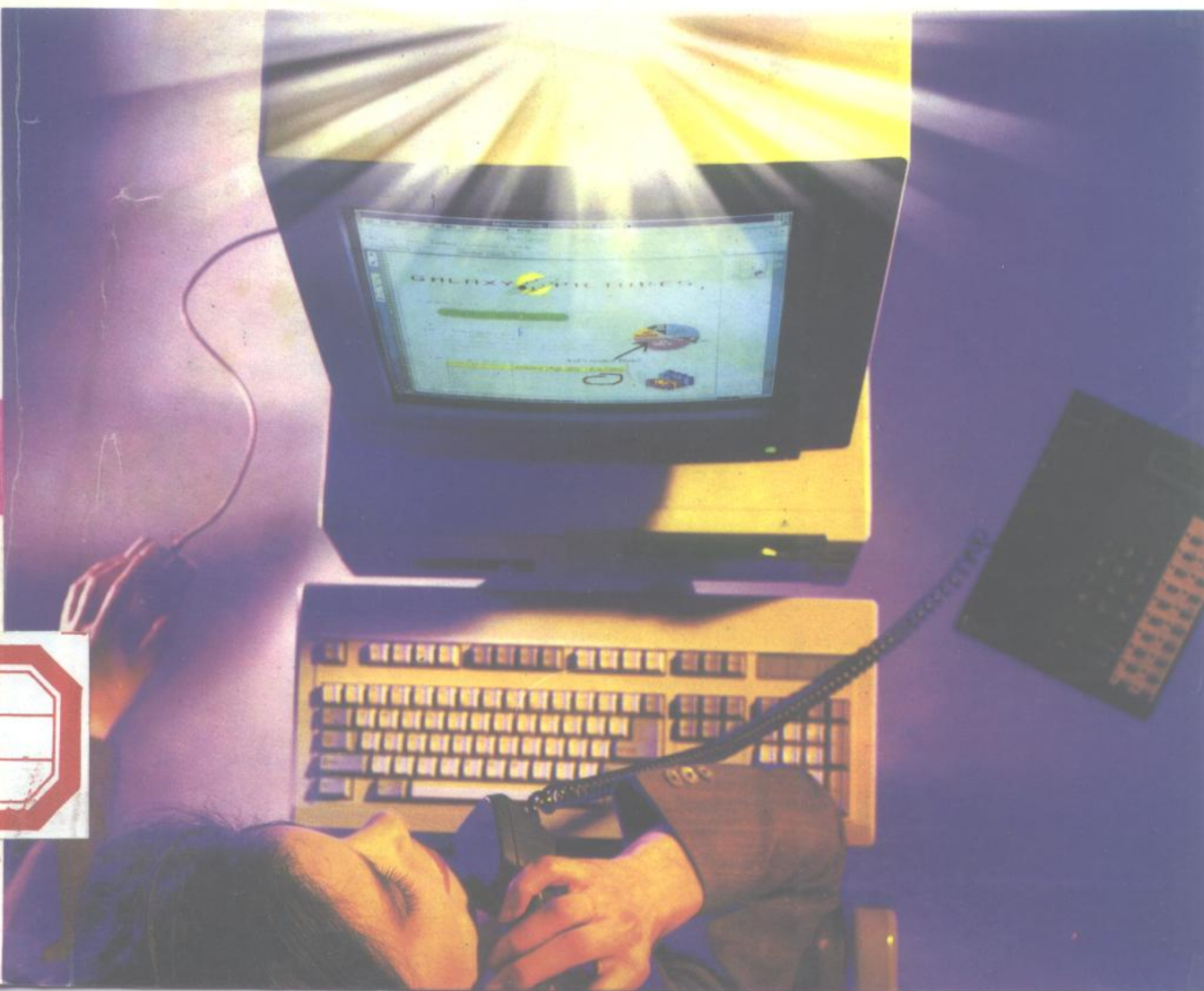


多媒体计算机技术丛书 张福炎主编

连续色调静止图像的 压缩与编码—— JPEG

蔡士杰 岳华 刘小燕 编译 张福炎 审校



南京大学出版社

TN 919.8

406320

C03

多媒体计算机技术丛书

张福炎 主编

连续色调静止图像的压缩与编码

——JPEG

蔡士杰 岳 华 刘小燕 编译
张福炎 审校



南京大学出版社

1995·南京

(苏)新登字 011 号

内 容 简 介

多媒体技术是 90 年代计算机研究、开发及应用领域的一个热点,而图像压缩技术是多媒体技术中的关键技术之一。本书系统地介绍了由国际标准化组织(ISO)制订的连续色调静止图像的数字压缩与编码标准(即 JPEG 标准),对有关概念、算法、处理及压缩数据文件格式等进行了较详细的叙述。

