



泛视频质量评价模型分析及应用

丁华¹, 王昕², 魏峥¹, 罗传飞¹, 宋皓¹, 施唯佳¹

(1. 天翼数字生活科技有限公司, 上海 200085; 2. 中国电信股份有限公司, 北京 100045)

摘要: 泛视频是指以视频为信息展现和交互形式的业务, 主要分为传统视频业务、新兴视频业务和衍生视频业务三类, 然而如何准确地进行视频播放质量评估却一直是一个难题。提出了一种泛视频播放质量评价模型及参数, 主要包括视频切换时延、视频起播时延以及视频播放卡顿次数和时长等, 其已在现网的 IPTV 天翼高清质量监测系统中得到了应用, 实践表明应用效果良好。所提研究成果可进一步应用于视联网产品领域等所有泛视频应用场景中。

关键词: 泛视频; 视频质量评估; 视频切换时延; 视频起播时延; 视频播放卡顿次数和时长

中图分类号: TP399

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2023171

Analysis and application of pan-video playback quality evaluation model

DING Hua¹, WANG Xin², WEI Zheng¹, LUO Chuanfei¹, SONG Hao¹, SHI Weijia¹

1. E-surfing Digital Life Technology Co., Ltd., Shanghai 200085, China

2. China Telecom Co., Ltd., Beijing 100045, China

Abstract: Pan-video refers to the business that takes video as the form of information presentation and interaction. It is mainly divided into three categories: traditional video business, emerging video business and derivative video business. However, how to accurately evaluate the quality of video playback has always been a problem. A pan-video playback quality evaluation model and parameters were proposed, which mainly included: video switching delay, video start delay, video playback pause times and duration, etc. It has been applied in the IPTV quality monitoring system of the current network, and the practice shows that the application effect is good. The proposed research results can be further applied to all pan-video application scenarios in the field of video network products.

Key words: pan-video, video playback quality evaluation, video switching delay, video start delay, video playback pause times and duration

0 引言

随着互联网和多媒体技术的蓬勃发展, 人类

社会进入泛视频媒体时代, 可以说“无处不视频”。长视频多元突破、短视频快速增长、网络直播带动大小屏深度融合, 在线观看视频已经成为人们

日常生活中主要的娱乐方式之一。用户对于观看体验的追求也不断提高,视频播放质量的优劣,是影响用户观看体验的关键因素,因此如何有效地评估视频的质量具有重要意义。人眼对视频的主观感知,往往在清晰度、流畅度、延时度 3 个方面比较敏感,另外视频色彩饱和度、画面亮度对比度、视频背景画面干净度和音视频唇音同步度也是影响主观感知的因素。泛视频播放质量评价一直是困扰人们的难题,人们一直致力于寻找一种能够全面、准确地评估视频质量的模型和方法。由此,在分析研究相关视频质量评估的方法和问题之后,本文提出了一种泛视频播放质量评价模型及参数的方法,主要涉及视频切换时延、视频起播时延、视频播放卡顿次数和时长 4 项指标参数。本文提出的模型及参数,已在电信运营商交互式网络电视(internet protocol television, IPTV)质量评估系统中得到应用,其视频质量评价效果良好,基本与人眼主观体验结果一致。

1 视频质量评估面临的问题

服务质量(quality of service, QoS)是广泛采用的服务度量标准之一,主要包括网络的吞吐率、时延、丢包率、抖动、误码率等,显然这些指标仅反映了服务技术层面的性能,甚至是网络传输层面的性能,忽略用户主观因素,因此它不能直接反映用户对服务的认可程度。用户体验质量(quality of experience, QoE)是一种以用户认可程度为标准的服务评价方法,它综合了服务层面、用户层面、环境层面的影响因素,直接反映了用户对服务的认可程度^[1]。涉及 QoE 的量化问题,目前较广泛采用的是国际电信联盟(ITU)建议的“平均意见得分(mean opinion score, MOS)”,它将 QoE 的主观感受分为 5 个层次^[2-3]。

按照是否有用户直接参与评价及是否给出 QoE 与其影响因素之间的关联模型,可以将现有

的评价方法分为主观评价方法、客观评价方法及主客观结合的方法。主观评价方法,是让用户直接对所使用的业务做出评价,因此它最准确,一般可以作为评价其他评价方法的标准。由于主观评价方法需要考虑太多的因素,实施步骤较复杂、代价大,没有得到广泛的应用。客观评价方法,通过计算输出序列相对于输入序列的失真程度,对业务做出评价,分为全参考、部分参考、无参考 3 种类型。客观评价具有简单、实用的特点,所以它是目前 QoE 评价方法研究中重要的一部分,但值得注意的是,为了更准确地反映用户对业务或者服务的主观感受,客观评价方法需要考虑人类的感知模型,这就给 QoE 的客观评价带来了一定的难度^[1,4]。

