

# 多媒体笔记

---

by 吴纵宇

## chap0 课程介绍

---

三次实验：ps、pr或ae、大作业汇报

## chap1 多媒体技术概要

---

你理解的多媒体是什么？

- 多媒体是融合两种或两种以上媒体的人-机互动的信息交流和传播媒体
  - 多媒体是信息交流和传播的媒体，从这个意义上说，多媒体和电视、报纸、杂志等媒体的功能是一样的
  - 多媒体是人-机交互媒体，这里指的“机”，主要是指计算机、(智能)手机或是由微处理器控制的其他终端设备。计算机的一个重要特性是“交互”，使用它容易实现人-机交互功能，这是多媒体和模拟电视、报纸、杂志等传统媒体大不相同的地方
  - 多媒体是以数字形式而不是模拟信号形式存储和传输的
  - 传播信息的媒体的种类很多，如文字、声音、电视图像、图形、图像、动画等。虽然融合任何两种或两种以上的媒体就可以称为多媒体，但通常认为多媒体中的连续媒体（声音和电视图像）是人-机互动的最自然媒体

什么是互联网思维

- 提出互联网思维的是百度公司创始人李彦宏。互联网思维，就是在（移动）互联网+、大数据、云计算等科技不断发展的背景下，对市场、用户、产品、企业价值链乃至对整个商业生态进行重新审视的思考方式。互联网时代的思考方式，不局限在互联网产品、互联网企业。这里指的互联网，不单指桌面互联网或者移动互联网，是泛互联网，因为未来的网络形态一定是跨越各种终端设备的，台式机、笔记本、平板、手机、手表、眼镜，等等。互联网思维是降低维度，让互联网产业低姿态主动去融合产业。互联网思维的六大特征：大数据、零距离、趋透明、慧分享、便操作、惠众生。

互联网思维带来的三大改变

- 改变一：万倍增速（时间缩短）
  - 互联网将彻底改变人们的速度观念。
  - “瞬间即时”的微信、微博和APP？

- 改变二：万倍扩容（空间放大）
  - 互联网将大大拓宽人类的生存空间。
  - 二十四版、四十八版、一百二十八版的报纸，怎么拼得过互联网无边无际的海量信息？
  - 超市建一个仓库、十个仓库、百个仓库……怎能及互联网商场上的无所不包的海量库容？过去谁能想像，整天坐在家中，手机一点就可以将全家吃喝拉撒统统搞定？
- 改变三：万众参与（人人互动）
  - 互联网将地球村民紧紧连在一起。
  - 过去谁能想像，一个热点话题，可以令全球几十万甚至几百万人同时讨论争得面红耳赤？

## 产品结构规划

引流产品、利润产品，形象产品的设计原则和方法

- 引流产品：负责带来流量
- 利润产品：负责贡献利润
- 形象产品：负责拔高形象

这三种产品各自扮演不同角色，完成不同任务，共同搭建了企业的产品结构金字塔，如下图所示

image-20220514211223068

无需学习的自然交互。

被动交互正在兴起。

协同交互

## 多媒体技术

是指能够在同一环境中获取、处理、存储、输出文本、图形、图像、音频和视频等多种媒体信息，并在这些媒体间建立逻辑联系，集成为一个具有交互性的系统的技术。

## 媒体

- 承载信息的载体
- 信息的传播和存储形式是多种多样的，即信息表现的媒体是多种多样的

## 媒体(medium)－分类方法1

CCITT(国际电联)定义五种类型：

- 感觉媒体
  - 感觉媒体指的是能直接作用于人们的感觉器官，从而能使人产生直接感觉的媒体。如文字、数据、声音、图形、图像等。在多媒体计算机技术中，我们所说的媒体一般指的是感觉媒体。
- 表示媒体
  - 表示媒体指的是为了传输感觉媒体而人为研究出来的媒体，借助于此种媒体，能有效地存储感觉媒体或将感觉媒体从一个地方传送到另一个地方。如语言编码、电报码、条形码等。
- 显示媒体
  - 表现媒体指的是用于通信中使电信号和感觉媒体之间产生转换用的媒体。如输入、输出设备，包括键盘、鼠标器、显示器、打印机等。
- 存储媒体
  - 存储媒体指的是用于存放表示媒体的媒体。如纸张、磁带、磁盘、光盘等。
- 传输媒体
  - 传输媒体指的用于传输某种媒体的物理媒体。如双绞线、电缆、光纤等。

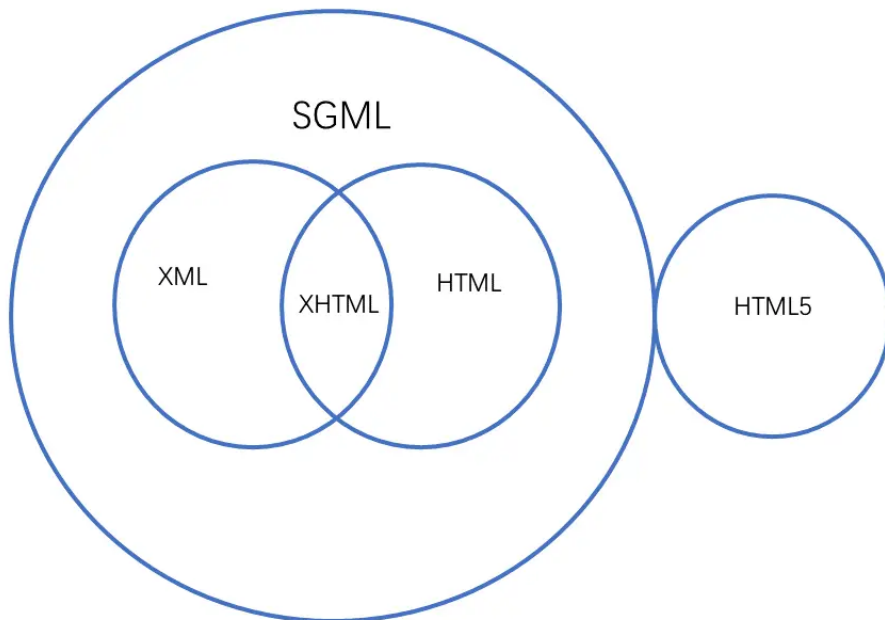
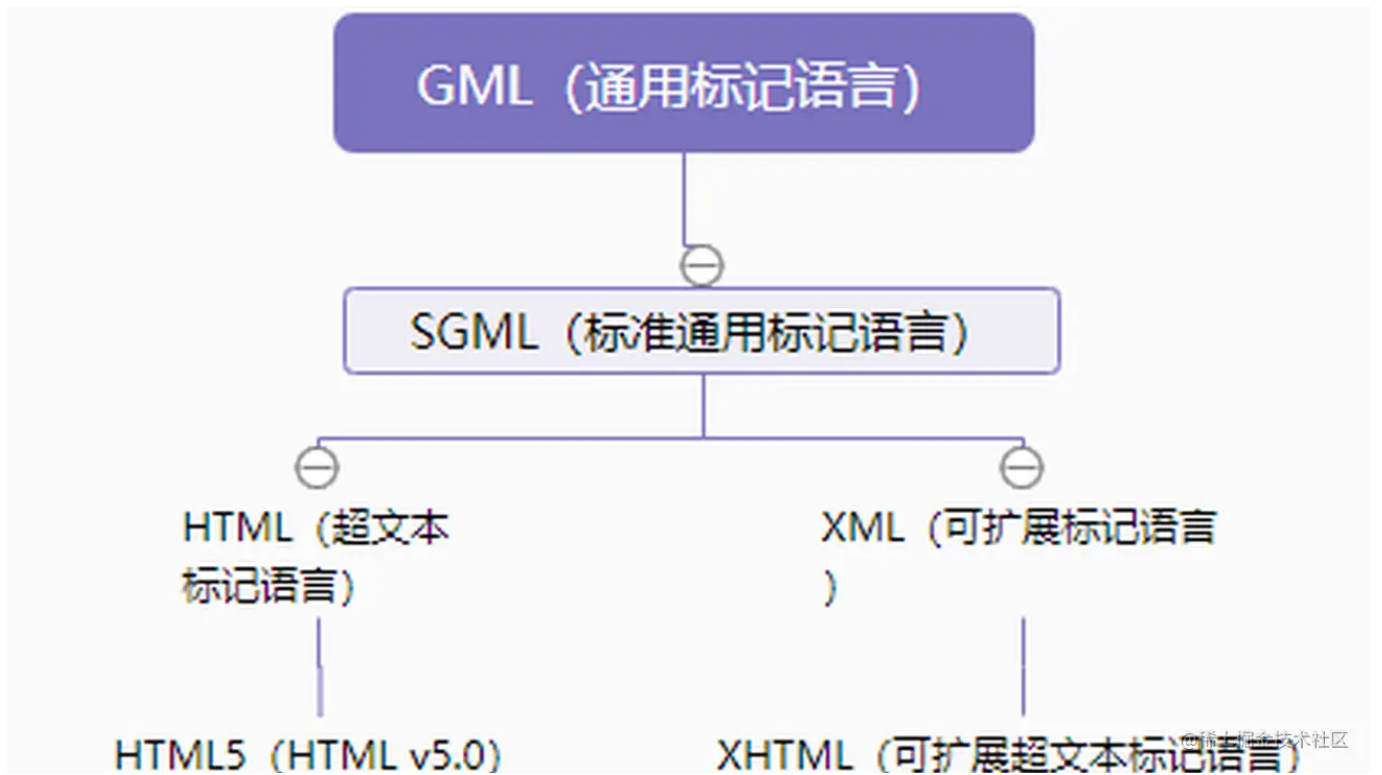
## 媒体(medium)－分类方法2

