

基于线性回归分析的部分参考 视频质量评估方法*

张兆林, 史浩山, 万 帅
(西北工业大学 电子信息学院, 西安 710072)

摘 要: 提出了一种基于线性回归分析的部分参考视频序列质量评估算法。该算法主要以原始视频的帧间差异作为参考信息, 通过分析编码前后视频时域活动性的变化和单帧比特数的变化关系来客观评价视频序列的编码质量。该方法只需少量原始参考信息, 算法简单。通过对标准视频序列的仿真实验, 该算法可有效评价不同视频的编码质量, 使用该质量评估方法测得的失真视频客观质量评分与其主观质量评分有较好的一致性。

关键词: 数字视频; 客观质量评估; 部分参考; 线性回归分析

中图分类号: TP301.6 **文献标志码:** A **文章编号:** 1001-3695(2012)04-1581-03

doi:10.3969/j.issn.1001-3695.2012.04.107

Reduced-reference video quality assessment method based on linear regression analysis

ZHANG Zhao-lin, SHI Hao-shan, WAN Shuai

(School of Electronics & Information, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: This paper proposed a reduced-reference (RR) video quality assessment method based on linear regression analysis. The proposed method utilized the difference between successive frames of the original video as the reference. It estimated the video quality on the basis of the change in the temporal domain and the difference of bits for coding each frame between the original video and the received video. The proposed method is computationally efficient. Experimental results show that the quality scores computed by the proposed algorithm are well correlated with the mean opinion scores associated to the subjective assessment.

Key words: digital video; objective quality assessment(OQA); reduced-reference; linear regression analysis

0 引言

视频质量评估在数字电视、视频点播(VOD)、视频会议、IPTV 和可视电话等多媒体通信系统中发挥着重要的作用。通常人们将视频质量评估的方法分为两大类: 主观质量评估(subjective quality assessment, SQA)和客观质量评估(OQA)。主观质量评估可以得到平均判分 MOS, 但必须提供严格的测试环境, 考虑大量的影响因素和可能性, 因而实现起来步骤复杂、代价昂贵、实时性差。

对于一个视频质量客观评估方法来说, 根据失真视频与其相应的原始参考视频的比较程度, 可分为全参考模型(full-reference, FR)、部分参考模型(RR)和无参考模型(no-reference, NR)。在多媒体通信系统中, 接收端往往无法获得原始视频。虽然 NR 模型在不需要任何原始视频信息的情况下, 可以方便地应用于移动终端、TV 机顶盒、视频电话、远程实时监控等, 但在精确性上要次于 FR 和 RR 模型。所以, 以最小的传输代价来传输部分参考信息的 RR 模型具有较高的应用价值。

1 国内外研究现状

从现有的国内外研究状况来看, RR 模型主要研究途径是在原始图像与失真图像之间统计和分析某些图像特征的改变来进行质量评估, 也就是说如何有效地提取能表征图像间变化的特征参量是 RR 模型研究的关键。

在以往的研究成果中, 有的是通过提取图像的谐波强度信息, 如在文献[1]中就是通过对图像空域梯度谐波的分析来评估视频质量; 有的是在小波域来建立统计模型, 如文献[2]; 还有利用植入数字水印技术作为部分参考信息^[3]。在文献[4]中则是通过对图像的兴趣点和显著物体的分析来评估图像质量。此外还有通过对已有的 FR 模型进行改进使其转变为 RR 模型, 如文献[5,6]就是利用原始视频的部分信息将全参考的 SSIM 方法变为 RR 模型。这些方法虽然在测量视频的编码失真方面取得了很好的效果, 但在部分参考信息的生成、提取以及分析方面较为复杂, 编/解码端或收/发端的相互协调工作要求较高。

本文提出了一种部分参考的视频序列质量客观评价方法。

收稿日期: 2011-07-06; **修回日期:** 2011-09-02 **基金项目:** 国家自然科学基金资助项目(F010204)

作者简介: 张兆林(1978-), 男, 河南濮阳人, 讲师, 博士研究生, 主要研究方向为视频质量评估、误码控制(zl84@sina.com); 史浩山(1946-), 男, 河南开封人, 教授, 博导, 主要研究方向为移动通信技术、数据通信与计算机网络等; 万帅(1979-), 女, 河南洛阳人, 副教授, 博士, 主要研究方向为视频质量评估、率失真优化和误码控制等。

该方法只需获得原始视频的帧间差异作为参考信息,通过分析视频序列编码后的单帧平均最大比特数的变化和编码前后平均帧间差异的变化关系,就可以客观评价视频序列的质量。本方法运算简单、参考信息数据量小、获取方便、评价灵敏度和应用灵活度高。

