

结合视觉感知的对象基嵌入式图像编码中的码率控制方法

李动节^{1,2}, 朱仲杰^{1,2}, 王玉儿¹

(1. 浙江万里学院 宁波市 DSP 重点实验室, 浙江 宁波 315100; 2. 郑州大学 物理工程学院, 河南 郑州 450001)

摘要: 结合人眼视觉感知特性, 提出一种面向对象基嵌入式编码的码率控制算法。首先结合人眼视觉感知特性估计各对象的重要性并确定其编码优先级; 然后依据优先级对视觉对象进行位平面建模和熵编码, 得到各自独立的码流; 最后基于率失真优化理论在给定码率下对对象码流进行优化截取和重装。仿真实验结果表明, 所提算法可以对不同重要对象进行差异化编码和传输, 与基于 PCRD 算法的码率控制方法相比, 新算法能提高重构图像的整体视觉效果。

关键词: 对象基编码; 嵌入式编码; PCRD 算法; 视觉感知特性; 码率控制

中图分类号: TP391

文献标识码: B

文章编号: 1000-436X(2012)04-0115-06

HVS-based rate-control scheme for object-based embedded image coding

LI Dong-jie^{1,2}, ZHU Zhong-jie^{1,2}, WANG Yu-er¹

(1. Ningbo Key Lab. of DSP, Zhejiang Wanli University, Ningbo 315100, China;

2. Physical Engineering College, Zhengzhou University, Zhengzhou 450001, China)

Abstract: A new rate control algorithm for object-based embedded coding was proposed by incorporating the characteristics of human visual systems (HVS). Firstly, the importance and coding priority of each visual object were estimated. Then, bit-plane modeling and entropy coding were implemented for each object based on the coding priority and its corresponding bit stream was outputted. Finally, bit streams of visual objects were truncated and reassembled based on the rate-distortion optimization principle under the given bit rate. Experimental results reveal that the proposed algorithm can encode and transmit different important objects with different strategies. Compared with the PCRD algorithm, the proposed algorithm can improve the overall visual quality of the reconstructed image.

Key words: object-based coding; embedded coding; PCRD algorithm; human visual system; rate control

1 引言

嵌入式编码能够实现一次压缩, 而解压重建出多幅不同质量和分辨率要求的图像, 近年来受到广泛关注, 是当前图像编码领域的一个研究热点。其中,

EBCOT 算法是最新的一种基于小波变换的嵌入式图像编码方法^[1,2]。它具有较高的压缩效率并支持 SNR 可伸缩性和分辨率可伸缩性编码。EBCOT 算法一般包括 2 个阶段: T1 编码和 T2 编码。在 T1 阶段, 主要完成小波变换、位平面建模和自适应二进制算术编

收稿日期: 2010-09-19; 修回日期: 2011-04-19

基金项目: 国家自然科学基金资助项目 (60902066, 60872094, 60832003); 浙江省自然科学基金资助项目 (Y107740); 浙江省教育厅科研基金资助项目 (Y201016875)

Foundation Items: The National Natural Science Foundation of China (60902066, 60872094, 60832003); The Natural Science Foundation of Zhejiang Province (Y107740); Science Research Foundation of Education Department of Zhejiang Province (Y201016875)

码。在 T2 阶段，根据给定码率基于率失真优化理论对压缩码流进行优化截取和重新组装以得到最终压缩码流。但传统 EBCOT 算法是以码块为基本编码单元，对图像内所有码块采用相同的编码策略，没有考虑图像中不同对象的视觉感知特性。然而，一般来说，图像中不同对象对人眼的视觉重要性是不相同的，在实际应用中，观察者常常只对图像的某些或某类对象感兴趣。例如，在医学研究中，医生只对病理特征对象感兴趣；在遥感图像分析中，观察者只对特定目标感兴趣等。因此，在图像编码中，如果考虑人的视觉感知特性，将图像分割成具有不同视觉兴趣的对象，以视觉对象为编码单元，对不同视觉对象采用不同的编码策略进行独立编码，就可以提高编码后重构图像的主观视觉感效果。

