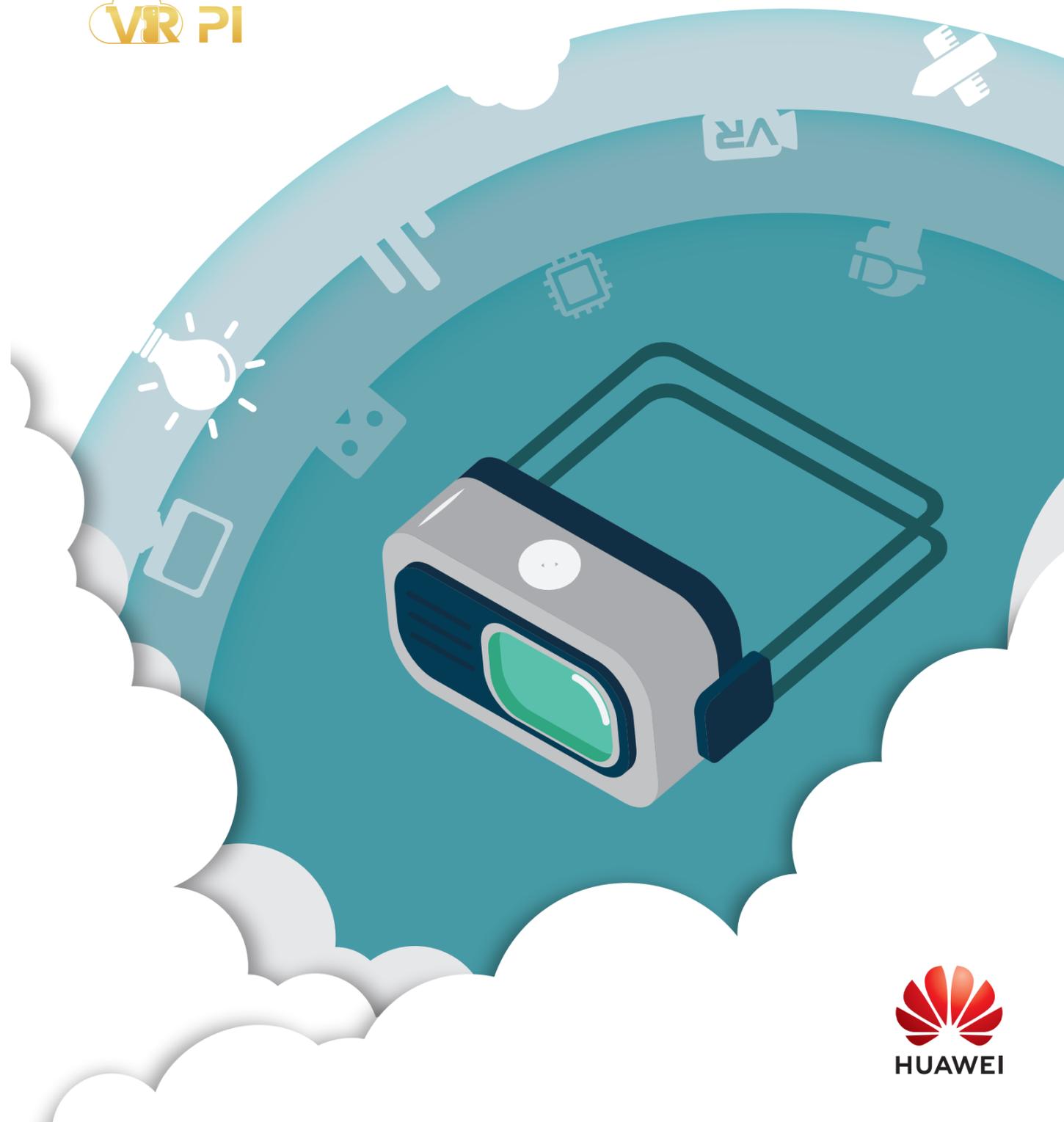


华为技术有限公司
深圳龙岗区坂田华为基地
电话: +86 755 28780808
邮编: 518129
www.huawei.com

云VR临场感指数 Cloud VR PI (π)

Cloud X业务体验模型系列



Wireless X Labs是一个全新的平台,旨在将电信运营商、技术供应商和垂直行业的合作伙伴聚集在一起,探索未来的移动应用场景,推动业务和技术创新,建立一个开放的生态系统。Wireless X Labs建立了三个实验室,探索三个主要领域:人与人之间的连接,垂直领域的应用和家庭应用。

商标声明

HUAWEI, HUAWEI 是华为技术有限公司商标或者注册商标,在本手册中以及本手册描述的产品中,出现的其它商标,产品名称,服务名称以及公司名称,由其各自的所有人拥有。

免责声明

本文档可能含有预测信息,包括但不限于有关未来的财务、运营、产品系列、新技术等信息。由于实践中存在很多不确定因素,可能导致实际结果与预测信息有很大的差别。因此,本文档信息仅供参考,不构成任何要约或承诺,华为不对您在本文档基础上做出的任何行为承担责任。华为可能不经通知修改上述信息,恕不另行通知。

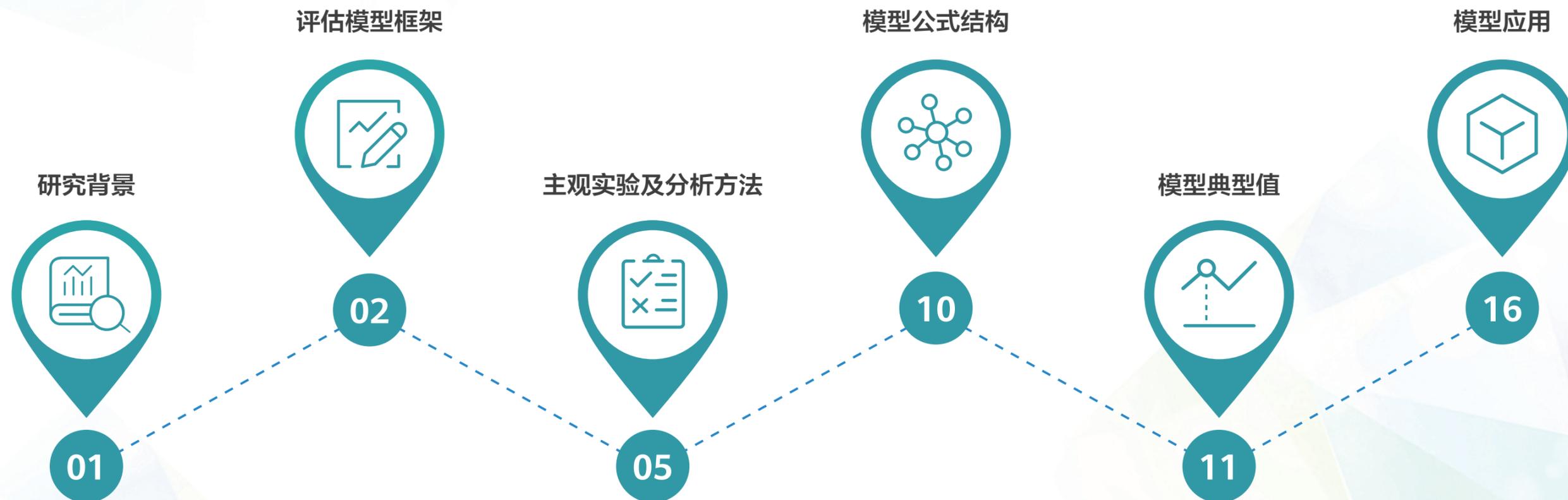
版权所有© 华为技术有限公司 2019。保留一切权利。

非经华为技术有限公司书面同意,任何单位和个人不得擅自摘抄、复制本手册内容的部分或全部,并不得以任何形式传播。



CONTENTS

目录



执行摘要:

