



# 基于智能终端的视频通信业务服务质量评测模型研究

韩建亭,张 夙

(中国电信股份有限公司上海研究院 上海 200122)

**摘 要:**运营商正在发展基于智能终端的视频通信业务,分析了视频通信业务的服务质量特点以及影响视频通信服务质量的关键因素,在此基础上,对视频主客观质量评估方法进行了分析,建立了视频通信业务服务质量的评测模型并进行了实验室优化验证,为运营商有针对性地提升视频通信业务的服务质量,提供了终端产品测试依据和评估方法。

**关键词:**视频通信;视频质量评估;智能终端;评测模型

**doi:** 10.3969/j.issn.1000-0801.2013.04.006

## Study on the Assessment Model of Video Communication Quality Based on Smart Terminal

Han Jianting, Zhang Su

(Shanghai Research Institute of China Telecom Co., Ltd., Shanghai 200122, China)

**Abstract:** Nowadays operators are focusing on developing the video communication services based on smart terminals. The characteristics and key affecting factors of video communication quality were discussed. Furthermore, according to the analysis of the methods of subjective and objective video quality assessment, an assessment model of video communication service quality was proposed and verified by experimental data, which can provide test reference and evaluation methods for operators to improve the quality of video communication service.

**Key words:** video communication, video quality assessment, smart terminal, assessment model

### 1 引言

随着光网城市的建设和 3G 无线网络的发展,通信网络更高的数据传输速率和数据业务支撑能力,为通信业务从传统的话音转向多媒体应用,提供了基础和条件。近几年智能终端产业链的迅速发展壮大,使得用户通过手机、平板电脑、智能 IPTV 机顶盒等,使用各类多媒体应用成为可能。典型的多媒体应用(如流媒体点播、视频监控、可视电话等)取得了迅猛发展,跨网络、多终端的融合通信业务正逐渐成为移动互联网的核心和热点业务。

随着新应用的层出不穷,用户对业务端到端的服务质

量也越来越关注。运营商为了提升竞争力,需要提供更加优质且有质量保障的端到端视频通信业务。影响视频通信业务质量的关键因素有哪些,对视频通信业务如何进行评价,如何测试,是目前运营商需要重点关注的课题。

而目前,业界对视频通信服务质量评价方法的研究较少,且无相关的标准指导,导致评测视频通信业务多以主观为主,缺乏科学性和规范性。本文结合视频质量的主观和客观评估方法,研究影响视频通信服务质量的关键指标,对几种基于智能终端的视频通信产品的视频服务质量进行了评测和技术验证,并对测试条件、测试方法、评分准则、数据处理等进行了规定,建立了基于智能终端的视频服务质量评



测模型,为规范视频通信业务质量的评测奠定基础。

## 2 视频通信与流媒体视频服务的差异

IPTV、互联网视频等流媒体视频服务与面向用户的端到端视频通信服务,对音视频的质量要求不同。视频通信(包括视频会议、视频电话等)业务必须支持实时的编码和解码,对时延和时延抖动的要求更为苛刻。从体验的角度分析,差异主要体现在以下几个方面。

### (1)双向性

参与通话的双方或多方都要进行媒体流的发送和接收,需要进行实时的编码和解码;流媒体视频服务基本上都是单向的,用户侧只需进行音视频的接收和解码,不需要编码。

### (2)业务具备强交互性

视频通信的双方在话音和视频方面都要求具备实时的交互性和互动性,流媒体的交互主要体现在点播和播放控制方面。

### (3)对视频质量的连续性要求不强

视频通信的视频主要是人体肢体语言动作,相比流媒体支持播放的球赛、飞行、车辆等视频,在连续性方面要求较低。

### (4)对音视频质量要求相对较低

视频通信主要用于通话双方的实时音视频交互,偶尔出现停顿或断续是可以容忍的,对画质和音质的要求相对流媒体播放的音视频画质,可以降低。

尽管随着IPTV等视频播放业务的兴起,国内外对流媒体视频服务质量的研究已经比较广泛,但视频通信业务对服务质量的差异性,决定了评价视频通信服务业务的服务质量模式,不能够采用流媒体服务质量的评测方法。

