

考虑丢包特性的无参考网络视频质量评估模型

刘河潮, 杨付正, 常义林, 元辉

(西安电子科技大学综合业务网理论及关键技术国家重点实验室, 710071, 西安)

摘要: 针对相同丢包率条件下不同的包丢失集中度对视频质量影响不同的问题, 提出一种考虑丢包特性并且不需要对视频解码的无参考网络视频质量评估模型. 首先依据码流信息在压缩域计算视频的时域复杂度, 然后在压缩域计算包丢失集中度和包丢失率对视频质量的影响因子. 考虑到人眼视觉系统的空域掩盖效应和时域掩盖效应, 进而提出由于丢包引起的视频质量下降的视频质量评估模型. 该模型不需要对视频进行完全解码, 具有资源开销小、实时性好的特点, 适合对实时传输的视频流进行质量评估. 实验结果表明, 使用该视频质量评估模型测得的视频包丢失质量评分与其主观质量评分有很好的一致性, 相比于国际标准 G. 1070 中的丢包视频质量评价模型, 无参考网络视频质量评价模型与主观质量评价的相关性平均提高了 6.18%.

关键词: 视频质量评估; 丢包特性; 包丢失集中度; 时域复杂度

中图分类号: TN919.85 **文献标志码:** A **文章编号:** 0253-987X(2012)02-0130-05

A No-Reference Assessment Model for Quality of Networked Video Based on Features of Packets Loss

LIU Hechao, YANG Fuzheng, CHANG Yilin, YUAN Hui

(State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China)

Abstract: A novel no-reference video quality assessment model is proposed based on the effect of concentrative degree of packet loss ratio and features of packet ratio on video quality. Without decoding, the temporal complexity is calculated in terms of the effect of packet loss ratio on video quality. The concentrative degree of packet loss ratio is calculated in compression domain to determine the effect of concentrative degree of packet loss ratio on video quality. Then, the present video quality assessment model is proposed by taken spatial and temporal error concealment artifacts of human visual system, packet loss ratio, and the degree of concentration packet loss ratio into account. The proposed model can assess video qualities without fully decoding networked video, thus reduces the cost and time delay. Experimental results show that the scores obtained by the proposed model are highly consistent with the subjective assessment scores. Compared with the video quality evaluation model in G. 1070, a gain of 6.18% in the average performance is achieved by the proposed model in terms of relativity to subjective results.

Keywords: video quality assessment; feature of packets loss; concentrative degree of packet loss ratio; temporal complexity

宽带网络的迅速发展使得基于互联网协议 (Internet Protocol, IP) 网络的实时流媒体服务得

收稿日期: 2011-04-11. 作者简介: 刘河潮(1980—), 男, 博士生; 常义林(通信作者), 男, 教授, 博士生导师. 基金项目: 国家自然科学基金资助项目(60902081, 60902052); 高等学校学科创新引智计划资助项目(B08038); 西安电子科技大学基本科研业务费资助项目(72105457).

网络出版时间: 2011-11-09

网络出版地址: <http://www.cnki.net/kcms/detail/61.1069.T.20111109.0842.001.html>

到了广泛的应用,而其中的多媒体通信正逐渐成为通信网的主要业务,因此对视频质量进行实时和准确的评估对视频的编解码发展、网络协议规划、网内质量监控和确保终端用户质量等至关重要^[1-2]. IP网络是一个异构、时变、缺乏服务质量(QoS)控制的网络,它的设计初衷是为了传输数据,但并不适合传输连续的流媒体.虽然它的传输效率高,但带宽波动范围大,易发生丢包现象.视频通信,如会议电视、可视电话等,主要内容是图像和视频媒体.视频通信的数据是经压缩编码后的数据,而压缩后的数据对丢包造成的误码非常敏感,使得图像的恢复质量比较差.同样的丢包率对视频质量造成的损伤也是不一样的,由于有些丢包只影响一帧的视频质量,有些会持续影响后面多帧的视频质量,有些丢包发生在场景的活动区域,有些发生在静止区域,这些不同情况的丢包并不都会被人类视觉所感知到,因此准确评估由于丢包所引起的视频质量下降比评估静止图像更加困难^[3].

视频质量评估可以分为主观视频质量评估和客观视频质量评估,这两种视频质量评估各有自己的优点和缺点.主观视频质量评估是通过一组用户观看被测视频来给出视频质量等级^[4].如果主观测试实验能够正确地设计并执行,这种方法是最为准确的,但是主观测试费时、费力,结果稳定性差,可移植性也差.客观视频质量评估就是利用影响视频质量各种因素建立一个数学模型,最终由计算机根据该模型客观地给出评分来评定视频的质量.根据可利用到的原始视频信息的多少,客观质量评价方法可分为^[5-6]:全参考质量评价(full-reference)、半参考质量评价(reduced-reference)和无参考质量评价(no-reference).

