

## 第一章

### 1.1 媒体及媒体分类

**数据定义：**数据是对事实、概念或指令的一种特殊表达形式，这种特殊的表达形式可以用人工的方式或者用自动化的装置进行通信、翻译转换或进行加工处理

数据包含：数字、文字、符号、图形、图像、声音、动画

**信息定义：**信息是对人有用的数据，这些数据将可能影响到人们的行为与决策。

**媒体定义：**指的是用于分发信息和展现信息的手段、方法、工具、设备或装置。

**媒体的分类：**感觉媒体、表现媒体、表示媒体、存储媒体、传输媒体

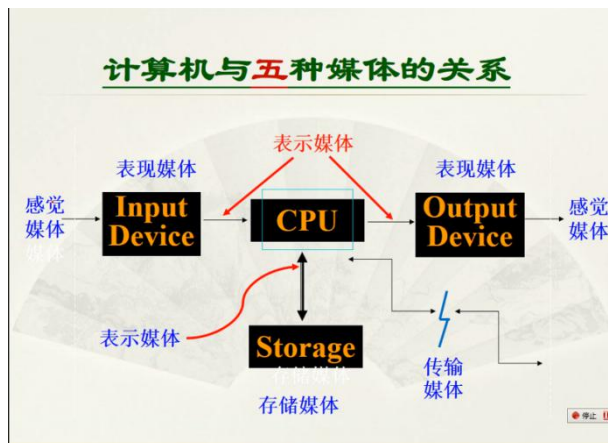
**感觉媒体：**能使人类听觉、视觉、嗅觉、味觉、触觉器官直接产生感觉（感知信息内容）的一类媒体，其中视觉 83%，听觉 11%

**表示媒体：**为了使计算机能有效地加工、处理、存储、传输感觉媒体而在计算机内部采用的特殊表示形式，即声、文、图、活动图像等的二进制编码表示

**存储媒体：**用于存放表示媒体以便计算机随时加工处理的**物理实体**

**表现媒体：**用于把感觉媒体转换成表示媒体、表示媒体转换为感觉媒体的**物理设备**

**传输媒体：**用来将表示媒体从一台计算机传送到另一台计算机的**通信载体**



### 1.2 多媒体及其特征

**多媒体定义：**字面上理解：对各种感觉媒体进行信息处理的技术与方法的一个领域。定义（狭义）：多媒体是利用计算机及其他设备交互处理多种媒体的方法和手段。

**多媒体计算机技术的定义：**计算机综合处理多种媒体信息，使多种信息建立逻辑连接，集成为一个系统并具有交互性。

**多媒体计算机技术的四个主要特性：**多样性、交互性、集成性、实时性

**多样性：**信息载体的多样性，处理设备的多样性，创作形式和工具的多样性

**交互性：**多媒体的交互性是指用户可以与计算机的多种信息媒体进行交互操作从而为用户提供了更加有效地控制和使用信息的手段，交互性是区分是否是多媒体的重要特征

**集成性：**是指以计算机为中心综合处理多种信息的媒体，它包括信息媒体的集成和处理这些媒体的设备的集成（1+1>2）

**实时性：**主要针对声音、活动的视频图像及动画等这些强实时性的信息媒体

## 第二章

### 2.1 声音的基本特性

**音频信号的物理参数:** 频率、振幅、相位

**单音:** 振幅和频率不变的声音信号

**复音:** 声音信号由不同的振幅与频率合成而得到的

复音中的最低频率称为复音的**基频** (**基音**)

