

DCT 算法及其与小波编码在图像处理中的比较

罗开仲, 黄士坦, 杨华民

(西安微电子技术研究所, 陕西 西安 710065)

摘 要: 由于数字媒体的不断普及, DCT 编码在 JPEG, MPEG, H. 26x 等编码标准中有着越来越广泛的应用。但随着 Shapiro 关于小波的嵌入式零树小波编码以及 Said 等人提出的集合分裂等级树编码的成功应用, 图像及视频信号的压缩编码迎来了一个新的阶段, 这显然对传统 DCT 编码提出了挑战。然而, Xiong 等人利用嵌入式 DCT 块变换之间直流系数(DC)的高度相关性也能获得较高的编码性能, 因为对 DCT 后的系数进行重组, 同样具有小波多分辨率图像的特性。文中主要提出 DCT 与 DWT 的比较, 以及基于 DCT 与小波变换在图像中的性能比较。

关键词: 离散余弦变换; 小波变换; 图像编码; MPEG; JPEG

中图分类号: TN919.81

文献标识码: A

文章编号: 1673-629X(2006)09-0079-03

DCT Arithmetic and Its Comparison with Wavelet Transform Coding in Image Manipulation

LUO Kai-zhong, HUANG Shi-tan, YANG Hua-min

(Xi'an Microelectronic Technology Research Institute, Xi'an 710065, China)

Abstract: DCT transforming coding has been widely used coding standards such as JPEG, MPEG and H. 26x due to the development of digital media. The compressing codings for image video signals have been advanced to a new stage since the EZW had been introduced by Shapiro and the SPIHT applied by Said et al, which are a competitive challenge against the conventional DCT's coding. However, DCT's coefficients have a similar multiresolutional attributes after reorganizing its datum like wavelet transforming. The paper mainly tell about the comparison of DCT and DWT in the image manipulation.

Key words: DCT; DWT; coding of image; MPEG; JPEG

1 DCT 的发展

离散余弦变换(Discrete Cosine Transform, DCT)是由 Ahmed 和 Rao 提出的, 它的提出虽然比 FFT (Fast Fourier Transform) 晚, 但其性能更接近于理想的 KL 变换, 并且由于 KLT 到目前为止没有发现有效的快速算法, 所以 DCT 在信号处理中得到了广泛应用。DCT 的快速算法是在继承完善 DFT (Discrete Fourier Transformation) 的基础上不断进步的, 如一个 8×8 图像块的 DCT (或逆 DCT) 按定义计算, 需要 1024 次乘法和 896 次加法运算, 根据 DCT 的正交对成性, 使用蝶形算法可以大大减少加乘次数^[1]。由于目前集成电路技术实现一次加、乘运算的时间几乎相等, 因此能将原来进行正交变换需要占系统总时间的 30% 以上降为原来时间的 $1/5 \sim 1/3$ 。

常规的 2D DCT 都是由可分离的 1D DCT 组成, 表 1 列出了 1×8 输入信号的计算量。采用递归 2D DCT 以及下标映射法来设计 2D DCT 已经比用传统的行列法运算

量大为减少, 特别是当 2D DCT 用于图像或视频编码系统中时, 就是根据信号的自然属性和视觉特性 HVS 先将 DCT 后的系数按 Zigzag 扫描, 然后进行只取前若干个较大的低频系数作为特征分量, 并用零代替后面高频系数的截断 DCT。在实际编码中, DCT 正是以它的“提取特征成分的能力和运算速度之间的最佳平衡”来显示出其应用优越性的。3D DCT 以及在利用彩色图像中的 RGB 分量上全部信息同时进行变换压缩, 则不仅可以有效地去除彩色图像中的各种冗余信息, 还可在 3D DCT 压缩编码中实现信息隐藏。另外, 实际传输的甚低码率, 如在 H. 263 中, 且图像(块)间平缓区域占多数, 则可通过快速运动补偿(减小时间冗余)来得到帧间(帧内直接 DCT)的 8×8 预测误差图像块 $x(m, n)$, 若满足 $\sum_{x=0}^7 \sum_{y=0}^7 \text{abs}(x(m, n)) < 16Q$ (Q 为量化参数), 即判断该块为零的 DCT 系数, 由于这样可以避免计算该 DCT 块的所有频率系数, 因而提高了实时应用的能力^[2]。总之, DCT 各种快速算法的进步有力地促进了基于 DCT 编码的进展, 主要表现在基于区域的 DCT 方法上和对变换后的系数进行策略性重组及其系数的高效应用上, 譬如层次嵌入式 EZH DCT (Embedded Zerotree based on Hierarchical DCT)、区域支撑 RS DCT

