

# 全景视频 QoE 评价研究进展

杜丽娜<sup>1,2</sup> 卓力<sup>1,2</sup> 李嘉锋<sup>1,2</sup>

(1. 北京工业大学计算智能与智能系统北京市重点实验室, 北京 100124;  
2. 北京工业大学信息学部微电子学院, 北京 100124)

**摘要:** 随着 5G 移动通信技术、高性能计算、传感技术的不断进步,全景视频受到了越来越多的关注。全景视频通过头戴显示设备,可以为用户提供远超出平面视频的逼真的立体视觉感知,具有良好的发展前景。为用户提供良好的体验质量(Quality of Experience, QoE)是视频服务提供商吸引和留住用户,在激烈的市场竞争中取得成功的关键。与平面视频相比,全景视频的数据量倍增,对视频数据的采集、编码、传输和存储均提出了更高的要求。因此,如何在网络传输带宽、存储资源有限的情况下,保证用户的 QoE 就成为工业界和学术界共同关注的研究热点问题。本文对全景视频 QoE 评价进行了综述,首先对全景视频在采集、拼接、投影、编码、传输、解码、反投影、渲染等各个环节可能存在的失真进行分析,总结归纳了用户 QoE 的各种影响因素,如人的因素、系统因素、情境上下文和视频内容特性等;在此基础上,从影响因素和建模方法等多个方面归纳了全景视频 QoE 评价模型的研究进展,及其在码率自适应、资源优化分配和码率控制等方面的应用情况;最后,介绍了具有代表性的全景视频 QoE 评价数据集以及常见的 QoE 模型性能评价准则,并探讨了 QoE 评价模型目前存在的问题和未来的研究方向。

**关键词:** 全景视频; 用户质量体验; 主客观影响因素

**中图分类号:** TP391 **文献标识码:** A **DOI:** 10.16798/j.issn.1003-0530.2022.09.006

**引用格式:** 杜丽娜,卓力,李嘉锋. 全景视频 QoE 评价研究进展[J]. 信号处理, 2022, 38(9): 1831-1842. DOI: 10.16798/j.issn.1003-0530.2022.09.006.

**Reference format:** DU Lina, ZHUO Li, LI Jiafeng. Survey on QoE evaluation of panoramic video[J]. Journal of Signal Processing, 2022, 38(9): 1831-1842. DOI: 10.16798/j.issn.1003-0530.2022.09.006.

## Survey on QoE Evaluation of Panoramic Video

DU Lina<sup>1,2</sup> ZHUO Li<sup>1,2</sup> LI Jiafeng<sup>1,2</sup>

(1. Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing University of Technology, Beijing 100124, China; 2. Information Department, Beijing University of Technology, Beijing 100124, China)

**Abstract:** With the continuous advancement of the 5th Generation (5G) mobile communication technology, high-performance computing and sensing technologies, panoramic video has attracted more and more research attention. Through head-mounted display devices, panoramic video can provide users with realistic and immersive visual perception far exceeding that of flat video, and has a promising prospect. Providing users with a good Quality of Experience (QoE) is the key for video service providers to attract and retain users and succeed in the fierce market competition. Compared with the usual flat video, the data volume of panoramic video usually increases exponentially, which will put forward higher requirements for the acquisition, coding, transmission and storage of video data. So how to ensure users' QoE with limited network transmission bandwidth and storage resources has become a hot research topic in both academia and industry recently. In this paper, the QoE evaluation of panoramic video is mainly reviewed. Firstly, the possible distortion types of

panoramic video in the process of video acquisition, stitching, projection, encoding, transmission, decoding, back projection, rendering, and so on, are analyzed comprehensively, and a variety of subjective and objective influencing factors of the users' QoE are summarized, such as human factors, system factors, context, and video content characteristics and so on. On this basis, the research advances of the QoE evaluation model of panoramic video are highlighted from the aspects of influencing factors, modeling methods and so on. And the main applications of the QoE evaluation model of panoramic video in the Adaptive Bit Rate (ABR) algorithm, resource optimization allocation, bit rate control, etc, are analyzed. Finally, several representative QoE evaluation datasets of panoramic video and commonly used performance evaluation metrics are introduced. The existing problems and the future tendency of the QoE evaluation model of panoramic video are also further discussed.

**Key words:** panoramic video; quality of experience; subjective and objective influencing factors

## 1 引言

近年来,随着5G移动通信和多媒体技术的迅猛发展,移动视频通信逐步朝着“立体式、沉浸感、全场景”的方向演进,全景视频通信引起了人们的广泛关注。全景视频又称为360°视频、全向视频或者球形视频,用户通过头戴式的显示设备(Head Mounted Display, HMD)来观看视频,享受身临其境的观看体验。全景视频可以广泛应用于医疗保健、教育培训、汽车制造、营销广告、物流运输、军事国防、地产以及旅游业等众多领域,也是元宇宙在实现沉浸式体验过程中必不可少的组成部分。

与平面视频相比,全景视频的数据量倍增,对视频数据的采集、编码、传输和存储均提出了更高的要求。一般说来,4K分辨率是全景视频最基础的画质要求,需要8K甚至更高的分辨率才能让用户获得高质量的视觉体验<sup>[1]</sup>。Netflix推荐的超高清视频流传输连接速度为25 Mbps<sup>1</sup>,但目前全球只有不到10%的网络带宽能达到这个要求<sup>[2-3]</sup>。在带宽资源有限、时变的网络环境中传输全景视频时,频繁的视角切换、网络波动会导致播放卡顿、质量切换等现象时有发生,进而影响用户的观看体验质量(Quality of Experience, QoE)。因此,如何对全景视频QoE进行评价就成为设计资源优化分配策略、提升用户体验质量必须解决的首要问题,目前已经成为工业界和学术界共同关注的研究热点。

国际电信联盟<sup>[4]</sup>对QoE定义为“用户使用一项应用或服务时感到的快乐或烦恼程度”。Perkis等人<sup>[5]</sup>对全景视频QoE定义为“它是根据用户的个性和当前状态,实现其对应用程序或服务的效用和/或

享受的期望的结果”。可以看出,QoE是一种以“人”为中心的评价准则,在QoE建模时既要考虑主观因素,又要考虑客观因素,这使得全景视频QoE评价极具挑战性。具体表现在以下几个方面:

(1)对于全景视频来说,眩晕感和临场感是影响用户观看体验的主要因素,用户的兴趣、爱好等也对观看体验有着一定的影响,但是这些主观因素因人而异,往往难以度量和表达,给QoE建模带来了极大的困难;

(2)视频内容纷繁复杂,场景的丰富程度、运动复杂程度等都对用户QoE有着直接的影响,如何对视频复杂的内容特性进行有效表达也是一个难点;

