中 关 村 现 代 信 息 消 费 应 用 产 业 技 术 联 盟 团 体 标 准 T/INFOCA 005-2022

移动直播视频体验质量指标与评测-观众端

System architecture of assessment for QoE of live streaming video services in mobile network-Audience

2022 年 10 月12日发布 2022年10 月12 日实施

中关村现代信息消费应用产业技术联盟 发布

目 次

[目 次 2](file:///C%3A%5C%5CUsers%5C%5Cw00467326%5C%5CDesktop%5C%5C20220805%5C%5C%E7%A7%BB%E5%8A%A8%E7%9B%B4%E6%92%AD%E8%A7%86%E9%A2%91%E4%BD%93%E9%AA%8C%E8%B4%A8%E9%87%8F%E6%8C%87%E6%A0%87%E4%B8%8E%E8%AF%84%E6%B5%8B-%E8%A7%82%E4%BC%97%E7%AB%AF_V1.23--%E6%89%B9%E6%B3%A8%E7%89%88.docx%22%20%5Cl%20%22_Toc111562684)

[前 言 0](#_Toc111562685)

[移动直播视频体验质量指标与评测-观众端 1](#_Toc111562686)

[1 范围 1](#_Toc111562687)

[2 规范性引用文件 1](#_Toc111562688)

[3 术语、定义和缩略语 2](#_Toc111562689)

[3.1 缩略语 2](#_Toc111562690)

[3.2 术语和定义 3](#_Toc111562691)

[4 移动直播QoE评估架构 3](#_Toc111562692)

[4.1 关键指标 3](#_Toc111562693)

[4.2 移动直播QoE评估架构 4](#_Toc111562694)

[5 移动直播质量指标构成 6](#_Toc111562695)

[5.1 观众端视频质量输入参数 6](#_Toc111562696)

[5.2 观众端音频质量输入参数 7](#_Toc111562697)

[5.3 观众端呈现体验质量输入参数 7](#_Toc111562698)

[5.4 观众端交互体验质量输入参数 8](#_Toc111562699)

[5.5 观众端QoE模型输出参数 9](#_Toc111562700)

[6 移动直播观众端用户体验评估算法 9](#_Toc111562701)

[6.1 总体模型综述 9](#_Toc111562702)

[6.2 模型四大模块综述 10](#_Toc111562703)

[6.2.1. 视听体验 10](#_Toc111562712)

[6.2.2. 呈现体验 13](#_Toc111562713)

[6.2.3. 交互体验 15](#_Toc111562714)

[附录A](#_Toc111562715)[ITU-T P.1203操作模式 16](#_Toc111562715)

[附录B 关于PPD计算方法 16](#_Toc111562716)

前 言

本标准是针对移动终端（如智能手机，平板电脑）场景下直播业务观众端QoE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。

本标准按照GB/T 1.1-2009给出的规则起草。

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： 中关村现代信息消费应用产业技术联盟、华为技术有限公司、上海兆言网络科技有限公司、 TCL华星光电技术有限公司、北京邮电大学、联通研究院、中国信息通信研究院、北京市博汇科技股份有限公司、深圳市中兴微电子技术有限公司、聚好看科技股份有限公司、西安电子科技大学、中恒达（北京）软件测评科技有限公司、北京快手科技有限公司、中国电信集团有限公司、中国移动通信有限公司研究院、优酷网络技术（北京）有限公司、中移（杭州）信息技术有限公司、德科仕通信（上海）有限公司等。

本标准主要起草人： 陈红、宋祖平、武亮平、段涛、黄卫东、闫石、贾武、王亚军、张家斌、孔德辉、刘帅、杨付正、陈劼联、陈戈、马英武、张世俊、李静、杨琛、吴雪波、姜卫平、史东平、钟琳、尹之帆、黄挺、张春蕾、吴敬芳、郑川川、余大力、刘璇。

移动直播视频体验质量指标与评测-观众端

1 范围

本标准规定了在固定和移动宽带网络中基于移动终端（智能手机，平板电脑）进行实时音视频互动直播时的QoE评估模型，分析了影响实时音视频互动直播QoE的主要因素和参数，定义了实时音视频互动直播QoE的评估计算方法。本标准中实时音视频互动直播是指主播端和观众端或者主播端与主播端需要频繁地进行实时音视频互动的直播。

本标准适用于公用电信网、公用互联网等环境下实时音视频互动直播的QoE进行综合评估，以及对影响实时音视频互动直播QoE的因素进行定位和分析。

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。

ITU-R BT.500-11，Methodology for the subjective assessment of the quality of television pictures

ITU-T P.913 (2021) Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment.

