

# 一种基于 DCT 域的图像快速检索技术

## Image Retrieval Technique in the DCT Domain

秦 军 (解放军理工大学通信工程学院研二队)

罗国明 (解放军理工大学通信工程学院电信工程系)

**摘要:** 图像视频压缩标准的出现,使得图像数据格式普遍为压缩格式,从而促进了压缩域内图像检索技术的迅速发展。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 域进行特征提取,能够避免繁琐的压缩数据编、解码操作,大大减少处理时间和处理量,节省内存空间,使得图像检索更加高效。本文首先说明了 DCT 变换的定义和 DCT 域的矩阵运算,最后着重介绍了基于 DCT 域图像检索技术的一种图像内容快速检索方法。

**关键词:** DCT 压缩域处理 特征提取 纹理描述 图像检索

### 1 引言

随着图像、视频压缩标准(JPEG、MPEG)和压缩技术的不断发展,图像和视频数据也已经达到了海量存储,这也使得数字图像检索技术日益显示出其重要性。在不久的将来,图像、视频检索可应用于多媒体信息系统、数字图书馆、电影行业、视频点播等多方面。然而,现在采用的图像检索方式主要还是基于图像内容的检索方式。媒体压缩标准的出现,使得图像数据格式普遍为压缩格式,由于传统的图像数据处理方式是在原始数据域或者是将压缩数据解码以后进行处理,而压缩域的图像数据处理则是直接在压缩数据上进行处理,这样不仅能避免编解码相关的复杂操作,减少数据处理量,节省内存空间,而且还能充分利用压缩域数据的性质进行高效、准确的数据检索。

常见的数字压缩技术有 DFT、DCT、小波变换等。离散余弦变换 DCT 技术有很多优点,首先,它是正交变换,可以将  $8 \times 8$  图像的空间域转换为频率域,只需要用少量的数据点表示图像;其次,DCT 产生的系数很容易被量化,因而能获得好的块压缩;最后,DCT 有快速算法,算法性能很好,易于在软硬件中实现,而且 DCT 算法是对称的,利用逆 DCT 算法可以用来解压缩图像,因此 DCT 是现在应用最广泛的多媒体数字压缩技术。图像、视频压缩国际标准 JPEG、MPEG 和 H. 26L 等都采用了 DCT 技术,研究 DCT 域的数字处理技术具有典型

性和现实性,利用 DCT 的特征提取技术进行信息检索也具有很大的吸引力和可行性,也是近年来国际上有关领域的研究热点之一。

### 2 DCT 变换的定义

一维  $N$  点序列  $x(n)$  的 DCT 变换  $X_c(k)$  定义为:

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

$$= 0, 1, \dots, N-1 \quad (1)$$

DCT 反变换定义为:

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

$$1, \dots, N-1 \quad (2)$$

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

在 JPEG、MPEG 等图像、视频压缩标准中所采用的二维 DCT 定义为:

$$X_{ij}(u, v) = \frac{2}{N} C(u) C(v) \left\{ \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} x(i, j) \cos \frac{(2i+1)\pi u}{2N} \cos \frac{(2j+1)\pi v}{2N} \right\} \quad i, j, u, v = 0, 1, \dots, N-1$$

(3)

因为是对  $8 \times 8$  的图像子块进行 DCT 变换,所以式中的  $N=8$ 。DCT 变换是一种线性正交变换,变换各矢量间单位正交,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,由于 DCT 变换计算量适中,性能优于 DFT 等其他变换,去相关压缩能力强,故现行的图像视频标准多是对  $8 \times 8$  的数据块进行 DCT 操作。

### 3 DCT 域的矩阵运算

由于 DCT 为线性变换,因此 DCT 变换可以表示为矩阵乘法的形式来实现,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,所以简单 DCT 块内滤波操作可以借用矩阵乘法来实现。

对二维 DCT 变换,把空域的二维数据  $x(i, j)$  和二维 DCT 数据  $X_c(u, v)$  表示为二维  $N \times N$  维矩阵  $x$  和  $X_c$  后, DCT 变换便能表示为  $X_c = T x T^{-1}$  的矩阵乘法形式,其中  $T$  为 DCT 变换矩阵,具有  $T^{-1} = T$  的性质。DCT 反变换可表示为  $x = T^{-1} X_c T$ 。

对于空域的一个  $N \times 2N$  的图像块  $p = [g^1 | g^2]$ ,若希望每行隔点抽样为一个新的  $N \times N$  图像块  $g$ ,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

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

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

$$\begin{aligned} G_c &= T [g^1 | g^2] S T^{-1} \\ &= T [g^1 | g^2] T_2^{-1} \times T_2 S T^{-1} \\ &= [G_c^1 | G_c^2] T_2 S T^{-1} \quad (4) \\ &= [G_c^1 | G_c^2] F \end{aligned}$$

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

### 4 一种基于 DCT 压缩域的快速图像检索方法

由于现行国际标准在编码时都使用类似的 DCT 算

法,这里有必要简单回顾一下基于 DCT 的 JPEG 标准算法,在 JPEG 编码中,编码分为计算 DCT、量化和变长编码 3 个步骤,首先将原始图像分割为互不重叠的  $8 \times 8$  的像素块,然后对每个块进行二维 DCT 变换,并用 JPEG 建议的符合视觉特点的量化表进行量化,使每一个子块生成一个 DC 系数和 63 个 AC 系数,接着对量化后的系数用 Zig-Zag 顺序重新排序,形成一个一维的量化后系数序列,最后对每一个像素块的 DC 系数用 DPCM 编码,而对剩余的 AC 系数则综合利用 Huffman 编码和游程长编码来进行压缩。

