

# JPEG2000 图像压缩过程及原理概述

刘方敏 吴永辉 俞建新

(南京大学计算机科学与技术系 南京 210093)

(南京大学计算机软件新技术国家重点实验室 南京 210093)

**摘 要** 随着多媒体技术的广泛应用,图像压缩界呼唤一种新的国际标准,以满足用户对更高压缩效率和对压缩图像的互动性和可伸缩性的要求, JPEG 2000 就是在这种背景下应运而生的.文中重点讨论 JPEG2000 PART 1 的图像压缩过程和对应的原理技术,说明它的几个重要特征是如何实现的;并对最近即将标准化的 JPEG2000 PART 2 作了一个简要的介绍,通过实验数据再次说明了 JPEG2000 的先进性.

**关键词** JPEG2000, EBCOT, 离散小波变换, 感兴趣区  
**中图法分类号** TP391

## Image Compression Process and Principle of JPEG2000

Liu Fangmin Wu Yonghui Yu Jianxin

(Department of Computer and Technology, Nanjing University, Nanjing 210093)

(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

**Abstract** As a new international standard for image compression JPEG2000 came into being to meet the users' newest demands of better performance and functionality. In this paper, the process, fundamental theories and involved technologies of image compression in JPEG2000 PART 1 are presented. Moreover, the realization of some important functions in JPEG2000 is also illustrated. At the end of this paper, the authors make a brief introduction to JPEG2000 PART 2, which will be standardized soon, and verify the advancement of JPEG2000 by experimental data.

**Key words** JPEG2000, EBCOT, DWT, ROI

### 1 引言

随着时间的推移,人们对图像压缩处理的要求越来越高,旧的标准(如 JPEG)难以适应最新的形势.在这种背景下,近几年来 JPEG 委员会的专家们一直致力于制订出一个全新的图形压缩国际标准,即 JPEG2000.该标准共有 8 个部分,其中 PART 1 在 2000 年 12 月份成为了国际标准,它的目标是提供一个最小化的无知识产权问题的 JPEG2000 可用系统;PART 2 目前仍处于 International Standard Under Publication 阶段,与 PART 1 比较,它采用的技术更加复杂,性能也更加优良,不过其中一些技术涉及到了知识产权问题(IPR),并且它的解码器并不能解码采用 PART 1 技术得到的码流;其余 6 个部分的标准化也在积极进行中,重点将放到诸如移动图像,标准的高质量实施,参考软件实施方案的提供,针对扫描传真等应用的图像格式,安全性,互动性以及传输协议,三维编码等问题上.

JPEG2000 的优秀之处并不仅仅体现在它对图像压缩性能提高上,更为重要的是,它提供了一种全新的图像再现形式.为了具体地理解这一点,先看一下 JPEG2000 PART 1 所支持的特征:

- (1)改进的压缩性能;
- (2)有损压缩和无损压缩的能力;
- (3)多解析度支持;
- (4)可嵌入的码流;
- (5)感兴趣区 (Region Of Interest, ROI);
- (6)对错误的鲁棒性;
- (7)对码流的随机访问处理;
- (8)灵活的文件格式;
- ……

这些特征的灵活运用,可以在达到高效压缩的同时,满足在移动和网络环境下对互操作和可伸缩性的要求<sup>[1]</sup>.而它们的成功实现则在于 JPEG2000 采用了众多的最新技术,如离

原稿收到日期: 2001-07-02;修改稿收到日期: 2002-05-30.刘方敏,男,1979年生,硕士研究生,主要研究方向为多媒体技术.吴永辉,男,1979年生,学士,工程师,主要研究方向为多媒体技术、数据库技术.俞建新,男,1953年生,高级工程师,主要研究方向为多媒体技术、计算机图形学.  
©1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

散小波变换 (Discrete Wavelet Transform, DW T), EBCO T (Embedded Block Coding with Optimal Truncation)编码算法,带 deadzone的量化器等. JPEG2000的需求针对性以及技术先进性保证了它光明的应用前景.

2 JPEG2000 PART 1的压缩过程及原理

图 1所示为一个典型 JPEG2000 PART 1的压缩过程.



图 1 JPEG2000 PART 1的压缩过程

2.1 数据预处理

PART 1的预处理一般包括三种操作: 区域划分,降低量级,分量变换.

2.1.1 区域划分

区域划分是指将图像划分为大小相等的若干区域,对每一区域独立进行压缩处理.区域划分的目的在于要降低压缩过程所需的内存资源,如果内存足够,这一步可以忽略. PART 1要求划分的区域是互不重叠的,因为这种划分是最简单的,但由此而产生的一个缺点就是边缘像素环境信息的缺乏,这个问题在 PART 2中得到了解决(见第 5节).

2.1.2 降低量级

降低量级是将采样精度为  $P$  的无符号整数减去  $2^{P-1}$ ,使原来范围为  $[0, 2^P - 1]$  的样本移位到  $[-2^{P-1}, 2^{P-1} - 1]$  这个关于 0 对称的范围内.这一步在简化对数值溢出等问题处理的同时,不会影响编码的效率.

2.1.3 分量变换

分量变换指对具有多个分量的图像先通过某种变换降低这几个分量之间的相关性,提高压缩效率.目前 PART 1中主要是对 RGB分量采用 ICT或 RCT,将色彩信息转换为频道信息.

ICT(Irreversible Color Transform)定义如下

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.16875 & -0.33126 & 0.500 \\ 0.500 & -0.41869 & -0.08131 \end{pmatrix} \times \begin{pmatrix} R \\ G \\ B \end{pmatrix}.$$

RCT(Reversible Color Transform)定义如下

$$Y = \left\lceil \frac{-R + 2G + B}{4} \right\rceil, U = R - G, V = B - G.$$

解压缩时需进行相应的逆变换以恢复原分量的值.

