

多媒体与智能交互复习

重要考点

- 点位图图像大小 = 像素 * 颜色深度/8
- JPEG 算法中，颜色子采样和量化（去除高频信号）是有损的
- 色彩空间的几种常见名词的比较

1. RGB vs. CMYK:

- a. **RGB:** 用于显示设备，通过光的加法混色。
- b. **CMYK:** 用于印刷，通过颜料的减法混色。

2. sRGB vs. AdobeRGB:

- a. **sRGB:** 较小色域，适用于互联网和消费级设备。
- b. **AdobeRGB:** 较大色域，适用于专业图像处理和打印。

3. LAB vs. RGB/CMYK:

- a. **LAB:** 设备无关，基于人眼感知，色域最大，用于高精度颜色处理。
- b. **RGB/CMYK:** 设备相关，分别用于显示和打印。

4. CIE:

- a. **CIE XYZ:** 标准色彩空间，作为其他色彩空间的基础，用于颜色转换和标准化。

- 声音文件大小

声音文件的数据量 = 采样频率 × 量化精度 × 声道数 × 时间 / 8

样频率单位： - Hz

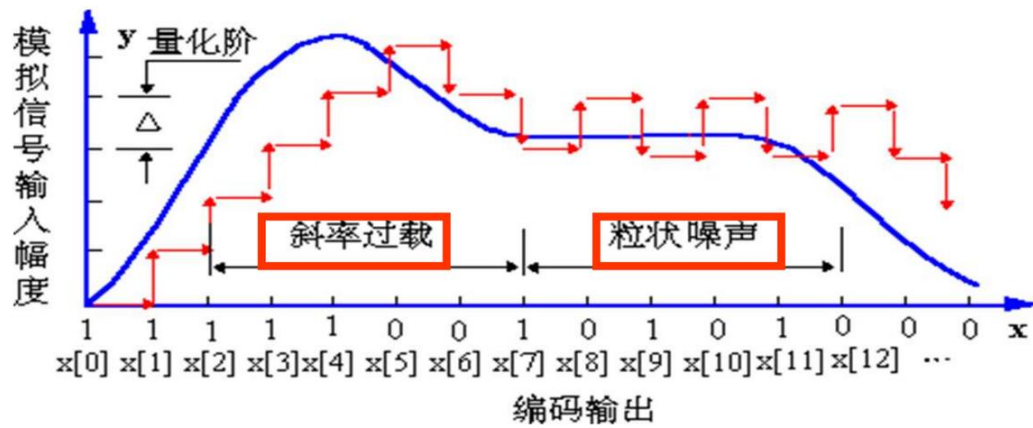
时间单位： - 秒

数据量单位： - 字节

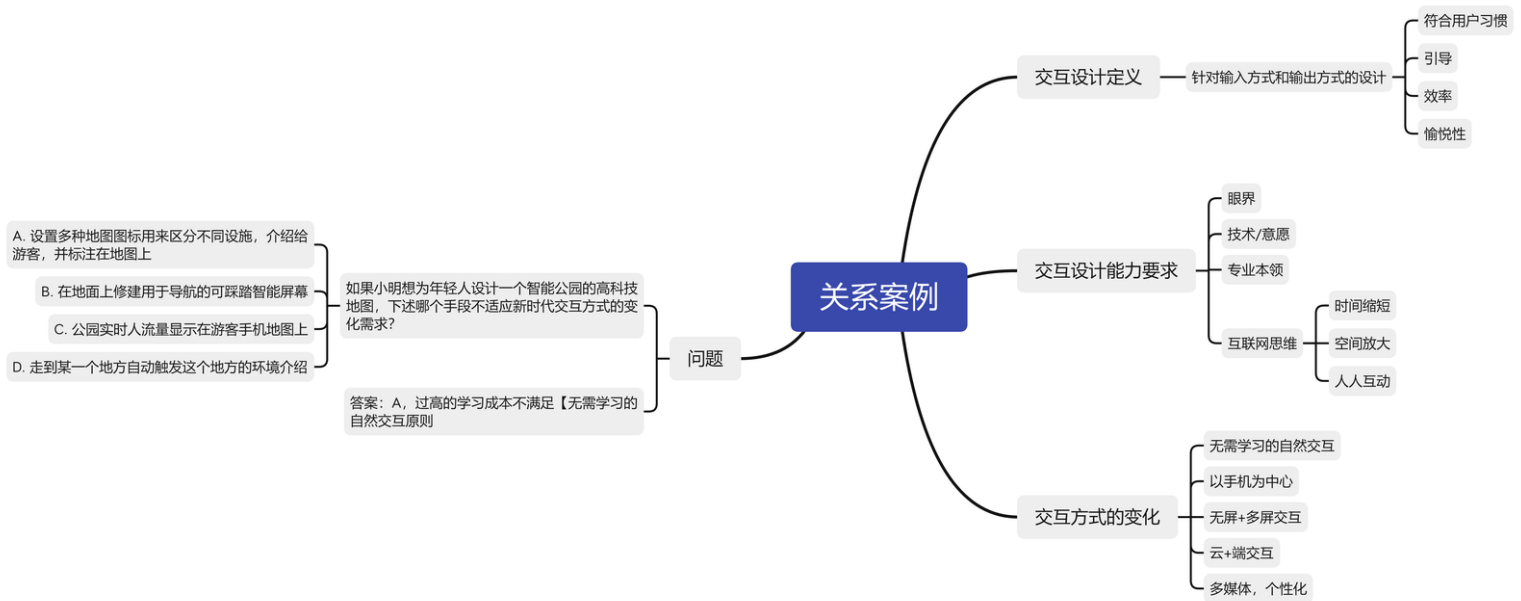
-
- 心理声学模型 = 听阈 + 频率遮蔽 + 时间遮蔽
- MPEG 利用听阈（排除听阈外的声音）和遮蔽效应（排除被遮蔽的声音）
- DM 的主要问题