本书可供从事多媒体技术研究、开发的人员参考,也可作为大专院校有关课程的教学参考书。

JS179 / 14

多媒体计算机技术丛书
连续色调静止图像的压缩与编码
——JPEG

蔡士杰 岳 华 刘小燕 编译
张福炎 审校

*

南京大学出版社出版

(南京大学校内 邮政编码:210093)

江苏省新华书店发行 常熟市印刷二厂印刷

*

开本:787×1092 1/16 印张:14.125 字数:349 千

1995 年 6 月第 1 版 1995 年 6 月第 1 次印刷

印数:1—5000

ISBN 7-305-02804-8/TP·127

定价:19.80 元

《多媒体计算机技术丛书》编委会

主 编 张福炎

编 委	蔡士杰	余崇智	唐 棠	程景云
	宋顺林	潘金贵	李存珠	黄宜华
	袁春风	张志力	武港山	

《多媒体计算机技术丛书》序

多媒体技术近年来发展异常迅速,它已成为计算机研究、开发及应用领域的一个热点。多媒体技术是计算机、影像处理、语音技术、通讯技术等高度结合的一个产物,它是推动当今计算机技术发展的巨大动力之一。多媒体技术的广泛应用,必将为计算机产业的大发展提供机会。美国苹果电脑公司总裁 John Scully 预言:“多媒体计算机机会像 PC 机在 80 年代独领风骚那样,改变 90 年代世界的面貌。”

多媒体把文字、图形、图像、动画、音频和视频集成到计算机中,使人们能更加自然、更加“人类化”地使用信息。人机交互的这种变化极大地拓展了我们的信息空间,满足了人们把多媒体信息做统一处理的需要。

在普通的 386 以上的 PC 机上,装上声音卡、视频卡,配上 CD-ROM 驱动器等所谓的多媒体升级套件,就变成了一台多媒体 PC 机,简称 MPC。如有可能,再挂上 B-ISDN 网(宽带综合业务数字通信网),在多媒体 Windows 的支持下,和音响设备、电视设备连在一起,我们便可欣赏、编辑、创作立体声音乐、电视图像等,使你进入一个多媒体世界。

我国从事计算机研究、开发及应用的广大人员对多媒体技术极为关注,对它的发展与应用表现出很大的热情,但由于这一技术较新,又是多学科交叉,目前有关这方面的技术资料、教材、参考书等还不多见。为帮助国内读者学习和掌握多媒体技术,推动和促进多媒体技术的开发与应用,南京大学多媒体计算机研究所张福炎教授组织编写了这套《多媒体计算机技术丛书》,并由南京大学出版社出版,该丛书包括:

1. 《多媒体计算机技术概论》;
2. 《CD 光盘的原理与应用》;
3. 《数字声音处理及其在多媒体计算机中的应用》;
4. 《连续色调静止图像的压缩与编码——JPEG》;
5. 《运动图像及其伴音的压缩与编码——MPEG》;
6. 《多媒体计算机软件开发环境与工具》;
7. 《多媒体计算机软件写作工具及应用》;
8. 《多媒体计算机人机接口》;
9. 《计算机动画技术》。

根据这一领域技术发展的需要,我们还将不断地充实、完善本丛书的内容,力求及时反

映国内外的最新发展动态和研究成果，以飨读者。

多媒体技术研究是一项跨世纪的重大工程，只要我们深入持久地开展研究，联系实际应用，加强国际合作与交流，就一定能取得可喜的成果，必将为我国的经济建设作出更大的贡献。

《多媒体计算机技术丛书》编委会

1994年10月

前 言

多媒体计算机技术是 90 年代计算机发展中的一个热点。图像数据的压缩和编码是多媒体技术中的关键技术之一。为了促进我国多媒体计算机技术的研究 开发和应用,我们编译了本书,并作为《多媒体计算机技术丛书》之一,献给广大读者。

