

基于视觉特性的图象分割编码算法

黄继武 Yun Q. Shi* 戴宪华

(汕头大学电子工程系, 汕头 515063)

*(Department of ECE, New Jersey Institute of Technology, Newark, NJ 07102, USA)

摘要 提出了一种考虑人眼视觉系统特点的四叉树分割编码算法。在分析视觉系统对比度特性的基础上,首先提出了一个基于视觉敏感度的四叉树分割标准。该标准性能优于常用的块方差。基于这一标准,我们把原始图象分割为 32×32 、 16×16 的均匀块和 8×8 的待分类块。对 8×8 块进一步应用一个基于对比度和多项式近似误差的二级分类器。最终,所有图象块被分为如下三类:具有不同大小的均匀块、 8×8 平滑块和 8×8 非平滑块。这三类图象块分别采用块均值、二维一阶多项式和DCT编码进行近似。实验结果证实了所提出的算法性能大大优于JPEG国际标准。

关键词 图象压缩 四叉树 图象分割 对比度掩蔽 视觉系统

0 引言

作为一类重要的图象压缩方法,基于四叉树的分割编码一直吸引着人们的注意。近年来,压缩域(compressed domain)内进行的和基于内容(content-based)的数据操纵在多媒体信息处理领域中备受重视,四叉树编码由于其所具有的层次数据结构也更加体现出其优点。据报道,为了提高压缩效率,基于四叉树分割编码的某些实验正在进行以便采用在将于1998年底推出的MPEG-4国际标准中^[1]。

在四叉树分割中,根据某一标准,一个块如果被判定为不一致(包含不同的区域),则它被分割为4个子块。该过程重复进行,直到图象块被认为是一致的。块均匀度是测量块一致性的主要方法。块方差(测量块内像素与块均值的背离程度)则是衡量块均匀度的最常用标准^[2,3]。除此之外,边缘信息也经常使用^[4,5]。

众所周知,均方误差(Mean Squared Error (MSE))有时与视觉系统(Human Vision System (HVS))的主观视觉效果并不匹配。由于块方差实际上就是均方误差,因此块方差并不是衡量区域均匀度的理想标准。基于这种考虑,在分析视觉系统对比度特性的基础上,我们提出了一个新的四叉树分割标准。该标准考虑了对比度掩蔽(contrast

masking)特性,利用视觉敏感度来测量一个区域的均匀度,提高了分割过程与视觉系统的匹配程度。应用这一标准,我们提出了一个基于四叉树的分割和分类混合编码算法。实验结果证实了所提出的算法性能大大优于JPEG国际标准,特别是在低比特率的情况下。

本文第1节概述了所提出的算法。第2节介绍了基于对比度掩蔽特性的分割和分类方法。第3节描述了经分割和分类后各类图象块的编码。第4节给出了实验结果和结论。

1 算法概述

本文所提出的编码算法分3个步骤:分割、分类和编码。算法原理如图1所示。

原始图象首先被分裂为 32×32 的图象块。对每一图象块,我们根据对比度掩蔽特性(contrast masking)计算其视觉敏感度(visual sensitivity),作为对图象块均匀度的测量标准。如果一个块的均匀度小于某一门限,该块被分为四个子块,直到块大小为 8×8 。为了对 8×8 图象块有效地编码,一个二级的分类器被用来对这些 8×8 块进行分类。分类器的第一级利用视觉敏感度作为分类标准,把所有 8×8 块分为均匀块和非均匀块。第二级则采用多项式近

似误差作为分类参数,进一步把非均匀块分为平滑块和非平滑块。因此,原始图象最终被分割为 32×32 、 16×16 和 8×8 的均匀块、 8×8 的平滑块和 8×8