ITU-T P. 1203 (2017) Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport.

ITU-T P. 1203. 1 (2017) Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport – Video quality estimation module.

ITU-T P. 1203.2 (2017) Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport – Audio quality estimation module.

ITU-T P. 1203. 3 (2017) Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport – Quality integration module.

3 术语、定义和缩略语

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

QoE 用户体验质量 Quality of User Experience

PPI 每英寸所拥有的像素个数 Pixels Per Inch

PPD 角分辨率 Pixels Per Degree

BPP 像素深度 Bits Per Pixel

MOS 主观体验评分 Mean Opinion Score

ITU 国际电信联盟 International Telecommunication Union

 IPTV IP电视 Internet Protocol Television

 OTT 互联网向用户提供各种应用服务 Over The Top

3.2 术语和定义

下列术语和定义适用于本文件。

3.2.1

用户体验质量 Quality of User Experience

用户体验移动直播的服务过程中相应的视听体验质量、交互体验时延、呈现体验的质量水平。

3.2.2

实时音视频互动直播 Real-Time Communication Interactive live streaming

直播中主播端和观众端频繁进行音视频实时互动的直播场景。

3.2.3

推流 Push Stream

推流指的是把采集阶段封包好的内容传输到服务器的过程。

3.2.4

拉流 Pull Stream

拉流是指播放端用户通过拉流地址从服务器拉取直播内容的过程。。

3.2.5

卡顿 Freeze

在直播过程中，出现音频或视频的播放不连续、不流畅、画面滞帧等现象。

3.2.6

花屏 Blurred Screen

在直播过程中，画面出现图像紊乱，异常颜色的方块图等部分画面无法有效显示的问题。

3.2.7

输出采样间隔 Output Sampling Interval

指解析后的视频或音频每隔1秒钟（不考虑卡顿）输出一次，且音频质量（QA）和视频质量（QV）的输出采样间隔必须与质量整合模块（QIgt）期望的输入相匹配。

3.2.8

 主播连麦PK Anchors are Connected to the Microphone PK

 指主播与其他主播间进行连麦互动PK。

3.2.9

 音画同步时延 Audio-to-Video Synchronization Delay

 直播中观众端播放器接收的音频流声音和观看到的画面的之间时间戳的差值

4 移动直播QoE评估架构

直播可以通过运营商的IPTV业务进行直播，也可以通过互联网服务商的固定或移动OTT视频业务系统实现，本标准主要聚焦于通过互联网服务商的固定或移动OTT视频业务系统实现的直播场景观众端的QoE。

实时音视频互动直播的过程为主播端将产生的音视频流实时地发送到实时音视频云平台，观众实时接收主播的音视频流进行观看。实时音视频互动直播中主播和观众需要频繁地进行音视频互动，或者主播之间进行连麦PK，互动性比较强，对端到端的时延要求比较高。直播中的用户角色分为主播和观众。

4.1 关键指标

#### 4.1.1 观众端关键体验指标

直播中观众端实时地接收主播端产生的音视频流进行观看。直播中衡量观众端体验的主要指标有音频体验质量、视频体验质量、交互体验质量、观众端呈现体验质量。

音频体验质量，取决于观众端音频的清晰度、流畅度等因素，可以音频采样率、声道数、码率、编码方式、编码参数、信号层质量等客观指标度量。

视频体验质量，取决于观众端视频的清晰度、流畅度等因素，涵盖了客户端视频的分辨率、帧率、码率、编码和终端多个维度的指标，可以编码参数、分辨率、帧率、码率、信号层质量等客观指标度量。相对于PPI而言，PPD可以更好地适应曲面屏、非平面的折叠屏等终端形态的演进需要,并且PPD可以考虑到观看距离的影响因素。因此通过PPD来综合衡量视频分辨率、屏幕分辨率、屏幕尺寸等参数对观看清晰度的影响。