(3)全景视频包括采集、拼接、渲染等多个环节,每个环节均会引入失真,而失真类型众多且相互影响,进一步增加了QoE建模的难度。

近年来,有不少学者开展了全景视频QoE评价的研究工作,并取得了一定的进展。国内外目前已经有了关于全景视频QoE评价模型的综述,比如,Chiariotti等人<sup>[6]</sup>从全景视频的编码和压缩、QoE评价模型、显著性测量和视窗预测以及全景视频的码率自适应算法等几个方面总结了全景视频的研究进展。Sara等人<sup>[7]</sup>从QoE的影响因素、主观和客观测量方法等方面综述了QoE评价模型所面临的困难和挑战。Gao等人<sup>[8]</sup>从理论和技术角度研究了全景视频QoE和传统QoE之间的差异,并对沉浸式体验的影响因素进行了探索和分类。本文则是对全景视频QoE的影响因素、QoE评价模型及其在码率自适应、资源优化分配和码率控制等方面的应用情况等进行了综述。

剩余部分安排如下:第2节分析全景视频QoE

<sup>1</sup>Netflix. (2014). Internet Connection Speed Recommendations. [Online]. Available: <https://help.netflix.com/en/node/306>

的各种影响因素;第3节总结归纳全景视频 QoE 评价模型的研究进展;第4节介绍全景视频 QoE 评价模型的应用;第5节介绍常用的全景视频数据集以及 QoE 评价模型的性能评价标准;第6节探讨了全景视频 QoE 评价模型未来可能的研究方向。

## 2 全景视频 QoE 影响因素分析

全景视频从视频采集到播放过程中涉及多个环节,包括采集、拼接、投影、编码、传输、解码、反投影、渲染等<sup>[9]</sup>,如图1所示。采集时,需要多个摄像机在同一时间、同一位置分别拍摄不同角度具有部分重叠的多个视频,然后将其进行拼接(缝合),构成完整的全景视频。之后,再将全景视频投影成平面视频,编码压缩为视频流,存储于服务端。用户在观看视频时,客户端依据设计的码率自适应算法来选择合适的视窗和质量,对到达客户端的视频进行解码、反投影、渲染后,再进行播放。可以看出,全景视频从采集到播放的各个环节都有可能存在失真,对用户的 QoE 造成影响。具体原因如下:

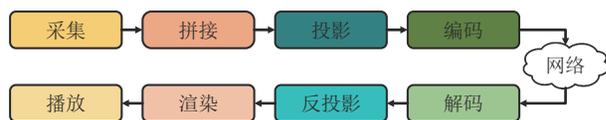


图1 全景视频从采集到播放的整体过程

Fig. 1 The overall process of panoramic video from acquisition to playback

(1)采集:摄像头会受到各种光学畸变的影响,例如色差、莫尔效应、噪声和运动模糊等<sup>[10]</sup>。

(2)拼接:由于不同摄像机之间可能存在光学畸变和运动不连续,使得拼接过程非常具有挑战性<sup>[11]</sup>。除此之外,视频拼接过程可能需要对多摄像机装置中的一些传感器限制和摄像机之间的一致性进行补偿。因此,现有的大多数方法生成的全景视频可能会出现模糊、可见接缝(由于对颜色和亮度不连续的不同曝光)、重影、边缘错位、图像结构信息缺失和几何变形等失真情况<sup>[12]</sup>。

(3)投影:由于缺乏面向全景视频的编码算法,因此对全景视频进行编码前,需要将全景视频投影为平面视频。由于球面域上的相邻区域可能不是平面域上的相邻区域,因此不同的投影方式可能意味着不同的几何变形和不连续区域,而编码效率与

投影方式密切相关<sup>[13]</sup>。

(4)编码:目前大部分的全景视频都采用 HTTP 的自适应流媒体(HTTP Adaptive Streaming, HAS)技术进行传输<sup>[14-15]</sup>。为便于传输,需要采用编码器(如 H. 266 及其改进方法<sup>[16]</sup>)将视频以不同的比特率进行编码,并将编码后的视频切割成视频片段存储于服务端,而低质量的视频会影响用户的观看体验。

(5)传输:在带宽资源有限、时变的网络环境中传输高质量全景视频是一项极具挑战性的任务<sup>[17]</sup>。虽然人们提出了众多基于视窗的全景视频传输机制<sup>[18]</sup>或者基于 Tile 的全景视频传输机制<sup>[19]</sup>,然而网络波动引起的质量切换和播放卡顿、中断等现象难以避免,给用户的观看体验造成严重影响。

(6)解码、反投影、渲染和播放:解码、反投影、渲染和播放等过程需要大量的显示和计算资源,在处理不属于用户视窗的视频内容时还会浪费大量的计算资源。除此之外,HMD 设备的性能也会影响用户的观看体验。

3D 视频与单通道视频的本质区别就在于 3D 视频添加了深度信息,能够产生立体视觉感受,使其对自然场景的表征更具真实感<sup>[20]</sup>。对 3D 视频而言,影响用户 QoE 的主要因素包括视频质量、深度感知以及视觉舒适度<sup>[21-22]</sup>。将平面视频拼接和投影之后可以获得全景视频,而视窗是其特有的概念,是用户观看全景视频时视角的直接反映,在观看全景视频期间,不同的视窗位置会带来不同的失真并影响用户的 QoE。因此,与单通道视频和 3D 视频的区别在于,全景视频 QoE 评价会着重考虑视窗内视频的质量、卡顿和质量切换等影响因素,而往往忽略视窗外的视频。作为全景视频不同于平面视频的观看感受<sup>[7]</sup>,临场感(Presence)和眩晕感(Cybersickness/motion sickness)同样会直接影响用户的 QoE。

临场感是指用户在观看视频时“身临其境”的感受<sup>[23]</sup>。Lee 等人<sup>[24]</sup>将临场感定义为“虚拟物体以感官或非感官的方式被体验为真实物体时的一种心理状态”。眩晕感是指用户在沉浸式体验过程中由于运动感知和视觉输入的不匹配导致的头晕或恶心等不良感受<sup>[25-26]</sup>。眩晕感与视频内容场景的丰富程度<sup>[27]</sup>、视频的运动复杂程度<sup>[28]</sup>等因素有很大关系。研究表明,几何失真<sup>[29]</sup>也会影响用户的

QoE。

Tran 等人<sup>[30]</sup>将 QoE 的影响因素分为用户属性、系统属性和内容属性等三类。Vlahovic 等人<sup>[7]</sup>将 QoE 的影响因素分为人的因素、系统因素以及情境上下文等三类。Fan 等人<sup>[31]</sup>将 QoE 的影响因素分为人的因素、内容因素和情境上下文和系统因素等四类。上述的分类方法分别从不同的角度归纳了全景视频的 QoE 影响因素,不够全面。

本文从各方面综合考虑,将全景视频 QoE 的影响因素归纳为四类,包括人的因素、系统因素、情境上下文以及视频内容特性,如图 2 所示。这些因素均会在一定程度上对用户的 QoE 造成影响。各影响因素的具体表述如下:

(1)人的因素:包括人眼视觉特性以及人的性别、年龄、习惯以及兴趣爱好等;

(2)系统因素:包括视频编解码器(如量化参数、编码码率、分辨率等)、QoS(Quality of Service)(如传输带宽、时延、误码率等)、采用的拼接、投影和渲染方法等;

(3)情境上下文:包括用户所处环境、网络连接、采用的观看设备以及观看区域等;

