

数字图像压缩技术 实用教程

王新年 张涛 编著



Digital Image Compression Technique

数字图像压缩技术 实用教程

王新年 张涛 编著



机械工业出版社
China Machine Press

本书主要介绍图像信号的压缩编码原理、静态图像编码方法、视频图像编码方法,以及数字图像与视频压缩编码的国际标准。主要内容包括量化、熵编码、预测编码、变换编码、矢量量化编码、运动补偿等编码技术以及 JPEG、JPEG-2000、MPEG-x 和 H.26x 等现行国际编码标准及其应用案例。对于编码标准,重点介绍其中采用的编码方法和应用,技术细节不做重点讲述。

本书以通俗易懂的方式介绍编码技术,注重介绍各项技术产生的背景和思路,并辅以 Matlab 和 C++ 源码、课后习题、课程设计和应用案例,启发读者了解各项技术的基本原理、应用方法和适用场合。

本书还设计了一些图像压缩实验,目的在于让读者熟悉视频编码的整个过程,并掌握其中的关键算法。图像压缩实验与各章课后课程设计互相补充,相得益彰,对加深图像编码理解和做与图像压缩相关的实际应用系统大有裨益。

本书可作为本专科以及高职院校相关课程的教材,也可供从事数字图像压缩研究和开发的工程技术人员参考。

版权所有,侵权必究。

本书法律顾问 北京市展达律师事务所

图书在版编目(CIP)数据

数字图像压缩技术实用教程/王新年,张涛编著. —北京:机械工业出版社,2009.8
(数字媒体专业规划教材)

ISBN 978-7-111-27808-5

I. 数… II. ①王… ②张… III. 数字图像处理—数据压缩—高等学校—教材
IV. TN919.8

中国版本图书馆 CIP 数据核字(2009)第 125160 号

机械工业出版社(北京市西城区百万庄大街 22 号 邮政编码 100037)

责任编辑:李俊竹

北京市荣盛彩色印刷有限公司印刷

2009 年 10 月第 1 版第 1 次印刷

184mm×260mm·12.5 印张

标准书号:ISBN 978-7-111-27808-5

定价:25.00 元

凡购本书,如有倒页、脱页、缺页,由本社发行部调换
本社购书热线:(010) 68326294

—○ 前 言 ○—

20 世纪 80 年代以来, 图像压缩技术随着多媒体和通信技术的发展, 在科学研究、工业生产、医疗卫生、教育、娱乐等方面得到了广泛的应用, 如高清晰电视、数字电影、远程医疗、视频监控、视频电话、网络电视、DVD、手机电视以及刚开通的 3G 手机等。

图像压缩技术的应用已经成为一个产业, 社会对这方面的人才需求越来越大。各高职、大专院校为了适应这种需求, 相继开设了数字图像压缩相关课程, 但由于读者专业背景和基础的不同, 对教材的要求也不尽相同。

本书旨在以简明易懂的方式介绍图像压缩技术, 并结合算法程序源码、习题和课程设计以及实际案例, 介绍各项技术产生的背景和思路, 启发读者了解各项技术的基本原理、应用方法和适用场合。力求使数学和信号处理基础较弱的学生在较短的学时内, 掌握图像编码技术基本思想, 并可根据需要自行研究相关技术应用细节。基于同样的目的, 本书没有介绍对数学基础要求相对较高的小波变换、分形等方法。

本书包括如下章节:

第 1 章介绍图像压缩的必要性、可行性、发展史以及图像压缩应用系统, 回答了为什么要进行图像压缩、为什么能够进行压缩和怎么用图像压缩的问题。

第 2 章介绍图像压缩的基本原理, 为第 3 章图像压缩技术做准备。

第 3 章介绍图像压缩的基本方法, 回答了怎么进行图像压缩的问题。

第 4 章介绍图像压缩国际标准, 回答了图像压缩技术在实际应用中所采用的技术问题。

第 5 章介绍图像压缩应用案例, 回答了图像压缩技术在我们日常生活中有哪些用处以及怎么用的问题。

第 6 章介绍图像压缩实验, 实验内容以实现一个基本的静态图像编码系统和一个视频编码系统为主线, 目的在于让读者熟悉视频编码的整个过程, 并掌握其中的关键算法。

第 7 章介绍典型编解码算法的 C++ 实现。

为了便于读者理解和验证所讲述的各个算法, 每个算法都配以 Matlab 和 C++ 源码, 这些源码尽量按照算法描述的格式编写, 并没有采用相应工具箱函数, 也没有刻意进行优化。目的在于: 1) 希望读者对算法先理解再优化; 2) 对于熟悉 Matlab 语言的读者, 这些源码可直接上机调试运行, 在此基础上优化; 对于不熟悉 Matlab 语言的读者, 可将这些源码当做描述算法的伪码, 并很容易改成 C、C++ 语言等 (在编写源码的过程, 作者尽量以类似 C 语言的形式给出, 目的也在于此)。

