

视频服务用户体验评估算法和参数-直播

The evaluation method and parameter for the User Experience of video service

2017年9月29日发布

2017年10月1日实施

中关村现代信息消费应用产业技术联盟 发布

目 次

前言.....	II
1 范围.....	1
2 规范性引用文件.....	1
3 缩略语、术语和定义.....	2
3.1 缩略语.....	2
3.2 术语和定义.....	3
4 视频服务用户体验指标体系.....	4
4.1 关键指标.....	4
4.2 指标框架.....	5
4.3 参数采集及指标处理.....	5
5 视频服务用户体验指标构成.....	6
5.1 交互体验质量参数.....	6
5.2 观看体验质量参数.....	7
5.3 视频体验质量参数.....	8
5.4 音频体验质量参数.....	9
5.5 其它业务 KQI 指标.....	9
5.6 其它网络 KPI 指标.....	10
6 视频服务用户体验评估算法.....	10
6.1 总体模型综述.....	10
6.2 模型四大模块综述.....	12
6.2.1 视频体验质量 Q_s	12
6.2.2 音频体验质量 Q_a	19
6.2.3 交互体验质量 Q_i	20
6.2.4 观看体验质量 Q_v	20

前 言

本标准是针对视频服务用户体验评估系列标准之一，本标准预计的架构如下：

- 视频体验评估的需求和场景
- 视频服务用户体验评估的系统架构
- 视频服务用户评估算法和参数-直播
- 视频服务用户评估算法和参数-点播
- 视音频业务主观评价技术专家组章程

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：中国电信集团公司

本标准主要起草人：中国信息通信技术研究院、国家新闻出版广电总局广播电视规划院、中国移动集团公司、中国联通集团公司、华为技术有限公司、腾讯公司、爱奇艺公司、优酷土豆股份有限公司、网宿科技股份有限公司、上海交大、博汇科技有限公司

视频服务用户体验评估算法和参数-直播

1 范围

本标准规定了应用于娱乐视频服务业务用户体验质量的评估场景和模型,分析了影响娱乐视频服务用户体验质量的关键因素,定义了用于评估视频服务用户体验质量的参数和技术方法。

本标准适用于对电信运营商、广电运营商、互联网视频服务商和其它相关厂商提供的娱乐视频服务用户体验质量进行综合评估,还适用于产业各方对影响视频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。

[ITU-T G.1081] IPTV性能监控点

[ITU-T Y.1991] IPTV的术语和定义

[ITU-T G.1080] IPTV服务的体验要求质量

[ITU-T P.912] 识别任务的主观视频质量评定方法

[ITU-R BT.1788 多媒体应用中视频质量主观评定的方法论

[RFC 4445] A Proposed Media Delivery Index (MDI), April 2006

[ETSI TR101 290 V1.2.1] Digital Video Broadcasting (DVB); Measurement Guidelines for DVB Systems, May 2001

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

[ITU-T P.1201] Recommendation ITU-T P.1201 (2012), Parametric non-intrusive assessment of audiovisual media streaming quality.

[ITU-T P.1201.1] Recommendation ITU-T P.1201.1 (2012), Parametric non-intrusive assessment of audiovisual media streaming quality – lower resolution application area.

[ITU-T P.1201.2] Recommendation ITU-T P.1201.2 (2012), Parametric non-intrusive assessment of audiovisual media streaming quality – higher resolution application area.

[ITU-T P.1202.1] Recommendation ITU-T P.1202.1 (2012), Parametric non-intrusive bitstream assessment of video media streaming quality – lower resolution application area.

[ITU-T P.1202.2] Recommendation ITU-T P.1202.2 (2013), Parametric non-intrusive bitstream assessment of video media streaming quality – Higher resolution application area.

[ITU-T P.1401] Recommendation ITU-T P.1401 (2012), Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models.

YD/T 1823-2008 IPTV业务系统总体技术要求

YD/T 2367-2011 IPTV质量监测系统技术要求

GYT 262-2012节目响度和真峰值音频电平测量算法

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。

BAR	花屏面积占比	Block Area Ratio
BTV	广播电视	Broadcast TV
CDN	内容分发网络	Content Delivery Network
CBR	恒定比特率编码	constant bit-rate
DPI	深度包检测	Deep Packet Inspection
FCC	快速频道切换	Fast Channel Change
FTP	文件传输协议	File Transfer Protocol
FR	全参考	Full Reference
GOP	图像组	Group of Pictures
HG	家庭网关	Home Gateway
HTTP	超文本传输协议	HyperText Transfer Protocol
IGMP	因特网组管理协议	Internet Group Management Protocol
IP	互联网协议	Internet Protocol
IPTV	IP 电视	Internet Protocol Television
ICP	因特网内容提供商	Internet Content Provider
HDR	高动态范围	High-Dynamic Range
HTTP	超文本传输协议	Hypertext Transfer Protocol
HLS	Apple 的动态码率自适应技术	HTTP Live Streaming
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
MOS	主观体验评分	Mean Opinion Score
MV	运动矢量	Motion Vector
NR	无参考	No Reference
OTT	互联网向用户提供各种应用服务	Over The Top
OQA	客观质量评估	Objective Quality Assessment
PPI	每英寸所拥有的像素数目	Pixels Per Inch
PON	无源光纤网络	Passive Optical Network
QP	量化参数	Quantization Parameter
QoE	体验质量	Quality of Experience
QoS	服务质量	Quality of Service
RR	部分参考	Reduced Reference
RTCP XR	实时传输控制协议扩展报告	Real-time Transport Control Protocol Extended Report
RTSP	实时流协议	Real Time Streaming Protocol
RTT	往返时延	Round-Trip Time
SMTP	简单邮件传输协议	Simple Message Transfer Protocol
sQuality	质量评分	score of Quality

sInteraction	互动体验评分	score of Interaction
sView	播放体验评分	score of View
sInteractionDeg	互动体验损伤	sInteraction Degradation
sViewDeg	播放体验损伤	sView Degradation
SQA	主观质量评估	Subjective Quality Assessment
SYN	同步信号	Synchronous
TCP	传输控制协议	Transmission Control Protocol
TR	花屏时长占比	Time Ratio
UDP	用户数据报协议	User Datagram Protocol
VC	视频内容复杂度	Video Complexity
vMOS	视频体验评分	video Mean Opinion Score
VBR	可变比特率编码	variable bit-rate

