



普通高等教育“十一五”国家级规划教材
电子信息科学与工程类专业精品教材

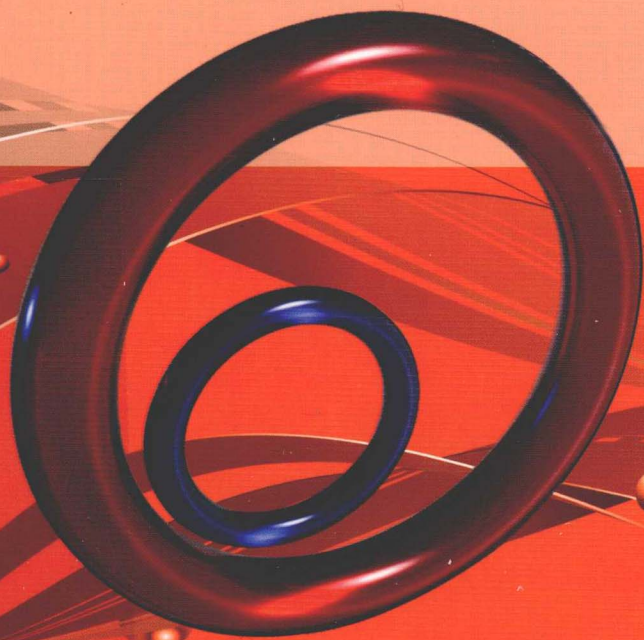


数据压缩 (第三版)

Data Compression

Third Edition

© 吴乐南 编著



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY <http://www.phei.com.cn>



普通高等教育“十一五”国家级规划教材
电子信息科学与工程类专业精品教材

数据压缩

(第三版)

吴乐南 编著

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书系“普通高等教育‘十一五’国家级规划教材”。第三版是在前两版的基础上的修订、补充和更新,继续讲述数据压缩的基本理论、实用技术、先进标准和具体应用。

全书共8章。第1章“绪论”,点明数据压缩的两个基本问题:“为什么压缩”和“压缩什么”。第2章“信源的数字化与压缩系统评价”和第3章“理论极限与基本途径”,是数据压缩的理论基础。第4章“统计编码”不但是无失真数据压缩的基本手段,也常成为其他压缩技术的组成部分,讨论了行之有效的数据和文件压缩算法。第5章“预测编码”和第6章“变换编码”,是限失真信源编码的主要原理和方法,结合语音、图像和电视等具体信源,介绍了有关算法和标准。第7章“分析—综合编码”和第8章“视频编码标准与进展简介”,引述了音像压缩的新方法和国内外新标准,供有志深入者参考。

本书的基本内容是高等学校电子信息科学与工程类专业本科生的教材或教学参考书,全书可用于研究生“数据压缩”或“信源编码”课程的教学参考,同时也适合从事数字通信、广播电视、消费电子、多媒体技术、遥感遥测、安防监控和计算机应用等工作的科技人员自学。

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

数据压缩 / 吴乐南编著. —3版. —北京:电子工业出版社,2012.8
电子信息科学与工程类专业精品教材
ISBN 978-7-121-17756-9

I. ①数… II. ①吴… III. ①数据压缩—高等学校—教材 IV. ①TP274

中国版本图书馆 CIP 数据核字(2012)第 172968 号

责任编辑:陈晓莉

印 刷: 涿州市京南印刷厂
装 订:

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

开 本: 787×1 092 1/16 印张:15.75 字数:443 千字

印 次: 2012 年 8 月第 1 次印刷

定 价: 32.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

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。

服务热线:(010)88258888。

第三版修订前言

本教材自从 2005 年 10 月出版第二版以来,在 2006 年入选“普通高等教育‘十一五’国家级规划教材”。由于我国高等教育的发展,教学和教材“精品化”的要求,特别是本学科领域认识的深化,技术的突破,需求的牵引,产品的创新,市场的竞争,以及对于标准的重视与前瞻,都持续地对本教材内容的修订与更新提出了紧迫的要求和更高的标准,因此,电子工业出版社和读者也要求编者在保持教材总体格局的前提下,对第二版继续进行修订、补充和更新。

第三版的章节总体结构同第二版,除修订了第二版的部分错误和疏漏外,基本保持了原作的框架风貌与教材特点。在内容的增删与更新方面主要进行了以下安排。

第 1 章: 仅略做更新。

第 2 章: 结构不变,做了少量的订正、润色与更新,列出了参考文献。

第 3 章: 在少量订正和润色的基础上,通过新增 3.3.5 节“基本途径之六——利用方差变换”、3.5 节“分布式信源编码”和 3.6 节“压缩感知”(可在教学中酌情处理,均以“*”号表示),将理论研究的新进展凝练为 4 条新的数据压缩基本途径,并列出了参考文献。

第 4 章: 做了少量的订正、润色与更新,扩充了参考文献。重点针对 4.4 节“游程编码”,删除了内容陈旧且不实用的 4.4.1 节“基本方法分析”,更新了 4.4.2 节(原 4.4.3 节)“连续色调图像的二维编码”,以反映国内外视频压缩标准中的新进展。

