

# 基于上下文建模的分类排序小波图像编码算法

周映虹 马争鸣

(中山大学电子与通信工程系 广州 510275)

**摘要** 该文提出了基于上下文建模的分类排序小波图像编码算法。该方法使用与 JPEG2000 标准相同的上下文模型,以最小化相对熵为准则用动态规划获得优化的上下文分类。在重要性编码中结合优化的上下文分类进行排序编码。实验数据表明该算法的压缩能力优于 JPEG2000,尤其适于低码率压缩。

**关键词** 上下文建模,上下文排序,上下文分类,重要性编码

中图分类号: TN911.73

文献标识码: A

文章编号: 1009-5896(2006)12-2405-04

## Context-Based Quantization and Sorting in Wavelet Image Coding

Zhou Ying-hong Ma Zheng-ming

(Electronics Department, Zhongshan University, Guangzhou 510275, China)

**Abstract** A context-based quantization and sorting algorithm used in wavelet image coding is proposed. The context template in the scheme is the same as JPEG2000's. The optimum context-based quantization is achieved by the means of dynamic programming for minimum relative entropy, and the pixel sorting coding is combined with optimum context quantization in significance coding. The results indicate the compression performance of the proposed algorithm is better than that of JPEG2000, especially at low bit-rate compression.

**Key words** Context modeling, Context sorting, Context quantization, Significance coding

### 1 引言

近年来,集优秀的压缩性能、质量和多分辨率可分级等特性的小波图像编码算法得到了飞速发展。其中以国际标准 JPEG2000<sup>[1]</sup>为主要代表。JPEG2000 采纳 Embedded Block Coding with Optimized Truncation(EBCOT)<sup>[2]</sup>作为核心编码器。与上个世纪九十年代备受关注的零树结构的 Embedded Zerotree Wavelet Coding<sup>[3]</sup>、集合分裂结构的 Set Partitioning in Hierarchical Tree<sup>[4]</sup>相比, EBCOT 完全摒弃了父辈和子辈之间的相关性,使用独立码块的组织编码方式,提供了较零树编码算法更为强悍的压缩性能和多样化的应用特性。

如何充分利用子带内以及子带间的相关性并按照率失真下降最陡的原则来编码,属于小波系数的分类排序技术,是有损压缩中的关键技术。EBCOT 使用两层编码:第 1 层编码是对所有比特平面进行基于上下文建模的熵编码;由于第 1 层编码不能保证按照重要性斜率单调下降来组织码流,第 2 层编码运用 Post-Compression Rate-Distortion optimization(PCRD-opt)<sup>[2]</sup>对第一层压缩后码流的优化截取。但 EBCOT 的上下文建模分类方案不是最优的<sup>[5]</sup>;且要计算每个截断点的率失真斜率,耗费为数不少的辅助信息来组织码流。文献[6]在线估算系数的率失真斜率期望值,计算量非常大。文献[7]提出基于一阶上下文建模的分类排序,距离已知

的重要系数越近的未知系数越先被重要性编码,但其上下文模型是一阶的,利用邻域相关性不够充分,且与 EBCOT 一样没有考虑上下文分类优化问题。为此,本文提出一种改进的基于上下文建模的分类排序小波图像编码算法,在重要性编码中结合优化的上下文分类方案进行排序编码,实验结果显示在复杂度没有增加的前提下,该方法取得了比 JPEG2000 稍高的 PSNR 值,尤其适合于低码率编码。

### 2 基于上下文建模的优化分类

根据信息论,在系数统计相关这一前提下对随机流  $X_n, X_{n-1}, \dots, X_1$  进行熵编码所需的最少平均信息量为

$$\begin{aligned} H(X_n, \dots, X_1) &= - \sum_{x_n, \dots, x_1} p(x_n, \dots, x_1) \log_2 p(x_n, \dots, x_1) \\ &= \sum_{i=1}^n \sum_{x_{i-1}, \dots, x_1} p(x_{i-1}, \dots, x_1) H(X_i/x_{i-1}, \dots, x_1) \\ &= - \sum_{i=1}^n \sum_{x_{i-1}, \dots, x_1} p(x_{i-1}, \dots, x_1) \\ &\quad \cdot \left( \sum_{x_i} p(x_i/x_{i-1}, \dots, x_1) \log_2 p(x_i/x_{i-1}, \dots, x_1) \right) \\ &= - \sum_{i=1}^n \sum_{x^{i-1}} p(x^{i-1}) \cdot \left( \sum_{x_i} p(x_i/x^{i-1}) \log_2 p(x_i/x^{i-1}) \right) \end{aligned} \quad (1)$$