- (1) **视觉类媒体**：通过视觉传达信息的媒体。如文字、点阵图像、矢量图形、动画、视频等；
- (2) **听觉类媒体**：通过声音传达信息的媒体。如波形声音、语音、音乐、
- (3) **触觉类媒体**：指环境媒体，不通过视觉和听觉，仍可以接受和传递信息。如皮肤可感受环境温度、湿度、压力，身体可感受振动、运动、旋转等。
- (4) **嗅觉类媒体**正在研究之中。
- (5) **味觉类媒体**正在研究之中。

**超文本**：文本的一种非线性结构的设想

**超媒体**：多媒体+超文本

**超文本语言**



- 标准通用标记语言(SGML)
- 超文本标记语言(HTML)
- 虚拟现实造型语言(VRML)
  - 虚拟现实建模语言。是一种用于建立真实世界的场景模型或人们虚构的三维世界的场景建模语言，也具有平台无关性。是目前Internet上基于 WWW的三维互动网站制作的主流语言。本质上是一种面向web，面向对象的三维造型语言，而且它是一种解释性语言。

VRML的对象称为结点，子结点的集合可以构成复杂的景物。结点可以通过实例得到复用，对它们赋以名字，进行定义后，即可建立动态的VR（虚拟世界）。VRML不仅支持数据和过程的三维表示，而且能提供带有音响效果的结点，用户能走进视听效果十分逼真的虚拟世界（如简易迷宫、国际象棋）。用户使用虚拟对象表达自己的观点，能与虚拟对象交互，为用户对具体对象的细节，整体结构和相互关系的描述带来的新的感受。

## 元宇宙是第三代互联网（Web 3.0）

image-20220515122228658

### 新的关于“媒体”的概念？

**新媒体**是基于数字技术、网络技术及其他现代信息技术或通信技术的，具有互动性、融合性的介质形态和平台。在现阶段新媒体主要包括网络媒体、手机媒体及其两者融合形成的移动互联网，以及其他具有互动性的数字媒体形式。

#### 网络新媒体，

**移动新媒体：**是所有具有移动便携特性的新兴媒体的总称，包括手机媒体、平板电脑、掌上电脑、psp、移动视听设备（如MP3、MP4、MP5）等

**跨媒体：**跨媒体传播是指信息在不同媒体之间的流布与互动，它至少包含两层含义：其一是指相同信息在不同媒体之间的交叉传播与整合。其二是指媒体之间的合作、共生、互动与协调。

**融媒体：**融媒体是指各种媒体呈现出多功能一体化的趋势，即充分利用互联网这个载体，把广播、电视、报纸这些既有共同点，又存在互补性的不同媒体，在人力、内容、宣传等方面进行全面整合，实现资源通融、内容兼容、宣传互融、利益共融的新型媒体。

**自媒体：**自媒体是指普通大众通过网络等途径向外发布他们本身的事实和新闻的传播方式

**流媒体：**是指将一连串的媒体数据压缩后，经过网上分段发送数据，在网上即时传输影音以供观赏的一种技术与过程，此技术使得数据包得以像流水一样发送；如果不使用此技术，就必须在下载使用前下载整个媒体文件

#### 多媒体的应用：

- **家庭：**家庭成员的自学教育，娱乐。智能家居
- **通信业：**信息点播主要有桌上多媒体通信系统和交互电视ITV两种形式。
- **教育**
  - 信息查询和检索(虚拟图书馆):
    - 在学校或家中，可以使用多媒体虚拟图书馆快速地查询所要的图书及技术资料。
  - 交互式学习

- 既可以为学生提供传统的授课方式，也可以提供交互式学习方式和个人自学为主的个别化学习方式

- 远程教学系统

- 商业和企事业
- 出版业
- 医疗卫生\*\*

## 军事

## 虚拟现实

## 多媒体技术发展趋势

- 多媒体技术的发展更趋于多种业务融合和网络化
- 多媒体终端设备的嵌入式和智能化
- 多媒体技术在PC-TV融合方向的发展趋势
- 多媒体技术在安全领域和军事方面的应用和发展
- 媒体输入输出

## 多媒体技术特点

- (1) 集成性 能够对信息进行多通道统一获取、存储、组织与合成。
- (2) 控制性 多媒体技术是以计算机为中心，综合处理和控制多媒体信息，并按人的要求以多种媒体形式表现出来，同时作用于人的多种感官。
- (3) 交互性 交互性是多媒体应用有别于传统信息交流媒体的主要特点之一。传统信息交流媒体只能单向地、被动地传播信息，而多媒体技术则可以实现人对信息的主动选择和控制。
- (4) 非线性 多媒体技术的非线性特点将改变人们传统循序性的读写模式。以往人们读写方式大都采用章、节、页的框架，循序渐进地获取知识，而多媒体技术将借助超文本链接（Hyper Text Link）的方法，把内容以一种更灵活、更具变化的方式呈现给读者。
- (5) 实时性 当用户给出操作命令时，相应的多媒体信息都能够得到实时控制。
- (6) 信息使用的方便性 用户可以按照自己的需要、兴趣、任务要求、偏爱和认知特点来使用信息，任取图、文、声等信息表现形式。
- (7) 信息结构的动态性 “多媒体是一部永远读不完的书”，用户可以按照自己的目的和认知特征重新组织信息，增加、删除或修改节点，重新建立链

多媒体系统是指利用计算机技术和数字通讯网技术来处理和控制多媒体信息的系统。从广义上分，多媒体系统就是集电话、电视、媒体、计算机网络等于一体的信息综合化系统。

## 多媒体计算机的基本配置（及可选配置）

一般来说，多媒体个人计算机（MPC）的基本硬件结构可以归纳为七部分：

- ① 至少一个功能强大、速度快的中央处理器（CPU）；
- ② 可管理、控制各种接口与设备的配置；
- ③ 具有一定容量（尽可能大）的存储空间；
- ④ 高分辨率显示接口与设备；
- ⑤ 可处理音响的接口与设备；
- ⑥ 可处理图像的接口设备；
- ⑦ 可存放大量数据的配置等；

多媒体软件一般分为多媒体系统软件、多媒体工具软件、多媒体应用软件。

## chap4 无损数据压缩v网课练习版

---

对同一个信息，若使用的描述方法不同，则形成记录的数据量可能完全不同

信息≠数据

最小数据量

无损压缩和有损压缩（important）

信息量大的符号，信源的熵越大

概率大用较短编码，概率小用较长编码

香农-范诺编码

Huffman编码(最小生成树，哈夫曼树)

