

文章编号:1672-058X(2012)09-0044-07

基于小波变换的无参考视频质量评价*

姚 杰¹, 谭建明¹, 陈 婧¹, 黄剑锋²

(1. 重庆通信学院 军事网格实验室, 重庆 400035; 2. 解放军理工大学 通信工程学院, 南京 210007)

摘 要:无参考视频质量评价算法均是针对特定的失真类型, 但一个失真视频中往往有多种失真类型存在, 仅对某种特定的失真进行评价无法全面反映出视频质量; 鉴于此, 基于小波变换提出了一种能够对多种失真类型进行综合评价的图像质量评价算法; 最后, 通过亮度和运动估计帧加权的方式将上述方法扩展到视频质量评价中; 实验证明了本方法具有较高的评价准确性。

关键词:视频质量评价; 支持向量机; 小波变换

中图分类号:TP391

文献标志码:A

通常来说, 人是视频应用的最终消费者, 因此评价视频质量最准确的办法是主观质量评价。但是根据 ITU-T BT500. 11^[1]所规定的方法进行主观质量评价, 不但要花费大量的人力和物力, 而且无法嵌入到视频应用系统中自动进行质量评价, 可用性很小。因此, 如何获得高效的客观视频质量评价算法已经成为目前亟待解决的问题。无参考算法视频质量评价算法不需要参考视频, 根据人类视觉特性直接对失真视频进行质量评价, 应用范围极广, 但是目前的无参考算法均是针对特定的失真类型而设计的。针对 JPEG 压缩引起的图像块效应, 文献[2]通过计算图像傅里叶域的边界强度来量化图像的块效应失真程度。针对 JPEG2000 压缩引起的振铃效应, 通常通过边界探测的办法来度量块边界的扩散程度, 而这种扩散程度在很大程度上能反映图像质量^[3]。针对图像的模糊失真, 文献[4]度量了模糊可检测的概率, 文献[5]通过阈值的方法确定了刚可觉察的模糊与图像质量之间的关系, 文献[6]利用基于变化率的奇异值分解方法来度量噪声鲁棒的模糊失真。

可以看出, 目前无参考图像质量评价算法都假设失真类型是已知的, 而且仅仅适用于一种特定的失真类型。在实际应用中, 许多情况下并不能确定图像中存在的失真类型, 并且一幅失真图像可能经历了多种失真过程, 而传统算法无法对存在多种失真的图像进行有效的质量评价。基于上述问题, 提出了一种能够对多种失真进行有效评价的无参考图像质量评价算法, 然后通过亮度和运动估计帧加权的方法将图像质量评价扩展到视频质量评价中。

1 失真图像的小波域特性

小波变换是进行图像分析的一种常用方法, 与传统的傅里叶变换相比, 小波变换不但能反映出图像中存在何种频率, 还能确定每种频率在图像中的空间位置。将图像通过合适的小波基变换到小波域可以将图像中低频的均值部分和各个子带高频的细节部分区分开, 便于对图像进行分析。

收稿日期:2012-03-01; 修回日期:2012-03-20.

* 基金项目: 国家创新基金(11c26215115768); 重庆市重点攻关基金项目(cstc2011ab2064).

作者简介: 姚杰(1986-), 男, 江苏盐城人, 硕士研究生, 从事图像质量评价研究.

Daubechies 9/7 小波基是进行图像压缩和图像分析时常用的小波基。研究表明,使用 Daubechies9/7 小波基对图像进行小波变换后的子带响应服从一定的统计规律。如图 1 所示,图像使用 Daubechies 9/7 小波基以 3 种尺度,每种尺度在 3 个方向进行小波变换后,被分成了 9 个子带和 1 个均值的组合。同时,图 1 还给出了标号为 2 和 3 的两个子带相应的系数柱状图。可以直观看出,系数服从一定的统计规律。

