

# 基于小波零树的图像感知熵编码研究<sup>1</sup>

袁杰辉 魏政刚 蔡元龙

(西安交通大学电信学院图像所 西安 710049)

**摘 要** 本文以小波作为编码框架,通过对人工合成图像的主观实验,得出小波变换域中各不同频带的视觉误差阈值函数,进而提出了一个新的量化编码方案。该方案基于小波域非均匀零树结构量化和高频低能系数的视觉非均匀量化。对实际图像的感知熵和几乎透明的情形作了研究,与 JPEG 标准也作了比较。实验表明该方案较充分、自适应地考虑人眼视觉冗余,具有稳健、实际可行的特点。对自然图像,在十倍压缩比左右可有视觉无失真质量。

**关键词** 感知编码,感知熵,小波变换,零树

**中图分类号** TN911.21

## 1 引 言

传统的信源编码理论存在两点不足。其一是没有达到目标码率的方法;其次是信源模型忽视了图像信号的真实情况信号(如非高斯性等),以及信息接受者并非以均方误差(MSE)为准则来评价原始与编码信号的相似程度。与传统方法不同,感知编码强调了压缩算法是使信号失真在感知意义上的最小化。如果在信号空间内的所有点上,噪声的分布都恰好处在临界可察觉失真水平上,则所对应的编码率是视觉完好的最低下限,称为“感知熵”,它比 Shannon 信息熵要低得多。而且对于一定过阈值的情形,从感知的角度仍是视觉最佳的。

感知编码的目的,在于尽可能地利用人的视觉系统潜力来达到更高的编码增益。而编码器对视觉模型框架的应用程度决定了其感知编码的效率。例如在离散余弦变换方法中,由于其所固有的空间信息淡化的缺点,局域最优化只能通过缩小方块尺寸达到,而尺寸大小选择是编码效率和局部掩盖效应的折衷问题。此外,由于人眼对图像中线性特征十分敏感,这种固有缺陷产生的方块效应,对主观降质有很大影响。小波/子带编码器,由于其独特的类视觉多通道倍频程方式分解、良好局域特性,为感知编码器设计提供了一种十分良好的选择。

研究设计具有高性能、稳健、实际可行的动态感知编码系统是十分棘手的难题<sup>[1]</sup>,其原因在于 HVS、图像对象的复杂性和编码策略本身的限制<sup>[2-5]</sup>。为克服前人的缺点,本文提出了一个新的一般化有效量化编码方案。

## 2 小波变换域的感知误差阈值函数

### 2.1 小波变换图像分解与合成

离散小波二维分解过程是:信号首先以水平一维方式进入滤波器组 G、H,其中 G、H 分别为低、高通滤波器。滤波器的输出经下采样再以垂直方式进入下一级 G、H,输出经下抽样分别为 LL、LH、HL 和 HH。LL 子图像可以以这种分解方式继续迭代进行下去,总的分解次数称为分解层数。信号的重构则为分解的逆过程。将分解层数记为  $\{0, 1, 2, 3, \dots\}$ ,  $\{0, HL\}$  即表示第 0 层的 HL 方向频带。

<sup>1</sup> 1997-07-02 收到, 1998-05-21 定稿  
国家自然科学基金资助课题

## 2.2 小波变换系数的亮度与对比度掩盖

图像经小波变换之后,再经量化和熵编码便达到数据的压缩。我们需要确定的是各频带小波系数的感知误差阈值 (Perceptual Error Threshold: PET) 分布,以使量化所引起的视觉误差在临界可觉察失真 (JND) 水平上。一般而言,掩盖效应主要包含频带特性、方向特性、亮度掩盖和对比度掩盖等基本方面。小波变换除了去相关的作用的同时,本身还提供了类视觉的空频分解,在此主要考虑亮度和对比度掩盖。亮度阈值的检测主要依赖于区域的平均亮度,例如亮度越大则阈值越高。对比度阈值指误差在不同对比度信号下的灵敏度。Peli<sup>[4]</sup> 针对复杂图像提出了区域带限对比度,这种定义是基于图象的多通道分解,与小波分解有着十分相似之处。因此亮度和对比度掩盖与平均亮度和系数幅度有关。在此,平均亮度的定义与 Peli 的类似。

## 2.3 主观测试图形与条件

采用阶跃函数作为基本测试图形。在黑色屏幕中央生成直径约 10cm 的一个圆,其中心垂线左右分别为阶跃的不同亮度以测试 HL 频带行为。以圆心 45 度方向分别为阶跃的两种亮度则用于测试 HH 频带的频带行为。

