



高等学校信息工程类“十二五”规划教材

# 媒体信号编码

◎主编 姚英彪 易志强

◎主审 田 斌

姚英彪  
易志强  
田斌



西安电子科技大学出版社  
<http://www.xduph.com>

**21** 世纪高等学校信息工程类“十二五”规划教材

# 媒体信号编码

主编 姚英彪 易志强  
主审 田 斌

本书受浙江省科协青年科技人才培养工程项目资助

西安电子科技大学出版社

## 内 容 简 介

本书是作者根据近几年的媒体信号编码教学实践及音视频等媒体信号的压缩编码国际标准编写而成的,试图系统地介绍媒体信号编码技术的基本概念、原理和技术标准。

本书共9章,分三大部分:媒体信号编码的基本理论、媒体信号常用编码方法和媒体信号国际编解码标准。第一部分包括前3章内容,主要介绍媒体信号数字化方法、媒体编码系统的评价方法、媒体信源编码基本理论。第二部分包括中间3章内容,主要介绍媒体信号的熵编码、预测编码和变换编码方法。第三部分包括最后3章内容,主要介绍语音、音频、图像和视频的编码理论和典型编解码标准。

本书既可以作为高等院校有关专业(如数字通信、信号与信息处理、计算机应用等)的高年级本科生的教材或教学参考书,也适合于从事多媒体通信、广播电视、消费类电子、媒体信号处理和计算机应用等工作的科技人员参考使用。

★本书配有电子教案,需要者可登录出版社网站,免费下载。

### 图书在版编目(CIP)数据

媒体信号编码/姚英彪,易志强主编. —西安:西安电子科技大学出版社,2011.12

21世纪高等学校信息工程类“十二五”规划教材

ISBN 978-7-5606-2619-2

I. ①媒… II. ①姚… ②易… III. ①多媒体—编码理论—高等学校—教材

IV. ①TP37 ②0157.4

中国版本图书馆CIP数据核字(2011)第124817号

策 划 马乐惠

责任编辑 任倍萱 马乐惠

出版发行 西安电子科技大学出版社(西安市太白南路2号)

电 话 (029)88242885 88201467 邮 编 710071

网 址 www.xduph.com 电子邮箱 xdupfxb001@163.com

经 销 新华书店

印刷单位 陕西光大印务有限责任公司

版 次 2011年12月第1版 2011年12月第1次印刷

开 本 787毫米×1092毫米 1/16 印张18

字 数 421千字

印 数 1~3000册

定 价 30.00元

ISBN 978-7-5606-2619-2/TP·1287

**XDUP 2911001-1**

\*\*\* 如有印装问题可调换 \*\*\*

本社图书封面为激光防伪覆膜,谨防盗版。

# 前 言

近年来,随着计算机、多媒体和网络通信技术的飞速发展,数字音视频技术逐渐代替了模拟音视频技术。然而,对于数字化后的音视频等媒体信号,如果没有有效的压缩编解码方案,海量的数据将给数字媒体信号存储和传输带来巨大的压力,这就促进了各种媒体信号压缩编码算法的发展。

编写本书的目的,是试图系统地介绍媒体信号编码的基本概念、原理和技术标准。本书可供高等院校有关专业(如数字通信、信号与信息处理、计算机应用等)的高年级本科生以及上述专业的教师和相关工程技术人员使用。

考虑到不同层次读者的要求,编写本书时,笔者力图做到以下3点:

(1) 内容紧凑。每章首先给出基本概念和基本理论,然后给出实例和解答,叙述简明扼要,由浅入深。

(2) 覆盖面广。本书对语音、音频、图像和视频编码的基本理论和方法都有所涉及,对与其相关的最新内容也有所提及。

(3) 实践性强。本书简要介绍了应用中有代表性的媒体编解码标准,对实践开发具有参考意义和实用价值。

本书内容分为三大部分。第一部分介绍媒体信号编码的基本理论;第二部分介绍媒体信号编码常用的方法;第三部分介绍代表性的媒体信号国际编解码标准。

第一部分内容包括第1~3章。第1章主要介绍语音、音频、图像和视频等媒体信号编码的必要性和可能性。第2章主要介绍媒体信号数字化方法和媒体编码系统的评价方法。第3章主要介绍信源编码基本理论,包括信源熵、信源无失真/限失真编码定理、率失真理论等。

第二部分内容包括第4~6章。第4章介绍在媒体信号编码中常用的熵保持编码方法。第5章介绍媒体信号编码中使用的预测编码方法,特别是视频图像的块匹配运动估计/运动补偿方法。第6章介绍变换编码方法,主要包括KLT、DCT和小波变换。

第三部分内容包括第7~9章。第7章介绍语音信号的统计特性、常见处理方法和常用的语音编码器。第8章介绍音频的感知编码理论和MP3编解码标准。第9章介绍图像视频的编解码方法和H.264编解码标准。

本书主要由姚英彪副教授和易志强博士编写,由姚英彪同志统稿。西安电子科技大学的田斌教授对本书提出了宝贵的修改意见,在此表示衷心的感谢。书中参考了大量的文献和资料,在此也对所有参考文献和资料的作者深表谢意。

