

常用多媒体 文件格式与 压缩标准解析

姜楠 王健 编著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

常用多媒体文件格式与 压缩标准解析

姜楠 王健 编著

電子工業出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书面向与多媒体数据打交道的科研人员和工程技术人员,以及其他对多媒体数据感兴趣的人员。本书从实际应用角度出发,先对多媒体数据的表示理论做了简单介绍,然后详细讲解目前科学研究和工程中常用的多媒体文件格式和压缩标准。

我们的编写目的并不是让读者通读全书内容,而是要让它成为对读者的工作和学习有所帮助的参考书,可以随时查阅相关内容,提高工作和学习效率。

书中各章相互独立,每章介绍一种文件格式或者压缩标准,分别介绍了图像、音频、视频和文本四种多媒体数据,包括 BMP, GIF, JPEG, JPEG 2000, PSD, MPEG-2, MPEG-4, AVI, WAVE, MIDI, PDF 等文件格式或者压缩标准。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

常用多媒体文件格式与压缩标准解析/姜楠,王健编著. —北京:电子工业出版社,2005.6
ISBN 7-121-01171-9

I. 常… II. ①姜… ②王… III. 多媒体—文件—数据压缩—标准 IV. TP274-65

中国版本图书馆 CIP 数据核字(2005)第 042834 号

责任编辑:王春宁

印 刷:北京市李史山胶印厂

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

经 销:各地新华书店

开 本:787×1092 1/16 印张:19.5 字数:494 千字

印 次:2005 年 6 月第 1 次印刷

印 数:4 000 册 定价:29.00 元

凡购买电子工业出版社的图书,如有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系。
联系电话:(010) 68279077。质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

前 言

多媒体数据处理是一门实用的学科,同时又需要一定的理论基础。计算机及相关学科的出现使得多媒体数据处理迅速发展并得到普遍应用。目前,多媒体数据处理已经成为工程学、计算机科学、信息科学、统计学、物理、化学、生物学、医学甚至刑事侦察、娱乐等社会科学领域中各学科学习和研究的对象,多媒体数据处理的需求与日俱增。在信息社会中,多媒体数据处理无论是在理论上还是在实践上都存在着巨大的潜力,已经有越来越多的科研机构 and 人员投入到多媒体数据处理的方方面面的研究中来。

要很好地利用数字多媒体数据,首先要把数据保存下来,目前常用的保存方法是按照一定的格式要求将数据保存成文件的形式。因此,无论是实际使用,还是理论仿真,只要想对数字多媒体数据进行利用或者处理,就要对多媒体数据的文件格式有所了解。本书就是解析多媒体文件格式和压缩标准的一本参考书。

本书面向与多媒体数据打交道的科研人员和工程技术人员,以及所有对多媒体数据感兴趣的人员。本书从实际应用角度出发,介绍了目前科学研究和工程中常用的多媒体文件格式和压缩标准。

本书的编写目的并不是让读者通读全书内容,而是要让它成为对读者的工作和学习有所帮助的参考书,可以随时查阅相关内容,提高工作和学习效率。

书中各章相互独立,先利用一章的篇幅简单介绍多媒体数据的表示理论,后面每章介绍一种文件格式或者压缩标准,分别介绍了图像(第2章到第10章)、音频(第11章到第13章)、视频(第14章到第19章)和文本(第20章到第21章)四种多媒体数据。包括BMP, GIF, JPEG, JP2, PSD, WAVE, MIDI, MPG, AVI, MOV, PDF等文件格式。

本书在写作风格上尽量做到深入浅出,用最通俗的语言把问题说明白、讲清楚。而且针对每种文件格式,都给出一个简短的文件实例,一个字节一个字节地分析数据的作用和含义。编者长期从事数字水印方面的研究,经常要和多媒体数据打交道,以自己的实际工作经验知道,无论是通过书籍还是网络资料得到的文件格式描述,都存在格式描述和实际文件对不上的问题,必须反复推敲,仔细消化才能理解其中的意思,给工作和学习带来了很大不便。为了彻底改变这种状况,决定写一本有关这方面的书。

