

数据压缩技术 及其应用

〔美〕托马斯·林奇 编著
吴家安 杜淑玲 译

人民邮电出版社

数据压缩技术及其应用

〔美〕 托马斯 J·林奇 编著
吴家安 杜淑玲 译

人民邮电出版社

Data Compression Techniques and Applications

by Thomas J. Lynch

1985

内 容 提 要

本书较全面、系统地阐述了数据压缩的理论、技术和应用。

本书分为三篇：数据压缩的理论基础；数据压缩技术；数据压缩的应用。全书共有十七章，主要内容有：数据压缩引言、熵压缩、冗余度压缩、变换编码、预测编码、非冗余度采样编码、时间及二进制信源编码、缓冲器控制、传输差错的影响、系统设计、语言压缩、遥测数据压缩、电视压缩、图象压缩和数据库压缩等。

本书内容丰富、取材新颖、阐述清晰、结构合理，包括有80年代以来数据压缩的一些新的研究课题和研究成果，是一本很好的数据压缩理论、技术及应用的参考书。

本书适合用做大专院校有关专业的高年级学生和研究生教材，也可供从事通信、雷达、计算机、遥测、遥控、电视、导航、医学电子学、地震数据处理等技术工作的广大科技工作者和大专院校师生自学或教学参考。

数据压缩技术及其应用

[美] 托马斯 J·林奇 编著

吴家安 杜淑玲 译

*

人民邮电出版社出版

北京东长安街27号

广 益 印 刷 厂 印 刷

新华书店北京发行所发行

各地新华书店经售

*

开本：850×1168 1/32 1989年6月 第一版

印张：12 20/32 页数：202 1989年6月北京第1次印刷

字数：327 千字 印数：1-2 500 册

ISBN 7-115-03835-X/TN.196

定 价：5.70元

译者序

大规模集成电路与计算技术的飞速发展和广泛应用使得信号的数字处理、数字存贮和数字传输日益显示出巨大的优越性，数字化技术的应用范围也迅速扩大到各个技术领域。因此，尽量减少数据信息的物理容积，持续时间和占有带宽以节省数据在处理、存贮和传输中的开销就显得非常重要了，在这种客观需求的推动下“数据压缩”作为一门独立的学科迅速地发展起来了。可以相信，今后数据压缩技术将获得更加广泛地应用，并为各有关领域带来经济上的效益。

为了促进我国数据压缩技术的发展，满足科研、生产和学习的需要，我们翻译了托马斯J.林奇的《数据压缩技术及其应用》(DATA COMPRESSION TECHNIQUES AND APPLICATIONS)一书。这本书内容丰富、取材新颖，并能理论联系实际。全书分为三篇：数据压缩的理论基础；数据压缩技术；数据压缩的应用。其主要特点简述如下：

1. 阐述了有关数据压缩的基本理论，对各种数据压缩技术的原理作了较详细的介绍，并介绍了现代数据压缩应用的大部分实际课题。

2. 给出了较多的应用实例，主要有：语言压缩、电视压缩、图象压缩、数据库压缩、远距引导飞行器电视压缩、遥测数据压缩、地震数据压缩、X射线图象压缩、报纸模版压缩等。这些实例可以做为学习和系统设计的参考。

3. 介绍了数据压缩技术的一些新课题，主要有：矢量编码、二进制信源编码、线性预测编码、数据库压缩和远距引导飞行器电视压缩等。

4. 每章之后都列出了精选的参考文献,供读者进一步学习使用。同时附有对各文献的评述及内容提要。

5. 一些章节还附有一定数量的例题和习题,供读者自学练习使用,书末附有习题解答。

通过我们的阅读感到这本书内容全面、系统,阐述较为深入,是一本难得的关于数据压缩原理及应用的参考书。限于我们的水平,虽然在翻译过程中对原书中的错漏进行了更改和补充、并适当地加了注释,但译文中仍难免有错误或不当之处,希望读者不吝指正。

