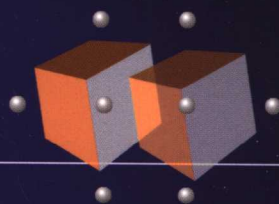




北京市高等教育精品教材立项项目



高等学校电子信息科学与工程类专业教材

数字视频图像处理

全子一 主编

门爱东 杨波 编



电子工业出版社

PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

内 容 简 介

本书不同于一般数字图像处理教程，它基于数字图像压缩编码技术的发展，视频图像处理技术在通信、广播电视及计算机领域突飞猛进的发展与应用（如数字电视，网络视频，移动多媒体通信，可视电话会议等），内容涉及广泛，主要包括生物医学、机器人、遥感以及计算机等领域的数字图像处理。本书根据以上领域视频图像发展的需要，力求系统地揭示这些应用技术原理的数学理论。全书分为8章。为了有助于深入理解课程的数学原理，书后附有大量的习题。

本书可作为通信、广播电视及计算机应用等相关专业高年级本科生和研究生的数字图像处理的教材，也可供相关领域的研究人员及工程技术人员学习参考。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有，侵权必究。

图书在版编目（CIP）数据

数字视频图像处理/全子一主编. —北京：电子工业出版社，2005.5

北京市高等教育精品教材立项项目·高等学校电子信息科学与工程类专业教材

ISBN 7-121-01146-8

I. 数… II. 全… III. 视频信号—数字图像处理—高等学校—教材 IV. TN941.1

中国版本图书馆 CIP 数据核字（2005）第 038509 号

责任编辑：杨丽娟 特约编辑：明足群

印 刷：北京牛山世兴印刷厂

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

经 销：各地新华书店

开 本：787×980 1/16 印张：16.5 字数：361 千字

印 次：2005 年 5 月第 1 次印刷

印 数：4 000 册 定价：26.00 元

凡购买电子工业出版社的图书，如有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系。联系电话：（010）68279077。质量投诉请发邮件至 zltz@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

前 言

我教授数字图像处理课已经 25 年。这门课程究竟应该主要包括什么内容？如何处理理论与实践（应用）的关系？是两大主要问题。

对于第一个问题，国外有不少经典教材，但课程内容各异。有两种不同的图像处理范围界定方法。一种用处理的输入和输出内容都是图像这一特点来界定图像处理的范围。它处理的对象是图形（包括字符）、图像（包括活动图像），特点是处理多维信息。另一种是输入图像，但输出是从这些图像中提取的特征（如边缘、轮廓以及不同物体的标识等），以及模式识别。国外大多数图像处理教材，这两种兼而有之，而且基于数字计算机作为处理手段。它的应用范围包括机器人、生物医学图像处理、天体图像处理、雷达和声纳图像处理、广播电视等。

这 20 多年来，数字图像处理的发展可以说一日千里，特别在通信领域、广播电视领域，以及计算机领域、因特网领域。它们的发展和进入千家万户。在信息化的今天，上述领域的发展和需求，从广度和深度上毫不夸张地说，远远超过其他领域对数字图像处理的应用。

因此，本书内容的定位主要是针对上述发展最活跃领域的要求，处理输入和输出都是图像（包括活动图像）的多维信息。因此，本书定名为《数字视频图像处理》。

视频图像处理主要讨论图像表示（模型）、处理技术和通信（传输）。具体内容是信源的压缩编码、接收解码与显示，以及各种传输技术（包括三网合一的传输）。因篇幅关系，本书着重讨论图像的信源及信宿处理。多维信号处理核心问题是压缩编码。视频图像处理还要解决实时图像处理。因此，它的发展又带动了硬件（大规模集成电路）和软件的大发展。

目前国际上对不同应用，已规范了上十种图像压缩编码国际标准。ISO 及 ITU 国际标准是图像处理原理的最后结果的集合，很少阐述原理本身。作为教材，我觉得应介绍视频图像处理的基本原理和算法，以及构成系统。

本书包括如下 8 章。

第 1 章是数字图像处理基础。它主要介绍人眼及视觉感知特性（因人眼是评价处理图像的最终评判者），建立多维图像（特别是电视图像）与一维信号的关系，图像的傅氏分析和矩阵表示、随机图像处理（包括随机模型）等基本分析工具。

