

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。

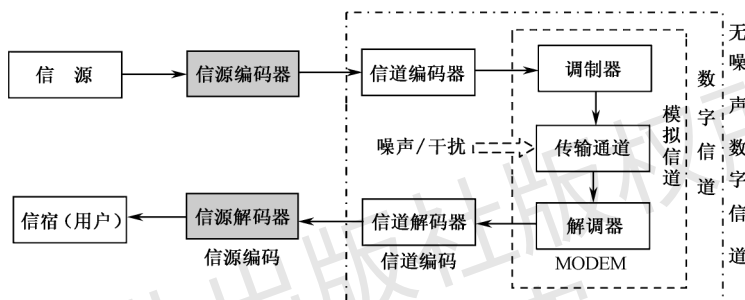


图 1.1 数字传输系统模型

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(Ultaspectral)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。

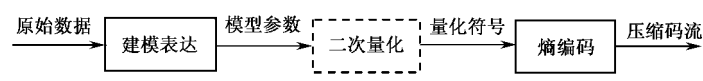


图 1.2 数据压缩的一般步骤

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)。

• 3 •

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

摇晃则可在一定程度上挤掉这种“外在冗余度”——空气所占据的存储空间,使茶叶排列得更紧密(“数据紧缩”——同一个茶叶罐可装得更多)。其次我们设想原来的这些茶叶受了潮,体积自然会有所膨胀,即“数据”具有“内在冗余度”——水分。显然,水分与空气一样,都是我们所不需要的,只会浪费茶叶罐的“存储容量”。如果设法去掉水分,比如将茶叶烘干,那么其体积就会收缩(“冗余度压缩”),同一个罐子还可装得更多。注意,此时茶叶的形状仍然保持完整(“无失真压缩”)。不难想像,如果我们将罐内的茶叶压碎,显然还可以再多装一些,将干茶叶完全压成粉末,一定可以装得最多。但这时的茶叶已不成其为“叶”,只能算做“粉”(“有失真压缩”),而这种压缩已“不可逆”,即我们无法将茶叶末再“还原”成原来形态各异的茶叶。

由此我们已建立了一个初步的概念,即:

- ① 有冗余度就可以压缩(罐中有空气,茶叶含水分);
- ② 压缩只能在一定限度内可逆(茶叶倒出来仍然完整);
- ③ 超过此限度,必然带来失真(茶叶会破碎);
- ④ 允许的失真越大,压缩的比例也可以越大(如果不计较茶叶形状,则压得越碎,同一只罐装得也越多)。在第 3 章我们将会明白,这个“限度”,就是数据的“熵”,这也就是为什么有失真压缩又称为熵压缩的道理。

这些概念直观上不难理解,可却是本学科的一些基本结论。

1.3.4 实用的数据压缩技术

为了去除数据中的冗余度,常常要考虑信源的统计特性,或建立信源的统计模型。因此,许多实用的冗余度压缩技术均可归结于一大类统计编码方法。此外,统计编码技术在各种熵压缩方法中也经常会用到。

熵压缩主要有两大类型:特征抽取和量化,如表 1.2 所示。

特征抽取的典型例子如指纹的模式识别:一旦抽取出足以有效表征与区分不同人指纹的特征参数,便可取其取代原始的指纹图像数据。类似的例子如石油勘探信号的处理:对浩瀚的地震勘探数据进行处理的最末有用结果,有时仅仅是得到了一张地层剖面图。显然,这一大类数据压缩技术是根据特定应用背景而专门设计的,其目的仅仅是为了保护信源中某些感兴趣的内容,根据压缩后的数据已无法在较低的失真度下重现信源的原始风貌,故不在本书讨论之列。

表 1.2 数据压缩技术的简单分类

数据 压 缩	冗余度压缩 (熵编码)	统计编码	霍夫曼编码、游程编码、二进制信源编码等			
			算术编码			
			基于字典的编码;LZW 编码等			
	熵压缩	其他编码	完全可逆的小波分解+统计编码等			
		特征抽取	分析/综合编码		子带、小波、分形、模型基等	
					其他	
		量化	无记忆量化		均匀量化、Max 量化、压扩量化等	
			有记忆量化	序列量化	预测编码	增量调制、线性预测、非线性预测、自适应预测、运动补偿预测等
					其他方法	序贯量化等
			分组量化		直接映射	矢量量化、神经网络、方块截尾等
				变换编码		正交变换;KLT、DCT、DFT、WHT 等
						非正交变换
						其他函数变换等

对于实际应用而言,量化是更为通用的熵压缩技术;除了直接对无记忆信源的单个样本做所谓零记忆量化外,还可以将有记忆信源的多个相关样本映射到不同的空间,去除了原始数据中的相关性后再做量化处理。由此又引出了预测编码和变换编码这两类最常见的实用压缩技术。另外,在特征抽取与量化相结合的基础上,又发展出一类高效的分析/综合编码技术。而一个实用的高效编码方案常常要同时综合考虑各类编码技术之所长。换句话说,常见信源的标准压缩方法,常常是表 1.2 中熵压缩和熵编码的若干方案的“混合编码”。