本书的写作灵感来自于编者在北京邮电大学信息安全中心以及北京邮电大学数字内容研究中心学习期间的工作,在此要真诚地感谢导师杨义先教授和钮心忻教授,没有他们以及他们领导的信息安全中心和数字内容研究中心,就不可能有对数字水印的深入研究。杨老师和钮老师高屋建瓴的学术眼光,兢兢业业的工作精神,时刻指引着编者走向前方。

编 者

2005年3月

目 录

第 1 章 多媒体数据表示理论	(1)
1.1 图像表示	(1)
1.1.1 图像的基本概念	(1)
1.1.2 位图和矢量表示	(3)
1.1.3 灰度图像和彩色图像	(4)
1.1.4 图像的颜色模型	(5)
1.2 声音表示理论	(8)
1.2.1 声音基础	(9)
1.2.2 声音的数字化	(9)
1.2.3 常见的声音格式	(11)
1.3 视频表示	(12)
1.4 常用编码和压缩方法	(13)
1.4.1 行程长度编码	(13)
1.4.2 霍夫曼编码	(13)
1.4.3 LZW 编码	(14)
1.4.4 算术压缩	(15)
1.4.5 有损压缩	(15)
第 2 章 BMP 图像文件格式	(16)
2.1 简介	(16)
2.2 BMP 文件格式	(16)
2.2.1 概要	(16)
2.2.2 BMP 中包含的域	(16)
2.3 域的详细描述	(20)
2.3.1 图像的宽度域和高度域	(20)
2.3.2 分辨率域	(20)
2.3.3 色深域	(20)
2.3.4 压缩域	(21)
2.3.5 调色板域	(23)
2.3.6 重要调色板域	(23)
2.4 举例	(23)
第 3 章 GIF 图像文件格式	(26)
3.1 简介	(26)
3.2 GIF 文件格式详解	(26)
3.2.1 文件头	(27)
3.2.2 逻辑屏幕描述块	(27)
3.2.3 全局调色板数据	(29)

3.2.4	图像描述块	(29)
3.2.5	局部调色板数据	(32)
3.2.6	图像压缩数据	(32)
3.2.7	图形控制扩展块	(32)
3.2.8	无格式文本扩展块	(34)
3.2.9	注释扩展块	(35)
3.2.10	应用程序扩展块	(36)
3.2.11	GIF 文件结束块	(37)
3.2.12	快速参考表	(37)
3.3	举例	(37)
第 4 章	JPEG 图像压缩标准	(40)
4.1	JPEG 压缩标准	(40)
4.2	JPEG 模型	(41)
4.2.1	编码器	(41)
4.2.2	解码器	(41)
4.2.3	交换格式	(42)
4.2.4	JPEG 用到的主要技术	(42)
4.2.5	有损压缩和无损压缩	(43)
4.3	色度空间	(44)
4.4	DCT 编码	(45)
4.5	量化	(48)
4.6	之字序列	(49)
4.7	熵编码	(51)
4.7.1	熵编码的中间符号表示	(51)
4.7.2	熵编码	(51)
4.8	其他的 JPEG 选项	(55)
4.8.1	递进模式	(55)
4.8.2	层次模式	(55)
4.8.3	无损模式	(55)
第 5 章	JPEG 图像文件格式	(56)
5.1	JPEG 简介	(56)
5.2	JPEG 文件格式	(56)
5.3	举例	(61)
第 6 章	JPEG 2000 图像压缩标准	(64)
6.1	原因和概念	(64)
6.2	JPEG 2000 压缩技术	(67)
6.3	预处理过程	(68)
6.3.1	图像分块	(68)
6.3.2	直流系数搬移	(70)