3.2 术语和定义

下列术语和定义适用于本文件。

3.1.1

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。

3.1.2

业务质量 quality of service

指网络利用各种技术，为指定的业务提供服务的能力。

3.1.3

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。

3.1.4

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。

3.1.5

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。

3.1.6

2k

2K分辨率指屏幕横向像素达到2000以上，主流2K分辨率为2560x1440。我国数字影院放映机主要采用这种分辨率。许多高端手机屏幕也开始使用这种分辨率。其他的2048×1536(QXGA)2560×1600(WQXGA)，2560×1440(Quad HD)也被注释成为不同的2K。

3.1.7

4K

4K分辨率是指水平方向每行像素值达到或者接近4096个，多数情况下特指4096*2160分辨率。而根据使用范围的不同，4K分辨率也有各种各样的衍生分辨率，例如Full Aperture 4K的4096*3112、Academy 4K的3656*2664以及UHDTV标准的3840*2160等，都属于4K分辨率的范畴。

3.1.8

视频直播 Video streaming

根据用户发出播放请求，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 视频服务用户体验指标体系

4.1 关键指标

用户使用视频服务的过程往往包括启动终端/播放器，接入服务系统，选择观看内容，进行播放、观看等操作。因此影响网络娱乐视频服务用户体验的因素是多方面的，在系统正常为用户提供服务的过程中，关键的因素主要包括视频源质量、音频源质量、用户的交互体验和观看体验。

视频源质量，取决于视频源的清晰度、流畅度、保真度（色调、对比度）等因素，涵盖了视频源的分辨率、帧率、码率、内容、编码和终端六个维度的指标，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。

音频源质量，取决于音频源的清晰度、流畅度、保真度等因素，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。

用户的交互体验，取决于系统对用户交互操作的响应速度，涵盖了平台、网络、终端性能指标，可以用户一次观看行为中对EPG操作的响应、视频初始加载时长、频道切换时长，快进快退的响应等客观指标度量。

观看体验，取决于观看过程中出现的节目信号质量，影响因素包括花屏、马赛克、卡顿、声画不同步等，可以通过视音频信息的传输性能和质量损伤等客观指标进行衡量。

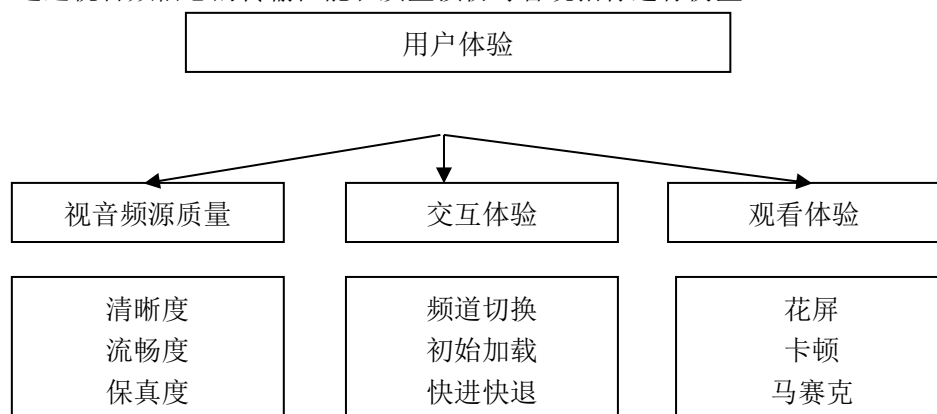


图1 影响用户体验的因素

为了科学的评估和量化视频服务的用户体验质量，把评估场景分为会话场景（表征用户一次完整的观看行为的体验质量）和实时场景（表征用户实时观看过程中的体验质量）。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视频服务的真实感知体验。

交互体验和观看体验客观指标对用户视频体验质量的影响权重随着观看行为的持续而弱化，当花屏/卡顿事件出现时，视频的质量体验立即下降，当花屏/卡顿结束，恢复正常播放时，用户体验就会逐渐缓慢恢复。假如后续一直能持续正常播放，实时的质量体验会逐渐恢复到正常值。用户的视频体验质量得分，综合了视频源质量、交互体验质量、观看体验质量，是真实反映用户体验感知的平均主观评分(MOS)。

4.2 指标框架

视频服务质量受到终端、网络、平台、片源等多方面因素影响。具体的指标体系如下图所示：

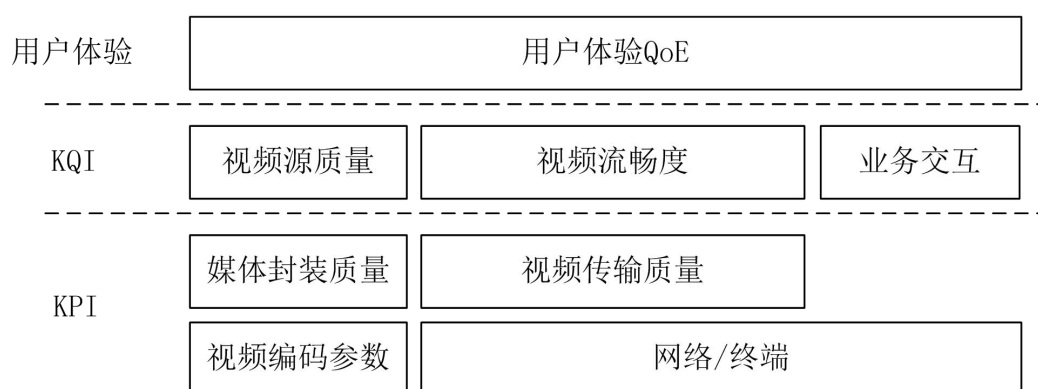


图2 视频质量指标框架

数据源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指标一般指基于网元设备的关键指标，主要包括终端性能、网络传输质量、视频传输质量、音频体验质量、媒体的封装质量和视频的编码参数。

KQI指标从业务和应用层面出发反映业务使用体验，视频播放的流畅度，即视频在播放过程中是否有出现缓冲卡顿、花屏和马赛克的情况；视频源质量，包括视频源本身的编码格式、视频清晰度、媒体封装等；业务交互体验是视频在观看过程中业务交互的情况包括直播业务的交互成功率以及交互时延等。KQI指标一般通过KPI指标处理得到。

QoE指标从用户层面出发，可以定义为终端用户对视频业务的总体主观感知。视频业务的感知体验就是反映用户观看视频业务体验的综合评分，用U_MOS表示。

定义 $U_MOS=f(Q_s, Q_a, Q_i, Q_v)$ ，即用户体验综合评分为（视频质量（ Q_s ），音频质量 Q_a ，交互质量（ Q_i ），观看质量（ Q_v ））的函数关系。

