高等学校教材

信息论与编码

徐家品　编著

高等教育出版社
HIGHER EDUCATION PRESS

高等学校教材

信息论与编码

Xinxilun yu Bianma

徐家品　编著

高等教育出版社·北京
HIGHER EDUCATION PRESS　BEIJING

内容简介

本书着重讲授以经典信息论内容为主的信息论基础；信息论的起源、发展及研究的内容；香农信息论的三大基本概念：信息熵、信道容量和信息率失真函数以及与之对应的三大定理；解决信息传输系统有效性、可靠性和安全性的三类编码；网络信息论与网络编码。

本书围绕信息传输系统这一工程应用背景，特别强调概念阐述清楚、突出重点、深入浅出。逻辑关系上，力求通过建立信息与编码理论的分析方法，在写法上条理清楚，实例丰富翔实，注意循序渐进、难度适中，并注重理论对实际应用的指导作用，引导读者理解信息论与编码的基本方法，培养抽象分析能力和系统工程概念。本书在体系结构上力求科学性、先进性与实用性相统一，内容安排注重体现系统性和相对的完整性。每章后配有习题。

本书可作为高等学校通信工程、电子信息工程、信息工程、电子信息科学与技术和信息安全等专业高年级本科生和研究生教材及教学参考书，不同专业可根据不同的学时数在内容上有所取舍。本书也可以作为信息、通信、电子工程相关专业技术人员的参考书。

图书在版编目（CIP）数据

信息论与编码／徐家品编著. —北京：高等教育出版社，2011.3
ISBN 978-7-04-031600-1

Ⅰ.①信… Ⅱ.①徐… Ⅲ.①信息论-高等学校-教材②信源编码-编码理论-高等学校-教材③信道编码-编码理论-高等学校-教材 Ⅳ.①TN911.2

中国版本图书馆 CIP 数据核字（2011）第 007472 号

策划编辑 吴陈滨　**责任编辑** 王丹丹　**封面设计** 张　楠　**责任绘图** 尹　莉
版式设计 余　杨　**责任校对** 杨雪莲　**责任印制** 朱学忠

出版发行	高等教育出版社	**购书热线**	010-58581118
社　　址	北京市西城区德外大街 4 号	**咨询电话**	400-810-0598
邮政编码	100120	**网　　址**	http://www.hep.edu.cn http://www.hep.com.cn
经　　销	蓝色畅想图书发行有限公司	**网上订购**	http://www.landraco.com http://www.landraco.com.cn
印　　刷	北京鑫海金澳胶印有限公司	**畅想教育**	http://www.widedu.com
开　　本	787×1092　1/16	**版　　次**	2011 年 3 月 第 1 版
印　　张	18.75	**印　　次**	2011 年 3 月第 1 次印刷
字　　数	450 000	**定　　价**	29.50 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换。

版权所有　侵权必究

物料号　31600-00

前　言

物质、能量和信息是构成任何系统的三大要素。信息虽然是无形的和抽象的，但它却是系统的灵魂。

1948 年，香农(C. E. Shannon)发表了他的划时代文章——《通信的数学理论》，宣告了一门崭新的学科——信息论的诞生。经过无数科技工作者 60 多年来的努力奋斗，人们在信息的度量、信息传输特性、纠错编码与压缩编码性能极限等理论问题及各种纠错编码和信源压缩编码方法、信息传输容量的研究方面，都取得了重大突破，有力地促进了通信与信息技术的飞速发展。

通信技术的发展得益于通信理论的正确指导和通信工程关键技术的不断突破。从理论的角度来看，通信的两大基本问题是信息传输的可靠性和有效性。科技工作者们不仅在数学上已严格地证明了香农编码定理，而且发现了各种具体可构造的有效编码理论和方法，可以趋近香农给出的极限。现在经由高斯(Gaussian)信道几乎可无差错地实现信息传输，其信息传输率可达信道容量的 80%(A. J. Viterbi, 1998)。对于非白高斯信道，香农的注水定理和多载波调制(MCM)技术可实现有效、可靠的通信，并已接近于理论极限值。最近出现的 CDMA、MCM、编码正交频分复用(Coded Orthogonal Frequency Division Multiplexing, COFDM)、分组编码调制(Block coded Modulation, BCM)、格码调制(Trellis Coded Modulation, TCM)、Turbo 码、低密度校验(Low Density Parity Check, LDPC)码、各种均衡技术、空时编码、对消技术以及信息存储编码调制技术都充分体现了香农定理的作用。