$X^{i-1} = X_{i-1}, \dots, X_1$  称为当前符号  $X_i$  的上下文。上下文建模就是根据上下文来精确估计  $p(x_i/x^{i-1})$ 。一个有效的上下文建模方案主要有 3 个基本问题要解决: (1) 确定上下文; (2)

2005-11-25 收到, 2006-06-26 改回  
广东省科技计划项目(2004B10101031, 2005B10303001)和珠海市科技计划项目(PC20051017)资助课题

估计条件概率；(3)上下文量化，也称为上下文分类。

一般是根据以往符号  $X_{i-1}, \dots, X_1$  与当前符号  $X_i$  在时间或空间上的远近主观地来挑选上下文。本文采用与 JPEG2000 相同的上下文模型，见图 1。估计条件概率，主要有两种方法：(1)假设图像是严格平稳、各态历经的信源，其条件概率用前面符号的累积概率来代替；(2)概率建模，假设条件概率服从某种概率模型。本文采用方法(1)。

$D0$	$V0$	$D1$
$H1$	$X$	$H0$
$D2$	$V1$	$D3$

图1 JPEG2000 的上下文模型  
Fig.1 The context model of JPEG2000

过多的条件概率参数  $p(x_i/x^{i-1})$  需要估计而图像样本有限，因此容易导致“上下文稀释”。上下文量化的目的就是为合并相类似统计特性的上下文，同时尽量减少由于  $p(x_i/f_{\Pi}(x^{i-1}))$  造成的比特消耗。文献[8]将上下文空间采用 fisher 判别准则投影到直线上进行上下文量化；文献[5]用互信息的原则说明了 JPEG2000 的上下文量化方案不是最优的；文献[9]采用 Generalized Lloyd Algorithm 算法来寻找上下文量化最优方案。以相对熵为原则，下面给出上下文量化最优化的相关数学依据。

设  $C = \{ (x_{i-1}, \dots, x_1) \mid x_j \in A, j=1, \dots, i-1 \}$  表示当前随机变量  $X_i$  的上下文  $X^{i-1}$  空间， $A$  表示有限的字母表。上下文量化就是把上下文空间  $C$  分割成  $K$  个互不相交的子空间： $\Pi = \{C_1, \dots, C_K\}$ ， $C = \bigcup_{k=1}^K C_k$ ，这里  $K$  称为上下文量化的阶数。定义变换： $f_{\Pi}: C \rightarrow \{y_1, \dots, y_k\}$ ，如果  $(x_{i-1}, \dots, x_1) \in C_k$ ，则  $y_k = f_{\Pi}(x_{i-1}, \dots, x_1)$ 。若用  $-\log_2 p(x_i/y_k)$  代替  $-\log_2 p(x_i/x_{i-1}, \dots, x_1)$  对  $x_i$  进行熵编码，必会造成额外的比特开销。以相对熵为原则，推导其额外比特开销的统计平均值为：

$$\begin{aligned}
 \Delta_{\Pi} &= H(X_i/Y) - H(X_i/X_{i-1}, \dots, X_1) \\
 &= \sum_{k=1}^K p(y_k) H(X_i/y_k) - \sum_{k=1}^K \sum_{(x_{i-1}, \dots, x_1) \in C_k} p(x_{i-1}, \dots, x_1) \\
 &\quad \cdot H(X_i/x_{i-1}, \dots, x_1) \\
 &= \sum_{k=1}^K \sum_{(x_{i-1}, \dots, x_1) \in C_k} p(x_{i-1}, \dots, x_1) \sum_{y_k} p(x_i/x_{i-1}, \dots, x_1) \\
 &\quad \cdot \log_2 \frac{p(x_i/x_{i-1}, \dots, x_1)}{p(x_i/y_k)} \\
 &= \sum_{k=1}^K \sum_{(x_{i-1}, \dots, x_1) \in C_k} p(x_{i-1}, \dots, x_1) d(p(\bullet/x_{i-1}, \dots, x_1), p(\bullet/y_k)) \\
 &\geq 0 \tag{2}
 \end{aligned}$$

