

图像压缩中的快速方向离散余弦变换*

陈波⁺, 王红霞, 成礼智

(国防科学技术大学 理学院, 湖南 长沙 410073)

Fast Directional Discrete Cosine Transforms Based Image Compression

CHEN Bo⁺, WANG Hong-Xia, CHENG Li-Zhi

(School of Science, National University of Defense Technology, Changsha 410073, China)

+ Corresponding author: E-mail: chenbo0354@yahoo.com.cn

Chen B, Wang HX, Cheng LZ. Fast directional discrete cosine transforms based image compression. *Journal of Software*, 2011, 22(4): 826-832. <http://www.jos.org.cn/1000-9825/3805.htm>

Abstract: The traditional discrete cosine transform (DCT) can only sparsely represent the horizontal and vertical edges in images. The computation complexity of directional prediction DCT (DPDCT), which has the ability to represent more directions, is much higher. To overcome these shortcomings, the fast directional discrete cosine transforms (FDDCT) is proposed in this paper, in which the transformation is performed on the predefined direction mode. Compared with DPDCT, no interpolation is needed in FDDCT, so FDDCT can sparsely represent the anisotropic edges in much faster images. A special lifting algorithm is designed between adjacent blocks to ensure the perfect reconstruction, which compacts energy in edges lying across the blocks. The experimental results show that the computation of FDDCT is no more than 1.4 times that of DCT's. Coding with the same set partition method, PSNR compressed images that are combined with FDDCT are 0.4~1.6dB higher than those with DCT and DPDCT. Also, the edges and the details in the images are much clearer and less distortion exists.

Key words: image compression; sparse representation; directional transform; discrete cosine transform; zero tree coding

摘要: 传统的二维 DCT(discrete cosine transform)无法稀疏表示除水平或垂直方向以外的边缘,而具有强方向表示能力的方向预测离散余弦变换(directional prediction DCT,简称 DPDCT)计算复杂度又过高.针对这些问题,提出了一种快速方向离散余弦变换(fast directional discrete cosine transform,简称 FDDCT).该算法沿给定的方向模式进行变换,避免了 DPDCT 中的插值运算,可以快速、稀疏地表示图像中各向异性边缘信息.此外,FDDCT 通过设计块边界提升,在进一步集中边缘能量的同时保证了算法的完全重构.实验结果表明,FDDCT 计算复杂度不超过 DCT 的 1.4 倍;采用同样的编码方法,基于 FDDCT 的压缩图像与基于 DCT 以及 DPDCT 的压缩图像相比,峰值信噪比可提高 0.4dB ~1.6dB,而且边缘细节更加清晰、完整.

关键词: 图像压缩;稀疏表示;方向变换;离散余弦变换;零树编码

中图法分类号: TP391 文献标识码: A

* 基金项目: 国家自然科学基金(61072118); 国防科学技术大学预研基金

收稿时间: 2009-02-15; 修改时间: 2009-07-06; 定稿时间: 2009-11-26

图像压缩中采用的传统二维变换由一维变换的张量积得到,即对图像的行、列依次进行一维变换,无法稀疏表示除水平或垂直方向以外的边缘,从而导致高倍压缩图像中各向异性边缘特征的畸变与丢失.如何将图像中的方向信息应用到图像压缩的变换过程中,是图像压缩领域中具有挑战性的问题.

近年来出现了许多以增强二维变换的方向表示能力为目的的新型变换技术,其中,多尺度几何分析(multiscale geometric analysis,简称 MGA)^[1]方法作为小波变换的改进,在图像去噪、增强等领域得到了成功应用,但由于 MGA 方法变换系数的冗余度过高,故而较难应用到图像压缩中.另一类可以统称为方向预测变换(directional prediction transform,简称 DPT)^[2-5].DPT 基于小波变换(discrete wavelet transform,简称 DWT)或离散余弦变换(discrete cosine transform,简称 DCT)可提升分解^[6-8]的性质,选取图像边缘方向上的像素作为提升步骤中的预测系数,使得绝对值较大的高频系数的数量明显减少.DPT 中一般首先采用率失真优化^[8]或基于能量最小的原则^[9]来确定边缘方向,再沿图像边缘方向进行提升运算.与 MGA 相比,DPT 同样具备较强的方向表示能力.但与多数 MGA 需要过采样不同,DPT 无冗余,仅需额外存储少量方向信息,从而实现了图像边缘的高效稀疏表示.与传统 DWT 或 DCT 压缩方法相比,基于 DPT 压缩后图像的峰值信噪比(PSNR)可提高 1dB~2dB.