第 5 章: 结构不变,订正了印刷错误,更新了参考文献,主要将 5.7.4 节由原来的“H. 264 的宏块划分”更改为“预测块划分与亚像素精度”,以侧重基本原理的理解。

第 6 章: 结构不变,订正了印刷错误,并配合正文内容增补了少量参考文献,以反映最新的研究进展。

第 7 章: 除了 7.2 节外,其他结构不变,仅做少量润色和文献更新。重点改写了 7.2 节“宽带声音的子带编码”;扩充了 7.2.1 节“宽带音频编码的特点”;将原 7.2.4 节“音响信号压缩的分析模型”作为新的 7.2.2 节,并删除了其中一些陈旧或非标准的例子,以体现原理在标准前面介绍的教学合理性;将原 7.2.2 节“宽带音频编码的标准考虑”和原 7.2.3 节“MPEG 音频编码标准的发展”的内容,合并调整为新的 7.2.3 节“宽带音频编码的 MPEG 标准”,并通过新的 7.2.4 节相应地增补“宽带音频编码的中国标准”;将原 7.2.5 节“动态码位分配及其 MPEG 实现”展开为新的 7.2.5 节“MPEG—1 音频算法”和新的 7.2.6 节“MPEG—2 AAC 音频算法”,使得各节之间的内容更加均衡,更便于教学内容的模块化组织,也能够更清晰地在整体上与新增的 7.2.7 节“DRA 音频算法”和 7.2.8 节“SVAC 音频算法”这两个新近公布的国家标准相对应,有利于读者的查找和对比。本章新增了大量的参考文献以供有兴趣的读者深入和溯源。

第 8 章: 名称由原来的“视频压缩编码的国际标准简介”更改为“视频编码标准与进展简介”,以涵盖已实施的中国标准和新的标准化进展;将已经陈旧的原 8.1 节“H. 261 建议”、原 8.2 节“MPEG—1 视频压缩标准”、原 8.3 节“MPEG—2 视频压缩标准”和原 8.4 节“H. 263 建议”大幅删减,同时并入原 8.5.2 节“现有视频编码标准的共性技术”和原 8.5.3 节“经典标准算法的不足”的内容,整体浓缩为新的 8.1 节“视频压缩编码国际标准的发展”;将原 8.5 节“MPEG—4 视频压缩标准概述”的其他内容,与原 8.6 节“MPEG—4 基于内容的编码”整体合

并,作为新的 8.2 节“MPEG—4 基于内容的编码”;将原 8.7 节“H. 264/AVC 视频压缩标准”作为新的 8.3 节而适度扩充,删除了非正式出版的参考文献,补充了最新的研究进展;将原 8.8 节“AVS 视频压缩标准简介”充实为新的 8.4 节“AVS 视频压缩标准”;增加了全新的 8.5 节“SVAC 视频压缩标准”、8.6 节“H. 265 视频编码标准简介”和 8.7 节“立体视频编码技术介绍”,补充了大量的近期参考文献。

习题或思考题:仅有个别更新。

缩写词索引:继续充实。

参考文献:统一在书后按章列出,并在正文中有所标记,需要时按章查找即可。为了在有限的篇幅内给出更大的信息量,编者尤其注重给出国内近期的有关博士论文,这样读者在查找时可通过每篇博士论文的综述及后面上百篇的参考文献,追溯到更多的源头期刊和会议论文。这种方式应该有利于不同层次的读者各取所需。谨在此对所有文献的作者表示深深的谢意和敬意。

本教材的参考学时数和对教学内容的取舍仍可参照第一版的“前言”灵活安排。

本教材为“江苏省高等教育精品课程”配套教材,开通的教学网站提供本课程教学团队的全程教学录像和全部课件,需求者请登录:www.hxedu.com.cn 索取。

本教材 3.6 节由戚晨皓讲师编写,8.7 节由罗琳副教授编写;陈阳教授提供了对于第二版的详细勘误;一些兄弟院校的任课教师也提出了许多宝贵意见,谨在此表示衷心的感谢。第三版时至今日才与读者见面,概因作者的耽搁,为此深感内疚。纵然时间太紧、发展太快可作为客观托辞,但跟踪吃力、缺乏动力却是真实写照,为此特别要对第三版的责任编辑陈晓莉女士多年来的耐心、理解、宽容和帮助表示真诚的敬意与谢意。本次修订仍遗留下许多作者自感仓促成文、未及深究以致难以把握的遗憾,恳请广大读者不断批评指正。

吴乐南

2012 年 7 月 18 日于南京

第二版前言

本教材是根据原电子工业部《1996—2000 年全国电子信息类专业教材编审出版规划》，由通信与信息工程专业教学指导委员会编审、推荐出版的。自从 2000 年 6 月第一版出版以来，虽然时间只过去了 52 个月，但我国高等教育的发展，教学要求的提高，特别是本学科领域技术的进步，产品的普及，市场的扩张，标准的换代，以及新的应用需求（如无线多媒体）的牵引，都相应地对本教材内容的更新提出了紧迫的要求和更高的标准。正是在这样的背景下，电子工业出版社要求编者在保持教材总体格局不应有大变化的前提下，对第一版进行修订、补充和更新。