本书共九章。第一、二两章介绍有关的基本概念;第三章给出 JPEG 压缩数据格式;第四、五章讨论编、解码的控制过程和原理;第六至第九章分别叙述了四种模式的编、解码操作。附录中绘出了 8 位精度亮度和色度的典型哈夫曼表、算术编码测试数据序列以及 JPEG 压缩与解压缩模块源程序(只含哈夫曼编码)。

本书由蔡士杰、岳华、刘小燕合作编译,蔡士杰统编全稿,张福炎教授作了全面的审校。

由于编者的经验与水平有限,书中难免还存在一些缺点和错误,殷切希望广大读者批评指正。

编 者

1994 年 10 月

目 录

第一章 概论	(1)
1.1 引言	(1)
1.2 编码器与解码器	(1)
1.3 对编码器的要求	(2)
1.4 对解码器的要求	(2)
1.5 JPEG 压缩算法分类	(3)
1.6 无失真编码	(3)
1.7 基于 DCT 的编码	(4)
1.8 编码处理模式	(5)
1.9 熵编码选择	(6)
第二章 源图像编码处理的约定和有关概念	(7)
2.1 源图像	(7)
2.1.1 源图像取样与二维矩阵	(7)
2.1.2 取样精度	(8)
2.1.3 数据单元	(8)
2.2 源图像数据的编码次序	(8)
2.2.1 最小编码单元	(8)
2.2.2 多分量的定义	(8)
2.2.3 非交叉次序和交叉次序	(9)
2.2.4 不完整 MCU 的补全	(10)
2.2.5 图像、帧和扫描	(11)
2.3 DCT 压缩	(11)
2.3.1 平移	(11)
2.3.2 FDCT 计算中取样的方向	(11)
2.3.3 FDCT 和 IDCT	(11)
2.3.4 DCT 系统的量化和反量化	(12)
2.3.5 差分 DCT 编码	(12)
2.3.6 Z 形序列	(13)
2.4 点变换	(13)
2.5 几种编码处理的基本特点	(14)
第三章 JPEG 压缩数据格式	(15)
3.1 引言	(15)
3.2 压缩数据格式中的基本术语	(15)
3.2.1 参数	(15)
3.2.2 标记	(15)
3.2.3 标记的赋值	(16)
3.2.4 标记段	(17)
3.2.5 熵编码数据段	(17)

3.3	压缩图像数据格式中的语法定义	(17)
3.4	顺序 DCT、累进 DCT 和无失真模式的交换格式语法	(18)
3.4.1	高层语法	(18)
3.4.2	帧标题的语法	(19)
3.4.3	扫描标题语法	(20)
3.4.4	表定义和杂标记段语法	(22)
3.5	层次式编码处理交换格式语法	(27)
3.5.1	高层次模式语法	(27)
3.5.2	DHP 段语法	(27)
3.5.3	EXP 段语法	(28)
3.6	简化格式	(28)
3.6.1	压缩图像数据的简化格式	(28)
3.6.2	表说明数据的简化格式	(28)
3.7	总结	(29)
第四章	编码和解码的控制过程	(32)
4.1	编码控制过程	(32)
4.1.1	图像的编码控制过程	(32)
4.1.2	帧数据的编码控制过程	(33)
4.1.3	扫描的编码控制过程	(33)
4.1.4	重启段的编码控制过程	(33)
4.1.5	MCU 的编码控制过程	(34)
4.2	解码控制过程	(34)
4.2.1	图像的解码控制过程	(34)
4.2.2	帧压缩数据的解码控制过程	(35)
4.2.3	扫描的解码控制过程	(35)
4.2.4	重启段的解码控制过程	(36)
4.2.5	MCU 的解码控制过程	(37)
第五章	算术编码与哈夫曼编码	(38)
5.1	算术编码基本原理	(38)
5.2	二进制算术编码原理	(38)
5.2.1	区间的递归划分	(38)
5.2.2	自适应二进制算术编码	(39)
5.3	二进制算术编码的实现	(39)
5.3.1	编码器的寄存器约定	(39)
5.3.2	编码约定和近似计算	(40)
5.3.3	过程 Code __ 0(s) 和 Code __ 1(s)	(40)
5.3.4	编码器的概率估计	(43)
5.3.5	编码器的重规范化	(45)
5.3.6	编码器的初始化	(48)
5.3.7	编码的终止	(49)
5.4	二进制算术解码	(50)

