多媒体技术

Multimedia (Y)
Technology

一、多媒体基本概念

河南大学计算机与信息科学学院 刘扬

2017/11/28 上午11时7分9秒

目录



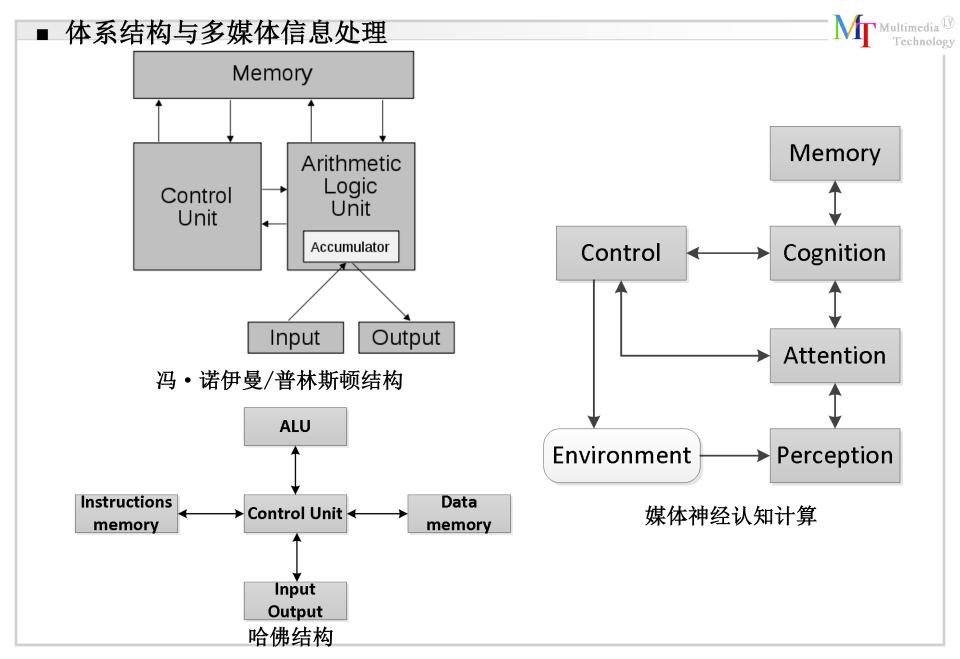
- ■前言
- 1、数据、信息与媒体
- 2、媒体的分类
- 3、多媒体技术的概念
- 4、多媒体技术的特性
- 5、多媒体的关键技术
- 6、多媒体技术的行业应用
- 7、多媒体技术国内外现状和趋势

前言



- 如果说Von Neumann体系结构的计算机是模拟"大脑"的信息运算功能,则多媒体技术的目的是解决计算机接受信息的"眼睛和耳朵"和表达信息"的手和喉舌",其本质是智能计算的I/0多媒体系统涉及的领域非常的宽:
 - 信息技术:信息采集(语音输入、图像输入)、信息传输 (多媒体通信)、信息存储(多媒体数据库、大容量存储 器)、信息处理(多媒体计算、情感计算、机器学习、图 像识别、语音识别)、信息再现(语音合成、图像生成)
 - 文学艺术: 认知心理学、语言学、绘画、音乐、影视录音出版等
 - · 多媒体仅仅是个开端, 高级多媒体技术将具备AI性能。

前言



1. 数据、信息与媒体



- 数据是为了在计算机中对 事实、概念或指令进行描述的一种 特殊格式,这种特殊的表达形式(格式)可以用计算机及其相 关设备自动地进行传输、翻译(转换)或加工处理。
- Data is a special format for describing fact, concept, or instruction in computers. This format could be transmitted, translated, or processed automatically by computer and related devices. (ISO)
- 数据有一定的格式,数据格式的规定就是数据的语法.
- 数据有一定的含义,数据的含义就是它的语义.
- 通过解释、推论、归纳、分析、综合等方法,从数据所获得 的有意义的内容称为信息。

数据、信息与媒体



- 媒体又称媒介、媒质,指的是用于表示、存储、分发、传输和展现数据(信息)的手段、方法、工具、设备或装置。媒体是信息的载体。具体在计算机科学中有两种意义:
 - 存储信息的实体 (如磁介质存储器,半导体存储器,光学存储器)
 - 信息的载体 (如数字、文字、声音、图形、图像)
- In general, one describes medium as a means for representation, storage, distribution, transmission and presentation of data.

