# 虚拟现实(VR)体验标准 技术白皮书

The Technology White Paper of VR User Experience Assessment Standard

#### 执行摘要:

本白皮书在 VEA 体验联盟秘书处及工作组总体规划,在核心贡献单位华为 XLabs 实验室、西安电子科技太学多媒体实验室、上海交通大学图像通信实验室、首都师范大学机器视觉实验室负责具体组织和实施,并在广大成员单位积极参与主观测试下,领域专家和工程师系统研究 VR 体验关键影响因素,构建了评估模型框架,并通过符合 ITU 规范的主观实验方法,阶段性地建立起 VR 体验评估模型。评估模型从视听沉浸体验质量、呈现体验质量和交互体验质量的角度,对分辨率、帧率、码率、FOV、MTP、自由度等 20 多个指标进行定义、建模和量化,为满足运营商、设备制造商、内容提供商等在不同层面、不同角度的体验评估需求,提供不同的评估模块,是 VR 体验评估标准化进程的重要里程碑。当然,现阶段受限于主观测试数据规模和指标测量手段,后续还需要 VEA 体验联盟和产业伙伴继续通力合作,提升模型准确性,并对模型实施更新迭代。

#### 目录

1	背景和范围		2
2	缩略语、术	语和定义	3
	2.1 缩略	语	3
	2.2 术语	和定义	4
3	VR 服务用	<sup>白</sup> 体验指标体系	5
	3.1 VR	系统简介	5
	3.2 关键	指标	6
		框架	
	3.4 参数	采集及指标处理	9
4	VR 服务用	<b>ウ体验指标构成</b>	
		体验质量输入参数	
		体验质量输入参数	
	4.3 交互	体验质量输入参数	11
	4.4 VR_	MOS 综合模型输出参数	12
5	VR 主观实	<b>检及分析方法</b>	12
	5.1 主观	实验方法	
	5.1.1	通用测试条件	
	5.1.2	特殊测试方法	
	5.1.3	主观实验过程	
	5.1.4	主观评分标准	
		本验主观实验平台	
		)测试范围构建	
		多自由度交互体验质量测试	
		实验测试数据示例	
6		<b>白体验评估模型</b>	
		模型综述	
		三大模块综述	
		沉浸体验质量 Q <sub>ImE</sub>	
	6.2.2	呈现体验质量 QPE	
	6.2.3	74—11 327X— 4///	
7	VR 体验评价	古模型的典型值	30
8	问题和展望		34
0	会安立試		25

# 1 背景和范围

近年来,随着虚拟现实(VR,virtual reality)技术的不断进步和完善,相关产业得到了蓬勃的发展。如今,虚拟现实技术已经进入到教育、娱乐、军事、医疗、环保、交通运输、公共卫生等各种与人们生产、生活息息相关的领域当中。相比传统的视频业务,VR 具有多视角,交互性强等优点,为用户提供了一种全新的视觉体验,具有极大的应用价值和商业潜力。



图 1 全球 VR 市场规模预测

然而,目前 VR HMD 仍存在屏幕分辨率低,佩戴舒适性较差,MTP 较长(或在终端 ATW/ASW 机制下存在黑边)等影响用户体验的诸多问题。在推进用户体验主观测试的同时,VEA 视频体验联盟成员单位也实施了 4K 显示器和 4K VR HMD 的用户选择意愿调研。调研结果表明,有 55.8%的用户选择使用显示器观看视频,影响用户选择的主要原因则是 VR HMD 佩戴舒适性较差和画面清晰度低。

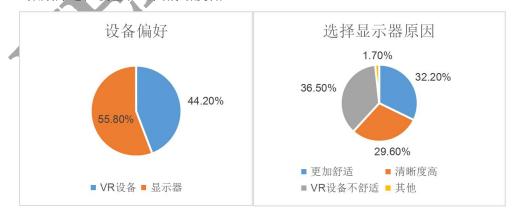


图 2 VR 设备与 4K 显示器的用户选择偏好及原因调研

VR 设备、内容和平台自身在不断地发展完善的同时,把 VR 体验评估模型作为重要的牵引工具,对用户体验现状进行评估,识别体验瓶颈进而促进优化,可以有效地推动 E2E 产业发展。因此,VR 体验评估模型具有重要的现实价值。

本白皮书系统分析了影响 VR 视频和游戏业务用户体验的主要影响因素,基于模型完整性和建模可行性,构建评估模型框架,并基于符合 ITU 规范的主观实验方法,阶段性地建立 VR 用户体验评估模型。模型不仅可对 VR 服务的用户体验进行综合评价,也可对不同的影响因素进行定量分析。

# 2 缩略语、术语和定义

# 2.1 缩略语

以下缩略语适用于本白皮书。

以「	「缩略语适用于本白皮书。	
缩略语	中文全称	英文全称
VR	虚拟现实	Virtual Reality
HMD	头戴式可视设备	Head Mount Display
GPU	图形处理器	Graphics Processing Unit
ITU	国际电信联盟	International Telecommunication Union
KPI	关键性能指标	Key Performance Indicator
KQI	关键质量指标 🔵	Key Quality Indicator
OTT	互联网服务商向用户提供的应用服务	Over The Top
MOS	主观体验评分	Mean Opinion Score
ACR	绝对等级评价法	Absolute Category Rating
MTP	移动视觉感知延迟	Motion To Photons (Latency)
PPD	角分辨率	Pixels Per Degree
DOF	自由度	Degree Of Freedom
Br	码率	Bitrate
FR	帧率	Frame Rate
QP	压缩量化参数	Quantization Parameter
BPP	每像素编码比特数	(coded) Bits Per Pixel
TCP	传输控制协议	Transmission Control Protocol
UDP	用户数据报协议	User Datagram Protocol
FEC	前向纠错	Forward Error Correction
QoE	体验质量	Quality of Experience
QoS	服务质量	Quality of Service
RTT	往返时延	Round-Trip Time
ATW	异步时间扭曲	Asynchronous Time Warp
ASW	异步空间扭曲	Asynchronous Space Warp

## 2.2 术语和定义

下列术语和定义适用于本白皮书。

#### 2.2.1

用户体验质量 quality of user experience 用户感受到的服务过程中硬件操作、服务操作和内容展示的质量水平。

#### 2.2.2

业务质量 quality of service

指网络能够利用各种技术,为指定的网络业务提供更好的服务能力。

#### 2.2.3

用户体验关键参数 key parameter of user experience

对用户感受到的服务过程中硬件操作、服务操作和内容展示的质量水平有关键影响的因素,可以被表示为客观的体验质量参数。

#### 2.2.4

业务性能参数 performance parameter of service

服务过程中可以衡量服务实现和内容展示质量水平的因素,可以被表示为客观的体验质量参数。

#### 2.2.5

网络性能参数 performance parameter of network 对业务完成和业务质量程度有直接影响的网络关键性能指标。

#### 2.2.6

2K 分辨率

指水平方向每行的像素数达到 2000 以上,主流 2K 分辨率为 2560×1440 (Quad HD)。2048×1536 (QXGA),2560×1600 (WQXGA) 和 2048×1080 也属于 2K 分辨率范畴。

#### 2.2.7

4K 分辨率

指水平方向每行的像素数达到或接近 4096,多数情况下特指 4096×2160。 4096×3112 (Full Aperture 4K),3656×2664 (Academy 4K)和 3840×2160 (UHD TV)也属于 4K 分辨率范畴。

#### 2.2.8

VR 视频点播 VR Video on Demand

根据用户发出 VR 视频点播请求,系统检索出存放在片源库中的节目信息,以视频和音频流的方式,通过传输网络传送到用户终端。

#### 2.2.9

VR 视频直播 VR Video Streaming

根据用户发出 VR 视频直播请求,系统将实时的视频和音频流,通过传输网络传送到用户终端。

#### 2.2.10

#### PLUGE Picture Line-Up Generation Equipment

用于调整图像监视器的黑电平和对比度的灰度测试规范。

#### 2.2.11

D65 标准光源。

# 3 VR 服务用户体验指标体系

## 3.1 VR 系统简介

Cloud VR 是目前重要的 VR 系统架构发展趋势,Cloud VR 系统构成如图 3 所示,主要包括三个部分: VR HMD,传输网络和云主机。其基本原理是,VR HMD 捕获用户动作后,将动作数据经传输网络上传到云主机,云主机完成逻辑计算、渲染、编码压缩,然后将视频数据流经传输网络下发到 VR HMD,最后由 VR HMD 对视频数据流进行解码并显示到屏幕。

云渲染、视频流化、数据传输和 HMD 刷新显示的串行处理过程容易导致 MTP 增大,目前 Cloud VR 多采用动作预测和端云异步渲染技术予以缓解。对于异步渲染,云端根据终端上传的姿态与位置信息渲染基础画面,终端则根据实时的姿态和位置信息,对本地已有的基础渲染画面采取旋转、平移、扭曲等二次渲染措施,生成实时画面。二次渲染算法一般包括 ATW 异步时间扭曲和 ASW 异步空间扭曲,前者进行头部姿态变化的偏移处理,后者进行头部位置变化的偏移处理,二次渲染视域中超出原有基础渲染画面的部分则会显示为黑边或拖影,但此时用户的感知 MTP 会较低,所以在 VR 用户体验评估中,应对系统是否采用动作预测或端云异步渲染技术予以充分考虑。