同时，在实际应用中，由于受存储空间和传输带宽的限制，往往要在给定码率下对图像进行编码和传输。因此，码率控制是图像编码中的一个关键问题。传统 EBCOT 算法采用压缩后率失真优化 (PCRD) 算法对 T1 编码后的码流进行优化截取和重装^[3,4]。但 PCRD 算法存储量和计算量比较大，系统实时性差。因此，许多学者针对这方面进行研究，提出了一些改进方法，主要包括：基于率失真斜率模型的算法^[5]、基于预测的实时截断算法^[6,7]、基于动态调整的控制算法^[8,9]等。这些算法在减少编码计算量和存储量方面有一定的贡献，但它们仍然是基于块基的码率控制算法，没有考虑人的视觉感知特性和图像重要对象的优先编码和传输问题，不能应用于对象基的编码算法。

本文在分析 EBCOT 算法的特点和人眼视觉特性的基础上，提出一种面向对象基嵌入式编码的码率控制算法。算法首先基于平均熵理论估计各视觉对象的重要性并确定编码优先级，依据优先级对视觉对象进行编码，得到各自独立的码流。最后，以

小波子带为码率控制单元，在给定码率下以对对象码流进行优化截取和重装。实验结果表明，所提算法可以对不同重要对象进行差异化编码和传输，与基于 PCRD 算法的码率控制方法相比，新算法能提高重构图像的整体视觉效果。

2 对象基嵌入式图像编码

本文首先提出了对象基嵌入式图像编码方案 (EOCOT)。该方案将图像分割成具有不同视觉兴趣的对象，对不同视觉对象进行独立编码，生成各自独立的码流。EOCOT 算法系统框架如图 1 所示。算法首先将原图像分割为不同的视觉对象 $O_i (i = 1, 2, 3, \dots, n)$ ，然后对图像进行 DC 变换和三级离散小波变换 (DWT)，得到各视觉对象在各子带内的小波系数分布。接下来以每个视觉对象为基本编码单元，按照分辨率由低到高的顺序依次对各子带内所有视觉对象进行独立位平面建模和二进制算术编码，得到各个对象的嵌入式码流。最后根据给定码率、对象优先级等参数指标对所有对象压缩码流进行优化截取和重新组装，得到给定码率下的最终压缩码流。

3 结合视觉感知的对象基码率控制新算法

在实际应用中，由于受存储空间和传输带宽的限制，往往要在给定码率下对图像进行编码和传输。因此，码率控制是图像编码中的一个关键问题。本文基于 EOCOT 算法的特点，结合人眼视觉感知特性，提出一种新的嵌入式码率控制算法。算法基于平均熵理论估计各视觉对象的重要性和编码优先级，依据优先级对视觉对象进行编码和码率控制，实现重要对象的优先编码和传输，从而提高重建图像的质量。算法的具体流程如图 2 所示。整个算法可以分为预处理、子带目标码率分配、熵编码和率失真优化截取 4 个主要步骤。对每个主要步骤