第 2 章是视频取样。它主要研究多维取样与其频谱的关系、取样速率变换。

第 3 章讲述的图像量化是图像处理中失真的主要来源。其主要内容包括最佳量化设计及矢量量化原理。

第 4 章介绍线性预测。在压缩编码中，普遍采用预测技术（特别是双向预测）降低码率。本章介绍它们的原理与算法。

第 5 章是图像正交变换。介绍二维正交变换，包括绝大多数压缩编码标准采用的 DCT 变换和最佳 KLT 变换，以及变换特性。

第 6 章为信源压缩编码。它主要介绍基于 DCT 的信源压缩编码，研究编码的系统部件（变换、量化、熵编码）的算法与设计，研究运动补偿压缩编码的原理。

第 7 章是 MPEG-2 传送系统。它主要研究 TS 码流分组传送，以及它们如何多路复用；还描述了定时与同步的时间戳传送原理。

第 8 章是小波变换及其压缩编码。本章的内容含小波分析和小波图像压缩。这是一个挑战性的课题。傅氏变换适用于平稳信号的谱分析，小波变换还适用于非平稳信号的谱分析。而图像信号是非平稳信号，因此基于小波变换的图像压缩具有比基于 DCT 变换的图像压缩更佳的效果。比如，能量高度的集中，更能匹配人眼视觉特性，易于实现嵌入式码流等，因而压缩效率更高，传输抗干扰性和灵活性好。这在 JPEG-2000 静止图像国际压缩标准，以及 FBI 指纹图像压缩中得到了证明。

数字图像处理是一门用数学较多、有较强的理论基础的学科。因此，本书选编了大量习题，用以深入理解和巩固正文所述的原理和方法。

这本教材成书是以历年讲稿为基础的。本着与时俱进的原则，结合视频图像处理的需要，对原讲稿做了较大的修订，并增加了新的内容。

本书可作为通信、广播电视及计算机应用等相关专业研究生教材，部分章节省略后也可作为高年级本科生教材。

要把一堆讲稿变成一本教科书，需大量的细致工作。首先，感谢门爱东教授在百忙之中对全书初稿的审校。其次，感谢研究生对成书的贡献，如朱江对数字图像基础、图像正交变换及其压缩编码等章的部分文字整理；廖宁对取样、量化、预测、MPEG-2 复接、小波变换及其压缩编码等章的文稿整理；陈秋华对绪论及全书习题的文稿整理；王晶晶对全书习题的校正，以及其他在读研究生所做的贡献，同时还要感谢前几届研究生对形成初稿的贡献。

最后，全书统稿及定稿由我负责，廖宁协助我做了大量具体工作，深表感谢。

金子一

2005 年 3 月于北京邮电大学

目 录

| | |
|-----------------------------------|----|
| 绪论 | 1 |
| 0.1 概述 | 1 |
| 0.1.1 计算机领域 | 1 |
| 0.1.2 通信领域 | 2 |
| 0.1.3 广播领域 | 2 |
| 0.2 声音和图像的数字化的 | 3 |
| 0.2.1 信息量 | 3 |
| 0.2.2 压缩技术 | 4 |
| 0.2.3 多媒体存储 | 6 |
| 0.2.4 多媒体通信 | 7 |
| 0.3 多媒体业务 | 8 |
| 0.4 硬件实现 | 9 |
| 第1章 数字图像基础 | 13 |
| 1.1 人眼的构造 | 14 |
| 1.2 视觉心理学——视觉感知特性 | 15 |
| 1.2.1 光强的感知特性 | 15 |
| 1.2.2 空间频率的感知特性 | 18 |
| 1.2.3 时间的感知特性 | 22 |
| 1.3 黑白视觉的数学模型 | 23 |
| 1.4 二维空间频率与视频时间信号谱 | 24 |
| 1.4.1 二维空间频率 | 24 |
| 1.4.2 视频时间信号及频谱 | 25 |
| 1.5 线性移不变滤波器 | 27 |
| 1.5.1 用 $\delta(m,n)$ 表示 $x(m,n)$ | 27 |
| 1.5.2 线性系统 | 27 |
| 1.5.3 移不变系统 | 28 |
| 1.5.4 线性移不变 (LSI) 系统 | 28 |
| 1.5.5 离散卷积的直接运算 | 28 |
| 1.5.6 离散卷积的快速算法 | 29 |
| 1.6 傅里叶变换 | 30 |