5.4.1	二进制算术解码原理	(51)
5.4.2	解码器的寄存器约定	(51)
5.4.3	解码约定和近似计算	(51)
5.4.4	解码过程	(51)
5.4.5	解码器的重规范化	(53)
5.4.6	解码器初始化	(55)
5.5	哈夫曼编码	(55)
5.5.1	哈夫曼编码原理	(55)
5.5.2	哈夫曼表说明	(56)
5.5.3	生成哈夫曼表说明的过程	(57)
5.5.4	从交换格式中的哈夫曼表说明到码表和码长的转换	(60)
5.5.5	字节内的位次序	(62)
第六章	基于 DCT 的顺序操作模式	(63)
6.1	基于 DCT 的顺序编码过程	(63)
6.1.1	基于 DCT 的顺序控制子程序和编码模型	(63)
6.1.2	基本哈夫曼编码子程序	(64)
6.1.3	扩充到 4 套哈夫曼表的 8 位精度、顺序 DCT 编码过程	(68)
6.1.4	采用算术编码的顺序 DCT 编码过程	(68)
6.1.5	采用哈夫曼码和 12 位精度的顺序 DCT 编码	(76)
6.1.6	采用算术编码和 12 位精度的顺序 DCT 编码过程	(77)
6.2	基于 DCT 的顺序解码过程	(77)
6.2.1	基于 DCT 的顺序控制子程序和解码模型	(77)
6.2.2	基本哈夫曼解码	(78)
6.2.3	扩充到 4 组哈夫曼表的 8 位精度的顺序 DCT 解码过程	(83)
6.2.4	使用算术编码的顺序 DCT 解码过程	(83)
6.2.5	哈夫曼编码和 12 位精度的顺序 DCT 解码过程	(87)
6.2.6	使用算术编码和 12 位精度的顺序 DCT 解码过程	(88)
第七章	基于 DCT 的累进操作模式	(89)
7.1	基于 DCT 的累进编码过程	(89)
7.1.1	基于 DCT 的累进操作的控制子程序和编码模型	(89)
7.1.2	使用哈夫曼编码的累进编码过程	(91)
7.1.3	使用算术编码的累进编码过程	(97)
7.2	DCT 的累进解码	(99)
第八章	无失真操作模式	(100)
8.1	无失真编码过程	(100)
8.1.1	无失真编码器的控制子程序	(100)
8.1.2	无失真编码的编码模型	(100)
8.2	无失真解码过程	(103)
8.2.1	无失真解码器的控制子程序	(103)
8.2.2	无失真解码模型	(103)

第九章 层次操作模式	(104)
9.1 引言	(104)
9.2 层次模式的编码	(104)
9.2.1 层次式图像编码的控制过程.....	(104)
9.2.2 差分帧的编码控制过程.....	(106)
9.2.3 差分帧的编码模型.....	(106)
9.2.4 差分帧的熵编码器的修改.....	(106)
9.3 层次模式的解码	(107)
9.3.1 层次模式下图像解码的控制过程.....	(107)
9.3.2 差分帧的解码控制过程.....	(108)
9.3.3 差分帧的解码模型.....	(108)
9.3.4 差分帧熵编码解码器的修改	(108)
附录	(109)
I. 8 位精度亮度和色度的典型哈夫曼表	(109)
1. DC 系数的典型哈夫曼表	(109)
2. AC 系数的典型哈夫曼表	(110)
3. DC 差分编码的典型哈夫曼表的说明	(118)
4. AC 差分编码的典型哈夫曼表的说明	(118)
II. 算术编码测试数据序列	(120)
III. 部分 JPEG 源程序	(133)
参考文献	(214)

第一章 概 论

1.1 引 言

多媒体计算机技术是 90 年代计算机发展中的一个热点。多媒体技术是将正文、图形、音频、视频、动画、通讯及计算技术结合在一起的一种技术。由于音频、视频、动画的数据量十分巨大,因此,音频数据和视频数据的压缩存储技术与解压缩还原技术,成了多媒体技术中的关键技术之一。