图像压缩中常用的离散余弦变换与离散小波变换均有对应的 DPT 算法.与 DWT(2 带滤波器)不同,DCT(N 点 DCT 为 N 带滤波器)的提升步骤中需要预测与更新的系数未必相邻^[8],当变换块方向与相邻块方向不一致时,多个提升运算需要同时更新同一系数,若采用传统提升算法,则变换无法完全重构.此外,现有方向预测 DCT(directional prediction DCT,简称 DPDCT)^[5]通过插值获取非整数采样点(像素点)的幅度来实现沿任意方向的变换.若插值核长度为 L ,则 DPDCT 的计算量约为 DCT 的 L 倍(块方向选择过程中还需进行多次 DPDCT).对于结构简单、存在诸如二进制变换^[8]等高速实现方案、特别适合对运算量敏感的计算环境(如实时处理)的类 DCT 变换而言,过高的计算复杂度必然会影响到 DPDCT 的实际使用.

针对 DPDCT 的缺点,本文提出了快速方向离散余弦变换(fast directional discrete cosine transform,简称 FDDCT)并应用于图像压缩.该变换借鉴 H.264 视频压缩标准^[9]中的帧内预测方法,可以避免方向变换中的插值运算,计算量与 DCT 相当.文中提出的块边界提升结构,还在充分去除块间方向冗余的同时保证了变换完全重构.

本文第 1 节分析 FDDCT 的相关理论并给出具体算法.第 2 节给出图像压缩的实验结果.

1 快速方向 DCT

1.1 方向预测

DPDCT 利用 DCT 矩阵可以提升分解的性质来改善变换的方向表示能力^[5].图 1 即为 8 点二进 DCT^[8]的提升格式实现流程,其中, $s_i(i=0, \dots, 7)$ 对应的归一化运算可与图像压缩中后续量化过程合并,其余提升系数均为分母为 $2^n, n \in \mathbb{Z}^+$ 的有理数,计算过程中仅包括加法与移位操作,适合快速实现.DCT 中每一个基本的提升运算都可以表示为 $x_m = x_m + a \cdot x_n$,其中: x 为图像像素的灰度值, m, n 为像素所在的行数; a 为提升系数,即图 1 中方框内的有理数.记这一提升运算为 $S(x[m], x[n], a)$,则 DCT 变换矩阵可以分解为 $S^1 S^2 S^3 \dots S^N$.

在对图像进行 DCT 列变换时, x_m 和 x_n 均位于图像的同列,如图 2(a)所示.而在 DPDCT 中, x_n 未必与 x_m 位于同一列,其位置由图像的边缘方向 θ 确定,如图 2(b)所示.当对图像进行 DPDCT 列变换时,首先依据能量最小原则确定图像边缘方向 θ ,然后对选定方向线与图像行的交点进行提升运算.因为 θ 可以为任意角度,所以方向线与图像行的交点可能为非采样点(非原图像像素,如图 2(b)所示中的 x_n 点).这些非采样点的灰度信息需要通过临近像素点插值得到,如文献[5]采用 8 点 sinc 插值,从而导致 DPDCT 的计算量增长到 DCT 的 8 倍以上.需要说明的是,由于图像分辨率有限以及方向信息占用存储空间等限制,DPDCT 对边缘方向进行了离散化,仅使用了 9 种不同的方向.