2 部分参考视频序列质量评估模型

目前广泛应用的视频压缩编码标准 MPEG-1/2/4, ITU-T H. 261, H. 263 和 H. 264 等在帧间编码时都采用了运动预测与运动补偿技术来压缩时域冗余,从本质上来讲是通过有限的比特数来表示视频流在时域上的差异。因而一个编码视频序列随着压缩比的提高,不但单帧的比特数随之减少,其前后帧之间的差异也会显著降低。因此可以通过对视频序列编码前后的比特数变化和帧间差异变化的关系来客观评价视频序列的质量。

2.1 编码帧与相邻帧的帧间差异

常用的评价两幅图像差异的方法有绝对误差和 (sum of absolute difference, SAD)、均方误差 (mean squared error, MSE) 和峰值信噪比 (peak signal-to-noise ratio, PSNR), 在本文中采用峰值信噪比的方法来表示编码帧与相邻帧的帧间差异。由于未编码的数字视频大多采用 YUV 格式,而人类视觉系统对于亮度分量 Y 的敏感程度要大于色度分量 U 和 V,因而 4:2:0 的采样率被广泛应用于数字视频领域。为了减少运算量,只计算两帧单个像素的亮度分量的差异。MSE 的表达式为

$$MSE = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (L_n(i, j) - L_{n-1}(i, j))^2 \quad (1)$$

其中: $L_n(i, j)$ 和 $L_{n-1}(i, j)$ 分别表示当前帧和前一帧图像在 (i, j) 位置处的像素亮度值, M, N 分别为图像的行、列像素数。因此两帧之间的 PSNR 表达式可由式(2)来表示:

$$PSNR = 10 \lg \frac{255 \times 255}{MSE} \quad (2)$$

可以看出,两帧的 PSNR 越大,表明视频序列的纹理复杂度越低或运动区域变化范围和强度越小。随视频序列压缩比的提高,图像的纹理和边缘细节丢失也就越严重,导致编码视频的前后帧逐渐趋同,使其帧间 PSNR 值也逐步增大。如图 1 所示,随着码率的降低,flower 序列的帧间亮度 PSNR 值明显增大。随后将原始视频每一帧与其前一帧的帧间差异 PSNR_s 作为部分参考信息伴随编码视频流一同传输到接收端。

2.2 单帧编码比特数与帧间 PSNR 的回归分析

目前广泛使用的视频压缩编码方法都采用了运动预测和运动补偿来消除时域冗余,再加上变换编码和量化的非线性,编码后的当前帧的比特数与其前一帧之间的差异 PSNR 无法用确定的函数关系来表示。但是对于同一场景的视频序列来说,单帧图像的平均亮度、色度、基本结构等存在较大相关性,因此把它的每一帧间编码帧比特数作为随机变量,相对于帧间编码帧与前一帧之间的差异 PSNR 的每一个确定值都有其分布,并且数学期望也存在,因而帧间编码帧比特数的数学期望可以表示成帧间编码帧与前一帧之间差异 PSNR 的函数,可以用线性回归方程来确定两者之间的关系,如式(3)所示:

$$\hat{\text{bit}} = \hat{\alpha} - \hat{\beta} \times PSNR \quad (3)$$

其中: $\hat{\alpha}$ 和 $\hat{\beta}$ 是一元线性回归模型未知参数 (α, β) 的估计量,

PSNR 是视频序列的帧间编码帧与前一帧亮度分量的峰值信噪比。这里需要说明的是,进行线性回归分析的目的是得到每一帧间编码帧编码比特数的数学期望随帧间亮度分量 PSNR 的变化趋势,而不是通过帧间亮度分量 PSNR 值来估计某一帧的编码比特数,因此对于有些纹理复杂的视频序列,即使回归的显著水平不高,也不影响预测。如果在编码序列中同时存在 P 帧和 B 帧时,应独立选取 P 帧或 B 帧进行回归分析,并在后续统计窗口中持续使用。

图 2 给出了 foreman 序列前 100 帧在不同比特率时帧间亮度差异 PSNR 值和 P 帧编码比特数的回归直线的图形。由于码率的不断降低,编码后帧间差异的细节丢失也就越严重。可以看出随着比特率的降低,回归直线斜率的绝对值显著减小,说明随码率的降低,描述单位帧间 PSNR 变化所需的比特数也随之减少。

