一种基于 SPIHT 算法的感兴趣区域编码新方法*

李晓飞[1]，马大玮[1,2]，范小麟[1]

(1. 重庆通信学院 研究生管理大队，重庆 400035；2. 重庆大学 通信工程学院，重庆 400044)

摘　要：在研究 JPEG2000 中的感兴趣区域(ROI)编码算法优缺点的基础上，结合嵌入式编码理论和人眼视觉特性，利用等级树分集(SPIHT)编码算法的特性，提出了一种高效的 ROI 编码方法，压缩后的码流具有嵌入式特点，支持渐进传输。实验结果表明：尽管在低码率下，该算法整体峰值信噪比较低，但主观视觉效果明显好于 SPIHT 算法，结果令人满意。

关键词：感兴趣区域；SPIHT 算法；小波变换；人眼视觉特性；渐进传输

中图法分类号：TP919.81　　**文献标识码**：A　　**文章编号**：1001-3695(2007)02-0189-03

New Region of Interest Coding Method Based on SPIHT Algorithm

LI Xiao-fei[1], MA Da-wei[1,2], FAN Xiao-lin[1]

(1. *Graduate Management Team, Chongqing Communication College, Chongqing* 400035, *China*; 2. *College of Communications Engineering, Chongqing University, Chongqing* 400044, *China*)

Abstract: In this paper, both the merits and defects of the Region Of Interest (ROI) coding algorithm in JPEG 2000 standards are investigated at first. Then, according to the characteristics of the embedded bitplane coding and the human visual system, a highly efficient ROI coding algorithm based on the Set Partitioning in Hierarchical Tree (SPIHT) algorithm is proposed. By using this algorithm, the compressed bitstreams are all embedded and suitable for progressive transmission. Finally, experiments show that the new algorithm provides much better effect of the subjective visual sensation and has satisfactory results, although the PSNR performance of the proposed algorithm at low rates is inferior to the original SPIHT algorithm.

Key words: Region Of Interest(ROI); SPIHT Algorithm; Wavelet Transform; Human Visual System; Progressive Transmission

近年来，随着多媒体技术和 Internet 的应用与发展，促使人们寻求更高压缩比的图像压缩编码方法。由于基于感兴趣区域的编码方法能够充分利用人眼的视觉特性，已成为图像编码的研究热点。

在远程医疗、图像检索及无线通信等应用领域，人们往往只对一幅图像中的部分区域感兴趣，这部分区域被称为感兴趣区域(Region Of Interest, ROI)。因此，可以对 ROI 采用低压缩比的有损压缩甚至无损压缩；而对背景区域(Back-Ground, BG)则采用高压缩比的有损压缩，这样能解决压缩比和重建图像质量的矛盾，这就是 ROI 编码技术。

新一代静止图像国际压缩标准 JPEG2000 正式将 ROI 编码技术纳入其中，并规定了两种标准算法，即基于比例移位法(Scaling-based Method)和最大位移法(Maxshift Method)。

本文在研究 JPEG2000 的两种标准 ROI 编码算法优缺点的基础上，结合嵌入式编码理论，利用 SPIHT 图像编码算法，提出了一种高效的 ROI 编码方法，改善了主观视觉效果，压缩后的码流还具有嵌入式的特点，支持渐进传输。

1　JPEG2000 的 ROI 编码算法分析

JPEG2000 中的基于比例移位法，采取预先选择一个适当的比例因子 S，在编码时将 BG 系数的幅值均缩小 2^S 倍，从而使得 ROI 系数优先被编/解码。但仍有一些 ROI 系数会与 BG 系数一起编码、传输和解码，这取决于比例因子 S 的大小。因此在渐进解码中，比例因子 S 起到了控制 ROI 和 BG 相对质量的作用。最大位移法可以看作上述方法的特殊情形，经过平移一个特定值使得缩小后的 BG 系数最大值小于 ROI 系数的最小值，即使 ROI 比特平面与 BG 比特平面无重叠。因为比例因子 S 不是预先选定的，因此算法存在无法调节 ROI 与 BG 相对重要程度的缺点。