基于 DCT 压缩域的快速图像检索方法首先对整个图像的 DCT 系数求均值,然后舍去高频系数,再对中低频系数求取对数,以生成纹理描述符,而查询图像和目标图像的相似度则用欧氏距离来测量。

#### 4.1 纹理特征描述

这种纹理描述方法适用于任何形式的图像和视频数据,对原始图像,则将其划分为许多  $8 \times 8$  的像素块并对每个子块进行 DCT 变换,从而得到整个图像的 DCT 系数,对于 MPEG 视频流可以采用帧间编码计算 DCT 系数。纹理描述符就是基于这些 DCT 系数构造的。

设原始图像序列为  $f(m, n), 0 \leq m < M, 0 \leq n < N$ , 则原始图像的分辨率为  $M \times N$ 。图  $f_k(i, j), 0 \leq i < 8, 0 \leq j < 8, 0 \leq k < \frac{M}{8} = K, 0 \leq l < \frac{N}{8} = L$  代表一个像素块,下标  $k, l$  表示一个特定像素块的索引,则像素块的总数目为  $KL$ ,且有关系式  $m = 8k + i, n = 8l + j$ 。

设  $(k, l)$  子块的 DCT 变换为  $F_k(n, v), n, v$  代表频率指数,则整个图像的 DCT 系数的平均值为:

$$F(u, v) = \frac{1}{KL} \sum_k \sum_l |F_k(u, v)| \quad (5)$$

计算结果共有  $8 \times 8 = 64$  个 DCT 系数,但考虑到人眼特性,所以高频系数对于特征提取意义不大,可以将其掩蔽从而减少计算量,经实践验证得到典型的掩蔽矩阵为  $M_{sk}$ ,如表 1。

经掩蔽后整个图像基于 DCT 的特征描述符为:

$$\begin{aligned} F(p) &= \log \left| \sum_{\text{all}(n,v)} F(u, v) \delta(M_{sk}(n, v) - p) \right|, 1 \\ &\leq p \leq P = 52, \delta(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases} \quad (6) \end{aligned}$$

表 1

1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	
32	33	34	35	36	37	38	
39	40	41	42	43	44		
45	46	47	48	49			
50	51	52					

## 4.2 特征描述符与相似度计算

基于相似性特征检索系统中,原始图像纹理描述符数据表存储在图像索引文件中。进行特征匹配时,查询图像  $q(m,n)$  和目标图像  $f(m,n)$  分别生成各自的 DCT 特征描述符  $Q(p)$  和  $F(p)$ ,特征描述符可由  $P$  维矢量集表示为:  $\{F(1), F(2), \dots, F(P)\}$ , 其中每一个矢量量化值都不超过 8bit。当选定不同的  $P$  值时,可以将检索分为多个检索等级。查询图像和目标图像之间的相似度则为:

$$d(q, f) = \sum_{p=1}^P \frac{|Q(p) - F(p)|}{\sigma_{Q(p)}} \quad (7)$$

其中  $Q(p)$  和  $F(p)$  分别表示查询图像和目标图像的第  $p$  个 DCT 特征系数,  $\sigma_{Q(p)}$  则表示查询图像  $Q(p)$  在特征数据表中的标准 DCT 系数偏移量。

通过建立以  $d(q, f)$  为基数的相似度统计模型,可以进行高效而准确的图像检索,与其他检索方法相比它具有算法复杂度很小,检索准确率高,纹理描述符容量小,特征提取快等特点,更适用于实时系统的检索。

## 5 结束语

DCT 变换是一种应用广泛的数据压缩方法,利用 DCT 良好的数学变换性质能够在 DCT 域进行高效的图像信息检索,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,将其用于图像检索能够充分发挥检索准确率高,检索速度快,检索计算复杂度低,计算量小等优势。

未来 DCT 域数据处理技术的开展,首先仍必须围绕提高效率这一中心问题展开,包括计算复杂度的精

简和系统资源的节省。传统空域图像处理技术在推广到 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,巧妙的优化算法是 DCT 域图像处理技术需要解决的难题之一。虽然基于 DCT 域的图像检索技术还处于比较初级的阶段,但随着今后高速海量数据处理场合的增多和研究的不断深入,相信基于 DCT 域的图像检索技术必将得到国际有关领域的高度重视。

## 参考文献

- 1 Jiang J, Armstrong AJ, Feng GC (2002) Direct content access and extraction from JPEG compressed images. *Pattern Recognition*
- 2 Mandal MK, Idris F, Panchanathan S (1999) A critical evaluation of image and video indexing techniques in the compressed domain. *Image Vision Comput J*
- 3 Armstrong A, Jiang J (2001) An efficient image indexing algorithm in JPEG compressed domain. (ICCE'2001), June 2001
- 4 Berman AP, Shapiro LG (1999) A flexible image database system for content-based retrieval.
- 5 N. Merhav and V. Bhaskaran, "Fast algorithms for DCT-domain image down-sampling and for inverse motion compensation," *IEEE Transactions on Circuits and Systems for Video Technology*, June 1997.
- 6 B. K. Natarajan and B. Vasudev, "A fast approximate algorithm for scaling down digital images in the DCT domain," *IEEE Int. Conf. Image Processing*, Oct. 1995.
- 7 Jeon, J. Park, and J. Jeong, "Huffman coding of DCT coefficients using dynamic codeword assignment and adaptive codebook selection," *Signal Processing; Image Commun.*, 1998.
- 8 [美] Ralf Steinmetz, Klara Nahrstedt, 多媒体原理(第一册)媒体编码及内容分析(英译版), 电子工业出版社, 2003。