

JPEG2000 中重要性编码及上下文建模的改进

周映虹 马争鸣

(中山大学信息科学与技术学院, 广州 510275)

摘要 众所周知,基于上下文建模的算术编码是 JPEG2000 标准中的关键技术,而对于重要性编码, JPEG2000 中采用的上下文模型是码块内的 8 邻域系数,由于其根据经验将所有上下文分成 9 类,且所有比特平面都使用这套固定的上下文分类方案。因此针对这种上下文分类方案的不足,对 JPEG2000 中重要性编码及其上下文建模方式进行了改进,即首先建立比原来的 3×3 更大的上下文模型;然后提出了一种基于重要上下文的扫描方式;最后基于新的扫描方式根据最小码长的准则进行上下文量化的优化,考虑到不同子扫描之间的统计差异,对不同的扫描子过程使用不同的上下文模型方案。实验结果表明,与 JPEG2000 的无损压缩相比,新的重要性编码算法的无损压缩结果(平均比特率)较 JPEG2000 提高了 1.312%。

关键词 JPEG2000 标准 重要性编码 上下文建模 上下文量化

中图法分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2008)08-1402-09

An Improvement of Significance Coding and Context Modeling in the JPEG2000 Standard

ZHOU Ying-hong, MA Zheng-ming

(School of Information Science and Technology of Sun Yat-Sen University, Guangzhou 510275)

Abstract Context-based arithmetic coding is the key component of the new JPEG2000 standard. For significance coding, JPEG2000 designs empirically context model using eight adjacent coefficients and classifies all context events into 9 conditional contexts and every bit plane use the same context classification maps. This paper addresses the improvement of significance coding and context modeling for significance coding in the JPEG2000 which enlarges the context template which was 3×3 square region in JPEG2000. We propose a new scan method based on significant context and exploit optimization techniques for context quantization by minimum description length considering the difference of statistic among sub scans. Finally, we establish different context models for different sub scans. Our experimental results show that the lossless compression performance of our scheme is improved 1.312% upon that of JPEG2000.

Keywords JPEG2000 standard, significance coding, context modeling, context quantization

1 引言

JPEG2000^[1]已被采用为最新的图像压缩国际标准,与以往基于离散余弦变换的 JPEG 标准不同, JPEG2000 使用了具有时频局部特性的小波变换,并

对小波系数使用了以下两层编码结构:第 1 层是内嵌比特平面编码;第 2 层是根据第 1 层编码的码率失真结果来组织码流。由此可见,小波变换后的图像,从无损压缩的角度看来,起关键作用的实际上是 JPEG2000 的第 1 层编码,也就是基于自适应上下文建模的比特平面内嵌编码。JPEG2000 的每个比特平

基金项目:广东省科技计划项目(2004B10101031, 2005B10303001)

收稿日期:2006-10-14;改回日期:2007-03-02

第一作者简介:周映虹(1978 ~),女,2007 年于中山大学电子与通信工程系获无线电物理专业博士学位。主要研究方向为小波图像压缩。E-mail: iris_zyh@126.com

面编码有重要性信息的编码、符号编码和精细化信息的编码等 3 个子过程。这 3 个子过程当中,重要性信息编码相关性最强,其占据了压缩任务的主要部分。

JPEG2000 采纳 EBCOT^[2] 作为核心编码器,其重要性信息的上下文建模方案是利用了当前系数的 8 邻域系数的重要性信息作为条件上下文和分成 9 类条件上下文类这两方面经验设计的上下文分类方案。问题是,在给定的上下文模型下, JPEG2000 的上下文分类是否最优? 如果不是最优,那么存在的更优分类方式是什么? 除了 8 邻域的上下文模型,是否还有更合适的上下文模型能提供更佳的压缩性能? 文献 [3]、[4] 分别以最小码长和最大互信息的准则对 JPEG2000 的 8 邻域上下文模型的分类方案进行了优化分析。根据文献 [3] 提供的无损压缩的测试图像实验结果,优化的分类方案的压缩效率仅仅比 JPEG2000 方案提高了 0.3%。根据文献 [4] 提供的有损压缩图像的峰值信噪比,优化的分类方案在相同压缩比下,比 JPEG2000 压缩图像的峰值信噪比还要低 0.02dB。因此,文献 [3]、[4] 都有相同的结论,即 JPEG2000 的上下文模型“接近最优”。尽管如此, JPEG2000 的上下文建模方案还是存在不少缺点,其一方面表现在,基于当前系数的 3×3 的上下文模型较小,这说明所能依据的上下文信息不够多;另一方面, JPEG2000 对所有比特平面都采用同一套上下文分类方案,这种方式对非平稳的随机序列是不恰当的。

本文提出了一种对 JPEG2000 重要性编码的上下文建模进行改进的方案,即首先,建立比原来 3×3 更大的上下文模型;然后根据重要小波系数的分布特性,先扫描靠近已发现的重要系数的邻域系数,再扫描较远的邻域系数,这样,具有不同聚集程度的小波系数被区分开来,从而减缓了因概率抖动而带来的影响。另外,对不同聚集程度的区域使用不同的上下文分类表,再针对这种扫描方式,以最小码长为准则来对上下文进行优化量化。对比实验表明,新的上下文模型和扫描方式由于运用了优化的上下文分类方案,因此其压缩结果(平均比特率)较 JPEG2000 提高了 1.312%。

2 JPEG2000 中的重要性编码及上下文模型

若小波系数幅值 w 用二进制表示,则其第 i 个二进制位是 w_i ($i=0,1,\dots,r$),其中 $i=r$ 表示系数的

最重要位(the most significant bit, MSB)。 $w_{\dots i}$ 表示从 MSB 到 i 位的关于幅值 w 的二进制信息。关于 w 在比特平面 2^i 的重要信息定义为

$$\text{sig}(w) = \begin{cases} 1 & w_{\dots i} > 0 \\ 0 & \text{其他} \end{cases} \quad (1)$$