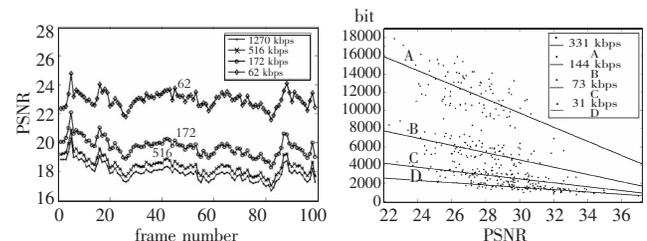


图 1 flower 序列帧间亮度差异 PSNR 值随码率变化示意图

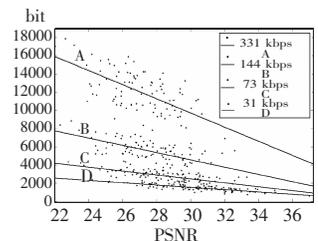


图 2 不同比特率时 foreman 序列帧间亮度差异 PSNR 值和 P 帧编码比特数的回归直线

2.3 编码前后单帧平均比特数的变化

由式(3)可以看出,当 PSNR = 0 时, bit_{\max} 值最大,即回归直线在纵轴上的截距。虽然在实际中一般不会出现帧间亮度分量 PSNR 为零的情况,它表示的意义是两帧亮度分量差异最大时,这段视频序列编码后所能达到的最大平均比特数,也就是前后两帧完全不相关时的编码比特数。由于横轴的帧间 PSNR 是以 dB 的形式表示,为使纵轴与横轴在数量级上相匹配,将纵轴以 10 为底取对数进行放缩,再将未压缩时原始数字视频单帧亮度分量比特数以 10 为底取对数,因此编码前后平均比特数的变化 bit_c 可以用式(4)来表示。

$$\text{bit}_c = \lg \text{bit}_s - \lg \text{bit}_{\max} \quad (4)$$

其中: bit_s 为未压缩时原始数字视频单帧亮度分量的比特数,可通过已编码的视频序列的分辨率和图像精度得到,如一帧分辨率为 352×288 的 CIF 彩色图像的亮度分量用 8 bit 来表示,则 $\lg \text{bit}_s = \lg 352 \times 288 \times 8 \approx 5.91$ 。由于 4:2:0 的采样率广泛应用于视频会议、数字电视和 DVD 存储等消费类应用,使用这种采样率时,编码后的帧间编码帧的单帧比特数远小于编码前单帧比特数。因此这里只采用了视频序列单帧亮度分量未编码时的比特数来计算单帧编码前后平均比特数的变化。

2.4 计算编码前后单帧平均比特数变化与平均帧间差异变化

在接收端,采用与式(2)相同的方法,计算出解码后视频序列的帧间差异,并求平均;再将接收到的原始视频的帧间差异求平均;然后将两者相减,得到编码前后视频序列的平均帧间差异变化,如式(5)所示。

$$\Delta PSNR_c = \Delta PSNR_d - \Delta PSNR_s \quad (5)$$

其中: $\Delta PSNR_d$ 是解码后视频流中帧间编码帧与其前一帧的平均帧间差异, $\Delta PSNR_s$ 是这些帧间编码帧在原始视频流中与前

一帧的平均帧间差异。因此一段视频序列的质量可以表示为

$$Q = \frac{\text{bit}_c}{\text{APSNR}_c} \quad (6)$$

可见 Q 值越大,表明用来表示编码前后单位帧间差异变化的平均比特数越多,视频序列的质量也就相应越高,反之亦然。由于 Q 的取值在 $(0, \infty)$,将 Q 值作为正切值,求出相应的角度值,这样就得到了最终视频质量的评估值 V_q 。 V_q 的取值在 $(0, 90)$ 之间,值越大表明测试的视频序列质量越高,值越小表明测试的视频序列质量越低。

$$V_q = \text{degree}(\arctan(Q)) \quad (7)$$

3 实验验证及结果分析

为了验证本文所提出的部分参考视频质量评估方法的有效性,选取了五个标准的 CIF 格式测试序列:flower、foreman、akiyo、football、highway,采用 H. 264 压缩编码标准 baseline 编码,使用的编码参考软件为 JM 10.2。每个测试序列只编码前 100 帧,编码帧的类型为 IPPP...,使用固定的量化参数,QP 变化范围为 28~48,变化步长为 2。

从图 3 可以看出,本文提出的视频序列质量评估方法可以随码率的降低,针对不同的视频序列有效地评价视频序列的编码质量。在实际中,纹理复杂度低的视频序列空间冗余度高,可以获得较大的压缩比率,因此对于纹理复杂度较低的 akiyo、highway、foreman 序列,即使在较低的码率时也具有较高的质量;但纹理复杂度低的视频序列编码质量对码率变化较为敏感,随码率的不断降低各种失真效应显著,视频质量下降迅速;而纹理复杂度较高的 flower、football 序列视频质量随码率变化并不显著,这符合人类视觉特性(HVS)中的纹理掩盖效应。从图中还可看出在比特率较低时,各视频序列的质量均加速下降,与人的实际主观感受相一致。

图 4 是归一化后的采用 SSIM、PSNR 方法和本文方法对 flower 序列在不同码率时进行质量评估的评估值。从图中可以看出,随码率的降低,本文方法的视频质量评估值的减小幅度明显大于全参考的视频质量评估方法 SSIM 和 PSNR,因此本文方法对由编码比特率变化引起的视频质量变化具有更高的评估灵敏度。

