

中关村现代信息消费应用产业技术联盟  
团 体 标 准

T/INFOCA 2-2019

---

虚拟现实用户体验评估标准

The evaluation standard for the user experience of VR service

(草案)

2019-10-16 发布

2019-10-18 实施

中关村现代信息消费应用产业技术联盟

发布

---

## 目 次

目 次	I
前 言	III
虚拟现实用户体验评估标准	1
1 范围	1
2 规范性引用文件	1
3 缩略语、术语和定义	1
3.1 缩略语	1
3.2 术语和定义	3
4 VR 业务类型划分	3
5 各类 VR 业务应用场景与用户体验评估需求	5
5.1 VR 直播视频业务	5
5.1.1 应用场景	5
5.1.2 视频体验评估需求	6
5.2 VR 点播视频业务	6
5.2.1 应用场景	6
5.2.2 视频体验评估需求	7
5.3 VR 游戏业务	7
5.3.1 应用场景	7
5.3.2 游戏体验评估需求	8
5.4 交互类 VR 教育业务	9
5.4.1 应用场景	9
5.4.2 VR 教育体验评估需求	9
6 VR 服务体验评估架构	9
6.1 VR 系统简介	9
6.2 关键指标	10
6.3 模型架构	11
7 VR 服务体验指标构成及参数采集	13
7.1 沉浸体验质量输入参数	13
7.2 观看体验质量输入参数	14
7.3 交互体验质量输入参数	14
7.4 VR_MOS 综合模型输出参数	15
7.5 输入参数采集方法	15
7.5.1 客户端采集	16
7.5.2 网络节点采集	18
7.5.3 服务端采集	18
8 VR 服务用户体验评估算法	19
8.1 总体模型综述	19
8.2 模型三大模块综述	19

8.2.1 沉浸体验质量 $Q_{ImE}$ .....	19
8.2.2 观看体验质量 $Q_{PE}$ .....	22
8.2.3 交互体验质量 $Q_{InE}$ .....	25
9 参考文献.....	26

## 前 言

本标准是针对虚拟现实服务用户体验评估的规范，本标准的内容框架如下：

- 一虚拟现实体验评估的需求和场景
- 一虚拟现实用户体验评估的系统架构
- 一虚拟现实用户体验评估算法和参数

本标准参考国内和国际相关标准，并结合国内网络的实际情况制定。

注意本文件的某些内容可能涉及专利，本文件的发布机构不承担识别这些专利的责任。

本标准由中关村现代信息消费应用产业技术联盟提出并归口。

本标准起草单位： 华为技术有限公司

本标准主要起草人：中国信息通信技术研究院、中国电信集团公司、华为技术有限公司、西安电子科技大学、中国联通集团公司、中国移动集团公司、首都师范大学、北京未来媒体科技股份有限公司、终极（北京）科技有限公司、创维集团有限公司、Pico 等

# 虚拟现实用户体验评估标准

## 1 范围

本标准一方面规定了VR服务用户体验评估的需求和场景，定义了VR用户体验评估模型的系统架构；另一方面分析了影响VR服务用户体验的主要因素和参数，定义了VR服务用户体验的评估计算方法。

本标准适用于对电信运营商、广电运营商、VR头戴设备提供商、云VR服务提供商、互联网VR视频服务商、互联网VR游戏服务商和其他相关厂商提供的VR服务（如VR点播/直播视频，VR游戏等）的用户体验质量进行综合评估，还适用于产业各方对影响VR服务质量的因素进行量化和分析。

## 2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

[3GPP TR 26.929 V0.4.0] VR用户体验的QoE参数和指标

[3GPP TR 26.918 V15.2.0] 3GPP上的VR媒体服务

[3GPP TS 26.118 V15.2.0] VR流应用的配置文件

[ITU-T / SG 12 G.QoE-VR based on TD326-R1] 1、识别VR业务及关键技术。2、识别影响VR QoE的关键质量因素。3、定义描述VR QoE质量的度量方法。4、研究VR质量的客观和主观评价方法论。5、输出保证良好VR QoE质量需求的建议。

[ITU-T / SG 12 G.360-VR] 研究具体如何评测360度全景视频的体验，包括主观测试方法，以及客观模型。

[IEEE P 2048.1] 该标准规定了VR与AR设备的种类和定义。

[IEEE P 2048.2] 该标准规定了沉浸式视频种类和质量指标。

[IEEE P 2048.3] 该标准规定了沉浸式视频文件和流媒体的格式，以及此类格式的功能性和交互性。

[IEEE P 2048.4] 该标准规定了在VR中个人身份认证的条件和方式。

[IEEE P 2048.6] 该标准规定了在VR APP中使用沉浸式用户界面的条件和方式及沉浸式用户界面中的功能性和交互性。

[IEEE P 2048.9] 该标准规定了沉浸式音频的分类和质量指标。

[IEEE P 2048.10] 该标准规定了沉浸式音频文件和流媒体的格式，以及此类格式所支持的功能性和交互性。

## 3 缩略语、术语和定义

### 3.1 缩略语

以下缩略语适用于本文件。

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
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
HTTP	超文本传输协议	Hyper Text Transfer Protocol
RTMP	实时消息传输协议	Real Time Messaging 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
CDN	内容分发网络	Content Deliver Network
CR	核心路由器	Core Router
BNG	宽带网络业务网关	Broadband Network Gateway
OLT	光线路终端	Optical Line Terminal
ONT	光网络终端	Optical Network Terminal
gNodeB	5G 无线接入网基站	5G New Radio Base Station
TCO	总拥有成本	Total Cost of Ownship

## 3.2 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。

## 4 VR 业务类型划分

目前VR服务应用场景较多，包含影视、游戏、教育、旅游、营销等。从VR服务面向的对象看，主要分为面向个人消费者和面向行业用户两种情况。面向个人消费者的VR服务包括VR影视、VR游戏、VR直播等；面向行业用户的VR服务包括VR教育、VR医疗、VR旅游、VR工业等。从VR服务的交互特征看，可以分为VR弱交互业务和VR强交互业务。VR弱交互业务以VR视频业务为主，包含巨幕影院、VR 360° 视频、VR直播等，用户可以在一定程度上选择视点和位置，但用户与虚拟环境中的实体不发生实际的交互；VR强交互业务以VR游戏为代表，包含VR游戏、VR教育、VR健身、VR社交等，用户可通过交互设备与虚拟环境进行互动，虚拟空间图像对交互行为作出实时响应。从VR服务的技术特征看，可以分为VR游戏、VR直播、VR视频、VR全景图、VR仿真等。

结合VR服务面向的对象、交互特征和技术特征3个方面的特点，本文将VR业务划分为5个大类，即VR点播视频、VR直播视频、VR游戏、VR强交互应用和VR行业应用，如表1所示。

表 1 VR 服务应用场景与业务类型的对应关系

序号	应用场景	说明	面向对象	技术特征	交互特征	VR 业务分类
1	VR 巨幕影院	为海量片源赋予全新体验	个人消费者	VR 点播视频	弱交互	VR 点播视频
2	VR 直播	最易商业变现的场景，如：新闻发布会、媒体见面会、沙龙等	个人消费者	VR 直播视频	弱交互	VR 直播视频
3	VR 360° 视频	身临其境的观影体验	个人消费者	VR 点播视频	弱交互	VR 点播视频
4	VR 游戏	用户最有付费意愿的场景	个人消费者	VR 游戏	强交互	VR 游戏
5	VR 教育	教育创新方向	行业用户	VR 仿真（或 VR 点播/直播视频）	强交互	VR 行业应用
6	VR 营销	零售营销新模式	个人消费者	VR 全景图	强交互	VR 强交互应用
7	VR 电竞馆	大空间、异地、多人合作的竞技游戏	行业用户	VR 游戏	强交互	VR 行业应用
8	VR 健身	让运动更有乐趣	行业用户	VR 仿真	强交互	VR 行业应用
9	VR 音乐	一场视听双重沉浸的盛宴	个人消费者	VR 仿真	强交互	VR 强交互应用
10	VR K 歌	主角视角的沉浸式 K 歌体验	个人消费者	VR 仿真	强交互	VR 强交互应用
11	VR 医疗	医学训练治疗利器	行业用户	VR 仿真	弱交互	VR 行业应用
12	VR 旅游	影响未来旅游业发展的九大科技趋势之一	行业用户	VR 仿真（或 VR 全景图）	强交互	VR 行业应用
13	VR 房地产	房地产模型视觉呈现新方式	行业用户	VR 仿真（或 VR 全景图）	弱交互	VR 行业应用
14	VR 社交	下一代社交网络平台	个人消费者	VR 全景图	强交互	VR 强交互应用
15	VR 购物	新零售时代网购“第三屏”	个人消费者	VR 全景图	强交互	VR 强交互应用

