

# 静态彩色图像的多维 DCT 变换压缩

邵凌一 李久贤 余加兵

(东南大学自动控制系, 南京 210096)

**摘要** 为了在高信噪比条件下来对静态彩色图像进行高倍压缩, 提出了一种具有自适应功能的多维离散余弦变换图像编码算法。该算法首先将图像分割成互不覆盖的  $8 \times 8$  子块; 然后抽取每个子块的统计特征, 再根据子块的方差和均值来判断子块内部信息的丰富程度, 其中对于内部信息丰富的子块, 可利用三维 DCT 编码来消除块内相邻像素间、色彩空间 Y、U、V 各分量间的相关性, 而对于内部信息不丰富的子块, 则利用四维 DCT 编码来消除块内相邻像素间、色彩空间 Y、U、V 各分量间、相邻子块间的相关性; 最后对变换系数进行量化编码。实验结果证明, 此算法可以提供比 JPEG 更高的压缩率和更高的平均峰值信噪比。

**关键词** 图像处理 离散余弦变换 DCT 图像压缩

中图法分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2004)07-0865-04

## Still Color Image Compression Based on Multi-D DCT

SHAO Ling-yi, LI Jiu-xian, YU Jia-bing

(Department of Automatic Control, Southeast University, Nanjing 210096)

**Abstract** An adaptive multi-dimensional discrete cosine transform (multi-D DCT) algorithm are proposed in order to obtain high compression for still color image under high signal-to-noise ratio. In this algorithm, the original image is splitted into non-overlapped  $8 \times 8$  blocks. Then the statistic characteristic of each block is calculated and assess the abundance of inner information in each block based on averages and variances. To the block of which the inner information is abundant, correlations between adjacent pixels of each block and YUV channels of color space are exploited through 3-dimensional discrete cosine transform (3-D DCT); To the block of which the inner information is not, correlations among adjacent pixels of each block, YUV channels of color space and adjacent blocks are exploited through 4-dimensional discrete cosine transform (4-D DCT). The transform coefficients are coded by quantization further. Experimental results have proved that the proposed algorithm can achieve higher compression ratio and The  $\overline{PSNR}$  (average Peak Signal-Noise-Ratio) of reconstructed image is higher than by JPEG.

**Keywords** Image processing, discrete cosine transform (DCT), Image compression

## 1 引言

图像数据压缩一直是图像通信中的重点和难点, 而且随着 Internet 的普及及图像应用范围的不断扩大, 对图像的压缩效果、压缩比、编解码的时间都提出了新的要求。如今已提出了多种压缩算法, 其中 DCT (离散余弦变换) 由于性能接近 K-L 变换, 且存在多种快速算法, 其变换后的系数容易编码, 同时能够有效避免 Gibb 现象, 因此被广泛应用于当前

的国际编码标准 (JPEG、MPEG、H. 261) 中。

然而传统的基于 DCT 的编码标准, 对于静态图像在以下几方面仍然没有完全消除数据之间的相关性, 从而限制了压缩比和峰值信噪比等性能的进一步提高:

(1) 这些编码方法大多针对灰度图像, 而对于彩色图像的压缩编码, 一般都是先将 R (红) G (绿) B (蓝) 数据转换成 Y (亮度) U (色度) V (色差) 数据, 然后将灰度图像的编码方法独立应用于各个分量上。实际上, 由于 RGB 各颜色之间的相关性很强,

尽管将其变换成 YUV 数据后相关性有所削减,但仍有极大的视觉冗余,因此,如果分别对 YUV 各分量进行压缩编码,就不能充分利用其潜在的相关性,也就限制了压缩比的进一步提高。

(2) 采用传统的 DCT 编码,当进行高压缩比压缩时,则重建图像的相邻块间有明显的方块效应。这是因为每块主要由少数低频系数表示,且各块之间相互独立处理,没有考虑到块之间的相关性和块边界的不连续性,从而限制了压缩比的进一步提高。