2. 媒体的分类



- 5 classes of media (CCITT)
 - perception medium (感觉媒体)
 - representation medium(表示媒体)
 - storage medium (存储媒体)
 - presentation medium (展现媒体)
 - transmission medium(传输媒体)

媒体的分类



- ①感觉媒体,即能使人的听觉、视觉、嗅觉、味觉和触觉器 官产生感觉而感知信息内容的一类媒体,如声音、文字、图 画、气味等,它们是人类有效使用信息的形式。
- ②表示媒体,为了使计算机能有效地处理、存储和传输感觉 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。
- ③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,如磁盘、光盘、半导体存储器等。
- ④展现媒体,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,前者是计算机的输入设备,后者是计算机的输出设备。
- ⑤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,如电话线、双绞线、同轴电缆、光缆、无线电、微波和红外线等。

媒体与多媒体



- 感觉媒体是人与人、人与计算机交换信息所使用的媒体;
- 表示媒体是计算机内部、计算机与计算机之间交换信息所使用的媒体;
- 多媒体技术所指的媒体是:感觉媒体及其相应的表示媒体。
- 表示媒体:

声: 语音《声响《音乐等》

图:图像、图形等

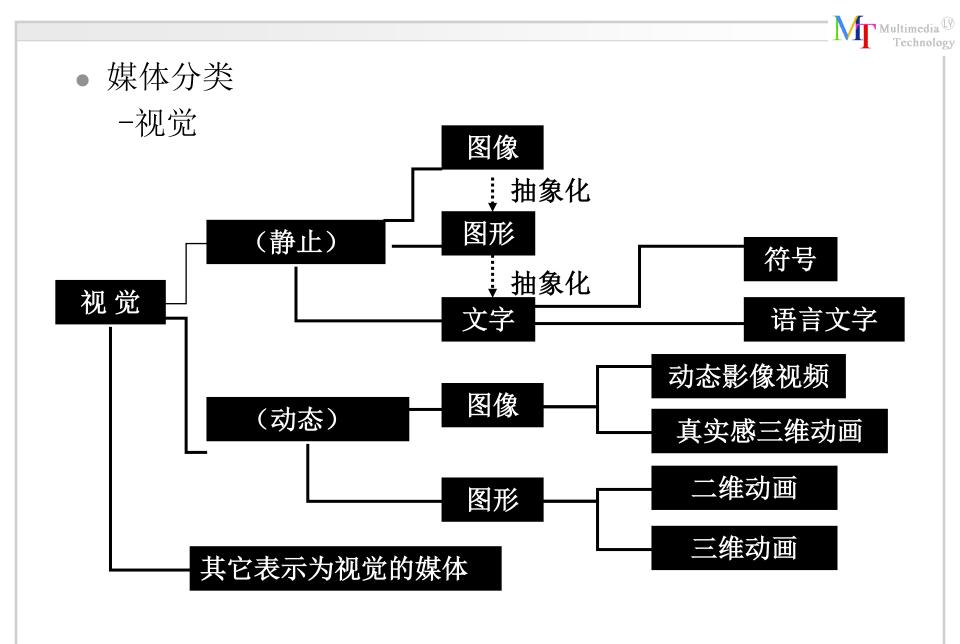
像: 动画、视频等

文:文字、符号、数值等

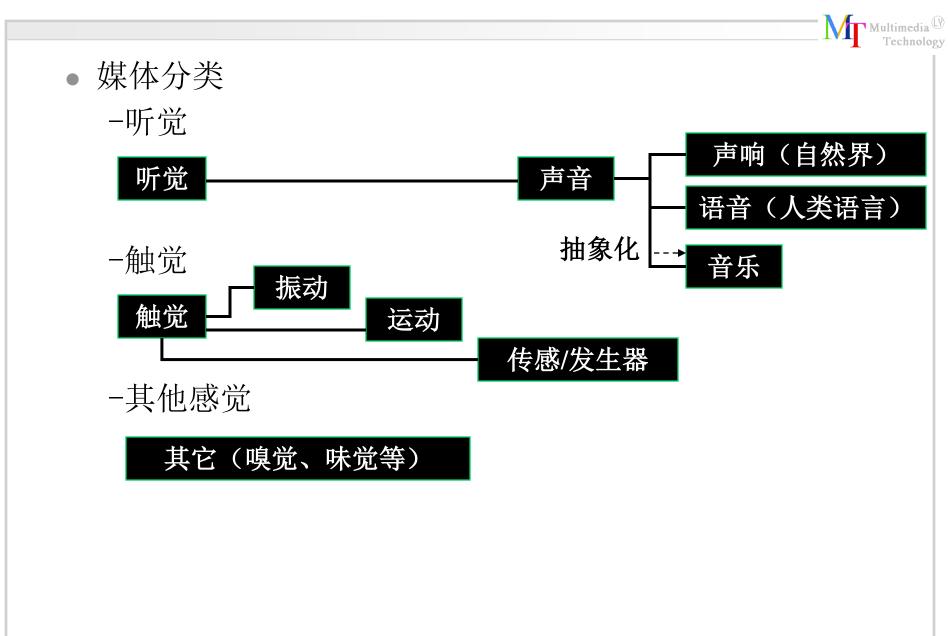
其他: 动作媒体、触觉媒体等



媒体与多媒体



媒体与多媒体



感觉媒体

表示媒体



类型 名称 分类

视觉媒体 — 景象 < 文字

景象

触觉媒体一力,运动,温度

嗅觉媒体 — 气味

味觉媒体 — 滋味

时间属性 空间属性 生成属性

图像 (image)

图形 (graphics)

文本 (text)

视频(video)

动画 (animation)

波形声音 (Wave)

合成语音(Speech)

MIDI音乐(Music)