第二版的章节总体结构同第一版，除修订了第一版的部分错误和疏漏外，基本保持了原作风貌，但在章、节、段的目录结构上，取消了原来某些章节中的第 4 级段落，使得各段（即第 3 级目录）的内容更均衡一些，以便教学内容的组织和选取。在内容的增删与更新方面主要进行了以下安排。

第 1 章与第 3 章不变，仅略做润色。

第 2 章：将原 2.2 节“量化”拆分为“标量量化”和“矢量量化”两节。

第 4 章：扩充了 4.2 节“霍夫曼编码”；调整了原来的 4.3 节“游程编码”，即将原 6.3.4 节“JPEG 基本系统的编码”中有关编码的内容按照“难点分散”的原则，作为新的 4.4.3 节“连续色调图像的二维编码”调整过来。另外，新增了 4.3 节“Golomb 编码与通用变长码”，以反映国内外视频压缩标准中的最新进展。

第 5 章：调整了 5.3 节，即将原来分散在原 5.3 节“语音信号的预测编码”与原 7.2 节“宽带声音的子带编码”中有关听觉感知方面的内容加以集中，形成新的 5.3 节“音频信号与听觉感知”；重新改写了原 5.3 节“语音信号的预测编码”，并作为新的 5.4 节；在原 5.4 节“静止图像的预测编码”中补充了国际标准 H. 264/AVC 和国家标准 AVS 视频中新出现的帧内预测模式；将原 5.5 节“活动图像的预测编码”拆分为新的 5.6 节“视频信号与视觉感知”和 5.7 节“活动图像的预测编码”两节；前者并入了原 8.1.1 节和原 8.2.1 节的“输入图像格式”，以突出重点，便于查找；后者则新增了 5.7.4 节“H. 264 的宏块划分”。

第 6 章：在 6.2 节中新增了在 H. 264 中出现的“基于 DCT 的整数变换”（6.2.5 节），并改写了 6.3 节。

第 7 章：裁减了 7.2 节；重新改写了原 7.3 节“小波变换编码”：理论部分作为新的 7.3 节“小波分析简介”可在教学中酌情处理（以“*”号表示），应用部分则作为新的 7.4 节“静止图像的小波变换编码”，重点引入了新的 JPEG 2000 标准；而对原 7.4 节（现 7.5 节）“从波形基编码到模型基编码”（仍加以“*”号表示）则进行了较大的扩充，主要是在模型基编码方面，意在分散 MPEG-4 视频标准介绍中的难点，并将这些体现先进技术潮流但又尚未普及实用的内容呈现给读者评判和思考。

第 8 章：主要基于技术发展、产品更新和即将开始的主流应用方面的考虑，简化或裁减了涉及 H. 261、MPEG-1、MPEG-2 和 H. 263 的前 4 节；拆分并充实了原 8.5 节“MPEG-4 视频压缩标准”；新增了 8.7 节“H. 264/AVC 视频压缩标准”和 8.8 节“AVS 视频压缩标准简介”。除了展现国际上的新标准，也想为我国自己的标准做一点宣传。

另外，各章都增添了一些习题或思考题，而书末的“习题答案”和“缩写词索引”也有相应的

充实。

关于参考文献,除了个别再次引用并标明出处者,第二版一般只列入本次修订所新增的新文献,而读者需要追溯的其他参考文献可在本教材的第一版及编者的旧作《数据压缩的原理与应用》(电子工业出版社,1995年2月第一版)找到。目的是既压缩篇幅,又为有研究兴趣的读者留下查找的线索;同时为方便查找,第二版的参考文献是按章列出并在正文中有所标记(当然又带来少量的重复)。编者谨在此对所有文献的作者表示深深的谢意和敬意。

本教材的参考学时数和对教学内容的取舍仍可参照第一版的“前言”灵活安排。

最后,编者真诚感谢电子工业出版社文宏武社长的鼓励和李秦华编辑的前期帮助,特别是对本次修订版的责任编辑陈晓莉女士的理解和宽容表示衷心的感谢。但是,尽管编者对于完稿期一拖再拖,仍感到时间太紧,发展太快,跟踪太吃力,力又不从心!使得本次修订仍遗留下许多连编者自己都深感遗憾之处,但愿能在“第三版”中加以弥补,也恳请广大读者不断批评指正。

吴乐南

2005年3月20日于南京

第一版前言

本教材是根据原电子工业部《1996—2000 年全国电子信息类专业教材编审出版规划》，由通信与信息工程专业教学指导委员会编审、推荐出版的。本教材由东南大学吴乐南担任主编，北京大学徐孟侠教授担任主审，电子科技大学李在铭教授担任责任编委。