## 3 影响视频通信服务质量的关键因素

视频通信的服务质量包括音频质量和视频质量两方面。对于音频服务质量,由于目前VoIP技术的发展较为成熟,已有非常成熟的评估模型,并且相较于视频码流,音频码流所需网络带宽较小,目前各类视频通信终端在网络条件极其恶劣的条件下,都会采取优先保障音频服务质量的策略,因此本文的研究范围仅限于视频服务质量。

视频通信的服务质量,主要由视频通信的终端能力和网络服务质量两个方面所决定。

### 3.1 终端能力

#### (1)终端处理能力

实现实时视频通信特别是高清格式,需要对原始视频

信号进行较大的压缩,而算法压缩比越高,复杂度就会相应提高,对终端处理芯片的计算能力要求也就越高。因此终端视频处理能力的提升,是实现高效压缩算法的硬件保障。对于PC以及平板电脑而言,处理能力很强,终端性能不会成为影响视频质量的瓶颈;对于手机而言,主频在500~800 MHz的智能手机都可承载视频通信业务;对于机顶盒而言,需要具备支持实时视频编解码能力的处理芯片。

#### (2)摄像头性能

摄像头作为一种视频输入设备,其品质决定了原始视频的质量,也就直接影响了接收端的视频输出。像素值是判断摄像头性能优劣的重要指标。目前市场上常见的摄像头都是采用USB接口的数字摄像头,像素值最低为30万左右。像素越高,图像质量越清晰,意味着同一幅图像所包含的数据量越大,在编解码器性能相同的前提下,对网络带宽的要求就越高。此外,摄像头的视频捕获能力是另外一个重要指标,通常需要依靠硬件与软件的结合达到满足标准速率要求的捕获指标。

#### (3)编解码器性能

视频编解码器的性能是影响视频质量的关键因素。最优的编码器设计必须权衡压缩效率及可用的计算能力,即在计算能力有限的情况下,获得最佳压缩效率。目前,图像压缩编码方式主要有H.263、H.264、MPEG系列、AVS等标准,其中,H.264能够在低带宽下实现高清晰的动态图像效果,而且编码时延小,作为新一代视频编解码标准,其优势非常明显。

## 3.2 网络服务质量

视频通信的服务质量与网络的性能有直接的关系,影响较大的因素包括分组丢失、端到端时延、时延抖动和带宽4个方面。

#### (1)分组丢失

分组丢失是影响视频质量的关键因素。发送端和接收端之间数据分组数目的差值即网络传输丢失的分组数目。分组丢失直接对视频数据分组所在帧的解码产生影响,进而对相关帧的解码产生影响,形成传播损伤,因此分组丢失对视频服务质量的影响是非常大的,如图像会出现马赛克、拖尾、停顿等。

#### (2)端到端时延

端到端时延分为固定时延和可变时延,前者包括编解码器引入的时延和打包时延,后者包括承载网上的传输时延、服务处理时延、去抖动时延。根据不同的网络负载状况,端到端时延会发生变化。时延只能使视频数据分组整

体到达接收端的时间推迟,并不会影响数据分组之间的相对时间,因此时延对视频服务质量没有太大的影响,即在保证音视频同步并且时延在一个合理值范围内的前提下,时延对视频通信服务质量的影响可以忽略。

(3)时延抖动

抖动一般指视频流中两个连续数据分组的端到端时延的差值,使得视频流到达接收端的时间不恒定,如果抖动时间超出一定的范围,解码器通常会丢弃此数据分组,这将直接影响解码。因此抖动对需要规则化传输数据分组的视频通信服务质量有着显著的影响,具体表现为图像出现凝固或者快进。

(4)带宽

视频数据量比较大,而IP网络本身无法提供带宽保证,因此在特定分辨率下,媒体带宽需求存在最小值,低于这个最小值,就会由于网络拥塞而导致分组丢失、时延、抖动,影响视频服务质量,因此为了保证客户端的可用性,视频码流不可以过高;另一方面,为了保障图像质量,码流亦不可以过低,因此视频码流必须在可用性和视频服务质量之间进行折中考虑。