如果考虑中分辨率全屏幕(Full Screen, 640×480)、真彩色(True Color, 24 位)、全动作(Full Motion, 25—30 帧/秒)的动画,则播放 1 秒钟视频画面的数据量为:

$$640 \times 480 \times 24 \times 30 = 221184000 \text{ 位} = 27648000 \text{ 字节}$$

即使是降低彩色逼真性要求,改用 8 位彩色,每秒显示帧数改为 25 帧,数据量可达:

$$640 \times 480 \times 8 \times 25 = 61440000 \text{ 位} = 7680000 \text{ 字节}$$

如此庞大的数据量带来两个方面的问题:其一是巨大的存储容量要求问题。即使使用 600MB 的光盘存储,也只能存放供播放 21.8 秒的全屏幕、真彩色、全动作的视频数据或 78 秒的降低彩色位数及播放频率的视频数据。第二个问题是读出速度的需求问题。目前光盘的读出速度约为 150KB/s,而一帧上述画面的数据量分别是 921.6KB 或 307.2KB。也就是说,1 秒钟从光盘上读出的数据还远远不够形成一帧图像。如果要保持 25 帧/秒的播放速度,8 位彩色,则 1 秒钟从光盘读取的数据只能提供屏幕上 77×77 的小区域的动画显示。如果保持 30 帧/秒的播放速度、24 位全彩色,则 1 秒钟从光盘上读取的数据只能提供给 40×40 的屏幕区域的动画显示。

由此可见,压缩与解压缩问题不解决,则多媒体计算机的实用化得不到保证。为此,国际上有关专家对图像压缩方法进行了多年的研究。国际标准化组织(ISO)在 1990 年成立了专门的工作小组(ISO/IEC JTC1/SC2/WG10),与在它之前成立的 ISO/IEC JTC1/SC2/WG8 中的一个特别小组 Joint Photographic Expert Group (JPEG)合作制订了静态图像压缩编码标准:Digital Compression and Coding of Continuous Still Images, 也称为 JPEG 标准。

本书对 JPEG 的有关概念、算法、处理及压缩数据文件格式等进行较详细的介绍。

1.2 编码器与解码器

图像的压缩就是将图像的位图信息(Bit Map)转变成另一种能将数据量缩减的表达形式,该操作称为编码(Encode)。在显示时则重新将编码数据形式转回到位图形式,该操作称为解码(Decode)。编码的压缩比是指编码前后的数据量的比值。

编码器(Encoder)是编码过程的具体体现。如图 1.1 所示,一个编码器输入原图像数据和表说明,经过一组指定的过程处理后产生并输出压缩的图像数据。

解码器(Decoder)是解码过程的具体表现。如图 1.2 所示,一个解码器以压缩的图像数据和相应的表说明作为输入,经过一组指定的过程处理后产生并输出重建的数字图像数据。

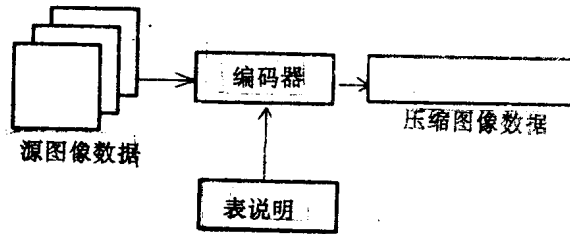


图 1.1 编码器

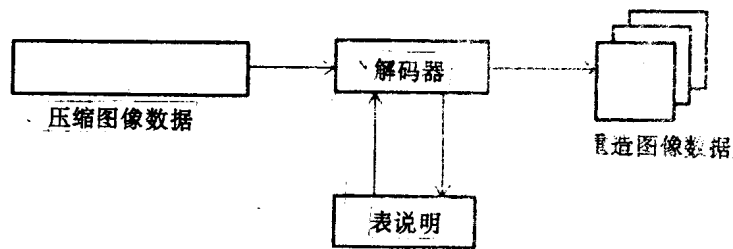


图 1.2 解码器

对同一图像来说,在对其压缩数据解码时使用的表说明(Table Specification)必须和编码时所用的表说明一致,所以如果将压缩图像数据交给另一应用程序去解码,则必须将编码时所用的表说明与压缩图像数据一起交给另一应用程序。如图 1.3 所示,在应用系统 A 中,编码生成的压缩图像数据要连同编码时所用的表说明按压缩图像数据交换格式组织成文件一起送给应用系统 B,才能在 B 系统中被解码并重建原图像。