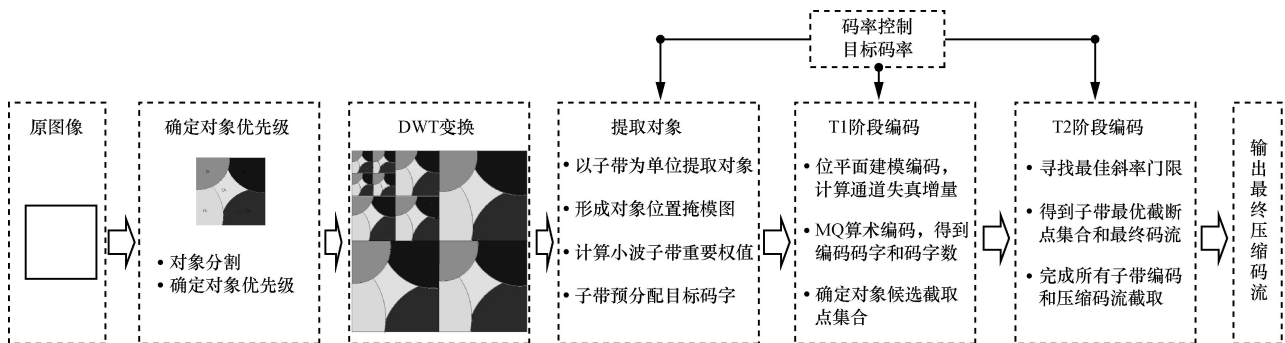


图 1 EOCOT 系统

简介如下。

1) 预处理

在预处理阶段，首先将图像分割成不同的视觉对象 $O_i (i=1,2,3,\dots,n)$ 。本文使用语义对象分割算法进行分割^[10]。然后，基于平均熵理论对视觉对象进行平均熵估计以确定其编码优先级。

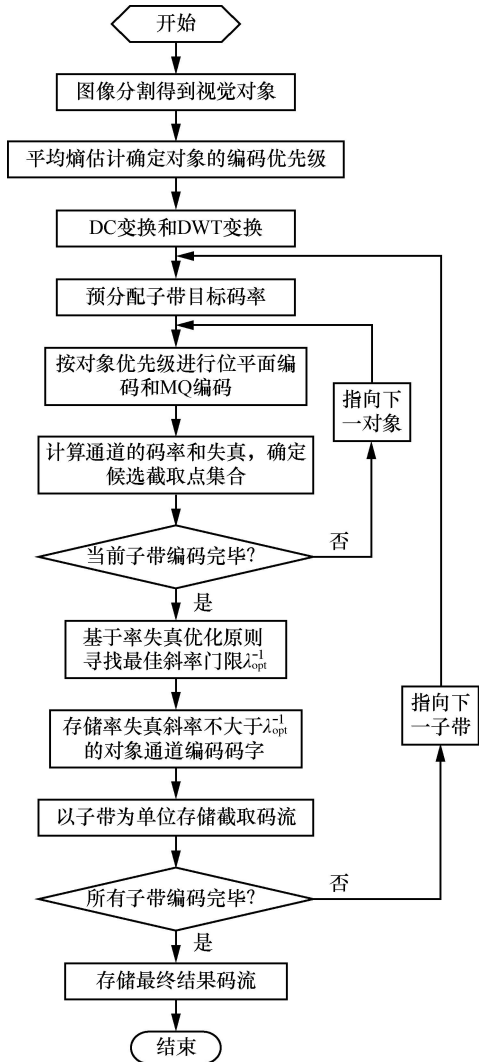


图2 码率控制算法流程

视觉生理学和视觉心理学的研究成果显示，图像编码中的主要人眼视觉特性包括^[11]：人眼对图像平滑区信息的失真比较敏感；人眼对图像边缘区信息的失真很敏感；人眼对图像纹理区信息的失真不敏感。而图像的平滑区、边缘区、纹理区的敏感程度与其信息熵紧密相关。因此，各视觉对象的重要性可以通过分析其信息熵来确定。理论上，信息熵可以根据香农定理进行精确计算，但计算量大。在实际中通常采用近似方法进行熵

估计，比较实用的熵估计方法有2种^[12]：一种是基于预测模板的熵估计法，一种是基于边缘检测的熵估计法。前者是在图像域加一个预测模板，对像素预测值与原值进行差值，然后用累加的差值表示图像的熵值大小，该方法运算量小，适合实时操作。后者是在边缘检测的基础上进行熵估计，检测出的边缘越丰富，其图像熵就越大。该方法运算量也很大，不宜实时操作。所以本文采用第一种方法，预测公式为

$$\hat{x}_{m+1,n+1} = \begin{cases} \min(x_{m,n+1}, x_{m+1,n}), x_{m,n} \geq \max(x_{m,n+1}, x_{m+1,n}) \\ \max(x_{m,n+1}, x_{m+1,n}), x_{m,n} \leq \min(x_{m,n+1}, x_{m+1,n}) \\ x_{m,n+1} + x_{m+1,n} - x_{m,n}, \text{其他} \end{cases} \quad (1)$$