网络视频质量评估已经成为网络视频业务发展的迫切需求,人们一直期望通过对视频特征或网络传输分析客观评价视频质量,然而视频特征信息获取、计算复杂度、实时性要求,以及分组丢失、时延、抖动等网络因素都为有效评价网络视频质量增加了困难^[5-6]。另外,移动网络视频质量评价方法,也还不能提供足够可靠的评价结果。在单向传输的视频质量评价中,无论主观方法还是客观方法都需要进一步研究^[7]。以上分析表明,人们之前基于网络传输参数或视频数据信息,提出过很多经验公式或模型评估视频质量,但往往效果不理想,无法准确评价。基本上无法与人眼主观体验感知相一致,有时甚至会出现视频质量评估值较高,人眼主观感知却很差的情况,或者反之。目前的客观评价方法,虽然可以精确地分析码流传输质量或视频图像特征质量,但其往往存在无法与主观感受效果匹配一致的问题。主观评价方法虽然可以按照人眼的主观感受划分出非常好、好、一般、差、非常差 5 个等级,但是实际操作时会出现等级分无法有效评定的问题,例如,视频卡顿多少次可认为是列入“一般”等级,这个没有客观标准,往往会因人而异^[8-10]。本文提



出的一种泛视频播放质量评价模型及参数的方法,属于客观评价方法类型,不采用主观评价方法的五分评分评价,取而代之的是评价视频时延、视频卡顿等参数,实践表明,评估效果基本上与人眼主观体验相一致,很好地解决了当前诸多视频质量评价方法所面临的问题。

2 IPTV 视频质量评估系统设计

近年来,人们一直试图通过主观评价和客观评价等方法评估 IPTV 网络的数字视频质量^[11]。例如,文献[12]针对 CBR 编码、MPEG.2 TS 封装、H.264Mainprofile@level4.1/1 920×1 080@25 f/s 的视频进行了主观评价。文献[13]提出了在机顶盒内置质量监测工具,监测 IPTV 业务的丢包率、媒体流速率、业务互动等指标。进一步地,人们还进行了三网融合下视频业务质量评估体系,以及优化编码参数以期提升 IPTV 高清频道的视频质量等研究工作^[14-15]。目前,国内互联网视频产业界尚不存在评估用户体验质量的一致标准,不同视频服务运营商对于用户体验的认识在角度和深度上存在较大差异。同时,视频服务运营商普遍缺少对用户行为规律的深刻研究和认识,往往投入了大量资金和成本进行系统和平台的管理和优化之后,用户体验质量却没有得到明显的改善,而造成资源的极大浪费^[16]。

本文针对某电信运营商的 IPTV 视频质量评估系统进行了应用设计,IPTV 视频质量评估系统示意图如图 1 所示。分析表明,该运营商的南方省机顶盒应用的是定制 IPTV 视频播放器,可升级机顶盒底层硬件视频编/解码芯片的库文件,并对播放器进行适配开发,增加接口向其他类似探针的应用提供相关采集数据。另外,该电信运营商的北方省机顶盒应用的是基于开放互联网视频服务(over the top, OTT)的第三方视频播放器,无法开发和适配相关接口向探针提供采集数据。根据用户体验关注的重点,结合实际应用效果,本

文将 IPTV 视频质量指标重点聚焦在视频流畅度和视频延时度两个方面。本文提出的新型视频质量评价模型设计方案,旨在通过视频播放器接口或网络数据抓包两种模式,分别采集分析视频切换时延、视频起播时延、视频播放卡顿次数和时长 4 项视频质量指标参数,以及计算用户观看视频业务体验的综合评分 U_MOS 值和网络健康指数 h_mos 值。机顶盒探针将采集计算的视频质量数据,及时上报到中心平台,平台对分省的 IPTV 视频优良率进行统计分析和考核处理。本文提出的视频播放质量评价模型及参数方法,很好地解决了电信运营商的全网智能机顶盒视频质量评估问题。