注意到 DPDCT 中计算量增加是由插值引起的,若方向线上的点均为图像中已有的采样点,则可以避免这一问题.因此,本文采用了类似 H.264 压缩标准帧内压缩中的预测模式.列变换过程中,沿事先选定的若干方向模

式中定义的方向线进行 DCT,方向线与图像行的交点均为图像采样点.FDDCT 列变换(变换块大小为 8)中使用的 5 种方向模式如图 3 所示,从 -45° (模式(a))逐步过渡到 45° (模式(e)).

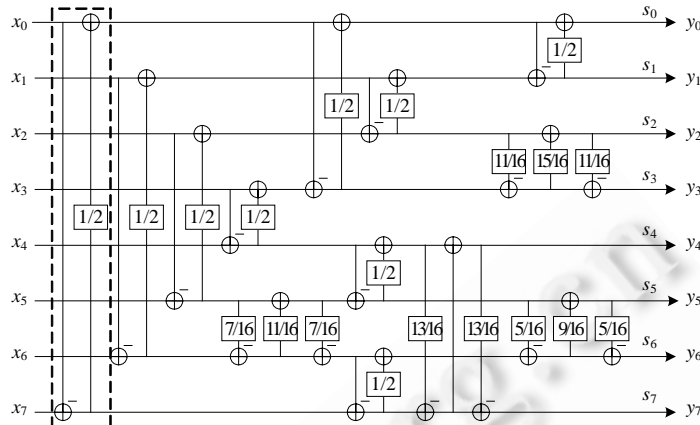


Fig.1 Flow chart of binary directional DCT
图 1 二进 DCT 提升格式流程图

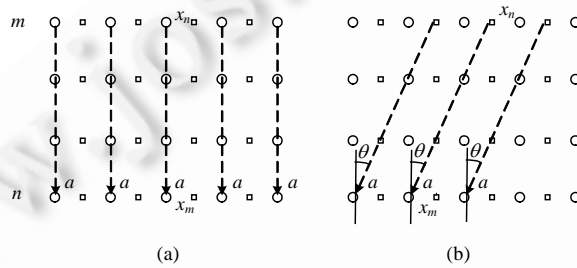


Fig.2 Directional lifting (the dashed line is directional line, \circ is the pixel, and \square is the middle point)
图 2 方向提升意图(虚线为方向线, \circ 表示图像像素, \square 表示相邻像素的中点)

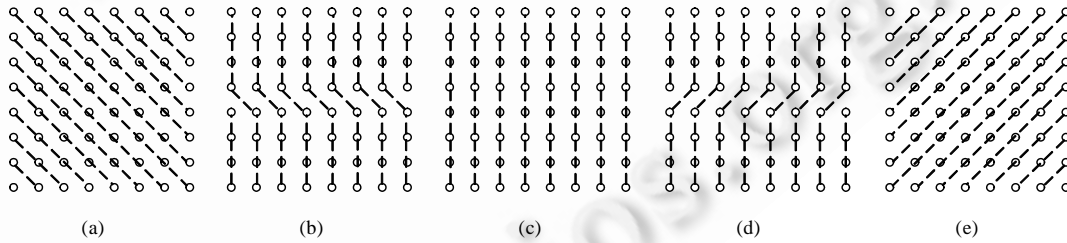


Fig.3 Five direction modes in FDDCT
图 3 FDDCT 中的 5 种方向模式

1.2 块边界的提升结构

与 DPDCT 中的方向线相比,图 3 中定义的方向模式可以避免插值运算,但在模式(a)~模式(e)中均存在长度小于 8 的方向线,DPDCT 中采用的传统等长 DCT 在这里不再适用.采用类似形状自适应 DCT^[10]中的变长 DCT 变换可以解决此问题.但是在变长 DCT 中,由于 DCT 低频系数的加权因子由信号长度来决定,当列变换长度不等时,会引起平均加权效应,从而破坏行变换后 DCT 系数的稀疏性.若通过低频校正^[10]来消除平均加权效应,变换的完全重构性质则无法保证,且增加了额外的计算量.并且,当 B_i^* 的长度不等于 2^n (n 为自然数)时,DCT 不存在

快速算法,计算量进一步增加.该方法与我们设计低计算复杂度方向变换的目标相违背.

