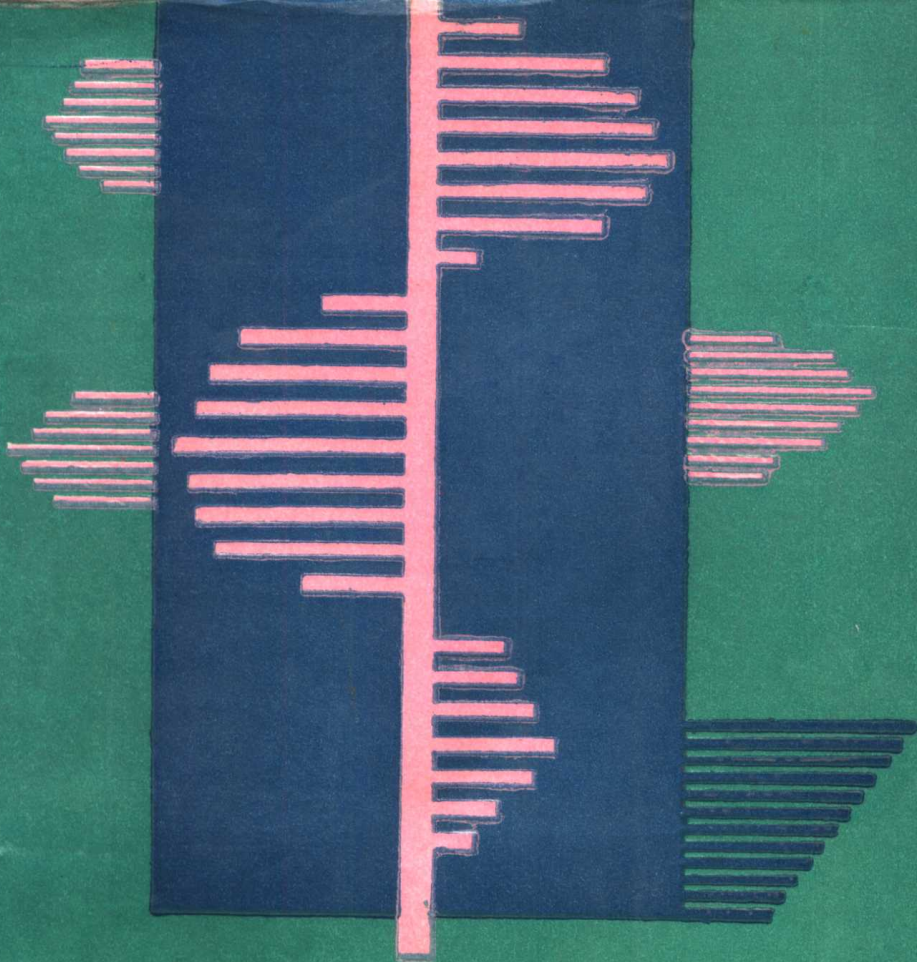




语音与图象的 波形编码原理及 应用

〔美〕N. S. 杰因特 著
〔德〕彼得·诺尔

钱亚生 诸庆麟 译 廖世静 审校

人民邮电出版社

语音与图象的波形编码 原理及应用

【美】 N.S.杰因特 著
【德】 彼得·诺尔

钱亚生 译
诸庆麟

廖世静 审校

人民邮电出版社

内 容 简 介

本书全面地、系统地介绍了波形编码的理论及其在语音与图象信号编码中的应用。它是论述语音和图象波形编码的一本专著。

全书共分十二章：第一章简要地概述了波形编码的基本问题——比特率、编码质量及复杂度；第2章扼要地介绍了语音与图象波形的统计特性及其模型；第3章论述了限带波形的抽样和重建问题；第4章详尽地阐述了量化问题；这是本书的重点；第5章介绍了PCM脉码调制的应用及国际标准的PCM系统；第6章系统地叙述差分脉冲调制DPCM系统，重点阐述了预测理论；第7章讨论噪声反馈编码，以达到改进编码系统主观质量；第8章详细地论述了各种增量调制系统的原理及其特点；第9章介绍了延迟判决编码；第10章讨论了高效图形编码—游程长度编码；第11章介绍了新颖的子带编码原理及其应用；第12章简要地讨论了语音与图象的变换域编码问题。