香农信息论不仅建立了信源和信道编码定理，给出了有效性的极限，而且为人们明确地指出了实现有效而可靠通信的必由之路是数字化和编码。通过 60 多年的努力，人们不仅在理论上发展了香农信息论，而且在实际上逐步实现了某些信道下的香农理论所指出的理想传输。信息论这一抽象而完美的理论，在几十年后会有如此巨大丰富的技术成果，实在令人惊叹！

人类已进入 21 世纪，数字化、信息化、网络化正在冲击、影响、改变社会生活的各个方面。预计到 2025 年，所有的传输都将数字化，灵巧的个人终端将为人们提供各种各样的服务，个人终端将通过几十米至几千米的无线信道与光纤等主干网连通，通向世界，构成了人类生存的信息环境，即信息空间(Cyberspace)。人们除了要有能力在物理空间中生存外，还必须学会能在无形的数字化信息空间中生存。人们的一切活动都将在信息空间中进行竞争和接受检验，许多有形的东西开始向数字的、无形的方向转变。人类活动将无一不受到信息网络的挑战，无一不在信息技术这一最新的高科技生产力的作用下迅速变化，也将因此对人类社会的发展产生巨大的影响。

为适应信息技术发展的新形势，应广大师生的要求，编者结合这些年在教学中的使用情况和科研体会，在已使用 6 年的《信息论与编码》讲义的基础上编写了本书。本书共 8 章，包括香农信息论的基本内容及主要结论；压缩编码的基本原理；纠错原理、方法及其在现代通信系统中的应用等内容。在传统内容的基础上充实了信息理论特别是编码技术发展的新成果和应用。内容涵

盖了通信中有关信息处理的基本原理和方法。

信息论与编码作为一门专业基础课程列入了许多高等学校的教学计划，考虑到不同高等学校和不同专业教学计划的具体要求会有些差异，本书力求兼顾各相关专业对信息论与编码教学内容的偏重和取舍，并将其作为本书的编写特色。书中加强了编码的一些基本技术和方法在通信、计算机网络、数字音像和信息处理等工程实践中的运用内容。考虑到学以致用的原则，本书在编写上加重了信源编码和信道编码部分中实用基础编码技术的分量。为了满足网络和无线通信的飞速发展对网络信息论与网络编码知识的需求，加入“网络信息论与网络编码初步”一章。通信系统的保密性和信息安全问题越来越受到关注，因此，本书加入“通信系统的保密与安全”一章。这些章节供教师在教学时根据专业培养方案的特点选用。

四川大学锦江学院杨雪梅老师参与了编写工作，并编写了本书的第7、8章。王刚、王元龙、汪先涛、李宗、李韬、张江平和杨成恩等研究生参与了资料收集和文字校对工作。

承蒙北京邮电大学吴伟陵教授对本书的审阅，他宝贵而富有建设性的建议和意见，对保证本书质量起到了重要的作用，在此表示衷心的感谢！本书在编写过程中得到了四川大学电子信息学院通信教研室老师们的帮助，在出版过程中得到了高等教育出版社的大力支持，在此表示衷心的感谢！

受视野和知识水平所限，书中难免出现谬误与疏漏，编者殷切恳请读者对此批评指正。

编著者

2010年5月于成都望江楼

试读结束，需要全本PDF请购买 www.ertongbook.com

目　录

第1章 绪 论

近半个世纪以来，以通信理论为核心的经典信息论，以信息技术为载体，它的内涵和外延不断迅速发展，信息论与编码理论所涉及的内容早已超越狭义的通信工程领域，迈入了信息科学的广阔天地。

本章先引入信息的概念，进而讨论信息论与编码理论的形成和发展，并对它研究的对象和内容进行描述。

1.1 信息的基本概念

1.1.1 信息的一般概念

信息(Information)这个词以及与信息这个词相关的词汇层出不穷，遍布报刊、电视和网络中，几乎伴随在当今人类生活的各个角落，深刻地影响着人类世界。信息是信息论中最基本、最重要的概念，那么，究竟什么是“信息”呢？

信息是作用于人类感觉器官的东西；

信息就是情报；

信息是事物之间的差异；

信息是物质和能量在时间和空间中分布的不均匀性；

信息就是知识；

信息是负熵；

数学家认为“信息是让概率分布发生改变的东西”；

哲学家认为“信息是物质的意识成分按完全特殊的方式融合起来的产物”。

很显然，这是人们从不同的角度对于信息的多种理解与描述。

1928年，美国数学家哈特莱(R. V. L. Hartley)在一篇题名为《信息传输》的论文中写到：“信息是选择的自由度”。

1948年，美国数学家香农在创立了信息论的著名长篇论文《通信的数学理论》中写到：“信息就是一种消息”。

1950年，美国数学家、控制论的主要奠基人维纳(N. Wiener)在《控制论——动物和机器中通信与控制问题》一书中写到：“信息既不是物质又不是能量，信息就是信息”。尽管这句话在当初