MP3音乐

表示媒体的时间属性



- time-independent(discrete)media(离散媒体,值不会随着时间而变化)
 - Text, graphics, static image
- time-dependent (continuous) media (连续媒体,值随着时间而变化)
 - Speech, music, video, animation

表示媒体的空间属性



■ 1D媒体

• 单声道speech , music

■ 2D媒体

• 双声道music, text , image , 2D graphics

■ 3D及多维媒体

• 3D立体声music , 3D graphics , 全景图像

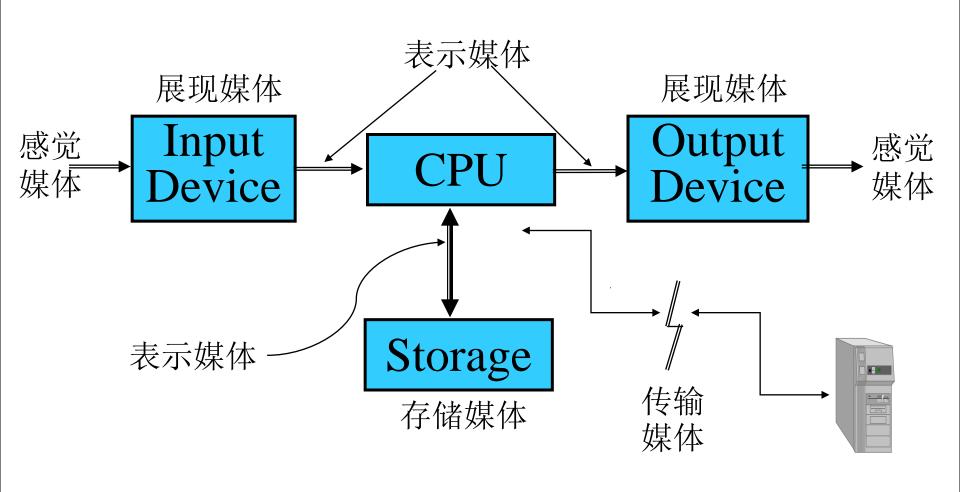
表示媒体的生成属性



- 自然媒体(使用数字化方法从现实世界获取的媒体)
 - 取样声音,扫描图像,数字视频
- 合成媒体(使用计算机创建的媒体)
 - 合成语音, MIDI, Graphics, Animation

计算机与5种媒体的关系







■ 1. 文本

- 文本是计算机文字处理程序的基础
- 文本数据可以在文本编辑软件里制作
- 文本文件分类
 - 非格式化文本
 - 格式化文本文件文件 带有各种文本排版信息等格式信息的文本文件 段落格式、字体格式、文章的编号、分栏、边框 文字的变化: 格式(style)、字的定位(align)、字体(font)、字的大小(size)

文本编码类型



■ 简体中文

- 936 GB2312 简体中文(GB2312)
- GBK GB2312-1980的扩展(包括简体和繁体)
- 10008 x-mac-chinesesimp 简体中文(Mac)
- 52936 hz-gb-2312 简体中文(HZ)
- 54936 GB18030 简体中文(GB18030)

■ 繁体中文

- 950 big5 繁体中文(Big5)
- 20000 x-Chinese-CNS 繁体中文(CNS)
- 20002 x-Chinese-Eten 繁体中文(Eten)

■ Unicode

- 1200 utf-16 Unicode
- 1201 unicodeF Unicode (Big-Endian)
- 12000 utf-32 Unicode (UTF-32)
- 65000 utf-7 Unicode (UTF-7)
- 65001 utf-8 Unicode (UTF-8)



■ 图形

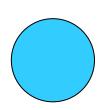
• 图形的概念

一般指用计算机绘制的画面,如直线、圆、短形、短弧、矩形、匠弧、矩形式是一贯。图形的格式是一组描述点、线、面等几何图形的大小、形状及其位置、维数的指令集合,在图形文件中只记录生成图的算法和图上的某些特征点,也称矢量图。

line (x1,y1,x2,y2,color) circle (x,y,r,color)

- 存储格式:
 - -点:记录点坐标和属性代码:
 - -线:记录两个或一系列采样点的坐标,并加属性代码;
 - -面:记录边界上一系列采样 点的坐标,由于多边形封闭 ,边界为闭合环,加面域属 性代码

图形通常由点、线、面、体、方程或 分析表达式的系数等几何形状参数和 灰度、色彩、线型、线宽、材质等非 几何属性参数组成。

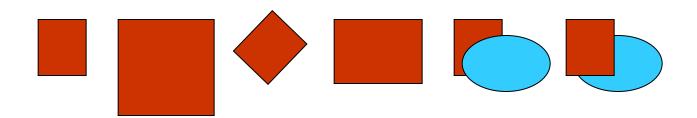


对象	值	位置
点	6	X,Y单点
线	1 8	X1, Y1; X2, Y2; Xn, Yn串
面	2 4	X1, Y1; X2, Y2; Xn, Yn闭合环



■ 图形

- 图形的特征
 - 图形是对图像进行抽象的结果(人工或自动);
 - 图形的矢量化使得有可能对图中的各个部分分别进行 控制(放大、缩小、旋转、变形、扭曲、移位等)
 - 图形的产生需要计算时间