复音中还存在一些其他频率, 是复音中的次要成分, 通常称为**谐音**

**人耳可听域:** 20~20,000Hz

**声音三要素:** 音调、音色、响度

**影响音质的因素:** 频带宽度、动态范围、信噪比

**音频信号数字化的过程:** 采样、量化、编码

采样频率与声音频率的关系: 根据 Nyquist 理论, 只有采样频率高于声音信号最高频率的**两倍**时, 才能不失真。

采样量化参数与文件大小: **文件大小=采样频率\*(量化精度/8)\*声道数\*时间/8** 目的是为了将字转换为字节

数字音频文件格式: .wav .cda .mp3 .RAM .WMA

### 2.2 电子合成音乐——MIDI

**乐器数字接口标准 MIDI,** 是规定计算机音乐程序、电子合成器和其他电子设备之间交换信息与控制信号的**通信标准**。

**MIDI 文件:** MIDI 文件本身不是声音文件, 而是指令, 通俗的讲是乐谱的数字描述。

**MIDI 端口:** MIDI IN: 接受来自其他 MIDI 设备的 MIDI 信息。

MIDI OUT: 发送本设备生成的 MIDI 信息到其他设备

MIDI Thru: 将从 MIDI In 端口传来的信息转发到相连的另一台 MIDI 设备上

MIDI 合成器在 PC 的声卡上

FM 合成: 频率调制合成 波表合成: 乐音样本合成

**MIDI 文件的特点:** 用乐谱指令代替声音文件 数据量小 编辑灵活 MIDI 音质与音源质量有很大关系 MIDI 在音质上还不能与真正的乐器完全相似, 无法模拟自然界中其他非乐器类声音 MIDI 音乐系统由计算机、带有 MIDI 接口的声音卡、具有 MIDI 输入和输出端口的键盘、用来输出键盘声音的音响设备(音箱和放大器)和音序软件等软、硬件组成。

### 2.3 语音合成与语音识别

**音频信号处理的过程:** 输入信号-编码器-传输、存储-解码器-输出信号

**波形编译码器:** 不利用生成话音信号的任何知识而企图产生一种重构信号, 它的波形与原始话音波形尽可能地一致。

**模型(音源)编译码器:** 企图从话音波形信号中提取生成话音的参数, 使用这些参数通过话音生成模型重构出话音。

**混合编译码:** 企图填补波形编译码和音源编译码之间的间隔 (**Abs 编译码器**)

**语音识别系统**主要包含: 特征提取、声学模型、语言模型以及字典与解码四大部分

计算机文语转换系统简称 TTS 系统

## 第三章

### 3.1 色彩的原理

电脑用的颜色模型是 RGB（红绿蓝）

**混色规律：**减性混合：参与搭配的颜色越多，明度越暗，三种原色混合为黑色

加光混合：参与搭配的颜色越多，明度越亮，三种原色混合为白色

中性混合：实质上并没有进行颜色的混合，而是由于眼睛的视觉特征造成的

**色彩三要素：**亮（明）度、色相、饱和度

**真彩色：**指一幅彩色图像的每个像素的 R、G、B 三个分量直接决定显示设备的基色强度。RGB 8:8:8 也称为“全彩色”。真彩色颜色深度 24 位。

**伪彩色：**伪彩色图像的每个像素的颜色不是由其各个基色分量的数值直接决定，而是通过一个像素值——彩色对照表，查找出其对应的 R、G、B 值，由此产生的彩色，称为“伪彩色”。

### 3.2 图形与图像的基本属性和类型

**矢量图形**的特点：图形的清晰度与分辨率无关。容易实施各种变换（移动、旋转、放大、缩小、扭曲）而不变形。存储数据量小。不适合表现人物、风景等复杂细节的景物

**位图图像**的特点：图像的清晰度与分辨率有关。不易进行变换和修改。存储数据量大。可直接、快速地在计算机屏幕上显示。适合表现场景复杂，有丰富层次和颜色细节的画面。

**图形：**用数学公式或模型来描述

**图像：**是自然界中多姿多彩的景物和生物，通过视觉感官在大脑中留下的印记。图像由基本显示单元“像素”构成。像点由若干个二进制位进行描述，二进制位代表图像颜色的数量，二进制位与图像之间存在严格的“位映射”关系，具有位映射关系的图叫作“位图”。

**图像的属性：**图像分辨率：分辨率的单位：dpi（每英寸显示的像点数）

**颜色深度：**表示一个像素的二进制位数，以 bit 为单位。

图像的获取过程：采样、分色、取样、量化、编码

彩色图像的像素通常由 3 个彩色分量组成，假设 3 个分量分别用 n, m, k 个二进位表示，则可表示  $2^{n+m+k}$  次方种不同的颜色。灰度图像和黑白图像的像素只包含 1 个亮度分量。