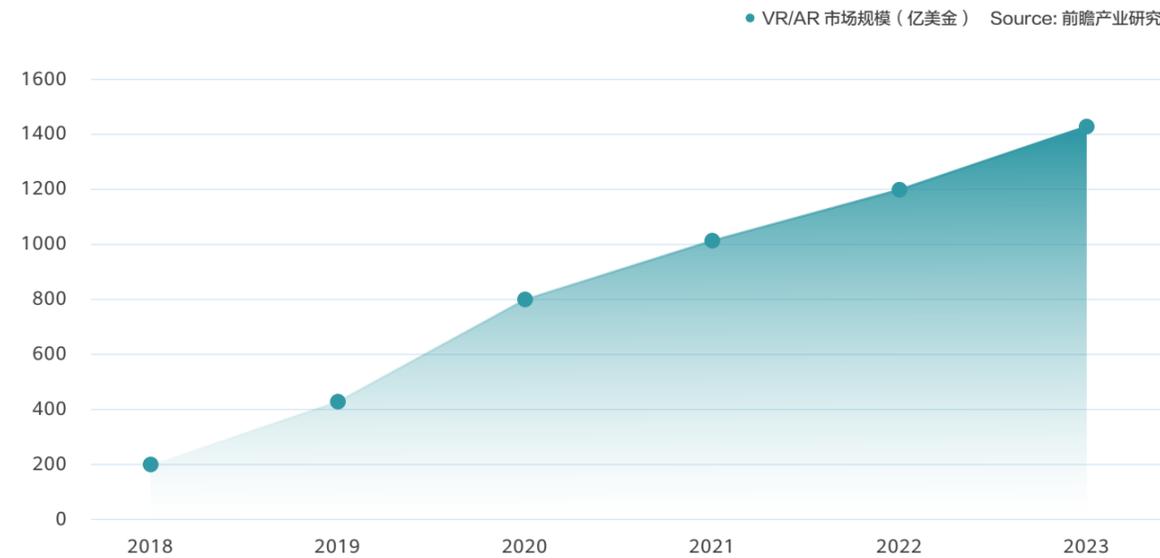
本文描述了云 VR (Cloud VR) 业务体验评估的需求背景, 评估模型框架, 主观实验和分析方法, 模型公式结构, 模型典型值和模型应用等。其中 VR 设备的高 PPD 等效测试, 云 VR 多自由度交互真实感测试等系业界首创, 从而使得 VR 业务体验评估模型不仅适用于当前现状评估, 而且对于未来 VR 业务的发展具有前瞻性指导。值得一提的是, 本评估模型算法将以免费 SDK 的形式发布和共享, 欢迎 VR 产业界的伙伴们积极试用并共同促进其发展完善。



RESEARCH BACKGROUND 研究背景

近年来，虚拟现实技术VR (Virtual Reality)，已经逐步进入到教育、娱乐、医疗、环保、交通运输、公共卫生等众多领域，具有极大的应用价值和商业潜力。相关调查机构公布的数据也表明，VR业务在未来若干年，将迎来一个快速发展期。

图1-全球VR/AR市场规模预测



同时，可以看到当前算力、存储云化的产业发展步伐越来越快，VR可以借助云端算力进行实时画面渲染。进一步地，5G网络的超大带宽和超低延时，保证了云端渲染画面的可获得性，并在随时随地接入、类本地操作方面，给用户体验带来极大提升。Cloud VR业务作为Cloud X系列业务之一，是需求、技术、成本和体验等方面综合推动、演进的必然业务形态。

然而，针对Cloud VR业务，目前业界还没有明确的体验评估模型，无法从用户体验出发，有效牵引E2E产业发展，并作为5G网络建设的参考依据。可以说，如何准确有效地评估和预测用户在应用Cloud VR业务时的体验是亟待解决的问题。

为此，华为X Labs联合国内高校知名领域专家，系统研究Cloud VR相关体验影响因素，构建评估模型框架，并通过符合ITU规范的主观实验和数据训练，建立起体验评估模型。模型的不同评估模块，可以满足运营商、设备制造商、内容提供商在不同层面、不同角度的体验评估需求。

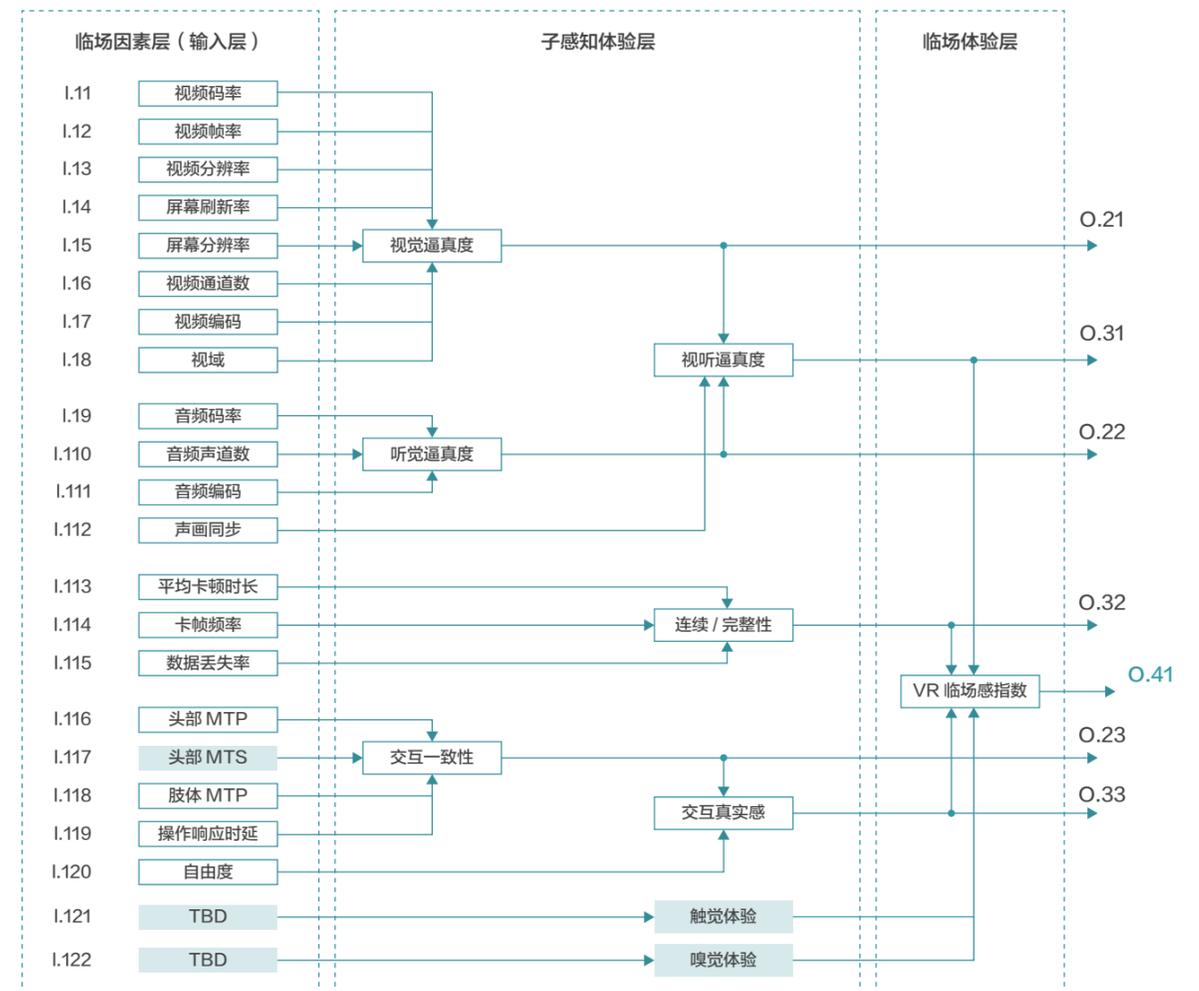


EVALUATION MODEL FRAMEWORK 评估模型框架

相比传统的视频业务，VR能够凭借其可自由切换的视角、丰富的动作交互等特性，为用户提供全新的业务体验。研究人员通常使用临场感 (Presence) 这一概念来表征用户沉浸于虚拟环境，身临其境的感受。