算术编码

RLE(run length encoding)

- 行程编码又称为游程编码，是一种统计编码
- 具有相同颜色并且是连续的像素数目称为行程长度

词典编码



- 用代码代替字符与字符串
- **词典是动态生成的**：词典开始为空，随着文件读入，将逐步增加条目，扩大词典尺寸。生成词典并按照词典把字符串信息转换成代码输出的过程，就是信息压缩过程。而根据词典，把得到的代码还原成字符串原样的过程，是解码过程。
- **词典不必传送**：解压程序依据词典的初始化信息和陆续接收到的代码将动态地生成与发送方一样的词典。
- 主要分为两类：
  - **\*\*第一类算法**：\*\*查找正在压缩的字符序列是否在以前输入的数据中出现过，若出现过，就用已经出现过的字符串替代重复的部分，输出是指向早期出现过的字符串的“**指针**”。如LZ77, LZSS
  - **\*\*第二类算法**：\*\*从输入的数据中创建一个“短语词典”。编码数据过程中当遇到已经在词典中出现的“短语” (它可以是任意字符的组合)时，编码器就输出这个词典中的短语的“**索引号**”，而不是短语本身。如LZ78, LZW
- **LZ77**
  - 相关术语：
    - 输入数据流、字符、编码位置、前向缓冲存储器、窗口、指针
  - **输出格式：“(Pointer, Length) Characters\*\*\*”**
- **LZSS算法**
  - 如果匹配串的长度比指针本身的长度长就输出指针，否则就输出真实字符。
- **LZ78算法编码思想(第二类)**
  - 不断地从字符流中提取新的缀-字符串(新“词条”)，用码字表示这个“词条”
  - 用**码字去替换字符流**，从而达到压缩目的。
  - 不断地从字符流中**提取新的缀-字符串(新“词条”)**，然后用**码字表示这个“词条”**。
  - LZ78编码器的**输出是码字-字符\*(W,C)对\***。
- **LZW算法编码（重点）**
  - 增加了一个术语—前缀根(Root)，它是由单个字符串组成的缀-字符串(String)
  - **只输出代表词典中的缀-字符串(String)的码字(Code Word)**（所以词典不能为空）
  - 所有可能出现的**单个字符都事先包含在词典中**
  - 词典也叫转换表、字符串表或编码对照表
  - LZW编码依靠词典来完成



- **LZW算法译码（重点）**

- 额外术语：当前码字、当前缀-符串、先前码字、先前缀-符串

## chap5 彩色数字图像基础

---

重复的数据，数学方法，**无损压缩技术**

超过极限的数据，部分去掉，**有损压缩技术**

实际的图像压缩是综合使用各种有损和无损压缩技术来实现的

人的视网膜有对红绿蓝颜色敏感程度不同的**三种**锥体细胞和杆状细胞

对不同**频率的光**的感知程度不同，对不同**亮度**的感知程度不同；

自然界中的任何一种颜色都可以有红绿蓝(RGB)三种颜色之和来确定，构成一个三位的矢量空间。

### 什么是数字图像

- **数字图像**，又称**数码图像**或**数位图像**，是二维图像用有限数字数值像素的表示。由数组或矩阵表示，其光照位置和强度都是离散的。数字图像是由模拟图像数字化得到的、以像素为基本元素的、可以用数字计算机或数字电路存储和处理的图像。

两个描述颜色的词：**颜色模型**和**颜色空间**

### 颜色模型

- 使用简单方法描述所有颜色的一套规则和定义，用数值指定颜色
- **RGB** (显示彩色图像用)
  - **有源物体**：一个能发出光波的物体称为有源物体，它的颜色由该物体发出的光波决定，使用RGB相加混色模型
  - 色彩三要素HSL：
    - 色调Hue
    - 饱和度Saturation
    - 亮度Lightness
- **CMY**(打印彩色图像用)
  - 青色cyan、品红magenta、黄色yellow、**K**：**Key plate** 因实务上多使用黑色，所以也可以简单视为black
  - **无源物体**：一个不发光波的物体称为无源物体，它的颜色由该物体吸收或者反射哪些光波决定，用CMY相减混色模型。减少了为视觉系统识别颜色所需要的反色光。

## 图像三个基本属性

分辨率、像数深度、真伪彩色和直接色

### 图象分辨率

- 屏幕分辨率：也称为显示分辨率，衡量显示设备再现图像时所能达到的精细程度
- 图像分辨率：也称为空间分辨率和像素分辨率，是图像精细化的度量程度

### 像数深度

- 指存储每个像素使用的位数，用来度量图像的分辨率，决定每个像素可能有的颜色数，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

### 真彩色

- 组成一个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，这样产生的彩色为“真彩色”。
- RGB8:8:8表示每个颜色用八位，称为真彩色或全彩色，**24位，16M，16777216**

### 伪彩色

- 每个像素的颜色值不是由每个基色分类的数值直接决定的，而是当做索引去查找颜色。

### 直接色

- 每个分量通过基色单独做变换

### 图像的种类

- 矢量图
  - 用一系列计算机指令来表示一幅图，如点线面以及其组合
  - 特点
    - **图形缩放后不变形**：矢量图形是以指令集合的形式描述的，图形缩放并不改变计算的公式，所以图形也就不会变形。
    - **占用的存储容量小**：由于矢量图形只是存储的图形指令，所以所需的存储空间很小。
    - **灵活性好**：由于构成矢量图形的各个部分是相对独立的，局部的处理不会影响其他部分。
    - **表现能力有一定局限**：对于复杂而又不易用数学方法描述的对象，矢量图形就无能为力。显示速度比点位图慢

- 位图

- 一幅彩色图像可以看成由许多的点组成的
- 文件字节数=图像分辨率X图像深度/8
- 特点
  - **逼真**：现实世界的任何物体，都是由许多微小的单元组成，点阵图由许多微小的像素构成，可以很真实的表现现实世界的任何对象。
  - **占用的存储容量大**
  - **缺乏灵活性**

- 灰度图

- 灰度图是只有明暗不同的像素而没有彩色像素组成的图像
- 按照灰度等级的数目来划分；只有黑白两种颜色叫单色图像

- 彩色图像

- $\gamma$ (伽马)校正

- 许多部件的光电转换特性是非线性的。输出和输入不成正比，但是他们之间有一个关系  
$$\text{输出} = \text{输入}^\gamma$$
- 为了在不同环境下观看到“原始场景”可在适当的地方加入 $\gamma$ 校正

- 直方图

- 二维的坐标:**横轴**代表的是图像中的亮度，由左向右，从全黑逐渐过渡到全白；**纵轴**代表的则是图像中处于这个亮度范围的像素的相对数量。
- 图片的“X光片”:直方图X横轴从左到右用0到255共256个数值来记录明度信息，其中0代表纯黑色的区域，255表示纯白色，128代表中性灰色。

- 曲线

- 曲线上比较陡直的部分代表图像对比度较高的部分。相反，曲线上比较平缓的部分代表图像对比度较低的区域。曲线向上弯曲会使图像变亮

## JPEG压缩编码

- JPEG是由ISO和IEC两个组织机构联合组成的专家组，开发**两种**基本的压缩算法
  - 一种是采用以离散余弦变换(Discrete Cosine Transform, DCT)为基础的**有损**压缩算法
  - 一种是采用以预测技术为基础的**无损**压缩算法
- 为了在保证图像质量的前提下进一步提高压缩比，JPEG 2000(简称JP 2000)标准中将采用小波变换(Wavelet)算法

- JPEG算法主要计算步骤
  - 正向离散余弦变换(FDCT)。
  - 量化(Quantization)。
  - Z字形编码(Zigzag Scan)。
  - 使用差分脉冲编码调制(Differential Pulse Code Modulation, DPCM)对直流系数(DC)进行编码
  - 使用行程长度编码(Run-length Encoding, RLE)对交流系数(AC)进行编码
  - 熵编码

