019.9008748]
- Zhang B, Yan Z Y, Wang J, Luo Y Y, Yang S and Fei Z S. 2018a. An audio-visual quality assessment methodology in virtual reality environment//*Proceedings of the IEEE International Conference on Multimedia & Expo Workshops*. San Diego, CA, USA: IEEE: 1-6. [DOI: 10.1109/ICMEW.2018.8551522]
- Zhang B, Zhao J Z, Yang S, Zhang Y, Wang J and Fei Z S. 2017. Subjective and objective quality assessment of panoramic videos in virtual reality environments//*Proceedings of the IEEE International Conference on Multimedia & Expo Workshops*. Hong Kong, China:

- IEEE: 163-168 [DOI: 10.1109/ICMEW.2017.8026226]
- Zhang C F and Liu S G. 2022. No-reference omnidirectional image quality assessment based on joint network//Proceedings of the 30th ACM International Conference on Multimedia. Lisboa, Portugal: ACM: 943-951 [DOI: 10.1145/3503161.3548175]
- Zhang W, Zou W J, Yang F Z, L  v  que L and Liu H T. 2019. The effect of spatio-temporal inconsistency on the subjective quality evaluation of omnidirectional videos//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Brighton, UK: IEEE: 4055-4059 [DOI: 10.1109/ICASSP.2019.8682221]
- Zhang Y, Zhang L, Hamidouche W and Deforges O. 2020. A fixation-based 360 benchmark dataset for salient object detection//Proceedings of the IEEE International Conference on Image Processing. Abu Dhabi, United Arab Emirates: IEEE: 3458-3462 [DOI: 10.1109/ICIP40778.2020.9191158]
- Zhang Y X, Liu Z Z, Chen Z Z, Xu X Z and Liu S. 2021. No-reference quality assessment of panoramic video based on spherical-domain features//Proceedings of the Picture Coding Symposium. Bristol, United Kingdom: IEEE: 1-5 [DOI: 10.1109/PCS50896.2021.9477498]
- Zhang Y X, Wang Y B, Liu F Y, Liu Z Z, Li Y M, Yang D Q, et al. 2018b. Subjective panoramic video quality assessment database for coding applications. IEEE Transactions on Broadcasting, 64(2): 461-473 [DOI: 10.1109/TBC.2018.2811627]
- Zhang Z C, Wang J Y, Wen F R, Guo Y J, Zhao X Y, Fang X Y, et al. 2025a. Large multimodal models evaluation: a survey. Science China Information Sciences, 68: 221301 [DOI: 10.1007/s11432-025-4676-4]
- Zhang Z C, Zhou Y J, Li C Y, Zhao B X, Liu X H and Zhang G T. 2025b. Quality assessment in the era of large models: A survey. ACM Transactions on Multimedia Computing, Communications and Applications, 21(7): 1-13 [DOI: 10.1145/3722559]
- Zhang Z C, Wu H N, Zhang E, Zhai G T and Lin W S. 2024. Q-Bench++: a benchmark for multi-modal foundation models on low-level vision from single images to pairs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(12): 10404-10418 [DOI: 10.1109/TPAMI.2024.3445770]
- Zhao J Z, Zhang B, Yan Z Y, Wang J and Fei Z S. 2017. A study on the factors affecting audio-video subjective experience in virtual reality environments//Proceedings of the International Conference on Virtual Reality and Visualization. Zhengzhou, China: IEEE: 303-306 [DOI: 10.1109/ICVRV.2017.00070]
- Zheleva A, Durnez W, Bombeke K, Wallendael G V and Marez L D. 2020. Seeing is believing: the effect of video quality on quality of experience in virtual reality//Proceedings of the Twelfth International Conference on Quality of Multimedia Experience. Athlone, Ireland: IEEE: 1-4 [DOI: 10.1109/QoMEX48832.2020.9123075]
- Zheng X L, Jiang G Y, Yu M and Jiang H. 2020. Segmented spherical projection-based blind omnidirectional image quality assessment. IEEE Access, 8: 31647-31659 [DOI: 10.1109/ACCESS.2020.2972158]
- Zhou M L, Chen L, Wei X K, Liao X R, Mao Q, Wang H Q, et al. 2023a. Perception oriented u-shaped transformer network for 360-degree no-reference image quality assessment. IEEE Transactions on Broadcasting 69, 2: 396-405 [DOI: 10.1109/TBC.2022.3231101]
- Zhou W and Wang Z. 2023b. Blind omnidirectional image quality assessment: integrating local statistics and global semantics//Proceedings of the IEEE International Conference on Image Processing. Kuala Lumpur, Malaysia: IEEE: 1405-1409 [DOI: 10.1109/ICIP49359.2023.10222049]
- Zhou W, Xu J H, Jiang Q P and Chen Z B. 2022a. No-reference quality assessment for 360-degree images by analysis of multifrequency information and local-global naturalness. IEEE Transactions on Circuits and Systems for Video Technology, 32(4): 1778-1791 [DOI: 10.1109/TCSVT.2021.3081182]
- Zhou Y, Gong W K, Sun Y J, Li L D, Wu J J and Gao X B. 2023c. Pyramid feature aggregation for hierarchical quality prediction of stitched panoramic images. IEEE Transactions on Multimedia, 25: 4177-4186 [DOI: 10.1109/TMM.2022.3171684]
- Zhou Y, Sun Y J, Li L D, Gu K and Fang Y M. 2022b. Omnidirectional image quality assessment by distortion discrimination assisted multi-stream network. IEEE Transactions on Circuits and Systems for Video Technology, 32(4): 1767-1777 [DOI: 10.1109/TCSVT.2021.3081162]
- Zhou Y, Wang Y, Li L D, Gao C Q and Lu Z L. 2022c. Research progress in objective quality evaluation of virtual reality images. Journal of Image and Graphics, 27(8): 2313-2328 (周玉, 汪一, 李雷达, 高陈强, 卢兆林. 2022. 虚拟现实图像客观质量评价研究进展. 中国图象图形学报, 27(8): 2313-2328) [DOI: 10.11834/jig.210949]
- Zhou Y F, Yu M, Ma H L, Shao H and Jiang G Y. 2018. Weighted-to-spherically-uniform ssim objective quality evaluation for panoramic video//Proceedings of the IEEE International Conference on Signal Processing. Beijing, China: IEEE: 54-57 [DOI: 10.1109/ICSP.2018.8652269]
- Zhu X L, Duan H Y, Cao Y Q, Zhu Y X, Zhu Y C, Liu J, et al. 2024. Perceptual quality assessment of omnidirectional audio-visual signals//Proceedings of the CAAI International Conference on Artificial Intelligence. Singapore: Springer: 512-525 [DOI: 10.1007/978-981-99-9119-8_46]
- Zhu X L, Duan H Y, Cao Y Q, Zhu Y C, Zhu Y X, Liu J, et al. 2025a. Subjective and objective audio-visual quality assessment for omnidirectional videos. IEEE Transactions on Image Processing, 34: 6506-6523 [DOI: 10.1109/TIP.2025.3613957]
- Zhu Y X, Duan H Y, Zhang K W, Zhu Y C, Zhu X L, Teng L, et al. 2025b. How does audio influence visual attention in omnidirectional