评估模型框架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，使测试环节测试量可控，并有利于分析、建立评估模型。具体评估模型框架如图 2所示。其中，临场因素层为模型输入层，包含了主观实验系统中可提取、量化的因素。子感知体验层则包含视听逼真度、连续/完整性和交互真实感3个方面。

图2-VR业务用户体验评估模型框图



模型输入参数及相应缩写如表 1 所示:

表1-评估模型输入参数

ID	参数	描述	缩写
I.11	视频码率	视频的平均码率 (比特每秒, bps)	Br
I.12	视频帧率	每秒内的视频帧数 (fps)	FR
I.13	视频分辨率	视频水平和垂直方向的像素点个数	R_h, R_v
I.14	屏幕分辨率	单目屏幕水平方向的像素点个数	R_{sh}
I.15	屏幕刷新率	屏幕每秒钟刷新次数	RR
I.16	视频通道数	单目视频 (1) 或立体视频 (2)	ST
I.17	视频编码	H.265/HEVC, H.264/AVC, VP9	<i>Video Codec</i>
I.18	视域	水平方向单目视场角	FoV_h
I.19	音频码率	音频的平均码率 (千比特每秒, Kbps)	ABr
I.110	音频声道数	立体声音 (2), 空间化声音 (8)	SP
I.111	音频编码	AAC-LC, Opus	<i>Audio Codec</i>
I.112	声画同步	声画同步时延 (秒, s)	t_{syn}
I.113	平均卡顿时长	单次观看过程中卡顿的总时长, 包括初始缓冲时长 (秒, s)	T_r
I.114	卡顿频率	单次观看过程中卡顿发生的频率 (即卡顿次数 / 观看时长)	RF
I.115	数据丢失率	应用层数据丢失率 (%)	ppl
I.116	头部 MTP	头部转动与画面刷新之间的滞后时长 (ms)	t_{hd}
I.117	头部 MTS	头部转动与音频方位变化之间的滞后时长 (ms)	t_{ad}
I.118	肢体 MTP	肢体运动与画面中肢体移动之间的滞后时长 (ms)	t_{bd}
I.119	操作响应时延	用户操作指令与画面中操作响应之间的滞后时长 (ms)	t_{od}
I.120	自由度	虚拟现实系统可操作的维度	DOF

评估模型输出及相应缩写, 如表 2 所示:

表2-评估模型输出

ID	描述	值域	缩写	全拼
O.21	视觉逼真度	1~5 分	MOS_{VF}	<i>Visual Fidelity</i>
O.22	听觉逼真度	1~5 分	MOS_{AF}	<i>Acoustic Fidelity</i>
O.23	交互一致性	1~5 分	MOS_{IC}	<i>Interaction Consistency</i>
O.31	视听逼真度	1~5 分	MOS_{AVF}	<i>Acoustic-visual Fidelity</i>
O.32	连续 / 完整性	1~5 分	MOS_{RP}	<i>Rebuffering / Packet Loss</i>
O.33	交互真实感	1~5 分	MOS_{IR}	<i>Interaction Realism</i>
O.41	VR 临场感指数	1~5 分	$VR PI$	<i>Virtual Reality Presence Index</i>





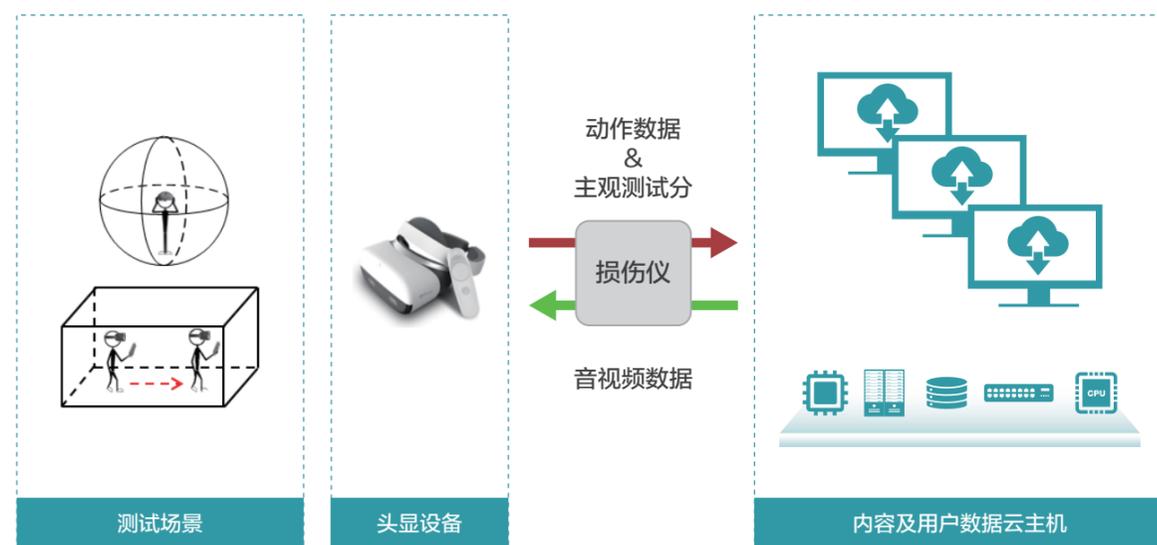
SUBJECTIVE EXPERIMENTAL METHODS 主观实验及分析方法

主观实验平台

VR体验主观实验平台可定量控制临场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3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头显分发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头戴设备组成，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3-虚拟现实主观实验平台的基本构成



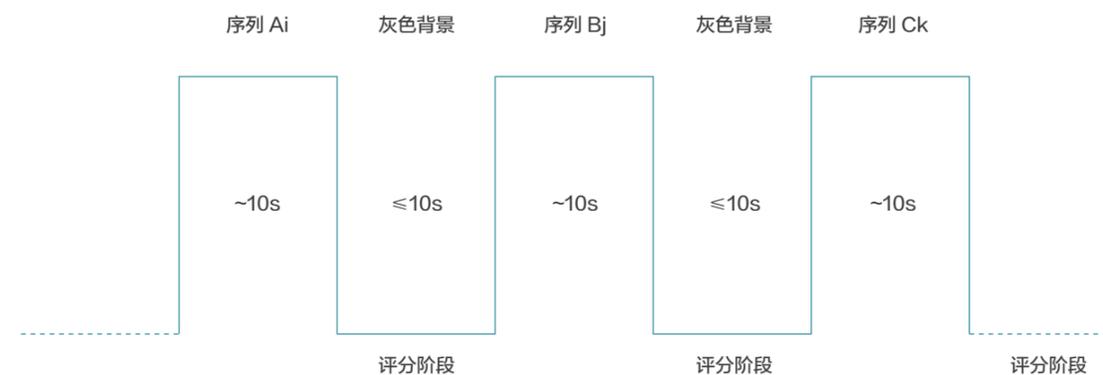
主观实验平台涉及的软硬件主要包括：1) HTC Vive Pro；2) HTC Vive；3) Pico Neo；4) 2K显示器（刷新率144Hz）；5) 4K显示器（刷新率90Hz）；6) 高性能云主机；7) VR播放器；8) VR打分软件；9) 6 DOF头显及手柄的测试与控制软件。

