

# 基于人眼视觉系统的自适应小波图像编码算法<sup>\*</sup>)

王向阳 杨红颖 左珂可

(辽宁师范大学计算机与信息技术学院 大连 116029)

(计算机软件新技术国家重点实验室(南京大学) 南京 210093)

**摘要** 本文提出了一种基于自适应小波分解与人眼视觉系统的低比特率图像压缩编码算法,该算法具有以下特点:(1)以子带能量为判别标准,能够根据图像内容进行自适应形式的小波分解;(2)建立了人眼视觉系统的掩蔽模型,并据此给出了全新的小波系数自适应量化策略。仿真实验表明:本文算法是一种高效的图像压缩算法,不仅其压缩效果明显优于 JPEG2000、EZW、FWP 等小波域图像压缩算法(特别是低比特率下),而且可广泛适用于不同特征的数字图像。

**关键词** 图像压缩,自适应小波变换,人眼视觉特性(HVS),低比特率

## A New Image Coding Based on Adaptive Wavelet Transform and Human Visual System

WANG Xiang-Yang YANG Hong-Ying ZUO Ke-Ke

(School of Computer and Information Technique, Liaoning Normal University, Dalian 116029)

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

**Abstract** In this paper, a new image coding algorithm with low bit-rate based on adaptive wavelet transform and human visual system is presented. Unlike many previous wavelet coders, the adaptive wavelet transform based on sub-image energy is utilized, and the adaptive wavelet coefficient quantization strategy based on novel visual masking model is adopted in our compression algorithm. The experiment results show that the new image compression scheme performs better than that of the state-of-art image coders (JPEG2000, EZW, FWP) reported in the literature, especially for low bit-rate.

**Keywords** Image compression, Adaptive wavelet transform, Human visual system (HVS), Low bit-rate

### 1 前言

小波(Wavelet)变换是 20 世纪 80 年代中期出现的新时频域信号分析工具。自 1989 年 S. Mallat 首次将小波变换引入图像处理以来,小波变换以其优异的时频局部能力及良好的去相关能力在图像压缩编码领域得到了广泛应用,并取得了良好的效果<sup>[1-4]</sup>。然而,理论分析和实验结果表明,现有小波域图像压缩编码算法普遍存在如下不足:

- 所采用的传统塔式小波分解仅仅递归分解了低频子带,未能对高频子带实施有效的分解处理,如图 1 所示。这种塔式分解形式不适合于某些特征图像(如指纹图像、遥感图像等)。尽管应用小波包分解<sup>[5]</sup>能够在一定程度上解决以上问题,但却导致了图像编码速度大大降低。

- 进行低比特率图像压缩时将出现较为严重的边缘模糊现象,即所谓 Gibbs 效应。

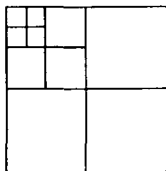


图 1 图像 3 级小波分解

本文提出了一种基于自适应小波分解与人眼视觉系统的低比特率图像压缩编码算法,该算法可以在一定程度上克服现有小波域图像编码方案所存在的不足。

### 2 基于图像内容的自适应小波分解

所谓基于图像内容的自适应小波分解,就是根据图像自身特征按照子带图像能量自适应地对 LL、LH、HL、HH 子带进行分解。一般说来,特征图像(如指纹图像、遥感图像等)的能量往往只集中在少数的若干频段,因此只需要对这些频段的子带信号进行分解。

为了实现图像的自适应小波分解,需要设计某种准则来决定对于一个子带图像的分解是否进行。考虑到计算量问题,本文将子带图像的能量作为衡量其重要性的准则,并将子带图像(含有  $M \times N$  个像素)的能量定义为(假设图像为  $I = \{x(i, j), 1 \leq i \leq M, 1 \leq j \leq N\}$ ):

$$E_o = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N |x(i, j)|^2 \quad (1)$$

其中,  $x(i, j)$  代表图像  $I$  的第  $i$  行、第  $j$  列像素灰度值。

则基于图像内容的自适应小波分解主要步骤如下:

(1) 按照式(1),计算给定图像  $I$  的能量  $E_o$ 。

(2) 选用合适的小波基,将上述图像  $I$  分解成 4 个子带  $l_{i,\theta}$  (其中,  $l$  表示小波分解层次,  $\theta \in \{LL, LH, HL, HH\}$  表示子带的方向)。

<sup>\*</sup>) 本文得到辽宁省自然科学基金(20032100)和计算机软件新技术国家重点实验室(南京大学)开放基金(A2004-05)资助。王向阳 教授,主要研究方向为多媒体数据压缩技术、多媒体版权保护技术。

(3) 仿式(1)计算各个子带图像  $I_{l,\theta}$  的能量  $E$ 。

(4) 如果子带图像  $I_{l,\theta}$  的能量  $E$  小于某个阈值(可以取为  $cE_0$ ,  $c$  是小于 1 的常数)或子带图像的最小尺寸为  $16 \times 16$ , 则终止分解。

(5) 如果子带图像  $I_{l,\theta}$  的能量  $E$  大于阈值  $cE_0$ , 则对子带图像  $I_{l,\theta}$  按照以上步骤再进行分解。

|    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|
|    | 1  | 4  | 5  | 16 | 17 | 20 | 21 |
| 2  | 3  | 6  | 7  | 18 | 19 | 22 | 23 |
| 8  | 9  | 12 | 13 | 24 | 25 | 28 | 29 |
| 10 | 11 | 14 | 15 | 26 | 27 | 30 | 31 |
| 32 |    | 33 | 34 | 45 | 46 | 49 | 50 |
|    |    | 35 | 36 | 47 | 48 | 51 | 52 |
| 37 | 38 | 41 | 42 | 53 |    | 54 |    |
| 39 | 40 | 43 | 44 |    |    |    |    |

图 2 Barbara 的自适应小波分解

图 2 为标准图像 Barbara 的自适应小波分解结果。不难看出, 该图像自适应小波分解算法不仅充分地去除了冗余信息, 而且尽可能地保留了有用信息。

值得注意的是, 本文所提出的自适应小波分解算法与图像内容有关, 其分解结果没有固定的统一形式, 故需要采用一定的数据结构(例如四叉树等)来记录相应的小波分解结果, 以便于解码。图 3 为标准图像 Barbara 自适应小波分解结果的四叉树表示。



图 3 分解结果的四叉树表示

### 3 人眼视觉系统与自适应量化

现有小波域图像压缩算法都是在 MSE 准则下完成小波系数量化工作的, 由于 MSE 准则仅仅根据小波系数幅值来确定其在编码中的重要性, 未能很好地刻画图像边缘和平坦区域的差别, 故必然导致低比特率图像压缩时出现较为严重的边缘模糊现象, 即所谓 Gibbs 效应。本文提出了基于人眼视觉掩蔽特性的自适应小波系数量化方法, 可在一定程度上解决 Gibbs 效应问题, 该自适应量化方法的基本思想为: (1) 对最低频子带 LL (即图 2 阴影部分) 的小波系数直接进行 DPCM 编码<sup>[6]</sup>; (2) 计算高频子带内小波系数的视觉掩蔽特性值并依此确定自适应量化步长, 然后按照符合人眼视觉特性的 Zigzag 扫描顺序 (即图 2 中的高频子带 1, 2, ..., 53, 54) 完成高频子带内小波系数的标量量化工作。

长期以来, 通过对人眼某些视觉现象的观察, 并结合视觉生理、心理学等方面研究成果, 发现了人眼视觉系统 (HVS) 的各种掩蔽 (Masking) 特性, 如能在量化 (标量量化) 高频子带小波系数时充分利用人眼的视觉掩蔽特性, 则可以在主观感觉图像失真相同的条件下允许更大的实际失真, 从而根据 Shannon 率失真理论可以采用更低的编码比特率而使图像的主观质量保持不变。