本教材的参考学时数为 48 学时(编者一直采用 32 学时)，其主要内容大致可分为三部分。第一部分(第 1 至第 3 章)阐述数据压缩的基本概念、前提条件、评价准则、理论极限和主要途径，是全书的理论基础；第二部分(第 4 至第 6 章)讨论统计编码、预测编码和变换编码等数据压缩的经典方法和实用技术，是全书的主体；第三部分(第 7 和第 8 章)介绍了基于子带、小波、分形和模型的数据压缩理论、方法和进展，并作为综合应用的实例，给出了有关视频压缩编码的若干国际标准，是全书的深入和拓展。教材主要依据编者原作《数据压缩的原理与应用》(电子工业出版社，1995 年 2 月第一版)并针对本科教学要求，分散难点(将有关标准所涉及的原理分散到各章)，突出重点(注重概念，将必要的推导和证明留作习题)，压缩篇幅(删除过时或未广泛应用的内容)，更新内容(增加 H. 263、MPEG-4 和 MPEG-2 AAC 的原理)。各章均有少量习题或思考题，书后附有参考文献、习题答案及全书缩写词的索引。

使用本教材时应注意根据不同的教学要求对内容进行适当取舍，灵活安排讲课时数。例如，教学学时为 32/48 学时，从第 1 章至第 8 章所用的参考学时数依次为 2/2、3/4、5/6、6/8、5/8、5/6、2/4 和 2/6(均不包括加“*”的章节)，留出 2/4 学时做实验。为利于自学或跨专业选修，本教材只要求读者具有普通工科大学生的概率论与线性代数基础。本教材也包含了一些较深入的内容，可用于研究生教学或参考。

本教材由吴乐南编写，其中第 7.3.1 至 7.3.4 节的内容及部分参考文献由王桥副教授提供。参加审阅工作的还有陈天授等同志，都为本书提出了许多宝贵意见，在此表示诚挚的感谢。由于编者水平有限，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

2000 年 1 月

目 录

第 1 章 绪论	(1)
1.1 什么是数据压缩	(1)
1.2 数据压缩的必要性	(2)
1.3 数据压缩技术的分类	(3)
1.3.1 数据压缩的一般方法	(3)
1.3.2 可逆压缩	(4)
1.3.3 不可逆压缩	(4)
1.3.4 实用的数据压缩技术	(5)
1.4 数据压缩的标准和应用	(6)
习题与思考题	(7)
第 2 章 信源的数字化与压缩系统评价	(8)
2.1 取样	(8)
2.1.1 取样定理	(8)
2.1.2 内插恢复	(9)
2.1.3 其他表述	(10)
2.2 标量量化	(11)
2.2.1 量化误差	(11)
2.2.2 均匀量化	(12)
2.2.3 最佳量化	(13)
2.2.4 压扩量化	(14)
2.3 矢量量化	(15)
2.3.1 基本原理	(15)
* 2.3.2 码书的设计	(17)
2.4 信号压缩系统的性能评价	(18)
2.4.1 信号质量:客观度量	(19)
2.4.2 信号质量:主观度量	(21)
2.4.3 比特率	(23)
2.4.4 复杂度	(23)
2.4.5 通信时延	(24)
2.4.6 编码与数字通信系统的性能空间	(24)
习题与思考题	(25)
第 3 章 理论极限与基本途径	(27)
3.1 离散无记忆信源	(27)
3.1.1 自信息量和一阶熵	(28)
3.1.2 基本途径之一——概率匹配	(28)
3.2 联合信源	(30)
3.2.1 联合熵与条件熵	(30)
3.2.2 基本途径之二——对独立分量进行编码	(32)
3.3 随机序列	(33)

3.3.1	极限熵	(33)
3.3.2	基本途径之三——利用条件概率	(33)
3.3.3	基本途径之四——利用联合概率	(35)
3.3.4	基本途径之五——对平稳子信源进行编码	(35)
3.3.5	基本途径之六——利用方差变换	(36)
3.4	率失真理论	(36)
3.4.1	率失真函数的基本含义	(36)
3.4.2	离散信源的率失真函数	(38)
* 3.5	分布式信源编码	(39)
3.5.1	无损编码——Slepian-Wolf 理论	(39)
3.5.2	基本途径之七——利用联合解码	(40)
3.5.3	有损编码——Wyner-Ziv 理论	(41)
3.5.4	基本途径之八——利用边信息解码	(41)
* 3.6	压缩感知	(42)
3.6.1	模型	(42)
3.6.2	感知矩阵	(43)
3.6.3	重建算法	(43)
3.6.4	基本途径之九——利用压缩感知	(44)
	习题与思考题	(44)
第4章	统计编码	(45)
4.1	基本原理	(45)
4.1.1	文件的冗余度类型	(45)
4.1.2	编码器的数学描述	(46)
4.1.3	变长码的基本分析	(47)
4.1.4	唯一可译码的存在	(49)
4.1.5	唯一可译码的构造	(50)
4.2	霍夫曼编码	(51)
4.2.1	霍夫曼码的构造	(51)
4.2.2	信源编码基本定理	(52)
4.2.3	截断霍夫曼编码	(54)
4.2.4	自适应霍夫曼编码	(55)
4.3	哥伦布编码与通用变长码	(56)
4.3.1	一元码	(56)
4.3.2	哥伦布编码	(57)
4.3.3	指数哥伦布码	(58)
4.3.4	通用变长码	(58)
4.4	游程编码	(59)
4.4.1	二值图像的游程编码	(59)
4.4.2	连续色调图像的二维编码	(62)
4.5	算术编码	(65)
4.5.1	多元符号编码原理	(66)
4.5.2	二进制编码	(67)
4.5.3	二进制解码	(69)
4.5.4	$Q(s)$ 的确定与编码效率	(70)

