

K-L 变换编码、DCT 变换编码、小波编码 在图像处理中的比较

连军江^{1, 2}, 李万社²

(1.伊犁师范学院奎屯校区 文理系, 新疆 奎屯 833200; 2. 陕西师范大学 数学与信息科学学院, 陕西 西安 710062)

摘 要: 讨论了 K-L 变换、DCT 变换和小波变换的基本原理, 对比了三种变换图像编码的性能, 指出了它们图像编码各自的优缺点.

关键词: K-L 变换; DCT; 小波变换; 图像编码

中图分类号: O157.4 **文献标识码:** A **文章编号:** 1673—999X (2007) 04—0019—05

离散余弦变换(Discrete Cosine Transform, DCT)是 1974 年由 Ahmed^[1]、Natarajan 和 Rao 等人提出的, 至今已有 30 多年的历史. 它的提出虽然晚于 FFT(Fast Fourier Transform), 但其性能更接近于理想的 K-L^[2]变换(Karhunen-Loeve Transform, KLT), K-L 变换是针对一类广泛的随机图像提出来的, 当对图像施加了 K-L 变换以后, 由变换结果而恢复的图像是将原图像在统计意义上的最佳逼近. 进一步研究表明, 对于常用的马尔可夫过程数据模型, 当相关系数 $r=1$ 时, K-L 变换退化为经典的 DCT 变换, 但由于 KLT 到目前为止还没有发现有效的快速算法, 所以 DCT 编码迅速发展成为 JPEG、MPEG、H.26x 等图像、视频编码标准中的核心. 而小波是近十几年来才发展并迅速应用到图像和语音分析等众多领域的数学工具, 是继 110 多年前建立傅立叶分析之后的一个重大突破, 特别是 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^{[3][4]}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, 这显然对传统的 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. 本文对比了三种变换图像编码的性能, 指出了它们图像编码各自的优缺点.

点.

1 变换编码的一般形式与意义

信源编码实质是一种变换, 就是将像素信息重组为更为紧凑的形式, 并产生一系列表示像素值的数字信号, 使之能在信道中更有效地传送. 图像通信^[5]要求是将图像信息清晰准确传至目的地, 而图像信息是高维信息, 内容复杂, 数量巨大, 如果直接传输, 由于受到信道和存储设备限制, 可能造成图像失真, 传输速度不佳, 甚至根本无法传输. 因此, 图像通信中关键技术是图像压缩编码. 而变换编码是目前已有的各种国际图像压缩编码标准中普遍采用的一种编码方法, 例如, 国际静止图像压缩编码标准 JPEG, 国际活动图像压缩编码标准 MPEG, 国际会议电视图像压缩编码标准 H.261 和极低比特率活动图像压缩编码标准 H.263 以及已知的各国高清晰的图像压缩编码方案. 在变换编码中, 编码是在图像方块基础上进行, 这导致变换编码的一个固有缺点——方块效应, 人眼对此非常敏感. 为解决此问题, 一个解决方案是对重叠相邻方块的部分数据点再做变换, 以此减少量化误差, 但这会降低效率. 为克服不足, Precen 和 Bradley^{[6][7]}提出了一种修正的 DCT(MDCT:Modified DCT). 它

收稿日期: 2007 - 09 - 03

基金项目: 国家自然科学基金资助项目 (10571113) .

作者简介: 连军江(1972—), 男, 陕西省澄城县人, 伊犁师范学院奎屯校区文理系教师, 陕西师范大学数学与信息科学学院在读硕士研究生, 主要研究方向为智能信号处理.

利用时域混叠消除(TDAC:Time Domain Aliasing Cancellation)技术,可以在不牺牲效率的情况下,达到减少量化误差的目的.显然,如果变换系数选择的恰当,所得变换系数间的相关性要明显小于原始像素之间的相关性,从而达到降低冗余度的目的.

2 三种变换编码的基本理论

2.1 K-L 变换编码的基本理论

K-L 变换是遥感图像增强和信息提取中用得最多的线性变换,是对原波段图像进行波谱信息的线性投影变换,在尽可能不减少信息量的前提下,将原图像的高维多光谱空间的像元亮度值投影到新的低维空间,减少特征空间维数,达到数据压缩、提高信噪比、提取相关信息、降维处理和提取原图像特征信息的目的,并能有效地提取影像信息.它可使原来多波段图像经变换后提供出一组不相关的图像变量,最前面的主分量具有较大的方差,包含了原始影像的主要信息,所以要集中表达信息,突出图像的某些细部特征,可采用主分量变换来完成. K-L 变换的基本原理如下所述:

任意时刻一个 $N \times N$ 抽样图像 $f_i(x, y)$, 可用 N^2 维矢量 X_i 来表示

$$X_i = [x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{iN^2}]^T, i = 1, 2, \dots, N$$

