

# DCT 域图象处理和特征提取技术

刘 艳 李宏东

(浙江大学信电系, 杭州 310027)

**摘 要** 现今,大量的图象与视频信息都是以压缩数据格式进行存储和传输的, DCT(Discrete Cosine Transform 离散余弦变换)是目前应用最为广泛的多媒体数据压缩技术之一。直接在 DCT 域实现如视频编辑、特征提取等传统空域处理技术,能够避免繁琐的压缩数据编、解码操作,减少处理时间和数据处理量,节省内存空间,这一技术对于高速海量的数据处理场合,如 Internet 信息检索、视频编辑和检索、远程监视图象的理解等,是很有吸引力的,因此其是近年来国际上有关领域的研究热点之一,本文对近年来文献中所见的 DCT 域图象处理和特征提取技术进行了回顾和综述,并在此基础上对其发展方向进行了探讨。

**关键词** 计算机图象处理(520·6040) 综述 DCT 图象编码 压缩域处理 视频编辑 特征提取

**中图分类号**: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2003)02-0121-08

## Image and Video Processing Techniques in the DCT Domain

LIU Yan, LI Hong-dong

(Dept. of Information and electronics, Zhejiang University, Hangzhou 310027)

**Abstract** Nowadays volumes of image and video data are stored and transferred in the form of compressed data structure. The Discrete Cosine Transform(DCT) is one of the most widely used digital compression techniques for multimedia. Data processing such as video editing or feature extraction directly in the DCT domain other than in the pixel domain will save the time needed for the compressed data to be coded or decoded as well as the memory for buffering by processing less data. Such techniques are desirable for the in-time massive data processing applications as Internet multimedia retrieval, video editing, therefore have been hotly discussed during recent years. In this paper, we present a review of the data procession techniques in the DCT domain proposed in the literatures, and try to figure out the advance direction in the future.

**Keywords** Computer image processing, Summarize, DCT, Image coding, Compressed domain processing, Video editing, Feature extraction

## 0 引 言

如今,大量的数字图象和视频信息都是以压缩数据格式进行存储和传输的,传统的图象处理和视频编辑技术是在空间域进行的,一般可分为 3 步,即解码压缩数据、空域处理、再编码还原为压缩格式,传统方案在实现中存在两个主要问题,一是编解码过程的计算量大,二是在整个处理过程中涉及到压缩域和空域数据,数据量大,需要耗费较大的内存空间。

压缩处理的思想是避开解码和再编码的过程,直接在压缩域进行图象或视频处理操作,这样不仅能避免解码、编码相关复杂操作,减少数据处理量,节省内存空间,而且还可能利用压缩域数据的性质,取得更好的数据处理效果。这一技术对于高速海量的数据处理场合,如 Internet 信息检索、视频编辑和检索、远程图象(如军事侦察航片,卫片)的理解等,是很有吸引力的,因此是近年来国际上有关领域的研究热点之一。

常见的数字压缩技术有 DFT、DCT、小波变换

等. 离散余弦变换 DCT 技术是现在应用最广泛的多媒体数字压缩技术. 图象、视频压缩国际标准 JPEG、MPEG 和 H. 261 等都采用了 DCT 技术. 统计数据显示, 一般  $8 \times 8$  的 DCT 系数中, 只有约 10% 的数据为非零(典型为 7~16 个), 而且即便是采用了快速算法, 在传统图象处理场合中, DCT 与 IDCT 的计算量也要各占总计算量的 40%, 总共是 80%<sup>[1-3]</sup>. 研究 DCT 域的数字处理技术具有典型性和现实性. DCT 为线性正交变换, DCT 域的处理方法还可以为其他线性正交变换域的数据处理方法, 如 DFT 域、小波变换压缩域等提供借鉴.

## 1 DCT 变换的定义

一维  $N$  点序列  $x(n)$  的 DCT 变换<sup>[3]</sup>  $X_c(k)$  定义为

$$X_c(k) = \frac{2}{N} C(k) \sum_{n=0}^{N-1} x(n) \cos \frac{(2n+1)\pi k}{2N}$$

$$n, k = 0, 1, \dots, N-1$$

DCT 反变换定义为

$$x(n) = \sum_{k=0}^{N-1} C(k) X_c(k) \cos \frac{(2n+1)\pi k}{2N}$$

$$n, k = 0, 1, \dots, N-1$$

其中,  $C(0) = \frac{1}{\sqrt{2}}, C(k) = 1, k = 1, 2, \dots, N-1$ .

在 JPEG、MPEG 等图象、视频压缩国际标准中所采用的二维 DCT 变换定义为

$$X_c(u, v) = \frac{2}{N} C(u) C(v) \left[ \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x(i, j) \times \right.$$

$$\left. \cos \frac{(2i+1)\pi u}{2N} \cos \frac{(2j+1)\pi v}{2N} \right]$$

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

且  $N=8$ , 即对  $8 \times 8$  的图象子块进行 DCT 变换.