#### 3.1 人眼视觉掩蔽特性分析

为了很好地刻画及利用视觉掩蔽特性, 本文算法考虑了人眼视觉系统 (HVS) 如下几个方面的主要特性, 并据此给出了小波域视觉掩蔽特性值计算模型。

(1) 人眼视觉系统对于不同方向、不同分解层次中高频子带图像的噪声不是非常敏感, 特别是对  $45^\circ$  方向的高频子带噪声更不敏感 (如 HH 子带)<sup>[7]</sup>。设不同分解层次  $l \in \{1, 2, 3, 4\}$ 、不同方向  $\theta \in \{LH, HL, HH\}$  高频子带对噪声的视觉掩蔽因子为  $\lambda(l, \theta)$ , 则  $\lambda(l, \theta)$  可由下式估计:

$$\lambda(l, \theta) = \begin{cases} \sqrt{2} & \text{if } \theta = HH \\ 1 & \text{if } \theta = LH \text{ 或 } HL \end{cases} \begin{cases} 1.00 & \text{if } l=1 \\ 0.32 & \text{if } l=2 \\ 0.16 & \text{if } l=3 \\ 0.10 & \text{if } l=4 \end{cases} \quad (3)$$

(2) 人眼对不同亮度区域噪声的视觉敏感性不同, 通常对中等灰度最为敏感, 在中等灰度很宽的范围中 Weber 比保持常数 0.02, 并向低灰度和高灰度两个方向呈非线性下降, 如图 4 所示<sup>[8]</sup> (本文算法中, 我们不妨假设这种非线性为关于灰度的二次曲线)。设不同亮度对噪声的视觉掩蔽因子为  $\Lambda(l, \theta, x, y)$ , 则有

$$\Lambda(l, \theta, x, y) = 1.0 - \frac{0.5}{128^2} (AL(l, \theta, x, y) - 128)^2 \quad (4)$$

其中,  $AL(l, \theta, x, y)$  表示小波系数  $I(l, \theta, x, y)$  所处的背景亮度的平均值:

$$AL(l, \theta, x, y) = \frac{1}{4} \sum_{i=0}^{1} \sum_{j=0}^{1} I_l^u(i + \frac{x}{2^{l-1}}, j + \frac{y}{2^{l-1}}) \quad (5)$$

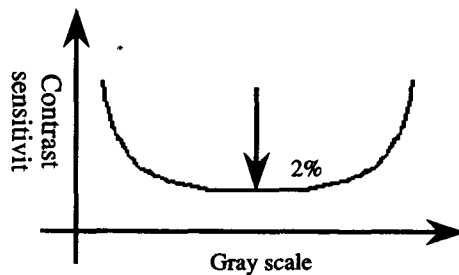


图 4 对比敏感度

显然, 当背景是中等亮度 (128) 时, 视觉掩蔽因子  $\Lambda(l, \theta, x, y)$  取最大值 1.0; 而当背景亮度偏离中间值时, 偏离越大, 视觉掩蔽因子  $\Lambda(l, \theta, x, y)$  取值越小。

(3) 人眼对图像平滑区域噪声较敏感而对纹理区域噪声较为不敏感, 即人眼对高频子带的高频分量失真往往比低频子带的高频分量失真更敏感。设像素点  $I(l, \theta, x, y)$  所在区域纹理掩蔽因子为  $\Xi(l, \theta, x, y)$ , 则有

$$\Xi(l, \theta, x, y) = \sum_{k=1}^{l-1} \frac{1}{16^k} \sum_{\theta \in \{LH, HL, HH\}} \sum_{i=0}^{1} \sum_{j=0}^{1} [I_{k+1}^u(j + \frac{x}{2^k}, i + \frac{y}{2^k})]^2 + \frac{1}{16^{l-1}} \text{Var}\{I_l^u(1 + j + \frac{x}{2^{l-1}}, 1 + i + \frac{y}{2^{l-1}}); i=0, 1; j=0, 1\} \quad (6)$$

通过上述分析与推导, 可以得到小波域视觉掩蔽特性值  $T(l, \theta, x, y) = \lambda(l, \theta) \Lambda(l, \theta, x, y) \Xi(l, \theta, x, y)$ 。显然, 该视觉掩蔽特性值  $T(l, \theta, x, y)$  不仅考虑了不同亮度的对比度掩蔽特性、不同纹理的掩蔽特性, 而且还考虑了人眼视觉系统对于不同分辨率、不同方向的敏感特性。