的非平滑块。针对这三类图象块的不同特点,块均值、二维一阶多项式和 DCT 被分别用来编码这些图象块。

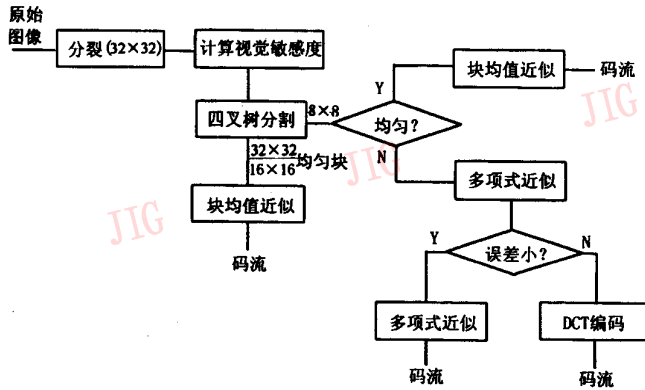


图 1 所提出编码算法的框图

2 基于对比度掩蔽特性的分割和分类

根据 Weber 定律^[6],在均匀背景 I 下,物体的可见性检测门限(对比度敏感门限) ΔI 为

$$\Delta I = 0.02 \cdot I \quad (1)$$

上式表明,背景越亮,可见性检测门限越高。

一个象素值变化的图象块可以被看为在均匀背景上迭加一个变化的信号。该信号幅度必须达到一定的强度(检测门限)才能为视觉系统所看到。因此,如果迭加的的信号幅度小于视觉系统的检测门限,该图象块可以被认为是均匀的。背景越亮,可允许迭加的信号越强。这种现象称为对比度掩蔽^[7]。换言之,对于被认为均匀的图象块,亮度较高时,允许块内的象素变化也更大些。

近年来,对视觉系统的进一步研究表明:对比度敏感门限与背景亮度的关系更接近指数规律。一些更准确的对比度敏感函数(contrast sensitivity function)已被提出^[8-10]。根据文献[9],对比度敏感门限可被表示为

$$\Delta I = I_0 \cdot \max\{1, (I/I_0)^\alpha\} \quad (2)$$

其中 I_0 为当 $I=0$ (背景亮度为零)时的对比度敏感门限。 α 为常数,根据视觉生理实验,其值为 $0.6 \sim 0.7$ 。

对于图象 $f(x, y)$ 中大小为 $n \times n$ 的块 B_k ,根据对比度敏感门限与背景亮度的非线性关系,我们定义了如下衡量块均匀度的参数:

$$d(B_k) = \frac{1}{n^2} \sum_{(x,y) \in B_k} w(m_k) \cdot \frac{|f(x,y) - m_k|}{m_k} \quad (3)$$

其中 m_k 为 B_k 的均值。 $w(B_k)$ 为考虑方程(2)后的加权因子,由下式确定:

$$w(m_k) = (1/m_k)^\alpha \quad (4)$$

方程(3)有二点不同于块方差:一是我们考虑了视觉系统的特性,采用了对比度而不是均方误差来测量图象块内信号变化的可见度;二是考虑了对比度敏感门限与背景亮度的非线性关系,引入了根据块平均亮度调整的加权因子 $w(B_k)$ (当采用方程(1)作为对比度视觉模型时, $w(B_k) \equiv 1$)。定义了块均匀度测量标准后,四叉树分割过程可表示如下:

若 $d(B_k) < T_1$, 则 B_k 被认为是均匀的;否则, B_k 不均匀。 T_1 为门限。

实验表明,利用 $d(B_k)$ 作为块均匀度测量标准用于四叉树分割,比利用块方差具有更好的主观视觉效果。

在四叉树分割中,树的深度的增加将导致辅助信息的增加。考虑到实际图象中出现较大面积的均匀区域的情况不多,我们把四叉树根节点对应的图象块大小限制为 32×32 (即一幅图象对应许多棵子树)。原始图象被分裂为 32×32 的图象块后,对每一图象块应用上述分割过程,直到图象块为均匀块或者块大小达到 8×8 。考虑到图象块尺寸太小对压缩率不利,而且为了增加树的深度,对 8×8 的块不再细分。