说明：1) *PPD*：Pixels Per Degree，角分辨率或空间分辨率，指视角中的平均每 1° 夹角内填充的像素点的数量。对于VR头显，*PPD*数值越大，就说明对细节的显示越精细，用户对显示画面的感受就越清晰；2) 在目前商用VR头显支持的*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*和帧率对视频质量的影响。

主观打分方法

根据 ITU-T P.913，使用单刺激法（Single Stimulus Methods, SSM）进行主观实验。视频质量测试序列的播放采用随机不重复播放的方式，测试人员在播放序列的间隔进行评价，如图4所示。在头戴设备上的评分中，采用的策略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，来播放下一个序列。

图4-SSM主观实验序列及评分过程示意图



测试人员评分采用 P.913 标准中建议的 ACR (Absolute Category Rating) 5 分制评分模式，各分值的含义如图5所示。

图5-ACR 5分制评分尺度

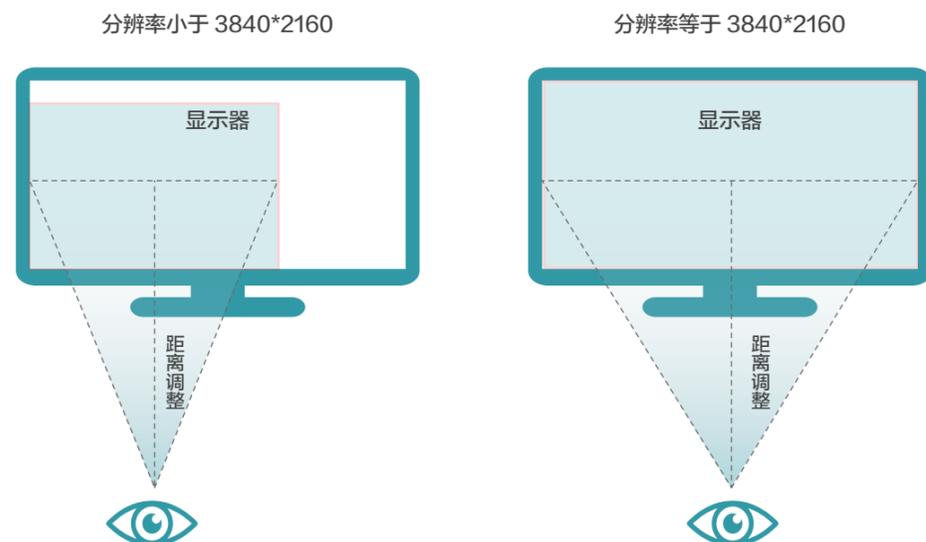


关键测试与分析方法说明

1. 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*，也即平均每度视角内像素点的数量，作为像素密度的度量。假设头显视域为 110° ，单眼分辨率 1K 至 10K，才能构建 10 至 90 *PPD* 的测试场景，但目前商用 VR 头显单眼分辨率最高仅为 2K，支持的 *PPD* 较低。为了全面获取 *PPD* 对视频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，测试数据和基于头显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 6 所示。

图6-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 3 所示。

表3-32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

2. 游戏多自由度交互真实感测试

测试以快速击破气球为任务要求，开发构建如下测试场景，用于测试用户的交互真实感：

- (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 4 定义了自由度场景编号，作为评估模型输入。

表4-自由度场景编号

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



主观实验测试数据示例

图 7 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 $BPP_{压后} \approx 0.21$)，在帧率为 30fps 时，不同 PPD 下的视频质量评分结果。

从图中可以看出，主观测试人员所感知的视频质量随像素密度 PPD 的增高而增高，且在 PPD 达到 60 以后，视频质量受 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) 的视频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7-PPD与视频质量的关系



注：1) BPP为每像素编码比特数 (bit)；2) H.264量化参数QP是量化步长Qstep的序号。对于亮度 (Luma) 编码而言，量化步长Qstep共有52个值，QP取值0~51。QP取最小值0时，表示量化最精细；相反，QP取最大值51时，表示量化最粗糙；3) $BPP_{压后} = \text{压缩前每像素比特位} / \text{压缩率}$ 。

MODEL FORMULA STRUCTURE 模型公式结构

VR 临场感指数评估模型,公式结构如下:

$$VR PI = \min(\max((MOS_{AVF} - 1) \cdot (1 - v_{53}(5 - MOS_{IR}) - v_{54}(5 - MOS_{RP})) + 1, 1), 5) \quad 1$$

式中，VR PI是临场感指数， MOS_{AVF} 是视听逼真度， MOS_{IR} 为交互真实感， MOS_{RP} 为视听连续/完整性， v_{53} 和 v_{54} 为模型系数。VR临场感以视听逼真度 MOS_{AVF} 为基础，交互真实感 MOS_{IR} 和连续/完整性 MOS_{RP} 则作为体验损伤因素，如交互时延、UDP传输丢包率的增加，会导致体验受损。

子体验项和相关入参的关系，可见下面表达式说明：

$$MOS_V = f_1(\text{Video Codec}, Br, FR, (R_h, R_v), ST) \quad 2$$

$$MOS_{VF} = f_2(MOS_V, FoV_h) \quad 3$$

$$MOS_{AF} = f_3(\text{Audio Codec}, ABr, SP) \quad 4$$

$$MOS_{AVF} = f_4(MOS_{VF}, MOS_{AF}, t_{svn}) \quad 5$$

$$MOS_{IR} = f_5(DOF, t_{hd}, t_{bd}, t_{od}) \quad 6$$

$$MOS_{RP} = f_6(T, RF, ppl) \quad 7$$



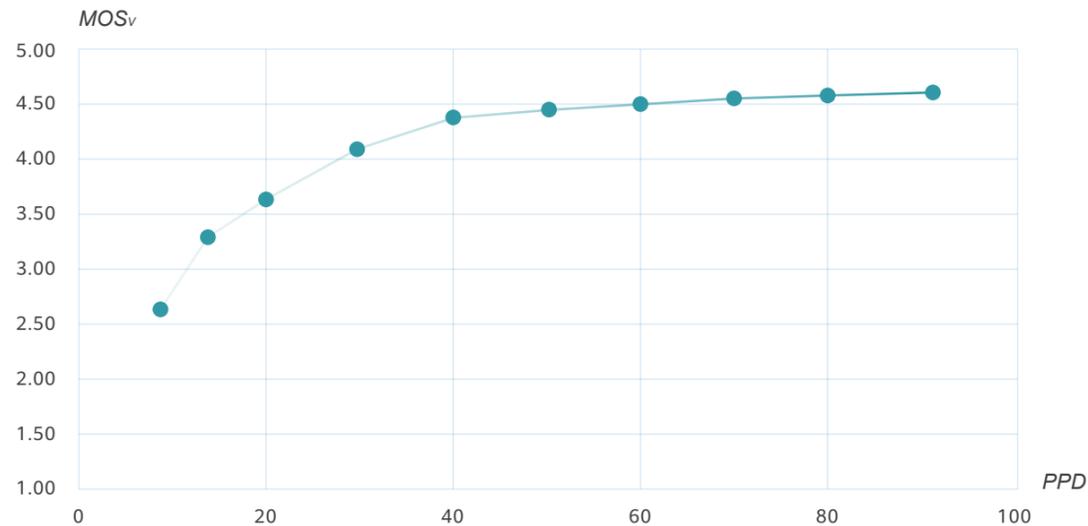
MODEL TYPICAL VALUE 模型典型值

我们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入参对相应体验评估模块的影响。

1. PPD 对视频质量 MOS_v 的影响典型值

基于 H.264 编码, BPP 为 0.1, 帧率 120fps 的条件, PPD 对视频质量 MOS_v 的影响典型值, 如图 8 所示。可以看到, PPD 为 15 时, 视频质量 MOS_v 约为 3.28 分; PPD 为 20 时, 视频质量 MOS_v 约为 3.67 分; PPD 为 60 时, 视频质量 MOS_v 约为 4.51 分, 且 PPD 大于 60 以后视频质量 MOS_v 增长变得平缓。