受到了人们的嘲笑和批评,但正是这句话揭示了信息与物质和能量具有不同的属性,即信息是独立于物质和能量之外存在于客观世界的第三要素。

这仅仅是有了信息的简明释义,我们感受到了信息的存在,但没有信息的准确定义,仍然感觉它难以捉摸。

尽管直至今天尚没有确切的信息定义,但随着科学技术的蓬勃发展,人们对信息的理解也在不断深入,越来越多的人认同:"信息是物质的一种普遍属性,是事物运动状态及运动状态变化的反映"。可以根据不同的约束条件,从不同层次对信息进行描述和定义。在无约束条件的层次上来描述,定义事物的信息是该事物运动的状态和状态改变的方式。这称为"本体论"的信息定义。这个层次上定义的信息是最普遍、最广义和使用范围最广的信息。

以"认识主体"作为约束条件,定义为信息是认识主体(动物或机器)所感知的或所表述的相关事物的运动状态及其变化方式的形式、含义和效用。这称为"认识论"的信息定义。这里认识主体所感知的东西是外部世界向认识主体输入的信息,而认识主体所表述的东西则是其向外部世界输出的信息。这个层次的信息定义其适用范围要窄一些,但内涵丰富。主体具有了感知能力和目的性,可以感知事物的运动状态和变化的内在含义及外在表现形式,能够判断其效用价值。在这个层面来研究信息,要同时研究"含义、形式和效用"3个要素,三者之间相互依存、不可分割。

作为技术术语被广泛使用的"信息",是指那些可收集、检测、识别、提取、转换、存储、传输、处理、检索、分析和利用的对象,是信息的具体表现形式,这些表现形式其实是信息的载体。信息本身既看不见,又摸不着,没有形状,没有重量,没有体积,信息依存于一定的物质形式——信息的载体而存在。计算机通过处理信息的载体来处理信息;通信系统把电磁波——信息的载体传输到信宿来实现信息的传输。所以,作为技术术语的"信息"实际上是指记号、符号、信号等表现信息所用的形式和载体,应用中要把信息的载体和信息的具体内容区分开。

一位美国科学家说过:"没有物质的世界是虚无的世界;没有能源的世界是死寂的世界;没有信息的世界是混乱的世界。"

花朵开放时的色彩可以引来昆虫为其授粉,色彩是一种信息;

成熟的水果会产生香味,诱来动物,动物食后为其传播种子,果香是一种信息;

药有苦味,让人难以吞咽,药味是一种信息;

听老师讲课可以得到许多知识,知识也是信息。

总之,信息处处存在,人的眼、鼻、舌、耳、身都能感知信息。

色彩⟸视觉
果香⟸嗅觉
药味⟸味觉
知识⟸听觉
冷热⟸触觉

1.1.2 信息的特点和性质

根据对信息的一般认识和理解,信息有以下主要特征。

信息来源于物质，又不是物质本身；它从物质的运动中产生出来，又可以脱离原物质而相对独立地存在。

信息与能量息息相关，但又与能量有本质的区别。

信息可以被认识主体获取和利用。

信息具有知识的本性，但又比知识的内涵更广泛。

信息来源于精神世界，但又不局限于精神领域。

根据上述特征和信息的基本定义，可以导出信息的一些重要性质。

存在的普遍性：物质运动及其变化是世界的本质属性，信息与物质运动状态及其变化有着不可分割的内在联系。因此，信息具有存在的普遍性。

有序性：信息是认识主体所感知的或所表述的事物运动状态及其变化方式的形式、含义和效用。信息可以消除事物运动状态和方式的不确定性，增强其有序性。要使系统从无序到有序，必须从外界获取信息。这是特别有价值的信息特性。

