Week1:

Digitization: sampling & quantization

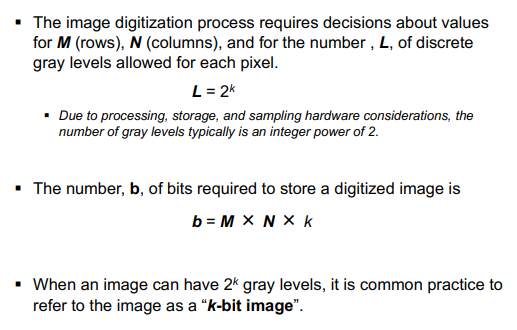
Visual perception: 1. Mach bands: human over/undershoot intensity boundary

2. simultaneous contrast

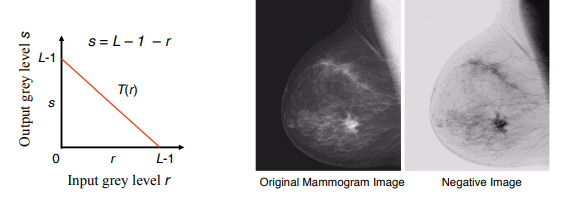
3. optical illusion: edge detection(not understood)

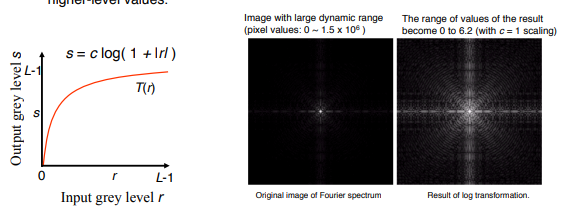
Application: anti aliasing, motion blur, image compression: psycho-visually redundant

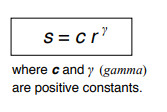
False conturing: not enough gray levels.

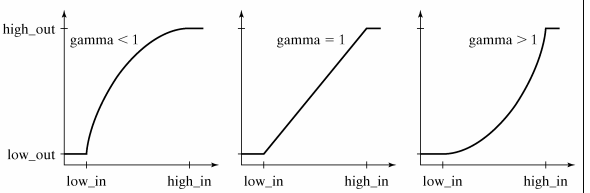


Linear operation: T(af+bg) = aT(f) + bT(g) , fg are image, ab are scalar

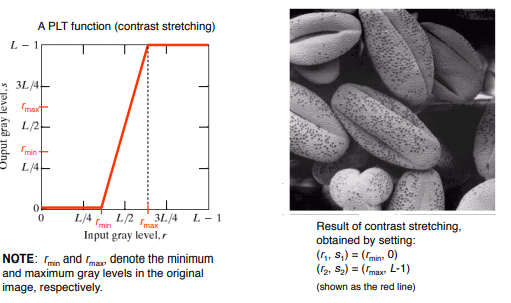
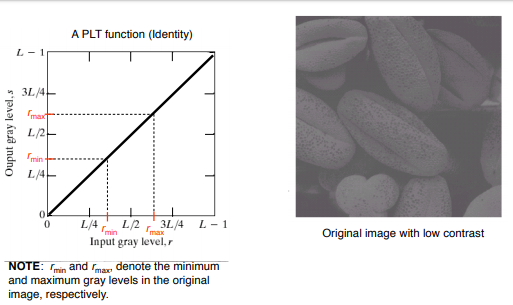
Image negative: 

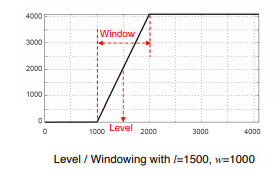
Log transform: expand dark area,while compress light area

Power-law transform:  contrast manipulation.