而对于强交互类业务,一般无法通过动作预测和端云异步渲染技术解决潜在的操作响应 时延长的问题,如射击类游戏,无法预判用户扣动扳机的时间,仍然需要 5G 网络高带宽、 低时延能力的支持。

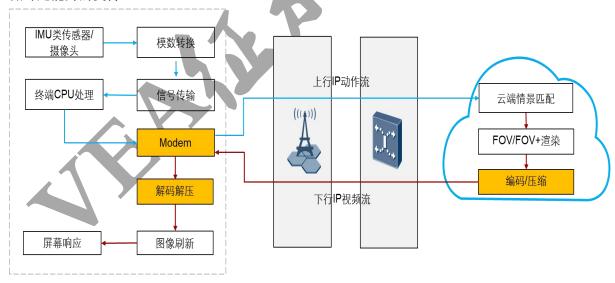


图 3 Cloud VR 系统构成

## 3.2 关键指标

用户应用 VR 视频(从视频数据传递方式的角度,分为直播和点播,从 VR HMD 的呈现方式角度,分为全景视频、FOV 视频和巨幕视频)或游戏服务的过程,通常包括启动 VR

HMD,接入服务系统,选择内容,实施播放、观看、操控等过程,影响用户体验的因素众多,在现阶段不具备条件将触觉体验,嗅觉体验纳入评估范畴的情况下,影响用户体验的关键维度主要包括沉浸体验、呈现体验和交互体验。

沉浸体验,取决于用户的视听质量,其中的重要因素一视觉质量,则进一步取决于视频或游戏画面的清晰度、流畅度、保真度(色调、对比度),特别地,作为 VR 用户的视觉体验评价指标,视觉质量还受立体感和视场角等因素影响。具体影响该 KQI 指标的 KPI 涵盖由视频源和 HMD 能力决定的分辨率、帧率、码率、内容、编码等维度。

呈现体验,通常受视频观看或游戏过程中传输带宽、时延抖动、丢包而导致的黑边、花屏、卡顿等问题影响,可以用数据传输性能和质量损伤等客观指标,通过建模衡量。

交互体验,取决于系统对用户操作的响应速度、交互维度丰富度等客观指标,涵盖平台、 网络、终端性能指标。

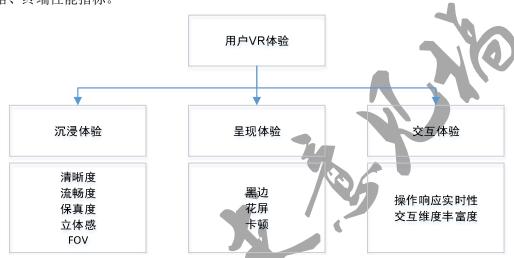


图 4 影响用户 VR 体验的因素

为了科学地评估和量化 VR 视频或游戏业务的用户体验质量,把评估场景分为会话场景(表征用户一次完整的视频观看或游戏的体验质量)和实时场景(表征用户实时观看或游戏过程中某一时刻的体验质量)。对用户在一定周期内(周、月、年)会话场景下的体验质量进行统计分析,可以反映用户使用 VR 视频或游戏服务的真实感知体验。

## 3.3 模型框架

VR 业务的综合体验受到 HMD、网络、平台、内容等多方面因素影响,具体的评估模型框架,在充分考虑模型完整性和建模可行性的基础上,可采用分级映射的方式构建。基于分级映射,可避免大量参数交叉测试,使主观测试环节测试量可控,并有利于分析、建立评估模型。具体的 VR 业务用户体验评估模型框架,如图 5 所示。评估模型框架将内容质量和HMD 规格/性能合并测试、研究其对用户视听体验的影响,将这两方面相应指标(如内容分辨率和 HMD 支持的最大分辨率)的低值,作为评估模型的实际输入参数。

- 1) 用户的视频体验(包含 VR 立体视觉和视场角的影响)和音频体验构成沉浸体验质量。在不受卡顿、花屏等体验损伤因素影响的情况下,沉浸体验质量是用户观看 VR 视频的主要体验关注方向,用来表征用户所见、所听的沉浸式感受:
- 2) 呈现体验质量主要用来表征受网络数据传输质量影响的体验损伤,可对网络数据传输质量提出明确要求:
- 3) 交互类业务,特别是交互游戏业务,对交互维度丰富程度以及操作响应的实时性提出更为明确的体验要求,在沉浸体验的基础上,用户进一步关注临场体验,也即

临场感。综合沉浸体验质量、交互体验质量和呈现体验质量的临场感,体验联盟定义为 VR\_MOS。

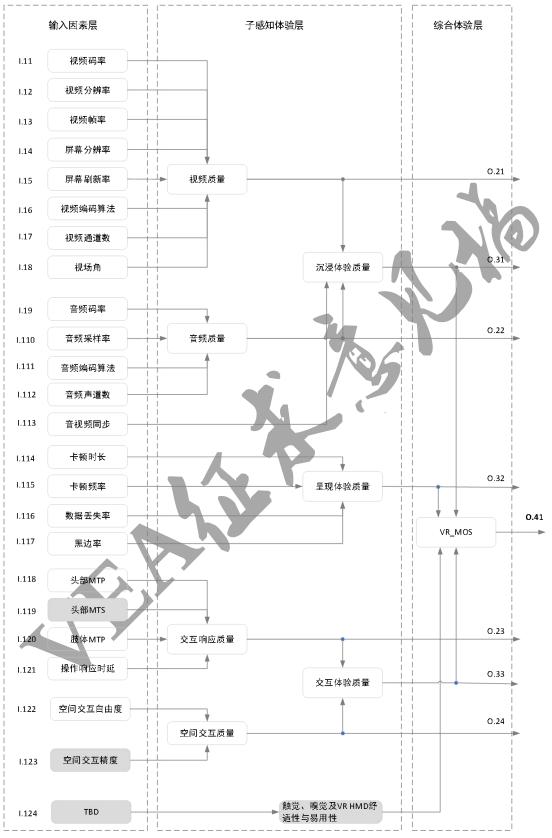


图 5 VR 视频/游戏业务用户体验评估模型框图

注:

除虚拟现实的触觉体验和嗅觉体验外,VR HMD 的佩戴舒适性,现阶段也暂未系统地纳入评估模型。主要考虑 HMD 重量是可量化指标外,其他影响因素,如材料,人体工程学设计,佩戴方式等,均难以客观量化并作为模型入参。但作为研究的一部分,成员单位测试研究了HMD 的重量和佩戴舒适性的关系,见章节 5.5 描述。

## 3.4 参数采集及指标处理

由于 VR 视频或游戏服务最终通过终端呈现给用户,是离用户最近的环节,因此评估影响 VR 视频/游戏服务的参数来源可以是来自终端的网络抓包数据、终端系统信息相关数据及播放器反映的视频信息相关的数据(对于 VR 游戏则需要 GPU 云主机配合提供音频/视频编码方式、音频声道数等参数)。

参数的获取可以通过终端如 VR HMD、VR 播放器等软硬件设备,或者通过网络探针的方式部署在网络各个位置,从而协助进行故障定位和诊断。当参数从终端获取时,计算出的业务质量精度高,但需要终端设备的支持;而从网络探针获取时,因为网络设备能力受限,无法使用比特流模型和混合模型,且无法采集到交互体验数据,因此计算出的业务质量精度较低。

获取的数据经过处理可以得到 KPI 指标 (网元相关指标) 和 KQI 指标 (业务相关指标),KPI 指标和 KQI 指标就组成了 VR 视频/游戏业务质量分析的 QoS 指标。QoE 指标是基于用户感知层面的,基于 KQI 指标从一定程度上可以反应用户对业务的使用体验。需要说明是,部分指标的测量,如 MTP,需要在渲染主机和 HMD 终端进行定制开发,才能实现相关指标的测量。