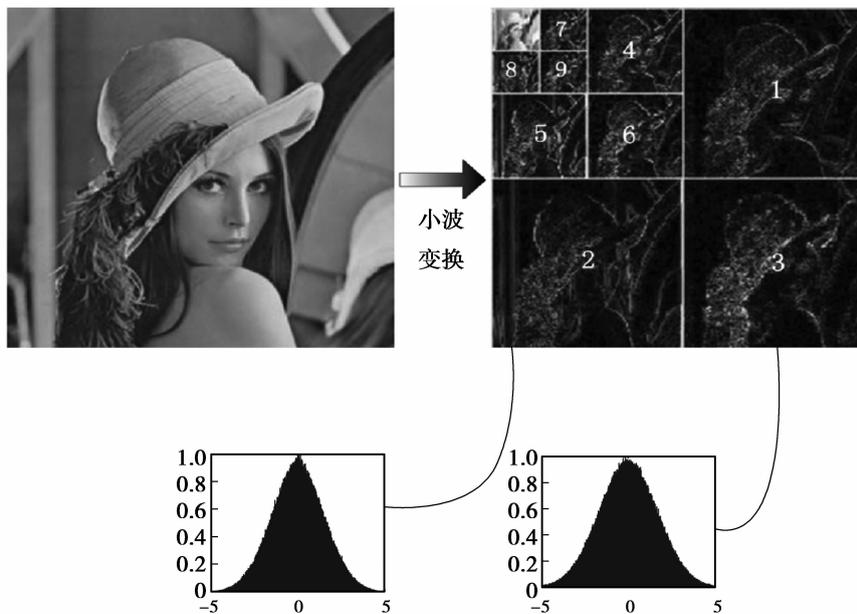


图 1 图像小波变换及其子带系数分布示意图

文献[7]指出,使用 Daubechies9/7 小波基对失真图像进行小波变换后,每个子带的系数都服从广义高斯分布。因此,用相应的广义高斯分布系数就能唯一地刻画该子带的失真特性。广义高斯分布的定义如式(1)所示:

$$f(x; \mu, \sigma^2, \gamma) = a e^{-(b|x-\mu|)^\gamma}$$

$$a = \frac{b\gamma}{2\Gamma(\gamma^{-1})}$$

$$b = \frac{1}{\sigma} \sqrt{\frac{\Gamma(3\gamma^{-1})}{\Gamma(\gamma^{-1})}}$$

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt (x > 0)$$

其中, μ , σ^2 和 γ 分别为广义高斯分布的均值,方差和形状系数。由于小波基是个带通滤波器,故均值 $\mu = 0$ 。因此,对于每个子带的系数分布情况用 σ^2 和 γ 两个参数就可以完全描述。文献[8]给出了一种求 σ^2 和 γ 的简便方法。

文献[7]还指出,失真会对原始图像的子带响应分布产生影响,且不同的失真产生的影响也不同,同一失真类型不同失真程度产生的影响也不同。因此,可以将失真对子带系数产生的影响进行量化,用子带响应分布的系数来描述失真的类型及程度。如图 2 所示,对原始(org)图像采用 JPEG 压缩(jpeg), JPEG2000 压缩(jp2k),白噪声(wn),高斯模糊(blur)和快褪色(ff)5 中失真方式进行失真处理后,同一个子带的响应在不同失真的影响下产生了不同的分布(经过了归一化处理)。基于这一点,提取一个 18 维向量[3 个方向(水平,垂直和对角线) \times 3 种尺度 \times 2 个系数]作为特性向量来描述图像所遭受的失真,该特性向量可以明显地区分不同的失真所产生的影响,且对同一类型的不同程度的失真也能进行一定区分。因此,可以通过该向量来确定图像遭受了何种类型的失真以及遭受该失真的程度如何。

