

图象小波变换编码的双金字塔数据结构

陈顺三 扬立 谭玫芳

(北京电力专科学校,北京 100044) (武汉水利电力大学,武汉 430072)

摘要 利用小波变换将图象分解成一组多分辨率子图象,提出了一种改进的编码方案。利用子图象的灰度统计分布特性和人眼视觉特性的分配比特数;采用双金字塔数据结构对图象数据进行编码,充分利用子图象的相似性,进一步减少了数据量。10幅奶牛图象的平均压缩比达到26.2,平均信噪比为30.6dB,具有较好的图象恢复质量。

关键词 小波变换,图象编码,金字塔数据结构

1 图象小波变换

H_r, H_c 分别表示水平方向和垂直方向离散的低通滤波器, G_r, G_c 分别表示水平方向和垂直方向离散的高通(实际为带通)滤波器,将其按下式作用于分辨率水平为 J 的图象 C_J :

$$\begin{cases} C_{J+1} = H_r H_c C_J \\ d_{J+1}^1 = H_r G_c C_J \\ d_{J+1}^2 = G_r H_c C_J \\ d_{J+1}^3 = G_r G_c C_J \end{cases} \quad (1)$$

将得到分辨率水平为 $J+1$ 的子图象 $C_{J+1}, d_{J+1}^1, d_{J+1}^2, d_{J+1}^3$, 这就是图象的 Mallat 分解算法^[1]。图1为将 C_0 进行三次分解得到图象结构。

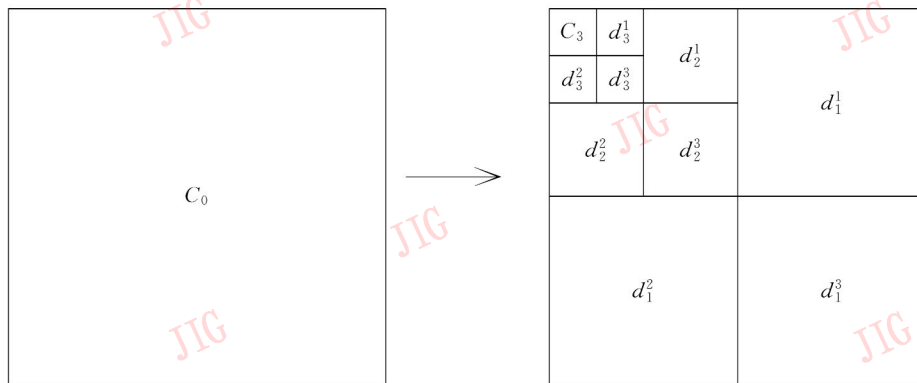


图1 三次分解后的图象结构

Fig. 1 The image structure after decomposing three times

C_0 是离散量化的原始图象, C_3 称为“模糊象”, 是经过三次低通滤波后的子图象; d_1^1, d_2^1, d_3^1 表示经过三次水平方向高通滤波后的系列子图象; d_1^2, d_2^2, d_3^2 表示经过三次垂直方向高通滤波后的系列子图象; d_1^3, d_2^3, d_3^3 表示经过三次对角线方向高通滤波后

的系列子图象。图象的 Mallat 重构算法:

$$C_J = H_r^* H_c^* C_{J+1} + H_r^* G_c^* d_{J+1}^1 + G_r^* H_c^* d_{J+1}^2 + G_r^* G_c^* d_{J+1}^3 \quad (2)$$