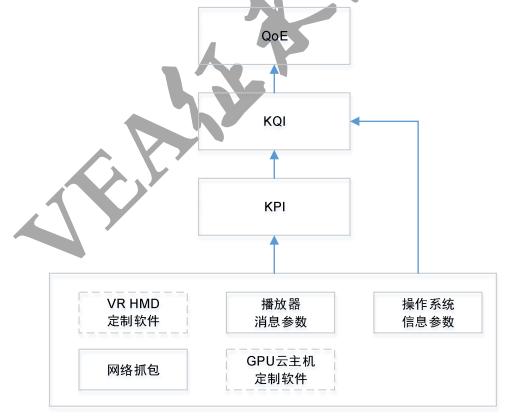


图 6 VR 视频/游戏数据采集及处理流程

# 4 VR 服务用户体验指标构成

# 4.1 沉浸体验质量输入参数

VR 视频源/游戏的沉浸体验质量输入参数说明,如表 1 所示。

表 1 (VR 视频源/游戏的) 沉浸体验质量参数

	25 N/I			支体验质重参数 「共立人名
序号	参数	定义	缩写	英文全称
4.1.1	视频码率	单位时间传输的视频 数据量(比特每秒, bps)	Br <sub>v</sub>	Video Bitrate
4.1.2	视频帧率	VR 视频/游戏每秒的 帧数(fps)	FR	Video Frame Rate
4.1.3	视频(水 平)分辨 率	VR 视频/游戏水平方 向的像素点个数	R <sub>h</sub>	Resolution (Horizontal)
4.1.4	视频 (垂 直)分辨 率	VR 视频/游戏垂直方 向的像素点个数	$R_{v}$	Resolution (Vertical)
4.1.5	屏幕分辨 率	VR HMD 单目屏幕水平方向的像素点个数	R <sub>Sh</sub>	Screen Resolution (Horizontal)
4.1.6	屏幕刷新 率	VR HMD 屏幕每秒刷 新的次数(Hz)	RR	Screen Refresh Rate
4.1.7	视频通道 数	单目视觉( <b>1</b> )或立体 视觉( <b>2</b> )	NVC	Number Of Video Channel
4.1.8	视频编码	视频编码标准,如 H.265/HEVC, H.264/AVC,VP9	Codec√	Video Codec
4.1.9	视场角	VR HMD 水平方向单 目视场角	FoV <sub>h</sub>	Field Of View
4.1.10	音频码率	单位时间传输的音频 数据量(千比特每秒, Kbps)	Bra	Audio Bitrate
4.1.11	音频声道 数	立体声音( <b>2</b> )或空间 化声音( <b>8</b> )	NAC	Number Of Audio Channel
4.1.12	音频编码 算法	音频编码标准,如 AAC-LC,Opus	Codeca	Audio Codec
4.1.13	音频采样 率	音频采样的频率 (Hz)	Sra	Audio Sampling Rate
4.1.14	音视频同 步	VR 视频/游戏的声音 和画面之间的同步时	T <sub>asyn</sub>	Asynchronous Time

	延		
	処		
<b>I</b>			

# 4.2 呈现体验质量输入参数

VR 视频/游戏的呈现体验质量输入参数说明,如表 2 所示。

表 2 (VR 视频/游戏的)呈现体验质量参数

序号	参数	说明	缩写	英文全称
4.2.1	卡顿时长	在观看 VR 视频或游戏过程中,发生卡顿的平均时长(秒,s)	程中,发生卡顿 T <sub>r</sub>	
4.2.2	卡顿频率	在观看 VR 视频或游戏过程中,卡顿次数/观看视频或游戏的总时长	RF	Re-buffering Frequeny
4.2.3	数据丢失率	应用数据丢失率 (%)	PLR	Packet Loss Rate
4.2.4	黑边率	头部转动或移动时, 二次渲染画面中黑 边的占比(%)	P <sub>black</sub>	Black Edge Rate

# 4.3 交互体验质量输入参数

VR 视频/游戏的交互体验质量输入参数说明,如表 3 所示。

表 3 (VR 视频/游戏的)交互体验质量参数

序号	参数	说明	缩写	英文全称
4.3.1	头部 MTP	头部转动到画面刷新之间 的滞后时长(ms)	MTP <sub>h</sub>	Head MTP Latency
4.3.2	头部 MTS	头部转动到音频方位变化 之间的滞后时长(ms)	MTS <sub>h</sub>	Head MTS Latency
4.3.3	肢体 MTP	肢体运动到画面中肢体移 动之间的的滞后时长(ms)	MTP <sub>b</sub>	Body MTP Latency
4.3.4	操作 响应 时长	用户操作与操作响应之间 的滞后时长(ms)	Tor	Operation Response Delay
4.3.5	空间 交互 自由 度	VR 系统支持交互、操作的 空间维度数,可通过支持交 互的节点数(如,头、手、 躯干等)和各节点支持的交 互维度数联合表征	DOF	Degree of freedom
4.3.6	空间 交互 精度	头部跟踪、动作捕捉、手势 识别、方向追踪等交互行为 定位的空间误差	SIA	Spatial Interaction Accuracy

# 4.4 VR\_MOS 综合模型输出参数

VR\_MOS 代表 VR 视频/游戏的整体临场感体验,输出参数说明,如表 4 所示。

表 4	模型输出参数
4X <b>T</b>	1 <del>111</del> 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

序号	参数	说明	缩写	英文全称
4.4.1	视频质量	对视频画面的真	$Q_V$	Visual Quality
		实感质量的评分		, ,
		对音频真实感质		
4.2.2	音频质量	量的评分	$Q_A$	Audio Quality
4.2.3	六五响应氏县	对交互时延质量	0	Interactive Response
4.2.3	交互响应质量	的评分	$Q_{IR}$	Quality
4.2.4	空间交互质量	对自由度和交互	Qsı	Spatial Interaction
4.2.4	工門又互灰里	精度质量的评分	<b>Q</b> SI	Quality
		VR 的画面和音频		Immersion
4.2.5	沉浸体验质量	的逼真度,沉浸感	$Q_{lmE}$	Experience Quality
		质量的评分		Experience Quality
		网络传输质量,如	$Q_{PE}$	
4.2.6	呈现体验质量	丢包、时延,对 VR		Presenting
7.2.0	主况件视从重	体验质量影响的		Experience Quality
		评分		
		对 VR 内容和设备	Q <sub>InE</sub>	Interaction
4.2.7	交互体验质量	的交互真实感进		Experience Quality
		行的评分		Experience Quality
4.2.8	VR MOS	用户 VR 体验质量	VR_MO	Virtual Reality MOS
7.2.0	V1\_IVIO3	综合评分	S	Virtual reality MOS

# 5 VR 主观实验及分析方法

## 5.1 主观实验方法

目前 VR 视频/游戏评价,特别是涉及用户体验的评价,最终都需要落地到主观评价方法上。主观评价方法采用的测量能够更直接地预测可能观看受测系统的人的反应。就此而言,可以认为用客观方法可能无法全面地描述系统的特性;因此,有必要用主观测量作为客观测量的补充。

VR 视频/游戏的主观评价方法主要依据 ITU-R BT.500、ITU-R BT.1788 等标准,提供符合标准的测试环境,测试者根据评价标准对测试素材的体验质量给出主观评价。对测试者的打分进行有效处理,得到测试视频序列的主观分值,进而建立用户体验评估模型。

#### 5.1.1 通用测试条件

#### 5.1.1.1 测试环境

在测试过程中,应满足如下环境要求,以达到最佳测试效果:

a) 温度 25℃~28℃

b) 湿度 45%~65%

c) 显示器亮度和对比度 满足 PLUGE 规范要求

d) 背景 满足 D65 规范要求

#### 5.1.1.2 测试素材

在时间域方面,主观测试选用 8 个 4K 分辨率(3840\*1920),120fps 的 5 秒原始 YUV 序列,通过抽帧将视频处理至 60fps、30fps 和 15fps 不同帧率的视频序列,建立帧率与画面质量 [1]间的关系。在空间域方面,将原始序列下采样为不同的分辨率、不同码率和像素密度,建立其与画面质量间的关系。

注:[1] 关于画面质量的定义,请参考本白皮书 6.2.1.1 小节,不同。

#### 5.1.1.3 测试者

#### 1. 视力和色弱

在测试阶段开始之前,应对观察者进行筛选,使之对 Snellen 氏 E 字视力表或 Landolt 氏 C 字视力表具有(校正至)正常的视敏度,并采用专门选定的表(例如石原氏色盲检查表)使之具有(校正至)正常的彩色视觉。应使用至少 15 位观察者,所需评价者的数目取决于所用测试程序的感受性和信度,并取决于所评估的影响的预期范围。

2. 性别比例

测试者男女性别比例建议为50%:50%,可视实际情况适当变动,但男女性别比例不高于60%:40%。

- 3. 年龄范围
  - 测试者年龄分布在 20~50 岁之间。
- 4. 职业分布

测试者包括学生、教师、公司职员、政府人员等。

5. 测试经历

建议测试者对 VR 视频/游戏业务有一定的了解,无频繁参加 VR 体验主观测试的经历, 在过去 6 个月中,没有参加过 VR 体验主观测试。

#### 5.1.1.4 评价须知

应向评价者仔细介绍评价方法、容易产生的损伤类型或质量因素、分级量表、顺序及定时。应采用训练序列说明要评价的损伤的范围和类型,所用图像不同于测试中要用的图像,

但具有可相比较的感受性。

#### 5.1.1.5 测试阶段

演示的播放应采用随机顺序,但测试条件的顺序应加以安排,使得疲倦或适应对分级的 影响在不同测试阶段之间得以平衡掉。为检查相干性,有些演示可以在不同的测试阶段予以 重复。

#### 5.1.2 特殊测试方法

由于目前商用 VR HMD 支持的 PPD 较低,为了全面获取 PPD 对视频的影响,所以利用 4K 显示器进行等效测试,第 5.3 节给出了详细的等效测试方案。