Multimedia (P) Technology

■ 3.图像

- 图像的概念
 - -图像是指由输入设备捕捉的实际场景画面,或以数字化形式存储的任意画面。
 - -静止的图像是一个矩阵 ,由一些排成行列的点组 成,这些点称之为像素点 (pixel),这种图像称为 位图(bitmap)。
- 存储格式: P(A,R,G,B)





■图像

- 图像的主要技术参数
 - -分辨率 屏幕分辨率

计算机显示器屏幕显示图像的最大显示区

图像分辨率 数字化图像的大小 像素分辨率

像素的宽高比,一般为1:1





■图像

- 图像的主要技术参数
 - 色彩空间 RGB、CMYK、Lab、HSX
 - 图像灰度 每个图像的最大颜色数

1位:黑白2色;

4位: 16色;

8位: 256色;

24位: 真彩色

- 图像文件大小

(高×宽×灰度位数)/8 字节





■图像

- 图形与图像的关系
 - -图形是矢量概念,图元;图像是位图概念,象素;
 - -图形显示图元顺序;图像显示象素顺序;
 - -图形变换无失真;图像变换有失真;
 - -图形以图元为单位修改属性、编辑;图像只能对象素或图块处理;
 - -图形是对图像的抽象,但在屏幕上两者无异

Multimedia Technology

■ 4. 视频

• 视频的概念

视频源于电视技术,它由连续的画面组成。这些画面以一定的速率连续地投射在屏幕上

, 使观察者具有图像连续运动的感觉。









■ 视频

- 模拟视频的制式
 - -PAL制(625/50)

每秒25帧,水平扫描线为625条,水平分辨率240~400个像素,隔行扫描。62年诞生于德国,应用于中国、新加坡、欧洲地区等(Pal-B、D、G、H、I、N、NC)。

-NTSC制 (525/60)

每秒30帧,水平扫描线为525条,水平分辨率240~400个像素,隔行扫描。53年诞生于美国,应用于美国、日本、台湾等(NTSC-M、NC、Japan等)。



■ 视频

• 视频的制式

-SECAM制(625/50)

每秒25帧,水平扫描线为525条、水平分辨率625。由法国人 提出,应用于俄罗斯、法国、非洲地区等。

• 视频的数字化

指在一段时间内以一定的速度对视频信号进行捕获并加以采样后形成数字化数据的处理过程。



■ 视频

- 数字视频的技术参数
 - 颜色空间
 YUV和YIQ (Y为亮度, UV和IQ为色差)
 YCbCr (Y为亮度, CbCr为色差)
 可以通过坐标变换而相互转换
 - 帧速

每秒钟显示的帧数目, 计量单位为帧率 (fps) 视频根据制式的不同: 30帧/秒 (NTSC)、25帧/秒 (PAL)



■ 视频

- 视频的技术参数
 - -数据量 帧速×每幅图像的数据量(不压缩)
 - -图像质量 与数据源质量有关

RGB视频>YUV视频>Y/C视频(S视频)>复合视频与视频数据压缩的倍数有关

注1: UV两个色差信号可合成为一个色度信号C,进而形成Y/C记录方式,这种方式得到的视频称为S视频。

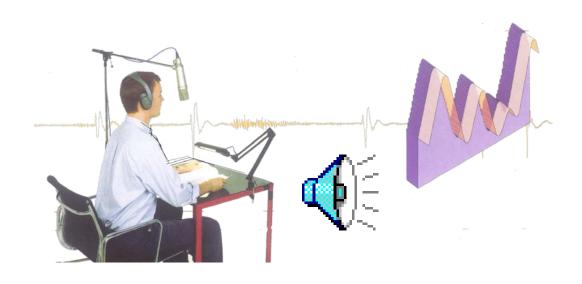
注2: 把亮度Y与色度C进一步混合在一起, 便形成复合视频。



■ 5. 音频

- 分类
 - -波形声音

包含了所有的声音形式,它可以把任何声音都进行采样量化,并恰当地恢复出来。





■音频

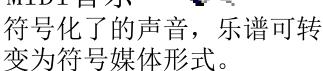
分类



-语音

人的说话声虽是一种特殊的媒体,但也是一种波形,所以和波形声音的文件格式相同。

-MIDI音乐



存储格式: (键(音调)、 通道号、音量、音色、音长 和击键强度)





■ 音频

- 数字声音波形质量的主要技术参数
 - 采样频率 等于波形被等分的份数,份数越多,质量越好 11.025KHZ、 22.05KHZ、44.1KHZ
 - 采样精度 每次采样信息量 8位、16位
 - 通道数 声音产生的波形数 单声道、立体声道、5.1声道
- 存储格式: (L,R) 数据量计算:● 存储格式: (L,R) (采样频率×采样精度×通道数 ×时间)/8 字节