| | | |
|-------------------|-----------------------------|-----------|
| 1.6.1 | 离散信号的展开——傅里叶变换 | 30 |
| 1.6.2 | 二维离散图像信号的傅氏变换及其特性 | 32 |
| 1.7 | 矩阵表示 | 34 |
| 1.7.1 | 二维图像的矩阵向量表示 | 34 |
| 1.7.2 | 卷积的矩阵向量积表示 | 35 |
| 1.7.3 | 图像在计算机中的表示及处理 | 37 |
| 1.8 | 随机过程及其分类 | 38 |
| 1.8.1 | 一维随机过程 | 39 |
| 1.8.2 | 二维随机过程——随机场 | 40 |
| 1.8.3 | 随机过程的统计特性 | 41 |
| 1.8.4 | 用矩阵表示的统计特性 | 45 |
| 1.8.5 | 随机过程的分类 | 45 |
| 1.9 | 随机场的相关模型 | 48 |
| 1.9.1 | 可分离模型 | 48 |
| 1.9.2 | 各向同性模型 | 48 |
| 1.10 | 随机场的线性系统模型——一阶因果模型 | 49 |
| 1.10.1 | 线性均方估值的正交原理 | 50 |
| 1.10.2 | 自回归 AR 模型 (Auto Regressive) | 50 |
| 1.10.3 | AR 模型的应用 | 53 |
| 1.10.4 | 滑动平均 MA 模型 (Moving Average) | 54 |
| 1.10.5 | 自回归滑动平均过程 (ARMA) | 55 |
| 1.11 | 图像的逼真度 | 56 |
| 第 2 章 视频取样 | | 57 |
| 2.1 | 正取样及其频谱 | 58 |
| 2.2 | 从正取样信号恢复原始图像 | 60 |
| 2.3 | 限带信号的二维取样定理 | 62 |
| 2.4 | 随机场取样 | 63 |
| 2.5 | 斜取样及恢复 | 64 |
| 2.6 | 常见取样网格 | 67 |
| 2.7 | 取样和显示的实际限制 | 70 |
| 2.7.1 | 取样孔径 $p_s(x, y)$ | 70 |
| 2.7.2 | 显示孔阑 $p_d(x, y)$ (内插函数) | 71 |
| 2.7.3 | Moire 效应和平坦场响应 | 71 |
| 2.8 | 视频取样率转换的基本原理 | 73 |
| 2.8.1 | 一维信号的内插 | 73 |

| | | |
|------------|-------------------------------|------------|
| 2.8.2 | 一维信号的下取样 | 75 |
| 2.8.3 | 等价结构 | 77 |
| 2.8.4 | 用抽取和内插实现多相分解 | 79 |
| 第3章 | 图像的量化 | 81 |
| 3.1 | 标量量化的原理 | 82 |
| 3.2 | 量化失真的客观度量 | 83 |
| 3.3 | Lloyd-Max 量化器——最佳均方量化器 (MMSE) | 83 |
| 3.3.1 | 求解最佳均方量化器 | 83 |
| 3.3.2 | 最佳均方量化器 (MMSE) 的特性 | 84 |
| 3.4 | 均匀量化器 | 86 |
| 3.4.1 | 定义 | 86 |
| 3.4.2 | 颗粒失真和过载失真 | 87 |
| 3.5 | 高分辨率量化近似解 | 90 |
| 3.6 | 最佳量化器的设计 | 92 |
| 3.6.1 | 最佳化设计的理论基础 | 92 |
| 3.6.2 | 迭代法 | 93 |
| 3.6.3 | 查表法 | 94 |
| 3.7 | 矢量量化 (Vector Quantization) | 95 |
| 3.7.1 | 定义 | 95 |
| 3.7.2 | 矢量量化器原理 | 96 |
| 3.7.3 | 矢量量化的应用范围 | 97 |
| 3.7.4 | 矢量量化的性能测度 | 97 |
| 3.7.5 | 最佳矢量量化器及其设计 | 98 |
| 第4章 | 线性预测 | 101 |
| 4.1 | 估值原理基础 | 102 |
| 4.1.1 | 观察随机矢量预测第二个随机矢量 | 102 |
| 4.1.2 | 最佳线性预测 | 102 |
| 4.2 | 带有限存储器的线性预测器 | 104 |
| 4.3 | 前向预测和后向预测 | 106 |
| 第5章 | 图像正交变换 | 111 |
| 5.1 | 一维离散时间信号的展开与分块正交变换 | 112 |
| 5.1.1 | 一维离散时间信号的展开 | 112 |
| 5.1.2 | 一维分块变换 | 113 |
| 5.2 | 二维正交变换和酉变换 | 114 |

| | | |
|--------------|---------------------------|------------|
| 5.2.1 | 二维正交展开 | 114 |
| 5.2.2 | 可分离酉变换 | 115 |
| 5.2.3 | 二维图像正交分解与基图像 | 116 |
| 5.3 | 酉变换的特性 | 118 |
| 5.3.1 | 能量保持与旋转 | 118 |
| 5.3.2 | 能量集中与去相关作用 | 119 |
| 5.4 | DFT 变换 | 121 |
| 5.5 | DCT 变换 | 122 |
| 5.5.1 | 正交分解 | 122 |
| 5.5.2 | DCT 变换与 DFT 变换的关系 | 124 |
| 5.5.3 | DCT 的快速算法 | 125 |
| 5.5.4 | DCT 变换的能量集中作用 | 126 |
| 5.6 | K-L 变换 | 126 |
| 5.6.1 | K-L 的定义 | 126 |
| 5.6.2 | K-L 的特性 | 127 |
| 第 6 章 | 基于 DCT 的压缩编码 | 129 |
| 6.1 | 相关的信息论基础知识 | 130 |
| 6.1.1 | 信息熵 (Entropy) 和压缩比 | 130 |
| 6.1.2 | 率失真函数 $D(R)$ | 131 |
| 6.2 | 信源压缩方法分类 | 133 |
| 6.3 | 基于线性变换的压缩编码 | 133 |
| 6.3.1 | K-L 变换 | 134 |
| 6.3.2 | 离散余弦变换 (DCT) | 134 |
| 6.3.3 | 带熵编码的标量量化器 | 135 |
| 6.3.4 | 最佳比特分配 | 136 |
| 6.3.5 | 编码增益 G_T | 137 |
| 6.3.6 | 区域编码、门限编码和游程编码 | 138 |
| 6.3.7 | 变长编码 (VLC) | 140 |
| 6.3.8 | 量化的块效应和变换系数加权 | 145 |
| 6.4 | DPCM 编码 | 147 |
| 6.4.1 | 差分量化 | 147 |
| 6.4.2 | 闭环预测量化 | 148 |
| 6.4.3 | DPCM 编/解码器 | 148 |
| 6.4.4 | DPCM 的性能 | 149 |
| 6.4.5 | 预测系数的确定 | 150 |

