

# 基于三维离散余弦变换的彩色图象压缩编码

桑爱军 陈贺新

(吉林大学通信工程学院测控与通信仪器系, 长春 130022)

**摘要** 离散余弦变换对图象信号有近似最优的去相关能力,但多维的变换公式一直没有给出,为此深入研究了三维离散余弦变换,提出了任意尺寸的三维函数  $f(x,y,z)$  的正交离散余弦变换公式,克服了以前系数的取值必须相等的缺点,并将之应用于彩色静止图象的压缩编码中,使得彩色图象的 R、G、B 3 帧可以作为一个整体同时进行变换,极大地去除了图象 R、G、B 3 帧间的相关性.理论分析和实验结果表明,在大幅度地增加压缩比的同时,峰峰信噪比也有明显提高,并且与国际标准 JPEG、MPEG 有很好的兼容性.

**关键词** 三维离散余弦变换 彩色图象 压缩编码

中图分类号: TN919.81 文章标识码: A 文章编号: 1006-8961(2002)12-1269-05

## Color Image Coding Based on Three-dimensional DCT

SANG Ai-jun, CHEN He-xin

(School of Communication Engineering, Jilin University, Changchun 130022)

**Abstract** Discrete Cosine Transform(DCT) can decrease the relativity among the images, but we didn't know the multidimension DCT accurately. In this paper, three-dimensional Discrete Cosine Transform(DCT) has been researched, its calculate method, transform and inverse transform were given, and any sizes of three-dimensional function  $f(x,y,z)$  DCT is proposed, that must be cube before. As we known, it is identity physics model unify reflection into that of color image's three frame, possess same texture, same fringe and same gray change, the information of the each frame almost could reflect the source image except color, the vision similitude are very strong. In other methods, R, G, B data were transformed into Y, Cr, and Cb to decrease the color information. It is clear that if we encode Y Cr Cb data separately, among each component potential relativity can't be decreased enough. So we regard R, G, and B as a whole, transform them in the same time with three-dimensional DCT. It can decrease relativity between frames than before. Simulation results demonstrate good reconstruction image quality and high compression ratio. It has good compatible with JPEG and MPEG standard.

**Keywords** Three-dimensional DCT, Color image, Compression coding

## 0 引言

图象由于具有直观、具体和生动等特点,并包含十分丰富的信息而被人们所重视.虽然彩色图象更符合人们的视觉习惯,但是在大量的有关图象压缩的文章中,人们研究较多的却是灰度图象和彩色运动图象,而对于彩色静止图象,则在将 R、B、G 信号转换成 Y、U、V 信号后,进行亚采样,然后对每个分

量再分别应用灰度图象的研究成果进行处理<sup>[1]</sup>.其实,彩色图象的 R、G、B 各颜色分量间的相关性很强,转换后生成的 Y、U、V 的相关性有所削弱,但仍有很大的视觉冗余.彩色图象的 R、G、B 3 帧是对同一个物理模型的统一反映,具有相同的纹理、边缘和灰度变化梯度,每一帧都几乎可以反映除颜色外的全部信息,具有很强的视觉相似性.如果分别对 Y、U、V 各分量进行压缩编码,就不能充分利用彩色图象中的各分量间潜在的相关性,从而也就限制了压

基金项目:国家自然科学基金(60172046)

收稿日期:2002-01-30; 改回日期:2002-05-13

缩比和峰峰信噪比等性能的进一步提高。

自从 1974 年 Ahmed 等人提出离散余弦变换 (DCT) 以来<sup>[2]</sup>, DCT 技术在图象压缩编码方面得到了广泛应用, 国际标准 JPEG, MPEG 中也采用二维的 DCT. 这是因为图象含有大量的低频信息, 而 DCT 对具有 Gauss-markov-1 统计特性的信号有近似最优的去相关性能力. 虽然在压缩效率上略逊于具有最好压缩能力的 K-L 变换, 但其可做到的高效处理性是 K-L 变换所难以比拟的, 所以其被广泛应用于图象编码方面<sup>[3]</sup>.

本文发展了三维 DCT 技术, 提出了任意尺寸的三维函数  $f(x, y, z)$  的正交离散余弦变换公式, 克服了以前系数的取值必须相等的缺点, 将彩色图象的 R、G、B 3 帧看作一个整体, 分别对行、列、纵三维进行离散余弦变换. 之前人们只研究了方阵的三维变换, 所以三维离散余弦变换只应用于  $8 \times 8 \times 8$  的变换中, 例如灰度图象的运动序列等, 限制了 DCT 的应用. 本文研究了任意尺寸的三维离散余弦变换, 对于彩色图象可以进行任意尺寸的变换.