#### 3.2 高频子带小波系数的自适应量化

量化高频子带小波系数是整个图像压缩过程的一个关键

环节。由于量化步长与量化器设计是否合理,将直接影响整个压缩系统的性能(量化误差主要来源于量化步长与量化器),所以量化步长与量化器的设计选取必须从整体上综合考虑映射变换、量化以及熵编码器之间的相互配合。

为了充分利用人眼视觉掩蔽(Masking)特性,本文将量化步长  $S(l, \theta, x, y)$  选取为

$$S(l, \theta, x, y) = S_0 T(l, \theta, x, y) \quad (7)$$

其中,  $l \in \{1, 2, 3, 4\}$  表示分解层次,  $\theta \in \{LH, HL, HH\}$  表示不同方向,  $(x, y)$  表示小波系数位置,  $S_0$  为基础量化步长,  $T(l, \theta, x, y)$  为小波域视觉掩蔽特性值。

另外,本文采用了如下自适应线性量化器,量化定义为对小波变换系数除以自适应量化步长,四舍五入取整。基本量

化公式和基本反量化公式分别为式(8)和式(9):

$$C(l, \theta, x, y) = [I(l, \theta, x, y) - S(l, \theta, x, y)/2]/S(l, \theta, x, y) \quad (8)$$

$$I(l, \theta, x, y) = C(l, \theta, x, y) * S(l, \theta, x, y) \quad (9)$$

其中,  $I(l, \theta, x, y)$  是自适应小波变换后的高频子带小波系数,  $C(l, \theta, x, y)$  为小波系数  $I(l, \theta, x, y)$  量化后的结果。  $S(l, \theta, x, y)$  表示基于人眼视觉掩蔽特性的自适应量化步长。

#### 4 低比特率自适应图像压缩编码算法框架

基于内容的低比特率自适应图像压缩编码过程如图5所示。

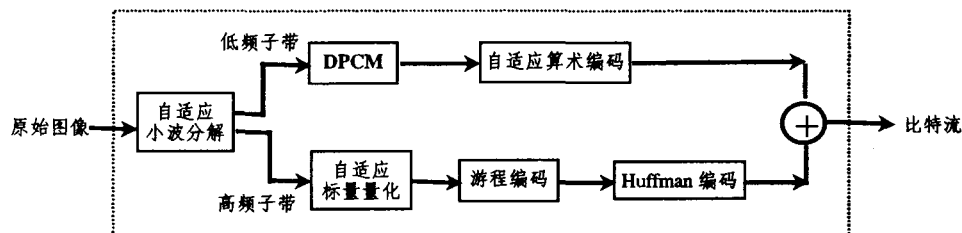


图5 低比特率自适应图像压缩算法流程图

可见,该自适应图像压缩编码算法主要包含三个部分:(1)对原始图像进行自适应小分解;(2)对最低频子带系数直接进行DPCM编码,然后再进行自适应算术编码;(3)对高频子带系数进行自适应标量量化编码,然后再进行游程编码与Huffman编码。

#### 5 仿真实验

为了验证本文自适应图像压缩算法的高效性,以下在

Pentium4/1.8GHz计算机上,以  $512 \times 512 \times 8\text{bit}$  标准图像Lena、Barbara和Mandrill为例进行了仿真实验,并与JBEG2000<sup>[9]</sup>、EZW<sup>[1]</sup>、FWP<sup>[5]</sup>(小波包算法)等压缩方案进行了比较,仿真实验结果如表1~表2和图6所示。其中,小波分解重构采用了常见的双正交9/7小波滤波器<sup>[10]</sup>,基础量化步长  $S_0$  选取为4。

