

基于人眼感知预测的 JPEG2000 码率控制算法

刘洁瑜 张德运

(西安交通大学电子与信息工程学院 西安 710049)

摘要: 基于小波域的 JPEG2000 压缩编码算法, 对重建图像造成的失真是结构信息的丢失, 从而使人眼的图像感知质量下降。为了解决这个问题, 该文提出一种基于人眼感知预测的 JPEG2000 码率控制算法(SIRA), 首先提出了一种度量图像感知质量下降的参数模型, 然后建立了一个能在编码之前预测 JPEG2000 压缩图像感知质量下降的单端预测模型, 基于预测模型实现了 JPEG2000 标准的码率分配。仿真实验结果验证了模型的正确性及算法的有效性。

关键词: JPEG2000; 码率控制; 结构信息; 预测

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2008)05-1198-05

Efficient Rate Control Algorithm Based on Human Perception Prediction for JPEG2000

Liu Jie-yu Zhang De-yun

(School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: The image distortion from JPEG2000 is the loss of structure information, which makes perceived quality dropping in particular at the lower bit rate. So, a rate control algorithm for JPEG2000 is proposed based on the prediction of perceptual structural information distortion without the reference image. The model of measuring the perceived image quality dropping is proposed at first. Then a prediction model of perceived quality dropping for JPEG2000 is constructed. Based on the model efficient rate control is realized. Numerical experiments verify the validity of the proposed method.

Key words: JPEG2000; Rate control; Structure information; Prediction

1 引言

码率控制算法的性能直接影响着图像/视频编码后的重建质量和系统的传输延迟, 因此成为图像/视频编码与通信中的研究核心问题之一^[1]。JPEG2000 是基于小波域的压缩编码算法, 其对重建图像造成的失真是结构信息的丢失, 即模糊(blur)和振铃效应(ringing), 尤其在低码率压缩时。这是因为量化让高频子带的部分小波系数变成 0, 从而使边缘信息丢失。JPEG2000 压缩标准中的码率控制算法是基于率失真理论的优化截取算法(PCRD)^[2], 只是用重建图像的 MSE 值最小作为码率优化函数, 而不管图像的结构信息如何, 结果会使人眼的感知质量下降。文献[3, 4]中提出的码率控制算法同样是依照率失真原理, 根据不同的方式寻找最优截断点, 优点是在不同程度降低了运算量和存储空间, 但其 PSNR 值比标准中的 PCRD 算法还降低了。文献[5]提出了一种优先扫描码率分配算法(PSRA), 首先将整个图像中所有可行截断点按照 S_i^m 斜率进行降序排序, 然后根据这个顺序进行选择, 即编码通道数据从最陡的斜率开始到最平缓的斜率一位一位编码, 直到达到目标码率, 后面的没有编码的通道就

舍弃。这种算法在降低 PCRD 算法复杂度的同时对重建图像质量有提高, 但重建图像的人眼感知质量还不是很理想。

我们知道, HVS(Human Visual System)对排列在一个方向或另一个方向的强梯度的存在特别敏感, 因为这是人眼边缘检测的机制。如果在小波域中能够预测各码块的人眼感知质量, 这样在编码时就可以将更多的码率分配给会引起严重结构失真的码块, 从而做到在编码时就防止尖锐的人眼感知质量的下降, 可以大大提高重建图像的质量。

基于上述分析本文建立一个能在编码之前预测 JPEG2000 压缩图像感知质量下降的模型, 基于预测模型提出 JPEG2000 码率控制新算法, 这样能在编码时就得到高质量的重建图像。

2 基于人眼感知预测的 JPEG2000 码率控制算法

2.1 图像人眼感知质量的度量

仅使用质量下降的重建图像或原始图像按一定方法来评估质量, 称为基于预测的单端无参考帧方法。无参考帧的方法在基于块失真的图像压缩质量预测中取得了很显著的效果^[6, 7], 这是因为块失真只出现在水平、垂直方向, 图像失真出现很有规律。而模糊和振铃是围绕着边缘出现, 过分依赖图像的内容, 使得在无参考帧的条件下预测难度较大。