4.3 参数采集及指标处理

由于视频服务最终通过终端呈现给用户，是视频服务系统中离用户最近的环节，因此评估影响视频服务质量的参数来源可来自终端的网络抓包数据、终端系统信息相关数据以及播放器反映的视音频信息相关的数据。

参数的获取可以通过终端如机顶盒、播放器等设备，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指标就组成了视频业务质量分析的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。终端获取的参数经过数据处理也可以获取用户的观看视频行为的指标。

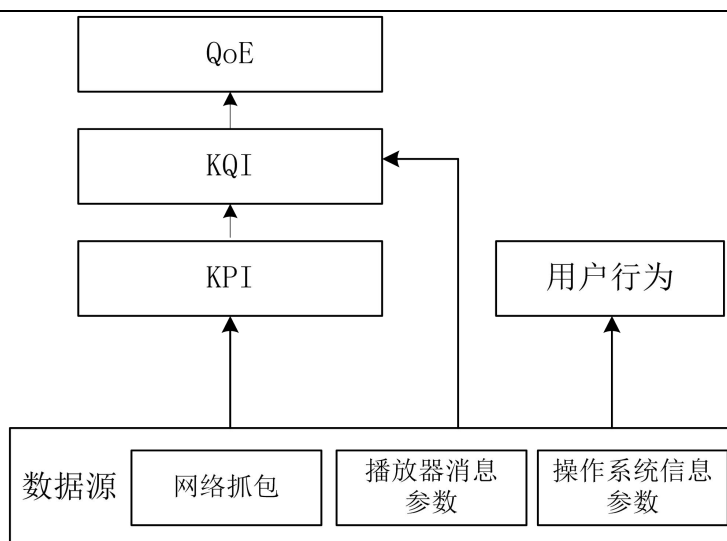


图3 数据处理流程

5 视频服务用户体验指标构成

5.1 交互体验质量参数

表1 交互体验

序号	参数	数据源	说明	备注
5.1.1	直播请求次数	终端播放器 终端网络抓包	直播请求事件的次数	
5.1.2	直播请求成功次数	终端播放器 终端网络抓包	直播请求成功的次数	
5.1.5	EPG 请求次数	终端网络抓包	EGP 请求事件的次数	
5.1.6	EPG 请求成功次数	终端网络抓包	EGP 请求成功的次数	
5.1.7	索引文件请求次数 (HLS)	终端播放器	索引文件请求事件的次数	
5.1.8	索引文件请求成功次数 (HLS)	终端播放器	索引文件请求成功的次数	
5.1.9	直播请求时延	终端播放器	发起直播业务请求到解出第一个 I 帧的时间间隔	$t_{zapping}$
5.1.12	EPG 请求时延	终端网络抓包	发起 HTTP GET 请求到收到响应数据之间的时间间隔	
5.1.13	异常中断的次	终端播放器	业务在使用过程中发生异常退	

	数		出观看的次数	
5.1.14	快进、快退、 暂停响应时间	终端播放器	业务使用过程中发起快进、快退、暂停的响应时间	

5.2 观看体验质量参数

表2 观看体验

序号	参数	数据源	说明	备注
5.2.1	缓存卡顿次数	终端播放器 (解码器)	在观看视频过程中发生缓冲区清空引起视频缓冲停顿的次数	<i>BufferFrequency</i>
5.2.2	缓存卡顿总时长	终端播放器	在观看视频过程中发生缓冲区清空引起视频缓冲停顿的时间累计	<i>BufferLength</i>
5.2.3	缓存卡顿持续时长	终端播放器	在观看视频过程中发生缓冲区清空引起视频缓冲停顿的时长	<i>BufferStartTime</i>
5.2.4	画面花屏次数	终端播放器 (解码芯片)	视频播放期间, 画面花屏次数,	<i>V_PLEF</i> (解码芯片提供或者通过丢包估算)
5.2.5	画面花屏总时长	终端播放器 (解码芯片)	视频播放期间, 画面花屏时长, 单位毫秒(ms),	中间变量, 用于计算 <i>V_NDF</i> (解码芯片提供或者通过丢包估算)
5.2.6	画面平均花屏面积百分比	终端播放器 (解码芯片)	画面发生花屏期间, 平均花屏面积百分比(%), 取值范围为[1%, 100%]; 其他情况下为0	<i>V_IRpF</i> (解码芯片提供或者通过丢包估算)
5.2.7	屏幕大小	终端系统软件	终端屏幕大小	可以通过 <i>HDMI EDID</i> 接口获取 <i>ScreenSize</i>
5.2.8	屏幕分辨率	终端系统软件	终端屏幕分辨率	可以通过 <i>HDMI EDID</i> 接口获取 <i>Screendpi</i>

5.3 视频体验质量参数

表3 视频质量

序号	参数	数据源	说明	备注
5.3.1	视频码率	终端播放器	所有视频的码率 RTSP 高清标清 4K HLS 索引文件申明的 编码速率	<i>BitRate</i>
5.3.2	视频帧率	终端播放器	视频的帧率	<i>VideoFrameRate</i>
5.3.3	视频分辨率	终端播放器	视频的分辨率	<i>Resolution</i>
5.3.4	编码方式	终端播放器	视频的编码方式： H.264、H.265、 MPEG-2;	<i>CodecType</i>
5.3.5	编码参数	终端播放器	每帧视频的类型： I、P、B_ref、B_noref	注 1：采集以 GOP 为单元， 输入视频片段内各帧的关 键参数。 注 2：计算 Skip 模式和非 Skip 模式比例时，H.264、 H.265 下分别以 MB、CU 为单位
			帧级比特大小	
			帧级 QP 三元组： 最大，最小，平均值	
			最大参考帧数目	
			帧级 MV 三元组： 最大，最小，平均值	
帧级 Skip 与非 Skip 模式比例，I 帧为 0				
5.3.6	信号层参数	源端检测	块状度	
			模糊度	
		终端播放器	对比度	
			噪点度	
			色彩度	
5.3.7	编码方式		编 码 方 式： CBR,VBR	

