

基于 DCT 和 DWT 图像编码方法的比较^{*}

A Comparison between Image Coding Methods based on DCT and DWT

王相海

(辽宁师范大学计算机与信息技术学院 大连116029)

(南京大学计算机软件新技术国家重点实验室 南京210093)

Abstract In the paper, the image coding methods based on DCT and DWT are analyzed, and the coding engines of the JPEG and JPEG-2000 are introduced. Finally, the performance comparisons of the some important DCT-based and DWT-based image coders are given, and the main functionalities of the JPEG and JPEG-2000 are analyzed.

Keywords DCT, DWT, JPEG, JPEG-2000

1. 引言

JPEG 标准^[1]已运作了十余年,是目前世界上最为流行的静态图像压缩标准之一, JPEG 算法以 DCT (Discrete Cosine Transform)^[2]作为其核心内容,采用此标准当压缩比达到20倍时图像基本不出现可见的失真,然而,随着图像压缩比的进一步提高, JPEG 算法通常会出现较为强烈的“块噪音”现象,而这种失真又是人眼所非常敏感的。基于 DWT (Discrete Wavelet Transform)^[3]的图像编码技术借助小波变换,将图像信号分解为具有不同空间分辨率、频率特性和方向特征的子带信号,实现了低频长时特征和高频短时特征的同时处理,使得信号的分解更适合于人的视觉特性和数据压缩的要求,为此新颁布的 JPEG-2000^[4]将其作为基本方法。基于此,正像文[5]所指出的,许多人在文献中通常将基于小波的最佳的编码方案与基于 DCT 的最差的编码方案相比,得出相应的一些结论。这往往会给人们带来一种错觉,好像凡是基于小波的图像编码一定会优于基于 DCT 的编码,并且 JPEG 也因 JPEG-2000 的出现而完全过时。事实上,这种提法并不完全正确。

本文首先对基于 DCT 和 DWT 的编码技术进行了分析,同时简单介绍了基于 DCT 的 JPEG 和基于 DWT 的 JPEG-2000 的编码机制;在此基础上对基于 DCT 和 DWT 的一些典型的图像编码方法的编码效果进行了比较;最后讨论和对比了 JPEG 和 JPEG-2000 的一些主要功能。

2. DCT 与 JPEG

2.1 DCT

1974年 N. Ahmed 等人首次提出了基于 DCT 的信号编码机制^[2]。DCT 可以看成是 Fourier-Cosine 级数的离散版本,它与 DFT (Discrete Fourier Transform)密切相关,可以采用 FFT (Fast Fourier Transform)算法在 $O(n \log n)$ 内完成其计算。然而,与 DFT 不同, DCT 是一个实值变换,并且利用其很少的变换系数便可实现对一个信号很好的逼近。二维 DCT 的一般定义如下列(1)、(2)式,其中(1)为 DCT 的正变换,(2)为反变换。

有关 DCT 及其应用的进一步讨论可参见文[6]。

$$C(n, m) = k(n)k(m) \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} f(x, y) \cos(\pi n \frac{2x+1}{2N}) \cos(\pi m \frac{2y+1}{2N}) \quad (1)$$

$$f(x, y) = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} k(n)k(m) C(x, y) \cos(\pi n \frac{2x+1}{2N}) \cos(\pi m \frac{2y+1}{2N}) \quad (2)$$

2.2 JPEG

1992年基于 DCT 的 JPEG 算法被确定为 JPEG 国际标准^[1],它是国际上彩色、灰度、静止图像的第一个国际标准。JPEG 标准对图像的编码提供了四种操作方式:顺序编码、累进编码、无失真编码和分层编码。图1和图2分别给出了 JPEG 标准中基于灰度图像的编码器和解码器的基本框图,框图中各项更为具体的讨论可参见文[6,7]。