4 视频通信服务质量评价方法

视频质量评价可分为两大类:视频主观质量评价(subjective quality assessment)和视频客观质量评价(objective quality assessment)。视频主观质量评价就是直接利用人对视频的主观感受给出视频质量的度量,视频客观质量评价则是利用评估模型给出的量化指标自动计算视频质量。

4.1 视频主观质量评价方法

主观质量评价方法得到的结果是视频质量主观感受的最直接反映,符合视频通信服务的最终目的,因此,ITU制定了多个针对视频主观测试的标准,如ITU-R BT.500-10规定了电视图像质量的主观评测方法,ITU-T P.910、P.911规定了多媒体应用中的视频主观测试方法,ITU BT.1129和BT.1788规定了部分主观视频测试方法。

P.910中描述的4种多媒体应用的测试方法包括:绝对种类评分(absolute category rating,ACR)、带隐藏参考的绝对种类评分(absolute category rating with hidden reference,ACR-HR)、损伤种类评分(degradation category rating,DCR)、成对比较法(pair comparison method,PC)。

主观评价的评分标准常常采用五分制,分别与5个等级相对应,见表1。

表1 MOS评分等级

| 等级序号 | 主观感觉       |
|------|------------|
| 1    | 完全察觉不到     |
| 2    | 能察觉到但不令人讨厌 |
| 3    | 稍微令人讨厌的    |
| 4    | 令人讨厌的      |
| 5    | 非常令人讨厌的    |

主观质量评价虽然能够准确反映视频质量,但也有以下不足之处:

- 主观质量评价必须在严格的测试环境中进行,实现起来步骤复杂,成本较高,并且具备较强的个体差异性;
- 主观质量评价耗时长,实时性不高,不利于实时视频通信中视频服务质量的评价。

在本文实验中,主观测试中使用ACR方法,受测者观看通话视频,使用滑动设备进行评分,在无参考视频的情况下该方法比较合适。

4.2 视频客观质量评价方法

由于视频主观质量评价存在诸多不足,因此在实际应用中常采用视频客观质量评价方法。根据评估过程中是否需要将源视频文件作为参考,可以把视频客观质量评价方法分为全参考(full reference,FR)、部分参考(reduced reference,RR)、无参考(no reference,NR)视频客观质量评估方法3种。其中,无参考方法不需要将源视频文件作为参考,通过提取并分析失真视频中的失真特征信息(如模糊效应、分块效应、噪声效应等)对视频质量给出评分。由于这种方法不需要源视频,因此非常适用于实时视频通信业务视频服务质量的评价,但无参考方法对某些失真的敏感程度较低,无法检测设计时没有考虑到的失真类型。

由于人类视觉系统(human visual system,HVS)对不同类型失真的敏感程度不同,因此客观评价的结果往往与视频的主观质量不完全一致,不能很好地反映视频通信业务的服务质量。

综合上述分析,本文在构建视频通信服务质量模型时,采取主、客观评估方法相结合的方式,即先由客观评估方法获取大量实验数据构建模型,再由主观评估方法对模型进行验证和修正。

5 构建视频通信服务质量模型

现有的视频主客观评价标准大多基于传统电视图像,对新兴的多媒体通信应用考虑不足,相关评价指标和方法



都不完善,需要进一步加以研究。ITU-T G.1070 及 Amendment(2009 年修正案)建立了对于小格式(VGA 及以下分辨率)音视频点对点通信的 QoE 模型。G.1070 标准利用多个音视频编码参数和网络参数预测视频电话的质量,从统计的角度回归相关参数,利用这些参数预测视频电话的质量。关于 HD 分辨率的 QoE 模型 G.OMVAS,目前正在研究中,SG12(Study Group 12)计划在 2014 年通过最终版本。G.1070 以及 G.OMVAS 的基本模型结构如图 1 所示。