吴家安 杜淑玲

1986年7月于西安

原 序

数据压缩技术现在已不再是理论上的新奇问题，而是许多数据系统的实际需要。随着电子技术的进步，以前只有用若干大型通用数字计算机才能完成的许多数据压缩技术，现在用硬件就能实现了。

最近一两年，许多部门对于将数据压缩技术用于实际数据系统和通信系统的兴趣与日俱增，例如，在商业界为了传输而进行的报纸模版的压缩，在军事上为了远距离引导飞行器而进行的电视压缩，许多政府机关如美国国家航空和航天管理局(NASA)进行的宇宙飞船的图象压缩。在每一种可能的数据压缩应用中，都需要搞清下述一些问题，即可以利用什么样的压缩技术，这些压缩技术是怎样进行工作的，每一种压缩技术又是如何实现的。

本书将力求为从事实际工作的通信工程师、计算机科学家、信息科学家和数据系统管理人员提供有关这方面的资料，使之成为他们的原始参考材料。

使用本书不需要具有数据压缩方面的实际经验，因为书中对每种压缩技术都分别进行了透彻的阐述；而且本书还给出了必要的理论基础，因而也不需要另外的参考书。

本书在内容编排上分为三篇：第一，理论基础；第二，数据压缩技术；第三，数据压缩的应用。在理论基础这一篇中，给出了必要的理论，以使得读者能够理解后面各章中所使用的概念和术语。此外，还阐述了量化的基本数据压缩技术和最佳信源编码（例如霍夫曼码）。在数据压缩技术这一篇中，讨论了变换编码、预测编码、非冗余度采样编码和二进制信源编码。在这一篇中还包括有时间编码、缓冲器控制、差错控制等具体实现的考虑和系统

设计。在数据压缩的应用这一篇中，阐述了在话音传输、遥测技术、电视、图象和数据库等实际领域内数据压缩的应用。所给出的具体应用的例子都是有代表性的和直观的，但并不是详尽无遗的。

为了便于读者学习和自我练习，在理论基础和数据压缩技术这两篇中给出了许多例题和习题。在每章的末尾都给出了加有注释的参考文献，这些文献可以指导读者在理论上进一步提高，同时为实践方法提供例证。这些参考文献是经过精选的，因而将会加强各章内容，但是这些参考文献本身并不构成一个关于数据压缩的完整的文献目录。

本书中的材料已由作者在相同名称的课程中使用过，该课程是佐治亚州州立华盛顿大学工程继续教育教程的一部分。

托马斯 J. 林奇

目 录

第一篇 数据压缩的理论基础

第一章 数据压缩导言	3
1.1 什么是数据压缩?	3
1.2 数据压缩的分类	4
1.3 本篇内容的编排方法	5
关于进一步阅读的建议	9
第二章 熵压缩	10
2.1 引言	10
2.2 连续数据源和离散数据源	11
2.3 信息和熵的概念	13
2.3.1 离散信源的熵	13
2.3.2 连续信源的熵	15
2.4 速率失真的概念	16
2.5 量化	22
2.5.1 零记忆量化	23
• 输入—输出特性	24
• 均匀量化器	25
• 非均匀量化器	27
• 压扩	30
2.5.2 分组量化	31
2.5.3 序列量化	32
2.5.4 量化器和速率失真边界	36
2.5.5 用 SNR 对无记忆信源的量化器进行比较	39

2.6 其它熵压缩技术	41
本章要点	42
关于进一步阅读的建议	42
第三章 冗余度压缩	48
3.1 引言	48
3.2 信源模型	48
3.2.1 统计独立采样	49
3.2.2 马尔柯夫相关采样	49
3.3 信息、熵和冗余度	50
3.3.1 信息和熵	50
3.3.2 冗余度	53
3.4 压缩比	55
3.4.1 压缩比的上界	55
3.4.2 逼近上界的方法	56
3.5 最佳信源编码	56
3.5.1 香农-范诺码	57
3.5.2 霍夫曼码	58
3.5.3 莫尔斯码	60
3.6 其它冗余度压缩技术	60
3.6.1 非冗余度采样码	61
3.6.2 二进制信源码	61
3.6.3 信道噪声的影响	62
本章要点	62
关于进一步阅读的建议	63

