

第1章介绍

1. 为什么要进行视频压缩？

- 未经压缩的数字视频的数据量巨大
- 存储困难
 - 一张DVD只能存储几秒钟的未压缩数字视频。
- 传输困难
 - 1兆的带宽传输一秒的数字电视视频需要大约4分钟。

一分钟的视频数据量

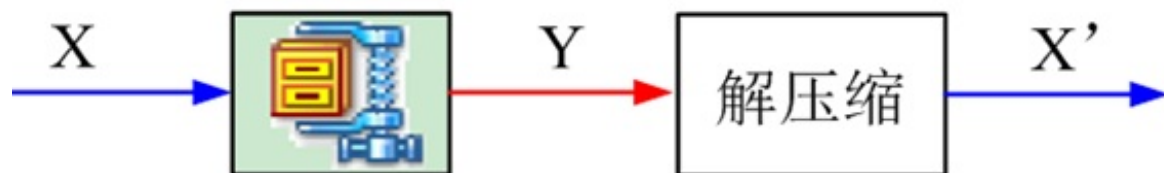
Frame Rate	Frame Size	Bits / pixel	Bit-rate (bps)	File Size (Bytes)
30 frames/sec	176 x 144 pixels	12	9,123,840	68,428,800

2. 为什么可以压缩

- 去除冗余信息
 - 空间冗余：图像相邻像素之间有较强的相关性
 - 时间冗余：视频序列的相邻图像之间内容相似
 - 编码冗余：不同像素值出现的概率不同
 - 视觉冗余：人的视觉系统对某些细节不敏感
 - 知识冗余：规律性的结构可由先验知识和背景知识得到

3. 数据压缩分类

- 无损压缩 (Lossless)
 - 压缩前解压缩后图像完全一致 $X=X'$
 - 压缩比低(2:1~3:1)
 - 例如：Winzip, JPEG-LS
- 有损压缩 (Lossy)
 - 压缩前解压缩后图像不一致 $X \neq X'$
 - 压缩比高(10:1~20:1)
 - 利用人的视觉系统的特性
 - 例如：MPEG-2, H.264/AVC, AVS

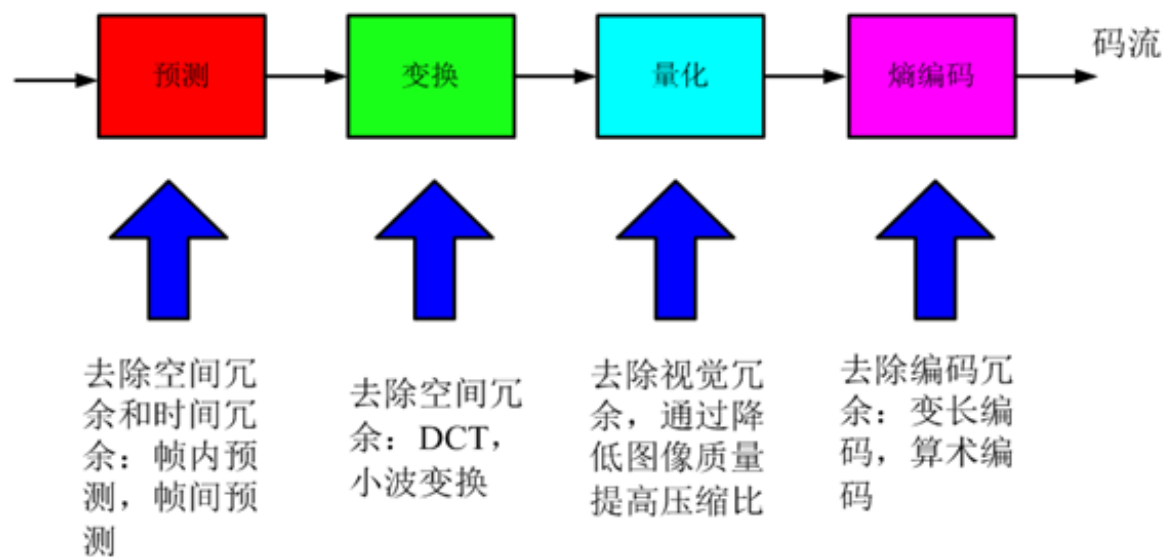


4. 编解码器

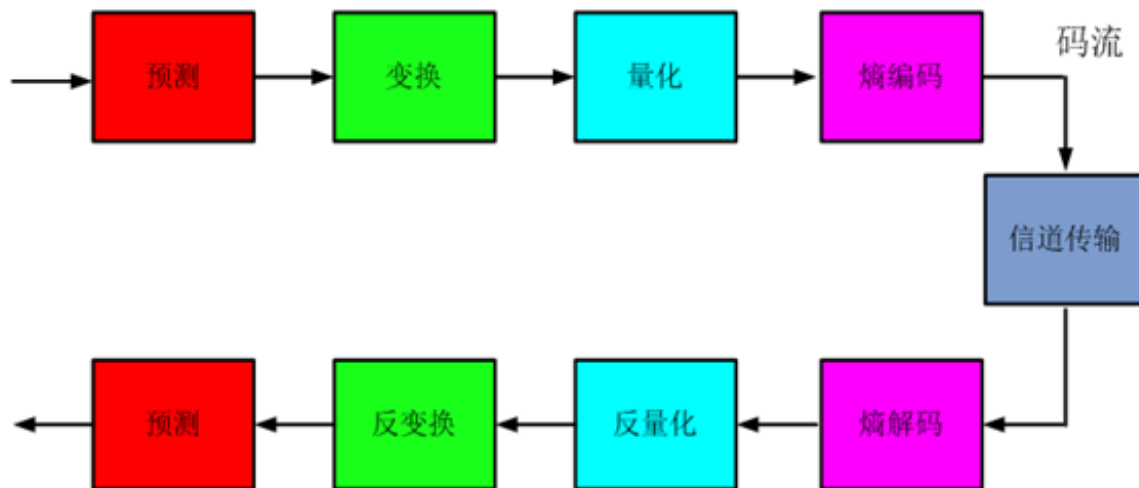
- 编码器 (Encoder)
 - 压缩信号的设备或程序
- 解码器 (Decoder)
 - 解压缩信号的设备或程序
- 编解码器 (Codec)
 - 编解码器对