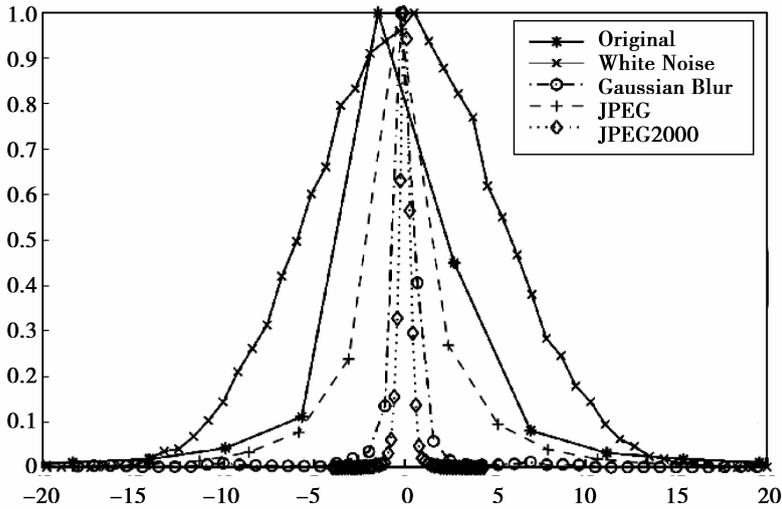


图 2 图像同一子带受不同失真影响分布变化图

2 无参考图像质量评价

2.1 算法总体描述

给定一幅失真图像,首先通过训练后的支持向量机计算出失真类型集中的每种失真类型在当前图像中发生的概率。失真类型集包含 JPEG 压缩, JPEG2000 压缩, 白噪声, 高斯模糊和快褪色 5 中失真类型。这些失真类型是根据 LIVE 测试图片库中的失真类型确定的,在此也将用该测试库中的图片作为训练图片。失真图片中每种失真类型发生的概率记为 $p_i \{i=1,2,3,4,5\}$ 。针对每种失真类型通过训练后的支持向量回归从 18 维的特性向量空间映射到质量分值空间,计算出相应的质量分值,并记为 $q_i \{i=1,2,3,4,5\}$ 。令 Q 表示当前图片的总体质量分值,那么 $Q = \sum p_i \cdot q_i$ 。整个算法流程如图 3 所示。

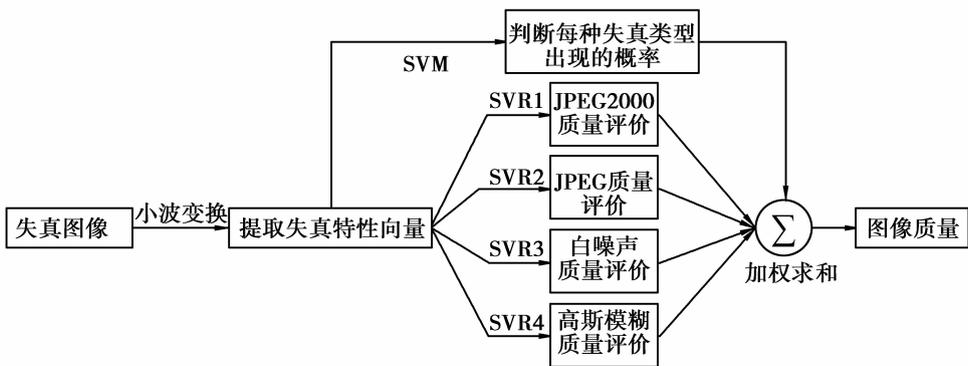


图 3 无参考图像质量评价算法流程图

2.2 计算失真类型概率

由于支持向量机具有良好的高维空间分类特性和收敛性,同时避免了过训练产生的误分类,我们选择具有放射基函数核 RBFK(Radial Basis Function Kernel)的支持向量机 SVM 作为失真类型分类器。需要强调的是,由于一幅图像可能受到多种失真,此处并不是将图像归于一类特定的失真,而是计算出各种失真类型在当前图像中出现的概率。对于 SVM 中需要设定的两个参数(C, γ),在此采用作用在整个训练集上的五交错法来确定。