$\text{sig}(w) = 1$ 表示重要,相反, $\text{sig}(w) = 0$ 表示不重要。重要性信息编码就是从所有系数的最大 MSB 位所在的比特平面开始扫描,直至所有系数的 MSB 位都被确定下来。下文讨论的重要性编码的对象均指二进制随机序列。

小波变换后, JPEG2000 以条带的扫描方式来扫描码块里的小波系数,即采用每 4 行逐列的扫描方式(见图 1)。JPEG2000 是使用 8 邻域系数的重要性信息作为原始的上下文模型(见图 2),其中未经分类的上下文一共有 256 种。JPEG2000 的重要性编码分为以下 2 个子过程:首先是 ZC(zero-coding)子过程,即该子过程将在之前的比特平面上的系数确定为非重要的,且其 8 邻域范围内至少存在一个已知重要的邻域系数,这样的系数用于当前比特平面的重要性确认;然后是 CleanUp(CP)子过程,该子过程用于对剩余未知重要性的系数进行重要性确认。除了最高比特平面只有一个 CP 子过程外,其余比特平面均包含以上 2 个子过程。

D_0	V_0	D_1
H_1	X	H_0
D_2	V_1	D_3

图 1 JPEG2000 的 8 邻域上下文模型

Fig. 1 8-connected neighborhood context of JPEG2000

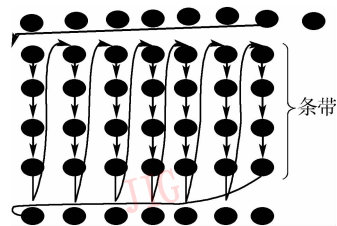


图 2 JPEG2000 的条带扫描

Fig. 2 Strip scanning of JPEG2000

由于实际编码中的样本数有限,使得高阶统计与编码代价存在矛盾, JPEG2000 将原始的 256 个不同上下文 c 分类成 9 类上下文 A_{JPEG2000} ,并根据高频小波系数的 3 个不同方向的子带建立了 3 个不同的上下文映射关系,即

$$A_{\text{JPEG2000}} = \{c: Q_{\text{HL, LH, HH}}(c) = m, m = 1, 2, \dots, 9\} \quad (2)$$

JPEG2000 的上下文映射索引表可参考文献 [4]。JPEG2000 的上下文分类方案设计是基于经验的。文献 [3]、[4] 专门研究了 JPEG2000 的重要性信息的上下文建模。文献 [4] 以分类后的条件熵与分类前最接近为准则来考察 JPEG2000 的上下文分类方案。文献 [3] 根据不同的比特平面和子带的重要小波系数的概率特性的不同, 将比特平面分成以下 7 类: 类 0, 1, 2, 3 分别对应于比特平面 2^0 、 2^1 、 2^2 和 2^3 , 类 4 代表比特平面 2^4 到 2^{r-2} (r 是 MSB 比特平面), 类 5 表示比特平面 2^{r-1} , 类 6 指的是比特平面 2^r ; 假设图像进行 Z 级小波变换, 其中最大分辨率的高频子带表示为 HL_1 、 LH_1 和 HH_1 , 最小分辨率的高频子带表示为 HL_Z 、 LH_Z 和 HH_Z , 那么所有子带分成以下 4 类: 类 0 是 HH_1 , 类 1 是 LH_1 以及 HL_1 , 类 2 为 $HH_{Z \dots 2}$, 类 3 为 $LH_{Z \dots 2}$ 以及 $HL_{Z \dots 2}$ 。文献 [3] 将这 28 种不同的情况下的 256 种上下文纹理信息, 并且将能量信息作为原始的上下文空间, 使用最小描述码长 (minimum description length, MDL) 准则对其进行了上下文初始化的优化以及上下文分类的优化, 但最终的压缩文件的尺寸仅仅较 JPEG2000 提高了非常微小的 0.3%。

3 扫描路径

一般地, 由于量化后的小波系数要通过扫描才可将原来 2 维的小波系数矩阵变成 1 维的序列, 再进行熵编码, 最后形成比特流, 因此, 扫描过程难免或多或少地割裂了原先 2 维小波图像的相关性。好的扫描方式应体现在产生的随机序列的相关性强, 即能尽可能保留原来 2 维矩阵的 2 维自相关。由于 JPEG2000 的条带扫描其实是采用局部线扫描方式, 因此其相关性比整逐行或整逐列扫描要强些。另外有一种是如 Hilbert-Peano 曲线的基于空间填充曲线 (space-filling curve, SFC) [5] 的扫描方式。基于 Hilbert-Peano 曲线的扫描, 其分维数逼近 2, 因此相比起分维为 1 的线扫描来说, 能提供更好的自相关序列 [6]。图 3 是对灰度图像 Lena (512 × 512) 和 Barbara (512 × 512) 进行 5 级 5/3 小波提升格式后, 再在比特平面 2^3 、 2^4 、 2^5 和 2^6 进行重要性编码所产生的 8 个序列的自相关函数 $R(n)$ 的平均统计。图 3 所示的是条带扫描、Hilbert 曲线 (Hilbert curve, HC) 扫描以及基于本文提出的新算法

扫描分别得到的重要信息序列的自相关函数 $R(n)$ 。图 3 说明 HC 的一阶自相关性较基于局部线扫描的自相关性高; 条带扫描的四阶自相关性较 HC 扫描的四阶自相关性强。 $n > 4$ 时, HC 扫描较条带扫描自相关性强。图 3 也暗示了小波系数存在一定程度的非线性相关性。

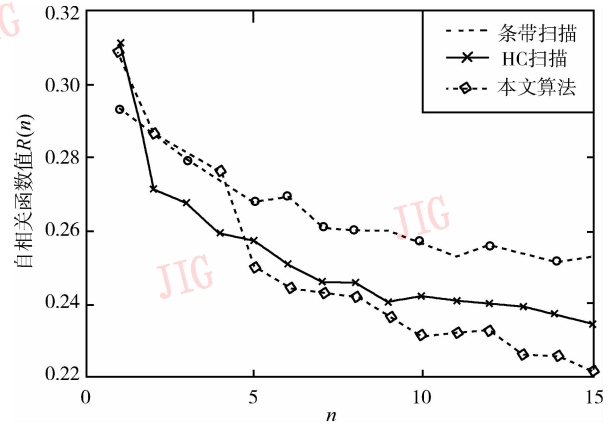


图 3 条带扫描、HC 扫描和本文算法的平均自相关函数
Fig. 3 The average autocorrelation of stripe scanning, HC scanning and proposed scanning method

