H.264/AVC 网络视频编码失真评估的比特流层模型

苏洪磊，李兵兵，杨付正
（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：为了实现对采用H.264/AVC标准编码的网络视频质量的实时监测，提出一种评估视频编码质量的比特流层模型.该模型无需完全解码，只通过简单解析视频的量化参数、编码比特率以及运动矢量等信息评估视频流质量.首先，通过主观评估实验分析确定量化参数与视频编码失真的基本关系模型，然后利用量化参数和编码比特率预测视频的空间复杂度，利用运动矢量信息预测视频的时间复杂度，并结合空域掩盖效应和时域掩盖效应建立起一个能够反映人视觉特性的视频质量评估模型.实验结果表明，与解析码流的无参考网络视频质量评估模型相比，利用该模型得到的客观质量分数与主观质量分数的皮尔森相关系数提高了0.016 0，均方根误差下降了0.079 7.

关键词：视频质量评估；网络视频；编码失真；比特流层模型；H.264/AVC标准
中图分类号：TN919.8 **文献标志码**：A **文章编号**：0253-987X(2012)06-0036-06

A Bitstream-Layer Model for Assessing Coding Distortion of H.264/AVC Networked Video

SU Honglei, LI Bingbing, YANG Fuzheng
(State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China)

Abstract: A bitstream-layer model for video coding distortion assessment is proposed to realize real-time quality monitoring for H.264/AVC networked video. The model realizes real-time quality assessment for networked video without resorting to full decoding, but by simply analyzing the payload information such as the quantization parameters, bit-rate and motion vectors. Firstly, the relationship between the coding distortion and quantization parameters for different video sequences is modeled. Then, the quantization parameters and bit-rate are employed to predict the spatial complexities, and the motion vectors are employed to predict the temporal complexities. After empirical data analysis, a mathematical model for video quality assessment is established taking account of the spatial and temporal masking effect of the human visual system. Experimental results show the advanced performance of the proposed model. A comparison with the NR QANV-PA model shows that the proposed model gets an increment about 0.016 0 in the Pearson Correlation Coefficient between the objective scores and mean opinion scores and a decrement about 0.079 7 in the root mean square error.

Keywords: video quality assessment; networked video; coding distortion; bitstream-layer model; H.264/AVC standard

---

收稿日期：2011-10-17. 作者简介：苏洪磊(1986—)，男，博士生；杨付正(通信作者)，男，副教授. 基金项目：国家自然科学基金资助项目(60902081，60902052)；国家国际科技合作专项资助项目(2010DFB10570)；中央高校基本科研业务费专项资金资助项目(720048885)；高等学校学科创新引智计划资助项目(B08038).
网络出版时间：2012-03-21 网络出版地址：http://www.cnki.net/kcms/detail/61.1069.T.20120321.1447.001.html

多媒体通信的广泛应用带来了网络视频业务的高速增长，有限的网络带宽和用户对视频高质量要求的矛盾已经显现. 如何对网络视频的质量进行实时有效的评估，进而及时改变编码和传输策略[1]，使用户获得更好的视觉体验，对于网络服务商就显得尤为重要.

主观质量评估虽然可以正确反映视频质量，但其操作复杂，成本高，并且不能满足实时性要求，因此客观质量评估在网络视频业务中具有重要的作用. 根据输入参数的不同，客观质量评估又可分为参数规划模型[2-4]、包层模型[5-7]、比特流层模型[8-10]、媒体层模型[11-12]和混合型模型[13-14]. 参数规划模型利用网络参数和应用参数进行网络视频质量评估，这些参数包括比特率、帧率、丢包率以及时延信息等. 然而，参数规划模型只能利用统计信息得出视频质量的平均值，不能准确地评估出单一网络视频的质量. 包层模型通过分析数据包头得到的信息进行网络视频质量评估，由于无需介入载荷信息，因此包层模型非常适用于载荷信息被加密的网络视频. 然而，包层模型仅可利用网络视频流的包头信息，评估的准确度不高. 媒体层模型需要对视频码流进行完全解码，得到像素域信息，评估时考虑到了影响人类视觉感知的因素. 混合型模型可以利用上述 4 种模型中所用到的所有信息来进行网络视频质量评估，所以使用该模型可以得到较高的评估性能. 但是，媒体层模型和混合层模型需要对视频流进行完全解码，不能满足实时性要求.