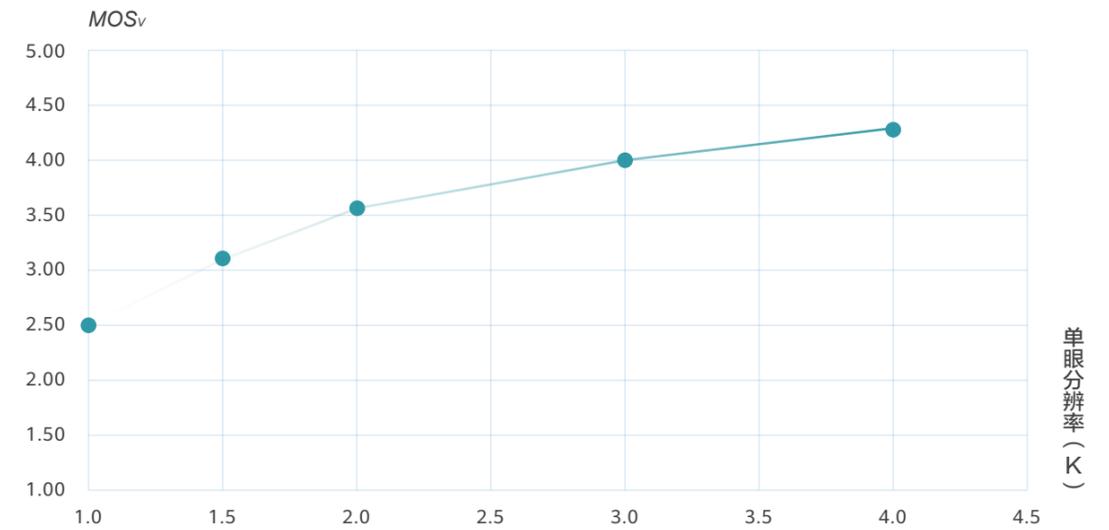
图8-PPD与视频质量 (MOS_v) 的关系



2. 单眼分辨率对视频质量 MOS_v 的影响典型值

基于 H.264 编码, BPP 为 0.1, 帧率 120fps, FOV 为 110° 的条件, 单眼分辨率对视频质量 MOS_v 的影响典型值, 如图 9 所示。可以看到, 单眼分辨率 1.5K 对应的视频质量 MOS_v 约为 3.13 分; 单眼分辨率 2K (目前商用高性能头显可达的分辨率) 对应的视频质量 MOS_v 约为 3.55 分。

图9-单眼分辨率与视频质量 (MOS_v) 的关系



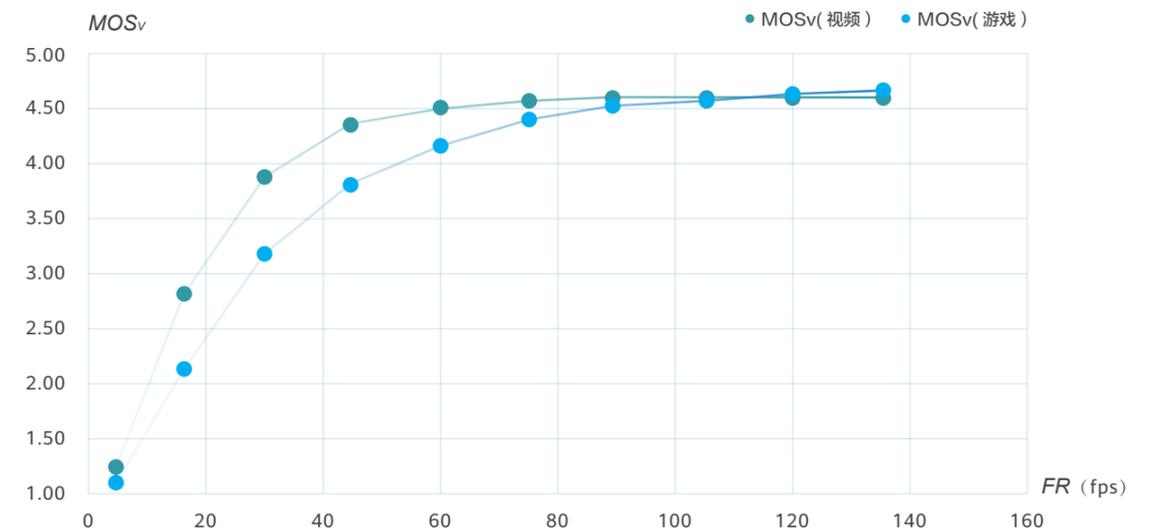
3. 帧率对视频质量 MOS_v 的影响典型值

在 H.264 编码, BPP 为 0.1, PPD 为 120 的条件下, 帧率对视频质量 MOS_v 的影响典型值, 如图 10 所示。

1) 游戏场景: 帧率为 30fps 时, 视频质量 MOS_v 约为 3.15 分; 帧率为 60fps 时, 视频质量 MOS_v 约为 4.15 分; 帧率为 90fps 时, 视频质量 MOS_v 超过 4.52 分, 帧率超过 90fps 后, 视频质量的增长趋势趋于平缓。

2) 视频场景: 帧率为 30fps 时, 视频质量 MOS_v 约为 3.88 分; 帧率为 60fps 时, 视频质量 MOS_v 约为 4.51 分; 帧率为 90fps 时, 视频质量 MOS_v 达到 4.61 分。

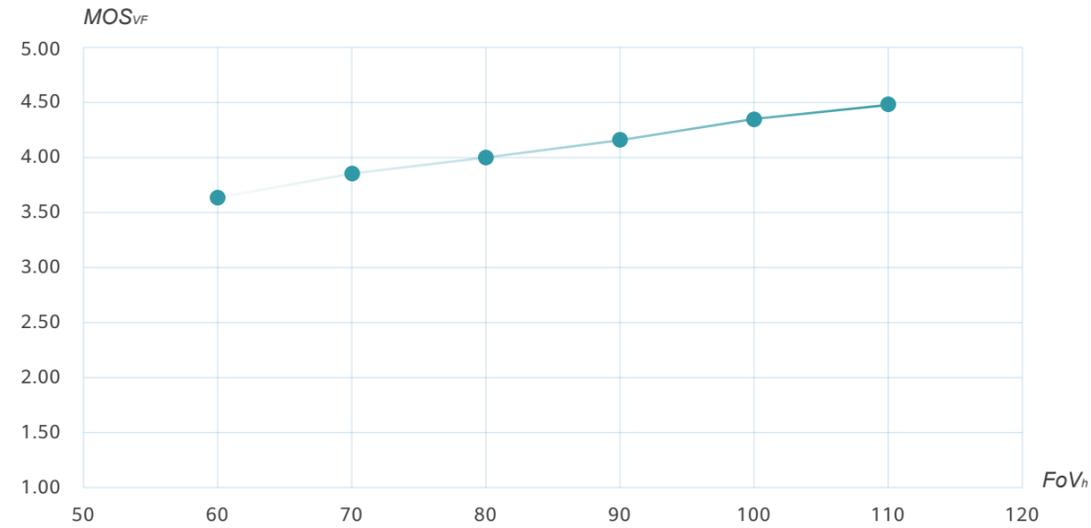
图10-帧率 (FR) 与视频质量 (MOS_v) 的关系



4. FOV对视觉逼真度 MOS_{VF} 的影响典型值

在视频质量的基础上，视觉逼真度主要受视域的影响。在H.264编码，BPP为0.1，帧率120fps，PPD为120的条件下，水平视域FOV（双眼均具备相应视域）对视觉逼真度 MOS_{VF} 影响的典型值如图11所示。可以看到，在水平视域 60° 到 110° 的范围内（目前可支持的测试范围），视觉逼真度 MOS_{VF} 呈线性增长。