**位图图像数据量：**垂直方向的分辨率\*水平方向的分辨率\*颜色深度/8（byte）

**图像常用后缀名：**.jpg .gif .bmp .png

**视觉暂留现象：**当被观察的物体消失后，影像仍在大脑中停留一段时间，这种现象被称为视觉暂留现象。

**动画：**动画由多副连续画面组成，当画面快速、连续地播放时，由于人类眼睛存在“视觉暂留效应”而产生动感。

全动画：24 帧/s 半动画：6 帧/s 电影中动画的播放速度是 24 帧/秒

计算机动画的类型：从动画生成的机制：**帧动画和造型动画。**

动画的技术参数：帧速度、数据量、图像质量

### 3.3 视频信号及数字化

电视信号为 25 帧/s 的 PAL 制（中国、西欧）或 30 帧/s 的 NTSC（美国、日本）。

隔行扫描：每一帧由两次扫描完成，奇数场和偶数场。逐行扫描：每一帧由一次扫描完成。

YUV 模型: Y 表示光源的亮度, U 和 V 表示色差的两个分量, 红色分量和蓝色分量。PAL 彩色电视制式中采用 YUV 模型来表示彩色图像。

视频信号的采样: 要满足奈奎斯特采样定理。采样频率必须是行频的整数倍。要满足两种扫描制式。

动画视频文件类型: .avi .wmv .mov .ram

## 第四章

### 4.1 多媒体数据压缩技术概述

压缩编码的理论基础是**信息论**。

信息量增加, 数据量增加。数据量增加, 信息量不一定增加。

信息论认为: 若信源编码的熵大于信源的实际熵, 该信源中一定存在冗余度。

从信息论的角度看, 压缩就是去掉信息中的冗余, 即保留不确定的信息, 去除确定的信息(可推知的), 也就是用一种更接近信息本质的描述来代替原有冗余的描述。这个本质的东西就是**信息量(即不确定因素)**。

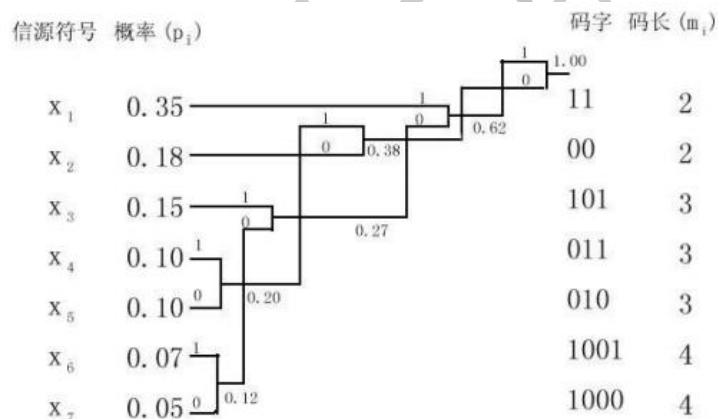
数据冗余的类型: 空间冗余、时间冗余、信息熵冗余、视觉冗余、听觉冗余、结构冗余、知识冗余。

**衡量压缩技术的指标:** 压缩比(输入数据和输出数据比), 压缩和解压速度, 还原的效果; **压缩方法:** 无损压缩和有损压缩, 软硬件开销。

### 4.2 数据压缩的基本原理和方法

**统计编码:** 统计编码属于**无损编码**。主要有哈夫曼编码、算术编码。

**哈夫曼编码**不唯一, 没有错误保护功能。需要知道先验概率。



**平均码长:**

$$R = \sum_{i=1}^n m_i \cdot p_i$$

$$= 2(P_1 + P_2) + 3(P_3 + P_4 + P_5) + 4(P_6 + P_7)$$

$$= 2(0.35 + 0.18) + 3(0.15 + 0.10 + 0.10) + 4(0.07 + 0.05)$$

$$= 2.59$$

**算术编码**也是一种对错误敏感的编码方法。例题 BAD。

**RLE 编码**：游程编码。只存储一个像素的颜色值以及具有相同颜色值的像素个数或具有相同颜色值的行数。

**词典编码：无损编码**。依据数据本身包含有重复代码的特性。

第二类词典编码包含一个编码词典。这种编码方法也被称为 LZW 编码。

LZ77 算法属于第一类编码思想。

## LZ77算法

位置	字符	步骤	位置	匹配串	字符	输出
1	A <sub>1</sub>	1	1	—	A <sub>1</sub>	(0,0)A <sub>1</sub>
2	A <sub>2</sub>	2	2	A <sub>1</sub>	B	(1,1)B <sub>3</sub>
3	B <sub>3</sub>	3	4	—	C <sub>4</sub>	(0,0)C <sub>4</sub>
4	C	4	5	B <sub>3</sub>	B <sub>6</sub>	(2,1)B <sub>6</sub>
5	B	5	7	AB	C <sub>9</sub>	(5,2)C
6	B					
7	A					
8	B					
9	C					