相对性：不同的感知者，对于同一个事物，所能获取的信息不一定是相同的，即信息的获取可能因人而异。

可度量性：信息的多少可以用数学方法进行度量。

可共享性：信息既不是物质，也不是能量，可以同时被分配给不同获取者，并被同时使用。当信息被他人共享后，原有的信息并不会因此而丢失。

可压缩性：信息可以进行加工、整理、归纳、概括、浓缩和处理，从而变得更加精练，即信息可以被压缩。

可存储、传输与携带性：信息总是依存于信息载体。因此，信息也可随信息载体一起被存储、传输和携带。

时效性：事物总是不断运动，运动也总是在不断变化，总是经历发生、发展、衰退的过程。信息也在时效上跟随其变化，因而有着时效性。

可扩充性：信息不是一成不变的，随着时间推移和事物不断的发展变化，信息也在不断地扩充。

可扩散性：信息可以在一定时间内进行较大范围内的传播扩散，比如电台广播、电视广播、网上发布等。

1.1.3 香农信息的定义

香农将通信系统概括为图 1-1 所示的框图。在通信系统中，其传输形式是消息，而消息传输过程最基本、最普遍的特点是：收信者在没有收到消息之前，是不知道消息的具体内容的，无法判断发送者发来的是描述何种事物运动状态及其变化的具体消息，也更无法判断描述的是这种状态还是那种状态。即使是收到了消息，由于信道干扰的存在，收信者也不能判断所得到的消息是否正确和可靠。客观事物的运动状态和变化总是不规则的、随机的，在没有获取信息之前，收信者总是存在着不知、不确定或疑义，通过消息的传递，收信者知道了消息的具体内容，原先的不知、不确定或疑义消除或部分消除。消息传递是一个从不知到知的过程，是一个从不确定到确定的过程。

试读结束，需要全本PDF请购买 www.ertongbook.com

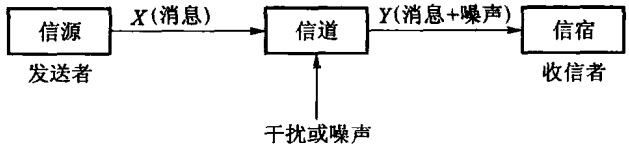

图 1-1 通信系统框图

如果消息是清楚明确的,传递过程中没有差错,收信者收到消息后,原来的不确定性消除了,收信者就获得了所有的信息。如果传递中存在干扰和噪声,消息会变得模糊不清,收信者收到消息后,原先的不确定性一点都没有消除,收信者就没有获得信息。如果有噪声和干扰使消息发生部分差错,收信者原先的不确定性消除了部分,那么,收信者就获得了一部分信息。

由此可见,通信是一个消除不确定性的过程,是一个获取信息的过程。不确定性消除越多,获取的信息就越多;原先的不确定性完全消除,就获取了全部信息;原先的不确定性被部分消除,就获取了部分信息;原先的不确定性完全没有消除,就没有获取任何信息。

信息是事物运动状态或存在方式的不确定性的描述。这就是香农信息的定义。

用数学语言来描述,不确定性就是随机性,具有不确定性的事件就是随机事件。因此,可以运用研究随机事件的数学方法,即概率论和随机过程来测度不确定性的大小。香农信息论主要研究的是概率信息,本书也以概率信息为主要研究对象。

1.2 信息论与编码理论的形成与发展

1928 年,哈特莱首先提出了用对数度量信息的概念。哈特莱的工作给香农很大的启示,他在 1941—1944 年对通信和密码进行深入研究。1948 年,香农在《通信的数学理论》的论文中,用概率测度和数理统计的方法系统地讨论了通信的基本问题,得出了几个重要并带有普遍意义的结论。香农理论的核心是:在通信系统中采用适当的编码后能够实现高效率和高可靠性的信息传输,并得出信源编码定理和信道编码定理。

从数学观点看,这些定理是最优编码的存在定理。但从工程观点看,这些定理不是结构性的,不能从定理的结果直接得出实现最优编码的具体途径。然而,它们给出了编码的性能极限,在理论上阐明了通信系统中各种因素的相互关系,为人们寻找最佳通信系统提供了重要的理论依据。

当已知信源符号的概率特性时,可计算它的信息熵,用熵来表示每个信源符号所载有的信息量。编码定理不但证明了必然存在一种编码方法使码的平均长度可任意接近但不能低于信息熵,而且还阐明达到该目标的途径就是使概率与码长匹配。信源编码定理出现后,编码方法就趋向于合理化。从无失真信源编码定理出发,1948 年,香农在论文中提出并给出了简单的编码方法(香农编码),1952 年,费诺(R. M. Fano)提出了一种费诺码,同年,霍夫曼(D. A. Huffman)构造了一种霍夫曼码,并证明了它是最佳码。

霍夫曼码是有限长度的块码中最好的码,亦即代码总长度最短的码。1956 年,麦克米伦(B. McMillan)首先证明了唯一可译变长码的克拉夫特(Kraft)不等式。霍夫曼码在实际中已有很好的应用,但它仍存在一些块码及变长码所具有的缺点。例如,概率特性必须精确地测定,它若略