其中， $x_{m,n}$ 为图像域内点 $[m,n]$ 处的像素值， $\hat{x}_{m+1,n+1}$ 为点 $[m+1,n+1]$ 处的预测像素值。如果预测的像素为图像的第一行或第一列，预测值为前一个像素点的值。此时，对象的平均熵为

$$H_{O_i} = \frac{1}{M} \sum_{m \in O_i} \sum_{n \in O_i} |\hat{x}_{m,n} - x_{m,n}| \quad (2)$$

其中， M 为对象 O_i 包含的像素点数。整个图像的平均熵为

$$H = \sum_{O_i \in O} H_{O_i} \quad (3)$$

令 P_{O_i} 表示对象 O_i 的优先级，由式(4)给出：

$$P_{O_i} = \frac{H_{O_i}}{H}, \sum_{O_i \in O} P_{O_i} = 1 \quad (4)$$

2) 子带目标码率分配

为了给予带分配目标码率，首先计算小波子带重要性权值 $W[j]$ ：

$$W[j] = \frac{\sum_{(m,n) \in j} x_{m,n}}{\sum_{(m,n) \in N} x_{m,n}} \quad (5)$$

其中， j 为子带标号， N 为图像所有像素点集合， $x_{m,n}$ 为点 $[m,n]$ 的像素值。然后，基于图像总目标码率 R_{target} ，预分配当前子带的目标码率：

$$R_{\text{target}}^j = W[j] \times R_{\text{target}} \quad (6)$$

3) 熵编码

对于带分配目标码率后，对于带内的视觉对象按照优先级依次进行位平面建模和二进制算术编码，并确定候选截取点集合。对每个视觉对象，按

照从最高位平面到最低位平面的顺序进行位平面建模。在每个位平面内依次按照重要性传播通道、幅值细化通道和清除通道的顺序进行扫描编码。为减少对非编码对象区域零位的编码，只对编码对象的位置信息掩模图内的系数进行扫描编码。在进行通道编码时，每编码一个系数，查失真估计表得到该位的编码失真增量^[13]，累加通道内的所有系数的失真增量得到该通道失真增量。

位平面建模后只是得到每个小波系数比特的上下文和二进制比特符号，并没有实现压缩。要实现压缩必须对二进制符号及其上下文进行进一步的熵编码。本文以对象通道为单位进行算术编码。完成一个对象的所有通道编码后，采用只剔除当前奇异点的方法确定候选截取点集合^[14]，同时存储率失真斜率。

4) 率失真优化截取

设子带 j 内对象 O_i 在 T1 编码产生的内嵌比特流的码率截止到 R_i^m , ni 是某个截取点，则子带 j 内总的码率为

$$R_j = \sum_i R_i^m \quad (7)$$

设对象 O_i 的系数在恢复图像中产生的失真为 D_i^m ，同时假设对象小波系数的失真测度是加性的，即

$$D_j = \sum_i D_i^m \quad (8)$$

其中， D_j 为子带 j 的失真大小。通常失真 D_i^m 可以用加权均方差 (MSE) 进行计算：

$$D_i^m = w_j^2 \sum_{m,n \in O_i} (\hat{V}_i[m,n] - V_i[m,n])^2 \quad (9)$$

其中， $\hat{V}_i[m,n]$ 为重建的视觉对象 O_i 的系数值， w_j 为子带 j 的加权系数。

率失真优化的目标是在满足 $R_j \leq R_{\text{target}}^j$ 的条件下寻找一组 $\{ni\}$ ，使得总失真 D_j 最小。如果采用拉