DCT 变换是一种线性正交变换, 变换核各矢量间单位正交. 二维 DCT 变换是对称可分离的, 即二维 DCT 变换可分解为行列方向的两次一维 DCT 变换. DCT 变换计算量适中, 性能优于 DFT 等其他变换, 去相关压缩能力接近 KLT 最佳变换, 故在数据压缩领域有广泛应用. 现行的图象视频标准多是对  $8 \times 8$  的数据块进行 DCT 操作.

## 2 DCT 域的性质和数学运算

利用 DCT 的性质, 可以直接在 DCT 域中施加多种数学运算. 它们主要有乘法-卷积运算、线性运

算、矩阵运算、几何变换运算等. 这些 DCT 域中的数学运算关系是 DCT 域数据处理的理论基础.

### 2.1 乘法-卷积关系

在文献[3]、[4]中探讨了 DCT 域与时域的乘法-卷积关系. 乘法-卷积关系是基于对  $x(n)$  的镜像对称扩展, 即定义

$$\hat{x}(n) = x(n), n = 0, 1, \dots, N-1$$

$$\hat{x}(n) = x(-1-n), n = -N, \dots, -N+1, \dots, -1$$

此时  $x(n)$  的 DCT  $X_c(k)$  与  $\hat{x}(n)$  的 DFT  $\hat{X}_i(k)$  间有关系式

$$X_c(k) = 2C(k) \exp(-j\pi k/2N) \hat{X}_i(k)$$

$$k = 0, 1, \dots, N-1$$

同样文献[4]中证明, 若对  $C(k)$  进行扩展, 将其在  $k=0$  到  $N-1$  区域外的数据规定为 0, 则

对  $Y_c(k) = X_c(k) H_c(k)$ , 有关系式

$$y(n) = \hat{x}(n) \otimes \hat{h}(n) \otimes z(n)$$

其中,  $z(n) = 4 \left\{ \left[ \frac{1}{2\sqrt{2}} - 1 \right] + \cos[\pi(2n-1) \times (N-1)/4N] \frac{\sin(2n-1)\frac{\pi}{4}}{\sin(2n-1)\frac{\pi}{4N}} \right\}$ ,  $H_c(k)$  为滤波器频率

响应,  $Y_c(k)$  和  $y(n)$  分别为对  $X_c(k)$  滤波后的频域和时域响应,  $h(n)$  为滤波器的冲击响应,  $\hat{x}(n)$  和  $\hat{h}(n)$  分别为  $x(n)$  和  $h(n)$  的镜像对称扩展,  $\otimes$  为圆周卷积符号. 文献[5]中利用此性质在忽略  $z(n)$  的条件下, 直接在 DCT 域中进行了滤波操作, 由于忽略了  $z(n)$ , 滤波结果中有明显伪边缘.

文献[3]在文献[4]的基础上, 对乘法-卷积关系进行了改进, 将  $C(k)$  在  $-N$  到  $-1$  的值扩展为  $C(-k)$ , 即为 1. 当  $h(n)$  为实数且偶对称时,  $h(n)$  的 DFT 将仍然是实数且偶对称.  $h(n)$  的此性质与人的视觉系统传输特性相符.

此时若有

$$Y_c(k) = X_c(k) H_c(k), k = -N, -N+1, \dots, N-1$$

则

$$y(n) = \hat{x}(n) \otimes \hat{h}(n), n = 0, 1, \dots, N-1.$$

上述性质可推广到二维.

文献[3]中的 DCT 域乘法-卷积关系不但在形式和计算量上得到了简化, 还由于滤波器  $h(n)$  与人的视觉系统传输特性一致, 改善了滤波效果. 文献[4]中对二维图象的滤波试验表明, 滤波效果是令人满意的.

### 2.2 线性性质与运算

图象压缩标准 JPEG 大致可以分为以下几个步骤:先对图象  $h(i, j)$  的  $8 \times 8$  空域块进行 DCT 变换,再用量化表  $q(u, v)$  对 DCT 系数  $H[u, v]$  进行量化,得到量化后的 DCT 系数  $H_Q[u, v]$

$$H_Q[u, v] = \text{IntegerRound} \left[ \frac{H[u, v]}{q(u, v)} \right]$$

然后对矩阵  $H_Q[u, v]$  进行 Zig-Zag 扫描,再进行 RLE 行程编码(Run-Length-Encoded),得到 RLE 数据结构块;最后用算术编码或 Huffman 编码压缩 RLE 数据块。

根据 DCT 变换的线性性质,文献[6]中推导了时域中的点运算与 DCT 系数间的关系,从而可直接对 JPEG 中的 RLE 数据块进行操作,同时利用量化后的 DCT 数据零值多的特点来进一步提高运算效率。这些 DCT 域的处理算法可用于图象渐变,添加字幕等视频编辑场合,与传统方法相比,其运算速度可以快 50~100 倍。时域中的点运算与 DCT 系数间的关系主要有:

(1) 常数乘 在空域中,图象  $h(i, j)$  和  $f(i, j)$  间若有关系式  $h(i, j) = af(i, j)$ ,  $a$  为常数,则对量化后的 DCT 系数,相应有关系式:

$$H_Q[u, v] = \frac{aq_F(u, v)}{q_H(u, v)} F_Q[u, v]$$

其中,  $H_Q[u, v]$ 、 $F_Q[u, v]$  为量化后的 DCT 系数,  $q_H(u, v)$ 、 $q_F(u, v)$  分别为其量化表。