本文使用传统的自适应算术编码 [7] 作为熵编码器, 它依赖对过去序列的累积统计来估计当前的重要性概率。假设 $\{X_1, X_2, \dots, X_n\}$ 为一随机变量序列, 那么对该序列进行熵编码需要的最小理想平均码长为

$$H(X_1 X_2 \dots X_n) = \sum_{i=1}^n H(X_i | \mathbf{X}^{(i-1)}) \quad (3)$$

其中, $\mathbf{X}^{(i-1)} = \{X_1 X_2 \dots X_{i-1}\}$ 。编码时, 扫描得到的序列分别为 $X_1 = x_1, X_2 = x_2, \dots, X_{i-1} = x_{i-1}$, 即 $\mathbf{x}^{(i-1)} = \{x_1 x_2 \dots x_{i-1}\}$ 。然而, 由于序列的真实条件概率 $p(X_i | \mathbf{x}^{(i-1)})$ 是难以获得的, 因而编码器常常使用 $p(X_i | \mathbf{x}^{(i-1)})$ 的估计值 $\hat{p}(X_i | \mathbf{x}^{(i-1)})$ 来驱动算术编码, 则实际的平均码长为

$$\begin{aligned} & - \sum_{i=1}^n \sum_{X_i} p(X_i | \mathbf{x}^{(i-1)}) \text{lb}(\hat{p}(X_i | \mathbf{x}^{(i-1)})) \\ & = \sum_{i=1}^n H(X_i | \mathbf{X}^{(i-1)}) + \sum_{i=1}^n D(p(X_i | \mathbf{x}^{(i-1)}) || \hat{p}(X_i | \mathbf{x}^{(i-1)})) \end{aligned} \quad (4)$$

可见, 要使平均码长尽可能少, 就需要最小化 $H(X_i | \mathbf{X}^{(i-1)})$ 和 $D(p(X_i | \mathbf{x}^{(i-1)}) || \hat{p}(X_i | \mathbf{x}^{(i-1)}))$ 两部分。下面先讨论如何安排扫描, 使得算术编码能更逼近估计概率, 以减少编码过程中产生的估计误差, 即式 (4) 的第 2 部分, 再讨论式 (4) 的第 1 部分。

图 4 是 Barbara(512 × 512) 经过 5 级 5/3 小波提升格式变换后在比特平面 2^5 上的小波系数图。其中灰色点表示在比特平面 2^5 之前的比特平面上被确定为重要系数的点,黑色点表示出现在比特平面 2^5 上的新的重要系数点,而白色区域则表示比特平面 2^5 上确定的不重要系数点。图 5 是使用 JPEG2000 重要性编码(ZC 和 CP 过程)对图 4 进行扫描产生的序列。图 6 是采用 Hilbert curve(HC)扫描替代条带扫描对图 4 进行扫描产生的重要性信息。为了保持原有图像的分辨率,图 5 ~ 图 7 都是将比特平面 2^5 上不需要扫描的已知重要系数(灰色点)放置在最前面。由图 5、图 6 可见, JPEG2000 的条带扫描及 HC 扫描两种方式,从重要系数出现可能性较大的区域过渡到重要系数出现可能性较小的区域的交叉次数较多,即产生的序列的重要性概率抖动得非常明显。

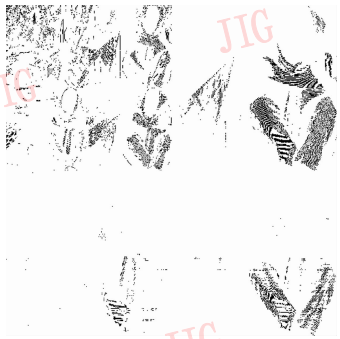


图 4 Barbara 图在比特平面 2^5 上的小波系数图

Fig. 4 Wavelet coefficients of 2^5 bit plane of image Barbara

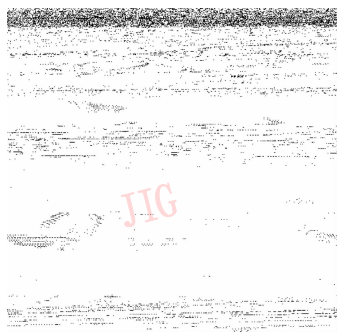


图 5 Barbara 图在比特平面 2^5 上的 JPEG2000 条带扫描结果

Fig. 5 The result based stripe scan of JPEG2000 of

2^5 bit plane of image Barbara

从图 4 可见,子带内仍存有部分非线性相关性,这主要表征在重要系数有聚集的趋势,即在新发生的重要系数的邻域周围,非常有可能也出现新的重要系数。尽管 HC 或局部条带的扫描方式利用了局部相关性,但由于在没有扫描到新重要系数的位置

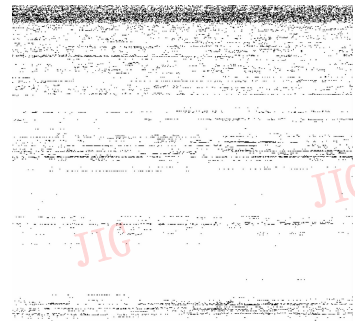


图 6 Barbara 图在比特平面 2^5 上 Hilbert curve 的扫描结果

Fig. 6 The result based on Hilbert curve scan of 2^5 bit plane of image Barbara

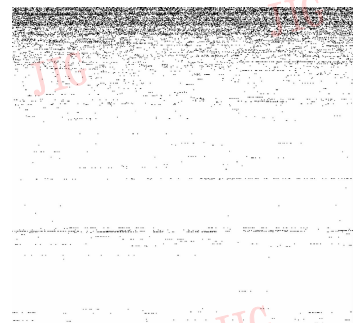


图 7 Barbara 图在比特平面 2^5 上新扫描算法的扫描结果

Fig. 7 The result based on proposed scan algorithm of

2^5 bit plane of image Barbara

之前,是难以预知其邻域系数的重要性概率的,同时也对该重要系数距离较近的子带内的邻域区域的重要性概率与距离更远的邻域区域的重要性概率之间有差异失去辨别能力,因而易产生概率抖动,进而导致 $D(p(X_i | \mathbf{x}^{(i-1)}) || \hat{p}(X_i | \mathbf{x}^{(i-1)}))$ 的增加,也使得编码损耗加大。究其原因,主要由于条带模式或者 HC 扫描模式对小波系数的特性利用不够所致。