本书各章的课后习题与课程设计不是课本内容的简单重复, 而是从实际应用中提炼出的问

题，是课本内容必要的补充。各章课后课程设计与实验互相补充，相得益彰，对加深图像编码理解和做与图像压缩相关的实际应用系统大有裨益。题号前带有*的习题为课程设计。

图像压缩是一门实践性很强的课程，实验与课程设计都要求上机编程实现，这样才能够加深理解。对于编程语言，建议在算法验证阶段最好使用 Matlab 或 Mathematics，待算法验证可行后，再用 C++ 之类的语言转成应用程序，发布出去。这样才可能把更多精力投入算法设计，而不是编程上。

本书是作者在所教授的“数字图像处理”和“图像压缩与传输”两门课程讲义的基础上总结出来的。在本书编写过程中，参考和引用了大量专家和学者的研究成果、著作和论文，以及麻省理工、斯坦福、卡内基梅隆等大学相关课程的课件，还有一部分内容来源于互联网和听课学生的一些建议。在此，向所有文献著作者和资料提供者表示深切的谢意和敬意。

在本书编写过程中，作者始终希望以浅显的语言和较少的数学知识介绍各种方法的产生背景，从特例逐步推广到一般，启发读者理解各项编码技术的基本原理，但有时期望与实际效果有一定差距，在此希望各位读者见谅。

本书的编写对作者来说既是挑战又是学习和提高的机会。由于作者的视野和水平有限，编写时间仓促，书中难免存在错误和不足之处。在此真诚希望各位前辈、专家和读者给予批评指正，并希望通过交流使本书内容得以完善。

梁德群教授、齐国清教授、王海姣老师还有徐鹏在本书编写过程中给予很多帮助，在此表示感谢。

在本书立项及出版过程中，得到了机械工业出版社华章分社的编辑们的热情鼓励、支持和帮助，在此表示最真诚的感谢。

—○ 教 学 建 议 ○—

教学内容	学习要点及教学要求	课时安排	
		信息类本科专业	非信息类
第1章 绪论	●掌握数据、数据量、存储空间、带宽以及数据与信息之间的关系。 ○理解数据压缩过程、图像压缩的必要性。 ●掌握图像压缩的可行性。 ○了解图像压缩的发展史。 ●掌握图像通信系统的模型。 ○了解图像压缩的应用案例。	2	2
第2章 图像压缩的基本原理	●掌握信源、编码器、信息传输模型、自信息和信息熵基本概念。 ●掌握自信息和信息熵的计算方法。 ○了解基本编码定理。 ○理解人眼的视觉特性。 ●掌握图像离散余弦变换及其性质。 ●掌握量化及量化误差基本概念和目的。 ○理解量化分类。 ●掌握均匀量化和非均匀量化。 ○理解矢量量化。 ●掌握压缩率、视频码率以及客观保真度准则。 ○了解主观评价准则。 ○了解算法复杂度。	6	8
第3章 图像压缩方法	●掌握图像编解码过程。 ○理解图像压缩方法分类。 ●掌握熵编码（行程编码、赫夫曼编码、算术编码、词典编码）方法基本原理、优缺点及适用条件。 ●掌握预测编码、变换编码方法的基本原理、优缺点及适用条件。 ○理解矢量编码方法。 ●掌握空间下采样和色度抽样的基本原理。 ○理解时间采样。	12	14

(续)

教学内容	学习要点及教学要求	课时安排	
		信息类本科专业	非信息类
第4章 图像压缩国际标准	○了解图像压缩国际标准发展史。 ○理解主流静止图像和视频图像编码标准。 ●掌握 JPEG 基本系统。 ○了解 JPEG2000 的特性。 ●掌握视频编码的基本技术。 ○理解 H.26x 和 MPEG-x 系列国际标准优缺点及所采用的技术。	4	4
第5章 图像压缩应用案例	○理解本地和远程视频监控系统主要实现的功能及其区别。 ○理解本地和远程视频监控系统主要硬件组成和软件功能。 ○理解手机彩信的收发过程。 ○理解数字电视系统的基本原理。 ●能够根据用户图像传输方面需求,提供相应技术支持和解决方案。	8	8
第6章 图像压缩实验	○实验一 图像压缩实验基础知识(2学时)。 ●实验二 图像行程编码和赫夫曼编码(4学时)。 ●实验三 图像二维 DCT 变换和量化(4学时)。 ●实验四 图像 DPCM 编码(2学时)。 ○实验五 运动估计(2学时)。 ○实验六 基本 JPEG 算法(4学时)。 ○实验七 基于运动补偿的视频编码(4学时)。	22	0
第7章 典型编解码算法的 C++ 实现	本章内容是前六章所讲算法的 C++ 语言实现,主要供学习者参考。讲授内容及课时由教师根据学生对 C++ 语言的熟悉程度以及学生实验所采用的编程语言而定。	不做要求	0
	教学总学时建议	理论: 32 实验: 22	理论: 36 实验: 0

说明:

- ① 教学要求分三个层次: ●掌握、○理解、○了解。
- ② 本书为信息类本专科专业课程教材,计划总学时为 54 学时,其中理论授课为 32 学时,实验为 22 学时。若计划学时数为 72 学时,则第 2 章和第 3 章分别增加 4 和 8 学时,实验一、实验五和实验六分别增加 2 学时;若计划学时数为 36 学时,则将第 2 章、第 3 章、第 5 章分别减少 2、2、4 学时,并且实验只做到实验四即可。不同专业根据不同的教学要求和计划教学学时数,可酌情对讲授内容进行适当取舍。
- ③ 非信息类本专科专业使用本教材可适当降低教学要求。实验环节可省略。
- ④ 本教材理论授课学时数包含习题课、课堂讨论等必要的课内教学环节。

—○ 目 录 ○—

前言

教学建议

第1章 绪论	1
1.1 图像压缩的必要性	4
1.2 图像压缩的可行性	4
1.3 图像压缩的发展史	8
1.4 图像压缩应用及案例	9
1.5 习题与课程设计	11
第2章 图像压缩的基本原理	13
2.1 信息论基础	13
2.1.1 信源、编码器	13
2.1.2 信息传输模型	14
2.1.3 信息的度量	14
2.1.4 基本编码定理	17
2.2 视觉特性	17
2.3 图像变换	21
2.3.1 一维离散余弦变换	23
2.3.2 二维离散余弦变换	25
2.3.3 二维离散余弦变换的性质	26
2.4 量化和编码	28
2.4.1 量化和编码定义	28
2.4.2 量化误差	30
2.4.3 量化分类	31
2.5 图像压缩算法质量评价	36
2.5.1 压缩比	36
2.5.2 视频码率	37
2.5.3 图像保真度	37
2.5.4 算法复杂度	39

2.6 习题与课程设计	40
第3章 图像压缩方法	43
3.1 引言	43
3.2 熵编码	45
3.2.1 行程编码	45
3.2.2 赫夫曼编码	49
3.2.3 算术编码	57
3.2.4 词典编码	62
3.3 预测编码	70
3.4 变换编码	83
3.5 矢量编码	88
3.6 基于下采样与插值的图像压缩	93
3.6.1 空间下采样	93
3.6.2 色度抽样	94
3.6.3 时间采样	97
3.7 各种编码方法总结	97
3.8 习题与课程设计	98
第4章 图像压缩国际标准	101
4.1 图像压缩国际标准发展史	101
4.2 静止图像编码标准	104
4.2.1 JPEG 标准	105
4.2.2 JPEG 2000 标准	108
4.3 视频图像编码标准	109
4.3.1 视频编码技术	111
4.3.2 视频编码标准概述	116
4.4 习题与课程设计	122
第5章 图像压缩应用案例	123
5.1 本地数字视频监控系统	123

5.2 远程视频监控系统	129	实验六 基本 JPEG 算法	145
5.3 手机彩信	133	实验七 基于运动补偿的视频 编码	146
5.4 数字电视	137		
5.5 习题与课程设计	141		
第 6 章 图像压缩实验	142	第 7 章 典型编解码算法的 C++ 实现	147
实验一 图像压缩实验基础知识	142	7.1 数据结构	147
实验二 图像行程编码和赫夫曼 编码	144	7.1.1 信号类声明与定义	147
实验三 图像二维 DCT 变换和 量化	144	7.1.2 图像读取与显示类声明 与定义	153
实验四 图像 DPCM 编码	145	7.2 编解码算法函数	158
实验五 运动估计	145	7.3 编解码算法函数实现	160
		参考文献	188

第1章 绪 论

随着信息技术的发展和社会的进步,人类对信息的要求越来越丰富,人们希望无论何时何地都能够方便、快捷、灵活地通过文字、语音、图像以及视频等多媒体进行通信。在数字通信中,多媒体需要大量的数据来表征,而传输如此大量的数据需要很大的网络带宽,这成为多媒体通信业务顺利开展的巨大障碍。如何在现有网络带宽的情况下,满足人们日益增长的需求呢,这就需要数据压缩。

压缩顾名思义就是缩小体积,例如,压缩饼干是将饼干中无用的水分挤掉,从而减少物体所占的空间。在日常生活中,我们经常用“人大”两个字来代替“中国人民大学”。用“头悬梁锥刺股”来描述一个人认真苦读,这6个字的成语可以顶替几百字来描述一个人勤奋好学、认真苦读的情景。这些都是一种压缩形式。在信息科学中,由于所有的信息都以二进制形式存放,空间以比特或位来度量,数据压缩的目的就是减少表达信息所需要的数据量(即位数)。

为便于理解,在具体讲述数据压缩之前,先介绍以下常用的术语:

信息:半个多世纪以来,信息论及信息科学得到了突飞猛进的发展,信息的概念日益深入人心,已经成为人们使用频率最高的词语之一,但是在辉煌的背后却始终有一个难题困扰着信息科学,这个难题就是信息的定义。因为在不同的场合,它被赋予了不同的含义。在本书中,我们采用的定义是:信息是能够接收、传递和理解的消息。

信源:是指消息的来源。如果信源输出的消息是以取值离散的符号形式出现,其不同符号数可以是有限个或可列无限个,称其为离散信源。如果信源输入的消息的取值是连续的,则称其为连续信源。

