

TURING

图灵计算机科学丛书

享誉世界的
数据压缩经典著作

Introduction to Data Compression, Fourth Edition

数据压缩导论

(第4版)

[美] Khalid Sayood 著
贾洪峰 译

内容全面新颖

示例精彩丰富

理论联系实践

方便学以致用



人民邮电出版社
POSTS & TELECOM PRESS

TURING

图灵计算机科学丛书

Introduction to Data Compression, Fourth Edition

数据压缩导论

(第4版)

[美] Khalid Sayood 著
贾洪峰 译

人民邮电出版社
北 京

图书在版编目 (C I P) 数据

数据压缩导论 : 第4版 / (美) 萨尤得 (Sayood, K.)
著 ; 贾洪峰译. — 北京 : 人民邮电出版社, 2014. 1
(图灵计算机科学丛书)
书名原文: Introduction to data compression,
fourth edition
ISBN 978-7-115-33600-2

I. ①数… II. ①萨… ②贾… III. ①数据压缩
IV. ①TP274

中国版本图书馆CIP数据核字(2013)第266577号

内 容 提 要

本书是数据压缩方面的经典著作, 目前已经出到了第4版。数据压缩技术及其应用从未停止前进的步伐, 图像、语音、文本、音频、视频等新的应用领域层出不穷。本书也与时俱进, 不断更新, 这一版基本涵盖了数据压缩领域的最新发展。书中首先介绍了基本压缩方法(包括无损压缩和有损压缩)涉及的数学知识, 然后从无损压缩开始, 依次讲述了霍夫曼编码、算术编码以及词典编码技术等。对于有损压缩, 还描述了标量、矢量以及微分编码和分形压缩技术。本书在上一版的基础上, 扩展讨论了基于小波的压缩技术, 特别是越来越流行的 JPEG 2000 标准, 还增加了范式霍夫曼码以及更多有关二进制算术编码的信息。

本书非常适合从事数据压缩相关工作的专业技术人员、软硬件工程师、学生等阅读, 数字图书馆、多媒体等领域的技术人员也可参考。

-
- ◆ 著 [美] Khalid Sayood
译 贾洪峰
责任编辑 刘美英
执行编辑 李 静
责任印制 焦志炜

- ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京鑫正大印刷有限公司印刷
- ◆ 开本: 787×1092 1/16
印张: 35.25
字数: 902千字
印数: 1—3 500册
2014年1月第1版
2014年1月北京第1次印刷
著作权合同登记号 图字: 01-2013-3667号



定价: 129.00元

读者服务热线: (010)51095186转600 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京崇工商广字第 0021 号

站在巨人的肩上
Standing on Shoulders of Giants



www.ituring.com.cn

版 权 声 明

Introduction to Data Compression, Fourth Edition

Khalid Sayood

ISBN: 978-0-12-415796-5.

Copyright © 2012 by Elsevier. All rights reserved.

Authorized Simplified Chinese translation edition published by the Publisher and Co Publisher.

Copyright © 2014 by Elsevier (Singapore) Pte Ltd.

All rights reserved.

Published in China by POSTS & TELECOM PRESS under special arrangement with Elsevier (Singapore) Pte Ltd. This edition is authorized for sale in China only, excluding Hong Kong SAR, Macao SAR and Taiwan Province. Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书简体中文版由 Elsevier (Singapore) Pte Ltd. 授权人民邮电出版社在中华人民共和国境内 (不包括香港特别行政区、澳门特别行政区和台湾地区) 出版与发行。未经许可之出口, 视为违反著作权法, 将受法律之制裁。

本书封底贴有 Elsevier 防伪标签, 无标签者不得销售。

前言

数据压缩已经成为信息革命的一门支撑技术，这场革命改变了我们的生活，而在此过程中，数据压缩也变得几乎无处不在（尽管经常是隐形的）。从MP3播放器到智能手机，再到数字电视和数字电影，数据压缩几乎成了所有信息技术的必备要素。压缩技术越来越多地融入日常生活，也表明这一技术已经相当成熟和稳定。这种成熟还有另外一种体现：本书各版本之间的差别微乎其微。本书第2版增加了在第1版出版之后出现的新技术，第3版增加了一章的篇幅，对于第2版中讨论得不够全面的音频压缩进行了补充。而第4版准备扩展讨论的是基于小波的压缩技术，特别是越来越流行的JPEG 2000标准。这一版共用两章讨论基于小波的压缩技术，其中一章专门讨论基于小波的图像压缩算法。我们还补充了在前几版中省略的一些细节，比如对范式霍夫曼码的介绍和更多有关二进制算术编码的信息。此外，还介绍了一些由互联网推动的技术，比如，在互联网环境中使用的语音编码算法。