16	VR 工程	升级工业制造	行业用户	VR 仿真（或 VR 全景图）	弱交互	VR 行业应用
----	-------	--------	------	-----------------	-----	---------

## 5 各类 VR 业务应用场景与用户体验评估需求

本章将简要描述VR服务的主要应用场景和各场景对于用户体验评估的需求。

在VR用户体验评估技术成熟后，与网络质量相关的部分指标（如时延）可进一步分解为对链路上各段网络的要求，对于网络的规划、建设、运维将具有指导意义。

### 5.1 VR 直播视频业务

#### 5.1.1 应用场景

VR 直播视频是娱乐性视频的一种重要表现形式，可以通过网络信号，在线收看体育赛事、重大活动和新闻等，属于实时性强但交互性较弱的业务，用户不能控制观看内容的进度。

与传统 OTT 直播视频相比，VR 直播视频具有以下特点：1）视频采集需要专业的全景视频拍摄设备，且多颗摄像头（一般 6 颗及以上）采集的原始视频需要逐帧进行亮度色彩调整、对齐、畸变矫正、投影到球面等一系列拼接与投影变换处理，对服务器的性能要求极高。2）VR 直播视频的传输带宽是传统 OTT 直播视频的至少 3 倍以上，原因是 VR 原始视频的分辨率主流是 4K（至少是 2K）、帧率一般为 30fps 及以上，均明显高于传统 OTT 直播视频的相应情况。3）需要专门的 VR 观看设备，而传统 OTT 直播视频只需要普通智能手机即可。4）上述多方面因素造成 VR 直播视频的平台成本和用户 TCO 都要比前者高得多。

VR 直播视频系统的架构如图 1 所示。其中从本地 PC 服务器到云平台的 RTMP 推流可以通过 5G/4G 或宽带网络实现，其典型码率为 6~20Mbps（H.265 编码）；而从 CDN 分发节点到用户的直播拉流也可以通过 5G/4G 或宽带网络实现，其典型码率则为 3~10Mbps（H.265 编码）。值得一提的是，终端用户可以通过手机（或 PC）+VR 眼镜，也可以直接通过 VR HMD 方式进行观看。

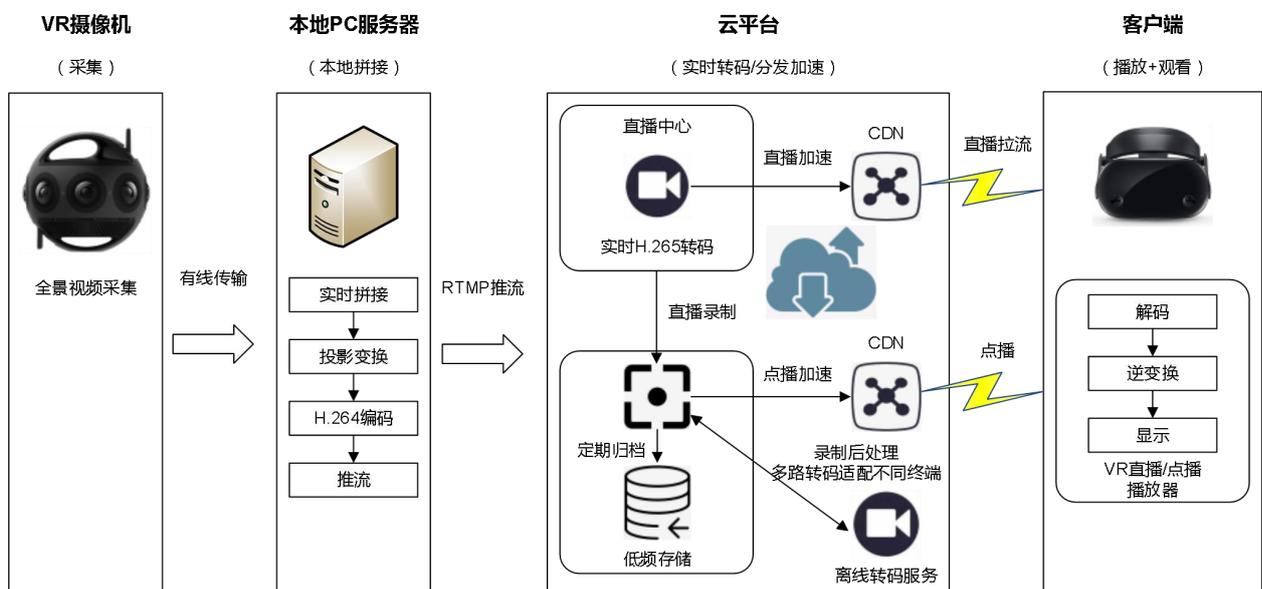


图 1 VR 直播视频架构示意图

### 5.1.2 视频体验评估需求

目前的 VR 直播视频传输方式下，整个传输链条上可能是基于 TCP 或 UDP 传输协议，即可能存在重传或丢包带来的质量劣化（如卡顿或花屏），该情况对其他 VR 视频业务如回看（实际上就是 VR 视频点播业务）也是同样存在的。在 VR 视频传输过程中，需要考虑卡顿或花屏的影响；为了反映各环节的问题，VR 直播视频体验评估需要支持云管端部署，以实现故障定界定位。

用户对 VR 直播视频最直观的体验是画面质量：VR 直播由于是实时流媒体，故当出现丢包时，如被丢弃的报文包含重要帧如 I 帧，则会出现短暂卡顿现象，而当被丢弃的报文仅有 B 帧或 P 帧，则会出现面积和范围不等的花屏现象，这是影响 VR 直播体验的关键因素。

## 5.2 VR 点播视频业务

对于普通用户的娱乐类 VR 视频服务，通常称为 VR 视听点播服务。它是由内容提供商和服务提供商面向广大用户提供的 VR 视频业务，用户可在提供的节目导航系统中选择观看。由 VR 视听点播服务平台提供的内容可以是 VR 全景图或视频信息，媒体资源可被存储在网络的任何位置上。VR 视听点播服务平台整合媒体资源，进行数字版权管理（DRM），并通过电信网络提供导航服务和内容交付。

用户可以在固定或移动的位置如家庭、办公室，或者移动的火车上来使用 VR 视听点播服务。用户可以使用手机（或 PC）+VR 眼镜、VR HMD 设备，通过流媒体通信能力接入宽带网络获得服务。

### 5.2.1 应用场景

VR 点播视频允许用户根据自己的需求选择相应的 VR 视频进行收看，而不是被动地收看某个特定的广播视频。VR 点播视频的实现过程如下：当用户发出 VR 点播视频请求时，VR 视听点播服务平台就会根据点播信息，将存放在片源库中的节目信息检索出来，以流媒体文件通过高速传输网络传送到用户终端。VR 点播视频系统的架构如图 2 所示。