2.3 映射质量分值

对于给定的属于某种特定失真类型的图像,将其 18 维的失真特性向量映射为一个质量分值。为了实现这种映射,采用支持向量回归的原理,并用支持向量机 ν -SVM 去实现这种支持向量回归。由于不同失真类型的失真特性向量差异较大,如果不加区分地直接进行映射训练,训练后支持线性回归的准确率将会很低。因此,对于 5 种不同失真类型选用 5 个支持向量回归分别进行训练,这样针对每种特定的失真类型及相应的失真程度,映射准确率就会大大提高。

采用 LIVE 图像测试库中所有已经进行分类的失真图片及其相应的由主观质量评价给出质量分数来训练 ν -SVM,使其学习从失真特性空间到质量分数空间的映射关系。经过训练之后,该回归就会对一个确定失真类型和相应失真程度的图像给出相应的质量分值。同样地,在 ν -SVM 中需要确定 3 个参数 (ν, C, γ) ,其中, (C, γ) 仍然采用作用在整个训练集上的五交错方法确定,而 ν 一般取默认值 0.5,并且在试验中也发现,改变 ν 的值对回归准确度影响很小。

3 基于亮度和运动估计的帧加权

视频是由连续的帧序列组成的,每 1 帧又是 1 张静态图片,所以图像质量评价可以作为视频质量评价的基础。但是,由于在视频中帧一般是以每秒 25—30 张的速度连续播放的,人在观看时能明显感觉到视频中物体的运动具有连续性,并能判断出运动的大小和方向。而在静态图片中,物体运动的大小和方向是无法准确获取的,如果简单地将每一帧的质量进行求和平均作为整个视频的质量显然是不合理的。另外,根据人类视觉系统的亮度敏感性原理,视频中黑暗的区域通常不会引起观察者的注意,因此应当赋予较小的权值,而视频中亮度较大的区域容易引起观察者的注意,应当赋予较大的权值。基于上述两点,计算当前帧的平均亮度并利用平均运动向量长度进行运动估计,提出了基于亮度和运动估计的帧加权求和方法。

假设当前帧的样本均值(亮度)为 μ ,令 ω 表示亮度加权系数为,则:

$$\omega = \begin{cases} 0 & \mu \leq 40 \\ (\mu - 40)/10 & 40 < \mu \leq 60 \\ 1 & \mu > 60 \end{cases}$$

在实验中发现,当视频中有非常大的整体运动情况发生时,算法的结果并不稳定。因此,给予那些运动较大的帧较小的权值来保证整个算法的稳定性。令 m 代表当前帧运动向量的平均长度, M 代表当前帧的运动强度, K_m 是反映运动强度的常数,那么:

$$M = \frac{m}{K_m}$$

令 W 表示当前帧的运动加权系数,则:

$$W = \begin{cases} 1 & M \leq 0.8 \\ (1.2 - M)/0.4 & 0.8 < M \leq 1.2 \\ 0 & M > 1.2 \end{cases}$$

对与一个有 F 帧组成的视频序列,令 ω_i 表示第 i 帧的亮度加权系数, W_i 表示第 i 帧的运动加权系数, Q_i 表示通过上述无参考质量评价得到的第 i 帧的图像质量, Q_F 表示该视频序列的质量,则:

$$Q_F = \frac{\sum_{i=1}^F \omega_i W_i Q_i}{\sum_{i=1}^F \omega_i W_i}$$

4 实验与分析

4.1 支持向量机及支持向量回归训练

训练集采用美国德克萨斯大学 LIVE 实验室的图像测试库。库中有 29 幅原始图像,通过 JPEG2000, JPEG, 白噪声, 高斯模糊和快褪色 5 中失真方式分别生成 226, 232, 173, 173 和 173 幅失真图像,并对每幅失真图像进行了主观质量评价,给出了相应的差异主观意见分值 DMOS(Differential Mean Opinion Score)。支持向量机和支持向量回归采用台湾林智仁的 libSVM 软件包实现,版本为 2.91,两者的核函数都选用放射基核函数。