2.2离散小波变换 (DWT)

2.2.1 离散小波变换的一般过程

预处理后的数据将进行离散小波变换,以进一步降低数据之间的相关性.与 JPEG采用的离散余弦变换(DCT)相比, DWT具有很好的局部性,能够针对不同类型特点的图像中的不同区域采用不同的空频分辨率,从而有可能取得更好的压缩比,而且它还可以提供实现无损压缩的机制.

简单来讲,一维 DWT即是对源信号进行了一系列的高通和低通滤波,并在每次滤波后将数据采样频率降为原来的一半,以保证每次小波变换后得到的系数与源信号数目相同.每次的低通滤波输出保存了源信息的低频信息,它是一个以更低分辨率对源信号的再现,集中了源信号中的大部分能量;而高通滤波输出保存的则是源信号的高频信息,如边界、材质等,其中所含能量很少.一次低通滤波后的信号往往还存在着大量的相关性,为提高压缩性能,仍需要对它再次滤波,直至信号之间相关性达到可以忽略的程度为止.高通滤波后的信号由于能量很小,再对它进行滤波往往是不划算的,因而一般不再对它滤波.这种滤波方式(如图 2所示)被称为 dyadic 分解,它是 JPEG2000 PART 1唯一支持的分解方式.

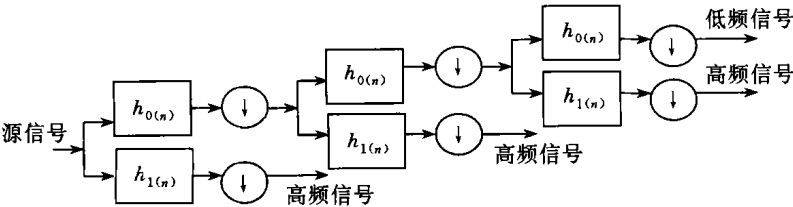


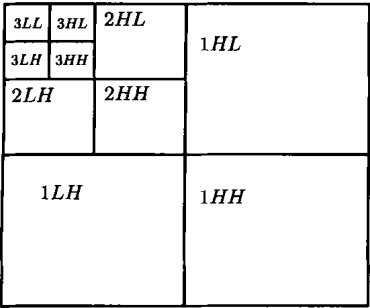
图 2 一维 DWT(↓在图中代表采样频率减半)

二维 DWT是对一维 DWT的简单扩充,通过将行信号和列信号与高通滤波器  $h_1(n)$ 和低通滤波器  $h_0(n)$ 进行不同的组合,源图像被划分为 4 个子带.其中唯一的一个低频子带仍可以继续分解,形成如图 3所示的结构.

有几个细节问题需要注意: (1) 对边界信号进行小波变换前,由于缺乏周边信号的信息,需先对信号进行扩展.具体的扩展方式受到滤波器的影响,一般采取对称扩展的方式(如图 4所示). (2) DWT 内在的提供了一个解决多分辨率

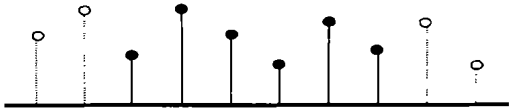
问题的机制,对一个  $N$  层小波分解的图像来说,它可以提供  $N+1$  种分辨率供用户选择.因为每层分解中的  $kLL$  都是对源图像的一个低分辨率再现. (3) DWT是针对整块区域的,而 DCT是针对某个  $8 \times 8$  块的,所以当压缩率超过一定限度时, JPEG 2000可以消除 JPEG不可避免的“马赛克”现象.关于 DCT和 DWT的差异,有一个形象的比喻:同是压缩一堵墙, DCT是将每块砖敲碎了重新组装再放回原处,而 DWT是将整堵墙敲碎了再重新组织. (4) JPEG2000 PART 1所支持

的滤波器组只有两种,一种是具有较高有损压缩性能的浮点型的(9,7)滤波器组,另一种是可同时满足有损无损压缩需求的整数型的(5,3)滤波器组(无损压缩的介绍见第2.2.2节).



注: $kHL(k=1,2,3)$ 代表在第 $k$ 层分解中对行信号进行高通滤波,对列信号进行低通滤波后间隔采样的结果

图3 二维DWT



注:实线为源信号;虚线是扩展的信号

图4 对一维信号的对称扩展

2.2.2基于卷积的滤波和基于提升小波的滤波

JPEG2000支持两种滤波方式:基于卷积的滤波和基于提升小波的滤波.基于卷积的滤波是传统方法,它将扩充后的源信号与 $h_0(n), h_1(n)$ 作基于内积的点运算,其缺点在于无法即时用离散小波变换系数替换对应点的源信号数据,因此要占用更多的内存;而基于提升小波的滤波可完全消除这一缺点,它的基本思想是先将源信号划分为奇数信号集合 $\{d_i^0\}$ 和偶数信号集合 $\{s_i^0\}$ ,然后对它们交替进行预测和更新的操作,依次得到 $\{d_i^1\}, \{s_i^1\}, \{d_i^2\}, \{s_i^2\}, \dots$ 其中 $\{d_i^n\}, n=0,1,2, \dots, N$ 保存的是源信号的高频滤波系数,  $\{s_i^n\}$ 保存的是低频滤波系数,并且数值上与相应的基于卷积的滤波系数相等.预测和更新操作的公式分别为

$$d_i^n = d_i^{n-1} + \sum_k P_n(k) s_{i-k}^{n-1}, n \in [1, 2, \dots, N],$$
$$s_i^n = s_i^{n-1} + \sum_k U_n(k) d_{i-k}^n, n \in [1, 2, \dots, N].$$

其中,  $P_n(k), U_n(k)$ 分别为预测和更新操作的权重,  $N$ 代表某次奇数(偶数)信号的获取所涉及的上一层的偶数(奇数)信号的个数,它的数值因所采用的滤波器族的不同而有所差异, JPEG2000 PART 1中 $N$ 取值为2.图5所示为 $N=2$ 时的一个基于提升小波的滤波过程.