5. 压缩系统的组成

(1) 编码器中的关键技术



(2) 编解码中的关键技术



6. 编解码器实现

- 编解码器的实现平台：
 - 超大规模集成电路VLSI
 - ASIC , FPGA
 - 数字信号处理器DSP
 - 软件
- 编解码器产品：
 - 机顶盒
 - 数字电视
 - 摄像机
 - 监控器

7. 视频编码标准

编码标准作用：

- 兼容：
 - 不同厂家生产的编码器压缩的码流能够被不同厂家的解码器解码
- 高效：
 - 标准编解码器可以进行批量生产，节约成本。

主流的视频编码标准：

- MPEG-2
- MPEG-4 Simple Profile
- H.264/AVC
- AVS
- VC-1

标准化组织：

- ITU : International Telecommunications Union
 - VECG : Video Coding Experts Group
- ISO : International Standards Organization
 - MPEG : Motion Picture Experts Group

8. 视频传输

- 视频传输：通过传输系统将压缩的视频码流从编码端传输到解码端
- 传输系统：互联网，地面无线广播，卫星

9. 视频传输面临的问题

- 传输系统不可靠
 - 带宽限制
 - 信号衰减
 - 噪声干扰
 - 传输延迟
- 视频传输出现的问题
 - 不能解码出正确的视频
 - 视频播放延迟

10. 视频传输差错控制

- 差错控制（Error Control）解决视频传输过程中由于数据丢失或延迟导致的问题
- 差错控制技术：
 - 信道编码差错控制技术
 - 编码器差错恢复
 - 解码器差错隐藏