在进行支持向量机训练之前先对所有的失真图像用 Daubechies 9/7 小波基变换到小波域,提取其 18 维的特性向量。这样,对支持向量机来说,形成了一个具有 5 个标签,977 个样本,每个样本具有 18 个属性的训练集。为了便于验证训练后的支持向量机的分类准确性,将训练集分为两部分。每种失真类型的前 80% 的图片组成实际训练集,后 20% 作为测试集,用来对训练好的支持向量机进行分类准确性测试。支持向量机中的参数 C 和 γ 通过在训练集上的五交错法得到的值为 2 和 0.02。经测试,训练后的支持向量机分类准确率为 81.414 9%。

进行支持向量回归训练时,用每幅失真图像的特性向量和相应的 DMOS 分值作为训练集训练支持线性回归。具体训练方法与支持向量机相似,参数 C 和 γ 通过在训练集上的五交错法得到的值为 2.3 和 0.03, ν 的取值为默认的 0.5。如果存在针对某一失真类型性能较好的无参考算法,可以用其替换支持向量回归中针对该类型的质量分值映射,以提高评分的准确性。例如,用文献[1]中的算法替换支持向量回归中针对 JPEG 压缩的分值映射,相信会取得较好的效果。在此侧重于描述整体算法的流程和原理,并没有过多关注质量分值映射可选的替代方法。

4.1 视频质量评价测试

视频质量评价实验采用 LIVE 实验室的视频测试序列库。序列库中包含了 10 个典型的自然场景,每个场景用 MPEG-2 压缩、H.264 压缩、模拟 IP 传输和模拟无线传输 4 种视频失真方式进行失真处理,生成了 15 个失真序列,并进行了主观质量测试,给出了每个失真序列的 DMOS 分值。

本实验中,将 K_m 的值设置为 16,并采用目前被普遍认可的两个指标来度量算法的性能。一个是算法结果和 DMOS 之间的秩相关系数 SROCC(Spearman Rank Order Correlation Coefficient),用来度量算法结果和主观分值之间的单调性;另一个是经过非线性回归处理后的算法结果和 DMOS 之间的线性相关系数 LCC(Linear Correlation Coefficient),用来度量算法的准确性。在进行非线性回归处理时,采用 VQEG 的报告^[9]中提供的方程:

$$Q' = \beta_1 + \frac{\beta_2 - \beta_1}{1 + e^{-\left(\frac{Q - \beta_3}{\beta_4}\right)}} \quad (9)$$

其中 Q 代表算法结果, Q' 代表经过处理后的算法结果。为了获得以上方程中能使算法结果和预测值方差最小的 $(\beta_1, \beta_2, \beta_3, \beta_4)$,采用 Matlab 中的 nlinfit 函数进行非线性最小二乘法优化。为了便于数据收敛,在进行优化之前用线性重调节的方法对算法的结果进行处理。最后,SROCC 和 LCC 在 Q' 和 DMOS 之间进行计算。为了比较算法的性能,同时使用全参考算法峰值信噪比 PSNR 和结构相似度 SSIM^[10]进行同样的视频质量评价测试,用来对比无参考算法与全参考算法之间的性能差异。

4.3 结果与分析

图 5、图 6 和图 7 给出了 PSNR, SSIM 和本算法的测试散点图,图中曲线是散点的最佳拟合曲线。表 1 给出了以上 3 种算法的 SROCC 和 LCC。由散点图可以直观地看出算法在单调性和准确性上都要好于全参考的 PSNR,但与全参考的 SSIM 算法相比还存在差距。通过表 1 的具体数据可以看出,算法的性能远好于 PSNR,并且接近全参考的 SSIM 算法。