收稿日期: 2005-11-22

作者简介: 罗开仲(1977-), 男, 陕西西安人, 硕士研究生, 研究方向为嵌入式系统结构、视频压缩算法; 黄士坦, 博士生导师, 主要从事空间计算机技术的研究。

(Region Support DCT)、形态 MR DCT(Morphological Representation of DCT)以及整型 DCT 等均强调变换方法的改进,其中 EZH DCT 与 MR DCT 是利用 DCT 系数的重组来得到类似小波多分辨率图像特性。另外,还有利用 DCT 域系数的分布特性及 HVS 特性来实现信息隐藏、内容检测、块虚像检测与消除,以及 DCT 域图像大小的调整,等等。这些无疑丰富发展了传统 DCT 的理论与实践,使之有了更大的应用与发展空间。

表 1 1×8 DCT 的计算量比较

	Wang	Loeffler	Chen	Suehiro	Vetterli
乘法	11	11	16	12	12
加法	29	29	26	29	29

2 DCT 的意义及其与 DWT 的比较

2.1 DCT 的实践意义

DCT 是酉变换的一种,其变换前后的信号熵和能量不变,且在时域(或空域)中的 n 维信号 X ,其变换核可分离的正、逆 DCT 分别定义如下:

$$F(u_1, \dots, u_n) = \sqrt{\frac{2^n}{N_1 \cdots N_n}} c(u_1) \cdots c(u_n) \sum_{t_1=0}^{N_1-1} \cdots \sum_{t_n=0}^{N_n-1} X(t_1, \dots, t_n) \cos \frac{(2t_1+1)u_1\pi}{2N_1} \cdots \cos \frac{(2t_n+1)u_n\pi}{2N_n} \quad (1)$$

$$X(t_1, \dots, t_n) = \sqrt{\frac{2^n}{N_1 \cdots N_n}} c(u_1) \cdots c(u_n) \sum_{u_1=0}^{N_1-1} \cdots \sum_{u_n=0}^{N_n-1} F(u_1, \dots, u_n) \cos \frac{(2t_1+1)u_1\pi}{2N_1} \cdots \cos \frac{(2t_n+1)u_n\pi}{2N_n} \quad (2)$$

其中, $c(0) = \frac{1}{\sqrt{2}}$, 其他 $c(\cdot)$ 为 1; $N_1 \cdots N_n$ 为相应维的大小。通常大小为 $N \times N$ 像素的 2 维图像,其 DCT 的各行(列)基矢量均是归一正交的(为了保证变换前后熵或范数相等)。由于图像信号的统计特性接近一阶马氏链,因此 DCT 基矢量与之非常匹配,且不仅 DCT 基矢量的设计在多数情况下符合 HVS 特性,而且编解码电路简单。综合 DCT 编解码性能及解相关性等方面因素,它仅次于理想的 KLT,所以 DCT 在众多变换(离散傅氏变换、斜变换、Haar 变换、KL 变换、沃尔什变换和 DCT 等)中脱颖而出。经过 DCT 后的系数主要值集中在左上角低频区域,其中 $F(0,0)$ 直流值最大。由于变换后系数矩阵右下角绝大多数高频系数趋向于零,因此为结合 HVS 特性,采用(有死区的)量化、Zigzag 扫描,截断 DCT 的变字长压缩编码创造了条件,更重要的是为数字化图像视音频信号的压缩找到了切实可行的方法。尽管带限的频率域模型存在频谱的截断误差,但由于其求解精度和抗噪声能力较好,因而有助于问题更加精确的描述和求解,这也正是频率方法研究的意义所在。