注意到,造成方向线长度不等的根本原因在于 DCT 变换是分块进行的,且相邻块间的变换过程完全独立,从而使得跨块边界的方向线被截断.如果采用相邻块内像素来补充较短的方向线(如图 4 所示),则可以避免该问题.与独立块变换相比,这样做的好处还在于能够利用边缘方向相近的相邻块间的相关性,充分集中块边界处沿边缘方向的能量,进一步减少相邻变换块间方向性冗余.但是我们发现,若相邻变换块的方向模式不相同,则变换过程中存在单个像素点被多条方向线占用的情况(如图 4 所示中多边形框中黑色实心圆被 4 条方向线占用),因此需要重新设计块边界 DCT,以保证变换可完全重构.

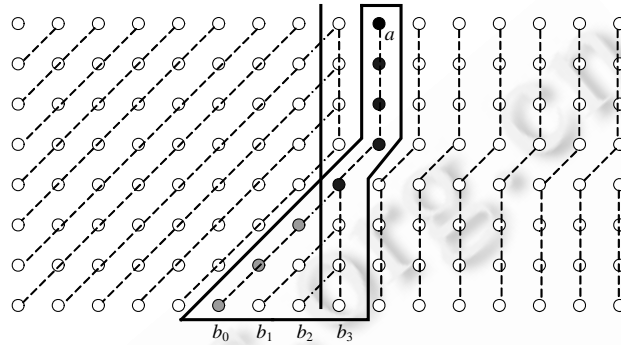


Fig.4 Direction line across blocks
图 4 块边界方向线示意图

基于 DCT 可提升分解的性质,设计块边界 DCT 算法只需考虑任意一个提升变换过程即可.不失一般性,以图 4 多边形框中像素点 $a, b_i, i=0, \dots, 3$ 对应的提升变换过程(即图 1 中虚线框内部分)为例,相应的块边界提升变换可分为预测(提升)与更新(对偶提升)两步实现:

$$\begin{aligned} \text{预测: } b_i^1 &= b_i - a, i = 0, \dots, 3 \\ \text{更新: } a^1 &= a + \frac{1}{2} \sum_{i=0}^3 w_i b_i^1 \end{aligned} \quad (1)$$

其中,输入为 a, b_i ,输出为 a^1, b_i^1, w_i 为加权因子.逆变换过程为

$$\begin{aligned} a &= a^1 - \frac{1}{2} \sum_{i=0}^3 w_i b_i^1 \\ b_i &= a + b_i^1, i = 0, \dots, 3 \end{aligned} \quad (2)$$

显然,此过程可完全重构.

在对图像进行 FDDCT 的过程中我们发现,上述块边界提升仅适用于相邻变换块方向模式相近的情况.当两者方向模式差别较大,如当前变换块方向模式为模式(a),而相邻块模式为模式(c)~模式(e)时,与 DCT 相比,FDDCT 系数的稀疏性一般不能得到有效改善.此时,块变换仍独立进行.若块内方向线到达块边界时长度小于 8,则沿块边界取像素点补足,保证变换等长.当用于补足长度的块边界像素点被多条方向线占用时,采用块边界提升即可.根据上述块边界处理方案,这里得到的图像块的方向 DCT 仅与其相邻图像块相关,因此可以采用文献[5]中的动态规划方法选择最优方向模式,这里不再赘述.