### 1 三维离散余弦变换

三维离散余弦变换的概念提出较早, 基本上是从一维到二维再到三维的概念推广而来<sup>[1]</sup>, 但它的定义一直以来比较含糊, 或者只给出  $8 \times 8 \times 8$  的变换公式.

#### 1.1 定义

对于函数  $a(x, y, z)$  ( $0 \leq x \leq M-1; 0 \leq y \leq N-1; 0 \leq z \leq L-1$ ), 三维离散余弦变换

$$p(u, v, w) = \sqrt{\frac{2^3}{LMN}} c(u)c(v)c(w) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \sum_{z=0}^{L-1} a(x, y, z) \times \cos\left[\frac{(2x+1)\pi u}{2M}\right] \cos\left[\frac{(2y+1)\pi v}{2N}\right] \cos\left[\frac{(2z+1)\pi w}{2L}\right]$$

$(u = 0, 1, \dots, M-1; v = 0, 1, \dots, N-1; w = 0, 1, \dots, L-1)$

三维离散余弦逆变换

$$a(x, y, z) = \sqrt{\frac{2^3}{LMN}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \sum_{w=0}^{L-1} c(u)c(v)c(w) p(u, v, w) \times \cos\left[\frac{(2x+1)\pi u}{2M}\right] \cos\left[\frac{(2y+1)\pi v}{2N}\right] \cos\left[\frac{(2z+1)\pi w}{2L}\right]$$

$(x = 0, 1, \dots, M-1; y = 0, 1, \dots, N-1; z = 0, 1, \dots, L-1)$

其中

$$c(u) = \begin{cases} \frac{1}{\sqrt{2}} & u = 0 \\ 1 & \text{其他} \end{cases}$$

$$c(v) = \begin{cases} \frac{1}{\sqrt{2}} & v = 0 \\ 1 & \text{其他} \end{cases}$$

$$c(w) = \begin{cases} \frac{1}{\sqrt{2}} & w = 0 \\ 1 & \text{其他} \end{cases}$$

显然

$$p(u, v, w) = \sqrt{\frac{2}{L}} c(w) \sum_{z=0}^{L-1} \left[ \sqrt{\frac{2}{N}} c(v) \sum_{y=0}^{N-1} \left[ \sqrt{\frac{2}{M}} c(u) \sum_{x=0}^{M-1} a(x, y, z) \times \cos\left[\frac{(2x+1)\pi u}{2M}\right] \cos\left[\frac{(2y+1)\pi v}{2N}\right] \right] \cos\left[\frac{(2z+1)\pi w}{2L}\right] \right]$$

即为三维可分离的离散余弦变换, 由一维的可逆性即可以证明. 显然, 正反变换的系数可以取不同的值, 比如  $\frac{2^3}{LMN}, 1$  等等, 只要满足乘积为  $\frac{2^3}{LMN}$  即可. 在本文中, 为了对称性, 取系数相等.

#### 1.2 二维特例

当  $f(x, y, z)$  3 个参数中有一个为常数时, 三维离散余弦变换即成为通常的二维 DCT:

$$p(u, v) = \frac{2}{\sqrt{MN}} c(u)c(v) \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} a(i, j) \times \cos\left[\frac{\pi(2i+1)u}{2M}\right] \cos\left[\frac{\pi(2j+1)v}{2N}\right]$$

$$a(i, j) = \frac{2}{\sqrt{MN}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} p(i, j) c(u)c(v) \times \cos\left[\frac{\pi(2i+1)u}{2M}\right] \cos\left[\frac{\pi(2j+1)v}{2N}\right]$$

$i, u = 0, 1, \dots, M-1; j, v = 0, 1, \dots, N-1.$

其中,

$$c(u) = \begin{cases} \frac{1}{\sqrt{2}} & u = 0 \\ 1 & \text{其他} \end{cases}$$

$$c(v) = \begin{cases} \frac{1}{\sqrt{2}} & v = 0 \\ 1 & \text{其他} \end{cases}$$

同样, 可以将其表示为如下矩阵形式

$$P_{M \times N} = C_{M \times M}^T A_{M \times N} C_{N \times N}$$

其中,  $C_{N \times N}$  中

$$c_{ij} = \begin{cases} \sqrt{\frac{2}{N}} \cos \frac{i\pi(2j+1)}{2N} & i \neq 0 \\ \frac{1}{\sqrt{N}} & i = 0 \end{cases}$$

$C_{M \times M}$  中

$$c_{ij} = \begin{cases} \sqrt{\frac{2}{M}} \cos \frac{i\pi(2j+1)}{2M} & i \neq 0 \\ \frac{1}{\sqrt{M}} & i = 0 \end{cases}$$

可以证明,逆变换为

$$A_{M \times N} = C_{M \times M}^T P_{M \times N} C_{N \times N}$$

$M=N=8$  时,即为 JPEG 中使用的  $8 \times 8$  DCT 变换.