每章后面都有大量练习题，可帮助读者理解书中所论述的内容。最后还有七个附录，阐述了与书中内容有关的信息论、线性系统、率失真函数、编码器性能的主观度量等知识。

本书可供高等学校有关专业的教师、研究生和大学本科高年级学生作为教材或参考书，也可供从事通信、语音与图象信号处理的科研人员参考。

语音与图象的波形编码原理及应用

〔美〕 N.S.杰因特

著

〔德〕 彼得·诺尔

钱亚生

诸庆麟 译

廖世静 审校

责任编辑：王晓明

人民邮电出版社出版

北京东长安街27号

河北省邮电印刷厂印刷

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

1990年 6 月 第一版

印张：27 页数：432

1990年 6 月河北第1次印刷

字数：718

千字

印数：1-2 000 册

ISBN7-115-03684-5/TN·411

定价：12.00元

DIGITAL CODING OF WAVEFORMS:
Principles and Applications
to Speech and Video

N. S. JAYANT
Bell Laboratories, Inc.
Murray Hill, NJ

PETER NOLL
Technical University of Berlin

PRENTICE-HALL, INC. Englewood Cliffs, New
Jersey 07632

译 者 序

语音信号和图象信号的波形数字编码是一门正在飞速发展的新技术。在最近二十多年内，它取得了重大的进展。在国际学术刊物发表了几千篇有关的学术论文，并出现了许多新型的实用化系统。目前，这项技术正朝着进一步降低波形编码的数码率和实用化的方向不断向前发展。

人类正面临着信息化社会的新时代。现代社会日益迫切需要对电话、电报、传真、可视会议电话、计算机数据以及图象等多种形式的信息进行传输、存储和交换。数字化语音和图象可以与数据一起在综合数字通信网中进行传输和交换。波形编码技术是新一代的综合数字网的重要基础。因此，这项技术对发展我国的数字通信系统将一定会起重要的作用。

本书是第一本全面地、系统地叙述波形编码理论及其应用的专著。它有以下三个特点。(1)书中既介绍了语音的编码问题又同时介绍了图象的编码问题。深入地讨论了波形编码的基本原理和各种算法。虽然已有单独论述语音或图象编码的著作，但同时论述这两个专题的专著还是第一本。(2)书中叙述了语音和图象编码的大量实际应用的例子，包括国际标准化数字通信系统以及300多张图表和曲线，生动形象地反映了波形编码的最新技术成就，如32kb/s·ADPCM、子带编码、噪声反馈编码以及变换域编码等语音与图象编码器。(3)本书的内容安排以及编写方式具有教科书的特点：循序渐进、深入浅出。各章末均附有练习题，全书共有150道习题和应用计算机的专题。书末还附有七个附录，摘录了常用的基本理论公式和编码器的主客观测试方法。

- 千比特这一单位的英文符号在本书中沿用了原版书中的符号。

鉴于本书的以上特点，因此本书可作为从事数字通信、语音与图象信号处理的科研人员以及大专院校师生参考书；也可作为有关专业的研究生、高年级大学生学习数字通信或语音图象信号处理课程的教科书。

本书的二位作者杰因特博士和彼得·诺尔博士曾长期在美国贝尔实验室声学研究室从事语音信号的波形编码与传输的研究工作，曾经发表过几十篇论文并取得了多项专利。书中有不少章节直接反映了作者的研究成果。本书已列入奥本海姆(A. Oppenheim)教授主编的著名数字信号处理丛书之内。它是一本波形编码的重要著作，因此，我们希望本书译本的出版，能够有助于国内有关专业人员了解波形编码这项新技术，并促进这项技术的发展。

本书的序言、第一章至第八章由钱亚生翻译，第九章至第十二章及附录均由诸庆麟翻译。全书翻译过程中得到清华大学冯重熙教授的大力支持和帮助，特此表示衷心感谢。

译者 1985年4月

序 言

波形的数字编码，现在已经不是一个崭新课题了。它已经有四十多年的发展历史，尤其是在最近的二十年内，获得空前规模的巨大进步。在这一时期内，发表的论文达到了几千篇之多。虽然，研究的重点可能产生重要的转移，论文的内容可能由理论性文献转到实用型的方案上，但是，可以预期到，还会继续有大量的学术论文发表。如同其他技术领域一样，可以预期在波形的数字编码的领域内，将会继续产生重大的创新，在降低比特率方面，能取得比本书中所介绍的各种方法更加令人敬佩的成果。尽管如此，可以实事求是地说，我们对于语音和图象编码的理解已经达到相当成熟的程度，这就使得编写一本象本书这种类型的书籍的时机已经成熟了。