6.3.3	成分变换	(70)
6.4	核心过程	(71)
6.4.1	小波变换	(71)
6.4.2	量化	(73)
6.4.3	嵌入零树小波编码	(74)
6.5	比特流的构成	(75)
6.6	打包和层次	(76)
6.7	新标准的特点	(77)
6.7.1	ROI	(77)
6.7.2	可测性	(78)
6.7.3	容错能力	(79)
6.7.4	视觉频率加权	(80)
6.7.5	新文件格式	(80)
第 7 章	JP2 图像文件格式	(81)
7.1	文件格式组织	(81)
7.2	框的定义	(82)
7.3	JP2 框	(83)
7.3.1	JP2 签名框	(84)
7.3.2	轮廓框	(84)
7.3.3	JP2 头框	(84)
7.3.4	邻近码流框	(88)
7.3.5	IPR 框	(88)
7.3.6	XML 框 (多个)	(88)
7.3.7	UUID 框 (多个)	(88)
7.3.8	UUID INFO 框 (多个)	(88)
7.4	举例	(89)
第 8 章	TIFF 图像文件格式	(92)
8.1	文件结构	(92)
8.2	文件格式详解	(93)
8.2.1	文件头	(93)
8.2.2	图像文件目录 (IFD)	(93)
8.2.3	标记指针 (标记字段)	(94)
8.3	各类指针详解	(96)
8.3.1	基本标记指针	(96)
8.3.2	信息字段	(101)
8.3.3	传真标记字段	(102)
8.3.4	文档存储和检索字段	(102)
8.4	压缩	(103)
8.4.1	压缩代码 2 和 32771	(103)

8.4.2	压缩代码 5	(105)
8.4.3	压缩代码 32766	(105)
8.4.4	压缩代码 32809	(106)
8.4.5	压缩代码 32901	(106)
8.4.6	压缩代码 32900	(106)
8.5	TIFF 基本格式	(107)
8.5.1	黑白图像 (Bilevel Images)	(107)
8.5.2	灰度图像 (Grayscale Images)	(108)
8.5.3	调色板图像 (Palette-color Images)	(109)
8.5.4	RGB 彩色图像 (RGB Full Color Images)	(109)
8.6	CMYK 图像 (CMYK Images)	(110)
8.7	举例	(110)
第 9 章	PSD 图像文件格式	(112)
9.1	文件格式	(112)
9.2	文件头	(113)
9.3	色彩模式数据	(114)
9.4	图像资源	(114)
9.4.1	图像资源的结构	(114)
9.4.2	基本资源	(115)
9.5	层和遮罩信息	(116)
9.5.1	层	(117)
9.5.2	全局层遮罩	(121)
9.6	图像数据	(121)
9.7	举例	(122)
第 10 章	PCX 图像文件格式	(124)
10.1	PCX 文件结构	(124)
10.1.1	文件头	(124)
10.1.2	翻译数据	(126)
10.1.3	位图数据	(127)
10.1.4	调色板	(127)
10.2	图像压缩数据	(128)
10.3	PCX 压缩方法	(133)
10.4	举例	(134)
第 11 章	RIFF 文件格式	(136)
11.1	RIFF 文件格式	(136)
11.2	文件头	(137)
11.3	RIFF 文件块	(137)
11.4	列表	(138)
11.5	RIFF 框架 (RIFF Forms)	(138)

第 12 章 WAVE 音频文件格式	(140)
12.1 WAVE 文件概述	(140)
12.2 文件头	(142)
12.3 文件块	(143)
12.3.1 格式块	(143)
12.3.2 数据块	(144)
12.3.3 提示块	(145)
12.3.4 播放表块	(146)
12.3.5 关联数据列表	(147)
12.3.6 标签块	(147)
12.3.7 注释块	(148)
12.3.8 标签文本块	(149)
12.3.9 采样器块	(149)
12.3.10 采样循环结构体	(151)
12.3.11 工具块	(151)
12.4 微软 PCM 格式	(152)
12.5 举例	(153)
第 13 章 MIDI 音频文件格式	(156)
13.1 文件头块	(156)
13.2 文件轨迹块	(157)
13.3 MIDI 指令	(157)
13.3.1 MIDI 事件指令	(157)
13.3.2 系统信息指令	(162)
13.4 举例	(163)
第 14 章 MPEG-2 视频压缩标准	(166)
14.1 标准简介	(166)
14.2 MPEG 视频模型	(168)
14.2.1 MPEG 视频压缩	(169)
14.2.2 相关概念	(170)
14.2.3 运动补偿预测	(173)
14.2.4 MPEG 帧图像的类型	(174)
14.2.5 数据压缩	(176)
14.2.6 MPEG-2 视频特性	(179)
14.3 MPEG 音频模型	(181)
14.3.1 声音编码	(181)
14.3.2 MPEG 音频编码层	(182)
14.3.3 MPEG-2 音频特性	(185)
14.4 MPEG-2 比特流处理	(188)
14.4.1 传输流	(188)