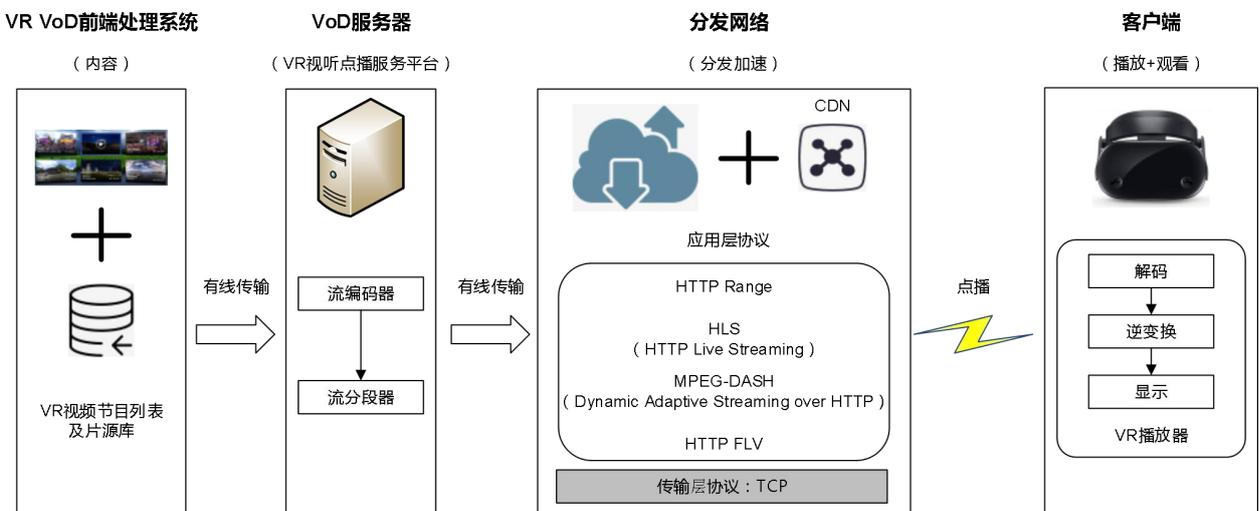


图 2 VR 点播视频架构示意图

VR 点播视频系统结构由 VoD 前端处理系统、VoD 服务器、分发网络和客户端组成。从业务流看，用户在客户端启动播放请求，并通过网络发送给 VoD 服务器；经过请求验证后，VoD 服务器把存储系统中可访问的节目名准备好，使用户可以浏览到所喜爱的节目菜单；用户选择节目后，VoD 服务器从存储子系统中取出节目内容，并传送到客户端进行播放。

### 5.2.2 视频体验评估需求

影响 VR 点播视频体验的因素包括视频源质量、交互体验和观看体验 3 个方面。视频源质量主要取决于分辨率、编码、码率等参数；交互体验即用户的互动过程体验，对于 VR 点播视频，从用户点击“开始”按键到视频开始播放的时延对于用户体验影响较大；至于观看体验，鉴于点播类视频业务一般通过 TCP 协议承载，在 VR 视频播放过程中由于网络性能波动导致视频卡顿，会对用户体验产生较差的影响。因此研究 VR 点播视频的体验评估时需要考虑上述因素。

VR 点播视频在进行体验评估时，还应同时支持加密和非加密两种场景。对于非标准分辨率的视频质量，也能够准确评估；还须考虑用户操作响应对视频体验造成的影响，如快进快退、拖动等。

## 5.3 VR 游戏业务

VR 游戏是指利用 VR 技术让玩家走进虚拟的游戏世界，拥有沉浸的视听感受，并通过身体的运动进行游戏，成为自己梦寐以求的英雄角色。云 VR 游戏，通过传统 VR 游戏结合云计算技术，将内容上云、渲染上云，可以有效降低用户侧对终端的要求，从而降低消费门槛。

云 VR 游戏是 VR 典型应用，交互性强，沉浸感强，是最能吸引用户的业务之一。按照交互特征看，有 3 DOF 的轻量级休闲游戏和 6 DOF 的重度大型游戏，可以分别满足轻度休闲玩家和高端核心玩家的需求。

### 5.3.1 应用场景

云 VR 游戏的系统架构如图 3 所示。云 VR 游戏的系统结构由内容层、平台层、网络层和终端层组成。其中，内容层负责向平台层提供 VR 游戏内容，包含 VR 游戏内容提供方和 VR 游戏内容聚合方；平台层负责对 VR 游戏业务进行渲染，包含实例运行、逻辑计算、实时渲染、实时推流等功能，为每个用户在云端生产一个应用实例，对其 VR 画面进行实时渲染、编码和推流；网络层一般包括骨干网、城域网和接入网，负责为 VR 游戏业务提供大带宽、低时延的稳定传输；终端层主要实现 VR 游戏用户的动作输入、捕获及上报，游戏内容呈现，用户鉴权等功能。

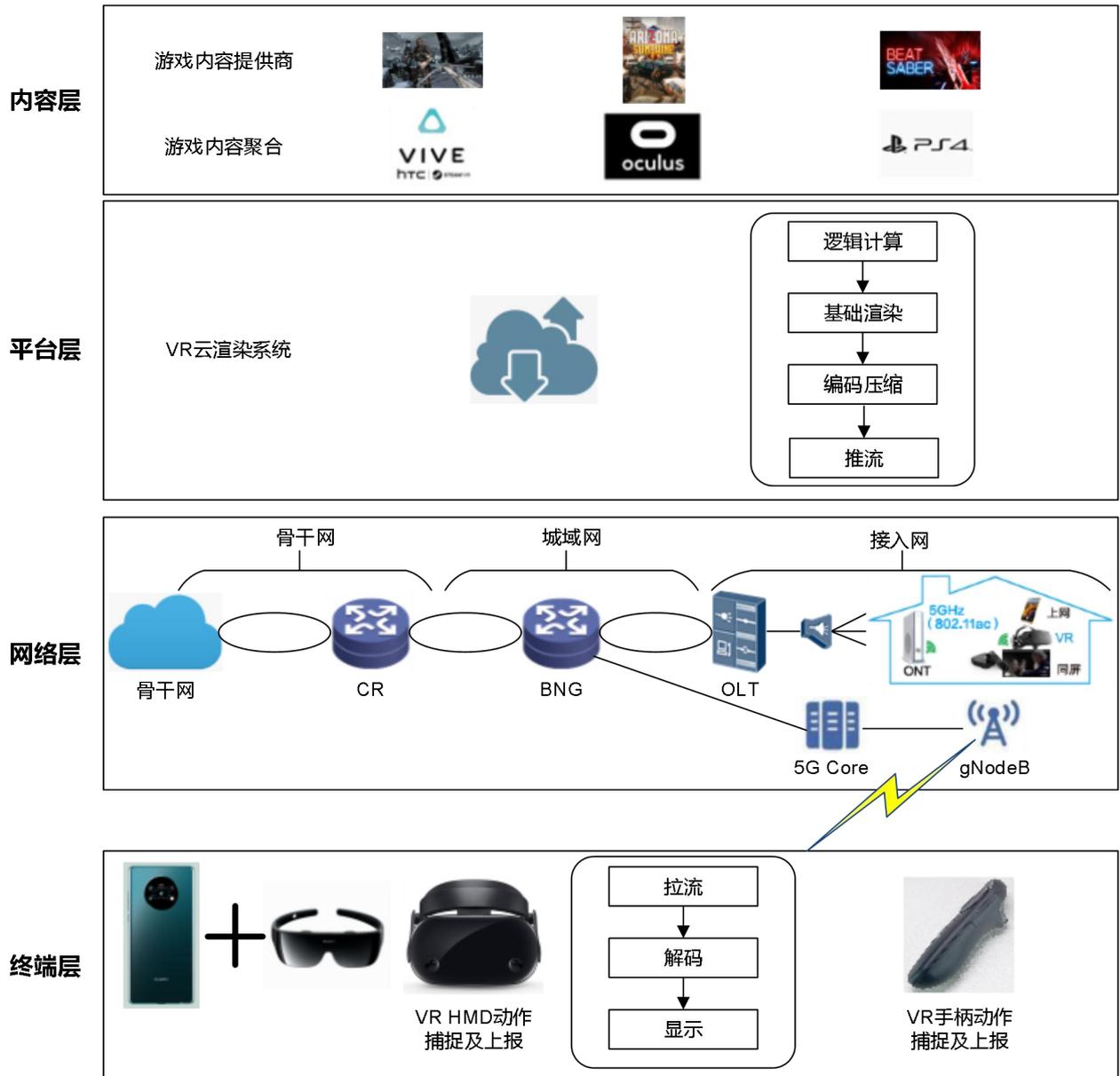


图3 云VR游戏系统架构示意图

### 5.3.2 游戏体验评估需求

影响VR游戏体验的因素包括游戏质量、交互体验和观看体验3个方面。游戏质量主要取决于游戏的分辨率、编码、码率等参数。交互体验即用户的互动过程体验，对于VR游戏，一方面从用户发出动作（如头部运动，手柄操控等）到游戏画面刷新的时延对于用户体验影响较大；另一方面动作交互的定位精度以及交互动作本身的维度数（如3DOF或6DOF）也会对用户体验产生影响。此外，在玩VR游戏时，用户经常有左右转头观看画面的需求，在转头过程中视场角范围的边缘处常常出现黑色或类似“拖影”的区域，即黑边现象。当黑边发生频率较大，或黑边的面积占比过大时，就会对用户体验产生较大影响。至于观看体验，在网络传输过程中由于游戏业务主要通过UDP协议承载，根据解码器对帧数据丢失处理机制的不同，现象可能是卡顿也可能是花屏。因此研究VR游戏的体验评估时需要考虑上