由于 X_i 是二进制随机变量,其概率单形体是 1 维区间,因此与 $p(X_i | \mathbf{x}^{(i-1)})$ 之间最小的相对熵不是位于最接近 $p(X_i | \mathbf{x}^{(i-1)})$ 的左边,就是落在最接近 $p(X_i | \mathbf{x}^{(i-1)})$ 的右边。若序列中的随机变量之间的概率分布特性越接近,则有利于算术编码估计概率,而且其产生的抖动影响越小。尽管人们还无法确切预知其概率分布特性,但根据小波系数的聚簇趋势,在已发现的重要小波系数空间邻近周围发现新的重要系数的可能性非常大,且普遍具有随着与这些已知重要系数的邻域距离增长,重要小波系数的出现频率逐步下降的趋势^[8]。据此,本文提出了一种新的多扫描方式:根据与已知重要系数的距离从近到远扫描,即大体上按照重要的小波系数出现的频率呈单调下降

的趋势来确定被扫描的顺序,这样不仅可以使得产生的序列抖动的次数下降,同时这种扫描方式可以明显区分出发生重要性概率不同的区域,从而有利于提高算术编码估计概率的准确度。图 8 是扫描依赖的上下文模型。其中 X 表示已知的重要系数,数字 1 到 6 表示其邻域系数与该重要系数之间的相应“距离” $d=1,2,\dots,6$ 。子带扫描顺序如图 9 所示。每个子带里除了 MSB 比特平面一次性按照 HC 进行清除扫描外,其余比特平面都按照 Scan_Dist 函数的流程进行扫描。重要系数链表 LIS 在扫描之前,先初始化为空链表,并将扫描过程中确认为新的重要系数加入 LIS 链表的末尾;然后即可借助 LIS 来访问其邻域的未知系数。重要性编码流程如下:

(1) 初始化重要系数链表 LIS 为空表。初始比特平面 $t_n = 2^{\lfloor \log_2(\max |w|) \rfloor}$ 。将数组 idx 初始化为 0, $\text{idx}[d-1]$ 表明基于距离等于 d 的邻域查询访问到 LIS 的第 $\text{idx}[d-1]$ 个已知的重要系数。

(2) 若 LIS 非空,分别执行依据 LIS 从 $d=1$ 到 $d=6$ 的邻域未知系数的访问,并按以下次序调用函数 Scan_Dist:

Scan_Dist(d , TRUE);其中 d 表示距离, d 依顺序从 1 到 2, \dots 到 6,先扫描完距离 $d=1$ 的邻域,再自增扫描完 $d=2,3,\dots,6$ 。TRUE 表示距离为 d 的首次扫描访问。

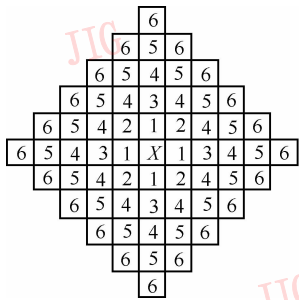


图 8 新扫描算法采用的上下文模型

Fig. 8 Context template of proposed scan algorithm

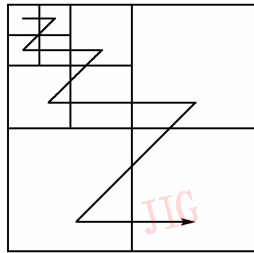


图 9 子带扫描顺序

Fig. 9 Scan path for coding the subbands

(3) 清除扫描过程,按照图中的子带扫描顺序扫描每个子带,子带里按照 HC 来扫描。一旦发现新的重要系数,则同样按距离从小到大依次执行以下函数:

Scan_Dist(d , FALSE);其中 d 表示距离,此处 d 从 1,2, \dots , 4,5,6,即基于新重要系数的邻域进行距离从小到大的扫描。FALSE 表示距离为 d 的扫描是次轮扫描。区分首轮和次轮扫描,是因为距离

d 的首轮聚簇性往往比次轮聚簇性强很多,这样可便于后续采用不同的概率估计模型。一旦发现新重要系数,则将该系数加入 LIS 的末端,并对其邻域进行从距离 1 到 6 的扫描,这是考虑到发现的新重要系数邻域发生重要系数的可能性较该像素点没有发现重要系数的高些,为了让概率变化次数减少,应对新重要系数邻域再行扫描,而不是直接扫描一次性访问完剩余的系数。函数 Scan_Dist 如下:

```
// global variable nIMP is the number of known significant
coeffs. Global array idx indicates the index
// scanned in LIS based on d
Function Scan_Dist (unsigned d, Boolean first_run)
{
    unsigned d_emb;
    while( idx[ d ] < nIMP )
        /* 若 LIS 中仍有已知重要系数没有被基于距离 d
        访问,则从 LIS 中取第 idx[ d ] 个系数出来,采用图中模型将该
        系数放置在“X”的位置,并对其标注为 d 的邻域且未被编码的
        点进行重要性编码 */
        for each coeff w_dis marked with “d” ( Fig. 8 ), the
        neighborhood coeffs of the idx[ d ] th X from LIS
        if w_dis have not been coded
            code( w_dis, first_run ); /* 对其进行重要性编码,
            变量 first_run 为 TRUE 表示首轮,为 FALSE 为次轮,是为了获
            取对应上下文模型 */
            if w_dis is a new significant coeff
                put w_dis to the end of LIS and then nIMP + +
                for ( d_emb = 1; d_emb < d; d_emb + + )
                    Scan_Dist( d_emb, FALSE );
                End for
            End for
            idx[ d ] + + ; move on to next node of LIS
        } // END while
    } // END function
```