格朗日准则，此时等价于使式 (10) 最小化：

$$\sum_i (R_i^m + \lambda D_i^m) \quad (10)$$

其中， λ 为拉格朗日乘子。率失真优化截取的关键是寻找一个合适的 λ 使得式 (10) 最小，并满足 $R_j \leq R_{\text{target}}^j$ 。本文采用二分法寻找最佳斜率门限 $\lambda_{\text{opt}}^{-1}$ 。子带内所有对象的率失真斜率不小于 $\lambda_{\text{opt}}^{-1}$ 的编码通道码字被包含进最终码流，而其余编码通道的码字则被丢弃。完成一个子带编码和最终码流截取后，以子带为单位存储对象最优截断码流。然后，进行下一个未编码子带的编码和优化截取，直到完成整幅图像所有子带的编码。最终码流按子带分辨率从低频到高频，按优先级由高到低对对象码流依次存储。最低分辨率子带码流出现在结果码流的起始部分。在子带内，视觉对象的码流按优先级大小排序，最高优先级的对象码流出现在子带码流的起始部分。

4 实验结果

为验证本文算法的可行性和有效性，基于 VC6.0 软件平台，进行了仿真测试，并与 PCRD 算法结果进行对比分析。测试图像从 Google 图像库中随机搜索得到，离散小波变换采用三级双正交 9/7 小波。实验中重点分析不同码率下图像的重构质量，采用信噪比增量指标，定义如下：

$$\Delta psnr = psnr2 - psnr1 \quad (11)$$

其中， $psnr1$ 为 PCRD 算法下重构图像的峰值信噪比， $psnr2$ 为本文算法重构图像的峰值信噪比。

图 3~图 6 和表 1 及表 2 给出了部分实验结果。其中， $\overline{\Delta psnr}$ 表示不同压缩比下的平均 $\Delta psnr$ 。图 3 和图 4 为不同码率控制下，测试图像在 PCRD 算法和本文算法下的重构图像质量对比结果。图 5 和图 6 为不同码率控制下，测试图像的重要视觉对象在 PCRD 算法和本文算法下的重构图像质量对比结