2.2 与 DWT 变换的比较

20 世纪 80 年代出现的小波变换编码,特别是 20 世纪 90 年代 Shapiro 的 EZW 以及 Said 等人关于 SPIHT 的成功应用,对传统的 DCT 编码提出了最严峻的挑战。然而, Xiong 等人就基于离散小波变换(DWT)与 DCT 编码

的性能作了系统深入的比较,根据文献[3]和[4],表 2 列出了各种算法编码重建图像的比较结果。通过对表 2 中数据的分析容易得出这样一个结论:在率失真方面,压缩效率(比特率)并不正比于视觉质量(峰值信噪比)的降低。也就是说, DCT 仍然能产生高压压缩效率和高峰值信噪比,因此 DCT 仍是未来图像视频压缩编码中不可缺少的主要工具之一。借助于小波变换来压缩图像,其编码质量并非任何情况下均占优,譬如“采用内嵌 DCT 子带编码算法”对某些图像进行压缩时还优于 SPIHT^[5],且多数 DCT 的计算较小波变换简单,例如,时域信号的卷积对应于 DCT 频域的直接乘积,而小波变换则无此计算之简单性。

表 2 典型图像(Lena 和 Barb)DCT 与 DWT 编码恢复图像的 PSNR 比较 单位: dB

比特率 (bpp)	JPEG		Improved-JPEG		EZ-DCT		SPIHT	
	Lena	Barb	Lena	Barb	Lena	Barb	Lena	Barb
0.25	31.60	25.20	32.30	26.70	32.25	26.83	34.11	27.58
0.50	34.90	28.30	35.90	30.60	36.00	30.82	37.21	31.39
0.75	36.60	31.00	38.10	33.60	38.06	33.70	39.04	34.25
1.00	37.90	33.10	39.60	35.90	39.62	36.10	40.40	36.41

3 DCT 具有小波多分辨率图像分解的特性

研究表明,有多种方法能使 DCT 产生类似于小波的多分辨率图像分解特性。常规的 2D DCT 通过系数重组就可得到一种多分辨率的图像,即可先将一幅图像分成 $N \times N$ 块($N = 2^L$),然后变换到 DCT 域中的每个块,其可被看成是尺度为 L 的具有 $3L+1$ 子带分解的系数树。若对所有 DCT 块的相同频率子带进行分组,并把它们放到对应位置上,则 $N=8$ 时,3 层 10 子带 DCT 系数重组如图 1 所示。图 1 中 0 代表子带 0 组(由所有 DCT 直流系数组成), $\dots,9$ 代表子带 9 组(即所有对角子带分量),其重组后的 DCT 系数可以进一步利用 DCT 编码器来获得更优的编码性能。在基于层式 DCT 的嵌入式(EZH DCT)编码中^[6],采用的就是这种形式的分解,结果比 EZ DCT 编码更令人满意(如表 3 所示)。

表 3 几种压缩编码算法重建图像 PSNR 比较 单位: dB

比特率 (bpp)	MR DCT		EZH DCT	EZ DCT
	Lena	Barb	Lena	Lena
0.25	32.55	26.84	32.40	32.27
0.50	35.99	30.61	36.60	35.98
0.75	38.07	33.59	38.10	38.04
1.00	39.49	35.89	39.60	39.60

