

GY

中华人民共和国广播电视和网络视听行业标准

GY/T XXX—XXXX

VR 视听内容质量主观评价方法

Subjective evaluation method of VR audiovisual content quality

(报批稿)

XXXX - XX - XX 发布

XXXX - XX - XX 实施

国家广播电视总局 发布

目 次

前言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	1
5 主观评价条件	2
6 评价员	2
7 评价前准备	2
8 评价实施	3
9 主观评价评分内容	4
9.1 概述	4
9.2 视听质量评分	4
9.3 观看舒适度评分	5
10 数据统计与筛选	6
10.1 通用要求	6
10.2 平均分的统计	6
10.3 标准偏差的统计	6
10.4 置信区间的统计	7
10.5 数据的筛选	7
参考文献	8

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请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专利的责任。

本文件由全国广播电视和网络视听标准化技术委员会（SAC/TC 239）归口。

本文件起草单位：国家广播电视总局广播电视规划院、国家广播电视总局广播电视科学研究院、中广电广播电影电视设计研究院有限公司、北京体育大学、咪咕文化科技有限公司。

本文件主要起草人：郑涛、商鹏、宁金辉、王惠明、魏娜、饶丰、王嘉、赵珊珊、赵琳琳、罗姣姣、王雷、汪芮、张乾、陈玲玉、孙琳、欧臻彦、许帅、刘汉源、刘文翰、邸杨骅、杨威、倪俊杰、王倩男、于明、郑欣、杨春。

VR 视听内容质量主观评价方法

1 范围

本文件规定了对VR视听内容的视听质量和观看舒适度进行主观评价的方法，包括主观评价条件、人员要求、评价流程、数据分析等。

本文件适用于对VR视听内容的视听质量和观看舒适度进行主观评价，也适用于VR系统和设备的研发、测试和运行维护。

2 规范性引用文件

本文件没有规范性引用文件。

3 术语和定义

下列术语和定义适用于本文件。

3.1

主观评价室 subjective assessment room

进行图像质量主观评价或声音质量主观评价的专用受控环境。

3.2

虚拟现实头戴式显示设备 virtual reality head-mounted display device

VR设备

佩戴在用户头部并提供给用户虚拟现实感觉和体验的显示设备。

[来源：GB/T 38259—2019，3.2]

4 缩略语

下列缩略语适用于本文件。

ACR 绝对类别评分 (Absolute Category Rating)

AR 增强现实 (Augmented Reality)

CMP 立方体投影 (Cube Map Projection)

EAC 等角方块映射 (Equi-Angular Cubemap)

ERP 等距长方形映射 (Equi-Rectangular Projection)

HDR 高动态范围 (High Dynamic Range)

LCD 液晶显示 (Liquid Crystal Display)

OLED 有机发光二极管 (Organic Light Emitting Diode)

VR 虚拟现实 (Virtual Reality)

3DoF 三自由度 (Three Degrees of Freedom)

5 主观评价条件

开展VR视听内容质量主观评价的主观评价室应具备安静的环境，环境噪声不应超过37dBA；为避免评价员转动身体或头部时碰撞周边物体，应以评价员为中心预留不小于4m²（2m×2m）的空地，且主观评价室净高度应不低于2.5m；在该空地内，评价员可选择站立或使用旋转座椅就座的方式进行评价，以保证其可以自由旋转观看VR视频。

为保障VR视听内容评价时的沉浸感，评价员应佩戴虚拟现实头戴式显示设备（以下简称“VR设备”）进行主观评价。

VR设备视音频处理能力应符合表1的规定。

表1 VR设备视音频处理能力

技术指标		参数
设备显示性能	亮度	LCD屏幕：≥120cd/m ² OLED屏幕：≥80cd/m ²
	色域覆盖率	≥80% sRGB@CIE1976 ≥70% NTSC@CIE1976
	对比度	LCD屏幕：≥500:1 OLED屏幕：≥1500:1
视频处理能力	视频分辨率	不低于被评价视频的分辨率
	帧率	不低于被评价视频的帧率
	动态范围	HDR
	幅型比	2:1、16:9
	解码能力	8K VR视频码率不低于300Mbps
音频处理能力	声道	立体声或环绕声或三维声
	解码能力	不低于64kbps
渲染交互	逆映射格式	ERP或CMP或EAC
	自由度	头部3DoF及以上

6 评价员

评价员的选用原则如下：

- 常规性 VR 视听内容宜选用非专业评价员来进行主观评价；
- 当需要精确判断时，应由受过专业评价训练的专业评价员来进行主观评价。

评价员总人数应不少于15人，选用要求如下：

- 应包含不同性别、不同受教育程度的观众，年龄应覆盖 18~35、36~50、51~65 三个年龄段；
- 应具有一定分析判断能力；
- 应能接受和掌握评价方法和要求；
- 视力（含矫正视力）应为 1.0 以上，无色盲、干眼症等眼部疾病症状；
- 观看 30s 以内的短视频时，应无强烈的眩晕、恶心等生理不适感。
- 评价组织者不应作为评价员。