(4)视频内容特性:包括纹理丰富程度、运动剧烈程度和复杂程度、场景复杂程度等。

可以看出,全景视频的 QoE 因素众多,既包含主观因素,又包含客观因素,这些因素不是相互独

立的,而是相互关联、相互影响。QoE 建模时应综合考虑多种影响因素,才能获得理想的性能。

### 3 全景视频 QoE 评价模型

QoE 建模的本质在于找到一种映射关系  $Y=f(X)$ ,其中  $X$  是 QoE 的各种影响因素; $Y$  是用户的 QoE,目前普遍采用平均意见得分(Mean Opinion Score, MOS)进行度量<sup>[32]</sup>。QoE 建模流程图如图 3 所示。建立全景视频 QoE 评价模型时,首先需要对各种影响因素进行表征和量化,之后将其组成特征参数矢量,最后采用最优化的传统数学方法或者机器学习方法建立特征参数矢量与 MOS 之间的映射关系<sup>[33]</sup>,用于预测用户的 QoE。

表 1 所示的是目前比较具有代表性的全景视频 QoE 评价模型,表中给出了这些模型所考虑的影响因素、QoE 度量、建模方法等。

从表 1 中可以看出:

(1)在影响因素方面,早期的全景视频 QoE 建模借鉴了平面视频的建模思路,考虑的影响因素比较简单,主要包括视频质量、卡顿、质量切换等,因此建模的准确性较差;

(2)近年来全景视频建模时则更多地考虑了“人”的因素,比如性别、临场感和眩晕感、人眼的视觉特性等,因此建立的模型能更好地评估人的主观感受,评价结果与人的主观感受相一致;

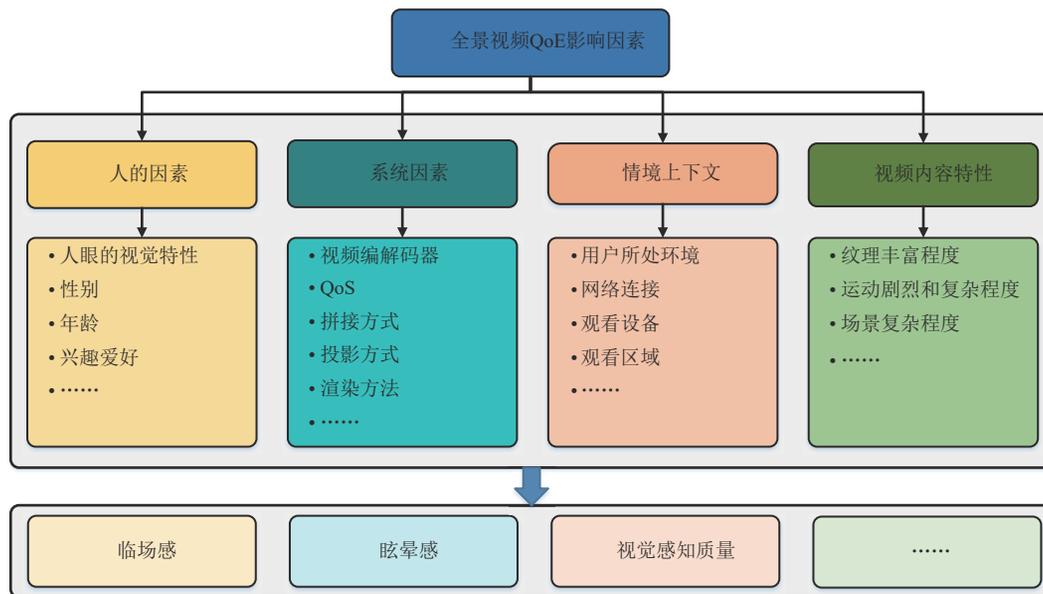


图 2 全景视频 QoE 影响因素

Fig. 2 Influencing factors of QoE in panoramic video

表 1 全景视频 QoE 评价模型汇总表

Tab. 1 The summary of QoE evaluation model for panoramic video

QoE 模型	提出时间	影响因素	建模方法	QoE 度量	数据集 (来源/数量)
Xie 等人 <sup>[37]</sup>	2019	量化参数、分辨率、质量切换	传统数学方法	MOS(1~5)	自建/1608
Yao 等人 <sup>[38]</sup>	2019	投影方式、量化参数、SI和TI	线性回归	MOS(0~9)	自建/36
Fei 等人 <sup>[40]</sup>	2019	视频码率、丢包率、时延、临场感和眩晕感	传统数学方法	MOS(1~5)	自建/468
Anwar 等人 <sup>[44]</sup>	2020	视频质量、质量切换、卡顿	贝叶斯回归	MOS	自建/208
Fan 等人 <sup>[31]</sup>	2022	视频质量、编码码率、临场感、性别、质量切换、TI和SI等	机器学习方法	MOS(0~9)	自建/48
EyeQoE <sup>[45]</sup>	2022	眼睛的注视和扫视等线索	机器学习方法	MOS(1~5)	EyeQoE/70

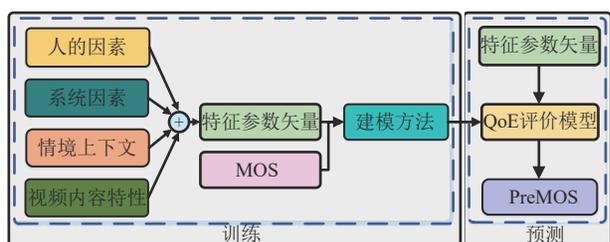


图 3 QoE 建模流程图

Fig. 3 The framework of QoE evaluation model

(3)考虑视频的内容特性会显著提升全景视频 QoE 建模的准确性,这是因为视频内容纷繁复杂,对人眼的主观感知有着直接的影响。尤其是利用深度卷积神经网络可以有效提取视频的时空特征,虽然建模时影响因素特征向量的维度会大幅增加,但是准确性也随之显著提高。

(4)建模工具则从早期的线性模型逐步转向机器学习方法,利用数据驱动的方式进行建模,建模的准确性也得以大大提升。但是需要大规模的数据集作为支撑,这也是目前全景 QoE 建模需要着重解决的问题。

2017年,Tran 等人<sup>[34]</sup>通过主观实验的方式系统地研究了视频编码参数、内容特性和 HMD 设备类型等因素对用户 QoE 的影响。之后,作者进一步深入探究了感知质量、临场感、眩晕感和可接受性等因素对用户 QoE 的影响,同时还考虑了量化参数(Quantization Parameters, QP)、视频的运动复杂程度、渲染设备和渲染模式(虚拟现实和非虚拟现实)四个关键的影响因素。研究表明:(1)当空间分辨率在 4K 到 2.5K 之间时,QoE 的差异较小。然而,当分辨率进一步降低或 QP 增加时,全景视频的质量会迅速下降;(2)用户的 QoE 与视频内容有很

大的关系。即使是拥有相同 MOS 值的两个视频,由于视频内容的不同,两个视频的码率之间仍然会有很大的差异;(3)视频的运动复杂程度对感知质量、临场感和眩晕感有着显著的影响;(4)当视频质量较高时,相比于平面视频来说,用户更喜欢选择 VR 模式进行观看。

