第一章：

1. 音调、音强、音色 分别对应频率、幅度、谐波

耳蜗、内耳耳蜗

1. 数字化实际上是将模拟音频信号经过A/D转换形成的数据。

音频数字化：取样、量化、编码

1. 奈奎斯特采样定理，为了不让音频失真，采样频率应该 f>=40khz
2. 波形音频的存储量M=采样频率\*采样位数\*声道数/8（字节/秒）
3. 数字音频一般分为波形音频与midi音频，他们的的区别：

音频是波形，传输的是声音信号，而MIDI只是控制音高、节奏与响度的指令。符号化的，MIDI像是一本乐谱，把需要演奏的音符、音符的力度、表情记在一个MIDI文件里。

与波形声音相比，MIDI不是声音数据而是指令，所以数据量要少得多。30分钟的音乐， 用MIDI文件记录只需200KB，用 16位CD品质的未压缩WAV文件记录需317MB

MIDI可以与其他波形声音配合使用，形成伴乐的效果。而两个波形声音一般是不能同时使用的

对MIDI的编辑也很灵活，用户可以自由地改变音调、音色等 属性，直到自己想要的效果

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它非乐曲类声音

 MIDI音频文件协议以及音频形态与存储方法。MIDI音频文件中不包含音频的波形，存储空间小，MIDI音频文件包括描述 声音的键、通道号、音量、音色、音长和击键强度等

第二章

1.色彩的心理属性：

颜色需要考虑三种心理属性：色调、亮度和饱和度，色调和饱和度统称为色度

2.图形和图像的区别 （表示方法）

图形指由轮廓线条构成的矢量图,矢量图形与分辨率无关,矢量图（vector-based image）是 用一系列计算机指令来描述和记录一幅图，这幅图可分解为一系列子图如点、线、面等.

图像由点阵组成的点位图（光栅）构成，各像素都有颜色和强度,位图图像与分辨率有关。 位图（bit-mapped image），是用像素点来描述或映射的图，也即位映射图。位图在内存中 也就是一组计算机内存地址位（bit）组成，这些位定义图像中每个像素点的颜色和亮度。 图像是由像素点阵构成的位图，其重要参数是图像色彩深度和分辨率。图像文件字节数B=（ H×W×位图深度）/8，按照图像存储文件格式常见有BMP，GIF，JPG，TIF等

图像的基本属性：

图像的属性包含分辨率、像素深度、真/伪彩色、图像的表示 法和种类等

分辨率

-显示分辨率：显示分辨率是指显示屏上能够显示出的像素数目。例如，显示分辨率为640×480表示显示屏分成 480行，每行显示640个像素，整个显示屏就含有307200个 显像点。屏幕能够显示的像素越多，说明显示设备的分辨 率越高，显示的图像质量也就越高

-图像分辨率：图像分辨率是指组成一幅图像的像素密度 的度量方法。对同样大小的一幅图，如果组成该图的图像 像素数目越多，则说明图像的分辨率越高，看起来就越逼 真。相反，图像显得越粗糙。分辨率 越高，像素就越多。

像素深度：是指存储每个像素所用的位数，它也是用来度量图像的分 辨率。像素深度决定彩色图像的每个像素可能有的颜色数，或者确定 灰度图像的每个像素可能有的灰度级数。往往把像素深度说成是图像 深度。表示一个像素的位数越多，它能表达的颜色数目就越多，而它 的深度就越深。例如，RGB 5∶5∶5表示一个像素时，用2个字节共16位表示，其中R，G，B各占5 位，剩下一位作为属性位。在这种情况下，像素深度为16位，而图像 深度为15位。 属性位用来指定该像素应具有的性质。用RGB 5∶5∶5表示的像素共16位，其最高位(b15)用作属性位，并把它称为 透明(Transparency)位，记为T

真彩色(true color) -真彩色是指在组成一幅彩色图像的每个像素值中，有R，G，B三个基色分量，每个基 色分量直接决定显示设备的基色强度，这样产生的彩色称为真彩色。真彩色图通常是 指RGB 8:8:8，即图像的颜色数等于224，也常称为全彩色(full color)图像。

伪彩色(pseudo color) -伪彩色图像的含义是，每个像素的颜色不是由每个基色分量的数值直接决定，而是把 像素值当作彩色查找表(color look-up table，CLUT)的表项入口地址，去查找一个显 示图像时使用的R，G，B强度值，用查找出的R，G，B强度值产生的彩色称为伪彩色。