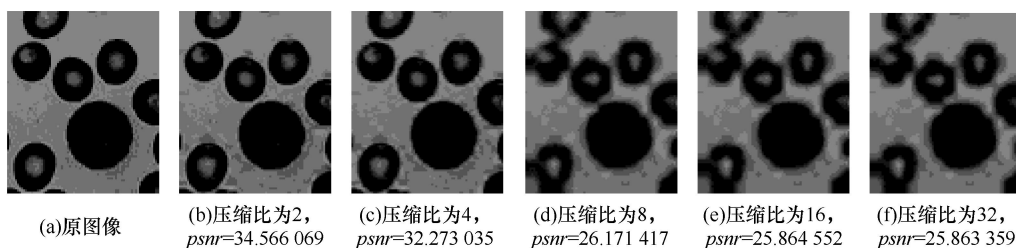


图 3 图像 Cell 在不同压缩比下 PCRD 算法重构图像

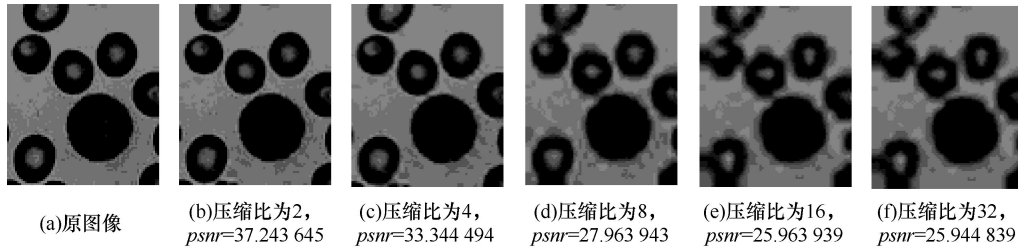


图 4 图像 Cell 在不同压缩比下用本文算法重构图像

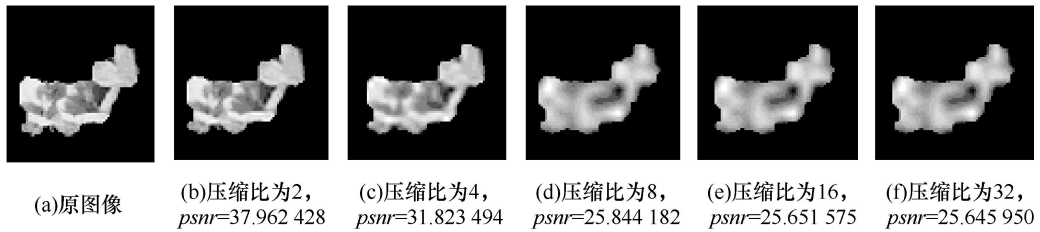


图 5 PCRD 算法下 Flower 中重要视觉对象在不同压缩比下的编码重建结果

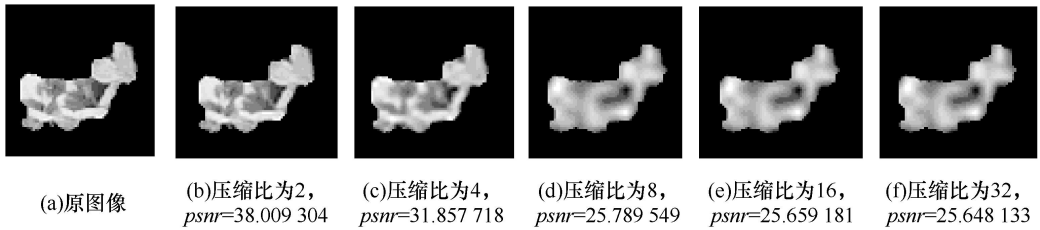


图 6 本文算法下 Flower 中重要视觉对象在不同压缩比下的编码重建结果

表 1 部分测试图像的重建图像质量对比实验结果

压缩比	Cell		Satellite		Flower	
	$\Delta psnr$	$\overline{\Delta psnr}$	$\Delta psnr$	$\overline{\Delta psnr}$	$\Delta psnr$	$\overline{\Delta psnr}$
2	2.677 6		0.196 7		0.057 0	
4	1.071 5		0.067 0		0.288 9	
8	1.792 5	1.144 5	0.022 8	0.055 4	0.010 5	0.077 4
16	0.099 4		-0.013 8		0.020 0	
32	0.081 5		0.004 1		0.010 4	

表 2 部分重要视觉对象重构图像质量对比结果

压缩比	Lake 中重要对象		Satellite 中重要对象		Flower 中重要对象	
	$\Delta psnr$	$\overline{\Delta psnr}$	$\Delta psnr$	$\overline{\Delta psnr}$	$\Delta psnr$	$\overline{\Delta psnr}$
2	1.146 3		1.578 2		0.046 9	
4	0.403 3		0.030 9		0.034 2	
8	0.178 7	0.345 6	0.042 2	0.329 2	-0.054 6	0.013 3
16	-0.001 3		-0.011 0		0.037 6	
32	0.001 1		0.005 7		0.002 2	

果。表 1 为 2 种算法在不同码率控制下，重构图像的信噪比比较结果；表 2 为 2 种算法在不同码率下，重要视觉对象重构图像的信噪比比较结果。由实验结果可以看出，本文算法能在给定码率下，可以根

据需要优先编解码重要视觉对象，与 PCRD 算法相比，提出的码率控制算法能够提高解码图像总体视觉效果。同时，仿真实验结果显示，本文提出的 EOCOT 对不同测试图像的平均编码时间比 EBCOT

节省 16%左右。

5 结束语

在分析研究现有块基嵌入式编码码率算法的基础上, 结合人眼视觉感知特性, 提出一种面向对象基嵌入式编码的码率控制算法。算法以视觉对象为基本编解码单元, 可以根据视觉重要性对不同重要的视觉对象进行差异化编码, 可以优先编码和传输重要视觉对象。算法可以在给定码率下对图像进行精确地码率控制, 与传统的 PCRD 码率控制算法相比, 新算法能提高重构图像的整体视觉效果。

参考文献:

- [1] ZHU W W, LIU L Z. Optimization of the EBCOT algorithm in JPEG 2000[J]. Computer Applications, 2008, 28(6):210-212.
- [2] TAUBMAN D. High performance scalable image compression with EBCOT[J]. IEEE Transactions on Image Processing, 2000, 9(7): 1158-1170.
- [3] HSIN H C, SUNG T Y. A simple rate distortion estimation for embedded block coding[A]. Proceedings of the 9th WSEAS International Conference on Multimedia Systems & Signal Processing[C]. Wisconsin, USA, 2009.43-48.
- [4] TAUBMAN D, MARCELLIN M. JPEG 2000:Image Compression Fundamentals, Standards and Practice[M]. Massachusetts, USA: Kluwer Academic Publishers, 2002.
- [5] VIKRAM K N, VASUDEAVN V, SRINIVASAN S. Rate-distortion estimation for fast JPEG 2000 compression at low bit-rates[J]. Electronics Letters, 2005, 41(1):16-18.
- [6] AMINGOU A, FATEMI O. A novel efficient rate control algorithm for hardware implementation in JPEG 2000[A]. Proceedings of the International Conference on Acoustics Speech and Signal Processing[C]. Philadelphia, PA, USA, 2005.21-24.
- [7] 徐勇, 徐智勇, 赵汝进等. JPEG 2000 的一种编码前码率分配算法[J].光电工程, 2008, 35(10):81-85.
XU Y, XU Z Y, ZHAO R J, *et al.* Pre-coding RDO rate allocation algorithm for JPEG 2000[J]. Opto-Electronic Engineering, 2008, 35(10): 81-85.
- [8] 郑启东, 刘鹏. 基于线性预测的动态阈值 JPEG 2000 码率控制算法[J]. 浙江大学学报(工学版), 2008, 42(8):1335-1339.
ZHENG Q Z, LIU P. Dynamic threshold JPEG 2000 rate control algorithm based on linear prediction[J]. Journal of Zhejiang University (Engineering Science), 2008, 42(8):1335-1339.
- [9] YEUN G Y M, AUO C. Efficient rate control for JPEG 2000 image coding[J]. IEEE Transactions on Circuits and System, Video Technology, 2005, 15(3):335-344.
- [10] ZHU Z J, WANG Y E, JIANG G Y. Statistical image modeling for semantic segmentation [J]. IEEE Trans on Consumer Electronics , 2010, 56(2): 777-782.
- [11] 王向阳, 杨红颖. 基于人眼视觉特性的快速图像编码算法[J]. 软件学报, 2003, 14(11):1964-1970.
WANG X Y, YANG H Y. A fast image coding algorithm based on human visual system[J]. Journal of Software, 2003, 14(11):1964-1970.
- [12] 孔繁铨, 李云松, 王柯伊等. 基于码率预分配的 JPEG 2000 自适应率控制算法[J]. 电子与信息学报, 2009, 31(1): 66-70.
KONG F Q, LI Y S, WANG K Y, *et al.* An adaptive rate control algorithm for JPEG 2000 based on rate pre-allocation[J]. Journal of Electronics & Information Technology, 2009, 31(1): 66-70.
- [13] 易爱清. JPEG 2000 中码率控制算法研究及硬件实现[D]. 西安: 西安理工大学, 2008.
YI A Q. Research on Rrate-Control Algorithm and Hardware Realization for JPEG 2000[D]. Xi'an: Xi'an University of Technology , 2008.
- [14] 王菊花. JPEG 2000 中 MQ 算术编译码器的研究[J].空间电子技术, 2003, 1:27-38.
WANG J H. Research on MQ arithmetic encoder and decoder for JPEG 2000[J]. Space Electronic Technology, 2003, 1:27-38.

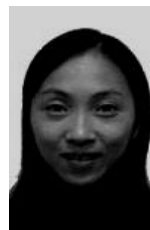
作者简介:



李动节 (1984-), 女, 河南郑州人, 郑州大学与浙江万里学院联合培养硕士生, 主要研究方向为 2D 与 3D 视频编码与传输。



朱仲杰 (1976-), 男, 安徽安庆人, 博士, 浙江万里学院教授, 郑州大学兼职研究生导师, 主要研究方向为 2D 与 3D 视频编码和传输、视频语义分析与检索。



王玉儿 (1979-), 女, 浙江宁波人, 浙江万里学院助理研究员, 主要研究方向为图像处理与传输。