(4) $t_n = t_n/2$,回到步骤(2)进行下一个比特平面的重要性编码。

新扫描算法产生的序列见图 7。从图 7 可见,新算法产生的序列,其重要系数出现的频率大体上呈单调下降的趋势,抖动影响小,由于不同概率特性的区域可明显区分开,因而有利于后续基于上下文建模的算术编码进行进一步压缩。

针对式(4)的第 1 部分,新扫描算法产生的随机序列,其所消耗的平均比特是否较条带扫描及 HC 扫描所消耗的平均比特小呢?由于新扫描算法运用了较大范围的邻域作为上下文信息来指导扫描,因此产生的随机序列具有较后两者自相关性强的优势。如

图 3 所示,新扫描算法产生的重要信息序列的自相关性较条带扫描以及 HC 扫描都强。随机序列编码时,如果从前面已编码的序列中获取的相关信息越多,则 $H(X_i | \mathbf{X}^{(i-1)})$ 就越小。因此,除了抖动小,有利于算术编码估计概率以外,新算法产生的序列还具有自相关性较强的特性,这就在很大程度上决定了对这样的随机序列进行熵编码消耗的平均比特是较小的。表 1 是经过 5 级 5/3 提升格式后,再采用 3 种扫描方式产生的重要性信息直接进入零阶自适应算术编码得到的编码结果。HC 扫描的平均码率较条带扫描的平均码率提高了 1.605%;新算法的平均码率较条带扫描的平均码率提高了 3.800%,可见,新算法的平均码率是 3 种扫描算法中最小的。

表 1 3 种扫描方式的编码结果

Tab. 1 The encoding results of three kinds of scanning methods

图像 (512 × 512)	3 种扫描方式编码结果 (Byte)		
	条带扫描	HC 扫描	新算法
Lena	73 033	71 967	71 007
Barbara	76 243	74 913	72 598
平均码率 (bpp)	2.277 771	2.241 211	2.191 238
(码率提高程度)	(0%)	(-1.605%)	(-3.800%)

4 上下文建模优化

下面将考虑如何在上述概率渐变的扫描方式上,尽可能用更小码长对扫描序列进行编码,即进一步讨论最小化式(4)的第 1 部分。理论上说,采用越高阶的最邻近空域或频域的上下文 $\mathbf{X}^{(i-1)}$ 的平均比特消耗 $H(X_i | \mathbf{X}^{(i-1)})$ 是最少的,然而实际编码中有限的样本与高阶上下文的矛盾则导致了所谓的“上下文稀释”的问题。为了缓解上下文稀释,由于具有相同或者相近概率特性的上下文可通过量化函数 G 被合并在一起,因此关于 $H(X_i | \mathbf{X}^{(i-1)})$ 的最小化就转化成关于 $H(X_i | G(\mathbf{X}^{(i-1)}))$ 的最小化问题,即上下文量化的优化问题。最新的上下文量化优化算法见文献[9]~[12]。下面首先对新扫描算法产生的上下文进行建模,再使用最小码长为目标函数进行上下文的优化分类。

4.1 上下文建模

直接使用图 8 的上下文模型作为原始的上下文空间不仅没有必要,也不切实际,因为上下文状态达 2^{60} 个之多,若用其作为上下文空间,则必然要耗费

巨大的内存空间去保存庞大的上下文映射表。本文结合纹理信息和能量信息,联合首轮和次轮扫描子过程来进行上下文建模,其所依赖的主要观察如下:

(1) 已知重要系数的 3×3 大小的邻域内出现的新的重要系数频率主要依赖 3×3 大小的邻域内的上下文信息,至于其 3×3 大小的邻域以外的上下文信息则可以忽略;

(2) 在当前比特平面之前就被确认的重要系数(称为过重要)距离最近的邻域周围出现重要系数的可能性较大;

(3) 若邻域内距离最近的已知重要系数大于 2,那么估计概率主要依赖距离最近上的已知重要系数的能量。

根据以上观察,每个扫描过程建立的上下文模型如下:

① 图 8 为距离 d 为 1 的首轮扫描 ($first_{run} == TRUE$) 建立的上下文模型。根据上述观察(1),当前待编码系数 d 为 1 的邻域距离内有已知重要系数,则当前待编码系数是重要的概率,其主要依赖 3×3 大小的邻域内的上下文信息,而在 3×3 大小的邻域以外的已知重要系数可以忽略,如此就可以大大减少上下文个数。下文中的 $\text{sig}(X)$ 表示对括号里的小波系数 X 求其在当前比特平面下的重要性。 3×3 大小的邻域内的上下文纹理信息主要是水平信

息 $\sum_{i=0}^1 \text{sig}(H_i)$ 、垂直信息 $\sum_{i=0}^1 \text{sig}(V_i)$ 、对角信息 $\sum_{i=0}^3 \text{sig}(D_i)$ 。考虑到观察(2),由于已知的过重要系数(即小波系数在比当前比特平面高的比特平面上就已被确定为重要系数)在其最近邻域周围出现新重要系数的可能性较大,因此对距离 d 为 1 的首轮扫描的上下文模型添加了过重要的上下文信息。其上下文模型如下:

$$C_1^{\text{first}} = \sum_{i=0}^3 \text{sig}(D_i) + 5 \times \left(\sum_{i=0}^1 \text{sig}(H_i) + H_{\text{over}} \right) + 20 \times \left(\sum_{i=0}^1 \text{sig}(V_i) + V_{\text{over}} \right) - 5 \quad (5)$$

其中, $H_{\text{over}} = \begin{cases} 1 & \text{如果 } H_0 \text{ 或 } H_1 \text{ 是过重要} \\ 0 & \text{其他} \end{cases}$