式中 x_{ij} 为第 i 次实现的矢量 X_i 的第 j 个分量. 鉴于图像信号是随机变量,度量随机变量之间的相关程度可用协方差(任意两个时刻上获得的随机变量的统计相关特性)矩阵表示,协方差矩阵是图像统计特性的重要反映. 矢量 X 的协方差矩阵 C_X 为

$$C_X = E[(X - M_X)(X - M_X)^T]$$

式中 $M_X = E(X)$, E 代表数学期望,此时根据 C_X 的特征值所求的特征矢量构成的矩阵 A , 用 A 对 C_X 实施以下变换 $C_Y = AC_X A^T$, 由矩阵理论知,上式变换的结果是典型的对角矩阵,对角矩阵的非零数可以按递减大小排列,显然 A 是正交矩阵,有 $A^{-1} = A^T$, 因此实施下列变换必是正交变换,即 $Y = A(X - M_X)$, 满足此条件的变换即为 K-L 变换.

容易从 K-L 变换的过程和结果得到两个重要结论: K-L 变换后的 Y 各元素之间是不相关的,易验

证 $E(Y) = E[A(X - M_X)] = E[A(X)] - AM_X = 0$, 通过变换消除了 $N \times N - N$ 各分量,如果取特征值前面 k 个较大者,则对 $Y = A(X - M_X)$ 实施逆变换,还

原后的 $X' = A_k^{-1} Y + M_X$ 的均方误差 $\sum_{j=k+1}^{N^2} \lambda_j$ 最小,即

还原的图像逼真度最佳. 由图像的统计特性决定变换矩阵是 K-L 变换的显著特征,也正是因为这一点, K-L 变换很难在实际中得到应用.

2.2 DCT 变换编码基本理论

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

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

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

$$\text{其中 } C(k) = \begin{cases} 1/\sqrt{2}, & k = 0 \\ 1, & k = 1, 2, \dots, N-1 \end{cases}$$

DCT 的各行(列)基矢量均是归一正交的(为了保证变换前后熵和范数相等). 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,而且编码电路简单. 综合 DCT 编码特性及相关性等方面因素,它仅次于理想的 K-L 变换,所以 DCT 在众多变换(DFT、斜变换、Haar 变换、KLT 和沃尔什变换等)中脱颖而出. 经过 DCT 变换后的系数主要值集中在左上角低频区域,其中直流值最大. 由于变换后系数矩阵右下角大多数高频系数趋向于零,这样就为结合 HVS 特性,采用(有死区)量化、Zigzag 扫描、截断 DCT 变字长压缩编码创造了条件,更重要的是为数字图像/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,该方法同样有助于问题更加准确地描述和求解,这也正是频率方法研究的意义所在.

2.3 小波编码的基本理论

小波分析是把一个信号分解成由原始小波经过移位和缩放后的一系列小波,因此小波是小波变

换的基函数. 将任意 $L^2(R)$ 空间中的函数 $f(t)$ 在小波基下进行展开, 称这种展开为函数 $f(t)$ 的连续小波变换(Continue Wavelet Transform, CWT), 其表达式为

$$WT_f(a, \tau) = \langle f(t), \varphi_{a, \tau}(t) \rangle = \frac{1}{\sqrt{a}} \int_R f(t) \overline{\varphi\left(\frac{t-\tau}{a}\right)} dt$$

式中 $a > 0$ 是尺度因子, τ 是移位因子. 当 $\varphi(t) \in L^2(R)$ 满足下面的可容许条件

$$C_\varphi = \int_R \frac{|\Phi(\omega)|^2}{|\omega|} d\omega < \infty \quad (\text{式中 } \Phi(t) \text{ 是 } \varphi(t) \text{ 的傅立叶变换}),$$

则存在连续小波变换的逆变换(ICWT)

$$\begin{aligned} f(t) &= \frac{1}{C_\varphi} \int_0^{+\infty} \frac{da}{a^2} \int_{-\infty}^{+\infty} WT_f(a, \tau) \varphi_{a, \tau}(T) d\tau \\ &= \frac{1}{C_\varphi} \int_0^{+\infty} \frac{da}{a^2} \int_{-\infty}^{+\infty} WT_f(a, \tau) \frac{1}{\sqrt{a}} \varphi\left(\frac{t-\tau}{a}\right) d\tau. \end{aligned}$$

小波变换是一种多分辨率的时频联合分析方法, 将任意信号表示成小波的叠加, 信号的叠加表示将信号分解为不同尺度级, 在每一尺度级, 该信号又在这一尺度级对应的分辨率下被分解, 尺度级对应着频率, 频率越高, 对应的分辨率越高. 其应用于图像压缩领域的出发点是多尺度信号分析和基于小波正交基族的信号分解, 该变换可将图像信号分解成不同空间分辨率、不同频率特征和方向特征子图像信号, 便于在失真编码中综合考虑人的视觉特性(HVS), 同时也利用图像的渐进传输, 其重构图像能有效克服 DCT 方法的方块效应和边缘震铃效应.