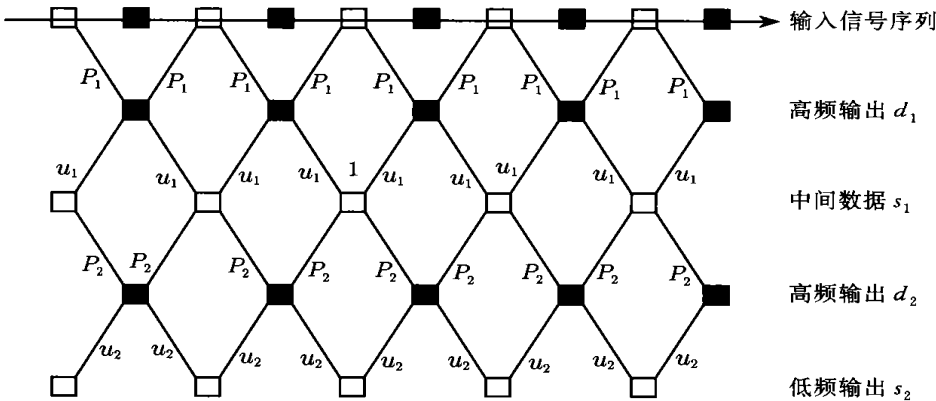


图5 基于提升小波的滤波过程(相当于两层分解的基于卷积滤波)

基于提升小波的滤波的另一个特点就是计算简单,不论是分解或重构都不涉及复杂的内积运算,而且它还提供了无损压缩的能力.对于基于卷积的滤波,即使是采用整数型的(5,3)滤波器组,随着分解层数的加深,由于计算机无法为完整的表示滤波系数提供足够的精度,从而出现信息丢失;而基于提升小波的滤波则可以和量化器结合提供一个整数到整数的压缩框架,以实现无损压缩<sup>[2]</sup>.

一个可能的基于提升小波滤波的无损压缩系统的构成为正变换

$$y(2n+1) = x(2n+1) - \left\lfloor \frac{x(2n) + x(2n+2)}{2} \right\rfloor,$$
$$y(2n) = x(2n) + \left\lfloor \frac{y(2n-1) + y(2n+1) + 2}{4} \right\rfloor.$$

逆变换

$$x(2n) = y(2n) - \left\lfloor \frac{y(2n-1) + y(2n+1) + 2}{4} \right\rfloor,$$

$$x(2n+1) = y(2n+1) + \left\lfloor \frac{x(2n) + x(2n+2)}{2} \right\rfloor.$$

这个系统得到的滤波系数全为整数,正、逆变换简单,并且逆变换可完整地恢复出正变换之前的结果,因而可实现无损压缩.

2.3 量化

JPEG2000的量化与JPEG量化基本相同:总体上都是采用均匀量化;不同子带的量化步长一般不同. JPEG2000的量化器的一个特殊之处在于它引入了一个“deadzone”的概念,在PART 1中,对每个子带的量化器,deadzone的宽度都是其它步长的2倍,如图6所示.这意味着如果采用步长为 $\Delta_b$ 的这种量化器并得到长为 $M_b$ 的量化索引值,那么在解码过程中,我们可通过选用步长为 $\Delta_b \times 2^{M_b - M_0}$ 的量化器来恢复原来的索引值最重要的前 $M_0$ 位.这样做的优点就是为我们提供信噪比分级的手段:先用较小的步长对信号进行细致的量化,然后依据用户的需求,采用不同的宽步长(一般为量化步长的2

的幂次方倍)进行逆量化,仅解码原索引中重要性较高的若干位,以提供给用户不同质量的图像。

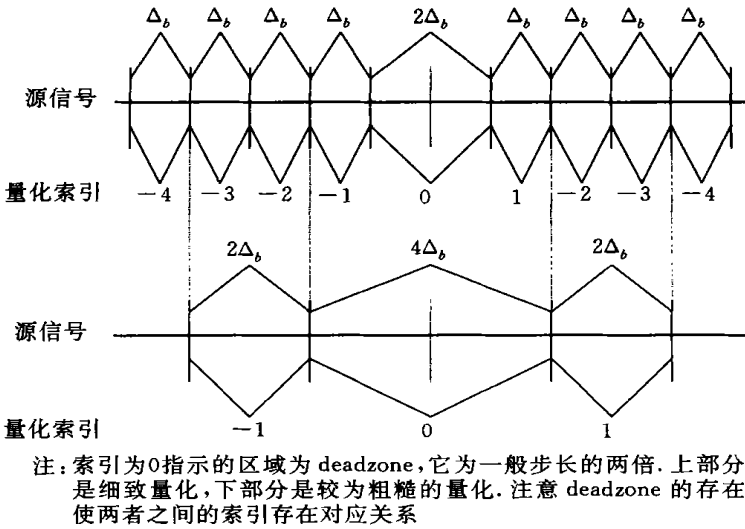


图 6 步长为  $\Delta_b$  的量化器

JPEG2000量化器的另一个不同在于,解码时量化索引的逆量化值可取量化器允许范围中的某个值而不是仅局限在中值点.如果取值策略正确,将有助于提高解码性能。

量化器步长是因子带而异的,目前有两种考虑因素:一种是人类视觉对子带信号的敏感性(HVS的属性);另一种是依据不同子带的均方误差对重建后图像的总均方误差的贡献大小来决定量化步长.量化器步长的传输也有两种方式:一种是类似于 JPEG 中的  $q$ -table,显式地传给解码器;另一种仅传输某子带的量化器步长  $\Delta_b$ ,其它子带步长由  $\Delta_b$  计算出。