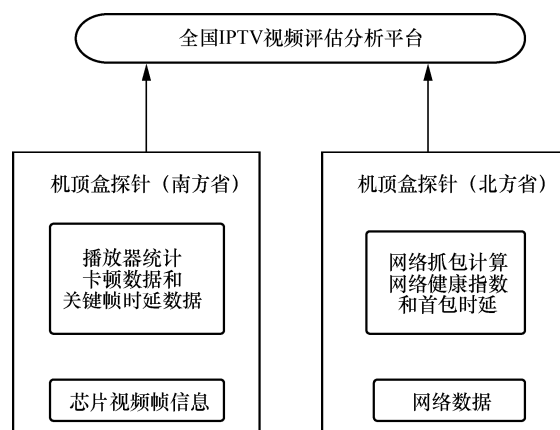


图1 IPTV 视频质量评估系统示意图

本文提出的一种泛视频播放质量评价模型及参数的方法,相较于传统的视频质量衡量标准,有明显的改进。一方面,其所采用的方法和参数简便实用、全面可靠;另一方面,虽然属于客观评价方法类型,但是评估效果与人眼体验结果高度一致。一般地,视频质量主要体现在清晰度、流畅度、延时度 3 个方面,然而目前互联网平台服务的视频清晰度(如分辨率等)往往都比较高,因此视频质量评估可以仅针对流畅度、延时度进行分析。研究表明,视频延时度主要体现在视频直播切换和视频点播起播阶段,视频流畅度主要反映视频播放阶段的卡顿情况。在视频切换或起

播阶段, 当视频关键帧出现后人眼将能观察到视频画面, 本文通过播放器接口可以准确地采集视频关键帧的时延, 其与人眼主观体验的视频时延高度一致。同时, 视频编/解码芯片上报的视频帧信息, 经由播放器解码后将直接送至显示器视频显示, 若 500 ms 内显示缓存没有接收到视频数据, 本文中探针算法模块将及时进行视频播放卡顿次数的统计, 其也与人眼体验的视频卡顿感知基本一致。

3 播放器接口采集视频质量指标的实现方案

3.1 IPTV 播放器采集视频质量参数的接口要求

南方省 IPTV 机顶盒底层硬件视频编/解码芯片的视频帧信息数据, 上报到播放器。播放器通

过智能机顶盒进程间数据通信 Android 接口定义语言 (Android interface definition language, AIDL) 方式, 将接收的帧信息数据和生成的视频质量指标数据, 及时传递给机顶盒质量探针, 两者之间的数据通信采用 JSON 格式进行传递。播放器采集视频质量指标参数见表 1。

3.2 播放器计算视频时延和视频卡顿方案描述

(1) 视频播放时延

IPTV 播放器向平台请求视频开始计时 (第一时刻), 平台下推视频到机顶盒编/解码芯片, 芯片解析后按接口定义要求每秒上报视频帧信息, IPTV 播放器收到芯片上报的视频关键帧为止 (第二时刻), 两个时刻的时间差值即视频播放时延。一般地, 在视频切换或起播阶段, 只有当第

表 1 播放器采集视频质量指标参数

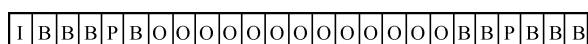
序号	分类	名称	描述	备注
1	video switching delay	switchChannel_time	视频切换时延 (播放器) (单位: ms)	播放器接收视频切换指令与第一个视频关键帧之间的时延
2	video start delay	Latency	视频起播时延 (播放器) (单位: ms)	播放器接收视频起播指令与第一个视频关键帧之间的时延
3.1	frame_info	frame_type	帧类型 (I/P/B 帧)	
3.2	frame_info	frame_rate	视频帧率 (单位: f/s)	
3.3	frame_info	frame_size	帧大小 (单位: KB)	
3.4	frame_info	frameqpmin	每秒的帧级 QP 的最小值	量化参数: QP
3.5	frame_info	frameqpmax	每秒的帧级 QP 的最大值	
3.6	frame_info	frameqpavg	每秒的帧级 QP 的平均值	
3.7	frame_info	framemvmin	每秒的帧级 MV 的最小值	运动矢量: MV
3.8	frame_info	framemvmax	每秒的帧级 MV 的最大值	
3.9	frame_info	framemvavg	每秒的帧级 MV 的平均值	
3.10	frame_info	frame_skip	每秒的帧级 skip 和非 skip 的比例, 1 帧为 0	
3.11	frame_info	decoder_buffer	解码器缓存状态: 0—OK 1—Overflow; 上溢 2—UnderFlow; 下溢 3—解码出现错误帧	芯片每秒上报一次帧信息。下溢表示: 该帧数据为空 (3.3~3.11 的帧数据中, 前 8 位均为 0, 第 9 位为 2)
4.1	video playback pause	FreezeCount	卡顿次数 (播放器)	统计每秒视频帧信息中 decoder_buffer=2 或 3 的个数, 超过 500 ms (1 s 内的半数帧以上) 即为卡顿一次
4.2	video playback pause	lagtv	卡顿时长 (播放器) (单位: ms)	通过发生卡顿的帧数进行时长换算
5	U_MOS	MOS	视频 MOS 值 (5 分制)	通过芯片帧信息数据计算, 每 5 min 计算一次



