

小波图像编码器研究进展^{*}

Wavelet Image Coder Research Evolution

周兵 李波 李炜

(北京航空航天大学计算机系 北京100083)

Abstract The research of wavelet coders has been being active since Mallat found the fast algorithm of wavelet transform. The novel compression standards JPEG2000 and MPEG-4, both using wavelet transform, will be finished in this year. There are two main factors which have relations with the performance of the wavelet image coders. One is the construction and selection of the Wavelet basis; the other is the quantization and coding of AC coefficients. After introducing the architecture of a general wavelet image coder, this paper discusses the selection of wavelet basis and the quantization of AC coefficients. At last, the paper introduces a few image coders which are suitable for the internet applications.

Keywords Wavelet, Image coder, Zerotree coder, Vector quantization, JPEG, JPEG2000

1 概述

根据解码后数据是否可以全部恢复,图像压缩可以划分为无损编码和有损编码。有损编码产生的图像还原以后,得到的图像与原图像存在一定的误差,但视觉效果一般是可以接受的。根据香农定理,无损压缩有一定的极限,所以当前图像编码研究主要集中在有损编码。根据编码原理,信息编码可以划分为预测编码、变换编码、信息熵编码、子带编码、分形编码、结构编码和基于知识的编码^[19,20]。

预测编码是针对统计冗余进行压缩。统计冗余反映在图像中,表现为相邻像素之间存在相关性,因而一个像素可以由与其相邻并且已经编码的像素来进行预测,最后编码传输的数据是量化后的差值,常用的预测编码方法有 PCM、DPCM 和 ADPCM 等。

变换编码也是针对统计冗余进行压缩。基本思想是通过变换来消除图像中存在的高度相关性,首先对空域图像作线性变换得到一组变换系数,然后对这组系数进行量化、编码和传输。典型的变换有离散 Fourier 变换(DFT)、离散余弦变换(DCT)、KL 变换(KLT)、Walsh 变换(WT)、斜变换(SLT)、Haar 变换以及小波变换。

信息熵编码根据信息熵原理,用短的码字表示出现概率大的位串或符号,常用的统计编码有 Huffman 编码、游程编码和算术编码。

子带编码是将图像数据变换到频域后,按频率分频带,然后用不同的量化器进行量化,从而达到优化组合。子带编码利用滤波器组,通过重复卷积的方法,经亚采样将输入信号分解为高频分量和低频分量,然后分别对高频和低频分量进行量化和编码。解码时,高频分量和低频分量经过插值和共轭滤波合成原信号。

分形编码利用图像内部子块之间的自相似性,在编码时将信号分解为若干分形子图,提取其迭代函数系统代码(IFS 代码),然后编码传输这些 IFS 代码,在恢复时由该代码重构各子图。

结构编码首先将图像中的纹理、边缘和轮廓结构特征提取出来,然后分别对它们进行编码。解码时,根据这些结构信息进行合成,从而恢复出原图像。

基于知识的编码在编码时通过各种分析手段,建立一定的模型来描述待编码的图像,同时提取模型的特征与状态参数并进行编码。解码时,依据这些参数,利用模型及相关知识恢复信源。

综合考虑多种因素,如去除信息相关的能力、算法的速度和实现复杂性以及与人类视觉的相似性等,在各种编码方案中基于小波的图像编码器具有较明显的优势,已成为近10年来静态图像编码研究的热点,其中零树小波编码已在 MPEG-4 中用于静态纹理对象的编码,基于小波变换的位平面编码已作为 JPEG2000 的图像压缩标准。

小波变换从子带变换发展而来,但具有更多的优越性,它不但为多分辨率分析、时-频分析和子带编码建立了统一的分析方法,而且提供了更合理的表示框架。

2 小波图像编码器的结构

静态图像的相关性既可能表现在像素的空间位置上,也可能表现在图像区域的空间结构上,信息压缩的前提就是信号之间存在相关,目的在于通过某些变换可以消除这些相关。

DCT 变换是经典图像编码器最广泛使用的 Fourier 类型的变换,它将时域或空域描述的信号变换到正交矢量空间(频域)中描述,在频域中这些信号存在很少的相关或者不存在相关。经过 DCT 变换后,原来分散的能量集中到少数低频系数,在允许一定误差时,可以忽略系数小于某个值的频率分量,从而使位率降低。

小波变换比 Fourier 类型的变换具有更好的时-频局部化特性,信号经小波变换后的能量也集中在少数变换系数上,但小波变换比 DCT 变换产生的变换系数更易于编码,不但能够解决低位率下的图像方块效应,而且能获得相同位率下更高的压缩效率。