| | | |
|--------------|--------------------------------------|------------|
| 6.4.6 | 无损预测编码 | 151 |
| 6.5 | 运动补偿预测编码 | 151 |
| 6.5.1 | 时域 DPCM 分析 | 151 |
| 6.5.2 | 运动补偿原理 | 152 |
| 6.5.3 | 块匹配法 (BMA) | 153 |
| 6.5.4 | 全搜索 | 154 |
| 6.6 | 图像编码的国际标准 | 155 |
| 6.6.1 | 压缩编码的预处理 | 155 |
| 6.6.2 | 公共图像格式 | 157 |
| 6.6.3 | 压缩编码国际标准 | 158 |
| 第 7 章 | MPEG-2 传送系统 | 163 |
| 7.1 | TS 码流和 PS 码流 | 164 |
| 7.2 | Buffer 的定时与同步 | 166 |
| 7.3 | PES 包和时间戳 (Timestamp) | 168 |
| 7.4 | 复用/去复用 | 169 |
| 7.5 | 传送流系统目标解码器 T-STD | 171 |
| 7.6 | 传送包的复接调度 | 173 |
| 第 8 章 | 小波变换与图像压缩 | 177 |
| 8.1 | 傅里叶变换的能力及局限性 | 178 |
| 8.2 | 短时傅里叶变换 (STFT) | 182 |
| 8.3 | 连续小波变换 CWT 及离散小波变换 DWT | 186 |
| 8.4 | 分段近似 (Piecewise Approximation) | 189 |
| 8.5 | 多分辨率分析 (MRA) | 192 |
| 8.6 | 信号编码 | 197 |
| 8.7 | 离散小波变换系数的计算 | 201 |
| 8.8 | 基于小波变换的压缩编码 | 212 |
| 8.8.1 | 二维小波变换 | 213 |
| 8.8.2 | 基本的阈值编码方法 | 215 |
| 8.8.3 | EZW 编码 | 217 |
| 8.8.4 | SPIHT 编码 | 224 |
| 习题 | | 229 |
| 第 1 章习题 | | 229 |
| 第 2 章习题 | | 232 |
| 第 3 章习题 | | 235 |

| | |
|------------------------|------------|
| 第 5 章习题 | 239 |
| 第 6 章习题 | 240 |
| 第 8 章习题 | 242 |
| 附录 | 247 |
| 附录 1 光度量单位 | 247 |
| 附录 2 信号与空间 | 247 |
| 附录 3 小波变换基的正交性证明 | 251 |
| 参考文献 | 253 |

绪 论

0.1 概述

随着多媒体技术的发展，多媒体技术能够提供更方便的服务和更容易的使用环境，例如多媒体教学综合系统、多画面互动电视服务、三维家庭影院等。多媒体技术源于通信领域、计算机领域和广播领域中对声音和图像的应用，如图 0.1 所示。在基于文本的传统通信业务中加入声音和图像可使其更生动、更吸引人。因此，计算机、通信和广播领域交融产生的多媒体技术使我们的生活发生巨大的变化。多媒体技术的核心是声音和图像的数字化，因为压缩后数字化的声音和图像可以很容易地利用计算机和数字通信网络传输。目前正在兴起的在因特网上传输数字电视，是发展中的耀眼光环。

然而，由于三个领域的背景和历史存在巨大差异，它们对声音和图像数字化的要求是不一样的。而且，对声音和图像直接数字化产生的文件，其数据量要远远大于传统的基于文本的文件。因此，声音和图像的大容量存储、高速传输网络和压缩技术在多媒体技术中起着重要的作用。另外，要推动多媒体技术的发展，压缩、传输和存储硬件的低成本实现也是必需的。从这个意义上讲，低功耗、低成本的 VLSI 技术也是很重要的。下面对多媒体信号处理技术的背景和它的硬件实现作一下简要的回顾。