(2) 常数加 空域中,图象  $h(i, j)$  和  $f(i, j)$  间若有关系式  $h(i, j) = f(i, j) + \beta$ ,  $\beta$  为常数,则对应量化后的 DCT 系数间有关系式

$$H_Q[u, v] = \frac{q_F(u, v)}{q_H(u, v)} F_Q[u, v] + \frac{8\beta}{q_H(u, v)} \delta(u)\delta(v)$$

其中,  $\delta(0) = 1$ , 对  $u \neq 0, \delta(u) = 0$ 。

当  $q_F/q_H = \gamma$  时,上式简化为

$$H_Q[0, 0] = \gamma F_Q[u, v] + \frac{8\beta}{q_H(0, 0)}$$

$$H_Q[0, 0] = \gamma F_Q(u, v) \quad (u, v) \neq (0, 0)$$

(3) 点相加 空域中,图象  $h(i, j)$  和  $f(i, j)$ 、 $g(i, j)$  间若有关系式  $h(i, j) = f(i, j) + g(i, j)$ , 则对应量化后的 DCT 系数间有关系式

$$H_Q[u, v] = \frac{q_F(u, v)}{q_H(u, v)} F_Q[u, v] + \frac{q_G(u, v)}{q_H(u, v)} G_Q[u, v]$$

(4) 点相乘 图象  $h(i, j)$  和  $f(i, j)$ 、 $g(i, j)$  间若有关系式  $h(i, j) = af(i, j)g(i, j)$ , 其中  $a$  为常数,则量化后的 DCT 系数间有关系式

$$H_Q(u, v) = \sum_{\substack{v_1, v_2 \\ u_1, u_2}} F_Q(v_1, v_2) G_Q(w_1, w_2) W_Q(v_1, v_2, w_1, w_2, u_1, u_2)$$

其中,变换系数阵  $W_Q(v_1, v_2, w_1, w_2, u_1, u_2) = \frac{aq_F(v_1, v_2)q_G(w_1, w_2)}{64q_H(u_1, u_2)} W(u_1, v_1, w_1) W(u_2, v_2, w_2)$

$$W(u, v, w) = \sum_i C(i, u) C(i, v) C(i, w)$$

$$C(i, u) = A(u) \cos \frac{(2i+1)}{16} u\pi \quad A(0) = \frac{1}{\sqrt{2}},$$

$u \neq 0, A(u) = 1$ 。

由于在 Zig-Zag 扫描后,  $8 \times 8$  的 DCT 数组已变为一维数据流,故对 RLE 结构有下式

$$H_Q(z) = \sum_{x, y} F_Q(x) G_Q(y) W_Q(x, y, z)$$

其中,  $x, y, z$  分别为 DCT 系数 Zig-Zag 扫描后的序号,  $H_Q(z)$ 、 $F_Q(x)$ 、 $G_Q(y)$  和  $W_Q(x, y, z)$  分别为  $h(i, j)$ 、 $f(i, j)$  和  $g(i, j)$  量化后的 DCT 系数和变换系数阵  $W_Q(v_1, v_2, w_1, w_2, u_1, u_2)$  Zig-Zag 扫描后的表达式。

注意到对典型的 DCT 系数矩阵,  $F_Q(x)$ 、 $G_Q(y)$  的绝大多数数值都为零,而  $W_Q(x, y, z)$  只有四分之一的矩阵元素非零,非零元素也只有 5 种可能取值<sup>[6]</sup>,故上式的计算量并不会太大。

文献[1]中详细考察了  $F_Q(x)$ 、 $G_Q(y)$ 、 $W_Q(x, y, z)$  与  $H_Q(z)$  的关系,利用 DCT 的正交性和对称性,对上述算法进行了进一步的优化,使得在 DCT 域用卷积实现点相乘操作可比在空域的传统操作快 4~5 倍。

### 2.3 矩阵运算的实现

由于 DCT 为线性变换, DCT 变换可以表示为矩阵乘法的形式实现。矩阵乘法、加法的 DCT 变换和矩阵 DCT 变换后的乘法、加法是等同的。故简单的 DCT 域内滤波操作可以借用矩阵乘法来实现,如 DCT 域内的分辨率变换等。

对二维 DCT 变换,把空域二维数据  $x(i, j)$  和二维 DCT 数据  $X_c(u, v)$  表达为  $N \times N$  维矩阵  $x$  和  $X_c$ 。后, DCT 变换便能表示为  $X_c = TXT^{-1}$  的矩阵乘法形式,其中  $T$  为 DCT 变换矩阵,具有  $T^{-1} = T^T$  的性质。DCT 反变换可表示为  $x = T^{-1}X_cT$ 。

对于空域的一个  $N \times 2N$  的图象块  $p = [g^1 | g^2]$ , 若希望每行隔点抽样为一个新的  $N \times N$  图象块  $g$ , 则在空域可表达为  $g = pS = [g^1 | g^2]S$ 。其中  $S$  为  $2N \times N$  抽样矩阵

$$S = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

在 DCT 域内,由矩阵乘法的结合律, $g$  的 DCT 变换  $G_c$  有关系

$$G_c = T[g \quad | \quad g^T]ST^{-1} = T[g^T \quad | \quad g]T_2^{-1} \times T_2ST^{-1} \\ = [G_c^T \quad | \quad G_c^T]T_2ST^{-1} = [G_c^T \quad | \quad G_c^T]F$$

