

基于网络丢包的网路视频质量评估

卢刘明^{1),2)} 陆肖元¹⁾

¹⁾ (上海未来宽带技术及应用工程研究中心有限公司, 上海 200336)

²⁾ (中国科学院上海微系统与信息技术研究所, 上海 200050)

摘要 针对传统的视频质量评估模型对于在包交换网络中传输的压缩视频(网络视频)进行质量评估的局限性, 如对视频质量进行评估时实时性差、没考虑人眼的视觉感知等问题, 提出了一种基于网络丢包的网路视频质量评估模型。该模型采用了统计学理论, 它可根据视频编码、封装传输的特点, 通过量化的评价指标来评定视频质量受网络丢包的影响程度, 以评估网络视频质量。利用该模型进行网络视频质量评估, 不仅资源开销小、实时性好, 而且特别适合对实时传输的视频流进行视频质量评估。为了验证该模型的正确性与可靠性, 采用它与 MPQM 模型分别对受系列丢包率影响的视频序列进行了评估实验。实验结果表明, 该模型的评价结果能较好地吻合 MPQM 模型的评价结果, 且符合人眼的视觉效果。

关键词 网络丢包 网络视频 视频质量

中图法分类号: TN919.85 文献标识码: A 文章编号: 1006-8961(2009)01-0052-07

Quality Evaluation of Video Over a Packet Network Based on Packet Loss

LU Liu-ming^{1),2)}, LU Xiao-yuan¹⁾

¹⁾ (Shanghai Engineering Research Center for Broadband Technologies & Applications, Shanghai 200336)

²⁾ (Shanghai Institute of Microsystem and Information Technology (SIMIT), Chinese Academy of Sciences, Shanghai 200050)

Abstract To solve the limitation of the traditional evaluating models of video quality in accessing the quality of video over a packet network, such as unsuitable for on-line usage, not considering human visual perception, this paper introduces a video quality evaluating model based on packet loss. This model, which is inferred based on the statistics theory, can access the video quality over a packet network with a quantitative index according to the characteristics of video coding, packetizing and translation. It consumes little computing resources in quality evaluation and is suitable for real-time applications. In order to demonstrate the feasibility and accuracy of this model the simulations and experiments utilize this model and MPQM model, which objectively assesses video quality based on human visual properties, to evaluate a series of video sequences distorted with different packet loss rates and contrast their evaluation values. The experimental results have shown that the evaluation values of this model are approximately equal to MPQM values and reflect human visual perception.

Keywords packet loss, IPTV, video quality

1 引言

目前无论在 IT 业还是在研究界, 网络电视

(IPTV) 及其相关技术成了热门话题^[1]。网络电视要得到良好发展, 并成功运营, 如何确保为用户提供高质量的视频是其难题之一, 由于 IPTV 系统运营的是大规模流媒体业务, 而大规模流媒体业务又具

基金项目: 国家高技术研究发展计划(863)项目(2005AA103210)

收稿日期: 2006-12-30; 改回日期: 2007-10-17

第一作者简介: 卢刘明(1975 ~), 男, 在站博士后, 2006年在东华大学获控制理论与控制工程专业博士学位。主要研究方向为计算机网络、流媒体技术、视频质量评估。E-mail: lmlu@b-star.cn

有高宽带、实时性要求高、对丢包非常敏感的特性,而且由于现有的 IP 网络是一种尽力而为 (best effort) 的网络结构,因此如何建立有效的网络视频质量评估方法,在系统运营前分析数字视频采用哪种编/解码算法/格式,同时又可在系统运行时,对网络视频质量进行监控,以确保用户得到高质量的服务是 IPTV 良好发展与运营的关键。