2.4 自适应算术编码(第一层编码)和  
码流组织(第二层编码)

2.4.1 EBCOT算法

本小节介绍 JPEG2000的另一个核心内容——优化截取的嵌入式块编码(EBCOT)算法,它是基于小波变换的嵌入式编码的方法之一.所谓“基于小波变换的嵌入式编码”是指编码器将等待编码的、经过小波变换后的比特流按重要性不同进行排序,提供多个满足不同目标码率或失真度的截断点,使得解码器方能根据目标码率或失真度的要求在某一截断点结束解码,提供相应质量的图像.当前这种编码方法很多,如 EZW 算法, SPIHT 算法等.那么, JPEG2000为何选择 EBCOT 呢?

JPEG2000非常强调的一点是灵活性,因为它能保证用同一个图像满足不同用户的需求.多分辨率支持和多失真度支持都是灵活性的必然要求.多分辨率支持可通过 DWT 来实现,多失真度支持则可通过位平面编码(位平面指将所有数据相同位的数据抽取出来组织到一起)来解决.但是,将这两种技术结合起来并不能达到“1+1=2”的效果,原因在于位平面编码过程中往往用到了源信号数据之间的某种相关性来提高编码效率,即将不同位平面或者不同位置上相关性较强的数据组织到一起进行编码.这将和多分辨率支持所要求的数据按子带排列的要求产生冲突,结果是虽然可能取得

了多失真度的支持,但无法用同一码流取得多分辨率的支持, EZW 和 SPIHT 算法均存在着这一问题.而 EBCOT 中则不会出现这一问题,因为它将源图像分为若干大小相等的块,对每块独立进行编码,如果想要得到某分辨率下的图像就仅仅解码该分辨率下对应的块;如果想要得到某失真度下的图像就仅仅解码每块中对应应该失真度的部分.它可能产生的缺点是由于没有利用到源信号数据块之间的相关性消除冗余而造成压缩性能降低.事实表明,由于块编码可以对每一块进行独立优化,因此这种操作带来的压缩性能的提高可完全补偿上述损失。

在进行块编码时, JPEG2000 强调多截断点的支持,越多的截断点,表明图像可提供更多的质量选择.如果我们对每块仅仅进行位平面编码,那么对于数据最高位数为  $N$  的块,最多可得到的截断点有  $N$  个.很多时候这种截断是粗糙的而且截断点数目过少.为了获得更多的截断点, EBCOT 引入“片段化位平面”的概念,将编码的逻辑单位进一步细化,对每一个位平面再进行 3 次扫描,得到 3 个“片段化位平面”,它们是对原位平面的一个无重复划分.这样对某块  $B$  来说,可能的截断点可以有  $3N$  个(实际有  $3N - 2$  个,原因见第 2.4.2 节)。

对码流的组织, EBCOT 也有专门的论述.因为若没有合理的组织,单纯的嵌入式编码并不能保证对多失真度的支持.图 7 所示为一种最简单的码流组织方式,它的码流虽然是可嵌入的,但并不能支持多失真度。

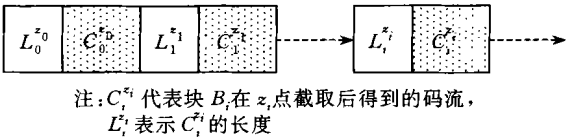


图 7 码流的一种最简单的组织方式

图 7 所示码流失败的原因在于它没有在码流的组织中加入有关图像质量分级的辅助信息,而 EBCOT 中采用了第

二层编码的概念对该信息的提供进行了规范化.

2.4.2 自适应算术编码

JPEG2000 PART 1的编码方法采用自适应算术编码而非 JPEG常用的霍夫曼编码.自适应算术编码与霍夫曼编码最大的不同在于它并不是对每个信号产生一个码值,而是对一个信号序列产生一个码值;它除了需要获得信号序列外,还要获得其中每个信号的预测概率.根据这个概率,将 [0, 1) 区间不断分割,最后得到的码值即是最后分割得到的小区间所对应的一个二进制表示.自适应性的获得在于不断更新信号的预测概率,使之总是趋近于实际.

位平面中点  $(u, v)$  的信号对应预测概率的获得需要考虑该点周围各点的信号分布以及该点在高位平面的信号分布.在 EBCOT 中,这些分布情况是由显著位来指示的.显著位在初始时(即对某块最高位平面编码时)全置为 0,当编码时得到非 0 数据时,它将立即由 0 变为 1.在编码的某个时刻显著位  $e(u, v)$  标志该点已编码数据的值若为 0,则  $e(u, v) = 0$ ;若大于 0,则  $e(u, v) = 1$ .

下面介绍对某块进行自适应算术编码所需的 3 次扫描中的具体内容(扫描的次序如图 8 所示,扫描每列高为 4 个像素,这也是块高度的最小值).