交互体验质量，取决于观众端端到端的时延（观众端端到端时延是指数据从主播端采集到观众端播放需要的时间）。观众端交互体验质量也包含观众端在直播间的跳转时延、观众进入直播间的首屏时延等。观众端在直播间的跳转时延是指在主播连麦PK场景下（如主播A与主播B连麦PK）主播A直播间的观众从主播A的直播间跳转到主播B的直播间所需的时间。观众进入直播间的首屏时延是指从观众端点击进入主播直播间到直播间首帧画面成功播放显示的时间。

观众端呈现体验质量，可以音频卡顿率、视频卡顿率、花屏、音画同步时延等指标进行衡量。



图1 影响直播观众端体验的关键因素

此外，视频采集损伤、观众端终端渲染均匀性、音频设备损伤、信号处理损伤等其它部

分的指标也对观众端用户体验有一定的影响。

为了科学的评估和量化移动直播观众端的用户体验质量，把评估场景分为会话场景（表征观众端用户一次完整的观看行为的体验质量） 和实时场景（表征观众端用户实时观看过程中的体验质量） 。

开展移动直播观众端用户体验评价综合了观众端视听体验质量、交互体验质量、呈现体验质量，是真实反映移动直播观众端用户体验感知的平均主观评分（MOS） 。

4.2 移动直播QoE评估架构

#### 4.2.1 观众端QoE评估架构

移动直播观众端QoE评估模型架构如图2所示：

1. 视听体验质量用于表征直播中观众端所见所听的主观感受。视听体验质量主要由视频质量和音频质量构成。视频质量主要与内容清晰度（受分辨率等因素影响）、内容流畅度（以帧率表征）和编码参数（如编码算法、码率）有关。音频质量主要与音频采样率、声道数（如单声道、双声道（立体声））、编码参数（如编码算法、码率）有关。
2. 呈现体验质量用于表征受网络传输质量影响的体验损伤，主要与直播过程中因传输带宽、时延抖动和丢包导致的卡顿、花屏和音画不同步等问题有关。
3. 交互体验质量用于表征直播中观众端交互行为的影响，主要与响应时延有关，如观众端端到端时延、观众在直播间跳转时延、观众进入直播间的首屏时延等。



图 2 移动直播观众端QoE评估模型架构

根据图2，除QIgt（Quality integration module，质量整合模块）之外的其他模块，模型输入和输出均比较清晰，此处不做展开。这里重点介绍一下比较复杂的QIgt模块。QIgt模块由QAVE、QPE和QInE 三个子模块构成。

* QAVE（Audio-Visual Experience Quality，视听体验质量）子模块，其输入是：1）每输出采样间隔的视觉体验质量（O.21）；2）每输出采样间隔的音频质量（O.22）。其输出是：基于每输出采样间隔的视听体验质量（O.31）；2）基于会话的视听体验质量（O.32）。关于QAVE子模块的更多细节请参考本标准6.2.1小节的相关描述。
* QPE（Presenting Experience Quality，呈现体验质量）子模块，其输入是：1）卡顿感知质量（QStall）；2）花屏感知质量（QBlock）；3）每输出采样间隔的音视频同步（）参数。其输出是：呈现体验质量（O.33）。关于QPE子模块的更多细节请参考本标准6.2.2小节的相关描述。
* QInE（Interactive Experience Quality，交互体验质量）子模块，其输入是：1）每输入采样间隔的交互响应参数（I.16）；其输出是：1）基于每输出采样间隔的交互体验质量（O.34）；2）基于会话的交互体验质量（O.35）。关于QInE子模块的更多细节请参考本标准6.2.3小节的相关描述。