正在开发中的新一代的通信和信息网络要求普遍应用数字语音、数字声频和数字图象。这些应用包括了传输、存储、可视电话会议以及许多种信息运算与显示的信号处理。上述各种系统中的基本问题就是符合信号质量等级指标的波形数字化技术。事实上，这就是本书的研究主题。

现在，已经出版了一些论述关于波形编码中若干具体专题的优秀书籍。这些书籍中包括了脉码调制和增量调制的一些开创性的著作。同样也还有一些深受读者欢迎的语音和图象处理的书籍，这些书籍中论述数字编码的专门章节编写得也很出色。但是，据作者所知，本书对波形编码原理和算法的叙述，无论从广度还是从深度来看，都可以称得上是第一本这种类型的书籍。除了本书的一位作者八年前汇编过的一本论文集外，本书也是第一本既论述语音系统又论述图象系统的论著。

由于还没有一本关于波形编码的专著出版过，于是就促使我们

编著了这本书。本书的出版使迄今为止分散在各个不同文献领域内的大量信息能够集中起来，这些文献包括了信息论、通信技术以及信号处理等多种领域。由于本书的二位作者十分熟悉语音编码问题，而对图象编码问题接触得不如语音编码那么多，所以要想在本书中同时阐述语音信号和图象信号的想法并不是一开始就很明确的。因为大多数编码算法的基本原理以及采用统计输入模型来分析编码器的概念；对语音波形和图象波形都是同样适用的，所以在编写本书的开始阶段，很快就明确了要在书中同时论述语音和图象波形的编码问题。实际上，通过引用这二方面的一些实例能够更好地来说明波形编码的原理，所以，本书中便采用了这种方法。

为了达到统一叙述的目的，本书中通常回避了一些只针对一种类型信号的特殊问题。但是，对于一些特别重要的和标准化的系统，如对数PCM语音编码器以及二电平的数字游程长度编码器则是例外。一般说来，对给定信号的编码算法的优化要考虑到给定信号的具体特性以及主观感觉上的具体问题，因而导致要求精确设计的编码结构，这些问题可能不一定在我们叙述的基本原理中加以讨论。因此，如果那些从事研究和设计的实际工作人员要探求上述问题的信息，除了阅读本书的有关章节外，可能还要补充阅读有关语音或图象的专门读物。为了满足这方面读者的要求，每章的最后都列出了一个相当详细的参考文献目录。当然，这个目录不可能包括全部参考文献。

语音的分析——合成声码器是一个完全超出本书范围的课题的实例。但是，本书中叙述的许多基本的信号处理算法同样可以直接应用于上述专门问题的研究中。量化和预测算法是这些基本算法的一些实例。

本书可能是唯一的一本大学的波形编码课程的教科书。本书第一次使编码可以成为完整的一学期的研究生课程，而不是把它只是作为信号处理或通信系统课程的一小部分来进行讲授。书中经过慎重选择的全部习题和应用计算机的练习题，把一些练习性的作业以

及各种基本方法的训练很好地结合起来。题解集可出售给讲授本课程的有关教授。目录中列出的带星号*的章节,表示在讲授课程中可以删去的一些专题或是可讲可不讲的内容。毋需多说,选用本书作为课程教科书的一些教师可以自由选定讲授的范围和重点。还可以用二学期的时间来教授本书。

本书作者已经把本书的部分内容用作一些大学正式课程的基本内容,并用作从事实际工作的有关专业人员的短训班课程和象贝尔实验室那样环境的专业研究人员的在职继续教育课程的基础。1983年初,本书首次在加利福尼亚大学圣达巴巴拉(*Santa Barbara*)分校用于讲授一个40学时的一学期的研究生课程。学习本课程的学生所具备的基础知识应是线性系统、概率论与数理统计的基本原理,以及为完成应用计算机的练习题而要求熟悉的编制程序的基本技能。

内 容 提 要

下面是全书各章的内容摘要。全书中大多数章节附有练习题。应用计算机的练习题单独列为一节。书末共有七个附录。

第一章给出了波形编码问题的定义,并概述了在降低比特率方面已取得的成果。介绍了数字化信号质量的客观度量及主观度量的准则。还讨论了根据三大类编码算法确定的编码器复杂度与编码延迟问题,书中后面各章详细地讨论了采用这些编码算法的一些例子。