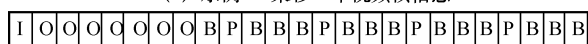
一个视频关键帧出现后,人眼才能观察到视频画面。由于视频播放器可以甄别视频关键帧,因此本方案可准确地获取 IPTV 直播(多播)频道切换时延、点播(单播)影视起播时延,其与人眼主观观察体验基本一致。

(2) 视频卡顿

人眼的视觉暂留现象,使静止图像连续出现时会有动画感;然而,当视频帧率低于 15 f/s 时,连续的运动视频往往就会有停顿的感觉。在视频编码序列中,当一个图像组(group of picture, GOP)中发现缓存下溢即出现空帧,且空帧数量超过一半时,该 GOP 就判断为卡顿组。视频帧信息结构示意图如图 2 所示,图 2(a) 示例 1 判断为卡顿 1 次/秒,因为 25 帧中有 13 帧为空帧;图 2(b) 示例 2 不判断为卡顿。



(a) 示例1: 某秒25个视频帧信息



(b) 示例2: 某秒25个视频帧信息

图2 视频帧信息结构示意图

3.3 基于芯片视频帧信息计算 U_MOS

宽带视频服务用户体验评估 U_MOS,按照其处理信号层级主要包含 Model0、Model1 和 Model2 的 3 层模型。其中 Model0 显式包含比特率(BitRate)、屏幕大小(ScreenSize)、像素密度单位(PPI)变量,其余变量在不同的应用场景(分辨率、编码器类型)下为定值。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。Model2 需要从原始或播放器连续采集视频帧的关键质量信息。本文探针的实验室实测数据主要参考行业标准及相关文献中 Model1 的模型算法^[17-19],取值为视频观看质量 QV,如式(1)所示。

$$QV = \min(\max(Qcod, 1), 5) \quad (1)$$

其中, Qcod 为视频序列的质量分析结果,取值范围为[1,5]。

$$Qcod = kfr_{\text{impact}} \cdot \exp(n_1 \cdot (QP_{\text{fr}_{\text{impact}}} + cpx_{\text{video}} + motion_{\text{impact}})) \quad (2)$$

其中, kfr 表示视频帧率与两个相邻视频关键帧(I 帧)之间的帧数的比值,如式(3)所示。

$$kfr_{\text{impact}} = n_2 \cdot kfr + n_3 \quad (3)$$

$$QP_{\text{fr}_{\text{impact}}} = n_4 + n_5 \cdot \left(\frac{\text{avg}QP_{\text{pic}}}{51} \right)^{n_6} + n_7 \cdot \frac{1}{fr} + n_8 \cdot \text{intraflicker} + n_9 \cdot (\max QP_{\text{pic}} - \min QP_{\text{pic}}) \quad (4)$$

其中, intraflicker 为一个布尔型变量,表示到目前接收到的帧为止,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。

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。

- 当前 I 帧 QP 值比前一帧 QP 值高 5。
- 当前 I 帧 QP 值比后一帧 QP 值高 5。

$$cpx_{\text{video}} = \min \left(\sqrt{\frac{br_v}{\text{AvgByte}_{1\text{-frame}}}} + n_{10} \cdot \text{skipRatio}, 1.0 \right) \quad (5)$$

$$motion_{\text{impact}} = n_{11} \cdot \text{avg}MV_{\text{pic}} \cdot \left(1 - \frac{fr}{30} \right) \quad (6)$$

其中, SkipRatio 为 Skip 模式和非 Skip 模式比例, br 为码率, fr 为帧率, AvgByte 为平均字节大小。视频观看质量 QV 计算式中涉及的各项系数见表 2。

4 网络数据抓包计算视频质量指标的实现方案

4.1 网络抓包计算视频质量的参数要求

北方省 OTT 机顶盒探针可以通过网络数据抓包方式,采集和分析相关视频质量指标参数。网络抓包计算视频质量指标参数见表 3。

表2 视频观看质量 QV 计算式中涉及的各项系数

n_1	n_2	n_3	n_4	n_5	n_6
1.764 17	$-1.875\ 08 \times 10^{-5}$	$9.251\ 12 \times 10^{-5}$	6.289 61	-2 407.725 558	20.236 81
n_7	n_8	n_9	n_{10}	n_{11}	
-11.374 05	0.069 90	0.003 28	-0.037 94	0.001 68	