其中

$$p(y_k) = \sum_{(x_{i-1}, \dots, x_1) \in C_k} p(x_{i-1}, \dots, x_1) \tag{3}$$

$$\begin{aligned}
 p(x_i/y_k) &= \sum_{(x_{i-1}, \dots, x_1) \in C_k} \left( \frac{p(x_{i-1}, \dots, x_1)}{\sum_{(x_{i-1}, \dots, x_1) \in C_k} p(x_{i-1}, \dots, x_1)} \right) \\
 &\quad \cdot p(x_i/x_{i-1}, \dots, x_1) \tag{4}
 \end{aligned}$$

$$\begin{aligned}
 d(p(\bullet/x_{i-1}, \dots, x_1) \parallel p(\bullet/y_k)) &= \sum_{x_i} p(x_i/x_{i-1}, \dots, x_1) \\
 &\quad \cdot \log_2 \frac{p(x_i/x_{i-1}, \dots, x_1)}{p(x_i/y_k)} \tag{5}
 \end{aligned}$$

式(5)表示用  $-\log_2 p(x_i/y_k)$  代替  $-\log_2 p(x_i/x_{i-1}, \dots, x_1)$  时所引起的相对熵，式(2)表明  $\Delta_{\Pi}$  是所有这些相对熵的统计平均。最好的上下文分类方法应是  $\Delta_{\Pi}$  最小的分类方法。由于上下文空间  $C$  的元素是有限的，当量化阶数  $K$  固定时，遍历所有可能的分割方法可找到使得  $\Delta_{\Pi}$  最小的分类方法。

由于上下文量化器的失真度量是相对熵，其量化器的胞腔是凸的。因此，对于二进制随机序列，假设上下文量化的阶数是  $K$ ，将  $p(1/x_{i-1}, \dots, x_1)$  或  $p(0/x_{i-1}, \dots, x_1)$  在一维上排序，那么最优的分类方案以在一维的区间范围内划分  $K$  个互不相交且其为原始区间范围的子区间的形式存在。如此，可用动态规划来寻找最优分类。假设排序后的上下文下标为  $i$ ，上下文空间一共有  $N$  个不同上下文， $\Delta_{\Pi}(1:N \Rightarrow K)$  指的是将排序后下标从 1 到  $N$  的上下文划分为  $K$  类造成额外的比特开销。根据式(6)用动态规划寻找使得式(6)最小的分类方案。

$$\begin{aligned}
 \Delta_{\Pi}(1:N \Rightarrow K) &= \min_{i=1, \dots, (N-K+1)} (\Delta_{\Pi}(1:i \Rightarrow 1) + \Delta_{\Pi}(i+1:N \Rightarrow (K-1))) \tag{6}
 \end{aligned}$$

对于上下文模型图 1，根据 8 邻域系数的重要与否重要性编码一共有 256 个不同的上下文。训练图像用 Boat(512×512×8)，woman(2048×2560×8)，peppers(512×512×8)，barbara(512×512×8)和mandrill(512×512×8)。训练时 JPEG2000 编码器<sup>[10]</sup>使用 D9/7 进行 5 级小波分解，在 128:1、64:1、32:1、16:1 和 8:1 等 5 种不同压缩比下平均统计 LH(LL 并入 LH，LH 表示水平低通垂直高通)、HL(水平高通垂直低通)和 HH(水平高通垂直高通) 3 个方向的  $p(x_{i-1}, \dots, x_1)$  和  $p(x_i/x_{i-1}, \dots, x_1)$ 。图 2 表明，将 256 个上下文分成 9 类平均比特损耗  $\Delta_{\Pi}$  接近零。表 1 是采用 JPEG2000 的 9 个上下文分类方案与优化后的 9 个上下文分类方案的  $\Delta_{\Pi}$  值的比较，可见理想情况下优化后的分类方案在对一个系数重要性编码比 JPEG2000 可以节省大约 0.009 的比特位。表 2 是训练图像统计的优化后的 9 个上下文类相应的  $p(1/y_i)$  以及不同上下文分类索引对应  $i$  的分数扫描过程  $S_i$ 。