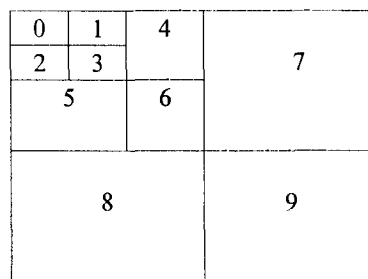


图 1 与 DCT 系数相对应的 3 层子带形式

层式 DCT 是另一种多分辨率分解形式,即先将原始图像经过 2D DCT 后的系数块均分为 4 个块,再把具有相同频带的子块按原来的空间位置组合成同频带子带,即得到 LL1、LH1、HL1 和 HH1,接着对低频 LL1 子带作逆 DCT,然后将得到的图像作第 2 层输入图像,并重复上述过程,直至低频 LLn 仅含直流系数为止,即得到类似小波变换的多分辨率图像。小波零树编码是根据同一类子带不同层系数的自相似性,通过引入零树概念来获得高的压缩比和良好的重构图像质量。在此不难得出,层式 DCT 编码较 EZ DCT 具有更高重构图像质量的原因是:

(1)经过系数块重组后,由于本层输入图像因 DCT 分块而产生的边界有一半落入更上一层输入图像的 DCT 分块内,于是原来由这些边界产生的方块效应被消除;

(2)与小波变换类似,其图像层数越高,对应的分辨率越低;

(3)分层子带变换(即将原信号按高、中、低频分成若干频带)中,由于高分辨率频带与低分辨率频带之间存在着相似性,因而可由低分辨率频带近似高分辨率频带,进而得到比原图像分辨率更高的图像。

因为相同方向的细节子图在不同的分辨率下是相似的,即高分辨率下的边沿特性可以看作是图像在小尺度下的边沿特性,而低分辨率下的边沿特性可以看作是图像在大尺度下的边沿特性,也就是不同尺度下的图像存在着相似,这一点可与小波多分辨率分析相对应。可进一步推断,如已知图 1 中 0 和 3 个方向上的细节子图 1、2、3,则可以恢复出更高分辨率的图像等。DCT 的多分辨率图像分解,为适应网络图像信号的可分级传输、内容检索等提供了条件。从人类视觉特性对图像进行由粗到细的理解过程看,用于图像压缩编码的任何变换应该具有多分辨率特性,像小波、分形、子带和 DCT 等,而不具有此特性的变换则一般不能直接应用,其生命力不强。

4 基于 DCT 和小波的图像编码的比较

在对图像编码中 DCT 和小波变换进行比较时,如果允许变换改变则固定量化器和熵编码器,这是在比较小波和 DCT 编码性能时能提供准确估计的惟一方法。当小波变换和基线 JPEG 量化器一起出现时,基线 JPEG 中的 DCT 就被 3 级的小波变换所代替。小波系数重新排列成小波块并且在量化和哈夫曼编码前经 Z 字形秩序扫描成矢量。在编码时,有时会采取一种逐次逼近的嵌入式编码,即系数的编码精度逐次提高,这样可以在比特率达到某种要求时结束编码或解码。如果固定 SPIHT 量化器,用它来量化 DCT 系数,就会得到基于 DCT 的嵌入式图像编码器^[4]。可以将一个 8×8 的 DCT 图像看作是 64 级子带的分解,然后把每个 8×8 的 DCT 块看作一个系数的 3 层树,如图 2 所示。将每块中的 64 DCT 系数分类后,就能够证实 DCT 系数间的父子关系如下:当 $1 \leq i \leq 63$ 时,系数 i 的父亲是 $[i/4]$,而当 $1 \leq j \leq 15$ 时,与系数 j 相关

的 4 个子系数的设置是 $\{4j, 4j+1, 4j+2, 4j+3\}$, 直流系数“0”是 DCT 系数树的根。在 512×512 的 Lena 图像中,两层的 8×8 DCT 变换结构相当于用 6 级的小波变换。基于 DCT 的编码器比基于小波的简单,虽然随着比特率的下降,性能差距加大了,但是对 Lena 用 DCT 代替小波变换造成的性能损失却不是太大。在基于 DCT 的嵌入式编码器中得到的 PSNR 比从 JPEG 和 Shapiro 的 EZW 编码器中要好得多。

0	1	4	5	16	17	20	21
2	3	6	7	18	19	22	23
8	9	12	13	24	25	28	29
10	11	14	15	26	27	30	31
32	33	36	37	48	49	52	53
34	35	38	39	50	51	54	55
40	41	44	45	56	57	60	61
42	43	46	47	58	59	62	63