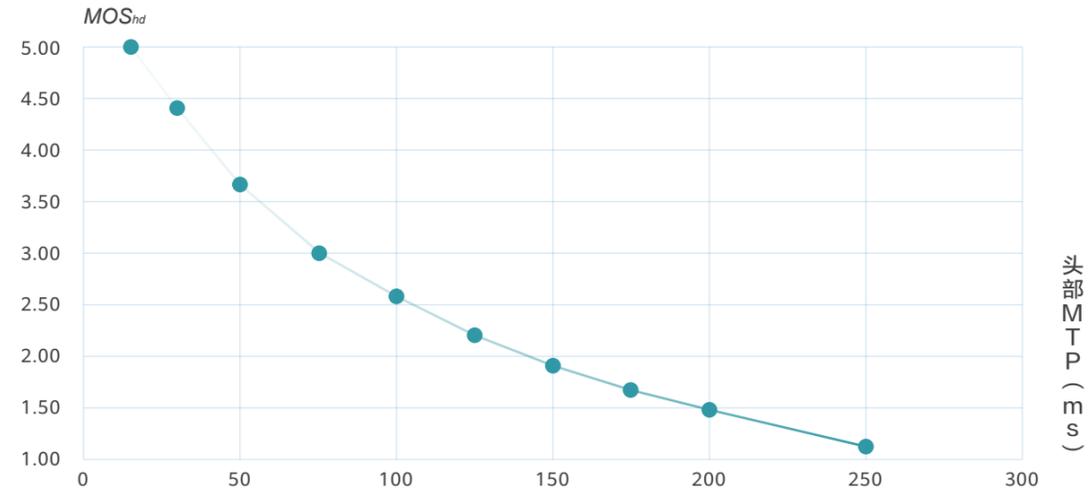
图11-视域 (FoV_h) 与视觉逼真度 (MOS_{VF}) 的关系



5. 头部 MTP 对交互一致性体验 MOS_{hd} 的影响典型值

头部 MTP 时延对头部动作的交互一致性体验 MOS_{hd} 的影响典型值如图12所示。当头部 MTP 时延小于等于 20ms 时， MOS_{hd} 不受损伤；当头部 MTP 时延等于 50ms 时， MOS_{hd} 为 3.63 分；当时延为 100ms 时， MOS_{hd} 为 2.55 分；当时延为 200ms 时， MOS_{hd} 下降为 1.47 分。

图12-头部MTP时延 (t_{hd}) 与头部交互一致性 (MOS_{hd}) 的关系

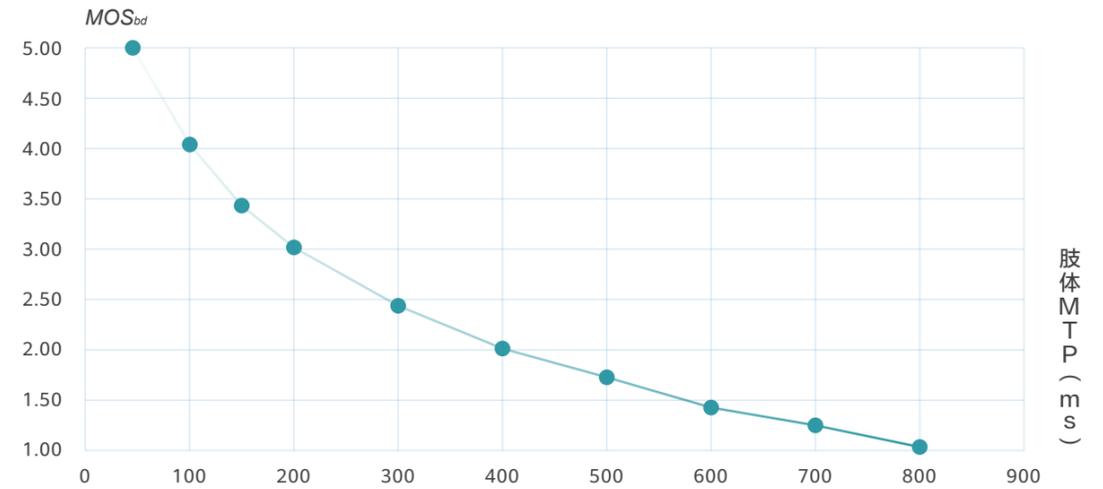


6. 肢体 MTP 对交互一致性体验 MOS_{bd} 的影响典型值

肢体MTP时延对肢体动作交互一致性体验 MOS_{bd} 的影响典型值如图13所示。当肢体MTP时延小于等于50ms时， MOS_{bd} 不受损伤；当肢体MTP时延等于100ms时， MOS_{bd} 为4.02分；当肢体MTP时延为300ms时， MOS_{bd} 为2.44分；当肢体MTP时延为500ms时， MOS_{bd} 下降为1.71分。

注：若肢体动作导致头部运动，如腿部的移动导致头部移动，则按头部 MTP 进行体验评估。

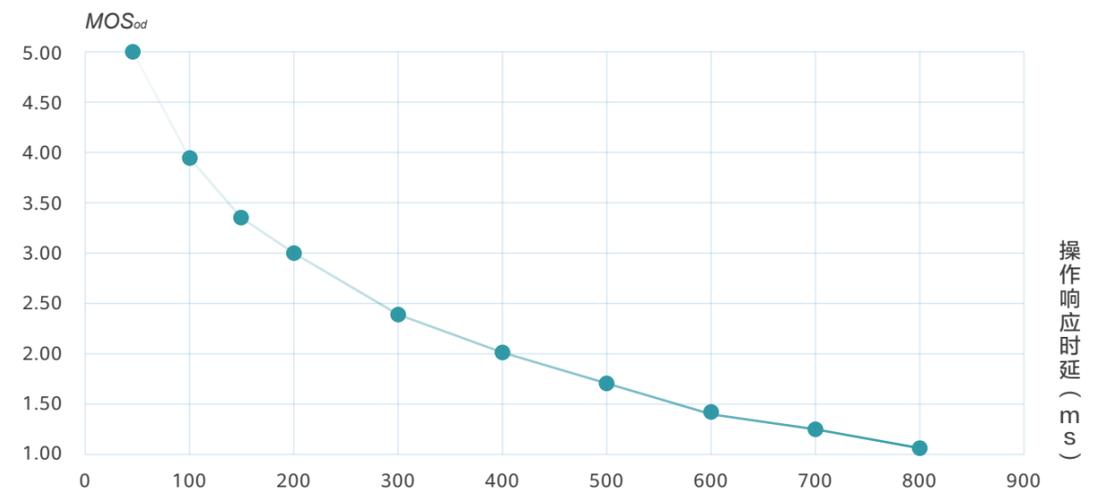
图13-肢体MTP时延 (t_{bd}) 与肢体交互一致性 (MOS_{bd}) 的关系



7. 操作响应时延对交互一致性体验 MOS_{od} 的影响典型值

操作响应时延对操作交互一致性体验 MOS_{od} 的影响典型值，如图14所示。当操作响应时延小于等于 50ms 时， MOS_{od} 不受损伤；当操作响应时延等于 100ms 时， MOS_{od} 为 3.95 分；当操作响应时延为 300ms 时， MOS_{od} 为 2.40 分；当操作响应时延为 500ms 时， MOS_{od} 下降为 1.70 分。