第二章叙述了与波形编码有关的语音和图象波形的一些特性。介绍了后面各章中将要引用的定量的统计模型,这些模型启发我们提出各种量化与编码算法,而且还可以根据率失真理论来判定各种算法的性能。在波形模型中采用的统计描述量包括有概率密度函数、自相关函数、功率谱密度以及后面称为谱平坦度的特性。本章中还讨论了语音和图象波形所特有的一些特性,如非平稳性和准周期性等。

第三章阐述了波形的抽样过程。抽样过程是数字编码系统的一

个基本的前提。这一章的大部分内容只是重新概述一些大家已经熟知的关于均匀时间抽样的定理与基本公式。本章还叙述了和语音及图象抽样有关的特殊问题，例如，讨论了具体的抗折叠运算或者预滤波运算的基本原理。在大多数语音和图象信号处理中都已假定抗折叠运算是不言而喻的。

第四章阐述了幅度量化问题。这是全书中篇幅最长的一章，这部分内容足够出版一个单行本。在这章中先详细叙述了均匀量化问题并导出了基本的 6dB/bit 信噪比的关系式，接着叙述了非均匀量化的极其重要的改进方法，非平稳输入的对数压扩和自适应量化问题。还介绍了量化误差的统计特性，它有助于深入理解习惯上假定的量化误差的一些特性，诸如误差概率密度分布的均匀分布性及其功率谱密度白噪声性等。这一章中还给出了变长度熵编码的优点，它使量化器的性能接近于率失真函数的边界值。具有类似的潜在能力的矢量量化器要在后面第九章内再进行介绍。

第五章论述脉码调制（PCM）。脉码调制是最早提出的数字编码的技术，而且这种技术已经被人们所熟知了。在语音与图象信号处理中，8比特PCM编码可以作为其他各种算法的相对比较标准。由于高质量语音通信中普遍采用8比特对数PCM，所以这一章大部分内容专门介绍这种PCM的标准规范。

第六章、第七章与第八章是相互紧密相关的。这三章分别介绍了差分PCM、噪声反馈编码和增量调制。

第六章开始，阐述了以线性预测理论为基础的差分PCM（DPCM），以及作为相关样值序列中去除多余度的度量值——预测增益的概念。一些例子说明了预测增益是和用于编码样值预测中的波形样值数目有关的。这一章的后面几节叙述了具有时变谱的非平稳过程的有效预测所要求的自适应算法，以及根据远间隔样值的预测算法，后者适用于对浊音语音和活动图象等准周期信号的预测。自适应预测是采用每个样值只编成4比特的高质量语音与图象编码系统的基础。采用自适应和基于远间隔样值预测的组合可以达

到降低比特率的最佳效果,而且这种组合可以在数码率低到每个样值只有1至2比特的情况下,进行语音和图象信号的高质量的数字化。

第七章阐述了开环DPCM及噪声反馈编码的概念,并且详细阐明了噪声谱整形问题。如在第六章中讨论过的那样,线性预测非常有效地利用了波形产生过程中固有的多余度。适当地设计DPCM系统中的预测误差量化器同样也可以利用人类感觉器官的一些特性,而且,还能得到在给定比特率时改善主观感觉质量的编码系统。噪声反馈编码为设计编码器提供了更多的自由度,而且使得比较直接地控制重建的噪声谱成为可能。这种噪声谱整形法能够很成功地获得DPCM型编码器在主观感觉质量上的最优化。

第八章介绍了增量调制,它是二电平量化器的DPCM的特殊情况。增量调制是一种最简单的数字编码方式。如同于PCM那样,增量调制在数字传输与信号处理中都已经获得广泛的应用。由于1比特量化器是很粗糙的,所以通常必须对输入波形进行过抽样。因而本章还讨论了预测增益与抽样频率之间的依赖关系。对于常见的一些输入功率谱类型,采用单积分电路或者一阶预测器的增量调制的信噪比为每倍频程下降9dB,即9dB/倍频程。本章在指出了双积分和 $\Delta\Sigma$ 增量调制的优点以后,全面地讨论了非平稳输入的几种自适应量阶算法。尤其是,音节自适应增量调制具有良好的抗传输误码的性能,这一点是众所周知的,本章内详细地介绍了这类通常称为连续可变斜率增量调制(CVSD)的编码器。