本学科涉及知识面广,学科发展迅猛,限于编者的水平,书中不妥之处在所难免,殷切希望各位读者批评指正。如有疑问,可发送邮件至 yaoyb@hdu.edu.cn。

编 者

2011年3月

# 目 录

|                               |    |
|-------------------------------|----|
| 第 1 章 绪论 .....                | 1  |
| 1.1 数字通信的一般模型 .....           | 1  |
| 1.2 常见的媒体信号 .....             | 3  |
| 1.3 媒体信号编码的必要性 .....          | 3  |
| 1.4 媒体信号压缩编码的分类 .....         | 4  |
| 1.5 媒体信号压缩编码标准 .....          | 6  |
| 习题与思考题 .....                  | 8  |
| 第 2 章 媒体信号分析及编码系统评价 .....     | 9  |
| 2.1 媒体信号的量化 .....             | 9  |
| 2.1.1 媒体信号的数字化 .....          | 9  |
| 2.1.2 量化的基本概念 .....           | 10 |
| 2.1.3 压扩量化 .....              | 12 |
| 2.1.4 矢量量化 .....              | 14 |
| 2.1.5 矢量量化的最优码书设计 .....       | 16 |
| 2.2 媒体信号的数字化 .....            | 17 |
| 2.2.1 声音信号的数字化 .....          | 17 |
| 2.2.2 图像信号的数字化 .....          | 18 |
| 2.2.3 图像的色彩空间 .....           | 19 |
| 2.2.4 计算机常见图像格式 .....         | 20 |
| 2.2.5 视频的 YCbCr 采样格式及制式 ..... | 22 |
| 2.3 媒体编码系统的性能评价 .....         | 23 |
| 2.3.1 编码质量的客观度量 .....         | 24 |
| 2.3.2 编码质量的主观度量 .....         | 26 |
| 2.3.3 其他性能指标 .....            | 31 |
| 2.3.4 媒体编码与通信系统的性能空间 .....    | 32 |
| 习题与思考题 .....                  | 33 |
| 第 3 章 信源编码理论 .....            | 34 |
| 3.1 离散信源的熵 .....              | 34 |
| 3.1.1 自信息量 .....              | 35 |
| 3.1.2 离散信源熵及其性质 .....         | 36 |
| 3.1.3 互信息及其性质 .....           | 40 |
| 3.1.4 有记忆信源的熵 .....           | 42 |
| 3.1.5 信源冗余度 .....             | 43 |
| 3.2 编码的基本概念 .....             | 44 |

|                               |    |
|-------------------------------|----|
| 3.2.1 编码的数学定义 .....           | 44 |
| 3.2.2 常用码的定义及分类 .....         | 45 |
| 3.2.3 平均码长与编码效率 .....         | 46 |
| 3.3 唯一可译码的判断与构造 .....         | 46 |
| 3.3.1 克劳夫特不等式 .....           | 47 |
| 3.3.2 唯一可译码的判断准则 .....        | 47 |
| 3.3.3 唯一可译码的构造 .....          | 48 |
| 3.4 无失真信源编码 .....             | 49 |
| 3.4.1 无失真定长编码定理 .....         | 49 |
| 3.4.2 无失真变长编码定理 .....         | 51 |
| 3.5 率失真函数与限失真信源编码 .....       | 52 |
| 3.5.1 失真函数 .....              | 52 |
| 3.5.2 率失真函数 $R(D)$ .....      | 53 |
| 3.5.3 率失真函数的性质 .....          | 55 |
| 3.5.4 率失真函数的计算及其指导意义与不足 ..... | 56 |
| 3.5.5 限失真编码定理 .....           | 58 |
| 习题与思考题 .....                  | 59 |
| <b>第4章 熵保持编码</b> .....        | 61 |
| 4.1 Huffman 编码 .....          | 61 |
| 4.1.1 Huffman 码的构造 .....      | 61 |
| 4.1.2 截断 Huffman 编码 .....     | 65 |
| 4.1.3 自适应 Huffman 编码 .....    | 67 |
| 4.2 游程编码 .....                | 67 |
| 4.2.1 二值图像的游程编码 .....         | 68 |
| 4.2.2 JPEG 图像量化系数的编码 .....    | 70 |
| 4.3 Golomb 编码与通用变长码 .....     | 74 |
| 4.3.1 一元码 .....               | 74 |
| 4.3.2 Golomb 编码 .....         | 75 |
| 4.3.3 指数 Golomb 码与通用变长码 ..... | 75 |
| 4.4 算术编码 .....                | 78 |
| 4.4.1 算术编码的起源 .....           | 78 |
| 4.4.2 算术编码的基本原理 .....         | 79 |
| 4.4.3 算术编码的码字计算与编码效率 .....    | 83 |
| 4.4.4 二元算术编码的实现 .....         | 83 |
| 4.5 字典码 .....                 | 88 |
| 4.5.1 LZ 码的基本概念 .....         | 88 |
| 4.5.2 LZW 算法 .....            | 90 |
| 习题与思考题 .....                  | 92 |
| <b>第5章 预测编码</b> .....         | 94 |
| 5.1 预测编码的基本原理 .....           | 94 |
| 5.1.1 DPCM 基本原理 .....         | 94 |
| 5.1.2 最佳线性预测 .....            | 95 |
| 5.1.3 常用预测编码方法 .....          | 96 |

