

Section 2: 多媒体压缩标准

• 概述:

- 空间冗余: 像内冗余, 图像(小图像块之间)的统计相关性
- 时间冗余: 帧间冗余, 视频序列中连续帧像素间的统计相关性
- 心理冗余: ①视觉对高频、对角线噪声不敏感 ②颜色: 视觉对图像亮度比颜色敏感
- 视频压缩标准 • 运动估计及补偿 • MPEG标准: MPEG-1/2 • 视频标准 • 数字音频 • 媒体存储 • 媒体显示

• 视频压缩标准

Recap: 视频或动态图像: 一系列带有音频的图像/帧;

视频压缩主要方法: 运动估计: 对宏块运动矢量的预测; ②运动补偿: 对宏块的预测值及实际值之差进行编码

• 宏块:

- 含4个亮度块(Y), 2个色度块(Cb, Cr), DCT预测及补偿的基本单元
- 1个块大小为8x8, 故宏块包含 16x16 Y分量及 8x8 Cb & Cr 分量

• 像内&帧间压缩:

- 像内压缩: ①将图像帧视为静止图像 ②减少空间冗余 ③采用基于JPEG的压缩
- 帧间压缩: ①使用时间的相关性 ②减少时间冗余 ③运动估计及补偿

• 视频分辨率:

- 标清SD: ① NTSC (480i) 720x480 或 640x480 ② PAL (576i) 720x576
- 高清HD: ① 720P: 1280x720 渐近 ② 1080i: 1920x1080 交错式 ③ 1080P: 1920x1080 渐近式 (全高清 FHD) [渐近 & 交错 扫描方法]
- 超清UHD: ① 4K: 4倍FHD ② 8K: 4倍4K

• 扫描方法:

- 渐近扫描方法: 按顺序扫描所有行
- 交错扫描方法: 分别按行和列进行扫描
- 色块下采样: 人对亮度分量比色度分量更敏感, 故对色度进行压缩
 - 4:4:4 未进行下采样
 - 4:2:2 下采样: 对每2个水平的Y及1个Cr, 1个Cb 采样
 - 4:2:0 下采样: 对每个2x2的Y及 1个Cr, 1个Cb 采样

• 运动估计及补偿

- 运动估计方法: ① 全查找 ② 三步查找 ③ 2D查找 ④ 基于查找 ⑤ 十字查找

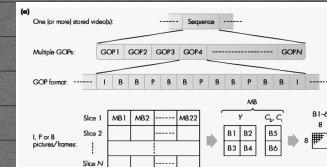
	方法	优点	缺点
全查找	在查找窗口中查找每个点	准确且能保持最佳压缩比	速度较慢且计算量大
三步查找	① 搜索9个点, 找到最小值点A ② 以A为中心, 查找距离减半, 搜索9个点, 找到最小值点B ③ 以B为中心, 查找距离减半, 搜索9个点, 找到最小值点C	速度快	不准确且压缩率低
2D查找	① 搜索5个点, 找到最小值点A, 以A为新一帧查找的中心 ② 若A位于之前帧的中心点或边界, 查找距离减半 若距离不是2x2搜索5个点; 否则搜索9个点	精度处于全查找和三步查找之间	准确度处于全查找和三步查找之间

• MPEG标准:

Recap: MPEG (Moving Picture Expert Group) 成立于1989年, 定义了隐式定义解码器的压缩数据流; 压缩算法及编码器由制造方确定

• MPEG

- 目的为压缩及解压视频信号
- MPEG 标准基于DCT编码及运动估计与补偿机制: ①DCT编码: 减少帧内冗余 ②运动估计与补偿: 减少帧间冗余
- 颜色视频源有3个分量: ① Y ② Cb ③ Cr ⇒ 常使用4:2:0格式采样
- GOP (Group of Picture) 视频序列被分为图像组, 每个GOP含 I, P, B 帧
 - I帧 (帧内编码帧): 编码时不考虑其它帧 ② Y, Cr, Cb 使用JPEG独立编码
 - P帧 (预测编码帧): 帧间编码, 使用之前的I帧或P帧进行运动估计及补偿后编码
 - B帧 (双向编码帧): 双向编码, 使用之前、之后或前后都有的帧进行编码
- 参数 M & N:
 - GOP大小为N: 两个I帧之间最短距离, 值多于12~15之间
 - 参数M: 两个最近I帧之间距离, 值多为1~3之间



• 参数M & N: 用户可选择的参数，在编码过程中使用；可用M,N权衡 疣码性能及错误比例

1. I,P,B帧编码

① I帧编码 frame \rightarrow blocks \rightarrow DCT \rightarrow R \rightarrow EE \rightarrow formatter \rightarrow bitstream

② P帧编码 参考I帧或P帧，差分值 $M_0 \rightarrow$ DCT \rightarrow R \rightarrow 变长编码 + 差分编码 (motion vector) \rightarrow EE \rightarrow bitstream

P帧的各blocks都有运动补偿，各block与I帧或前序P帧做比较；

P帧各block在速率窗内寻找参考，找到最匹配的块后确定 error 及 vector

vector 用来表示目标块与 block 之间的距离；error = target & reference 的差值；vector 用差分编码；error 以 (Y, Cr, Cb) 每片编码

③ B帧编码 以 P帧/I帧双参考，得 $M_0 \& M_1$, $M_1 - \frac{1}{2}(M_0 + M_1) \rightarrow$ DCT \rightarrow R \rightarrow RLE + vector (DE) \rightarrow EE \rightarrow bitstream