5.4 音频体验质量参数

表 4 音频质量

序号	参数	数据源	说明	备注
5.4.1	音频码率	终端播放器	音频数据每秒所需要的平均比特数	<i>BitRate</i>
5.4.2	音频采样率	终端播放器	音频数据每秒采样点数	<i>SampleRate</i>
5.4.3	音频声道数	终端播放器	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	<i>NumberOfChannels</i>
5.4.4	编码方式	终端播放器	音频的编码方式： 无损：PCM 有损：MP3、WMA、AAC、AC3	<i>CodecType</i>
5.4.5	编码参数	终端播放器	AAC 编码中的尺度因子	<i>scalefactor</i>
5.4.6	信号层参数	源端检测点 终端播放器	响度 动态范围 左右声道相位差 爆音 静音长度	
5.4.7	编码方式	终端播放器	编码方式：CBR,VBR	

5.5 其它业务 KQI 指标

索引获取时间

索引获取时间指标定义为终端从向视频调度服务器发起视频索引请求，到接收到视频调度服务器返回的视频索引信息的时长。

CDN地址获取时间

CDN地址获取时间指标定义为终端从向视频调度服务器发起分片CDN信息请求，到接收到视频调度服务器返回的分片CDN所在IP地址等信息的时长。

视频分片TCP建连时间

视频分片TCP建连时间指标定义为终端建立下载某个视频分片的TCP连接，从客户端发出第一个SYN报文到收到第一个SYN ACK报文的时间。

分片出错次数

分片出错次数指标定义为在单位观看时间内，由于视频CDN服务器自身问题主动返回ERROR或重定向等报文，导致视频播放中断的次数。

分片下载平均速率

分片下载平均速率指标定义为从终端向视频CDN服务器发起HTTP GET下载请求起，到分片文件完整地传输到终端为止，这期间的视频分片下载平均速率。

分片峰值下载速率

分片峰值下载速率指标定义为从终端向视频CDN服务器发起HTTP GET下载请求起，到分片文件完整地传输到终端为止，这期间的视频分片下载的峰值速率。

分片资源分布

分片资源分布指标是所有被测视频分片按所在服务器IP地址归属地进行分组统计并进行统计计算形成的数据记录表。

其它参数有待进一步补充。

5.6 其它网络 KPI 指标

网络KPI决定了用户终端到互联网内容的端到端实际可用吞吐量，影响互联网内容传输到用户终端的时间，从而影响用户体验QOE。影响用户业务体验的网络KPI指标主要考虑：

接入带宽

运营商提供给固定宽带用户的承诺接入带宽。

双向时延

客户端（端）与测量时播放的视频分片文件所在服务器之间的双向时延。该指标同时适用于固定宽带互联网用户和移动互联网用户。

丢包率

客户端与测量时播放的视频分片文件所在服务器之间的管道双向丢包率。该指标同时适用于固定宽带互联网用户和移动互联网用户。

其它参数有待进一步补充。

6 视频服务用户体验评估算法

6.1 总体模型综述

$U_MOS=f(Qs, Qa, Qi, Qv)$ ，即总体模型为四个模块（视频体验质量 Qs ，音频体验质量 Qa ，交互体验质量 Qi ，观看体验质量 Qv ）的函数关系，具体功能场景分为会话场景（用户一次观看行为时长大于1分钟），及实时场景（用于实时质量监控）。

$$U_MOS = (Q_{av} - 1) \cdot (1 - c_1 \cdot (5 - Qi) + c_2 \cdot (5 - Qv)) + 1$$

Q_{av} 是视音频总体体验质量：

$$Q_{av} = \alpha + \beta \cdot Qs + \gamma \cdot Qa \cdot Qs$$

$c_1 \sim c_2$ 分别是交互体验和观看体验的动态加权系数。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。

根据具体视频服务的特点，直播模式的应用场景下，可分别采用 RTSP/RTP 传输协议和 HTTP 传输协议，对应的终端播放也可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出现丢包时，解码器直接丢弃问题帧，体验表现为视频卡顿。在无丢包掩盖机制下，解码器直接播放问题帧，体验表现为花屏和视频卡顿）。具体影响实时和会话体验的计算变量如下表所示：

			实时(Instant)	会话(Session)
直播	RTSP/RTP	丢包掩盖	$Q_s, Q_a,$ $Q_i = Q_i_ZapInstant$ $Q_v = Q_v_STInstant$	$Q_s, Q_a,$ $Q_i = Q_i_ZapSession$ $Q_v = Q_v_STSession$
		无丢包掩盖	$Q_s, Q_a,$ $Q_i = Q_i_ZapInstant$ $Q_v = Q_v_STInstant + Q_v_BLInstant$	$Q_s, Q_a,$ $Q_i = Q_i_ZapSession$ $Q_v = Q_v_STSession + Q_v_BLSession$
	HTTP	$Q_s, Q_a,$ $Q_i = Q_i_LoadInstant$ $Q_v = Q_v_STInstant$	$Q_s, Q_a,$ $Q_i = Q_i_LoadSession$ $Q_v = Q_v_STSession$	

6.2 模型四大模块综述

6.2.1 视频体验质量 Q_s

序号	变量	物理意义	计算所需变量
1.a	<i>ScreenSize</i>	视频显示屏幕尺寸, 对角线尺寸 (英寸)	5.2.7
1.b	<i>Resolution</i>	视频分辨率	5.3.3
1.c	<i>BitRate</i>	视频码率	5.3.1
1.d	<i>CodecType</i>	视频编码类型	5.3.4
1.e	<i>VideoFrameRate</i>	帧率	5.3.2
1.f	<i>PPI</i>	电视机分辨率 (默认情况下, PPI 的分辨率与视频内容的分辨率一致)	5.2.8
1.g	<i>FrameType</i>	GOP 中视频的帧类型 (I/P/B)	5.3.5
1.h	<i>TotalBpF</i>	GOP 中各类型帧对应的比特率	
1.i	<i>Qp</i>	帧级 QP (最大值, 最小值, 均值)	
1.j	<i>MV</i>	帧级运动矢量 MV (最大值, 最小值, 均值)	
1.k	<i>SkipRatio</i>	Skip 模式和非 Skip 模式比例	
1.l	<i>MaxRef</i>	最大参考帧数目	
1.m	<i>Blockiness</i>	块度	5.3.6
1.n	<i>Blurriness</i>	模糊度	5.3.6
1.o	<i>Chroma</i>	色彩度	5.3.6

1.p	<i>Contrast</i>	对比度	5.3.6
1.r	<i>Noise</i>	噪点度	5.3.6

$$Q_s = f(\text{ScreenSize}, \text{VideoComplexity}, \text{Resolution}, \text{BitRate}, \text{CodecType}, \text{VideoFrameRate}, \text{Reserved}...)$$