^{*} 本课题受国家自然科学基金(69974005,69935010)、北京市科技新星计划和高等学校骨干教师资助计划资助。李波 教授,博士生导师。周兵 副教授。

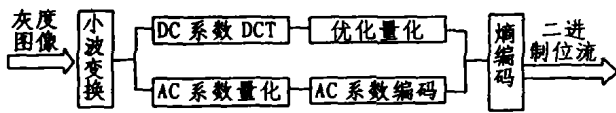


图1 典型小波编码器结构框图

自1989年 Mallat 将小波变换引入图像编码应用以来,基于小波的图像编码一直是最活跃的研究方向,并已纳入 ISO 标准 JPEG2000 和 MPEG-4。图1是小波图像编码器的典型系统结构。系统首先利用多级离散小波变换将原图像划分为低频分量(又称 DC 系数)和分别对应不同方向的高频分量(又称 AC 系数);接着对高、低频系数根据人类的视觉生理特性分别进行不同策略的量化和编码,一般对低频分量采用快速 DCT 变换,然后经过优化的非均匀量化后,再传送到熵编码模块形成压缩的二进制位流;对高频系数一般先进行均匀量化,然后再使用零树和位平面编码思想并结合熵编码生成压缩的二进制位流。

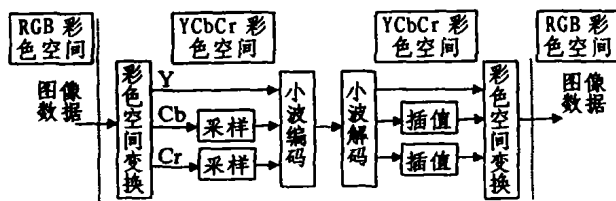


图2 彩色图像编、解码结构框图

在计算机中大多数彩色图像采用 RGB 彩色空间,但 RGB 三分量的细节、明暗度与原图都非常接近,如果采用三分量分别编码的方法,那么编码的总开销近似于编码一个分量开销的三倍,也近似于将原图转化为灰度图像后其压缩开销的三倍,可见使用 RGB 空间压缩彩色图像很难达到很高的压缩效率。而 YCbCr 彩色空间将原图的亮度信息和色度信息分离开, Y 分量中保存的是原图像的亮度信息,相当于原图的一个灰度级表示,而 CbCr 分量保存的是原图像的色度信息,内部不具有明显的明暗差别,也不具有原图像明显的轮廓。Y 分量的压缩开销相当于将原图像转化为灰度图像压缩的开销,而 Cb、Cr 分量可以在充分利用其相关性的基础上,使其压缩开销远远低于 Y 分量。因此使用 YCbCr 彩色空间来压缩图像,其效率远远高于 RGB 彩色空间。图2为彩色图像编、解码结构示意图。

自小波图像编码器出现以后的近十年来,基于小波的静态图像编码研究主要集中在小波的构造与选择、AC 系数的量化与编码器两个方面,前者偏重于基础理论研究,而后者主要是编码技巧的探讨,包括标量量化、矢量量化、网格矢量量化、零树方法、小波包方法、时-频局部方法、嵌入块独立编码等。

3 图像编码器的小波选取

3.1 小波与滤波器族

图像编码常用的小波有正交小波,双正交小波和小波包等 2-D 可分离离散小波,它们能通过相应的一维小波和尺度函数来构建。

正交小波是小波家族中产生 $L^2(\mathbb{R}^n)$ 正交基的小波,其中对图像编码最重要的一类小波是具有紧支撑正交小波。在离散小波变换中,紧支小波对应于有限脉冲响应(FIR)滤波器,

因而能够更有效地实现。Mallat 首先提出离散小波变换的快速算法, Daubechies 给出了构造紧支小波的系统化方法。Daubechies 小波和 Coifman 小波是最广泛使用的紧支小波。通常用一个与小波滤波器的长度成正比的整数表征小波,对紧支小波而言,小波滤波器的长度与小波的光滑度和正则性有关,光滑度和正则性可能对编码器的性能产生影响。

紧支小波的缺点在于它的非对称性,非对称性导致相关的 FIR 滤波器产生非线性相位。在使用非线性相位小波滤波器和有限长度数据计算 DWT 时,采用周期性包围“wraparound”扩展法可能在小波子带边缘处引起走样,如果采用线性相位小波和对称性“flip-over”延拓^[5],可以避免这些走样(除 Haar 小波以外)。只有放弃小波的紧支性和正交性才能获得小波滤波器的对称性。

双正交小波是正交同时具有紧支性和对称性的小波,大多数情况下采用双正交小波是因为它的对称性。当使用双正交小波时,计算 DWT 的镜像滤波器(QF)不再是正交的,然而它们与反向 DWT 用到的另一对滤波器是正交的,但这并没有丝毫增加计算的复杂性和实现的困难,所以很多小波图像编码器采用它。Villasenor 等人对双正交小波滤波器进行了大量的研究^[6]。