■ 6. 动画

• 动画的概念

运动的图画,实质是一幅幅静态图像的连续播放。 动画的连续播放既指时间上的连续,也指图像内容上的连续,即播放的相邻两幅图像之间内容相差不大。

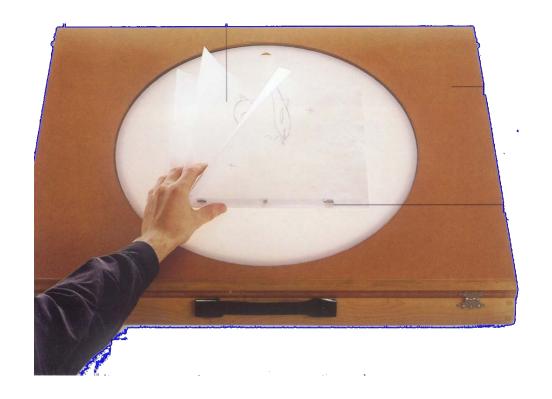
- 计算机设计动画方法
 - -造型动画
 - 帧动画





■ 动画

• 帧动画





■ 动画

• 帧动画





媒体的性质和特点



■ 各种媒体具有不同特点和性质

- 媒体是有格式的
- 不同媒体表达信息的特点和程度各不相同
- 媒体之间可以相互转换
- 媒体之间的关系也具有丰富的信息

■ 媒体具有空间性质

- 表现空间
- 媒体按相互的空间关系进行组织
- 视觉空间、听觉空间和触觉空间这3者既相互独立又需要相互结合

■ 媒体的时间性质

- 表现需要时间
- 媒体在时间坐标轴上的相互关系

媒体的性质和特点



- 媒体的时空综合和时空"上下文"
 - 媒体的时间关系存在于同步、实时等许多方面。空间和时间组成了一个三维的时空坐标系统。
 - 时间与空间的联系构成了媒体的时空"上下文"

■ 媒体的语义

- 媒体的语义是有层次的
- 抽象的程度不同,语义的重点也就不同(语音、图像〈文字、图形)

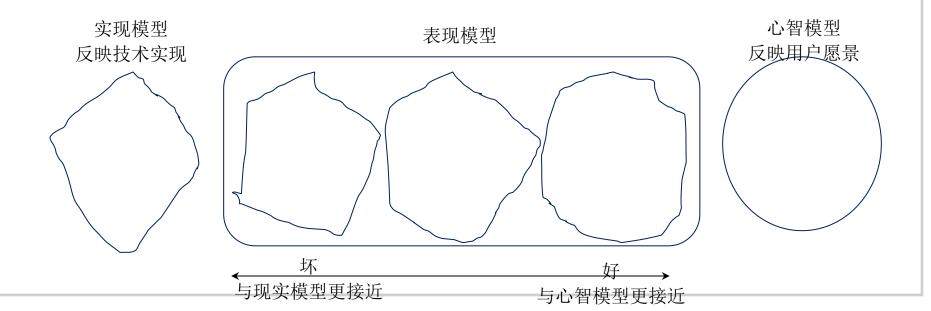
■ 媒体结合的影响

- 媒体结合是多层次的
- 媒体结合有利于信息接受和理解



■隐喻

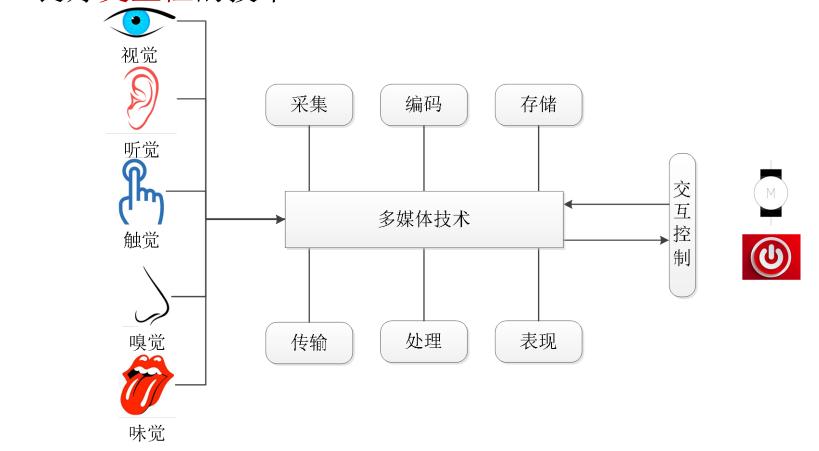
- 交互的概念模型,也称心智模型 (metal model)
- 实现模型(系统模型) (implementation model/System model)
- 表现模型(设计师模型) (represented model/designer's model)