目前, 对于模糊和振铃预测模型的建立大都是寻找和构建反映原图像局部特征的掩盖图像来作为另一种意义下的参考帧^[8,9]。而文献[10]中提出一种基于学习的图像质量盲评估算法, 取得效果较理想, 但它是在原始图像中对每个边缘点作为一个样本值, 运算复杂, 而将边缘点周围的像素值选为特征值也不十分理想。

二阶统计特征在图像的结构信息分析中应用较为广泛, 也被认为最为有效和可靠^[11]。在本文中, 定义结构信息失真度量时主要考虑该度量不但应该是有效的, 而且应该使建立预测模型更加容易和准确。一幅图像的结构信息包含在小波域的高频子带内, 因此本文定义结构信息强度 I 为结构信息失真的度量, 等于小波域中高频子带内各码块系数的标准差 σ , 即有

$$I_i = \sigma_i \quad (1)$$

构建 l 子带的结构信息强度向量 $\mathbf{IS}_l = (I_{l1}, I_{l2}, \dots, I_{lM})^T = (\sigma_{l1}, \sigma_{l2}, \dots, \sigma_{lM})^T$, 其中 M 为 l 子带中的码块个数。则通过 L 级小波分解后的一帧图像的结构信息失真定义为

$$BD = \frac{1}{3L} \sum_{l=1}^{3L} (E(\mathbf{IS}_l) \times \sqrt{\gamma_b}) \quad (2)$$

其中 $E(\mathbf{IS}_l)$ 为 l 子带的结构信息强度的均值, 即

$$E(\mathbf{IS}_l) = \frac{1}{M} \sum_{i=1}^M \sigma_{li} \quad (3)$$

式(2)中 γ_b 为合成小波滤波器的能量权值, 其定义与 JPEG2000 标准中量化步长选择中所定义的能量权值相同。从信号与系统的角度看, 对于一维内插滤波系统, 一级内插滤波的能量权值是该系统冲击响应的系数的平方和的 1/2; 二级内插系统的能量权值, 则为第一级系统的冲击响应内插后, 与第二级系统的冲击响应进行卷积, 是该卷积结果的系数平方和的 1/4。依次类推, 可以计算得出 N 级系统的能量权值。对于二维滤波系统, 能量权值则是行方向和列方向上的一维能量权值的积。离散小波变换实际上等价于用与小波对应的分解滤波器和合成滤波器对信号先内插后滤波。因此, 可以用上述方法计算得出 N 级小波变换所得到的不同子带的能量权值 γ_b ^[2]。

作为一个例子, 图 1 显示了 Foreman 序列(测试序列 1)及 Mother&daughter 序列(测试序列 2)分别作为测试图像帧的结构信息失真曲线。曲线中结构信息下降与实际感知的质量降低完全对应(图2)。由此可见, 结构信息失真的度量与人眼感知质量之间的相关性很高。另外, 感知质量的下降程度依赖于不同的图像内容。

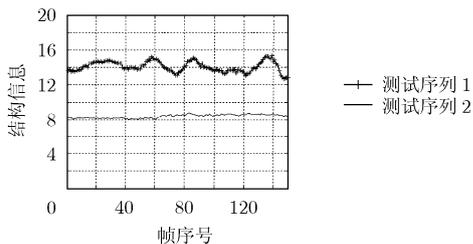


图 1 两个测试序列的结构信息失真度量(目标码率为 0.125bpp)



图 2 结构信息的人眼感知质量

2.2 图像人眼感知质量预测模型

建立预测模型的基本思想是采用机器学习的方法, 因为相似质量的图像在一定特征条件具有一些共同的规律, 而这些规律可以通过给定样本的学习来得到。具体过程分 3 步进行: 建立训练库; 选择分类器; 建立预测模型。