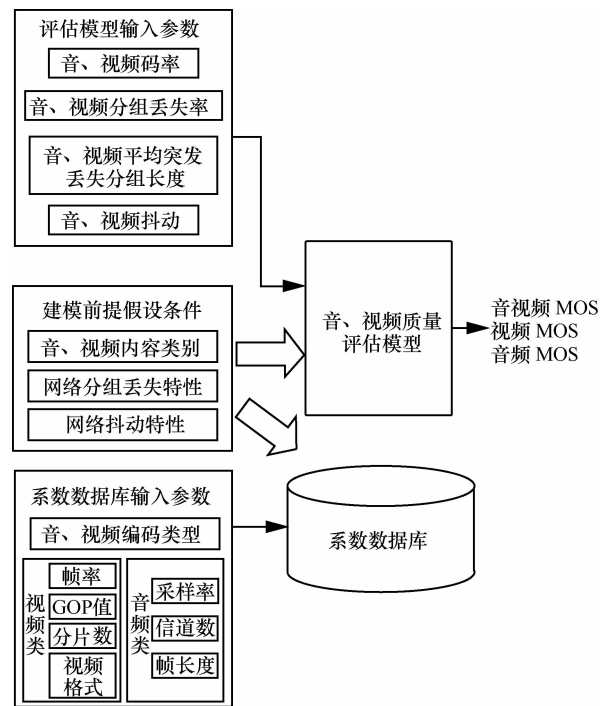


图 1 G.1070 提出的建议模型

G.1070 提出的建议模型表示,符合一定的假设条件时,通过输入一组音视频通话过程中提取的关键参数,如编解码、码率、分组丢失率、时延等,查找模型数据库中对应应该类型通话的系数值,根据这组系数值以及模型公式,估算本次音视频的通话质量。

模型数据库由网络侧参数(分组丢失率、抖动等)以及视频类参数(编解码类型、码率、帧率、视频格式等)组成,当对特定视频通信终端进行评测时,视频类参数已经确定,因此可以对模型进行简化,得到视频服务质量与网络参数之间的关系,如下所示:

$$MOS_{Video}=F(PLR,Jitter,Delay) \quad (1)$$

其中, $MOS_{Video}$  为视频 MOS 值,采用五分制; $PLR$  为分

组丢失率, $Jitter$  为抖动, $Delay$  为时延,需要注意的是,由于分辨率会对 MOS 值造成影响,因此应针对客户端提供的不同分辨率格式(如 CIF、VGA、720P 等)分别建模。

## 6 实验环境与实验方法

基于上述分析,需要在实验室搭建如图 2 所示的实验环境,模拟不同网络条件(包括分组丢失、抖动等)下,通过视频质量的主观及客观评价方法,评估各类智能终端的视频通信客户端产品的能力。通过实验,采样网络参数与视频质量的关系数据,分析模型原型、拟合系数,进一步建立视频通信产品的服务质量评估模型。

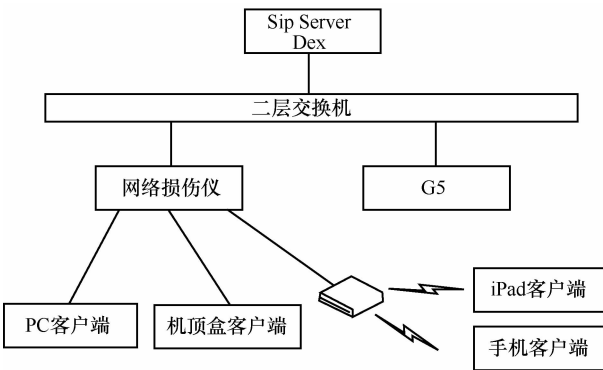


图 2 实验总体框架

实验采用的主要测试工具见表 2。其中,G5 和 Dex 是 Hammer 测试仪表,可以在局域网环境内分别模拟 IMS 核心网网元设备和视频通信终端,与被测视频通信客户端互通,并对视频质量进行客观评分,G5 集成的客观评估算法是一种无参考的视频客观质量评估算法(VQmon);网络损伤仪用来模拟不同的网络条件,包括分组丢失率、抖动、带宽等。采用 G5 测试仪表,与 H.264 终端互通,对视频质量做出评价。