传统的视频质量评估方法是基于视频编/解码算法的,这种方法仅通过它来评价由该算法产生的视频质量的优劣,而没有考虑用于传送视频的网络性能。这些评价方法主要分为主观评价与客观评价两类^[2]。其中,主观评价方法^[3]是人为地评价总体的视频质量,最常用的是由 ITU 推荐的 MOS (mean opinion score)。但是主观评价方法实时性不好,通常不能用于实时视频通信中的视频质量评估。客观评价方法一般通过比较原始图像序列与经编/解码后的图像序列来评价视频质量,常用的是峰值信噪比 (PSNR) 和均方误差 (MSE)^[4],由于它们是在像素级上进行操作,其对视频时域和空域上的失真非常敏感,而人的视觉系统对某些类型失真的敏感程度就没有这么高,所以得到的评价结果不一定与视频的主观质量评价完全一致,而且该方法需要原始参考视频,这在实时视频监控中较难做到。同时,在客观评价方面,Christian 等人根据人类视觉特性提出了一种基于人类视觉系统 (HVS) 的视频质量评价模型——MPQM (moving picture quality metric)^[5],该模型的主要特点是用一个 3 维滤波器组来模拟 HVS 的多通道特性^[6],但由于频率分解增加了算法的运算量,因此计算时间较长。此外, Mohamedden 等人还提出了一种无参考的实时视频质量评估方法^[7],它把所有影响视频质量的因素 (包括编/解码算法的影响以及网络性能等因素等) 利用神经网络进行训练来获得影响视频质量的因素与视频质量的近似关系,用来评价视频质量,但是由于神经网络计算复杂,其针对任何一个影响视频质量因素的不同数据都需要重新进行计算,所以可行性较差。Reibman 等人提出了基于均方误差的无参考网络视频质量评价方法^[8],但它得出的评价结果与主观的视频质量评价结果有较大差距。Verscheure 等人提出的网络视频质量评估方法^[9],由于没有考虑组层 (GOP) 结构、封装情况等因素,因此得出的评价结果误差较大。

为此,本文提出一种通过网络丢包情况来评估

网络视频质量的方法,它是通过分析视频传输的网络丢包情况,根据视频编/解码算法及组层结构、帧频、封装情况,利用可量化的评价指标来评定视频质量受网络损伤的影响程度,以评估网络视频的质量。由于影响网络视频质量的因素错综复杂,而本文则旨在关注网络性能对网络视频质量产生的影响,同时由于网络性能中的抖动、延时所导致的效果就是产生丢包,所以统一采用丢包率来衡量网络的性能。

2 网络视频对丢包的敏感性

2.1 视频的数据结构及其传输

网络视频形成过程如图 1 所示,原始图像序列通过编码器编码,并打包封装后在 IP 网络上进行传输,解码器收到视频流后通过解包并进行解码,最后输出解码后的视频序列。



图 1 网络视频的形成过程

Fig. 1 Production of IPTV video

在目前常用的 MPEG-2^[10-11] 编码方案中视频序列经多层划分,建立了由组层、帧 (画面)、片、宏块等构成的数据结构 (如图 2 所示)。其中,帧分为 I 帧 (帧内编码帧)、P 帧 (预测编码帧)、B 帧 (双向预测编码帧) 3 种类型。画面又被划分为若干个片,由于一幅画面内的任何一片均可以单独解码,而不

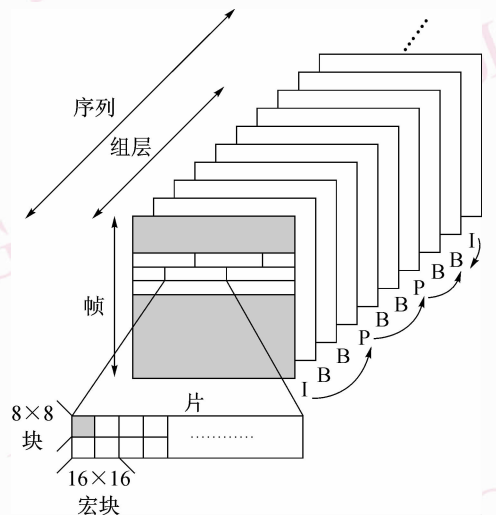
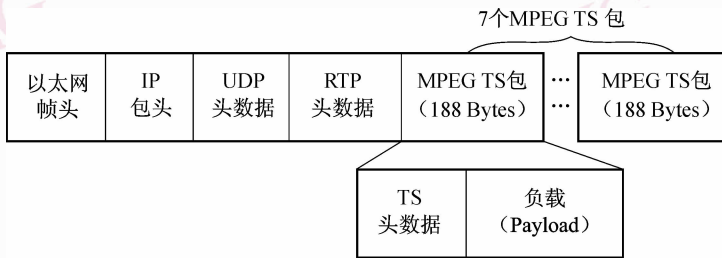