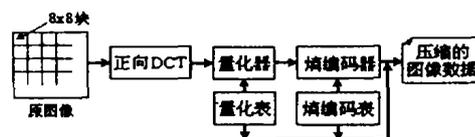


图1 JPEG 编码器框图



图2 JPEG 解码器框图

3. DWT 与 JPEG-2000

3.1 DWT

小波变换是80年代后期发展起来的应用数学分支,因其具有良好的空间-频率局部化性质,特别适用于对非平稳信号的分析,因而很快被 I. Daubechies 和 S. Mallat 引入到信号处理领域^[3,8,9]。下面对小波变换的基本情况进行介绍,更为详尽的讨论可参见文[10,11]。

定义1(小波函数) 设 $\Psi(t) \in L^2(\mathbb{R})$, 即 $\Psi(t)$ 为一平方

^{*}南京大学计算机软件新技术国家重点实验室资助项目,王相海 博士,副教授,主要研究领域为计算机图形学、图像及多媒体信息处理。

可积函数,若其 Fourier 变换 $\Psi(\omega)$ 满足条件 $\int_{\mathbb{R}} (|\Psi(\omega)|^2/\omega) d\omega < \infty$, 则称 $\Psi(t)$ 为小波母函数。

由该定义容易得到 $\int_{-\infty}^{\infty} \Psi(t) dt = 0$, 即 $\Psi(t)$ 为具有正负交替的震荡波形。

定义2(连续小波基函数) 将小波母函数 $\Psi(t)$ 进行伸缩和平移可得: $\Psi_{a,\tau} = a^{-1/2} \Psi((t-\tau)/a)$, 其中 $a > 0, \tau \in \mathbb{R}$, 通常称 $\Psi_{a,\tau}(t)$ 为依赖于参数 a, τ 的小波基函数, a, τ 分别被称为尺度因子和平移因子。

定义3(连续小波变换) 将 $L^2(\mathbb{R})$ 空间中任意函数 $f(t)$ 在小波基下展开, 称这种展开为函数 $f(t)$ 的连续小波变换(简记为 CWT), 其表达式为:

$$WT_f(a, \tau) = \langle f(t), \Psi_{a,\tau}(t) \rangle = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} f(t) \overline{\Psi\left(\frac{t-\tau}{a}\right)} dt$$

小波变换与 Fourier 变换最重要的不同是小波基具有尺度因子 a 、平移因子 τ 两个参数, 这样小波变换是一种变分辨率的时频联合分析方法, 当分析低频(对应大尺度)信号时, 其时间窗很大; 当分析高频(对应小尺度)信号时, 其时间窗减小。这恰恰符合实际问题中高频信号持续时间短、低频信号持续时间长的自然规律。

定义4(离散小波) 小波基函数 $\Psi_{a,\tau}(t)$ 的 a, τ 限定在一些离散点上取值, 通常取 $a = a_0^m$ (m 为整数且 $a_0 \neq 1$), 此时所得的小波基函数被称为离散小波基函数。

在实际工程应用中, 诸如奇异性检测、图像处理等方面, 一般取 $a_0 = 2$, 此时经过归一化后可得:

$$\text{离散小波基函数: } \Psi_{m,n}(t) = 2^{-\frac{m}{2}} \Psi(2^{-m}t - n), (m, n \in \mathbb{Z});$$

离散小波变换(DWT): 对 $\forall f \in L^2(\mathbb{R})$, 函数 $f(t)$ 的 DWT 为: $WT_f(m, n) = \int_{\mathbb{R}} f(t) \overline{\Psi_{m,n}(t)} dt, (m, n \in \mathbb{Z});$

离散小波变换的逆变换: 函数 $f(t)$ 的逆小波变换为:

$f(t) = \sum_{j,k} \langle f, \Psi_{j,k} \rangle \tilde{\Psi}_{j,k}(t)$, 其中 $\tilde{\Psi}_{j,k}(t)$ 为 $\Psi_{j,k}(t)$ 的对偶函数, 此时基函数 $\{\Psi_{j,k}\}_{j,k \in \mathbb{Z}}$ 需构成一个“标架”, 即对 $\forall f \in L^2(\mathbb{R}), \exists A, B \in \mathbb{R}^+$, 使得:

$$A \|f\|^2 \leq \sum_{m,n} |\langle f, \Psi_{m,n} \rangle|^2 \leq B \|f\|^2.$$

此外, S. Mallat 在研究图像处理问题时提出了多分辨率分析理论(MRA)^[4]. MRA 不仅为 $L^2(\mathbb{R})$ 空间正交小波基的构造提供了一个简便方法, 而且为小波的分解与重构提供了快速算法, 即 Mallat 算法。多分辨率分析在正交小波变换理论和实际工程应用中具有重要的地位。