7 评价前准备

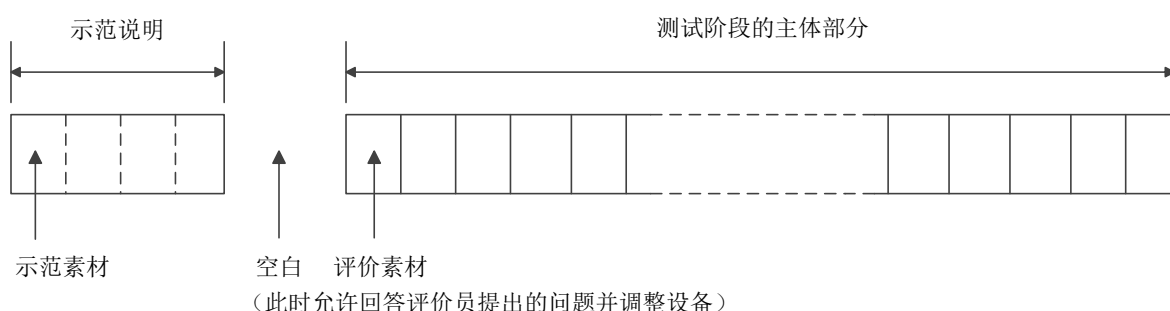
在评价开始前，评价组织者应对评价员进行VR设备使用培训。培训内容应包括：

- 指导评价员正确佩戴VR设备；
- 指导评价员正确运用手柄或手势对VR设备进行操作；
- 向评价员说明使用VR设备过程中可能出现的不适症状；
- 明确告知评价员，若在评价过程中出现严重不适，立即停止测试。

8 评价实施

8.1 周期说明

8.1.1 评价周期流程见图1。



注1：示范说明中，每格代表一个示范素材；测试阶段主体部分中，每格代表一个评价素材。

注2：测试阶段主体部分包括视听质量评价和观看舒适度评价。

图1 评价周期流程

8.1.2 在每个评价周期开始前，评价组织方应向评价员详细、准确地说明：

- a) 评价流程；
- b) 评价方法；
- c) 打分要求；
- d) 测试图像；
- e) 评价时间长度。

8.1.3 单个评价周期（包括示范说明）的总时长应不超过50min。

8.1.4 评价应按下列顺序进行：

- a) 视听质量评价；
- b) 观看舒适度评价。

8.1.5 若评价所需总时长超过8.1.3的规定时，应划分为若干个符合8.1.3要求的评价周期。

8.1.6 相邻两评价周期间，评价员休息时间应不少于15min，以消除疲劳影响。

8.1.7 评价员应分别独立评价。

8.2 素材播放与评分

8.2.1 一个评价周期由“多次播放-评分”过程组成。

8.2.2 每次播放结束后，评价员应对素材整体呈现效果和观看舒适度分别进行评分。评分完成后，方可进入下一个素材播放和评分。

8.2.3 相邻两素材播放的间隔时间应大于10s。

8.2.4 每个素材应循环播放，确保评价员能转动头部或眼睛完成对视频全部画面的观看。

8.3 设备操作与素材顺序

8.3.1 进入评价周期后，评价员不应擅自操作VR设备。

8.3.2 示范阶段：评价员应在引导员的操作指引下，通过手柄或手势选择示范素材，用于稳定判别力；示范阶段结果不纳入正式统计。

8.3.3 正式测试阶段：评价员应按照引导员的操作指引，依次播放待测试素材。

8.3.4 如需比较不同测试素材之间的差异，评价员可在该轮测试结束后重复观看相应素材。

8.4 音频设置

播放所有素材时，应打开音频开关，确保评价员在观看VR视频画面的同时可完整收听到素材音频内容。

9 主观评价评分内容

9.1 概述

VR视听内容质量主观评价包括视听质量评分和观看舒适度评分。评价时引导员协助评价员完成打分，具有AR功能的VR设备可通过AR摄像头在纸质评分表上进行评分；不具备AR功能的VR设备通过向引导员口头报数或在VR设备中的打分程序中打分。

9.2 视听质量评分

视听质量评分主要对视听内容本身呈现总体效果进行评分，分值越高质量越高，包括视听内容的画面质量、声音质量和沉浸感三个维度，每个维度均按照ACR，使用单刺激连续质量评分表进行评分。单刺激连续质量评分表见图2。