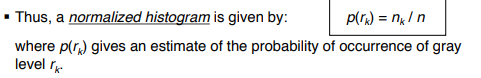


Piecewise linear transformation: contrast-streching, gray-level slicing

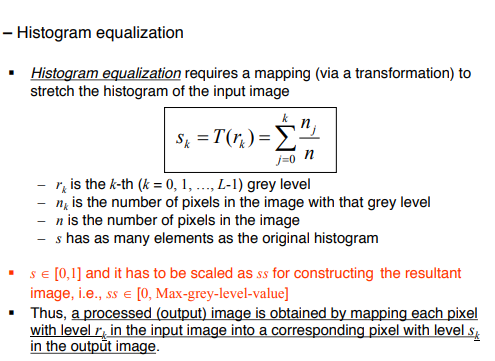




Histogram:

Normalized histogram: divide each value by total number of pixels. 

Histogram processing: The image that has been histogram equalized **always has pixels that reach the brightest grey level**

(各等级数量除以总数再乘上最大灰度等级得到ss)

Spatial filtering: 先取和，再记得除以filter的总数

Median filtering: 先排序范围内像素， 再把目标替换为中位数

----------------------------------------------

Week2 morphlogical image processing

Translation： 位移

Reflection： 180度旋转（所有点坐标\*-1）

Dilation： 先旋转180度再计算，erosion不用

Boundary extraction： 原图减去erosion后的图片得到边框

Week3： color models

HSV：立起来的rgb方块， 对角线是0，0，0-255，255，255 全是灰色

Brightness : O = I \* x

Gamma correction: O = I ^ (1/gamma)

Contrast: O = (I-a) \* b

Pan: 位移；

OVER： O=（A\*M）+【（1-M）\*B】 M是蒙版 把蒙住的A放在B上（覆盖）

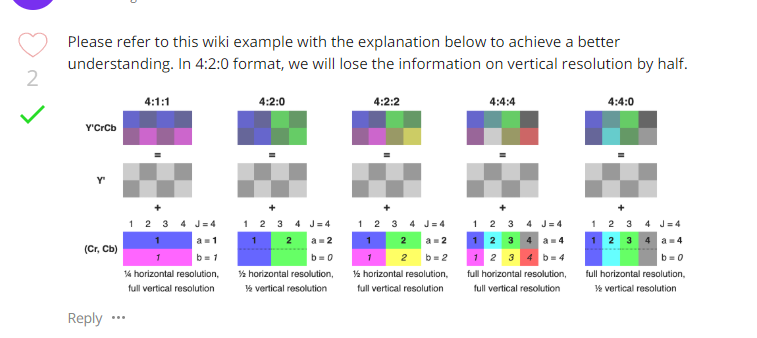
MIX： O = （MV\*A）+【（1-MV）\*B】 MV是系数（混合）

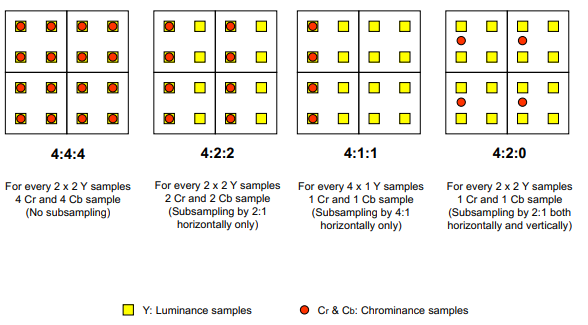
Week4 video data processing

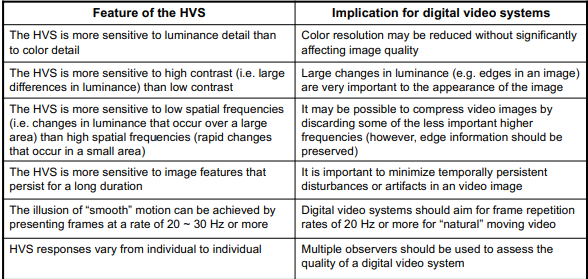
Spatial sampling: non-interlace（一行采样）；interlace（奇偶采样）