## chap2 数字声音及MIDI简介

---

声音是携带信息的及其重要的媒体。物体在外力的作用下产生振动，对周围的空气产生变化的压强，并以类似波的运动形式在空气中传播，当声波传播到人的耳朵中而引起耳膜振动时，人所感觉到的就是声音

### 具有普通波的所有特性

- ①在不同介质的界面上能产生反射和折射，对各向同性介质的界面，遵守反射定律和折射定律；②通常的线性波叠加时遵守波的叠加原理③两束或两束以上的波在一定条件下叠加时能产生干涉现象；④波在传播路径上遇到障碍物时能产生衍射现象（见光的衍射）；⑤横波能产生偏振现象

声音的几个概念：

①声音的**强弱**：声波压力大小。

②音调的**高低**：声音频率。

③声音的**信号**：分类。

### 模拟信号和数字信号

将具有一定振幅和频率的连续变化的模拟声音信号，通过A/D转换器，以一定的频率对模拟音频信号截取一个振幅值，并用指定字长的二进制位表示，从而将连续的模拟音频信号转变成能被计算机处理的，离散的数字音频信号

为什么要使用数字信号的三点：

1)数字信号计算是一种精确的运算方法，它不受时间和环境变化的影响；

2)表示部件功能的数学运算相对容易实现；

3)对数字运算部件进行编程，如欲改变算法或改变某些功能，还可对数字部件进行再编程。

## 数字化过程的两个步骤：采样和量化

- 均匀采样；线性量化或者非线性量化

声音数字化需要回答两个问题：

- 每秒钟需要采集多少个声音样本，也就是**采样频率**
  - **奈奎斯特理论**：采样频率须大于等于声音信号**截止频率**（最高频率）的**2倍**，这样就能把以数字表达的声音还原成原来的声音，称为**无损数字化**
- 每个声音样本需要用多少位表示，也就是**量化精度**
- **均匀量化（线性量化）**
  - 量化值是均匀分布的
  - 量化阶距：有限个小幅度（量化阶距）的集合
- **非均匀量化（非线性量化）**

## 声音质量与数据率

- 与声音质量有关的重要因素
  - 采样频率
  - 量化精度（位数）
  - 声道数
- 声音文件的数据量 =  $(\text{采样频率} \times \text{量化精度} \times \text{声道数} \times \text{时间}) / 8$

采样频率单位：Hz，时间单位：秒，数据量单位：— 字节

## 根据声音的频带把声音质量分为五个等级

- 电话
- 调幅广播AM
- 调频广播FM
- 光盘CD
- 数字录音带DAT

## 常见声音文件的格式

- .WAV
- .VOC
- .AU

- .MID
- .SND
- mp3
- wma

## AV VCS: 语音变声器

### 声音质量的度量

- 客观质量度量: 信噪比SNR
- 主观质量度量: 主观平均判分法MOS

### 电子乐器数字接口(Musical Instrument Digital Interface, MIDI)

- 一套指令的约定, 不是声音信号。

### 波形音频和MIDI音频

#### 产生MIDI的方法

- 频率调制(FM)合成 (简称调频合成)
  - 理论上可以合成任何声音
  - 但是声音不够真实
- 乐音样本合成 (简称波表合成)
  - “波表查找”技术产生音乐
  - 波表技术以高精度录下真实乐器的声音, 并将录制的各种实际乐器的数字化声音存储在只读存储器(ROM)中
- **DIN连接器**是由德国标准化学会所制定的一项连接器标准。由于有太多DIN连接器标准, 如果不附加编号 (如"DIN 41524连接器"), DIN连接器一词并不代表任何一种连接器。
- **MIDI接口**
  - 使用屏蔽式双绞电缆, 电缆最大长度约15米
  - 所有DIN连接器只用了五个引线中的三个
  - 引线2用于屏蔽, 引线4和5用于传送信号电流
- MIDI可以有各种不同的连接方式, 链和环就是两种基本的形式
- 一个通道指一条独立的信息传输路线。单个物理MIDI通道分成16个逻辑通道, 每个通道访问一个独立的逻辑合成器, 可充当一种乐器的角色

### 声音如何存储

#### 领域1 语音/声音生成

deepfake是换脸的

Modulate.ai生成声音

领域2语音识别

领域3降噪效果和实时处理性能

- 与视觉方面的背景虚拟化相比，实时视频会议的噪声抑制更有难度

用声音隔空取物！利用声镊技术可以分离血细胞和肿瘤细胞

## chap3 话音编码

---

话音信号有高度的周期性

话音在短时间周期里(20ms数量级) 可以认为**基本不变准定态(Quasi-Stationary)**

话音编码器正是**利用这种周期性来压缩数据**。话音编码器就是企图揭示这种周期性，目的是为了减少数据率而又尽可能不牺牲声音的质量

**声音的数字化过程**：采样、量化、编码

**音频编码**

- 目的：减少存储空间，提高数据传输率，通过编码实现对音频数据的压缩
- 考虑的因素：不同声音信号的特点；系统的传输率；存储容量的限制；声音恢复的质量要求；系统实现的代价；声音的应用场合等
- 音频信号的处理过程：采样和编码、存储和传输、解码、播放

三大类编码方式

- 波形编译码器(Waveform Codecs)
- 音源(参数)编译码器(Source Codecs)
- 混合编译码器(Hybrid Codecs)

一般来说，波形编译码器的话音质量高，但数据率也很高；音源编译码器的数据率很低，产生的合成话音的音质有待提高；混合编译码器使用音源编译码技术和波形编译码技术，数据率和音质介于它们之间。

**脉冲编码调制PCM**

PCM是对音频信号采样、量化后不作其它处理直接以数字化格式存储。PCM可以选择不同的采样频率，但**并不对音频信号进行压缩**只要采样频率足够高，量化位数足够多，就能使解码后恢复的