数据:是使用约定俗成的符号对客观事物的数量、属性、位置及其相互关系进行的抽象表示,以便于在相关领域进行保存、传递和处理。在数字计算机中,数据就是0和1的组合序列,例如用“01000000”表示大些字母“A”,而用“01100000”表示小写字母“a”。

数据量:表达每个数据所需符号个数的总和,数学描述为:

$$Q = \sum_{i=1}^N l(d_i) \quad (1-1)$$

其中 $l(d_i)$ 为表达第 i 个数据所需要的符号个数, N 为数据个数。

在计算机科学中,数据用二进制数0和1表示,每个0或1称为一个位或比特(bit),因此数据量的单位是比特(bit,缩写为b),可以推出上式中 $l(d_i)$ 的另一种表示形式:

$$l(d_i) = \log_2 d_i \quad (1-2)$$

存储空间：在计算机科学中，存储空间是指存储数据所需要的字节数，每个字节由 8 位组成，字节 (Byte) 缩写为 B。存储空间和数据量数值上相等。

带宽：通常指信号所占据的频率区间的宽度。在用来描述信道时，带宽是指能够有效通过该信道的信号的最大频带宽度。对于模拟信号而言，带宽又称为频宽，以赫兹 (Hz) 为单位。在数字通信中，带宽指的是比特率 (bit rate)，即每秒钟传输数据的位数，单位是 bps (bits per second)，有时也指每秒传送的字节数，单位是 Bps (Bytes per second)。数据量与带宽的关系：数据量 = 带宽 × 传输数据所需要的时间。

数据和信息的关系：数据是信息的表达方式，是信息的载体。不同的数据集可以表示相同的信息，例如“我饿了”和“I am hungry”，不同的数据所承载的信息却是一样的。

由于不同的数据集可以表示相同的信息，我们就希望传输或存储数据量少的数据集，这样可以在带宽或存储空间有限的情况下，传输或存储更多的信息。如何获得信息的较少数据量载体呢？这就是数据压缩的任务。

那么什么是数据压缩呢？数据压缩 (data compress) 就是以较少位数的数据来表示原始数据信息的过程，如图 1-1 所示。

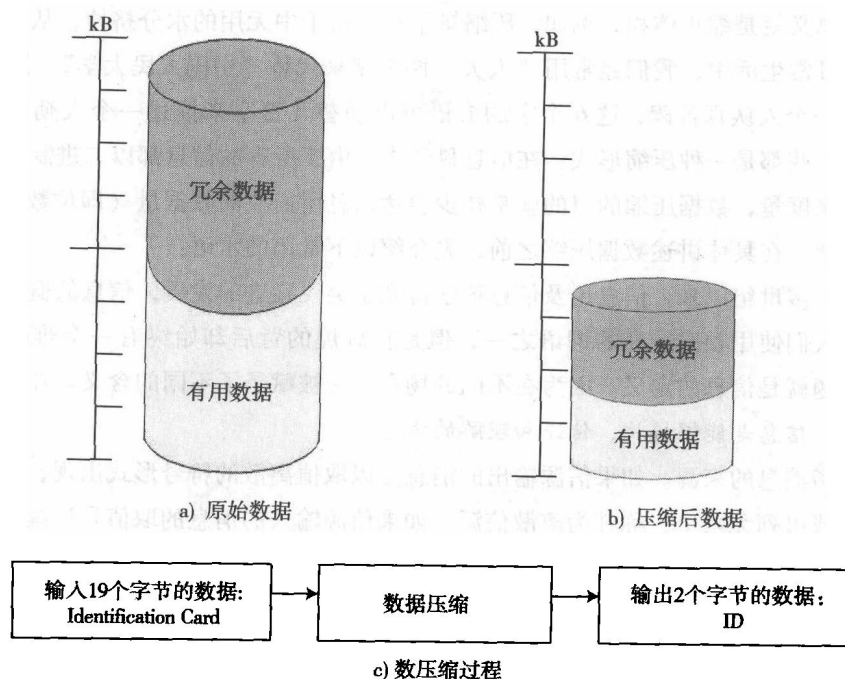


图 1-1 数据压缩演示

从压缩后的数据恢复出原始数据的过程，称为解压缩 (decompress)。在信息科学中，数据压缩有时又称为编码 (encode)；解压缩称为解码 (decode)。执行这些功能的模块分别称为压缩器 (compressor) 或者编码器 (encoder)，解压缩器 (decompressor) 或者解码器 (decoder)。将编码器和解码器通称为 Codec。

数据压缩技术就是研究如何对数据进行压缩的技术，通常情况下，数据压缩技术包括两部分：压缩和解压缩，如图 1-2 所示。

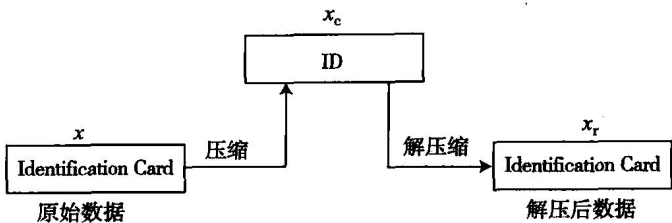


图 1-2 数据压缩和解压缩

在图 1-2 中，输入数据 x 经过数据压缩算法进行压缩，形成 x 的较少位数表示 x_c ， x_c 可以用于存储或传输。当需要了解 x 的信息时，需要对 x_c 解压缩，形成数据 x_r ，若 x 与 x_r 完全相等，则这种数据压缩方法称为无损压缩，否则称为有损压缩。