表 2 测试工具描述

| 模块名称      | 模块描述                                                                                                        |
|-----------|-------------------------------------------------------------------------------------------------------------|
| Dex       | Sip Server,模拟 IMS 服务器                                                                                       |
| G5        | 测试仪表,与视频通信客户端互通,对通话质量进行客观评分                                                                                 |
| 家庭网关      | 接入视频通信客户端                                                                                                   |
| 网络损伤仪     | 模拟不同网络条件,包括分组丢失率、抖动、时延、带宽等                                                                                  |
| 各类视频通信客户端 | 模拟不同分辨率,所有客户端均采用 H.264 编解码(X264 编解码器);机顶盒为 Android 系统,内置海思芯片,采用罗技 C525 摄像头;PC 为 Windows XP 系统,采用罗技 C910 摄像头 |

实验时,改变网络损伤仪的设置,分别针对分组丢失、抖动等,控制网络损伤,将测试仪表 G5 仿真视频通信终端,与被评测的视频通信客户端通信,以获取客观 MOS 评分;为了减少误差,同样的网络条件下,进行 10 次通话,每次通话时长为 3 min,对 10 个 MOS 值取平均作为最终 MOS 值。改变网络参数,得到多组数据后,通过观察数据关系估计模型原型并得到系数。

基于大量测试数据,研究关键指标和视频服务质量的关系,通过对多个客户端的评测,不断修正评测方法和评测指标,最终为不同的分辨率提出视频质量评测模型。

7 视频通信服务质量的非线性模型建立

考察分组丢失率和视频质量的关系时,时延、抖动均设置为 0,不限带宽,改变分组丢失率取值。实验时,主要针对 PC 客户端(分辨率为 720P、CIF)和机顶盒客户端(分辨率为 VGA)进行测试,结果如图 3 所示。

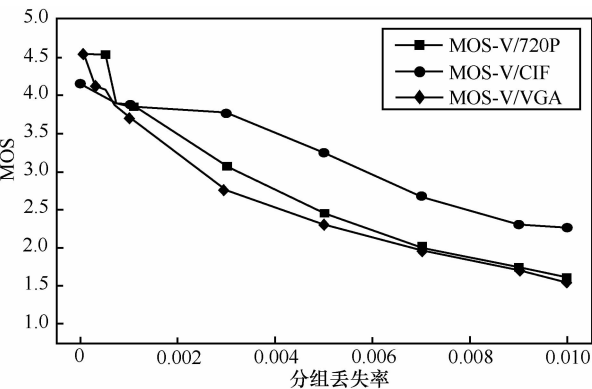


图 3 720P、VGA、CIF 对应的 MOS 与分组丢失率曲线

从图 3 中可以看出,随着分组丢失率的增大,MOS 值不断减小,即视频质量随之下降;当分组丢失率大于 1%时,MOS 值已降至 3.0 甚至 2.0 以下,并且基本不再变化,视频质量已经远远不可接受,继续增大分组丢失率,不再有意义。同时,主观观测的结果表示,分组丢失一旦出现,图像就会出现马赛克,在 MOS 值高于 4.0 时,视频质量基本不受影响;在 MOS 值低于 4.0 但高于 3.5 时;视频质量受到的影响较小,MOS 值低于 3.5 时,视频质量受到的影响较大。

建模时,模型应在满足准确性的前提下尽量保持简单易用。从图 3 可见,分组丢失率对视频质量的影响呈非线性形态,因此可采用多阶非线性回归方程对模型进行分析,回归方程形式如下:

$$y=y_0+A \cdot \exp(R_0 \cdot x) \tag{2}$$

其中, $x$  代表分组丢失率, $y$  代表 MOS 值, $y_0$ 、 $A$ 、 $R_0$  为待定系数。模型的准确度可以由拟合优度( $R^2$ )检验, $R^2$  越接近 1,则表示模型的拟合程度越高。

$$R^2=1-\frac{\sum_{i=1}^n(\hat{y}_i-y_i)^2}{\sum_{i=1}^n(y_i-\bar{y})^2} \tag{3}$$