上述两种方法的移位过程如图 1 所示，其中 MSB 是最重要的比特平面，LSB 是最不重要的比特平面。

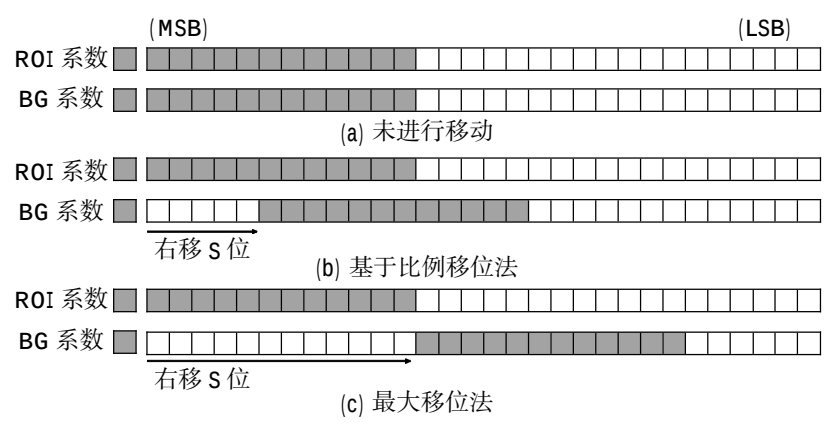


图 1　JPEG2000 中定义的标准 ROI 编码

在实际工作中，基于比例移位法需要导出一个含有 ROI 形状信息的比特掩模，并利用其标明每一个子带中哪些节点坐标上是 ROI 系数。由于码流置入了 ROI 的坐标信息将使比特流增大，解码器仍需再次求出 ROI 掩模，因此增加了计算的复

收稿日期：2005-11-30；修返日期：2006-01-12

基金项目：重庆市自然科学基金资助项目(CSTC2005BB2206)

杂度和存储要求。而最大移位法中 ROI 系数的最小值大于 BG 系数最大值,故无须传送 ROI 的坐标信息,解码器也无须计算 ROI 掩模。

2 SPIHT 图像编码算法

SPIHT 算法是 Said 和 Pearlman 在嵌入式零树小波算法(Embedded Zerotree Wavelet, EZW)基础上提出的一种高效的图像编码算法。SPIHT 算法充分考虑了图像经过小波分解后,能量集中在最高分解级的低频子带及各子带之间具有空间自相似性等特点,在继承了零树思想的基础上,设计了一种新的空间方向树结构来更有效地组织小波系数,使每个树包括了同一空间位置所有子带中的小波系数。同时通过对树的划分,将尽可能多的非重要系数汇集在一个子集中,用一个单位符号表示。在 SPIHT 编码过程中,将小波的重要信息保存在三个列表中:不重要像素列表(LIP)存放不重要小波系数的位置坐标;重要像素列表(LSP)存放重要小波系数的位置坐标;不重要集合列表(LIS)存放不重要集合的位置坐标。

SPIHT 采用 2 的幂次方作为初始阈值,每次扫描后阈值减半,经过多次扫描重要系数是小波变换系数整数部分的二进制表示,这是一种典型的比特平面编码。由于 SPIHT 采用这种从最重要的比特平面开始每次编码一个比特平面的编码方法,可以得到完全嵌入的具有连续可分级特性的输出码流,有利于实现精确的码率控制,支持渐进传输。

3 基于 SPIHT 的 ROI 编码

3.1 初始算法

本文结合 JPEG2000 的 ROI 编码算法思想,将 SPIHT 算法应用于包含 ROI 的图像编码中,其算法步骤如下:

(1) 指定 ROI 并确定该区域点的坐标集;

(2) 用小波变换对整个图像进行塔式分解;

(3) 确定 ROI 区域,生成 ROI 掩模;

(4) 将 ROI 系数平面左移 S 位,即 ROI 系数乘以因子 2^S;