本书的篇幅再次因为新增了这些内容而有所增加，但书的主旨没有改变，仍然是对数据压缩技术与理论进行概论性介绍。先是对大多数流行压缩技术进行讲解，然后介绍如何利用这些技术来压缩图像、语音、文本、音频和视频。我们知道大家不希望这本书太厚。当你打开本书，开始阅读某一章时，希望你能很轻松地理解其中的内容。如果有些地方讲解得不够清楚，请将具体问题发送到我的电子邮箱（sayood@datacompression.unl.edu），我会尽力提供帮助（当然，留为家庭作业的问题和项目还是要靠你自己来完成）。

读者对象

如果你正在设计用于实现压缩算法的软件或硬件，或者需要与从事这些设计的人们进行交流，或者正参与多媒体应用的开发，而且拥有电子工程学、计算机工程学或计算科学方面的一些背景知识，本书定会对你有所帮助。书中提供了大量示例以帮助读者自学，还讨论了各种多媒体标准。本书的目的并不是要介绍所有细节，以实现某种标准，而是希望提供一些信息，帮助读者理解和遵循这些标准。最高权威永远都是标准文档。

作为教材

之所以要编写本书，是因为我们需要有一本内容全面的教材，可以向电子工程、计算机工程或计算机科学系高年级本科及研究生介绍数据压缩方面的知识。本书几乎每章最后都会给出

习题和项目计划。出版社提供了解决方案指南。我们还在datacompression.unl.edu上提供了各课程首页的链接，包含大量很有价值的项目思路和相关资料。

对于为时一个学期的课程来说，本书中的资料有些多。不过，很多章节都标注了星号，如果对这些内容取舍得当，各种着重介绍不同压缩技术的课程都可选用本教材。如果强调无损压缩，可以讲授前7章的大部分内容。随后，为了简单体验一下有损压缩，可以介绍第9章的第1节至第5节，接下来讲解第13章中JPEG的内容，还有第19章，其中讨论了多媒体通信中使用的视频压缩。如果课程更侧重音频压缩，那应当介绍第14章和第17章，而放弃第13章和第19章。这时，可以根据班内学生的背景知识，将第12章留作背景阅读材料。如果课程更强调有损压缩，讲授内容应当涵盖第2章、第3章的前两节、第4章的第4节和第6节（并且简单概述第2节和第3节）、第8章、第9章的节选部分，以及第10章至第16章。然后，可以根据课时情况和师生的兴趣，讲授其余三章的部分内容。我发现安排学期项目总是很有帮助的，在这种项目里，学生可以根据自己的兴趣爱好进行研究，学到一些自己感兴趣但课堂上又没有讲授的内容。

学习方法

本书讨论了无损压缩、有损压缩两类压缩技术，以及它们在图像、语音、文本和视频压缩方面的应用。在介绍各种无损与有损编码技术时，没有讲过多的理论知识，仅限于将相关技术联系在一起所需要的内容。而且这些必备理论知识没有提前放在一起进行研究，而是等到真正用到时才介绍。因此，本书共有三章数学预备知识，其中每一章给出的数学材料都是理解和体会后续压缩技术所必需的。

尽管本书是一本导论性的教科书，但导论一词见仁见智。我们尝试采用一种双轨制方法，以满足不同读者的需求。如果我们认为某些材料有助于增进对相关主题的理解，但省略该内容又不至于严重影响理解技术，就用星号（★）标注这些章节。如果读者的主要目的是理解各种压缩技术的工作原理，特别是将本书用作自学教材时，我们建议跳过这些带星号的部分，至少在第一次阅读本书时跳过它们。如果读者需要采用一种偏理论的学习方法，那应当学习这些带星号的章节。除了带星号的章节之外，我们尽量减少与数学相关的内容。

用好示例

对我来说，如果能看到示例，理解起来会更容易一些。所以，我采用大量示例来解释概念。如果读者发现某些概念理解起来有点困难，那多在示例上花点时间应当会有帮助。

从整体上来说，压缩仍然是一种技艺，而要提高对一种技艺的熟练程度，需要对其过程有“感觉”。对于本书讨论的大多数技术，我们都给出了软件实现以及大量数据集。这些软件和数据集可以从datacompression.unl.edu上获得。这些程序是用C语言编写的，在多种平台上进行了测试。它们应当可以在大多数UNIX机器上运行，稍加修改后，也可以在其他操作系统上运行。

强烈建议读者以自己喜欢的数据来运用和修改这些程序，以理解压缩中涉及的一些主题。一种非常有用而又可以实现的目标是：读者在学习完本书之后，自行开发压缩软件包。这样可以很好地体会不同压缩方法中涉及的各种权衡。我们已经尽可能多地给出各种技术的对比。但是，不同数据有各自的特性。在给定条件下究竟应当选用哪种压缩技术，最好的办法就是试用一下。