比特流层模型不仅可以使用包层模型中可获取的参数信息，并且可以通过对视频流的简单解析得到帧类型、量化参数以及运动矢量等相关信息，通常具有较好的性能，而且该模型不对视频流进行完全解码，计算复杂度低，因此适合于对大量视频流的实时评估. 我们曾在文献[8]中提出了一种视频质量评估方法，该方法通过解析数据包头及视频帧头，得到量化参数、帧类型、比特率、帧播放时间以及丢包个数、丢包位置等信息，进而建立了一种基于率失真函数的视频质量评估的比特流层模型. 该模型基于音视频编码标准 MPEG4[15] 而建立，但随着多媒体通信的发展，视频编码专家组(VCEG)和动态图像专家组(MPEG)联合制定了目前性能最好的视频编码标准 H.264/AVC[16]. H.264/AVC 具有很好的压缩性能，在各种网络视频业务中具有良好的表现，并且其应用越来越广泛，因此针对于采用 H.264/AVC 标准的网络视频进行实时正确地评估日趋重要.

本文首先确定了量化参数与视频编码失真的基本关系模型，然后结合空域掩盖效应和时域掩盖效应，提出了一种适用于 H.264/AVC 网络视频编码失真评估的比特流层模型.

1　编码失真与量化参数

量化是视频编码失真的根本原因，因此量化步长是反映编码失真的重要信息. H.264/AVC 使用标量量化器并将变换与量化联合进行[16]. 参照量化步长与量化参数的索引对照表，两者关系可近似表示为

$$Q_s = 2^{(Q_p-4)/6} \tag{1}$$

式中：Q_s 为量化步长；Q_p 为量化参数. 由式(1)可知，量化参数决定了量化误差，进而影响了视频的编码失真. 量化参数大，则视频的量化误差大. 然而，人类视觉系统对于视频的感知是非均匀和非线性的. 为了研究 Q_p 与人眼感知的编码失真之间的关系，本文选取不同内容特性的视频序列 Mobile(具有较高的空间复杂度)、Soccer(具有较高的时间复杂度)和 News(具有较低的空间复杂度和时间复杂度)，对各视频序列在固定 Q_p 模式下进行编码. 实验按 Q_p 升序在区间[20,48]内每隔 2 个值选取 1 个量化参数，然后对生成的各视频序列进行主观测试实验，得到平均主观质量分数 V_{MOS}(mean opinion score, MOS). 图 1 为不同内容特性的视频序列的 V_{MOS} 随 Q_p 变化的示意图，可以看出，量化误差引起的失真在一定范围内不能被人眼察觉，随着量化参数的增大，量化误差会引起越来越明显的方块效应，使得主观质量分数下降.

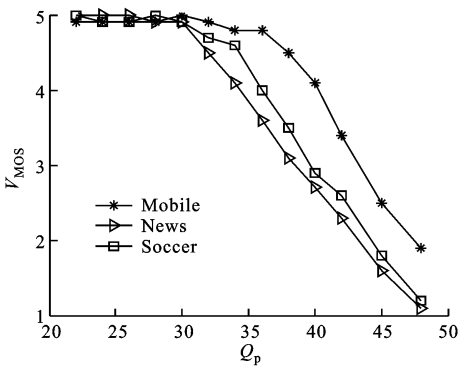


图 1　不同内容特性的视频序列 V_{MOS} 与 Q_p 的关系

从图 1 还可以发现，不同的视频序列其 V_{MOS} 下降的起始 Q_p 不同，并且随着 Q_p 的增大，各 Q_p 对应的 V_{MOS} 也各不相同. 这是因为视频内容对编码失真

产生了符合人类视觉系统特性的空间掩盖效应和时间掩盖效应[17]. 例如:Mobile 序列具有较高的空间复杂度,当出现编码失真时,其较丰富的空间纹理特征对编码失真具有空域掩盖效应,削弱了人眼对方块效应的感知,从而使得 V_{MOS} 较大;对于时间复杂度较高的 Soccer 序列,当出现编码失真时,复杂的运动特性同样削弱了人眼对编码失真的感知,从而得到较大的 V_{MOS}. 因此,空间复杂度和时间复杂度成为预测视频编码失真的重要因素.