注:

- 1) 计算所需变量中的符号解读: 5.1.3 表示正文 5.1 节中第 3 个指标。1.a 表示附录章节 1 的第一个变量。
- 2) Q_s 模块按照其处理信号层级主要包含 Model0, Model1, Model2 三层模型。其中 Model0 显式包含 *BitRate*, *ScreenSize*, *PPI* 变量, 其余变量在不同的应用场景下(分辨率、编码器类型)为定值。Model1 需要从编码数据包及比特流中采集视频帧关键编码信息(详见 5.3.5), Model2 需要从原始或播放器连续采集视频帧的关键质量信息(详见 5.3.6), Model1 和 Model2 中采集到的 KPI 相关信息到 Q_s 的具体计算公式将在后续版本中定义。

ParamedicModel (mode 0)

$$Q_s' = f(PPI, \text{ScreenSize}) = c_7 \cdot \left(1 - \frac{1}{1 + \left(\frac{PPI}{c_8 \cdot \text{ScreenSize}^{c_9}}\right)^{c_{10}}}\right) \quad \backslash * \text{MERGEFORMAT (0-0-1)}$$

$$Q_s = Q_s' - \frac{Q_s' - c_{11}}{1 + \left(\frac{\text{BitRate}}{c_{12}}\right)^{c_{13}}} \quad \backslash * \text{MERGEFORMAT (0-0-2)}$$

在不同尺寸的屏幕显示视频时, 需要用有效显示 PPI 及屏幕尺寸 (ScreenSize) 来评估在特定设备上播放视频时用户的体验质量。有效显示 PPI 由原视频的分辨率, 屏幕分辨率以及屏幕尺寸计算得到。

其中, Q_s' 表示显示质量, 即在固定 PPI 和屏幕尺寸条件下(如 720P 视频在 42" 电视播放)能够取得的最高质量得分; c_{11} 表示相应最低分值, 本文中取 1.0, 以上的系数均为大数据观测统计所得。

根据不同的编码方式(H.265、H.264、MPEG-2)有不同的对应 c_{12} 和 c_{13} 取值。

Bitstream Model (Mode 1)

在不同尺寸的屏幕显示视频时, 需要用有效显示 PPI 及屏幕尺寸 (ScreenSize) 来评估在特定设备上播放视频时用户的体验质量, 公式引用 * MERGEFORMAT (0-0-3);

$$Q_s' = f(PPI, \text{ScreenSize}) = c_7 \cdot \left(1 - \frac{1}{1 + \left(\frac{PPI}{c_8 \cdot \text{ScreenSize}^{c_9}}\right)^{c_{10}}}\right) \quad \backslash * \text{MERGEFORMAT (0-0-4)}$$

视频部分的质量公式为:

$$Q_s = Q_s' - (5 - Q_{cod}) * \frac{(Q_s' - 1)}{4} \quad \backslash * \text{MERGEFORMAT (0-0-5)}$$

Q_{cod} 为视频序列的质量分析结果, 取值范围 [1, 5], 通过线性映射到取值区间 [1, Q_s];

Mode 1 视频源质量可以表示为如下函数映射关系(注: Mode 1 参数统一由字母 $\{n_i\}$ 表示。):

$$Qcod = f(\text{FrameType}, \text{TotalBytesPerFrame}, Qp, MV, \text{SkipRatio})$$

$$= f \left(\begin{array}{l} \text{FrameType}, \text{TotalBytesPerFrame}, \\ [Qp_max, Qp_avg, Qp_min], \\ [MV_max, MV_avg, MV_min], \\ \text{SkipRatio} \end{array} \right)$$

* MERGEFORMAT (0-0-6)

$$Qcod = kfr_{impact} \cdot \exp \left(n_1 \cdot \left(QP_fr_{impact} + cpx_{video} + motion_{impact} \right) \right)$$

* MERGEFORMAT (0-0-7)

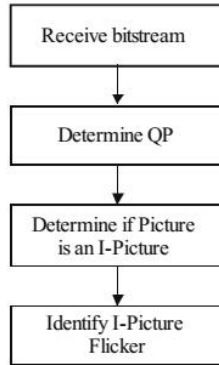
$$kfr_{impact} = n_2 \cdot kfr + n_3 \quad \backslash * \text{MERGEFORMAT (0-0-8)}$$

其中 kfr 表示视频帧率与两个相邻 I 帧之间的帧数的比值。

$$QP_fr_{impact} = n_4 + n_5 \cdot \left(\frac{avgQP_{pic}}{51} \right)^{n_6} + n_7 \cdot \frac{1}{fr} + n_8 \cdot intraflicker \quad \backslash * \text{MERGEFORMAT (0-0-9)}$$

$$+ n_9 \cdot (maxQP_{pic} - minQP_{pic})$$

其中 $intraflicker$ 为一个布尔型变量,表示到目前接受到的帧为止,是否出现 I-帧闪烁现象,所谓 I-帧闪烁是指某一接收到的 I-帧的 QP 值较之前及之后的 I-帧的 QP 值相比,有明显突变(增加)的现象。通过如下算法确定。



若某一 I-帧 QP 值满足以下条件,即可判定为发生了 I-帧闪烁现象

1. 当前 I-帧 QP 值比前一帧 QP 值高 5
2. 当前 I-帧 QP 值比后一帧 QP 值高 5

$$cpx_{video} = \min \left(\sqrt{\frac{br_v}{AvgByte_{1-frame}}} + n_{10} \cdot skipRatio, 1.0 \right) \quad \backslash * \text{MERGEFORMAT (0-0-10)}$$

$$motion_{impact} = n_{11} \cdot avgMV_{pic} \cdot \left(1 - \frac{fr}{30} \right) \quad \backslash * \text{MERGEFORMAT (0-0-11)}$$

Hybrid Model (Mode 2)

$$Qs = f(\text{Blockiness}, \text{Blurriness}, \text{Contrast}, \text{Noise})$$

1) 块效应 Blockings: 块效应是指在编码过程中造成的块边界不连续的情况,块效应严重时,视频在将出现明显块状缺陷,影响视觉效果,降低观赏感受。

原理公式

$$blockiness = (-10.38 + 17.86 * blockiness_temp) / 2$$

$$blockiness_temp = \frac{globalInnerSum}{globalOuterSum}$$

具体计算过程:

将一帧视频分为 $L \times L$ 块, w_n , 其中 n 为块下标 ($1 \leq n \leq \lfloor \frac{M}{L} \rfloor + \lfloor \frac{N}{L} \rfloor$),

序号	公式	复杂度
(1)	计算每一分块最后一列与倒数第二列像素灰度值差值绝对之和 $verticalInnerSum$	$O(M*N)$
(2)	计算每一分块最后一列与下一分块第一列像素灰度值差值绝对之和 $verticalOuterSum$	$O(M*N)$
(3)	计算每一分块最后一行与倒数第二行像素灰度值差值绝对之和 $horizontalInnerSum$	$O(M*N)$
(4)	计算每一分块最后一行与下一分块第一行像素灰度值差值绝对之和 $horizontalOuterSum$	$O(M*N)$
(5)	$globalInnerSum = \sum_n verticalInnerSum + \sum_n horizontalInnerSum$	$O(\lfloor \frac{M}{L} \rfloor * \lfloor \frac{N}{L} \rfloor)$
(6)	$globalOuterSum = \sum_n verticalOuterSum + \sum_n horizontalOuterSum$	$O(\lfloor \frac{M}{L} \rfloor * \lfloor \frac{N}{L} \rfloor)$

其中,

M : 第 n 帧横向像素个数

N : 第 n 帧纵向像素个数

f : 视频序列的帧数

2) 模糊度 **Blurriness**: 模糊度是指对图像模糊程度的衡量, 模糊是指基于图像像素灰度的梯度幅度变化, 该变化可以表征图像边缘信息, 当梯度幅度过小时, 该条边缘将不够清晰明显, 影响视觉效果, 降低观赏感受。

原理公式

$$blurriness = 5 \cdot \left(\frac{\exp(-1.5 \cdot blur_temp + 2.87)}{1 + \exp(-1.5 \cdot blur_temp + 2.87)} \right)^{0.14}$$

$$blur_temp = \frac{edgewidth}{edgeNo}$$

具体计算过程:

视频序列复杂度: $O(f*M*N)$

将一帧视频分为 $L \times L$ 块, w_n , 其中 n 为块下标 ($1 \leq n \leq \left\lfloor \frac{M}{L} \right\rfloor + \left\lfloor \frac{N}{L} \right\rfloor$),

序号	公式	复杂度
(1)	对每一块 w_n , 估算标准差 $std_n = \sqrt{\frac{1}{Q-1} \left(\sum f(i,j)^2 - \frac{1}{Q} (\sum f(i,j))^2 \right)}, \quad Q = \left\lfloor \frac{M}{L} \right\rfloor + \left\lfloor \frac{N}{L} \right\rfloor$	$O\left(\left\lfloor \frac{M}{L} \right\rfloor * \left\lfloor \frac{N}{L} \right\rfloor\right)$
(2)	计算视频帧每一像素的水平方向及垂直方向梯度如下, 形成水平方向, 垂直方向梯度矩阵 $gradX = \{f(i,j) - f(i-1,j)\}$ $gradY = \{f(i,j) - f(i,j-1)\}$ $grad = gradX * gradY$	$O(M*N)$
(3)	$cutoff = 16 * \frac{\sum_i grad}{grad.height * grad.width}$	$O(1)$
(4)	通过如下方式判断像素点是否属于边缘: $x = grad(i,j) > cutoff$ $y = (gradX(i,j) \geq gradY(i,j)) \& \& (grad(i,j) > grad(i-1,j))$ $\& \& (grad(i,j) > grad(i+1,j))$ $z = gradY(i,j) \geq 0 \& \& (grad(i,j) > grad(i,j-1))$ $\& \& gradY(i,j) \geq 0 \& \& (grad(i,j) > grad(i,j+1))$ $if(x \& \& (y z))$, 则标记当前 $f(i,j)$, 为边缘	$O(M*N)$
(5)	通过如下方式计算边缘的宽度:	$O(M*N)$

	<pre> tempCenter = center if(gradX (i,center) < 0 gradX (i,center +1)< 0) { edgesNo++; while(- gradX (i,tempCenter+1) > std_n) { tempCenter++; } edgeswidth += tempCenter - center; tempCenter=center while(- gradX (i,tempCenter) > std_n) { tempCenter--;} edgeswidth += tempCenter - center; else if(gradX (i,center) >0 gradX (i,center +1)< 0) { edgesNo++; while(- gradX (i,tempCenter+1) > std_n) { tempCenter++; } edgeswidth += tempCenter - center; tempCenter=center while(- gradX (i,tempCenter) > std_n) { tempCenter--;} edgeswidth += tempCenter - center; </pre>		
(6)	$blur = \frac{edgewidth}{edgeNo}$		O(1)

其中, $f(m,n)$: 第 i 帧像素点 (m,n) 的灰度值

M : 第 n 帧横向像素个数

N : 第 n 帧纵向像素个数

$edgewidth$: 边缘宽度

$edgewidth$: 边缘数量

f : 视频序列的帧数

3) 对比度 Contrast: 合理的对比度可以显示生动、丰富的色彩, 展现更多的细节、更好的清晰度以及灰度层次。然而对比度过高则视频将产生失真感, 对比度过低画面则将表现为灰蒙蒙。不合理的对比度将影响人眼的主观感受

原理公式

$$C_n = \sqrt{\frac{1}{MN-1} \left(\sum f(i,j)^2 - \frac{1}{MN} (\sum f(i,j))^2 \right)}, \quad O(M*N)$$

具体计算过程:

每一帧复杂度: $O(M*N)$

视频序列复杂度: $O(f*M*N)$

其中,

M : 第 n 帧横向像素个数

N : 第 n 帧纵向像素个数

C_n : 第 n 帧的图像对比度

4) 噪点度 Noise: 噪点度定义对像素色度值的浮动的衡量, 该种浮动对图片整体质量无正影响且无固定规律。在被压缩视频中通常存在多种噪点类型。其中最常见的是量化噪声及蚊式噪声。其中量化噪声主要由于对像素值的量化引入, 其分布具有小范围随机特性, 且在整体图像上分布不统一。蚊式噪声主要表现为物体边缘的噪声以及运动纹理干扰。

原理公式:

$$N_n = \frac{\sum_i \sum_j (G_{i,j} - \{w_n\}_{i,j})}{G.height * G.width}$$

具体计算过程:

每一帧复杂度: $O(M*N)$

视频序列复杂度: $O(f*M*N)$

将一帧视频分为 $L \times L$ 块, w_n , 其中 n 为块下标 ($1 \leq n \leq \lfloor \frac{M}{L} \rfloor + \lfloor \frac{N}{L} \rfloor$),