如果将前面所介绍的小波称为第一代小波的话, 那90年代后期由 Sweldens 等提出的基于提升方案(Lifting Scheme)的小波^[12]通常被称为第二代小波。与第一代小波相比, 基于 Lifting Scheme 小波是一种较 Mallat 算法更快、更简便和更易操作的另类小波变换方法, 并且所有第一代小波均可通过 Lifting Scheme 小波的方法来实现变换。更为详尽的叙述可参见文[12~14]。

3.2 JPEG-2000

JPEG-2000 是一个新的静态图像编码标准^[2], 它已于2001年正式颁布。该标准适用于具有不同特性的各类静态图像, 同时, 它在一个统一的和集成的系统内支持不同的解码模式, 像客户/服务器模式、适时传输模式、有限存储器和带宽模式等等。

JPEG-2000 编码采用的是以 DWT 为基础的分块处理方

式, 限于篇幅, 下面仅给出 JPEG-2000 编码的总体过程, 更详细的讨论可参见文[4]。

JPEG-2000 编码总体经历如下过程:

- 将图像分解为分量图像(Components);
- 将图像和分量图像分解为一个长方形的瓦片(Tiles)。瓦片分量是原始图像或重构图像的基本单位;
- 对瓦片分量进行小波分解, 得到各分解层, 这些分解层由一个个子带构成;
- 对每个子带进行量化并将其划分为一个个码块;
- 按码块的扫描顺序提取码块系数的位平面(bit-planes)并进行熵编码;
- 如有 ROI(Region Of Interest)编码的要求, 则优先对相关的系数进行编码。

JPEG-2000 所规定的码流格式和解码方式可支持多种可分级解码, 诸如分辨率可分级、质量可分级以及 ROI 可分级等等, 同时还可以在在一定程度上直接对编码信息进行直接访问。

4. 基于 DCT 和 DWT 图像编码的比较

4.1 基于 DCT 和 DWT 典型编码方法的编码效果比较

借助 DWT, 图像信号可以被分解成许多具有不同空间分辨率、频率特性和方向特征的子带信号, 实现低频长时特征和高频短时特征的同时处理, 从而使得图像信号的分解更有利于产生高效的数据压缩方法。除了利用图像小波变换的频率特性进行编码外, 人们还对小波变换的空间压缩特性和系数分布的相似性进行了分析, 提出了具有高效、简捷和可精确控制压缩比等特点的零树小波图像编码方法, 比如 EZW^[15]、SPIHT^[16] 和 EBCOT^[17] 等等。近年来, JPEG 中所体现的基本 DCT 的编码模式也得到不断的改进, 出现了许多有效的编码方法, 比如优化量化矩阵^[18]、优化阈值选取^[19] 以及联合二者^[20] 的编码模式等等; 也出现了基于 DCT 系数的零树编码方法^[21,22]。文[5,21]中对基于 DCT 和 DWT 的典型的编码方法进行了讨论, 结合二文我们给出如下比较结果。

表1 基于 DCT 和 DWT 典型图像编码效果比较

| Rate | PSNR(dB) | | | | | | | | | |
|------|------------------|---------|--------------------|---------|------------|---------|---------|---------|-----------|---------|
| | Baseline JPEG[6] | | Optimized JPEG[22] | | EZ-DCT[23] | | EZW[17] | | SPIHT[18] | |
| b/p | Lena | Barbara | Lena | Barbara | Lena | Barbara | Lena | Barbara | Lena | Barbara |
| 0.25 | 31.6 | 25.2 | 32.30 | 26.70 | 32.25 | 26.83 | 33.17 | 26.77 | 34.11 | 27.58 |
| 0.50 | 34.9 | 28.3 | 35.90 | 30.60 | 36.00 | 30.82 | 36.28 | 30.53 | 37.21 | 31.39 |
| 0.75 | 36.6 | 31.0 | 38.10 | 33.60 | 38.06 | 33.70 | — | — | 39.04 | 34.25 |
| 1.00 | 37.9 | 33.1 | 39.60 | 35.90 | 39.62 | 36.10 | 39.55 | 35.14 | 40.40 | 36.41 |