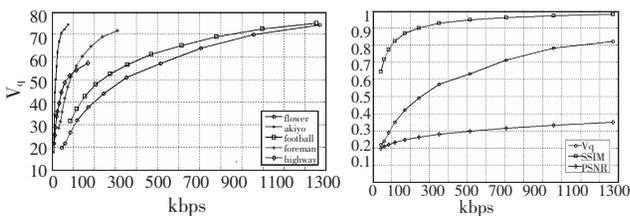


图 3 本文方法对不同视频序列进行质量评估的结果

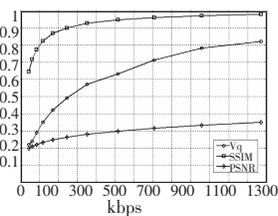


图 4 归一化后不同方法对 flower 序列的质量评估结果

此外,采用以下三个尺度来量化模型的性能^[7]:均方误差(RMSE)、Pearson 相关系数、Spearman 秩相关系数。表 1 给出了归一化后该模型与全参考的 SSIM、PSNR 方法的各性能参数。可以看到,使用本文提出的客观视频质量评估方法得到的客观评分与主观质量评分相比,有最小的 RMSE,反映了本文方法具有很好的预测精度;较大的 Pearson 相关系数和 Spearman 秩相关系数反映了本文方法在预测单调性方面优于 PSNR,而略差于 SSIM。

表 1 几种视频质量评估方法的性能比较

方法	RMSE	Pearson	Spearman
SSIM	0.27	0.84	0.92
PSNR	0.32	0.64	0.68
V_q	0.17	0.80	0.81

图 5 给出了本文评估方法、SSIM、PSNR 的评分分别与主观评分的对应分布。图 5 中的每个数据点对应一个受测失真视频序列,纵坐标为采用客观质量评估方法测得的客观评分,横坐标表示视频序列的主观质量评分。从数据点的分布情况可以看出,本文提出的质量评估方法所测得的客观质量与相应的主观质量一致性略差于全参考的 SSIM 方法,但优于 PSNR。

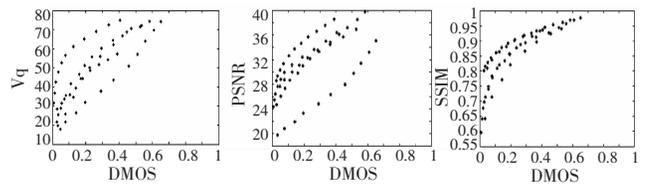


图 5 不同客观评估方法的评分与主观评分的对应分布

4 结束语

本文提出的部分参考视频质量评估方法不需要在发送端和接收端对单帧视频进行复杂的图像分析或特征提取,方法简单、参考信息数据量小、获取方便、评价的灵敏度和应用的灵活度都较高。由于本方法应用了统计学原理,因此实验研究表明,在一个场景内提取的帧间编码帧数及样本数应大于 30 帧,在实际应用中这需要与视频质量评价的实时性相权衡,并且场景切换后应重新开始统计,因而本方法更适用于场景变换不大的视频会议、视频监控以及视频通话等实时多媒体通信领域。

参考文献:

- [1] GUNAWAN I P, GHANBARI M. Reduced-reference video quality assessment using discriminative local harmonic strength with motion consideration[J]. IEEE Trans on Circuits and Systems for Video Technology, 2008, 18(1): 71-82.
- [2] MASRY M, HEMAMI S S, SERMADEVI Y. A scalable wavelet-based video distortion metric and applications[J]. IEEE Trans on Circuits and Systems for Video Technology, 2006, 16(2): 260-273.
- [3] KAYVANRAD M H, SODAGARI S, AVANAKI A N, et al. Reduced reference watermark-based image transmission quality metric [C]// Proc of the 3rd International Symposium on Communications, Control, and Signal Processing. 2008: 526-531.
- [4] NAUGE M, LARABI M C, FERNANDEZ C. A reduced-reference metric based on the interest points in color images [C]// Proc of the 28th Picture Coding Symposium. 2010: 610-613.
- [5] REHMAN A, WANG Zhou. Reduced - reference SSIM estimation [C]// Proc of the 17th International Conference on Image Processing. 2010: 289-292.
- [6] TAGLIASACCHI M, VALENZISE G, NACCARI M, et al. A reduced-reference structural similarity approximation for videos corrupted by channel errors[J]. Multimedia Tools and Applications, 2010, 48(3): 471-492.
- [7] VQEG draft final report from the video quality experts group on the validation of objective models of video quality assessment [EB/OL]. (2003-12-25). <http://www.its.bldrdoc.gov/vqeg>.