小波包是一组超完备的多分辨率正交或双正交基,小波包基可由生成小波的同样的 QF 对来生成。从子带编码的观点来看,任何与满子带树拥有相同根节点的子树,都代表由同一个 QF 对产生的小波包基。很明显,可根据某些准则从完备的小波基中选择一个最好基。Coifman 等^[4]提出了基于熵的最好基选择算法,该算法收敛于一个最小熵基。不过,Coifman 的熵指的是对系数能量压缩的测量,而不是一般定义下的熵。另一个选择最好基算法是由 Ramchandran 等人提出的在率失真意义下的最好基选择算法。如果主要关心有损压缩,则在给定率失真最小化的最好基似乎要优于 Coifman 的最小熵基。采用最好小波包基的理论基础是,以花费函数来评价,它至少与基于小波基的分解一样好。

但是,使用小波包基需要付出一定代价,首先,需要一些附加位来编码优化的基结构;另外,基于小波包的图像编码器在编码和解码两个阶段所花费的时间不对称,编码的速度要慢一些,因为要搜索最好基,所以计算代价更高一些。

多小波是指使用多个小波进行张成和平移共同产生正交 $L^2(\mathbb{R}^n)$ 基,与单个小波相比,正交多小波滤波器的长度可能更短,具有更多的消失矩,和更好的对称性,近来有文献报道用多小波对图像进行分析。

3.2 小波基的选取

任何实正交的小波对应的滤波器族均能实现图像的分解与合成,然而对同一幅图像,用不同的小波基进行分解所得到的数据压缩效果是不同的。我们希望经小波分解后得到的三个方向的细节分量具有高度的局部相关性,而整体相关性被大部分甚至完全解除。由于小波变换是将原始图像与小波基函数以及尺度函数进行内积运算,1989年 Daubechies 基于离散滤波器迭代的方法构造了紧支集的规范正交小波,因而内积运算转换为信号和离散滤波器的卷积运算,小波变换中的小波基的选择转换为正交镜像滤波器 QMF 的选择,对小波基的选择应考虑以下因素^[1]:(1)小波基的正则性阶数;(2)待处理图像与小波基的相似性;(3)滤波器长度和计算复杂度;(4)小波变换的分解层数(变换级数);(5)小波变换的能量集中特性好;(6)小波变换的边界问题。

研究表明,小波的正则性对小波图像压缩的性能有重要影响。在对变换系数进行量化时,可能导致误差,压缩比越高则误差越大,这些误差在重建滤波器后表现为重建图像的失真,失真的大小除与变换系数误差的大小成正比外,还与重建滤波器的单位脉冲响应有关,重建滤波器的正则性越差,则重建图像的变化就越不光滑,对人眼的刺激就越大,因而视觉效果就越差。实验表明,小波函数正则性 r 越高,图像压缩后重建的效果越好。设 FIR 滤波器的长度为 L ,则对长度 $L \leq 10$ 的情况,图像压缩后重建效果随 r 的增加而改善的速度较快,而当 $L > 10$ 后,增加 r 和 L 对图像压缩重建后的峰值信噪比 PSNR 改善效果就不明显了。滤波器的长度 L 越大,滤波所需要的时间也越长,对图像处理这样大的数据量运算而言,这是需要考虑的一个重要因素。因此由较小的 L 设计具有尽可能大的正则性阶数 r 的滤波器将有利于图像数据的压缩处理。

当然,对大多数情况“正则性阶数越高图像压缩效果越好”,在某些情况下也存在正则性阶数小的小波基对图像压缩效果优于正则性阶数高的小波。例如,对电视图像信号发生器产生的棋盘信号或方格信号(主要用于电视调试),在相同数码率下,用 Haar 小波进行压缩的效果就优于用 Daubechies 小波 ($N > 2$) 得到的效果,这主要是所采用的正交基——Haar 小波函数的基本图像与待编码图像的结构有一定的相似性。基函数的基本图像与待编码图像的结构越相似,压缩效果就越好,这说明压缩效果的好坏除与小波函数的正则性阶数有关外,还与图像本身的结构有关。

尽管到目前为止,人们仍没找到一个小波基能对各种类型的图像都呈现出最佳的能量压缩性能,但通过大量实验,研究人员还是找到一些对各种情况(各种图像、各种位率)总体上相对较好的小波基,如 JPEG2000 推荐的几种小波滤波器:

(1) Daubechies (9,7) 滤波器,很多文献^[6,7]已经证明(9,7)小波滤波器在低位率时有最好的性能,(9,7)小波滤波器的缺点是硬件实现相对复杂一些;

