

基于三维离散余弦变换的静止彩色图象压缩

丁绪星 朱日宏 李建欣 杜小祥

(南京理工大学电光学院, 南京 210094)

摘要 为了既能获得较高的压缩比,又能保证较好的图象质量,提出了一种基于三维离散余弦变换(3D DCT)的静止彩色图象编码方法。在此算法中,首先将原始图象分割成互不覆盖的 8×8 子块;然后根据局部能量的大小将各子块分为4类,再把属于同一类的子块组合到一起,形成三维立方体并对三维立方体进行三维离散余弦变换,从而提高压缩效率。计算机模拟结果表明,此算法在压缩比高于40时,其平均峰值信噪比约为29dB,解压后的图象质量有较高的保真度,在同等情况下,其比JPEG标准提高约5dB。

关键词 计算机图象处理(320·6040) 图象压缩 三维离散余弦变换

中图分类号: TN919.81 **文献标识码**: A **文章编号**: 1006-8961(2003)06-0637-04

Still color Image Compression Based on 3-D DCT

DING Xu-xing, ZHU Ri-hong, LI Jian-xin, DU Xiao-xiang

(School of Electronic Engineering and Photoelectric Technology, Nanjing University of Science & Technology, Nanjing 210094)

Abstract In this paper, A three dimensional discrete cosine transform(3D DCT) coder for still color images is presented, which can not only achieve high compression ratio but also ensure good image quality. In the proposed method, the original image is splitted into non-overlapped 8×8 blocks and segments subblocks of the image into four different classes based on local energy magnitude. Then those subblocks with the same class are gathered to form 3D cuboids, and employs 3D DCT to the 3D cuboids. This improves the efficiency of compression. Computer simulation results show that the \overline{PSNR} (average Peak Signal-Noise-Ratio) of reconstructed images achieve 29dB with high fidelity when the compression ratios are higher than 40. Compared to JPEG, it advances more than 5dB in the same situation. The proposed method achieves higher fidelity of decoded images than by JPEG.

Keywords Computer image processing, Image compression, 3D discrete cosine transform

0 引言

DCT(离散余弦变换)的性能接近K-L变换,且有多种快速算法,变换后的系数便于编码;DCT将二维图象从空域映射到频域,在空域中像素能量的分布是均匀的,而在频域中能量则集中到少数低频系数,即具有良好的能量压缩和空间去相关特性。另外,DCT还具有算法简单,易于硬件实现等优点,所以被广泛应用于当前的国际编码标准(如JPEG, MPEG, H. 261等)中。

传统的基于DCT的方法,压缩一个典型的8bit

彩色图象,一般在压缩比小于10时,解压的图象无明显方块效应,但在高压缩比时,重建图象的相邻块将有明显的方块效应。因为每块主要由少数低频系数表示,且各子块是独立处理的,没有考虑块间的相关性和块沿边的不连续性,从而限制了压缩比的提高。目前有许多技术可以减小方块效应,如基于MRF的混合变换编码算法^[1]就很有效。传统的去除方块效应的方法采用后处理技术,即在DCT后再进行低通滤波^[2]、中值滤波^[3]、POCS^[4]、CLS^[5]、基于区域的迭代方法^[6]等。但这些编码效率低,同时还会带来其他的人为效应,如跨越块边界的信号变模糊、振铃效应等。当图象子块间变化较平缓时,这些方法

能达到较高的压缩比,但当图象子块间的变化剧烈时,将失去作用。

一般情况下,3D DCT 主要用来压缩视频信号。基于 3D DCT 的彩色图象压缩编码方法是基于频域,对图象子块分类,用三维压缩的概念压缩二维静止彩色图象,实验结果表明,此方法压缩效率高,同时在图象子块间的差别较大以及较高压缩比的情况下也无明显的方块效应。

1 3D DCT 编码方法

用 3D DCT 来压缩具有相同特征的图象子块,其编码系统模型如图 1 所示。

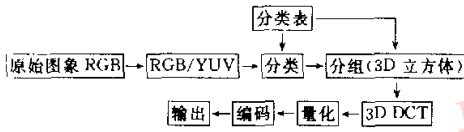


图 1 3D DCT 压缩编码系统模型

根据人类视觉特性可知,人眼对亮度信号和色度信号的敏感程度不同,即对亮度信号敏感,而对色度信号不敏感,所以先将 R、G、B 图象转换为 Y、U、V 图象。其中,亮度(Y)信号携带的图象信息所占比重最大,而色度(U、V)信号相对地要小得多,因此,R、G、B 到 Y、U、V 的空间转换实际上是能量集中的过程。之后再进行分类、分组、3D DCT、量化、编码等操作。