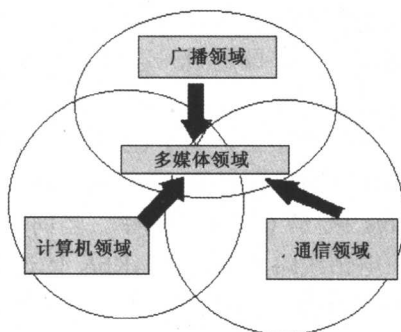


图 0.1 多媒体领域的产生

0.1.1 计算机领域

计算机中的数据一开始时仅是计算用的数值。现在，文档（包括文字、表格和图表等）数据也是计算机中的数据。随着个人计算机（PC）的普及，文档数据也越来越多。例如，如果不用 PC 或工作站中的文字处理软件的话，要出版书籍和文章是十分困难

的。近来，声音和图像也被引入计算机数据中。这意味着可以像用文字处理软件处理文字一样，用软件处理声音和图像。这使得使用“剪切”和“粘贴”来处理声音和图像成为可能，个人可以在 PC 上轻松制作生动的多媒体演示稿。这就是计算机多媒体技术的最初应用。

事实上，近来的高速处理器已经达到了对压缩视频信号进行实时软解压（解码）的水平。尽管高清晰度视频实时软压缩仍然超出微处理器的处理能力，但是随着 VLSI 技术的发展，大型处理器体积不断缩小，成本不断降低，达到了 PC 处理器水平，从而加速了多媒体技术的应用。

0.1.2 通信领域

虽然多媒体技术的技术革命是在计算机领域里出现的，但是多媒体的概念一开始是在通信领域里引入的。在通信领域里，世界各国都建设了很好的数字通信线路，用于 PCM（脉冲编码调制）语音数据和计算机数据的传输。PCM 数字语音编码有巨大的意义，因为通过数字化可以实现高质量语音的远距离传输；因为存在信号衰减、热噪声和串话干扰，模拟传输的质量很难保证。当数字语音通信普及后，自然而然地想到利用语音通信信道传输计算机数据，用于代替建设新的计算机网络。多路复用系统提供很好的灵活性和低成本的数据传输。

在多媒体网络中应注意两点。第一点是在数字语音传输中包括计算机数据，即同时存在实时语音数据和非实时数据。这就是计算机领域、通信领域和广播领域最初的交融，三个领域的边界消失，但是实时和非实时数据的差异带来实时操作的一些困难。这从因特网上传输视频信号就可以看到，实时视频传输对预留带宽的要求是很难满足的。服务质量（QOS）将是下一代多媒体网络的主要问题。

第二点是通信网络的低成本实现，即通过一条线路进行通信、计算机和广播传输，特别是利用一条用户线扩展带宽。若将模拟电话用户线改为数字多媒体线路，这种低成本实现将带来个人的多媒体应用。事实上，由于在声音频带的调制器上使用先进的信号处理技术，传统单用户线的传输速率提高很快。PC 调制解调器（MODEM）可以在电话线路上达到 28.8Kb/s、33.6Kb/s 或 57.6Kb/s 的速率。当向电信运营商申请 ISDN（综合业务数字网）服务后，可以在电话线路上达到 144Kb/s 的速率。而申请开通 ADSL（非对称数字用户线）业务后，用户得到的传输速率进一步提高，可以高达几兆比特每秒，但这依赖于用户线路的质量。要收看到一定质量的视频节目，要求起码有 1Mb/s 的速率。

0.1.3 广播领域

声音和视频是电视广播的主要内容。然而，电视广播信号的数字化是近年才开始的（除了在广播工作站编辑和存储而对信号数字化外）。在 MPEG-2 标准颁布后，广播数字化首先在有线电视和卫星电视系统内得到实现。数字化可以在保证传输质量的前提下

增加有线或卫星电视的频道数，因为 MPEG-2 能够达到很高的压缩比，而且同轴电缆调制解调器的价格能够达到消费者可以接受的水平。由于有线电视网络是树状结构的，模拟有线电视系统对不同距离的用户其质量不同，而数字传输则能够保证网络中所有位置用户的接收质量。由于数字电视广播有许多优点，所以地面电视广播也向数字化方向发展。目前已有美国的 ATSC、欧洲的 DVB-T 和日本的 ISDB-T 三个地面数字电视广播国际标准，我国也正在制定自己的地面数字电视标准。