4.5.5 算术码评述	(71)
4.6 基于字典的编码	(72)
4.6.1 LZ 码基本概念	(72)
4.6.2 LZW 算法	(72)
4.6.3 通用编码评述	(75)
习题与思考题	(75)
第 5 章 预测编码	(78)
5.1 DPCM 的基本原理	(78)
5.2 最佳线性预测	(79)
5.2.1 MMSE 线性预测	(79)
5.2.2 预测阶数的选择	(81)
5.3 音频信号与听觉感知	(82)
5.3.1 语音信号的时域冗余度	(82)
5.3.2 语音信号的频域冗余度	(83)
5.3.3 单音的听觉感知	(84)
5.3.4 多音的掩蔽效应	(85)
5.4 语音信号的预测编码	(86)
5.4.1 技术与标准的沿革	(86)
5.4.2 LPC 语音合成模型	(88)
5.4.3 线性预测合成—分析编码	(90)
5.5 静止图像的预测编码	(92)
5.5.1 帧内预测器的设计	(92)
5.5.2 JPEG 的无损压缩模式	(93)
5.5.3 JPEG—LS 压缩标准	(94)
5.5.4 H. 264 和 AVS 的帧内预测模式	(95)
5.6 视频信号与视觉感知	(95)
5.6.1 电视信号概述	(96)
5.6.2 数字电视的编码参数	(96)
5.6.3 CIF 格式与 SIF 格式	(99)
5.6.4 电视图像信号的时间冗余度	(99)
5.6.5 人的视觉感知特性	(101)
5.7 活动图像的预测编码	(102)
5.7.1 帧间预测编码的发展	(102)
5.7.2 二维运动估计的基本概念及方法	(103)
5.7.3 块匹配运动估计	(105)
5.7.4 预测块划分与亚像素精度	(106)
习题与思考题	(108)
第 6 章 变换编码	(110)
6.1 基本原理	(110)
6.2 离散正交变换	(111)
6.2.1 基本概念	(111)
6.2.2 KL 变换	(113)
6.2.3 图像编码中的正交变换	(114)
6.2.4 DCT	(115)

6.2.5 基于 DCT 的整数变换	(117)
6.3 图像的正交变换编码	(118)
6.3.1 变换矩阵的选择	(118)
6.3.2 变换域系数的选择	(120)
6.3.3 系数的量化	(121)
6.3.4 JPEG 的操作模式和数据组织	(123)
6.3.5 JPEG 的系统描述	(125)
6.4 MDCT	(126)
* 6.5 深化认识	(127)
习题与思考题	(128)
第 7 章 分析—综合编码	(130)
7.1 子带分析	(130)
7.1.1 子带编码的主要特点	(130)
7.1.2 整数半带滤波器组	(132)
7.1.3 二维子带分解	(133)
7.1.4 正交镜像滤波器组	(133)
7.2 宽带声音的子带编码	(135)
7.2.1 宽带音频编码的特点	(135)
7.2.2 音响信号压缩的分析模型	(137)
7.2.3 宽带音频编码的 MPEG 标准	(138)
7.2.4 宽带音频编码的中国标准	(142)
7.2.5 MPEG—1 音频算法	(144)
7.2.6 MPEG—2 AAC 音频算法	(146)
7.2.7 DRA 音频算法	(148)
7.2.8 SVAC 音频算法	(148)
* 7.3 小波分析简介	(150)
7.3.1 基本观念	(151)
7.3.2 小波基的选择	(154)
7.3.3 第一代小波构造的统一框架	(156)
7.3.4 第二代小波构造的统一框架	(157)
7.3.5 提升格式的特点	(159)
7.4 静止图像的小波变换编码	(160)
7.4.1 图像 DWT 系数的零树结构	(160)
7.4.2 图像 DWT 系数编码的 SPIHT 算法	(162)
7.4.3 JPEG 2000 的发展历程	(164)
7.4.4 JPEG 2000 特征集	(165)
7.4.5 JPEG 2000 图像编码算法	(167)
* 7.5 从波形基编码到模型基编码	(169)
7.5.1 基于信源模型的图像编码技术分类	(170)
7.5.2 分形图像编码简介	(170)
7.5.3 模型基图像编码的基本思想	(173)
7.5.4 MPEG—4 中的人脸模型化定义	(175)
7.5.5 模型基辅助的视频混合编码示例	(177)
习题与思考题	(180)