近年来面向网络多媒体消费电子产品的快速发展,使得网络视频质量的评估变得越来越受到人们的青睐并成为近几年的研究热点^[7-10].文献^[7-9]所提出的方法主要考虑平均丢包对视频质量的影响,并且利用的是均方误差(mean squarer error, MSE)来评价视频的质量,并没有考虑人类视觉系统(human visual system, HVS)的特性.文献^[10]虽然考虑了HVS特性,但是没有考虑相同丢包率情况下不同的包丢失集中程度对视频的质量影响.事实上,网络视频质量的下降不仅受丢包率的影响,还受到包丢失的集中程度和视频内容复杂程度等因素的影响.有鉴于此,本文在研究了丢包率对视频质量的影响的同时,研究了相同丢包率情况下不同的丢包集

中度对视频质量的影响,由此提出一种考虑包丢失集中度的影响和视频内容复杂度的无参考视频质量评估模型,该模型还考虑到人眼视觉系统的空域掩盖效应和时域掩盖效应对视频质量的影响.实验结果表明,使用该方法测得的视频质量评分和主观质量评分有很好的 consistency.

1 基于包丢失率的模型

1.1 丢失率对视频质量的影响

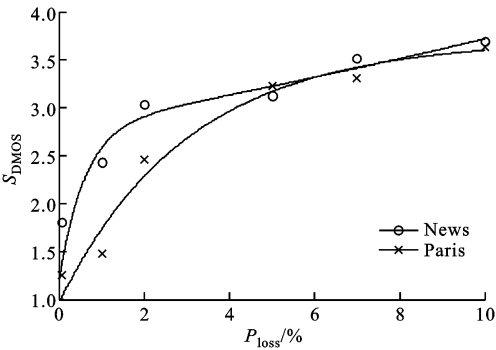
由于网络的时变不稳定特性,编码压缩后的视频在通过IP网络传输时,包丢失、时延和抖动等是不可避免的,时延和抖动可以等效成包丢失.在编码方式一定情况下,包丢失是引起网络视频失真主要原因,因此包丢失率是反映视频由于包丢失引起失真的重要信息.网络视频传输过程中,许多因素都会导致数据包被丢弃,如信道带宽大小、数据包在什么时间发送、链路是否拥塞等.为了评估不同丢包率下的视频恢复质量,首先要建立一个丢包模型,模拟网络传输中丢包的发生.

本文在实验中选择独立丢包信道模型.它假定每个数据包在网络上传输时被丢弃的概率是独立的.在这个模型中,只有两种结果——丢包和不丢包,假设丢包的概率用 P_{loss} 表示,则不丢包的概率为 $(1-P_{\text{loss}})$.所以该模型也称为贝努利模型.当发包总数为 N 时,则丢包的总数为 NP_{loss} .丢包概率可以通过丢包总数/发包总数来计算.该模型比较简单且计算的复杂度不高.

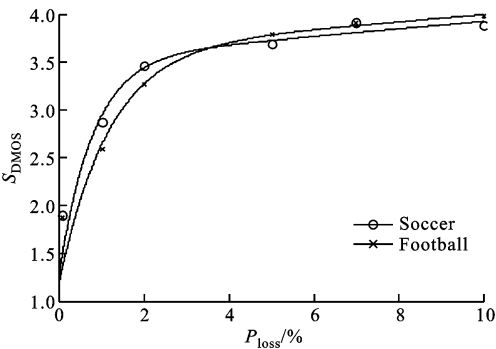
为了研究网络视频由于丢包造成的视频质量损伤与恢复视频质量之间的关系,本文进行了大量的实验,首先得到了视频的主观质量.实验采用独立丢包信道模型,采取均匀丢包策略产生丢包,视频选用CIF标准视频序列,分辨率为 352×288 像素,帧率为30帧/s,编码速率为512 kb/s,视频编码器是MPEG-4.在这些主观视频质量实验中,用解码后视频的平均意见得分 R_{MOS} (mean opinion score, MOS)表示视频的绝对等级分值 R_{ACR} (absolute category rating, ACR).由包丢失引起的视频质量下降分值用 S_{DMOS} (differential MOS, DMOS)表示.