移动直播客户端的QoE指标从用户层面出发，可以定义为终端用户观众端对直播业务的综合主观感知，反映移动直播的体验综合评分，用MOS表示。

定义MOS = f (*QAVE*, QInE, QPE)，即用户体验综合评分为（视听体验质量*QAVE*，呈现体验质量QPE，交互体验质量 QInE ）的函数关系。

5 移动直播质量指标构成

5.1 观众端视频质量输入参数

观众端视频质量提取的输入参数I.11表征直播中观众端视频的质量，I.GEN则主要表征观众端终端的显示能力（如屏幕分辨率，刷新率等）。

表1 观众端视频质量输入参数

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 参数 | 数据源 | 定义/取值 | 备注 |
| I.11 |
| 5.1.1 | 视频码率 | 移动直播APP  | 视频单位时间取样的视频数据量（Kbps），直播业务观众端视频的接收码率 | VideoBitrate，缩写为*BrV* |
| 5.1.2 | 视频帧率 | 移动直播APP | 视频每秒的帧数（fps）直播业务观众端观看直播视频的帧率 | VideoFrameRate，缩写为*FRV* |
| 5.1.3 | 视频水平分辨率 | 移动直播APP | 视频水平方向的像素点个数直播业务观众端观看直播视频水平分辨率 | VideoHorizontalResolution，缩写*Rh* |
| 5.1.4 | 视频垂直分辨率 | 移动直播APP | 视频垂直方向的像素点个数直播业务观众端观看直播视频垂直分辨率 | VideoVerticalResolution，缩写为*Rv* |
| 5.1.5 | 编解码算法 | 移动直播APP | 直播业务，观众端视频的编解码算法。如H.264、H.265、VP9等 | VideoCodecAlgorithm，缩写为*CodecV* |
| 5.1.6 | 信号层参数 | 移动直播APP | 块状度 |  |
| 模糊度 |
| 对比度 |
| 噪点度 |
| 色彩度 |
| I.GEN |
| 5.1.7 | 屏幕水平分辨率 | 终端OS | 终端屏幕水平方向的像素点个数直播业务观众端移动设备的屏幕水平分辨率 | ScreenHorizontalResolution，缩写为*RSh* |
| 5.1.8 | 屏幕垂直分辨率 | 终端OS | 终端屏幕垂直方向的像素点个数直播业务观众端移动设备的屏幕垂直分辨率 | ScreenVerticalResolution，缩写为*RSv* |
| 5.1.9 | 屏幕尺寸 | 终端OS | 终端屏幕对角线的长度（inch）直播业务观众端移动设备的屏幕尺寸 | ScreenSize，缩写为*ScrSize* |
| 5.1.10 | 屏幕刷新率 | 终端OS | 终端屏幕每秒刷新的次数（Hz）直播业务观众端移动设备的屏幕刷新率 | ScreenRefreshRate，缩写为*ScrRR* |
| 5.1.11 | 观看距离 |  | 观众端用户观看直播时距离终端屏幕的最佳观看距离（cm） | DistanceToScreen，缩写为*D2Scr* |

5.2 观众端音频质量输入参数

音频质量提取的输入参数是I.12，表征直播中观众端音频的质量。

表2 观众端音频质量输入参数

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 参数 | 数据源 | 定义/取值 | 备注 |
| I.12 |
| 5.2.1 | 音频码率 | 移动直播APP | 单位时间传输的音频数据量（Kbps）直播过程中观众端的音频码率 | AudioBitrate，缩写为*BrA* |
| 5.2.2 | 音频采样率 | 移动直播APP | 音频数据每秒采样点数 | 直播过程中观众端音频采样率 |
| 5.2.3 | 音频声道 数 | 移动直播APP | 声道数是声音录制时的音源数量或 回放时相应的扬声器数量。直播过程中观众端音频声道数 | NumberOfChannels，缩写为*NoCA* |
| 5.2.4 | 编解码方式 | 移动直播APP | 音频的编解码算法。涵盖主流的音频编解码算法，如AAC、Opus等。  | AudioCodecAlgorithm，缩写为*CodecA* |
| 5.2.5 | 信 号 层 参 数 | 移动直播APP | 响度动态范围左右声道相位差爆音静音长度 |  |