## 5.1.3 主观实验过程

本次主观评价实验使用单刺激法(Single Stimulus Methods,SSM)进行测试。在 SSM 中,待测试条目/序列以随机的方式单独进行,每一观察者最好采用不同的随机顺序,测试者只对测试条目/序列的质量进行评分。测试条目/序列采用不重复的方式展示给测试者,各测试条目/序列只展示一次。

具体地,在沉浸体验质量的主观实验中,测试人员在播放序列的间隔进行评分,如下图 7 所示。在 VR HMD 上的评分中,采用的方法是由测试人员自行控制评分时长,即评分完毕后点击"下一个"按钮来播放下一个测试序列。

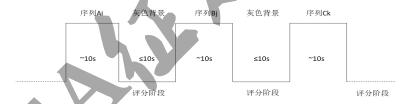


图 7 SSM 主观实验序列及评分过程示意图

# 5.1.4 主观评分标准

ITU-R BT.500 规定了视频质量和损伤的 5 级评分标度, VR 视频/游戏体验量化感知评分(MOS 值)也可采用类似的 5 级评分标度,评估人员根据视频质量体验情况,对视频质量和损伤进行适当的等级评分。5 级评分标度将主观感知评分分为 5 级,范围从 1~5,5分最高,1 分最低。具体评分表如表 5 所示。

表 5 ITU-R BT.500 定义的 5 级评分量表

	五级量表	
质量	损伤	
5 优 4 良 3 中 2 差 1 劣	<ul><li>5 不可察觉</li><li>4 可察觉,但不讨厌</li><li>3 稍微讨厌</li><li>2 讨厌</li><li>1 很讨厌</li></ul>	

## 5.2 VR 体验主观实验平台

VR 体验主观实验平台可定量控制输入因素层的各输入参数,为主观测试者提供不同的 VR 体验经历,并提供主观评分功能。图 8 给出了主观实验平台的基本结构,逻辑上包括如下 3 个模块:

- 1) VR 内容服务模块。该模块实现向 VR HMD 分发 VR 内容的功能。
- 2) 终端服务及技术参数统计模块。该模块由 VR HMD 组成,根据不同的实验目的, 向测试人员呈现不同质量和自由度的测试场景。同时,该模块还向用户数据云主机 反馈相关技术参数。
- **3)** 主观体验统计模块。该模块部署于云主机,负责记录并统计测试人员给出的主观体 验评分数据

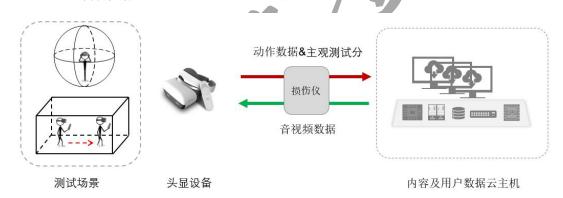


图 8 VR 主观实验平台的基本构成

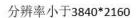
主观实验平台涉及的软硬件主要包括: 1) HTC Vive Pro; 2) Pico G2; 3) Pico Neo; 4) 2K 显示器(刷新率 144Hz); 5) 4K 显示器(刷新率 90Hz); 6) 高性能 GPU 云主机; 7) VR 播放器; 8) VR 打分软件; 9) 6 *DOF* HMD、6 DOF 手柄及其测试与控制软件。注:

- 1) PPD: 指视场角中的平均每 1° 夹角内填充的像素点的数量。PPD 数值越高,说明画面显示越精细、清晰,用户针对画面的视觉体验就越佳。
- 2) 在目前商用 VR HMD 支持的 *PPD* 和刷新率较低的情况下, 2K 和 4K 显示器主要用于等效测试 *PPD* 和帧率对画面质量的影响。

## 5.3 PPD 测试范围构建

为了研究分辨率对用户体验的影响,引入指标 PPD,也即平均每度视场角内像素点的

数量,作为像素密度的度量。假设 VR HMD 视域为 110°,单目分辨率 1K 至 10K,才能构建 10 至 90 PPD 的测试场景,但目前商用 VR HMD 单目分辨率最高仅为 2K,支持的 PPD 较低。为了全面获取 PPD 对画面质量影响的测试数据,实验选用了 32 英寸 4K 分辨率的显示器进行 PPD 等效测试,测试数据和基于 VR HMD 的测试数据一并作模型训练数据。测试视频的播放方式为视频像素和屏幕像素 1:1 播放(若视频像素低于屏幕像素,视频画面不会铺满屏幕),再通过调整用户观看距离,控制用户实际观看测试序列的 PPD,如图 9 所示。



分辨率3840\*2160

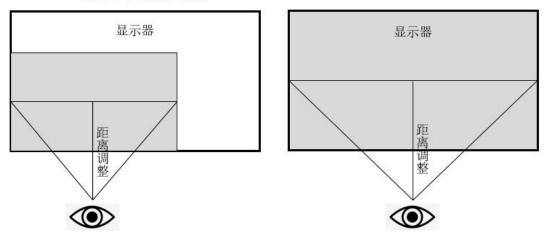


图 9 等效 PPD 测试方法

等效 PPD 和视频分辨率、观看距离的关系,如表 6 所示

表 6 32 寸 4K 显示器的等效 PPD 对照表

测试序列分辨率	观看距离(cm)	FOV	PPD
734*413	36.3	73.3°	10
1100*619	27.2	73.3°	15
1467*825	36.3	73.3°	20
2200*1238	27.2	73.3°	30
2934*1650	36.3	73.3°	40
3840*1920	56.7	64°	60
3840*1920	90.3	42.7°	90

## 5.4 游戏多自由度交互体验质量测试

游戏多自由度交互体验质量测试,以快速击破气球为任务要求,开发构建如下测试场景:

- 1) 头部 3 *DOF*、手部 3 *DOF* 自由度,可射线击破气泡,设置不同头部、肢体和操作的交互时延(主观实验平台固有 MTP 时延约为 27ms):
- 2) 头部 6 DOF、手部 3 DOF 自由度,可射线击破气泡,设置不同头部、肢体和操作的交互时延;
- 3) 头部 6 *DOF*、手部 6 *DOF* 自由度,可射线或触碰击破气泡,设置不同的头部、肢体和操作交互时延。

表 7 定义了自由度场景编号,作为评估模型输入。

表 7 自由度场景编号

实验编号	1	2	3
	头部 3 DOF +手部 3	人 人部 6 <i>DOF</i> +手部 3	头部 6 DOF +手部 6
应用描述	DOF +操作 1 DOF 游	DOF +操作 1 DOF 游戏	DOF+操作1DOF游
	戏		戏
自由度场景	7	10	13
编号	<i>I</i>	10	13

## 5.5 主观实验测试数据示例

#### 1. PPD 对画面质量的影响的主观测试数据

图 10 展示了系列测试序列中,高画质测试序列(QP=22, 压缩前每像素比特位为 12,压缩率为 57,则压缩后  $BPP_{\mathbb{K}^{\hat{g}_{E}}} \approx 0.21^{[2]}$ ),在帧率为 30fps 时,不同 PPD 下的画面质量评分结果。

从图中可以看出,主观测试人员所感知的画面质量随像素密度 PPD 的增高而增高,且在 PPD 达到 60 以后,画面质量受 PPD 的影响明显趋缓。此外,在相同 PPD 条件下,普通视频画面(如 Beauty~YachtRide)的画面质量要高于游戏画面(如 Battlefield),原因在于普通视频通常采用动态模糊技术,帧率 30fps 时,便可感觉比较流畅,而游戏画面系实时生成,帧率 30fps 时,仍能觉察出画面的跳跃感。

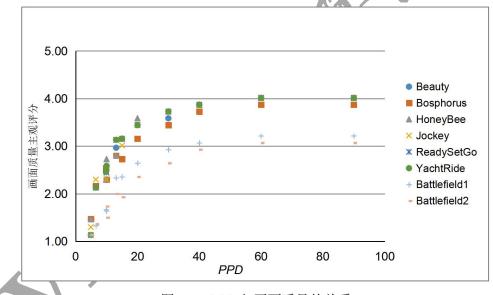


图 10 PPD 与画面质量的关系

注:

[2] BPP 为每像素编码比特数,其计算方法为:  $BPP_{\mathbb{E}^{\#_{E}}}$  = 压缩前每像素比特位 / 压缩率。
2. VR 头戴设备重量对佩戴舒适性的影响的测试数据

VR头戴设备的佩戴舒适性是影响用户综合体验的因素之一,佩戴舒适性和头戴设备的重量、佩戴方式、触面材料及人体工程学设计等因素相关。但仅设备重量可作为模型的量化入参,所以,现阶段并未将佩戴舒适性系统地纳入体验评估模型。设备重量和佩戴舒适性的主观测试结果如下图 11 所示,可以看到,当头戴设备的重量比较重时,佩戴舒适性受其他因素影响较大。