3. 加光混合（色光三基色红(R)、绿(G)、蓝(B)），减光混合（颜料三原色青(C)、紫(M)、黄(Y)）：

加光混合：

R+G→Y

R+B→M （紫色）

G+B→C （青）

R+G+B→W（白）

减光混合：

C+M→B

M+Y→R

Y+C→G

Y+C+M→K（黑）

4.颜色模型

 RGB是根据三基色原理，用基色光 单位来表示光的量，则在RGB色彩 空间，任意色光F都可以用R、G、 B三色不同分量的相加混合而成

F＝r [ R ] + g [ G ] + b [ B ]

 CMYK是根据三原色原理，由于彩 色墨水和颜料的化学特性，用等 量的CMY三基色得到的黑色不是真 正的黑色，因此在印刷术中常加 一种真正的黑色（black ink）， 所以CMY又写成CMYK。

蒙赛尔：是基于人类对颜色的感觉三个基本特征描述的颜色模型，它们都是以色调为基础的色彩空间

Lab：L\*a\*b 颜色模型由亮度或光亮度分量 (L) 和两个色度分量组成；两个分量即 a 分量（从 绿到红）和 b 分量（从蓝到黄）。

在颜色模型中，L\*a\*b具有最宽的色域 ，它包括RGB和CMYK色域中的所有颜色 。通常RGB色域包含能在计算机显示器 或电视屏幕（发出红、绿和蓝光）上 所有能显示的颜色。因而一些诸如纯 青或纯黄等颜色不能在显示器上精确 显示。

 CMYK色域较窄，仅包含使用印刷色油 墨能够打印的颜色。当不能被打印的 颜色在屏幕上显示时，它们称为溢色 ──即超出CMYK色域之外。

第三章

1.动态视觉

动态视觉特征是视频处理的依据，也是视频的色彩空间模型 的基础。

以一定的速率把胶片投影到银幕上才能有运动 的视觉效果，这种现象是由视觉残留造成的，动画和电影利用的正是 人眼这一视觉残留特性

2.色彩空间：不同于图像（rgb）

YIQ适用于NTSC彩色电视制式 

YUV适用于PAL和SECAM彩色电视制式 

YCrCb适用于计算机用的显示器 

YUV与YIQ模型 

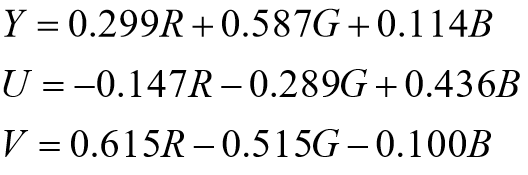
在彩色电视制式中，使用YUV和YIQ模型来表示彩色图像。 在PAL彩色电视制式中使用YUV模型，Y表示亮度，UV用来表 示色差，U、V是构成彩色的两个分量；在NTSC彩色电视制 式中使用YIQ模型，其中的Y表示亮度，I、Q是两个彩色分 量。 

YUV/YIQ特点 

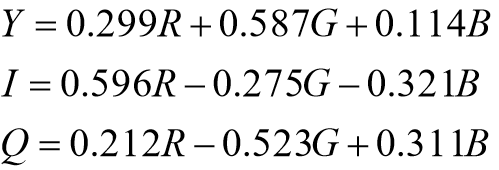
亮度信号(Y)和色度信号(U、V)是相互独立的 

可以利用人眼的特性来降低数字彩色图像所需要的存储容 量。

YUV与RGB彩色空间变换

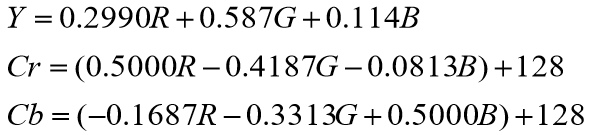


YIQ与RGB彩色空间变换



YCrCb与RGB彩色空间变换

 数字域中的彩色空间变换与模拟域的彩色空间变换不同。 它们的分量使用Y、Cr和Cb来表示，与RGB空间的转换关系 如下：



3.