表3 网络抓包计算视频质量指标参数

序号	分类	名称	描述	备注
1	Video switching delay	Switch Channel_time1	视频切换时延 (探针首包) (单位: ms)	探针网络抓包, 从类似 igmp leave 的视频切换消息开始, 到收到第一个网络数据包的时间
2	video start delay	Latency1	视频起播时延 (探针首包) (单位: ms)	探针网络抓包, 从类似 DESCRIBE rtsp:// 的视频起播消息开始, 到收到第一个网络数据包的时间
3	h_mos	StreamHealth	网络健康指数 (5 分制)	视频播放期间, 主要基于 MDI 的 MLR 和 DF 统计, 每 5 min 计算一次

4.2 网络抓包计算视频时延方案描述

直播切换时延的计算要求是从网络抓包解析到多播互联网组管理协议 (internet group management protocol, IGMP) 的 igmp join 消息开始, 到收到第一个数据包的时间为止。点播起播时延的计算要求是从网络抓包解析到单播实时流传输协议 (real time streaming protocol, RTSP) 的 DESCRIBE 消息开始, 到收到第一个数据包的时间为止; 如果存在 302 跳转, 以第一个请求时间为起始时间。

由于网络抓包只能定位数据首包、无法精确定位视频关键帧, 因此探针网络抓包计算的直播切换时延和点播起播时延, 往往计算数值比较小。一般地, 视频切换时延、视频起播时延的发生过程和时刻大致如下: 用户遥控器按键操作 (T_1)、播放器发起视频请求 (T_2)、机顶盒发送网络包请求协议 (T_3)、平台返回网络视频流首包 (T_4)、播放器获取视频关键帧 (T_5)、视频输出到显示器 (T_6)。其中, $T_1 \sim T_2$ 、 $T_2 \sim T_3$ 、 $T_5 \sim T_6$ 的 3 个区间时长基本上是相对不变的; $T_3 \sim T_4$ 、 $T_4 \sim T_5$ 两个区间的时长是不断变化的, 主要由网络传输和系统响应能力决定。前文所述的播放器计算到关键帧的时延区间为 $T_2 \sim T_5$, 此处所述的网络抓包计算到首包的时延区间为 $T_3 \sim T_4$ 。

4.3 优化拟合网络健康指数公式计算 h_mos

互联网的技术规范 RFC 4445 中的媒体传输质量指标 (media delivery index, MDI) 是 IP 视频流传输质量的技术标准, 是行业内公认的网络指标参数方法, 被用来衡量网络媒体流抖动和丢包率, 其包含两个参数: 延迟因子 (DF) 和媒体丢包速率 (MLR) [20]。

利用 MDI, 本文提出了一种网络健康指数的计算 h_mos 值的模型算法, 如式 (7) 所示。

$$h_mos = (5 - (\text{int}(\text{delay} / a)) \times b - ((\text{float})\text{lost} / c)) \times (\text{pow}(d, e \times \text{lost} \times f)) \quad (7)$$

其中, delay 为 MDI 中的 DF, lost 为 MDI 中的 MLR, a 、 b 、 c 、 d 、 e 、 f 是系数。

上述经验表达式中的网络健康指数模型系数需要进行优化, 目标是使 h_mos 数值曲线与卡顿曲线变化趋势相似, 网络健康指数曲线优化逼近卡顿曲线变化趋势示意图如图 3 所示。在实验室条件下, 特定机顶盒既可以带有播放器接口和芯片帧信息卡顿模块, 同时又可以植入网络健康指数模块。在测试过程中, 将网络健康指数 h_mos 曲线先在 $x=2.5$ 的直线上水平翻转, 并与卡顿曲线标幺量纲统一处理, 利用单纯形优化、最小二乘算法使两曲