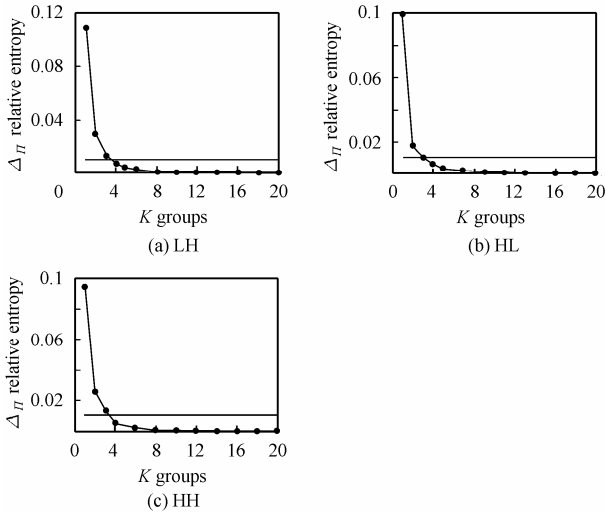


图 2 3 个方向子带的上下文分成  $K$  类的相对熵统计值  $\Delta_{17}$ ，水平线是 JPEG2000 对应的  $\Delta_{17}$ ，曲线代表优化方案

Fig.2 The Curve stands for the relative entropy  $\Delta_{17}$  obtained by optimally context classifying the 256 contexts into  $K$  groups, low horizontal line stands for  $\Delta_{17}$  obtained by JPEG2000, for LH, HL, HH subband, respectively

表 1 JPEG2000 与优化后的 9 个上下文分类方案的  $\Delta_{17}$

Tab.1 The relative entropy obtained by JPEG2000, by optimally context classifying 256 contexts into 9 groups, respectively

	LH	HL	HH
JPEG2000	0.01	0.01	0.01
用式(6)优化后	0.00076	0.00089	0.00074

表 2 优化后的 9 个上下文类下  $p(1/y_i)$  及对应排序子过程

Tab.2 The conditional probability of significance coefficient under the nine optimal context groups and the corresponding ordered subpasses, respectively

上下文分类号下标	对应排序子过程	LH 子带	HL 子带	HH 子带
0	$S_0$	0.029285	0.062604	0.045317
1	$S_1$	0.09278	0.089764	0.10903
2	$S_2$	0.11613	0.11445	0.14719
3	$S_3$	0.1656	0.16054	0.19916
4	$S_4$	0.23527	0.24209	0.259
5	$S_5$	0.3342	0.31905	0.31205
6	$S_6$	0.4016	0.38533	0.40282
7	$S_7$	0.49989	0.4555	0.48707
8	$S_8$	0.62905	0.57499	0.59411

### 3 基于上下文建模的排序

若失真度量是 MSE，根据文献[6]，则比特平面位编码中对任意未编码系数  $x_{b,i,j}$  ( $b$  表示子带下标， $(i, j)$  表示在子带  $b$  中的坐标) 进行重要性编码的率失真斜率的期望值为

$$\frac{E(\Delta D)}{E(\Delta R)} = \frac{\Delta D_{\text{sig}}}{1 + [H(p_{b,i,j}(1))/p_{b,i,j}(1)]} \quad (7)$$

其中系数确定为重要的失真下降值为  $\Delta D_{\text{sig}}$ ， $p_{b,i,j}(1)$  是  $x_{b,i,j}$  在阈值  $T$  下为重要的概率。对于  $p_{b,i,j}(1) \in (0, 1)$ ，

$$\begin{aligned} & \frac{d[H(p_{b,i,j}(1))/p_{b,i,j}(1)]}{d(p_{b,i,j}(1))} \\ &= \frac{-p_{b,i,j}(1)\log_2 p_{b,i,j}(1) - (1-p_{b,i,j}(1))\log_2(1-p_{b,i,j}(1))}{p_{b,i,j}(1)} \\ &= \frac{\log_2(1-p_{b,i,j}(1))}{p_{b,i,j}^2(1)} < 0 \end{aligned} \quad (8)$$

则  $E(\Delta D)/E(\Delta R)$  是  $p_{b,i,j}(1)$  的单调递增函数。因此，按照  $p_{b,i,j}(1)$  从大到小的次序来安排系数的编码次序，也正是按照最陡的率失真斜率期望来安排系数的编码次序，此为排序编码出发点所在。