(5) 按照 SPIHT 算法思想,对处理后的系数分别进行 ROI 编码和非 ROI 编码,循环编码至用户定义的阈值比特平面 R。

解码过程为编码的逆过程。

3.2 改进的 ROI 编码机制

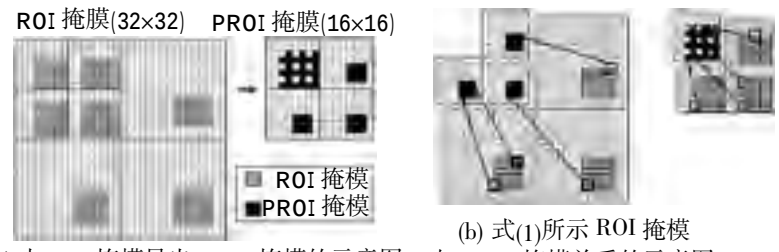
3.2.1 优化 ROI 掩模

依照上文提出的 ROI 编码算法,在执行重要性检测之前,均需要判断每个节点是否对重建 ROI 有贡献,以便将 ROI 系数传给解码器。如果节点不是 ROI 系数,暂时跳过节点检测,此后仍需重复执行此过程。然而,每个节点不管其子孙中是否含有 ROI 系数都要执行重要性检测。如果其子孙中不存在 ROI 系数,则表明它们对于重建 ROI 是毫无意义的。

因此,我们定义 PROI 掩模来记录那些子孙中至少存在一个 ROI 系数的节点坐标。此时若空间方向树中的某个节点不属于 PROI 系数,编码器将不再检测其子孙,从而提高了编码效率。除了最高频子带以外,ROI 掩模与 PROI 掩模的关系可以表示为

$$\mathrm{PROI}(i,j) = \bigcup_{(k,l)\in O(i,j)} \mathrm{ROI}(k,l) \qquad (1)$$

其中,PROI(·)和 ROI(·)分别表示 PROI 系数和 ROI 系数的坐标集合;$O(i,j)$ 表示节点(i,j)四个儿女的坐标集合,∪表示逻辑或运算。PROI 掩模的导出如图 2 所示。



(a) 由 ROI 掩模导出 PROI 掩模的示意图　(b) 式(1)所示 ROI 掩模与 PROI 掩模关系的示意图

图 2　PROI 掩模的导出

3.2.2 分割 LIP 和 LIS

在引入 PROI 掩模后,编码器首先执行节点检测以判断节点是否为 ROI 系数,如果不是将转入子孙检测;如果子孙检测判断出节点不是 PROI 系数,则跳出 ROI 编码。因此,LIP 保存了三种类型的记录:重要的非 ROI 系数、不重要的非 ROI 系数和不重要的 ROI 系数。在接下来的比特平面编码中,排序过程只对不重要的 ROI 系数进行检测;而非 ROI 系数无论重要与否都被保留在 LIP 中。当 ROI 编码结束后,编码器从最大阈值开始对 LIP 中的记录进行重复扫描。

在以上的 ROI 编码过程中,以不同阈值扫描所产生的多种记录都被大量添加至 LIP 中,使得编码效率降低。因此,根据添加记录时的阈值 2^n 将 LIP 分割成 LIP(n),将 LIS 分割成 LIS(n)。当编码器达到用户定义的比特平面 R,ROI 编码结束。编码器开始从最大阈值比特平面 $n_{\max}$到 R 执行非 ROI 编码,其具体过程为:当执行到阈值比特平面 n 的重要性检测时,只需对所有当 $n\le p\le n_{\max}$时的 LIP(p) 和 LIS(p) 执行检测。当非 ROI 编码结束后,由于所有节点已进行了重要性检测,故编码器将 LIP(n) 合并成为一个 LIP;LIS(n) 合并成为一个 LIS。

3.3 改进的 ROI 编码算法流程

为了更准确地描述 ROI 编码算法,引入如下两个定义:

定义 1　给定节点(i,j)的系数为 $c_{i,j}$,阈值为 2^n,定义重要函数 $S_n(i,j)$ 为

$$S_n(i,j) = \begin{cases} 1, & |c_{i,j}\ge 2^n| \\ 0, & \text{其他} \end{cases} \qquad (2)$$

定义 2　节点坐标的集合:

$O(i,j)$ ={节点(i,j)的儿女坐标}

$D(i,j)$ ={节点(i,j)的所有子孙坐标}

$H(i,j)$ ={所有根节点的坐标}

$L(i,j) = D(i,j) - O(i,j)$ ={除去儿女后节点(i,j)的所有子孙坐标}

此外,将 LIS 中的两个集合 $D(i,j)$ 和 $L(i,j)$ 分别归类为 A 类和 B 类。具体的算法流程描述如下:

(1) 初始化

令 $n = \lfloor \log_2(\max_{(i,j)}\{|c_{i,j}|\}) \rfloor$,并传送 $n_{\max} = n$;

令 LSP = ϕ, LIP = $\{(i,j) \mid (i,j)\in H\}$, LIS = $\{D(i,j) \mid (i,j)\in H\}$。

(2) 排序过程

(2.1) 对每个$(i,j)\in$LIP(p),其中 $n\le p\le n_{\max}$:

若$(i,j)\in$ROI,输出 $S_n(i,j)$;

若 $S_n(i,j)=1$,将(i,j)移到 LSP 中,输出 $c_{i,j}$的符号。

(2.2) 对每个$(i,j)\in$LIS(p),其中 $n\le p\le n_{\max}$:

①若(i,j)为 A 类

若$(i, j) \in$ PROI, 输出 $S_n(D(i, j))$。

若 $S_n(D(i, j)) = 1$, 则对每个$(k, l) \in O(i, j)$ 且$(k, l) \in$ ROI:

若 $S_n(k, l) = 1$, 将(k, l) 移到 LSP 中, 输出 $c_{k,l}$的符号;

若 $S_n(k, l) = 0$, 将(k, l) 加到 LIP(n) 末尾。

若 $L(i, j) \neq \phi$, 将(i, j) 作为 B 类移到 LIS(n) 末尾并转到②步。

若 $L(i, j) = \phi$, 将(i, j) 从 LIS(p) 中移出。

若$(k, l) \notin$ ROI, 将(k, l) 加到 LIP(n) 末尾。

②若(i, j) 为 B 类

若$(i, j) \in$ PROI, 输出 $S_n(L(i, j))$, 若 $S_n(L(i, j)) = 1$:

将每个$(k, l) \in O(i, j)$ 作为 A 类移到 LIS(n) 末尾;

将(i, j) 从 LIS(p) 中移出。

(3) 细化过程

对每个$(i, j) \in$ LSP, 除了在(1) 中最后一次排序加入的那些点以外, 若$(i, j) \in$ ROI, 输出 $|c_{i,j}|$的第 n 个最重要的比特。

(4) 循环过程

令 $n = n - 1$, 返回(2), 直到 $n = R$。

4　实验结果

实验采用 Lena(512 ×512, 8bpp) 标准测试图像(图 3), (9/7) 小波, 五级小波变换, $S = 6$, $R = 4$, 选定的 ROI 形状为矩形, 其左上角和右下角顶点分别为(200, 200) 和(370, 390)。此种情况下, ROI 为位于两顶点间的人脸区域, 易于导出掩模。

(a) 原始图像　　(b) 标定 ROI 的原图

图 3　测试图像 Lena(512×512, 8bpp)

整个图像的实验比较结果如图 4 所示, ROI 的实验比较结果如图 5 所示。表 1 为整个图像在不同码率下的编码结果比较, 表 2 为 ROI 在不同码率下的编码结果比较。

(a) SPIHT 算法结果(0.1bpp)　(b) 本文算法结果(0.1bpp)　(c) SPIHT 算法结果(0.5bpp)　(d) 本文算法结果(0.5bpp)

图 4　整个图像的实验比较结果

(a) SPIHT 算法结果(0.1bpp)　(b) 本文算法结果(0.1bpp)　(c) SPIHT 算法结果(0.5bpp)　(d) 本文算法结果(0.5bpp)