■ DM的主要问题：

- 1) **斜率过载**：输出代码不能保持跟踪输入信号的快速变化。
- 2) **粒状噪声**：输入信号与预测值的差值接近0的区域，输出出现随机交变的“0”和“1”的现象。



○ 智能交互



基础理论技术

基础理论技术

什么是多媒体

- 定义
 - 融合两种及以上的人机交互的信息交流和传播媒体
- 分类方法
 - 国际电联定义的五种类型
 - 按照感官分类

更多多媒体

- 超媒体 = 多媒体 + 超文本
- 新媒体 = 数字技术 + 网络 + 所有人传播
- 跨媒体 = 信息在不同媒体间的流动与互动
- 网络媒体，移动媒体，自媒体，流媒体。。。

发展趋势

- 基于内容的处理
 - 图像理解
 - 语音识别
 - 全文检索
- 多媒体+网络
- 多媒体+只能交互
- 虚拟现实

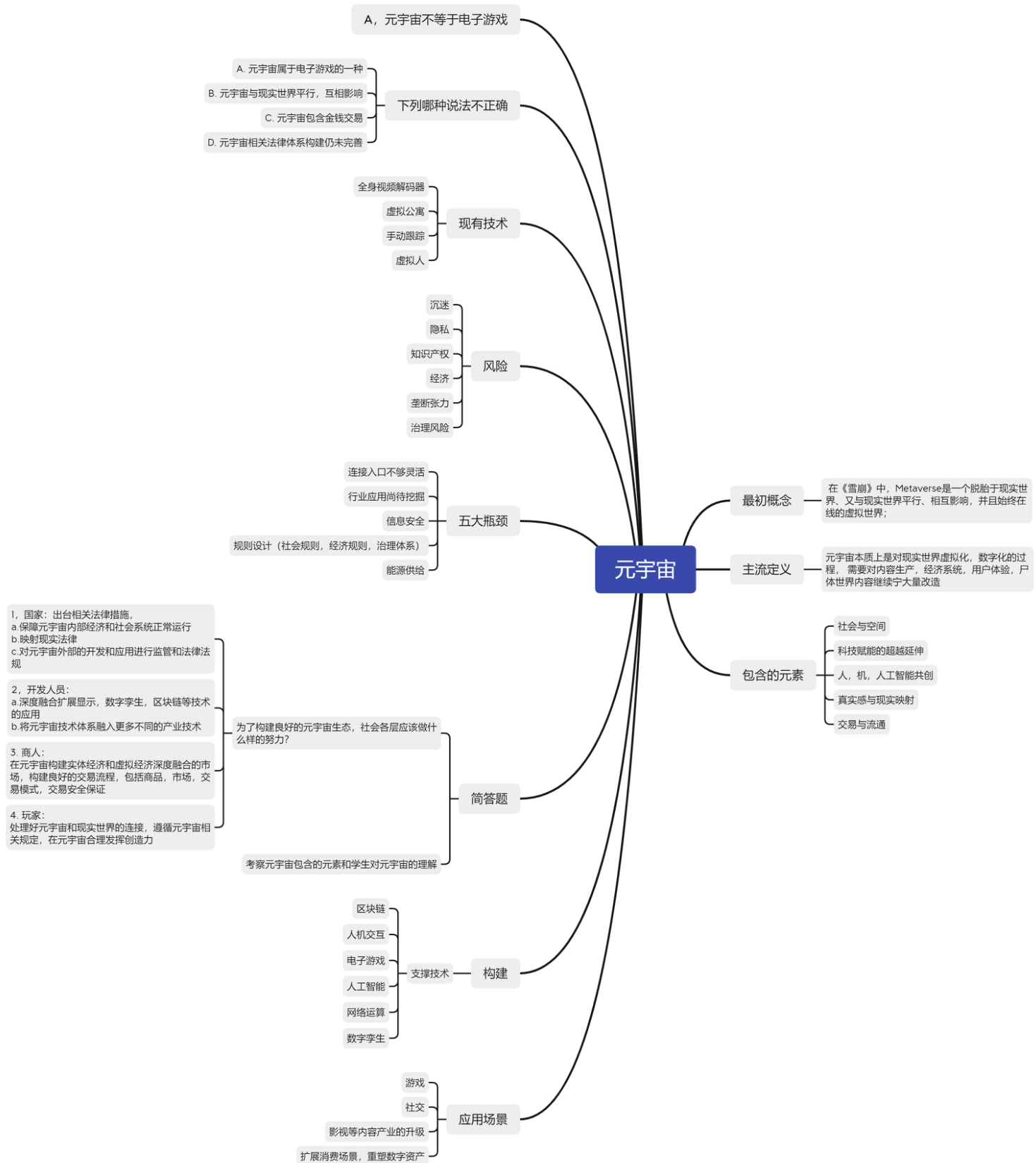
问题

“对来自微博、多媒体分享网站和新闻网站中的数据进行分析，从各关联数据中获取共同表达的高层语义”属于哪种媒体重点研究的范畴？

答案： C，这是研究不同来源数据之间的跨越形式，是跨媒体重点研究的范畴

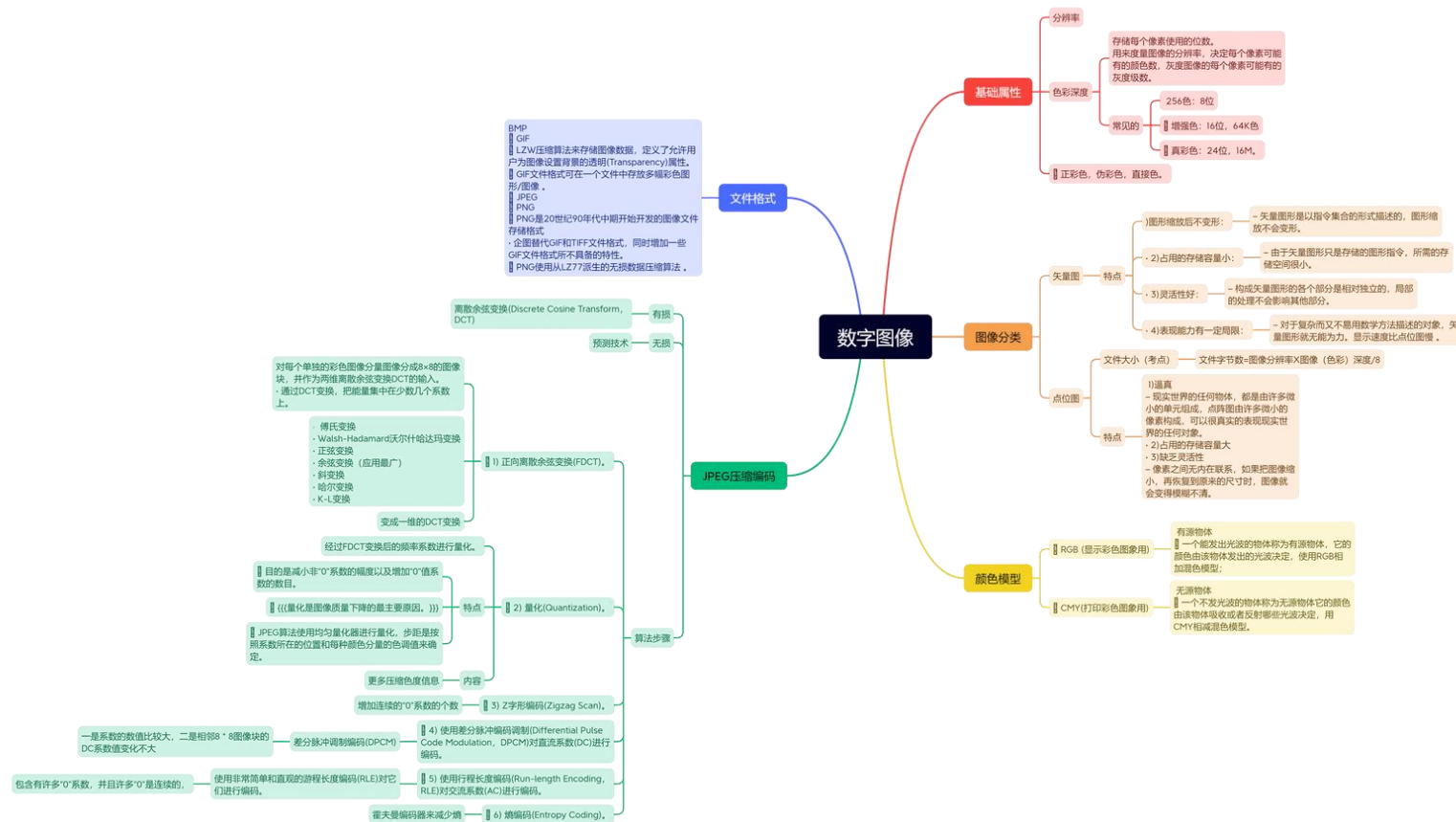
- 超媒体
- 新媒体
- 跨媒体
- 自媒体

元宇宙



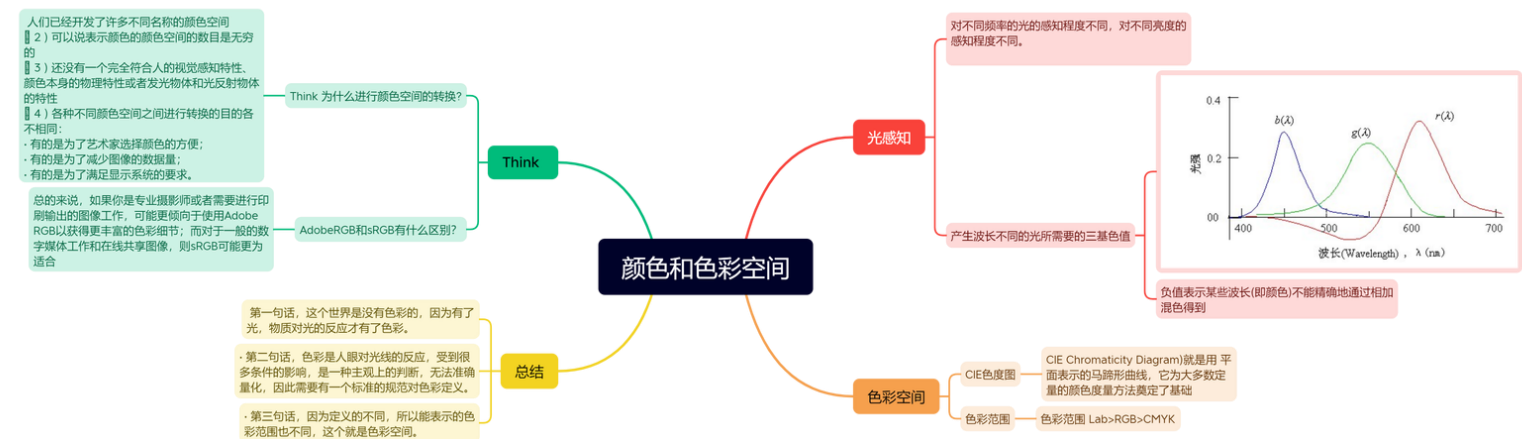
图像

数字图像



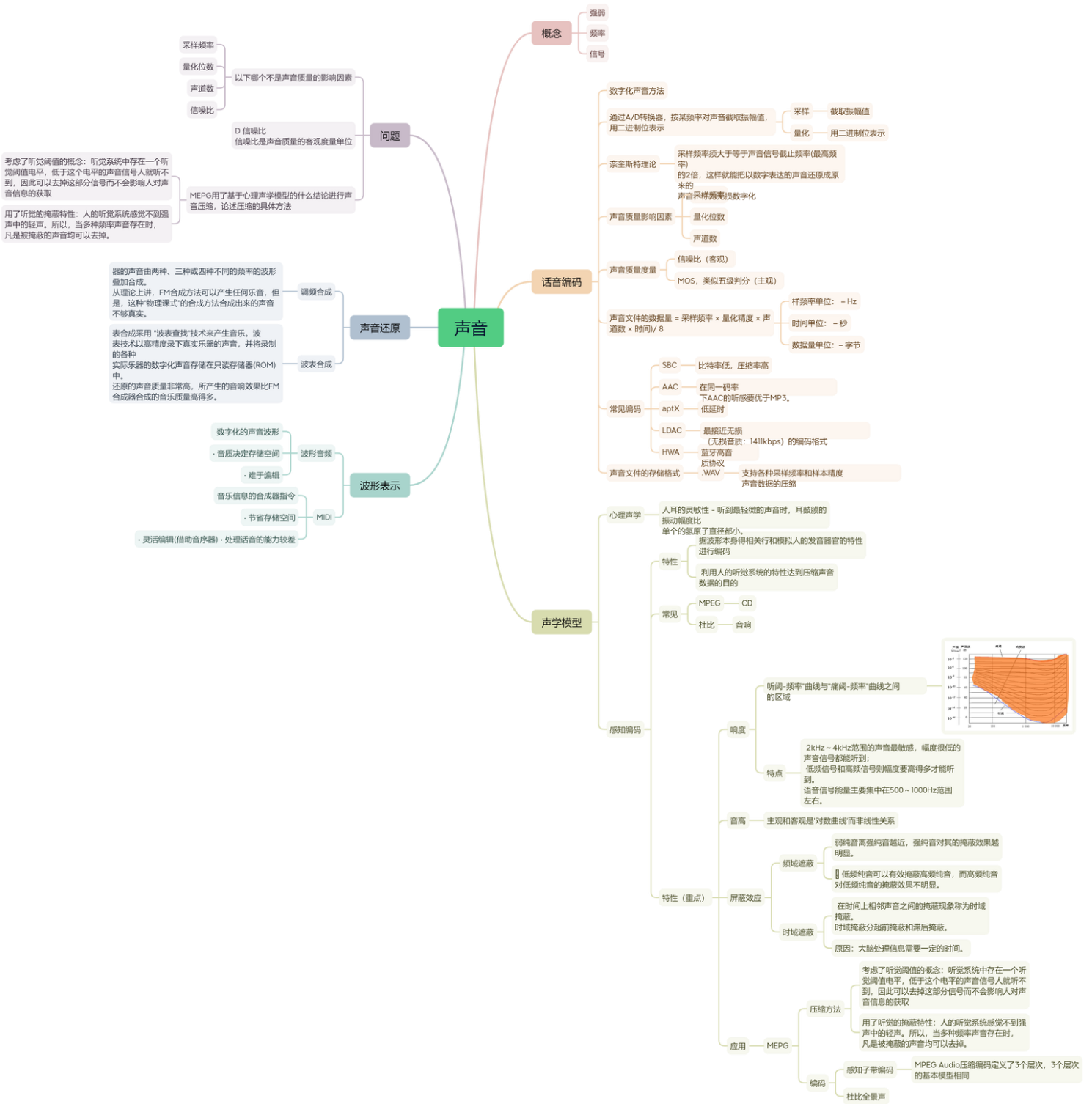