分割过程完成后,我们应用一个二级分类器对

8×8 图象块进行分类,以便对 8×8 图象块自适应编码,提高压缩率。分类器的第一级采用和上述分割过程相同的标准,把所有 8×8 块分为均匀和非均匀二类。即对于所有的 8×8 块 B_k , 如果 $d(B_k) < T_2$, 则 B_k 被分类为均匀块;否则, B_k 为非均匀块。 T_2 为门限。

非均匀块被分类器的第二级进一步分为平滑块和非平滑块。在一维情况下,一阶多项式常被用来拟合变化比较平缓的数据。这种方法可推广到二维场合。因此,二维一阶多项式近似误差可作为衡量数据是否平滑的标准。基于这一思路,我们将每一非均匀块用如下多项式近似:

$$g_k(x, y) = a_k \cdot x + b_k \cdot y + c_k \quad (5)$$

系数 a_k, b_k 和 c_k 可采用最小均方误差 (Minimum Mean Squared Error (MMSE)) 准则求解。用 $g_k(x, y)$ 近似 B_k 引起的误差 (MSE) 为:

$$MSE(B_k) = \| f(x, y) - g_k(x, y) \|_{(x,y) \in B_k} = \frac{1}{n} \sum_{(x,y) \in B_k} [f(x, y) - g_k(x, y)]^2 \quad (6)$$

如果 $MSE(B_k) < T_3$, B_k 被分类为平滑块;否则, B_k 为非平滑块。 T_3 为门限。

经过上述分割和分类后,原始图象被分为如下三个子集:

- \mathcal{S}_1 : 32×32、16×16 和 8×8 的均匀块;
- \mathcal{S}_2 : 8×8 平滑块;
- \mathcal{S}_3 : 8×8 非平滑块。

即

$$f(x, y) = \bigcup_{i=1}^3 \mathcal{S}_i \quad (7)$$

3 编 码

分割和分类后,原始图象被分为具有不同局部性质的图象块。根据它们所具有的特点,用不同的编码器编码不同的图象块。

(1) 均匀块 \mathcal{S}_1 。均匀块采用块均值近似。即

$$\mathcal{E}(B_k) = \{m_k\}, \quad \forall B_k \in \mathcal{S}_1 \quad (8)$$

$\mathcal{E}(\cdot)$ 代表编码器函数。

由于均匀块内的像素值差异很小,利用块均值近似后所引起的视觉误差将很小。

(2) 平滑块 \mathcal{S}_2 。由第 2 节可知,平滑块被定义

为那些用二维一阶多项式近似后均方误差较小的块(除均匀块之外)。这也暗示:对平滑块来说,二维一阶多项式是一种有效的编码方法^[11-13];而对那些像素值变化较为激烈的块,则采用其它编码方法,以便在压缩率和图象质量上得到较好的折衷。

用二维一阶多项式来编码平滑块可表示如下:

$$\mathcal{E}(B_k) = \{a_k, b_k, c_k\}, \quad \forall B_k \in \mathcal{S}_2 \quad (9)$$

其中 a_k, b_k 和 c_k 为方程 (5) 的系数,采用 MMSE 准则求解^[13]:

$$a_k = \frac{\sum_{(x,y) \in B_k} (x - x_0) \cdot f(x, y)}{\sum_{(x,y) \in B_k} (x - x_0)^2}$$

$$b_k = \frac{\sum_{(x,y) \in B_k} (y - y_0) \cdot f(x, y)}{\sum_{(x,y) \in B_k} (y - y_0)^2}$$

$$c_k = m_k$$

其中 (x_0, y_0) 表示块 B_k 的中心坐标。对 8×8 块,若左上角坐标为 (x_1, y_1) , 则 (x_0, y_0) 为 $(x_1 + 3.5, y_1 + 3.5)$ 。