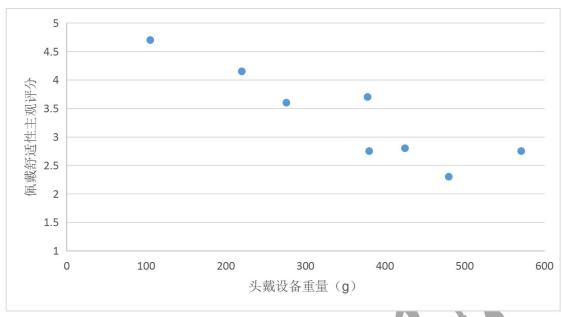


图 11 头戴设备重量与佩戴舒适性的关系

# 6 VR 服务用户体验评估模型

# 6.1 总体模型综述

 $VR\_MOS = f(Q_{ImE}, Q_{PE}, Q_{InE})$ ,即 VR 临场体验的综合评价模型为三个子模块(沉浸体验质量  $Q_{ImE}$ ,呈现体验质量  $Q_{PE}$ ,交互体验质量  $Q_{InE}$ )的函数关系,具体评估场景分为会话场景(用户一次 VR 视频观看 VR 游戏行为的时长大于 1 分钟)及实时场景(用于实时质量监控)。

$$VR\_MOS = \min\left(\max\left(\left(Q_{lmE} - 1\right) \times \left(1 - v_{59} \times \left(5 - Q_{lnE}\right) - v_{60} \times \left(5 - Q_{PE}\right) - v_{61} \times Q_{lnE} - Q_{PE}\right) + 1,\right),\right)$$

式中 V59~ V61 是模型系数, 其取值分别为:

<b>V</b> 59	V <sub>60</sub>	<b>V</b> 61
0.25	0.25	0.045

根据 VR 视频/游戏采用的传输协议/算法,对应的终端视频播放/游戏可以分为采用丢包掩盖机制和无丢包掩盖机制(在丢包掩盖机制下视频/游戏出现丢包时,解码器直接丢弃问题帧,体验表现为视频/游戏卡顿。在无丢包掩盖机制下,解码器直接播放问题帧,体验表现为视频/游戏花屏和卡顿)。具体影响实时和会话体验的计算变量如表 8 所示。

表 8 模型计算变量

传输协议/算法	有无丢包掩盖	实时(Instant)	会话 (Session)
ТСР	有/无	$Q_{ImE}$ $Q_{PE} = Q_{PE}$ _Instant $Q_{InE} = Q_{InE}$ _Instant	$Q_{ImE}$ $Q_{PE} = Q_{PE}$ _Session $Q_{InE} = Q_{InE}$ _Session

UDP w/o FEC	有/无	$Q_{ImE}$ $Q_{PE} = Q_{PE}$ _Instant $Q_{InE} = Q_{InE}$ _Instant	$Q_{ImE}$ $Q_{PE} = Q_{PE}$ Session $Q_{InE} = Q_{InE}$ Session
UDP+FEC (with low <i>PLR</i> )	有	$Q_{ImE}$ $Q_{PE} = Q_{PE}$ _Instant $Q_{InE} = Q_{InE}$ _Instant	$Q_{ImE}$ $Q_{PE} = Q_{PE}$ Session $Q_{InE} = Q_{InE}$ Session
UDP+FEC (with staged High <i>PLR</i> )	无	$Q_{lmE}$ $Q_{PE} = Q_{C}_{lnstant} + Q_{lnstant}$ $Q_{lnE} = Q_{lnE}_{lnstant}$	$Q_{lmE}$ $Q_{PE} = Q_{C}_{Session} + Q_{l}_{Session}$ $Q_{lnE} = Q_{lnE}_{Session}$
UDP+QUIC like	有/无	$Q_{ImE}$ $Q_{PE} = Q_{PE}$ _Instant $Q_{InE} = Q_{InE}$ _Instant	$Q_{lmE}$ $Q_{PE} = Q_{PE}$ Session $Q_{lnE} = Q_{lnE}$ Session

# 6.2 模型三大模块综述

# 6.2.1 沉浸体验质量 Q<sub>IME</sub>

序号	变量	物理意义	计算所需变量
1.a	Br <sub>v</sub>	视频的平均码率	4.1.1
1.b	FR	每秒内的视频帧数	4.1.2
1.c	Rh	视频水平方向的像素点个数	4.1.3
1.d	R <sub>v</sub>	视频垂直方向的像素点个数	4.1.4
1.e	R <sub>sh</sub>	单目屏幕水平方向的像素点个数	4.1.5
1.f	RR	屏幕每秒钟刷新次数	4.1.6
1.g	NVC	单目视觉(1)或立体视觉(2)	4.1.7
1.h	Codec <sub>v</sub>	W频编码标准,如 H.265/HEVC,H.264/AVC,	
1		VP9	4.1.8
1.i	FoV <sub>h</sub>	水平方向单目视场角	4.1.9
1.j	Bra	音频的平均码率	4.1.10
1.k	NAC	立体声音(2)或空间化声音(8)	4.1.11
1.1	Codeca	音频编码标准,如 AAC-LC,Opus	4.1.12
1.m	Sra	音频采样率,如 48KHz	4.1.13
1.n	T <sub>asyn</sub>	音视频同步时延	4.1.14

$$Q_P = f_1(Codec_v, Br_v, FR, (R_h, R_v))$$

$$Q_{V} = f_{2}(Q_{P}, NVC, FoV_{h})$$
(3)

$$Q_A = f_3(Codec_a, Br_a, NAC) \tag{4}$$

$$Q_{ImE} = f_4(Q_V, Q_A, T_{asym}) \tag{5}$$

注:

- 1) 计算所需变量中的符号解读: 4.1.1 表示正文 4.1 节中第 1 个指标。1.a 表示附录章 节 1 的第一个变量。
- 2) 由于本白皮书中的视频质量,特指在 VR 中包含了受视场角和立体感影响的视觉体验质量,但为了更好的观察帧率,分辨率等指标对传统视频的画面质量的影响,研究人员保留画面质量 Q<sub>0</sub> 的概念,并作为重要的一个建模环节。
- 3)根据 ITU 视频体验建模规范,本白皮书主要基于 Mode0 方式,也即基于码流层信息的轻量级建模方式,构建画面质量  $Q_p$ 的评估模型。

#### 6.2.1.1 画面质量 Q<sub>P</sub>

画面质量主要考虑了视频编码、VR HMD 分辨率和视频帧率的影响,其评估模型为:

$$Q_{P} = \begin{cases} \min \left( \max \left( (v_{1} \times \exp(v_{2} \times BPP) + v_{3}) \times (1 + v_{4} - v_{4} / (1 + (PPD/v_{5})^{v_{6}})) \times \right), 5 \right), FR \leq RR \\ (v_{7} \times \exp(v_{8} \times FR) + v_{9}), 1 \end{cases} \\ \min \left( \max \left( (v_{1} \times \exp(v_{2} \times BPP) + v_{3}) \times (1 + v_{4} - v_{4} / (1 + (PPD/v_{5})^{v_{6}})) \times \right), 5 \right), FR > RR \end{cases}$$

$$(6)$$

$$(7)$$

$$(8)$$

$$(7)$$

$$(8)$$

$$(7)$$

$$(7)$$

$$(8)$$

$$(7)$$

$$(8)$$

$$(7)$$

$$(8)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(1)$$

$$(9)$$

$$(9)$$

$$(1)$$

$$(9)$$

$$(1)$$

$$(9)$$

$$(1)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$(9)$$

$$BPP = \frac{BPF}{R_h \times R_v} \tag{7}$$

$$BPF = \frac{Br_v \ (bps)}{FR \ (fps)} \tag{8}$$

全景视频 PPD 计算如下:

$$PPD = \begin{cases} \frac{R_{h}}{360} & R_{h} \le R_{Sh} \times \frac{360}{FOV} \\ \frac{R_{Sh}}{FOV} & R_{h} > R_{Sh} \times \frac{360}{FOV} \end{cases}$$
(9)

FOV 视频 PPD 计算如下:

$$PPD = \begin{cases} \frac{R_h}{FOV} & R_h < R_{sh} \\ \frac{R_{sh}}{FOV} & R_h \ge R_{sh} \end{cases}$$
 (9')

N	
$\sum \left(1-P_{black} ight)$	
$FOV = \frac{i=1}{N} \times FoV_h$	(10)
/ <b>V</b>	

式中,FOV为单目视场角, $V_1 \sim V_9$ 为模型系数,其取值分别如下:

业 务 类型	编码 器	V <sub>1</sub>	V <sub>2</sub>	<i>V</i> <sub>3</sub>	V <sub>4</sub>	<b>V</b> <sub>5</sub>	<b>V</b> <sub>6</sub>	<b>V</b> <sub>7</sub>	V <sub>8</sub>	<b>V</b> 9
游戏	H.265	-0.3616	-20.31	0.8632	3.305	11.816	1.82	-1.3	-0.033	1.44
	H.264	-0.3616	-14.21	0.8632	3.305	11.816	1.82	-1.3	-0.033	1.44
	VP9	-0.3616	-17.26	0.8632	3.305	11.816	1.82	-1.3	-0.033	1.44
视频	H.265	-0.3616	-20.31	0.8632	3.305	11.816	1.82	-1.39	-0.06	1.408
	H.264	-0.3616	-14.21	0.8632	3.305	11.816	1.82	-1.39	-0.06	1.408
	VP9	-0.3616	-17.26	0.8632	3.305	11.816	1.82	-1.39	-0.06	1.408