述因素。

## 5.4 交互类 VR 教育业务

VR 教育是指通过 VR 技术来构建虚拟学习环境，如虚拟实验室、宇宙中的天体运动、生物中的微观世界等，将抽象、不易理解的知识以形象、生动、直观的形式呈现，学习者使用 VR 设备就可以进入虚拟的课堂中沉浸式地体验知识，而不再枯燥地死记硬背。

### 5.4.1 应用场景

云 VR 教育的系统架构与云 VR 游戏类似，可以参考本文 5.3.1 节。

### 5.4.2 VR 教育体验评估需求

云 VR 教育的体验评估需求与与 VR 游戏类似，可以参考本文 5.3.2 节。

## 6 VR 服务体验评估架构

### 6.1 VR 系统简介

VR 端到端系统一般由终端、网络和内容提供端 3 部分组成，内容提供端可以是本地，也可以是云主机或服务器，如图 4 所示。终端在获取用户动作或指令后，将其发送到内容提供端，内容提供端通常在完成逻辑计算、渲染、编码压缩等步骤后，将相应内容数据传送到终端，终端再对内容数据进行解码，并显示到屏幕。

随着云端渲染技术及 5G 网络的发展，Cloud VR 是目前重要的 VR 系统架构发展趋势，但云渲染、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。

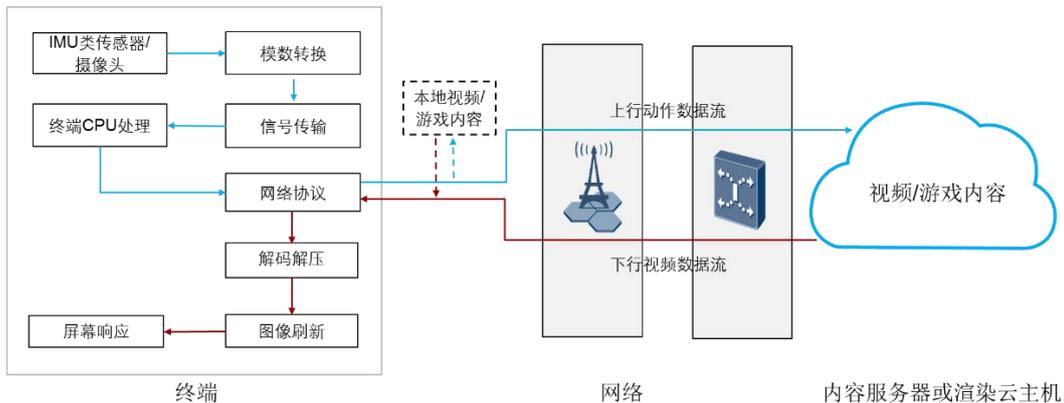


图4 VR系统架构

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。

## 6.2 关键指标

终端、网络、平台和内容，任何一个方面的能力受限都会影响VR服务的用户体验。VR视频类业务和强交互类业务，有类似的视听体验需求，关键KPI指标基本相同：

- 1) VR服务视听体验的上限由内容源决定。内容源质量主要影响因素有：a) 内容分辨率 - 衡量图像/视频的精细程度；b) 内容帧率 - 以帧为单位的图像连续呈现的频率，帧率越高则流畅度越好；c) 码率 - 视频文件在单位时间内使用的数据流量，在视频编码算法一定的条件下，码率越高则视频画面更清晰，画质更高。
- 2) VR服务的视听体验还与终端设备的解码和显示性能密切相关。VR HMD的性能主要由3个参数来表征：a) PPD - 衡量VR HMD屏幕显示清晰度；b) 刷新率 - VR HMD屏幕每秒从GPU获取新图像的次数；c) 视场角。上述3个参数的数值越大则表明该VR HMD的性能越佳。
- 3) VR服务的视听体验也受平台处理能力的影响。如VR交互游戏用户过多，超出平台的渲染处理能力，可能导致GPU渲染的帧率降低，从而影响游戏的流畅度体验。
- 4) VR服务的观看体验损伤与传输网络的性能有关。传输网络的带宽、丢包、时延及抖动都可能导致VR服务的体验受损。

此外，VR设备直接由用户佩戴在头上使用，其佩戴舒适度体验主要与4个因素相关：a) 重量和尺寸；b) 发热和散热；c) 面部贴合度；d) 透气性。

强交互类业务在交互体验上，主要受交互维度数、交互响应时延以及交互定位精度的影响。交互维度数目前主要指HMD和操控手柄支持的交互维度数，后续随着应用场景的发展与变化，可能需要有更多的肢体节点支持多维度交互。交互响应时延则主要包括：a) 头部MTP时延；b) 肢体MTP时延（可以是多节点）；c) 操作性响应时延（如扣动扳机）。

但是现阶段并不具备条件将所有影响因素及其相关KPI指标都纳入体验评估范畴，从指标可测量的角度，目前将影响VR服务用户体验的关键因素划分为沉浸体验、观看体验和交互体验3个方面，具体见6.3节。

### 6.3 模型架构

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，采用分级映射的方式构建，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 服务用户体验评估模型框架，如图 5 所示。

- 1) 视频质量（包含 VR 立体视觉和视场角的影响）和音频质量构成沉浸体验质量。在不受卡顿、花屏等体验损伤因素影响的情况下，沉浸体验质量用来表征用户观看 VR 视频时所见、所听的沉浸式感受，相关 KPI 指标涵盖由内容源和 HMD 能力决定的分辨率、帧率、码率、编码参数等。评估模型将内容源和 HMD 能力两方面相应指标（如内容源分辨率和 HMD 支持的最大分辨率）的较低值，作为其实际输入参数。
- 2) 观看体验质量用以表征受网络数据传输质量影响的体验损伤，与 VR 业务应用过程中，因传输带宽、时延抖动及丢包导致的黑边、卡顿和花屏等问题有关。
- 3) 交互体验质量主要受系统支持的交互维度数以及交互响应时延的影响。交互类业务，特别是交互游戏业务，对相关时延指标提出更为严格的要求。
- 4) 在沉浸体验的基础上，交互类业务用户进一步关注交互体验，即临场感。综合沉浸体验质量、交互体验质量和观看体验质量的临场感，VEA 视频体验联盟将其定义为 VR\_MOS。