图 2 8×8 的 DCT 图像示意图

5 基于 DCT 和小波变换的视频编码的比较

标准的视频编码方法是建立在将 DCT 应用到运动补偿图像的框架上的。基于小波的视频编码一直以来都是一个很有兴趣的研究主题。在标准方法中的一个观点就是用小波变换代替 DCT。在基于 DCT 和小波的 2 维变换的视频编码方法中,运动补偿图像同很多自然的静止图像相比有很大的不同,即在运动补偿后空间系数会减少。在数字视频应用的巨大潜力下,很多研究人员都把小波应用到对运动补偿图像的编码中。对 MPEG-4 引用了 Samoff 公司提出的基于小波的零树熵(ZTE)视频编码器输出的结果后,把它们与校验模型(VM)的输出相比较,可以看出:在相同的比特率和帧频下,基于小波的 ZTE 编码器在 VM 上产生的客观性能与基于 DCT 的编码器相当。目前,在 3 维子带/基于小波的视频编码中已经有了很多研究。Choi 和 Wood 用 3 维子带方法和分等级的各种块基的 MC 方案已经得出了比 MPEG-1 更好的结果。3 维的 SPIHT 视频编码器比 H.263 编码器的性能稍微差一些,但是当不用 MC 时,其复杂性要少很多。为了比较 DCT 和小波变换对于视频编码的不同点,需要建立一个 3 维的 DCT 基的 SPIHT 视频编码器,这样可以看作是 SPIHT 视频编码器的基于 DCT 变量。当比较 DCT 基和小波基的视频编码的 PSNR 差距时,发现它们的差别很小,比实际的图像要少很多。

6 结 论

分析图像和视频压缩算法应当从全面的系统观点来考虑,经过静止图像和视频的基于 DCT 和小波的编码比较后,可以得知在图像编码中的主要因素是量化器和熵编码器而不是小波变换和 DCT 的差别。对于静止的图像来

(下转第 84 页)

类,要找到一个合适的阈值,在这里采用的是一个线性支持向量机。还给出了对原始数据直接应用非线性支持向量机分类的结果,其核函数及其参数是用十折交叉验证找出的使训练错误率最低的多项式核 $K(x_i, x_j) = (1 + x_i \cdot x_j)^d, d = 4$ 。

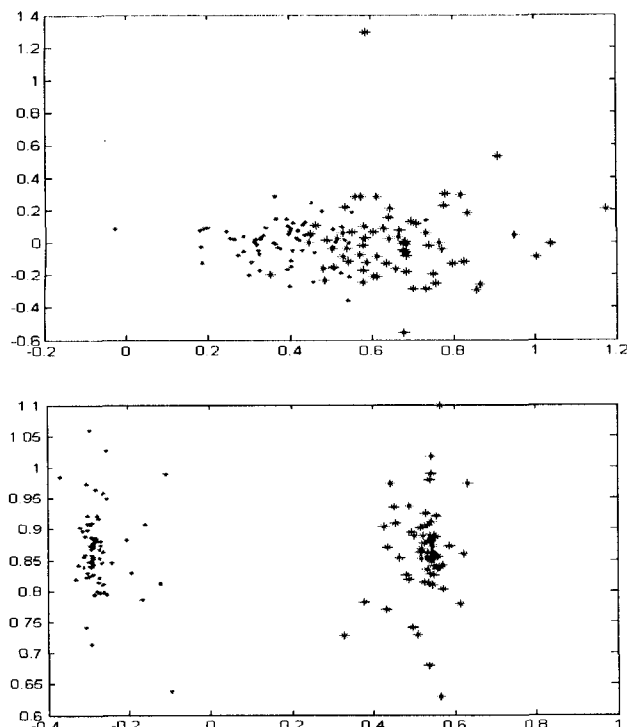


图 1 训练数据的投影分布

表 1 是对 KFDA 与 FDA 识别结果的比较,从表中可以看到,KFDA 比传统 FDA 识别效果有显著提高(提高了 5.7 个百分点),二者均比非线性支持向量机识别率高很多。KFDA 识别率比 BCI - competition 2003 公布的参赛者用其它方法获得的最佳识别率仅低 4 个百分点,而比第二名还高了一个百分点^[8],这说明 KFDA 确实是一种有效的特征提取方法。