|                                 |     |
|---------------------------------|-----|
| 5.2 语音信号的线性预测编码 .....           | 99  |
| 5.2.1 基于语音短时和长时相关性的语音生成模型 ..... | 99  |
| 5.2.2 LPC 声码器 .....             | 100 |
| 5.2.3 特征参数的提取 .....             | 100 |
| 5.3 JPEG 图像无损/近无损预测编码 .....     | 102 |
| 5.4 活动图像的预测编码 .....             | 104 |
| 5.4.1 图像信号的统计特征分析 .....         | 104 |
| 5.4.2 电视图像中典型景物的类型 .....        | 106 |
| 5.4.3 帧间预测编码方法 .....            | 107 |
| 5.5 图像的运动估计/运动补偿 .....          | 109 |
| 5.5.1 块匹配准则 .....               | 109 |
| 5.5.2 块匹配运动估计/运动补偿面临的主要问题 ..... | 110 |
| 5.5.3 全搜索法 .....                | 111 |
| 5.5.4 快速运动估计算法 .....            | 112 |
| 5.6 基于内插预测的图像编解码技术 .....        | 118 |
| 5.6.1 基于内插预测的图像编解码方案概述 .....    | 118 |
| 5.6.2 APT 图像编解码方案 .....         | 119 |
| 5.6.3 APT 的优势及其局限性 .....        | 123 |
| 习题与思考题 .....                    | 124 |
| <b>第 6 章 变换编码</b> .....         | 126 |
| 6.1 变换编码的基本概念和原理 .....          | 126 |
| 6.1.1 变换编码的概念与历史 .....          | 126 |
| 6.1.2 正交变换与正交矩阵 .....           | 127 |
| 6.1.3 正交变换的压缩原理 .....           | 127 |
| 6.2 KLT 编码 .....                | 128 |
| 6.3 DCT 编码 .....                | 130 |
| 6.3.1 DCT 概述 .....              | 130 |
| 6.3.2 DCT 定义 .....              | 131 |
| 6.3.3 快速算法 .....                | 133 |
| 6.3.4 基于 DCT 的图像编码 .....        | 134 |
| 6.3.5 修正离散余弦变换(MDCT) .....      | 138 |
| 6.4 时-频局部化对变换的要求 .....          | 140 |
| 6.5 小波变换 .....                  | 143 |
| 6.5.1 连续小波变换 .....              | 143 |
| 6.5.2 离散小波变换 .....              | 146 |
| 6.5.3 第二代小波变换 .....             | 149 |
| 6.5.4 图像的小波变换 .....             | 151 |
| 6.6 DCT 与小波变换的性能比较 .....        | 156 |
| 习题与思考题 .....                    | 159 |
| <b>第 7 章 语音编码技术</b> .....       | 161 |
| 7.1 语音信号统计特性 .....              | 161 |
| 7.1.1 语音信号的产生 .....             | 161 |
| 7.1.2 语音信号的时域统计特性 .....         | 162 |

|                                               |     |
|-----------------------------------------------|-----|
| 7.1.3 语音信号的频域统计特性 .....                       | 163 |
| 7.2 语音信号处理 .....                              | 163 |
| 7.2.1 语音信号的数字生成模型 .....                       | 163 |
| 7.2.2 语音信号的预处理 .....                          | 165 |
| 7.2.3 短时平均能量、幅度和过零率 .....                     | 165 |
| 7.2.4 短时自相关函数和平均幅度差函数 .....                   | 167 |
| 7.3 语音波形编码器 .....                             | 169 |
| 7.3.1 ADPCM 与 G.721 语音编码器 .....               | 169 |
| 7.3.2 SBC 与 G.722 编码器 .....                   | 172 |
| 7.4 语音参数编码器 .....                             | 173 |
| 7.4.1 参数编码的类型 .....                           | 174 |
| 7.4.2 LPC-10 声码器 .....                        | 174 |
| 7.4.3 LPC-10e 声码器 .....                       | 177 |
| 7.5 语音混合编码器 .....                             | 178 |
| 7.5.1 语音混合编码器的基本原理 .....                      | 178 |
| 7.5.2 码激励线性 CELPC 预测编码 .....                  | 180 |
| 7.5.3 美国 EIA/TIA 标准 8 kb/s 的 VSELPC 声码器 ..... | 181 |
| 7.5.4 G.728 建议的 16 kb/s 的 LD-CELPC 声码器 .....  | 182 |
| 7.5.5 G.729 建议的 8 kb/s 的 CS-ACELP 声码器 .....   | 186 |
| 7.6 变速率语音编码 .....                             | 191 |
| 7.6.1 变速率语音编码的必要性和可能性 .....                   | 191 |
| 7.6.2 变速率语音编码关键技术 .....                       | 192 |
| 7.6.3 常见变速率语音编码器 .....                        | 193 |
| 习题与思考题 .....                                  | 194 |
| <b>第 8 章 音频编码技术</b> .....                     | 195 |
| 8.1 音频编码概述 .....                              | 195 |
| 8.1.1 音频编码技术分类 .....                          | 195 |
| 8.1.2 几种常用的音频编码标准 .....                       | 197 |
| 8.1.3 音质比较 .....                              | 200 |
| 8.2 感知音频编码理论 .....                            | 201 |
| 8.2.1 绝对听觉阈值 .....                            | 201 |
| 8.2.2 关键子带 .....                              | 202 |
| 8.2.3 同时掩蔽 .....                              | 203 |
| 8.2.4 非同时掩蔽 .....                             | 205 |
| 8.2.5 感知熵 .....                               | 205 |
| 8.2.6 心理声学模型 2 .....                          | 206 |
| 8.3 SBR 频带复制技术 .....                          | 208 |
| 8.4 子带编码与滤波器组 .....                           | 209 |
| 8.4.1 子带编码技术 .....                            | 209 |
| 8.4.2 时频分析滤波器组设计原则 .....                      | 211 |
| 8.4.3 音频编码中常见的滤波器组和窗函数 .....                  | 212 |
| 8.4.4 前回声失真和前回声控制 .....                       | 214 |
| 8.5 MP3 音频标准分析 .....                          | 216 |