其中,  $F = T_2ST^{-1}$ ,  $T_2 = \begin{bmatrix} T & 0 \\ 0 & T \end{bmatrix}$

### 2.4 几何变换

由 DCT 的定义式直接推导出对 DCT 块内数据在空域逆时针旋转  $90^\circ$ 、 $180^\circ$ 、 $270^\circ$  的操作,以及关于垂直、水平中轴与对角线对调数据的操作,在 DCT 域内都可以通过交换数据位置和数据符号的简单操作来实现<sup>[7-9]</sup>。设变换前空域数据为  $f(i, j)$ , 变换后数据为  $g(i, j)$ , 相应的 DCT 数据为  $F_c(u, v)$ ,  $G_c(u, v)$ ,  $i, j, u, v = 0, 1, \dots, 7$ , 则有:

(1) 列对调  $g(i, j) = f(7 - i, j)$ , 对应 DCT 域:  $G_c(u, v) = \cos(u\pi)F_c(u, v)$ , 将奇数行的 DCT 系数符号改变即可。

(2) 行对调  $g(i, j) = f(i, 7 - j)$ , 对应 DCT 域:  $G_c(u, v) = \cos(v\pi)F_c(u, v)$ , 将奇数列的 DCT 系数符号改变即可。

(3) 以主对角线为轴对调  $g(i, j) = f(j, i)$ , 对应 DCT 域:  $G_c(u, v) = F_c(v, u)$ , DCT 系数行列交换。

(4) 以辅对角线为轴对调  $g(i, j) = f(7 - i, 7 - j)$ , 对应 DCT 域  $G_c(u, v) = \cos(u\pi)\cos(v\pi)F_c(u, v)$ , DCT 系数行列交换, 行列号奇偶不同的系数改变符号。

(5) 旋转  $90^\circ$ , 对应 DCT 域  $G_c(u, v) = \cos(v\pi)F_c(u, v)$ , DCT 系数行列交换, 再改变奇数列系数符号。

(6) 旋转  $180^\circ$ , 对应 DCT 域  $G_c(u, v) = \cos(u\pi)\cos(v\pi)F_c(u, v)$ , DCT 系数行列号奇偶不同的系数改变符号。

(7) 旋转  $270^\circ$ , 对应 DCT 域  $G_c(u, v) = \cos(u\pi)F_c(v, u)$ , DCT 系数行列交换, 再改变奇数行系数符号。

空域常见的几何变换,如任意角度旋转、剪裁、透视变换等,都可分解为提取行或列,再移位、重组的简单操作。上述 DCT 几何变换性质与 DCT 矩阵操作性质相结合后,只需要 4 种提取行,6 种移位的矩阵操作便可在 DCT 域内实现空域的几何变换<sup>[8,9]</sup>。对存在非整数位移量的情况,还需要对操作结果进行相应的插值滤波(低通),以消除伪边缘效应。对变换后的 DCT 数据进行适当截断,可达到一定的抗伪边缘效应效果。

由于矩阵操作的复杂性,在 DCT 域实现空域通用几何变换的方法<sup>[8,9]</sup>,除某些特定的几何变换外,在计算复杂度上虽然并没有多少改进,但却揭示了特殊几何变换与 DCT 系数间的简单对应关系,具有通用性,可以为纹理等具有几何对称性质图案的 DCT 域特征提取提供参考。

### 2.5 反变换连续化

DCT 和 IDCT 可以视为离散线性变换形式

$$DCT: G(m) = T\{g(n)\} = \sum_{n=0}^{N-1} g(n)f_n(m)$$

其中,  $f_n(m)$  为 DCT 正变换核:

$$IDCT: g(n) = T^{-1}\{G(m)\} = \sum_{m=0}^{N-1} G(m)r_m(n)$$

其中,  $r_m(n)$  为 DCT 反变换核。

文献[10]中证明,在对称延拓的条件下,对 IDCT 可以将 DCT 反变换核  $r_m(n)$  连续化为  $r_m(x)$ , 从而将  $g(n)$  连续化为  $\hat{g}(x) = \sum_{m=0}^{N-1} G(m)r_m(x)$ 。通过数学推导,即得到以下的 DCT 微分、积分、移位、放缩等公式:

(1) 微分

$$T\{g'(n)\} = \sum_{n=0}^{N-1} g'(n)f_n(m) \\ = \sum_{n=0}^{N-1} f_n(m) \sum_{p=0}^{N-1} G(p)r'_p(n) \\ = \sum_{p=0}^{N-1} G(p) \sum_{n=0}^{N-1} f_n(m)r'_p(n)$$

(2) 移位

$$T\{g(n+a)\}(m) = \sum_{n=0}^{N-1} \left( \sum_{p=0}^{N-1} G(p)r_p(n+a) \right) f_n(m) \\ = \sum_{p=0}^{N-1} G(p) \sum_{n=0}^{N-1} f_n(m)r'_p(n+a)$$