JPEG2000 使用第 2 层编码来对有限的离散的截断点计算其率失真斜率并以此来组织码流，优化截断点，由于是在比特平面编码后进行的，该算法可称为离线率失真优化。文献[6]提出一种在线率失真优化策略：在编码过程中对每个系数估计其率失真斜率期望值，对于大于和等于当前率失真斜率阈值的系数先行扫描。前者需要在编码时耗费运算时间和辅助信息来组织码流，后者计算量更大；两者都将系数的分类与排序割裂开来。文献[7]提出基于一阶上下文建模的分类排序，不需要在编码时估计率失真斜率期望值和重组码流，但提出的上下文模型仅仅是一阶的，没有充分利用其邻域系数的相关性来进行分类排序，分类的过程中没有考虑上下文量化优化的问题。因此，本文将基于上下文建模的优化分类和排序结合起来考虑：采用表 2 训练的 9 类上下文分类优化方案作为上下文分类方案，并根据该表的训练重要性概率从大到小安排基于相应上下文分类的排序编码次序，其扫描先后次序为： $S_8, S_7, S_6, S_5, S_4, S_3, S_2, S_1, S_0$ 。该方法可避免在线或是离线计算率失真斜率，无需用边信息去说明，只需要训练出优化的上下文分类方案。JPEG2000 的重要性编码有 Significance pass 过程和 Cleanup pass 两个过程，相比下该方法分子过程数是 9，提高了率失真性能。

### 4 算法描述

下面给出基于上下文建模的分类排序算法的详细描述：

(1) 采用提升格式 Daubechies 9/7 核进行五级变换，标量化，实验使用的量化步长与 JPEG2000<sup>[10]</sup> 一致。

(2) 初始化  $n = \lfloor \log_2 \max(|c|) \rfloor$ ， $T = 2^{\lfloor \log_2 \max(|c|) \rfloor}$ 。从最高比特平面开始扫描， $n$  等于该方向的最大 Most Significant Bit 位则开始对该方向子带进行比特平面编码。

(3) 每个子带每个比特平面位被分成了 10 个子扫描过程，逐一从 LL 到 HL 到 LH 到 HH 子带，从粗尺度子带到细尺度子带，单个子带按照光栅方式逐行扫描，处理完一个子过程接着进行下一个子过程：见上下文模型图 1，未编码系数根据其上下文找到所属的分类下标号  $i$ ，先扫描子带中上下文属于下标号 8 的所有系数，再扫描子带中上下文所属下标号大于和等于 7 的所有系数，依此类推，每次扫描  $S_i$  只扫描下标号大于和等于  $i$  的系数，一直扫描到下标号 0 的所有系数。对上一个比特平面确定为重要系数的系数在当前比特平面位要进行精细扫描  $M$ 。10 个子扫描先后次序为

$S_8, S_7, S_6, S_5, S_4, S_3, S_2, S_1, M, S_0$ 。

(4)  $n = n + 1$ ,  $T_n = T_{n-1} / 2$ 。重复进行(3)至码流达到目标码率。

需要说明几点: (1) 排序扫描过程中, 当前系数根据其上下文所属的上下文类, 选择该类的上下文算术编码器, 因此, 重要性扫描过程中用了 9 个上下文算术编码器; (2) 一旦确定为重要系数的系数, 马上启用符号编码, 符号编码、精细编码和 JPEG2000 一样, 使用相同的上下文索引表; (3) 对于子扫描  $S_0$ , 和 JPEG2000 一样, 如果和当前系数在同一列的相邻的 4 个系数其上下文都属于上下文类 0, 则进入游程模式。

## 5 实验结果与分析

为了检验算法的有效性, 本文选用 Lena(512×512×8), Goldhill(512×512×8)和 Cafe(2048×2560×8)图像进行实验。表 3 给出各个算法对 Lena 的仿真结果。其中 case1 编码器除了没有进行排序扫描, 其余都和本文算法(case3)一样, 结果说明结合上下文分类和排序较单一上下文优化分类编码性能好。与文献[7]仿真其 3×3 上下文模型的结果(case2)相比, 本文算法(case3)PSNR 稍好 0.15dB(64:1)。与 JPEG2000 相比, 本文算法 PSNR 值提高了 0.1~0.42dB。图 3 是 Goldhill 和 Cafe 在不同码率下用本文算法减去 JPEG2000 重建图像的 PSNR 差值。结果显示本文算法对纹理复杂的 Cafe 图表现稍逊, 对图像细节中等复杂的 Goldhill 图像, 本文算法是行之有效的, 低码率压缩较高码率好。

表 3 4 个算法对 Lena 图在不同码率下的重建图像 PSNR 值(分贝)  
Tab.3 Peak Signal Noise Ratio of image Lena reconstructed at bit rates for four different schemes

	128: 1	64: 1	32: 1	16: 1	8: 1
JPEG2000 <sup>[10]</sup>	27.952	30.982	34.114	37.286	40.385
Case1	28.276	31.038	33.912	36.902	40.038
Case2	28.276	31.018	34.150	37.266	40.392
Case3	28.376	31.171	34.262	37.385	40.500