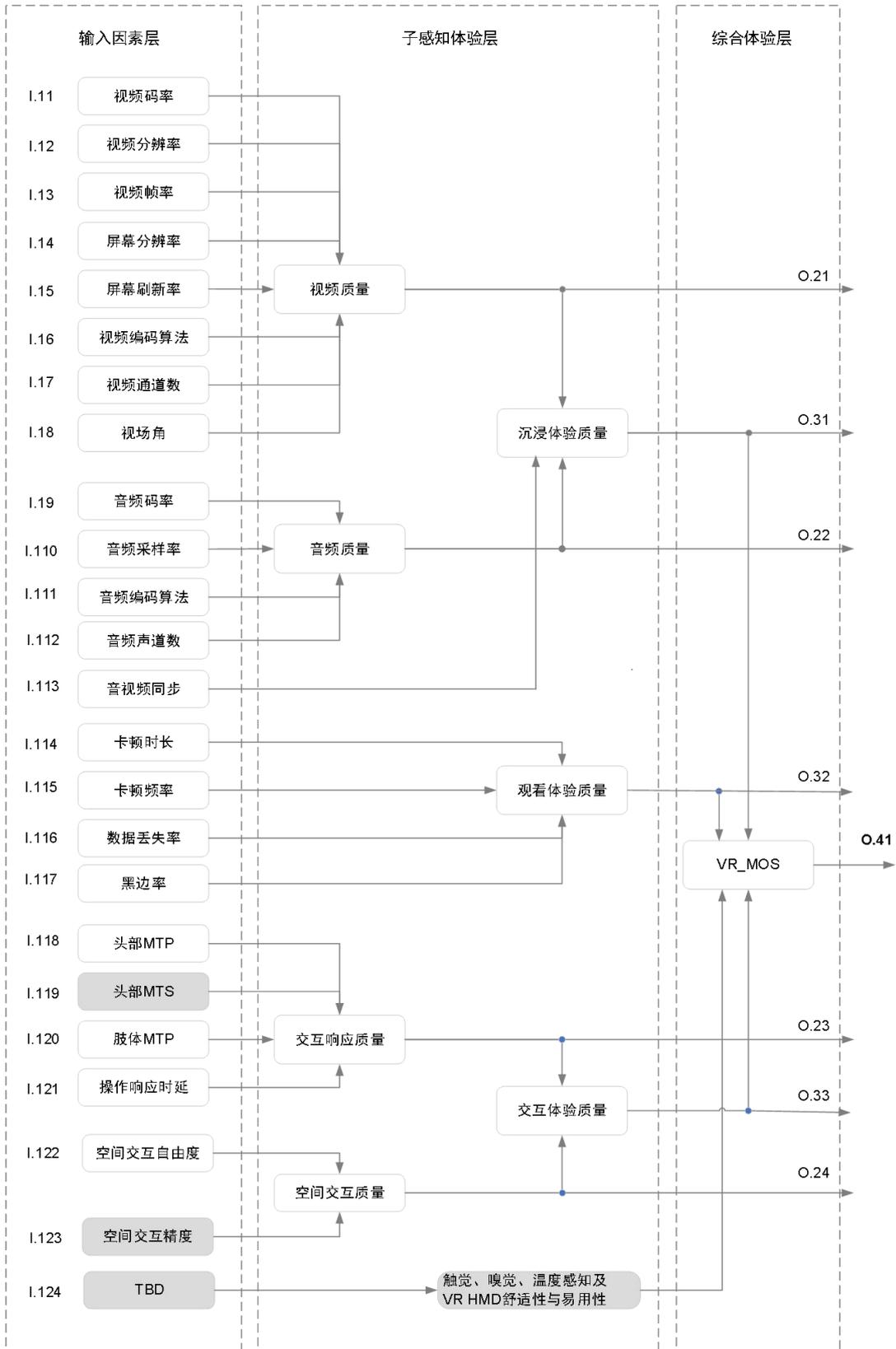


图 5 VR 服务用户体验评估模型框图

注:

除触觉、嗅觉和温度感知外，VR HMD 的佩戴舒适性，现阶段也暂未系统地纳入评估模型。主要原因是除 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。

## 7 VR 服务体验指标构成及参数采集

### 7.1 沉浸体验质量输入参数

VR 服务的沉浸体验质量输入参数说明，如表 2 所示。

表 2 沉浸体验质量参数

序号	参数	定义	缩写	英文全称
7.1.1	视频码率	单位时间传输的视频数据量（比特每秒，bps）	$Br_v$	<i>Video Bitrate</i>
7.1.2	视频帧率	VR 视频/游戏每秒的帧数（fps）	$FR$	<i>Video Frame Rate</i>
7.1.3	视频（水平）分辨率	VR 视频/游戏水平方向的像素点个数	$R_h$	<i>Resolution (Horizontal)</i>
7.1.4	视频（垂直）分辨率	VR 视频/游戏垂直方向的像素点个数	$R_v$	<i>Resolution (Vertical)</i>
7.1.5	屏幕分辨率	VR HMD 单目屏幕水平方向的像素点个数	$R_{Sh}$	<i>Screen Resolution (Horizontal)</i>
7.1.6	屏幕刷新率	VR HMD 屏幕每秒刷新的次数（Hz）	$RR$	<i>Screen Refresh Rate</i>
7.1.7	视频通道数	单目视觉（1）或立体视觉（2）	$NVC$	<i>Number Of Video Channel</i>
7.1.8	视频编码算法	视频编码标准，如 H.265/HEVC, H.264/AVC, VP9, AVS2/3	$Codec_v$	<i>Video Codec</i>
7.1.9	视场角	VR HMD 水平方向单目视场角	$FoV_h$	<i>Field Of View</i>
7.1.10	音频码率	单位时间传输的音频数据量（千比特每秒，Kbps）	$Br_a$	<i>Audio Bitrate</i>
7.1.11	音频声道数	立体声音（2）或环绕声/空间化声音（8）	$NAC$	<i>Number Of Audio Channel</i>
7.1.12	音频编码算法	音频编码标准，如 AAC-LC, Opus, DD+	$Codec_a$	<i>Audio Codec</i>
7.1.13	音频采样率	音频采样的频率（Hz）	$Sr_a$	<i>Audio Sampling Rate</i>

7.1.14	音视频同步	VR 视频/游戏的声音和画面之间的同步时延 (秒, s)	$T_{asyn}$	<i>Asynchronous Time</i>
--------	-------	------------------------------	------------	--------------------------

## 7.2 观看体验质量输入参数

VR 服务的观看体验质量输入参数说明, 如表 3 所示。

表 3 观看体验质量参数

序号	参数	说明	缩写	英文全称
7.2.1	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)	$T_r$	<i>Re-buffering Time</i>
7.2.2	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	$RF$	<i>Re-buffering Frequency</i>
7.2.3	数据丢失率	应用数据丢失率 (%)	$PLR$	<i>Packet Loss Rate</i>
7.2.4	黑边率	头部转动或移动时, 二次渲染画面中黑边的占比 (%)	$P_{black}$	<i>Black Edge Rate</i>

## 7.3 交互体验质量输入参数

VR 服务的交互体验质量输入参数说明, 如表 4 所示。

表 4 交互体验质量参数

序号	参数	说明	缩写	英文全称
7.3.1	头部 MTP	头部转动到画面刷新之间的滞后时长 (ms)	$MTP_h$	<i>Head MTP Latency</i>
7.3.2	头部 MTS	头部转动到音频方位变化之间的滞后时长 (ms)	$MTS_h$	<i>Head MTS Latency</i>
7.3.3	肢体 MTP	肢体运动到画面中肢体移动之间的滞后时长 (ms)	$MTP_b$	<i>Body MTP Latency</i>
7.3.4	操作响应时长	用户操作与操作响应之间的滞后时长 (ms)	$T_{or}$	<i>Operation Response Delay</i>
7.3.5	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	$DOF$	<i>Degree of freedom</i>

7.3.6	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	<i>Spatial Interaction Accuracy</i>
-------	--------	---------------------------------	-----	-------------------------------------

#### 7.4 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 5 所示。

表 5 模型输出参数

序号	参数	说明	缩写	英文全称
7.4.1	视频质量	对视频画面的真实感质量的评分	Q <sub>V</sub>	<i>Visual Quality</i>
7.4.2	音频质量	对音频真实感质量的评分	Q <sub>A</sub>	<i>Audio Quality</i>
7.4.3	交互响应质量	对交互时延质量的评分	Q <sub>IR</sub>	<i>Interactive Response Quality</i>
7.4.4	空间交互质量	对自由度和交互精度质量的评分	Q <sub>SI</sub>	<i>Spatial Interaction Quality</i>
7.4.5	沉浸体验质量	VR 的画面和音频的逼真度，沉浸感质量的评分	Q <sub>ImE</sub>	<i>Immersion Experience Quality</i>
7.4.6	观看体验质量	网络传输质量，如丢包、时延，对 VR 体验质量影响的评分	Q <sub>PE</sub>	<i>Presenting Experience Quality</i>
7.4.7	交互体验质量	对 VR 内容和设备的交互真实感进行的评分	Q <sub>InE</sub>	<i>Interaction Experience Quality</i>
7.4.8	VR_MOS	用户 VR 体验质量综合评分	VR_MOS	<i>Virtual Reality MOS</i>

#### 7.5 输入参数采集方法

本采集方法中的输入参数主要依靠 VR HMD 侧安装 VR 指标测试 APP 和服务器侧打点获取；或者针对合作伙伴 VR HMD 设备提供商的 VR APP 深入定制，通过 VR APP+VR\_MOS SDK 的形式获取。其中服务器为用户自行部署的云服务器（如华为云），在 VR 体验过程中主要承担云渲染等功能。具体指标测试系统如图 6 所示。