|                              |            |
|------------------------------|------------|
| 8.5.1 MP3 音频压缩码流的组成 .....    | 216        |
| 8.5.2 MP3 音频压缩码流的比特池技术 ..... | 221        |
| 8.5.3 MP3 音频压缩码流解码流程 .....   | 222        |
| 习题与思考题 .....                 | 224        |
| <b>第9章 图像视频编码技术</b> .....    | <b>225</b> |
| 9.1 视觉生理与心理学 .....           | 225        |
| 9.1.1 失真可察觉门限 .....          | 225        |
| 9.1.2 运动察觉和时间掩盖 .....        | 227        |
| 9.1.3 视觉心理理论运用 .....         | 228        |
| 9.2 视频压缩编码技术 .....           | 229        |
| 9.2.1 视频压缩编码概述 .....         | 229        |
| 9.2.2 视频标准化组织 .....          | 231        |
| 9.2.3 视频编解码标准发展 .....        | 233        |
| 9.3 H.264 视频压缩标准概述 .....     | 237        |
| 9.3.1 编/解码器框图 .....          | 237        |
| 9.3.2 档次与级别、宏块和片 .....       | 238        |
| 9.3.3 技术特点 .....             | 240        |
| 9.4 H.264 视频压缩标准关键模块 .....   | 242        |
| 9.4.1 帧内预测 .....             | 242        |
| 9.4.2 帧间预测与运动补偿 .....        | 246        |
| 9.4.3 整数变换 .....             | 249        |
| 9.4.4 量化 .....               | 253        |
| 9.4.5 CAVLC .....            | 256        |
| 9.4.6 CABAC .....            | 261        |
| 9.4.7 H.264 编码的码率控制 .....    | 266        |
| 9.4.8 去方块滤波 .....            | 269        |
| 习题与思考题 .....                 | 275        |
| <b>参考文献</b> .....            | <b>276</b> |

## 第1章 绪 论

近年来,随着数字通信、计算机、信号处理、微电子相关技术的发展和广泛应用,语音与音频、图像与视频等媒体信号的编解码技术也随之迅速发展,取得了许多突破性成果,极大地促进了多媒体数字通信技术的普及和发展。

本章主要介绍数字通信的一般模型,媒体信号压缩编码的作用、必要性及意义,压缩编码分类、标准及应用等有关知识。通过对这些基本概念和基本知识的讲解,使读者对现代多媒体信号压缩编码有一个初步的了解。

### 1.1 数字通信的一般模型

人类社会是建立在信息交流的基础上的,信息交流是推动人类社会文明、进步与发展的巨大动力。根据香农信息论的定义,信息是事物运动状态或存在方式的不确定性的描述。人们通过对周围世界的观察得到各种信息,信息是抽象的意识或知识。

通信的本质就是信息交流,即通信系统上传输的本质内容是信息。但是为了传输信息,必须在发送端将信息转化为具体的消息。消息是包含信息的语言、文字、数据、图像、符号等。在通信中,消息特指担负着传递信息任务的符号及其序列。消息是具体但非物理的。为了在实际的通信系统中传输消息,就必须把消息加载(调制)到具有某种物理特征的信号上去。也就是说,信号是信息的载体,它是物理性的,如电信号、光信号等。接收端收到载有信息的物理信号后,经过处理变成语言、文字、图像、符号等形式的消息,人们再从中得到有用信息。图1-1是比较完整的数字通信系统的一般模型,现实生活中的数字通信系统基本上都可以用此模型来概括。这个模型将数字通信系统抽象为信源、编码、信道、解码(译码)、信宿等五个模块。