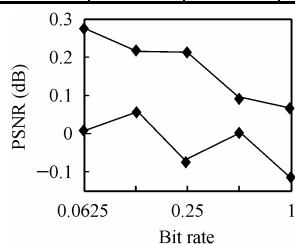


图 3 本文算法较之 JPEG2000 对 Goldhill(下曲线)、Café(上曲线)图在不同码率下的 PSNR 差值

Fig.3 The comparison in Peak Signal Noise Ratio for image Goldhill and Cafe between JPEG2000 and proposed algorithm, upper curve stands for image Cafe, lower curve stands for image Goldhill

本文算法本质上是 8 阶条件概率模型, 上下文量化分类得到优化, 子扫描过程提高到了 10, 这无异乎截断点多了, 率失真凸性在不进行后处理时得到了改善。除了训练优化的上下文分类时需要一定的运算复杂度, 编码时根据训练好的分类表来编码, 编码复杂度没有增加。但 PSNR 幅度仍有提高空间。深究其因主要有三: 训练上下文优化分类方案并不

能精确估计单个图像的  $p(x_i/x^{i-1})$ , 因此根据训练得到的最优方案很受训练图像本身限制; 编码时算术编码是用累积概率  $\hat{p}(x_i/f_{\Pi_{\min}}(x^{i-1}))$  来驱动编码的,  $\hat{p}(x_i/f_{\Pi_{\min}}(x^{i-1}))$  与  $p(x_i/f_{\Pi_{\min}}(x^{i-1}))$  之间的差距导致编码的总体损耗实际是  $\Delta_{\Pi_{\min}} + d(p(X_i/f_{\Pi_{\min}}(x^{i-1}))\|\hat{p}(X_i/f_{\Pi_{\min}}(x^{i-1})))$ , 而算术编码器带来的损耗如何最小化仍没有考虑; 该上下文优化分类与排序是在随机符号流平稳遍历的假设前提下得到的, 而符号流一般是非平稳的。如何克服以上的这些缺陷, 仍需继续深入研究。

## 6 结束语

本文提出结合上下文建模的优化分类和排序的算法, 实验数据表明, 在复杂度没有增加的前提下, 其压缩能力比 JPEG2000 稍优, 尤其适于低码率编码。主要特点有: 以相对熵最小为原则优化上下文分类, 充分利用高阶相关性将上下文分类与排序结合起来考虑, 编码时无需额外边信息; 子带单独编码, 相对零树编码器而言能避免子带间的错误繁衍。

## 参考文献

- [1] ISO/IEC 15444-1:Information technology-JPEG 2000 image coding system-Part1: Core coding system, 2000.
- [2] Taubman D. High performance scalable image compression with EBCOT. *IEEE Trans. on Image Processing*, 2000, 9(7): 1158-1170.
- [3] Shapiro J M. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans. on Signal Processing*, 1993, 41(12): 3445-3462.
- [4] Said A, Pearlman W A. A new, fast, and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans. on Circuits and Systems for Video Technology*, 1996, 6(3): 243-250.
- [5] Zhen Liu, Karam L J. Mutual information-based analysis of JPEG2000 contexts. *IEEE Trans. on Image Processing*, 2005, 14(4): 411-422.
- [6] Li J, Lei S. Rate-distortion optimized embedding. Picture Coding Symposium. Berlin, Germany, 1997: 201-206.
- [7] Peng Kewu, Kieffer J C. Embedded image compression based on wavelet pixel classification and sorting. *IEEE Trans. on Image Processing*, 2004, 13(8): 1011-1017.
- [8] Wu Xiaolin. Context quantization with Fisher discriminant for adaptive embedded wavelet image coding. Data Compression Conference, Snowbird, USA, 1999: 102-111.
- [9] Chen Jianhua. Context modeling based on context quantization with application in wavelet image coding. *IEEE Trans. on Image Processing*, 2004, 13(1): 26-32.
- [10] Adams M D. Official reference implementation of the JPEG-2000 Part-1 codec. [Online]http://www.ece.uvic.ca/~mdadams/jasper/. 2005. Available.

周映虹: 女, 1978 年生, 博士生, 研究方向为图像与视频压缩。  
马争鸣: 男, 1957 年生, 教授, 研究方向为通信与信息系统。