有变化,就需要更换码表,以及对于二元信源,常需要多个符号合起来编码,才能取得好的效果等。因此,在实用中常需要做一些改进,同时也就有研究非块码的必要。

算术码就是一种非块码,它是从整个序列的概率匹配来进行编码的。此概念也是香农首先提出的,后经许多学者改进,已进入实用阶段。1968 年前后,埃利斯(P. Elias)发展了香农 - 费诺码,提出了算术编码的初步思路。而里斯桑内(J. Rissanen)在 1976 年给出和发展了算术编码,1982 年,他和兰登(G. G. Langdon)一起将算术编码系统化,并省去了乘法运算,使其更为简化、易于实现。

对概率特性未知或不确知的信源进行有效的编码,上述方法已无能为力。20 世纪 70 年代就有学者提出通用编码。要测定信源的精确概率特性,尤其对高阶条件概率是非常困难的,并且有时信源的概率特性根本无法测定,或是否存在也不知道。例如地震波信号就是如此,因为无法取得大量实验数据。当信源序列是非平稳的时,其概率特性随时间而变更,要测定这种信源的概率特性也近乎不可能。因此,总希望能有一种编码方法,通用于各类概率特性的信源。

1977 年,由齐弗(J. Ziv)和兰佩尔(A. Lempel)提出了 LZ 算法,它是适用于通用信源的编码算法之一。1978 年,他们又提出了改进算法,而且齐弗也证明此方法可达到信源的熵值。1990 年,贝尔(T. C. Bell)等在 LZ 算法基础上又做了一系列变化和改进。现在,LZ 码已广泛应用于文本的数据压缩中。

通用编码中最困难的问题是准则问题。这与概率匹配问题不同,此时已不能确定最佳的标准。当概率特性已知时,信源编码定理给出了极限值,达到这个界的就是最佳码。当概率特性未知时,就无法确定这个界。一般认为它的概率特性是存在的,只是未能测量而不确知而已。这样就可与该概率特性下的极限熵相比较,来确定某种通用编码是否渐近最佳。由此可见,通用编码不但在实用上而且在理论上都需要进一步探讨。

前面介绍的无失真信源编码适用于离散信源或数字信号,不适用于连续信源或模拟信号,如语音、图像等信号的数字处理。因为连续信源的每个样值所能载荷的信息量无限大,而数字信号的值则为有限,对连续信源不引入失真是不可能的。并且连续信源所对应的信宿一般是人,当失真在某一限度以下时是不易被感觉到的,因此是容许的。

限失真信源编码的研究较信道编码和无失真信源编码落后 10 年左右。1948 年,香农在其论文中已体现出了关于率失真函数的思想。1959 年,他发表了"保真度准则下的离散信源编码定理",首先提出了率失真函数及率失真信源编码定理。1971 年伯格尔(T. Berger)的《信息率失真理论》(《Rate Distortion Theory》)是一本较全面地论述有关率失真理论的专著。率失真信源编码理论是信源编码的核心,是频带压缩、数据压缩的理论基础。连续信源编成代码后就无法无失真地恢复原来的连续值,此时只能根据率失真理论进行限失真编码。从率失真函数 $R(D)$ 出发的限失真编码定理虽给出了最佳编码的存在性,也就是在保证平均失真小于允许失真 D 的情况下,最佳码的码率可以压缩到略大于 $R(D)$,但未能给出像概率匹配那样具体的编码途径。限失真编码实际上就是最佳量化问题。最佳标量量化通常不能达到率失真函数所规定的值。后来就提出矢量量化,就是多个信源符号合成一个矢量并对它进行编码。从理论上说,在某些条件下,用矢量量化来编码可达到上述的 $R(D)$ 值,但在实现上还是非常困难的,有待进一步的研究改进。

对于有记忆信源,条件熵必然不大于无条件熵,而且常远小于后者。这就是说,解除符号之间的相关性可进一步压缩码率。以上这些编码方法,都以无记忆信源为目标,对有记忆信源尚可