画面质量 QP公式具有以下特性:模型参数覆盖范围广,包括了未来 VR HMD 和全景视频水平的像素密度、码率和帧率;兼顾摄像机拍摄视频及计算机渲染游戏;编码码率增加,画面质量逐渐增大;画面质量随 PPD 增加而增大且当大于一定 PPD 后变化平缓。

#### 6.2.1.2 视频质量 Q√

视频质量  $Q_V$ 与画面质量  $Q_P$ 、VR HMD 的水平视场角之间的关系可描述如下:

$$Q_{V} = \max\left(\min\left(v_{10}Q_{P} + v_{11}FoV_{h} + v_{12}, 5\right), 1\right)$$
 (11)

式中, V<sub>10</sub>~V<sub>12</sub>为模型系数, 其取值分别如下:

	<b>V</b> 10	V11	V <sub>12</sub>
单目视觉	0.595	0.020	-0.735
立体视觉	0.655	0.016	-0.342

## 6.2.1.3 音频质量 QA

当音频采用频率为 48KHz,采样位深为 16,且音频声道数分别为立体声或空间化音频时,音频质量评估模型为:

$$Q_{A} = v_{16} \times \left(1 + v_{13} - \frac{v_{13}}{1 + \left(\frac{Br_{a}}{v_{14}}\right)^{v_{15}}}\right) + v_{17}$$
(12)

式中, V13~V17为模型系数, 其取值分别如下:

	V <sub>13</sub>	V <sub>14</sub>	V <sub>15</sub>	V <sub>16</sub>	V <sub>17</sub>
立体声	4	47.100	2.134	0.81	0.3

空间化音频	4.2	42	1.25	0.96	0.04

#### 6.2.1.4 沉浸体验质量 QIME

沉浸体验质量兼顾了下述三个方面:视频质量,音频质量,以及音视频同步的影响作用, 其评估模型为:

$$Q_{ImE} = \max \left( \min \left( \max \left( v_{18} Q_V + v_{19} Q_A + v_{20} Q_V Q_A + v_{21}, 1 \right), 5 \right) \cdot \min \left( v_{22} \cdot \exp \left( v_{23} \cdot T_{asyn} \right) + v_{24}, 1 \right), 1 \right)$$
(13)

式中, V18~ V24为模型系数, 其取值分别如下:

V <sub>18</sub>	V <sub>19</sub>	V <sub>20</sub>	V <sub>21</sub>	V <sub>22</sub>	V <sub>23</sub>	V <sub>24</sub>
0.9534	0.1954	-0.01747	-0.3466	1.156	-3.72	0.141

## 6.2.2 呈现体验质量 QPE

序号	变量	物理意义	计算所需变量
2.a	Tr	单次 VR 视频观看或 VR 游戏过程中卡顿的总时长,包括初始缓冲时长	4.2.1
2.b	RF	单次VR视频观看或VR游戏过程中卡顿发生的频率(即卡顿次数/观看时长)	4.2.2
2.c	PLR	应用数据丢失率	4.2.3
2.d	P <sub>black</sub>	一次 VR 业务体验过程中因用户头部转动而当前 画面未就绪是导致的黑边在视口内的最大占比	4.2.4

$$Q_{PE} = f_5(T_r, RF, PLR, P_{black})$$
(14)

## 6.2.2.1 视听连续性 Qd

视听连续性损伤主要表现为卡顿和因头部转动中当前画面未就绪而加入的黑边。视听连续性  $Q_c$  的公式为:

$$Q_{C} = \max \left( \left( v_{42} \times \ln \left( T_{r} + v_{43} \right) + v_{44} \right) \times \left( v_{45} \times \ln \left( RF + v_{46} \right) + v_{47} \right) \times \left( v_{48} \times \exp \left( v_{49} \times P_{black}^{v_{50}} \right) + v_{51} \right), 1 \right)$$

$$\tag{15}$$

$$RF = \frac{N_r + \mu \times N_{initial}}{T} \tag{16}$$

$$T_r = \frac{\mu \times t_{initial} + \sum_{i=1}^{n} t_i}{N_{initial} + N_r}$$
(17)

其中, $N_r$ 为卡顿次数; T为 VR 视频/游戏单次体验的总时长;  $t_i$ 为第 i 次卡顿的时长;

 $t_{initial}$ 为初始缓冲时长(不存在初始缓冲时该值为 0),当存在初始缓冲时  $N_{initial}$ =1,当不存在 初始缓冲时  $N_{initial}$ =0。 $P_{black}$ 为终端异步渲染导致的黑边在视场角内的占比。此外, $v_{42} \sim v_{47}$  为模型系数,其取值分别如下:

V <sub>42</sub>	V <sub>43</sub>	V <sub>44</sub>	V <sub>45</sub>	V <sub>46</sub>	V <sub>47</sub>	V <sub>48</sub>	V <sub>49</sub>	V <sub>50</sub>	V <sub>51</sub>
-0.3707	0.1408	1.842	-0.4741	1.565	2.167	-0.4	0.4231	0.3267	1.4

#### 6.2.2.2 视听完整性 Q/

视听完整性损伤包括因使用 UDP 传输丢包产生画面破损(花屏)和因头部转动中当前 画面未就绪而加入黑边所造成的体验损伤。

当基于 UDP 传输且不使用 FEC 时,视听完整性 Q,的公式为:

$$Q_{I} = \max\left(\left(v_{52} \times \exp\left(-\frac{PLR}{v_{53}}\right) + v_{54}\right) \times \left(v_{48} \times \exp\left(v_{49} \times P_{black}^{v_{50}}\right) + v_{51}\right), 1\right)$$
 (18)

其中,*PLR* 是数据丢失率(如丢失率为 0.3%则公式中为 0.3)、*P<sub>black</sub>* 表示一次体验过程中 因用户头部转动而当前画面未就绪时导致的黑边在视口内的占比。

当使用 FEC 时,视听完整性 Qi的公式则为:

$$Q_{I} = \max \left( \left( v_{55} \times \exp(-v_{56} \times FL_{FEC}) + v_{57} \right) \times \left( v_{48} \times \exp(v_{49} \times P_{black}^{v_{50}}) + v_{51} \right), 1 \right)$$
 (18')

其中, $FL_{FEC}$ 是 FEC 纠正失败率。此时视频编码码率  $Br_v$ 的计算需要考虑 FEC 的数据占比,即有:

$$Br_{v}' = Br_{v} \times (1 - R_{FEC}) \tag{19}$$

式中, $Br_{v}$ 为除去 FEC 冗余后实际的视频码率。 $R_{FEC}$ 为 FEC 的冗余数据占比。需要提醒的是,当使用 FEC 时需要将公式(8)中的  $Br_{v}$ 替换为  $Br_{v}$ '。

此外, V52~V57为模型系数, 其取值分别如下:

<b>V</b> 5	2	V53	<b>V</b> 54	<b>V</b> 55	V56	<b>V</b> 57
3.8	в с	.7788	1.2	3.98	0.33	1.02

#### 6.2.2.3 呈现体验质量 QPE

当传输协议/算法基于无丢包掩盖机制时,呈现体验损伤可能同时包括卡顿、花屏和黑边(此时同时存在  $Q_c$ 和  $Q_l$ 子体验项),也可能仅包括花屏和黑边(此时仅存在  $Q_l$ 子体验项,

典型情况是华为云 VR,系基于 UDP+FEC 传输协议);当传输协议/算法基于有丢包掩盖机制时,呈现体验损伤仅包括卡顿和黑边(此时仅存在 Qc 子体验项,典型情况是世博云 VR,系基于 TCP 传输协议)。

视听连续性的实时评分可以根据公式(15)计算:

$$Q_C \quad STInstant = Q_C \tag{20}$$

视听完整性的实时评分可以根据公式(18)或公式(18)计算:

$$Q_{I} = BLInstant = Q_{I}$$
 (21)

呈现体验的会话评分是根据当前实时评分和上一时刻的会话评分,通过α滤波函数计算 当前会话评分,α滤波系数受实时采样片段时长和已播放 VR 视频/VR 游戏时长影响。

当存在 Qc子体验项时,具体公式如下:

$$Q_{C} \_STSession_{n} = \alpha \times Q_{C} \_STSession_{n-1} + (1-\alpha) \times Q_{C} \_STInstant$$
 (22)

 $Q_{C} \_STSession = \alpha_{1} \times \exp(\beta_{1} \times StallingRatio) + \alpha_{2} \times \exp(\beta_{2} \times StallingRatio)$  (23)

其中,这里采用一种变通处理方法,将视听一致性  $Q_c$  映射到 0 和 1 之间,并替换掉公式(23) 中的 *StallingRatio* 变量。从而有:

$$Q_{C} = STSession = \alpha_{1} \times \exp\left(\beta_{1} \times \left(\frac{Q_{C}}{5}\right)\right) + \alpha_{2} \times \exp\left(\beta_{2} \times \left(\frac{Q_{C}}{5}\right)\right)$$
(24)