3. 多媒体技术的概念

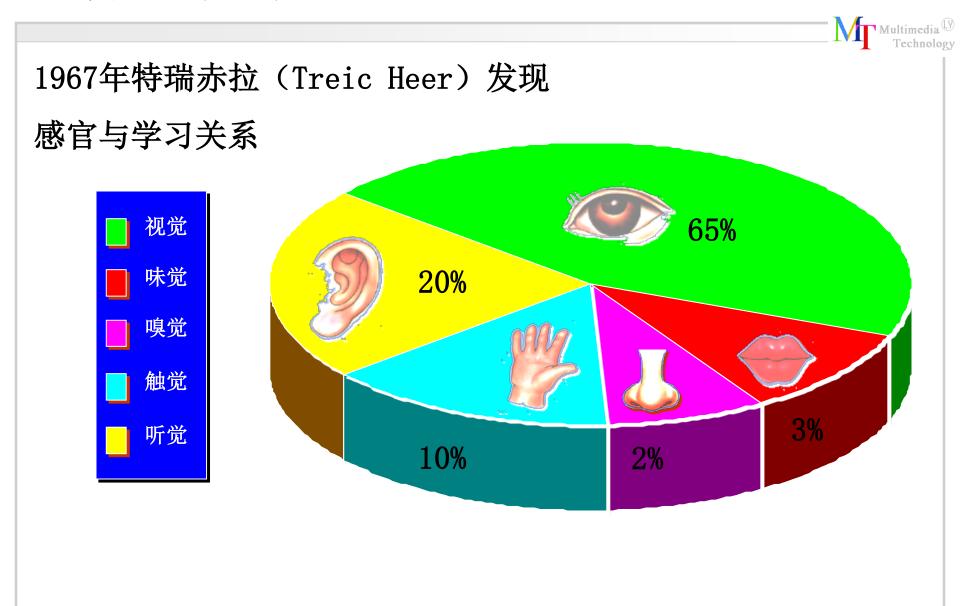
Multimedia (1)
Technology

■ **多媒体技术:** 以数字化为基础,能够对<mark>多种媒体</mark>信息进行采集、编码、存储、传输、处理和表现,综合处理多种媒体信息并使之建立起有机的逻辑联系,集成为一个系统并能具有良好交互性的技术。





- 多媒体技术强调的是交互式综合处理多种媒体(感觉媒体) 的技术。从本质上来说,它具有三种最重要的特性:
 - 媒体的多样性:信息载体的多样性是相对于计算机而言的 ,指的就是信息媒体的多样化,有人称之为信息多维化。 如文字、声音、视频、图形、图像、动画等。多媒体信息 必须通过**采样、量化**以数字的形式而不是以模拟信号的形式存储和传输的,其中的连续媒体(声音和电视图像)是人与机器交互的最自然的媒体。
 - 多媒体技术要求系统处理信息多维化,通过信息采集、 传输、处理与展现,使人机交互过程更加广阔,满足人类 感官空间全方位的信息需求。

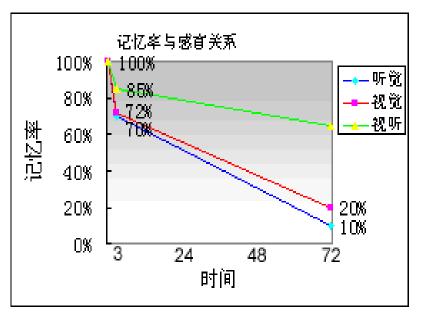




• 媒体的集成性(综合性): 多种不同媒体综合地表现某个内容,取得更好的效果;

- 多媒体技术**有机集成**文字、文本、图形(静止图形和动画)、图像、视频、语音等多种媒体。集成包括媒体信息集成、软件工具集成和硬件设备集成,使计算机具备"

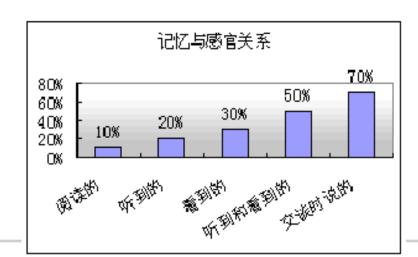
五官"。







- 处理的交互性: 使人们获取和使用信息的过程中具有细粒度的控制和操纵能力。
 - 通过各种媒体信息,使参与的各方的进行编辑,控制和传输。交互性将向用户提供更加有效的控制和干预信息的手段和方法,也为应用开辟了更加广阔的领域。交互可以让用户自由控制干预信息处理,增加对信息的注意力和理解,延长信息保留时间。当我们完全地进入到一个与信息环境一体化的虚拟信息空间自由遨游时,这才是交互式应用的高级阶段,这就是虚拟现实(Virtual Reality)







- 也就是说
 - 多媒体是人-机交互式媒体,交互性是最为重要的特性。
 - 多媒体技术的精髓是集成和交互。
 - 其综合了计算机技术、通讯技术、视听技术以及文学、音乐、绘画、摄影等艺术于一体。
 - 多媒体既是软件和硬件的系统集成,也是信息的集成,并且信息之间通过逻辑有机地连接在一起,可以交互控制。