表1 不同码率下重构图像的峰值信噪比(dB)

| 比特率    | Lena     |       |       |       | Barbara  |       |       |       | Mandrill |       |       |       |
|--------|----------|-------|-------|-------|----------|-------|-------|-------|----------|-------|-------|-------|
|        | JPEG2000 | EZW   | FWP   | 本文    | JPEG2000 | EZW   | FWP   | 本文    | JPEG2000 | EZW   | FWP   | 本文    |
| 0.0625 | 27.39    | 27.54 | 29.08 | 28.81 | 22.89    | 23.10 | 24.84 | 24.52 | 20.29    | 20.67 | 21.30 | 21.17 |
| 0.125  | 30.23    | 30.23 | 31.86 | 31.62 | 24.79    | 24.03 | 25.89 | 25.48 | 21.13    | 21.08 | 21.63 | 21.53 |
| 0.25   | 33.16    | 33.17 | 34.81 | 34.50 | 27.28    | 26.77 | 28.50 | 28.37 | 22.62    | 22.51 | 23.08 | 22.96 |
| 0.5    | 36.41    | 36.28 | 37.84 | 37.59 | 30.99    | 30.53 | 32.36 | 32.02 | 24.80    | 24.64 | 25.20 | 24.95 |
| 1.0    | 39.27    | 39.55 | 40.94 | 40.67 | 35.79    | 35.14 | 37.01 | 36.94 | 28.41    | 28.13 | 28.73 | 28.59 |

表2 不同码率下编/解码时间对照(标准图像Lena)

| 比特率    | 编码时间(ms) |      |      | 解码时间(ms) |      |      |
|--------|----------|------|------|----------|------|------|
|        | EZW      | FWP  | 本文   | EZW      | FWP  | 本文   |
| 0.0625 | 594      | 2125 | 892  | 601      | 934  | 673  |
| 0.25   | 748      | 3576 | 1211 | 747      | 1290 | 828  |
| 1.0    | 1194     | 4678 | 1519 | 1789     | 2367 | 1360 |

**结束语** 本文提出了一种基于内容的低比特率自适应图像压缩编码算法,该算法首先以子带能量为判别标准,根据图像内容对其实施自适应小分解,再对最低频子带系数直接进行DPCM编码及自适应算术编码,最后结合人眼视觉掩蔽特性自适应确定量化步长,并据此对高频子带系数进行标量量化及熵编码。实验结果表明:(1)相同压缩比下,本文算法的图像复原质量与FWP(小波包)算法接近,而明显优于JPEG2000和EZW算法。(2)本文算法的编解码速度略慢于JPEG2000和EZW算法,但明显快于FWP(小波包)算法。

#### 参考文献

- Shapiro J M. Embedding image coding using zerotrees of wavelet coefficients[J]. IEEE Trans. on Signal Processing, 1993, 41(12): 3445~3462
- Said A, Pearlman W A. A new, fast, and efficient image codec based on set partitioning in hierarchical trees[J]. IEEE Trans on Circuits and Systems for Video Technology, 1996, 6(3): 243~250
- 王向阳, 杨红颖. 基于人眼视觉特性的快速图像编码算法[J]. 软件学报, 2003, 14(11): 1964~1970
- 王向阳, 杨红颖. 基于高频子带自适应扫描的零树图像编码算法[J]. 计算机研究与发展, 2003, 30(3): 469~475
- Meyer F G, Averbuch A Z, Strömberg J-O. Fast Adaptive Wavelet Packet Image Compression[J]. IEEE Trans. on Image Processing, 2000, 9(5): 792~800
- 徐孟侠. 图像编码基础[M]. 北京: 人民邮电出版社, 1984
- Barni M, Bartolini F, Piva A. Improved wavelet-based watermark-

ing through pixel-wise masking[J]. IEEE Transactions on Image Processing, 2001, 10(5): 783~791

- 8 沈兰荪, 卓力, 田栋, 汪孔桥. 视频编码与低码率传输[M]. 北京: 电子工业出版社, 2001
- 9 Santa-Cruz D, Ebrahimi T, Askelof J, et al. Christopoulos. JPEG2000 still image coding versus other standards[A]. In:

Proc. of SPIE's 45th annual meeting, Applications of Digital Image Processing XXIII, San Diego, CA, 2000. 446~454

- 10 Antonini M, Barlaud M, Mathieu P, et al. Image coding using wavelet transform[J]. IEEE Trans on Image Processing, 1992, 1(2): 205~220



(a) EZW 算法复原图像

(b) FWP 算法复原图像

(c) 本文算法复原图像

图6 标准图像 Barbara 重构复原效果对照(0.25bpp)