第 8 章 视频编码标准与进展简介	(181)
8.1 视频压缩编码国际标准的发展.....	(181)
8.1.1 ITU-T H.26x 系列.....	(181)
8.1.2 MPEG-x 系列.....	(186)
8.1.3 现有视频编码标准的共性技术.....	(189)
8.1.4 早期标准算法的不足.....	(190)
8.2 MPEG-4 基于内容的编码.....	(191)
8.2.1 基本描述.....	(191)
8.2.2 视频验证模型.....	(193)
8.2.3 视频对象的分割.....	(194)
8.2.4 视频对象编码.....	(196)
8.3 H.264/AVC 视频压缩标准.....	(197)
8.3.1 基本框架.....	(197)
8.3.2 视频编码技术特征.....	(198)
8.3.3 数据传输技术特征.....	(200)
8.3.4 性能测试.....	(200)
8.4 AVS 视频压缩标准.....	(201)
8.4.1 标准化过程.....	(201)
8.4.2 特色技术.....	(202)
8.4.3 性能与应用.....	(203)
8.5 SVAC 视频压缩标准.....	(204)
8.5.1 标准化过程.....	(204)
8.5.2 技术特点.....	(205)
8.5.3 性能评测.....	(206)
8.6 H.265 视频编码标准简介.....	(207)
8.6.1 视频编码标准划代.....	(207)
8.6.2 H.265/HEVC 新技术预览.....	(207)
8.7 立体视频编码技术介绍.....	(208)
8.7.1 基本原理.....	(209)
8.7.2 多视角编码.....	(210)
8.7.3 标准化进展.....	(211)
习题与思考题.....	(213)
附录 A 习题答案	(214)
附录 B 缩写词索引	(221)
参考文献	(230)

第 1 章 绪 论

折中折中在深入本书讨论之前,先简单了解一下有关数据压缩的基本概念、定义、必要性、一般方法、技术分类和应用前景的概况,有助于提高我们的学习兴趣。

1.1 什么是数据压缩

人类社会已进入信息时代。而信息的本质,要求交流和传播。即使是最高度的机密,也需要有解密的使用者。否则,不能称之为信息。于是需要将信息从“这里”传输到“那里”——典型的“通信”概念;或者将信息从“现在”传输到“将来”——所谓“存储”问题。这两种物理过程,均可用图 1.1 所示的一个统一的数字传输系统模型来概括。

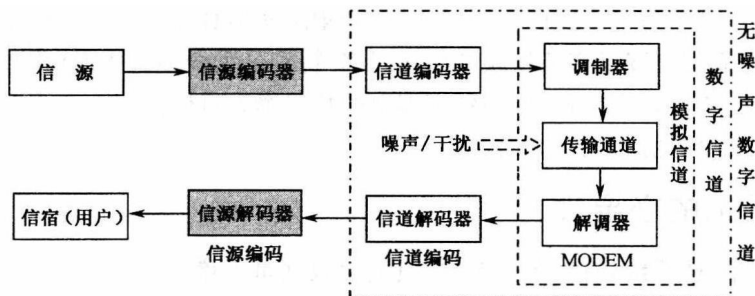


图 1.1 数字传输系统模型

图 1.1 中信源是数字的(或对模拟信号数字化的),而实际的物理传输通道却多为模拟信道,为此发送端通过调制器而接收端借助解调器实现数字序列在模拟信道上的传输。调制器加上解调器(modulator+demodulator)合称调制解调器(modem),简称调制器,它把物理上的模拟信道,转化为一个实际的数字信道。

之所以称“实际的”,是因为在传输过程中信道噪声和干扰在所难免,有可能使解调后的传输信息出错。为此,图 1.1 中的信道编码器加上信道解码器(coder+decoder)合称信道编解码器(codec)的任务,就是尽量使数字信息在传输过程中不出错或少出错,即使出了错也要能自动检错和尽量纠错。如果信道编码和信道解码(统称信道编码)足以保证数字序列的无误差传输,则图 1.1 中的信道 codec 就把由 modem 和模拟信道构成的实际数字信道,改造成一个理想的无噪声数字信道。因此,信道编码主要解决传输的可靠性问题。

而图 1.1 中的信源编码和信源解码即为本书所要研究的内容,统称为信源编码(阴影部分)。信源编码器加上信源解码器也合称信源 codec,它通过对信源的压缩、扰乱、加密等一系列处理,力求用最少的数码最安全地传递最大的信息量。因此,信源编码面对的是传输的安全性和有效性。

信源编码和信道编码都是信息科学的重要分支。信源编码最初研究密码和压缩编码两大类编码方法,但随着理论基础和社会需求的细化,信息安全甚至密码学本身都已发展成为独立的学科。由此,信源编码主要解决传输的有效性问题,它构成了数据压缩的理论基础。那么,什么是数据压缩呢?