同轴电缆调制解调器也可提供 VOD（视频点播）业务，但要增加 STB（机顶盒）设备。

0.2 声音和图像的数字化

0.2.1 信息量

声音和图像数字化的根本问题在于它们巨大的信息量，让我们来看一下各种媒体的信息量。一个字母可由一个字节的 ASCII 码表示，那么一页纸若有 50 行，每行 60 个字母，则一页字母需要 3KB，一本 330 页的书需要的存储量为 1MB。这个容量几乎相当于一张 1.44MB 软盘的大小。用于立体声重放的 Hi-Fi（高保真）声音包含两路信号（左和右声道）。在 CD（高密度光盘）里，对每个声道的信号进行频率为 44.1kHz 的取样；在 DAT（数字声音磁带）里，进行频率为 48kHz 的取样。这样的取样频率能保证最高频率为 20kHz 的声音信号的重现。假若每个取样值表示为 16b 数字形式，即每个取样值 2B（字节），则 1 秒钟内立体声信号需要 200KB。这就意味着每 5 秒钟高保真声音信号就产生相当于一本 330 页书的信息量。同样，对于视频信号，NTSC 制电视每秒钟需要处理 30 幅（帧）图像，每幅图像的格式为 720×480 像素，每个像素表示为 24b R/G/B 信号（每个分量 8 比特）或 16b 亮度/色差信号（每像素的亮度和两个色差信号各用 8 比特表示，但每个取样点的亮度都有效，而色差信号是各点有效的），则求得 1 秒 NTSC 信号至少需要 20MB，相当于 20 本上述的书。HDTV 信号每幅图像为 1920×1080 像素，每秒 60 帧。这样的话，每秒钟的信息量达到 240MB。从图 0.2 可以看出，与文本数据相比，声音和图像的数据量要大得多。因此，要像处理文本数据一样来处理声音和图像，压缩技术是必需的。

早在 20 世纪 80 年代起，用于存储或重放数字声音信号的高密度光盘（CD）就出现在消费市场中，但是数字视频信号长期以来只局限在专业领域内。现在，采用 MPEG-2 压缩技术的数字视频光盘（DVD）出现了。一般来说，没有经过压缩的数字视频信号是不如模拟信号经济的。先进的压缩技术、大容量存储技术和高速通信网络的出现，为多媒体技术的实现提供了条件。上述技术都是以数字信号处理（DSP）为基础的。因此，人们对多媒体信号处理及 VLSI 实现非常感兴趣。

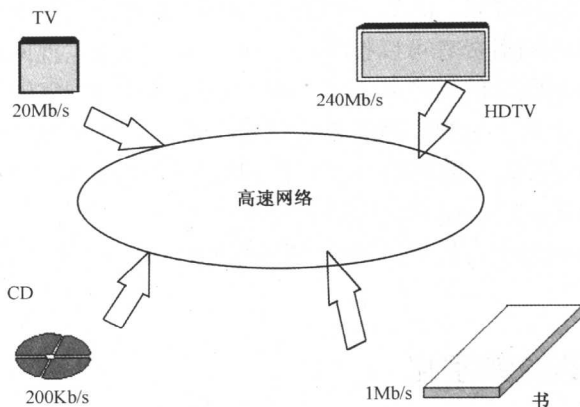


图 0.2 每种媒体所需的信息量

0.2.2 压缩技术

图 0.3 表示计算机领域、通信领域和广播领域交融对图像和声音的要求。由于这三个领域以前一直是独立发展的，而多媒体技术要支持这三个领域。所以，它们之间存在冲突，体现于数字视频格式和功能上。

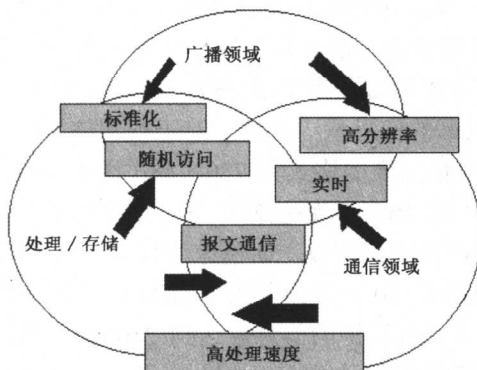


图 0.3 各个领域的要求

20 世纪 60 年代，根据对取样数据的非线性压缩扩展的研究，通信领域首先应用压缩技术进行 PCM 数字语音的传输。同期，NASA 太空卫星在从太空向地球传送月亮表面图片时，应用了图像压缩技术。

其后，在两个电视台之间利用 45Mb/s 的高速 PCM 骨干网传输实时电视节目时，出现了视频压缩技术。对于电视广播来说，最重要的是要尽可能地达到 NTSC 制或 PAL 制的质量。以 PAL 为例，这意味着视频信号每秒有 25 帧图像，每帧图像为 720×576 像素。相应地，电视会议和可视电话系统独立地出现在通信界，因而它们主要考虑通信