对大量主观测试实验结果的研究发现,由于包丢失引起的视频的主观质量下降与包丢失率的关系可以用包丢失率的指数函数形式来估计,本文给出了几种代表在空域和时域复杂度都不同的视频序列News、Paris、Mother-Daughter、Mobile、Soccer、Football的 S_{DMOS} 与丢包率的关系图,如图1所示.由图1可

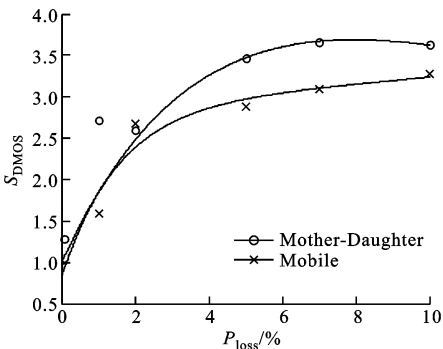
知,随着丢包率的增大,由于网络包丢失引起的网络视频质量的失真也同时增大,呈指数增加,且随着丢包率的增加而增速减慢并趋于稳定,不超过 4.



(a)Paris 和 News



(b)Soccer 和 Football



(c)Mother-Daughter 和 Mobile

图 1 丢包率与视频质量下降分值的关系

此外通过分析发现,对于不同的视频序列,其主观质量下降曲线的梯度各不相同.这是因为人类视觉系统的感知能力不仅与丢包率有关,还受视频内容的空域复杂度和时域复杂度的影响.

1.2 包丢失集中度对视频质量的影响

上面研究了均匀丢包策略情况下,丢包率对视频质量的影响,发现随着丢包率的增加,视频质量下降增大,但是增速减慢并趋于稳定.然而,在实际 IP 网络中时常发生的是非均匀的丢包,那么在随机丢包情况下,丢包率对视频质量的影响又是如何呢?

通过研究发现,在同样的丢包率情况下,若丢包集中程度不同,对视频的质量影响也不同,包丢失越集中则对视频的质量影响越大.在同样的丢包率下,包丢失宽度越小则包丢失集中度越高.包丢失集中度可以表示为

D_{pl} = \frac{c(kN - 1)}{kN(L_{loss} + c)} \tag{1}

式中:D_{pl}为包丢失集中度;N 为一段视频丢失包的总数;L_{loss}为丢包跨度,指的是在单位时间内第一个包丢失到最后一个包丢失的距离;c 和 k 为待定常量.显然,在相同的包丢失跨度内,丢包数量越多,则包丢失集中度越高;丢包数量越少,则包丢失集中度越低.在丢失一个包的情况下 L_{loss}=0.

1.3 时域复杂度

目前已经提出的与视频内容相关的因子主要有以下 3 种.

(1)初始均方误差(initial mean squared error, IMSE),是指无误码的恢复图像与经过误码掩盖后的恢复图像之间的 MSE.

(2)残余能量(residual energy),是指经过运动补偿以后的残差信息的能量(所有 DCT 系数的平方和).如果丢失了一帧数据,即使采用的误码掩盖技术能够准确的预测丢失的运动矢量,但恢复视频还是与原始视频存在差异,残余能量就是用来衡量这种差异大小的.

(3)运动相关因子(motion-related factors),主要是运动矢量(motion vector, MV).

由于前 2 种因子的计算都要用到原始视频数据,因此不适合用于在网络中间节点对视频质量的监控,应用性受到限制;运动矢量虽然不需要原始视频,但是必须对视频进行完全解码后才能得到,由于视频的加密或者版权问题,因此无法对视频完全解码,应用同样也受到限制.由于以上原因,本文采用时域复杂度来表征视频内容.

一般说来,对于视频的时域复杂度,可以用运动补偿后的残差信号的方差或视频的运动矢量来表示.视频时域复杂度越高,则残差信号像素的方差或运动矢量就越大.但是,在实际使用时,由于版权、加密或者计算复杂度的要求而不能对视频完全解码,因此也就不能得到像素的 MSE、方差或运动矢量值.

在无参考视频质量评估中,不对视频流进行完全解码,只对码流的包头信息进行解析进而确定视频的时域复杂度.通过对包头解码后得到的信息包括:编码量化参数 Q_p(quantitative parameter, QP)、

帧类型、编码速率、丢包的数量、丢包的位置和每一帧在显示器上的显示时间. 获取上述信息后,本文采用文献[10]的时域复杂度模型

$$\sigma = Q_P(aR + b) \tag{2}$$

式中: σ 表示视频的时域复杂度; Q_P 是编码量化参数; R 为编码速率; a 和 b 为待定模型参数.