式中 $H_r^*, H_c^*, G_r^*, G_c^*$ 分别是 H_r, H_c, G_r, G_c 的共轭滤波器。

2 小波系数的比特分配

2.1 阈值取样

图 2 所示的灰度分布图和图 4(b)所示的小波

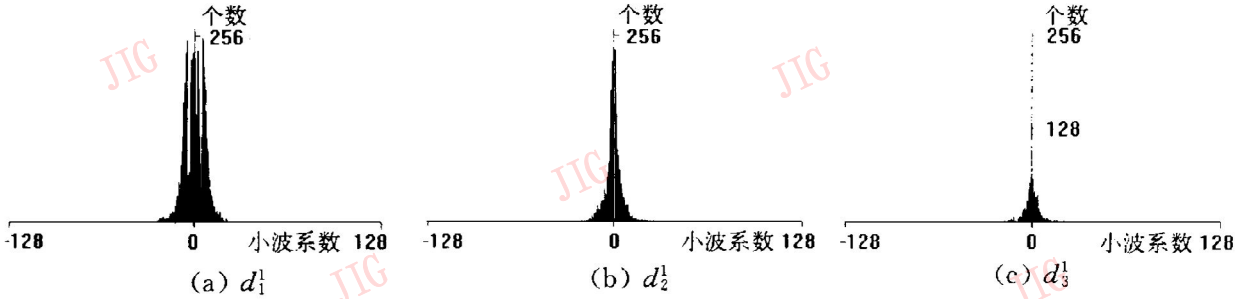


图 2 水平方向子图象的灰度分布

Fig. 2 The gray distribution of horizontal sub-images

阈值取样数据大大减少,有利于进行编码压缩。由于人眼对低频失真比较敏感,对高频失真不是十分敏感,人眼具有对光的方向检测的功能,对对角线方向不甚敏感,所以对不同的子图象采用不同的比特数编码。

2.2 最佳比特数分析

设图象具有 N 级灰度,记作 $\{h_1, h_2, h_3, \dots, h_N\}$, 将各个灰度级出现的概率记作 $\{p_1, p_2, p_3, \dots, p_N\}$, 该信源表示为:

$$X = \left\{ \begin{matrix} h_1, h_2, h_3, \dots, h_N \\ p_1, p_2, p_3, \dots, p_N \end{matrix} \right\} \quad (4)$$

该图象的熵为:

$$H(X) = - \sum_{i=1}^N p_i \log_a p_i \quad (5)$$

熵 $H(X)$ 描述了该图象的平均信息量。

如果给第 i 个灰度分配 N_i 个比特,则图象的平均比特数为:

$$\bar{N} = \sum_{i=1}^N p_i N_i \quad (6)$$

图象无失真编码效率为:

$$\eta = \frac{H(X)}{\bar{N} \log_a n} \quad (7)$$

取编码字符集 $A = \{0, 1\}$ 时, $n = 2$; 以比特为单位时, \log_a 的底 $a = 2$; 当编码效率为 100% 时, 由式(6)得到平均比特数:

$$\bar{N} = H(X). \quad (8)$$

为了简化计算,对同一子图象的全部象素分配同一比特数,即该子图象的最佳比特数。表 1 列出了

分解图表明小波系数大部分接近零,而且大致呈对称分布。如果选定一阈值 $Th > 0$, 将小波系数 $D(x, y)$ 按下式处理:

$$D(x, y) = \begin{cases} D(x, y) & |D(x, y)| > Th \\ 0 & |D(x, y)| \leq Th \end{cases} \quad (9)$$

10 幅奶牛图象各子图象比特数的均值。

表 1 各子图象最佳比特数的均值

Table 1 The mean optimal bit per pixel in the sub-images

子图象	d_1^1	d_1^2	d_1^3	d_2^1	d_2^2	d_2^3	d_3^1	d_3^2	d_3^3
比特数	2.5	2.8	1.1	4.9	4.9	4.4	5.6	5.6	4.2

2.3 比特分配

为了提高处理速度,实际只选用 1, 2, 4, 8 个比特数进行等长编码。本编码是一种失真编码,但表 1 可作为实际分配比特数的依据。

(1) C_3 为图象的最低频率分量,它集中了图象的大部分能量,因此分配 8 比特;

(2) d_1^3 表示对角线方向的最高频率分量,人眼对它的失真不敏感。表 1 给出的平均比特数为 1.1, 实际分配 0 比特;

(3) d_1^1, d_1^2 分别为垂直和水平方向的最高频率分量,表 1 给出的平均比特数分别为 2.5 和 2.8, 实际分配 2 比特;

(4) d_2^1, d_2^2 分别为垂直和水平方向的次高频率分量,表 1 给出的平均比特数均为 4.9, 实际分配 4 比特; d_2^3 为对角线方向次高频分量,表 1 给出的平均比特数为 4.4, 实际分配 2 比特;