Temporal sampling： 帧率采样

Digital video standard: ccir-601: YCrCb , luminance and chrominance are together component,4:2:2 format

Sif: luminance & chrominance divided, 4:2:0

;;cif;;qcif

Human visual system: 

Optical flow: 观察与运动不一定一致

Week5 audio data representation & coding

Sound: periodic & non-periodic. Music & speech

Sampling theorm: 取样率是音频最高频率的二倍（20khz人类上限nyquist rate就取样40khz）

Sampling size： 一般16bit 最低8bit

音频文件大小： 取样率r 取样大小s 每秒音频就占据rs/8个比特

Pcm编码： 直接取整 用于manipulating ；；non-linear pcm 编码： 小声音占据更多资源（对数编码） 用于storing and transferring

DPCM编码： 储存前一个取样和本体的差值（2,6）得到（2，-4）

Predictor based 压缩： 记录预测值与实际值的差值

Perceptually based压缩： 利用人类听觉， 使用mask阈值仅保留高于阈值频率声音

Subband coding： 把音频分成高频低频分别压缩

Progressive 压缩： 粗压缩再解压然后与原版/上级差值做差， 重复三次以上， 然后根据需求选择差值等级

Week6 computer graphics

2d图像： path：线的集合；；path和shape可以stroke/filled；；transformation

Local coordinate system: 以物品自己为原点和轴

Spline： B-spline： 控制点不在线上， Bezier spline： 控制点在线上但用切线， NURBS：控制点不在线上且只用重量

Extrusion： 2d变3d延伸 ；； lathe： 车床；； sweep： 面变体扫；；skinning： 包裹

Mesh tessellation： 分割多边形增加模型面数

Subdivision surface： 不断的平滑边缘

渲染技巧： wireframe和hiddenline；；

Flat rendering： 只考虑中心法线；；smooth rendering： 考虑每个顶点；；phong rendering： 考虑高光；；ray tracing： 考虑入射光

材质bump map： 添加凹凸

材质的缺点： 1 消耗内存 2 需要高分辨率 3 难以正确覆盖

Procedural材质的优点： 不消耗内存， 3d形态， 容易变化

Week7 computer animation

Rotoscope： 描电影 用于手画动画；；Stop-motion： 一幅一幅拍照 ；；Cutout： 纸盘画

3d动画步骤： 生产前： 编故事 生产中： 建模动画渲染贴图 生产后： 组合 后效 发售

动画控制： explicitly declared ： 直接控制关键帧

Live action、analyzing： 动作描绘；；procedual： 用于fx；；constraint based： 控制限制区域；；kinematics： 控制关节（fk： 控制关节 ik： 给输入输出自己画关节）；；dynamic： 考虑物理定律；；facial animation： 面部感应