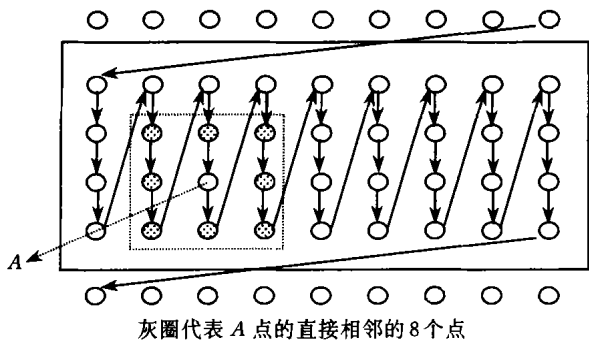


图 8 编码的扫描次序

(1) “显著性传播”扫描

本次扫描的编码对象是当前状态位为 0,且其直接相邻 8 个系数的状态位至少有一个为 1 的样本,即在当前位平面中最有可能成为“显著”(显著位为 1)的那些“不显著”样本.根据文献 [3],它们单位数据包含的信息量最大,而 EBCOT 的编码原则是将最有价值的信息优先编码,使得在文件大小一定的情况下取得最小的失真度.这些样本直接相邻的 8 个样本的状态位作为“上下文”提供给编码器,编码器根据“上下文”选择一种预测概率对其进行算术编码<sup>[4]</sup>.  $e(u, v)$  要即时反映  $(u, v)$  点最近被编码位的变化,它将影响下面待扫描点的“上下文”,进而影响编码,这也是本次扫描名称中“传播”的由来.另外,在某样本变为显著时,要对它的正负号立即进行编码,编码过程类似于上述的数值编码.

(2) “数值范围重定义”扫描

对当前状态位为 1,但没有被第一次扫描编码过的样本进行编码,这些样本的高位字节是“显著”的,包含的信息量也较大.编码过程类似于“显著性传播”扫描.

(3) “清除”扫描

将所有剩下的未编码的样本进行编码.由于本次编码的

数据都是“不显著”的,所以可采用一种称为“run 模式”的编码模式.其原理是将一列 4 个样本作为一个整体进行编码,如果这 4 个样本同是“不显著”的(即都属于本次扫描),就可以仅用一个二进制符号来代表;如某列中至少含有一个“显著”样本,在标记了第一个“显著”样本的位置后,立即停止“run 模式”,改用前两次的编码方式进行编码.当在编码过程中某样本成为“显著”时,同样需对其进行正负号编码<sup>[5]</sup>.

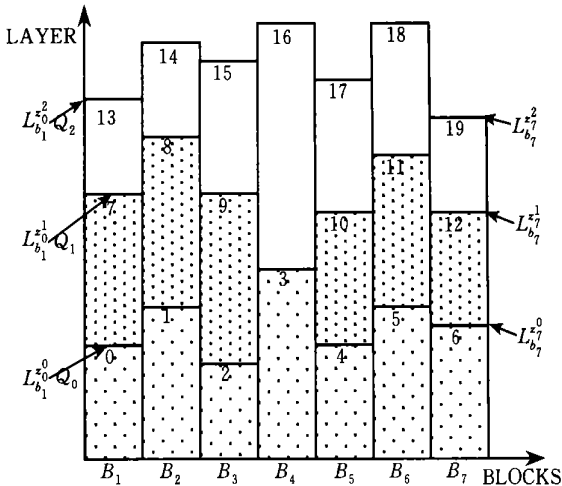
由于在初始化编码时,所有的状态位均被置为 0,因此没有前两次扫描.这样,对量化索引最高位为  $N$  的块,可能的截断点有  $3N - 2$  个.

2.4.3 码流的组织(第二层编码)

以上三种扫描的次序是固定的,它反映的核心思想是把位平面上不同位置的样本按其包含的信息量分级,然后依次编码.这种编码方式是码流可支持多失真度的基础,但是采用这种编码方式得到的码流并不保证一定支持多失真度的图像再现,图 7 所示为一个反面的例子.

JPEG2000 采用层来组织码流,它的功能是提供有关图像质量分级的信息.第一层  $Q_0$  摘取了每一块中的某一部分,并满足当总长度受限的情况下,各块失真度之和最小,即在  $\sum_i l_i^0 \leq l_{\max}^0$  时,  $D^0 = \sum_i D_i^0$  最小( $z_i^0$  指块  $i$  对应  $Q_0$  层的截断点;  $l_i^0$  指块  $i$  中对应  $Q_0$  层的部分长度;  $D_i^0$  指块  $i$  中对应  $Q_0$  层的部分失真度).而接续的某层  $Q_n$  则取每一块中的长为  $l_i^n - l_i^{n-1}$  的附加部分,并满足在  $\sum_i l_i^n \leq l_{\max}^n$  情况下,  $D^n = \sum_i D_i^n$  最小.

图 9 所示为一个 3 层图像码流的组织.