数据可以是文字、语音、图像、视频、图形、动画以及各种采集或生成的数字序列。各种类型数据有其共性也有其特性，因此所采用的压缩方法不尽相同，本书主要介绍图像和视频的压缩方法。

在实际应用中，数字图像在数学上可以描述为一个二维矩阵，如图 1-3 所示。

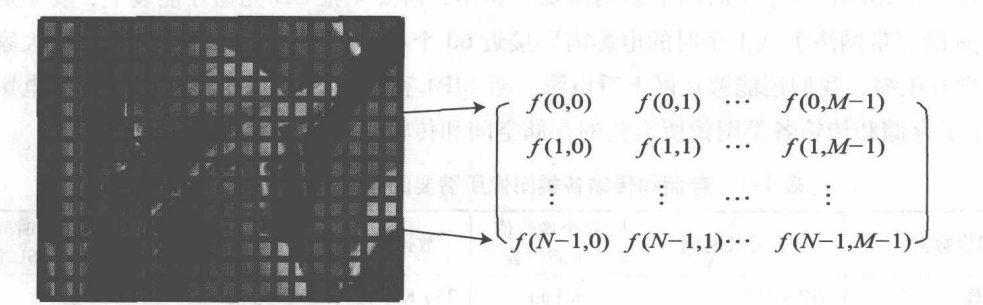


图 1-3 数字图像及其矩阵表示

矩阵中每一个元素称为像素 (pixel 或 pel)，每一个像素值称为图像的亮度值，表示图像的颜色信息。通常情况下，若图像是灰度图像，则该值是一个 8 位的无符号整数，若是彩色图像，则该值由三个无符号的 8 位整数组成 (如 RGB 分量)。矩阵的行号和列号分别对应图像垂直方向 (y) 和水平方向 (x) 的坐标。 M 和 N 分别表示图像的宽度和高度。

图像是传感器 (如数码照相机) 在某一瞬间拍摄的平面图像，而视频则是传感器 (如数码摄像机) 在一段时间内所拍摄的平面图像序列，是二维图像在时间轴上的有序排列，用来描述景物的状态和变化过程，是运动着的图像。图像与视频的关系如图 1-4 所示。

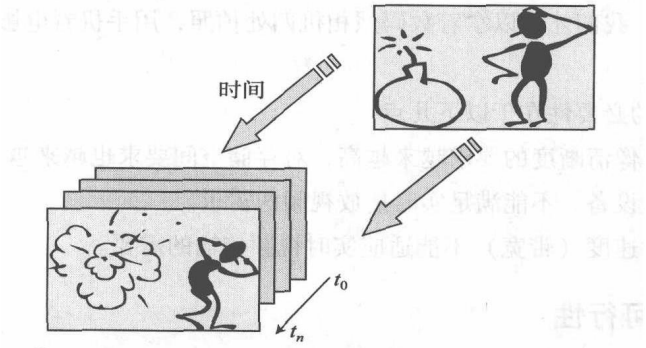


图 1-4 图像与视频的关系

视频中的每一幅图像称为一帧 (frame)，一秒的视频通常由 25 帧 (英国、西欧国家、中国等国家采用的 PAL 制式和法国、俄罗斯采用的 SECAM 制式) 或 30 帧 (美国、加拿大、日本、中美洲、南美洲等采用的 NTSC 制式) 组成。因此，习惯上图像概念也包含视频。

图像信息具有直观、形象、易懂和信息量大的特点，是人类最丰富的视觉来源。“百闻不如一见”说的就是图像这个特点。但由于图像信号的内容复杂，数据量非常大，使图像信息交流很困难。如何在保证图像质量的前提下用最少量的数据实现图像信息的存储、记录和传输，达到优质、经济、可靠的要求，这就是研究图像压缩的目的。

1.1 图像压缩的必要性

为什么要对图像数据进行压缩呢？答案很简单，因为图像数据所占存储空间很大。大家可以考虑一下，要存一秒的电影需要多少字节呢？假设每秒电影由 25 帧 (PAL 制式) 图像组成，每幅图像的宽度和高度分别为 352 和 288，那么一秒的电影所需要的空间为： $352 \times 288 \times 3 \times 25 = 7603200\text{B} \approx 7.2\text{MB}$ ，一个小时的电影则需要 2.6GB，需要 4 张 CD 光盘才能装下，要是通过网速为 1Mbps 的宽带网络下载 1 小时的电影需要接近 60 个小时的时间。通过这些数字，大家可以想象，若没有压缩，我们还能够在网上看电影，听 MP4 和视频聊天吗？更别提高清电影了！表 1-1 列出了存储和传输各类图像所需要的存储空间和传输时间，供大家参考。