Singla 等人<sup>[35]</sup>研究了不同 HMD(Oculus DK2 和 HTC VIVE)设备对用户眩晕感的影响。实验结果表明,与 Oculus DK2 相比,HTC VIVE 可以提供更好的 QoE;而与 HMD 相比,视频内容和分辨率对用户 QoE 的影响更大。

2018年,Xie 等人<sup>[36]</sup>研究了不同视频质量对视窗内外区域的影响,建立了 QP、分辨率和 MOS 之间的关联关系模型,具体表达式为 QP 和分辨率两个指数函数的乘积。作者还进一步研究了视频质量切换对用户 QoE 的影响,通过调整视频的量化参数和分辨率产生多组质量由高到低的失真视频,建立了多种质量切换方式的指数模型,并将其相乘作为最终的 QoE 模型<sup>[37]</sup>。

2019年,考虑到视频的内容特性,Yao 等人<sup>[38]</sup>探究了投影方式、QP、时间信息(Temporal Information, TI)和空间信息(Spatial Information, SI)等多种影响因素对 QoE 的影响,将这些影响因素相结合构成特征参数矢量,采用线性回归的方式建立了特征参数矢量和 MOS 之间的映射关系。Mahmoudpour 等人<sup>[39]</sup>考虑到人类视觉系统会对运动信息做出强烈响应这一特性,结合视频质量和抖动等两种因素,建立了一种线性 QoE 评价模型。由于考虑的影响因素较少,模型的准确性有待于进一步的提升。

Fei等人<sup>[40]</sup>建立了一种包含传输失真的QoE评价模型。作者首先建立了一个包含压缩失真和传输失真的数据集,该数据集包含9个原始视频以及经过不同网络环境传输得到的468个失真视频。在主观评价部分,观看者对沉浸感、眩晕感和整体QoE等多个方面的体验进行打分。建模时,首先根据IQX假说(QoS参数和QoE之间存在指数关系<sup>[41]</sup>),将视频码率、丢包率和时延分别与沉浸感和眩晕感建立指数模型,之后,将沉浸感和眩晕感等影响因素输入INN(Improved two-step Neural Network)中与MOS建立映射关系。数据集通过仿真网络中设置时延、丢包、带宽等参数来构建,换言之,该数据集不是采用HAS技术进行传输所获得的真实失真视频。

Singla等人<sup>[42]</sup>进一步研究了带宽、时延和分辨率等因素对用户QoE的影响以及不同时延对用户眩晕感的影响。实验结果表明,47毫秒是可接受的最大网络时延,且随着使用时间的增加,用户的眩晕感会随之增强,因此在测试过程中进行适当休息是有必要的。另外,在观看过程中,眩晕感与性别有着密切联系,相比于男性,女性更容易产生眩晕感,因此质量体验更差。

2020年,Anwar等人<sup>[43]</sup>评估了三种不同码率(1 Mbps、5 Mbps和15 Mbps)下,各种卡顿对用户QoE的影响以及卡顿和码率之间的相互作用,采用贝叶斯推理方法建立了全景视频QoE模型。实验结果表明,当视频质量接近最高和最低时,视频中多次卡顿的负面影响更甚。该作者进一步研究了量化参数、分辨率、渲染设备、性别、用户兴趣和用户对虚拟现实视频的熟悉程度等六个QoE影响因素对感知质量的影响。具体而言,首先建立了量化参数、分辨率、渲染设备与感知质量之间的映射关系,用于预测用户的感知质量;另外,建立了性别、用户兴趣和用户对虚拟现实视频的熟悉程度与眩晕感之间的映射关系,用来预测用户的眩晕感。建模时,采用逻辑回归、K近邻算法、决策树和SVR等四种机器学习方法建立感知质量与MOS之间的映射关系<sup>[44]</sup>。结果表明,相比于其他三种建模方法,逻辑回归可以获得更优的性能。

2022年,Fan等人<sup>[31]</sup>首先建立了一个包含压缩失真和质量切换等多种失真的全景视频数据集,之

后提取了视频帧质量、编码码率、临场感、SI和TI等多种影响因素构成了特征参数矢量,采用随机森林、Gradient boosting和支持向量回归(Support Vector Regression, SVR)等不同的机器学习方法建立了特征参数矢量和MOS之间的映射关系。实验结果表明,视频帧质量、编码码率等因素对用户QoE影响最大,而眩晕感与性别等因素有关。

Zhu等人<sup>[45]</sup>提出了一种新的基于视觉行为的QoE评价模型EyeQoE,并公布了名为EyeQoE的数据集。该模型利用新的视觉感知模式提取主观和客观影响因素来进行QoE建模。具体而言,作者首先将眼睛的注视和扫视等线索构建成一个图形结构数据,该数据既保留了时域特征,又保留了内容依赖性;之后,设计了一种基于图卷积神经网络的分类器,通过从图形结构数据中提取的内在特征来进行QoE建模。实验结果表明,该模型可以获得更准确的预测性能。

Hooft等人<sup>[46]</sup>探究了网络带宽、视窗预测、码率分配等因素对自适应点云流媒体感知质量的影响,建立了采用不同分发方案和点云视频内容生成的点云视频PSNR、SSIM等客观指标与MOS之间的关联关系。实验结果表明,传输过程中加入视窗预测可以提高用户的QoE;客观指标与MOS之间有很强的线性关系,这种关系还受点云视频内容的影响。

立体视频是利用人眼双目视差原理,双眼各自独立地接收来自同一场景特定摄像点的左右图像,形成双目视差,大脑依据双目视差获得图像的深度信息,使欣赏到的图像有强烈深度感、逼真感,达到超强的立体视觉效果<sup>[47-48]</sup>。立体全景视频能够给观看者提供全方位、无比真实的视觉体验。因此部分研究者对立体全景视频的QoE评价进行了研究。

兰等人<sup>[49]</sup>提出了一种基于多智能体强化学习的立体全景视频非对称传输自适应流方法,该方法首先采用视点预测概率模型来获取左右视点每个瓦片的观看概率,然后,设计了一种立体全景视频码率选择框架,以选择合适的码率,进而提高用户的QoE。其QoE评价模型为视窗内视频的质量、卡顿时长和质量切换的线性模型。

综上所述,现有的全景视频QoE评价模型着重考虑视频质量、时延、质量切换、性别、临场感和眩

晕感等各种影响因素,通过数据驱动的方法建立复杂影响因素与 MOS 之间的映射关系,虽然取得了一定的进展,但是还存在以下不足:

(1)大多在自建的数据集上进行建模和验证,数据通过仿真的方式生成,且数据规模有限。

(2)建模时考虑的影响因素不够全面,往往只关注了少数几个影响因素。导致建立的 QoE 模型的准确性比较差,适用范围有限。

(3)视频内容特性往往用 SI 和 TI 进行表达,难以有效刻画视频复杂的时空特性<sup>[50-51]</sup>。

(4)研究中考考虑的失真类型较少,多为压缩失真,对传输、渲染等其他环节导致的失真考虑得较少。事实上,在复杂多变的网络环境下传输高码率的全景视频时,可能会导致卡顿、质量切换等,也是影响用户 QoE 的重要因素。