多媒体是技术与应用发展的必然





5、多媒体的关键技术



- 视频和音频数据压缩和解压缩技术
- 多媒体计算机硬件体系结构专用芯片(压缩和解压缩技术,图像特技、图 形处理、语音处理芯片、音频与视频信号的获取与量化专业芯片等)
- 多媒体计算机软件系统的核心AVK/AVSS(视频/音频核心部件<支撑系统>)
- 多媒体数据的模型、组织与管理(海量的多媒体数据的管理需要00DBMS技术、光存储技术(CD、VCD、DVD)和超媒体技术,研究开发多媒体数据库管理系统(MDBMS))
- 多媒体信息的展现与交互(展现: Presentation, 表现各种媒体)
- 多媒体通信与分布处理技术(宽带网络实现多媒体分布数据的信息共享、 数据计算、通信、交换和协同操作)
- 多媒体特征信息和语义特征的提取、理解、识别、认证和检索等(基于内容的检索〈Content Based Retrieval, CBR〉如基于内容的音频和图像检索)
- 多媒体技术标准研究

6、多媒体技术的行业应用



- 多媒体家电(信息家电: ITV、可视电话)、多媒体家庭教育
- 多媒体教学: 多媒体技术是教学工具的一次飞跃。早期使用教学幻灯、投影、录音、录象等, 多媒体技术则提供交互的学习方式。多媒体分布式网络技术则使远程教学得以实现, "网络大学"已成现实。
 - 虚拟实验室:运用多媒体技术可以进行模拟实验和演示, 如核实验、股票交易、物理、化学和生物等实验
 - 虚拟图书馆: 多媒体咨询服务,资料图书检索。
- 分布式多媒体技术: WEB技术、多媒体会议系统、多媒体视频点播系统(VOD)、多媒体监控检测系统、多媒体通信系统、远程教学系统
- 多媒体工业控制制造技术:产品设计虚拟制造、生产现场监控
- 多媒体远程医疗技术: 多媒体医疗数据库、多媒体医疗专家 系统

虚拟远程教室



■ 多媒体白板

其作用为多媒体流的播放器和会议工具,从媒体服务器上检索课件并向用户播放,教师讲课和师生交互的视频/音频通道,把讲课记录为多媒体的同步流,此多媒体流可被索引,编辑和在课后被检索





ARTEMIS

| Real-time Tele-Educational Management Services) 是综合的远程教育管理与信息服务。 网络播放教师讲课的*实时系统*,同时,ARTEMIS 步方式的管理与协作服务。

!基本上是参照现在普通的数学模式而定义的。 中,参与者的角色是不同的,在我们的模型中, 置的。同时,在同一教室中,同时只能有一个人 言人的图像,如果可能的话。发言权的控制掌握 现有的模型中,我们不支持参与者间进行私有的 确实可以通过文字信息的方式进行交流(Text)

"组成:多媒体板,视听子系统,Mini-media服 塞

Multimedia (1) Technology

■ 视听技术发展简史

● 1839年 法 达盖尔 照相术 🖥

1843年 英 贝恩(Bain) 传真

1876年 美 贝尔 电话

1895年 意 马可尼 无线电

1893年 美 爱迪生 电影

1904年 英 弗来明 电子二极管

1907年 美 福莱斯 电子三极管

1925年 英 贝尔德(Baird) 电视

1945年 匈 冯. 诺依曼 计算机原理

1946年 美 电子数字计算机

1948年 美 贝尔实验室 晶体管(10年时间)

1971年 美 Intel 微处理器4004 微机

(速度与价格的的摩尔法则: 18月2倍)



















Cambricon NPU



IBM TrueNorth



谷歌张量处理器 Tensor Processing Unit(TPU)



英伟达GPU



中星微嵌入式NPU





■ 多媒体数据库(信息存储)

- 海量的多媒体数据的存储技术,基于多媒体对象的高速度检索,研究开发多媒体数据库管理系统(MDBMS)其典型应用是视频、音频、图形、图像等数据的存储和基于内容的检索,可能还涉及到语音和图像的模式识别技术以及人工智能技术。通过扩展RDBMS,采用大二进制对象(BOLB)类型存储多媒体数据。
 - -根据媒体数据存放位置和形式,MDBMS一般结构形式有
 - -联邦型(分离)
 - -集中统一型(集中)
 - -客户/服务器型(中间件和专用SERVER)
 - -超媒体型(WEB)等
- 多媒体通信和分布式多媒体技术(信息传输)
 - 典型的应用有ITV、VOD、视频会议、可视电话、移动通信等技术,具体向宽带和窄带两个方向发展

超文本



- 超文本(hypertext),一种把文本中遇到的相关文本组织在一起的方法,实际上,这个词的真正含义是"链接"的意思,用来描述计算机中的文件的组织方法。超文本是一种数据库方法,提供沿着链访问数据的新方法;超文本也是表达思想的方法,是一种思想工具,类似人工智能的语意表达方法;超文本也是一种接口模型,通过控制按钮访问下级数据。典型的应用是帮助文件
 - 超文本是一种新型信息管理技术,以结点〈Node〉为单位组织信息,通过链〈Link〉连接构成表达特定内容的信息网络〈Network〉。
 - -结点〈Node〉:表达信息的基本单位(信息块)。
 - -链〈Link〉: 表达结点间关系(导航),结点间有索引链和结构链两种。索引链实现结点中和域的连接,通过索引链可实施对相关信息的查找以及交叉访问;结构链对层次性信息进行操作,连接不同层次的父子结点
 - -网络〈Network〉: 由结点和链构成的有向图。系统具备检索、浏览和注释能力。