第九章阐述了延迟判决编码,这类算法通过某些方法来预测或者说考虑要进行数字化的实际序列的特性。根据定义可知,这种方法的机理会牵涉到明显的编码延迟问题以及搜取最佳输出序列或其中一部分输出的问题,这些输出序列是存储在诸如树型、格点型或者码本中,或者就在其中产生的。以码本为基础的编码算法通常被称为矢量量化器。即使输入是一个相互独立的相同分布的随机变量序列,延迟判决编码也还是有用的,这是这章中一个使人很感兴趣的结论。实际上,对于象伽玛分布那样的随机变量输入信号来

说,若采用普通的标量量化器,则其性能是很差的,而这种方法却提供了能达到率失真函数界值的重要途径,延迟判决编码对于相关序列的编码也是一种十分有用的方法。在这种情况下,有二种方法可供选择。第一种方法是以对相关输入序列直接进行延迟判决编码为基础的;第二种方法是以采用类似于差分编码去除多余度的方法为基础,然后,再对预测误差的波形进行延迟判决编码。语音编码中的一些实例说明了这些方法在数码率为每样值1比特左右的情况下的效率。延迟判决编码同样也为实现分数数码率——如每个样值0.5比特的编码技术奠定了必要的基础。

第十章探讨了一类非常特殊的二值的图形或者传真图形的波形的编码问题。许多重要的二值图象沿着水平扫描线方向上具有由全黑或全白的象素组成的许多长簇游程段。游程长度编码的概念是简单的,是一种对于这类特殊输入波形高效率无误差的编码方法。考虑到对任一给定的游程长度的出现概率是游程长度的一个非线性显函数,这种方法通常要求采用变长度编码。游程长度编码方法同样也能够利用输入图形中的垂直相关性或者说是行间相关性,以便在相应的二维算法中获得较高的编码效率。国际上,已经把几种游程长度编码方法标准化了。因此,这章里较详细地介绍了这些国际标准。一些较高效的游程长度编码算法用来对标准的图形文件进行编码,平均来说,每个象素只要0.1比特。

第十一章与第十二章重新回到有非零误差的多电平输入信号的编码问题。与前面几章的预测编码方法不同,本书的最后二章内,采用了频域或者变换域的方法。

在第十一章的子带编码方法中,语音输入先要分割成若干个相邻的子带信号,然后把各子带信号分别进行编码。由于分割成子带后可以采用可变的比特分配方法,所以对于那些设想为比较重要的子带可以采用较多个量化电平。比特分配可以反映不同子带内语音功率的差异,而且,或者说也可以表示这些子带内出现的噪声的主观感觉上的差别。第十一章的重点放在相对较简单的语音子带编码

上, 其中采用的子带数目比较小, 比特分配算法又是时不变的。

第十二章介绍了变换域编码问题。这一章把子带编码方法推广到更加一般化的情况, 就是把输入信号在频域内分割成非常多的子带, 或者更一般来说, 非常多的变换域分量。实际上, 这些分量就是对输入序列块进行线性运算后求得的变换系数, 输入序列块的长度和所要求的变换系数的数目是相同的。变换域编码的性能随着输入序列的长度增加而得到进一步改善。变换域编码的最佳性能可以证明是和差分或者预测编码系统的最佳性能相等的。但是, 在重建误差序列的主观感觉特性以及传输误码对主观感觉的影响方面, 变换域编码和预测编码是完全不同的。虽然, 实用的变换域编码采用次优化的变换, 但它容易实现。这一章内详细讨论了这些次优化变换及其特性的一些例子。这章的最后几节集中介绍了它们在语音和图象编码中的应用。对于语音编码, 重点介绍了跟踪输入信号时变谱特性的自适应比特分配的算法。对于图象编码, 另一个重点是在二维或三维编码中需要有效的方法。在上述二种情况下, 变换域编码可以在大约每个样值 2 比特的比特率情况下进行非常高质量的数字化, 这一结果是可以和自适应预测编码的最佳情况相比拟的。

目 录

译者序

序言

第一章 引论	(1)
1-1 波形的数字编码	(1)
1-2 主观质量、比特率和编码器的复杂度	(11)
1-3 从信息论得出的理论限值	(25)
第二章 波形的特性	(33)
2-1 概述	(33)
2-2 语音和图象波形的特性	(35)
2-3 随机波形的数学分析	(52)
2-4 具有特殊相关性与功率谱的信源模型	(74)
*第三章 限带波形的抽样和重建	(104)
3-1 概述	(104)
3-2 确定性低通信号	(107)
3-3 带通信号	(116)
3-4 随机过程	(118)
3-5 二维和三维信号	(120)
3-6 抽样引起的误差——折叠效应	(124)
3-7 离散时间误差和连续时间误差	(134)
第四章 量化	(139)
4-1 概述	(139)
4-2 量化误差方差的计算	(143)
4-3 均匀量化	(149)