费用,而不是图像质量。使用一路 64Kb/s 的 PCM 信道或 2.048Mb/s 的一次复用 PCM 线路传输可视电话或电视会议的费用是可以接受的。因此,在这些应用中,压缩算法着重于降低码率,因而会降低图像分辨率和每秒钟的帧数。ITU-T (国际电信联盟电信标准化组) 建议的国际标准视频压缩算法 H.261 采用的图像格式为 CIF (共同中间格式) 和 QCIF (四分之一 CIF), 它们分别对应于 PAL 制的四分之一和十六分之一分辨率。它们的帧数可以比传统的每秒 30 帧少一些。例如,一个 QCIF 格式每秒 7.5 帧的运动图像可以用 64Kb/s 的速率传输。另一个重要的因素是通信系统要求有短的编码时延。在建立 H.261 标准时,可以清楚地看到通信和电视广播规范的截然不同。ITU-R (国际电信联盟无线通信组) 决定建立一个电视广播标准,因此出现了 G.723 建议。这两个标准的基本压缩算法几乎是一样的:带 MC (运动补偿) 的 DPCM (差分 PCM) 与 DCT (离散余弦变换) 的混合编码方案。

从广播领域的观点来说,广播数字视频节目的根本要求在于压缩技术。从计算机的角度来看,最重要的要求是视频和音频文件的随机访问,因为计算机用户往往想要访问某个特定时间的视频和音频内容,而不是从头开始。很不幸,这种功能没有被考虑进 ITU-T H.261 建议和 ITU-R G.723 建议中。因此,ISO (国际标准化组织) 和 IEC (国际电工委员会) 决定联合制定适用于广播、通信和计算机三个领域的国际标准。基于 H.261 和 G.723 算法,加上扩展功能后得到了 MPEG-1/2 标准,它们称为 MPEG 通用算法。既然是通用算法,则对于不同的应用需要设置几个参数。例如,图像的分辨率可由“级”(level) 选择,算法可由“类”(profile) 选择。

MPEG 的制定具有重要的意义。原来基于计算机的 MPEG-2 标准被 ITU-T 作为 H.262 通信标准。1996 年, MPEG 组织获得广播界 Emmy 奖。这些事情说明 MPEG 将三个领域结合在一起了。因此,研究音频和视频压缩算法、压缩函数(如 DCT)的实现、运动补偿及变长编码(无失真编码)等意义重大。

在多媒体存储系统和多媒体通信网络中,除了音频和视频压缩算法外,编解码纠错也是非常重要的。因为,在存储系统和通信网络中是不可能完全准确的,在压缩数据序列中是存在错误的,尽管错误率非常低。由于数据压缩是必需的,那么,一个比特的错误就可能对解码处理造成很大的损伤。图 0.4 表示压缩和纠错间的关系,压缩部分看做信源编码,纠错部分看做信道编码。因为压缩去除信源的冗余数据,而纠错为了避免由于信道引起的音视频错误,加入了一些冗余信息。纠错应该对由于外部噪声产生的随机错误和由于持续干扰产生的脉冲噪声都有效。

在许多时候,两个不同的纠错编码器以串联的形式出现。第一个是分组纠错码,它只能纠正一个分组内的错误。为了处理的方便,压缩后的数据传输有一定的格式,例如 ATM (Asynchronous Transfer Mode, 异步传输模式) 信元格式、高速网络中的包格式等。分组纠错码只用于有数据格式的情况下。Reed-Solomon 编码是这种编码的一个例子。

在压缩数据格式化和分组纠错编码后,它们被串行地存储或传输。存储系统或传输

系统并不关心它们的内容，它们只是一个简单的比特流。这时，用卷积编码纠正信道错误。由于信息被卷积成数据串，而解码过程是去卷积过程，因而数据变得相当复杂。维特比（Viterbi）解码器是卷积码的高效解码。

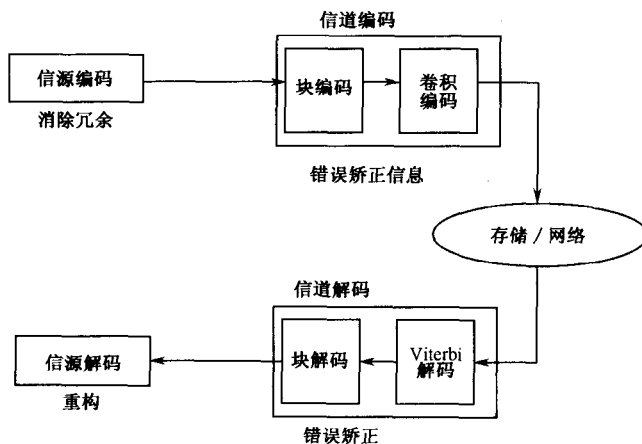


图 0.4 多媒体系统的纠错编解码

在信源标准编码算法中，常常用 DCT（离散余弦变换）变换进行信源编码。这是因为对强相关的图像来说，余弦变换的性能非常接近最佳 KL 变换。但是，一种新的变换——小波变换具有与人眼相似的特性：多分辨率分解。在 MPEG-2 标准化期间，一些研究机构和公司提议用小波变换代替 DCT 变换。虽然为了与 MPEG-1 的过渡，这个建议没有被 MPEG-2 采用，但是其后的 MPEG-4 在处理静止图像上采用了小波变换。