14.4.2	传输流解码	(189)
14.4.3	程序流	(191)
14.4.4	程序流解码	(191)
14.5	MPEG-2 视频流处理	(192)
14.5.1	编码过程	(193)
14.5.2	MPEG-2 编码码流	(194)
14.5.3	解码过程	(195)
14.5.4	特点	(195)
第 15 章	MPG 文件格式	(197)
15.1	MPEG 音频层 1/2/3 的帧头	(197)
15.2	MP3 音频文件格式	(200)
15.2.1	MP3 文件结构	(201)
15.2.2	MP3 文件中的帧格式	(201)
15.2.3	ID3V1	(203)
15.2.4	ID3V2	(205)
15.3	MPEG 视频头信息	(208)
15.4	举例	(210)
第 16 章	MPEG-4 视频压缩标准	(211)
16.1	标准简介	(211)
16.2	基本功能与组成	(212)
16.2.1	MPEG-4 标准提供的功能	(212)
16.2.2	AV 对象 (AVO)	(213)
16.2.3	系统构造	(214)
16.2.4	MPEG-4 的结构和句法	(215)
16.2.5	传输多媒体集成框架	(217)
16.2.6	递交多媒体集成框架应用界面	(218)
16.2.7	数据平面	(219)
16.2.8	缓冲区管理和实时识别	(220)
16.3	MPEG-4 音频	(220)
16.3.1	音频编码	(220)
16.3.2	参数编码器	(221)
16.3.3	码激励线性预测 (CELP) 编解码器	(223)
16.4	MPEG-4 视频	(224)
16.4.1	概述	(224)
16.4.2	视频对象	(224)
16.4.3	多路复接和传送	(226)
16.4.4	视频图像对象平面	(227)
16.4.5	视频编码方案	(228)
16.4.6	形状编码	(230)

16.4.7	运动信息编码	(230)
16.4.8	纹理编码	(231)
16.4.9	Sprite 编码	(231)
16.4.10	视频分辨率可变编码	(232)
16.4.11	视频解码	(233)
第 17 章	MP4 视频文件格式	(235)
17.1	描述结构	(236)
17.1.1	文件结构	(237)
17.1.2	对象结构	(237)
17.1.3	元数据和媒体数据结构	(237)
17.2	框定义	(237)
17.2.1	文件类型框	(239)
17.2.2	视频框	(240)
17.2.3	媒体数据框	(240)
17.2.4	视频头框	(241)
17.2.5	轨迹框	(242)
17.2.6	媒体框	(243)
17.2.7	视频媒体头框	(243)
17.2.8	声音媒体头框	(243)
17.2.9	块偏移框	(244)
17.2.10	自由空间框	(245)
17.2.11	用户数据框	(245)
17.3	扩展	(246)
17.3.1	对象	(246)
17.3.2	存储格式	(246)
第 18 章	AVI 视频文件格式	(247)
18.1	AVI 文件格式	(247)
18.2	文件头	(249)
18.3	“hdrl”列表	(249)
18.3.1	“avih”块	(249)
18.3.2	“strl”子列表	(250)
18.4	“movi”列表	(251)
18.5	可选索引块	(252)
18.6	举例	(252)
第 19 章	MOV 视频文件格式	(254)
19.1	原子	(254)
19.2	文件格式	(256)
19.3	视频原子	(257)
19.3.1	视频头原子	(257)