1.3 对编码器的要求

一个编码器是一种或多种编码处理的具体表现,它至少要满足下列两点之一:

- ①编码器必须以合适的精度把源图像数据转换为符合压缩图像数据交换格式的压缩数据。
- ②编码器必须以合适的精度把源图像数据转换为符合简化格式的压缩图像数据。

1.4 对解码器的要求

一个解码器应做到下列三点:

- ①它应能以适当的精度将交换格式中的压缩图像数据在指定表说明数据的配合下解码重构图像数据。
- ②可接受并存储与简化格式的压缩图像数据一致的任何表说明数据,供解码过程使用。
- ③在建立表说明后能以适当精度将任何符合简化格式的压缩图像数据解码并重建原图像数据。

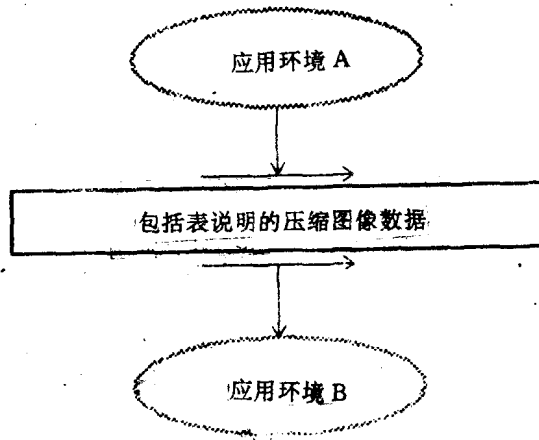


图 1.3 压缩图像数据的交换格式

1.5 JPEG 压缩算法分类

JPEG 压缩算法分为两大类:无失真压缩 (Lossless Compression) 和有失真压缩 (Lossy Compression)。使用无失真压缩算法将源图像数据转变为压缩数据,该压缩数据经对应的解压缩算法处理后可获得与源图像完全一样的重建图像。有失真压缩算法基于离散余弦变换 (DCT),所生成的压缩图像数据经解压缩生成的重建图像与源图像在视觉上保持一致。一般说来,压缩比越大,视觉上的一致性越差。

1.6 无失真编码

无失真编码过程由图 1.4 给出。预测器最多可组合三个相邻位置 a、b 和 c 的取样值来形成 x 位置的预测值 (图 1.5),然后用 x 位置的实际值去减去预测值以形成差值。差值再经哈夫曼 (Huffman) 编码或算术 (Arithmetic) 编码进行无失真的熵编码 (Entropy)。

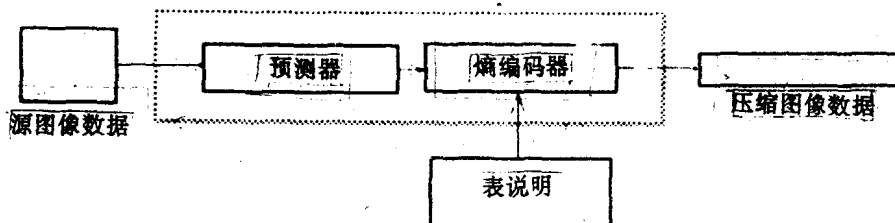


图 1.4 无失真编码简图

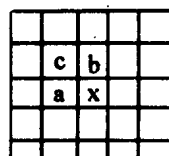


图 1.5 使用三相邻取样值作预测

1.7 基于 DCT 的编码

基于 DCT 的编码过程由图 1.6 给出。图中仅给出了单分量图像的简单情况。在分量时，对每一分量的处理独立地进行。

在编码过程中，输入的分量取样值按 8×8 分块，每一块经正向离散余弦变换 (FDCT — Forward DCT) 转换成 64 个按序排列的 DCT 系数，其中第一个称为直流系数 DC，其余 63 个称为交流系数 AC。

然后，按照图 1.6 中表说明中的对应的 64 个值来对这 64 个系数进行量化 (Quantization)。量化表并不由标准给定，而是由应用根据图像特征、显示设备及观察条件来指定。