图 2 MPEG-2 视频序列编码的数据结构

Fig. 2 Encoding data structure of MPEG-2 video sequences

需要其他片的信息,因此片能提供对数据错误的抗干扰能力。每片开始时,用于预测离散余弦变换(DCT)的直流(DC)系数的预测值和运动矢量的预测值均需被复位为一个固定的数,并且片中的第1个宏块的位置可以从片头得到。每一片又划分为若干个宏块,宏块是运动补偿和量化器标尺变化的基本单位。

目前常用的视频数据在网络上传输时的封装结构如图3所示。传输时,首先,经编码后的视频数据形成视频原始流(ES),然后被打包成大小可变的原始流分组(PES),最后将视频原始流分组再组合成固定长度的传送流(TS)包。在对TS进行RTP打包时,一般7个MPEG TS包打成一个实时传输协议(RTP)包,完成IP封装后即形成一个IP-TS包。



注: IP-网际协议,UDP-用户数据报协议,RTP-实时传输协议,TS-传输流

图3 视频数据封装结构

Fig. 3 Encapsulation structure of video data

2.2 丢包对网络视频的影响

外界干扰、黑客破坏等原因,最终都会导致网络视频流丢包的产生,即视频流信息的丢失。网络视频不但对丢包特别敏感,且丢包对视频质量的影响程度取决于所丢掉的信息。语法信息(诸如头信息与系统信息)的丢失对视频质量的影响不同于语义信息(诸如运动向量、DCT系数等等)的丢失对视频质量的影响,而且质量降低的程度还取决于丢失的语义信息在视频画面中的位置。一般而言,语法信息的丢失所产生的影响比语义信息的大,因为语法信息的丢失可能导致整个数据块无法解码,而语义信息的丢失则只能影响与其相关的视频内容。

丢包所产生的误差包括起始误差以及它们在时空上传播而形成的误差,例如:(1)由于丢包将导致一个或更多片的丢失,通常对于固定长度的封装结构,一个包或者包含一部分的片内容,或者包含一个片结束部分的内容以及另外一个片开始部分的内容,或者包含帧头的信息,因此,一个包的丢失将导致以下3种情况的发生:丢失一个片;丢失两个片;丢失整个帧,而它们形成的初始误差还需根据解码器所采用的误差掩盖策略来确定;(2)由于视频编/解码算法的特性,致使初始误差将在时间与空间上进行传播,其传播程度是由丢失信息中宏块的类型以及丢包后接收到的宏块的运动信息决定。图4描述了在不同类型的画面中网络丢包与可视化信息之

间的映射关系。由于采用了有差别的编码、可变长的编码,因此数据丢包将在单个帧中蔓延,直到下一个重新同步点(例如帧头或片头)开始。这种现象称为空间误差传播,它可损害任何类型的帧。当丢包在参考帧(I帧或P帧)中发生时,损坏的宏块会影响以该错误宏块作为参考的后来帧中的非帧内编码的宏块。这种现象称为时间误差传播,它是由于帧之间的预测而产生的。

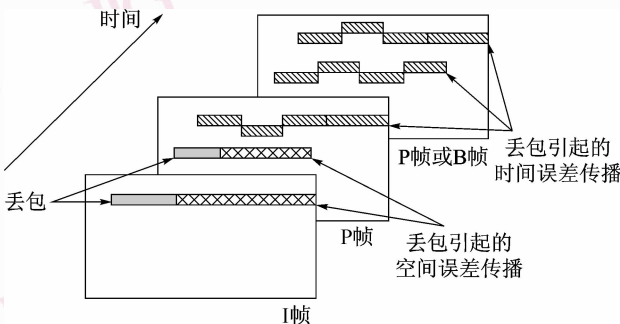


图4 MPEG-2 丢包引起的误差传播

Fig. 4 Error propagation produced by packet loss in MPEG-2

3 基于丢包现象的网络视频质量评估模型

影响网络视频质量的因素错综复杂,主要包括视频源/电视制式、视频编/解码处理、视频数据封装

处理、网络性能、视频图像复杂程度等,而本文则旨在关注网络性能对网络视频质量产生的影响,由于网络性能中的抖动、延时所导致的效果就是产生丢包,所以本文统一采用丢包率来衡量网络的性能。建立网络视频质量评估模型的目的是为了利用可量化的评价指标来评定视频质量受网络损伤的影响程度,以评估网络视频的质量。