内容与组织

各章大体内容如下。第2章介绍理解无损压缩需要的数学预备知识。第3章、第4章专门讨论编码算法，包括霍夫曼编码、算术编码、Golomb-Rice码和Tunstall码。第5章和第6章介绍许多流行的无损压缩方案及其应用。这些方案包括LZW、ppm、BWT和DMC等。第7章介绍一些无损图像压缩算法以及它们在许多国际标准中的应用。这些标准包括JBIG标准及各种传真标准。

第8章专门介绍有损压缩的数学预备知识。量化是大多数有损压缩方案的核心，第9章和第10章专门对此进行了研究，第9章讨论的是标量量化，第10章讨论矢量量化。第11章讨论差分编码技术，具体来说就是差分脉冲编码调制（DPCM）和增量调制（ Δ 调制）。这一章还将讨论CCITT G.726标准。

第12章讲的又是数学预备知识，目标是为理解后面四章奠定必要的数学基础，相应内容包括变换、子带和基于小波的压缩技术。和前面介绍数学预备知识的那两章一样，并非所有人都需要这一章介绍的所有内容。第13章介绍JPEG标准，第14章介绍CCITT G.722国际标准，第16章介绍EZW、SPIHT和JPEG 2000。

第17章专门讨论音频压缩，将介绍各种MPEG音频压缩方案，包括人们熟悉的MP3方案。

第18章讨论的技术涉及分析要压缩的数据和生成向接收器发送数据的模型。接收器利用这一模型合成数据。这些分析/合成和合成分析方案包括用于低速率语音编码的线性预测方案，还有分形压缩技术。这一章介绍美国联邦LPC-10标准。码激励线性预测（CELP）是合成分析方案的一个流行示例。这一章讨论三种基于CELP的标准（美国联邦标准1016、国际标准G.728和宽带语音压缩标准G.722.2）和2.4 kbit/s混合激励线性预测（MELP）技术。这一章还将介绍三种在当前因特网应用中进行语音压缩的语音压缩标准：因特网低比特率编解码器、ITU-T G.729标准和SILK。

第19章研究视频编码，通过介绍各种国际标准（包括H.261、H.264和各种MPEG标准）来介绍流行的视频编码技术。

个人观点

对我来说，数据压缩并不只是对数据的处理，它还是找出数据中已有结构的过程。在9世纪，诗人奥玛·海亚姆（Omar Khayyam）写道：

移动的手指书写着，且写
且行；任你虔诚，或是智慧，
都无法诱其回头，消去半行，
任你流尽泪水，也难洗去片言。

——《鲁拜集》

这段内容没有几行文字，但解释起来却需要相当长的篇幅。它们表达了人类共有的体验，所以我们可以头脑中再现诗人在几个世纪之前试图传递的东西。要理解这些文字，不仅要通晓这门语言，而且还要有一个贴近诗人所在时代的生活模型。诗人的天才之处就在于他找到了这样一个现实模型，几个世纪之后，在各种不同文化的人中仍然占有重要地位。所以，寥寥数语可以唤起无限遐思。

数据压缩技术并不追求如此宏大的目标，将它与诗歌相提并论可能有些过于冒昧了。但这两者之间又的确有许多相似之处。在数据压缩中，首先要为各种数据中的不同结构建立模型，然后利用这些模型（可能还需要最终感知这些数据的一些应用环境），以更紧凑的方式来表示这些数据。这些结构可能是非常简单的图案，只要画出数据的图形就能看到；也可能比较复杂，需要更为抽象的方法才能体会到。信息常常包含在数据的结构中，而不是包含在数据本身中，数据压缩技术的发展过程与这些结构的发现分不开。

在Douglas Adams的*The Long Dark Teatime of the Soul*一书中，主人公发现如果他以某种特定方式歪一下自己的脑袋，就能进入瓦尔哈拉殿堂（相当破旧的那种）。要发现数据中的某些结构，有时也需要我们以某种特定方式歪一下脑袋。歪脑袋的方式有无数种，为了避免弄疼脖子，最好知道一些能在一般情况得出有利结果的歪头方式（如果真的弄疼了脖子，那我们的类比就显得荒谬至极了）。本书的目标之一是为读者的深入探索提供一个参考系，希望读者能够在探索过程中感受到和我一样多的愉悦。

致谢

本书的写作过程充满乐趣。因为大家给予我的帮助，我的工作变得相当轻松，而且大大提升了最终书稿的质量。而向大家的帮助表示感谢，这本身就是一件很开心的事。

本书第1版获益于多位专家提出的认真而详尽的批评意见，他们是：IBM公司的Roy Hoffman、加州大学圣克鲁兹分校的Glen Langdon、加州州立理工大学的Debra Lelewer、华盛顿大学的Eve Riskin、柯达公司的Ibrahim Sezan、罗德岛大学的Peter Swaszek。他们详细审阅了第1版的全部或大部分内容。纽约科技大学的Nasir Memon、当时在S3的Victor Ramamoorthy、杜比公司的Grant Davidson、当时在盖布泽土耳其科学技术研究理事会（TÜBITAK）的Hakan Çağlar、加州大学圣巴巴拉分校的Allen Gersho审阅了部分手稿。