数据压缩,就是以最少的数码表示信源所发的信号,减少容纳给定消息集合或数据采样集合的信号空间。

所谓信号空间即被压缩对象,是指:

- ① 物理空间,如存储器、磁盘、磁带、光盘、USB 闪存盘(U 盘)等数据存储介质;
- ② 时间区间,如传输给定消息集合所需要的时间;
- ③ 电磁频段,如为传输给定消息集合所要求的频谱、带宽等。

也就是指某信号集合所占的空域、时域和频域空间。信号空间的这几种形式是相互关联的:存储空间的减少也意味着传输效率的提高与占用带宽的节省。这就是说,只要采用某种方法来减少某一种信号空间,就能压缩数据。从这个意义上讲,通过选择不同 modem 的调制与解调方式,可以在同样的频宽上传送更高的数码率,提高了单位频带的利用率,因而也具有频带压缩之功效。但这属于传输信道的频带压缩技术,与信源本身无关,不在本书讨论。只是有必要指出:调制/解调既影响频带利用的有效性,又涉及信息传输的可靠性。

究竟采用什么方法,以及压缩哪一种信号空间,要根据实际需要与技术条件决定。最初,人们关心提高电话信号传输带宽的利用率,继而希望图文传真提速,接着要求减小数据存储空间,而近年来则更紧迫于降低高清视频的传输码率,特别是无线多媒体通信的传输带宽。近代信源编码的理论与方法,主要也以压缩数字编码的数码率为目标。因此在今天,“数据压缩”与“信源编码”已是两个具有相同含义的术语了。

1.2 数据压缩的必要性

采用数字技术(或系统)有许多优越性,但也使数据量大增。表 1.1 列举了一些常见的数字化音、视频格式。如果对每个取样(取样频率为 f_s)的幅度值用 R 位二进制编码(叫做 R 比特^①)表示,就得到数字信号的传输速率或比特率 I ,即

$$I = f_s \times R(\text{bit/s, b/s 或 bps}) \tag{1.2.1}$$

表 1.1 数字化音、视频格式

数字音频格式	频带范围(Hz)	取样频率(kHz)	样本精度(bit)	声道数	原始码率(Kb/s)
电话	300~3400	8	8	1	64
调幅(AM)广播	50~7000	11.025	16	1	176.4
调频(FM)广播	20~15000	22.03	16	2	705.6
激光唱盘(CD)	20~20000	44.1	16	2	1411.2
数字录音带(DAT)	20~20000	48	16	2	1536
数字电视格式	每秒帧数	图像分辨率(像素)	样本精度(bit)	亮度信号原始码率(Mb/s)	
CIF 格式的亮度信号	30	352×288	8	24.33	
CCIR 601 的亮度信号	30/25	720×480/720×576	8	82.944	
HDTV 亮度信号一例	60	1920×1080	8	995.3	

此即为该信号在通信线路上每秒钟应传送的位数,或者保存 1 秒钟信号样值所需占用的存储容量。传输速率 I 也可以用每秒千位(Kb/s)、每秒兆位(Mb/s)甚至每秒吉位(Gb/s)来表示。当信号带宽给定从而 f_s 为已知且不变时,传输速率就简单地由每样值的位数 R (或 bit/样值)来确定。在有关编码的文献及本书中,比特率(或数码率、码率、速率、数据率)用来表示 I 和 R ,具体指哪一个可从其量纲看出,不会混淆。一般传输时多用 I ,而存储时只用 R ,因为此时不再涉及时间。

① 比特(bit; binary digit)的涵义即为二进制数字。

【例 1-1】从传输角度看:数字电话的取样率最低,按每一取样用 8 位压扩量化(见 2.2.1 节),通常其数码率也需要 $I=8\times 8=64\text{Kb/s}$;一路广播级的彩色数字电视,若按 4:2:2 的分量编码标准格式(见 5.5.1 节),用 13.5/6.75/6.75 MHz 频率采样,每像素(pixel; picture+element,常简写为 pel)用 8 位编码,数码率为

$$I = (13.5 + 6.75 + 6.75) \times 8 = 216 \text{ Mb/s}$$

若实时传送,需占用上述数字话路 3375 个。若能将其压缩到原来的 1/3,即可同时增开 2250 路数字电话;而一路 4:2:2 格式的高清晰度(或高分辨率)电视(HDTV: High Definition Television),数码率更可高达

$$I = 1920 \times 1080 \times 60 \times 2 \times 8 = 1990 \text{ Mb/s}$$

【例 1-2】从存储角度看:一幅 512×512 像素、8bit/pel 的黑白图像占 256 B^①;一幅 512×512 像素、每分量 8 位的彩色图像则占 $3 \times 256 = 768\text{B}$;一幅 $2230 \times 2230 \times 8\text{bit}$ 的气象卫星红外云图占 4.74MB,而一颗卫星每半小时即可发回一次全波段数据(5 个波段),每天的数据量高达 1.1GB。但这与光谱分辨率在 $\lambda/100$ (λ 为光波长)的高光谱(Hyperspectral)甚至达到 $\lambda/1000$ 的超高光谱(Ulttraspectral)遥感图像的海量数据相比,却又差多了。

【例 1-3】海洋地球物理勘探遥测数据,是用 60 路传感器,每路信号按 1kHz 频率采样、16 位模—数转换器(ADC)量化而得,每航测 1km 就需记录 1 盘 0.5 英寸的计算机磁带,而仅仅一条测量船每年就可勘测 15000km!