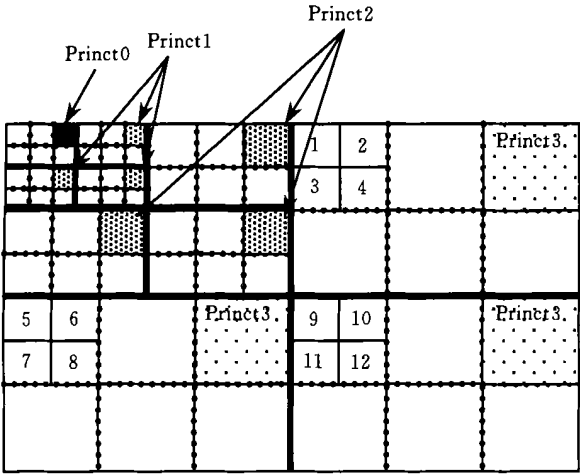
注:  $B_i$  代表某块,数字代表对应码流的排列顺序;  
 $Q_1$  完全不需要  $B_4$  的贡献

图 9 3 层图像码流的组织

JPEG2000 为了更好地表达这种分层的思想,又引入了几种中间逻辑结构,即 Princt 和 包. Princt 指某一分辨率下空间某连续区域在所有子带中对应块的集合;而包则包括由某个 Princt 中所有块的一次或者几次扫描后得到的码流中的一个连续片段,它是编码的逻辑单位,并采用包头封装了

必要的逻辑信息.而层就是以包为单位组成的.为了在不同的分辨率下明确指示同一块空间区域, JPEG2000采用了 Canvas 坐标系,它在实质上与分形编码中的码表是一样的.在这种坐标系中,空间的矩形区域是由该矩形左上角的坐标和它的长宽来指定的,不同分辨率下的同一块空间区域的左上角坐标和长度可以相互换算.例如,在分辨率为  $64 \times 64$  的图像中,左上角顶点坐标  $(x_0, y_0)$  为  $(31, 31)$ , 长宽  $(x_{size}, y_{size})$  为  $(32, 32)$  的矩形区域,在分辨率为  $32 \times 32$  时  $(x_0, y_0)$  为  $(15, 15)$ ,  $(x_{size}, y_{size})$  为  $(16, 16)$ .除了定位空间区域外, Canvas 坐标系还有利于裁剪操作,因为对坐标原点以上部分和以左部分的裁剪,并不影响剩余部分的坐标,无需重新编码.

图 10所示为 Princt和块、子带之间的关系.



注:粗线代表不同子带的边界,虚线代表不同 Princt 的边界,实线代表不同块的边界;数字代表块在 Princt 中的排列顺序

图 10 块、Princt、子带之间的关系

综上所述,包需要 4 个参数来指定,即分量 (Component), 分辨率 (Resolution), 层 (Layer), Princt. 这样,包的组织就可以有若干种顺序,如 Layer-Resolution-Component-Princt, Resolution-Princt-Component-Layer, Component-Princt-Resolution-Layer 等,而每一种组织顺序意味着对某类应用更好的支持.如 Component-Princt-Resolution-Layer 的组织顺序对图像库的应用是很有利的,因为可以方便地得到不同质量的图像.与 JPEG 标准相比,这种以包为基础的编码和解码方式非常灵活,从一种应用到另一种应用只需对包进行重新排列,并不需要重新解码、逆变换这样的复杂操作.

3 JPEG2000 PART 1的一些重要功能

本小节着重介绍基于以上压缩过程的 JPEG2000 一些重要功能及其实现.

3.1 多失真度支持

多失真度支持的实现方法如下:

方法 1. 量化时采用一个很小的步长,对变换系数进行最细致的量化,而在逆量化时采用不同的宽步长(一般为量化步长的 2 的幂次方倍),舍弃量化索引中不重要位上的信

息,满足不同用户对图像质量的要求.

方法 2. 充分利用 EBCOT 中层的概念,对某一特定目标码率,仅解码相关层所含的包,从而达到在一定文件大小的限制下,对图像的最佳再现.实施这一方法的前提是层组织的正确性, JPEG2000 采用拉格朗日 R-D 优化方法确保了这一点.第 2.4.3 节对层的描述说明的即是该方法的一般过程,具体实现见文献 [3].

3.2 感兴趣区 (ROI)

在实际应用中,用户可能需要图像中的某些特定区域能以较其它部分更高的质量进行编码、解码,或者在图像显示过程中对这些区域实行优先显示, JPEG2000 通过提供一种称为感兴趣区的功能来满足用户的这种需求.

ROI 的压缩过程如下: (1) 确定图像中用户感兴趣的区域. (2) 用 ROI 掩膜指明这些区域.在背景区域,其值为 0;在 ROI 区域,其值为 1. (3) 在编码前,将 ROI 内的系数加倍 (或将背景系数减倍) 以提高 ROI 内数据 (或降低背景数据) 编码精度.其余过程同第 2 节所述.在解压缩时,要将 ROI 区域信息以及系数加倍 (或背景系数减倍) 信息传递给解码器,以恢复原系数值.

JPEG2000 具体实现 ROI 采用的 Maxriff 方法与上文略有不同,它先找到背景区域变换系数的最高位数  $s$ ,而后将所有背景系数除以 2,确保即使 ROI 内系数最小的值也比背景系数大,使得解码时先解码 ROI 内系数,再将背景系数扩大 2 倍后进行解码.它的优点在于无需显式地传递 ROI 区域信息给解码器.

3.3 容错性

JPEG2000 最大的容错性是由 EBCOT 提供的. EBCOT 采用的独立编码确保某块内的错误不会扩散到其它块中,并且在进行某“片段化位平面”扫描前,“上下文”对应的概率分布都要重置,以确保上一个“片段化位平面”扫描的错误不会影响到下一次扫描的结果 [6].

除此之外, JPEG2000 还从码流组织上提供了容错的手段.如对某位平面编码, JPEG2000 规定在每一次 Cleanup 扫描过后,编码器都应在其后加入“1010”标志;解码时如未能正确取得此标志,说明该位平面的码流中存在错误.同时在第二层编码中, JPEG2000 对每一个包进行了记数,确保包头和包的数据能够正确对应起来,提高鲁棒性.