Presented with xmind

颜色空间

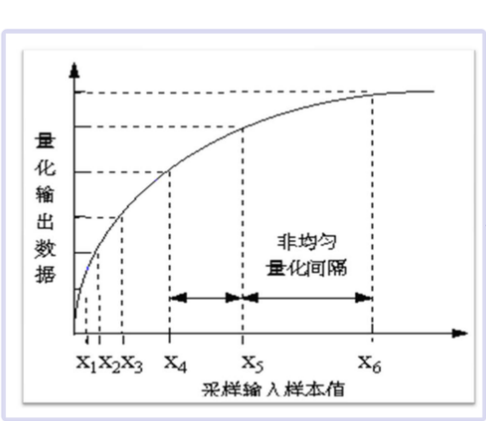


Presented with xmind

声音



话音编码



话音编码

非均匀量化

子带编码

子带编码将整个音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。

低频	高频
小-量化阶 多-量化级数	大-量化阶 少-量化级数
低频率带中，为了保护音调和均衡级的地位，故数采用较小的量化阶，较多的量化级数。	音中的摩擦音和类似噪声的声音，通常出现在高频子带中，对它分配较少的级数。

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。
- 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。
- 基本思想
 - 1. 差分量化：利用相邻样值之间的差值进行量化。
 - 2. 自适应量化：使量化器自动适应输入电平的变化，随时估计输入信号的时变幅度，以修正量化阶距，当输入信号与预测信号的差值小时，使用小的量化阶距，差值大时，使用大的量化阶距。

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。
- 特点
 - 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。
- 均匀量化

增量调制(DM)
自适应增量调制(ADM)