**预测编码**：根据离散信号之间存在一定关联性的特点，利用前面的一个或多个信号预测下一个信号，然后对实际值与预测值的差进行编码。属于**有损编码**。

预测编码中典型的压缩方法：**增量调制**（一般增量调制（DM）、自适应增量调制（ADM）），**差分脉冲调制（DPCM）**，**自适应差分脉码调制（ADPCM）**。

预测编码分为线性预测编码和非线性预测编码两类。DPCM 和 ADPCM 都属于线性预测编码。

**例**：若 DPCM 系统的输入为：

{0 1 2 1 1 2 3 3 4 4 .....}

则编码过程如下：

$x(n)$	0	1	2	1	1	2	3	3	4	4	...
$\tilde{x}(n)$	0	0	1	2	1	1	2	3	3	4	...
$d(n)$	0	1	1	-1	0	1	1	0	1	0	...

**变换编码**：利用图像块内像素值之间的相关性，对数据进行某种形式的正交变换，从一种信号空间变换到另一种信号空间（变换域或频域），使变换后的数据的幅度范围比原先的小，然后对变换后的数据进行编码，达到压缩目的。**变换本身不进行数据压缩**。

常用的变换编码有：KLT、DCT（**离散余弦变换编码**（经过正向 DCT 变换，矩阵元素两个足码之和较小者（即矩阵左上角部分）代表低频成分，大者（即矩阵右下角部分）代表高频成分。由于大部分图像区域中相邻像素的变化很小，所以大部分图像信号的能量都集中在低频成分，高频成分中可能有不少数值为 0 或接近 0 值））、WHT、DFT

DCT（离散余弦变换编码）属于**有损编码**。

#### 4.3 图像数据的压缩标准

静止图像的压缩标准: JPEG

动态图像的压缩标准: MPEG

JPEG 包括两种基本的压缩算法: 采用**离散余弦变换 (DCT)**为基础的有损压缩算法与以预测技术为基础的无损压缩算法(**哈夫曼编码**)。

MPEG 压缩算法采用了两个基本技术: 一是在空间方向上, 图像数据压缩采用 JPEG 压缩算法技术, 减少空间冗余。二是在时间方向上, 采用基于 16\*16 宏块的运动补偿来减少时间冗余。

**MPEG-1 VCD mp3** 为了减少时间冗余量, MPEG 将 1/30 秒时间间隔的帧序列电视图像, 以三种类型的图像格式表示: **内码帧 (I)** (不需要参考其他画面、独立进行压缩编码的画面)、**预测帧 (P)** (需要参考前面已编码的 I 画面和 P 画面进行预测编码的画面) 和**插补帧 (B)** (既参考前面的 I 画面和 P 画面、有参考后面的 I 画面和 P 画面进行双向预测编码的画面), **D 帧** (仅使用自身信息进行编码、用作正向快速搜索的画面)。

**运动补偿的两种算法:** 运动补偿预测法, 运动补偿插补法。

**MPEG-2** 主要用于 DVD 及数字电视

**MPEG-4 mp4** 是以**内容为中心**的描述方法。在信息描述中首次采用了**对象**的概念。基于内容的视频编码过程可由三步完成: **VO 的形成、编码、复合**

**MPEG-7 多媒体内容描述界面**, 目标是扩展现在有限的查询能力, 包括静止的画面、图形、声音、运动视频以及上述元素如何结合成多媒体信息(“剧情”、合成信息), 还包括上述一般形式中较为特殊的情况, 如面部表情、人物特点等。对信息的描述方法不依赖于材料的表示形式。