改进。最简单的方法是多个符号合成为一个新符号,并设新符号组成的序列是独立序列,这样就可用上述方法进行编码。这种方法并不理想。合并的符号数少时,新符号之间的相关性不能解除;合并的符号数多时,复杂性将大为提高,而且对实时处理十分不利。因此,曾提出许多解除相关性的编码方法。比较有效的有预测编码和变换编码:前者利用前几个符号来预测后一个符号的值,预测值与实际值之差,亦即预测误差作为待编码的符号,这些符号之间的相关性就大为减弱,这样可提高压缩比;后者是样值空间的变换,例如从时域变到频域,在某些情况下,可减弱相关性,取得良好的压缩比。预测编码和变换编码已在实际中有所应用。从理论上说,把有记忆信源转换成无记忆序列,尚无理想的方法,更没有不十分复杂而能实际应用的方法。以上简述了根据香农两个编码定理发展起来的各种信源编码方法,也就是从概率论形成的语法信息出发,去掉冗余而达到压缩码率的目的。

现在,编码理论与技术不仅在通信、计算机以及自动控制等电子学领域中得到直接的应用,而且还广泛地渗透到生物学、医学、生理学、语言学、社会学和经济学等各领域。在编码理论与自动控制、系统工程、人工智能、仿生学、电子计算机等学科互相渗透、互相结合的基础上,形成了一些综合性的新兴学科。尤其是随着数学理论,如小波变换、分形几何理论、数学形态学以及相关学科,如模式识别、人工智能、神经网络、感知生理心理学等的深入发展,世界范围内的有关专家一直在寻求现有压缩编码的快速算法,同时,又在不断探索新的科学技术在压缩编码上的应用,因此新颖高效的现代压缩方法相继产生。

小波变换(WT)的小波函数系的时宽、带宽面积很小,且在时间轴和频率轴上都很集中,也就是说展开系数即小波变换系数的能量集中,并且不同频带之间的小波变换系数相关联。据此,有人提出零树(Zerotree)矢量量化方法,它可以达到几百倍的压缩比,且可按不同的压缩比编码,灵活性大。在压缩编码向着智能化和高速化方向发展的今天,神经网络和模型基编码成为当今研究的热点之一。神经网络(Neural Networks,NN)之所以很适合编码,是因为神经网络具有大规模并行处理及分布式信息存储的优势,有良好的自适应性、自组性和容错性,有很强的学习功能、联想记忆功能。神经网络的强映射能力和非线性特性,使它可以通过学习具有相当接近输入信号特征空间基带的能力。因此,用来解决最佳变换的方法是很有效的。应用 BP 算法的多层非线性感知网曾成功地用于 DPCM 编码。利用 Kohonon 的自组织映射进行矢量量化的码本设计取得了极大的成功。用 SOFM 算法所生成的码本就很少依赖于初始码本,且生成的码本的拓扑结构能用来进一步提高编码效率和降低计算复杂度。然而,现有的一些用于编码的神经网络模型都是在模拟人脑功能的思想下建立的,没有考虑信源的特点和肉眼的视觉机理,因此压缩效果不太理想。从理论上讲,神经网络可以模拟肉眼的信息处理过程。这种模拟不限于网络结构方面,还包括网络的学习机制;但大多数神经网络的学习算法中,使用的只是均方误差或 P 阶矩误差失真准则,也没有考虑人类视觉系统的特性。另外,神经网络还未能发挥其强大的信息表征和处理功能,这些与神经网络的理论研究还很不成熟、尚未形成完整的理论体系有关,有待于进一步研究。而基于模型基的编码(Model Based Coding)策略着重利用景物中的物体结构模型,在一定程度上利用了景物的三维信息。也就是说,它使用结构化的信源模型来表示信源信号,其主要优点是用结构的方式来描述信源内容。它的应用领域自然有别于波形编码。模型编码的关键之处就是如何建模。可以建立三维(3D)模型,也可以建立二维(2D)模型。

3D 模型又可分为基于语义的模型(Semantic Based Model)(即参数化的模型)和面向物体的

模型(Object Oriented Model),但是,建模的问题还有待于深入研究。基于语义的方法可以有效地利用景物中已知物体的知识,以实现非常高的压缩比,但它也仅能处理已知的物体,并需要较复杂的信源分析与识别技术。而面向物体的方法可以处理一般的对象,已知的或未知的,显然有更广泛的应用前景,但其未能充分利用景物的知识,或只能在低层次上运用物体知识,编码效率也就无法同前者相比拟。

在研究信源编码的同时,另外一部分科学家从事信道编码(纠错码)的研究工作。这一工作已取得了很大的进展,并已经形成一门独立的分支——纠错码理论。1950 年汉明(R. W. Hamming)发表的论文《检错码与纠错码》是开拓编码理论研究的第一篇论文。这篇论文主要考虑在大型计算机中如何纠正所出现的单个错误。