后效： color image processing和数码特效 可以增加亮度，减少对比度，移除错误，增加细节，组合图片

Week8 image data analysis and retrieval（图像检索）

Text based： 通过描述 需要人工标记

Content based： 通过内容 先保存多维feature vector， 用户再提供相似范例用来检索

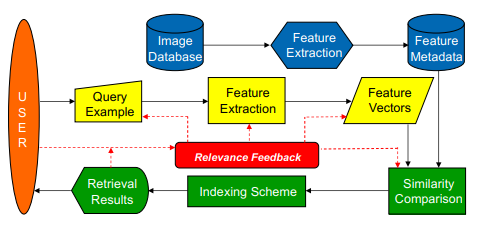
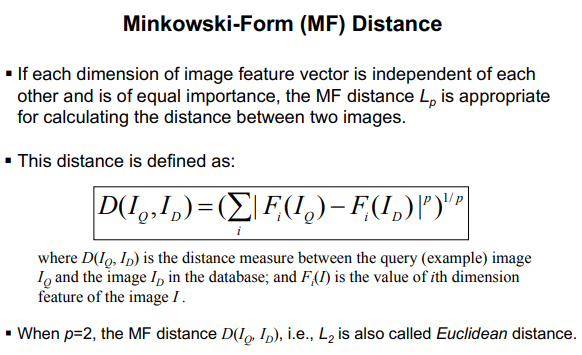
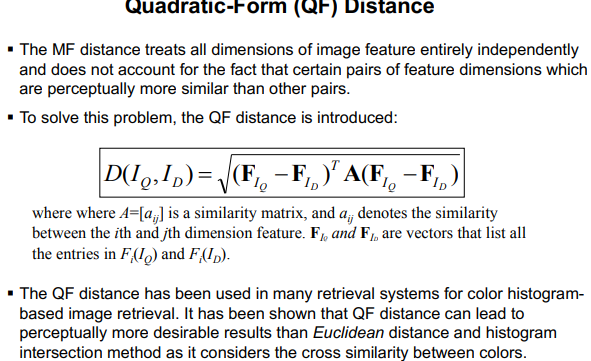


Image content包含visual 和semantic content， visual包括general和domain specific

Visual content descriptor： 分global（全图）和local（区域）

最常用的general content： 颜色 包括color moment色矩, color histogram, color coherence vector 和color correogram

相似度判断： 



Indexing scheme： 用pca/kl/神经网络减少维数 再用r树/inear quad树/kdb树/grid file建立目录

General content： 材质： structural辨别材质安放规则对规律有效 和statistical辨别图片亮度数据

General content： shape： 一般用于辨别区域和物体；；方法：boundary based，region based，geometry based，structure based

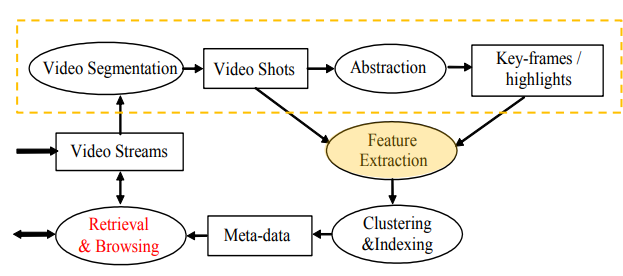
Week10 Video Data Analysis and Retrieval

Content based video retrieval(CBVR): 用于新闻，广告，音乐，电影

Cases： 1 用户已知剧情并确认在数据库里 2 用户已知剧情但是不知道数据库 3 搜关键字

Cbir vs cbvr： temporal info（时序信息）

Access video content： analysis（镜头边界，关键帧等）， representation（结构，帧序）， browsing（内容）， retrieval（检索）



镜头边界： pixel based易受noise影响，使用intensity stat来测量；也可以transform based或histogram based

关键帧： temporal，clustering，curve splitting

Video content representation/feature：

Sequential key frame representation： 平铺关键帧

Shot based rep： 包括镜头的移动和颜色变化

Scene based rep： 场景分类 更semantic

Temporal slice based： 一维扫描（一条线）沿时间变化的合集

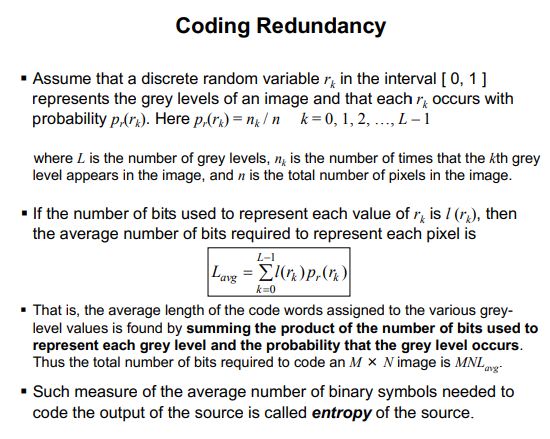
应用： 电影分类

Object based： 根据物品分类

Week11 Image Coding & Compression

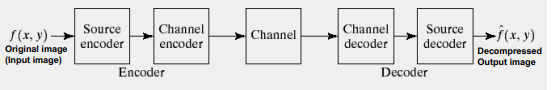
Data redundancy： 若n1n2代表相同物品 那n1的冗余度就等于1-1/（n1/n2）