由此可见,信息时代带来了“信息爆炸”。数据压缩的作用及其社会效益、经济效益将越来越明显。反之,如果不进行数据压缩,则无论传输或存储都很难实用化。而数据压缩的好处就在于:

- ① 较快地传输各种信源(降低信道占有费用)——时间域的压缩;
- ② 在现有通信干线上开通更多的并行业务(如电视、传真、电话、可视图文等)——频率域的压缩;
- ③ 降低发射功率(这对于依靠电池供电的移动通信终端,如手机、个人数字助理(PDA)、无线传感器网络(WSN)等尤为重要)——能量域的压缩;
- ④ 紧缩数据存储容量(降低存储费用)——空间域的压缩。

1.3 数据压缩技术的分类

数据压缩的分类方法繁多,尚未统一。从考察其一般方法入手,可望得到更本质的认识。数据压缩的一般步骤如图 1.2 所示。

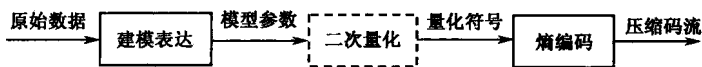


图 1.2 数据压缩的一般步骤

1.3.1 数据压缩的一般方法

所谓“数据”,通常是指信源所发信号的数字化表示或记录。而本书所谓的“数据压缩”,则是考虑以更少的数码来“进一步”地“表示”这样的原始数据。因此,任何数据压缩方法,都可以抽象成图 1.2 所示的 3 个主要步骤(有些步骤可以没有)。

① 遵循计算机工程习惯,本书用大、小写字母“B”和“b”分别表示字节(Byte)和二进制位(bit)。

① 建立一个数学模型,以便能更紧凑或更有效地“重新表达”规律性不那么明显(或本质性不那么突出)的原始数据;

② 设法更简洁地表达利用该模型对原始数据建模所得到的模型参数(或新的数据表示形式)。由于这些参数可能会具有无限的(或过高的)表示精度,因此可以将其量化为有限的精度——为区别于对原始信号数字化时已进行过的一次量化过程,故称为“二次量化”;

③ 对模型参数的量化表示或消息流进行码字分配,以得到尽可能紧凑的压缩码流。此时的编码要求能“忠实地”再现模型参数的量化符号,故称为“熵编码”。

显然,在这“三步曲”中,如果没有“②”且建模表达是一个可逆过程,则从压缩后的码流中就可能完全恢复原始数据;否则,由于“二次量化”的存在,便无法完全再现原始数据。由此,能够取得一致的分类方法,就是将数据压缩分为在某种程度上可逆的与实际上不可逆的两大类,这样更能说明它们的本质区别。而如果综合考虑图 1.2 的“三部曲”,则可将不可逆压缩理解为“混合编码”,因为通常的不可逆压缩过程中总是包含着可逆的编码技术(“两部曲”)。

1.3.2 可逆压缩

可逆压缩也叫做无失真、无差错编码(error free coding)或无噪声(noiseless)编码,而不同专业的文献作者还采用了另外一些术语,如冗余度压缩(redundancy reduction)、熵编码(entropy coding)、数据紧缩(data compaction)、信息保持编码(lossless, bit-preserving),等等。香农(C. E. Shannon)在创立信息论时,提出把数据看做是信息和冗余度的组合。冗余度压缩的工作机理,是去除(至少是减少)那些可能是后来插入数据中的冗余度,因而始终是一个可逆过程。本书也更多地使用了冗余度压缩的术语。

【例 1-4】在一个数据采集系统中,如果信号在一段长时间内不变,则许多连续采样值将是重复的。若能去除这些重复数据,便可得到冗余度压缩。显而易见的方法就是计算两个不同采样值间重复采样的数目(叫做游程),然后将变化的采样值与该重复数目一起发送。显然,这种压缩技术总是可以根据压缩后的数据恢复原来的数据——没有丢失信息。这样做并未涉及数据在物理媒质上的具体存储表示,而有时这种表示本身也会引入额外的冗余度。

【例 1-5】工程上常用 12 位 ADC 采集数据。为了能高速采样并便于处理,往往就用一个字(2Byte)来保存一个样值,这就使得每一样值额外增加了 4 位冗余度,若改用 3 个字(48bit)存 4 个数据,即可消除这一额外冗余度,使数据存储更加紧凑。

这种对于数据外在冗余度的压缩常称为数据紧缩,其原理是直观的,效果是显然的,无须多加讨论。冗余度压缩是针对数据内部的多余信息进行研究,例如对例 1-4 中的重复数据采用不同的表示方法。虽然也有人不加区分地混用“compression”和“compaction”两术语,但读者应当注意“压缩”与“紧缩”二词的细微差异。

1.3.3 不可逆压缩

不可逆压缩就是有失真(lossy)编码,信息论中称熵压缩(entropy compression)。

【例 1-6】为了简单地实现熵压缩,在监测采样值时设置某个门限:只有当采样值超过该门限时,才传输数据。如果这种事件不常出现,就会实现信号空间的较大压缩,但实际的原始采样值就不可能恢复——丢失了信息。

【例 1-7】设想将茶叶(“数据”)倒入一个铁罐(“存储器”)的情形:当罐子装满后,如果这时轻轻地将茶叶罐颠一颠、摇一摇,那么一定可以再多装些,这是因为原来茶叶之间有空隙。而