对于第2版，北得克萨斯州大学的Steve Tate、新墨西哥州立大学的Sheila Horan、Oerlikon Contraves Group的Edouard Lamboray、加拿大蒙特利尔大学的Steven Pigeon、Raytheon Systems公司的Jesse Olvera审阅了全部手稿。土耳其海峡大学的Emin Anarim以及Hakan Çağlar在我撰写小波一章时提供了帮助。Mark Fowler全面审阅了第12章至第15章，纠正了错误与遗漏问题。Tim James、Devajani Khataniar和Lance Pérez还审阅了第2版的新增部分。Chloeann Nelson在阻止我使用分裂不定式^①的同时，还尽力使本书的前两版变得通俗易懂。第3版获益于Rob Maher（现在蒙大拿州）的批评，他慷慨地腾出时间，帮助我完成了音频压缩一章。

自第1版出版以来，有很多读者提出意见和批评，感谢你们。尤其要感谢Roberto Lopez-Hernandez、Dirk vom Stein、Christopher A. Larrieu、Ren Yih Wu、Humberto D'Ochoa、Roderick Mills、Mark Elston和Jeerasuda Keesorth，感谢他们指出书中的错误，并提出改进建议。我还要感谢给我发来批评意见的各位老师，特别是田纳西大学的Bruce Bomar、得克萨斯大学阿灵顿分校的K.R. Rao、密苏里大学罗拉分校的Ralph Wilkerson、迪尤肯大学的Adam Drozdek、华盛顿大学的Ed Hong和Richard Ladner、科罗拉多矿业大学的Lars Nyland、克罗地亚萨格勒布大学的Mario Kovac、加拿大阿卡迪亚大学的Jim Diamond、以色列本古里安大学的Haim

^① 分裂不定式是指在to与动词之间插入副词，属于英文中不是错误的错误。——译者注

Perlmutter。来自特拉华大学的Paul Amer是我最早、最忠实、最受欢迎的批评者之一。深深感谢他的盛情！

Frazer Williams和Mike Hoffman在内布拉斯加大学与我同系共事，他们审阅了本书第1版。Mike一直为我提供指导，审阅了本书每一版的新增章节，这一版也不例外。我严重依赖于他的洞察力与批评意见，如果没有他，我可能会迷失。能拥有像他这样才华出众、慷慨宽宏的朋友，实属幸事！

这一版中的改进和修改有很多要归功于纽约州立大学宾汉姆顿分校的Mark Fowler和法国国立巴黎高等矿业学校的Pierre Jouvelet。大量新内容是因为Mark认为应当增加而增加的。他在这些修改的计划与实施期间都提供了指导。在我收到的本书批评意见中，Pierre的意见最为详尽。他熟悉压缩技术的各个方面，而且愿意分享它们，从而极大地提升了本书的质量。小波图像压缩一章获益于亚利桑那大学Mike Marcellin的审阅。Mike同意在期末最为繁忙的时候审阅这一章，相信所有从事教学工作的人们都会赞赏这一玉成之举。Mike就是一块美玉。Pat Worster编辑了许多章节，试着教会我如何正确使用分号，让我适当多用一些逗号。本书的阅读体验能有大的改进，正是因为她的关注。有了她的帮助，本来应当可以打造一本完美的教科书。而它之所以未臻完美，正反映了我的不足。

原来在Morgan Kaufmann出版社工作的Rick Adams使我相信必须对本书进行修订了。Andrea Dierna接手处理本书，也要接手与它那个执拗的作者打交道，但她在非常短的时间里完成了审核和修订，让工作开展起来。Meagan White的任务是让本书达到出版标准，这可不是什么值得羡慕的好活，她还得允许我不断打乱她的进度表。Danielle Miller是一位永远都彬彬有礼的项目管理员，尽管她不得不与一位总爱拖延的作者打交道，但却能让整个项目按计划进行。Charles Roumeliotis是文字编辑，他发现了大量我永远不会发现的错误，我和读者都应当好好感谢他。

本书中的大多数示例都是在一间由Andy Hadenfeldt建立的实验室中生成的。James Nau慷慨地付出自己的时间，帮助我从无数的软件泥淖中脱身。在我惊慌失措时，一封邮件或一个语音留言就能把他找来。我实验室（它有一个很恰当的名字——Occult 信息实验室）的同事们在大大小小各个方面为我提供帮助。Sam Way曾试着教我Python（但他失败了），并帮助处理示例。Dave Russell已经在授课中使用过本书，他为我提出了非常有益的批评意见，只是出于对长者的尊重，他的批评总是非常委婉。与Ufuk Nalbantoglu就数据压缩抽象理论进行的讨论，帮助我理清了思路。