图 5　ROI 的实验比较结果

表 1　整个图像在不同码率下的编码结果比较

码率(bpp)	整个图像的 PSNR(dB)	
	SPIHT 算法	本文算法
0.1	29.772	28.415
0.5	37.092	36.905

表 2　ROI 在不同码率下的编码结果比较

码率(bpp)	ROI 的 PSNR(dB)	
	SPIHT 算法	本文算法
0.1	28.605	34.403
0.5	36.302	46.993

从上面的实验结果可以看出:

(1) 当码率为 0.1bpp 时, 虽然 SPIHT 算法编码重建得到的恢复图像整体质量较好, 但人脸区域的细节信息丢失较多, 脸部较为粗糙, 睫毛出现黏连; 而本文算法的恢复结果中, 背景区域虽然模糊, 但人脸区域却保持了较高的清晰度。尽管图 4(b) 的 PSNR 不高, 比图 4(a) 要低 1.357dB, 但作为 ROI 的人脸区域却比 SPIHT 算法中高出 5.798dB。因此图 4(b) 的整体视觉质量要明显好于图 4(a)。

(2) 由于本文算法保持了 SPIHT 算法中的嵌入式特点, 使得当码率提高为 0.5bpp 时, 产生的恢复图像整体质量只低 0.187dB, 背景区域效果两者几乎相同。同时, 当 ROI 编码执行到用户定义的比特平面 R 后, 移位后 ROI 系数的大部分比特平面已完成编码, 所以图 4(d) 视觉质量很高, 与原始图像效果相当。而 SPIHT 算法得到人脸区域的 PSNR 与本文算法差别达 10.691dB, 因此图 4(c) 的主观效果稍差。

5　结论

本文提出的编码方法充分考虑了人眼的视觉特性。编码时, 结合 JPEG2000 中的 ROI 编码算法思想, 将 ROI 和 BG 加以区分, 对 ROI 优先进行编码并组织码流, 获得了有效的编码效果。同时, 压缩后的码流具有嵌入式特点, 支持渐进传输。该算法是一种高效的 ROI 图像编码算法, 易于嵌入基于小波变换的图像编/解码方法中, 在医学诊断图像、网络传输及无线通信等领域, 该算法具有参考和实用意义。

参考文献:

[1]　ui-Sung Kang, Ho-Joong Choi, Sung-Jea Ko. Progressive Region of Interest Coding Using an Improved Embedded Zerotree Wavelet Coding[C]. Korea: Proceedings of the IEEE, 1999. 609-612.

[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]　D Nister, C Christopoulos. Lossless Region of Interest Coding[J]. Signal Process, 1999, 78(1): 1-17.

[4]　Charilaos Christopopulos, *et al*. Efficient Methods for Encoding Regions of Interest in the Updating JPEG2000 Still Image Coding Standard[J]. IEEE Signal Processing Letters, 2000, 7(9): 247-249.

[5]　K H Park, H W Park, Region-of Interest Coding Based on Set Partitioning in Hierarchical Trees[J]. IEEE Trans. Circuits Systems Video Technol., 2002, 12(2): 106-113.

[6]　A Said, W Pearlman. A New, Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees[J]. IEEE Trans. Circuits Systems Video Technol., 1996, 6(3): 243-250.

[7]　沈兰荪, 卓力, 田栋, 等. 视频编码与低速率传输[M]. 北京: 电子工业出版社, 2001. 85-105.

[8]　吴乐南. 数据压缩原理与应用(第 2 版)[M]. 北京: 电子工业出版社, 2003. 406-413.

作者简介:

李晓飞(1982-), 男, 山东菏泽人, 硕士研究生, 主要研究方向为图像压缩、图像处理等; 马大玮(1972-), 男, 四川盐亭人, 教授, 博士, 主要研究方向为图像通信、视频压缩处理等; 范小麟(1982-), 男, 重庆人, 硕士研究生, 主要研究方向为数据压缩与编码等。