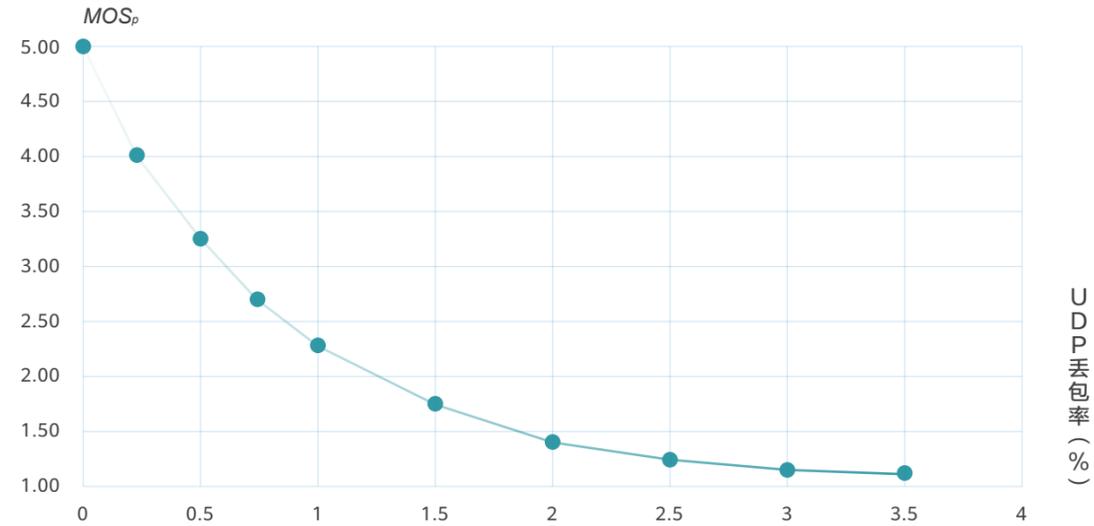
图14-操作响应时延 (t_{od}) 与操作交互一致性 (MOS_{od}) 的关系



8.UDP 丢包率对视听完整性 MOS_P 的影响典型值

VR视频数据基于UDP传输时，UDP包丢失率对视听完整性 MOS_P 的影响典型值如图 15所示。当UDP丢包率为0.25%时， MOS_P 为3.98分；当UDP丢包率为0.5%时， MOS_P 为3.23分；当UDP丢包率为1%时， MOS_P 为2.25分。

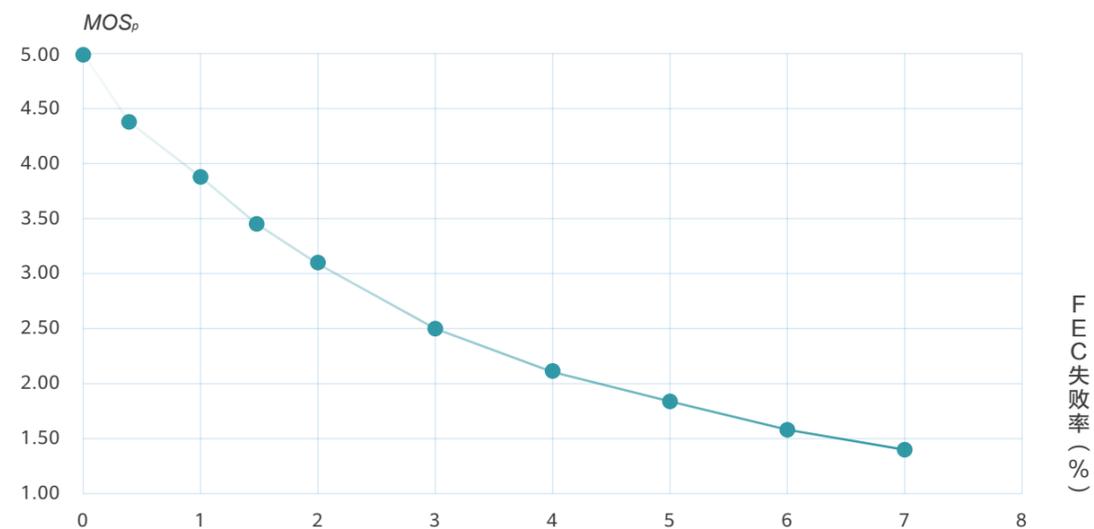
图15-UDP包丢失率 ($pp\%$) 与视听完整性 (MOS_P) 的关系



9.FEC 失败率对视听完整性 MOS_P 的影响典型值

视频数据基于 UDP+FEC 模式传输时，FEC 失败率对视听完整性 MOS_P 的影响典型值，如图 16 所示。FEC 失败率为 0.5% 时， MOS_P 为 4.39 分；FEC 失败率为 1% 时， MOS_P 为 3.88 分；FEC 失败率为 2% 时， MOS_P 为 3.08 分。

图16-FEC失败率与视听完整性 (MOS_P) 的关系



MODEL APPLICATION 模型应用

1. 不同分辨率全景视频片源的视频质量 MOS_V 评估

在 H.264 编码， BPP 为 0.1， FOV 为 110° ，帧率 30fps，全景视频为 3D，左右眼视频数据压缩时无相互参考，整体码率为单眼视频码率的 2 倍的条件下，不同分辨率的全景视频片源对应的 PPD 、码率及视频质量 MOS_V ，如表 5 所示，8K 全景视频的视频质量 MOS_V 仅处在可接受水平。

表5-不同分辨率全景视频体验得分

全景水平分辨率	视域内等效 PPD	码率 (Mbps)	视频质量 (MOS_V)
1920	5.33	10.55	1.54
2560	7.11	18.75	1.81
3840	10.67	42.19	2.29
7680	21.33	168.75	3.13

2. 主流商用 VR 头显最佳体验评估

在 H.264 编码， BPP 为 0.1，无连续 / 完整性体验损伤的条件下，目前商用的高性能 VR 头显可达到的视频质量和 VR 临场感指数典型值，如表 6 所示。

表6-不同分辨率全景视频体验得分

设备	性能规格	业务类型	视频质量 (MOS_V)	临场感指数 ($VR PI$)
HTC vive Pro	单眼分辨率 1440*1660， FOV 为 110° ，刷新率 90Hz，头 6 DOF + 手 6 DOF	视频	3.04	3.15
		游戏	2.98	3.07
Pico Neo 商业版	单眼分辨率 1440*1660， FOV 为 101° ，刷新率 90Hz，头 6 DOF + 手 6 DOF	视频	3.17	3.1
		游戏	3.11	3.02
小米 VR 一体机	单眼分辨率 1280*1440， FOV 为 100° ，刷新率 72Hz，头 3 DOF + 手 3 DOF	视频	2.98	2.99
		游戏	2.83	2.5
小米 VR 一体机 + NOLO	单眼分辨率 1280*1440， FOV 为 100° ，刷新率 72Hz，头 6 DOF + 手 6 DOF	视频	2.98	2.99
		游戏	2.83	2.88

3. 不同单眼分辨率头显在不同网络状况下的体验对照表

基于如下条件，不同单眼分辨率的头显，在不同网络状况下的游戏业务体验对照，如表 7 所示。

- 1) 基于 H.264 编码, BPP 为 0.21。
- 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。
- 6) 单眼分辨率水平垂直像素数比: 9:10。