$V_{\text{over}} = \begin{cases} 1 & \text{如果 } V_0 \text{ 或 } V_1 \text{ 是过重要} \\ 0 & \text{其他} \end{cases}$

C_1^{first} 的可能取值是 $[0, 74]$,也就是说,一共有 75 个不同的状态,包含了绝大部分的纹理信息。属

于 $d=1$ 的次轮扫描 ($first_{run} = FALSE$), 由于没有过重要的邻域上下文, 故其上下文建模为以下 40 种不同上下文状态:

$$C_1^{second} = \sum_{i=0}^3 \text{sig}(D_i) + 5 \times \left(\sum_{i=0}^1 \text{sig}(H_i) \right) + 15 \times \left(\sum_{i=0}^1 \text{sig}(V_i) \right) - 5 \quad (6)$$

② 由于在 d 为 2 的子过程中, $\sum_{i=0}^1 \text{sig}(H_i) = 0$ 、 $\sum_{i=0}^1 \text{sig}(V_i) = 0$ 、 $H_{over} = 0$ 、 $V_{over} = 0$, 因此这 4 类上下文都不予考虑, 根据观察(1), 主要考虑对角信息 $\sum_{i=0}^3 \text{sig}(M_i = 2)$; 根据观察(3), 则考虑 d 为 3 的上下文的能量信息 $\sum_{i=0}^3 \text{sig}(M_i = 3)$, 而其余的 d 大于 3 的上下文信息则忽略不计。 d 为 2 的子过程首轮和次轮的上下文模型都有 20 种上下文状态, 即

$$C_2 = \sum_{i=0}^3 \text{sig}(M_i = 3) + 5 \times \left(\sum_{i=0}^3 \text{sig}(M_i = 2) \right) - 5 \quad (7)$$

其中, $M_i = 2$ 表示图 8 中标注为“2”的邻域位置点, 同理, $M_i = 3$ 表示图 8 中标注为“3”的邻域位置点。

③ 根据观察(3), 距离 d 为 3 的子扫描过程的首轮和次轮上下文模型采用距离 d 为 3 以及 d 为 4 的能量信息 $\sum_{i=0}^3 \text{sig}(M_i = 3)$ 和 $\sum_{i=0}^{11} \text{sig}(M = 4)$, 一共有以下 16 种上下文状态, 即

$$C_3 = -1 + \sum_{i=0}^3 \text{sig}(M_i = 3) + \sum_{i=0}^{11} \text{sig}(M = 4) \quad (8)$$

④ d 为 4 的首轮和次轮上下文模型主要考虑 d 为 4 的能量信息 $\sum_{i=0}^{11} \text{sig}(M = 4)$, 有以下 12 种上下文状态, 即

$$C_4 = -1 + \sum_{i=0}^{11} \text{sig}(M_i = 4) \quad (9)$$

⑤ $d=5$ 、 $d=6$ 的首轮和次轮概率没有明显区别, 由于其重要性概率都非常小, 估计概率主要依赖 d 的大小, 因此 $d=5$ 和 $d=6$ 以及 d 大于 6 以外情况的各自使用一个上下文。

由此可见, 一共使用了 $75 + 40 + 20 \times 2 + 16 \times 2 + 12 \times 2 + 1 + 1 + 1 = 214$ 种不同的上下文状态。这 214 个上下文状态主要由 155 个 3×3 大小的邻域内(包括 $d=2$)的上下文信息以及 59 个 3×3 大小的邻域外的上下文信息组成。在每个子过程的扫

描过程中, 要为每个待编码的系数计算所属的上下文索引, 即在式(5)~式(9)中, 选择满足所属子扫描的上下文模型公式来计算所属的上下文索引, 以便将原始上下文空间映射为这 214 个上下文状态中的某个状态。如距离 $d=1$ 的首轮扫描的上下文映射(式(5)), 需要执行 10 次加法, 2 次乘法; 而 $d=1$ 的次轮扫描的上下文映射(式(6)), 需要执行 8 次加法, 2 次乘法; 根据式(7), 距离 d 为 2 的子过程首轮和次轮的上下文映射需要执行 8 次加法, 1 次乘法; 根据式(8), 距离 d 为 3 的首轮和次轮的上下文映射需要执行 16 次加法; 根据式(9), 距离 d 为 4 的首轮和次轮的上下文映射需要执行 12 次加法; $d=5$ 和 $d=6$ 以及距离 d 大于 6 以外的情况, 由于各自使用一个上下文状态, 因此都不需要执行加法和乘法。实际上, 214 个上下文状态对于有限的样本而言, 仍然过多, 下面阐述使用最小化码长准则来对 214 个上下文状态进行优化分类。

4.2 最小化描述码长准则优化上下文量化

若上下文空间 C 是 d 维, 那么量化阶数为 M 的上下文量化 $Q(c)$ 可以描述为将原始的 d 维上下文空间划分为 M 个互不相交且和为原始 d 维空间的集合:

$$A_m = \{c | Q(c) = m\} \quad m = 1, \dots, M \quad (10)$$

这实际是一个矢量量化问题^[12]。若待编码的随机矢量为 X , 则根据熵函数的凸性, 量化后 $H(X|Q(C)) \geq H(X|C)$ 。因此, 若以最小相对熵为准则, 则量化的对象实际就成了与相应集合对应的具有凸集性质的概率分布函数^[12]。由于对于二进制随机变量 X 来说, 概率单形体实际是 1 维的, 因此就可利用标量量化技术或者动态规划来实现条件熵最小的上下文空间优化分类^[11]。由于条件熵总是随着阶数 M 的增加而下降, 但实际码长损耗并不是随着 M 增加而下降, 因此以最小码长为准则就成为最小条件熵之外的另一个选择。

本文使用上下文下 0 和 1 相应出现的频率来作为估计的概率(见式(11))。 $\{n_0, n_1\}$ 是 c 下 0 和 1 相应出现的次数。

$$\begin{cases} \hat{p}(0|c) = \frac{n_0 + \delta}{n_1 + n_0 + 2\delta} \\ \hat{p}(1|c) = \frac{n_1 + \delta}{n_1 + n_0 + 2\delta} \end{cases} \quad (11)$$