目前,由于图像压缩技术已经由原来单一的图像分块压缩发展成为能够根据图像中的对象、背景、复杂度等的不同进行自适应编码,因而大大提高了压缩比和图像还原质量。例如,目前在 MPEG-4、H264 标准中均采用了基于对象的压缩编码技术,即在编码前,首先要通过对视频序列进行分析来从原图像中分割出各个视频对象(video object),然后分别对每个视频对象进行单独编码。如今,在 DCT 的变换编码中,如何根据对象来进行自适应编码已经成为研究的热点。在文献[1]中,作者通过计算图像子块的均值,并根据均值之间的差值大小来区分不同子块之间的相关性,从而实现了自适应编码。这种方法是一种非常有效的根据对象进行编码的方法,本文即是在此基础上,对此算法进行了改进。实验证明,改进后的算法不仅能够达到更高的压缩比,同时对图像的细节还可以产生很好的还原。

## 2 多维 DCT 变换

在文献[1]的算法中,由于考虑到了块之间的相关性,从而消除了块间的冗余;但其并没有考虑到 YUV 各分量之间也有很大的冗余,而通过消除 YUV 分量间的相关性,则可以进一步提高图像的压缩比,同时可保证图像还原后的质量。具体的压缩算法如下:

(1) 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,对色度信号不敏感。由于将 RGB 图像转换成 YUV 图像,实际上是能量集中的过程,所以在传输亮度信号时,要保证其清晰度,而对色度信号可适当降低清晰度。

$$Y = R \times 0.299 + G \times 0.587 + B \times 0.114$$

$$U = R \times -0.169 + G \times -0.332 + B \times 0.500 + 128$$

$$V = R \times 0.500 + G \times -0.149 + B \times -0.0813 + 128$$

因此,本文提出的算法重点是在压缩过程中,在对亮度信号 Y 进行有效压缩的同时,尽量降低亮度分量 Y 的失真度。由于 RGB 数据到 YUV 数据的转换去除了彩色图像各分量之间的部分冗余,从而为下面的进一步压缩奠定了基础。

(2) 抽取图像统计特征 首先把 YUV 图像分割成互不重叠的  $8 \times 8$  图像子块,然后将 Y(亮度)分量图像分块,并计算每个子块的方差  $M_{i,j}$  及平均值  $A_{i,j}$  ( $i, j$  为图像子块坐标)。之所以要计算图像子块的方差是为了区分图像子块是复杂块,还是平坦块;计算平均值是为了判别相邻子块之间的相似程度。对于 U(色度)、V(色差)分量,只需要计算每个子块的方差即可。

(3) 归类 对于 Y(亮度)分量的图像,首先设定方差阈值  $M_{t,y}$  作为判断子块属性的标准,若  $M_{i,j} > M_{t,y}$ ,则标记为复杂块,否则标记为平坦块;然后按行扫描各块,对于复杂块,归入类  $S_{complex}$ ;对于平坦块,若相邻两个子块的像素值的算术平均值之差小于阈值  $\zeta$ ,则说明这两块之间具有很大的相似度,应将其归为相同类。其算法流程图如图 1 所示。