#### 4 全景视频 QoE 评价模型的应用

全景视频数据量庞大,给压缩、传输和存储带来了很大的压力。如何在网络带宽有限和网络状况时变的情况下,提高用户的 QoE,就成为全景视频通信需要重点解决的关键问题。

为此,研究者们开展了 QoE 驱动的自适应传输和资源优化等方面的研究,这些研究从两个角度开展工作:一是以 QoE 作为奖励函数,在网络可用资源约束条件下,使得 QoE 最大化;二是在保证用户具有一定 QoE 的前提下,使得系统资源利用率最大化。除了用作奖励函数,QoE 模型也常用于评价算法性能的优劣。

现有的面向全景视频的码率自适应(Adaptive Bit Rate, ABR)算法普遍借鉴了平面视频 ABR 算法中的 QoE 评价模型,并对其进行改进。考虑到算法对实时性有严格的要求,ABR 算法中普遍采用线性 QoE 模型。Mao 等人<sup>[52]</sup>提出的 Pensieve 是最具代表性的平面视频 ABR 算法之一,其 QoE 评价模型如下:

$$QoE = \alpha \sum_{n=1}^N Q_n - \beta \sum_{n=1}^N T_n - \gamma \sum_{n=1}^N |Q_{n+1} - Q_n| \quad (1)$$

其中, $R_n$ 表示第  $n$  个视频片段的码率, $Q_n=q(R_n)$ ,表示视频质量, $T_n$ 表示卡顿时长,最后一项表示质量切换, $\alpha, \beta, \gamma$ 表示每一项的权重。

Park 等人<sup>[53]</sup>提出了一种全景视频 ABR 算法,其

中采用的 QoE 评价模型考虑了视频片段的码率、卡顿时长、质量切换以及 Tile 的丢失等四种影响因素,建立了一种线性模型。Fu 等人<sup>[54]</sup>提出的 ABR 算法中,QoE 评价模型为视频质量、空间质量变化、时间质量变化和卡顿时长等四种影响因素的线性加权。Yang 等人<sup>[55]</sup>提出的 ABR 算法中,QoE 评价模型为 Tile 内的码率、Tile 外的码率、卡顿时长以及质量切换的线性加权。

上述 QoE 评价模型大多为线性模型,由于考虑的影响因素比较少,其性能有待进一步提高。

针对上述问题,Yuan 等人<sup>[56]</sup>提出了一种基于 Tile 的全景视频 ABR 算法。该算法首先采用高斯模型预测用户的观看视窗;然后采用 Zipf 模型确定各个视窗的优先级;最终,基于不同视窗内容的率失真特性,设计了两阶段优化算法确定视窗质量的组合方式。其奖励函数为 QoE 模型,该模型考虑了视频质量(PSNR)、质量切换、卡顿时长等三种因素,通过线性加权得到。实验结果表明,该算法在避免视角切换延迟的同时,节约了带宽,并解决了局部视窗请求时,空间-时间质量不可控的难题,保障了用户观看视角的空间-时间质量一致性。

Shen 等人<sup>[57]</sup>设计了一种面向 QoE 的全景视频码率分配策略,该方法的主要思想是保证视觉重要区域具有较高的质量,而非重要区域则可以具有较低的质量。具体过程如下:首先计算每个视频片段的显著性值及其显著性权重系数,之后将显著性权重系数与视频质量进行加权,将加权后的视频质量与质量切换、卡顿时长等影响因素进行线性加权构成 QoE 评价模型,用于指导全景视频传输策略的设计。实验结果表明,设计码率分配策略时考虑视频显著性信息可以有效提高用户的 QoE。

上述的研究工作表明,QoE 驱动的自适应传输策略和资源优化分配策略可以充分利用有限的网络资源,提升用户 QoE。

#### 5 常用公共数据集和性能评价指标

数据集是开展 QoE 评价研究不可或缺的工具之一,主要用于 QoE 建模,并与其他算法进行性能比较。但是目前公认的数据集普遍是针对平面图像/视频的,在主观实验的基础上,有的学者建立了全景视频 QoE 评价数据集,但是数据规模有限、失

真类型较少,而且传输失真多通过手工或者仿真产生,与实际的传输情况存在一定的差异。目前公开的全景视频的QoE评价数据集非常稀缺,且尚未有公认的权威数据集<sup>[58]</sup>。

下面首先介绍目前常用的数据集,然后介绍QoE模型的性能评价指标。

### 5.1 数据集

表2总结了几个常用的全景视频QoE评价数据集,表中给出了各个数据集的发布时间、失真类型、分辨率、原始视频的帧率以及观看设备等信息。其中数据集大小这一列表示原始视频和失真视频的数量,对于原文中未说明原始视频数量的以失真视频总数代替。

2017年,Duan等人<sup>[59]</sup>提出的IVQAD数据集是较早的全景视频数据集。在该数据集中,作者构建了包含10个原始视频和150个失真视频的数据集。作者认为码率、帧率和分辨率是最主要的影响视频质量的因素。视频使用MPEG-4进行编码,分辨率分别设置为4096×2048、2048×1024和1024×512。在每种分辨率下,设置不同的码率和帧率以模拟不同的带宽要求。

Zhang等人<sup>[60]</sup>提出的数据集共包括16个原始视频,通过24种方式对这些视频进行处理,最终得到384个失真视频。为了模拟多媒体业务中的应用,使用了三个编码器和六个不同的码率。通过使用FFMPEG软件,将高斯噪声和盒形模糊添加到两种不同码率(10 Mbps和3 Mbps)的视频中,并采用

H.264编码器,模拟网络数据包的丢失情况。

为了研究帧率对用户QoE的影响,Zhang等人<sup>[61]</sup>和Lopes等人<sup>[62]</sup>根据JVET(Joint Video Exploration Team)<sup>2</sup>提供的10个8K和4K原始视频,分别自建了不同的数据集。Zhang等人<sup>[61]</sup>将帧率设为60 fps和30 fps,通过下采样,将原始视频的分辨率降为3600×1800,每个原始视频分别对应5个失真视频。Lopes等人在8192×4096和3840×1920的分辨率下设置了更多的帧率,分别是7.5 fps、10 fps、15 fps、30 fps以及60 fps。

2022年,Zhu等人<sup>[45]</sup>公布了目前最新的全景视频QoE评价数据集EyeQoE。该数据集中的源视频都是从YouTube和Vimeo两大全景视频平台下载的。原始视频的分辨率为4K,帧率为25 fps。这些视频涵盖自然、体育和城市景观等各种类型。每个视频的持续时间均为25 s,没有明显的场景变化。每个失真视频都包含两种类型的失真,分别是压缩失真和传输失真。对于前者,使用H.264编码将4K原始视频压缩为较低的分辨率,如2K、1080p和720p。对于后者,添加了卡顿帧来模拟三个不同版本的卡顿:8次卡顿,每次持续s,4次卡顿,每次持续2 s,以及2次卡顿,每次持续4 s。