3.4 增量调制(DM)与自适应增量调制(ADM)

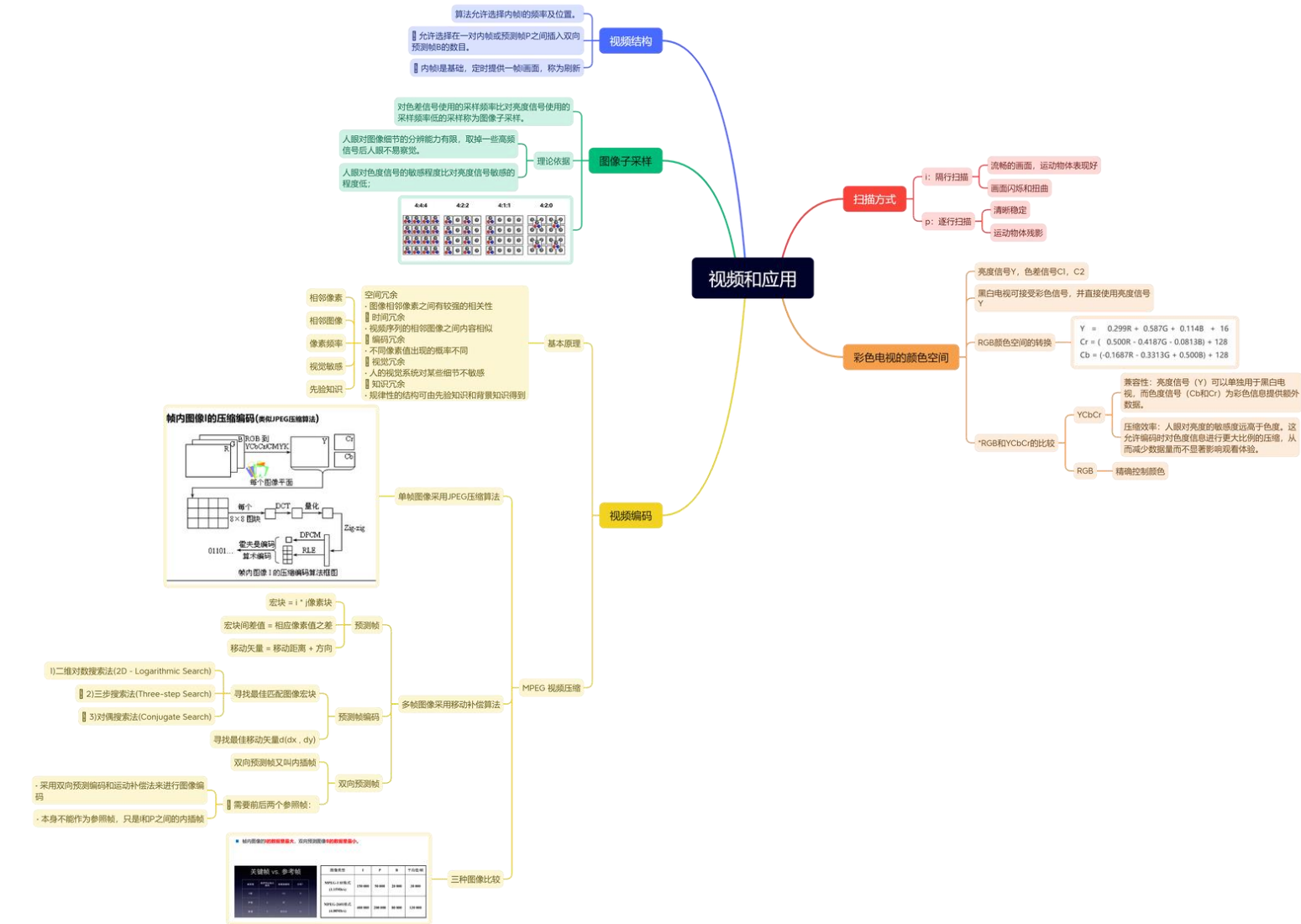
- 增量调制(Δ调制)
 - 也称Δ(delta modulation, DM)，调制，是一种预测编码技术。
- 基本思想：
 - 根据当前信号预测下一个信号，产生一个预测值，对实际值与预测值之差进行编码。
 - 输入信号 **实际值** - **预测值**：x[n] = 1
 - 输入信号 **实际值** - **预测值**：x[n] = 0
- 编码长度：**1位**

$y[i+1] = y[i] \pm \Delta$
+：当前值 > 预测值
-：当前值 < 预测值

- 主要问题
 - 1) 斜率过载：输出代码不能保持跟踪输入信号的快速变化。
 - 2) 粒状噪声：输入信号与预测值的差值接近0的区域，输出出现随机交变的“0”和“1”的现象。