当存在 Q<sub>1</sub>子体验项时,具体公式如下:

$$Q_{I} = BLSession_{n} = \alpha \times Q_{I} = BLSession_{n-1} + (1 - \alpha) \times Q_{I} = BLInstant$$
 (25)

$$Q_1$$
\_BLSession =  $\alpha_3 \times \exp(\beta_3 \times BlockingRatio) + \alpha_4 \times \exp(\beta_4 \times BlockingRatio)$  (26)

考虑到 VR 视频/游戏的花屏比例(BlockingRatio)很难测量,一般可以基于丢包率、FEC 纠错失败率等参数进行估算。具体做法是,将视听完整性  $Q_i$ 映射到 0 和 1 之间,并替换公式(26)中的 BlockingRatio 变量即可。从而有:

$$Q_{i} \_BLSession = \alpha_{3} \times \exp\left(\beta_{3} \times \left(\frac{Q_{i}}{5}\right)\right) + \alpha_{4} \times \exp\left(\beta_{4} \times \left(\frac{Q_{i}}{5}\right)\right)$$
(27)

最后,呈现体验质量评估模型为:

$$Q_{PE} = \begin{cases} Q_{C} & TCP \\ Q_{I} & UDP \text{ without FEC} \\ Q_{I} & UDP \text{ with FEC} \end{cases}$$

$$(28)$$

注:

HMD 在本地采用会导致黑边(或拖影)的 ATW 或 ASW 机制,本质上来说是 HMD 对 MTP 时延高,画面帧率不足问题采取的体验优化措施。所以,在用户体验评估时,若考虑黑边对用户体验的影响,则不考虑 MTP 对用户体验的影响。

## 6.2.3 交互体验质量 QINE

序号	变量	物理意义	计算所需变量
3.a	MTP <sub>h</sub>	头部转动与画面刷新之间的滞后时长	4.3.1
3.b	MTS <sub>h</sub>	头部转动与音频方位变化之间的滞后时长	4.3.2
3.c	MTP <sub>b</sub>	肢体运动与画面中肢体移动之间的滞后时长	4.3.3
3.d	Tor	用户操作指令与画面中操作响应之间的滞后时长	4.3.4
3.e	DOF	对象在空间内移动的方式	4.3.5

$$Q_{InE} = f_5(DOF, MTP_h, MTP_b, T_{or})$$
 (29)

## 6.2.3.1 VR 视频的交互体验质量 Q<sub>InE</sub>

对于 VR 视频,交互体验质量评估模型为:

$$Q_{hE} = \max\left(\min\left(v_{25} \times DOF + v_{26} - DMOS_{hm}, 5\right), 1\right)$$
 (30)

式中,DMOShm为头部 MTP 导致的交互一致性下降量

DMOS<sub>hm</sub> 与头部 MTP 的关系可以表示为:

$$DMOS_{hm} = \max\left(\min\left(v_{31} \times \ln\left(v_{32} \times MTP_h + v_{33}\right) + v_{34}, 4\right), 0\right)$$
 (31)

式中, V25~ V26 及 V31~ V34 为模型系数, 其取值分别如下:

	V <sub>25</sub>	V <sub>26</sub>	V <sub>31</sub>	V <sub>32</sub>	V <sub>33</sub>	V <sub>34</sub>
4	0.0667	4.3	1.563	0.046	0.01	0.058

# 6.2.3.2 VR 游戏的交互体验质量 Q<sub>InE</sub>

对于 VR 游戏,交互体验质量评估模型为:

$$Q_{InE} = \max\left(\min\left(\min\left(v_{27} \times \ln(DOF) + v_{28}, 5\right) - DMOS_{m}, 5\right), 1\right)$$
(30')

$$DMOS_{m} = \min \left( \max \left( DMOS_{hm}, DMOS_{om}, DMOS_{bm} \right) + v_{29} \times \frac{DMOS_{hm} \times DMOS_{om} \times DMOS_{bm}}{DMOS_{hm} + DMOS_{om} + DMOS_{bm} + v_{30}} \right) + v_{29} \times \frac{DMOS_{hm} \times DMOS_{om} \times DMOS_{bm}}{DMOS_{hm} + DMOS_{om} + DMOS_{bm} + v_{30}} \right) + v_{29} \times \frac{DMOS_{hm} \times DMOS_{om} \times DMOS_{bm}}{DMOS_{hm} + DMOS_{om} + DMOS_{bm} + v_{30}}$$

式中,*DMOS<sub>om</sub>* 为操作响应时延导致的交互一致性下降量; *DMOS<sub>bm</sub>* 为肢体 MTP 导致的交互一致性下降量; *DOF* 根据游戏自由度取值(7,10 或 13)。

DMOSom与操作响应时延的关系可以表示为:

$$DMOS_{om} = \max\left(\min\left(v_{39} \times \ln\left(T_{or} + v_{40}\right) + v_{41}, 4\right), 0\right)$$
(32)

DMOS<sub>bm</sub>与肢体 MTP 的关系可以表示为:

$$DMOS_{bm} = \max\left(\min\left(v_{35} \times \ln\left(v_{36} \times MTP_b + v_{37}\right) + v_{38}, 4\right), 0\right)$$
(33)

式中, V27~ V30 及 V35~ V41 为模型系数, 其取值分别如下:

V <sub>27</sub>	V <sub>28</sub>	V <sub>29</sub>	<b>V</b> 30	<b>V</b> 35	<b>V</b> 36	<b>V</b> 37	<b>V</b> 38	<b>V</b> 39	V40	V41
1.1	1.6	0.98	0.001	1.443	0.018	0.01	0.119	1.343	-7.905	-5.02

# 7 VR 体验评估模型的典型值

通过将模型中其他入参设置为较优值的方法,可比较直观地观察特定输入参对相应体验评估模块的影响。

1. PPD 对画面质量 QP的影响的典型值

在基于 H.264 编码,BPP 为 0.1,帧率为 120fps 的条件下,PPD 对画面质量  $Q_P$  的影响典型值,如图 12 所示。可以看到,PPD 为 15 时,画面质量  $Q_P$  约为 3.28 分;PPD 为 20 时,画面质量  $Q_P$  约为 3.70 分;PPD 为 60 时,画面质量  $Q_P$  约为 4.52,PPD 大于 60 以后画面质量  $Q_P$  增长变得平缓。

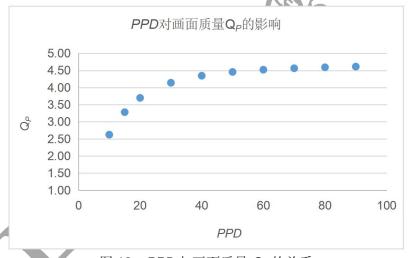


图 12 PPD 与画面质量 Qp的关系

2. VR HMD 单目分辨率对画面质量 QP的影响的典型值

在基于 H.264 编码,BPP 为 0.1,帧率为 120fps,FOV 为 110°的条件下,VR HMD 单目分辨率对画面质量  $Q_P$ 的影响典型值,如图 13 所示。可以看到,VR HMD 单目分辨率 1.5K 对应的画面质量  $Q_P$ 约为 3.13 分;VR HMD 单目分辨率 2K(目前商用高性能 VR HMD 可达的分辨率)对应的画面质量  $Q_P$ 约为 3.57 分。

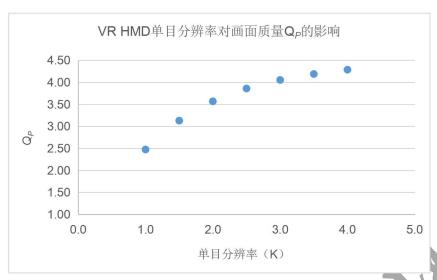


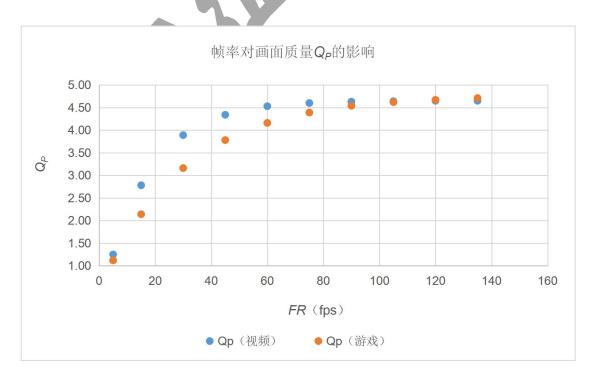
图 13 VR HMD 单目分辨率与画面质量 Qp 的关系

3. 帧率对画面质量 QP的影响的典型值

在 H.264 编码,BPP 为 0.1,PPD 为 120 的条件下,帧率对画面质量  $\mathbb{Q}_P$  的影响典型值,如图 14 所示。

1) 游戏场景: 帧率为 30fps 时,画面质量  $Q_P$ 约为 3.16 分;帧率为 60fps 时,画面质量  $Q_P$ 约为 4.16 分;帧率为 90fps 时,画面质量  $Q_P$ 超过 4.54 分,帧率超过 90fps 后,画面质量 的增长趋势趋于平缓。