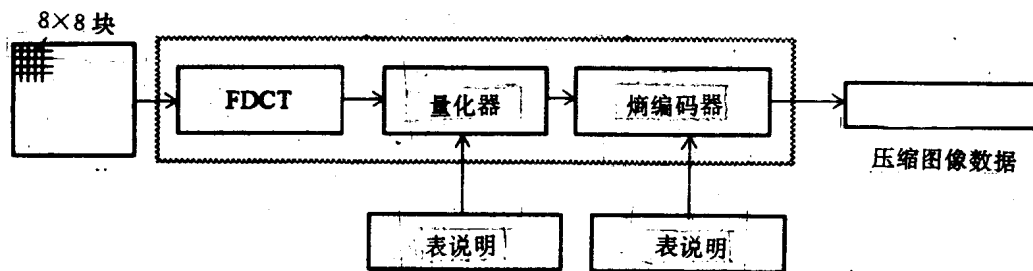


图 1.6 基于 DCT 的编码器简图

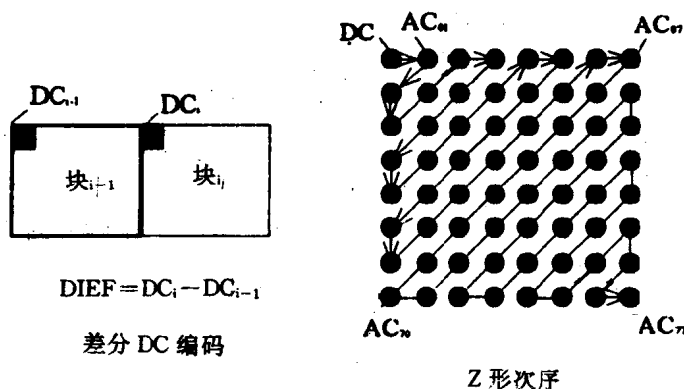


图 1.7 熵编码前量化系数准备

量化以后，直流分量 DC 和 63 个交流分量 AC 分别进行熵编码。前一块的 DC 用来作为对当前块的 DC 的量化预测，计算出差值进行编码。63 个 AC 系数并不使用这样的差分编码，而

是转换成一维的“Z”形序列(图 1.7)。

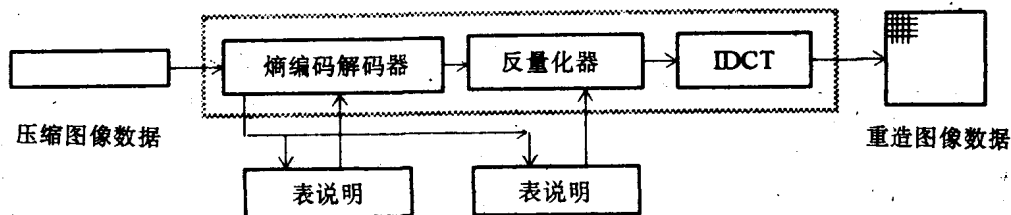


图 1.8 基于 DCT 的解码器简图

量化后的系数接着被送给熵编码过程,以便进一步压缩数据。这里可使用两种编码过程中的一个。如果使用哈夫曼编码,则要向编码器提供哈夫曼表说明。如果使用算术编码,则必须提供算术编码条件表说明。

解码是编码的逆过程。图 1.8 给出了解码的处理过程。熵编码解码器对量化的 DCT 系数熵编码的 Z 形序列解码。接着,在反量化之后,DCT 系数再经 DCT 的逆变换(IDCT—Inverse DCT)处理成为 8×8 的取样块。

1.8 编码处理模式

编码处理一共有四种模式:基于 DCT 的顺序模式(Sequential DCT-based),基于 DCT 的累进模式(Progressive DCT-based),无失真模式(Lossless)和层次模式(Hierarchical)。但在编码器实现时不一定要包括所有这四种模式。

在基于 DCT 的顺序模式中,所有 8×8 取样块从左到右,从上到下依次输入。每一块被量化后立即进行熵编码并作为压缩图像数据的一部分输出,从而尽可能地降低了对系数存储的要求。