2 空间复杂度和时间复杂度的预测

2.1 空间复杂度的预测

空间复杂度可以由变换块像素值的标准差来反映. 显然,标准差越大,空间复杂度就越大;标准差越小,则空间复杂度越小. 然而,比特流层模型无法获取像素域信息,因而无法直接得到视频帧的空间复杂度. 假设视频信号的 DCT 系数独立同分布且近似服从双边拉普拉斯分布[18],则当量化参数固定时,比特率与像素值区域标准差近似服从线性关系[8]

$$R_p = a\sigma_p + b \tag{2}$$

式中:R_p 为编码对应区域每个像素的平均比特率;σ_p 为区域标准差;a、b 为斜率与截距. 本文选取具有不同内容特性的测试序列:Container、Football、Foreman、Mobile、Mother&Daughter、News、Silent 和 Soccer,取各视频序列各 200 帧,按帧内编码模式固定 Q_s 进行编码,得到每帧编码的比特率和每帧内所有变换块内像素标准差的平均值,如图 2 所示. 图 2 中各测试点近似服从式(2)所示的线性关系. 从图 2 中还可以发现,Q_s 不同,相应的斜率与截距也各不相同. 将式(2)中自变量与因变量互换得

$$\sigma_S = f(Q_s)R + g(Q_s) \tag{3}$$

式中:σ_S 为视频帧的空间复杂度;R 为编码整个视频帧的比特率;$f(Q_s)$、$g(Q_s)$ 分别为斜率拟合函数和截距拟合函数. 显然 $f(Q_s)$ 和 $g(Q_s)$ 也分别是随 Q_s 变化的函数,如图 3 所示. 从图 3 中可看出 $f(Q_s)$ 与 Q_s 近似成线性关系

$$f(Q_s) = a_1 Q_s + b_1 \tag{4}$$

而 $g(Q_s)$ 与 Q_s 近似符合对数函数

$$g(Q_s) = a_2' \ln(Q_s) + b_2' \tag{5}$$

式中:a_2'、b_2' 为中间过渡常数. 将式(1)分别代入式(4)与式(5)得

$$f(Q_p) = a_1 2^{(Q_p-4)/6} + b_1 \tag{6}$$

$$g(Q_p) = a_2 Q_p + b_2 \tag{7}$$

再将式(6)、式(7)代入式(3),可得视频帧的空间复杂度为

$$\sigma_S = (a_1 2^{(Q_p-4)/6} + b_1)R + (a_2 Q_p + b_2) \tag{8}$$

式中:a_1、b_1、a_2 及 b_2 为待定常数,由最小二乘拟合得到.

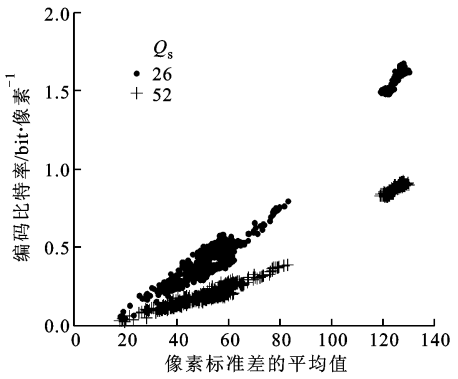
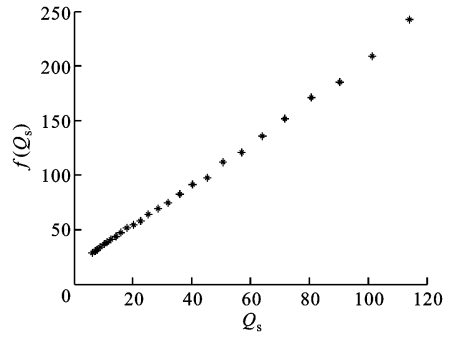
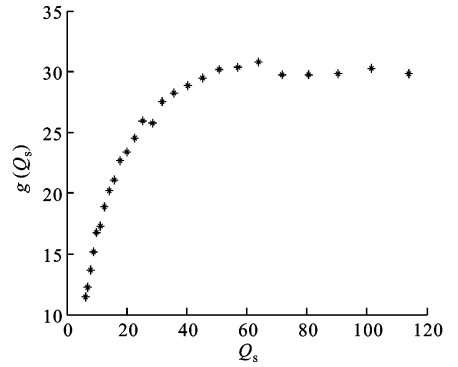