评分量化使用直尺对评分表进行度量，并转换为百分制，对评价员的评分值和评价员信息进行统计。VR视听内容视听质量评分维度和指标见表2。

单刺激连续质量评分表

	画面 质量	声音 质量	沉浸 感
优			
良			
中			
差			
劣			

姓名：
评价员 性别：
年龄：

图 2 评分表

表 2 VR 视听内容视听质量评分维度和指标

评价维度	评价指标	指标描述
画面质量	分辨率	视频画面的清晰度，包括细节表现、锐度等
	色彩还原度	视频画面色彩的准确性
	亮度对比度	视频画面的明暗对比和层次感
	画面流畅度	视频画面的流畅度
	拼接特性	视频画面的“缝合效果”、变形程度等
声音质量	清晰度	视频声音的清晰度
	声画同步	声音和视频的同步程度
沉浸感	空间感	对视频中空间大小、距离、深度等的感知程度
	场景真实度	视频场景的真实感
	交互自然度	通过手柄或手势对VR视频操作交互的自然程度，包括物理反馈、声音反馈、互动流畅度等

9.3 观看舒适度评分

每段测试素材播放并完成视听质量评分后，评价员应立即对该素材的观看舒适度进行评分，分值越高舒适度越差。评分维度见表3。

表 3 VR 视听内容的观看舒适度评分样表

观看舒适度维度	评分			
	无 (0分)	轻微 (1分)	中度 (2分)	严重 (3分)
眼酸涩、眼胀				
流泪				
视线模糊				
头痛				
头胀				
头晕 (睁眼时)				
头晕 (闭眼时)				
额前眉心不适				
恶心				
唾液分泌增加				
胃不适				
打嗝				
下意识抽搐				
心脏不适				
紧张出汗				
交互操作困难				
耳胀、耳鸣				

10 数据统计与筛选

10.1 通用要求

评价完成后，应对VR视听内容评分进行平均分统计，包括视听质量评分和观看舒适度评分。

10.2 平均分的统计

每一被测 VR 视听内容的视听质量评分和观看舒适度评分平均分 \bar{u}_k 按公式 (1) 进行计算。

$$\bar{u}_k = \frac{1}{N} \sum_{i=1}^N u_{ik} \quad \dots\dots\dots (1)$$

式中：

N ——评价员数量；

u_{ik} ——评价员*i*对被测 VR 视听内容*k*的评分。

10.3 标准偏差的统计

每一被测VR视听内容的标准偏差 S_k 按公式 (2) 进行计算。

$$S_k = \sqrt{\sum_{i=1}^N \frac{(\bar{u}_k - u_{ik})^2}{(N-1)}} \quad \dots\dots\dots (2)$$

式中：

N ——评价员数量；

\bar{u}_k ——每一被测VR视听内容的评分平均分；

u_{ik} ——评价员*i*对被测VR视听内容*k*的评分。

10.4 置信区间的统计

每一被测VR视听内容的95%置信区间为：

$$[\bar{u}_k - \delta_k, \bar{u}_k + \delta_k]$$

其中， $\delta_k = 1.96 \times \frac{S_k}{\sqrt{N}}$ 。

10.5 数据的筛选

对每一评价员的评分 u_{ik} 计算均值 \bar{u}_k 、标准偏差 S_k 和系数 β_{2k} 。 β_{2k} 按公式(3)进行计算。

$$\beta_{2k} = \frac{m_4}{(m_2)^2} \dots\dots\dots(3)$$

式中：

$$m_x = \frac{\sum_{i=1}^N (u_{ik} - \bar{u}_k)^x}{N}$$

对于每一评价员 i ，计算其 P_i 和 Q_i ，即：对于 $k = 1$ 至 K

若 $2 \leq \beta_{2k} \leq 4$ ，则：

若 $u_{ik} \geq \bar{u}_k + 2S_k$ ，则 $P_i = P_i + 1$ ；

若 $u_{ik} \leq \bar{u}_k - 2S_k$ ，则 $Q_i = Q_i + 1$ 。

否则：

若 $u_{ik} \geq \bar{u}_k + \sqrt{20}S_k$ ，则 $P_i = P_i + 1$ ；

若 $u_{ik} \leq \bar{u}_k - \sqrt{20}S_k$ ，则 $Q_i = Q_i + 1$ 。

若 $\frac{P_i}{K} > X$ 或 $\frac{Q_i}{K} > X$ ， X 为经验值0.2，则删除该评价员 i 。

参 考 文 献

- [1] GB/T 38259—2019 信息技术 虚拟现实头戴式显示设备通用规范
 - [2] GY/T 134—1998 数字电视图像质量主观评价方法
 - [3] GY/T 340—2020 超高清晰度电视图像质量主观评价方法 双刺激连续质量标度法
 - [4] GY/T 356—2021 VR视频系统节目制作和交换用视频参数值
 - [5] SJ/T 11965—2025 超高清虚拟现实显示设备技术规范
 - [6] ITU-R BT.500-15 Methodologies for the subjective assessment of the quality of television images
 - [7] ITU-T P.800.2 Mean opinion score interpretation and reporting
 - [8] ITU-T P.919 Subjective test methodologies for 360° video on head-mounted displays
-