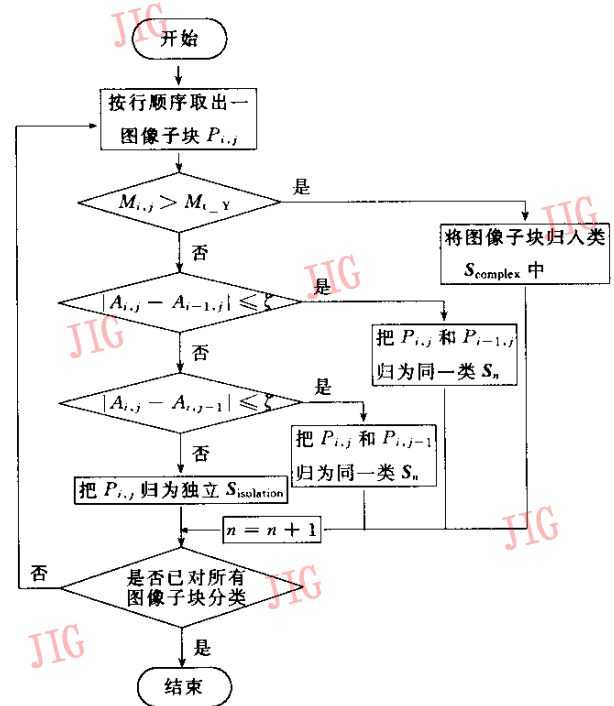


图 1 多维 DCT 变换压缩算法流程图

图 1 中,  $P_{i,j}$  为 Y 分量中第  $i, j$  图像子块;  $M_{i,j}$  为 Y 分量中第  $i, j$  子块的灰度方差;  $A_{i,j}$  为 Y 分量中第  $i, j$  子块的灰度平均值。

对于 U、V 分量,首先分别设定各分量方差阈值  $M_{t\_u}$ 、 $M_{t\_v}$ ;然后把 U、V 分量中子块灰度方差大于阈值的,都放入集合  $S_{\text{complex}}$  中,当处理完所有的分量以后,会形成以下集合:  $S_{\text{complex}}$ 、 $S_{\text{isolation}}$ 、 $S_1 \dots S_n$ ;接着根据集合将子块所属类型写入分类表,最后对分类表进行 Huffman 编码,这样它所占位数开销就很小。

(4) 自适应多维 DCT 编码 对于每幅静态彩色图像经过上述处理后,会形成 Y、U、V 3 幅灰度图像和一份分类表。然后编码就根据分类表,采取如下的自适应多维 DCT 编码策略:

对于  $S_{\text{complex}}$  集合中的图像子块,由于子块图像复杂,包含很多细节信息,因此即使子块的均值与相邻的子块很接近,也应该单独进行压缩,以避免损失细节内容;对于  $S_{\text{isolation}}$  集合中的图像子块,由于它与相邻的图像子块灰度均值相差很大,因此,也不具备与相邻子块连接起来进行压缩的条件。由此可见,对于  $S_{\text{complex}}$ 、 $S_{\text{isolation}}$  集合中的图像子块,只采用 3D-DCT,即可消除色度和亮度分量之间的相关性以及子块内部数据的相关性。

3D-DCT 正变换为

$$F(u,v,w) = \frac{8C(u)C(v)C(w)}{8 \times 8 \times 3} \sum_{i=0}^2 \sum_{j=0}^7 \sum_{k=0}^7 \{f(i,j,k) \times \cos[\frac{(2i+1)u\pi}{2 \times 3}] \cos[\frac{(2j+1)v\pi}{2 \times 8}] \cos[\frac{(2k+1)w\pi}{2 \times 8}]\}$$

3D-DCT 逆变换为

$$f(i,j,k) = \frac{8C(u)C(v)C(w)}{8 \times 8 \times 3} \sum_{u=0}^2 \sum_{v=0}^7 \sum_{w=0}^7 \{F(u,v,w) \times \cos[\frac{(2u+1)u\pi}{2 \times 3}] \cos[\frac{(2v+1)v\pi}{2 \times 8}] \cos[\frac{(2w+1)w\pi}{2 \times 8}]\}$$

其中

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

对于  $S_1 \dots S_n$ ,可分别对集合内的子块进行 4D-DCT,其增加的维数就是相同类型子块的个数。可见,如果一个集合内类型相同的子块数越多,则压缩比越高。对于灰度值变化不大的图像,若相同类型的子块数目多,则其压缩比也将得到很大提高;同时,将具有相类似特征的图像放在一起进行 DCT,还可以大大减少方块效应。

4D-DCT 正变换为

$$F(t,u,v,w) = \frac{16C(t)C(u)C(v)C(w)}{n \times 8 \times 8 \times 3} \sum_{h=0}^n \sum_{i=0}^2 \sum_{j=0}^7 \times \sum_{k=0}^7 \{f(h,i,j,k) \cos[\frac{(2h+1)u\pi}{2 \times n}] \times \cos[\frac{(2i+1)v\pi}{2 \times 3}] \cos[\frac{(2j+1)w\pi}{2 \times 8}] \cos[\frac{(2k+1)u\pi}{2 \times 8}]\}$$

(n 为相同类型子块的个数)