通过对每一个象素进行采样，并且按 照颜色或者灰度进行量化，就可以得到图像的数字化结果 。数字化的结果放在显示缓存区中，与显示器上的点一一 对应，这就是位图图像。

对彩色电视图像进行采样时，可以采用两种采样方法。

 一种是使用相同的采样频率对图像的亮度信号和色差信号进 行采样

 另一种是对亮度信号和色差信号分别采用不同的采样频率进 行采样。如果对色差信号使用的采样频率比对亮度信号使用 的采样频率低，这种采样就称为图像子采样

子采样也就是利用人的视觉系统这两个特性来达到压缩彩色电视信号。

 一是人眼对色度信号的敏感程度比对亮度信号的敏感程度低，利用这个特性可 以把图像中表达颜色的信号去掉一些而使人不察觉；

 二是人眼对图像细节的分辨能力有一定的限度，利用这个特性可以把图像中的 高频信号去掉而使人不易察觉

第四章：

1．三维计算机动画制作过程分以下几个阶段：

（1）造型:物体建模造型（利用点线面构造物体的轮廓）

（2）确定颜色和材质(纹理):真实感设计（光源和表面材质< 纹理>渲染、）

（3）设置灯光和摄像机位置

（4）描述和设置动画要求:动画设置（模型位移、变形、背 景和光色变化）

（5）图像绘制:图像生成（通过光线反射、投 射、慢射、折射和透视等，对场景进行推、拉、摇、移和 变焦等，以产生动画）

（6）输出动画结果配音等

2.视频的基本要素：空间分辨率、帧频

帧频：每一秒钟所扫描的帧数称为帧频

图像的获取过程

 取样( 每一帧图像离散为：行、象素 )

 彩色象素的分色处理（使用基色表示）

 彩色分量的量化处理（A/D转换）

获取设备：扫描仪、数码相机、

第五章：

1. 空间位置跟踪：空间六个自由度

Polhemus三维定位机构：六个自由度(X, Y, 和 Z 笛卡尔坐标和俯 仰角、偏行角、滚动角)

1. 三维立体图像的产生方法：

立体图像的产生 （1）旋转法 （2）投影变换法 （3）3D扫描法与立体摄象法

视差理论：

决定立体视觉的是视差。视差的种类大致分为四种：零视 差、正视差、负视差和发散视差。它们产生的立体感觉是 不同的。

视差（parallax）是投影到人眼视网膜上图像上两点间的 水平距离，正是这个距离产生了视觉上的立体感。注视远 处的物体时的视差与注视近处物体时的视差是不同的，因 此所谓的立体感正是这个不同距离作用的结果。

3DVA的基本理论

人类感知声源位置的最基本的理论是双工理论（1907年 Rayleigh提出Duplex理论：人通过双耳判断声源的方向和 远近，成为双耳定位）。该理论基于两种因素：两耳间声 音的到达时间差ITD（Interaural Time Differences）和 两耳间声音的强度差IID（Interaural Intensity Differences）。

 时间差是由于距离的原因造成的，当声音从正面传来时，距离相等，所以没有时间差；但若偏右3o，则到达右耳的 时间就要比左耳约早30ms，而正是这30ms，使得我们辨别 出了声源的位置。

 强度差是由于信号衰减造成的，信号的衰减是因为距离产 生的，在很多情况下是因为人的头部遮挡，使声音衰减， 产生了强度的差别，使得靠近声源一侧的耳朵听到的声音 强度要大于另一耳。

HRTF方法 

实现空间真实感的关键是建立起耳廓模型，这种方法被称 为“双耳相关函数法”，也即与头部有关的转移函数（ Head-Related Transfer Function，HRTF）。

头部对入射声波所引起的衍射作用，以及声音到达两耳时形 成的时间差和强度差使人耳能对声音方向定位。

 声波从声源到人耳鼓膜处的变化可以看作是人的双耳对声波 的滤波作用，人们将声波从自由场传到鼓膜处的变换函数称 之为与头部相关传递函数

测量多声源SPK在左右耳MIC对SPK的头部相关脉冲相应(Head Related Impulse Responses, HRIR),其相应变换即HRTF。HRTF依赖于声源方位角 、高度、距离和频率，而对远声场只与方位角、高度和频率有关，而且 每个人HRTF都不同。

1. 触觉媒体对应的基本描述方法：

振动、反馈

触觉刺激（物体的表面纹理等）、反作用力（ 推门的门重感觉）、运动感觉（摇晃、振动等）及温度、 湿度等环境信息

第六章：

1. 数据压缩技术的性能指标

1.压缩比：压缩性能常常用压缩比定义（输入数据和输出数 据比）

 例：512×480， 24bit/pixel(bpp)