2) 视频场景: 帧率为 30fps 时,画面质量  $Q_P$ 约为 3.89 分; 帧率为 60fps 时,画面质量  $Q_P$ 约为 4.53 分; 帧率为 90fps 时,画面质量  $Q_P$ 达到 4.63 分。



#### 图 14 帧率与画面质量 Qp的关系

#### 4. 头部 MTP 对交互响应质量的影响的典型值

头部 MTP 时延对头部动作的交互响应质量  $Q_{IR}$ 的影响的典型值如图 15 所示。当头部 MTP 时延小于等于 20ms 时, $Q_{IR}$ 不受损伤;当头部 MTP 时延等于 50ms 时, $Q_{IR}$ 为 3.63分;当时延为 100ms 时, $Q_{IR}$ 为 2.55分;当时延为 200ms 时, $Q_{IR}$ 下降为 1.47分。

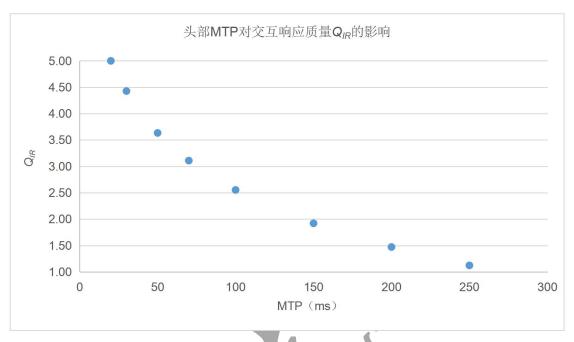


图 15 头部 MTP 与交互响应质量的关系

#### 5. UDP 丢包率对呈现体验质量的影响典型值

VR 视频数据基于 UDP 传输,且 VR HMD 不产生异步时间扭曲(ATW)中间帧时,UDP 包丢失率对呈现体验质量  $Q_{PE}$ 的影响的典型值,如图 16 所示。当 UDP 丢包率为 0.2%时, $Q_{PE}$ 为 4.14 分;当 UDP 丢包率为 0.5%时, $Q_{PE}$ 为 3.20 分;当 UDP 丢包率为 1%时, $Q_{PE}$ 为 2.25 分。

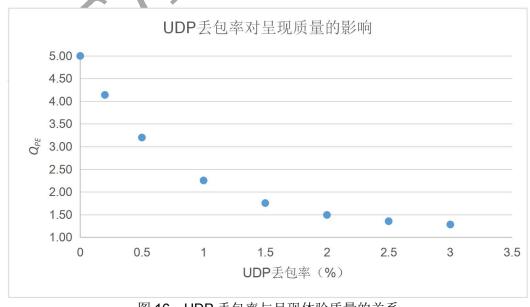


图 16 UDP 丢包率与呈现体验质量的关系

注: 若内容数据基于 TCP 传输,从带宽需求满足的角度,会有更高的丢包率控制要求。针对特定的数据传输带宽要求,不同的 TCP 传输控制算法,有不同的丢包率要求。

#### 6. 不同分辨率全景视频片源的画面质量评估

在 H.264 编码,*BPP* 为 0.1 (典型画质),帧率为 30fps,全景视频为 3D,左右眼视频数据压缩时无相互参考,整体码率为单目视频码率的 2 倍的条件下,不同分辨率的全景视频片源对应的 *PPD*、码率及画面质量,如表 9 所示,8K 全景视频的画面质量仅处在可接受水平。

 114347/11422/4/3										
全景水平分辨率	视域内等效 PPD	码率(Mbps)	画面质量 Q <sub>P</sub>							
1920	5.33	10.55	1.49							
2560	7.11	18.75	1.77							
3840	10.67	42.19	2.28							
7680	21.33	168.75	3.17							

表 9 不同分辨率全景视频体验得分

#### 7. 不同单目分辨率 VR HMD 在不同网络状况下的体验对照表 /

基于如下条件,不同单目分辨率的 VR HMD,在不同网络状况下的游戏业务体验对照,如表 10 所示。

- 1) 基于 H.264 编码, BPP 为 0.21 (此时画质较优)。
- 2) 水平视域 FOV: 110°。
- 3) 视频自由度: 头部 6DOF+手部 6DOF+手柄操作 1DOF
- 4) 音频编码算法: AAC-LC, 音频固定码率: 140Kbps。
- 5) 不考虑连续/完整性影响, VR HMD 设备无 ATW 机制。
- 6) 单目分辨率水平垂直像素数比: 9:10, 左右图像数据无相互参考压缩。

说明: 网络带宽需求通常为: 平均码率\*1.5。

表 10 不同单目分辨率 VR HMD 在不同网络状况下的体验对照

单目 视水 分率 Rh	PPD	帧率 <i>FR</i> (fps)	码率 (Mbps)	画面 质量 Q <sub>P</sub>	头部 MTP 时延 <i>MTP<sub>h</sub></i> (ms)	操作响 应时延 <i>T<sub>or</sub></i> (ms)	肢体 MTP 时延 <i>MTP</i> ₅ (ms)	交互体 验质量 Q <sub>InE</sub>	VR_MOS	
		60	60.08	3.05	20	50	50	4.42	3.04	
1.5K	13.64	13.64	7 00	00.00	.00 3.05	30	70	70	3.76	2.56
1.51		an	90.12	3.33	20	50	50	4.42	3.17	
		30.12	0.00	30	70	70	3.76	2.66		
		18.18	60	106.81	3.48	20	50	50	4.42	3.25
2K			J. <del>-1</del> U	30	70	70	3.76	2.72		
211	10.10		160.22	3.79	20	50	50	4.42	3.40	
			3.79	30	70	70	3.76	2.83		
		60 240.33	3.95	20	50	50	4.42	3.47		
3K	27.27		240.00	0.90	30	70	70	3.76	2.89	
		90	360.49	4.31	20	50	50	4.42	3.64	

					30	70	70	3.76	3.02
4K	36.36	60	427.25	4.18	20	50	50	4.42	3.58
					30	70	70	3.76	2.98
		90	640.87	4.56	20	50	50	4.42	3.76
					30	70	70	3.76	3.11
6K	54.55	60	60 961.30	61.30 4.38	20	50	50	4.42	3.68
		00			30	70	70	3.76	3.05
		90 1441.96 4.77	1441.06	1.00 4.77	20	50	50	4.42	3.86
			30	70	70	3.76	3.19		

# 8 问题和展望

考虑到 VR 业务体验评估模型不仅要适用于当前现状评估,而且还要对未来一段时间 VR 业务的发展,具有前瞻性指导意义,VEA 体验联盟的领域专家在业界首创和实现了 VR 设备的高 PPD 等效测试、多自由度交互体验质量测试等方案,并将相关测试结果融合到评估模型中。

但是,由于现阶段受限于主观测试数据规模,模型准确性或有待进一步提高,此外,随着应用场景的丰富和变化,指标测量技术的发展,评估模型也需要不断更新迭代,因此,后续仍然需要 VEA 体验联盟和产业伙伴通力合作,共同推进 VR 体验评估的标准化进程。

为了加强产业合作,评估模型算法以免费 SDK 形式在 VEA 体验联盟网站发布,实现共享,欢迎 VR 产业界的伙伴们踊跃使用。当然,为了确保产业伙伴能够正确集成 SDK 或直接引用模型算法,VEA 体验联盟将对相关评测工具进行认证和授权,产业伙伴可在联盟网站发起相关申请。在确保评估模型被正确使用后,产业伙伴的测试数据则可以作为模型优化的支撑数据,从而形成升级、迭代、应用的良性循环。

下载本白皮书、模型 SDK 或了解模型的认证授权等信息,可扫描下面二维码,访问 VEA 体验联盟官方网站。



# 9 参考文献

[1] 中国信息通信研究院,"视频服务用户体验评估算法和参数-点播/直播",链接: http://www.gxkjcg.cn/standard/html/?140.html

[2] 华为 X Labs, "云 VR 临场感指数 Cloud VR PI(π)",链接:

https://www.huawei.com/en/industry-insights/outlook/mobile-broadband/xlabs/insights-whitepapers/cloud-vr-presence-index

[3] ITU, "ITU-R BT.1788 主观质量评价标准中文版",链接:

 $\underline{https://www.itu.int/dms\_pubrec/itu-r/rec/bt/R-REC-BT.1788-0-200701-I!!PDF-C.pdf}$ 

[4] ITU, "ITU-R BT.500-13 建议书", 链接:

https://www.itu.int/dms\_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-I!!PDF-C.pdf

[5] Song W , Cheng D , Deng Z , et al. Design and assessment of a wide FOV and high-resolution optical tiled head-mounted display[J]. Applied Optics, 2015, 54(28):E15.

[6] W. Song, Y. Wang, D. Cheng, and Y. Liu, "Light field head-mounted display with correct focus cue using micro structure array," Chin. Opt. Lett. 12(6), 060010 (2014).