声音信号有很高的质量

大信号—大阶距；小信号—小阶距

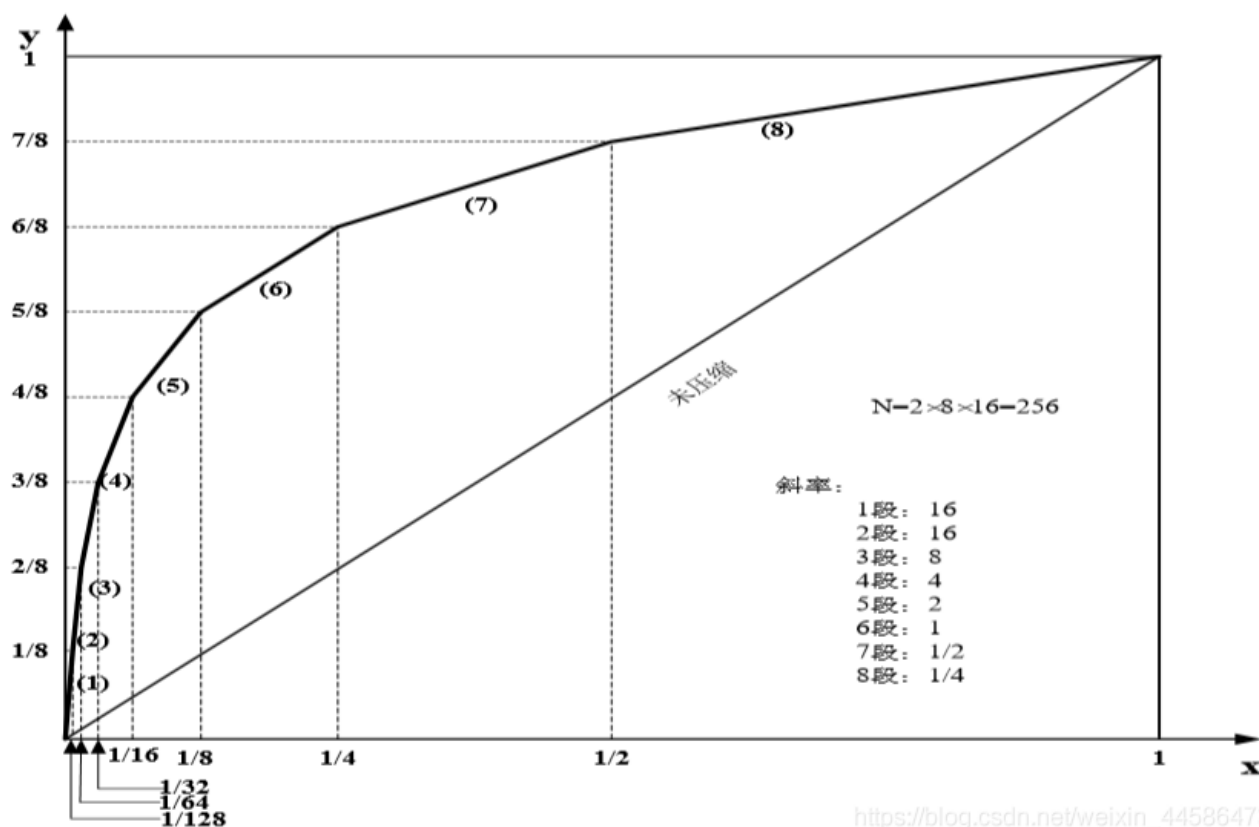
$\mu$ 律压扩算法和A律压扩算法：8位编码，采样频率8KHz

$\mu$ 律压扩算法

- 13位PCM码
- $F_{\mu}(x) = \text{sgn}(x) \ln \frac{(1 + \mu |x|)^{1/\mu}}{\ln(1 + \mu)}$ 
  - $\mu$ 为确定压缩率的参数，它反应最大量化间隔和最小量化间隔之比，其值在100-500之间
  - 对数PCM编码
  - |              |              |              |               |               |               |                |          |
|--------------|--------------|--------------|---------------|---------------|---------------|----------------|----------|
| <b>1/255</b> | <b>3/255</b> | <b>7/255</b> | <b>15/255</b> | <b>31/255</b> | <b>63/255</b> | <b>127/255</b> | <b>1</b> |
| 1/8          | 2/8          | 3/8          | 4/8           | 5/8           | 6/8           | 7/8            | 1        |
- 主要用在北美和日本等地区的数字电话通信中

A律压扩算法

- 14位PCM码
- 在 [0 - 1]区间不均匀地划分为8段，这不均匀的8段里面，我们再在每段里面均匀地划分出16小段
- 欧洲和中国大陆等地区的数字电话通信中



[https://blog.csdn.net/weixin\\_44586473](https://blog.csdn.net/weixin_44586473)

### 自适应脉冲编码调制(APCM)

- 根据输入信号幅度大小来改变量化阶距大小
- 量化阶大小的方法
  - 前向自适应(forward adaptation)
  - 后向自适应(backward adaptation)

### 差分脉冲编码调制(DPCM)

- 根据过去的样本值估算出预测值，然后对预测值和实际样本值的差值编码，从而降低音频数据的编码率，达到压缩目的。
- 存储或传送的是差值而不是幅度绝对值，降低了传送和存储的数据量。

### 自适应差分脉冲编码调制(ADPCM)

- 自动阶距+差值
- G.721 ADPCM压缩率 2:1

### 子带编码(SBC)

- 子带编码将整个音频频带分成若干个子带，分别对每个子带信号进行编码，然后由多路编码器将各子带编码复合后输出。
- 子带编码的优点
  - 对每个子带信号分别进行自适应控制
  - 根据每个子带信号在感觉上的重要性，对每个子带分配不同的位数，用来表示每个样本值

低频：小量化阶，多量化级数

高频：大量化阶，少量化级数

## G.722 SB-ADPCM编译码器

### 增量调制(DM)与自适应增量调制(ADM)

- **DM**（1位系统）存在的问题：
  - 斜率过载
  - 粒状噪声
- **自适应增量调制(ADM)**
  - 根据输入信号斜率的变化自动调整量化阶距的大小，以使斜率过载和粒状噪声都减到最小
  - 在检测到斜率过载时开始增大量化阶 $\Delta$ ，而在输入信号的斜率减小时降低量化阶 $\Delta$

G.721 推出于1984年，G.723 推出于 1988年。1990年二者整并成 G.726+

## chap6 颜色的度量体系

颜色是人的大脑对物体的一种感觉，而这种感觉又是带有极端的主观性，因此用数学方法来描述比较困难，目前好像还没有一种人类感知理论被普遍接受，因此还需要我们继续努力

- **颜色空间**：颜色空间是用空间中的点表示颜色的数学方法，人们用它来指定和产生颜色，使颜色形象化
- 为了适应不同的应用，人们已经构造了各种各样的颜色空间，如颜色空间的大小和类型有限的颜色空间；感觉上是线性的空间
- 设备相关和设备无关之分
- 现在采用的色彩空间变换有三种：YIQ，YUV 和 YCrCb。

### 颜色度量体系

- 颜色制或者叫做颜色体制， 实际上就是人们组织和表示颜色的方法。
- 组织和表示颜色的方法主要有两种：

- **颜色模型**(Color Model)

- RGB, HSB, CMY, CIE XYZ, CIELAB, CMYK和颜色的光谱描述方法都是颜色模型
- **颜色空间是颜色模型最普通的例子**：颜色空间在本质上是定义在某种坐标系统下的子空

间，空间中的每一个坐标表示一种不同的颜色

- **编目系统**(Cataloging System)

- **CIE色度图**(CIE Chromaticity Diagram)就是用平面表示的马蹄形曲线，它为大多数定量的颜色度量方法奠定了基础

### 三个特性

- **色调**：色调取决于可见光谱中的**光波的频率**，它是最容易把颜色区分开的一种属性。用于区别颜色的名称或颜色的种类。色调是视觉系统对一个区域呈现的颜色的感觉。
- **饱和度**：饱和度是相对于明度的一个区域的色彩，是指**颜色的纯洁性**，它可用来区别颜色**明暗的程度**。
- **明度**：**\*\*亮度（明度）是视觉系统对可见物体辐射或者发光多少的感知属性\*\***

### RGB模型

- RGB模型采用物理三基色，其物理意义很清楚，但它是一种**设备相关**的颜色模型。

### CIE 1931 XYZ

- 想象的相加基色**X, Y和Z**。
- **设备无关**的颜色系统
- X, Y和Z基色都是用**正数**去匹配所有的颜色，并且用**Y值**表示人眼对**亮度**的响应。
  1. 所有的X, Y和Z值都是正的，匹配光谱颜色时不需要一种负值的基色。
  2. 用Y值表示人眼对亮度(luminance)的响应。
  3. X, Y和Z是相加基色。每一种颜色都可以表示成X, Y和Z的混合。

### CIE 1931 xyY

- 从XYZ直接导出的一个颜色空间
- 使用**亮度Y**参数和**颜色坐标x,y**来刻画颜色

### CIE 1976 LUV

- 用数学公式把CIE 1931 XYZ中的x,y坐标变换到一个名u,v为的新坐标系

## CIE 1976 LAB

- CIELAB颜色系统**数学变换**后的颜色空间
- 使用最广泛的物体颜色度量方法，并作为度量颜色的国际标准。
  - CIE 1976 Lab\*颜色空间是 CIE 1931 XYZ颜色空间的一种**数学变换**的结果。
- 对色坐标(Opponent Color Coordinate)
  - 颜色不能同时是红和绿，或者同时是黄和蓝，但颜色可以被认为是红和黄、红和蓝、绿和黄以及绿和蓝的组合。

## 色彩空间，也叫色域

AdobeRGB 与 sRGB

## chap7 颜色空间变换

---

### 颜色空间的分类-感知角度

- 混合(Mixture)型颜色空间：按三种基色的比例合成颜色
- 非线性亮度/色度(Luma/Chroma)型颜色空间：用一个分量表示非色彩的感知，用两个独立的分量表示色彩的感知。当需要黑白图像时，这样的系统非常方便。
- 强度/饱和度/色调(Intensity/Saturation/Hue)型颜色空间