需要强调的是,提出的是无参考算法,在无法获得原始视频任何信息的情况下,通过一定的训练获得视频质量评价能力,选用的训练集不同取得的最终评价能力也会不同。因此,若选用涵盖图像失真类型更多,失真图像数量更大的训练集进行训练,相信算法的性能会更加理想。

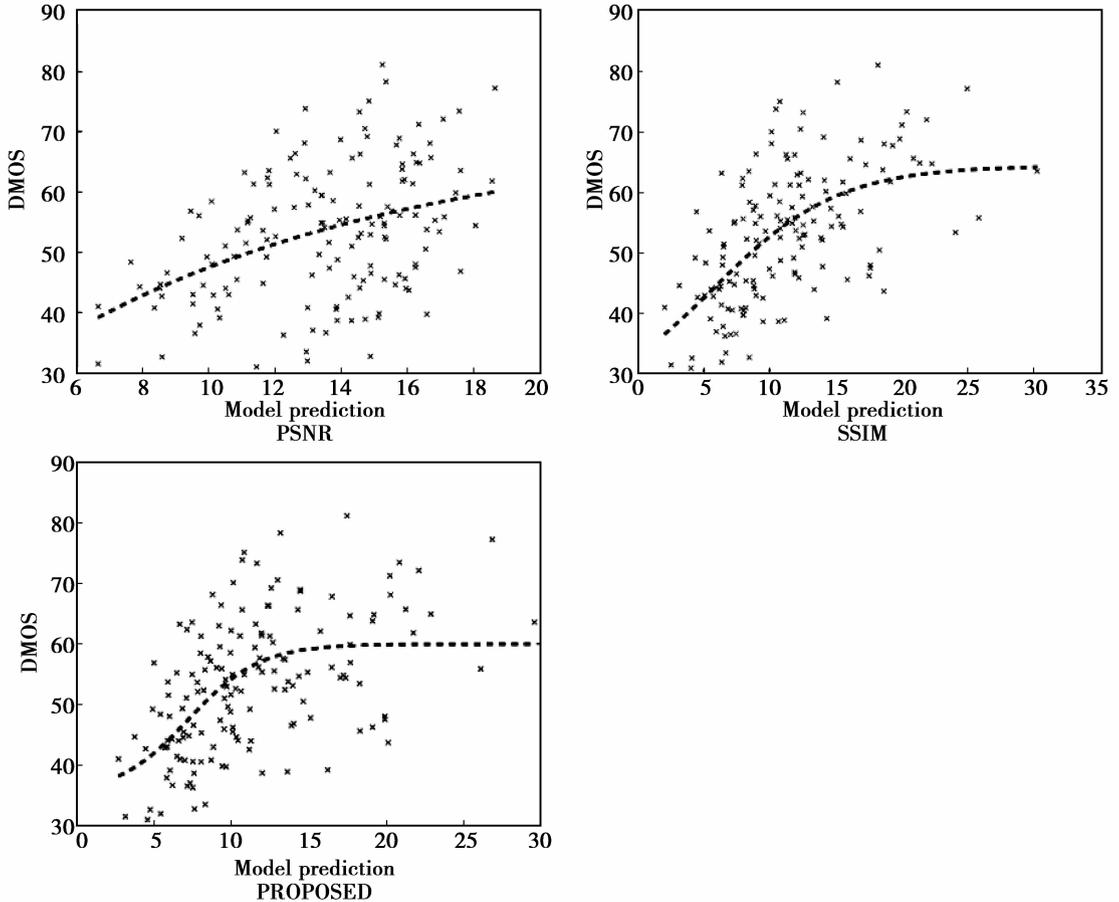


图4 PSNR、SSIM 和此处算法的散点比较图

表1 此处算法与 PSNR 和 SSIM 的性能指标比较

算 法	SROCC	LCC
PSNR(全参考)	0.368 4	0.403 5
SSIM(全参考)	0.525 7	0.544 4
PROPOSED(无参考)	0.476 9	0.508 2