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

4D-DCT 逆变换为

$$f(h,i,j,k) = \frac{16C(t)C(u)C(v)C(w)}{n \times 8 \times 8 \times 3} \sum_{t=0}^n \sum_{u=0}^2 \times \sum_{v=0}^7 \sum_{w=0}^7 \{F(t,u,v,w) \cos[\frac{(2t+1)u\pi}{2 \times n}] \times \cos[\frac{(2u+1)u\pi}{2 \times 3}] \cos[\frac{(2v+1)v\pi}{2 \times 8}] \cos[\frac{(2w+1)w\pi}{2 \times 8}]\}$$

### 3 量化编码

静态图像经过多维 DCT 变换后,还要通过量化器将变换系数进行压缩编码。静态图像压缩一般可采用文献[2]所用方法产生量化表,因为这种量化的方法均方误差小,且在高压缩比时,不会出现明显的方块效应。由于在量化后出现了多零值,因此对系数应采用游程编码,即首先顺序对子阵进行 Zig Zag 扫描,将多维矩阵转换为一维形式,再进行 Huffman 编码。

### 4 实验结果

为了验证本文算法的效果,在 CPU 为 P IV 1.7G,内存为 256M 的 PC 机上进行了实验,测试图像大小为 608×504 的 Windows 自带图像。实验有关结果数据如表 1 所示(表中  $C_r$  为压缩比,  $\overline{PSNR}$  为平均峰值信噪比)。

表 1 Windows 图像实验结果对比

	JPEG		本文算法		
	Y	U	Y	U	V
$C_r$	30.72		31.06		
$\overline{PSNR}$ (dB)	32.83	32.67	31.98	42.76	39.05
			38.82		

图 2 为本文算法与 JPEG 算法的实验仿真图像。实验表明,采用自适应的多维 DCT 技术,不仅提高了图像的压缩比,而且图像解码时无明显的方块效应,其质量明显高于 JPEG 标准解压后的图像。



图 2 不同算法压缩后恢复图像

表 2 Lena 图像实验结果对比

	文献[1]算法	本文算法
$C_r$	40.50	40.60
$\overline{PSNR}(\text{dB})$	28.81	30.17

## 5 结 论

多维 DCT 编码具有自适应的功能,由于它可将图像多方面的相关性去除,因此具有很高的压缩比,但由于其算法比较复杂,需要的计算量非常大,因此,在确定图像子块间的相关性时,可以采用设定几个类的阈值来把不同的子块归类到不同的类中,而不是采用计算相邻子块的均值差的方法来分类,进一步的实验表明,这种方法还可以有效降低压缩时间。同时,计算多维 DCT 可以采用文献[3]所提出的方法进行 DCT 分解,其乘法次数仅为行列法的 50%。另外,由于这种算法计算结构十分规则,模块性和对称性强,其流水线结构易于 VLSI 的并行实现,因此,可以通过硬件来实现。

### 参 考 文 献

- 1 丁绪星,朱日宏,李建欣等. 基于三维离散余弦变换的静止彩色图像压缩[J]. 中国图象图形学报,2003,8A(6):637~640.

表 2 为本算法的实验结果与文献[1]中算法的结果比较,由表 2 可见,本文的算法能够提供更大的峰值信噪比值。

- 2 Monro D M, Shrerlock B G. Optimal quantisation strategy for DCT image compression [J]. IEEE Proceeding Vision Image Signal Processing, 1996, 143(1):10~14.
- 3 吴红文,李久贤,夏良正. 一种新的二维离散余弦变换快速算法 [J]. 东南大学学报,1996,26(2):7~15.



邵凌一 1980 年生,2002 年获东南大学自动控制系学士学位,现为该系模式识别与图像处理方向硕士研究生。主要研究方向为数字图像处理、压缩编码和视频监控。

E-mail:slingyi@hotmail.com



李久贤 1961 年生,东南大学自动控制系教授。长期从事数字图像处理和模式识别方面的科研、教学工作。



余加兵 1979 年生,2002 年获东南大学自动控制系学士学位,现为该系模式识别与图像处理方向硕士研究生。主要研究方向为模式识别和数字图像处理。