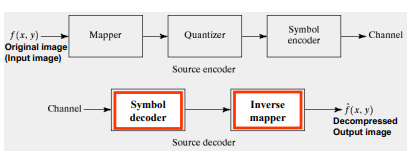
冗余来源： coding reduncancy， inter-pixel redundancy， phycho-visual reduncancy

把代表灰度的bit数量与灰度概率乘积，再取和 l（rk）是bit数量 ，Pr（rk）是灰度概率（灰度像素数/总像素数） 用symbol coder去掉

inter-pixel： 有的像素的值可以由相邻像素得到 （可以用阈值来去掉没用的东西） 用mapper去掉

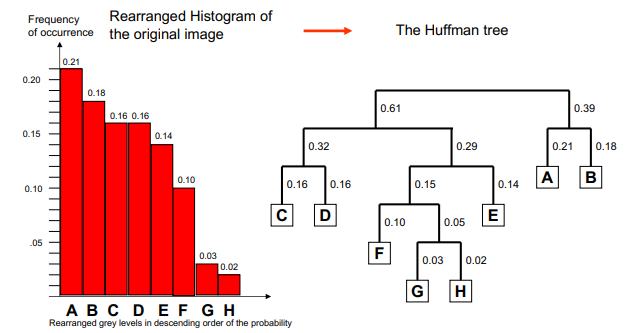
psycho-visual： 人类不用感知的东西 （灰度压缩）用quantizer去掉

图片压缩模型： 

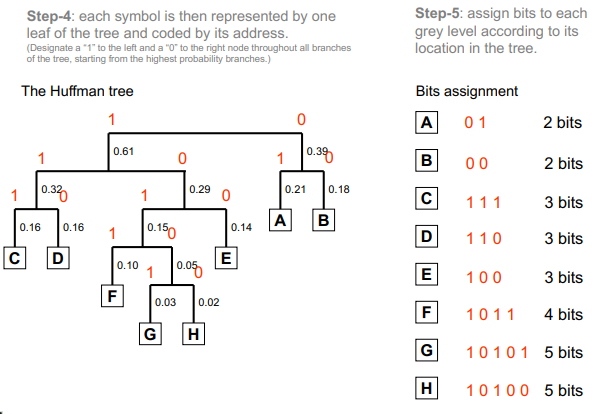


无损压缩： Huffman 编码： 减少编码所需符号

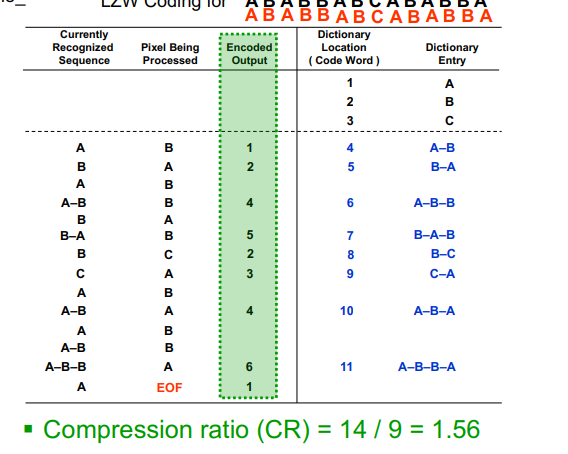
步骤： 1 获取histogram 2 从高到低重新排序 3 建立二叉树 （从最低概率开始 母节点是概率和 用概率和与剩下的最低概率形成新叉）