二维 FDDCT 对图像列、行分别进行变换,行、列变换过程类似.需要说明的是,由于模式(a)、模式(e)互为转置,二维 FDDCT 中,若首先进行列变换,则行变换时,方向模式仅包括模式(b)~模式(d)的转置,因此二维 FDDCT 中的方向模式可对应 8 个不同的方向.

$$DC_{FDDCT} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.125 & 0.125 & 0.125 & 0.125 \\ 0.125 & 0.125 & 0.125 & 0.125 & 0 & 0 & 0 & 0 \end{bmatrix}^T \begin{bmatrix} z_0^0 & z_1^0 & z_2^0 & z_3^0 & z_4^0 & z_5^0 & z_6^0 & z_7^0 \\ z_0^1 & z_1^1 & z_2^1 & z_3^1 & z_4^1 & z_5^1 & z_6^1 & z_7^1 \end{bmatrix} \quad (3)$$

公式(3)给出了方向模式(b)对应的 FDDCT 列变换中低通分量 DC_{FDDCT} 的 Z 变换表示,其中, $z_i(i=0, \dots, 7)$ 表示第 i 列,两个矩阵作点积运算.由于 FDDCT 中无需插值运算,与 DPDCT 的低通分量 DC_{DPDCT} (如文献[5]中的图 3(a)所示)相比, DC_{FDDCT} 系数矩阵中非零系数的个数仅有 8 个,与 DCT 相同,远小于 DC_{DPDCT} 系数矩阵中非零系数个数;同时,与 DC_{DPDCT} 中的浮点系数不同, DC_{FDDCT} 系数均可二进表示,适合快速实现. FDDCT 其他频带也具有类似的性质,因此,当图像块方向模式相同时, FDDCT 计算量与 DCT 相同,约为 DPDCT 的 1/8.

2 实验结果

我们分别将作者针对多带块变换提出的量化集合分裂编码方法^[11]应用于 DCT, DWT(JPEG2000 标准中采用的 CDF97 小波,变换层数为 4)与 FDDCT(在块方向信息编码过程中,首先使用临近块方向进行差分预测,然后对预测残差进行算术编码),对大量灰度图像进行压缩测试.部分标准测试图像的压缩结果(峰值信噪比(PSNR, 单位: dB))如图 5 所示.为了有利于快速实现,这里均采用二进变换.

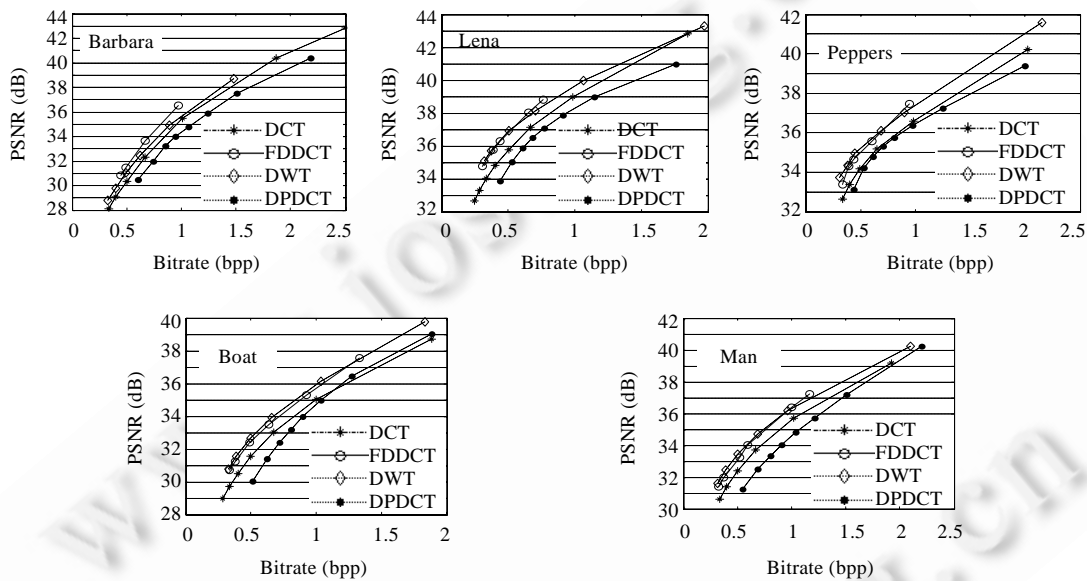


Fig.5 PSNR versus bit rate curves of different transforms

图 5 基于不同变换方法压缩图像峰值信噪比曲线(PSNR)

从图 5 可以看出:基于二进 FDDCT 与基于二进 DCT 的压缩方法相比,重构图像的 PSNR 提高约 0.4dB~1.5dB,更接近于 DWT;对于边缘较为丰富的图像,如 Barbara, PSNR 提高幅度较大.此外,图 5 中还给出了文献[5]中基于 DPDCT 的压缩结果,与其相比,本文方法的 PSNR 也提高了 1dB 以上.