1.4 数据压缩的标准和应用

数据压缩,可以说是一门既古老又年轻的学科。早在 1843 年出现的莫尔斯(S. Morse)电报码(见 4.1.3 节),就是最原始的变长码数据压缩实例。但由于技术实现上的障碍,长期以来主要处于理论研究和计算机仿真阶段。随着数字信号处理方法、计算机技术和微电子工艺的进步,特别是有关机构如国际标准化组织(ISO)、国际电工委员会(IEC)和国际电信联盟的电信标准部(ITU-T、ITU-TS 或 ITU-TSS)陆续制订的各种数据压缩与通信的标准和建议,极其有力地推动了标准化的数据压缩技术和高效的数字调制技术的迅速普及,使得图 1.1 所示的数字传输系统模型在通信、计算机、广播电视、光盘存储等各个领域都得到了成功的应用,直接引发了消费电子产品乃至整个电子信息领域的一场“数字化革命”,把我们带进了一个以网络化和多媒体化为主要技术特征的崭新时代。

标准化的数据压缩技术,为各种电子信息产品从模拟过渡到数字铺平了道路,其应用已随处可见,家喻户晓。表 1.3 是 1990 年以来已形成(或将形成)的主要数据压缩标准及其应用。这些标准的建立极大地推进了数据压缩技术的实用化、产业化,而全球性的技术竞争、标准开放和经济一体化潮流,反过来又强烈地刺激着信源编码理论研究的进一步拓展,因为任何一种新的数据压缩方法欲广为应用,其性能就必须比现有的标准方法更优异。

表 1.3 主要的数据压缩标准及其典型应用

标 准 号	俗 称	适用信源	典 型 应 用
ITU-T T. 82 ISO/IEC 11544	JBIG-1	二值图像、图形	G4 传真机、计算机图形
ISO/IEC 14492	JBIG-2	二值图像、图形	传真、WWW 图形库、PDA 等
ITU-T T. 81 ISO/IEC 10918	JPEG	连续色调静止图像	图像库、传真、彩色印刷、数码相机等
ITU-T T. 87 ISO/IEC 14495	JPEG-LS	连续色调静止图像	医学、遥感图像资料的无损/近似无损压缩
ISO/IEC 15444	JPEG 2000	连续色调静止图像	各种图形、图像(含计算机生成的)
ITU-T G. 723、G. 728 和 G. 729		语音	数字通信和电话录音等
ITU-T H. 261 建议	P×64	活动图像	ISDN 上的会议电视/可视电话
ITU-T H. 263 建议		活动图像	PSTN 上的会议电视/可视电话
ISO/IEC 11172	MPEG-1	活动图像及伴音	VCD、DAB、多媒体、VOD 等
ITU-T H. 262 ISO/IEC 13818-2	MPEG-2 视频	高质量活动图像	SVCD/DVD、VOD/MOD、多媒体视频游戏、DVB、DTV/HDTV 等
ISO/IEC 13818-3	MPEG-2 音频	高质量多声道声音	DAT、DCC、DAB 等及数字视频伴音
ISO/IEC 14496	MPEG-4	多媒体音像数据	WWW 上的视频、音频扩展
ITU-T H. 264 建议 (MPEG-4 Part 10)		各种活动图像	H. 261/263 和 MPEG-1/2 应用的替代

标 准 号	俗 称	适用信源	典 型 应用
中国的先进音视频编码系列标准 (GB/T 20090—2006)	AVS 标准	活动图像和音频	广播电视、音像产品和多媒体通信等
中国的安全防范监控数字视音频 编解码技术要求 (GB/T 25724 — 2010)	SVAC 标准	活动图像和音频	安防监控、网络摄像头等
ITU—T H. 265 建议	HEVC	活动图像	高清/超高清视频传输、流媒体等

本书首先用两章的篇幅对数据压缩的前提条件——“信源的数字化”、比较标准——“压缩系统评价”和“理论极限与基本途径”进行必要的讨论之后,将逐章对统计编码、预测编码、变换编码和分析/综合编码这几类实用的数据压缩技术加以详细介绍。它们分别以不同的理论准则为指导,并且把一些具体的压缩方法和压缩对象(数据类型)概括在一起。最后,通过对若干视频压缩国际标准的简单介绍,向读者展示完整的综合运用这些技术的实例。

习题与思考题

- 1-1 数据压缩的一个基本问题是“我们要压缩什么”,对此你是怎样理解的?
- 1-2 数据压缩的另一个基本问题是“为什么进行压缩”,对此你又是怎样理解的?
- 1-3 你如何理解信号的空域、时域和频域这几种空间形式是相互关联的?
- 1-4 利用数据压缩可以降低发射功率,道理何在? 如何进行?
- 1-5 你了解计算机中对于浮点数(或实型数据)是如何建模表达、量化与编码的吗?
- 1-6 数据压缩技术是如何分类的?
- 1-7 特征抽取和量化有什么不同吗?
- 1-8 请列举几种你所见到的采用了数据压缩技术的产品。
- 1-9 数据压缩技术为什么要标准化?
- 1-10 你还知道哪些有关数据压缩的标准?
- 1-11 你认为表 1. 3 中的数据压缩标准可以互换使用吗? 为什么?