### 颜色空间的分类-技术角度

- 1、RGB型颜色空间/计算机图形颜色空间。
- 2、XYZ型颜色空间/CIE颜色空间：
- 3、YUV型颜色空间/电视系统颜色空间：

### 进行颜色空间的转换原因

- 人们已经开发了许多不同名称的颜色空间
- 可以说表示颜色的颜色空间的数目是无穷的
- 还没有一个完全符合人的视觉感知特性、颜色本身的物理特性或者发光物体和光反射物体的特性
- 各种不同颜色空间之间进行转换的目的各不相同
  - 有的是为了艺术家选择颜色的方便
  - 有的是为了减少图像的数据量

- 有的是为了满足显示系统的要求
- 有些颜色空间之间可以直接变换
- 有些颜色空间之间不能直接变换。它们之间的变换需要借助其他颜色空间进行过渡。

## chap8 彩色数字电视基础

---

电视有黑白电视和彩色电视之分，彩色电视用Y、C1、C2彩色表示法分别表示

### 彩色电视制式

- 世界上流行的一共有三种：NTSC、PAL和SECAM，他们互不兼容
- **NTSC**(正交平衡调幅制)
  - 美国,加拿大等大部分西半球国家及日本,韩国,菲律宾,和中国的台湾地区采用
- **PAL** (倒相正交平衡调幅制)
  - 由于NTSC存在相位敏感造成彩色失真的缺点.德国于1962年制定了逐行倒相正交平衡调幅制  
(Phase-Alternative Line)彩色电视广播标准,德国,英国,中国,朝鲜等采用。
- **SECAM**(顺序传送彩色与存储制)
  - 法国，前苏联及东欧国家采用，世界上大约有65各国家和地区在试用。
- **扫描线**
  - 电视的画面清晰度是以水平清晰度作为单位。清晰度的单位是“**电视行(TVLine)**”也称**线**。
- **电视的扫描特性**
  - **隔行扫描i**：一帧图像由两部分组成:奇数场和偶数场(扫描的行数必须是奇数)
  - **逐行扫描p**:扫一遍就显示一幅完整的图像
  - 电视采用隔行扫描，计算机显示器一般采用逐行扫描
- **电视扫描术语**：
  - **场频/场速率** (field rate)  $f_f$
  - **帧频**  $f_F$
  - **行频**  $f_H$

- PAL/SECAM和NTSC的有效扫描行每帧分别为576行和480行，而在水平方向BT. 601规定三种电视制式的有效水平采样点都是每行720个有效采样点。其余在行消隐和场消隐以及在行正程扫描中多出的采样点丢弃，因此PAL制BT. 601数字视频的有效图像大小为720 x 576像素，而NTSC制数字视频的有效图像大小为720x480像素。
- 用RGB彩色空间表示的彩色信号需经过变换电路转换到一个亮度信号Y和二色差信号C1、C2的彩色空间，再进行调制，然后进行传输。
- C1、C2的含义与具体应用有关:PAL中，C1、C2表示U、V；NTSC中C1、C2表示I、Q。不同的彩色空间在不同的应用中也许会比原始的RGB彩色空间具有更有用的特性。
- 彩色电视中使用Y、C1、C2彩色空间的优点

### 分量数字化有两种方法

- 先转换再分离
- 先分离再转换

### 图像子采样

- 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的采样称为图像子采样
- 图像子采样的理论依据与用途
  - 人的视觉系统的特点(图像子采样的理论依据):
    - 人眼对色度信号的敏感程度比对亮度信号敏感的程度低；
    - 人眼对图像细节的分辨能力有限，去掉一些高频信号后人眼不易察觉。
  - 图像子采样的用途
    - 根据人的视觉系统的特点，可以通过图像子采样实现彩色电视图像数据的压缩
- 4:2:2，平均每个像素用2个样本表示
- 4:1:1，平均每个像素用1.5个样本表示
- 4:2:0，平均每个像素用1.5个样本表示

## chap9 MPEG简介

---

如今，随着互联网带宽的提升和高速移动设备的普及，信息的获取途径和传播途径也在与日增加，视频逐渐成为互联网多媒体消费的重要组成部分。

### 视频专用神经网络的发展方向

- 视频检索（Video Search）专用神经网络
  - 视频检索、视频推荐是大型视频分享网站所要面对的首要问题，给予用户良好的检索体验，并适当地推荐用户需要的视频内容，帮助用户快速地找到自己感兴趣的视频。



- 视频分割（Semantic Video Segmentation）专用神经网络
  - 视频分割的目标在于像素级别地分割出视频中的人、车等常见物体。而随着AR/VR技术的发展，像素级别的目标识别可以辅助许多不同的AR/VR相关应用，这也促进了视频分割的发展。
- 视频生成（Video Generation）专用神经网络
  - 随着用户越来越容易地不同设备上进行视频拍摄，非专业用户对视频的自动/半自动编辑、美化也逐渐有了更多的需求。因此，视频生成神经网络便可以帮助用户编辑创作自己的视频作品

MPEG标准阐明了声音和电视图像的编码和解码过程，严格规定了声音和图像数据编码后组成比特数据流的句法，提供了解码器的测试方法等，但没有对所有内容都作严格规定，尤其是对压缩和解压缩的算法，这样既保证了解码器能对符合MPEG标准的声音数据和电视图像数据进行正确解码，又给MPEG标准的具体实现留有很大余地。毫无疑问，MPEG就是多媒体标准

## MPEG与MPEG标准

- MPEG（Moving Picture Expert Group）活动图像专家组（运动图像专家组）的英文缩写。MPEG是1988年国际标准化组织(ISO)和国际电工委员会(IEC)联合成立的专家组，负责开发电视图像数据和声音数据的编码、解码和同步等标准。
- MPEG标准：MPEG开发的标准称为MPEG标准。

## MPEG视频的层次

### 标准之间存在差异

### 运动图像的压缩基本概念

- 运动图像数据压缩主要根据下面两点
  - 运动图像序列中有许多**重复的数据**，使用数学方法来表示这些重复数据就可以减少数据量
  - 人的眼睛对图像细节和颜色的**辨认有一个极限**，把超过极限的部分去掉，这也就达到了压缩数据的目的。
- 利用前一个事实的压缩技术就是无损压缩技术，利用后一个事实的压缩技术就是有损压缩技术。实际的图像压缩是综合使用各种有损和无损压缩技术来实现的

### 运动图像压缩技术

- 在空间方向上（静态图像）：图像数据压缩采用JPEG（Joint Photographic Experts Group）压缩算法来去掉冗余信息。

- 在时间方向上（图像变化）：图像数据压缩采用移动补偿（Motion Compensation）算法来去掉冗余信息。

## **MPEG标准列表**

- **MPEG-1**：运动图像数字压缩标准，1992年正式发布。（MP3， VCD）
- **MPEG-2**：运动图像数字压缩标准。（DVD， SVCD）
- **MPEG-3**：已于1992年7月合并到高清晰度电视（High-Definition TV， HDTV）工作组。
- **MPEG-4**：运动图像数字压缩标准及多媒体应用标准（1999年发布）。
- **MPEG-5**：直至1998年9月还没有见到定义。
- **MPEG-6**：直至1998年9月还没有见到定义。
- **MPEG-7**：多媒体内容描述接口标准。
- **MPEG-21**：多媒体框架和综合应用方面的框架。该标准致力于在大范围的网络上实现透明的传输和对多媒体资源的充分利用。

## **MPEG-1数字电视标准**

- 发布时间：1992年。
- 设计目标：处理标准图像交换格式(SIF)的电视

## **MPEG-2数字电视标准**

- 发布时间：1994年
- 设计目标：与数字电视广播有关的高质量图象和声音编码标准。能提供的位传输率至少达到4-9Mbits/sec间，最高可达15Mb/s
- 主要是用来定义电视图像数据、声音数据和其他数据的组合，把这些数据组合成一个或者多个适合于存储或者传输的基本数据流。