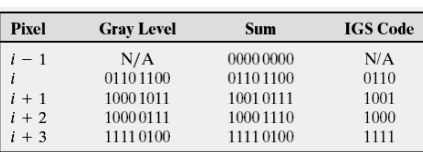


4 所有枝叶都是左为1右为0 ；；；5 根据灰度位置分配bit 并编写lookup table



LZW coding： 根据代码出现的规律去重新编写编码



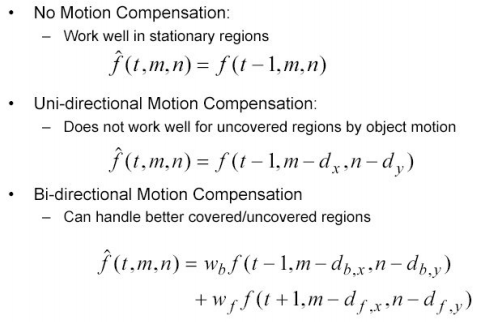
有损压缩： IGS quantization： 每个像素的8位灰度和前一个像素的后四位灰度bit取和，再留下前四位bit编码 

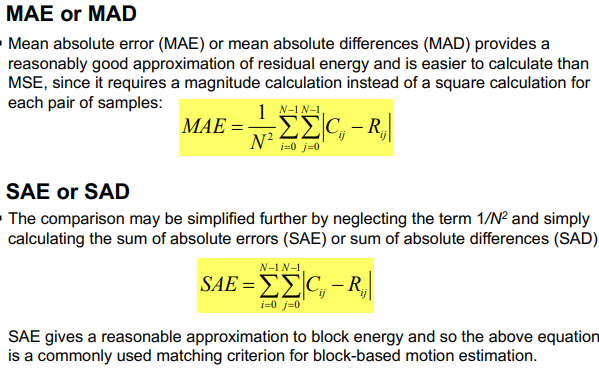
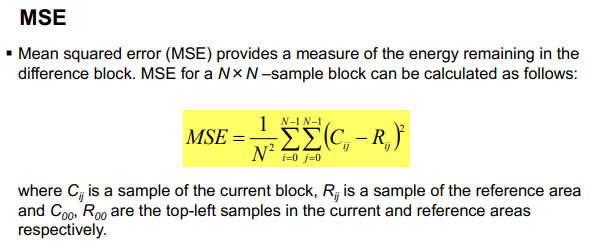
Week12 Video Coding & Compression

多一个temporal redundancy（gif只编码两帧的区别） 可用frame differencing

Motion compensation： 用镜头本体的移动补足两帧的区别 减少需要记录的差值

i-frame： 无依赖 p-frame： 从前一帧预测来 b-frame： 从前后预测来





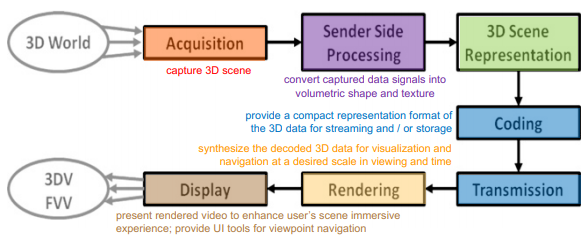
视频编码标准： ITU负责实时交换 ISO负责存储分放

Mpeg1： sub-pixel interpolation

Mpeg2： profile（不同媒体应用）和level（不同需求）

Mpeg4： object/toolkit coding， 支持低比特率。 Scalable coding： 减少复杂性，减少空间/时间分辨率， 减少质量

Week13 3DV and FVV： 3d video和free viewpoint video

、

Geometry based： 通过scene geometry表达 包括3d mesh，材质，但需要重建

Image based： 仅从2d图片中产生 不用建模但需要高分辨率和多重图片取样

通过multiview-camera和depth sensor获取对象3d布局

Sender side processing： 从采样中提取目标

Rendering： 同时渲染多维多时的图像 附加上scaling

Display： 1 传统的头瞄/摇杆操作 2 多用户直接置身于360度环境