5.3 观众端呈现体验质量输入参数

呈现体验质量提取的输入参数包括I.13、I.14和I.15，其中I.13表征卡顿事件信息，I.14表征花屏事件信息（基于丢包近似评估花屏感知）， I.15表征音画同步信息。

表3 观众端呈现体验质量输入参数

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 参数 | 数据源 | 定义/取值 | 备注 |
| I.13 |
| 5.3.1 | 音频卡顿 | 移动直播APP | 在直播过程中，观众端出现音频播放不连续、不流畅等现象。指从主播端到观众端端到端的音频卡顿。可以取音频卡顿的时长。 |  |
| 5.3.2 | 视频卡顿 | 移动直播APP | 在直播过程中，观众端出现视频播放不连续、不流畅、画面滞帧等现象。指从主播端到观众端端到端的视频卡顿。可以取视频卡顿的时长。 |  |
| I.14 |
| 5.3.3 | 数据包丢失率 | 移动直播APP | 在每一个输入采样间隔内，视频包丢失百分比（%） | PacketLossRatioInSamplingIntervalt，缩写为*pLRt* |
| 5.3.4 | 呈现损伤标识 | 移动直播APP | 根据终端对丢包处理是否有丢包掩盖机制，可能存在3种取值：0-仅卡顿（缺省），1-仅花屏，2-同时存在卡顿和花屏 | PresentImpairmentFlag，缩写为*PsIpmFlag* |
| I.15 |
| 5.3.5 | 音画同步时延 | 移动直播APP | 直播中观众端接收的音频流声音和观看到的画面的同步时延（ms） | AudioDelayRelativeToVideo，缩写为 |

5.4 观众端交互体验质量输入参数

观众端交互体验质量提取的输入参数是I.16，表征直播过程中观众端的交互行为的响应时延，涵盖观众端端到端的时延（端到端时延是指数据从主播端采集到观众端播放需要的时间），观众端交互体验质量也包含主播连麦PK场景下观众端在直播间的跳转时延、观众进入直播间的首屏时延。

表4 交互体验质量输入参数

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 序号 | 参数 | 数据源 | 定义/取值 | 备注 |
| I.16 |
| 5.4.1 | 交互行为标识 | 移动直播APP | 根据直播过程中的观众的交互行为，端到端时延、直播间跳转时延、进入直播间的首屏时延等。 | InteractionFlag，缩写为*ItractFlag* |
| 5.4.2 | 观众端端到端时延 | 移动直播APP | 从主播端采集到观众端播放需要的时间 |  |
| 5.4.3 | 观众在直播间跳转时延 | 移动直播APP | 主播连麦PK场景下，观众从一个主播直播间切换到另外一个主播直播间需要的时间 |  |
| 5.4.4 | 观众进入直播间的首屏时延 | 移动直播APP | 观众初始进入直播间到首帧出图的时间 |  |