1.4 网络丢包无参考视频质量评估模型建立

全参考视频评估和半参考视频质量评估需要完全解码或部分解码并且还要传输额外的信息,不仅耗时而且耗费带宽资源,使得在实际应用当中不是很现实,并且由于版权或者加密的原因,无法对视频进行完全解码,因此无参考的不需对视频进行完全解码的视频质量评价最适合实时的网络视频评估.

由 1.1 和 1.2 节的分析可知,不同的丢包率对视频的质量影响不同,在同样丢包率不同的丢包集中度和不同内容的视频对视频质量的影响也不同. 由此,本文提出一种结合视频内容、丢包率和丢包集中度的无参考视频质量评估模型

$$V_Q = V_C D_{pl} \left(1 - a_1 \exp \left(- \frac{a_2}{P_{loss}} \right) \right) \tag{3}$$

式中: V_Q 表示客观视频质量评分; V_C 表示在没有丢包的情况下,经过压缩后的视频质量评分; $1 - a_1 \exp \left(- \frac{a_2}{P_{loss}} \right)$ 表示丢包引起的视频质量下降因子,该因子由视频的时域复杂度决定,其中 a_1 和 a_2 是与视频的时域复杂度相关的函数

$$a_1 = a_3 + a_4 \ln(a_5 \sigma) \tag{4}$$

$$a_2 = a_6 + a_7 \sigma \tag{5}$$

a_3 、 a_4 、 a_5 、 a_6 和 a_7 为待定参数, σ 为视频的时域复杂度.

2 实验结果与分析

本文所有实验均采用 MPEG-4 编解码器,使用的标准视频序列格式为 CIF(352×288 像素),帧率为 30 帧/s, Q_P 设为 5、8、15、20、25、28、30 和 35. 丢包采用随机丢包策略,丢包率分别为:0%、0.5%、1%、2%、5%、7%、10%. 由丢包引起的初始误码在解码端使用零运动误码掩盖(zero-motion error concealment, ZMEC)方法处理,即丢失的宏块直接用与它最接近的参考帧同一位置上的宏块来代替,对于丢失数据帧则直接用前一参考帧代替. 分别使用本文提出的质量评估方法和国际标准 G. 1070 方法^[11]对所有失真视频进行质量评估,然后与视频的主观质量比较来评价该方法的性能. 使用本文提出的评估方法对网络视频评分时,模型参数分别取值

为: $c = -3.993\ 7$, $k = 0.050\ 3$, $a = 1\ 260$, $b = 0.003$, $a_3 = 0.755\ 3$, $a_4 = 0.031\ 6$, $a_5 = 0.038$, $a_6 = -0.956\ 9$, $a_7 = 0$.

本文的主观测试采用单激励方法(single stimulus methods, SSM),相关测试条件参照文献[12]中 VQEG 指定的标准执行. 使用主观平均分 S_{MOS} 表示视频质量的好坏,具体主观质量分级标准如表 1 所示.

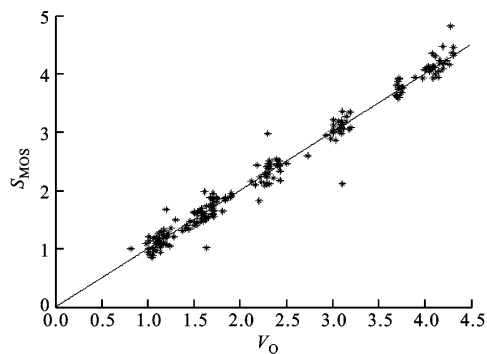
表 1 视频主观质量判分标准

S_{MOS}	评分尺度	质量等级
5	质量很好,没有察觉失真	很好
4	察觉失真但不反感	好
3	察觉失真且稍微反感	一般
2	察觉失真且令人反感	差
1	察觉失真且令人很反感	很差