超媒体



■ 超媒体

- 超媒体(hypermedia),不仅可以包含文字而且还可以包含图形、图像,动画、声音和电视片断,这些媒体之间也是用超级链接组织的,而且它们之间的链接也是错综复杂的
- 超媒体是超文本和多媒体技术的结合词,超媒体允许信息结点可以是文本、图形、图像、音频、视频、动画以至程序等。典型的应用是网页WEB技术。
 - 多媒体是超媒体系统中的一个子集,超媒体系统是使用超链接构成的全球信息系统,全球信息系统是因特网上使用TCP/IP协议和UDP/IP协议的应用系统
- 智能超媒体: 节点和链具备计算和推理能力。其推理模型一般 采用语义网模型和流控模型。

流媒体与智能流



- **流媒体:** 通过网络传输的文件,系统在播放之前并不整个的下载文件,而是在开始时有一段延迟,将先缓存部分数据到内存中,系统播放时直接从缓存读取数据。流媒体实现的关键技术就是智能流式传输。
 - 智能流是一种新的多媒体传输技术,它可以自动检测网络 状况并调整视频流的属性,以便达到最佳质量。之所以开 发这项技术,是因为不同 Internet 连接实际能达到的传 输量和连接速度有很大差异。这种差异不利于接收流式媒 体。智能流取决于以多比特率编码的内容。这就为提供了 多种速率的视频流,以便网络带宽不稳定时进行切换。

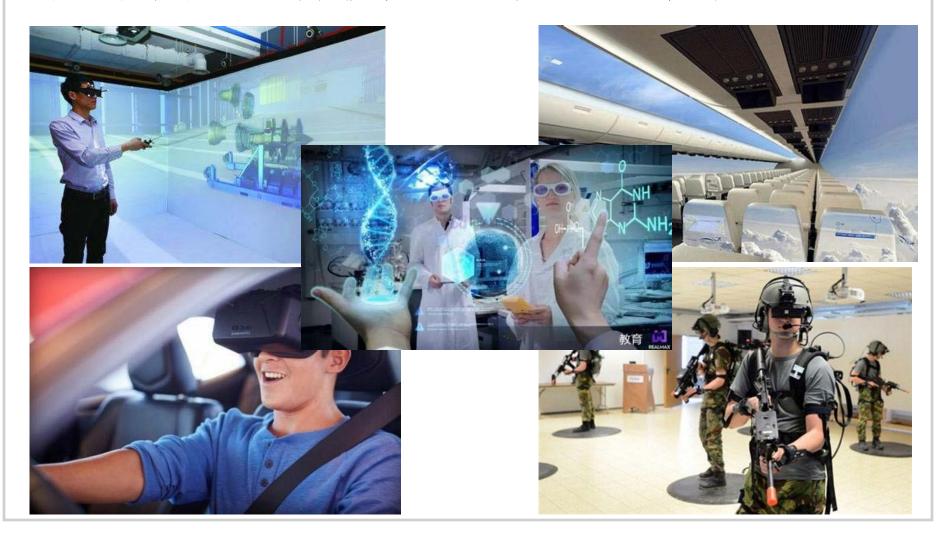


■ 虚拟现实技术

- 虚拟现实(Virtual Reality)技术是在计算机系统环境下,集视、听、说和触动等多种感觉器官功能于一体的仿真技术。利用VR技术可以带动各个领域的仿真技术的发展。VR综合应用计算机图像、模拟与仿真、传感技术和显示技术等。硬件设备有数据手套、轨迹追踪装置、语音识别以及摄象机等。
- VR为用户生成一个具有逼真视觉、听觉、触觉及嗅觉的虚拟世界(环境),用户可以根据自身的感觉,用人的自然技能与这个世界(环境)进行交互,参与其中的事件,让用户感觉如身临其境。
- VR具有3个最突出的特征(称为3 "I"特性)
 - -交互性(interactivity)
 - -沉浸感(illusion of immersion)
 - -想象力(imagination)
- 虚拟现实造型语言(Virtual Reality Modeling Language, VRML)是一种用来描述万维网页面上三维交互环境的文件格式。

Multimedia Technology

■ VR技术的应用: 娱乐、游览、推销、建筑、教学、军事、医疗、体育训练、驾驶模拟(包括车、船、飞机等)





■ 智能多媒体技术(信息处理)

- 语音合成(TTS)
- 语音识别(SR)
- 语言理解(NLI)
- 图像识别(OCR、FR)
- 图像生成(IR)
- 生物特征识别(BR)
- 多媒体信息检索(CBR)
- 情感计算(AC)



知识点小结



- 数据、信息与媒体,媒体的分类
- 多媒体技术的概念、特性
- 多媒体信息的采样、获取与量化
- 多媒体数据的存储(光盘、多媒体数据库)
- 多媒体数据的传输(网络多媒体、超媒体、流媒体)
- 智能多媒体的计算技术(语音合成与语音识别、图像识别与图像生成)
- 虚拟现实技术与VRML
- 多媒体技术的行业应用与关键技术