其中, δ 是小于或等于 1 的参数, 本文取 $\delta = 1/2$ 。根据文献[11], 对于随机序列 x_r , 理想的码长为

$$L(x_r) = - \sum_{i=0}^{r-1} \log(\hat{p}(x_i | c_i)) \quad (12)$$

上下文集合 A_m 下 0 和 1 的出现次数为 $\{N_0(A_m), N_1(A_m)\}$, $T_m = N_0(A_m) + N_1(A_m)$, 令 A_m 下的二进制序列与第 k 个出现的随机符号 $i (i=0, 1)$ 对应的出现次数为 $n_{i,k}$ 。根据文献[11], 那么该集合下的理想码长可证得为

$$\begin{aligned}
 L_m(N_0(m), N_1(m)) &= - \sum_{k=0}^{T_m-1} \text{lb} \left(\frac{n_{i,k} + \delta}{k + 2\delta} \right) \\
 &= \sum_{k=0}^{T_m-1} \text{lb}(k + 2\delta) - \sum_{k=0}^{T_m-1} \text{lb}(n_{i,k} + \delta) \\
 &= \sum_{k=0}^{T_m-1} \text{lb}_2(k + 2\delta) - \sum_{i=0}^1 \sum_{j=0}^{N_i(m)-1} \text{lb}_2(j + \delta) \\
 &= - \log \frac{\prod_{j=0}^{N_0(m)-1} (j + \delta) \prod_{j=0}^{N_1(m)-1} (j + \delta)}{\prod_{j=0}^{N_0(m)+N_1(m)-1} (j + 2\delta)} \quad (13)
 \end{aligned}$$

当令 $N_i(m) = 0$ 时, 则 $\prod_{j=0}^{N_i(m)} (j + \delta) = 1$ 。利用 Stirling's approximation 的 $n! \approx n^n e^{-n} \sqrt{2n\pi}$ 的码长为

$$\begin{aligned}
 L_m(N_0(m), N_1(m)) &\approx -N_0(m) \text{lb} \left(N_0(m) - \frac{1}{2} \right) - \\
 &N_1(m) \text{lb} \left(N_1(m) - \frac{1}{2} \right) + \left(N_1(m) + N_0(m) + \right. \\
 &\left. \frac{1}{2} \right) \text{lb} (N_1(m) + N_0(m)) \quad (14)
 \end{aligned}$$

$$\begin{bmatrix}
 L(1: 1 \rightarrow 1) & L(1: 2 \rightarrow 1) & \cdots & L(1: |C| - 1 \rightarrow 1) & L(1: |C| \rightarrow 1) \\
 0 & L(2: 2 \rightarrow 1) & \cdots & L(2: |C| - 1 \rightarrow 1) & L(2: |C| \rightarrow 1) \\
 0 & 0 & \ddots & \vdots & \vdots \\
 \vdots & \vdots & \ddots & L(|C| - 1: |C| - 1 \rightarrow 1) & L(|C| - 1: |C| \rightarrow 1) \\
 0 & 0 & \cdots & 0 & L(|C|: |C| \rightarrow 1)
 \end{bmatrix} \quad (17)$$

再根据式(17)计算 $M > 1$ 的分类消耗的最小码长(式(18)), 式(18)中的 X 表示不需理会。在对式(18)进行计算的过程中, 如果 $L(1: |C| \rightarrow K+1) \geq$

$$\begin{bmatrix}
 L(1: |C| \rightarrow 1) & X & X & X & X \\
 L(1: |C| \rightarrow 2) & L(2: |C| \rightarrow 2) & \cdots & L(|C| - 2: |C| \rightarrow 2) & L(|C| - 1: |C| \rightarrow 2) \\
 L(1: |C| \rightarrow 3) & L(2: |C| \rightarrow 3) & \cdots & L(|C| - 2: |C| \rightarrow 3) & 0 \\
 L(1: |C| \rightarrow 4) & L(2: |C| \rightarrow 4) & \cdots & 0 & 0 \\
 \vdots & \vdots & \vdots & \vdots & \vdots
 \end{bmatrix} \quad (18)$$

根据文献[11], 以最小条件熵为代价函数可满足“concave Monge property”, 以最小条件熵为代价函数的动态规划算法的计算代价为 $O(M|C|)$; 而最小描述码长准则却不满足“concave Monge property”, 其动态规划算法(式(18))的计算代价为 $O(M|C|^2)$ 。

其中, $\delta = 0.5$ 。合并上下文后的总体码长为

$$L = \sum_{m=1}^M L_m(N_0(m), N_1(m)) \quad (15)$$

那么, 最小化码长就是寻找一种分类方式 $Q(c) \rightarrow A_m, m = 1, \dots, M$, 使得 L 最小化。训练时, 可将统计获得的概率

$$\hat{p}(0|c) = \frac{n_0(c) + \frac{1}{2}}{n_0(c) + n_1(c) + 1}$$

按照升序排序, 而排序后, 所对应的上下文 $c_i, i \in [1, |C|]$ 则对应 $\hat{p}(0|c_i) \leq \hat{p}(0|c_{i+1})$, $|C|$ 表示上下文状态数。令 $q_{\min} = \min(\hat{p}(0|c)), q_{\max} = \max(\hat{p}(0|c))$, 那么就可可在区间 $[q_{\min}, q_{\max}]$ 上通过动态规划寻找 M 个子区间 $(q_{l-1}, q_l], l = 1, \dots, M$ 来使得码长 L 最小^[11]。令 $L_{\min}(i: |c| \rightarrow M)$ 表示将排序后对应的上下文 c_i 到 $c_{|C|-i+1}$ 一共 $|C| - i + 1$ 个上下文划分成 M 类所耗费的最小码长, 动态规划是依据以下关系来寻找最优方案:

$$\begin{aligned}
 L_{\min}(1: |C| \rightarrow M) &= \min_{i=1-|C|+M+1} (L(1: i \rightarrow 1) + \\
 &L_{\min}(i+1: |C| \rightarrow M-1)) \quad (16)
 \end{aligned}$$