19.3.2 调色板原子	(258)
19.4 轨迹原子	(259)
19.4.1 轨迹头原子	(259)
19.4.2 剪切原子	(260)
19.5 媒体原子和媒体信息原子	(260)
19.5.1 媒体原子	(260)
19.5.2 媒体头原子	(261)
19.5.3 媒体信息原子	(261)
19.6 媒体数据原子类型	(263)
19.7 举例	(263)
第 20 章 PDF 文本文件格式	(265)
20.1 PDF 文件格式简介	(265)
20.2 对象	(266)
20.2.1 布尔对象	(266)
20.2.2 数对象	(266)
20.2.3 字符串和文本对象	(267)
20.2.4 名称	(267)
20.2.5 矩阵	(268)
20.2.6 字典	(268)
20.2.7 流	(269)
20.2.8 空对象	(270)
20.2.9 迂回对象	(270)
20.3 文件结构	(270)
20.3.1 文件头	(271)
20.3.2 文件体	(271)
20.3.3 交叉参考表	(271)
20.3.4 文件索引	(273)
20.3.5 增量更新	(274)
20.3.6 加密	(274)
20.4 文档结构	(275)
20.4.1 目录	(276)
20.4.2 页树	(277)
20.4.3 页对象	(277)
20.4.4 大纲树	(278)
20.4.5 文章线索	(279)
20.4.6 目的文件	(281)
20.4.7 形式	(281)
20.5 常用数据结构	(281)
20.5.1 矩形 (Rectangle)	(281)
20.5.2 日期 (Date)	(281)

20.5.3 数树 (Number Tree)	(281)
20.6 线性 PDF	(282)
20.6.1 线性 PDF 文档结构	(282)
20.6.2 线索表	(283)
20.7 举例	(284)
第 21 章 RTF 文本文件格式	(286)
21.1 语法	(286)
21.1.1 RTF 语法	(286)
21.1.2 外部语法	(287)
21.2 RTF 文件	(288)
21.3 文件头	(288)
21.3.1 RTF 版本	(288)
21.3.2 字符集	(288)
21.3.3 默认字体	(288)
21.3.4 字体表	(289)
21.3.5 颜色表	(289)
21.4 文档域	(289)
21.4.1 信息组	(290)
21.4.2 文档格式属性	(290)
21.4.3 节文本	(291)
21.4.4 段文本	(292)
21.4.5 字符文本	(292)
21.4.6 文档变量	(292)
21.4.7 图片	(292)
21.4.8 书签	(293)
21.5 举例	(293)
参考文献	(295)

第1章 多媒体数据表示理论

1.1 图像表示

图像是一类重要的多媒体数据,人们获取的信息的 70%来自视觉系统,也就是说,在所有多媒体数据中,图像提供的信息量最多。但是,图像数字化之后的数据量非常大,在因特网上传输很费时间,在存储介质上存储时需要较多的存储区,因此就必须要对图像数据进行压缩。压缩的目的就是在不降低图像视觉质量的前提下,满足存储容量和传输带宽的要求,寻找更有效的方法,用比较少的数据量表示原始图像。

图像数据压缩主要是根据下面两个基本事实来实现的:一个是图像数据中有许多重复的数据,使用数学方法来表示这些重复数据可以减少数据量;另一个事实是人眼视觉系统对图像细节和颜色的辨认有一个范围,把超过范围的部分去掉,也能达到压缩数据的目的,而且这样人眼感觉不到图像的变化,不会影响图像的视觉质量。利用前一个事实的压缩技术就是无损压缩技术,利用后一个事实的压缩技术就是有损压缩技术。实际的图像压缩是综合使用各种有损和无损压缩技术来实现的。

为了了解人的视觉系统如何认识彩色,计算机系统中如何表示彩色图像,从而更有效地减少数据量,本部分将首先介绍表示数字彩色图像所需要的基本知识。

1.1.1 图像的基本概念

1. 视觉系统的特点

可见光是波长在 380~780nm 之间的电磁波,不同波长表现出不同的颜色。我们看到的大多数光不是一种波长的电磁波,而是由许多不同波长的电磁波混合而成的。研究表明,人的视网膜有对红、绿、蓝三种颜色敏感程度不同的三种锥体细胞,另外还有一种在光功率极低的条件下(也就是光线非常暗的情况下)才起作用的杆状体细胞,因此颜色只存在于眼睛和大脑。颜色是视觉系统对可见光的感知结果。