(3) 放缩

$$T\{g(an)\}(m) = \sum_{p=0}^{N-1} G(p) \sum_{n=0}^{N-1} r_p(an)f_n(m)$$

(4) 积分

$$T \left\{ \left[ \int g(x) dx \right] \right\}_{x=n} (m) = \sum_{p=0}^{N-1} G(\beta) \sum_{a=0}^{N-1} \left\{ \left[ \int r_p(x) dx \right] \right\}_{x=an} f_n(m)$$

其中,  $r'_p(n)$  为  $r'_p(x)$  在  $x=n$  处的值,  $r_p(an)$  为  $r_p(x)$  在  $x=an$  处的值。

这几个公式可直接应用于 DCT 域图象边缘提取等数据处理场合。

### 3 DCT 域的图象处理技术

上述 DCT 变换的各种数学变换性质是现有文献中 DCT 域图象处理技术的主要理论根据。现有文献中的 DCT 域图象处理技术大致可以分为 3 类: 第 1 类是 DCT 域的滤波, 包括分辨率变换, 这一类的处理主要是利用 DCT 的乘法-卷积特性或矩阵运算, 直接在 DCT 域内完成空域滤波等“点”操作, 需要严谨的数学推导以达到操作的精确性, 计算相对复杂; 第 2 类是视频编辑处理技术, 是直接利用 DCT 变换的线性性来完成简单的视频分割与编辑合成等的技术, 理论推导上较直观, 计算复杂度小; 第 3 类是 DCT 域的特征提取, 应用于基于内容的多媒体检索, 往往是通过 DCT 系数的组合来描述特征, 有效即可, 随意性大。

#### 3.1 DCT 域的滤波和分辨率变换

空域的滤波往往需要在邻域中进行, DCT  $8 \times 8$  的数据块分割破坏了邻域的连接性, 加大了 DCT 滤波的难度。直接利用 DCT 的乘法-卷积特性对 DCT  $8 \times 8$  压缩数据块进行滤波测试<sup>[4]</sup>, 由于设计的滤波器长度不大, 频率响应曲线平滑, 因此滤波效果较好, 没有明显的伪边缘和其他失真现象产生。此种滤波为 DCT 块内滤波, 在某些情况下会引入明显的伪边缘。为改善滤波质量, 文献[5]中提出了改进的 DCT 块间滤波。在对某个 DCT 块进行滤波的时候, 如同空域卷积那样, 在该 DCT 块的某个块邻域中进行。用 CCIR 推荐的应用于数字电视的低通滤波器作测试<sup>[3]</sup>, 效果很好。显然 DCT 块间滤波会增加计算量。

分辨率的改变在实际的应用中很常用(如 MPEG High 1 440×1 152→CCIR 601 720×576→MPEG CIF 352×288, CCIR 4:4:4→CCIR 4:2:2→MPEG 4:2:0, 视频浏览、网络传输等), 文

献[5]、文献[11]~文献[18]讨论了在 DCT 域直接实现分辨率改变的方法。DCT 域分辨率改变主要是基于矩阵运算。当降低分辨率时, 在隔点抽取样值之前, 还要低通滤波, 消除伪边缘。空域分辨率改变的矩阵虽然是稀疏矩阵, 但其 DCT 却不是, 故在 DCT 域直接操作时, 往往计算量较大, 需要利用一些性质优化算法来提高计算速度<sup>[13~15]</sup>。文献[14]、[15]设计了特殊滤波器, 使 DCT 域的变换矩阵稀疏化, 并以保留 DCT 低频成分为主。由于 DCT 系数能量往往主要分布在低频区域, 这样在提高计算速度的同时, 保证了图象质量。试验证明, 这种方法较空域常规方法节省时间, 且经过分辨率降低-回升操作的图象质量要明显优于常规的空域变换方法。大致说来, DCT 上的分辨率变换算法要比传统的空域方法节省约 10% 到 80% 的计算时间。

#### 3.2 视频编辑处理技术

视频编辑技术主要利用上面所述的 DCT 变换线性性质。文献[19]综述了 DCT 域可行的视频编辑技术, 如图象渐变、添加了幕、加膜处理、人物背景合成、变速处理等操作, 都可以通过简单的 DCT 线性变换和矩阵操作来实现, 且计算复杂度不大, 与传统的空域操作相比, 将节省 60%~75%(DCT) 或 10%~23%(MC-DCT) 的计算量。

视频处理技术主要有视频场景切换判定、边缘、颜色信息提取, 不同视频、图象格式间的数据转换等, 主要是利用 DCT 的低频系数或 DCT 系数的频带、方向性信息<sup>[23~26]</sup>(图 1)。

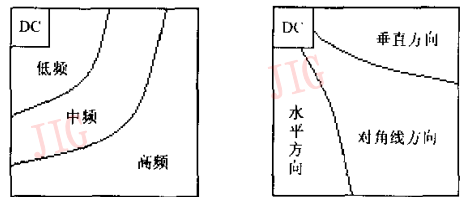


图 1 DCT 系数频带和方向信息

如根据 DCT 直流分量 DC 系数及其直方图<sup>[21,26]</sup>, 或频率带直方图<sup>[23]</sup>判别视频场景切换; DCT 系数所带的方向性特征可以用来初步判定边缘的存在与否和边缘方向。文献[26]从常见的 JPEG、MPEG-1, 2, H. 263 所采用的 Y, Cr, Cb 空间中, 通过简单变换提取 HIS 空间的伪 H、伪 S 信息, 结合 DCT 直流系数判定视频场景切换。文献[27]利用经过解码的颜色信息并结合 DCT 分块信息判

断人脸图案的存在.文献[28]介绍了一种不同视频、图象格式间DCT数据流转换的快速算法.在DCT域中进行运动估值<sup>[29]</sup>,提高了H.236的编码效率.