根据视频编/解码算法以及网络传输的特点,本文利用概率统计学的理论,通过组层结构、帧频、封装情况,按以下步骤建立网络视频丢包率与网络视频质量之间的关系。

(1)对于特定的组层结构、帧频、封装情况,根据网络视频丢包率 p (伯努利^[12]丢包),即可得出单帧零丢包、单帧丢一个包的概率。

(本节下面用到的符号说明: F_0 表示单帧零丢包的事件, F_1 表示单帧丢一个包的事件, F_2 表示每帧丢两个包以上的事件; S_I 表示I帧包含包数的平均值, S_P 表示P帧包含包数的平均值, S_B 表示B帧包含包数的平均值; P_0^{pr} 表示前向参考帧(pr)零丢包的概率, P_0^{sr} 表示后向参考帧(sr)零丢包的概率, P_1^{pr} 表示前向参考帧丢一个包的概率, P_1^{sr} 表示后向参考帧丢一个包的概率, N_P 表示一个GOP中包含的P帧的数目, N_B 表示一个GOP中包含的B帧的数目)

用 $P(I)$ 、 $P(P)$ 、 $P(B)$ 分别表示视频序列中出现I帧、P帧、B帧的概率,其中

$$P(I) = \frac{1}{1 + N_P + N_B}, P(P) = \frac{N_P}{1 + N_P + N_B},$$

$$P(B) = \frac{N_B}{1 + N_P + N_B}$$

(注:一般情况下,每个GOP包含一个I帧,本模型的推导就考虑了这种情况)。

①单帧无丢包的情况如下:

$$P(F_0) = P(I)P(F_0|I) + P(P)P(F_0|P) + P(B)P(F_0|B)$$

$$P(F_0|I) = (1-p)^{S_I}$$

$$P(F_0|P) = \frac{1}{N_P} \sum_{k=1}^{N_P} (1-p)^{S_I+kS_P}$$

$$P(F_0|B) = \frac{1}{N_B} \sum_{j=1}^{N_B} (P_{j,0}^{pr} P_{j,0}^{sr} (1-p)^{S_B})$$

其中,当参考帧 r (pr或sr)为I帧时,则 $P_0^r = (1-p)^{S_I}$;

当参考帧 r 为P帧时,设P帧在单个GOP序列中的排号为 k ,则

$$k=1 \text{ 时, } P_{k,0}^r = (1-p)^{S_I};$$

$$k>1 \text{ 时, } P_{k,0}^r = (1-p)^{S_I+kS_P}$$

②单帧丢一个包的情况如下:

$$P(F_1) = P(I)P(F_1|I) + P(P)P(F_1|P) + P(B)P(F_1|B)P(F_1|I)$$

$$= S_I p (1-p)^{S_I-1}$$

$$P(F_1|P) = \frac{1}{N_P} \sum_{k=1}^{N_P} (P_{k,1}^r (1-p)^{S_P} + P_{k,0}^r p (1-p)^{S_P-1} S_P)$$

其中, $P_{1,1}^r = S_I p (1-p)^{S_I-1}$, $P_{1,0}^r = (1-p)^{S_I}$;

$k>1$ 时,

$$P_{k,1}^r = P_{k-1,1}^r (1-p)^{S_P} + P_{k-1,0}^r p (1-p)^{S_P-1} S_P$$

$$P_{k,0}^r = P_{k-1,0}^r (1-p)^{S_P}$$

$$P(F_1|B) = \frac{1}{N_B} \sum_{j=1}^{N_B} (P_{j,0}^{pr} P_{j,0}^{sr} p (1-p)^{S_B-1} S_B + P_{j,1}^{pr} P_{j,0}^{sr} (1-p)^{S_B} + P_{j,0}^{pr} P_{j,1}^{sr} (1-p)^{S_B})$$

其中,当参考帧 r 为I帧时, $P_0^r = (1-p)^{S_I}$, $P_1^r = p(1-p)^{S_I-1} S_I$;

当参考帧 r 为P帧时,设P帧在单个GOP序列中的排号为 k ,则

$$k=1 \text{ 时, } P_{k,1}^r = S_I p (1-p)^{S_I-1}, P_{k,0}^r = (1-p)^{S_I};$$

$$k>1 \text{ 时, } P_{k,1}^r = P_{k-1,1}^r (1-p)^{S_P} + P_{k-1,0}^r p (1-p)^{S_P-1} S_P, P_{k,0}^r = P_{k-1,0}^r (1-p)^{S_P}.$$