• 凡是带*号的章、节，如略过不读，将不会影响本书的连贯性。

4-4	非均匀量化	(156)
4-5	对数量化	(171)
*4-6	熵编码的量化	(177)
*4-7	量化误差的统计特性	(191)
*4-8	粗量化器中采用的扰动法	(198)
*4-9	传输误码的影响	(213)
4-10	自适应量化	(228)
第五章	脉冲编码调制(PCM)	(270)
5-1	概述	(270)
5-2	语音、音频和视频波形的PCM编码	(274)
5-3	高质量的语音PCM编码系统	(278)
*5-4	传输误码的影响	(297)
第六章	差分PCM	(307)
6-1	概述	(307)
6-2	一阶、二阶和三阶线性预测器	(317)
6-3	N 阶线性预测器	(325)
6-4	低复杂度的DPCM系统	(341)
6-5	自适应预测	(355)
6-6	对周期信号的远间隔样值的预测	(382)
*6-7	传输误码的影响	(398)
*第七章	噪声反馈编码	(429)
7-1	概述	(429)
7-2	D° PCM	(435)
7-3	噪声反馈编码	(443)
第八章	增量调制	(456)
8-1	概述	(456)
8-2	单积分增量调制的量化噪声	(464)
*8-3	双积分、二阶预测以及总 和增量调制 (ΣDM)	(481)

8-4	自适应增量调制 (ADM)	(490)
•8-5	传输误码的影响	(512)
•第九章	延迟判决编码	(524)
9-1	概述	(524)
9-2	码本、树型及格点型编码器	(528)
9-3	搜索算法	(538)
9-4	无记忆信源的延迟判决编码; $R=1$	(544)
9-5	相关信源的延迟判决编码	(552)
•第十章	游程长度编码	(569)
10-1	概述	(569)
10-2	游程长度编码的原理	(575)
10-3	二值图形的游程长度编码	(582)
10-4	推广到多值输入信号	(587)
第十一章	子带编码	(595)
11-1	概述	(595)
11-2	传输速率 I 、信噪比 SNR 及其 相对于PCM的增益	(600)
11-3	整数子带的滤波器组	(605)
•11-4	正交镜象滤波器组	(607)
11-5	语音的子带编码	(612)
•11-6	传输误码的影响	(619)
第十二章	变换域编码	(625)
12-1	概述	(625)
12-2	$N=2$ 阶的线性变换	(628)
12-3	$N \times N$ 变换	(634)
12-4	最佳比特分配和区域抽样	(641)
•12-5	最佳变换 (KLT)	(656)
12-6	次优变换 (DHT, DWHT, DFT, DCT)	(669)

12-7	语音和图象信号的自适应变换域编码·····	(691)
*12-8	传输误码的影响·····	(707)
	计算机模拟练习题·····	(725)
*附录A	确定性信号与线性系统 ·····	(733)
A-1	富里叶表示·····	(733)
A-2	二维 (2-D) 信号·····	(736)
A-3	线性系统·····	(739)
A-4	连续时间信号和系统·····	(743)
*附录B	矩阵运算 ·····	(747)
*附录C	信源和信道的信息理论 ·····	(754)
C-1	无记忆离散幅度信源·····	(755)
C-2	有记忆离散幅度信源·····	(757)
C-3	离散幅度信源的无失真编码·····	(760)
C-4	无记忆连续幅度信源·····	(763)
C-5	有记忆连续幅度信源·····	(764)
C-6	信道容量·····	(766)
*附录D	失真率函数 ·····	(773)
D-1	概述·····	(773)
D-2	信源熵的计算公式概述·····	(774)
D-3	离散幅度信源·····	(776)
D-4	连续幅度信源·····	(778)
*附录E	编码器性能的客观度量 ·····	(795)
E-1	信号/噪声比 (SNR) ·····	(795)
E-2	分段信噪比 (SNRSEG) ·····	(796)
E-3	频域加权SNR及重建误差·····	(798)
E-4	SNRMAX、SNRMIN和其他 非最小均方误差度量·····	(800)
E-5	信号/非相关噪声的比值 (SNRu) ·····	(802)
*附录F	编码器性能的主观度量 ·····	(805)