在基于 DCT 的累进模式下,所有 8×8 的块仍然按上述顺序编码,但对图像需进行多趟扫描。这通过在量化和熵编码之间增加一个与图像一样大小的系数缓存来实现。当每一块量化后,它的系数存入该缓存。所有块均完成量化并存入缓存后分几次扫描来编码,每次只有部分被编码。因此解码时输出的图像也是分几次完成的。图 1.9 给出了顺序和累进两种模式下解码过程的对比,缓存中的量化后的系数被分批扫描的处理方式有二种,一种是频谱选择法,另一种是逐次逼近法。在频谱选择法中,每次扫描 Z 形序列中的指定段并进行编码,每一段位于 8×8 块的频谱高或低部分。在逐次逼近中,每一次的扫描系数的各位不全部参与编码,而是从高到低按指定位数逐次编码。两种过程都可以单独使用,也可灵活地组合使用。

在层次式模式中,一幅图像被编码成一些帧的序列。这些帧给出参考重建分量,用作后继帧的预测。对于给定分量,除了第一帧以外,差分帧是对源分量和参考重建分量的差值编码的结果。差值的编码可以仅仅使用基于 DCT 的处理,仅仅使用无失真处理,或使用基于 DCT 的但最后对每一分量进行无失真处理。向下取样和向上取样滤波器可能被用来提供如图 1.10 所示的空间分辨率的金字塔。另外,层次模式可以在指定空间分辨率下改善重建分量的质量。

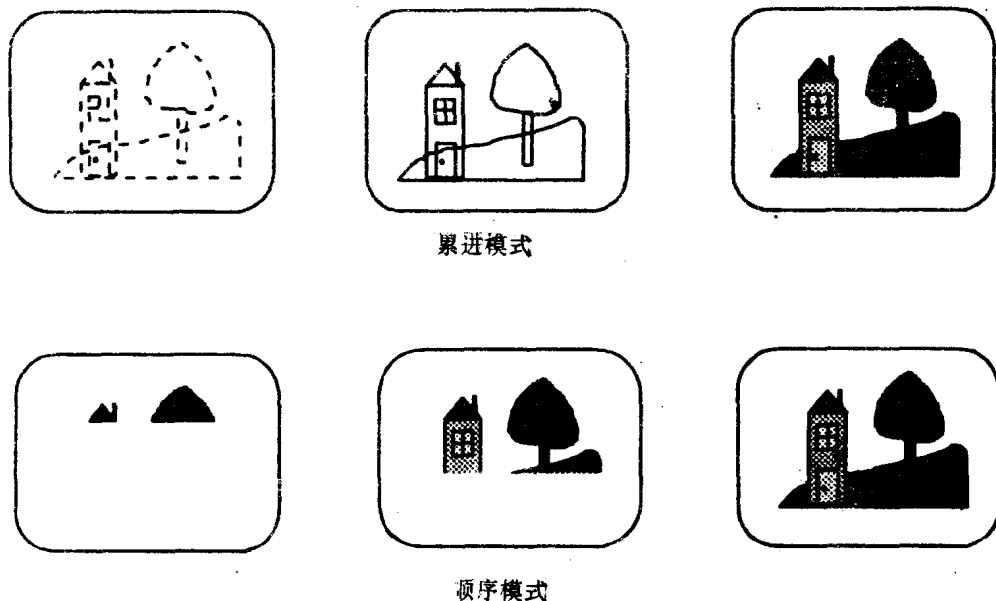


图 1.9 累进和顺序模式的显示过程

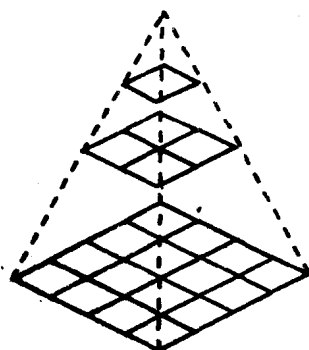


图 1.10 层次式多分辨率编码

层次模式提供了与基于 DCT 的累进模式类似的逐步表示,但适用于多种分辨率环境。层次模式还提供了逐步变换到无失真阶段的能力。

1.9 熵编码选择

熵编码过程有两种:哈夫曼编码和算术编码。哈夫曼编码使用图 1.1 和 1.2 中表说明所确定的哈夫曼表。算术编码过程使用由表说明指明的算术编码调整表。哈夫曼表没有规定缺省值,所以应用程序可以选择适合于其自己环境的哈夫曼表。算术编码调整有规定的缺省值。

JPEG 基本系统的顺序处理使用哈夫曼编码。扩充系统的基于 DCT 的处理和无失真处理,既可使用哈夫曼,也可使用算术编码。