### 3.3 DCT域特征提取

DCT域的特征提取,如图象边缘提取、纹理分类、区域分割,应该说现在还处于比较初级的阶段,还只是对DCT数据的简单直观应用.如用DCT系数的3个大致方向性分布特征,初步判定边缘的存在与否和边缘方向;用直流系数的统计直方图作为图象相似性测度和判据等.这里面既有现今基于内容的描述和检索正处于探索阶段的原因,也有DCT系数的 $8 \times 8$ 分块特征在某种程度上并不利于对整体信息表达的原因.利用DCT的几何变换性质,通过在DCT域计算图象边缘的强度、方向、位置等信息,在DCT域直接提取图象边缘<sup>[2,30]</sup>,文献[2]的边缘检测效果与Sobel边缘检测算子相近,文献[30]的计算量优于空域方法且抗噪声性能好.将解压缩和滤波卷积的操作合而为一,利用DCT系数零值多的特点和操作的对称性,在DCT域直接通过卷积滤波提取边界,计算速度可提高 $2 \sim 10$ 倍<sup>[31]</sup>.在纹理分类方面,由于DCT数据的频域特性和纹理的统计特性,DCT域的纹理分析大多通过统计直方图、计算不同频段内数据的低阶统计量等统计方法来实现<sup>[32-35]</sup>,具有快捷、方便、分辨力强、准确率高的特点.利用DCT直流系数所组成的低分辨率图象和图象纹理的DCT交流系数统计特征,对CT图象进行分类和检索<sup>[36]</sup>,试验证明,该方法高速、有效、实用性强.文献[37]~文献[44]中,还探讨了在DCT域中实现区域分割、文字提取、图象检索、目标识别、人脸识别等传统的图象处理方法.

DCT域的相关图象处理技术还有DCT域数字水印加密和信息隐藏等.由于其思想方法和上面所论及的DCT域内图象视频处理和特征提取方法有差异,在这里就不详细论述了,有兴趣者可参见文献[45]、[46]等.

## 4 结束语

DCT变换是一种应用广泛的数据压缩方法.利用DCT乘法-卷积运算、线性运算、矩阵运算、几何变换运算等良好的数学变换性质,能够在DCT域直接实现如滤波、视频编辑处理、特征提取等传统的空域数据处理技术,从而避免了传统空域操作中繁

琐的解DCT压缩数据,处理后再还原为DCT压缩数据的操作过程.其对数据量的减少,处理时间的节省,处理效率的提高,效果明显.这一技术对于如Internet信息检索、视频编辑、实时图象理解等实际应用场合,是很有意义的.

未来DCT域数据处理技术的发展,首先仍必须围绕提高效率这一中心问题展开.DCT域数据处理技术较之传统的空域处理技术的诱人之处是它的高效率,包括计算复杂度的精简和系统资源的节省.这是DCT域数据处理思想的出发点.传统空域图象、视频处理技术在推广到DCT域的过程中,DCT数据 $8 \times 8$ 的局部分块特征往往成为提高效率的绊脚石.如何利用特定问题的特点和DCT变换的性质,巧妙地优化算法,是DCT域视频处理技术需要解决的难点之一.文献[31]针对边界提取这一传统问题,很好地进行了算法合并和优化,取得了 $2 \sim 10$ 倍的加速比,是一个典型的例子.

在实际工作中,DCT域特征空间的建模将可能是一个很有潜力的发展方向.DCT数据的频域特征不仅使数据量得到了压缩,还为基于内容的目标表达,即特征提取提供了一个在空域描述之外,有可能更为有效的特征描述方式.DCT域特征提取过程,在广义上可以视作是一个将DCT数据映射到相应的特征空间的操作.根据提取特征的特点,在DCT域寻求直接建立特征空间的方法,或在DCT域进行直接建模,将是DCT域特征提取的一个可能发展方向.利用纹理的频域特征和统计分布特点计算纹理的DCT数据低阶统计量的方法<sup>[32-35]</sup>,或在DCT域中实现“特征脸”的尝试可以说是这一发展方向的雏形.空域特征提取和建模的新经验和新成果将会促进DCT域特征提取技术的发展.

DCT还有很多很好的性质,如其为仅次于KLT变换的最优变换,以及与KLT、FFT、DST等变换的密切关系,等等,都是可以用于数据处理场合的有力工具.现有的DCT性质以及前面所述的DCT域图象视频高效率数据处理方法,在实际工作中还未得到广泛的应用.DCT域数据处理的方法还可在一维的音频信号中应用.所有的这些,都预示着DCT域数据处理技术具有广阔的应用前景.