## 第五章

传统计算机与多媒体计算机区别: 在传统计算机的基础上, 加上了处理视频和音频的设备, 具有多媒体功能的个人计算机。

**MMX: 多媒体扩展。**

**MMX 的主要技术特点:** 新的紧缩数据类型 饱和型运算方式 新增 57 个 MMX 指令 单指令多数据技术 (SIMD) 与 IA 结构的全兼容性

**声卡工作原理:** 声卡由声源获取声音, 并进行模拟/数字转换或压缩, 而后存入计算机中进行处理。声卡还可以把经过计算机处理的数字化声音通过解压缩、数字/模拟转换后, 送到输出设备进行播放或录制。

**声卡的主要功能**

- 1、录制、回放数字声音文件
- 2、压缩和解压声音数据
- 3、语音合成技术
- 4、语音识别
- 5、音量控制
- 6、具有 MIDI 接口
- 7、CD-ROM 接口
- 8、游戏接口
- 9、支持全双工功能

**视频采集卡工作原理:** 视频信号源、摄像机、录像机或激光视盘的信号首先经过



A/D 变换, 送到多制式数字解码器进行解码得到 YUV 数据, 然后由视频窗口控制器对其进行剪裁, 改变比例后存入帧存储器。帧存储器的内容在窗口控制器的控制下, 与 VGA 同步信号或视频编码器的同步信号同步, 再送到 D/A 变换器变成模拟的 RGB 信号, 同时送到数字式视频编辑器进行视频编码, 最后输出到 VGA 监视器及电视机或录像机。

#### 视频卡的主要功能简介

- 全活动数字图像的显示、抓取、录制、支持 Microsoft Video for Windows;
- 可以从录像机 (VCR)、摄像机、ID、IV 等视频源中抓取定格, 存储输出图像;
- 近似真彩色 YUV 格式图像缓冲区, 并可将缓冲区映射到高端内存;
- 可按比例缩放、剪切、移动、扫描视频图像;
- 色度、饱和度、亮度、对比度及 R、G、B 三色比例可调;
- 可用软件选端口地址和 IRQ;
- 具有若干个可用软件相互切换的视频输入源, 以其中一个做活动显示。

**扫描仪工作原理:** 以平面式扫描仪为例, 把原件面朝下放在扫描仪的玻璃台上, 扫描仪内发出光照射原件, 反射光线经一组平面镜和透镜导向后, 照射到 CCD 的光敏器件上。来自 CCD 的电量送到模数转换器中, 将电压转换成代表每个像素色调或颜色的数字值。步进电机驱动扫描头沿平台作微增量运动, 每移动一步, 即获得一行像素值。扫描彩色图像时分别用红、绿、蓝滤色镜捕捉各自的灰度图像, 然后把它们组合成 RGB 图像。

**CCD:** CCD 是数字照相机的成像部件, 可以将照射于其上的光信号转变为电压信号。CCD 芯片上的每一个光敏元件对应将来生成的图像的一个像素, CCD 芯片上光敏元件的密度决定了最终成像的分辨率。

**触摸屏的种类:** 红外线触摸屏 电阻触摸屏 电容式触摸屏 表面声波触摸屏 近场成像触摸屏

#### 5.4 光存储技术

Philips 首先将激光的技术用于计算机存储技术。

**CD-ROM 的光道结构特点:** 光道为螺旋形。光道上各处的存储密度一致, 光盘转速与光头所处位置的半径成反比, 保证线速度恒定。

**CD-ROM 信息的记录:** 突变处代表“1”, 凹坑和凸区的长度代表“0”的个数。

**EFM 码:** 8 到 14 调制编码, 合并码: 000

**CD-ROM 驱动器的性能指标:** 容量, 平均存取时间, 数据传输率 (突发传输率、持续传输率 (单速 150KB/s))

DVD 单面单层: 4.7GB

### 4、DVD与VCD的技术比较

	VCD	DVD
图像压缩标准	MPEG-1	MPEG-2
图像传输率	1.15M bit/s	1-10M bit/s
分辨率	NTSC	主等级 720*480
	PAL	主等级 720*576
画面比	4:3	4:3 / 16:9
字幕	不可选	32 种语言可选

## 第六章

多媒体人机界面设计的原则

面向用户的原则

一致性原则

简洁性原则

适当性原则

顺序性原则

结构性原则

合理选择文本和图形

使用多窗口