(上接第183页)

解问题的过程。领域模型主要包含两方面的知识:(1)有关课程的内容,即叙述型知识;(2)有关应用这些知识来求解问题的知识,即过程型知识。知识表示方法采用语义网络、产生式规则等形式。

系统根据收集到的信息进行建模,通过学生模型表示学生的学习特征和学习状态;通过教师模型运用教学策略,对学生进行诊断和评估,选择或生成相应的教学信息并提示给学生。

基于知识的搜索与推理建立教师模型。教师模型的作用是结合教学策略和课程结构方面的知识,为学生选择问题,监督和评价学生的行为,当学生需要时为其选择适当的补习材料,是实现交互、进行教学的主要手段。教师模型的主要功能是:收集学生的各类数据,根据这些数据,分析学生当前所处的状态,通过诊断、评估,灵活地控制学生的学习内容和重组教学路径,决定教学策略和方法;根据教学效果进行自我修改与完善,从而不断地提高教学规则的可用性和功效性。

依据学生的行为特征与认知能力,建立学生模型。学生模型是众多学生的模拟仿真,是指在学习系统中的一种可靠的、既能反映(不同程度地)学生普遍行为特征和认知能力,又能反映学生个性特征的学习行为和认知结构的模型。学生模型的作用是准确反映学生的水平、学习能力等,利用学生模型动态可生成适合于个别化教学的策略与内容,达到实现个别化教学的目的。不同功能的学习系统,学生模型的内容及结构不尽相同,它的内容依系统的教学目标来变化,即它与专家模块(知识库)的内容和教师模块所选用的教学模式(教学推理机)紧密相关联。它所提供的数据是教学模型进行规则推导,进而提供自适应功能的依据。

整个教学过程根据学生的现实情况进行,系统自动跟踪学生的学习状态变化,根据学生的个人爱好、接受倾向、个人学习基础、个人能力,自动检索知识库并产生相应的教学信息,提供给学生,建立个性化的学习环境。

教学策略的自动调整对实现个别化教学非常关键。教学

策略智能推理机是体现教学系统智能性的最主要的模块,它根据学习者以往的学习记录及当前学习状态进行推理,再根据专家知识库提供的教学规则得出推导结果,从而选择一定的教学策略,最后采用自适应展示、自适应导航两种自适应技术将自适应效果以动态的形式展现给学生。教学策略智能推理机的实现还要涉及到诸如信息的输入输出、转换、对学生能力的预测、测试与评价等工作。

**结束语** Agent思想及网络计算环境为情感Agent的具体实现提供了框架模型,为远程教学提供新的理论依据。利用Agent思想分析远程教育系统的总体需求,设计解决方案,可充分体现教学过程的自主性;让计算机具有情感也就是让计算机具有智能,是人机和谐情感交互的重要前提,可使远程教育模式和手段更具人性化色彩,推动远程教育进入智能化的新阶段。

基于情感Agent构筑现代远程教学系统是一个值得长期研究的课题,是大有可为的,必将推动教育技术水平的全面提升和创新,将有巨大的发展潜力和应用价值,同时必将为远程教育产业带来良好的社会效益和巨大的经济效益。

## 参考文献

- 1 王志良. 人工心理学——关于更接近人脑工作模式的科学[J]. 北京科技大学学报, 2000, 22(5)
- 2 毛峡, 丁玉宽, 牟田一弥. 图像的情感特征分析及其和谐感评价[J]. 电子学报, 2001(12)
- 3 刘林森. 大脑何时接通电脑[J]. 世界科学, 2001(8)
- 4 Picard R W. Affective Computing[M]. MIT Press, London, 1997
- 5 胡包钢, 谭铁牛, 王珏. 情感计算——计算机科技发展的新课题[J]. 科学时报, 2000(3)
- 6 Wooldridge M, Jennings NR. Intelligent Agents: Theory and Practice[J]. The Knowledge Engineering Review, 1995, 10(2): 115~152
- 7 马希荣. 基于多Agent的分布式智能教学平台关键技术研究[D]. [南开大学博士学位论文]