(2) CRF13-7 滤波器,是整系数小波滤波器,具有较低的实现复杂性,适合于硬件实现;

(3) Legall I5-3 滤波器,是整系数滤波器,具有较低的实现复杂性,常用于无损压缩。

3.3 小波变换的层次数

由一维小波采用张量积构成的可分离小波变换,将原始图像分解成一个低频信号和三个方向的高频分量信号,低频信号又可进一步分解成四个子带,故总的子带数为 $3K+1$,其中 K 为分解层数,大部分文献报道 $K=3$,也有 $K=5$ 或 6 。究竟分解多少层可以满足要求,虽然要看图像的复杂程度和滤波器长度,但从子带信息来看,当一个子带分成四个子带的熵值的和应该是很小,否则就不值得再分解。如信号 f 的熵定义为

$$H(f) = - \sum_{i=1}^n p(f_i) \log_2 p(f_i)$$

$p(f_i)$ 为符号 f_i 出现的概率。如果给定子带 S_j 要进一步分解成 $S_{j+1}, W_{j+1}^1, W_{j+1}^2$ 和 W_{j+1}^3 四个子带,则将其熵分别记为 $H(S_j), H(S_{j+1}), H(W_{j+1}^1), H(W_{j+1}^2), H(W_{j+1}^3)$ 。由于熵与编码所需的平均比特数直接对应,则为了实现数据压缩,分解后的四个子带编码所需的平均比特数应小于分解前的子带 S_j 编码所需的平均比特数,也即应有 $\frac{1}{4} |H(S_{j+1}) + H(W_{j+1}^1) + H(W_{j+1}^2) + H(W_{j+1}^3)| < |H(S_j)|$

基于非分离二维正交滤波器的小波分解和上述由一维小波采用张量积构成的可分离小波分解不同,它是将每一层分解成一个低频子带和一个高频子带,而不是三个高频子带,所需的分解层数也应满足分解后的熵小于分解前的熵,有文献报道非分离小波分解一般分解四层。

4 量化

量化是有损图像压缩关键的一步。已有研究证明在相同条件下^[2],图像编码器中小波变换比 DCT 变换,最多能提高 1dB。也就是说,影响图像编码器性能的关键因素是量化方法和熵编码算法。

量化是用有限个状态表示一组连续值采样的过程。很明显量化必然带来某些信息损失,好的量化器就是以尽可能小的失真来表示原信号。为了实现这一目标,需要为输入信号找到一个概率模型,其策略是:对信号发生概率基本可以忽略的位置处,分配少量或不分配量化值,而信号发生概率较大的位置点,分配较多的量化值。

量化器由它的输入划分和输出量化值来确定,优化的量化器满足下列条件^[3]:

(1) 重心条件 (centroid condition): 给定编码器划分区,最好的解码器生成的量化值是划分区质量的中心;

(2) 最近邻域条件: 给定解码器量化值点,最好的编码器的划分区精确地落在量化值点的中点,也就是说,每一个 x 点以最临近的量化值点来分组。

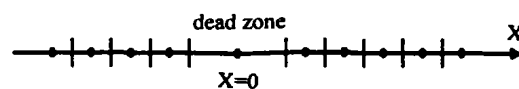


图3 截止区量化器,在 $x=0$ 附近有较大的划分区,其它位置采用均匀量化

4.1 均匀量化

均匀量化器的所有重构层都是等间距的,均匀量化是最简单的量化策略,但同时又使用最多的量化方法。在实际应用当中,通常对标准量化器稍加改进,即设 $|x| \leq \Delta$,量化输出为零,如图3所示。除 $[-\Delta, \Delta]$ 之外的每个量化带 $[i\Delta, (i+1)\Delta]$,其解码值都处于带的中心位置。H. Gish^[8]等已证明在非对称高位率的情况下均匀量化器是优化的,尽管实际应用中带有截止区的量化器仅是次优化的,但它们的性能几乎与 Lloyd-Max 编码器一样好^[9],另外截止区量化器还有低复杂性和对信源系数分布变化的鲁棒性。

4.2 矢量量化

矢量量化 (VQ) 是标量量化的推广,标量量化一次只对一个像素量化,而矢量量化一次对一个像素块即一个矢量进行量化。VQ 的基本结构实际上与标量量化相同,也是由编码器和解码器构成,编码器决定了输入矢量空间的划分,每一个划分赋予一个索引,称为码字,所有码字 (codeword) 的集合构成码书 (codebook)。解码器将索引映回到重现矢量,编码器和解码器的结合产生一个空间划分到离散矢量的映射。