3 三种变换编码的性能比较

3.1 K-L 变换编码与 DCT 变换编码比较

自然图像的电视信号统计特性可近似于一阶 Markov 过程, 其协方差矩阵又近似于下式的 Toeplitz 矩阵

$$\begin{bmatrix} 1 & p & p^2 & \dots & p^{N-1} \\ p & 1 & p & \dots & p^{N-2} \\ p^2 & p & 1 & \dots & p^{N-3} \\ \dots & \dots & \dots & \dots & \dots \\ p^{N-1} & p^{N-2} & p^{N-3} & \dots & 1 \end{bmatrix}$$

形式, 而 DCT 变换矩阵与 Toeplitz 矩阵的特征向量矩阵非常相似, 即很近似自然图像的 K-L 变换矩阵. 因此对于 K-L 变换矩阵可借助 Toeplitz 矩阵来求特

征矢量及其特征矢量矩阵, 但研究发现对于四阶以上的 K-L 变换矩阵, 就不能照搬这种方法来求特征矢量, 所以说 K-L 变换的难点就在于求它的特征矢量矩阵, 也正是由于 DCT 的基向量逼近 Toeplitz 矩阵特征向量, 才使得 DCT 更接近于理想的 K-L 变换. 因此在实际应用中, DCT 变换可以认为是一种最佳变换. 总之, 在去相关性上, DCT 不如 K-L 变换, 在复原图像质量上, DCT 也次之. 但当图像的足够分块数 N 确定后, DCT 变换矩阵随之确定, 并与图像的统计特性相接近, 能在实际中较多地应用二维 DCT 变换编码或三维变换压缩编码, 同时可运动补偿帧间(内)预测去除时域冗余度.

3.2 DCT 变换编码与小波变换编码的比较

3.2.1 反演存在的条件不同

由于 DCT 变换的变换核是可逆的, 所以任何一个经过 DCT 变换后的信号都存在逆变换. 而对于小波变换, 所采用小波必须满足“可容许条件”, 逆变换才存在.

3.2.2 能量集中是这两种变换共同的特点

变换域编码的作用就是将时域(或空域)中的信号变换到另外一个正交矢量空间(即变换域), 并使变换域中各信号分量之间相关性很小或互不相关, 从而与变换前相比, 其能量更加集中. DCT 变换是先将整体图像分成 $N \times N$ 像素块, 然后对 $N \times N$ 像素块逐一进行 DCT 变换, 由于大多数图像的高频分量较小, 相应于图像高频成分系数经常为零, 因此 DCT 变换使实际信号的能量主要集中在低频段, 在允许一定误差的条件下, 对其进行编码时, 可以忽略能量小于某值的频率分量, 使数码率降低, 但数码率很低时要产生令人眼无法忍受的方块效应. 小波变换具有很好的时频局部化特性, 信号经小波变换后的能量也集中在少数变换系数上, 合理利用其变换系数分布特点, 可克服 DCT 变换编码产生的方块效应, 获得较好的压缩效果.

3.2.3 图像压缩后失真度比较

图像经 DCT 变换后, 编码图像被划分成相互不重叠的 $N \times N$ 的子块, 每个子块通过前向通道 DCT(FDCT)变换成 n^2 个 DCT 系数, 其中包含一个直流 DC 系数和 $(n^2 - 1)$ 个交流 AC 系数, 对于这 n^2 个 DCT 系数, 根据人眼视觉特性, 通过设置不同视觉阈值或量化电平, 将许多能量较小的高频分量量化为零, 这样, 就增加了变换系数中“0”的个数, 而能量较大的系数保留下来, 从而实现了较高

效率的压缩. 由于高频带是图像中的边缘、轮廓和纹理等细节信息的体现, 而且在各个频带所表示的边缘、轮廓等信息的方向是不同的, 其中包括水平方向、竖直方向、对角线方向的信息, DCT 变换处理的图像不具备这样的方向选择性, DCT 变换编码没有保留原图像块的精细结构, 不能反映原始图像块的边缘、轮廓等信息(这是由于 DCT 缺乏时、频局域性造成的), 因此它不能很好地反映编码图像的细节. 而小波变换编码处理图像具有良好的方向选择性, 这与人眼的视觉系统十分吻合.