若要自适应地获得与最小码长对应的阶数 M , 则需先计算以下矩阵

$L(1: |C| \rightarrow K)$, 则令最佳阶数 $M = K$, 并退出计算, 最佳分类方案可以根据式(18)回溯得到。

本文采用文献[13]的 10 个图像为训练图像, 比文献[3]少独立考虑比特平面 2^3 和 2^{r-1} , 初始概率均为 0.5, 本文将比特平面分成 5 类, 其中类 0, 1, 2 分别对应比特平面 $2^0, 2^1$ 和 2^2 ; 类 3 对应比特平面 2^3 到 2^{r-1} (r 是 MSB), 类 5 表示 2^r ; 子带分类与文献[3]相

同,分成 4 类。那么一共有 20 张上下文索引表,每张优化表是采用上述最小化描述码长准则来对 4.1 节提出的 214 个上下文状态进行优化得到的。

5 实验结果

本文使用了 10 个常用的测试图像对新方案的编码性能进行了实验测试,并将新的重要性编码方案和 JPEG2000 的上下文模型方案在相同的实验条件下进行对比。实验时,每个图像在进行 5 级 5/3 小波分解后, JPEG2000 采用条带扫描以及其上下文模型对子带图像进行比特平面编码;改进方案则采用前面提出的新扫描方法以及优化的上下文模型对子带图像进行比特平面编码。编码结果表明,新方案比 JPEG2000 的平均比特率提高了 1.312%,压缩的结果见表 2。

表 2 JPEG2000 以及新算法的重要性编码的编码结果以及平均比特率

Tab. 2 Code lengths in bytes and average bit rate in bits per pixel for significance coding for JPEG2000 and proposed scheme

测试图像	编码结果(Byte)	
	JPEG2000	本文方案
Lena	71 528	70 507
Barbara	72 697	71 331
Mandrill	80 519	79 307
Peppers	74 112	73 547
Tiffany	69 423	68 541
Sailboat	75 708	74 673
Airplane	66 954	66 245
Woman	1 454 380	1 432 041
Bike	1 471 551	1 454 357
Cafe	1 579 681	1 561 873
平均比特率(bpp)	2.246 770	2.217 299(-1.312%)

6 结 论

综上所述,图像进行小波变换后,还需要使用某种扫描策略将 2 维的矩阵转化成 1 维的序列进行熵编码。那么,如何扫描才可以产生自相关性最强的序列,以便使算术编码器能根据以往出现过的符号来对当前符号的概率进行估计? 本文提出了一种利用已知的重要上下文进行多次扫描的新扫描算法,该算法产生的序列具有概率波动小、自相关性强的优势,同时新扫描算法可以对具有不同重要系数聚簇特性的子过程进行不同的上下文建模,以便最后

采用最小描述码长准则进行上下文量化优化。就平均码率而言,新方案比 JPEG2000 在同等情况下平均比特率提高了 1.312%。

由于 JPEG2000 标准只是明确定义了解码一致的比特流语法结构,其编解码技术仍留有很大的空间。用本文提出的清除扫描代替 JPEG2000 中的 CP 扫描,其余从距离 1 到 6 的子扫描作为 JPEG2000 中的 ZC 扫描,以便下一步将新方案嵌入到 JPEG2000 的框架中。

参考文献(References)

- Charilaos C. The JPEG2000 still image coding system: An overview [J]. IEEE Transactions on Consumer Electronics, 2000, 46(4): 1103 ~ 1127.
- Taubman D. High performance scalable image compression with EBCOT [J]. IEEE Transactions on Image Processing, 2000, 9(7): 1158 ~ 1170.
- Krivoulets A, Wu Xiao-lin, Forchhammer S. On optimality of context modeling for bit-plane entropy coding in the JPEG2000 Standard [A]. In: Proceedings of International Conference on Very Low Bitrate Video Coding[C], Madrid, Spain, 2003: 208 ~ 216.
- Liu Zhen, Karam Lina J. Mutual information-based analysis of JPEG2000 contexts [J]. IEEE Transactions on Image Processing, 2005, 14(4): 411 ~ 422.
- Wilhelm B, Tobias B, Murkus L, et al. Fast variable run-length coding for embedded progressive wavelet-based image compression [J]. IEEE Transactions on Image Processing, 2001, 10(12): 1781 ~ 1790.
- Matias Y, Shamir A. A video scrambling technique based on space filling curves [A]. In: Proceedings of International Conference on Advances in Cryptology-CRYPTO'87[C], Santa Barbara, California, USA, 1987: 398 ~ 417.
- Witten I, Neal R, Cleary J. Arithmetic coding for data compression [J]. Communications of the ACM, 1987, 30(6): 520 ~ 540.
- Peng Ke-wu, Kieffer John C. Embedded image compression based on wavelet pixel classification and sorting [J]. IEEE Transactions on Image Processing, 2004, 13(8): 1011 ~ 1017.
- Wu Xiao-lin. Context quantization with fisher discriminant for adaptive embedded wavelet image coding [A]. In: Proceedings of IEEE International Conference on Data Compression[C], Snowbird, Utah, USA, 1999: 102 ~ 111.
- Wu Xiao-lin, Chou P A, Xue X H. Minimum conditional entropy context quantization [A]. In: Proceedings of IEEE International Conference on Information Theory[C], Sorrento, Italy, 2000: 43 ~ 43.
- Forchhammer S, Wu Xiao-lin, Andersen J D. Lossless image data sequence compression using optimal context quantization [J]. IEEE Transactions on Image Processing, 2004, 13(4): 509 ~ 517.
- Xu Man-tao, Wu Xiao-lin, Fränti Pasi. Context quantization by kernel fisher discriminant [J]. IEEE Transactions on Image Processing, 2006, 15(1): 169 ~ 177.
- CIPR Still Images [EB/OL]. <http://www.cipr.rpi.edu/resource/stills/index.html>, 2006-10-10.