虽然矢量量化是非常强大的编码工具,但当维数增加时其计算复杂性和空间复杂性以指数速度增加。为了克服这一问题,研究人员提出了多种 VQ 设计方法,其中比较著名的有:树状 VQ、形状增益 VQ、分类 VQ、多台阶 VQ、格状 VQ、层状 VQ 等。

量化器的优化设计需要有关图像空间概率分布方面的知

识,虽然我们有可能得到低阶联合概率分布方面的经验知识,但对高阶情况却很难获得,并且随着模型维数的增加,每一子带可用于模型估计的数据量逐渐减少, pdf 估计的可靠性降低,这就是人们所知的“维数的灾难”。

在实用小波编码器设计中,人们已经找到很多利用小波变换系数自身结构特点高效的编码方法,这些编码方法结合最简单的均匀量化可以达到矢量量化的压缩效果,而且避免了矢量量化的计算复杂性。

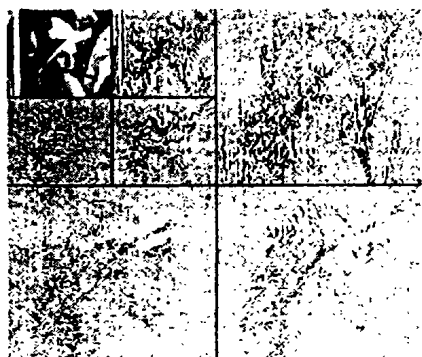


图4 Lena 图像二级变换系数的分布

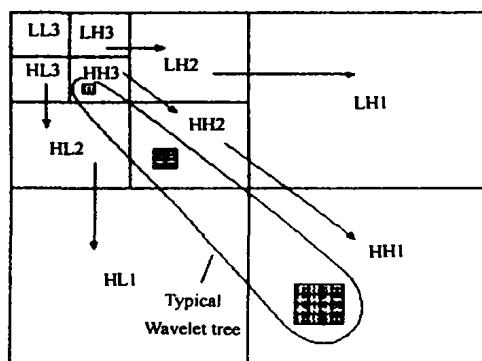


图5 小波系数零树结构

5 零树编码

从小波变换后的系数分布可知,小波变换具有较好的系数能量压缩特性.系数能量的压缩导致能更好地发挥标量量化的作用,图4是 Lena 图像二级变换系数的分布情况,观察其高频子带,实际上我们还能看到 Lena 结构的存在,特别是在精细尺度子带.从相关概念上讲,只要有轮廓结构存在,就说明有相关,那么仍有压缩的余地,所以高级小波压缩算法几乎都注重研究如何利用高频子带的结构特点,实现最有效的压缩.观察小波系数分布可以看出,高频带中大片区域存在较少或没有能量,并且含有能量的小部分区域在各个频带中的分布位置和形状非常类似.这些高频能量区域源于对图像信息欠佳的能量压缩,图像中的变化平坦或缓慢的区域可由低频小波基元素很好地描述,因此产生很高的能量压缩,但在边缘位置,低频小波基不能有效地描述这些信号,致使一些能量泄漏到高频区域,成为高频系数,这种情况类似地发生在小波变换的各个尺度上,因此表示图像边缘信息的高频系数具有同样的形状.零树编码器正是利用了这一特点,将跨带相关与零编码表示有机地结合起来,构成功能强大的压缩算法。

Lewis 和 Knowles 首先实现了零树编码的思想^[10],在他

们的算法中图像被表示为一个零树结构,如图5所示,树根节点代表低频带尺度的小波系数,树内节点对应于某一尺度的小波系数,树的高度反映尺度的大小,每一非树叶节点,都有4个子节点,它们分别对应具有同样空间位置像素点下一精细尺度的4个不同相位;树的底部为叶节点(没有子节点)。Lewis 和 Knowles 的零树模型是通过如下观察得到的:在大多数情况下,小波系数小时,它的子树系数也小.出现这种现象主要由于重要系数产生于边缘和纹理,而边缘和纹理是局部性的,因此 Lewis 和 Knowles 假定不重要的父节点总是对应不重要的子节点,包含不重要系数的树或子树可当作零树。

Lewis 和 Knowles 对小波系数进行量化,先设定一个量化门限(阈值),如果当前节点小于门限则认为它是不重要节点,忽略它的所有子节点,否则分别检查它的4个子节点并重复这一过程;如果当前节点为叶节点,则转到下一根节点,然后重复这一过程。

Lewis 和 Knowles 编码器实现数据压缩,除得益于小波变换很好的能量压缩特性之外,还利用了联合零值编码.小波域数据为行程编码提供了很大的便利,不但产生大量的零程,而且不必传输零程的程度,因为叶节点自动表示零程的结束.这种量化方法类似于 JPEG 的编码方法,实际上是一种混合标量/矢量量化策略,因为零树的表示符实际上代表一个矢量,该矢量由当前系数和紧随其后直到叶节点的零值构成。