(1) 训练库建立 训练样本由图像小波域中高频子带的码块 $B(i)$ 组成, $i = 1, 2, \dots, N$, N 为总样本数。 $B(i)$ 包含有无失真图像(不经过压缩的)中的相应码块, 定义为 $B^+(i)$, $i = 1, 2, \dots, N^+$; 还包含有失真图像(经过 JPEG2000 压缩的)中的相应码块, 定义为 $B^-(i)$, $i = 1, 2, \dots, N^-$, 其中 $N = N^+ + N^-$ 。定义由 $B^+(i)$ 组成的样本子集为 $S^+(i)$, $B^-(i)$ 组成的样本子集为 $S^-(i)$ 。选择样本的特征量 $f(i)$, $i = 1, 2, \dots, N$ 。用上一节中所定义的码块的结构信息强度加权值及码块小波系数的平均值来构成特征向量 $\mathbf{F}(i)$, 即有

$$\left. \begin{aligned} \mathbf{F}(i) &= [f_1(i) \ f_2(i)]^T \\ f_1(i) &= I_i \times \sqrt{\gamma_b} = \sigma_i \times \sqrt{\gamma_b} \\ f_2(i) &= \mu_i \end{aligned} \right\} \quad (4)$$

(2) 选择分类器 理论上讲, Bayes 分类器产生的分类误差是最小的, 因此本文选择 Bayes 分类器作待测信号的分类器。

(3) 预测模型建立 当给定新的输入 $\mathbf{F}(j)$ 时, 利用 Bayes 分类器建立如式(5)所示的结构失真预测模型:

$$D(j) = \frac{P(S^+ | \mathbf{F}(j))}{P(S^- | \mathbf{F}(j)) + P(S^+ | \mathbf{F}(j))} \quad (5)$$

其中 $P(S^+ | \mathbf{F}(j))$, $P(S^- | \mathbf{F}(j))$ 为给定输入 $\mathbf{F}(j)$ 时, 该输入属于 S^+ , S^- 的后验条件概率。利用 Bayes 法则, 有

$$P(S^+ | \mathbf{F}(j)) = \frac{P(\mathbf{F}(j) | S^+) \times P(S^+)}{P(\mathbf{F}(j) | S^+) \times P(S^+) + P(\mathbf{F}(j) | S^-) \times P(S^-)}$$

$$P(S^- | \mathbf{F}(j)) = \frac{P(\mathbf{F}(j) | S^-) \times P(S^-)}{P(\mathbf{F}(j) | S^+) \times P(S^+) + P(\mathbf{F}(j) | S^-) \times P(S^-)}$$

式(5)可变为

$$D(j) = \frac{P(\mathbf{F}(j) | S^+) \times P(S^+)}{P(\mathbf{F}(j) | S^+) \times P(S^+) + P(\mathbf{F}(j) | S^-) \times P(S^-)} \quad (6)$$

$P(\mathbf{F}(j) | S^+)$, $P(\mathbf{F}(j) | S^-)$ 为给定类型 S^+ , S^- , 输入 $\mathbf{F}(j)$ 分别属于相应类的的条件概率。

同文献[6]假设条件一样,认为训练样本的条件分布是高斯分布 $N(\mu, \sigma^2)$, 则可利用最大似然估计方法求得 $\hat{\mu}, \hat{\sigma}^2$, 可得: $N_{S^+}(\hat{\mu}^+, \hat{\sigma}^{+2}), N_{S^-}(\hat{\mu}^-, \hat{\sigma}^{-2})$, 而 $P(S^+) = N^+/N, P(S^-) = N^-/N$, 从而解得式(6)。则整幅图像的结构失真可表示为

$$\overline{BD} = \frac{1}{M} \sum_{i=1}^M D(i) \quad (7)$$

2.3 基于预测模型的码率控制算法(SIRA)