表 1-1 存储和传输各类图像所需要的存储空间和传输时间

图像数据	尺寸及时间	每个像素值 所需位数	数据量	带宽	传输时间 (1M ADSL 宽带)
灰度图像	512 × 512	8 bpp	256 KB	2 Mb 每幅	2 秒
彩色图像	512 × 512	24 bpp	768 KB	6 Mb 每幅	6 秒
医学图像	2048 × 1680	12 bpp	5.04 MB	40.3 Mb 每幅	40.3 秒
超清晰图像	2048 × 2048	24 bpp	12.29 MB	98.3 Mb 每幅	98.3 秒
1210 万像素数码照 相机拍摄的照片	4000 × 3000 (最大 有效像素)	24 bpp	34.3 MB	274.66 Mb 每幅	4.58 分钟
30 万像素的数码摄 像机拍摄的视频	640 × 480, 1 分钟 (25 帧每秒)	24 bpp	1.35 GB	180 Mb 每秒	2.92 小时
1080P 高清晰电影	1920 × 1080, 1 分钟, 60 帧每秒 (fps)	24 bpp	21.35 GB	2.85 Gb 每秒	47.46 小时

大家或许听说过 JPEG、MPEG-x、H.26x、MPx，这些都是图像或视频压缩的国际标准，正是因为有了这些标准，我们才可以拿着数码照相机四处拍照，用手机看电影、发彩信、听美妙的音乐，多惬意呀！

总之，图像压缩的必要性在于以下几点：

- 1) 随着对视频图像清晰度的要求越来越高，对存储空间要求也越来越大。
- 2) 低存储空间的设备，不能满足实时播放视频的需求。
- 3) 现有网络传输速度 (带宽) 不能适应实时视频传输的需要。

1.2 图像压缩的可行性

为什么图像能够被压缩呢？因为图像中存在着冗余的数据。需要说明的是数据和信息不是

一回事，数据是信息的表达方式，是信息的载体。数据所需要的存储空间要比信息大，数据中与信息无关的部分称为冗余。数据冗余并不是一个抽象的概念，它可以进行定量地描述。例如用 n_1 和 n_2 分别表示用于表达相同信息的 2 个数据集合中的信息载体单位的个数，那么数据集合 1 对数据集合 2 的相对数据冗余 R_D 定义如下：

$$R_D = \frac{n_1 - n_2}{n_1} = 1 - \frac{n_2}{n_1} = 1 - \frac{1}{C_R}$$

(1-3)

其中 C_R 称为压缩率：

$$C_R = \frac{n_1}{n_2}$$

(1-4)

若 C_R 为 10，则说明数据集合 1 的数据量是数据集合 2 的 10 倍，数据集合 1 相对于数据集合 2 有 90% 的数据是冗余的。

数据压缩或图像压缩方法就是研究如何减少或去掉数据中的冗余部分以减小数据的存储空间。图像压缩中数据冗余包括以下四种。

1. 编码冗余

举一个例子来说明。图 1-5 中右边表格中的数据是左边图像的数据表示。统计得到整幅图像有 64 个像素，其中灰度值是 3 的点 40 个，灰度值为 255 的点 21 个，灰度值是 150 的点 3 个。若每个灰度值用 8 位二进制表示，则共计需要 $64 \times 8 = 512$ 位的存储空间。若只是把三个数区别开，大家想想有必要每个点的值都用 8 位表示吗？确实没必要，我们可以用一位二进制 0 表示出现次数最多的 3，用一位二进制的 1 表示 255，用两位二进制的 10 表示出现次数最少的 150，则表示整幅图像只需要 $40 + 21 + 2 \times 3 = 67$ 位空间就可以了。图像数据的这种冗余称为编码冗余。图像数据存在编码冗余的原因在于各个灰度值在图像中出现的次数不一样。减少这种冗余的方法就是把出现次数多的灰度值用较少位数进行编码，而出现在次数少的灰度值用较多的位数进行编码，这样整幅图像的数据量就少了。后面介绍的赫夫曼编码和算术编码就是基于这种思想进行压缩的。

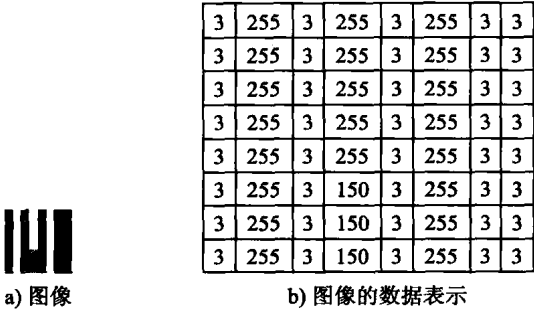


图 1-5 编码冗余示意

2. 空间冗余

空间冗余又称为像素冗余或几何冗余，它描述的是图像具有局部自相似这个特性，也就是说图像中的一点的颜色和周围邻域内的像素点的颜色值相等或接近，单个像素携带的信息相对较小。自然界中的人或物的图像一般都有这个特性。可以想象若人脸图像不具有自相似性，那么