图 2 像素区域标准差和平均编码比特率的关系



(a)式(3)斜率项的拟合结果



(b)式(3)截距项的拟合结果

图 3 对图 2 中 Q_s 所对应的测试点按式(3)进行最小二乘拟合的结果

2.2 时间复杂度的预测

帧间预测编码主要利用视频序列相邻帧间的相关性去除时域冗余来达到压缩数据的目的. H.264/AVC 利用已编码视频帧/场和基于块的运动补偿的预测模式进行帧间预测. 分块运动补偿的基本概念是将视频帧分成若干块或宏块,并设法在参考帧中某一给定的搜索范围内根据一定的匹配准则找到每个块或宏块在参考帧中的与当前块最相似的块,即

匹配块. 匹配块与当前块的相对位移即为运动矢量，匹配块与当前块的对应像素点的差值即为残差信号. 于是，视频压缩时只需保存运动矢量和残差信号就可以完全恢复出当前块. 因此，运动矢量作为视频编码的重要参数被放入视频码流中. 另一方面，运动矢量的大小可以有效地反映视频内容的运动特性.

在运动预测中，一个块或宏块的运动矢量由横向分量和纵向分量组成. 于是，运动矢量的大小表示为

$$|\boldsymbol{V}_{\mathrm{m}}|=(|\boldsymbol{V}_{\mathrm{m},x}|^2+|\boldsymbol{V}_{\mathrm{m},y}|^2)^{1/2} \tag{9}$$

式中：$\boldsymbol{V}_{\mathrm{m},x}$ 和 $\boldsymbol{V}_{\mathrm{m},y}$ 分别表示该块或宏块运动矢量的横向分量和纵向分量.

如图 4 所示，视频中的物体运动越剧烈，运动矢量就越大(如 Football). 相反地，物体运动较缓慢的视频序列(如 Container)，其运动矢量相对较小. 于是，在一定程度上，运动矢量可以用来描述反映视频内容运动特性的时间复杂度.

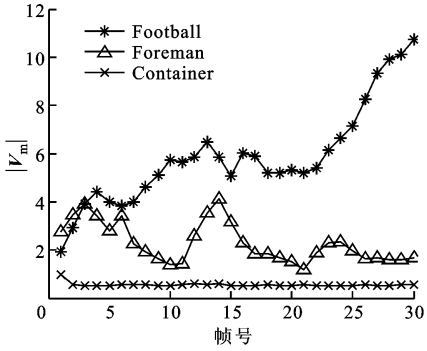


图 4　不同运动特性的视频序列的运动矢量大小(前 30 帧)

一般来说，运动矢量为 0 的块或宏块不会对编码失真产生时域掩盖效应. 因此，设整个视频帧内的非 0 运动矢量个数为 N，本文定义该帧的时间复杂度为

$$\sigma_{\mathrm{T}}=\frac{1}{N}\sum_{n=1}^{N}|\boldsymbol{V}_{\mathrm{m}}|_n \tag{10}$$

3　编码失真评估

如图 1 所示，对于某一个视频序列，当量化参数小于某值时，视频帧质量为最大值；当量化参数不小于该值时，视频帧质量随量化参数增大而呈近似线性关系下降. 这种规律可描述为

$$Q_{\mathrm{F},n}=\begin{cases}Q_{\mathrm{m}}, & 0\leqslant Q_{\mathrm{p},n}<T_n\\ Q_{\mathrm{m}}-a_5(Q_{\mathrm{p},n}-T_n), & T_n\leqslant Q_{\mathrm{p},n}\leqslant 51\end{cases} \tag{11}$$

式中：$Q_{\mathrm{F},n}$ 为第 n 帧质量；$Q_{\mathrm{p},n}$ 为编码第 n 帧的量化参数；Q_{m} 为视频帧质量的最大值；T_n 为分界点对应的量化参数值；a_5 为待定常数，由参数拟合实验得到.