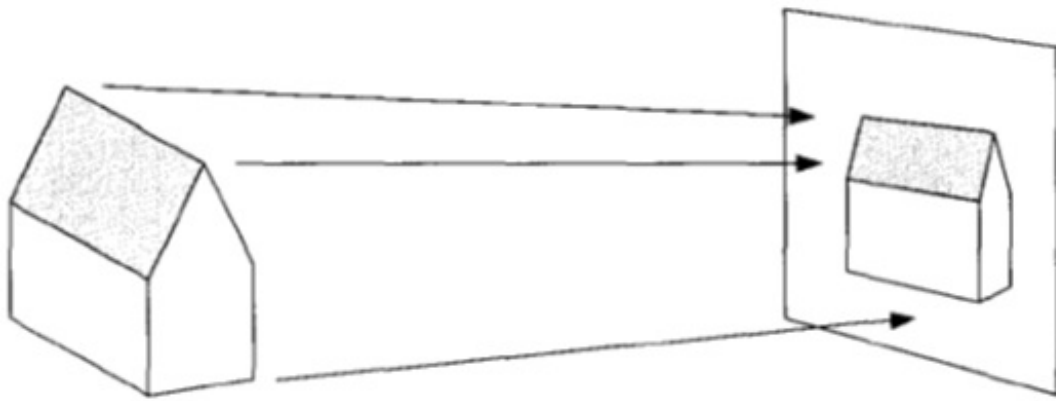
11. 视频传输的QoS参数

- 数据包的端到端的延迟
- 带宽：比特/秒
- 数据包的流失率
- 数据包的延迟时间的波动

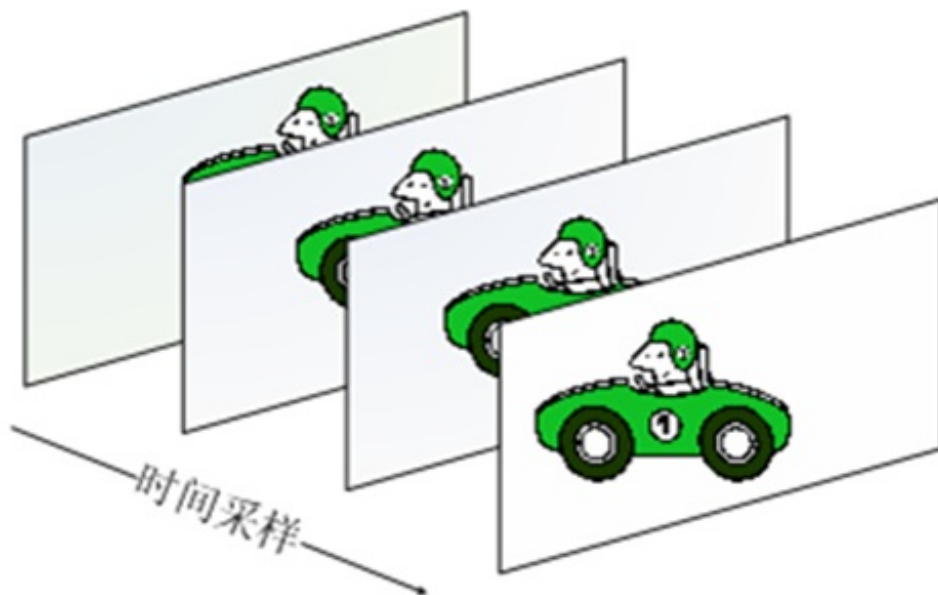
第2章 数字视频

1. 图像与视频

- 图像：是对视觉感知的物质再现。
- 三维自然场景的对象包括：深度，纹理和亮度信息
- 二维图像：纹理和亮度信息

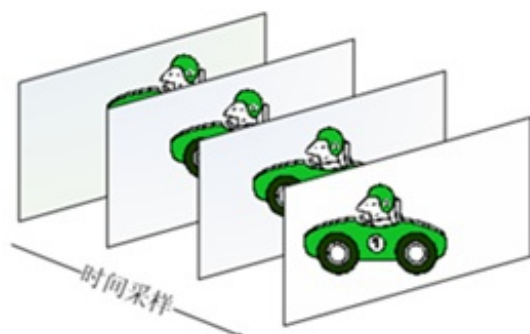
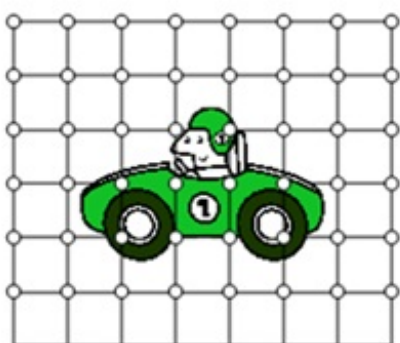


- 视频：连续的图像。
- 视频由多幅图像构成，包含对象的运动信息，又称为运动图像。



2. 数字视频

- 数字视频：自然场景空间和时间的数字采样表示。
 - 空间采样
 - 解析度 (Resolution)
 - 时间采样
 - 帧率：帧/秒



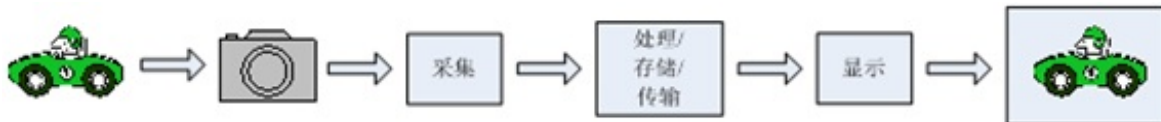
3. 空间采样

- 二维数字视频图像空间采样



4. 数字视频系统

- 采集
 - 照相机，摄像机
- 处理
 - 编解码器，传输设备
- 显示
 - 显示器

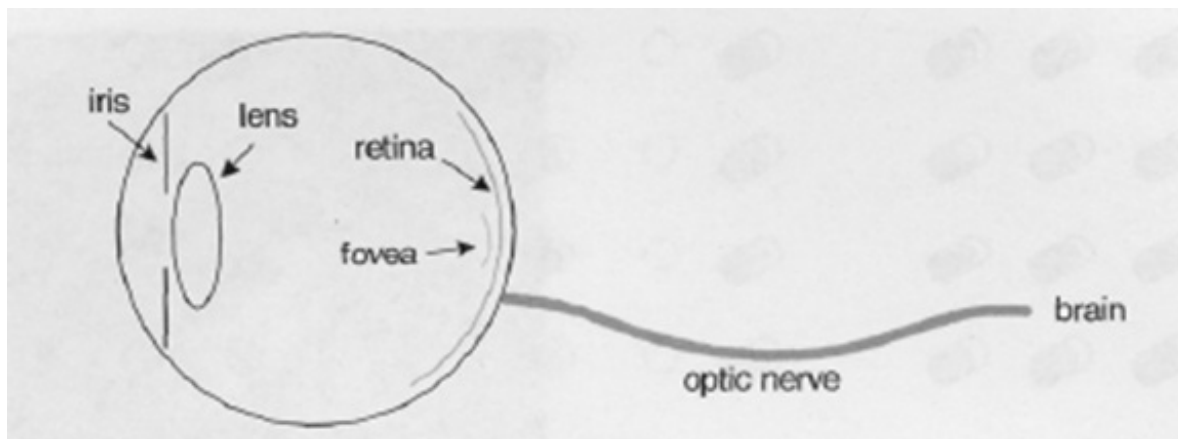


5. 人类视觉系统HVS

- HVS
 - 眼睛
 - 神经
 - 大脑

- HVS特点：

- 对高频信息不敏感
- 对高对比度更敏感
- 对亮度信息比色度信息更敏感



- 对运动的信息更敏感

6. 数字视频系统的设计应该考虑HVS的特点：

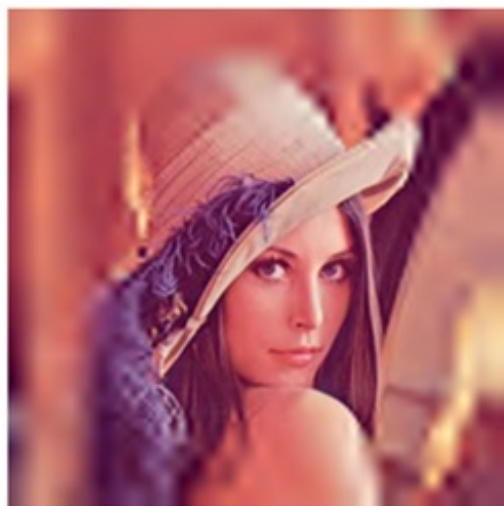
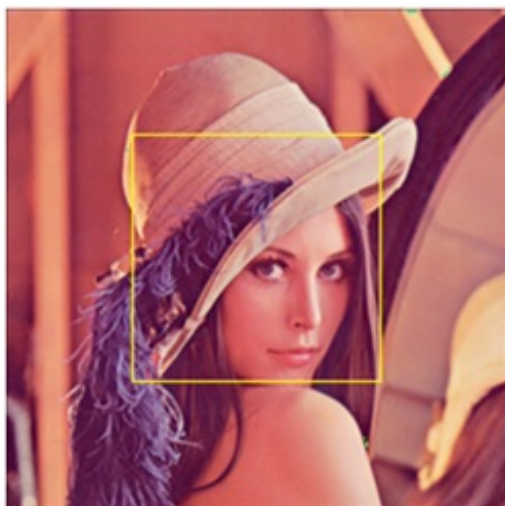
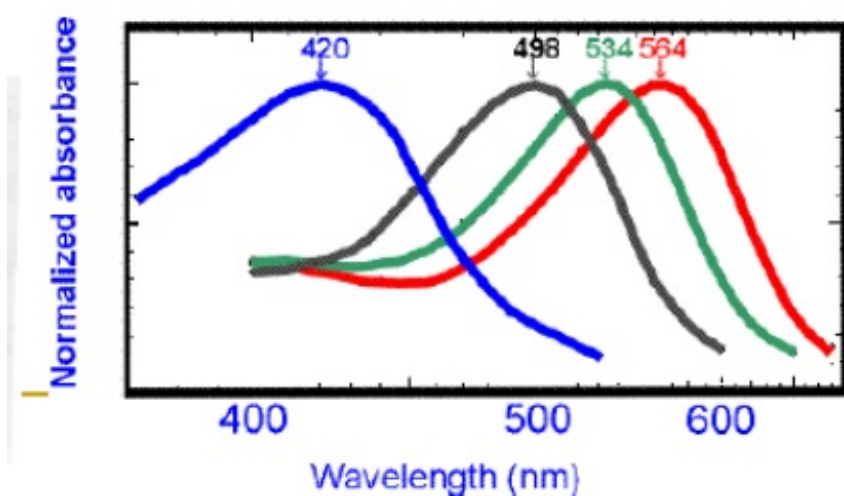
- 丢弃高频信息，只编码低频信息
- 提高边缘信息的主观质量
- 降低色度的解析度
- 对感兴趣区域（Region of Interesting, ROI）进行特殊处理