本书附带了一些数据集作为示例，我还要感谢这些数据集中的各位“模特”。图像中的人物分别是Sinan Sayood、Sena Sayood、Elif Sevuktekin。那个女声属于Pat Masek。

本书汇集了我这些年来学到的知识。能遇到这些老师非常幸运。David Farden，现在在北达科他州立大学工作，他引领我走进数据通信领域。Norm Griswold原来在得克萨斯农工大学工作，是他引领我走进数据压缩领域。Jerry Gibson现在在加州大学圣巴巴拉分校工作，他是我的博士生导师，帮我开始了自己的专业生涯。这个世界可能不会因此而感谢他，但我要感谢他！

我还从我在内布拉斯加大学和土耳其大学的学生那里学到很多东西。他们的兴趣和好奇心迫使我不断学习，始终与今天广泛的数据压缩领域保持接触。我从他们那里学到的东西，至少与他们从我这里学到的一样多。

如果没有美国宇航局的支持，这些学习过程大多都不会发生。戈达德宇宙飞行中心的Warner Miller和Pen-Shu Yeh、刘易斯研究中心的Wayne Whyte都是这些支持与思想的源泉。我衷心感谢他们给予我的指导、信任和友谊。

我的两个孩子Sena和Sina十分通情达理，并不埋怨我在工作中度过的夜晚和周末。在我首次撰写本书时，他们都还很小（有图片为证）。他们现在都长大了，在我眼里，他们一直都那么棒，而这些年来，“那本书”一直是他们的伙伴（偶尔也让他们不满过）。感谢他们的通情达理，感谢他们给予我的无穷乐趣。

本书之所以能够面世，最要感谢的人就是我的伴侣、我最亲密的朋友Füsun。有了她的支持和友爱，我可以放手去做一些事情，否则我可能根本不会考虑这些事情。她是我宇宙的中心，是我生命的色彩。在遇到她之后，我的所有努力都离不开她的鼎力支持，这本书同样如此。

目 录

第 1 章 引言	1	3.3 非二进制霍夫曼码★	49
1.1 压缩技术	3	3.4 自适应霍夫曼编码	51
1.1.1 无损压缩	3	3.4.1 更新过程	52
1.1.2 有损压缩	4	3.4.2 编码过程	54
1.1.3 性能的测量	4	3.4.3 解码过程	56
1.2 建模与编码	5	3.5 Golomb码	57
1.3 小结	8	3.6 Rice码	58
1.4 项目与习题	8	3.7 Tunstall码	60
第 2 章 无损压缩的数学预备知识	9	3.8 霍夫曼编码的应用	62
2.1 概述	9	3.8.1 无损图像压缩	62
2.2 信息论简介	9	3.8.2 文本压缩	63
2.3 模型	18	3.8.3 音频压缩	64
2.3.1 物理模型	18	3.9 小结	65
2.3.2 概率模型	19	3.10 项目与习题	66
2.3.3 马尔可夫模型	19	第 4 章 算术编码	68
2.3.4 复合信源模型	21	4.1 概述	68
2.4 编码	22	4.2 引言	68
2.4.1 独特可译码	22	4.3 为一个序列编码	69
2.4.2 前缀码	25	4.3.1 生成标签	70
2.4.3 Kraft-McMillan不等式★	25	4.3.2 解读标签	75
2.5 算法信息论	28	4.4 生成二进制码	77
2.6 最小描述长度原则	28	4.4.1 算术代码的独特性和效率	77
2.7 小结	29	4.4.2 算法实现	80
2.8 项目与习题	30	4.4.3 整数实现	84
第 3 章 霍夫曼编码	32	4.5 自适应算术编码	91
3.1 概述	32	4.6 二进制算术编码	91
3.2 霍夫曼编码算法	32	4.6.1 QM编码器	95
3.2.1 最小方差霍夫曼码	35	4.6.2 MQ编码器	95
3.2.2 范式霍夫曼码	37	4.6.3 M编码器	96
3.2.3 有限长度的霍夫曼码	39	4.7 霍夫曼编码与算术编码的对比	97
3.2.4 霍夫曼码的最优性★	41	4.8 应用	99
3.2.5 霍夫曼码的长度★	42	4.9 小结	99
3.2.6 扩展霍夫曼码★	44	4.10 项目与习题	100
3.2.7 霍夫曼码的实现	46	第 5 章 词典方法	102
		5.1 概述	102