然而，不同内容特性的视频序列对应的式(11)中的分界点不同. 对于空间复杂度或时间复杂度低的视频序列，其 V_{MOS} 下降时的 Q_{p} 较小，而随着空间复杂度或时间复杂度的增大，其对应的视频序列 V_{MOS} 下降时的 Q_{p} 越来越大. 设 T_0 为空间复杂度或时间复杂度较低的视频帧 V_{MOS} 下降时的 Q_{p}，则视频内容对人眼的掩盖效应可表示为

$$T_n=T_0\left(1+\left(\frac{\sigma_{\mathrm{S},n}}{a_3}\right)^{b_3}\right)\left(1+\left(\frac{\sigma_{\mathrm{T},n}}{a_4}\right)^{b_4}\right) \tag{12}$$

式中：T_0、a_3、b_3、a_4 及 b_4 为待定常数，由参数拟合实验得到；$\sigma_{\mathrm{S},n}$ 和 $\sigma_{\mathrm{T},n}$ 分别为第 n 帧的空间复杂度和时间复杂度. T_0、a_3、b_3、a_4 及 b_4 的拟合方法如下：将 Container、Football、Foreman、Mobile、Mother & Daughter、News、Silent 和 Soccer 等 8 个测试序列分别对各序列的 200 个视频帧进行固定 Q_{p} 的帧内编码模式和帧间编码模式编码，然后由主观测试实验得到各视频帧对应的 T_n，由式(8)和式(10)分别得到各视频帧 $\sigma_{\mathrm{S},n}$ 和 $\sigma_{\mathrm{T},n}$，最后由遍历搜索算法得到最小二乘意义下的各参数的最优值. 然而，对于采用帧内编码模式的视频帧(如 I 帧)，其码流载荷中无运动矢量信息，所以无法获得该帧的时间复杂度；对于采用帧间编码模式的视频帧(如 P 帧)，则无法获取空间复杂度. 通常情况下，视频序列相邻帧的相关性很强，空间复杂度和时间复杂度在相邻帧之间的变化不大，所以 I 帧的 $\sigma_{\mathrm{T},n}$ 由其临近的 P 帧的时间复杂度近似表示，而 P 帧的 $\sigma_{\mathrm{S},n}$ 则由其临近的 I 帧的空间复杂度近似表示.

得到视频序列各视频帧质量后，取各视频帧质量的平均值作为该视频序列的质量

$$Q=\frac{1}{N}\sum_{n=1}^{N}Q_{\mathrm{F},n} \tag{13}$$

至此，本文建立了一个网络视频编码失真评估的比特流层模型，如图 5 所示. 该模型通过对网络视频流的解析得到比特率、量化参数以及运动矢量，然后通过量化参数和比特率得到视频的空间复杂度，通过运动矢量得到视频的时间复杂度，最后利用量化参数与编码失真的关系以及视频内容对人眼感知的空间掩盖效应和时间掩盖效应得到视频质量.

4　实验结果

本文实验所采用的编解码器为 JM10.2[19] 软件

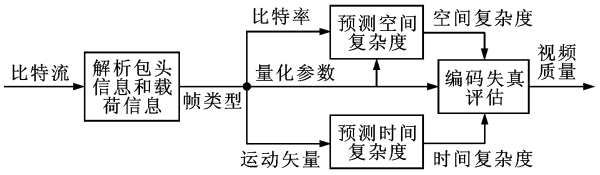


图 5 网络视频编码失真评估的比特流层模型

平台. 实验所用测试序列及其内容特性如表 1 所示. 视频序列格式为 CIF(common intermediate format),帧率设置为 25,I 帧间隔设置为 25,编码模式为“IPPP”. 编码比特率分别设置为 64、96、128、192 和 256 kb/s. 使用主观质量分数与本模型所得分数进行比较. 主观测试实验由 25 名非专业人员完成,相关测试条件参照文献[13]中的标准执行.

表 1 测试序列及其内容特性

视频序列	序列内容特性
Container	物体运动缓慢,很多细微运动
Football	剧烈运动,相机平动和转动
Foreman	运动方向复杂,场景转换
Mobile	空间纹理复杂
Mother&Daughter	运动缓慢,背景静止
News	前景运动缓慢,背景运动剧烈
Silent	运动物体单一,幅度较大
Soccer	物体和相机运动都剧烈

表 2 所示为实验参数,该实验所得结果适用于使用 JM10.2 软件平台生成的 H.264/AVC 视频序列,若用于由其他编码器生成的视频序列,实验参数需要调整.