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. Xiong^{[8][9]}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, 表 1 列出了算法编码重建图像比较结果. 通过对表中数据分析容易得出这样一个结论: 在失真率方面, 压缩率(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 仍然能够产生高压压缩率和高峰值信噪比, 因此 DCT 仍然是未来图像/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”对于某些图像压缩时还优于“SPIHT”, 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 1 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

3.2.4 多分辨率分析能力比较

小波变换具有多分辨率分析的能力, 算法是采用 Mallat^[10]提出的多分辨率塔形算法. 小波变换处理的图像各个频带分别对应了原图像在不同尺度和不同分辨率下的细节, 以及一个由小波变换分解级数决定的最小尺度、最小分辨率下对原始图像的最佳逼近. 研究表明^[11], 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: 一种是将常规的二维 DCT 通过系数重组来获得多分辨率的图像; 另一种是层式 DCT 编码分解形式. 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, 用于图像编码的任何变换都应该具有多分辨率分析特征, 在这一点上, 小波变换、DCT 变换均有相似之处, 但就目前来看, 小波变换的多分辨率分析能力比 DCT 表现得更有优势.

4 结论

分析图像和视频压缩方法应当以全面的系统观点来考虑, 三种不同的变换编码, 各自有不同的性能. DCT(与 KLT 相比)变换进行压缩的图像是那些信号带宽较窄的图像, 在 DCT 基础上的 ADCT(自

适应离散余弦变换)变换压缩效果更佳, 况且 K-L 变换本身也不是最佳上界. DCT 对变换信号的“紧凑有效性和独立于信号本身的变换以及变换后的系数游程编码”独具优势, 而小波仅在“变换”最为理想. 并且 DCT 经过 30 多年的理论与实践的丰富与发展, 它已在图像压缩上保证有较高的编码性能的同时, 从有损压缩发展到(近)无损压缩, 功能上也突破了原来的 JPEG 压缩编码, 实现了图像处理、分析及技术应用等, 形式上也从单纯的 DCT 走向与其它先进编码的融合. 基于 DWT 的编码在低编码压缩比的时候, 小波变换得到的图像质量好于传统基于 DCT 的 JPEG 编码方法, 而在高压压缩比时, 小波变换得到的图像质量降低很快, 需要有更好的系数组织和编码方式. 由于基于 DCT 变换编码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] N.Ahmed, T.Natarajan, K.T.Rao. Discrete cosine transform[J]. IEEE Trans. On Computers, 1974, C-23: 90-93.
- [2] 谷秋隆嗣, 编著. 快速算法与并行信号处理

- [M]. 薛培鼎, 徐国鼎, 译. 北京: 科学出版社, 2003.
- [3] 全子一. 图像信源压缩编码及信道传输理论与新技术[M]. 北京: 北京工业大学出版社, 2006.
- [4] Amir Said, William A Pearlman. A new fast and efficient image code based on set partitioning in hierarchical trees [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1996, 6(6): 243-250.
- [5] 何小海, 主编. 数字图像通信及其应用[M]. 成都: 四川大学出版社, 2006.
- [6] John P. Princen and Alan B. Bradley. Analysis/synthesis filter bank design based on time domain aliasing cancellation [J]. IEEE Trans. Acoust. Speech Sig. Proc, 1986, ASSP-34(5): 1153-1161.
- [7] A. W. Johnson, A. B. Bradley. Adaptive transform coding incorporating time domain aliasing cancellation [J]. Speech Comm, 1987(6): 299-308.
- [8] 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.
- [9] Xiong Z, Orchard M, Guleryuz O. A DCT-based embedded image coder [J]. IEEE Signal Processing Letters, 1996, 11(3):289-290.
- [10] [法] 马拉特(Mallat. S.). 信号处理的小波导引(原书第二版)[M]. 杨力华, 等译. 北京: 机械工业出版社, 2002.
- [11] 罗开仲, 黄土坦, 杨华民. DCT 算法及其与小波编码在图像处理中的比较[J]. 计算机技术与发展, 2006, 16(9).
- [12] Shaorong Chang, Lawrence Carin. A modified SPIHT algorithm for image coding with a joint MSE and classification distortion measure [J]. IEEE Transactions on Image Processing, 2006, 15(3): 713-725.
- [13] Byong-ki Lee, Yo-Sung Ho. Rate-distortion optimized zero-tree image coding using wavelet transforms[C]. Zagreb, Croatia. EC-VIP-MC 2003, 4th EURASIP Conference. 381-384.
- [责任编辑: 新柳]

K-L Arithmetic, DCT Arithmetic and Their Comparisons with Wavelet Transform Coding in Image Manipulation

LIAN Jun-jiang^{1,2}, LI Wan-she²

(1. Yili Normal University, Kuitun 833200, Xinjiang, China; 2. Shanxi Normal University, Xi'an 710062, Shanxi, China)

Abstract: Discussing the fundamental process about K-L transform, DCT transform and Wavelet transform, and making a contrast about the characteristics for three transform in the aspects of image code, this paper points out their own strengths and weaknesses for image code.

Key words: K-L transform; DCT; Wavelet transform; coding of image