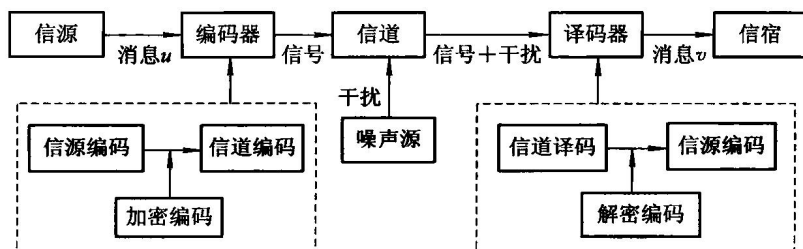


图 1-1 数字通信系统的一般模型

### 1. 信源

信源是信息的来源地，它的作用就是产生信息，但信息并不能直接输出。信源输出的信息是以符号形式出现的具体消息，这些消息承载了信息。信源输出的消息可以有多种形式，但可以归纳为两类：离散消息，例如由字母、文字、数字等各种符号组成的符号序列或单个符号；连续消息，例如语音、视频等在时间上连续的信号。

信源的核心问题是信源符号的概率如何分布，它包含多少信息，用香农信息论的术语来说，就是信源的熵是多少。

### 2. 编码器与译码器

编码是对消息符号进行处理的过程，或者说是把消息变换成信号的过程，而译码就是编码的反变换。根据编码器的实际功能，可将编码器分为三种：信源编码、信道编码和加密编码。

信源编码的作用有两个：一是把信源发出的消息转换成由码元符号（一般为  $m$  进制码元）组成的码元符号序列，形成基带信号；二是通过信源编码来压缩信源的冗余度，减少传输所需的码率，提高通信系统传输消息的有效性（这正是本书要讨论的内容）。信源编码的主要指标就是它的编码效率，即理论上所需的码率与实际达到的码率之比。

信道编码的作用是在信源编码器输出的符号序列上有目的地增加一些监督码元，使之具有检错或纠错的能力，其目的在于提高通信系统传输消息的可靠性。由于信道编码会增加冗余的监督码元，因而会增大码率，这正好与信源编码相反。

加密编码的目的在于隐藏码元符号中的信息内容，防止信号在传输过程中信息被泄露，从而提高通信系统传输信息的安全性。

译码器与编码器一一对应，其作用是对接收到的编码信号（已叠加了干扰）进行反变换，以尽可能准确地恢复原始的信源符号。

### 3. 信道与干扰源

信道是传输消息的通道，或者说是传输物理信号的设施和媒介。在狭义的通信系统中，实际信道有明线、电缆、波导、光纤、无线电波传播空间等等，这些都属于传输电磁波能量的信道。对广义的通信系统来说，信道还可以是其他的传输媒介。信道除了传送信号以外，还有存储信号的作用，如书写通信方式就是一例。信道的主要问题是单位时间内它能够不失真传送多少信息，即信道容量的大小。

信息在处理和传输的每个环节都可能会引入噪声和干扰，使得信号发生畸变。为了分析方便，一般把在系统其他部分产生的干扰和噪声也等效地折合成信道干扰，看成是由一个噪声源产生的，作用于所传输的信号上。这样，信道输出的是已叠加了干扰的信号。由于干扰或噪声往往具有随机性，所以信道的特性一般用概率空间来描述（如输入和输出之间的条件概率矩阵或者条件概率密度函数），而噪声源的统计特性又是划分信道的一个重要依据。

### 4. 信宿

信宿是信息传递的对象，即信息的目的地。信宿一般为接收信息的人或机器，它实际接收的是消息  $v$ ，该消息  $v$  可以与信源发出的消息  $u$  相同，也可以不同。当两者形式不同时， $v$  是  $u$  的一个映射。信宿关心的问题是能够从  $v$  中收到或提取多少有用信息。

## 1.2 常见的媒体信号

图 1-1 所示的数字通信系统传输的主要内容为媒体信号。媒体信号客观地表示了自然界和人类活动中的原始信息,是承载各种信息的载体,是信息的具体表示形式。媒体信号编码属于图 1-1 中的信源编码,其目的主要在于解决媒体信号传输过程中的有效性问题。

常见的媒体信号主要包括下面几种:

(1) 文字(Text): 文字是语言的书写符号系统,是记录语言的书写形式。

(2) 图像(Image): 图像一般指自然界中的客观景物通过某种系统映射的结果,可使人们产生视觉感受,如照片、图片等,一般采用位图形式进行存储。就色彩而言,图像一般分为单色图像、灰度图像和彩色图像三大类。

(3) 图形(Graphic): 图形是指采用某种算法语言或应用软件生成的矢量化图形,一般采用数学方法进行描述,具有体积小、线条圆滑变化等特点。

(4) 音频(Audio): 人类能够听到的所有声音信号的总称,其频率范围在 20 Hz~20 kHz 左右。语音是一类特殊的音频信号,单指由人发声器官发出的各种声音信号。