虽然 Lewis 和 Knowles 编码器产生主观上可接受的图像,但它的率失真性能不如 JPEG,造成这一情况的主要原因是该算法对系数的量化缺乏精密性.一个系数小于门限时,它的所有子节点都不重要,这只是一可能,而 Lewis 和 Knowles 的算法仅简单地假定小的父亲总产生小的儿子,一旦假定不成立,将大的系数量化为零就会导致很大的失真.克服这一缺点导致新一代零树编码器的产生。

6 Shapiro 和 Said-Pearlman 编码器

Lewis 和 Knowles 算法尽管具有许多后来的优秀编码器的基本思想,但它缺少后来的编码器用于指定系数重要性的映射表.重要性映射表可以看作一个二进制函数,由函数值确定每一个系数是重要的还是不重要的,如果系数指定为不重要,则可把它量化为0,否则系数被量化为非零值。

首先采用重要性映射表(significance map)的编码器是著名的 Shapiro 嵌入零树小波编码器 EZW (Embedded Zerotree Wavelet)^[11].Shapiro 算法不仅传送非零数据而且传送重要性映射表,映射表所需的位数可能占据编码输出的主要部分,尤其在甚低位率的情况下。

Shapiro 算法在重要性映射表中定义了三个符号:零树(ZT),孤立零(IZ)和重要值(SV).利用这一改进的零树结构和熵编码算法,Shapiro 编码器获得了很好的率失真性能.一旦不重要父节点含有一个或几个重要的子节点,则产生 IZ 符号,这一事件的概率较低,因而所需的位率相对较高,但为避免因丢失子树中重要信息导致较大的失真,这样做是值得的。

另外,Shapiro 还提出了嵌入编码的概念,产生嵌入代码的编码器称为渐进传输或逐步求精编码器.在逐步求精过程,首先传送图像的近似数据,然后随着更多信息位的传送逐步提高近似程度,这种算法对带宽受限的网络应用具有实际意义。

Shapiro 使用位平面方法生成嵌入代码,如图6所示,将图像的小波系数按它们的重要性(如图7的顺序)排成一个一维

数组, 先低频带后高频带, 每一系数都用二进制 $b_7, b_6, b_5, b_4, b_3, b_2, b_1, b_0$ 表示, 处于同一位号 b_i 的所有位构成一个位平面. Shapiro 在扫描位平面形成位片之前, 先将小波系数与一个门限 T 相比较, 如果系数大于 T , 则标记为重要系数, 否则就是不重要系数, 将其置为零. 在完成第一遍扫描之后, 将门限调整为 $T/2$, 重复上述过程, 直到所有位片编码完毕. 另外, 为了更好地实现嵌入, Shapiro 还将系数数值与系数的符号分开编码.

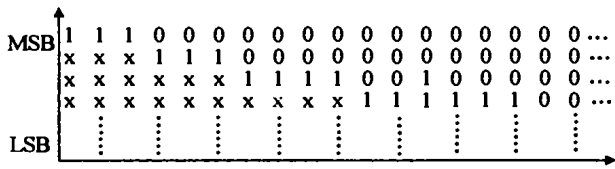


图6 小波系数位平面

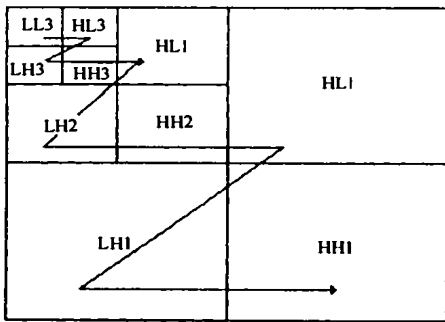


图7 小波系数的扫描顺序

从图6可以看出, 因为门限的限制, 位片中零的数目很大, 零树编码的作用就是避免传输这些零值, 一旦遇到零树符号, 我们就知道它的所有子节点也都为零, 因此不必传送附加的信息. 实际上, 零树是一种更聪明的行程编码, 它按另一种方式对系数进行排序, 生成更长的行长, 并且行程结束是自动隐含的, 因而不需传送行长信息.

对一个给定的系数, 可能需要多遍扫描直到到达某个门限之后才被标记为重要系数, 并且至此之后, 该系数不会再有很多的零树信息.