1952 年,费诺给出并证明了费诺不等式,并给出了关于香农信道编码逆定理的证明;1957 年,沃尔夫维兹(Wolfwitz)采用类似典型序列的方法证明了信道编码强逆定理;1961 年,费诺又描述了分组码中码率、码长和错误概率的关系,并提供了香农信道编码定理的充要性证明;1965 年,格拉格尔(R. G. Gallager)发展了费诺的证明结论,并提供了一种简明的证明方法;1972 年,阿莫托(S. Arimoto)和布莱哈特(R. Blahut)分别发展了信道容量的迭代算法。1948 年,香农首先分析并研究了高斯信道问题;1964 年,霍尔辛格(J. L. Holsinger)发展了有色高斯噪声信道容量的研究;1969 年,平斯克尔(M. S. Pinsker)提出了具有反馈的非白噪声高斯信道容量问题;1989 年,科弗尔(T. M. Cover)对平斯克尔的结论给出了简洁的证明。从能够纠正单个错误的汉明码过渡到能够纠正多个错误的所谓 BCH 码,整整经历了 10 年的时间。因此,可以说 20 世纪 60 年代是代数编码理论发展的鼎盛时期。20 世纪 70 年代出现了高帕码(Goppa Codes),从而又把编码理论推向了一个新的高峰。到了 20 世纪 80 年代,茨伐斯曼(Tsfasman)等人运用代数几何的方法推广了高帕码的思想,指出存在 $GF(m)$ 上的一列码。这一令人吃惊的结果给编码理论的进一步发展带来了新的希望。汉明码出现后,人们把代数方法引入到纠错码的研究,形成了代数编码理论。由此找到了大量可纠正多个错误的好码,而且提出了可实现的编译码方法。但代数编码的渐近性能很差,不能实现香农信道编码定理所指出的结果。因此,1960 年前后提出了卷积码的概率译码,并逐步形成了一系列概率译码理论。尤其以维特比(Viterbi)译码为代表的译码方法被美国卫星通信系统所采用,使香农理论成为真正具有实用意义的科学理论。香农 1961 年的论文《双路通信信道》开拓了网络信息论的研究。1970 年以来,随着卫星通信、计算机通信网的迅速发展,网络信息理论的研究异常活跃,成为当前信息论的中心研究课题之一。艾斯惠特(R. Ahlswede)和廖(H. Liao)分别于 1971 年和 1972 年找出了多元接入信道的信道容量区。接着,1973 年沃尔夫(J. K. Wolf)和斯莱平(D. Slepian)将它推广到具有公共信息的多元接入信道中。科弗尔(T. M. Cover)、艾斯惠特于 1983 年分别发表文章讨论相关信源在多元接入信道的传输问题。1972 年,科弗尔提出了广播信道的研究。伯格曼斯(P. Bergmans)(1973)、格拉格尔(1974)、科弗尔(1975)、马登(K. Marton)(1979)、伊·盖马尔(A. El Gamal)(1979)和范·德·缪伦(E. C. Vander Meulen)(1979)等分别研究了广播信道的容量区问题。近 20 多年来,这一领域研究活跃,使得网络信息论的存在理论已日趋完善。随着人类进入信息时代,信息的传递、存储和交换日益骤增。现代化的通信网、计算机信息网以及各种类型的数据库和电子数据交换系统,特别是因特网的迅速发展,使得信息的安全和保密问题与越来越多的人密切相关。个人、公司、集团、政府部门、军事部门的一些有价值的敏感信息在大规模分布式计算机网的环境下,一方

面为合法用户提供了极大的方便,另一方面也为非法用户提供了更多介入机密信息的机会。保密学是一门研究通信安全和保护信息资源的既古老而又年轻的科学和技术,它包括密码编码学和密码分析学两方面。

密码编码学是信息安全技术的核心,密码编码学的主要任务是寻求产生安全性高的有效密码算法和协议,以满足对消息进行加密或认证的要求。密码分析学的主要任务是破译密码或伪造认证信息,实现窃取机密信息或进行诈骗破坏活动。这两个分支既相互对立又相互依存,正是由于这种对立统一关系,才推动了密码学自身的发展。香农在1949年发表的《保密通信的信息理论》论文中,首先用信息论的观点对信息保密问题进行了全面的论述。由于保密问题的特殊性,直至1976年迪弗(Diffe)和海尔曼(Hellman)发表了《密码学的新方向》一文,提出了公开密钥密码体制后,保密通信问题才得到广泛研究。尤其当今,信息的安全和保密问题更加突出和重要。人们把线性代数、初等数论、矩阵等引入保密问题的研究,已形成了独树一帜的分支——密码学理论。