眼睛本质上是一个照相机,人的视网膜通过神经元来感知外部世界的颜色,每个神经元或者是一个对颜色敏感的锥体,或者是一个对颜色不敏感的杆状体。而红、绿、蓝三种锥体细胞在光线充足的正常情况下起主要作用。这就意味着,在计算机中人们可以只使用红、绿、蓝三种光点来表示图像的颜色,使数字图像看上去与自然图像没有什么差别,红、绿、蓝三种光点分别对应红、绿、蓝三种锥体细胞。

数字图像中的任何一种颜色都可以由红、绿、蓝(下面分别简称为 R 、 G 、 B)这 3 种原色值的混合来确定,每个原色都有一个取值范围,它们构成一个 3 维的 RGB 矢量空间。当三原色都取最大值时表现为白色,当三原色都取最小值时表现为黑色。

2. 分辨率

平时经常遇到两种分辨率:显示分辨率和图像分辨率。

显示分辨率是指显示器上能够显示出的像素数目,是显示器的一个衡量指标。例如,显示

分辨率为 640×480 表示显示屏分成 480 行, 每行显示 640 个像素, 整个显示屏就含有 $640 \times 480 = 307\,200$ 个显像点。屏幕能够显示的像素越多, 说明显示设备的分辨率越高, 显示的图像质量也就越高。

图像分辨率是指组成一幅图像的像素密度的度量方法。对同样大小的一幅图, 如果组成该图的图像像素数目越多, 则说明图像的分辨率越高, 看起来就越逼真。相反, 图像显得越粗糙。

在用扫描仪扫描彩色图像时, 通常要指定图像的分辨率, 用每英寸 (1 英寸=2.54cm) 多少个像素点表示, 英文简称为 dpi (dots per inch)。如果用 300dpi 的分辨率来扫描一幅 8 英寸 $\times$ 10 英寸的图像, 就得到一幅 2400×3000 像素的图像。分辨率越高, 像素点就越多。

图像分辨率与显示分辨率是两个不同的概念。图像分辨率是确定组成一幅图像的像素数目, 而显示分辨率是确定显示图像的区域大小。如果显示屏的分辨率为 640×480 , 那么一幅 320×240 的图像只占显示屏的 $1/4$; 相反, 2400×3000 的图像在这个显示屏上就不能显示一个完整的画面。

3. 像素深度

像素深度又叫色深, 是指存储每个像素所用的位数, 也就是用计算机中的多少个位来表示一种颜色。它决定彩色图像的每个像素可能的颜色数, 或者确定灰度图像的每个像素可能的灰度级数。

例如, 一幅彩色图像的每个像素用 R 、 G 、 B 三个分量表示, 若每个分量占用 8 位, 那么一个像素共用 24 位表示, 这就是说像素的深度为 24, 每个像素可以有 $2^{24} = 16\,777\,216$ 种颜色。表示一个像素的位数越多, 它能表达的颜色数目就越多, 它的深度就越深。

虽然像素深度可以很深, 但各种显示器的颜色深度受到硬件的制约却不能无限大, 而且人眼分辨颜色的能力也只是在一定范围之内, 超出了这个范围, 即使像素深度再加深, 人眼也感觉不出来。因此, 如果像素深度太大, 除了增加存储和传输的负担外, 没有什么用处。不过, 如果像素深度太浅, 那就会影响到图像的质量, 图像看起来让人觉得很粗糙, 也很不自然。计算机中常用的像素深度有 1 位、8 位、16 位、24 位等, 具体选用哪一种色深, 与图像的应用领域有关系, 后面将介绍。

在用二进制数表示彩色图像的像素时, 除 R 、 G 、 B 分量用固定位数表示外, 往往还增加 1 位或几位作为属性位。例如, RGB 图像分别用 5 位来表示时 (简记为 RGB 5:5:5), 一共占用 $5 \times 3 = 15$ 个位, 可以用 2 个字节共 16 位表示, 这样 R 、 G 、 B 分量各占 5 位, 剩下一位就可以作为属性位。在这种情况下, 我们说像素深度为 16 位。当然有的文献将这种情况的像素深度定义为 15 位, 这也是可以的, 本书采用第一种定义方式。