### 5.2 性能评价指标

常用的QoE模型性能评价指标主要包括皮尔逊线性相关系数(Pearson Linear Correlation Coefficient, PLCC)、斯皮尔曼秩相关系数(Spearman Rank Order Correlation Coefficient, SROCC)和肯德

表2 全景视频QoE评价数据集

Tab. 2 Overview of the panoramic video datasets for QoE evaluation

数据集名称	年份	评价标准	压缩失真	传输失真	数量	视频格式	分辨率	时长	帧率/fps	观看设备	是否公开
IVQAD <sup>[59]</sup>	2017	MOS	有	有	10/150	MPEG-4	4096×2048,2048×1024,1024×512	15 s	30	HTC VIVE	/
Zhang等人 <sup>[60]</sup>	2017	DMOS	有	有	16/384	H.265	4096×2048	10 s	30	HTC VIVE	是
Zhang等人 <sup>[61]</sup>	2018	DMOS	有	有	10/50	H.265	3600×1800	10 s	60或30	HTC VIVE	/
Lopes等人 <sup>[62]</sup>	2018	DMOS	有	有	6/36	HEVC	7680×3840,3840×1920,1920×960,960×480	10 s	7.5/10/15/30/60	Oculus Rift	/
Fei等人 <sup>[40]</sup>	2020	MOS	无	有	9/468	无	/	/	/	/	/
Fan等人 <sup>[31]</sup>	2022	MOS	有	有	6/48	/	3840×1920	20 s	30	HTC VIVE Pro	/
EyeQoE <sup>[45]</sup>	2022	MOS	有	有	70	/	2K,1080p,720p	25 s	200	HTC VIVE	是

<sup>2</sup>ITU. 2019. JVET-Joint Video Experts Team. <https://www.itu.int/en/ITU-T/studygroups/2017-2020/16/Pages/video/jvet.aspx>.

尔秩相关系数(Kendall Rank Order Correlation Coefficient, KROCC)。三个指标的取值范围均为 $[0, 1]$ ,其数值越接近于1,说明模型性能越好。PLCC和SROCC指标由以下公式计算获得:

$$PLCC = \frac{\sum_{i=1}^m (y_{pi} - \bar{y}_p)(y_i - \bar{y})}{\sum_{i=1}^m (y_{pi} - \bar{y}_p)^2 (y_i - \bar{y})^2} \quad (2)$$

$$SROCC = \frac{6 \sum_{i=1}^m d_i^2}{m(m^2 - 1)} \quad (3)$$

其中, $m$ 是数据集中的视频数量; $y_i$ 是第 $i$ 个视频的真实分数; $y_{pi}$ 是相应的预测分数; $\bar{y}$ 是数据集中真实分数的均值; $\bar{y}_p$ 是相应的预测分数的均值; $d_i$ 是相应视频的 $y_i$ 和 $y_{pi}$ 之间的排序等级差异。

KROCC的计算公式如下:

$$KROCC = \frac{P - Q}{K} \quad (4)$$

对于KROCC,将测试集中视频的真实分数 $y_i$ 和预测分数 $y_{pi}$ 组成 $k$ 个数据对,那么从 $k$ 个数据对中任取两组数据 $\{(y_{pi}, y_i), (y_{pj}, y_j)\}, i \neq j$ ,则共有 $K=k(k-1)/2$ 种可能。对于提取的两组数据对,若 $y_{pi} > y_i, y_{pj} > y_j$ 或 $y_{pi} < y_i, y_{pj} < y_j$ ,则称这样的数据对为同序数据对,以 $P$ 表示测试集中同序数据对的个数;若 $y_{pi} > y_i, y_{pj} < y_j$ 或 $y_{pi} < y_i, y_{pj} > y_j$ ,则称这样的数据对为逆序数据对,以 $Q$ 表示数据集中逆序数据对的个数。

表3汇总了目前比较具有代表性的全景视频QoE评价模型的性能,这些模型往往是在自建的数据集上进行测试。

表3 全景视频 QoE 评价模型性能

Tab. 3 The QoE evaluation model performance of panoramic video

QoE模型	提出时间	数据集大小	PLCC	SROCC
Xie等人 <sup>[37]</sup>	2019	1608	0.97	0.97
Yao等人 <sup>[38]</sup>	2019	36	0.71	0.77
Fei等人 <sup>[40]</sup>	2019	468	0.922	/
Anwar等人 <sup>[44]</sup>	2020	208	0.8905	0.8862
Fan等人 <sup>[31]</sup>	2022	48	0.98	0.81

## 6 总结与展望

QoE建模是全景视频通信中优化分配资源、提升用户体验的前提和关键。尽管全景视频QoE评价研究已经取得了一定的进展,但是现阶段开展全

景视频QoE评价的研究工作还不够深入,仍有很大的研究空间。在未来的工作中,可以开展以下方面的研究:

(1)大规模QoE评价视频数据集的建立。大规模数据集是开展QoE评价研究的基础和前提,但是目前权威的全景视频QoE评价数据集较少。与平面视频相比,全景视频涉及更多、更复杂的环节,主观测评不易设计,需要考虑的影响因素众多,且设备条件往往难以具备,这些都使得建立大规模的全景视频QoE评价数据集十分困难。

(2)视频内容特性的量化和有效表达。视频内容特性对用户的观看体验有着直接的影响,然而视频内容的多样性和失真类型的复杂性使得对视频内容特性的量化和表达十分困难。如何有效表达复杂的视频时空特性是一个值得深入研究的问题。

(3)人眼视觉特性的应用。QoE评价是一个以“人”为中心的评价准则,现有的QoE评价对人眼的视觉特性考虑得不够,未来的研究工作中可以更多地考虑人眼的视觉特性这一因素,使得评价结果更符合人眼的感知结果。

(4)观看体验的度量。临场感和眩晕感是全景视频的观看体验区别于平面视频之处。但是由于这种体验因人而异,主观性强,涉及的因素很多,因此难以有效度量和建模表达,目前研究中涉及这方面的内容相对较少。如何综合考虑各种因素,有针对性地全景视频的观看体验进行建模也是一个值得研究的内容。

## 参考文献

- [1] HOSSEINI M, SWAMINATHAN V. Adaptive 360 VR video streaming: divide and conquer[C]//2016 IEEE International Symposium on Multimedia (ISM). San Jose, CA, USA: IEEE, 2016: 107-110.
- [2] THOMPSON J, SUN J, MÖLLER R, et al. Q1 2017 state of the internet-connectivity report[J]. State of the Internet Report, 2017.
- [3] FAUTIER T. State-of-the-art virtual reality streaming: solutions for reducing bandwidth and improving video quality[J]. SMPTE Motion Imaging Journal, 2018, 127(7): 1-10.
- [4] ITU-T: Definition of quality of experience [EB/OL]. International Telecom Telecommunication Union, Liaison Statement, Ref: TD109rev2 (PLEN/1 2), 2007. Online:

- <https://www.itu.int/rec/T-REC-P.10-200701-S! Amd1> [2022-07-30].
- [5] BRUNNSTRÖM K, BEKER S A, DE MOOR K, et al. Qualinet white paper on definitions of quality of experience [EB/OL]. European Network on Quality of Experience in Multimedia Systems and Services, 2013. Online: <https://hal.archives-ouvertes.fr/hal-00977812/document> [2022-07-30].
- [6] CHIARIOTTI F. A survey on 360-degree video: coding, quality of experience and streaming[J]. *Computer Communications*, 2021, 177: 133-155.
- [7] VLAHOVIC S, SUZnjeVIC M, SKORIN-KAPOV L. A survey of challenges and methods for quality of experience assessment of interactive VR applications[J]. *Journal on Multimodal User Interfaces*, 2022: 1-35.
- [8] GAO Yun, WEI Xin, CHEN Jianxin, et al. Toward immersive experience: evaluation for interactive network services[J]. *IEEE Network*, 2022, 36(1): 144-150.
- [9] FAN C L, LO W C, PAI Y T, et al. A survey on 360° video streaming: acquisition, transmission, and display [J]. *ACM Computing Surveys*, 2020, 52(4): 1-36.
- [10] BOEV A, HOLLOSI D, GOTCHEV A. Classification of stereoscopic artefacts [J]. *Mobile3DTV Project report*, available online at <http://mobile3dtv.eu/results>, 2008: 1-60.
- [11] SZELISKI R. Image alignment and stitching: a tutorial [J]. *Foundations and Trends® in Computer Graphics and Vision*, 2007, 2(1): 1-104.
- [12] AZEVEDO R G de A, BIRKBECK N, DE SIMONE F, et al. Visual distortions in 360° videos[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 30(8): 2524-2537.
- [13] CHEN Zhouzhong, LI Yiming, ZHANG Yingxue. Recent advances in omnidirectional video coding for virtual reality: projection and evaluation [J]. *Signal Processing*, 2018, 146: 66-78.
- [14] BUDAGAVI M, FURTON J, JIN G, et al. 360 degrees video coding using region adaptive smoothing [C]//2015 IEEE International Conference on Image Processing (ICIP). Quebec City, QC, Canada: IEEE, 2015: 750-754.
- [15] FEUVRE J L, CONCOLATO C. Tiled-based adaptive streaming using MPEG-DASH [C]//Proceedings of the 7th International Conference on Multimedia Systems. Klagenfurt Austria: ACM, 2016: 1-3.
- [16] 岳洋洋, 曾焕强, 陈婧, 等. 采用联合特征的 H. 266/VVC 屏幕内容快速模式选择算法[J]. *信号处理*, 2021, 37(8): 1479-1486.
- YUE Yangyang, ZENG Huanqiang, CHEN Jing, et al. Fast mode selection algorithm for H.266/VVC screen content using joint features[J]. *Journal of Signal Processing*, 2021, 37(8): 1479-1486. (in Chinese)
- [17] YAQOUB A, BI T, MUNTEAN G M. A survey on adaptive 360° video streaming: solutions, challenges and opportunities [J]. *IEEE Communications Surveys & Tutorials*, 2020, 22(4): 2801-2838.
- [18] OZCINAR C, DE ABREU A, SMOLIC A. Viewport-aware adaptive 360° video streaming using tiles for virtual reality [C]//2017 IEEE International Conference on Image Processing (ICIP). Beijing: IEEE, 2017: 2174-2178.
- [19] JIANG Zhiqian, ZHANG Xu, XU Yiling, et al. Reinforcement learning based rate adaptation for 360-degree video streaming [J]. *IEEE Transactions on Broadcasting*, 2021, 67(2): 409-423.
- [20] 邓慧萍. 3D 视频的深度图优化与深度编码方法研究 [D]. 武汉: 华中科技大学, 2013.
- DENG Huiping. Research on depth map optimization and coding for 3d video [D]. Wuhan: Huazhong University of Science & Technology, 2013. (in Chinese)
- [21] LIU Yanwei, CI Song, TANG Hui, et al. QoE-oriented 3D video transcoding for mobile streaming [J]. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*. 2012, 8(3s): 1-20.
- [22] XU Yiwen, HUANG Jingquan, LIU Weiling, et al. QoE assessment and management of MV/3D video services [J]. *International Forum on Digital TV and Wireless Multimedia Communications*, 2020, 1390: 348-358.
- [23] MESTRE D, FUCHS P, BERTHOZ A, et al. Immersion et présence [J]. *Le traité de la réalité virtuelle*. Paris: Ecole des Mines de Paris, 2006: 309-338.
- [24] LEE K M. Presence, explicated [J]. *Communication Theory*, 2004, 14(1): 27-50.
- [25] REBENITSCH L, OWEN C. Review on cybersickness in applications and visual displays [J]. *Virtual Reality*, 2016, 20(2): 101-125.
- [26] XU Mai, LI Chen, LIU Yufan, et al. A subjective visual quality assessment method of panoramic videos [C]//2017 IEEE International Conference on Multimedia and Expo (ICME). Hong Kong, Hong Kong: IEEE, 2017: 517-522.
- [27] STANNEY K M, HALE K S, NAHMENS I, et al. What to expect from immersive virtual environment exposure: influences of gender, body mass index, and past experience [J]. *Human Factors: The Journal of the Human Fac-*