Said 和 Pearlman 在 Shapiro 算法的基础上提出了 SPIHT (Set Partitioning in hierarchical Trees) 图像编码算法^[12], 更注重一些细节的处理, 压缩性能比 Shapiro 提高 0.3~0.6dB. Shapiro 的算法只考虑单个零树的情况, 但在实际应用当中, 多个零树连续出现的概率很高, 以致于给它们分配特定的符号是值得的, Said 和 Pearlman 编码器实现了这一点, 这是其性能比 Shapiro 提高的主要原因.

7 频率自适应编码器和空域-频率自适应编码器

基于小波包变换的编码器称频率自适应编码器, 它按照高斯过程对变换编码进行分析, 以寻找最好基使增益最大化. Ramchandran 等^[13]提出使用率失真准则寻找最好小波基算法.

使用小波包变换产生增益的大小对图像结构的依赖性很高, 例如, 对 Lena 图像在位率为 1 时比使用小波变换 PSNR 提高大约 0.1dB, 在位率为 0.25 时为 0.25dB; 对 Barbara 图像在同样位率下 PSNR 提高约 2dB. 这是因为 Lena 图像频谱的

各个子带划分几乎是均匀、平滑的, 而 Barbara 图像的频谱在高频带具有很窄的尖峰.

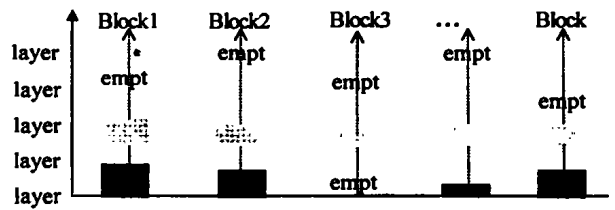


图8 描述不同子块对各位流层的字节贡献, 图块高度表示贡献的大小, empty 表示没有贡献

最好小波基算法不只局限于频域自适应变换, Smith 等^[14]提出了基于小波包的频率和空间自适应变换, 该算法通过一个二叉树分解过程在选定的图像区域中进行自适应频率分割, 实验结果表明, 空间-频率自适应编码对非平稳图像有很好的效果.

空间-频率自适应策略的主要缺点是预计算工作量很大, 另外为了控制计算复杂性必须限制于二进制划分, 但二进制划分不一定是最佳的^[15].

8 EBCOT 编码器

EBCOT (Embedded Block Coding with Optimized Transform) 属于熵编码方案的一种, 由于该方案不仅能获得较高的信息压缩性能和容易控制位率, 而且具有低存储器要求和适度的实现复杂性, EBCOT 已被 JPEG2000 采纳.

EBCOT^[16,17]的基本思想是, 将每一个子带划分为多个子块, 对这些子块使用相同的算法进行完全独立的编码, 生成可分离的位流, 而不必考虑子块之间的相关性, 并且允许在任意长度位置处截断位流.

在完成对整个图像压缩之后, 紧接着对所有子块进行一次后处理, 以确定达到特定的位率、失真和其它度量参数目标, 每一块的嵌入位流应该在哪一点处截断, 这样得到图像质量目标截断块第一个位流层; 调整目标位率、失真度等参数, 对所有子块进行满足新要求的截断点计算, 这样又得到第二个截断位流层; 如此重复, 随着位流层的增加图像的近似精度越来越高, 直到完全精确地描述原图像. 图8为所形成的部分位流层.

EBCOT 编码策略不仅适于小波变换而且适于小波包变换, 并且子块独立编码为算法的并行实现提供了可能. EBCOT 采用最简单的均匀量化, 落在 $-\sigma \sim \sigma$ (σ 为量化步长) 范围内的系数直接量化为零. EBCOT 熵编码采用改进的算术编码 MQ, MQ 能避免乘、除运算, 具有更低的计算复杂度, 并有利于算法的硬件实现. 另外, EBCOT 编码器还具有低存储器要求, 有效的位率控制, 高压缩性能, 以及可对特定区域有选择地解码等特点.

9 未来的发展

当前图像编码研究正从多个不同前沿继续发展. 在最基础的层次, 出现了对小波变换新的解释, 这一新的理论架构称为上举策略 (lifting scheme)^[18], 能够提供比基于标准 Fourier 方法更简单灵活的小波设计方法, 使用上举策略构造的非分

(下转第 64 页)

噪声后的结果, $psnr=16db$, $Sim=9.98$; 加20db Gauss 白噪声后, $psnr=22db$, $Sim=19.23$; 加30db Gauss 白噪声后, $psnr=18db$, $Sim=12.24$; 加50db Gauss 白噪声后, $psnr=14db$, Sim

$=7.39$ 。由图4~7看出, 虽然经过处理后的水印图像质量有严重下降, 但水印检测器响应值仍然高于6, 表明水印仍然存在。所以该算法对这些通常的图像处理具有较强的鲁棒性。