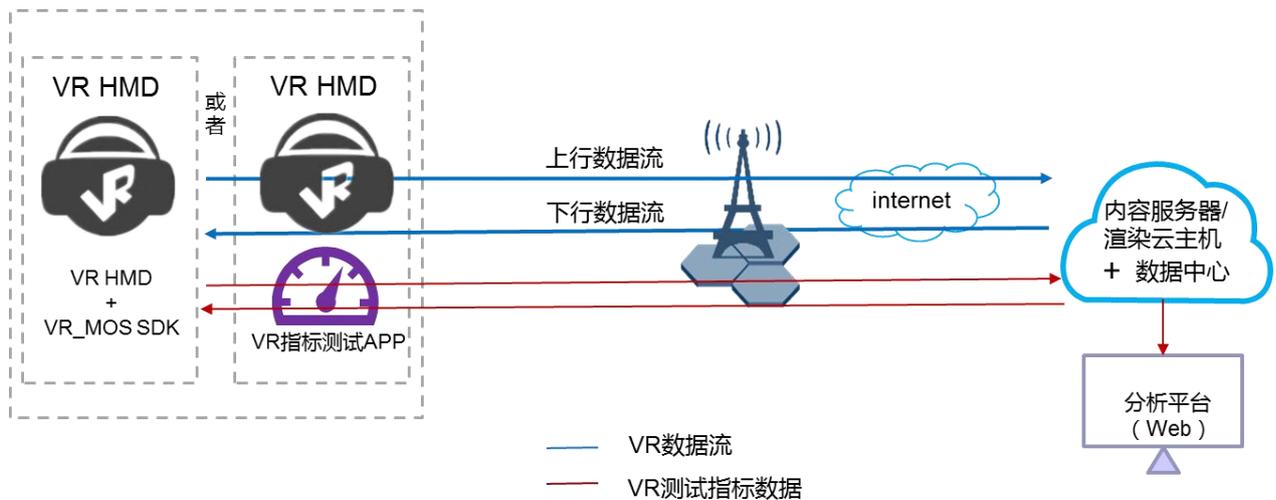


图6 VR指标测试系统架构示意

VR指标测试系统中各模块功能如下：

**VR指标测试APP：**软件内置于VR HMD中，使用手柄点击界面可选择测试的VR内容，并且开始测试，测试完成后可将测试结果显示在VR HMD中。其功能描述如下：

- 1) 收集VR HMD端获取到的VR体验指标，并上传数据到云端数据中心；
- 2) 给测试用户提供测试操作界面；
- 3) 显示测试结果（即数据中心汇总了客户端和服务端后总的结果）。

**VR APP+VR\_MOS SDK：**VR\_MOS SDK和合作伙伴VR HMD设备提供商的VR APP集成在一起，功能上与在VR HMD上安装VR指标测试APP相当。

**数据中心：**接收数据（包括VR HMD客户端和云主机端获取到的VR体验指标）和存储数据。

**内容服务器和渲染云主机：**云主机主要功能是云渲染，并且获取服务器端的体验指标。

**数据分析平台 (Web)：**实时展示VR业务体验指标，满足用户的业务监控、风险预警、地理信息分析等展示需求。

### 7.5.1 客户端采集

通过客户端可采集的输入参数，如表6所示。

表6 VR指标测试系统客户端采集的输入参数

参数	指标获取方法
视频帧率	可利用将画面渲染到屏幕的函数，统计渲染成功时的累计帧数，每秒统计一个帧率累计值。
屏幕分辨率	每个VR开发平台获取方式不同，可参考如下两个方法： 1. 若VR HMD能调用到Android接口，则通过Android平台接口获取高度和宽度的像素数。 2. 也可以通过建立VR HMD基本信息数据库，在读取此参数时通过查找数据库的方式获取。
屏幕刷新率	每个VR开发平台获取方式不同，可参考如下两个方法： 1. 若平台有调用VR HMD信息的接口，则通过此平台接口获取屏幕刷新率，如Oculus VR HMD可通过

	<p>ovrHmdInfo 读取刷新率。</p> <p>2. 可以通过建立 VR HMD 基本信息数据库（如 Pico neo 屏幕刷新率：90Hz；小米，Oculus Quest, Oculus Go 屏幕刷新率：60~72Hz，默认为 60Hz），在读取此参数时通过查找数据库的方式获取。</p>
视频通道数	<p>由 Server 侧或者 Client 侧获取均可，这里有两种方法可供参考：</p> <ol style="list-style-type: none"> <li>1. Server 侧对比左右眼画面，判断两幅画面是否一致，若一致判断为单目内容，不一致则判断为双目立体内容。</li> <li>2. 判断是否调用了定位接口，一般情况下调用了定位接口，则为双目立体内容，否则判断为单目。</li> </ol>
视场角	<p>可以通过 VR SDK 接口获取，如 Pico Native SDK 通过 void Pvr_GetFOV(float &amp;fov) 获得视场角。</p>
卡顿时长	<p>视频类采集此指标（主要是全景视频、FOV 全景视频），可基于 Client 侧开源播放器接口，获取单次卡顿时长（记为 <math>t_i</math>，即本次开始出现卡顿到本次卡顿结束的时间差）和卡顿次数。则卡顿时长 = <math>\sum(t_i) / \text{卡顿总次数}</math></p> <p>至于 CG 类游戏，此指标采集方法可参考表格下方的备注。</p>
卡顿频率	<p>视频类采集此指标（主要是全景视频、FOV 全景视频），可基于 Client 侧开源播放器接口，获取卡顿次数和播放时长。则卡顿频率 = <math>\text{卡顿总次数} / \sum(\text{播放时长})</math></p> <p>至于 CG 类游戏，此指标采集方法可参考表格下方的备注。</p>
黑边率	<p>黑边率 = 横向 FOV 上的黑边角度 (Error In Degree) / FOV 其中 Error In Degree 计算方法（以 Oculus 为例）如下：</p> <ol style="list-style-type: none"> <li>1. 在一帧数据的 TrackingInfo 生成时，记录此时 Oculus 返回的 HMD 预测方向四元组 OrientQuat_Predict。</li> <li>2. 在该帧数据渲染完毕准备提交 Oculus 显示时，记录此时 HMD 的方向四元组 OrientQuat_Current。</li> <li>3. 计算 OrientQuat_Predict 到 OrientQuat_Current 之间，在 Y 轴方向的角度差分量 Yaw。</li> </ol> <p>一秒内每一帧的黑边角度 (Yaw) 取算术平均值，作为该秒的 Error In Degree。</p>
头部 MTP	<p>1 帧的物理时延 = 终点时间戳 - 起点时间戳。其中：</p> <p>tracking: VR HMD 上以 72Hz 或 60Hz 的频率进行动作捕捉，此时间点为动作捕捉的起点。记录 tracking 的时间点，恰好在获取到 tracking 动作之后，在将 tracking 动作通过 UDP 发送之前，记录为 tracking 时间戳（即起点时间戳）。</p> <p>submit: 一帧画面正式 render 到 Oculus VR HMD 上完成的时间戳（即终点时间戳）。代码上是完成调用 vrapi_SubmitFrame2(Ovr, &amp;frameDesc) 的时间。理论上，该时间代表用户在屏幕上看到画面的时间，但存在较小的误差。</p>

	指标计算方法：头部 MTP = 1s 内所有帧的指标求和 / 1s 内所有帧的次数。
肢体 MTP	先判断 DoF，只有 6DoF 及以上才考虑此指标。其采集方法类似于头部 MTP。
操作响应时长	先判断 DoF，只有 6DoF 及以上才考虑此指标。其采集方法类似于头部 MTP。
空间交互自由度	不同的 VR HMD 有不同的获取方法： 1. 有些可直接通过 VR SDK 接口获取，如 Pico 通过 <code>public static int UPvr_GetControllerAbility(int hand)</code> 函数获取当前 Pico Neo 手柄支持 3DoF 还是 6DoF。 2. 对于不能直接获取的，可以建立 VR HMD 基本信息数据库（如：Oculus Go 只支持 3DoF，Oculus Quest 只支持 6DoF，Nolo VR 设备只支持 6DoF），在读取此参数时通过查找数据库的方式获取。