其中, $\hat{y}_i$ 代表样本回归拟合得到的 MOS 值, $y_i$ 代表测试得到的 MOS 值, $\bar{y}$ 代表实验测得的 MOS 值的平均值。

所得的不同视频分辨率对应的回归方程系数见表 3。

表 3 拟合得到的回归方程系数

| 分辨率   | 720P     | VGA        | CIF       |
|-------|----------|------------|-----------|
| $y_0$ | 1.185 99 | 1.309 6    | 1.305 35  |
| $A$   | 3.394 05 | 3.156 65   | 2.880 29  |
| $R_0$ | -202.17  | -247.972 8 | -98.187 8 |
| $R^2$ | 0.986 62 | 0.993 56   | 0.976 41  |

由拟合优度可见,3 种模型的拟合优度均大于 0.90,都能比较好地拟合实验中的数据。拟合效果如图 4~图 6 所示。

考察时延和视频质量的关系时,分组丢失率、抖动均设置为 0,不限带宽,改变时延的设置。实验结果显示,时延设置为 100 ms、200 ms、500 ms 甚至 1 s,MOS 值均无变化,主

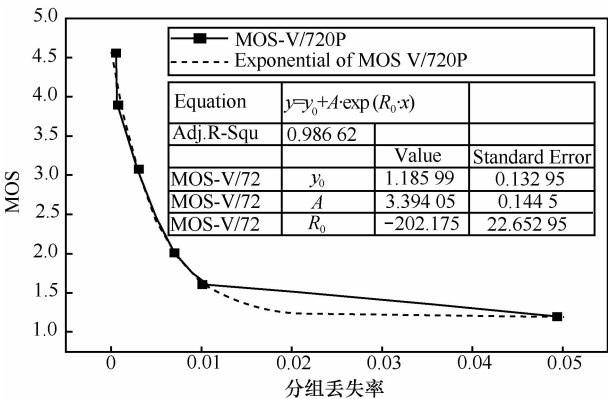


图 4 720P 的模型拟合效果

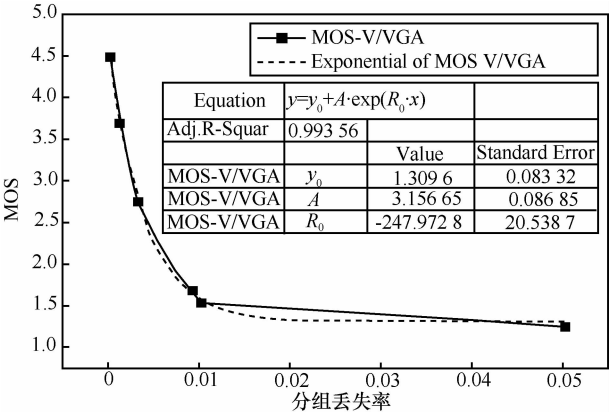


图 5 VGA 的模型拟合效果

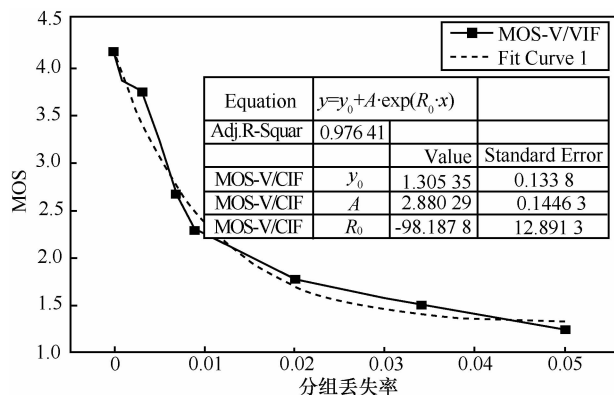


图6 CIF的模型拟合效果

观感受除了视频出现几秒时延之外,图像质量没有变化,而根据中华人民共和国通信行业标准《IP 网络技术要求——网络性能参数与指标》的规定,进行多媒体传输(视讯业务),网络时延上限不可超过 400 ms。实际网络情况的时延基本不会超过这个范围,可以验证时延对视频质量没有太大的影响。