7. RGB色彩空间

- 三原色：红（R），绿（G），蓝（B）。
- 任何颜色都可以通过按一定比例混合三原色产生。

- RGB色度空间

- 由RGB三原色组成
- 广泛用于



BMP，TIFF，PPM等

- 每个色度成分通常用8bit表示[0,255]

8. YUV色彩空间

- YUV色彩空间：
 - Y：亮度分量
 - UV：两个色度分量
 - YUV更好的反映HVS特点

9. RGB转化到YUV空间

亮度分量Y与三原色有如下关系：

经过大量实验后ITU-R给出了，，，

主流的编解码标准的压缩对象都是YUV图像



10. YUV图像分量采样

- YUV图像可以根据HVS的特点，对色度分量下采样，可以降低视频数据量。
- 根据亮度和色度分量的采样比率，YUV图像通常有以下几种格式：

$$Y = k_r R + k_g G + k_b B$$

$$k_r = 0.299$$

11. 通用的YUV图像格式

- 根据YUV图像的亮度分辨率定义图像格式

$$k_g = 0.587$$

12. 帧和场图像

- 一帧图像包括两场——顶场，底场

$$k_b = 0.114$$

13. 逐行与隔行图像

- 逐行图像：一帧图像的两场在同一时间得到， $t_{top} = t_{bot}$ 。

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1687 & -0.3313 & 0.5000 \\ 0.5000 & -0.4187 & -0.0813 \end{bmatrix} \begin{bmatrix} R \\ G+128 \\ B+128 \end{bmatrix}$$

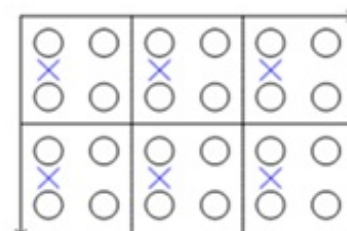
- 隔行图像：
一帧图像的两场在不同时间得到， $t_{top} \neq t_{bot}$ 。



4:4:4



4:2:2



4:2:0

14. 视频质量评价

- 有损视频压缩使编解码图像不同，需要一种手段来评价解码图像的质量。
- 质量评价：
 - 客观质量评价
 - 主观质量评价
 - 基于视觉的视频质量客观评价
- 客观质量评价：通过数学方法测量图像质量评价的方式。

• 优点：

- 可量化
- 测量结果可重复
- 测量简单

• 缺点：

- 不完全符合人的主观感知

格式	亮度分辨率
SQCIF	128x96
QCIF	176x144
CIF	352x288
4CIF	704x576
SD	720x576
HD	1280x720， 1920x1080