(3) 非平滑块 \mathcal{S}_3 。非平滑块采用 DCT 编码。每个非平滑块的 DCT 系数被计算并量化后,用锯齿形 (zigzag) 扫描方式转化为一维序列。即

$$F_k(u', v') = \mathcal{DCT} \{ f(x, y), (x, y) \in B_k \},$$

$$B_k \in \mathcal{S}_3, 0 \leq u', v' < 8$$

$$F'_k(l) = \text{zigzag} \{ \text{int}(F_k(u', v') / Q(u', v')) \}$$

其中, $Q(u', v')$ 为对应于空间频率 (u', v') 的量化系数,这里采用 JPEG 推荐的量化表^[14]; $\text{int}(\cdot)$ 为取整函数。

在 $F'_k(l)$ 中,前 L 个低频系数用游程编码 (run-length coding)。余下的非零系数则由系数值和位置表示。即

$$\mathcal{E}(B_k) = \{ F'_k(l), 0 \leq l < L \}$$

$$\cup \{ (F'_k(l), l), F'_k(l) \neq 0 \text{ and } l \geq L \},$$

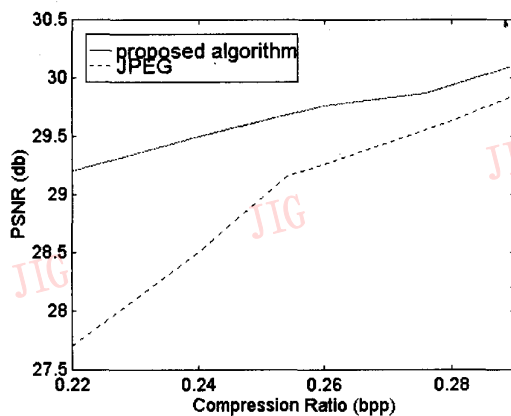
$$\forall B_k \in \mathcal{S}_3 \quad (10)$$

4 实验结果和结论

我们在“Lena”图象上测试了所提出算法的性能并与 JPEG 国际压缩标准进行了比较。JPEG 标准已推出多年,但新的标准尚未问世,因此 JPEG 仍具有比较意义。实验结果如图 2 所示。



(a)



(b)



(c)



(d)

图 2 实验结果

图 2(a)为 $512 \times 512 \times 8$ 的“Lena”原始图象。图 2(b)从峰值信噪比 (PSNR, Peak Signal-to-Noise Ratio) 的角度将所提出算法与 JPEG 进行了比较。在保持相同压缩率的情况下,本文所提出的算法的重建图象客观质量比 JPEG 明显改善。随着压缩率的进一步降低(低比特率),效果更加突出。图 2(c)和(d)给出了在 0.22 bpp 的压缩率下分别应用所提出算法和 JPEG 得到的重建图象。可以看到,所提出算法的重建图象的主观质量比 JPEG 好得多。其区别在高分辨率显示器上更加明显。重建图象的主、客观效果表明:所提出的算法压缩性能大大优于 JPEG。

在计算复杂性上,与 JPEG 相比,本文算法也有明显的降低。表 1 为二种算法对各类图象块所需要的大概计算量。

从表 1 可以看出,均匀度和多项式近似的计算量远小于 DCT,在所提出算法中,只有少部分图象块需采用 DCT 编码。因此,其计算复杂性比 JPEG 低

得多。

表 1 所提出算法与 JPEG 计算量的比较

图象块(8×8)	JPEG 计算量	所提出算法计算量
均匀	DCT 编码	均匀度计算
平滑	DCT 编码	均匀度计算 + 多项式计算 (含误差计算)
非平滑	DCT 编码	均匀度计算 + 多项式计算 (含误差计算) + DCT 编码

表 1 未考虑分割过程中 32×32 、 16×16 块均匀度的计算。由于均匀度计算量比 DCT 小得多,且对于 32×32 、 16×16 的均匀块,其子块无需再计算。可以认为,对于大多数实际图象,分割所需的计算量与表 1 相比可以忽略。

门限值 T_1 和 T_2 用来控制压缩率。在实验中,它们的取值分别为 $T_1, T_2 = 1.5 \sim 4.5$, $T_3 = 100 \sim 450$ 。