(5) 视频(Video): 视频信号是动态播放的图像。利用人眼的视觉暂留效应,当每秒钟播放 24 幅以上内容连续变化的静态图像时,人眼无法辨别单幅的静态画面,看上去是平滑连续的动态图像,即视频,如图 1-2 所示。



图 1-2 由连续变化的图像构成的视频

以上这些媒体对象都是现实生活中客观存在的信源,在进行存储和传输时,都牵涉到信源编码(通信专业术语)或数据压缩(计算机专业术语)问题,即如何以最少的数据码率表示信源所发出的信号,减少容纳给定信源消息集合或数据采样集合的信号空间。本书重点讨论音频、图像和视频这三种媒体信号的压缩编码问题,通信系统模型中的其他信源和其他编码问题不在本书讨论范围之内。

## 1.3 媒体信号编码的必要性

近年来,随着计算机、多媒体和网络通信技术的飞速发展,数字媒体技术逐渐代替了模拟媒体技术。例如,数字媒体技术已经在数字影音系统、高清晰度电视、数字音频广播、电话会议系统、无线通信与移动通信、消费类电子、互联网多媒体业务等领域中得到了广泛的应用。然而,数字化后的音视频等媒体信号,如果没有有效的压缩编码方案,海量的数据将给存储和传输带来巨大的压力,这就促进了各种媒体信源压缩编码算法的发展。

在语音、音频、图像和视频等媒体信号实时通信中,语音信号需要的传输带宽无疑是最低的。语音信号的带宽一般在 4 kHz 以下,对应的采样频率为 8 kHz,一般采用 8 位压

扩量化, 原始的未经压缩的数字信号的码率为  $I=8\times 8=64\text{ kb/s}$ 。在当代的数字通信系统中, 这个码率虽然不是很高, 不采用压缩编码技术也可以实现实时通信, 但是采用压缩编码技术后其效率会更高。比如, 在现代的移动通信系统中, 实际的语音信号经过压缩编码后码率一般都低于  $16\text{ kb/s}$ , 甚至更低。这样, 实时传输压缩后的语音信号的信道带宽仅为实时传输原始语音信号信道带宽的  $1/4$ , 甚至更低, 从而节省了信道资源。

从存储角度来看, 一幅  $2032\times 1354$  像素、每像素采用  $8\text{ bit}$  量化的灰度图像, 需要  $2.7\text{ MB}$  左右的存储空间。同样大小的彩色图像, 则需要  $8.1\text{ MB}$  左右的存储空间。一个存储容量为  $1024\text{ MB}$  的数码相机, 仅能拍摄存储  $126$  张未经压缩的  $2032\times 1354$  像素的彩色图片。实际上, 现在的数码相机一般都会采用 JPEG 静态图像压缩技术, 该技术能根据图像内容动态调整压缩比例, 但一般都能达到  $10:1$  的压缩比例。这就是说, 一个存储容量为  $1024\text{ MB}$  的数码相机, 至少能拍摄存储  $1200$  张以上  $2032\times 1354$  大小的彩色图片, 这对大多数应用已经足够。

在语音、音频、视频等媒体信号实时通信中, 高清晰视频(High Definition Television, HDTV)信号需要的传输带宽无疑是最高。HDTV 的图像大小一般为  $1920\times 1080$  个像素, 每秒钟  $60$  帧, 一般按  $4:2:2$  的格式采样亮度信号和色度信号(见 2.2.5 节), 每秒钟的数码率高达  $1.85\text{ Gb/s}$ 。这就是说, 一分钟的未经压缩编码的 HDTV 信号需要高达  $13.9\text{ GB}$  的存储空间, 而一部 HDTV 格式的电影一般都时长  $90$  分钟。可见, 如果没有视频压缩编码技术, 实时传输和存储未经压缩的 HDTV 信号几乎是不可能的。

由此可见, 海量的语音、音频、图像及视频等媒体数字信号, 如果不进行压缩编码, 则无论是传输或是存储这些信号都很困难。因此, 媒体信号压缩编码的作用是不可替代的, 它的社会效益、经济效益也是非常明显的。当媒体信号经过压缩编码后, 不仅其数据量大大减少了, 也带来了一些其他好处:

- (1) 节省媒体信号在各种信道上的传输时间, 使信息交流更加快速, 即时间域压缩;
- (2) 减少媒体信号在实时传输时的带宽要求, 或者增加现有信道可开展的业务, 即频率域压缩;
- (3) 减少媒体信号在通信时的能量消耗, 降低通信设备的发射功率等, 即能量域压缩;
- (4) 减少媒体信号存储时所需的内存空间, 即空间域压缩。