15. 客观评价的方法

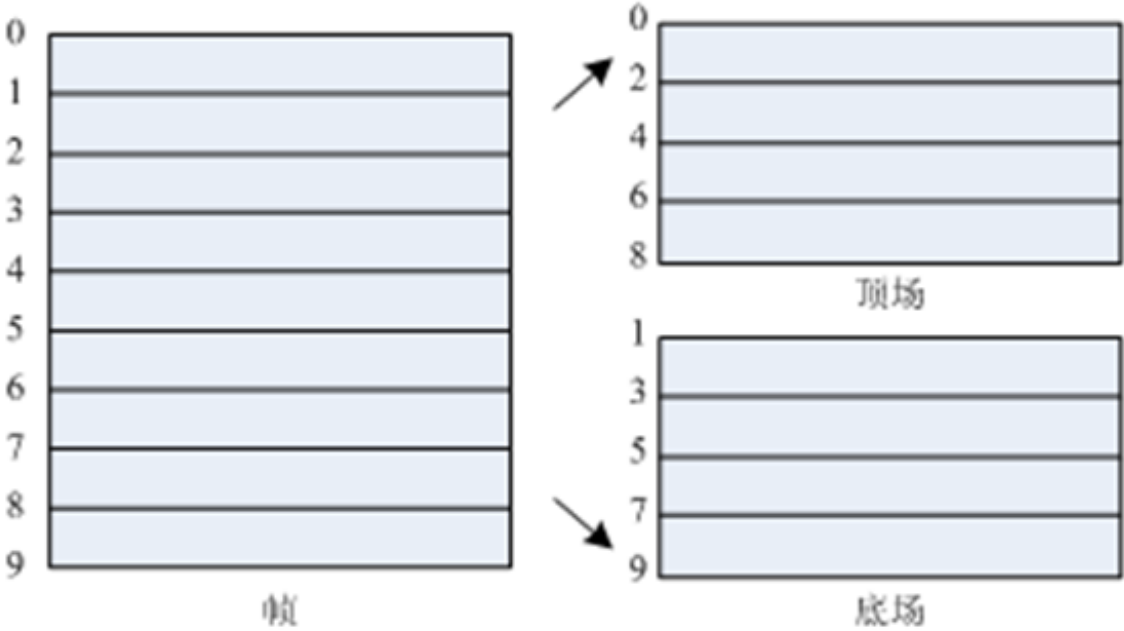
常用的客观评价方法：

16. 主观评价方法

- 主观质量评价：用人的主观感知直接测量的方式。
- 优点：
 - 符合人的主观感知

• 缺点：

- 不容易量化
- 受不确定因素影响，测量结果一般不可重复
- 测量代价高



常用主观评价方法

17. 基于视觉的视频质量客观评价方法

- 基于视觉的视频质量客观评价：将人的视觉特性用数学方法描述并用于视频质量评价的方式。
- 结合了主观质量评价和客观质量评价两方面优点。
- 常用方法：结构相似度（Structural SIMilarity，SSIM）方法。
- 将HVS的特征用数学模型表达出来。
- 未来重要的研究方向

1. 通信系统的组成

- 信源：产生消息
- 信道：传输消息
- 信宿：接收消息



原始图像

2. 基本概念

- 通信中对信息的表达分为三个层次：信号，消息，信息。



顶场



底场

- 信号：是信息的物理层表达，可测量，可描述，可显示。如电信号，光信号。



- 消息：是信息的载体，以文字，语言，图像等人类可以认知的形式表示。



- 信息：不确定的内容。

- 均方差 (Mean Square Error, MSE)

$$MSE = \frac{1}{NM} \sum_n \sum_m [P_o(n, m) - P_d(n, m)]^2$$

- 平均绝对值差 (Mean Absolute Difference, MAD)

$$MAD = \frac{1}{NM} \sum_n \sum_m |P_o(n, m) - P_d(n, m)|$$

- 信噪比 (Signal Noise Rate, SNR)

$$SNR = 10 \log_{10} \left[\frac{\frac{1}{NM} \sum_n \sum_m P_d(n, m)^2}{MSE} \right]$$

- 峰值信噪比 (Peak Signal Noise Rate, PSNR)

$$PSNR = 10 \log_{10} \left[\frac{(2^n - 1)^2}{MSE} \right]$$

3. 信息熵

信息的特点

信息的测量



自信息量

条件信息量

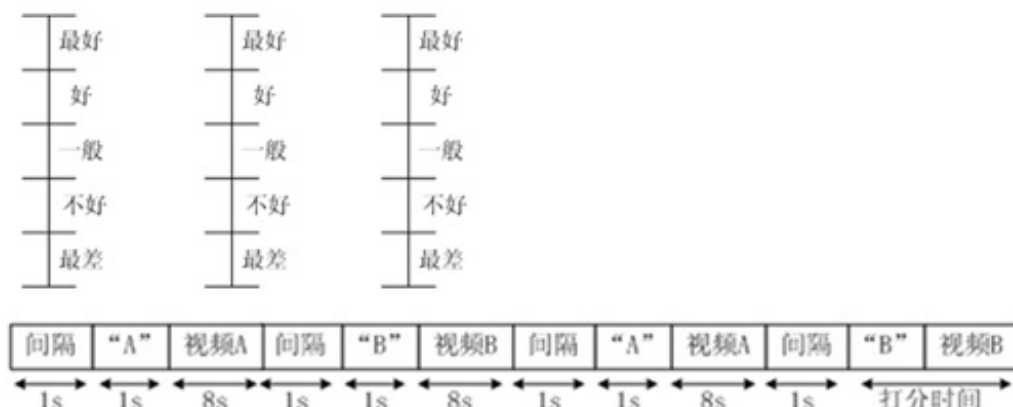
4. 信息熵

5. 条件熵和联合熵

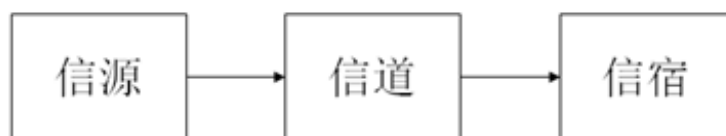
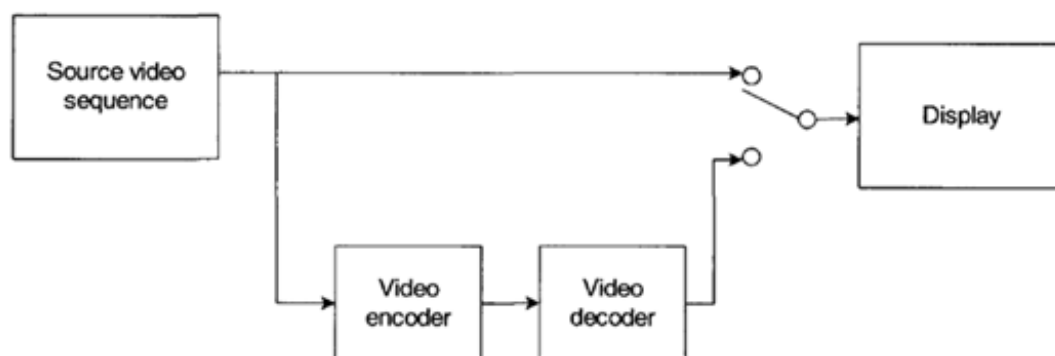
6. 熵的性质