第二篇 数据压缩技术

第四章 数据压缩技术导言	67
4.1 引言	67
4.2 四种基本数据压缩技术	68

4.2.1	变换编码	68
4.2.2	预测编码	69
4.2.3	非冗余度采样编码	69
4.2.4	二进制信源编码	69
4.3	实现的考虑	70
4.4	系统设计的考虑	71
第五章	变换编码	73
5.1	引言	73
5.2	四种变换	74
5.2.1	主要成分变换	75
5.2.2	傅立叶变换	80
5.2.3	阿达玛变换	84
5.2.4	哈尔变换	87
5.3	变换编码的量化	90
5.4	自适应变换编码	92
5.4.1	自适应采样	92
5.4.2	自适应变换	93
5.4.3	自适应量化	93
	本章要点	95
	关于进一步阅读的建议	96
第六章	预测编码	100
6.1	引言	100
6.2	预测	100
6.3	增量调制	102
6.3.1	普通增量调制	103
6.3.2	自适应增量调制	105
6.4	差分脉码调制(DPCM)	106
6.4.1	DPCM系统的工作原理	108
6.4.2	DPCM的设计	108

• DPCM预测器的设计	110
• DPCM量化器的设计	112
6.4.3 自适应DPCM	117
• 自适应预测	117
• 自适应量化	118
6.5 组至组压缩	120
本章要点	121
关于进一步阅读的建议	121
第七章 非冗余度采样编码	125
7.1 引言	125
7.2 多项式预测器	126
7.2.1 零阶预测器	128
• 行程编码	128
7.2.2 一阶预测器	132
7.2.3 多项式预测器的比较	134
7.3 多项式插入器	136
7.4 统计预测器	138
本章要点	139
关于进一步阅读的建议	139
第八章 时间编码及二进制信源编码	142
8.1 引言	142
8.2 时间信息的数学模型	143
8.2.1 统计独立符号模型	144
8.2.2 马尔柯夫模型	145
8.3 时间编码方法	148
8.3.1 序列时间码	149
• 二进制序列码	149
• 林奇—戴维森码	150
8.3.2 非冗余度采样时间(TNRS)码	153

• 标准 TNRS 码	153
• 改进的 TNRS 码(群集码)	153
8.3.3 行程码	155
• 定长码	155
• 霍夫曼(Huffman)码	155
• 二进制不连续 1 (BNO)码	156
8.4 时间编码方法的比较	159
8.4.1 统计独立符号模型	159
• 序列时间码	160
• 非冗余度采样时间(TNRS)码	161
• 行程码	161
8.4.2 马尔柯夫符号模型	163
• 序列时间码	163
• 非冗余度采样时间(TNRS)码	163
• 行程码	164
8.4.3 结论	168
8.5 二进制信源编码	174
8.5.1 序列码	174
8.5.2 TNRS 码	177
8.5.3 行程码	178
本章要点	179
关于进一步阅读的建议	179
第九章 缓冲器控制	184
9.1 引言	184
9.2 缓冲器的数学模型	185
9.3 缓冲器容量的估计	187
9.4 缓冲器上溢控制	188
9.5 缓冲器下溢控制	191
本章要点	191

关于进一步阅读的建议	192
第十章 传输差错的影响	194
10.1 引言	194
10.2 压缩比和失真	194
10.2.1 未压缩—未编码系统	196
10.2.2 压缩—未编码系统	198
10.2.3 压缩—编码系统	199
10.2.4 特性量度的使用	201
10.3 变换编码的差错控制	201
10.4 预测编码的差错控制	203
10.4.1 增量调制中噪声的影响	203
10.4.2 DPCM 中噪声的影响	206
10.4.3 组至组编码中噪声的影响	210
10.5 非冗余度采样编码的差错控制	210
10.5.1 一般情况	211
10.5.2 一个例子	214
本章要点	215
关于进一步阅读的建议	216
第十一章 系统设计	218
11.1 引言	218
11.2 熵压缩和冗余度压缩	218
• 压缩比	219
• 平均失真	220
• 复杂程度	221
11.3 数据压缩技术的特点	221
11.3.1 变换编码	222
11.3.2 预测编码	224
11.3.3 非冗余度采样编码	228
11.4 压缩算法选择	230