## **MPEG-3标准**

- 由于MPEG-2的出色性能表现，已能适用于HDTV，使得原打算为HDTV设计的MPEG-3，还没出世就被抛弃了

## **MPEG-4多媒体应用标准**

- 发布时间：1998年11月
- 设计目标：在异构网络环境下能够高度可靠地工作，并具有很强的交互能力和灵活性

## 内容基

- 与MPEG-1、MPEG-2标准的区别
  - 内容基 (Content – Based) 的压缩编码方法，突破了MPEG-1、MPEG-2以矩形块处理图像的方法。
  - 对一幅图像按内容分割成不同的子块，将感兴趣的物体从场景中截取出来，进行编码处理。
  - 有利于内容的交互处理，同时基于内容截取的子块内的信息相关性很强，可获得数据的高压缩率。
- MPEG-4引入了视听对象AO、VO的概念
  - VO的构成依赖于具体应用和系统所处的环境
  - 对低要求应用，VO可以是一个矩形帧，从而与原来的标准兼容
  - 对基于内容表示要求较高的应用，VO可能是场景中的某一物体或某一层面，也可能是计算机产生的二维、三维图像等。
- MPEG-4扩充了编码的数据类型，由自然数据对象扩展到计算机生成的合成数据对象，采用了合成对象/自然对象混合编码算法；

## DMIF (传输多媒体集成框架)

- 传输多媒体集成框架(Delivery Multimedia Integration Framework, DMIF)的会话协议，它用来管理多媒体数据流
  - 该协议在原则上与文件传输协议FTP(File Transfer Protocol)类似
  - 其差别是：FTP返回的是数据，而DMIF返回的是指向到何处获取数据流的指针
  - DMIF覆盖了三种主要技术：广播技术，交互网络技术和光盘技术

## MPEG-7多媒体内容描述接口

背景：信息存在的形式也不断丰富，除文本以外，声音、图形、图像等多媒体信息载体被广泛地应用。

### 需要解决信息定位问题

- 没有很好的声音图像信息的检索工具，是因为对这类信息还没有普遍认同的、高效的描述方法
- 旨在解决对多媒体信息描述的标准问题。
- 只有首先解决了多媒体信息的规范化描述之后，才能更好地实现信息定位

MPEG-7的工作于1996年启动。目的是制定一套描述符标准，用来描述各种类型的多媒体信息及它们之间的关系，一遍更快更有效的检索信息。而不是视听数据压缩标准。

- 多媒体材料可以包括静态图像、图形、3D模型、声音、话音、电视以及他们之间的额组合关系
- MPEG-7的“多媒体内容描述接口”通过给图像、声音和图像数据赋予定义的方法来提高再利用性。具体的讲就是给多媒体信息**附加上元数据**，以便可以按照各个画面的特征和重要度进行检索和处理。在这一方面多媒体研究所进行了大量的开发和研究。

MPEG-7将提供内容的描述而不是内容本身（MPEG-1，MPEG-2，MPEG-4），仅仅是已有3个标准的补充。

## MPEG-21(多媒体框架标准)

背景:多媒体的标准层出不穷，这些标准涉及到多媒体技术的各个方面。开发**胶水标准**

# chap14 MPEG声音

---

## 感知声音编码

人主要用耳朵接受声音

**心理声学**：把研究听觉系统如何感知声音波形的幅度和频率的科学称为心理声学。

## 听觉系统的感知三个特性(重点)

- **响度、音高（音调）、掩蔽效应**
- **对响度的感知**
  - **听阈**：当声音弱到人的耳朵刚好可以听到时的声音强度称为听阈。听阈是随频率变化。听阈-频率曲线
  - **痛阈**：当声音强到人的耳朵感到疼痛时的声音强度称为痛阈。痛阈也随频率变化。痛阈-频率曲线
  - **听觉范围**：“听阈-频率”曲线与“痛阈-频率”曲线之间的区域。
  - **主观感觉与频率紧密相关**：2kHz～4kHz范围的声音最敏感，幅度很低的声音信号都能听到；
- **对音高的感知**
  - 音高的**客观度量与主观度量**
    - 客观度量：振动速度(频率)表示声音的音高，单位Hz
    - 主观度量：用人耳对音高的主观感觉确定，单位Mel(美) $Mel=1000\log_2(1+f)$
- **掩蔽效应**

- 一种频率声音阻碍听觉系统感受另一种频率的声音
- 掩蔽效应是心理声学模型的基础
- 掩蔽分类
  - 频域掩蔽
    - 把20Hz~16kHz范围内划分成24个临界频带，单位为Bark
    - 1bark=一个临界频带的宽度.
      - $f < 500\text{Hz}$ 下,  $1\text{bark} = f/100$ .
      - $f > 500\text{Hz}$ 下,  $1\text{bark} = 9 + 4\log_2(f/1000)$
  - 时域掩蔽

## MPEG Audio与感知特性

- MPEG Audio是指MPEG1Audio, MPEG2Audio和MPEG-2AAC, 处理10~20000kHz范围里的声音数据。
- MPEG Audio使用心理声学模型来实现压缩声音数据的目的
- MPEG Audio主要采用感知自带编码技术
  - 时域-频域变换
  - 心理声学模型
  - 位分配、量化和编码
- **MPEG-1 Audio的编码层**
  - MPEG Audio压缩编码定义了**3个层次**, 3个层次的基本模型相同。
  - 层1最基础, 每个后继层次比前面的层次压缩比更高, 同时编码解码器也更复杂。
- **MPEG-2 Audio相对于MPEG-1所作的扩充**
  - 增加了16kHz, 22.05kHz和24kHz采样频率;
  - 扩展了编码器的输出速率范围: 由32~384kb/s扩展到8~640kb/s;
  - 增加了声道数: 支持5.1声道和7.1声道的环绕声。
- **MPEG-2 AAC**
  - 一种非常灵活的声音感知编码标准。

- 主要使用听觉系统的掩蔽特性来减少声音的数据量，并且通过把量化噪声分散到各个子带中，用全局信号把噪声掩蔽掉

- **MPEG Audio**采用两种感知编码

- 感知子带编码
- 杜比实验室开发的DolbyAC-3编码

## **MPEG-4 Audio**

### 主要特点

- 应用范围广泛，支持从话音到高质量的多通道声音，从自然声音到合成声音的处理。
- 特别是它能有效地表示音频对象特征
- 对自然声音的处理
  - 参数编码器
  - CELP编码器
  - 时间/频率编码器
- 对合成声音的处理
  - **TTS(文-语转换)系统**
    - 将文本形式的信息转换成自然语音，其最终目标是使计算机输出清晰而又自然的语音
    - 使计算机像人一样，根据文本的内容可带各种情调来朗读任意的文本
    - TTS是一个十分复杂的系统，涉及到语言学、语音学、信号处理、人工智能等诸多的学科