3.4 文件格式

每种图像文件格式的主要内容就是对颜色空间的定义和元数据的嵌入. JPEG2000 PART 1 中对此作了非常灵活的规定.

PART 1 提供了一种称为 JP2 的文件格式.它对颜色空间的定义有两种方式:一是强调互动性,仅提供了两种颜色空间,即 RGB 和灰度级;另一种则强调灵活性,提供了一种描述颜色空间变换的数据结构,允许用户自定义空间.对元数据的嵌入也有两种定义方式:采用 UUID (Universal Unique Identifier) 的方式和采用 XML 的方式.无论采用哪一种方式,元数据在文件中嵌入的位置都是随意的.

4 JPEG2000 PART 2

由于 JPEG委员会考虑到计算复杂度、编码效率以及知识产权的问题 (JPEG委员会希望 JPEG2000 PART 1没有涉及知识产权问题),因此很多技术并没有放到 PART 1中,而是被吸收到在 2001年 12月份成为国际标准的 PART 2中.总地来讲, PART 2是对 PART 1的一个深入和提高. PART 2中与 PART 1相关的一些技术特点简介如下:

(1)数据预处理

在 PART 1的区域划分中,划分出来的空间区域是不能重叠的;在 PART 2的区域划分中,划分出来的空间区域是可以重叠的,虽然这样会带来计算上的复杂度,但它消除了边缘信息的缺乏.

在 PART 1的量级降低中,针对某一个区域划分出来的区域,变换系数仅仅允许减去  $2^{-1}$  ( $s$  为该区域中变换系数的最高位);在 PART 2中,这一数值更加一般化,并不局限于某个值.

在 PART 1的分量转换中,仅允许对 RGB分量进行 ICT和 RCT两种转换;在 PART 2中,允许对多分量进行变换,并且提供了对多分量去除相关的一般方法.

(2)量 化

在 PART 1的量化中,各子带的量化器的结构都是一样的 (deadzone的宽度都是其它索引位置步长的 2倍);在 PART 2的量化中,各子带的 deadzone的宽度可以是其它索引步长的  $x$  倍 ( $1 \leq x \leq 2$ ),以更好地体现子带之间的差异性.

(3) DCT

在 PART 1的 DCT中,仅对  $kLL$  ( $k= 0, 1, 2, \cdots, N$ )继续分解以降低相关性;在 PART 2中,其它子带 ( $kLH, kHL, kHH, k= 0, 1, 2, \cdots, N$ )也可视需要深入分解.

在 PART 1中允许的滤波器族仅有两种:可回复的 (5, 3)滤波器族和不可回复的 (9, 7)滤波器族;在 PART 2中用户自定义的滤波器族是允许的.

(4) ROI

在 PART 1中的 ROI由于采用了 Maxriff方法,因此不能够支持 ROI与背景系数的多对比度;而在 PART 2中,通过显式地向解码器提供 ROI相关信息,允许 ROI与背景的多对比度.

(5)格 式

PART 1提供的格式为 JP2,而 PART2中提供的格式为 JPX,它提高了对颜色空间和元数据定义的灵活性,并可在一个文件中包含多幅不同图像 (类似于动态 gif).除此之外,它还定义了众多的元数据域来说明图像的历史、内容、特征、知识产权等信息.

5 JPEG2000和多级树集合分裂算法 (SPIHT)的性能比较

由于采取了众多的先进技术,理论上 JPEG 2000应提供更好的性能和更多的功能,下面通过几组对比数据加以验证.用作参考的 SPIHT算法是业界一个成熟的比较好的压缩算法,而且它同样采用离散小波变换.实验采用的压缩性能的度量是信噪比 (PSNR),

$$PSNR = -10 \lg \left( \frac{255^2}{MSE} \right).$$

JPEG2000采用 JPEG委员会提供的 JPEG2000 Verification Model的 C语言实现 (<http://www.jpeg.org>), SPIHT的实现则采用一个实际中广泛使用的系统 (<http://www.cipr.rpi.edu>):它包括 3种单色调编解码器 (实验中使用的是压缩性能最高的一个)和一种彩色图像编解码器.

为保证结果的客观性,实验采用 Waterloo Brag Zone推荐的标准测试图像系列 (<http://links.uwaterloo.ca/>),它将图像分为 3类: Grayset1(小灰度图像), Grayset2(中灰度图像), Colorset(彩色图像).对每一类图像,考虑到大小、自然、非自然等因素,抽取两个代表进行测试,如表 1所示.

表 1 不同目标码率下实验数据

类 属	图 片	编解码器	0.125 bpp	0.25 bpp	0.5 bpp	1.0 bpp
Grays et 1	Goldhill(N) (256× 256× 8)	JPEG 2000	24.68 dB	26.48 dB	29.03 dB	32.25 dB
		SPIHT	25.40 dB	27.10 dB	29.52 dB	32.86 dB
	Lena(N) (256× 256× 8)	JPEG 2000	24.59 dB	27.82 dB	31.68 dB	36.59 dB
		SPIHT	26.11 dB	28.96 dB	32.74 dB	37.67 dB
Grays et 2	Barb(N) (512× 512× 8)	JPEG 2000	25.05 dB	27.82 dB	31.42 dB	36.56 dB
		SPIHT	25.38 dB	28.13 dB	32.11 dB	37.45 dB
	Mountain(N) (640× 480× 8)	JPEG 2000	17.90 dB	19.15 dB	21.02 dB	24.10 dB
		SPIHT	18.05 dB	19.37 dB	21.44 dB	24.64 dB
Colorset 1	Frymire(A) (1118× 1105× 24)	JPEG 2000	15.52 dB	17.18 dB	19.63 dB	23.23 dB
		SPIHT	15.57 dB	17.38 dB	19.97 dB	23.55 dB
	Monarch(N) (768× 512× 24)	JPEG 2000	26.00 dB	29.23 dB	33.25 dB	37.91 dB
		SPIHT	27.01 dB	30.25 dB	34.22 dB	38.98 dB

注: N表示自然图像, A表示人工图像, dB为 PSNR单位

为获得对同一图像的不同目标码率再现, JPEG2000对同一码流进行了不同的截取;并且 JPEG2000提供多分辨率支持,通过对包的重排,将原来为失真度渐变的图像变为分

辨率渐变的图像.这些功能是 SPIHT无法提供的,所以在功能上, JPEG2000已是“稍胜一筹”了.