本文把视觉模型结合到二叉树分割编码中,提出了一个利用视觉系统特点的分割和分类混合编码算法。主要的贡献是:(1)提出了一个基于对比度掩蔽特性的二叉树分割标准。该标准的性能优于最常用的块方差。(2)在新的分割标准的基础上,提出了一个有效的编码算法。该算法结合了二叉树分割和分类的优点。在重建图象质量和压缩率上达到了较好的折衷,其压缩性能大大优于 JPEG,且保持较小的计算量。

参考文献

- 1 Brailean J, Li W, Ostermann J, Zhang Y -Q. MPEG-4 video coding standard. ISCAS Tutorial on MPEG-4, Chapter 3.1, June, 1997.
- 2 El-Sakka M R. A segmentation criterion for digital image compression. ICASSP'95, 1995, 4.
- 3 Vaisey J, Gersho A. Image compression with variable block size segmentation. IEEE Trans. Signal Processing, 1992, 40(8):2040~2060.
- 4 Vernazza G. A new splitting strategy for adaptive transform coding. Signal Processing, 1993, 34(3).
- 5 Lee C -H, Chen L -H. Novel image compression method using edge-oriented classifier and novel predictive noiseless coding method. IEE Proceedings: Vision, Image, and Signal Processing, Dec. 1997, 144(6): 361~368.
- 6 Gonzalez R C, Wintz P. Digital image processing(2nd). Addison-Wesley Publishing Co., 1987.
- 7 Watson A B. DCT quantization matrices visually optimized for individual images. SPIE: Human Vision, Visual Processing and Digital Display IV, 1913: 202~216, 1993.
- 8 Legge G E, Foley J M. Contrast masking in human vision. J Opt Soc Am, 1980, 70(12).
- 9 Watson A B. Efficiency of a model human image code. J Opt Soc Am, 1987, 4(12).
- 10 Paly S. The visible differences predictor: An algorithm for the assessment of image fidelity. Digital Images and Human Vision, M I T Press, 1993, 177~206.
- 11 Kunt M, Bqenard M, Leonardi R. Recent results in high-compression image coding. IEEE Trans CAS, 1987, CAS-34(11).
- 12 De Natale F G B, Desoli G S, Giusto D D, Vernazza G. Polynomial approximation and vector quantization: A region-based integration. IEEE Trans COM, Feb /Mar/Apr 1995, 43(2/3/4).
- 13 黄继武,戴汝为. 图象块编码—分类的方法. 中国图象图形学报, 1997, 2(12).
- 14 Wallace G. The JPEG still image compression standard. IEEE Trans on Consumer Electronics, 1992, 38(1): xviii-xxxiv.

黄继武 1962年10月出生。1982年1月本科毕业于西安电子科技大学电子对抗专业。1987年12月研究生毕业于清华大学通信与电子系统专业。现任汕头大学电子工程系副教授。主持完成多项研究课题,发表论文30多篇。目前主要研究方向为图象处理、图象数据压缩和图象水印技术。



A Segmentation-Based Image Coding Algorithm Using the Features of Human Vision System

Huang Jiwu, Yun Q. Shi* and Dai Xianhua

(Department of Electronic Engineering, Shantou University, Shantou 515063)

* (Department of ECE, New Jersey Institute of Technology, University Heights, Newark, NJ 07102, USA)

Abstract In this paper, a segmentation-based hybrid image coding algorithm is proposed which takes the features of the human vision system (HVS) into account. First, a novel criterion for quadtree-based segmentation using contrast masking is presented. This segmentation criterion performs better than the frequently used criterion, block variance. With this criterion, the original image is segmented into uniform blocks of 32×32 , 16×16 and blocks of 8×8 to be classified. Then a two-level classifier, based on contrast masking sensitivity and polynomial approximation error, is further applied to the blocks of 8×8 such that all blocks are classified into a few categories: uniform blocks of different sizes, smooth blocks of 8×8 and non-smooth blocks of 8×8 . At last, block mean, polynomial approximation and DCT coding are applied to the blocks accordingly. The experimental results demonstrate that the proposed algorithm performs much better than JPEG.

Keywords Image compression, Quadtree, Image segmentation, Contrast masking, Human vision system