## 8 实验室测评与验证

在实验室,通过对 1 000 多次视频通话的信息分析,提取分组丢失、抖动、有效码流、客观 MOS 评分等关键信息,建立视频通信服务质量数据仓库,得出评测模型。

为了验证模型的准确性,组织专家团队对视频服务质量进行主观评测,修正实验结果。评测时,采用 ITU-T P.910 中推荐的视频主观质量评估方法 ACR 作为研究和评价模型的基准,找到 8 名观测者为视频通信客户端的服务质量进行评分统计。

此外,对相同分辨率下的不同视频通信客户端进行测评,720P 分辨率下,两款视频通信客户端的 MOS 值随分组丢失率变化的曲线如图 7 所示。通过对比评估并结合视频通信的业务需求,可见客户端 Ju 更符合业务需求,客户端 Cai 需要按照评测模型进行优化和改进。随着视频通信

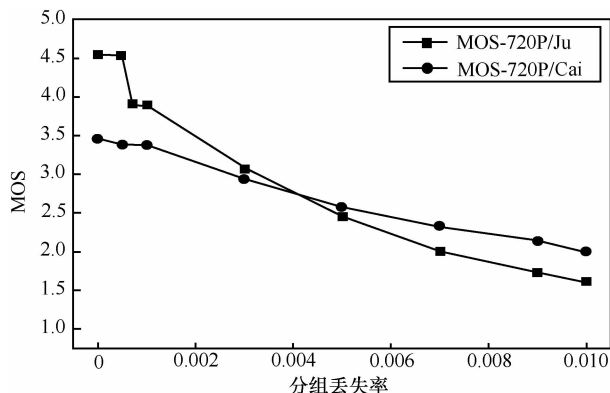


图7 720P 格式下两款客户端的 MOS 随分组丢失率变化曲线

客户端的版本更新和性能提升,可以持续积累实验数据,对评测模型进行不断修正。同时,评测模型也可以作为评价新入网视频通信客户端的标准。

## 9 结束语

本文围绕基于智能终端的视频通信业务视频服务质量,通过分析网络服务质量对视频服务质量的影响以及目前已有的视频主客观质量评估方法,制定了适合视频通信业务的测试标准,包括测试条件、测试方法、评分准则、数据处理等,建立了视频服务质量的评测模型并进行实验室优化验证,为运营商有针对性地提升视频通信业务的服务质量提供了测试依据和方法。由于目前测试仪表和测试手段的限制,抖动对视频质量的影响暂未有有效的测试结论,下一步将对该结论进行补充。

## 参考文献

- 1 ITU-T Recommendation P.910. Subjective Video Quality Assessment Methods for Multimedia Applications, 1996
- 2 ITU-R Recommendation BT.500-11. Methodology for the Subjective Assessment of the Quality of Television Pictures, 2002
- 3 ITU-T Recommendation G.1070. Opinion Model for Video-Telephony Applications, 2007
- 4 Study Group 12. Proposal on Opinion Model for Audio and Video Streaming Applications (G.OVAS), 2011
- 5 Hyun Jong Kim. A study on a QoS/QoE correlation model for QoE evaluation on IPTV service. Advanced Communication Technology (ICACT), the 12th International Conference, South Korea, 2010
- 6 IP 网络技术要求——网络性能参数与指标标准.YD/T 1171-2001, 2011
- 7 张大陆,张起强,胡治国等. IP 网络中视频流 QoE 评价模型. 计算机工程与应用, 2012(3)

## [作者简介]



**韩建亭**,男,硕士,中国电信股份有限公司上海研究院高级工程师、高级产品经理,主要从事家庭网络、固网移动融合、移动互联网融合通信等领域的技术研究和产品开发工作。

**张夙**,女,硕士,中国电信股份有限公司上海研究院产品开发工程师,主要研究方向为移动互联网应用、家庭宽带应用。

(收稿日期:2013-04-03)