序号	公式	复杂度
(1)	对每一块 w_n , 估算标准差 $std = \sqrt{\frac{1}{Q-1} \left(\sum f(i,j)^2 - \frac{1}{Q} (\sum f(i,j))^2 \right)}, Q = L*L$	$O\left(\left\lfloor \frac{M}{L} \right\rfloor * \left\lfloor \frac{N}{L} \right\rfloor\right)$
(2)	计算各块标准差统计直方图 $StdHisto_n(k) = StdHisto_n(k) + 1; \text{ if } std_n = k$	$O\left(\left\lfloor \frac{M}{L} \right\rfloor * \left\lfloor \frac{N}{L} \right\rfloor\right)$
(3)	Find $max_std = \max(StdHisto_n(k))$	$O\left(\left\lfloor \frac{M}{L} \right\rfloor * \left\lfloor \frac{N}{L} \right\rfloor\right)$
(4)	$var0 = max_std, val1 = max_std + 1, val2 = max_std - 1$	$O(1)$
(5)	$weighted_std = var0 * StdHiston(maxstd) +$ $var1 * StdHiston(maxstd + 1) +$ $var2 * StdHiston(maxstd - 1)$	$O(1)$
(6)	Find the blocks $\{w_n^{select}\}$ whose $std_{w_n} = \lfloor weighted_std \rfloor$ $\text{or } std_{w_n} = \lceil weighted_std \rceil$	$O\left(\left\lfloor \frac{M}{L} \right\rfloor * \left\lfloor \frac{N}{L} \right\rfloor\right)$

(7)	$\text{for } i = -1:1, j = -1:1$ $\text{mask}(i, j) = \frac{1}{2 * 3.14 * ([\text{weighted_std}])^2} \exp\left(\frac{-(i^2 + j^2)}{2 * ([\text{weighted_std}])^2}\right)$	O(M*N)
(8)	$\text{for } i = -1:1, j = -1:1$ $G = \{w_n^{\text{select}}(i, j)\} * \text{mask}(i, j)$	O(M*N)

其中, $Mask^*$: 与每帧图像做卷积的算子, 卷积窗口大小为 $3*3$

$\{w_n\}$: 提取出的符合要求的图像块集合

$G_n(x,y)$: 第 n 帧像素点 (x,y) 与算子做卷积后的计算值

N_n : 第 n 帧图像的噪点度

根据训练数据获得上述 4 个测度与 MOS 之间的分段对应关系 Q_{s_i} , 然后加权平均获得总的 Q_s :

$$Q_s = f(\text{Blockiness}, \text{Blurriness}, \text{Contrast}, \text{Noise}) = \frac{1}{4} \sum_i \alpha Q_{s_i}$$

$$i = \{\text{Blockiness}, \text{Blurriness}, \text{Contrast}, \text{Noise}\}$$

6.2.2 音频体验质量 Q_a

序号	变量	物理意义	计算所需变量
2.a	<i>BitRate</i>	音频数据每秒所需要的平均比特数	5.4.1
2.b	<i>SampleRate</i>	音频数据每秒采样点数	5.4.2
2.c	<i>NumberOfChannels</i>	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	5.4.3
2.d	<i>CodecType</i>	音频的编码方式	5.4.4
2.e	<i>Scalefactor</i>	AAC 编码中的尺度因子	5.4.5
2.g	<i>Loudness</i>	响度	5.4.6
2.h	<i>Dynamic Range</i>	动态范围	5.4.6
2.i	<i>Skewing</i>	左右声道相位差	5.4.6
2.j	<i>Sonic boom</i>	爆音	5.4.6
2.k	<i>Mutelength</i>	静音长度	5.4.6

$$Q_a = f(\text{Bitrate}, \text{SampleRate}, \text{NumberOfChannels}, \text{CodecType}, \text{Scalefactor}, \text{Loudness}, \text{DynamicRange} \dots)$$

注: 按照其处理信号层级主要包含 Model0, Model1, Model3 三层模型。其中 Model0 包含 BitRate、CodeType 等变量, Model1 包含 BitRate, SampleRate, Number of Channels, CodeType, Scalefactor 变量。Model2 则需要从原始音频或播放器连续采集音频 PCM, 获得 Loudness、Dynamic Range、Skewing 等关键质量信息。

Basic model (Mode0)

$$Q_{coda} = a_1 \cdot \exp(a_2 \cdot \text{bitrate}) + a_3 \quad \backslash * \text{MERGEFORMAT (1-4-1)}$$

其中根据不同的编码方式（E-AC3、AAC、MPEG1 Layer2、Layer3）有不同的对应 a_1 、 a_2 和 a_3 取值。总体音频质量如下表示：

$$Q_a = 100 - Q_{coda} \quad \backslash * \text{MERGEFORMAT (1-4-2)}$$

Q_a 可通过如下方式映射为 0~5 分 MOS 值：

```
MOS_max = 4.9;
MOS_min = 1.05;
if (Qa>0 & Qa<100)
```

$$MOS = 1 + \frac{MOS_max - MOS_min}{100} Q_a + Q_a \cdot (Q_a - 60) \cdot (100 - Q_a) \cdot 7.0E - 6$$

* MERGEFORMAT (1-4-3)

```
else if (Qa>=100)
MOS = MOS_max
else
MOS = MOS_min
end
```

Paramedic Model (Mode1)

$$Q_a = f(\text{BitRate}, \text{SampleRate}, \text{NumberofChannels}, \text{CodeType}, \text{Scalefactor})$$

Bitstream Model (Mode2)

$$Q_a = f(\text{Loudness}, \text{Dymanicrange}, \text{Skewing}, \text{Sonicboom}, \text{Mutelength})$$

6.2.3 交互体验质量 Q_i

序号	变量	物理意义	计算所需变量
2.a	$Q_{Zapping}$	直播时频道切换时长	2.c
2.c	$t_{zapping}$	直播时频道切换时延	5.1.9
2.e	t	当前已播放时长	系统自动记录
2.f	T	最大遗忘时长	系统参数