- 非负性：信源熵是非负值，即 $H(X) \geq 0$;
- 扩展性：信源熵 X 有 M 个符号，如果其中一个符号出现的概率趋于零，信源熵就等于剩余 $M-1$ 个符号的信源熵；
- 极值性（最大信息熵）：对于具有 M 个符号的信源，只有在所有符号等概率出现的情况下，信源熵达到最大值，即

● 双刺激连续质量评价（Dual Stimulate Continuous Quality Metric, DSCQM）



● DSCQM测试系统



$$H(x_0, x_1, \dots, x_{M-1}) \leq H(1/M, 1/M, \dots, 1/M) = -\sum_{i=0}^{M-1} 1/M \log 1/M = \log M$$

- 可加性： $H(XY) = H(X) + H(Y|X) = H(Y) + H(X|Y)$
- 熵不增：条件熵不大于信息熵 $H(X|Y) \leq H(X)$ ；
- 联合熵不大于各信息熵的和，即 $H(XY) \leq H(X) + H(Y)$ 。

7. 互信息量

8. 互信息

- 物理意义：
 $H(X)$ 是X所含的信息， $H(X|Y)$ 是已知Y的条件下X还能带来的信息量。那么两者之差就是由于知道Y使得X减少的信息量，也即由Y可以得到的关于X的信息量。

9. 各种熵的关系

11. 信源编码

- 信源编码：将消息符号转变成信道可传输的信息。
- 两个基本问题：
 - 用尽可能少的信道传输符号来传递信源消息，提高传输效率；
 - 减少由于信道传输符号的减少导致的失真。

12. 离散信源统计特性

13. 离散信源类型：简单无记忆信源和马尔可夫信源

14. 编码分类

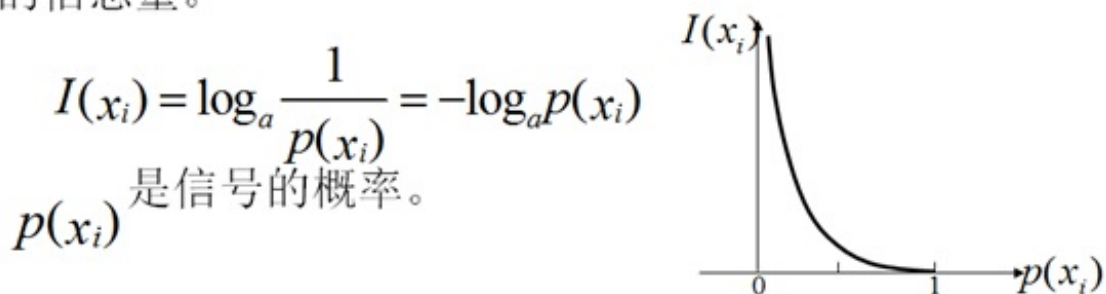
- 随概率P的递减性：概率越大，信息量越小

$$P(x) \uparrow, I[P(x)] \downarrow$$

- 可加性：两个独立消息的总信息量应是两个消息的信息量的和

$$I[P(x)P(y)] = I[P(x)] + I[P(y)]$$

- 信息量（自信息量）：随机事件发生后所带来的信息量。



- 一个符号出现的可能性越大，其信息量越少。

- 条件信息量：一个随机事件在另外一个随机事件已经发生的情况下再发生时所提供的信息量。

$$I(x_i | y_j) = -\log_a p(x_i | y_j)$$

- 当 x_i 与 y_j 相互独立时，有 $P(x_i | y_j) = P(x_i)$ ，此时， x_i 的条件信息量就等于自信息量，即 $I(x_i | y_j) = I(x_i)$ ；
- 一般情况下，有 $P(x_i | y_j) > P(x_i)$ ，由定义可知， $I(x_i | y_j) < I(x_i)$ ；
- 当 $x_i = y_j$ 时， $P(x_i | y_j) = 1$ ， $I(x_i | y_j) = 0$ 。
即事件 y_j 的发生使得事件 x_i 的不确定度等于零，不含有任何信息量。

- 等长码：在一组码字集合C中的所有码字 c_m ($m = 1, 2, \dots, M$)，其码长都相同，则称这组码C为等长码。
- 变长码：若码字集合C中的所有码字 c_m ($m = 1, 2, \dots, M$)，其码长不都相同，称码C为变长码。

15. 平均码长

16. 等长码与变长码比较

- 等长编码将信源输出符号序列的任意一种取值（概率可能不同）都编码成相同长度的输出码字，没有利用信源的统计特性；
- 变长编码可以根据信源输出符号序列各种取值的概率大小不同，将他们编码成不同长度的输出码字，利用了信源的统计特性。因此又称其为熵编码。

17. Huffman编码

- 熵：消息集合的不确定度。
- 信息熵：N个随机事件的平均信息量。

$$H(X) = E[-\log_a p(x_i)] = -\sum_{i=0}^{N-1} p(x_i) \log_a p(x_i)$$

$$\sum_{i=0}^{N-1} p(x_i) = 1$$

- 信息熵从平均意义上来表征信源总体信息的测度
- 例：天气预报，有两个信源

$$\begin{bmatrix} X_1 \\ p(x) \end{bmatrix} = \begin{bmatrix} \text{晴天} & \text{阴天} \\ 1/4 & 3/4 \end{bmatrix} \quad \begin{bmatrix} X_2 \\ p(x) \end{bmatrix} = \begin{bmatrix} \text{晴天} & \text{阴天} \\ 1/2 & 1/2 \end{bmatrix}$$