人脸就五彩斑斓，不能称为脸了。

如何利用图像的这个特性进行压缩呢？目前主要有以下四种方式：

1) 对于有连续相等灰度值的情况，采用灰度值和相等点数组组合的编码方式。如图 1-5 的第一列，有连续 8 个点的灰度值是 3，则将这 8 个点编码为“8（相等点的个数），0（灰度值为 3 的编码）”，这种编码方法就是行程编码的基本思想，可以计算，如果对图 1-5 采用这种思想进行编码的话，数据量会很小。

2) 对于灰度值相近的情况，可以利用对当前点与前一个点的差进行编码的方法。由于灰度值相近，差的绝对值就接近于 0，所占的位数就少了，从而达到压缩的目的。在解压缩时，把差加到前一点的灰度值上就得到当前点的灰度值了。例如对数列 [1, 2, 3, 4, 5, 6, 7, 8, 9] 进行压缩，如果采用上面介绍的方法效果就会不好，但对该数列的变形（当前点与前一个点的差）[1, 1(2-1), 1(3-2), 1(4-3), 1(5-4), 1(6-5), 1(7-6), 1(8-7), 1(9-8)] 进行压缩就非常方便了，这是差分编码的思想。在实际应用中，当前点的值可能与前几个点的灰度值相关，此时可以采用当前点与前几个像素点灰度值的加权差的 $e(i) = f(i) - \sum_{j=0}^k a_j f(i-j)$ 进行编码，这是预测编码的思想。

3) 图像像素间有较强的相关性，可以通过一种变换把这相关性去掉，让图像的信息集中在少数几个系数上，这样也可以减少冗余，达到对图像进行压缩的目的。在图 1-6 中，图 1-6b 是图 1-6a 的一种变换，可以看出图 1-6a 的信息只集中在图 1-6b 的两组对称的数据上了。变换域编码就是基于这种思想的，目前各种视频或图像的编码标准都采用了这种思想。

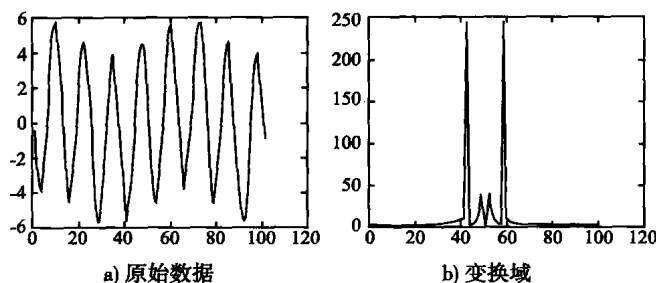


图 1-6 数据及其变换域表示

3. 心理视觉冗余

人观察图像的目的就是获得有用的信息，但人眼并不是对所有的视觉信息具有相同的敏感度，在实际应用中，人也不是对所有的信息具有相同的关心度。在特定场合，一些信息相对另外一些信息来说，就不那么重要，这些相对不重要的信息就是心理视觉冗余。科学实验表明，人眼的分辨力是有限的，人眼不能区别各种颜色或灰度级。对整幅图像而言，人眼能区别 40~60 个灰度级，而对图像的局部，人眼只能区别 32 个灰度级，其他的灰度级相对来说就是心理视觉冗余信息。图 1-5 和图 1-7a 中的图看起来是没有差别，但它们对应的数据却不一致。在这两幅图中，2、3 和 4 的灰度级与 3 相近；148、150 和 151 的灰度级与 150 相近；253、254 和 255 的灰度级与 255 相近。若将图 1-7b 中的相应灰度值都换成对应的 3、150 和 255，这样就利于压缩，数据量就小了，这就是心理视觉冗余。科学实验也表明人眼对亮度信息敏感，而对颜色信息相对不

敏感，因此，在压缩时颜色信息可以损失，在目前流行的视频压缩标准中，就采用了这种思想。图 1-8a 和图 1-8b 在视觉效果上相差无几，但它们的数据量却差了近一倍。

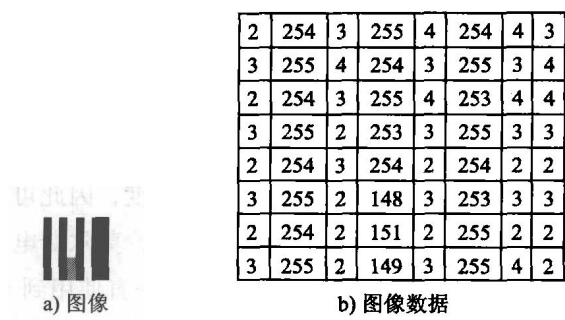


图 1-7 心理视觉冗余演示

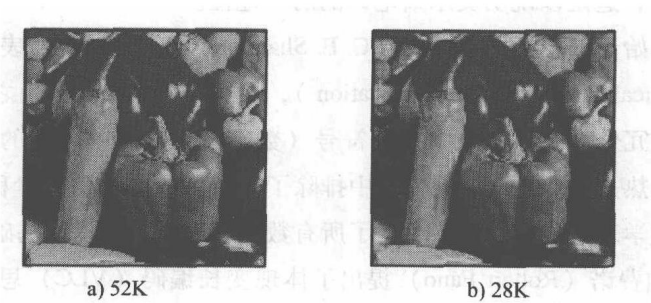


图 1-8 心理视觉冗余压缩演示

4. 时间冗余

视频是由一段有序的图像组成的，在这个序列中相邻两帧图像的拍摄间隔仅为 40 毫秒（PAL 制式），在这么短时间内，相邻两帧图像之间变化不大，具有较大的相关性，这反映为时间冗余。在图 1-9 中，相邻两帧图像差别很小，背景不变，只有小球和手在动，并且幅度很小。