从解压后的图像(如图 6 所示,为了显示清晰,仅截取了图像中有代表性的部分,并进行放大)可以看到,基于二进 FDDCT 的压缩方法主观质量最好.以 Barbara 为例,与 DCT 相比,基于 FDDCT 的压缩显著减少了黑白相间的线状纹理的模糊,块效应也得到一定抑制.同时,由于 FDDCT 中没有插值运算,避免了逆变换过程中系数量化误差的进一步扩散,因此,条纹比 DPDCT 压缩后的图像更加清晰.

考虑到 FDDCT 中的方向模式需根据图像自适应地选择,计算量与具体图像有关,难以给出精确的运算量估计,这里,我们通过统计大量标准测试图像的变换时间来进行比较.测试环境为 Pentium 4 2.93GHz CPU, Windows XP(SP2)操作系统,测试使用图像的大小为 512, 1 024, 2 048, 结果见表 1.本文提出的二进 FDDCT 所需平均时间与图像大小呈近似线性关系,均不超过 DCT 所需时间的 1.4 倍,约为 DWT 的 1/3,而 DPDCT 的计算时间为 DPDCT 的 5 倍以上(采用文献[5]中使用的 8 点插值).因此,本文提出的 FDDCT 实现了对图像各向异性边

缘信息的快速稀疏表示.

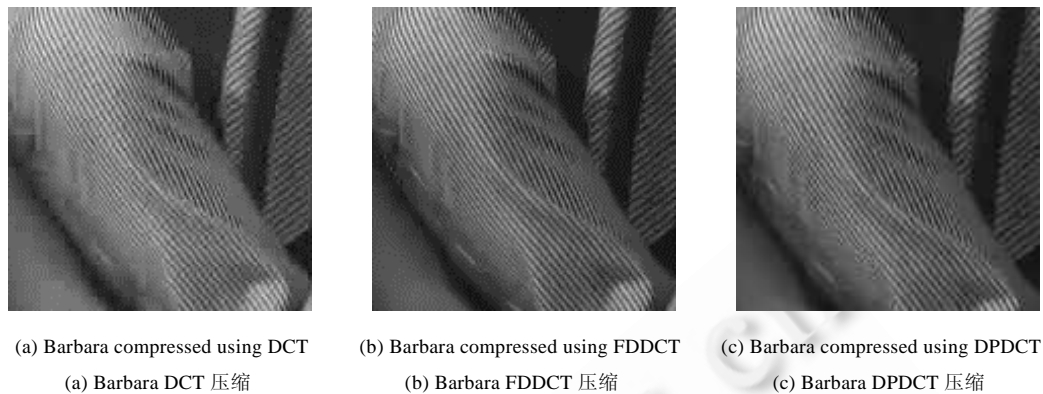


Fig.6 Parts of compressed Barbara (compression ratio: 16:1)

图 6 压缩图像局部放大(压缩比 16:1)

Table 1 Executing time (second) of different transforms (s)

表 1 FDDCT 与 DCT,DWT 的运算时间 (秒)

Size of images	Executing time of different transforms (s)			
	DCT	FDDCT	DPDCT	DWT
512×512	0.006 547	0.008 672	0.045 186	0.032 030
1024×1024	0.030 160	0.038 900	0.190 050	0.129 220
2048×2048	0.131 870	0.179 953	0.825 348	0.557 810

3 结 论

与 DWT 相比,DCT 及其改进变换的优势在于计算复杂度低、存储开销小.例如,针对下一代便携数码成像设备的 JPEG XR 压缩标准,采用的就是类 DCT 变换^[12].本文提出的 FDDCT 继承了 DCT 计算复杂度低的优点,同时方向表示能力更强,可以快速而稀疏地表示图像中各向异性边缘.实验结果显示,与 DCT 以及 DPDCT 相比,FDDCT 压缩图像质量均有较明显的提高;与 DWT 相比,FDDCT 压缩图像质量相近,但计算复杂度较低.由于其变换系数与 DCT 系数性质相近,FDDCT 可方便地应用于现有图像与视频压缩标准之中.