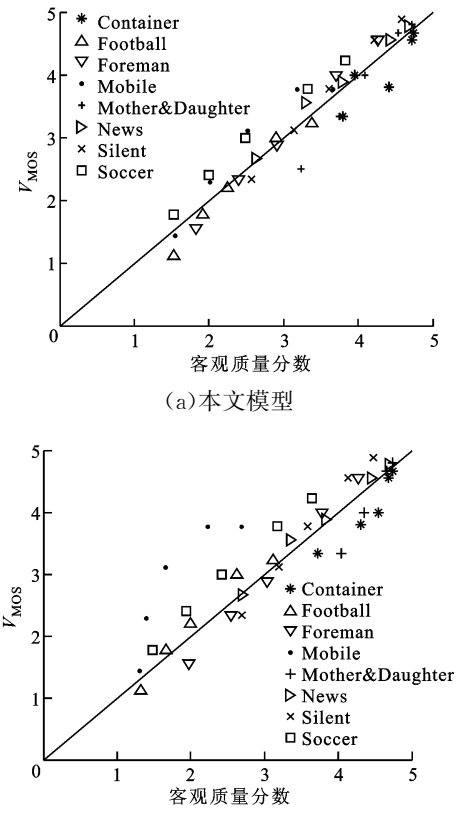
表 2 参数列表

a_1	b_1	a_2	b_2	a_3
1.91	15.92	0.79	−1.83	200
b_3	a_4	b_4	a_5	T_0	Q_m
0.55	45	1.70	0.23	20	4.75

表 3 给出了本文模型与 NR QANV-PA 模型[8]的性能参数比较. 本文使用主观质量分数和客观质量分数的皮尔森相关系数 V_{PCC}(pearson correlation coefficient,PCC)和均方根误差 V_{RMSE}(root-mean-squared error,RMSE)评价模型的性能. 从表 3 中可以看出,对于不同内容特性的视频序列,2 种模型都具有较高的 V_{PCC},说明 2 种模型对各种内容特性的视频都具有较好的预测单调性,但是对于大部分视频序列,由本文模型所得的 V_{RMSE} 比 NR QANV-PA 模型所得的 V_{RMSE} 小,说明本文模型具有更高的预测精度. 表 3 末项给出了 2 种模型的整体性能参数,与 NR QANV-PA 模型相比,本文模型所得的 V_{PCC} 提高了 0.016 0,同时 V_{RMSE} 下降了 0.079 7. 这说明在统计意义上,本文模型比 NR QANV-PA 模型具有更好的预测单调性和更高的预测精度. 图 6 给出了由 2 种模型所得的客观质量分数与主观质量分数的散点图,通过比较图 6a、图 6b,可以直观地验证上述结论.

表 3 2 种模型的性能参数比较

测试序列	本文模型		NR QANV-PA 模型	
	V_{PCC}	V_{RMSE}	V_{PCC}	V_{RMSE}
Container	0.872 8	0.351 3	0.938 8	0.384 4
Football	0.984 6	0.212 4	0.980 2	0.226 2
Foreman	0.998 9	0.222 4	0.999 0	0.274 1
Mobile	0.960 3	0.406 3	0.883 3	1.145 1
Mother&Daughter	0.996 8	0.383 3	0.999 1	0.600 0
News	0.995 0	0.147 3	0.995 2	0.116 4
Silent	0.998 1	0.243 4	0.997 2	0.324 3
Soccer	0.996 1	0.411 2	0.997 6	0.516 2
整体性能评分	0.964 4	0.2772	0.948 4	0.356 9



(a)本文模型

(b)NR QANV-PA 模型

图 6 2 种模型所得主客观质量分数散点图

5 结 论

本文提出了一种适用于 H.264/AVC 编码标准

的视频质量评估的比特流层模型. 该模型不仅分析了量化参数对编码失真的影响，同时考虑到视频内容对编码失真的空域掩盖效应和时域掩盖效应. 通过解析视频载荷信息，得到量化参数、比特率以及运动矢量等信息，并利用上述信息预测得到视频的空间复杂度和时间复杂度，进而建立起评估 H.264/AVC 网络视频编码失真的模型. 实验结果表明，使用该评估模型所得客观质量分数与主观质量分数具有很好的一致性.