注：

- 1) VR 游戏卡顿的定义：每间隔 T（例如 100ms）周期性监测 VR 游戏画面，如果连续 n（n≥2）次以上仍然是同一帧（帧 ID 相同，画面完全相同），说明发生了 1 次卡顿，且卡顿时长为 (n-1)\*100 ms。
- 2) 在单次玩 VR 游戏过程中（假设起始时间为 t1，此时卡顿总次数和卡顿总时长均初始化为 0；终止时间为 t2），每发生 1 次卡顿则卡顿总次数累加 1，与此同时卡顿总时长也相应进行累加。则相应指标的计算公式如下：

$$\text{卡顿时长 (T}_r\text{)} = \text{卡顿总时长} / 1000$$

$$\text{卡顿频率 (RF)} = \text{卡顿总次数} / (t_2 - t_1)$$

### 7.5.2 网络节点采集

关于通过网络节点采集的输入参数，有待进一步补充。

### 7.5.3 服务端采集

通过服务端可采集的输入参数，如表 7 所示。

表 7 VR 指标测试系统服务端采集的输入参数

参数	指标获取
视频码率	视频码率 = 音视频总的码率 - 音频码率。 音视频总的码率可通过 Server 端发送数据函数，统计每一个发包大小，叠加在一起，每秒统计一个累计值。
视频（水平）分辨率	读媒体信息，此信息为 Server 端编码的固定参数。
视频（垂直）分辨率	读媒体信息，此信息为 Server 端编码的固定参数。
视频编码算法	读媒体信息，此信息为 Server 端编码的固定参数。
音频码率	若流中能区分音视频流则参考视频码率获取方法；若无法区分音视频流，一般情况下可以赋值 140Kbps。

音频声道数	读取媒体信息，通过媒体信息中包括 Audio channel, channel layout 等信息判断。
音频编码算法	媒体信息查询，若音视频流是分开的，那么此信息为 Server 端编码的固定参数。
音频采样率	CG 类 Server 侧采样频率一般是固定的，目前建模使用 48KHz。
音视频同步	由 Client 侧和 Server 侧配合获取，通过判断一帧视频和一帧音频是否要在同一个时间渲染和播放。可在 Server 侧给音视频流打时间戳，在 Client 侧比较这两个时间戳的时延差。
数据丢失率	由 Server 侧和 Client 侧配合实现，一段时间内在 Server 侧统计发包数 (M)，在 Client 侧统计收包数 (N)，则计算公式为：数据丢失率 = (M-N) / M

## 8 VR 服务用户体验评估算法

### 8.1 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  $Q_{ImE}$ , 观看体验质量  $Q_{PE}$ , 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)。

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

式中  $v_{59} \sim v_{61}$  是模型系数, 其取值分别为:

$v_{59}$	$v_{60}$	$v_{61}$
0.25	0.25	0.045

特别地, 终端对丢包的处理可以分为有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。

### 8.2 模型三大模块综述

#### 8.2.1 沉浸体验质量 $Q_{ImE}$

序号	变量	物理意义	计算所需变量
1.a	$Br_v$	视频的平均码率	7.1.1
1.b	$FR$	每秒内的视频帧数	7.1.2
1.c	$R_h$	视频水平方向的像素点个数	7.1.3
1.d	$R_v$	视频垂直方向的像素点个数	7.1.4

1.e	$R_{sh}$	单目屏幕水平方向的像素点个数	7.1.5
1.f	$RR$	屏幕每秒钟刷新次数	7.1.6
1.g	$NVC$	单目视觉 (1) 或立体视觉 (2)	7.1.7
1.h	$Codec_v$	视频编码标准, 如 H.265/HEVC, H.264/AVC, VP9, AVS2/3	7.1.8
1.i	$FoV_h$	水平方向单目视场角	7.1.9
1.j	$Br_a$	音频的平均码率	7.1.10
1.k	$NAC$	立体声音 (2) 或环绕声/空间化声音 (8)	7.1.11
1.l	$Codec_a$	音频编码标准, 如 AAC-LC, Opus, DD+	7.1.12
1.m	$Sr_a$	音频采样率, 如 48KHz	7.1.13
1.n	$T_{asyn}$	音视频同步时延	7.1.14

$$Q_P = f_1(Codec_v, Br_v, FR, (R_h, R_v)) \quad (2)$$

$$Q_V = f_2(Q_P, NVC, FoV_h) \quad (3)$$

$$Q_A = f_3(Codec_a, Br_a, NAC) \quad (4)$$

$$Q_{ImE} = f_4(Q_V, Q_A, T_{asyn}) \quad (5)$$

注:

- 1) 计算所需变量中的符号解读: 7.1.1 表示正文 7.1 节中第 1 个指标。1.a 表示附录章节 1 的第一个变量。
- 2) 本文中的视频质量, 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, 但为了更好地观察帧率, 分辨率等指标对传统意义上画面质量的影响, 这里保留画面质量  $Q_P$  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, 本文主要基于 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。

### 8.2.1.1 画面质量 $Q_P$

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

$$Q_P = \begin{cases} \min \left( \max \left( \left( v_1 \times \exp(v_2 \times BPP) + v_3 \right) \times \left( 1 + v_4 - v_4 / \left( 1 + (PPD/v_5)^{v_6} \right) \right) \times \right), 5 \right), FR \leq RR \\ \min \left( \max \left( \left( v_7 \times \exp(v_8 \times FR) + v_9 \right), 1 \right), 5 \right), FR > RR \end{cases} \quad (6)$$

$$BPP = \frac{BPF}{R_h \times R_v} \quad (7)$$

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

全景视频  $PPD$  计算如下:

$$PPD = \begin{cases} \frac{R_h}{360} & R_h \leq R_{sh} \times \frac{360}{FOV} \\ \frac{R_{sh}}{FOV} & R_h > R_{sh} \times \frac{360}{FOV} \end{cases} \quad (9)$$

$FOV$  视频  $PPD$  计算如下:

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

$$FOV = \frac{\sum_{i=1}^N (1 - P_{black})}{N} \times FoV_h \quad (10)$$

注:

公式 (10) 中  $P_{black}$  的含义请参考本文 8.2.2 节。

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	$v_1$	$v_2$	$v_3$	$v_4$	$v_5$	$v_6$	$v_7$	$v_8$	$v_9$
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  $Q_P$  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,  $Q_P$  逐渐增大;  $Q_P$  随  $PPD$  增加而增大且当大于一定  $PPD$  后变化趋缓。

### 8.2.1.2 视频质量 $Q_V$

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:

$$Q_V = \max\left(\min(v_{10}Q_P + v_{11}FoV_h + v_{12}, 5), 1\right) \quad (11)$$

式中， $v_{10} \sim v_{12}$ 为模型系数，其取值分别如下：

	$v_{10}$	$v_{11}$	$v_{12}$
单目视觉	0.595	0.020	-0.735
立体视觉	0.655	0.016	-0.342

### 8.2.1.3 音频质量 $Q_A$

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音或空间化声音时，音频质量评估模型为：

$$Q_A = v_{16} \times \left( 1 + v_{13} - \frac{v_{13}}{1 + \left( \frac{Br_a}{v_{14}} \right)^{v_{15}}} \right) + v_{17} \quad (12)$$

式中， $v_{13} \sim v_{17}$ 为模型系数，其取值分别如下：

	$v_{13}$	$v_{14}$	$v_{15}$	$v_{16}$	$v_{17}$
立体声音	4	47.100	2.134	0.81	0.3
空间化声音	4.2	42	1.25	0.96	0.04

### 8.2.1.4 沉浸体验质量 $Q_{ImE}$

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：

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

式中， $v_{18} \sim v_{24}$ 为模型系数，其取值分别如下：

$v_{18}$	$v_{19}$	$v_{20}$	$v_{21}$	$v_{22}$	$v_{23}$	$v_{24}$
0.9534	0.1954	-0.01747	-0.3466	1.156	-3.72	0.141

### 8.2.2 观看体验质量 $Q_{PE}$

序号	变量	物理意义	计算所需变量
2.a	$T_r$	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	7.2.1
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）	7.2.2
2.c	$PLR$	应用数据丢失率	7.2.3
2.d	$P_{black}$	一次 VR 业务体验过程中因用户头部转动而当前画面未就	7.2.4