- 两个信源的信息熵分别为：

$$H(X_1) = \frac{1}{4} \log 4 + \frac{3}{4} \log \frac{4}{3} = 0.809$$

$$H(X_2) = \frac{1}{2} \log 2 + \frac{1}{2} \log 2 = 1$$

- 条件熵：条件信息量的数学期望。

$$H(X|Y) = E[I(x_i|y_j)] = \sum_{i=0}^{N-1} \sum_{j=1}^{M-1} p(x_i y_j) I(x_i|y_j)$$

$$= -\sum_{i=0}^{N-1} \sum_{j=1}^{M-1} p(x_i y_j) \log_2 p(x_i|y_j)$$

- 联合熵：多条消息的信息量的数学期望。

$$H(XY) = E[I(x_i y_j)] = -\sum_{i=0}^{N-1} \sum_{j=1}^{M-1} p(x_i y_j) \log_2 p(x_i y_j)$$

- Huffman编码：典型的变长编码。

- 步骤：

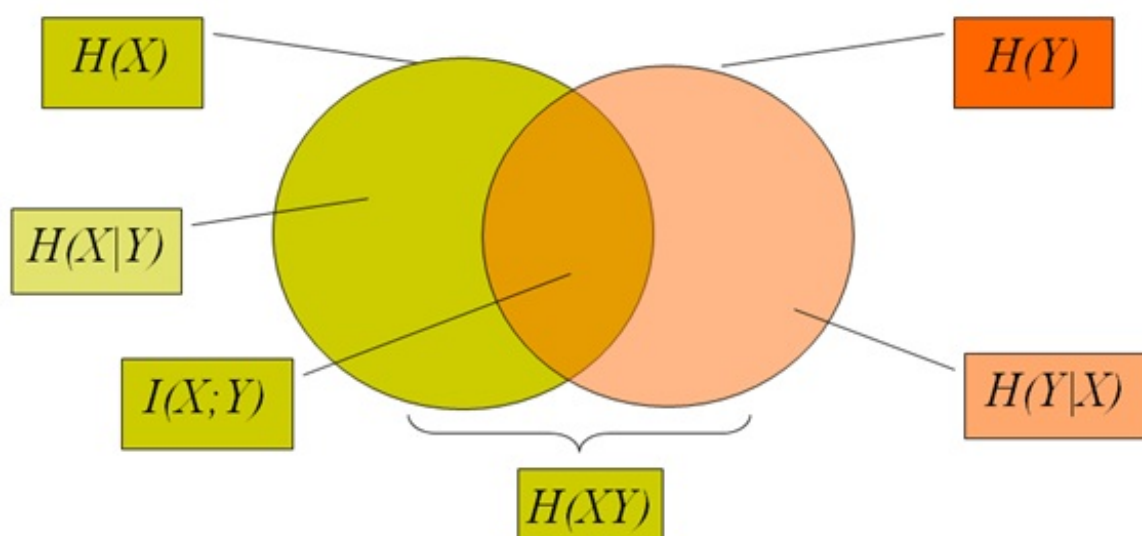
- 将信源符号按概率从大到小的顺序排列，假

- 事件 x_i 是否发生具有不确定性，用 $I(x_i)$ 度量。
- 接收到符号 y_j 后，事件 x_i 是否发生仍保留有一定的不确定性，用 $I(x_i|y_j)$ 度量。
- 观察事件前后，这两者之差就是通信过程中所获得的信息量，称为事件 x_i 和事件 y_j 之间的互信息量，用表示： $I(x_i; y_j)$

$$I(x_i; y_j) = I(x_i) - I(x_i|y_j) = \log \frac{p(x_i|y_j)}{p(x_i)}$$

- 互信息：

$$\begin{aligned} I(X; Y) &= \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} P(x_i y_j) \log \frac{P(x_i|y_j)}{P(x_i)} \\ &= \sum_{i=0}^{N-1} P(x_i) \log \frac{1}{P(x_i)} - \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} P(x_i y_j) \log \frac{1}{P(x_i|y_j)} \\ &= E \left[\log \frac{1}{P(x_i)} \right] - E \left[\log \frac{1}{P(x_i|y_j)} \right] \\ &= H(X) - H(X|Y) \end{aligned}$$



定 $p(x_1) \geq p(x_2) \dots \geq p(x_n)$

- 给两个概率最小的信源符

- 单消息离散信源：

- 只输出一个离散消息（符号）
- 统计特性

用符号可能取值范围 X 和符号取值 x_i 的概率 $P(x_i)$ 描述

$$\begin{bmatrix} X \\ p(x_i) \end{bmatrix} = \begin{bmatrix} x_1, & \cdots, & x_i, & \cdots, & x_n \\ p(x_1), & \cdots, & p(x_i), & \cdots, & p(x_n) \end{bmatrix}$$

其中, $0 \leq p(x_i) \leq 1, i = 1, 2, \cdots, n$, 且 $\sum_{i=1}^n p(x_i) = 1$

- 离散消息序列信源

- 输出一个离散消息序列
- 统计特性

L 个离散消息的序列可以表示成 L 维随机向量

$$X = (X_1, \cdots, X_i, \cdots, X_L)$$

其取值有 n^L 种可能, 构成取值集合 X^L , 则离散消息序列的取值集合及其概率为:

$$\begin{bmatrix} X^L \\ p(x) \end{bmatrix} = \begin{bmatrix} a_1, & \cdots, & a_i, & \cdots, & a_{n^L} \\ p(a_1), & \cdots, & p(a_i), & \cdots, & p(a_{n^L}) \end{bmatrix}$$

- 简单无记忆信源：离散信源输出符号彼此间相互独立, 而且所有符号服从同一种概率分布。

$$p(x) = p(x_1, \cdots, x_l, \cdots, x_L) = \prod_{l=1}^L p(x_l)$$

- 马尔可夫信源：输出符号间彼此相关, 且每个符号只与它前面的一个符号相关, 而这种相关性可以用符号间的转移概率来描述。

$$p(x_n | x_{n-1}) = p(x_n | x_{n-1} x_{n-2} \cdots)$$

号 $p(x_{n-1})$, $p(x_n)$ 各分配一个码位"0"和"1", 将这两个信源符号合并成一个新符号, 并用这两个最小的概率之和作为新符号的概率, 结果得到一个只包含 $(n-1)$ 个信源符号的新信源。称为信源的第一次缩减信源, 用 S_1 表示。