从比较结果可以看出, 尽管 DWT 较 DCT 具有更好的空间分辨率、频率特性和方向特性, 但仅从编码效果上看, 基于 DWT 的零树小波编码方案之所以有较好的效果, 除部分由于采用了 DWT 对图像进行分解外, 很重要一点是采用了“零树”量化模式, 如果对于图像的 DCT 系数采用相应的优化量化模式, 同样可以取得较好的编码小波, 甚至可以取得超过 EZW 的编码效果, 而 DCT 一般又较 DWT 有更小的计算复杂度。当然, DWT 好的性质为实现具有其它特性的编码机制提供了可能, 比如可分级性、鲁棒性、对码流的直接操作性等等。这也是 JPEG-2000 采用 DWT 作为其核心技术的一个重要原因。

4.2 JPEG 与 JPEG2000 的主要功能比较

以下我们对 JPEG 和 JPEG-2000 的主要功能进行讨论, 同时就二标准对这些功能的支持情况进行简单比较, JPEG-2000 与其它标准更为详尽的对比可参见文[23]。

总体而言, JPEG 是为了给出一个适用于连续色调图像的压缩方法, 它能适用于任何种类的连续色调的图像, 其长宽比不受限制; 它的计算复杂性是可控制的; 此外, JPEG 算法除了能够满足顺序编码和无失真编码操作外, 还能在一定程度上实现质量可分级(累进编码)和空间可分级(分层编码)的编码功能。

与 JPEG 相比, 基于 DWT 的 JPEG-2000 码流形式更加灵活, 操作更加方便, 从而为实现灵活多样的需求提供了可能。作为一个新的编码标准, JPEG-2000 的主要功能包括:

- 具有较好的低比特率压缩效果: 一般要求在比特率低于 0.25b/p 时仍能提供优秀的编码效果;
- 支持有损和无损的编码模式: JPEG-2000 渴望在对码流的质量可分级解码中能够自然实现无损解码;
- 支持大型图像的编码: 突破了 JPEG 所限制的 64k×64k 的界限;
- 支持有噪音环境下的图像传输: 支持位错误的鲁棒性, 这对无线通讯的应用是重要的;
- 支持各种类型的图像: 它们可以是自然图像、文本图像、计算机生成图像, 也可以是医学图像、遥感图像和合成图像等;
- 支持像素精度级的质量和分辨率的渐进传输, 该需求对于图像网络的多质量服务具有重要意义;
- 支持 ROI 编码: 允许图像的某一部分以更好的质量被解码, 这在一定程度上体现了基于内容的编码能力。

此外, JPEG-2000 还具有开放式的结构、实时的编解码能力、固定大小的存储空间和对码流直接访问等需求。

表2给出了 JPEG 和 JPEG-2000 所具有一些功能的情况, 其中“+”号表示支持该功能, “-”号表示不支持, “+”号越多, 表示所支持的程度越好。

表2 JPEG 与 JPEG2000 主要功能比较

| 功能 | JPEG | JPEG2000 |
|----------|-------|----------|
| 无损压缩 | + | +++ |
| 有损压缩 | +++ | +++++ |
| 渐进传输 | ++ | +++++ |
| ROI 编码 | - | +++ |
| 码流的直接访问 | - | ++ |
| 低复杂度 | +++++ | ++ |
| 错误鲁棒性 | ++ | +++ |
| 编码各种类型图像 | ++ | +++ |