11.5 级联编码器	131
本章要点	231

第三篇 数据压缩的应用

第十二章 数据压缩应用导言	235
12.1 引言	235
12.2 数据压缩技术简史	235
12.3 压缩什么和为什么压缩	236
12.4 第三篇的内容及叙述方法	237
12.4.1 语言压缩	238
12.4.2 遥测压缩	238
12.4.3 电视压缩	239
12.4.4 图象压缩	239
12.4.5 数据库压缩	240
第十三章 语言压缩	241
13.1 引言	241
13.2 语言作为数据源	241
13.3 波形编码	244
13.3.1 时域波形编码	245
• 脉冲编码调制(PCM)	245
• 差分脉码调制(DPCM)	246
• 增量调制 (DM)	247
• PCM、自适应 DPCM 和自适应 DM 的比较	249
• 自适应预测编码 (APC)	249
• 可变语言调节 (VSC)	249
13.3.2 频域波形编码	250
• 分波段编码 (SBC)	251
• 自适应变换编码 (ATC)	252
13.4 语声编码	255

13.4.1	声码器	255
13.4.2	线性预测编码器 (LPC)	260
	• LPC 参数	260
	• 对 LPC 参数的量化	262
	• 信道差错对 LPC 特性的影响	264
	• 波形编码器与 LPC 的级联	264
13.5	语言压缩技术的硬件实现	264
	本章要点	265
	关于进一步阅读的建议	265
第十四章	遥测数据压缩	274
14.1	引言	274
14.2	遥测的特点	275
14.2.1	信道带宽	275
14.2.2	时间相关性	275
14.2.3	使用分包的多路复用	276
14.2.4	差错控制	277
14.3	空间数据压缩	277
14.3.1	导弹/火箭遥测数据	277
14.3.2	卫星遥测数据	278
14.4	地震数据压缩	280
14.5	心电图压缩	282
	本章要点	285
	关于进一步阅读的建议	285
第十五章	电视压缩	290
15.1	引言	290
15.2	电视作为信息源	291
15.3	黑白电视压缩	292
15.3.1	变换编码技术	292
	• 二维变换	293

• 三维变换	293
15.3.2 预测编码技术	293
• 增量调制	294
• DPCM	295
• 二维 DPCM	296
• 帧间编码	296
15.3.3 混合编码	297
15.3.4 斜变换	298
15.3.5 伪随机技术	301
• 脉动	301
• 伪随机扫描	301
15.3.6 慢扫描技术	302
15.4 彩色电视压缩	303
15.5 远距引导飞行器 (RPV) 电视压缩	306
15.6 会议电信 (TELECONFERENCING) 电视压缩	307
15.7 可视图文 (VIDEOTEX)/图文电视 (TELETEXT) 压缩	309
15.8 高分辨力电视压缩	310
本章要点	311
关于进一步阅读的建议	311
第十六章 图象压缩	319
16.1 引言	319
16.2 二电平图象压缩	320
16.2.1 二电平传真压缩	320
16.2.2 报纸压缩	323
16.3 多电平图象压缩	324
16.3.1 多频谱图象压缩	324
16.3.2 X 射线图象压缩	327
16.3.3 多电平传真	329

本章要点	329
关于进一步阅读的建议	329
第十七章 数据库压缩	334
17.1 引言	334
17.2 什么是数据库	334
17.2.1 数据文件	335
17.2.2 索引和索引簿	337
17.3 文件压缩	338
17.3.1 非字母数字文件压缩	339
17.3.2 字母数字文件压缩	340
17.4 索引压缩	343
本章要点	346
关于进一步阅读的建议	347
附录 A 差错控制编码简介	351
A.1 差错控制编码和数据压缩	351
A.2 差错控制码的分类	351
A.3 分组码	352
A.4 卷积码	357
附录 B 习题答案	361
附录 C 术语汇编	366