- 将缩减信

● 等长码平均码长:

$$\bar{n} = \sum_{m=0}^{M-1} n_m p(c_m) = \sum_{m=0}^{M-1} n p(c_m) = n$$

● 变长码平均码长:

$$\bar{n} = \sum_{m=0}^{M-1} n_m p(c_m)$$

源 S_1 的符号仍按概率从大到小的顺序排列，重复步骤2，得到只含 $(n-2)$ 个符号的缩减信源 S_2 。

- 重复上述步骤，直至缩减信源只剩下两个符号为止，此时所剩两个符号的概率之和必为1。然后从最后一级缩减信源开始，依编码路径向前返回，就得到各信源符号所对应的码字。

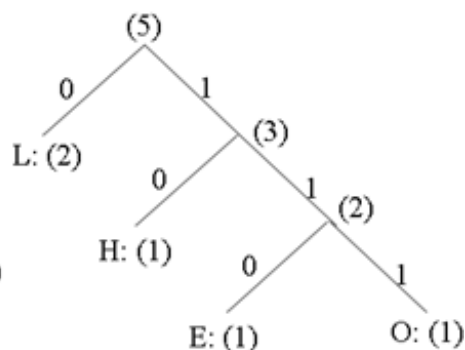
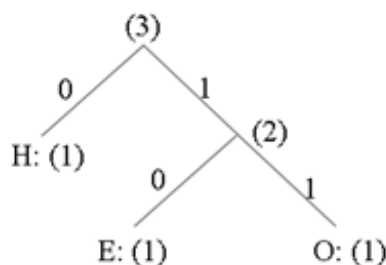
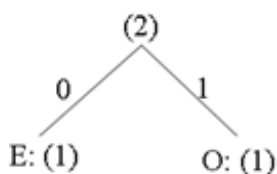
例子：对HELLO进行编码，符号频率统计：

18. 信道编码

符号	H	E	L	O
数目	1	1	2	1

编码过程：

- 排序：L(2), H(1), E(1), O(1)
- 自底向高构造二叉树



- 信道编码主要考虑如何增加信号的抗干扰能力，提高传输的可靠性，并且提高传输效率。
- 一般是采用冗余编码法，赋予信码自身一定的纠错和检错能力，使信道传输的差错概率降到允许的范围之内。

19. 信道类型

- 根据信道连续与否分类
 - 离散信道
 - 连续信道
 - 半连续信道
- 根据信道是否有干扰分类
 - 无干扰信道
 - 有干扰信道

- 根据信道的统计特性分类

- 无记忆信道
- 有记忆信道
- 恒参信道
- 变参信道
- 对称信道
- 非对称信道

20. 信道容量

- 在信息论中，称信道无差错传输的最大信息速率为信道容量。
- 仙农信道容量公式：
 - 假设连续信道的加性高斯白噪声功率为 N ，信道带宽为 B ，信号功率为 S ，则该信道的容量为

$$C = B \log_2 \left(1 + \frac{S}{N} \right)$$

- 由于噪声功率 N 与信道带宽 B 有关，则噪声功率 $N = n_0 B$ 。因此，仙农公式还可以表示为

$$C = B \log_2 \left(1 + \frac{S}{n_0 B} \right)$$

21. 香农信道容量公式的意义

- 在给定 B 和 S/N 的情况下，信道的极限传输能力为 C ，而且此时能够做到无差错传输。如果信道的实际传输速率大于 C 值，则无差错传输在理论上就已不可能。因此，实际传输速率一般不能大于信道容量 C ，除非允许存在一定的差错率。
- 提高信噪比 S/N （通过减小 n_0 或增大 S ），可提高信道容量 C 。特别是，若 $n_0 \rightarrow 0$ ，则 $C \rightarrow \infty$ ，这意味着无干扰信道容量为无穷大；
- 增加信道带宽 B ，也可增加信道容量 C ，但做不到无限制地增加。这是因为，如果 S 、 n_0 一定，有

$$\lim_{S \rightarrow \infty} C = \frac{S}{n_0} \log_2 e \approx 1.44 \frac{S}{n_0}$$

- 维持同样大小的信道容量，可以通过调整信道的 B 及 S/N 来达到，即信道容量可以通过系统带宽与信噪比的互换而保持不变。

22. 失真

- 失真：信源的消息经过编解码后不能完全复原
- 在实际的信源和信道编码中，消息的传输并不总是无失真的。
 - 由于存储和传输资源的限制
 - 噪声等因素的干扰

23. 率失真理论

- 仙农定义了信息率失真函数 $R(D)$
 - D 是消息失真
 - R 是码率
- 率失真定理：在允许一定失真度 D 的情况下，信源输出的信息率可压缩到 $R(D)$ 。

24. 失真函数

- 失真函数：信源符号 $X=\{x_1, x_2, \dots, x_n\}$ ，经信道传输接收端符号 $Y=\{y_1, y_2, \dots, y_n\}$ ，对于每一对 (x_i, y_j) 指定一个非负函数 $d(x_i, y_j)$ ，称 $d(x_i, y_j)$ 为单个符号的失真度或失真函数。对于连续信源连续信道的情况，常用 $d(x, y)$ 表示。
- 常用失真函数：

$$d(x_i, y_j) = \begin{cases} 0 & i = j \\ a & a > 0, i \neq j \end{cases} \quad d(x_i, y_j) = (y_j - x_i)^2$$

- 平均失真度：

$$\overline{D} = E[d(x_i, y_j)] = \sum_{i=1}^n \sum_{j=1}^m p(x_i, y_j) d(x_i, y_j)$$