(5) d_3^1, d_3^2 为垂直和水平方向的最低频率分量,表 1 给出的平均比特数均为 5.6, 实际分配 8 比特; d_3^3 为对角线方向的最低频率分量,表 1 给出的平均比特数为 4.2, 实际分配 4 比特。

3 双金字塔数据结构

经阈值取样后的每个子图象小波系数大部分为零,采用金字塔图象编码方法十分有效^[2];而同一方向的子图象具有相似的结构,物理上又形成了一个金字塔图象结构。对同一方向的系列子图象,用下述方法构造本文提出的双金字塔数据结构:

(1) 将子图象 d_3^i 按 4×4 大小划分成若干块,如果某块象素全为 0,则给该块标记为 0,否则标记为 1,8 个块的标记用一个字节表示,这样构成以块为单元的标记数据图象 I_0 ;

(2) 将 I_0 中标记为 1 的块展开为 2×2 大小的小块,如果该小块象素全为 0,则将其标记为 0,否则标记为 1,这样构成数据图象 I_1 ;

(3) 将 I_1 中标记为 1 的块展开为 2×2 大小的象素,如果象素值为 0,则给该象素标记为 0,否则标记为 1,这样构成数据图象 I_2 ;接下来对图象 I_2 中标记为 1 的象素用 8 比特或 4 比特进行编码;

(4) 子图象 d_2^i 和图象 I_2 是相似的,省去 1~3 步骤,直接将图象 I_2 中标记为 1 的点展开为 d_2^i 中 2×2 大小的象素,如果展开处的象素值为 0,则给该象素标记为 0,否则标记为 1,这样构成数据图象 I_3 ;对图象 I_3 中标记为 1 的象素用 4 比特或 2 比特进行编码;

(5) 子图象 d_1^i 和图象 I_3 是相似的,直接将图象 I_3 中标记为 1 的点展开为 d_1^i 中 2×2 大小的象素,如果展开处的象素值为 0,则给该象素标记为 0,否则标记为 1,这样构成数据图象 I_4 ;对图象 I_4 中标记为 1 的象素用 2 比特或 0 比特进行编码(见图 3)。