 输出15000byte 输入＝737280byte

 压缩比＝737280/15000＝49

 2．质量

 压缩方法：无损压缩（质量不变）和有损压缩 -有损压缩：失真情况很难量化，只能对测试的对象进行 估计。

 3．压缩解压速度 

在许多应用中，压缩和解压可能不同时用，在不同的位置 不同的系统中。所以，压缩、解压速度分别估计。  静态图像中，压缩速度没有解压速度严格；动态图像中， 压缩、解压速度都有要求，因为需实时地从摄像机或VCR中 抓取动态视频。 

4．硬软件系统 

有些压缩解压工作可用软件实现。设计系统时必须充分考 虑：

算法复杂 － 压缩解压过程长

算法简单 － 压缩效果差 

目前有些特殊硬件可用于加速压缩/解压。

硬接线系统速 度快，但各种选择在初始设计时已确定，一般不能更改。 因此在设计硬接线压缩/解压系统时必须先将算法标准化。

2.数据压缩与数据冗余

数据压缩技术分类 

根据解码后数据与原始数据是否完全一致可以分为两大类：

一类是熵 编码、冗余压缩法，也称无损压缩法、无失真压缩法；

二是熵压缩法 ，也称有损压缩法、有失真压缩法

数据冗余的类型

 空间冗余、时间冗余、视觉冗余、信息熵冗余、其它冗余：包括结构冗余、知识冗余等

3.连续色调静止图像的数字压缩和编码标准JPEG

JPEG标准划分为： 4种编码模式：

-无失真编码模式（lossless encoding mode）

-基于离散余弦变换(Discrete Cosine Transform,DCT)的 顺序编码模式（DCT based sequential encoding mode）

-基于DCT的渐进编码模式（DCT based progressive encoding mode）

-基于DCT的多分辨率编码模式（DCT based hierarchical encoding mode） 其中后三种模式采用基于DCT的有损编码算法

4. MPEG内容主要包括：MPEG视频组、MPEG音频组和 MPEG系统组三部分。

 MPEG视频组任务是研究视频信号压缩传输

 MPEG音频组任务是研究音频信号压缩传输

 MPEG系统组任务解决多道压缩数据流的同步及合 成问题

提高压缩比：

时间域：块基运动补偿技术缩减画面间时间冗余度

 空间域：JPEG域基变换压缩技术以缩减空间冗余度

mpeg-1,mpeg-2

第七章：

1. 识别方法

 单模态生物特征识别 -人脸识别、指纹识别、虹膜识别、掌纹识别、字符识别、唇动识别等

 多模态生物特征融合 -生物特征融合可分为数据融合、特征融合和决策融合等 -基于人脸、语音、唇动的识别系统

常用bit生物特征：指纹、掌纹、虹膜、脸像、声纹、签名笔 迹等

特征提取、训练模型、测试模型

2.那些可用于生物特征识别：人脸识别、指纹识别、虹膜识别、笔迹

语音合成，语音识别不属于BIT的内容

第八章

1.人机交互: 一门研究人类所使用的交互式计算系统的设计、实施、评 估及相关主要现象的学科

2.GUI：色彩是设计GUI重要环节（颜色与背景调和）

CUI：文本是最基本最明了的提示信息，它是CUI最主要的信息 交互手段，良好的文本的呈现受文字的大小、颜色、字体、 样式、位置等因素的影响

3.时空同步：

同步的基本概念

 多媒体是在不同应用环境中文本、图像、声音、视频等各种媒体的集 成。既然需要将这些媒体安排在一起表现，就有一个先来后到和空间 位置的关系。这个关系就是同步关系，系统对各个媒体对象按照这个 关系进行的控制过程，就是同步。

 同步的过程与时间和空间有着密切的关系，大多数同步都建立在时间 的基础上。

 基于时间的媒体（Time-based Media，时基媒体）或时间依赖媒体（ Time-dependent）： -媒体与时间有着强烈的依赖关系，在采样、传输和回放表现时更需 要以时间为参照系进行有序的组织。

时间关系是多媒体对象的本质特征。  视频、音频等基于时间的连续媒体必须要有严格的同步时间要求。  同步的过程与时间有着密切的关系，同步有应用同步、合成同步、现场同 步和系统同步。

4.多媒体存储技术有哪些，存在的问你题有哪些？

第九章：

1.多媒体数据库的模型以及常见的结构有哪些：

多媒体数据模型

NF2 数据模型、面向对象数据模型、超媒体数据模型 、信息元模型、表现与同步模型

 文献模型：其基本结构是层次状的，其主结构是树形的。

 专有媒体数据模型：象图像数据库、视频数据库、全文数据库等针对特定领域 的数据库，往往根据自己的需要建立符合自己特性的体系结构和数据模型，以 完成特定的任务。

多媒体数据库的一般结构形式

 一、联邦型 二、集中统一型 三、客户/服务器型 四、超媒体型

2.