## 1.4 媒体信号压缩编码的分类

### 1. 信号编码的一般模型

媒体信号经过数字化后, 转换成数据信号。媒体信号数据的压缩编码就是以尽量少的数据来表示原始媒体信号的信息, 其一般模型如图 1-3 所示。

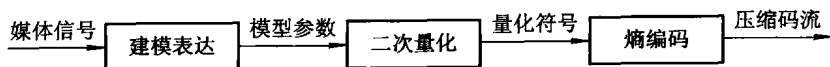


图 1-3 信号编码的一般模型

- (1) 建模表达。根据要解决的问题, 建立一个客观的数学模型, 以便能更有效或更紧

凑地重新表达原始数据信号,减少数据中的冗余,得到与原始数据信号相关但又不同的模型参数数据。

(2) 二次量化。用更简洁的数码表示得到的模型参数数据。对具体的应用来说,由于这些模型参数数据可能用无限(或过高的)表示精度,而实际应用不需要如此高的精度,所以可以重新量化这些模型参数,降低表示这些参数所需要的数据量。由于从模拟信号转换到数字信号时已经经过了一次量化,因此此处称为二次量化。

(3) 熵编码。对模型参数的量化表示或消息流进行码字分配,得到最后的压缩码流。此时一般要求编码后的码流能不失真地反映模型参数的量化符号,即能保持量化后信号的熵,所以称之为熵编码(具体参见第4章)。

在上述三个步骤中:步骤(1)既可以为可逆过程,也可以为不可逆过程;步骤(2)一般都为不可逆过程;步骤(3)一般为可逆过程。可逆或不可逆是指能否由经过处理后的数据无差错地恢复出处理前的数据。如果能无误差地恢复,则为可逆过程;反之则为不可逆过程。

## 2. 信号压缩编码分类

媒体信号的压缩编码方法可根据编码过程中是否采用不可逆过程分为两类,即可逆压缩编码和不可逆压缩编码。可逆压缩编码又叫做冗余度压缩、无失真编码、熵保持编码等,其核心是可以由压缩码流无失真、无差错地恢复出原始码流。不可逆压缩又叫做熵压缩、有失真编码,其核心是不能由压缩码流无差错、无失真地恢复出原始码流。香农在创立信息论时指出,数据是信息与冗余度的组合,熵编码的目的就是要去除数据中的冗余度。实际的媒体信号编码同时采用可逆压缩(无失真)编码技术和不可逆压缩(有失真)编码技术。

设想我们从森林将一批大小不一的圆木(数据)装上一辆卡车,并将其运输(通过信道传输)到木材加工厂。这时,我们的目标是降低运输成本,而主要手段是在同一辆车上装更多的圆木,同时尽量减轻每车圆木的重量(可以看到,装更多圆木和减轻每车圆木重量有时是矛盾的)。在这种情况下,我们判断有无失真的标准是圆木的形状是否发生变化。

通常,我们想多装一些圆木最简单的方法就是有序地摆放圆木。这是因为随机摆放圆木的话,圆木之间的间隙会很大(外在冗余度),有序摆放可以减少间隙(外在冗余度去除),从而可以在同一辆卡车上装载更多的圆木。其次,由于湿的圆木比较重,体积也有所膨胀(存在“内在冗余度”水分),这对降低运输成本都是不利的。因此我们可以将这些木材晾干(内在冗余度去除)后再运输。最后,如果木材加工厂允许我们对圆木进行少量的初加工,比如去掉树皮、将圆木加工成方木、将木材加工成标准长度等,并且这些加工过程中的废料都会被扔掉,则同一辆卡车上必然能装载更多的木材,同时能减轻每车木材的重量。显然,此时圆木的形状已经发生改变,比如从圆木变成方木(有失真压缩),而这种压缩已经“不可逆”,即我们无法再从方木“还原”成原来的圆木。

由这个例子我们可以给出一些压缩编码的基本概念和结论:

- ① 有冗余度就可以压缩(随机摆放圆木有较大空隙,湿的圆木内部有水分);
- ② 压缩只能在一定限度内可逆(圆木形状不发生改变);
- ③ 超过一定的限度,必然带来失真(如去掉圆木树皮、圆木变方木、加工圆木至标准长度等,这些措施都会使得圆木的形状发生改变);
- ④ 允许的失真越大,压缩的比例也越大(极端情况下,圆木因初加工被抛弃,没有圆木需要运输)。

1.5 媒体信号压缩编码标准

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(IUT - T)陆续制定了各种媒体信号的压缩编码标准和建议,使得图 1 - 1 所示的多媒体数字通信系统的发展有了统一的依据,并在通信、计算机、广播电视等各个领域都取得了巨大成功,人们的生活进入一个内容多媒体化、传输网络化和处理计算机化为主要特征的数字多媒体时代。

1. 音频编解码标准

音频编码始于 20 世纪 70 年代的脉冲编码调制(PCM),并按语音编码和音频编码两条路线发展。在音频编码方面,主要关键技术包括:心理声学模型、时频变换、窗切换、时域噪声整形、带宽扩展、立体声编码、空间音频编码,其发展历程如表 1 - 1 所示(详细内容见第 8 章)。在语音编码方面,主要关键技术包括:PCM、ADPCM、SB - ADPCM、线性预测、码激励线性预测(CELP)、代数码激励线性预测(ACELP)等,其发展历程如表 1 - 2 所示(详细内容见第 7 章)。