交互体验质量首先考虑频道切换和初始播放时延因素对体验的影响，其它因素可在后续进一步补充。

$Q_i_ZapInstant$ 、 $Q_i_ZapSession$

$$Q_{Zapping} = c_{15} \cdot \exp(c_{16} \cdot t_{zapping}) + c_{17} \cdot \exp(c_{18} \cdot t_{zapping}) \quad \backslash * \text{MERGEFORMAT (1-4-4)}$$

$$Q_i_ZapInstant = Q_{Zapping} \quad \backslash * \text{MERGEFORMAT (1-4-5)}$$

会话评分是在实时评分的基础上，根据遗忘曲线进行衰减：

$$Q_i_ZapSession = (\alpha_1 \cdot \exp(\beta_1 \cdot \frac{t}{T}) + \alpha_2 \cdot \exp(\beta_2 \cdot \frac{t}{T})) \cdot QZapping \quad \backslash * MERGEFORMAT (1-4-6)$$

6.2.4 观看体验质量 Q_v

序号	变量	物理意义	计算所需指标	计算所需公式
3.a	$Q_{Blocking}$	基于花屏占比进行评分	3.b、3.g、3.h	(3-1-2)
3.b	V_{AIRF}	视频帧的平均损伤率（面积占比）	3.c、3.d	* MERGEFORM
3.c	V_{IRpF}	视频帧的损伤率，出现损伤/降质的视频帧中，受影响的面积比例。 在头端通过播放器解视频流获得。如果头端不支持上报，或者在网络侧，可以通过解析视频的 IP 流，基于参数模型折算。详见 [ITU-T P.1201.1] 3.2.2 “Basic video parameter calculation module”	5.2.6	——
3.d	V_{NDF}	视频流中损伤/降质视频帧的数量	5.2.5、5.3.2	—
3.e	V_{TNF}	视频流中的视频帧的总数	5.3.9	——
3.f	V_{IR}	Block Area Ratio, 视频流损伤率（即花屏时长占比）	3.d、3.e	* MERGEFORM
3.g	V_{PLEF}	视频损伤/降质发生的次数	5.2.4	——
3.h	V_{CCF}	视频复杂度，通过对视频流的分析获得；当提取不到内容复杂度时，可用经验值替代	5.3.7	——
3.i	$Frequency$	视频观看过程中发生缓冲的次数	5.2.1	——
3.j	$Duration$	所有缓冲事件长度的平均值，所有的缓冲时长之和除以缓冲次数	3.i、3.k	¥* MERGEFOR
3.k	$BufferLength_i$	第 i 次卡顿的时长	5.2.2	——
3.l	$Interval$	多次缓冲情况下，缓冲间隔的平均值；（只有多于一个重缓冲事件发生（ $Frequency > 1$ ）时才使用）	3.i、3.k、3.m	¥* MERGEFOR

3.m	<i>BufferStartTime_i</i>	第 i 次卡顿发生的时间	5.2.3	—
-----	------------------------------------	--------------	-------	---

观看体验质量首先考虑Block和stalling因素对体验的影响，其它因素可在后续进一步补充。

$Qv_BLInstant$ 、 $Qv_BLSession$

$$Qv_BLInstant = 5 - QBlocking \quad \backslash * \text{MERGEFORMAT (1-4-14)}$$

$$QBlocking = f(BlockingTimeRatio, BlockingAreaRatio, Frequency)$$

$$= \frac{\left(\frac{V_AIRF \cdot V_IR}{c_{23} \cdot V_CCF + c_{24}} \right)^{c_{25}} \cdot \left(\frac{V_PLEF}{c_{26} \cdot V_CCF + c_{27}} \right)^{c_{28}}}{1 + \left(\frac{V_AIRF \cdot V_IR}{c_{23} \cdot V_CCF + c_{24}} \right)^{c_{25}} \cdot \left(\frac{V_PLEF}{c_{26} \cdot V_CCF + c_{27}} \right)^{c_{28}}} \quad \backslash * \text{MERGEFORMAT (1-4-15)}$$

$$\text{其中, } V_AIRF = \frac{\sum_{i=1}^{V_NDF} V_IRpF_i}{V_NDF}, \quad V_IR = \frac{V_NDF}{V_TNF}$$

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 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时长影响。具体公式如下：

$$Qv_BLSession_n = \alpha \cdot Qv_BLSession_{n-1} + (1 - \alpha) \cdot Qv_BLInstant \quad \backslash * \text{MERGEFORMAT (1-4-16)}$$

$$Qv_BLSession = \alpha_5 \cdot \exp(\beta_5 \cdot BlockingRatio) + \alpha_6 \cdot \exp(\beta_6 \cdot BlockingRatio) \quad \backslash * \text{MERGEFORMAT (1-4-17)}$$

$$BlockingRatio = BlovkingTimeRatio \cdot BlockingAreaRatio \quad \backslash * \text{MERGEFORMAT (1-4-18)}$$

$Qv_STInstant$ 、 $Qv_STSession$

$$Qv_STInstant = 5 - QStalling \quad \backslash * \text{MERGEFORMAT (1-4-19)}$$

$$QStalling = f(Duration, Interval, Frequency) = \frac{\left(\frac{Duration}{c_{29}} \right)^{c_{30}} \cdot \left(\frac{Interval}{c_{31}} \right)^{c_{32}} \cdot \left(\frac{Frequency}{c_{33}} \right)^{c_{34}}}{1 + \left(\frac{Duration}{c_{29}} \right)^{c_{30}} \cdot \left(\frac{Interval}{c_{31}} \right)^{c_{32}} \cdot \left(\frac{Frequency}{c_{33}} \right)^{c_{34}}} \quad \backslash * \text{MERGEFORMAT (1-4-20)}$$

其中，

$$Duration = \frac{\sum_{i=1}^{NRE} bufferLength_i}{Frequency} \quad \backslash * \text{MERGEFORMAT (1-4-21)}$$

$$\backslash * \text{MERGEFORMAT (1-4-22)}$$

$$Interval = \frac{\sum_{i=1}^{Frequency-1} [bufferStartTime_i - (bufferStartTime_{i-1} + bufferLength_{i-1})]}{Frequency - 1}$$

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 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时长影响。具体公式如下：

$$Qv_STSession_n = \alpha \cdot Qv_STSession_{n-1} + (1 - \alpha) \cdot Qv_STInstant \quad \backslash * MERGEFORMAT (1-4-23)$$

$$Qv_STSession = \alpha_7 \cdot \exp(\beta_7 \cdot StallingRatio) + \alpha_8 \cdot \exp(\beta_8 \cdot StallingRatio)$$

* MERGEFORMAT (1-4-24)

$$StallingRatio = StallingDuration / TotalPlayTime$$

* MERGEFORMAT (1-4-25)