5.2 引言	102	7.8 MRC-T.44	160
5.3 静态词典	103	7.9 小结	162
5.4 自适应词典	105	7.10 项目与习题	162
5.4.1 LZ77方法	105	第8章 有损编码的数学预备知识	164
5.4.2 LZ78方法	109	8.1 概述	164
5.5 应用	114	8.2 引言	164
5.5.1 文件压缩——UNIX compress	115	8.3 失真度标准	166
5.5.2 图像压缩——图形交换格式	115	8.3.1 人类视觉系统	168
5.5.3 图像压缩——可移植网络 图形	116	8.3.2 听觉	169
5.5.4 调制解调器中的压缩—— V. 42 bis	117	8.4 信息论回顾★	169
5.6 超越压缩——Lempel-Ziv复杂度★	118	8.4.1 条件熵	170
5.7 小结	120	8.4.2 平均互信息	171
5.8 项目与习题	120	8.4.3 微分熵	172
第6章 基于上下文的压缩	123	8.5 率失真理论★	175
6.1 概述	123	8.6 模型	181
6.2 引言	123	8.6.1 概率模型	181
6.3 部分匹配预测	124	8.6.2 线性系统模型	183
6.3.1 基本算法	124	8.6.3 物理模型	187
6.3.2 转义符号	129	8.7 小结	187
6.3.3 上下文的长度	130	8.8 项目与习题	188
6.3.4 排除原则	131	第9章 标量化	189
6.4 Burrows-Wheeler变换	131	9.1 概述	189
6.5 Buyanovsky关联编码器 (ACB)	135	9.2 引言	189
6.6 动态马尔可夫压缩	136	9.3 量化问题	189
6.7 小结	138	9.4 均匀量化器	193
6.8 项目与习题	138	9.5 自适应量化	201
第7章 无损图像压缩	139	9.5.1 前向自适应量化	201
7.1 概述	139	9.2.5 后向自适应量化	203
7.2 引言	139	9.6 非均匀量化	208
7.3 CALIC	141	9.6.1 pdf优化量化	209
7.4 JPEG-LS	144	9.6.2 压扩量化	212
7.5 使用条件平均值进行预测	146	9.7 熵编码量化	217
7.6 多分辨率方法	147	9.7.1 Lloyd-Max量化器输出的熵 编码	217
7.7 传真编码	151	9.7.2 熵约束量化★	217
7.7.1 游程长度编码	151	9.7.3 高速率最优量化★	218
7.7.2 CCITT第3组与第4组—— 建议T.4与T.6	152	9.8 小结	221
7.7.3 JBIG	154	9.9 项目与习题	221
7.7.4 MH、MR、MMR和JBIG的 对比	158	第10章 矢量量化	223
7.7.5 JBIG2-T.88	159	10.1 概述	223
		10.2 引言	223

10.3 矢量量化器相对于标量量化器的 优势	225	12.2 引言	281
10.4 Linde-Buzo-Gray算法	229	12.3 矢量空间	281
10.4.1 初始化LBG算法	233	12.3.1 点积或内积	282
10.4.2 空单元格问题	237	12.3.2 矢量空间	283
10.4.3 用LBG压缩图像	237	12.3.3 子空间	284
10.5 树状结构的矢量量化器	240	12.3.4 基	284
10.5.1 树状结构矢量量化器的 设计	243	12.3.5 内积的正式定义	285
10.5.2 剪枝树状结构矢量量化器	244	12.3.6 正交集与标准正交集	285
10.6 结构化矢量量化器	244	12.4 傅里叶级数	286
10.6.1 金字塔矢量量化	245	12.5 傅里叶变换	288
10.6.2 极矢量量化器和球面矢量 量化器	246	12.5.1 帕伐瓦尔定理	289
10.6.3 格型矢量量化器	246	12.5.2 调制性质	290
10.7 矢量量化的变体	250	12.5.3 卷积定理	290
10.7.1 增益-形状矢量量化	250	12.6 线性系统	290
10.7.2 去均值矢量量化器	250	12.6.1 时不变	291
10.7.3 分类矢量量化	251	12.6.2 传递函数	291
10.7.4 多级矢量量化	251	12.6.3 冲激响应	292
10.7.5 自适应矢量量化	252	12.6.4 滤波器	293
10.8 网格编码量化	253	12.7 采样	295
10.9 小结	256	12.7.1 理想采样——频域视角	295
10.10 项目与习题	257	12.7.2 理想采样——时域视角	296
第 11 章 差分编码	259	12.8 离散傅里叶变换	298
11.1 概述	259	12.9 Z变换	299
11.2 引言	259	12.9.1 查表法	302
11.3 基本算法	261	12.9.2 部分分式展开	302
11.4 DPCM中的预测	264	12.9.3 长除	306
11.5 自适应DPCM	268	12.9.4 Z变换的性质	306
11.5.1 DPCM中的自适应量化	269	12.9.5 离散卷积	307
11.5.2 DPCM中的自适应预测	269	12.10 小结	308
11.6 增量调制	272	12.11 项目与习题	308
11.6.1 常因子自适应增量调制 (CFDM)	273	第 13 章 变换编码	310
11.6.2 连续可变斜率增量调制	274	13.1 概述	310
11.7 语音编码	274	13.2 引言	310
11.8 图像编码	277	13.3 变换	313
11.9 小结	279	13.4 所关心的变换	317
11.10 项目与习题	279	13.4.1 Karhunen-Loève变换	317
第 12 章 变换、子带与小波的数学 预备知识	281	13.4.2 离散余弦变换	318
12.1 概述	281	13.4.3 离散正弦变换	320
		13.4.4 离散Walsh-Hadamard变换	320
		13.5 变换系数的量化与编码	321
		13.6 在图像压缩中的应用——JPEG	327
		13.6.1 变换	327
		13.6.2 量化	328