B帧编码由前后两帧共同决定；前后帧都会产生 vector 及 difference error (matrix)

使用目前 block 及处理中前后帧中最匹配的块值得到 difference error \rightarrow 构成 matrix

有最大差分 error 的 block 被选择及编码 error = target & reference 的差值；vector 用差分编码；error 以 (Y, Cr, Cb) 每片编码

编码中存在的问题：运算速度、复杂度、压缩率、应用度、耗电量、硬件

典型压缩比 I帧 10:1 ~ 20:1 P帧 20:1 ~ 30:1 B帧 30:1 ~ 50:1

2. MPEG-1

① 使用源输入格式 (source input format)，也称为 CCIR 601 颜色 TV 格式；MPEG-1 只支持非隔行视频

② MPEG-1 常见分辨率 25 帧/s 352x288 \Rightarrow PAL video ; 30 帧/s 352x240 NTSC video 都使用 4:2:0 采样

③ MPEG-1 释为 ISO/IEC 11172 标准 川江系统 / 视频 / 音频 / - 数据 / 软件

3. MPEG-2

MPEG-2 与 MPEG-1 相比：①比特率更高 > 4Mbps ② 引入 2:2 及 4:4:4 采样方式 ③ 支持隔行视频
④ 支持可拓展编码 ⑤ 有更高压缩率 ⑥ 有更多视频格式 ⑦ 配置文件的级别不同

4. 可拓展编码

可拓展编码可以使得 MPEG-2 和处在不同的比特率的网络，比特率可变，有噪声的网络传输

可拓展编码由 base layer & enhancement layer 组成：

{ base layer 独立编码，传输，解码 \rightarrow 获得基本视频质量
enhancement layer 由 base layer 及之前的 enhancement layer 决定编码，与 base layer 共同解码时会提高视频质量

可拓展编码：SNR 可拓展；enhancement layer 提供更高 SNR 质量

空间可拓展性：enhancement layer 提供更高清晰度

时间可拓展性：enhancement layer 提供更高帧率

混合可拓展性：将以上任意两者结合起来综合命令

数据分割 量化 DCT 系数被分为几部分，较低的 AC 系数用于 base layer 编码，较高的 AC 系数用于增强层编码

5 视频编码标准 MPEG 标准：MPEG-1, MPEG-2, MPEG4 AVC/H.264, H.265

JPEG:一般压缩比 10:1, MPEG-2:压缩比 30:1, H.264:一般压缩比为 50:1

6 数字化音频：常用 20 ~ 20000 Hz \Rightarrow 人类可以听到的声音

{ 模拟音频：自然界音频为模拟的，数据保真度及耐久性更低

数字音频：范围更大，抗噪能力更高，更精确的复制，耐久性更高，易于编辑，传输，分类及存储，能纠错

AD 转化 Analog \rightarrow 采样 \rightarrow 量化 \rightarrow 编码 \rightarrow digital

音频特性 speech: 50 ~ 10KHz ; Music: 15 ~ 20KHz, 采样频率为 f_{max} ，各样本对应比对数应在噪声容限内

样本比特 speech: 12 bits / sample ; music: 16 bits / sample

7. PCM 电话语音调制

公共电话交换网 PSTN 速率是 200 ~ 34 kHz，对应 $f_{max} = 6.8 \text{ kHz}$ 实际采样率 8 kHz，各样本用 8 bits 量化 \rightarrow 64 kbps 数据率 \rightarrow 数据串进线

8. 音频文件格式 ①无压缩：WAV, AIFF

②有损压缩：FLAC, ALAC(.m4a), MPEG-4...

③有损压缩：AAC, MP3 等；MP3 利用 MPEG-1/2 音频的第 3 层，压缩率为 4:1 ~ 20:1

9. 媒体存储 ①传统介质：硬盘，磁带；易于读写及漫改 ②光学存储 促进了多媒体技术：CD, VCD, DVD, 蓝光；易于存储，价格低且方便

光学存储利用螺旋状轨道反射激光的强度来读写信息

CD-ROM 批量存储音频数据及通用存储介质；有 2 种模式 ① 数据率 150 kbps，容量 650 Mbytes，存储计算机数据 ② 数据率 176 kbps，容量 747 Mbytes

VCD/Video CD 可利用光碟存储音频，利用 120mm 硬盘，利用 MPEG-1 压缩，比特率为单级 1150 kbps

DVD: digital versatile Disc; 数据存储容量更高，利用刻凹坑，提高存储密度；存储数据加密；利用 MPEG-2 压缩，数据率 98 Mbps

蓝光：利用高精度光盘，MPEG-2 / H.264 / VC-1 压缩及物理抹除，单面 25GB 容量，双面 50GB 容量；存储数据高清晰度及高帧

闪存：通过闪存存储 ROM；用于存储卡或驱动；又轻又方便又安静(?)

	CD	DVD	Blu-Ray Disc
Typical storage capacity	800 MB	4.7GB SL 27GB DL	27GB SL
Disc type (single-sided)	No	Single Sided	Single Sided
Track pitch	1.6 um	0.74 um	0.52 um
Minimum pit length (nm)	0.5 um	0.4 um	0.35 um
Wavelength of laser	780nm Red	650nm Red	405nm Blue
Data transfer rate/ speed	153.6 kbps	1350 kbps	4.5 Mbps
Thickness of Disc	1.2mm	1.2mm	1.1mm
Thickness of surface layer	0.6mm	0.6mm	0.1mm
Codecs	MPEG1	MPEG2	MPEG2/H.264