(a) 原始图像



(b) 加水印图像

图2 水印的不可见性比较



图5 剪掉四周 75%数据, 值留下中心 25%数据的结果



图6 7×7 中值滤波的结果



图4 JPEG 压缩后的水印图像, 压缩质量为 15%



图7 加 Gaussian 白噪声的结果

结论 本文利用图像小波变换系数的树结构关系和各层小波系数的特点, 将最低频逼近和中频系数分为两类: 一类对应于强纹理区域, 另一类对应于弱纹理区域, 根据人眼视觉特性, 在强纹理区域嵌入较强水印, 在弱纹理区域嵌入较弱的水印, 使水印具有不可见性和较好的鲁棒性。

参考文献

1 刘瑞琦, 谭铁牛. 数字图像水印研究综述[J]. 通信学报, 2000, 21(8): 39~48

2 Cox I J, Killian J, Leighton T. A secure, robust watermark for multimedia[J]. IEEE Trans. Image Processing, 1997, 6(12): 1673~1687
 3 Daubechies I. Ten Lectures on Wavelets, CBMS-NSF conference series in applied mathematics, SIAM Ed
 4 Shapiro J M. Embedded image coding using zerotrees of wavelet coefficients [J]. IEEE trans. On Signal Processing, 1993, 41(12): 3445~3462

(上接第61页)

离小波已证明有潜力改进编码器的性能。

基于小波变换的视频编码研究更具挑战性。在视频编码中将2-D小波变换与时域的运动估计预测结合起来取代当前的基于DCT的编码, 有许多人在从事这方面的研究。把视频信号看作3维数组, 从而使用3-D小波分析实现对其压缩也是可能的。这一方法遇到的困难源于离散小波变换的基本属性, 离散小波变换是一个空间变化算子, 而视频信号包含有目前离散小波变换不适合处理的时间变化特性。

参考文献

1 李建平, 等. 小波分析与信号处理. 重庆出版社, 1997
 2 Xiong Z, et al. A Comparative Study of DCT- and Wavelet-Based Image Coding. 见: 微软中国研究院论文集, Oct. 1999
 3 Tekalp A M. 数字视频处理. 电子工业出版社, 1998
 4 Coifmam R R, et al. Signal processing and compression with wavelet packets. In: Proc. of Intl. Conf. on Wavelet Applications, Toulouse, France, 1992. 8~13
 5 Mallat S G. A theory of multiresolution signal decomposition: the wavelet representation. IEEE Trans. Pattern Anal. And Machine Intell., vol. PAMI-11, July 1989
 6 Villasenor J D, Bellzer B, Liao J. Wavelet filter evaluation for image compression. IEEE Trans. On Image Processing, 1995, 4(8)
 7 Lu J, Algazi V R, Estes R R. A comparative study of wavelet image coders. Optical Engineering, 1996, 35(9): 2605~2619
 8 Gish H, Pierce J N. Asymptotically efficient quantizing. IEEE Transactions on Information Theory, 1968, IT-14(Sep.): 676~683

9 Farvardin N, Modestino J W. Optimum quantizer performance for a class of non-Gaussian memoryless sources. IEEE Transactions on Information Theory, 1984, 30(May): 485~497
 10 Lewis A S, Knowles G. Image compression using the 2-d wavelet transform. IEEE Transactions on Image Processing, April 1992
 11 Shapiro J. Embedded image coding using zero-trees of wavelet coefficients. IEEE Transactions on Signal Processing, 1993, 41(Dec.)
 12 Said A, Pearlman W A. A new, fast, and efficient image codec based on set partitioning in hierarchical trees. IEEE Transactions on Circuits and Systems for Video Technology, June 1996
 13 Ramchandrandran K, Vetterli M. Best wavelet packet bases in a rate-distortion sense. IEEE Transactions on Image Processing, 1992
 14 Smith J R, Chang S F. Frequency and spatially adaptive wavelet packets. In: Proc. ICASSP, May 1995
 15 Davis G, Nosratinia A. Wavelet-Based Image Coding: An Overview. Applied and Computational Control, Signals, and Circuits, 1998, 1(1)
 16 Taubman D. High performance scalable image compression with EBCOT. IEEE Transactions on Image Processing, 1998
 17 JPEG2000 Verification Model 5. 0 (Technical description). ISO/IEC JTC1/SC29/WG1 N1420. Aug. 1999
 18 Sweldens W. The lifting scheme: A new philosophy in biorthogonal wavelet constructions. in Wavelet Applications in Signal and Image Processing III, SPIE 2569, 1995
 19 李波, 汪海. 基于小波包变换的分层预测图像压缩算法. 计算机学报, 1999(8)
 20 程正兴. 小波分析算法与应用. 西安交通大学出版社, 1997