表7-不同单眼分辨率不同网络状况体验对照表

单眼视频 水平分辨率 R_h	PPD	帧率 (fps)	平均码率 (Mbps)	视频质量 MOS_V	视听逼真度 MOS_{AVF}	头部 MTP 时延 t_{hd} (ms)	操作响应 时延 t_{od} (ms)	肢体 MTP 时延 t_{bd} (ms)	交互真实感 MOS_{IR}	临场感指数 $VR PI$
1K	9.09	60	25	2.43	3.10	20	50	50	4.42	2.80
						20	70	70	3.90	2.52
						20	100	100	3.37	2.24
						30	50	50	3.85	2.50
						30	70	70	3.76	2.45
						30	100	100	3.14	2.13
						50	50	50	3.05	2.08
						50	70	70	2.91	2.01
						50	100	100	2.64	1.86
		90	38	2.65	3.23	20	50	50	4.42	2.91
						20	70	70	3.90	2.61
						20	100	100	3.37	2.32
						30	50	50	3.85	2.59
						30	70	70	3.76	2.54
						30	100	100	3.14	2.19
						50	50	50	3.05	2.15
						50	70	70	2.91	2.07
						50	100	100	2.64	1.91
1.5K	13.64	60	57	3.03	3.45	20	50	50	4.42	3.09
						20	70	70	3.90	2.77
						20	100	100	3.37	2.45
						30	50	50	3.85	2.74
						30	70	70	3.76	2.69
						30	100	100	3.14	2.31
						50	50	50	3.05	2.26
						50	70	70	2.91	2.17
						50	100	100	2.64	2.00

1.5K	13.64	90	86	3.30	3.61	20	50	50	4.42	3.23
						20	70	70	3.90	2.89
						20	100	100	3.37	2.54
						30	50	50	3.85	2.86
						30	70	70	3.76	2.80
						30	100	100	3.14	2.39
						50	50	50	3.05	2.34
						50	70	70	2.91	2.25
2K	18.18	60	102	3.43	3.68	20	50	50	4.42	3.29
						20	70	70	3.90	2.94
						20	100	100	3.37	2.59
						30	50	50	3.85	2.91
						30	70	70	3.76	2.85
						30	100	100	3.14	2.44
						50	50	50	3.05	2.38
						50	70	70	2.91	2.28
						50	100	100	2.64	2.10
		90	153	3.74	3.86	20	50	50	4.42	3.44
						20	70	70	3.90	3.07
						20	100	100	3.37	2.69
						30	50	50	3.85	3.04
						30	70	70	3.76	2.97
						30	100	100	3.14	2.53
						50	50	50	3.05	2.47
						50	70	70	2.91	2.37
						50	100	100	2.64	2.17
3K	27.27	60	229	3.89	3.95	20	50	50	4.42	3.52
						20	70	70	3.90	3.13
						20	100	100	3.37	2.74
						30	50	50	3.85	3.10
						30	70	70	3.76	3.03
						30	100	100	3.14	2.58
						50	50	50	3.05	2.51
						50	70	70	2.91	2.41
						50	100	100	2.64	2.21
		90	343	4.24	4.15	20	50	50	4.42	3.69
						20	70	70	3.90	3.28
						20	100	100	3.37	2.86
						30	50	50	3.85	3.24
						30	70	70	3.76	3.17
						30	100	100	3.14	2.68
						50	50	50	3.05	2.62
						50	70	70	2.91	2.51
						50	100	100	2.64	2.29

4K	36.36	60	407	4.12	4.08	20	50	50	4.42	3.63
						20	70	70	3.90	3.23
						20	100	100	3.37	2.82
						30	50	50	3.85	3.19
						30	70	70	3.76	3.12
						30	100	100	3.14	2.65
						50	50	50	3.05	2.58
						50	70	70	2.91	2.47
						50	100	100	2.64	2.26
	90	610	4.49	4.29	20	50	50	4.42	3.81	
					20	70	70	3.90	3.38	
					20	100	100	3.37	2.95	
					30	50	50	3.85	3.34	
					30	70	70	3.76	3.27	
					30	100	100	3.14	2.76	
					50	50	50	3.05	2.69	
					50	70	70	2.91	2.57	
					50	100	100	2.64	2.35	
6K	54.55	60	916	4.33	4.20	20	50	50	4.42	3.73
						20	70	70	3.90	3.32
						20	100	100	3.37	2.89
						30	50	50	3.85	3.28
						30	70	70	3.76	3.21
						30	100	100	3.14	2.71
						50	50	50	3.05	2.64
						50	70	70	2.91	2.53
						50	100	100	2.64	2.31
	90	1442	4.73	4.43	20	50	50	4.42	3.93	
					20	70	70	3.90	3.49	
					20	100	100	3.37	3.03	
					30	50	50	3.85	3.44	
					30	70	70	3.76	3.37	
					30	100	100	3.14	2.84	
					50	50	50	3.05	2.76	
					50	70	70	2.91	2.64	
					50	100	100	2.64	2.41	

在基于 5G 网络传输，并借助动作预测技术，无交互一致性体验损伤设想下，根据上表，可以观察当前的 VR 产业发展现状，并着眼用户体验进一步提升，规划未来的产业发展路径：

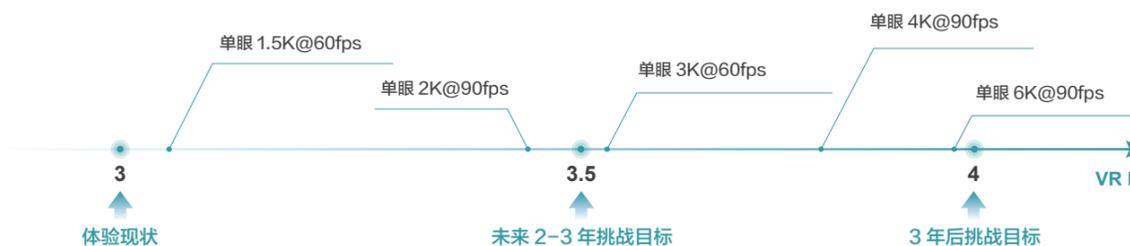
1) 综合考虑当前主流头显和内容的呈现能力，以单眼分辨率 1.5K，帧率 60fps 计算，目前可达到的临场感指数为 3.09 分，仅达到入门级“可接受”水平。

2) 临场感指数达到 3.5 分，可作为未来 2-3 年的体验提升目标，这需要单眼分辨率 2K，帧率 90fps。或者，单眼分辨率 3K，帧率 60fps。

3) 临场感指数达到 4 分，用户体验达到“好”的水平，则可作为 3 年后的挑战目标，这需要单眼分辨率 6K，帧率 90fps。而单眼分辨率 4K，帧率 90fps，临场感指数 3.81 分，可作为一个过渡目标。

当前 VR 体验现状与产业发展路径建议，可见下图 17 所示。

图17-VR体验现状及未来产业发展路径



CONCLUSION 结束语

5G 已来，我们相信，本课题的研究成果除了能够对实际用户 Cloud VR 业务体验评估提供重要参考外，也能在规划环节为预测 Cloud VR 业务体验提供有效的理论依据。为了加强产业合作，该评估模型算法以免费 SDK 形式发布和共享，欢迎 VR 产业界的朋友们积极试用，并提出优化意见或建议。

联系方式：XLabs@huawei.com

了解更多模型细节和进行体验预测，可扫描下面二维码，访问 Cloud X 业务体验评估模型介绍网页



本白皮书由华为 X Labs 编写。

由于 5G E2E 产业相关技术快速发展，本白皮书仅做参考，不能作为投资研究决策的依据，不能作为道义的、责任的和法律的依据或者凭证，无论是否已经明示或者暗示。发布方将随时补充、更正和修订有关信息，但不保证及时发布。对于本报告所提供信息所导致的任何直接的或者间接的投资盈亏后果不承担任何责任。

本白皮书版权为华为所有，未经书面许可，任何机构和个人不得以任何形式翻版、复制和发布。如引用发布，需注明出处，且不得对本报告进行有悖原意的引用、删节和修改。

相关问题或建议，请联系：XLabs@huawei.com