表 1 - 1 主要的音频编码算法和标准

| 时 间    | 编 码 算 法                              | 体 现 标 准            |
|--------|--------------------------------------|--------------------|
| 1982 年 | PCM                                  | CD                 |
| 1992 年 | 心理声学模型、PQMF 滤波器组、自适应量化、子带编码、时频变换、比特迟 | MPEG - 1 音频层 1/2/3 |
| 1997 年 | 心理声学模型、时频变换、窗切换、时域噪声整形               | MPEG - 2/4 AAC     |
| 1999 年 | 分级编码、多声道立体声编码、频谱复制                   |                    |
| 2004 年 | ACELP/TCX、带宽扩展、参数立体声                 | AMR - WB+          |
| 2006 年 | 空间音频编码(Spatial Audio Coding, SAC)    | MPEG Surround      |

表 1 - 2 主要的语音编码算法和标准

| 时 间    | 编 码 算 法                                                                                           | 体 现 标 准            |
|--------|---------------------------------------------------------------------------------------------------|--------------------|
| 1972 年 | 脉冲编码调制(Pulse-Code Modulation, PCM),每秒采样 8000 次;每次采样为 8 bit 个位,总共 64 kb/s                          | G. 711             |
| 1988 年 | SB - ADPCM 子带-自适应差分脉冲编码(Sub-Band Adaptive Differential Pulse Code Modulation)                     | G. 722             |
| 1988 年 | RPE - LTP,规则脉冲激励-长时预测-线性预测编码(Regular Pulse Excited-Long Term Prediction-Linear Predictive Coding) | GSM                |
| 1991 年 | VSELP 向量和激励线性预测编码(Vector Sum Excited Linear Prediction)                                           | IS - 54,JDC(美数字蜂窝) |
| 1996 年 | CS - ACELP 共轭结构-代数码激励线性预测编码 Conjugate Structure-Algebraic Code Excited Linear Prediction          | G. 729             |
| 2000 年 | 代数码激励线性预测(ACELP)                                                                                  | AMR - WB           |

## 2. 图像与视频编码

图像与视频编码最早出现于 1948 年 Shannon 和他的两个学生 Oliver 与 Pierce 联合发表的对电视信号进行脉冲编码调制的论文中,这标志着数字图像与视频压缩编码技术的开端。1969 年,在美国举行的首届“图像编码会议(Picture Coding Symposium)”表明图像编码以独立的学科跻身于学术界。近半个世纪以来,图像以视频压缩编码技术早已走出实验室,广泛应用于信息社会的各个领域,其典型代表如表 1-3 所示(详细内容见第 9 章)。

表 1-3 主要的媒体信号压缩编码标准及应用

| 标准名称                           | 关键技术                           | 适用范围         | 典型应用                          |
|--------------------------------|--------------------------------|--------------|-------------------------------|
| ISO/IEC 14492<br>(JBIG-2)      | 基于模型的编码、字典搜索、软模块匹配等技术          | 二值图像         | 传真、WWW 图像库                    |
| ISO/IEC 10918<br>(JPEG)        | 变换编码、Huffman 编码、预测编码           | 连续色调<br>静止图像 | 彩色图像、图像编辑、数码相机                |
| ISO/IEC 15444<br>(JPEG 2000)   | 小波变换、面向对象编码、优化截取嵌入块编码 EBCOT    | 连续色调<br>静止图像 | Internet、移动通信、医学图像、数字图书馆、电子商务 |
| ITU-T H.261                    | DCT、自适应量化、Huffman 编码、运动估计与补偿   | 活动图像         | ISDN 视频会议/可视电话                |
| ITU-T H.263                    | H.261 全部技术、双向预测、半像素运动估计、算术编码   | 活动图像         | PSTN 视频会议/可视电话                |
| ISO/IEC 11172<br>(MPEG-1)      | JPEG 全部技术、I/P/B 帧、双向预测、半像素运动估计 | 活动图像<br>及伴音  | 光盘存储、VCD、VOD、视频监控、网络音频 MP3    |
| ITU-T ISO/IEC<br>13818(MPEG-2) | MPEG-1 全部技术、帧/场、质量可扩展的编码、容错编码  | 活动图像<br>及伴音  | DVD、VOD、DVB、DTV/HDTV、AAC      |
| ISO/IEC 14496<br>(MPEG-4)      | MPEG-2 全部技术、对象编码、小波变换、动态网格编码   | 多媒体音<br>像数据  | Internet 与移动通信的视频、交互视频、可视编辑   |
| ITU-T H.264<br>(MPEG-4 Part10) | 分级编码、算术编码、整数变换、帧内预测、可变图像块、多帧参考 | 各种活动<br>图像   | H.261/H.263 和 MPEG-1/2 应用的替代  |
| 中国先进音视频编码<br>AVS               | 与 H.264 类似                     | 活动图像<br>及伴音  | H.264 应用的替代                   |



## 习题与思考题

- 1-1 媒体信号编码的作用是什么？
- 1-2 为什么媒体信号传输或存储时要进行压缩编码？
- 1-3 媒体信号压缩编码技术如何分类？
- 1-4 媒体信号压缩编码标准有哪些？
- 1-5 实际生活中，我们会碰到哪些媒体信号压缩的例子。