线之间的差值平方和达到最小。通过不断地对 h_mos 表达式优化, 最终确定出一组优化系数。

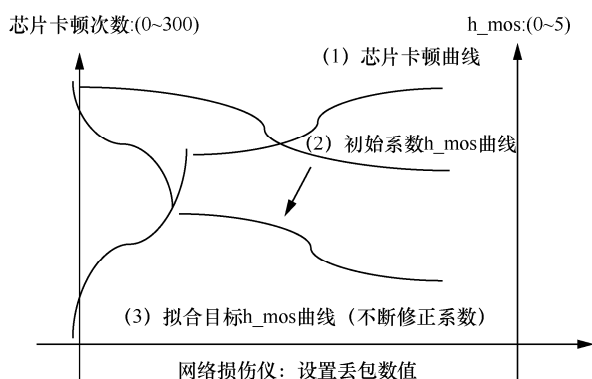


图3 网络健康指数曲线优化逼近卡顿曲线变化趋势示意图

网络健康指数曲线在与卡顿曲线拟合优化时, 可利用网络损伤仪进行控制, 以期在实验室连续控制不同的网络丢包的测试环境, 并记录横轴数据。在网络损伤仪的控制下, 卡顿数据曲线变化趋势将基本保持不变, 而 h_mos 数值曲线随模型系数改变而发生变化。机顶盒探针内置的网络健康指数算法运算模块, 将按规范 RFC 4445 中的 MDI 要求, 实时采集和计算网络的 DF 值和 MLR 值, 代入式 (7) 计算 h_mos 值。当单纯形优化算法优化搜索到某组模型系数时, 重新进行实验室网络损伤控制, 将得到新的 h_mos 曲线, 再次与卡顿目标曲线拟合优化, 直至优化得到最后一组系数 a 、 b 、 c 、 d 、 e 、 f 为止, 之后停止网络损伤控制测试操作。实验室某款机顶盒测试拟合后, 得到了一组优化系数: $a=50$ 、 $b=0.1$ 、 $c=50$ 、 $d=2.72$ 、 $e=-0.000\ 394$ 、 $f=50$ 。

特别地, 如果实验室没有上述特定机顶盒, 视频播放卡顿曲线也可以通过人眼观察方式记录形成, 统计记录周期和卡顿计次频率与上述一致, 分别为 5 min 和 1 s。

5 实验室测试的视频质量指标数据分析

5.1 终端侧采集的视频指标数据分析

实验室特定机顶盒同时具备播放器计算模

块功能 (包括直播切换开始到关键帧、点播起播开始到关键帧、帧信息卡顿数据和 U_MOS) 和网络抓包计算模块功能 (包括直播切台开始到数据首包、点播起播开始到数据首包、网络健康指数 h_mos)。实验室测试中, 特定机顶盒探针每 5 min 上报的数据与人眼实际主观感知对比见表 4。

(1) 播放器采集数据分析

播放器计算的直播切台时延和点播起播时延基本上与人眼感知时延接近, 实验效果良好。例如, 第二次测试的点播起播时延, 播放器采集的 Latency1 为 503 ms 时, 人眼实测感知时延为 500 ms 左右。第一次测试的直播切台时延, 播放器采集的 switchChannel_time 为 1 087 ms, 人眼实测感知时延为 1 000 ms 左右。

播放器计算的视频卡顿次数和卡顿时长, 基本上与人眼感知一致 (个别数据除外), 实验效果比较理想。例如, 10 次的视频播放测试中, 播放器采集的卡顿次数和卡顿时长基本与人眼实测感知一致; 但是第 5 次测试时, 播放器采集卡顿次数为 1 次、人眼实测感知为 13 次。

U_MOS 计算值在视频卡顿不是很严重时, 数值有波动; 但是在视频卡顿比较严重时, 数值变化不明显, 不能有效反映视频质量下降的情况。例如, 第 5 次测试时, 视频播放效果较差, 卡顿时长达到 13 000 ms, 但是 MOS (U_MOS) 值仍然高达 3.95 分。

(2) 网络抓包采集数据分析

探针网络抓包计算的直播切台时延和点播起播时延, 多数情况下与人眼感知时延的误差约为 500 ms。例如, 第 2 次测试的点播起播时延, 网络抓包采集的 Latency1 为 29 ms 时, 人眼实测感知时延为 500 ms 左右; 第 3 次测试的直播切台时延, 网络抓包采集的 switchChannel_time1 为 13 ms, 人眼实测感知时延小于 500 ms。

h_mos 计算值在视频卡顿不是很严重时, 数

表4 特定机顶盒探针每5 min上报的数据与人眼实际主观感知对比

参数/片源测试	1	2	3	4	5	6	7	8	9	10
switchChannel_time	1 087		251		375		1 467		3 419	
switchChannel_time1	52		13		21		7		23	
人眼感知视频切换时延(直播切台)	1 000		<500		<500		<1 500		3 000	
Latency		503		482		505		498		186
Latency1		29		19		17		26		51
人眼感知视频起播时延(点播起播)		500		500		500		500		<500
FreezeCount	1	1	1	1	1	2	2	2	1	1
lagtv	1 200	960	1 400	840	13 680	1 680	1 680	1 600	1 080	1 360
人眼感知卡顿次数	1	1	1	1	13	2	2	2	1	1
人眼感知卡顿时长	1 000	<1 000	1 000	<1 000	13 000	1 000	1 000	1 000	<1 000	1 000
MOS (U_MOS)	4.28	4.34	4.23	4.04	3.95	3.91	3.92	3.92	4.21	4.16
StreamHealth (h_mos)	5.00	5.00	5.00	5.00	2.00	5.00	5.00	5.00	5.00	5.00