13.6.3 编码	329	15.8 小结	400
13.6.4 格式——JFIF	332	15.9 项目与习题	401
13.7 MDCT在音频压缩中的应用	334	第16章 基于小波的图像压缩	402
13.8 小结	336	16.1 概述	402
13.9 项目与习题	337	16.2 引言	402
第14章 子带编码	339	16.3 嵌入式零树编码器	404
14.1 概述	339	16.4 多级树集合分裂	410
14.2 引言	339	16.5 JPEG 2000	415
14.3 滤波器	343	16.5.1 色彩分量变换	416
14.4 基本子带编码算法	348	16.5.2 分片	417
14.4.1 分析	349	16.5.3 小波变换	417
14.4.2 量化与编码	349	16.5.4 量化	418
14.4.3 合成	350	16.5.5 第I层编码	419
14.5 滤波器组设计★	350	16.5.6 第II层编码	425
14.5.1 降采样★	351	16.5.7 JPEG 2000比特流	426
14.5.2 升采样★	353	16.6 小结	430
14.6 使用两通道滤波器组的完美 重构★	354	16.7 项目与习题	430
14.6.1 两通道PR正交镜像滤波 器★	357	第17章 音频编码	432
14.6.2 功率对称FIR滤波器★	359	17.1 概述	432
14.7 M频带正交镜像滤波器组★	360	17.2 引言	432
14.8 多相分解★	362	17.2.1 频谱屏蔽	433
14.9 比特分配	366	17.2.2 时间屏蔽	434
14.10 在语音编码中的应用G. 722	368	17.2.3 心理声学模型	434
14.11 在音频编码中的应用——MPEG 音频	369	17.3 MPEG音频编码	435
14.12 在图像压缩中的应用	369	17.3.1 第I层编码	435
14.12.1 分解图像	371	17.3.2 第II层编码	437
14.12.2 对子带进行编码	373	17.3.3 第III编码——MP3	437
14.13 小结	374	17.4 MPEG高级音频编码	440
14.14 项目与习题	375	17.4.1 MPEG-2 AAC	441
第15章 小波	377	17.4.2 MPEG-4 AAC	444
15.1 概述	377	17.5 Dolby AC-3 (Dolby Digital)	445
15.2 引言	377	17.6 其他标准	446
15.3 小波	379	17.7 小结	447
15.4 多分辨率分析和尺度函数	383	第18章 分析/合成与合成分析方案	448
15.5 用滤波器实现	388	18.1 概述	448
15.5.1 尺度变换与小波系数	390	18.2 引言	448
15.5.2 小波族	392	18.3 语音压缩	449
15.6 双正交小波	394	18.3.1 信道声码器	450
15.7 提升	397	18.3.2 线性预测编码器(美国 政府标准LPC-10)	451
		18.3.3 码激励线性预测(CELP)	457