属性位用来指定该像素应具有的性质, 它在不同情况下的具体含义可以不同。

例如, 在 RGB 5:5:5 的情况下, 用 2 个字节共 16 位的最高位作为属性位, 并把它称为透明位, 记为 T 。 T 的含义理解为: 假如显示屏上已经有一幅图 (称做原图) 存在, 当原图或者原图的一部分要重叠在当前图的上面时, T 位就用来控制原图是否能看得见。如果 $T=1$, 表示在该像素点上的原图完全看不见, 也就是用当前图覆盖原图; 如果 $T=0$, 表示在该像素点上的原图能完全看见, 也就是用原图覆盖当前图。

在用 32 位表示一个像素时, 若 R 、 G 、 B 分别用 8 位表示, 剩下的 8 位常称为 α 通道 (Alpha Channel) 位, 或称为覆盖位、中断位、属性位, 用 A 表示。它的用法可用一个预乘 α 通道的例子说明。假如一个像素 (A, R, G, B) 的四个分量都用归一化的数值表示, 那么 (1,1,0,0) 表示的就是红色, 显示时不进行透明处理, 完全显示。当像素为 (0.5,1,0,0) 时, 还是表示红色,

但是需要用 α 通道(A)预乘RGB的值,这样颜色就变成(0.5,0,0)。这表示原来该像素显示的红色的强度为1,而现在显示的红色的强度降了一半。

用这种办法定义一个像素的属性在实际中很有用。例如在一幅彩色图像上叠加文字说明,而又不想让文字完全把图像覆盖掉,就可以用这种办法来定义文字像素。在图像产品生产中,也往往把数字电视图像和计算机生产的图像混合在一起,这种技术称为视图混合技术,它也采用 α 通道。

4. 真彩色和伪彩色

(1) 真彩色 (true color)

真彩色是指在组成一幅彩色图像的每个像素值中,有R、G、B三个基色分量,每个基色分量直接决定显示设备的基色强度,这样产生的彩色称为真彩色。例如用RGB 5:5:5表示的彩色图像,R、G、B各用5位,用R、G、B分量大小的值直接确定三个基色的强度,这样得到的彩色是真实的原图彩色。

如果用RGB 8:8:8方式表示一幅彩色图像,就是R、G、B都用8比特来表示,每个基色分量占一个字节,共3个字节,每个像素的颜色就是由这3个字节中的数值直接决定,它所能表示的颜色数目多达 $16\ 777\ 216$ 种,此种图像也常称为全彩色图像。用3个字节表示的真彩色图像所需要的存储空间很大,而人的眼睛是很难分辨出这么多种颜色的,因此在许多场合往往用RGB 5:5:5来表示,每个彩色分量占5个比特,再加1比特显示属性控制位共2个字节,生成的真颜色数目为 $2^{15}=32\text{K}$ 。

(2) 伪彩色 (pseudo color)

伪彩色图像是指每个像素的颜色不是由每个基色分量的数值直接决定,而是把像素值当做彩色查找表(Color Look-Up Table, CLUT)的表项入口地址,去查找一个显示图像时使用的R、G、B强度值,用查找出的R、G、B强度值产生的彩色称为伪彩色。这样生成的图像通常叫做调色板图像,彩色查找表又叫做调色板。在调色板图像中有两个矩阵,一个是图像矩阵,其中记录的是每个像素对应的CLUT表项入口地址,另一个是调色板矩阵,矩阵中的每个元素代表一种颜色。

彩色查找表CLUT是一个事先做好的表,表项入口地址也称为索引号。例如,16种颜色的查找表,0号索引对应黑色,...,15号索引对应白色,这种情况下调色板矩阵只有16个元素。