1.1 分类

此过程是分别将 Y、U、V 图象进行 2D DCT 后,根据局部能量的大小再把每个分量分为 4 类,即首先将 Y 分量(U、V 分量和 Y 分量作相同操作)分为互不重叠的 8×8 图象子块,然后再对每个子块进行 2D DCT。由于变换后图象的能量主要集中在子块的左上角,因此按图 2 在其中选择 6 个系数,按照下式

$$a = v_1^2 + v_2^2 + h_1^2 + h_2^2 + d_1^2 + d_2^2 \quad (1)$$

来计算分类值 a,如果 $a \leq A_1$,将其归为类型 1;如果 $A_1 < a \leq A_2$,将其归为类型 2;如果 $A_2 < a \leq A_3$,将其归为类型 3;其他则归为类型 4。其中阈值 A_1, A_2, A_3 是由重复试验后获得的经验值,可以进行自适应地调整,以获得最佳结果。每个子块所属类型的信息存储在分类表中,分类表按行编码,它在分类表中所占位数的开销仅为 $2/64 = 0.03125\text{bpp}$ 。

LC	v_1	v_2					
h_1	d_1						
h_2	d_2						

图 2 用于分类的系数选取

1.2 分组

分组的目的是根据分类表,将相同类的子块组合到一起形成 3D 立方体,作为 3D DCT 的输入数据。在解码端,立方体根据分类表重建,为了准确地重建,分类表的存储与传输必须准确无误,设分类表中子块间的相对位置如图 3 所示。

	$(k, j-1)$	
$(k-1, j)$	(k, j)	$(k+1, j)$
	$(k, j+1)$	

图 3 分类表中子块间的相对位置

输入为图象的分类表,其中的每个数据为一个 8×8 子块的分值,定义分组算法如下:

- (1) 按行扫描分类表,设扫描开始位置为 $C[0][0]$;
- (2) 当 $k \neq 0$ 且 $j \neq 0$ 时,根据位置 $C[k][j]$ 的分值 $V[k][j]$,进行如下操作:如果 $V[k-1][j] = V[k][j]$,则将位于 $C[k][j]$ 和 $C[k-1][j]$ 的子块归为同一组,并标记 $C[k][j]$ 子块已被分组;如果 $V[k][j-1] = V[k][j]$,则将位于 $C[k][j]$ 和 $C[k][j-1]$ 的子块归为同一组,并标记 $C[k][j]$ 子块已被分组;
- (3) 按行扫描分类表下一个分类值,重复第 2 步,直至所有的子块被分组。

一种典型的分组结果如图 4 所示(图中数字 1,

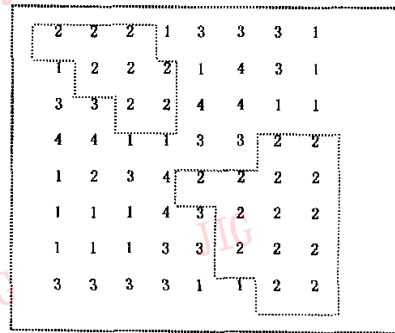


图 4 一种根据分类表的分组结果

2,3,4 分别表示类型 1,2,3,4)。在图 4 中,有许多相邻的图象子块属于类型 2,这样的集合有两个,按行扫描将属于类型 2 的子块聚集到一起形成一个 3D 立方体 $8 \times 8 \times 22$ (22 为立方体长度,即属于类型 2 的子块个数)。

1.3 3D DCT

对形成的 3D 立方体定义 3D DCT 正变换

$$F(u, v, \omega) = \frac{8C(u)C(v)C(\omega)}{n \times n \times m} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \left\{ f(i, j, k) \cos \left[\frac{(2i+1)u\pi}{2m} \right] \cos \left[\frac{(2j+1)v\pi}{2n} \right] \cos \left[\frac{(2k+1)\omega\pi}{2n} \right] \right\} \quad (2)$$

3D IDCT(三维 DCT 逆变换)

$$f(i, j, k) = \sum_{u=0}^{m-1} \sum_{v=0}^{n-1} \sum_{\omega=0}^{n-1} \left\{ C(u)C(v)C(\omega) \cos \left[\frac{(2i+1)u\pi}{2m} \right] \cos \left[\frac{(2j+1)v\pi}{2n} \right] \cos \left[\frac{(2k+1)\omega\pi}{2n} \right] \right\} \quad (2)$$

其中

$$C(u), C(v), C(\omega) = \begin{cases} 1/\sqrt{2} & u, v, \omega = 0 \\ 1 & \text{其他} \end{cases} \quad (4)$$

m 为 3D 立方体的长度。可见,长度越长,压缩比越高,尤其是对那些同一背景占大部分的图象和变化不大的图象,经过分组后,能获得较长的立方体,这不但提高了压缩比,而且大大减小了方块效应,这是因为把图象子块以三维立方体的形式整块进行 DCT,而不是传统的逐一进行 DCT。

在量化阶段,对 3D DCT 后的系数进行量化

$$F_q(u, v, \omega) = \frac{F(u, v, \omega)}{Q(u, v, \omega)} \quad (5)$$

其中, $F_q(u, v, \omega)$ 为量化后的系数, $F(u, v, \omega)$ 为量化前的系数, $Q(u, v, \omega)$ 为量化表中的量化因子。静态图象压缩编码失真的主要来源在量化阶段,本文采用文献[7]的方法产生量化表,因此种量化的均方误差小,在高压缩比时,不会出现明显的方块效应。