18.3.4 正弦编码器	459	19.10.13 增强型参考图片选择 模式	502
18.3.5 混合激励线性预测 (MELP)	461	19.11 ITU-T建议书H.264	502
18.4 宽带语音压缩ITU-T G.722.2	463	19.11.1 运动补偿预测	503
18.5 互联网应用的语音编码	464	19.11.2 变换	503
18.5.1 iLBC	464	19.11.3 帧间预测	504
18.5.2 G.729	468	19.11.4 量化	504
18.5.3 SILK	471	19.11.5 编码	506
18.6 图像压缩	473	19.12 MPEG-4第二部分	506
18.7 小结	478	19.13 数据包视频	507
18.8 项目与习题	479	19.13.1 ATM网络	508
第 19 章 视频压缩	480	19.13.2 ATM网络中的压缩问题	508
19.1 概述	480	19.13.3 数据包视频的压缩算法	509
19.2 引言	480	19.14 小结	510
19.3 运动补偿	481	19.15 项目与习题	510
19.4 视频信号表示	483	附录 A 概率与随机过程	511
19.5 ITU-T建议书H.261	488	A.1 概率	511
19.5.1 运动补偿	488	A.1.1 发生频率	511
19.5.2 环路滤波器	489	A.1.2 信任度量	512
19.5.3 变换	490	A.1.3 公理方法	513
19.5.4 量化与编码	490	A.2 随机变量	514
19.5.5 速率控制	491	A.3 分布函数	515
19.6 基于模型的编码	492	A.4 期望	517
19.7 非对称应用	493	A.4.1 均值	518
19.8 MPEG-1视频标准	494	A.4.2 二阶矩	518
19.9 MPEG-2视频标准H.262	496	A.4.3 方差	518
19.10 ITU-T建议书H.263	498	A.5 分布的类型	519
19.10.1 不受限运动矢量模式	500	A.5.1 均匀分布	519
19.10.2 基于语法的算术编码 模式	500	A.5.2 高斯分布	519
19.10.3 高级预测模式	500	A.5.3 拉普拉斯分布	519
19.10.4 PB帧模式和改进的PB帧 模式	500	A.5.4 伽玛分布	519
19.10.5 先进的帧内编码模式	500	A.6 随机过程	520
19.10.6 块效应消除滤波模式	501	A.7 项目与习题	521
19.10.7 参考图片选择模式	501	附录 B 矩阵概念简要回顾	522
19.10.8 时间、SNR和空间伸缩性 模式	501	B.1 矩阵	522
19.10.9 参考图片重新采样	501	B.2 矩阵运算	523
19.10.10 降低分辨率的更新模式	501	附录 C 根格	527
19.10.11 交替帧间变长编码模式	501	参考文献	529
19.10.12 改进量化模式	501	索引	544

第 1 章

引言

在过去 10 年里，我们一直在见证通信方式的变革（有人称之为“革命”），而且这一过程仍在持续。这一变革既包括因特网规模持续不断的增长，也包括移动通信的爆炸式发展，还体现在视频通信重要性的不断增加。在多媒体革命的所有这些领域中，数据压缩都是基础支撑技术之一。如果没有数据压缩算法，就不太可能将图像放在网站上，更不要说音频和视频了。如果没有压缩，移动电话传送的语音也不会变得越来越清晰。没有压缩，数字电视就不会出现。在很长一段时期里，数据压缩只是一小部分工程师和科学家的研究领域，而现在已经无处不在了。用手机打个电话，你会用到压缩。上网冲浪，你是在压缩技术的帮助下打发时间（或消磨时光）。在 MP3 播放器上听音乐，或是观看 DVD，这些娱乐方式也仰仗压缩技术。

那什么是数据压缩，我们为什么需要它呢？大多数人都听说过 JPEG 和 MPEG，它们是用来表示图像、视频和音频的标准。在这些标准中，数据压缩算法用于降低呈现图像、视频或音乐所需要的比特数^①。简言之，数据压缩就是以紧凑方式表示信息的技术或科学。人们识别出数据中存在的结构特征，并加以利用，生成这些紧凑表示方式。这里说的数据可以是文本文件中的字符，可以是语音或图像波形的样本，也可以是由其他过程生成的数列。之所以需要数据压缩，是因为人们以数字形式生成和利用的信息越来越多（所谓数字形式，就是用数据字节表示的数字）。表示多媒体数据所需要的字节数可能非常庞大。例如，要以无压缩数字方式表示 1 秒钟的视频（利用第 18 章介绍的 CCIR 601 格式），需要超过 20 MB，即 160 Mb。想想一部电影的长度有多少秒钟，就明白为什么需要压缩了。要表示 2 分钟无压缩 CD 品质的音乐（44 100 个样本/秒，16 比特/样本），需要超过 8400 万比特。从网站下载这样等级的音乐，会耗费相当长的一段时间。

人类活动对环境的影响越来越大，人们迫切希望更多地了解所处的环境，了解它是如何运转的，我们正在对它做些什么。世界各地的多个航空机构，包括欧洲航天局（ESA）、美国国家航空航天局（NASA）、加拿大航天局（CSA）和日本宇宙航空研究开发机构（JAXA），正在合作从事一项有关全球变化的项目，等到它完全运行时，每天会产生以太字节计的数据。新的基因排序技术使基因信息库的规模不断扩大，而新的医学扫描技术可能会生成拍字节^②的数据。

既然需要传输、存储的数据呈爆炸式增长，那为什么不致力于开发更好的传输与存储技术呢？其实这项工作一直在进行之中，只是取得的成果还不够。人们在这方面已经取得了很大的

① 本书译文中同时使用了“位”和“比特”两种译法，具体取决于上下文的中文习惯。——译者注

② 兆： 10^6 。吉： 10^9 。太： 10^{12} 。拍： 10^{15} 。艾： 10^{18} 。泽： 10^{21} 。尧： 10^{24} 。