- tors and Ergonomics Society, 2003, 45(3): 504-520.
- [28] BONATO F, BUBKA A, PALMISANO S, et al. Vection change exacerbates simulator sickness in virtual environments [J]. *Presence: Teleoperators and Virtual Environments*, 2008, 17(3): 283-292.
- [29] JONES J A, KRUM D M, BOLAS M. The effect of eye position on the view of virtual geometry [C]//2014 IEEE Virtual Reality (VR). Minneapolis, MN, USA: IEEE, 2014: 87-88.
- [30] TRAN H T T, NGOC N P, PHAM C T, et al. A subjective study on user perception aspects in virtual reality [J]. *Applied Sciences*, 2019, 9(16): 3384.
- [31] FAN C L, HUNG T H, HSU C H. Modeling the user experience of watching 360° videos with head-mounted displays [J]. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2022, 18(1): 1-23.
- [32] LI X, AGGARWAL S, SINGH A, et al. Enhancing quality of experience (QoE) assessment models for video applications [C]//6th Joint IFIP Wireless and Mobile Networking Conference (WMNC). Dubai: IEEE, 2013: 1-4.
- [33] ZHOU Wei, MIN Xionguo, LI Hong, et al. A brief survey on adaptive video streaming quality assessment [J]. *Journal of Visual Communication and Image Representation*, 2022: 103526.
- [34] TRAN H T T, NGOC N P, PHAM C T, et al. A subjective study on QoE of 360° video for VR communication [C]//2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSp). Luton: IEEE, 2017: 1-6.
- [35] SINGLA A, FREMEREY S, ROBITZA W, et al. Measuring and comparing QoE and simulator sickness of omnidirectional videos in different head mounted displays [C]//2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX). Erfurt, Germany: IEEE, 2017: 1-6.
- [36] XIE Shaowei, XU Yiling, QIAN Qiaojian, et al. Modeling the perceptual impact of viewport adaptation for immersive video [C]//2018 IEEE International Symposium on Circuits and Systems (ISCAS). Florence: IEEE, 2018: 1-5.
- [37] XIE Shaowei, XU Yiling, SHEN Qiu, et al. Modeling the perceptual quality of viewport adaptive omnidirectional video streaming [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 30(9): 3029-3042.
- [38] YAO Shunhuai, FAN Chingling, HSU C H. Towards quality-of-experience models for watching 360° videos in head-mounted virtual reality [C]//2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX). Berlin, Germany: IEEE, 2019: 1-3.
- [39] MAHMOUDPOUR S, SCHELKENS P. Omnidirectional video quality index accounting for judder [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 31(1): 61-75.
- [40] FEI Zesong, WANG Fei, WANG Jing, et al. QoE evaluation methods for 360-degree VR video transmission [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2020, 14(1): 78-88.
- [41] HOSSFELD T, TRAN-GIA P, FIEDLER M. Quantification of quality of experience for edge-based applications [M]//Managing Traffic Performance in Converged Networks: Volume 4516. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007: 361-373.
- [42] SINGLA A, GÖRING S, RAAKE A, et al. Subjective quality evaluation of tile-based streaming for omnidirectional videos [C]//Proceedings of the 10th ACM Multimedia Systems Conference. Amherst Massachusetts: ACM, 2019: 232-242.
- [43] ANWAR M S, WANG J, ULLAH A, et al. Measuring quality of experience for 360-degree videos in virtual reality [J]. *Science China Information Sciences*, 2020, 63(10): 202301:1-15.
- [44] ANWAR M S, WANG J, KHAN W, et al. Subjective QoE of 360-degree virtual reality videos and machine learning predictions [J]. *IEEE Access*, 2020, 8: 148084-148099.
- [45] ZHU Huadi, LI Tianhao, WANG Chaowei, et al. EyeQoE: a novel QoE assessment model for 360-degree videos using ocular behaviors [J]. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2022, 6(1): 1-26.
- [46] HOOFT J, VEGA M T, TIMMERER C, et al. Objective and subjective QoE evaluation for adaptive point cloud streaming [C]//2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX). Athlone, Ireland: IEEE, 2020: 1-6.
- [47] 张艳. 立体视频质量评价关键技术研究 [D]. 上海: 上海大学, 2013.  
ZHANG Yan. Research on the key technologies of stereo video quality assessment [D]. Shanghai: Shanghai University, 2013. (in Chinese)
- [48] 王腾飞. 面向 QoE 的立体视频差错控制技术研究 [D].

- 宁波: 宁波大学, 2014.
- WANG Tengfei. Research on QoE oriented technologies of stereoscopic video error control [D]. Ningbo: Ningbo University, 2014. (in Chinese)
- [49] 兰诚栋, 饶迎节, 宋彩霞, 等. 基于强化学习的立体全景视频自适应流[J]. 电子与信息学报, 2022, 44:1-8.
- LAN Chengdong, RAO Yingjie, SONG Caixia, et al. Adaptive streaming of stereoscopic panoramic video based on reinforcement learning [J]. Journal of Electronics & Information Technology, 2022, 44:1-8. (in Chinese)
- [50] DU Lina, ZHUO Li, LI Jiafeng, et al. Video quality of experience metric for dynamic adaptive streaming services using DASH standard and deep spatial-temporal representation of video [J]. Applied Sciences, 2020, 10 (5): 1793:1-15.
- [51] DU Lina, ZHUO Li, LI Jiafeng, et al. Deep spatial and temporal information based QoE evaluation model for HTTP adaptive streaming [C]//2021 5th International Conference on Computer Science and Artificial Intelligence. Beijing China: ACM, 2021: 394-401.
- [52] MAO Hongzi, NETRAVALI R, ALIZADEH M. Neural adaptive video streaming with Pensieve [C]//Proceedings of the Conference of the ACM Special Interest Group on Data Communication. New York, NY, USA: Association for Computing Machinery, 2017: 197-210.
- [53] PARK S, HOAI M, BHATTACHARYA A, et al. Adaptive streaming of 360-degree videos with reinforcement learning [C]//2021 IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa, HI, USA: IEEE, 2021: 1838-1847.
- [54] FU Jun, CHEN Xiaoming, ZHANG Zhizheng, et al. 360SRL: a sequential reinforcement learning approach for ABR tile-based 360 video streaming [C]//2019 IEEE International Conference on Multimedia and Expo (ICME). Shanghai, China: IEEE, 2019: 290-295.
- [55] YANG Shujie, HU Jialu, JIANG Ke, et al. Hybrid-360: an adaptive bitrate algorithm for tile-based 360 video streaming [J]. Transactions on Emerging Telecommunications Technologies, 2022, 33(4):1-16.
- [56] YUAN Hui, ZHAO Shiyun, HOU Junhui, et al. Spatial and temporal consistency-aware dynamic adaptive streaming for 360-degree videos [J]. IEEE Journal of Selected Topics in Signal Processing, 2020, 14(1): 177-193.
- [57] WANG Shen, DING Lianghui, ZHAI Guangtao, et al. A QoE-oriented saliency-aware approach for 360-degree video transmission [C]//2019 IEEE Visual Communications and Image Processing (VCIP). Sydney, Australia: IEEE, 2019: 1-4.
- [58] 陈思佳. 全景视频的质量评价和编码优化 [D]. 武汉: 武汉大学, 2020.
- CHEN Sijia. Quality assessment and coding optimization for omnidirectional video [D]. Wuhan: Wuhan University, 2020. (in Chinese)
- [59] DUAN Huiyu, ZHAI Guangtao, YANG Xiaokang, et al. IVQAD 2017: An immersive video quality assessment database [C]//2017 International Conference on Systems, Signals and Image Processing (IWSSIP). Poznań, Poland: IEEE, 2017: 1-5.
- [60] ZHANG Bo, ZHAO Junzhe, YANG Shu, et al. Subjective and objective quality assessment of panoramic videos in virtual reality environments [C]//2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). Hong Kong, Hong Kong: IEEE, 2017: 163-168.
- [61] ZHANG Yingxue, WANG Yingbin, LIU Feiyang, et al. Subjective panoramic video quality assessment database for coding applications [J]. IEEE Transactions on Broadcasting, 2018, 64(2): 461-473.
- [62] QUELUZ M P, LOPES F, ASCENSO J, et al. Subjective and objective quality assessment of omnidirectional video [C]//TESCHER A G. Applications of Digital Image Processing XLI. San Diego, United States: SPIE, 2018: 25.

### 作者简介



**杜丽娜** 女, 1992年生, 山西忻州人。北京工业大学博士生, 研究方向为 QoE 评价与自适应传输等。  
E-mail: dulina@emails.bjut.edu.cn



**卓力** 女, 1971年生, 江苏徐州人。博士, 北京工业大学教授、博士生导师, 研究方向为图像/视频的编码与传输、多媒体大数据处理等。  
E-mail: zhuoli@bjut.edu.cn



**李嘉锋** 男, 1986年生, 天津人。博士, 北京工业大学讲师、硕士生导师, 研究方向为计算机视觉、图像增强。  
E-mail: lijiafeng@bjut.edu.cn