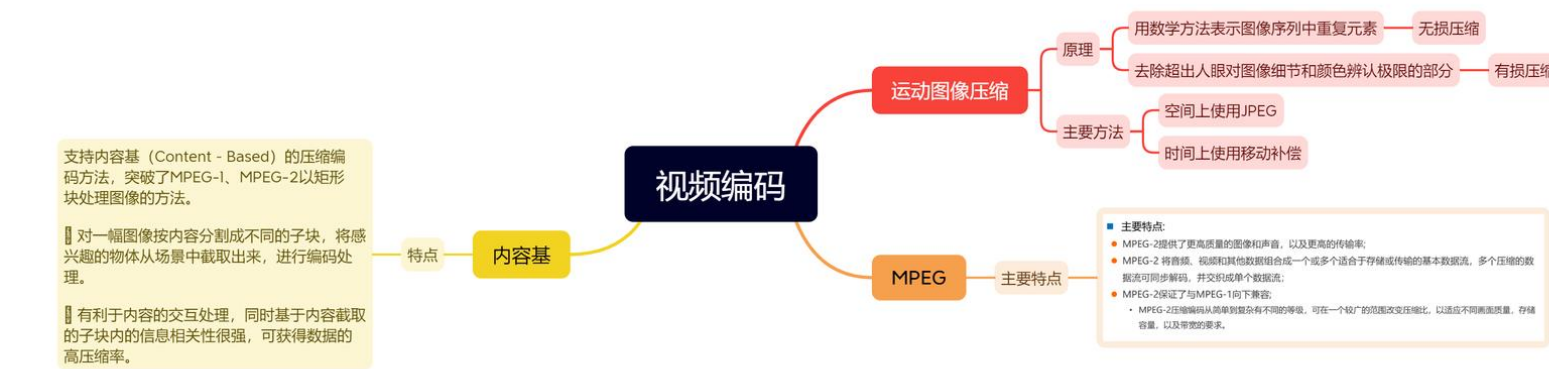
视频

视频基础



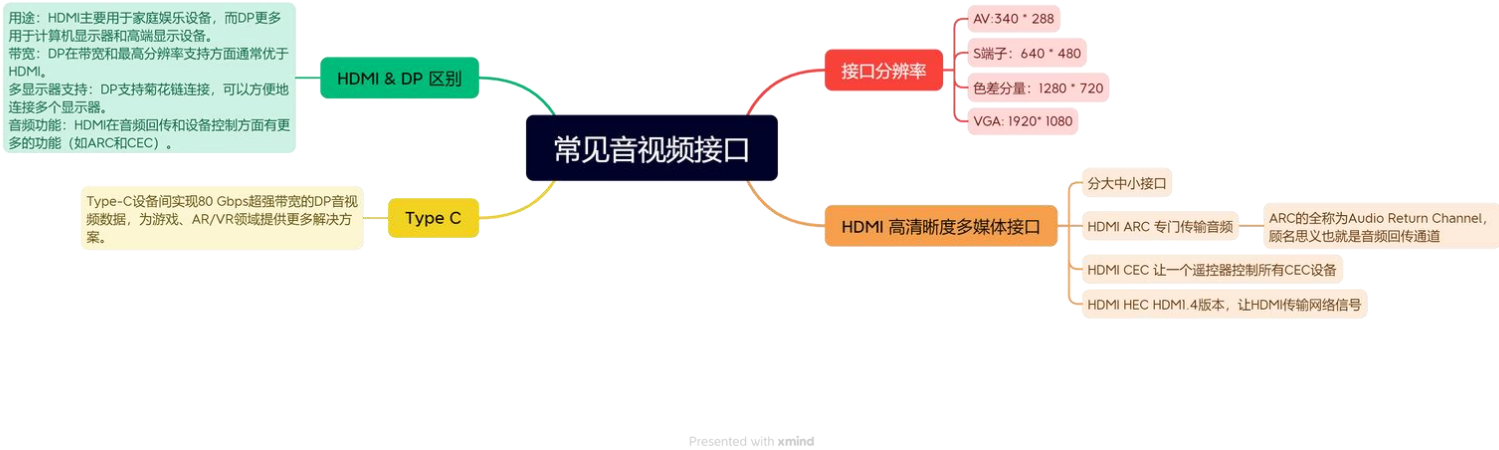
Presented with xmind

视频编码

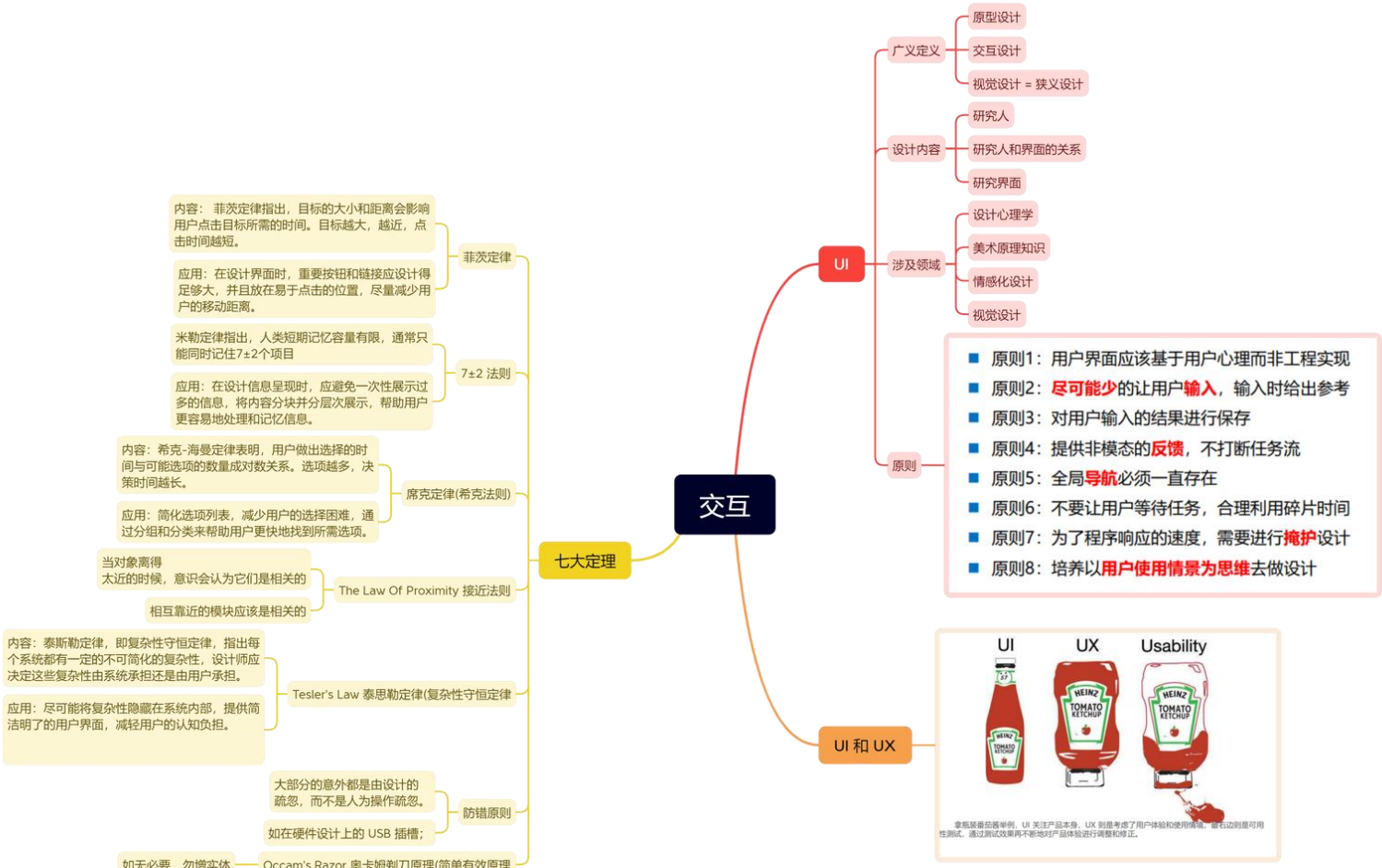


Presented with xmind

常见音视频接口



UI



构想一个带多模态的智能家居场景

以早晨起床为例子。

智能窗帘根据光线自动调节, 缓缓打开模拟自然唤醒。当用户的脑电波传感器检测到睡眠逐渐变浅时, 系统

开始准备唤醒程序。智能镜子显示天气预报和日程安排，同时播放舒缓的音乐.智能床垫通过监测用户的脑电波和睡眠质量，确定最佳的唤醒时间，并在轻柔振动中唤醒用户。用户可以通过语音指令唤醒语音助手，询问当天的天气、交通情况以及早餐建议。智能冰箱显示触摸屏幕，让用户查看存货和推荐菜谱，语音指令控制智能咖啡机，准备好用户喜爱的咖啡。