③单帧丢两个包以上的情况如下:

$$P(F_2|I) = 1 - P(F_0|I) - P(F_1|I)$$

$$P(F_2|P) = 1 - P(F_0|P) - P(F_1|P)$$

$$P(F_2|B) = 1 - P(F_0|B) - P(F_1|B)$$

$$P(F_2) = P(I)P(F_2|I) + P(P)P(F_2|P) + P(B)P(F_2|B)$$

(2)根据单帧丢一个包与丢两个包的劣化度,通过单帧零丢包、单帧丢一个包的概率,即可得出网络视频的质量。

设在无网络损伤情况下,视频质量为 Q_0 (由视频质量评分标准来评定,如表1所示),其值可随同流媒体包一起传输,设单帧丢一个包的劣化度为 D_1 ,一般取0.5~1;单帧丢两个包以上的劣化度为 D_2 ,一般取0.7~1,劣化度可根据画面内容的变化程度取值,变化程度大的取较大值,则受网络损伤后的视频质量 Q 计算如下:

$$Q = (P(F_0) + P(F_1)(1-D_1) + P(F_2)(1-D_2)) Q_0$$

由此公式可得到视频质量受网络损伤的影响程度,即 $P(F_0) + P(F_1)(1 - D_1) + P(F_2)(1 - D_2)$, 并可由未经网络损伤的视频质量 Q_0 来计算出网络视频的质量 Q 。(注:由于 $P(F_0) + P(F_1) + P(F_2) = 1$,且 $0 < D_1 < D_2$,所以若 $P(F_0)$ 越大,则 Q 越大,而 $P(F_1)$ 与 $P(F_2)$ 越大,则 Q 越小)。

表 1 视频质量评分标准

Tab. 1 Evaluation scores of video quality

质量分数	妨碍尺度	质量尺度
5	丝毫看不出图像质量变坏	很好
4	可看出图像质量变化但不妨碍观看	好
3	明显地看出图像质量变坏	一般
2	图像质量对观看有妨碍	差
1	图像质量对观看有严重妨碍	很差
0	觉察不到视频信号	无法觉察

4 实验验证

4.1 实验方案

为了确认网络视频质量评估模型的评估效果,本文通过实验对网络视频质量评估模型进行了验证,实验方案如图 5 所示。

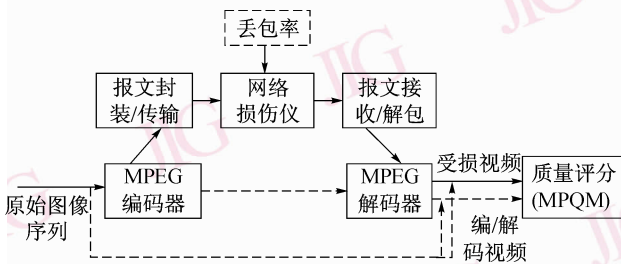


图 5 实验方案

Fig. 5 Experiment scheme

在实验中,本文采用 MPQM 模型^[5-6]对视频质量进行评分,即首先分析原始图像序列不经过网络传输而直接编解码后的视频质量的分值 Q_0 ,再分析它经网络传输后的视频质量情况;然后通过模拟不同的网络损伤来得出对应的视频质量分值 Q_1 、 Q_2 、 \dots ;最后把它们与根据网络视频质量评估模型得出的视频质量分值进行比较。

4.2 视频质量度量尺度

本实验先采用 MPQM 模型计算视频质量的实际值,然后把该值与本文所提模型的估计值进行比较。如图 6 所示,MPQM 模型是根据人类视觉系统

对不同方向、不同空间频率与时间频率带宽的信号(通道)具有不同敏感度的特性,先按照 5 种空间频率、4 种取向、2 种时间频率将原始图像序列和受损图像序列分解成若干个感知通道;然后在每一个通道中考虑对比敏感度和掩盖效应,以进行基于感知通道的图像序列损伤程度的计算,再把所有感知通道的数据组合起来进行综合感知;最后根据视频质量评分标准(详见表 1)得出受损图像序列的评分。

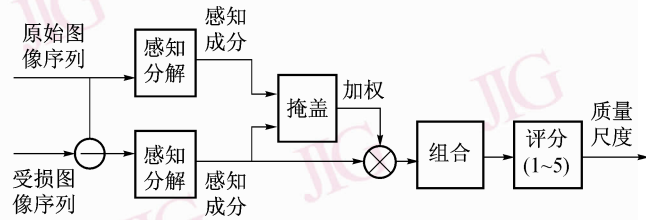
图 6 MPQM 模型结构^[5]