	绪时导致的黑边在视口内的最大占比	
--	------------------	--

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

### 8.2.2.1 视听连续性 $Q_C$

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：

$$Q_C = \max\left(\left(v_{42} \times \ln(T_r + v_{43}) + v_{44}\right) \times \left(v_{45} \times \ln(RF + v_{46}) + v_{47}\right) \times \left(v_{48} \times \exp\left(v_{49} \times P_{black}^{v_{50}}\right) + v_{51}\right), 1\right) \quad (15)$$

$$RF = \frac{N_r + \mu \times N_{initial}}{T} \quad (16)$$

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

其中， $N_r$ 为卡顿次数； $\mu=0.1$ ； $T$ 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 $N_{initial}=0$ 。此外， $v_{42} \sim v_{47}$ 为模型系数，其取值分别如下：

$V_{42}$	$V_{43}$	$V_{44}$	$V_{45}$	$V_{46}$	$V_{47}$	$V_{48}$	$V_{49}$	$V_{50}$	$V_{51}$
-0.3707	0.1408	1.842	-0.4741	1.565	2.167	-0.4	0.4231	0.3267	1.4

### 8.2.2.2 视听完整性 $Q_I$

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。

在典型VR业务流数据编码方案，编码结构为IPPP（该编码结构利于降低解码时延，但对丢包表现得更敏感），每帧分为10个Slice，帧率为60fps，GOP为60（对应1s）的情况下，当基于UDP传输且不使用FEC时，视听完整性 $Q_I$ 的公式为：

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

其中， $PLR$ 是数据丢失率（如丢失率为0.3%则公式中为0.3）。

在上述相同条件下，当使用FEC时，视听完整性 $Q_I$ 的公式则为：

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

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：

$$Br'_v = Br_v \times (1 - R_{FEC}) \quad (19)$$

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， $v_{52} \sim v_{57}$ 为模型系数，其取值分别如下：

$v_{52}$	$v_{53}$	$v_{54}$	$v_{55}$	$v_{56}$	$v_{57}$
3.95	0.052	1.05	3.98	0.33	1.02

### 8.2.2.3 观看体验质量 $Q_{PE}$

当终端对丢包的处理基于无丢包掩盖机制时，观看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  $Q_I$  子体验项），也可能仅包括花屏和黑边（此时仅存在  $Q_I$  子体验项，典型情况是华为云 VR，系基于 UDP+FEC 传输协议）；当终端对丢包的处理基于有丢包掩盖机制时，观看体验损伤仅包括卡顿和黑边（此时仅存在  $Q_C$  子体验项，典型情况是视博云 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：

$$Q_C\_STInstant = Q_C \quad (20)$$

视听完整性的实时评分可以根据公式（18）或公式（18'）计算：

$$Q_I\_BLInstant = Q_I \quad (21)$$

观看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 $\alpha$ 滤波函数计算当前会话评分， $\alpha$ 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  $Q_C$  子体验项时，具体公式如下：

$$Q_C\_STSession_n = \alpha \times Q_C\_STSession_{n-1} + (1 - \alpha) \times Q_C\_STInstant \quad (22)$$

$$Q_C\_STSession = \alpha_1 \times \exp(\beta_1 \times StallingRatio) + \alpha_2 \times \exp(\beta_2 \times StallingRatio) \quad (23)$$

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：

$$Q_C\_STSession = \alpha_1 \times \exp\left(\beta_1 \times \left(\frac{Q_C}{5}\right)\right) + \alpha_2 \times \exp\left(\beta_2 \times \left(\frac{Q_C}{5}\right)\right) \quad (24)$$

当存在  $Q_I$  子体验项时，具体公式如下：

$$Q_I\_BLSession_n = \alpha \times Q_I\_BLSession_{n-1} + (1 - \alpha) \times Q_I\_BLInstant \quad (25)$$

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

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:

$$Q_I\_BLSession = \alpha_3 \times \exp\left(\beta_3 \times \left(\frac{Q_I}{5}\right)\right) + \alpha_4 \times \exp\left(\beta_4 \times \left(\frac{Q_I}{5}\right)\right) \quad (27)$$

最后, 观看体验质量评估模型为:

$$Q_{PE} = \begin{cases} Q_C & TCP \\ Q_I & UDP \text{ without FEC} \\ Q_I & UDP \text{ with FEC} \end{cases} \quad (28)$$

注:

- 1) 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。
- 2) 公式 (22) - (27) 中, 与实时和会话评分计算相关的系数, 不作为本文研究的重点, 本处只用作逻辑关系说明, 具体可参见相关文档说明。

### 8.2.3 交互体验质量 $Q_{InE}$

序号	变量	物理意义	计算所需变量
3.a	$MTP_h$	头部转动与画面刷新之间的滞后时长	7.3.1
3.b	$MTS_h$	头部转动与音频方位变化之间的滞后时长	7.3.2
3.c	$MTP_b$	肢体运动与画面中肢体移动之间的滞后时长	7.3.3
3.d	$T_{or}$	用户操作指令与画面中操作响应之间的滞后时长	7.3.4
3.e	$DOF$	对象在空间内移动的方式	7.3.5

$$Q_{InE} = f_6(DOF, MTP_h, MTP_b, T_{or}) \quad (29)$$

#### 8.2.3.1 VR 视频的交互体验质量 $Q_{InE}$

对于 VR 视频, 交互体验质量评估模型为:

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

式中,  $DMOS_{hm}$  为头部 MTP 导致的交互一致性下降量。

$DMOS_{hm}$  与头部 MTP 的关系可以表示为:

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

式中,  $v_{25} \sim v_{26}$  及  $v_{31} \sim v_{34}$  为模型系数, 其取值分别如下:

V25	V26	V31	V32	V33	V34
0.0667	4.3	1.563	0.046	0.01	0.058

### 8.2.3.2 VR 游戏的交互体验质量 $Q_{InE}$

对于 VR 游戏，交互体验质量评估模型为：

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

$$DMOS_m = \min\left(\max(DMOS_{hm}, DMOS_{om}, DMOS_{bm}) + v_{29} \times \frac{DMOS_{hm} \times DMOS_{om} \times DMOS_{bm}}{DMOS_{hm} + DMOS_{om} + DMOS_{bm} + v_{30}}, 4\right) \quad (31')$$

式中， $DMOS_{om}$  为操作响应时延导致的交互一致性下降量； $DMOS_{bm}$  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）。

$DMOS_{om}$  与操作响应时延的关系可以表示为：

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

$DMOS_{bm}$  与肢体 MTP 的关系可以表示为：

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

式中， $v_{27} \sim v_{30}$  及  $v_{35} \sim v_{41}$  为模型系数，其取值分别如下：

V27	V28	V29	V30	V35	V36	V37	V38	V39	V40	V41
1.1	1.6	0.98	0.001	1.443	0.018	0.01	0.119	1.343	-7.905	-5.02

## 9 参考文献

- [1] 中国信息通信技术研究院，“视频服务用户体验标准 1.0”，链接：<http://www.gxkjcg.cn/standard/html/?140.html>
- [2] VEA 视频体验联盟，“虚拟现实（VR）体验标准技术白皮书”，链接：[www.chinavea.cn/upload/P24b193bf9ba6439ea0bf94cee2401598.pdf](http://www.chinavea.cn/upload/P24b193bf9ba6439ea0bf94cee2401598.pdf)
- [3] 3GPP，“QoE parameters and metrics relevant to the Virtual Reality (VR) user experience”，链接：[https://www.3gpp.org/ftp/Specs/archive/26\\_series/26.929/26929-g10.zip](https://www.3gpp.org/ftp/Specs/archive/26_series/26.929/26929-g10.zip)
- [4] 3GPP，“Virtual Reality (VR) media services over 3GPP”，链接：[https://www.3gpp.org/ftp/Specs/archive/26\\_series/26.918/26918-g00.zip](https://www.3gpp.org/ftp/Specs/archive/26_series/26.918/26918-g00.zip)
- [5] ITU，“ITU-R BT.1788 主观质量评价标准中文版”，链接：[https://www.itu.int/dms\\_pubrec/itu-r/rec/bt/R-REC-BT.1788-0-200701-1!!PDF-C.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.1788-0-200701-1!!PDF-C.pdf)
- [6] ITU，“ITU-R BT.500-13 建议书”，链接：[https://www.itu.int/dms\\_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-1!!PDF-C.pdf](https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-1!!PDF-C.pdf)
- [7] 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.