加入了预测后的码率控制算法的系统结构图如图 3 所示,说明了在码率控制和避免严重的图像人眼感知质量下降之间,人眼感知质量预测模型和率失真模型起到了桥梁作用。

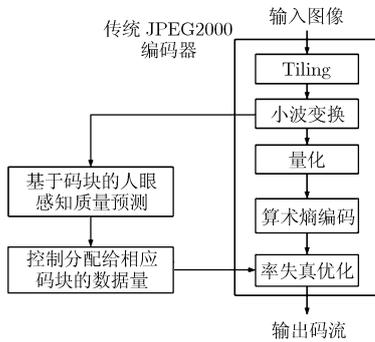


图 3 加入信息预测的 JPEG2000 编码器结构示意图

文献[5]中提出的 PSRA 算法中首先将整个图像中所有可行截断点按照 $S_i^{n_i}$ 斜率进行降序排序, 然后根据这个顺序进行码率分配。码块的扫描顺序是按由低频子带往高频子带逐次的顺序, 而同一子带中按码块空间位置逐次进行, 这也是按照越处于低频的子带系数含有图像越高的能量的原则进行的。本文借鉴 PSRA 优先扫描的思想实现码率控制, 具体按两步进行: 首先基于 PSRA 思想, 规定一个码块的扫描顺序, 本文按照码块所含结构信息的大小, 对高频子带内的码块重新排序。通过结构失真预测模型预测每个块结构信息, 将信息强度大的码块排在前面, 这样码块扫描时, 对低频 LL 子带码块逐次扫描按 PSRA 编码后, 就按信息强度大小顺序进行块扫描, 从而可以保证排在前面的数据块分配多的码率。

接下来对高频子带每个码块根据所预测的 $D(i)$ 调节所对应斜率, 从而在得到高感知质量的图像同时, 也不破坏

JPEG2000 码流的 SNR 渐进性等特性。加入提升因子 $\beta_i^{n_i}$ 的 $S_i^{n_i}$ 斜率可表示为

$$S_i^{n_i} = \frac{\Delta D_i^{n_i}}{\Delta R_i^{n_i}} \times \beta_i^{n_i} \quad (8)$$

$$\beta_i^{n_i} = \begin{cases} 1, & D(i) \leq BD_{th} \\ 1 + c(D(i) - D_{th}), & D(i) > BD_{th} \end{cases} \quad (9)$$

其中 BD_{th} 为一阈值, 即当 $D(i)$ 大于一定值时, 才对相应码块做提升处理, 否则斜率不变, BD_{th} 可通过训练库中的样本训练得到。c 为调节常数。

3 仿真实验结果及分析

(1) 图像人眼感知预测模型的评估 选择文献[12]中的 29 幅高分辨率的图像(典型的 768×512) 分别按 0.03bpp~1.25bpp 的码率进行 JPEG2000 图像压缩得到 319 幅测试图像, 其中包含 29 幅原图像。做 3 级小波分解, 码块大小为 64×64 , 计算相应码块的特征向量, 构成训练库。

我们评估式(6)所表达的预测模型, 通过如下定义的预测误差比进行:

$$p_e = \frac{|BD - \overline{BD}|}{BD} \quad (10)$$

其中 BD 通过式(2)获得, \overline{BD} 通过式(6)、式(7)获得。分别利用 150 帧的 Foreman 序列(cif 格式)、Mother&daughter 序列(cif 格式)和 Coastguard(qcif 格式)序列作为多帧测试图像, 得到如表 1 所示的平均预测误差比, 可以看出, 平均预测误差比都很小, 虽然本文的预测模型是从不同格式的训练样本中得到的, 但预测结果很准确, 说明本文的统计学习分类方法是有效的。

(2) 码率控制实验 如上节所述, 只当预测到将会造成严重的人眼感知质量下降时, 才会使用提升因子改变分配给相应块的码率。图 4 显示了用 PCRD 和本文提出的 SIRA 编码后重建的各序列结构信息失真曲线, 可以看出, 用 SIRA 编码的结构信息失真要小于 PCRD 的。

图 5 显示了不同算法得到的各序列相应的 PSNR 曲线, 用 PCRD 编码的序列的平均 PSNR 值分别是 25.3469dB, 32.3107dB 和 20.2157dB, 用 SIRA 编码的序列的平均 PSNR 是 25.5236dB, 32.5638dB 和 20.4648dB。因此, SIRA 方法不但提高了人眼的感知质量, 而且提高了 PSNR 大约 0.1767dB~0.25dB。图 6 所示为用 SIRA 和 PCRD 分别编码的码流中挑选的 3 帧图像, 可以看出, 用 SIRA 编码的帧的人眼感知质量比 PCRD 的要好的多。

表 1 平均预测误差比 p_e (bpp)