### 参考文献

- 1 Shen Bo, Sethi I K, Bhaskaran V. DCT domain alpha blending [A]. In: Proc. of IEEE International Conference on Image

- Processing (ICIP98)[C]. Chicago, Illinois, USA, 1998, 1: 857~861.
- 2 Shen Bo, Sethi I K. Direct feature extraction from compressed images[A]. In: Proc. SPIE Storage and Retrieval for Image and Video Database IV[C], San Jose, California, USA, 1996, 2670: 404~414.
  - 3 Chitprasert B, Rao K R. Discrete cosine transform filtering[J]. Signal Processing, 1990, 19(3): 233~245.
  - 4 Chen W H, Fralick S C. Image enhancement using cosine transform filtering[A]. In: Image Sci. Math. Symp. [C]. Monterey, CA, USA, November, 1976: 186~192.
  - 5 Neri A, Russo G, Talong P. Inter-block filtering and downsampling in DCT domain[J]. Signal Processing: Image Communication, 1994, 6(4): 303~317.
  - 6 Smith B C, Rowe L A. Algorithms for manipulating compressed images[J]. IEEE Computer Graphics and Applications, 1993, 13(5): 34~42.
  - 7 韩江, 全春来, 李进治等. DCT 域中特定几何变换的性质及其应用[J]. 系统工程与电子技术, 1995, 17(10): 34~40.
  - 8 Shen Bo, Sethi I K. Inner-block operations on compressed images[A]. In: Proc. ACM Intl. Conf. Multimedia'95 [C]. San Francisco, California, USA, 1995: 489~498.
  - 9 Shen Bo, Sethi I K. Algorithms in the JPEG discrete cosine transform compressed domain[J]. Journal of Electronic Imaging, 1996, 5(2): 182~190.
  - 10 Robert Reeves. New shift, scaling and derivative properties for the DCT[A]. In: Proc. of IS&T/SPIE Conference on Visual Communication and Image Processing'99 [C]. San Jose, California, USA, 1999, 3653: 418~428.
  - 11 叶楠, 郑志航. MPEG 码流中 I 帧 1/4 图象的快速近似提取方法[J]. 数字视频, 2001, (3): 69~70, 81.
  - 12 Hu Q, Panchanathan S. A comparative evaluation of spatial scalability techniques in the compressed domain[A]. In: Proc. of IEEE Conference on Electrical and Computer Engineering[C], Calgary, Alberta, Canada, 1996, 1: 474~477.
  - 13 Natarajan B K, Vasudev B. A fast approximate algorithm for scaling down digital images in the DCT domain[A]. In: Conf. Rec. IEEE Int. Conf. Image Processing(ICIP)[C]. Washington DC, USA, 1995, 2: 241~243.
  - 14 Dugad R, Ahuja N. A fast scheme for altering resolution in the compressed domain[A]. In: Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition[C], Fort Collins, Colorado, USA, 1999, 1: 213~218.
  - 15 Dugad R, Ahuja N. A fast scheme for downsampling and upsampling in the DCT domain[A]. In: Proc. of the IEEE 1999 Conference on Image Processing[C]. Kobe, Japan, 1999, 2: 909~913.
  - 16 Dugad R, Ahuja N. A fast scheme for image size change in the compressed domain[J]. IEEE Trans. on Circuit and Systems for Video Technology, 2001, 11(4): 909~913.
  - 17 Hu Q, Panchanathan S. Image/video spatial scalability in compressed images[J]. IEEE Transactions on Industrial Electronics, 1998, 45(1): 23~31.
  - 18 Merhav N, Bhaskaran V. Fast algorithms for DCT-domain image down sampling and for inverse motion compensation[J]. IEEE Trans. CSVT, 1997, 7(3): 468~476.
  - 19 Chang S-F, Messerschmitt D G. Manipulation and compositing of MC-DCT compressed video[J]. IEEE Journal of Selected Areas in Communications, Special Issue on Intelligent Signal Processing, 1995, 13(1): 1~11.
  - 20 Noguchi Y, Messer D G, Chang Shib-Fu. MPEG video compositing in the compressed domain[J]. IEEE Circuits and Systems, 1996, 43(2): 596~599.
  - 21 Divakaran A, Ito H, Sun H *et al.* Fade-in/out scene change detection in the MPEG-1/2/4 compressed video domain[A]. In: Part of the IS&T/SPIE Conference on Storage and Retrieval for Media Databases[C]. San Jose, California, USA, 2000, 3972: 518~523.
  - 22 Won C S, Park D K, Yoo S J. Extracting image features from MPEG-2 compressed stream[A]. In: Storage and Retrieval for Image and Video Data Base VI, [C]. San Jose, California, USA, 1998, 3312: 426~436.
  - 23 Bao O K C, Lay J A, Guan L. Compressed domain video parsing using energy histograms of the lower frequency DCT coefficients[A]. In: Part of the IS&T/SPIE Conference on Storage and Retrieval for Media Databases[C]. San Jose, California, USA, 2000, 3972: 293~300.
  - 24 Meng J, Chang S F. CVEPS a compressed video editing and parsing system[A]. In: ACM Multimedia'96 Proceedings[C]. Boston MA, USA, 1996: 43~53.
  - 25 Song B C, Ra J B. Fast edge map extraction from MPEG compressed video data for video parsing[A]. In: Part of the IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases W [C]. San Jose, California, USA, 1999, 3656: 710~720.
  - 26 Won C S, Park D K, Na I Y *et al.* Efficient color feature extraction in compressed video[A]. In: Part of the IS&T/SPIE Conference on Storage and Retrieval for Image and Video Databases W [C]. San Jose, California, USA, 1999, 3656: 677~684.
  - 27 Wang H, Chang S F. A highly efficient system for automatic face region detection in MPEG video[J]. IEEE Trans. on Circuits and Systems for Video Technology, 1997, 7(4): 615~628.
  - 28 Shibata Y, Chen Zh, Campbell R H. A fast degradation free algorithm for DCT block extraction in the compressed domain[A]. In: Proc. of the IEEE 1999 International Conference on Acoustics, Speech, and Signal Processing [C], Phoenix, Arizona, USA, 1999, 6: 3185~3188.
  - 29 曹宁, 吴敏. 基于 DCT 的频域块匹配算法及其实现[J]. 计算机工程, 2000, 26(11): 129~130.
  - 30 Kim T Y, Han J H. Model-based discontinuity evaluation in the