值变化波动不明显;然而在视频卡顿比较严重时,数值明显下降,较好地反映了视频质量下降的情况。例如,第5次测试时,视频播放效果较差,卡顿时长达到13 000 ms,StreamHealth (h_mos)值相应降低到2.00分。

5.2 平台侧IPTV播放优良率统计分析

平台优良率统计主要分为直播切台优良率、点播起播优良率以及视频播放卡顿优良率。单个机顶盒每5 min上报一次话单,如出现一次时延大于1 s或卡顿超过两次,则该话单为一次质差块。单个机顶盒中的质差块,如超过全天话单的5%,则该盒子确定为当天的质差用户。全省全天的优良率统计为 $(1 - \text{当天的质差用户数} / \text{当天全天的话单用户总数}) \times 100\%$ 。

平台在统计南方省的优良率时,直接采用终端探针上报的视频时延数据和视频卡顿数据计算;另外,由于U_MOS值不完全准确,因此在实际应用中并没有采用该结果。平台在统计北方省的时延优良率时,将在探针上报的视频首包时延数据的基础上加500 ms后再进行计算处理;平

台在统计视频播放卡顿优良率时,当h_mos值小于4.8分时记为质差话单,实验表明视频卡顿两次的临界点时往往h_mos值接近4.8分。

6 结束语

目前电信运营商正在积极践行网络强国、数字中国和“云改数转”战略,依托云网边端能力,构建数字生活业务新格局。同时,利用视频监控产品基础能力,全力打造视联网,业务覆盖数字家庭、智慧城市、智慧社区、数字乡村等应用领域。本文的研究成果已经在现网智能机顶盒内得到了规模应用,探针部署量超过了8 000万个。本文所提视频质量评价模型和方法的目标是力求评价结果与人眼主观体验感知相一致,可进一步应用于视联网产品领域等所有泛视频应用场景中。

参考文献:

- [1] 林闯,胡杰,孔祥震. 用户体验质量(QoE)的模型与评价方法综述[J]. 计算机学报, 2012, 35(1): 1-15.