5.5 观众端QoE模型输出参数

直播观众端QoE评估模型的输出信息，如表5所示。

表5 直播观众端QoE评估模型的输出参数

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 编号 | 参数 | 定义/取值 | 频次 | 备注 |
| 5.5.1 | O.21 | 对直播视频的画质评分。取值范围：1-5 | 直播每一个输出采样间隔（如：每1秒） | Visual Quality，缩写为*QV* |
| 5.5.2 | O.22 | 对直播视频的音质评分。取值范围：1-5 | 直播每一个输出采样间隔（如：每1秒） | Audio Quality，缩写为*QA* |
| 5.5.3 | O.31 | 对直播视频的视听质量的评分。取值范围：1-5 | 直播每一个输出采样间隔（如：每1秒） |  |
| 5.5.4 | O.32 | 对直播视频的视听质量的评分。取值范围：1-5 | 每一个直播会话 | Visual Audio Experience Quality，缩写为*QAVE* |
| 5.5.5 | O.33 | 网络传输质量（如带宽、丢包、时延）对直播视频体验质量影响的评分。取值范围：1-5 | 每一个直播会话 | Presenting Experience Quality，缩写为*QPE* |
| 5.5.6 | O.34 | 交互响应时延对直播体验质量影响的评分。取值范围：1-5 | 直播每一个输出采样间隔（如：每1秒） |  |
| 5.5.7 | O.35 | 交互响应时延对直播视频体验质量影响的评分。取值范围：1-5 | 每一个直播会话 | Interaction Experience Quality，缩写为*QInE* |
| 5.5.8 | *QStall* | 直播视频的卡顿感知质量。取值范围：1-5 | 每一个直播会话 | Stall Quality |
| 5.5.9 | *QBlock* | 直播视频的花屏感知质量。取值范围：1-5 | 每一个直播会话 | Block Quality |
| 5.5.10 | Qasyn | 直播视频的音画同步质量。取值范围：1-5 | 每一个直播会话 | Audio-to-video synchronization Quality |
| 5.5.11 | O.41 | 对直播视频体验质量的综合评分。取值范围：1-5 | 每一个直播会话 | Video Live broadcast MOS，缩写为MOS |

6 移动直播观众端用户体验评估算法

6.1 总体模型综述

即总体模型为三个模块（视听体验质量*QAVE*，呈现体验质量*QPE*，交互体验质量*QInE*）的函数关系，用于评估移动直播场景中观众端的用户体验 。

（1）

公式（1）中*v*1~*v*2 分别是呈现体验质量和交互体验质量的动态加权系数。初始权重系数值由大数据调研的结果得到，同时叠加惩罚因子，当分项的得分变化时，权值系数也会相应调整。

模型通过主观标注的打分数据库进行数据拟合或模型训练获得，数据库足够覆盖典型使用场景。

6.2 模型四大模块综述

1.
2.
3.
4.
5.
6. 1.
	2.

### 视听体验

根据ITU-T P.1203视频体验建模规范，本标准主要基于附录A所述的Mode 0方式，也即基于元数据信息的轻量级建模方式，构建视频质量O.21和音频质量O.22的评估模型。O.21和O.22的模型参数T为媒体长度，以秒为单位。如果同时考虑了音频和视频，并且它们的长度不相等，则应使用两个输入的较短持续时间。 较长的输入应在末尾被截断以匹配较短输入的持续时间。输入/输出采样的时间t的取值范围：0 – T，单位是秒。

6.2.1.1 视频质量

*QV*= *f* (*ScreenSize*, *ScreenRefreshRate*, *ScreenResolutionHorizontal*, *ScreenResolutionVertical*, *DistanceFromEyeToScreen*, *VideoResolutionHorizontal*, *VideoResolutionVertical*, *VideoBitrate*, *VideoFrameRate*, *VideoCodec*, *Bitdepth*,*TransferCharacteristics*,*ColorPrimaries*)

当基于Mode 0（从元数据中提取分辨率、码率、帧率、编码算法等信息）时，视频质量（O.21）主要考虑了移动直播观众端视频编解码、时域失真、因视频与屏幕分辨率不一致导致的空域失真等因素的影响。

（2）

（3）

（4）

（5）

（6）

（7）

（8）

（9）

公式（3）~（7）中根据不同的视频编码算法对应不同的模型系数v3~v15取值。

注：

1. 关于PPD计算方法的推导过程，可参考附录B。
2. 当视频和屏幕水平分辨率（即像素个数）不同时，公式（8）中应取二者的较小值，以便反映用户实际看到的水平像素个数。
3. 上述公式（8）PPD计算是以手机横屏观看直播为例，对于竖屏观看，PPD计算方法需要相应进行变更，上述公式中，公式如下所示：

6.2.1.2 音频质量

*QA*  *f (* *BrA*, *NoCA*, *CodecA* )

当基于Mode 0（从元数据中提取采样率、码率、声道数、编码算法等信息）时，音频质量（O.22）评估模型为：

（10）

 (11)