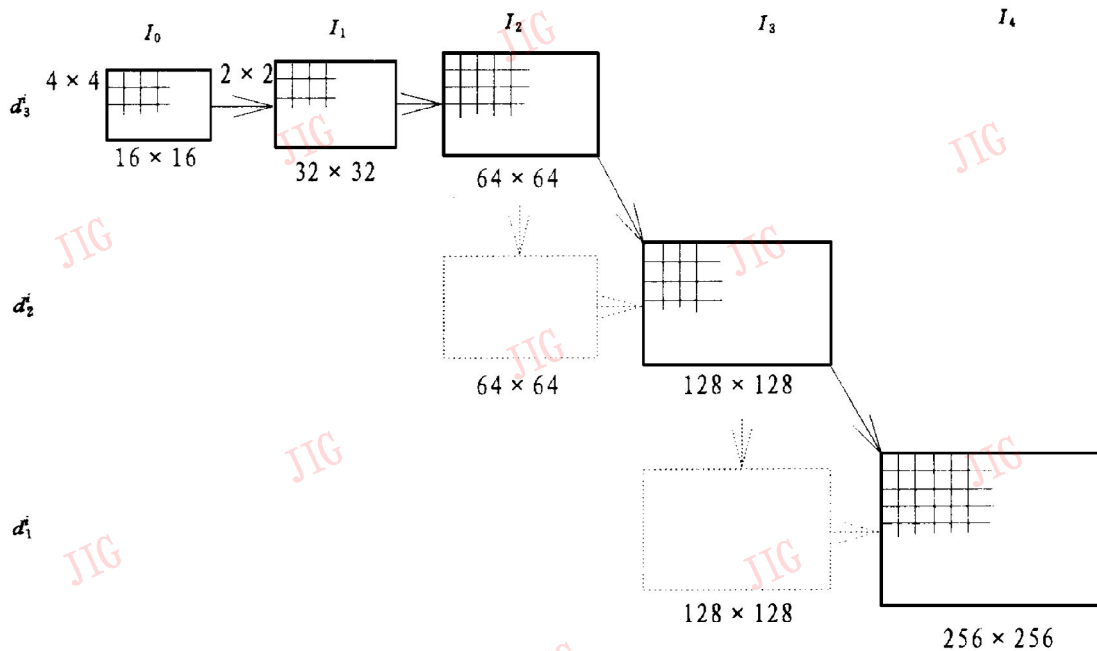


图 3 双金字塔数据结构编码过程

Fig. 3 The coding process of dual-pyramid data structure

C_3 子图象的相关性极强,采用 DCT 对其进行变换编码。 C_3 的大小为 64×64 ,考虑去相关的效果和方块效应,选取图象变换尺寸为 64×64 ;经变换后的图象能量集中在低频空间,故我们只对变换系

数阵的左上三角进行编码,其余作 0 处理,不编码。

编码图象文件包括文件头结构及各方向子图象数据,图象文件数据结构如下:

文件头	“模糊象” C_j 数据	水平方向子图象数据	垂直方向子图象数据	对角方向子图象数据
-----	----------------	-----------	-----------	-----------

文件头的长度为 128 字节,其结构如下:

0	4	5	95	105	106	123	125	127
图象尺寸	分解次数	小波系数	子图象比特数分配	C_j 编码方式	保留	水平方向子图象地址	垂直方向子图象地址	对角方向子图象地址

4 结果分析

试验采集的图象大小为 512×512 象素, 256 级灰度, 小波函数选用文献[3]采用的函数, 其参数见表 2。由于 $h(n)$ 是偶对称函数, 表中只列出 $n \geq 0$ 的一半。对 10 幅奶牛图象变换编码各子图象的平均比特数 (bpp) 见表 3; 表中 $bpp1$ 表示采用双金字塔编码方法编码的平均比特数, 其平均压缩比为 26.2; $bpp2$ 表示采用经典的金字塔编码方法编码的平均比特数, 其平均压缩比为 17.6。可见, 双金字塔编码方法和经典的金字塔编码方法相比压缩比有显著的提高。

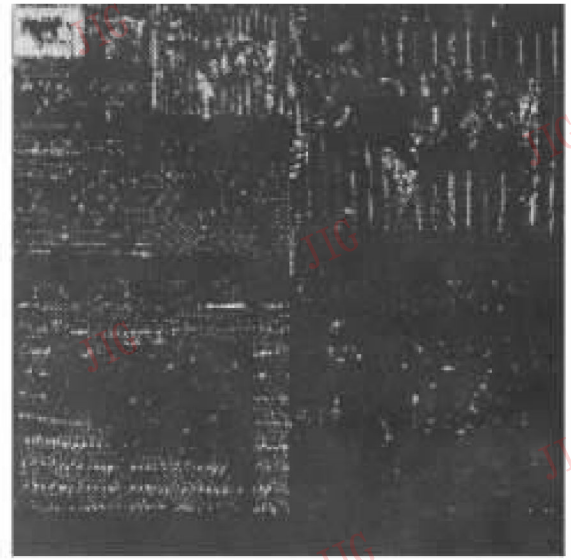
图 4(a) 表示原图象, 图 4(b) 表示三次分解后的全图象, 图 4(c) 表示采用金字塔编码方法编码后重构的图象, 图 4(d) 表示采用双金字塔编码方法编码后重构的图象。可以看出, 重构的图象和原图象在低频部分看不到失真, 而在高频部分(边缘)则变得模糊了; 图 4(d) 和 (c) 相比, 边缘更为模糊, 这是由于双金字塔编码方法利用图象子图象的相似性而产生的失真。从视觉效果来看, 图象重构的质量还是令人满意的。一般用峰值信噪比描述图象重构的质量:

PSNR =

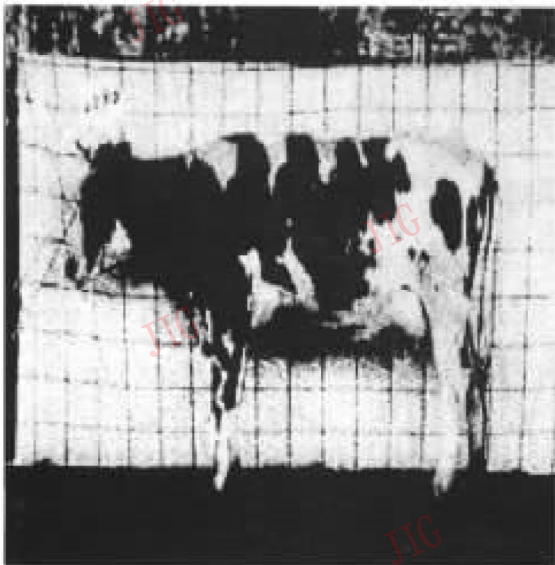
$$10 \lg \frac{(255)^2}{\frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [g(x, y) - f(x, y)]^2} \quad (9)$$



(a)



(b)



(c)



(d)

图 4 小波分解及重构的图象

Fig. 4 The images after decomposing and reconstructing

表 2 小波函数参数

Table 2 The parameters of wavelet functions

n	0	1	2	3	4	5	6	7	8	9	10	11
$h(n)$	0.542	0.307	-0.035	-0.078	0.023	0.030	-0.012	0.013	-0.006	0.006	-0.003	-0.002

表 3 各编码子图象比特数(bpp)的均值

Table 3 The mean bit per pixel in the coded sub-images

子图象	d_1^1	d_1^2	d_1^3	d_2^1	d_2^2	d_2^3	d_3^1	d_3^2	d_3^3	C_3
bpp1	0.273	0.251	0	0.558	0.554	0.305	1.646	1.619	1.186	1.240
bpp2	0.469	0.453	0	0.853	0.846	0.514	1.646	1.619	1.186	1.240

式中 $g(x, y)$ 表示重构的图象, $f(x, y)$ 表示原始的图象, $M \times N$ 为图象的尺寸。计算得到平均信噪比为 30.6dB。

5 结论

小波变换具有时间~频率定位能力并初步实现图象平稳成分和非平稳成分分离;低频成分精确定位于频域,且基本上是平稳的;高频成分定位于空间域且为非平稳的,它表现为图象的边缘和纹理。非平稳信号是信号处理的难点,由于图象高频成分被定位于空间域,因此很容易对其实现高效的编码。

本文根据人眼的视觉特性和图象非平稳的和局部化的特性,以及子图象小波系数的分布来分配比特数,并采用双金字塔数据结构对图象数据进行编码,在保持图象一定恢复质量的前提下,使图象压缩比得以提高。

参考文献

- 1 龙瑞麟. 高维小波分析. 北京:世界图书出版公司北京公司,1995.
- 2 Burt P J. The Laplacian Pyramid as a Compact Image Code. IEEE Trans. Commun. 1983, OMM-31, 532~542.
- 3 Mallat S. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE Trans. on PAMI, 1989. 11 (7): 674~693.



陈顺三,1985年获武汉水利电力大学水电厂自动化专业学士学位,1990获该校同专业硕士学位,1996年获北京农业工程大学工学博士学位。曾从事过动力系统的建模与仿真技术的研究,目前研究的方向有图象识别技术和图象数据库设计等。

A Dual-pyramid Data Structure for Image Coding Using Wavelet Transform

Chen Shunsan, Yang Li

Tan Meifang

(Beijing electrical power college, Beijing 100044) (Wuhan hydraulic & electrical university, Wuhan 430072)

Abstract wavelet transform is adopted to decompose an image into a group of multiresolution sub-images. An improved coding scheme is developed which have considered the gray distribution properties of sub-image and human being's vision, image data have been reduced farther by using dual-pyramid data structure for sub-image coding, which have great similarity with each other. The mean compression rate and signal-to-noise ratio for 10 cow's image is 26.2 and 30.6, respectively. The reconstructural quality of image is high.

Keywords Wavelet transform, Image Coding, Pyramid data structure.