- LIN C, HU J, KONG X Z. Survey on models and evaluation of quality of experience[J]. Chinese Journal of Computers, 2012, 35(1): 1-15.
- [2] ITU-T. Methods for subjective determination of transmission quality: P.800[S]. 1996.
- [3] ITU-R. Methodology for the subjective assessment of the quality of television pictures: BT.500-13[S]. 2012.
- [4] 李永强, 沈庆国, 朱江, 等. 数字视频质量评价方法综述[J]. 电视技术, 2006(6): 5.
- LI Y Q, SHEN Q G, ZHU J, et al. Overview of video quality evaluation methods[J]. Video Engineering, 2006(6): 5.
- [5] 杨付正, 万帅. 网络视频质量评估技术研究现状及发展动向[J]. 通信学报, 2012, 33(4): 107-114.
- YANG F Z, WAN S. Overview of state-of-the-art and future of networked video quality assessment[J]. Journal on Communications, 2012, 33(4): 107-114.
- [6] 刘河潮, 常义林, 元辉, 等. 一种网络丢包的无参考视频质量评估方法[J]. 西安电子科技大学学报(自然科学版), 2012, 39(2): 29-34.
- LIU H C, CHANG Y L, YUAN H, et al. No-reference video quality assessment over the IP network based on packet loss[J]. Journal of Xidian University (Natural Science), 2012, 39(2): 29-34.
- [7] 魏耀都, 谢湘, 匡镜明, 等. 移动视频质量评价方法及发展趋势[J]. 电信科学, 2010, 26(3): 57-63.
- WEI Y D, XIE X, KUANG J M, et al. Mobile video quality evaluation methods and its development[J]. Telecommunications Science, 2010, 26(3): 57-63.
- [8] 方玉明, 眭相杰, 鄢杰斌, 等. 无参考图像质量评价研究进展[J]. 中国图象图形学报, 2021, 26(2): 265-286.
- FANG Y M, SUI X J, YAN J B, et al. Progress in no-reference image quality assessment[J]. Journal of Image and Graphics, 2021, 26(2): 265-286.
- [9] 汪芮. 超高清视频图像质量主观评价方法探究[J]. 广播与电视技术, 2021, 48(6): 25-29.
- WANG R. Research on subjective image quality assessment method for UHD video[J]. Radio & TV Broadcast Engineering, 2021, 48(6): 25-29.
- [10] 王惠明, 欧臻彦, 王倩男, 等. 8K 超高清视频质量客观评测方法研究[J]. 现代电视技术, 2022(10): 116-121.
- WANG H M, OU Z Y, WANG Q N, et al. Research on objective evaluation methods for 8K teltra high definition video quality[J]. Advanced Television Engineering, 2022(10): 116-121.
- [11] 栗蔚, 魏凯. IPTV 视频质量评价技术综述[J]. 电信网技术, 2009(5): 24-29.
- LI W, WEI K. Overview of IPTV video quality measurement technologies[J]. Telecommunications Network Technology, 2009(5): 24-29.
- [12] 罗斯青, 肖晴, 胡源. 面向 IPTV 的 H.264 高清编码视频质量评估[J]. 电信科学, 2008, 24(5): 83-86.
- LUO S Q, XIAO Q, HU Y. A subjective quality assessment test for HD video based on H. 264 technology in IPTV[J]. Telecommunications Science, 2008, 24(5): 83-86.
- [13] 姚良. 影响 IPTV 业务质量的关键因素及质量提升措施探讨[J]. 电信科学, 2009, 25(3): 32-35.
- YAO L. Research on key factors effecting IPTV service quality and service improvement measures[J]. Telecommunications Science, 2009, 25(3): 32-35.
- [14] 姚良, 奚溪. 三网融合下视频业务质量评估体系的研究[J]. 电信科学, 2011, 27(3): 27-31.
- YAO L, XI X. Research of video service quality evaluation system in triple-play field[J]. Telecommunications Science, 2011, 27(3): 27-31.
- [15] 罗传飞, 张玮. IPTV 高清频道视频质量提升研究[J]. 电信科学, 2012, 28(12): 135-140.
- LUO C F, ZHANG W. Research of improving IPTV HD channels video quality[J]. Telecommunications Science, 2012, 28(12): 135-140.
- [16] 张玉冉. 互联网视频用户体验质量评估模型的研究与实现[D]. 北京: 北京邮电大学, 2015.
- ZHANG Y R. Research and implementation of evaluation model of Internet video user experience quality[D]. Beijing: Beijing University of Posts and Telecommunications, 2015.
- [17] 工业和信息化部. 宽带视频服务用户体验评估 直播算法和参数: YD/T 3778-2020[S]. 2020.
- Ministry of Industry and Information Technology. The live algorithm and parameter for the user experience assessment of broadband video service: YD/T 3778-2020[S]. 2020.
- [18] 工业和信息化部. 宽带视频服务用户体验评估 点播算法和参数: YD/T 3779-2020[S]. 2020.
- Ministry of Industry and Information Technology. The Vod Algorithm and parameter for the user experience assessment of broadband video service: YD/T 3779-2020[S]. 2020.
- [19] 贺甜甜, 刘彦凯, 宋利. 视频用户体验理论与实践[J]. 电信科学, 2017, 33(8): 67-76.
- HE T T, LIU Y K, SONG L. Theory and practice of quality of

experience in video services[J]. Telecommunications Science, 2017, 33(8): 67-76.

[20] RFC. A proposed media delivery index (MDI): 4445[S]. 2006.

[作者简介]



丁华（1969—），男，天翼数字生活科技有限公司高级工程师，主要研究方向为 IPTV 及探针、视频监控和视频通话等。



王昕（1977—），男，中国电信股份有限公司工程师，主要研究方向为宽带和基于高带宽接入的 IPTV 等。



魏峥（1983—），男，天翼数字生活科技有限公司工程师，主要研究方向为 IPTV 及探针、视频视觉和机器人智能等。

罗传飞（1978—），男，博士，天翼数字生活科技有限公司教授级高级工程师，主要研究方向为视音频信号处理、IPTV 视频质量指标和视联网等。

宋皓（1988—），女，天翼数字生活科技有限公司工程师，主要研究方向为 IPTV 及探针、计算机数据处理系统和云网 CDN 等。

施唯佳（1978—），男，天翼数字生活科技有限公司高级工程师，主要研究方向为 IPTV 及探针、融合视讯和视联网等。