5 结束语

本文的关键在于利用失真图像小波变换后提取的失真特性向量来描述图像中存在的失真类型及相应的失真程度。另外,由于视频的失真和图像的失真有许多相似之处,因此,在充分考虑视频的亮度和运动特性的基础上,通过一定的权值调整,可以将图像质量评价扩展为视频质量评价。提出的无参考算法与性能较好的全参考算法相比虽然还存在一定差距,但无参考算法不需要获得原始视频的任何信息,因此其适用范围要远远大于全参考算法,这对许多实际的视频应用来说是非常关键的。另外,算法有极大的可扩展性,若利用针对特定失真效果显著的无参考算法替代相应的支持向量回归中的质量分值映射,相信算法的性能会进一步提升。今后的工作重点主要有两个方面:一是充分挖掘图像小波域的特性,改进现有的用失真特性向量来描述失真的方

法,或是找到其它的失真描述方法,能够更加准确地刻画失真类型和失真程度;二是进一步研究人类视觉系统对视频感知的其他特性,找到更加合适的加权方法将图像质量评价扩展为视频质量评价。

参考文献:

- [1] ITU-R REC. BT500-11. Methodology for the Subjective Assessment of the Quality Television Picture [S]. The International Telecommunication Union, 2002
- [2] SUTHAHARAN S. No-reference visually signi? cant blocking artifact metric for natural scene images [J]. Signal Processing, 2009, 89(8):1647-1652
- [3] MARZILIANO P, DUFAUX F, et al. Perceptual blur and ringing metrics: Application to JPEG2000 [J]. Signal Processing: Image Communication, 2004, 19(2):163-172
- [4] NARVEKAR N, KARAM L. A No-reference Perceptual Image Sharpness Metric based on a Cumulative Probability of Blur Detection [C] //1st International Workshop on Quality of Multimedia Experience (QoMEX), 2009
- [5] FERZLI R, KARAM L. A No-Reference Objective Image Sharpness Metric Based on the Notion of Just Noticeable Blur (JNB) [J]. IEEE Transactions on Image Processing, 2009, 18(4):717
- [6] ZHU X, MILANFAR P. A no-reference sharpness metric sensitive to blur and noise [C]. //1st International Workshop on Quality of Multimedia Experience (QoMEX), 2009
- [7] MOORTHY A K, BOVIK A C. Statistics of natural image distortions [C] //IEEE International Conf. Acoustics, Speech and Signal Proc, 2010
- [8] SHARIFI K A, Leon-Garcia. Estimation of shape parameter for gen-eralized Gaussiandistributions in subband decompositions of video[J]. IEEE Tran. Circuits Syst. Video Technol, 1995, 5(1):52-56
- [9] Video Quality Expert Group. Report on the validation of objective quality metrics for video quality assessment [R]. Ann Marie Rohaly, et al. 2000. http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseI
- [10] WANG Z, BOVIK A C, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Process, 2004, 13(4):600-612
- [11] 余成波,熊飞,李洪兵,等. 大规模 WSN 分簇路由蚁群算法的研究[J]. 四川兵工学报, 2011, 32(11):50-53

Non-reference Video Quality Assessment Based on Wavelet Transform

YAO Jie¹, TAN Jian-ming¹, CHEN Jing¹, HUANG Jian-feng²

(1. Military Grid Laboratory, Chongqing Communication Institute, Chongqing 400035, China;

2. School of Communication Engineering, PLA University of Science and Technology, Nanjing 210007, China)

Abstract: All current video quality assessment algorithms are designed for specific distortion type, however, there are a variety of distortion types in a distorted video and the assessment only for specific distortion type cannot fully reflect video quality. In this view, based on wavelet transform, a non-reference image quality assessment algorithm for comprehensive evaluation on many distortion types was proposed. Finally, a luminance and motion estimation frame weighted approach was utilized to extend this algorithm to video quality assessment. The experiments show that the proposed algorithm has high assessment accuracy.

Key words: video quality assessment; support vector machine; wavelet transform