多媒体数据库可划分为四个层次 

第4层 概念数据模式层

第 3层 存取与存储数据模型层

第 2层 媒体支持层

第 1层 数据模型层

3. 多媒体信息检索(Multimedia Information Retrieval, MIR)

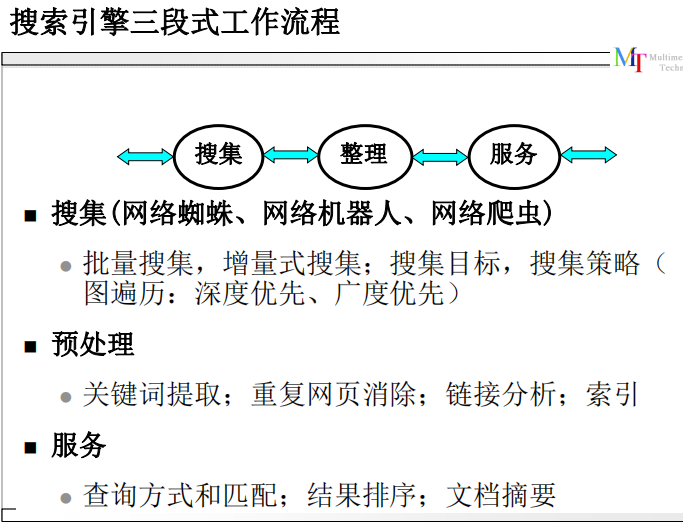
 指对多媒体数据中蕴涵的内容进行计算、分析和理解，以 提取媒体语义信息，便于用户对多媒体内容的快速访问一 种信息检索技术

1. cbr系统的4个部分

CBR系统的一般结构

-插入子系统 -特征提取子系统 -数据库子系统 -查询子系统

6.多媒体搜索引擎：



第十章：

1.超媒体的基本概念：

超媒体（Hypermedia） :多媒体超文本

 节点中的数据不仅仅可以是文字,而且可以是图形、图像 、声音、动画、动态视频，甚至计算机程序或它们的组合

 成为支持多媒体信息管理的天然技术；

 为强调系统是多媒体的而引入“超媒体

2.智能超媒体的2个特点

节点和链是构成超媒体的核心,智能超媒体扩展了普通超媒体 节点和链的类型,以适应于链的计算和推理,以及多媒体信息 的智能化表现。

3．多媒体服务质量(Quality of Service，QoS)是一种抽象概念，用于 说明网络服务的“好坏”程度。 

由于不同的应用对网络性能的要求不同，对网络所提供的 服务质量期望值也不同。这种期望值可以用一种统一的QoS 概念来描述。从支持QoS的角度，多媒体网络系统必须提供 QoS参数定义和相应的管理机制。

QoS参数

QoS是分布式多媒体信息系统为了达到应用要求的能力所 需要的一组定量的和定性的特性，它用一组参数表示

-QoS={吞吐量、延迟、延迟抖动、差错率与可靠性}

4. 流媒体播放方式

 按照通信方式(网络通信设计) - 单播 - 组播(多播)

 按照播放模式(Ｃ/Ｓ连接) - 点播 - 广播

基本理论与原理：

1. 多媒体与跨媒体概念
2. 多媒体智能处理基本流程（可以举例子）
3. 多媒体信息检索基本问题：语义鸿沟，为什么要做相似性检索
4. 怎样压缩，怎样提高压缩比（图像）
5. 数字图像，音频，动画，数字化所对应的基本属性与参数
6. 流媒体基本原理、常用的协议

流媒体（Stream Media）并不是单一的技术，它是融合很多网络技术之 后所产生的技术。它涉及到流媒体数据的采集、压缩、存储、传输以及 网络通信等多项技术。

 流媒体指在Internet/Intranet中使用流式传输技术的连续时基媒体，如 ：音频、视频或多媒体文件。

支持流媒体传输的网络协议

 实时传输协议RTP与实时传输控制协议RTCP

 实时流协议RTSP

 多用途因特网邮件扩展协议MIME

 资源保留协议RSVP