结束语 本文首先简单介绍了 DCT 与 JPEG 以及 DWT 与 JPEG-2000 的编码机制, 在此基础上, 对基于 DCT 和 DWT 典型算法的图像编码效果和 JPEG 与 JPEG-2000 的主要功能进行了比较。从分析中可以看出, 基于 DWT 的图像编码方法不应该完全取代基于 DCT 的编码方法, 特别在低于 20 倍压缩比的低复杂度应用领域。然而, 在要求低码率、可分级编码、ROI 编码以及对码流的直接操作等应用领域, JPEG-2000 将会展示其美好的应用前景。

参 考 文 献

1 Pennebaker W B, Mitchell J L. JPEG Still Image Data Compres-

sion Standard. New York: Van Nostrand Reinhold, 1992

2 Ahmed N, Natarajan T, Rao K R. Discrete cosine transform. IEEE Transactions on Computers, 1974, (23): 90~93

3 Daubechies I. Orthonormal bases of compactly supported wavelets. Communication of Pure and Applied Mathematics, 1988, 410(xII): 909~996

4 ISO/IEC JTC1/SC29/WG 1, ISO/IEC FCD 15444-1: Information technology JPEG 2000 image coding system: Core coding system [WG 1 N 1646R], march 2000. http://WWW.jpeg.org/FCD15444-1.htm

5 Xiong Z, et al. A Comparative Study of DCT- and Wavelet-Based Image Coding. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(5): 692~694

6 Rao K R, Yip P. Discrete Cosine Transforms - Algorithms, Advantages, Applications, Academic Press, 1990

7 钟玉琢, 蔡莲红, 李树青, 史元春. 多媒体计算机技术基础及应用. 北京: 高等教育出版社, 1999, 6

8 Mallat S. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, 11(7): 674~693

9 Mallat S. Multi-frequency channel decomposition of images and wavelet model. IEEE Transactions on ASSP, 1989, 37(12)

10 彭玉华. 小波变换与工程应用. 科学出版社, 1999

11 Chan Y T. Wavelet Basics. Kluwer Academic publishers, Norwell, MA, 1995

12 Sweldena W. The lifting scheme: A construction of second generation Wavelets[J], SIAM J. Math. Anal., 1997, 29(2): 511 ~ 546

13 Sweldena W, Schroder P. Building your own wavelets at home[A]. Wavelets in Computer Graphics. ACM SIGGRAPH Course Notes [R/OL], 1996

14 Sweldena W. The lifting scheme: A custom-design construction of biorthogonal wavelets[J]. Journal of Appl. And Comput. Harmonic Analysis, 1996, 3(2): 186~200

15 Shapiro J M. Embedded Image Coding Using Zerotrees of Wavelet Coefficients. IEEE Transactions on Signal Processing, 1993, 41(12): 3445 ~ 3462

16 Said A, Pearlman W. An New, Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees. IEEE Transactions Circuits and Systems for Video Technology, 1996, 6(3): 243 ~ 249

17 Taubman D. High performance scalable image compression with EBCOT. IEEE Trans. Image Proc., 2000, 9(7): 115 ~ 117

18 Wu S, Gersho A. Rate-constrained picture adaptive quantization for JPEG baseline coders. In Proc. ICASSP'93, 1993, 5: 389~392

19 Ramchandran K, Vetterli M. Rate-distortion optimal fast thresholding with complete JPEG/MPEG decoder compatibility. IEEE Trans. Image Processing, 1994, 3(9): 700~704

20 Crouse M, Ramchandran K. Joint thresholding and quantizer selection for transform image coding: Entropy-constrained analysis and applications to JPEG. IEEE Trans. Image Processing, 1997, 6(2): 285~297

21 Xiong Z, Guleryuz O G, Orchard M T. A DCT-based embedded image coder. IEEE Signal Processing Letters, 1996, 3: 289~290

22 Monro D M, Dickson G J. Zerotree Coding of DCT Coefficients. Proc. IEEE ICIP, 1997, 2: 625~628

23 D. S.-Cruz, Ebrahimi T. An analytical study of JPEG 2000 functionalities. In Proc. of the IEEE ICIP, 2000, 2: 49~52