(下转第 916页)

[2] Wu Yonghui, *et al.* Research for biological manufacturing engineering based on rapid prototyping[J]. Chinese Journal of Mechanical Engineering, 2000, 36(9): 61~ 64(in Chinese) (吴永辉,等.基于快速成型的生物制造工程研究[J].机械工程学报,2000,36(9): 61~ 64)

[3] Zollikofer C P E, *et al.* Tools for rapid prototyping in the biosciences[J]. IEEE Computer Graphics and Applications, 1995, 15(11): 48~ 55

[4] Barequet G, *et al.* Multilevel sensitive reconstruction of polyhedral surfaces from parallel slices[J]. The Visual Computer, 2000, 16(2): 116~ 133

[5] Ateshian G A, *et al.* A B-spline least-squares surface-fitting method for articular surfaces of diarthrodial joints[J]. Journal of Biomechanical Engineering, 1993, 115(4): 366~ 373

[6] Ke Yinglin. Study on the technology and applications of scattered data based geometry modeling [D]. Nanjing Nanjing University of Aeronautics & Astronautics, 1992(in Chinese) (柯映林.离散数据几何造型技术及其应用研究[博士学位论文].南京:南京航空航天大学,1992)

[7] Farin G. Curves and Surfaces for Computer Aided Geometric Design[M]. Boston Academic Press, 1988

[8] Wang Guozhao. Discrete intersection algorithm of B<sup>2</sup>zier curve/surface[J]. Journal of Zhejiang University, 1984(Special Issue: Computational Geometry): 108~ 119(in Chinese) (汪国昭. B<sup>2</sup>zier曲线曲面的离散求交方法[J].浙江大学学报,1984(计算几何专集): 108~ 119)

[9] Li Jijun, *et al.* The trimming of composite triangular B<sup>2</sup>zier surface[J]. Journal of Computer-Aided Design & Computer Graphics, 2000, 12(1): 65~ 69(in Chinese) (李际军,等.复合三角 B<sup>2</sup>zier曲面的裁剪[J].计算机辅助设计与图形学学报,2000,12(1): 65~ 69)

[10] Piegl L A. Fundamental Developments of Computer-Aided Geometric Modeling[M]. New York: Academic Press, 1993

[11] Ma Weiyin, *et al.* Parameterization of randomly measured points for least squares fitting of B-spline curves and surfaces [J]. Computer-Aided Design, 1995, 27(9): 663~ 675

[12] Greiner G, *et al.* Data dependent thin plate energy and its use in interactive surface modeling [J]. Computer Graphics Forum, 1996, 15(3): C175~ C185

[13] Terzopoulos D, *et al.* Dynamic NURBS with geometric constraints for interactive sculpting [J]. ACM Transactions on Graphics, 1994, 13(2): 103~ 136

(上接第 911页)

在压缩性能方面,绝大部分情况下 JPEG2000与 SPIHT的最优实现的表现相差无几,考虑到 SPIHT采用了最新的小波变换技术,可以认为 JPEG2000的压缩性能比当前大多数压缩算法高.同时,这也说明了 EBCOT 中的独立块编码技术带来的优点可完全弥补没有充分利用数据位置相关性所带来的损失.

## 6 结 论

JPEG2000是当今图像压缩界的一个新标准,它对一些新兴的图像压缩领域,如遥感、扫描、数字照相、医疗图像等提供了一个全新的解决问题的框架.这个框架的先进性和光明前景正为事实所证明.目前,在 JPEG2000的具体实施上,尚未出现大量的、免费的、高效的工具软件的支持.相信随着时间的推移,这个问题将得到逐步地解决.

## 参 考 文 献

[1] C Christopoulos, T Ebrahimi, A Skodras. The upcoming JPEG2000 standard[A]. In: Proceedings of Invited Tutorial to the 11th Portuguese Conference on Pattern Recognition (RECPAD 2000), Porto, Portugal, 2000. 359~ 366

[2] Majid Rabbani, Rajan Joshi. An overview of the JPEG2000 still image compression standard [J]. Signal Processing Image Communication, 2002, 17(1): 3~ 48

[3] David Taubman. High performance scalable image compression with EBCOT [J]. IEEE Transactions on Image Proceedings, 2000, 9(7): 1158~ 1170

[4] E Ordentlich, D Taubman, M Weinberger, *et al.* Memory efficient scalable line-based image coding [A]. In: Proceedings of the IEEE Data Compression Conference, Snowbird, 1999. 218 ~ 227

[5] David Taubman, Erik Ordentlich, Marcelo Weinberger, *et al.* Embedded block coding in JPEG2000 [J]. Signal Processing Image Communication, 2002, 17(1): 49~ 72

[6] ISO /IEC FCD15444-1-2000. JPEG2000 image coding system [S]