## **chap11 MPEG及视频直播等技术**

---

为什么巨大的原始视频可以编码成很小的视频呢？这其中的技术是什么呢？

- 视像数据之所以能被压缩，主要是数据中存在大量的冗余数据
- 包括时间冗余、空间冗余、结构冗余、视觉冗余和数据冗余

帧和帧间？

### **MPEG**

#### **MPEG Video的数据压缩依据和方法**

- **基本思想和方法要点：**
  - 在空间方面，对单帧图像数据，采用**JPEG压缩**算法来去掉帧内的冗余信息；

- 在时间方面，对连续的多帧图像数据，采用**移动补偿算法**来去掉帧间的冗余信息。
- **单帧**：JPEG压缩 **多帧图**：移动补偿算法
- **运动图像分类及排序**
  - **帧内图** (Intra Pictures **I**)
  - **预测图** (Predicted Pictures **P**)
  - **插补图** (Bidirectional Prediction **B**) 即双向预测图
- Group of Pictures(GoP)
- **内帧(I)**
  - 采用JPEG算法进行中等程度的压缩压缩
  - 内帧压缩算法在编码解码时不依赖其它任何帧，可作为随机访问点
  - 帧内图像I的压缩编码类似JPEG压缩算法)
- **预测帧(P)**
  - 以前面的参照帧(I或P)为基础进行预测编码和运动补偿法压缩存储，它本身又可作为后面预测帧的参照帧。
  - 编码与JPEG类似
  - 以图像宏块为基本编码单元，一个宏块定义为I×J像素的图像块
  - 以图像宏块为基本编码单元，一般16×16
  - **对P图像的压缩编码**
    - 当前要编码的图像宏块与参考图像的宏块之间的**差值**
    - **宏块的运动矢量**
  - **查找与参考宏块最佳的预测图像编码宏块**
    - **寻找最佳宏块的三种简化算法**
      - 二维对数搜索法(2D – Logarithmic Search)
      - 三步搜索法(Three-step Search)
      - 对偶搜索法(Conjugate Search)
- **双向预测帧(B)**
  - 双向预测帧又叫内插帧
  - 需要前后两个参照帧：
    - 采用双向预测编码和运动补偿法来进行图像编码
    - 本身不能作为参照帧，只是I和P之间的内插帧
  - **压缩比最高**
    - 内帧（I 帧）是基础
    - 定时提供一帧I画面，称为刷新
- 帧内图像的**I**的**数据量最大**，双向预测图像**B**的**数据量最小**
- **视频图像的结构**
  - 算法允许选择内帧I的频率及位置



- 允许选择在一对内帧或预测帧P之间插入双向预测帧B的数目。
- 内帧I是基础，定时提供一帧I画面，称为刷新
- 传统插帧方法优劣势一览
  - 帧采样
  - 帧混合
  - 运动补偿
  - 光流法
  - **Super SloMo：AI 插帧法**

## 视频直播技术

深度造假（deepfaking）

## chap 显示技术

---

### 显示器分类

- CRT（Cathode Ray Tube）显示器
- LCD（Liquid Crystal Display）显示器
- PDP（Plasma Display Panel）等离子显示器
- 电子发光显示器

**LCD是液晶显示屏Liquid Crystal Display**的全称，主要有TFT、UFB、TFD、STN等几种类型的液晶显示屏。

LED（Light-emitting diode）发光二极管屏幕，其实这些显示器的影像仍是以液晶产生，发光二极管只是当作光源，在技术上仍是LCD显示器，或叫LED背光液晶显示电视。

OLED（Organic Light-Emitting Diode）有机发光二极管，它和LED仅仅虽然只有一个字母之差，但实际上两者描述的是完全不同的事物。

### 背投

### 投影仪

### 等离子电视

### 液晶电视

### LED显示器

### OLED

曲面屏电视

视网膜屏

PPI

3D显示原理与技术

被动式3D和主动式3D。

人体学接口设备（英語：Human interface device，缩写HID

VR三个关键词

- 可交互
- 三维（3D）
- 沉浸

AR（增强现实）

MR混合现实

## chap您好！色彩！

---

色彩心理学

- 年龄
- 性别差异：男性与女性
- 色彩的文化差异

传统的配色方案类型

- 基本的12辐色轮是创建配色方案的重要工具。

原色

色彩冷暖

- 暖色系
- 冷色系
- 中性色系

调色，掌握相邻色和补色的关系是调色的重要基础

增加一种颜色：增加相邻色or减少互补色

三角色也是一组颜色。

分散的互补色

四方色

四方补色

## 接口

**VGA接口：**VGA接口也称为D-Sub接口。在CRT显示器时代，VGA接口是必备的。因为CRT是模拟设备，而VGA采用的也是模拟协议，所以它们理所当然被匹配来使用。VGA接口采用15针插针式结构，里面传输分量、同步等信号，是很多老显卡、笔记本和投影仪所使用的接口。后来液晶显示器出现，也带有VGA接口。显示器内置了A/D转换器，将模拟信号转换为数字信号在液晶显示器上显示。由于线材与信号干扰等一系列问题，VGA使用时一般仅能够达到1080p分辨率，在高分辨率下字体容易虚，信号线长的话，图像有拖尾现象。在数字设备高度发展的今天，VGA接口已逐渐退出舞台，一些显示器也不再带有VGA接口。明基专业显示器系列中，也仅有入门级的BL2420PT保有了VGA接口。

**DVI（Digital Visual Interface）接口：**DVI的种类非常多，接口上有DVI-A、DVI-D、DVI-I，又可分为单通道与双通道。**\*\*DVI-A（DVI-Analog）\*\***接口只传输模拟信号，实质就是VGA模拟传输接口规格，常用于转接显卡的DVI-I输出到VGA显示器接口。**\*\*DVI-D（DVI-Digital）\*\***接口是纯数字接口，不兼容模拟信号。**\*\*DVI-I（DVI-Integrated）\*\***接口，兼容DVI-I和DVI-D两种插头，兼容数字和模拟信号。它们的接口形状不同，如果接口不匹配就无法插入使用。



**HDMI（高清晰多媒体）接口：**HDMI接口在2002年提出，现在已经发展到HDMI 2.1标准，而且随着行业发展，HDMI 2.1标准已经能够支持4K 120Hz及8K 60Hz，支持高动态范围成像（HDR），可以针对场景或帧数进行优化，向后兼容HDMI 2.0、HDMI 1.4。最主要的是，它是视音频同时传输的。

**DP接口：**DP（Display Port）接口是HDMI接口的竞争对手。在一些显卡与显示器上，除了有HDMI接口，也带有DP接口。目前DP接口发展到1.4版本，能传输10bit的4K 120Hz视频，也可以

支持8K 60Hz视频。DP1.4兼容USB Type-C接口，这就意味着，我们可以使用DP1.4协议，在USB 3.1传输数据的同时，同步传输高清视频。从趋势来看，这对于推动周边影音及影像设备的升级，是非常有好处的。而且，DP接口也可以单根线同步传输声音，达到使用显示器声音的使用方式。

元宇宙是整合多种新技术而产生的新型虚实相融的互联网应用和社会形态，它基于扩展现实技术提供沉浸式体验，基于数字孪生技术生成现实世界的镜像，基于区块链技术搭建经济体系，将虚拟世界与现实世界在经济系统、社交系统、身份系统上密切融合，并且允许每个用户进行内容生产和世界编辑。元宇宙仍是一个不断发展、演变的概念，不同参与者以自己的方式不断丰富着它的含义

## 同学展示

---

Snapseed是iOS和安卓系统的一款图像处理软件和剪影。剪映是抖音官方推出的一款手机视频编辑剪辑应用。带有全面的剪辑功能，支持变速，多样滤镜效果，以及丰富的曲库资源。

VTube Studio是一款可用于iPhone/iPad,Android,macOS与Windows的应用。该软件使用你的手机或网络摄像头进行面部捕捉，并相应地对Live2D Cubism 模型进行动画处理。

Audacity是一个跨平台的声音编辑软件，用于录音和编辑音频。

Avatarify，是一款变脸软件，可以将别人的脸套在自己的表情上进行实时驱动。

DeepVocal（简称：DV）是由boxstar开发的免费，自由的歌声合成软件，同时也是一个全球虚拟歌手创造企划。

Snapchat（色拉布）是由斯坦福大学两位学生开发的一款“阅后即焚”照片分享应用。利用该应用程序，用户可以拍照、录制视频、添加文字和图画，并将他们发送到自己在该应用上的好友列表。这些照片及视频被称为“快照”（"Snaps"），而该软件的用户自称为“快照族（snubs）”。

PhotoZoom是一款新颖的、技术上具有革命性的对数码图片进行放大的工具。

Procreate是一款运行在iPadOS上的强大的绘画应用软件。LazyShare是一款简单的绘画素材分享软件

Lr 也同样不仅仅是一个照片编辑软件，它更是一个照片管理软件，它的特点是以工作流为主，我们可以一个操作就作用于选定的多张照片，可以一次性的处理很多照片，然后再分别对每一张照片进行细调整。

Adobe After Effects是一款特效制作软件，包括影视特效，栏目包装，动态图形设计等等。所以AE一般应用在原创视频的制作上。和三维软件结合使用，可以使作品实现更加绚丽的效果。