将量化后的系数按二维“之”字形扫描进行 Huffman 游程编码。

2 实验结果

在 CPU 为 1G, 内存为 256M 的 PC 机上进行实验,测试图象为 Lena, 大小为 512×512 ; 实验的有关数据如表 1 所示(表中 C_r 为压缩比, \overline{PSNR} 为平均峰值信噪比, T 为压缩时间)。本文的算法与 JPEG 相比稍复杂一些,需要的时间要长,实验时采用 2D DCT 快速算法,同时限制 3D 立方体的长度不要超过 20,因为长度过长会大大增加计算量,使压缩时间加长,内存开销过大。利用 3D DCT 技术,在高压缩比时,解压后的图象无明显的方块效应,其质量明显高于用 JPEG 标准解压后的图象(图 5)。

表 1 实验结果

Lena	JPEG	本文算法	JPEG	本文算法
C_r	40.46	40.50	80.40	80.60
\overline{PSNR} (dB)	23.33	28.81	15.50	27.20
T (s)	1.2	9.41	1.51	14.78

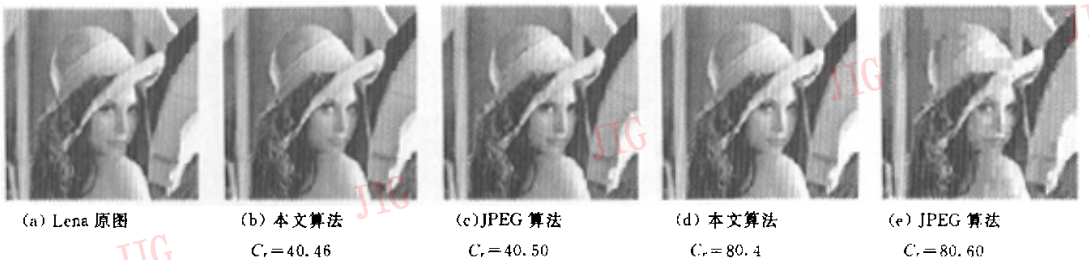


图 5 测试图象 Lena 原图和本文算法、JPEG 算法的解码图象

3 结 语

用 3D DCT 方法压缩图象充分利用了图象 3 个空间方向上的相关性,解压后图象的质量无疑比 2D DCT 的方法好,但是同时也带来了计算量大的缺点。

分类的阈值 A_1, A_2, A_3 的选取方法可以多种多样,通过不断试验能取得更好的效果。进一步的实验表明,在此方法中进行 2D DCT 时,采用和其他变换方法相结合,将会取得更好的压缩性能,此方法适合于对压缩时间要求不高的场合,下一步的工作是进一步改进算法,使压缩时间变短,然后用硬件来实现此算法。

参考文献

- 1 Merier Thomas, Ngan King N, Crebbin Gregory. Reduction of blocking artifacts in image and video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(3): 490~499.
- 2 Tzou K H. Post-filtering of transform-coded images [A]. In: Proc. SPIE Applications of Digital Image Processing XI [C], San Diego, CA, USA, 1998, 974: 121~126.
- 3 Hsu Y F, Chen Y C. A new adaptive separable median filter for removing blocking effects [J]. IEEE Transactions on Consumer Electron, 1993, 39(8): 510~513.
- 4 Zaklor A. Iterative procedures for reduction of blocking effects in transform image coding [J]. IEEE Transactions on Circuits Systems Video Technology, 1992, 2(3): 91~95.
- 5 Crouse M, Ramchandran K. Nonlinear constrained least squares estimation to reduce artifacts in block transform-coded images [A]. In: Proceedings IEEE International Conference Image Processing (ICIP'95) [C], Washington DC, 1995, 1: 462~465.
- 6 Merier T, Ngan K N, Crebbin G. A region-based algorithm for enhancement of images degraded by blocking effects [A]. In: Proceedings of IEEE Tencon'96 [C], Perth, Australia, 1996, 1: 405~408.
- 7 Monro D M, Shrerlock B G. Optimal quantisation strategy for DCT image compression [J]. IEEE Proc. Vis Image Signal Process, 1996, 143(1): 10~14.



丁绪星 1971年生,现为南京理工大学通信与信息系统专业博士研究生.从事的研究有图象压缩与传输、光纤通信技术、图象压缩算法的硬件设计与实现.



朱日宏 1964年生,博士,现任南京理工大学电光学院副院长、教授、博士生导师.长期从事光学图象实时处理、光通信技术、光电检测技术的科研、教学工作.



李建欣 1977年生,现为南京理工大学光学工程专业博士研究生.主要研究方向为图象压缩与传输、视频信号处理.

杜小祥 1976年生,现为南京理工大学通信与信息系统专业硕士研究生.主要研究方向为图象压缩及其硬件实现.