0.2.3 多媒体存储

CD-ROM 或 VCD，以及 DVD（数字视频光盘）等具有电影和视频所需的大存储量，它们将计算机和电视广播领域联系在一起。通过采用激光拾取数据，一个 12cm 直径的 CD-ROM 的存储量达到 780MB。CD-ROM 的码率通常为 1.128Mb/s，这不包括纠错信息的消耗。从 20 世纪 80 年代起，CD-ROM 的读取速度提高到原来的 16 或 32 倍，甚至更高的倍数，但是存储量依然没变。MPEG-1 标准要求 CD-ROM 能够存储一个小时 SIF（标准图像格式，与 CIF 几乎相同，NTSC/PAL 分辨率的四分之一）格式的电视节目，其中视频的码率为 1.15 Mb/s，音频的码率为 128Kb/s。

DVD（原来是数字视频光盘的缩写，现在指数字通用光盘）规范的存储能力提高到 4.7GB，而尺寸和原来的 CD 一样，存储能力提高的原因是激光波长的缩短和基于数字信号处理的精确机械控制。DVD 采用 MPEG-2 算法，用于提供完全与 NTSC 和 PAL 格式分辨率兼容的存储。视频压缩码率的平均值为 4Mb/s，最高值为 9Mb/s。VCD 和 DVD 采用 MPEG 标准是因为它的编码具有准随机访问的能力和快进/退功能。

0.2.4 多媒体通信

当计算机通过局域网（如以太网）连到一起时，计算机系统之间使用的通信方法是包通信。在宽带网络里，引入 ATM 通信，ATM 采用的是和包相似的信元结构。在包通信和信元通信网络中，当负荷很大时，实时通信有时会中断片刻。例如，当网络发生拥塞时，包会被系统自动放到输出缓冲区中，因而不能及时地传送。在 ATM 中，当交换系统缓冲区满后，到来的信元将被自动丢弃。近来，包通信和信元通信的热点在于如何利用因特网（Internet）传输视频信号。这里有三个主要的问题：视频传输带宽预留、连续视频数据的高容量传输和视频准确地重建 30 帧/秒同步。

然而，MPEG 需改善包或信元的丢失问题，MPEG 视频帧结构的准随机访问能力可以有效地防止包或信元丢失产生的错误扩散。另外，在 ATM 网络中，通过加入相位锁定装置，MPEG 传送层可以提供时钟的精确恢复。

在 MPEG-4 里讨论了一种更方便可靠的方法，即对视频对象进行编码处理。每个图像首先分成许多对象，然后对这些对象编码。当网络负荷变大时，只传输最重要的对象信息。

在宽带通信网络中，ATM 被用于骨干网上，骨干网主要用光纤线路传输。但是，光纤直接到达小型办公室或家庭的道路仍然很远，因为费用太高了。因此，将现有的用户线路数字化是一个很好的方法，这就产生了 ISDN（综合业务数字网络）标准。ISDN 提供了 128Kb/s 数据加上 16Kb/s 信息信道。这对于低码率的多媒体终端（例如使用 ITU-T H.263 视频编码和 G.723 语音编码，或使用 MPEG-4）来说是可以使用的，但是，这个码率对于发送和接收 MPEG-1/2 质量的视频信号就太低了。

在现有的用户线上进一步提高码率是可行的，如 xDSL，它使用了正交频分复用多路载波调制解调技术。这种技术克服了用户线上存在的传输损伤，如串话干扰和非平坦频谱特性等，从而达到几兆比特每秒（Mb/s）的高码率。通过这种方法，可以实现 MPEG-1 实时可视对话。

另一种方法是利用有线电视网络的同轴电缆。当音、视频数字化后，有线电视中自然要使用数字调制器。因为数字音、视频信号通常是压缩的，所以数字有线电视与模拟有线电视相比能提供更多的频道。一些新增加的频道可供通信或 WWW（万维网）应用。由于同轴电缆的特性，在同轴电缆调制解调器上使用的是 QAM（正交幅度调制）调制方法，这种方法原先用于数字微波通信或卫星通信上。

数字无线通信，包括微波通信和卫星通信是近来很重要的主题。其中，数字蜂窝系统最受关注。数字蜂窝系统通过很多小型蜂窝单元来覆盖服务区域，使用的是弱电磁波。由于电磁波很弱，因此相同的波长可在不同的蜂窝单元里重复使用而不会影响相邻的蜂窝单元。由于蜂窝单元覆盖范围很小，因此用户终端不需要通过天线发射高功率的电磁波。所以，数字蜂窝系统的终端可以做得很小，以便于携带。在降低呼叫拥塞方面，CDMA（码分多址复用）有一定的优势，因为过多的呼叫只会引起复用信道信噪比