这样就可以运用矩阵形式对任意尺寸的二维函数进行离散余弦变换,这对于图象变换来说是极为方便的.

### 2 彩色图象的三维表示及三维 DCT 变换

如前所述,一幅尺寸为  $I \times J$  的彩色图象,由 3 帧同样大小的 R、G、B 灰度图象组成,因此可用三维函数  $f(x, y, z)$  表示,其中,  $(1 \leq x \leq I, 1 \leq y \leq J)$  表示的是彩色图象的空间位置,  $1 \leq z \leq 3$  代表彩色图象的 3 个颜色分量.彩色图象的这种表示就把其 在各像素的位置关系以及相应的各个分量间的关系建筑在同一模型中,这样也就可以用三维函数的方法处理图象,并充分利用彩色图象在空间和各颜色分量间的相关性,实现进一步的能量压缩.事实上,

有很多人为了使用三维 DCT,将 8 帧连续的灰度图象划块,组成  $8 \times 8 \times 8$  的数据组进行通常意义下的三维离散余弦变换,这不仅限制了彩色图象的运用,而且限制了运动图象的时间取值长度.而运用本文中的任意尺寸三维离散余弦变换技术,却可以将变换进行得任意自如.

在 JPEG 和 MPEG 标准中,图象被划分成  $8 \times 8$  的子块,进行 DCT 变换.为了与其兼容及利用已有的科技成果,并充分考虑到方块效应及计算复杂度,在本文中采取了  $8 \times 8 \times 3$  的子块分割方法.

### 3 仿真实验

对  $256 \times 256$  像素的 pepper 和 dog 彩色图象进行压缩编码仿真实验.采用目前比较流行的矢量量化<sup>[4]</sup>来压缩系数.实验流程如图 1 所示,并与基本彩色系统的仿真结果及小波变换的结果进行了对比,实验数据及结果对比见表 1,重建图象见图 2.

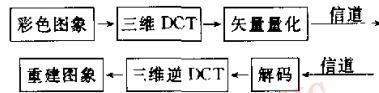


图 1

表 1 基本彩色系统与三维 DCT 变换的压缩效果比较

图象	压缩方法	$C_r$	PSNR				效果比较	
			$PSNR_R$	$PSNR_G$	$PSNR_B$	$\overline{PSNR}$	压缩比	$\overline{PSNR}$
pepper	基本彩色系统	32.38	14.65	31.04	34.78	31.31	-	-
	本文方法	42.67	28.34	27.74	28.40	28.16	2.91	-4.22
		38.4	33.35	32.99	33.01	33.12	2.62	0.74
dog	基本彩色系统	34.91	41.63	42.12	42.02	41.93	2.38	9.55
		13.42	29.50	32.03	29.53	30.35	-	-
	本文方法	42.67	27.18	27.54	28.45	27.72	3.18	-2.63
		38.4	30.81	31.13	32.20	31.41	2.86	1.06
		34.91	37.01	38.03	38.94	37.99	2.60	10.27
小波变换	23.3	32.76	34.59	31.98	33.11	-	-	
	34.8	30.89	32.24	30.21	31.11	-	-	

图象数据是从典型测试图象库中选取的  $256 \times 256 \times 24$ bit 的彩色图象.

压缩比为

$$C_r = (256 \times 256 \times 24) / b_{p0}$$

其中,  $b_{p0}$  为压缩编码后的总比特数.

图象质量的客观评价采用峰峰信噪比,即

$$PSNR = 10 \log \frac{255 \times 255}{MSE} \text{ (dB)}$$

其中,  $MSE$  是重建图象的均方误差.对于彩色图

象,重建图象有 R、G、B 3 帧,相应的峰峰信噪比表示为:  $PSNR_R, PSNR_G, PSNR_B$ . 平均峰峰信噪比为:

$$\overline{PSNR} = (PSNR_R + PSNR_G + PSNR_B) / 3$$

实验结果表明,由于将图象的 R、G、B 3 帧表示在一个函数中,运用任意尺寸的三维离散余弦变换进行处理,明显降低了颜色分量间的相关性,图象压缩比大幅度提高的同时,信噪比也明显提高.其不仅优于常用的基本彩色压缩方法,也优于运用小波变换对彩色静止图象每帧单独压缩编码的方法<sup>[5]</sup>.