Fig. 6 Structure of MPQM model

4.3 实验结果

实验采用的组层结构为 IBBPBBPBB, I 帧的平均包含的包数为 66, P 帧的平均包含的包数为 37; B 帧的平均包含的包数为 16。视频按表 1 所示的评分标准进行评分,若原始视频的视频质量为 4.67,则表明画面内容变化程度较大,此时取单帧丢一个包的劣化度为 0.9,单帧丢两个包以上的劣化度为 1。

实验结果数据的获取通过以下两个阶段:

(1) MPQM 值的获取

利用思博伦公司的视频测试系统(VTS)^[13]对受损视频序列的视频质量进行测量来求得 MPQM 值(注:VTS 是以 v-factor 指标进行视频质量测量,而 v-factor 则是基于 MPQM 模型的)。对于每一个以给定丢包率损伤的视频序列,先利用 VTS 每 10 s 测量一次来得到它的 MPQM 值,进而得到一系列的 MPQM 值,然后求出它们的平均值,以作为该视频序列的总体 MPQM 值,即实际值。

由图 7 可看出,丢包率分别为 0、1/100、1/500、1/1000、1/2000 时候的 MPQM 值曲线。由图 7 可见,丢包率越大,视频劣化程度越明显,但是对于某个时间段,丢包率大的帧可能视频质量更好,这是由于丢失的信息不同而产生的。实验中取各曲线中特定时间范围内各采样点的 MPQM 值的平均值作为受该丢包率损伤的视频序列视频质量的实

际值。

(2) 估计值的获取

可由第 3 节的网络视频质量评估模型来求得估计值。实验结果如图 8 所示,由图 8 可见,由网络视频质量评估模型得出的估计值与 MPQM 值近似相等。同时,由实验结果可知,由于网络视频质

量对丢包比较敏感,当丢包率(注:指视频数据的丢包率)达到 1/1000,眼睛即可以明显感觉到视频质量的劣化,当丢包率大于 1/800 时,视频质量劣化程度增大,因此,由网络视频质量评估模型得出的估计值可较好地反映出实际的视频质量情况。