表 1 两模型验证结果比较

方法	FDA	KFDA	非线性支持向量机
识别率	111/140=79.3%	119/140=85.0%	99/140=70.7%

(上接第 81 页)

说,小波变换和 DCT 之间的差别很小,对视频编码来说其差别则更小。而且 DCT 同样具有小波多分辨率图像的特性并且其计算较小波简洁,所以,DCT 变换具有很强的生命力,在未来的图像压缩中必将还能发挥重要的作用。

参考文献:

- [1] Cho N I, Lee S U. Fast algorithm and implementation of 2D DCT[J]. IEEE Transactions on Circuits and Systems, 1991, 38(3):297 - 305.
- [2] 钟伟才,刘 静,焦李成. H.263 中全零系数块预测的新方法[J]. 中国图形图像学报,2002,7A(12):1271 - 1273.

4 结 论

由于脑电信号的复杂性,有效的意识特征分散于大量的特征数据中,且具有较强的非线性。传统的 Fisher 判别只能局限于线性空间,在处理非线性问题时往往不能取得好的效果。核 Fisher 判别法是一种基于 Fisher 线性判别的非线性判别法。这种方法由于使用了核技巧,从而将 Fisher 判别推广到了非线性空间,有效地提高了判别分类效果。实验结果证明 KFDA 是一种非常有效的意识任务特征提取方法,非常值得推广应用到其它领域,而对 KFDA 的核函数进行深入研究也是很有价值的。

参考文献:

- [1] Eleanor A, Currana B, Stokesa M J. Learning to Control Brain Activity: A Review of the Production and Control of EEG Components for Driving Brain Computer interface (BCI) systems[J]. Brain and cognition, 2003, 51(3):326 - 336.
- [2] Schloegl A, Lugger K, Pfurtscheller G. Using Adaptive Autoregressive Parameters For a Brain - Computer - Interface Experiment[A]. The 19th Annual International Conference IEEE/EMBS [C]. Chicago, USA: [s. n.], 1997. 1533 - 1535.
- [3] Pfurtscheller G, Aranibar A. Event - related Cortical Desynchronization Detected by Power Measurements of Scalp EEG [J]. Electroencephalography and Clinical Neuro - physiology, 1977,42:817 - 826.
- [4] 薛建中,闫相国,郑崇勋. 用核学习算法的意识任务特征提取与分类[J]. 电子学报,2004,32(10):1749 - 1753.
- [5] 边肇祺,张学工. 模式识别[M]. 北京:清华大学出版社,2000.
- [6] Shawe - Taylor J, Cristianini N. Kernel Methods for Pattern Analysis[M]. 北京:机械工业出版社,2005. 48 - 68.
- [7] 苑诗松,王静龙,濮晓龙. 高等数理统计[M]. 北京:高等教育出版社,施普林格出版社,1997. 428 - 441.
- [8] Schlögl A. Outcome of the BCI - competition 2003 on the Graz data set[Z]. Berlin, Germany:Graz University of Technology,2003.

- [3] Xiong Z, Orchard M, Guleryuz O. A DCT - based embedded image coder[J]. IEEE Signal Processing Letters, 1996, 11(3): 289 - 290.
- [4] Xiong Z, Orchard R K. A comparative study of DCT and wavelet based image coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(8):692 - 695.
- [5] 陈 军,吴成柯. 基于 DCT 变换的内嵌静止图象压缩算法[J]. 电子学报,2002,30(10):1570 - 1573.
- [6] 赵德斌,张大鹏,高 文. 基于层次 DCT 变换的嵌入式图象编码方法[J]. 软件学报,2001,12(9):1287 - 1293.