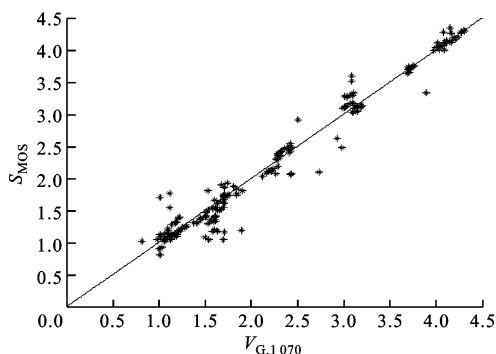
采用以下 4 个尺度来量化模型的性能:相关系数 C_{CC} (correlation coefficient, CC), 逸出度 R_{OR} (outlier ratio, OR), 均方根误差 E_{RMSE} (root mean square error, RMSE) 和 Spearman 秩相关系数 C_{SROCC} (Spearman's rank-order correlation coefficient, SROCC). 相关系数表示客观和主观分数的相关度,它给出对预测精度的评价,较大的相关系数值意味着较好的精度. 逸出度和均方根误差表示预测的一致性,逸出度和均方根误差的值越小,预测的一致性越高. Spearman 秩相关系数是另一个相关度指标,较大的 Spearman 秩相关系数意味较好的精度. 表 2 给出了 2 种质量评估模型的 4 个性能评价指标,由表 2 可知本文所提方法得到的视频质量和视频主观质量有很好的一致性. 为了更方便和清楚的说明本文提出方法的有效性,图 2 给出了本文模型和国际标准 G. 1070 模型的客观质量评分与主观评分的对应分布图,其中 $V_{G.1070}$ 表示用国际标准 G. 1070 方法得出的视频质量的客观分值,图中的每一个数据点对应一个受测失真视频序列,从散点图可以清晰看到本文所提方法可以准确地评估视频的丢包失真,本文提出的方法与相应的主观质量具有较好的一致性.

表 2 2 种质量评估模型的性能指标

模型	C_{CC}	E_{RMSE}	C_{SROCC}	R_{OR}
国际标准 G. 1070 模型	0.896 1	0.183 7	0.803 9	0.202 8
本文模型	0.957 9	0.081 7	0.902 6	0.092 2



(a) 本文模型



(b) 国际标准 G. 1070 模型

图2 2种模型的客观质量评分与
主观质量评分对应分布图

3 结 论

本文提出一种不对视频解码无参考的网络视频质量评估模型,并且考虑了视频内容、丢包率和丢包集中度对视频质量的影响.该模型不对视频进行完全解码,只是对接收到的码流进行简单的头分析,所以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.实验结果表明,使用该方法能够较准确地反映视频的主观质量.

参考文献:

- [1] 马震远,周杰,陈楚,等.一种分段测量保证 QoS 约束的任播通信模型[J].西安交通大学学报,2010,44(2):44-49.
MA Zhenyuan, ZHOU Jie, CHEN Chu, et al. An any cast communication model with QoS constraints based on segmental measurement [J]. Journal of Xi'an Jiaotong University, 2010, 44(2): 44-49.

- [2] MAISONNEUVE J, DESCHANEL M, HEILES J, et al. An overview of IPTV standards development [J]. IEEE Transactions on Broadcasting, 2009, 55(2): 315-328.
- [3] RAO K R, BOJKOVIC Z S. Packet video communications over ATM networks [M]. Upper Saddle River, NJ, USA: Prentice Hall, 1999: 119-121.
- [4] Telecommunication Standardization Sector. Subjective audiovisual quality assessment methods for multimedia applications [S]. Geneva, Switzerland: Telecommunication Standardization Sector, 1998.
- [5] WU Hongren, RAO K R. Digital video image quality and perceptual coding [M]. Boca Raton, FL, USA: CRC Press, 2006: 87-122.
- [6] Telecommunication Standardization Sector. User requirements for objective perceptual video quality measurements in digital cable television [S]. Geneva, Switzerland: Telecommunication Standardization Sector, 2000.
- [7] REIBMAN A R, VAISHAMPAYAN V A, SERMADEVI Y. Quality monitoring of video over a packet network [J]. IEEE Transactions on Multimedia, 2004, 6(2): 327-334.
- [8] TAO Shu, APOSTOLOPOULOS J, GUERIN R. Real-time monitoring of video quality in IP networks [J]. IEEE Transactions on Networking, 2008, 16(5): 1052-1065.
- [9] NACCARI M, TAGLIASACCHI M, TUBARO S. No-reference video quality monitoring for H. 264/AVC coded video [J]. IEEE Transactions on Multimedia, 2009, 11(5): 932-942.
- [10] YANG Fuzheng, WAN Shuai, XIE Qingpeng, et al. No-reference quality assessment for networked video via primary streaming analysis of bit stream [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(11): 1544-1554.
- [11] Telecommunication Standardization Sector. Opinion model for video-telephony applications [S]. Geneva, Switzerland: Telecommunication Standardization Sector, 2007.
- [12] VQEG. Multimedia group test plan 1.21 [DB/OL]. (2008-03-09) [2010-12-10]. http://www.its.bldrdoc.gov/vqeg/projects/multimedia/mm_phase_i_testplan_v1.21_final.doc.

(编辑 刘杨)