- DCT domain[J]. *Signal Processing*, 2001, 81(4): 871~882.
- 31 Shen B, Sethi I K. Convolution-based edge detection for image/video in block DCT domain[J]. *Journal of Visual Communication And Image Representation*, 1996, 7(4): 411~423.
- 32 Reeves R, Kubik K, Osberger W. Texture characterization of compressed aerial images using DCT coefficient[A]. In: *Proc. SPIE: Storage and Retrieval for Image and Video Databases V* [C]. San Jose, California, USA, 1997, 3022: 398~407.
- 33 Sim D G, Kim H K, Park R H. Fast texture description and retrieval of DCT based compressed images[J]. *IEEE Electronics Letters*, 2001, 37(1): 18~19.
- 34 Huang Y L, Chang R F. Texture features for DCT-coded Image Retrieval and Classification[A]. In: *Proc. of the IEEE 1999 International Conference on Acoustics, Speech, and Signal Processing* [C]. Phoenix, Arizona, USA, 1999, (6): 3013~3016.
- 35 Bae H J, Jung S H. Image retrieval using texture based on DCT[A]. *International Conference on Information, Communications and Signal Processing*. ICICS'97[C]. Singapore, 1997, 1065~1068.
- 36 苏北北, 黄天锡, 廖孟扬等. 基于JPEG的医学图象特征提取[J]. *中国图象图形学报(A版)*, 1999, 4A(7): 583~589.
- 37 Shuster M, Mottaleb M A. Exploiting the JPEG compression scheme for image retrieval[J]. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1996, 18(8): 849~853.
- 38 Ricard L, de Queiroz. *Processing JPEG-compressed images and documents*[J]. *IEEE Trans. On Image Processing*, 1998, 7(12): 1661~1672.
- 39 Eng H L, Ma K K. Unsupervised image object segmentation over compressed domain[J]. *IEEE Image Processing*, 2000, 9(3): 758~761.
- 40 Seales W B, Yuan C J, Hu Wei *et al.* Object recognition in compressed imagery[J]. *Image and Vision Computing*, 1998, 16(5): 337~352.
- 41 Chen B, Latifi S, Kanai J. Edge enhancement of remote sensing image data in the DCT domain[J]. *Image and Vision Computing*, 1999, 17(12): 913~921.
- 42 Tsapatsoulis N, Doulanis N, Doulanis A *et al.* Face extraction from non-uniform background and recognition in compressed domain[A]. In: *Proc. of the IEEE 1998 International Conference on Acoustics, Speech and Signal Processing* [C]. Seattle, WA, USA, 1998, 5: 2701~2704.
- 43 Eickeler S, Muller S, Rigoll G. High quality face recognition in JPEG compressed images[A]. In: *Proc. of the IEEE 1999 International Conference on Image Processing* [C]. Kobe, Japan, 1999, 1: 672~676.
- 44 Eickeler S, Muller S, Rigoll G. Recognition of JPEG compressed face images based on statistical methods[J]. *Image and Vision Computing*, 2000, 18(4): 279~287.
- 45 黄继武, Yun Q, SHI, 程卫东. DCT域图象水印: 嵌入对策和算法[J]. *电子学报*, 2000, 28(4): 57~60.
- 46 张璟, 刘云超, 陈一民等. 基于DCT变换的图象中的数据隐蔽技术及理论[J]. *计算机工程*, 2000, 26(12): 123~125.



刘艳 1977年生, 2000年7月获浙江大学信息工程学士学位, 现为浙江大学信息学院信电系信号与信息处理专业硕士研究生. 主要从事压缩编码、图象处理、多媒体信息检索等方面的研究.



李宏东 1972年生, 博士, 现为浙江大学信电系副教授, 研究方向为计算机视觉、图象理解、模式识别和机器人导航.