图 1-9 时间冗余示例

视频压缩主要是基于这种特性进行的，通常做法是：利用前面帧的图像或者后面帧来预测当前帧，然后对预测误差采用静态图像的压缩方法进行压缩。现行的各种视频标准都是基于这种思想的。

目前各种图像或视频压缩方法主要是围绕如何减少图像存在的四种冗余进行的，需要说明的是每种压缩方法并不是仅对一种冗余进行，而是对多种冗余进行的。

1.3 图像压缩的发展史

图像压缩的历史也是数据压缩的历史。数据压缩的历史可以追溯到 1838 年, 美国人莫尔斯 (Morse) 发明了电报, 并通过莫尔斯电码进行信息传输。莫尔斯码用点 (· “滴”) 和划 (— “答”) 的组合来表示字母和数字。滴答的个数与字母或数字出现的频率有关, 例如对于英语中使用频率高的字母 e 或 t 用一个“滴”或“答”表示, 而不常出现的字母 z 则用两个“答”和两个“滴”表示。这种编码方式可以有效缩短整篇电文的长度, 因此可以说莫尔斯电码是早期利用编码冗余进行数据压缩的应用。作为一种信息编码标准, 莫尔斯电码拥有其他编码方案无法超越的长久生命。在海事通信中莫尔斯电码作为国际标准一直使用到 1999 年。1997 年, 当法国海军停止使用莫尔斯电码前, 发送的最后一条消息是: “所有人注意, 这是我们在永远沉寂之前最后的一声呐喊!”, 这足以说明莫尔斯电码的历史地位。

现代数据压缩开始于信息论之父香农 (C. E. Shannon) 于 1948 年发表的论文“通信的数学理论” (A Mathematical Theory of Communication), 在这篇信息论的开山之作中, 香农指出: 任何信息都存在冗余, 冗余大小与信息中每个符号 (数字、字母或单词) 的出现概率或者说不确定性有关。香农借鉴热力学的概念, 把信息中排除了冗余后的平均信息量称为“信息熵”, 并给出了计算信息熵的数学表达式。信息熵奠定了所有数据压缩算法的理论基础。

1949 年, 香农和费诺 (Robert Fano) 提出了体现变长编码 (VLC) 思想的香农费诺编码方法, 但该编码方法实用性差。1951 年, 费诺的学生, 麻省理工学院的博士生, 赫夫曼 (David A. Huffman) 在进行老师布置的作业过程中, 对香农费诺编码方法进行优化, 提出了著名的赫夫曼编码方法, 从此赫夫曼算法及其变形在数据压缩领域得到广泛应用。今天, 许多知名的压缩工具和压缩算法 (如 WinRAR、gzip、JPEG 和 MPEG) 里都有赫夫曼编码的影子。

赫夫曼编码方法的领先优势在 20 世纪 70 年代末被两个算法打破, 其一是用一个小数表示一段信息的算术编码, 其二是用数据序列在词典中的条目号代表数据序列的词典编码。1976 年, J. Rissanen 提出了一种可以成功地逼近信息熵极限的编码方法——算术编码。1982 年, Rissanen 和 G. G. Langdon 一起改进了算术编码。之后, 人们又将算术编码与 J. G. Cleary 和 I. H. Witten 于 1984 年提出的部分匹配预测模型 (PPM) 相结合, 开发出了压缩效果近乎完美, 号称压缩效果天下第一的通用压缩算法。算术编码虽然可以获得最短的编码长度, 但其本身的复杂性阻碍了算术编码的应用, 直到在 JPEG2000 中才得以正式采用。

1977 年, Abraham Lempel 和 Jacob Ziv 独辟蹊径, 用最简单的查词典思想, 创造出了比赫夫曼编码更有效, 比算术编码更快捷的压缩算法, 20 世纪 80 年代中期, Terry Welch 对其进行了改进, 从此压缩算法进入 LZW 算法时代, 我们熟悉的 PKZIP、WinZIP、WinRAR、gzip 等压缩工具以及 ZIP、GIF、PNG 等文件格式都是 LZW 算法的杰作。

20 世纪 80 年代, 数字图像得以广泛应用, 各种压缩标准相继出现。1980 年, 国际电报电话咨询委员会 (CCITT) 针对传真类应用, 发布了压缩和传递二值图像的 CCITT Group 3 标准, 在该标准中采用了去除空间冗余的行程编码和去除编码冗余的赫夫曼压缩方法, 目前我们收发传真采用的还是该标准。采用行程编码的位图格式 (BMP), 采用 LZW 算法的图像交换格式 GIF