图1-1所示就是一个调色板图像,(a)是图像,(b)是它的调色板。

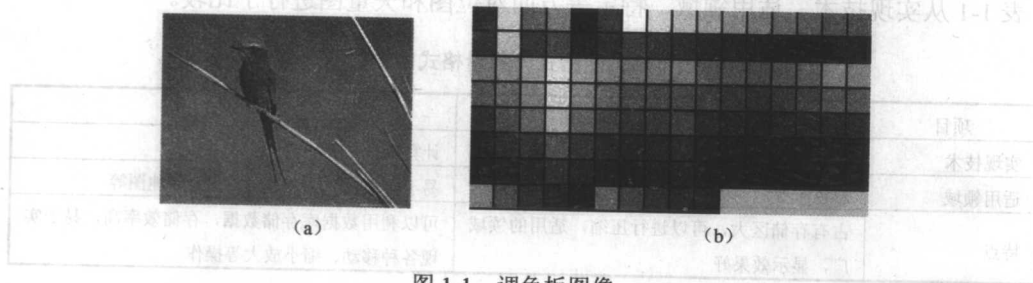


图1-1 调色板图像

1.1.2 位图和矢量表示

在计算机中,表示图像有两种常用的方法:一种叫做矢量图(vector based image)法;另一种叫位图(bit mapped image)法。

1. 矢量图

矢量图是用一系列计算机指令来表示一幅图，如画点、画线、画曲线、画圆、画矩形等分别对应不同的画图指令。这种方法实际上是用数学方法来描述一幅图，先将图像分割成许许多多的简单部分，如点、线、圆、矩形等，每一部分用一个数学表达式来表达，然后按照矢量图的语法编程，用相应的命令来表达。显示图像的过程可以看做按照绘图指令画图的过程。

矢量图有许多优点。例如，当需要管理每一小块图像时，矢量图法非常有效；目标图像的移动、缩小放大、旋转、拷贝、属性的改变（如线条变宽变细、颜色的改变）也很容易做到；相同的或类似的图可以把它们当做图的构造块，并把它们存到图库中，这样不仅可以加速画的生成，而且可以减小矢量图文件的大小。

缺点是，当图变得很复杂时，计算机就要花费很长的时间去执行绘图指令。此外，对于一幅复杂的彩色照片（如一幅真实世界的彩照），很难用数学来描述，因而就很难用矢量法表示。

2. 位图

位图法是把一幅彩色图分成许许多多的像素，每个像素用若干个二进制位来指定该像素的颜色、亮度和属性。因此，一幅图是由许许多多描述每个像素的数据组成的，这些数据通常称为图像数据，而这些数据作为一个文件来存储，这种文件又称为图像文件。

位图文件占据的存储器空间比较大。影响位图文件大小的因素主要有两个：图像分辨率和像素深度。分辨率越高，组成一幅图的像素就越多，图像文件也越大；像素深度越深，表达单个像素的颜色和亮度的位数就越多，图像文件也要变大。而矢量图文件的大小则主要取决于图的复杂程度。

3. 位图格式与矢量格式的比较

矢量图与位图相比，显示位图文件比显示矢量图文件要快；矢量图侧重于“绘制”、去“创造”，而位图偏重于“获取”、去“复制”；矢量图和位图之间可以用软件进行转换，由矢量图转换成位图采用光栅化（rasterizing）技术，这种转换也相对容易；由位图转换成矢量图用跟踪（tracing）技术，这种技术在理论上说是容易，但在实际中很难实现，对复杂的彩色图像尤其如此。

表 1-1 从实现技术、适用领域、特定等方面对位图和矢量图进行了比较。

表 1-1 位图格式与矢量格式的比较

项目	位图	矢量图
实现技术	像素数据表示图像	计算机指令表示图像
适用领域	彩色图像，各种数字设备采集的图像数据	易于用数学方法表示的图像，如地图等
特点	占有存储区大，可以进行压缩，适用的领域广，显示效果好	可以利用数据库存储数据，存储效率高，易于实现各种移动、缩小放大等操作

1.1.3 灰度图像和彩色图像

灰度图像是去掉颜色信息的图像，仅保留图像的亮度信息，这种图像只用黑、白以及介于黑白之间的各种灰色表示，不同深浅的灰度表示的就是图像的亮度。彩色图像是保留颜色