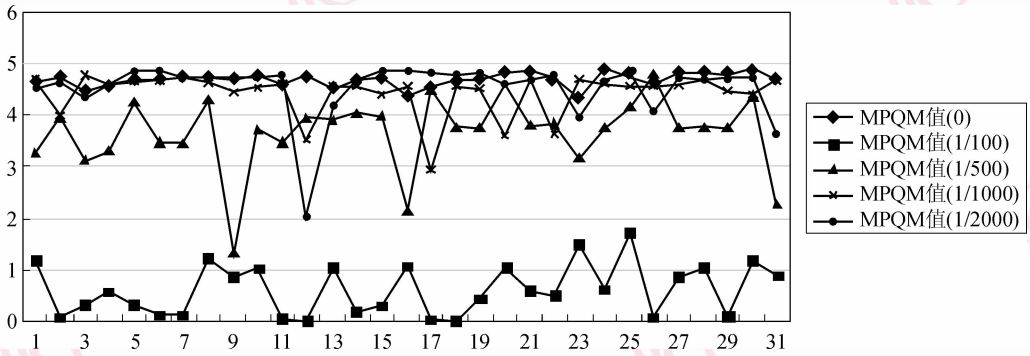


图 7 实验得到的 MPQM 值

Fig. 7 MPQM values produced in the experiment

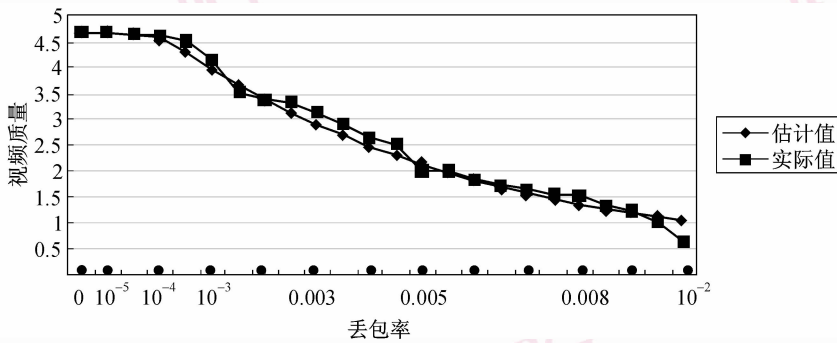


图 8 MPQM 模型的实际值与本模型的估计值的比较

Fig. 8 A contrast between MPQM values and the values produced from our model

5 结 论

针对传统的视频质量评估模型对包交换网络中传输的压缩视频进行质量评估中的局限性,如进行视频质量评估时实时性差、没考虑人眼的视觉感知等,提出了一种基于丢包现象的网络视频质量评估模型。该模型采用了统计学理论来得到,它可根据视频编码、封装传输的特点,通过可量化的评价指标来评定视频质量受网络丢包的影响程度,以评估网络视频质量。实验结果表明,该模型的评价结果能较好地吻合 MPQM 模型的评价结果,并能符合人眼的视觉效果。

参考文献 (References)

- 1 Hou Zi-qiang. Issues of IPTV's location and development [A]. In: The IPTV World Forum 2006[C], London, UK, 2006.
- 2 Tong Yu-bing, Hu Wei-wei, Yang Dong-kai, et al. A review on the video quality assessment methods[J]. Journal of Computer-Aided Design & Computer Graphics, 2006, 18(5):735-741. [佟雨兵,胡薇薇,杨东凯等. 视频质量评价方法综述[J]. 计算机辅助设计与图形学学报,2006,18(5):735-741.]
- 3 ITU_T Recommendation P. 901. Subjective Video Quality Assessment Methods for Multimedia Applications [EB/OL]. <http://www.itu.int>.
- 4 Wu H R, Lambrecht C, Yuen M, et al. Quantitative quality and impairment metrics for digitally coded images and image sequences [A]. In: Proceedings of Australian Telecommunication Networks &

- Applications Conference [C], Melbourne, Australia, 1996: 389-394.
- 5 Christian J. van Den Branden Lambrecht, Olivier Verscheure. Perceptual quality measure using a spatio-temporal model of the human visual system [A]. In: Proceedings of the SPIE [C], San Jose, CA, USA, 1996: 450-461.
- 6 Wang Nan-nan, Li Gui-lin. Video quality evaluation models based on human visual properties [J]. Journal of Image and Graphics, 2001, 6(6): 523-527. [王楠楠, 李桂苓. 符合人眼视觉特性的视频质量评价模型 [J]. 中国图象图形学报, 2001, 6(6): 523-527.]
- 7 Mohamed S, Rubino G, Cervantes F, *et al.* Real-time video quality assessment in packet networks: A neural network model [A]. In: Proceedings of International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA) [C], Las Vegas, Nevada, USA, June, 2001: 1071-1083.
- 8 Reibman Amy R, Vaishampayan Vinay A, Sermadevi Yegnaswamy. Quality monitoring of video over a packet network [J]. IEEE Transactions on Multimedia, 2004, 6(2): 327-334.
- 9 Verscheure Olivier, Frossard Pascal, Hamdi Maher. MPEG-2 video services over packet networks: Joint effect of encoding rate and data loss on user-oriented QoS [A]. In: Proceedings of the 8th International Workshop on Network and Operating Systems Support for Digital Audio and Video [C], Cambridge, UK, 1998: 257-264.
- 10 ISO/IEC JTC1/SC29 CD 13818-1, 2, 3. Information Technology- Generic Coding of Moving Pictures and Associated Audio Information- Part 1, 2 and 3 [S].
- 11 Zhong Yu-zhuo, Wang Qi, Zhao Li, *et al.* The International Standard of MPEG-2 Encoding and Compression of Moving Pictures and the New Development of MPEG [M]. Beijing: Tsinghua University Press, 2002. [钟玉琢, 王琪, 赵黎等. MPEG-2 运动图像压缩编码国际标准及 MPEG 的新进展 [M]. 北京: 清华大学出版社, 2002.]
- 12 William Feller (Translated by Hu Di-he). An Introduction to Probability Theory and Its Applications [M]. The Third Edition, Beijing: People Post Press, 2006: 112-182. [威廉·费勒著(胡迪鹤译). 概率论及其应用 [M]. 第 3 版, 北京: 人民邮电出版社, 2006: 112-182.]
- 13 Spirent Communications. Spirent Video Test System (VTS) [EB/OL]. <http://www.spirentcom.com/documents/3942.pdf>.