tional videos? database and model. IEEE Transactions on Image Processing, [DOI: 10.1109/TIP.2025.3567842]

Zou W J, Yang F Z, Zhang W, Li Y and Yu H P. 2018. A framework for assessing spatial presence of omnidirectional video on virtual reality device. IEEE Access, 6: 44676-44684 [DOI: 10.1109/ACCESS.2018.2864872]

Zou W J, Yang L H, Yang F Z, Ma Z B and Zhao Q Y. 2020. The impact of screen resolution of hmd on perceptual quality of immersive videos//Proceedings of the IEEE International Conference on Multimedia & Expo Workshops. London, UK: IEEE: 1-6 [DOI: 10.1109/ICMEW46912.2020.9105945]

Zou W J, Zhang W and Yang F Z. 2021. Modeling the perceptual quality for viewport-adaptive omnidirectional video streaming considering dynamic quality boundary artifact. IEEE Transactions on Circuits and Systems for Video Technology, 31(11): 4241-4254 [DOI: 10.1109/TCSVT.2021.3050157]

作者简介

刘学林,男,讲师,主要研究方向为视觉质量评价。E-mail: xuelinliu-bill@foxmail.com

方玉明,通信作者,男,教授,主要研究方向为人工智能、计算机视觉、信号处理。E-mail: leo.fangyuming@foxmail.com

鄢杰斌,男,副教授,主要研究方向为视觉质量评价。E-mail: jiebinan@foxmail.com

侯敬文,男,副教授,主要研究方向为视觉质量评价。E-mail: jingwen003@e.ntu.edu.sg

诸汉炜,男,教授,主要研究方向为视觉质量评价。E-mail: hanwei.zhu@ntu.edu.sg

睦相杰,男,讲师,主要研究方向为视觉质量评价。E-mail: xjsui@cityu.edu.mo