1.3 信息论与编码理论研究的内容

1.3.1 信息论与编码理论的研究对象

信息论的研究对象是广义通信系统。不仅电子的、光学的信号传递系统,任何系统,只要能够抽象成通信系统模型,如图1-2所示,都可以用信息论与编码理论研究。

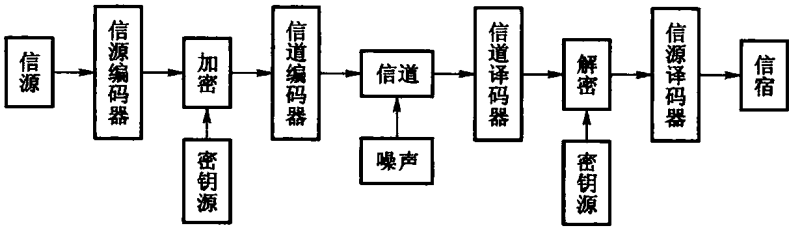

图1-2 通信系统模型

1. 信息源

信息源简称信源,是产生消息和消息序列的源。它可以是人、生物、机器或其他事物,是事物各种运动状态或存在状态的集合。信源可能出现的状态是随机的、不确定的,但又是有一定的规律性的。

2. 编码器

编码是将消息变换成信号的措施。编码器输出的是适合信道传输的信号,信号携带着信息,它是信息的载体。

编码器分为两种类型,即信源编码器和信道编码器。信源编码器对信源输出的消息进行适当的变换处理,目的是提高传输的效率。而信道编码器是为了提高信息传输的可靠性,对消息进行变换处理。

3. 信道

信道是通信系统把载荷着信息的信号从甲地传输到乙地的媒介。对于实际的通信系统，信道有明线、电缆、波导、光纤、无线电波传播空间等，这些都属于传播电磁波能量的信道。对于广义的通信系统，信道就是可以传输和存储信息的任何媒介。

信道要受到干扰或窜入噪声。为了研究方便，通常把系统中其他部分的干扰和噪声都等效折合到信道中，被看成是信道干扰信号，再用一个等效噪声源产生后，让其叠加于传输的信号上。这样信道的输出中除了信号还叠加有干扰信号。干扰和噪声往往具有随机性，因此，信道特性要用概率空间来描述，而噪声源的统计特性是划分信道类型的依据。

4. 译码器

译码是把信道输出的编码信号进行反变换。译码器也分为信源译码器和信道译码器。

5. 信宿

消息传送的对象，即接收消息的人或机器。

6. 加密

以计算机为核心的大规模信息网络，尤其是互联网的建立和发展，对信息传输质量的要求越来越高。不仅要求快速、有效、可靠地传输信息，而且还要求传输过程中保证信息的安全保密，不被篡改和伪造。因此，在编码环节中增加了加密编码，相应地在译码环节中加入了解密译码。

1.3.2 信息论与编码理论的研究范畴

关于信息论的研究范畴，一般有以下三种解释。

1. 狭义信息论

应用近代概率统计方法研究信息的基本性质及度量方法，研究信息传输、处理等一般规律的学科。

主要研究信息的测度、信道容量、信息率失真函数，与这三个概念相对应的香农三定理以及信源和信道编码。

2. 一般信息论

主要是研究信息传输和处理问题。除了香农基本理论之外，还包括噪声理论、信号滤波和预测、统计检测与估计理论、调制理论。后一部分内容以美国科学家维纳为代表。

虽然维纳和香农等人都是运用概率和统计数学的方法研究准确或近似再现消息的问题，都是通信系统的最优化问题。但他们之间有一个重要的区别。

维纳的研究对象如图 1-3 所示，重点是在接收端。研究消息在传输过程中受到干扰时，在接收端如何把消息从干扰中提取出来，并建立了最佳过滤理论（维纳滤波器）、统计检测与估计理论、噪声理论等。

香农的研究对象如图 1-4 所示，是从信源到信宿的全过程，是接收、发送端联合最优化问

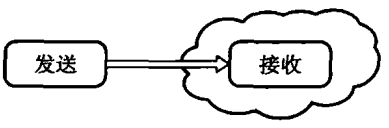

图 1-3 维纳的研究对象

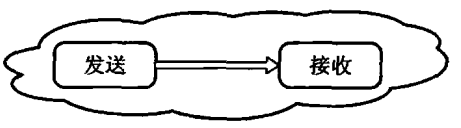

图 1-4 香农的研究对象

试读结束，需要全本PDF请购买 www.ertongbook.com