R 帧 PSNR=36.81



G 帧 PSNR=31.13



B 帧 PSNR=32.20

(c) dog 图象在压缩比 38:4 时的恢复图象



R 帧 PSNR=27.18



G 帧 PSNR=27.54



B 帧 PSNR=28.45

(b) dog 图象在压缩比 42:67 时的恢复图象



R 帧 PSNR=33.35

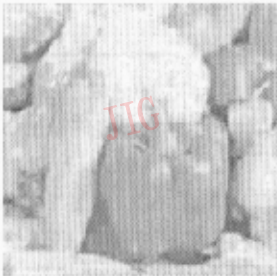


G 帧 PSNR=32.98



B 帧 PSNR=33.01

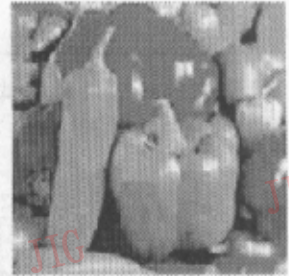
(c) pepper 图象在压缩比 38:4 时的恢复图象



R 帧 PSNR=28.34



G 帧 PSNR=27.74



B 帧 PSNR=28.40

(d) pepper 图象在压缩比 42:67 时的恢复图象

图2 不同压缩比下的恢复图象

## 4 结 论

离散余弦变换与图象的统计特性相吻合,是准最优的变换,由于有快速算法而得到了广泛的应用。彩色图象更接近人类的视觉习惯,本文研究了任意尺寸的三维离散余弦变换,将图象的空间位置与颜色分量表示于一个统一的三维函数中,并将之应用于彩色图象的压缩编码中,大幅度降低了图象 R、G、B 3 帧间的相关性。仿真实验结果证明,该方法优于基本颜色系统中的二维离散余弦变换的性能,能量进一步集中,压缩比和峰峰信噪比同时提高,而且与目前流行的 JPEG, MPEG 有较好的兼容性,具有良好的发展前景。

### 参 考 文 献

- 1 何振业. 多维数字信号处理[M]. 北京:国防工业出版社.
- 2 Ahmed N., Natarajan T., Rao K R. Discrete cosine transform[J]. IEEE Trans. on Computer, 1974, 1:90~93.

- 3 高文. 多媒体数据压缩技术[M]. 北京:电子工业出版社,1994.
- 4 Linde Y, Buzo A, Gray R M. An algorithm for vector quantizer design[J]. IEEE Trans. Commun., 1980, 28(1):84~95.
- 5 基于小波二重树编码的彩色图象压缩[J]. 电子科学学刊, 2000, 4:672.



**桑爱军** 1973 年生,1994 年获华中理工大学学士学位,1997 年获大连理工大学硕士学位,现在吉林大学任教,1999 年攻读博士学位。主要研究方向为多维信号处理、图象编码、变换理论等。



**陈贺新** 1949 年生,1982 年、1985 年、1989 年先后获吉林工业大学电子工程系工学学士、工学硕士、工学博士学位,现为吉林大学通信工程学院教授,博士生导师。主要研究兴趣是多维数字信号处理、人工神经网络、计算机视觉等。

## ELSA Gloria4 700XGL 专业绘图显卡

ELSA Gloria4 700XGL 专业绘图显卡采用 nVIDIA Quadro4 700XGL 芯片,搭配 64MB DDR-SDRAM,支持 VAG+VAG、VAG+DVI 模式的双屏幕输出显示,结合 nView 的多荧幕显示技术,可在极大程度上拓展用户的工作区域,并提供给设计师足够的创意与表现空间,非常适用于重视工作效率和采购成本的专业用户。

ELSA Gloria4 700XGL 专业绘图显卡采用的 nVIDIA 最新 Quadro4 700XGL 芯片拥有完整的可编程化双顶点着色引擎,能够使得影像效果逼真展现。同时,Quadro4 700XGL 芯片具有动态内存带宽管理功能,并能优化 OpenGL 和 DirectX 驱动程序的高效组合,其提供的强大闭塞面剔除功能和每秒处理 5 千 4 百万多边形生成、贴图及渲染功能,使得用户可以得到即时的动画效果,极大地提高了工作效率。