目标码率	0.05	0.1	0.25	0.5	0.75	1	1.25
Foreman	0.083	0.078	0.058	0.042	0.035	0.020	0.018
Mother&daughter	0.076	0.070	0.052	0.039	0.030	0.015	0.019
Coastguard	0.081	0.076	0.053	0.042	0.031	0.018	0.011

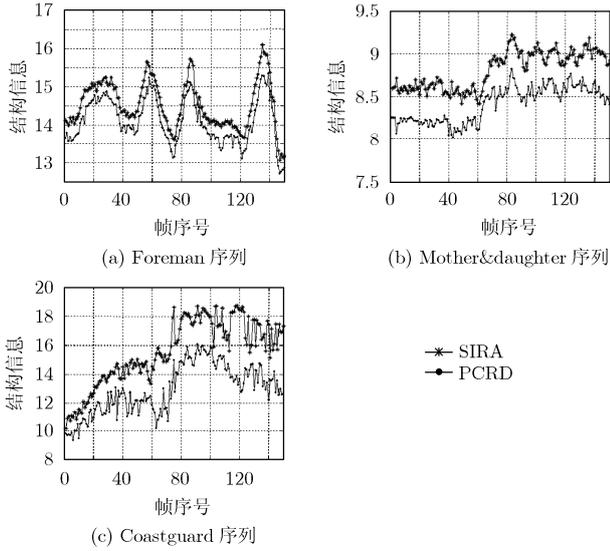


图 4 各序列的结构信息失真曲线(目标码率 0.125bpp)



图 6 不同码率控制算法的人眼感知对比

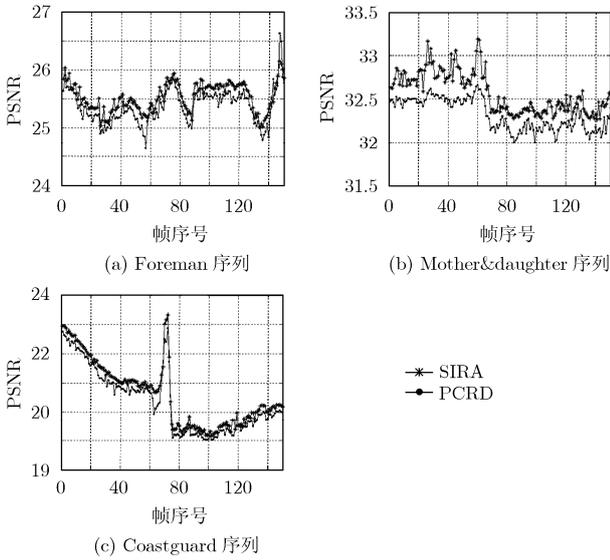


图 5 各序列的 PSNR 曲线(目标码率 0.125bpp)

作为码率控制算法, 相对于主观质量的提升而言, 对于目标码率的控制准确程度和码率的波动程度这两个指标也很重要。表 2 所示为测试序列在本文的 SIRA 算法与传统的 PCRD 算法所取的实际平均编码码率对比, 可看出在目标码率的控制准确程度上, 两种算法的性能是相当的。图 7 所示为不同目标码率下的每一帧的实际码率, 可以看出码率的波动范围很小。

4 结束语

针对 JPEG2000 的 PCRD 码率控制算法造成的重建图像人眼感知质量低的问题, 本文建立了图像小波域中人眼感知质量的预测模型, 基于预测模型提出了 JPEG2000 的 SIRA 码率控制新算法。SIRA 方法不但提高了人眼的感知质量, 而且在一定程度上提高了图像的 PSNR 值。

表 2 目标码率与实际控制码率的对比 (bpp)

目标码率		0.05	0.1	0.25	0.5	0.75	1	1.25
Foreman (平均值)	PCRD	0.0445	0.0949	0.2421	0.4946	0.7414	0.9939	1.2435
	SIRA	0.0443	0.0950	0.2418	0.4932	0.7409	0.9941	1.2408
Mthr&dotr (平均值)	PCRD	0.0442	0.0944	0.2420	0.4934	0.7435	0.9942	1.2442
	SIRA	0.0440	0.0939	0.2416	0.4962	0.7445	0.9935	1.2412
Coastguard (平均值)	PCRD	0.0499	0.0876	0.2241	0.4760	0.7342	0.9457	1.2273
	SIRA	0.0499	0.0869	0.2239	0.4763	0.7341	0.9461	1.2121