公式（11）中根据不同的音频编码算法和声道数对应不同的模型系数v16~v20取值。

6.2.1.3 视听体验质量评估

基于输出采样间隔的视听体验质量（O.31）兼顾了下述两个方面：视频质量*QV*（O.21），音频质量*QA*（O.22），其评估模型为：

（12）

公式（12）中*v*21~*v*24是模型系数。

最后，基于会话的视听体验质量（O.32，即*QAVE*）评估是在当前实时评分和上一时刻会话评分的基础上进行计算得到的。

（13）

公式（13）中是模型系数。

### 呈现体验

6.2.2.1 卡顿感知质量评估

在呈现体验中与卡顿相关的过程参数有2个，由原始入参通过一定运算得到：

* *StallNum*：卡顿次数，系原始入参
* *TotalStallLen*：卡顿时长的加权求和

1、TotalStallLen的计算

 为了计算*TotalStallLen*，必须首先为每个单独的卡顿事件i分配一个权重*w\_stalli*，具体取决于其在媒体会话中的位置，根据遗忘曲线进行衰减。

（14）

（15）

（16）

注：

1. 参数*StallNum*仅统计卡顿事件（不包括初始加载事件）。如果卡顿事件数为0，则*TotalStallLen*为0。

最后，卡顿感知质量（*QStall*）评估模型为：

（17）

（18）

公式（14）~（18）中~是模型系数。

6.2.2.2 花屏感知质量评估

在呈现参数提取模块中输出（花屏相关的）参数是*PLR*，系通过对*pPLRt*进行时变加权求和得到。其中，权重依赖于它在媒体会话中的位置，根据遗忘曲线进行衰减。

（19）

（20）

花屏感知质量（*QBlock*）评估模型为：

（21）

（22）

公式（19）~（22）中~是模型系数。

6.2.2.3 呈现体验质量评估

当终端对丢包的处理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（此时同时存在*QStall*和*QBlock*子体验项），也可能仅包括花屏；当终端对丢包的处理基于有丢包掩盖机制时，呈现体验损伤仅包括卡顿。

此外，考虑到音画同步的影响，故呈现体验质量（O.33）评估模型为：

（23）

（24）

公式（23）~（24）中是模型系数。

### 交互体验

对于移动直播，基于输出采样间隔的交互体验质量（O.34）的公式为：

（25）

公式（25）中，*DMOSom*(*ItractFlag*)为交互响应时延导致的交互一致性下降量。

其中，移动直播*DMOSom*(*ItractFlag*)与交互响应时延的关系可以表示为：

（26）

公式（26）中是模型系数。

最后，基于会话的交互体验质量（O.35）评估模型为：

（27）

（28）

（29）

公式（29）中是模型系数。

附录A ITU-T P.1203操作模式

根据ITU-T P.1203及P.1203.1/2规范，建议本标准的QV/QA评估模型基于Mode 0方式构建：

* Mode 0：从元数据中获得的信息，例如获取有关编解码器、编码和显示分辨率、码率、帧率，以及卡顿的信息。
* Mode 1：来自模式0的所有信息，以及基于数据包报头解析的其他视频和音频帧信息（例如视频帧的大小和持续时间，以及视频帧的类型（区分I帧和非I帧））。
* Mode 2：来自模式1的所有信息，以及高达2%（以字节为单位）基于深度数据包解析的整体媒体流信息和部分比特流解析。
* Mode 3：来自模式1的所有信息，以及基于比特流解析的完整媒体流信息。

不同操作模式之间的关系图4所示。



图 4 ITU-T P.1203.1操作模式

附录B 关于PPD计算方法

如图5所示，以手机屏幕横屏为例，PPD计算公式如下：

（1）

（2）

（3）

一般地，手机公开的基本参数是对角线尺寸*S*（单位：英寸），水平和垂直分辨率（像素个数），故根据公式（2）和（3）可知，

（4）

将公式（4）代入公式（1）可得：

（5）

同理可知，纵向PPD公式可得：

（6）

至此，公式（5）即为当横屏观看时手机屏幕PPD的最终计算公式，公式（6）即为当竖屏观看时手机屏幕PPD的最终计算公式。

此外，常见移动终端（智能手机、平板电脑）的舒适观看距离D典型值分别为30cm和41cm。



图 5 PPD计算示意图