References:

- [1] Jiao LC, Hou B, Wang S, Liu F. Image Multiscale Geometric Analysis. Xi'an: Xidian University Press, 2008 (in Chinese).
- [2] Piella G, Heijmans HJAM. Adaptive lifting schemes with perfect reconstruction. IEEE Trans. on Signal Processing, 2002,50(7): 1620-1630. [doi: 10.1109/TSP.2002.1011203]
- [3] Wang D, Zhang L, Vincent A. Curved wavelet transform for image coding. IEEE Trans. on Image Processing, 2006,15(8): 2413-2421. [doi: 10.1109/TIP.2006.875207]
- [4] Ding WP, Wu F, Wu XL, Li SP, Li HQ. Adaptive directional lifting-based wavelet transform for image coding. IEEE Trans. on Image Processing, 2007,16(2):416-427. [doi: 10.1109/TIP.2006.888341]
- [5] Xu H, Xu JH, Wu F. Lifting-Based directional DCT-like transform for image coding. IEEE Trans. on Circuits and Systems for Video Technology, 2007,17(10):1325-1335. [doi: 10.1109/TCSVT.2007.903552]
- [6] Sweldens W. The lifting scheme: A custom-design construction of biorthogonal wavelets. Applied and Computational Harmonic Analysis, 1996,3(2):186-200. [doi: 10.1006/acha.1996.0015]
- [7] Yan YS, Shi QY. Reversible DCT mapping integers to integers and lossless image compression. Journal of Software, 2000,11(5): 620-627 (in Chinese with English abstract). http://www.jos.org.cn/ch/reader/view_abstract.aspx?flag=1&file_no=20000506&journal_id=jos

- [8] Liang J, Tran TD. Fast multiplierless approximations of the DCT with the lifting scheme. IEEE Trans. on Signal Processing, 2001, 49(12):3032–3044. [doi: 10.1109/78.969511]
- [9] Int'l Telecommunication Union. Advanced Video Coding for Generic Audiovisual Services. Telecommunication Standardization Sector of ITU, ITU-T Rec. H.264 ISO/IEC 14496-10 (AVC), 2005.
- [10] Zeng B, Fu JJ. Directional discrete cosine transforms—A new framework for image coding. IEEE Trans. on Circuits and Systems for Video Technology, 2008,18(3):305–313. [doi: 10.1109/TCSVT.2008.918455]
- [11] Chen B, Cheng LZ, Wang HX. LBT based low complexity image compression method. In: Tang YY, Wang SP, Lorette G, Yeung DS, Yan H, eds. Proc. of the Int'l Conf. on Pattern Recognition, Vol.1. Hong Kong: IEEE Computer Society, 2006. 941–944.
- [12] JPEG 2000 Digital Cinema Successes and Proposed Standardization of JPEG XR. 2007. <http://www.jpeg.org/newsrel19.html>

附中文参考文献:

- [1] 焦李成,侯彪,王爽,刘芳.图像多尺度几何分析理论与应用.西安:西安电子科技大学出版社,2008.
- [7] 闫宇松,石青云.可逆的 DCT 整型变换与无失真图像压缩.软件学报,2000,11(5):620–627. http://www.jos.org.cn/ch/reader/view_abstract.aspx?flag=1&file_no=20000506&journal_id=jos



陈波(1981—),男,山西太谷人,博士,讲师,主要研究领域为图像压缩.



成礼智(1963—),男,博士,教授,博士生导师,主要研究领域为快速算法,图像处理,信息隐藏.



王红霞(1977—),女,博士,副教授,主要研究领域为多尺度分析理论,图像处理.