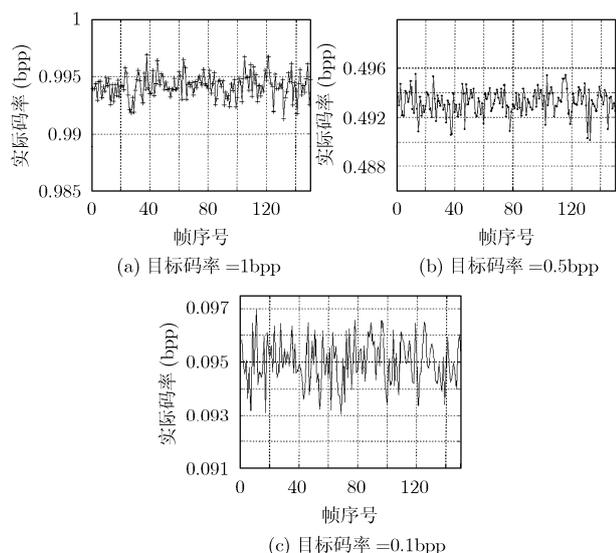


图7 Foreman 序列不同目标码率下每一帧实际输出码率曲线图

参考文献

- [1] Ortega A and Ramchandran K. Rate-distortion methods for image and video compression. *IEEE Trans. on Signal Processing*, 1998, 15(6): 23-50.
- [2] ISO/IEC 15444-1:2000. Information technology-JPEG 2000 image coding system-Part 1: core coding system [S]. 2000.
- [3] Yeung Y M, Au O C, and Chang A. Successive bit-plane rate allocation technique for JPEG2000 image coding. *IEEE ICASSP'03. Hong Kong*, 2003, Vol.3: 261-264.
- [4] Chang TeHao, Lian ChungJr, Chen HongHui, Chang JingYing, and Chen LiangGee. Effective hardware-oriented technique for the rate control of JPEG2000 encoding. *IEEE ISCAS'03. Bangkok, Thailand*, 2003, Vol.2: 684-687.
- [5] Yeung Y M and Au O C. Efficient rate control for JPEG2000 image coding. *IEEE Trans. on Circuits and Systems for Video Tech.*, 2005, 15(3): 335-344.
- [6] Meesters L and Martens J B. A single-ended blockiness measure for JPEG-coded images. *Signal Processing*, 2002, 82(3): 369-387.
- [7] Wang Z, Sheikh H R, and Bovik A C. No-reference perceptual quality assessment of JPEG compressed images. *International Conference on Image Processing [C]. Rochester, New York, USA*, 2002, Vol.2: 3-6.
- [8] Sheikh H R, Bovik A C, and Cormack L. No-reference quality assessment using natural scene statistics: JPEG2000. *IEEE Trans. on Image Processing*, 2005, 14(11): 1918-1927.
- [9] Ong EePing, Lin Weisi, and Lu Zhongkang, *et al.* No-reference JPEG-2000 image quality metric. *ICME'03. Baltimore, USA*, 2003, Vol.1: 545-548.
- [10] Tong Hanghang, Li Mingjing, and Zhang HongJiang, *et al.* No-reference quality assessment for JPEG2000 compressed images. *ICIP'04. Singapore*, 2004, Vol.5: 3539-3542.
- [11] Connors R W and Mcmillin C M, *et al.* Identifying a locating surface defects in wood: Part of an automated lumber processing system. *IEEE Trans. on PAMI*, 1983, 5(6): 573-583.
- [12] Sheikh H R, Wang Z, Cormack L, and Bovik A C. Live image quality assessment database. Available: <http://live.ece.utexas.edu/research/quality>.

刘洁瑜: 女, 1970年生, 博士生, 研究方向为图像、视频处理与通信技术。

张德运: 男, 1941年生, 教授, 主要研究方向为新一代互联网体系结构及高速网关键技术、智能化网络运行管理分析系统、网络安全技术和可靠性技术。