这种图形的设计基于以下两点: 其一是阶跃经小波变换,有一个相对很大的系数,能较孤立地研究单个系数的行为。另一个则是边缘的特殊性。边缘附近存在着明显的掩盖效应。但边缘的一致性被破坏时,其误差可视阈值又相当低。因此,阶跃图形可以将这两种重要视觉现象同时加以考虑和研究。表 1 列出了测试条件。

表 1 主观测试条件

相对观测距离 ( $d/h$ )	6
房间亮度	黑暗
屏幕最大亮度	58.6(cd/m <sup>2</sup> )
屏幕最小亮度	0.23(cd/m <sup>2</sup> )
观测时间	无限制
观察者人数	3

## 2.4 测试结果

测试阶跃信号分别为: {32, 0; 32, 24; 64, 0; 64, 32; 64, 52; 127, 32; 127, 63; 127, 95; 220, 32; 220, 63; 220, 127; 220, 191}。对测试图形先进行小波四层次分解,然后对各个层次不同方向的频带分别进行主观测试。具体过程为: 由于垂直方向的阶跃信号,小波系数只存在 HL 的方向的频带中,先将干扰信号加在 {0, HL} 频带上,而其他频带保持不变。干扰信号以 (+ + - - ……) 方式叠加在最大系数上,用以同时观测边缘扭曲失真。然后依次测试 {1, HL}、{2, HL}、{3, HL} 频带。考虑带间的相互综合影响,各个频带独立测试时所得的阈值共同乘以一个加权系数  $\alpha(0.5)$ 。此外实验还表明, LH 方向频带的 PET 行为与 HL 频带类似。

由于篇幅所限,表 2 仅列出 HL 频带的各层次正系数的 PET 函数。从实验结果中可以发现: (1) 视觉阈值与区域平均亮度近似存在线性关系,而线性函数受小波系数幅度大小调制。(2) Zigzag 效应所引起的另一方面是饱和效应的出现。(3) 系数的正负对 PET 有影响。

表 2 HL 各层次 (0,1,2,3) 正系数的 PET 函数。z 为 PET 值, x 为区域亮度, y 为小波系数幅度

0	1	2	3
$z(\max) = 7.0 \quad y \geq 28$ $z = 3.5 + 0.0156 * x$  $z(\min) = 4.0$	$z(\max) = 4.0 \quad y \geq 28$ $z = 1.0 + 0.0150 * x$  $z(\min) = 1.75$	$z(\max) = 2.5 \quad y \geq 28$ $z = 0.4 + 0.0110 * x, \quad y \geq 10$ $z = 0.0060 * x$ $z(\min) = 0.3$	$z(\max) = 1.0 \quad y \geq 20$ $z = -0.15 + 0.0055 * x, \quad y \geq 9$ $z = -0.06 + 0.0035 * x$ $z(\min) = 0.15$

### 3 基于感知的两级量化方案

零树结构的高效力是基于小波频带能量递减规律<sup>[6]</sup>, 但当系数变得较小时, 此规律将不存在, 零树的效力也将大大降低。其次, 从小波 PET 的分布来看, 低频部分的 PET 要比高频部分小很多。根据这些特点, 先采用非均匀零树结构对高频带的高能量部分以及低频带量化, 然后对高频部分剩余的低能系数在 PET 的控制下量化。具体的方案如下:

第一级 图像经五层次分解, 小波采用双正交 B97 小波。零树非均匀量化阈值各层次分布由高至低为 [1, 1, 2, 8, 16]。

第二级 第二级实际上是对第 1 层  $|X| < 8$ , 第 0 层  $|X| < 16$  的低能系数进行量化处理。本文的目的在于对视觉熵的研究, 因而据 PET 函数设计了三种量化方法, 分别对应于临界透明、几乎透明和视觉失真已开始明显等三种质量状态。

设量化阈值为  $\{T_{ij}; i=0, 1; j = -2, -1, 1, 2\}$ , 其中  $i$  表示层次号,  $j$  表示阈值序号。输出水平为  $\{S_{ij}; i=0, 1; j = -2, -1, 0, 1, 2\}$ ,  $S_{i0} = 0$ 。令  $X$  为小波系数,  $\bar{X}$  为  $X$  的量化值,  $T_{pet}$  为该系数上的 PET 值, 则量化规则如下:

如果  $(|X| < T_{pet})$  或  $(|X| < T_{i1})$ , 则  $\bar{X} = S_{i0}$ ; 否则如果  $(|X| < T_{i2})$ , 则  $\bar{X} = S_{i1}$ ; 否则  $\bar{X} = S_{i2}$ 。

表 3 列出了 HL 频带上的量化阈值和输出水平。HH 层的数值与 HL 层因其 PET 特点略有不同。系数经过这种视觉量化后, 便可按顺序采用二次 (Second Order) 算数编码。综上所述, 新的感知编码方案可如图 1(a)、1(b) 所示。

表 3 第二级中三种不同量化方法所对应的 HL(LH) 频带上量化阈值和输出水平

	方案 I		方案 II		方案 III	
	量化阈值	输出水平	量化阈值	输出水平	量化阈值	输出水平
第 1 层	-5, -2, 2, 5	-6, -3, 0, 3, 6	-8, -3, 3, 8	-4.5, 0, 4.5	-8, -4, 4, 8	-5, 0, 5
第 0 层	-16, -5, 5, 16	-10, 0, 10	-16, -6, 6, 16	-10, 0, 10	-16, -8, 8, 16	-11, 0, 11

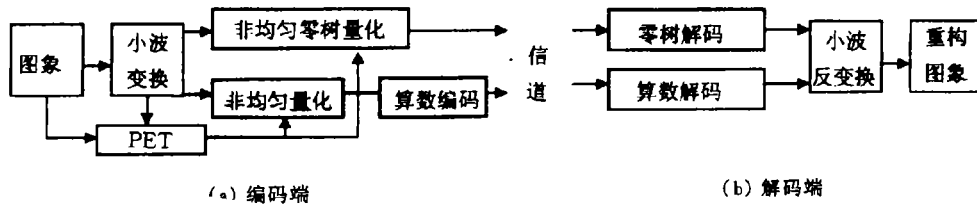


图 1 基于感知的两级编 / 解码方案

## 4 实验结果及讨论

对四幅具代表性内容(边缘、纹理、平缓)的图像进行了研究。图像为  $512 \times 512 \times 8\text{bit}$  大小的黑白图像。与 JPEG 标准也做了比较。表 4 列出了恢复图像的客观实验结果。主观测试结果 SCORE 也列在表 5 中。由于主要关心的是透明或几乎透明的情况,因此采用了与以往不同的主观评分标准,即 5: 没有差别, 4: 微弱差别, 3: 有差别, 2: 较大差别, 1: 很大差别。测试条件同表 1, 但测试者增加为 9 人。

表 4 本文方案中三种量化方法的 PSNR (dB) 和在相同码率下与 JPEG 的比较

图像	PSNR		码率 (bpp)	PSNR		码率 (bpp)	PSNR		码率 (bpp)
	I	JPEG		II	JPEG		III	JPEG	
Lena	36.96	35.45	0.654	36.62	34.69	0.536	35.98	34.03	0.458
Peppers	37.96	35.74	0.579	37.51	35.04	0.459	36.91	34.37	0.390
Scene	37.04	36.38	0.756	36.37	35.23	0.576	35.56	34.25	0.505
Girl	38.86	37.71	0.403	38.41	36.04	0.301	37.97	34.87	0.259

表 5 本文方案中三种量化方法的 S、SCORE 结果和在相同码率下与 JPEG 的比较

图像	S		SCORE					
	I	II	I	JPEG	II	JPEG	III	JPEG
Lena	4.97	4.27	5.0	5.0	4.3	4.1	3.6	2.9
Peppers	4.86	4.17	4.9	4.9	4.2	3.7	3.3	2.7
Scene	5.00	4.37	5.0	5.0	4.6	4.7	3.4	3.7
Girl	4.82	4.25	5.0	4.3	4.7	3.1	3.9	1.9

另一方面,针对 PSNR 的缺点,还给出了一种针对视觉熵新的质量测度下的结果<sup>[7]</sup>。对于视觉阈值水平上,注意到小波编码器质量的降低主要是由于细节的模糊所引起,同时再考虑到纹理性结构的掩盖效应,可得出一个能较好反映误差视觉效果的质量测度  $S$ :

$$S = 5.10 - 1.745(X_0 + 0.01X_1) + 0.775(Y_0 + 0.5Y_1),$$

式中  $X_i$ 、 $Y_i$  分别为小波域最高两个频带中反映模糊和掩盖效应的参数项。

从实验结果当中可得出以下几点:

(1) 在透明或几乎透明的条件下,小波域高频带低能量系数对视觉质量的影响不可忽视。它们所占的比特率对感知熵的贡献也很可观。通过对图像各个细节部分的仔细考察,总的说来对感知熵的估计,方案 I 可认为是一种稳健的和一般化的算法。方案 II 可以看作是几乎透明的情况。细节平滑在屏幕前可以察觉。但在 6 倍视距外,这种差别依然是较弱的。

(2) 对感知熵的估计,小波方法和 JPEG 方法性能是相近的,一般为十倍压缩比左右。但本文方法更具有一般适应性和实用性,而不象在 JPEG 中需要对每一幅图像通过不断调整质量控制系数  $Q$ , 以及长时沉闷的主观测试检验。另一方面在相同的码率下, JPEG 方法的稳健性稍差。即方块效应容易在平缓区出现,以及主观质量随码率的降低下降较快。在奔腾 100 计算机上,对典型图像 Lena (512), JPEG 编、解码时间分别均为 5s 左右,本文方法在编码时间上稍长,为 40 多秒;解码过程由于简单,只需 7s 多。

(3) PSNR 不能有效地反映出主观的质量评价来。比如在感知熵水平上,尽管 JPEG 方法与小波方法相比, PSNR 一般要低 1dB 多,但其主观感觉对某些图像并无大的差别。而测度  $S$  由于考虑了误差的视觉影响模式,更符合主客观实际结果。

## 5 结 论

本文依据小波的视觉误差阈值分布, 提出了一种新的有效可行的视觉熵两级编码方法, 这种有效性, 其一来自于小波域高效力的零树结构; 其次, PET 函数考虑了复杂图像的特点, 较好地反映了区域亮度和对比度掩盖效应, 因而对零树非均匀量化阈值分布确定和高频低能系数的非均匀量化, 较充分自适应地考虑人眼视觉容忍度。

## 参 考 文 献

- [1] Jayant N, Jonston J, Safranek R. Signal compression based on models of human perception. Proc. IEEE, 1993, 81(10): 1385-1421.
- [2] Safranek R. A perceptually tuned subband image coder with image dependent quantization and post-quantization data compression. Proc. ICASSP, 1989, 1945-1948.
- [3] Watson A B. Visual threshold for wavelet quantization error. SPIE vol.2657, 1996, 382-392.
- [4] Peli E. Contrast in complex images. J. Opt. Soc. Am. A, 1980, 7(10): 2032-2040.
- [6] Shapiro J. Embedded image coding using zerotrees of wavelet coefficients. IEEE Trans. on. SP., 1993, SP-41(12): 3445-3462.
- [7] 袁杰辉. 基于人的视觉特性和小波变换的图像压缩编码研究: [博士学位]. 西安: 西安交通大学, 1998
- [5] Daly S. The visible difference predictor: An algorithm for the assessment of image fidelity. In Digital Image and Human Vision, A. B. Watson, editor, Cambridge, Massachusetts: MIT Press, 1993.

## A STUDY ON THE PERCEPTUAL ENTROPY OF IMAGES USING WAVELET TRANSFORM

Yuan Jiehui    Wei Zhenggang    Cai Yuanlong

(Image Processing Center, Xi'an Jiaotong University, Xi'an 710049)

**Abstract** In this paper, perceptual error threshold (PET) functions in different wavelet subbands are obtained through subjective tests on artificial patterns. A new and generic two-stage coding scheme is thus proposed, based on the nonuniform zero-trees of wavelet coefficients and the visually nonuniform quantization of small coefficients in higher subbands. Researches have been done for the visual transparent and nearly transparent cases, in comparisons with JPEG standard. Experimental results showed the feasibility and robustness of the scheme. Transparent quality can be obtained around 10 compression ratios for natural images.

**Key words** Perceptual coding, Percetual entropy, Wavelet transform, Zero-tree

袁杰辉: 男, 1964 年生, 博士生, 从事有关图像处理及通信的研究开发与应用工作. 感兴趣的领域有数字图像通信、多分辨率信号处理、计算机视觉及多媒体应用等.

魏政刚: 男, 1974 年生, 博士生, 感兴趣领域为视觉模型、图像压缩和神经网络研究等.

蔡元龙: 男, 教授, 博士生导师, 主要感兴趣的领域为模式识别、图像处理、计算机视觉和图像通信等.