参考文献：

[1] 李鹏，常义林，冯妮娜，等. 用于视频传输的自适应分级正交幅度调制不等保护方法[J]. 西安交通大学学报，2011，45(4)：72-77.
LI Peng, CHANG Yilin, FENG Nina. Unequal error protection method for video transmission using adaptive hierarchical quadrature amplitude modulation [J]. Journal of Xi'an Jiaotong University, 2011, 45(4): 72-77.

[2] Telecommunication Standardization Sector. ITU-T Recommendation G. 1070 opinion model for video-telephony applications [S]. Geneva, Switzerland: Telecommunication Standardization Sector, 2007.

[3] JOSKOWICZ J, LOPEZ-ARDAO J C. Enhancements to the opinion model for video-telephony application [C]// Proceedings of the 5th International Latin American Networking Conference. New York, USA: ACM, 2009: 87-94.

[4] MOHAMED S, RUBINO G. A study of real-time packet video quality using random neural networks [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(12): 1071-1083.

[5] LIAO Ning, CHEN Zhibo. A packet-layer video quality assessment model based on spatiotemporal complexity estimation [EB/OL]. (2010-08-04)[2010-11-12]. http://dx.doi.org/10.1117/12.863473.

[6] GARCIA M N, RAAKE A. Parametric packet-layer video quality model for IPTV [C]// Proceedings of International Conference on Information Sciences Signal Processing and their Applications. Kuala Lumpur, Malaysia: ISSPA, 2010: 349-352.

[7] YAMAGISHI K, HAYASHI T. Parametric packet-layer model for monitoring video quality of IPTV services [C]// Proceedings of IEEE International Conference on Communications. Piscataway, NJ, USA: IEEE, 2008: 110-114.

[8] YANG Fuzheng, WAN Shuai, XIE Qingpeng, et al. No-reference quality assessment for networked video via primary analysis of bit-stream [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(11): 1544-1554.

[9] VERSCHEURE O, FROSSARD P, HAMDI M. User-oriented QoS analysis in MPEG-2 video delivery [J]. Real-Time Imaging, 1999, 5(5): 305-314.

[10] BRANDAO T, QUELUZ M P. No-reference image quality assessment based on DCT domain statistics [J]. Signal Processing, 2008, 88(4): 822-833.

[11] 卢刘明，陆肖元. 基于网络丢包的网络视频质量评估[J]. 中国图象图形学报，2009，14(1)：52-58.
LU Liuming, LU Xiaoyuan. Quality evaluation of video over a packet network based on packet loss [J]. Journal of Image and Graphics, 2009, 14(1): 52-58.

[12] 凌云，夏军，屠彦，等. 视觉感兴趣区的提取及其在视频图像质量评估中的应用[J]. 东南大学学报，2009，39(4)：753-757.
LING Yun, XIA Jun, TU Yan, et al. Detection of region of interest and its application in video image quality assessment [J]. Journal of Southeast University, 2009, 39(4): 753-757.

[13] VQEG. Hybrid perceptual/bitstream group TEST PLAN 2.2 [EB/OL]. (2011-01-31) [2011-02-24]. http://www.its.bldrdoc.gov/media/8244/vqeg_hybirl_testplan_2—_2.doc.

[14] DAVIS A G, BAYART D, HANDS D S. Hybrid no-reference video quality prediction [C]// Proceedings of The International Symposium on Broadband Multimedia Systems and Broadcasting. Piscataway, NJ, USA: IEEE, 2009: 1-6.

[15] MPEG. Coding of audio-visual objects: Part 2 visual [S]. Geneva, Switzerland: ISO/IEC JTC1/SC29 WG11, 2003.

[16] Telecommunication Standardization Sector. ITU-T Recommendation H.264 advanced video coding for generic audiovisual services [S]. Geneva, Switzerland: Telecommunication Standardization Sector, 2009.

[17] WINKLER S. Issues in vision modeling for perceptual video quality assessment [J]. Signal Processing, 1999, 78(2): 231-252.

[18] BERGER T. Rate distortion theory [M]. Upper Saddle River, NJ, USA: Prentice-Hall, 1971: 94-95.

[19] Heinrich Hertz Institute. VCEG reference software version JM 10.2. [CP/OL]. (2011-01-06) [2011-02-20]. http://iphome.hhi.de/suehring/tml/download/old_jm/jm10.2.zip.

（编辑　刘杨）