# 一、绪论

1. 什么是图像，模拟图像和数字图像的区别，分别用什么来描述
2. 数字图像0表示白还是黑
3. 什么是数字图像处理，可以分为哪几个层次
4. 数字图像处理的历史与数字计算机的发展密切相关，它必须依靠\_\_\_\_\_\_\_及\_\_\_\_\_\_\_、\_\_\_\_\_\_\_和\_\_\_\_\_\_\_\_等相关技术的发展。
5. 数字图像的应用领域：最主要的图像源是\_\_\_\_\_\_，其他主要的能源包括\_\_\_\_\_\_、\_\_\_\_\_\_和\_\_\_\_\_\_\_。
6. 伽马射线成像的主要用途包括\_\_\_\_和\_\_\_\_\_\_。
7. X射线成像在医学中的应用
8. 紫外光的应用
9. 可见光和红外波段成像的应用
10. 身份的特征识别技术包括
11. 数字图像处理的主要研究内容
12. 数字图像应用

# 二、基本概念

1. 两种颜色模型，哥哥模型的属性的物理含义
2. 马赫带效应：人类视觉系统有增强边缘对比度的机制
3. 主观亮度与实际亮度之间的关系
4. 采样与量化
5. 黑白图像，灰度图像，彩色图像
6. 灰度直方图的定义
7. 常见matlab处理图像函数：

读入图像，显示图像，保存图像，

# 三、图像几何变换

1. 图像的几何变换是指图像处理中对图像\_\_\_\_、\_\_\_\_、\_\_\_\_和\_\_\_\_，这些简单变换以及变换中灰度内插处理等。
2. 矩阵乘法，可以实现绕原点的比例缩放、反射、错切和旋转等各种变换，但不能实现图像的平移以及绕任意点的比例缩放、反射、错切和旋转等变换
3. 如何解决任意点的操作（齐次坐标表示法：以n+1维表示n维）
4. 图像的放大和缩小是如何实现的
5. 图像旋转的矩阵，指明是顺时针旋转还是逆时针旋转
6. 行插值和列插值的具体实现，是跟前一个一样还是后一个一样
7. 仿射变换的一般表达式，有几个自由度？平移、比例缩放和旋转变换都是一种称为仿射变换的特殊情况
8. 什么是透视变换？透视变换的一般表达式，有几个自由度？透视变换也是一种平面映射 ，并且可以保证任意方向上的直线经过透视变换后仍然保持是直线。
9. 灰度插值的几种方法：\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_\_、\_\_\_\_\_\_。分别给出各个插值的原理或者公式。
10. 四种插值，分别有什么优缺点，卷积插值有哪些模板

# 四、图像频域变换

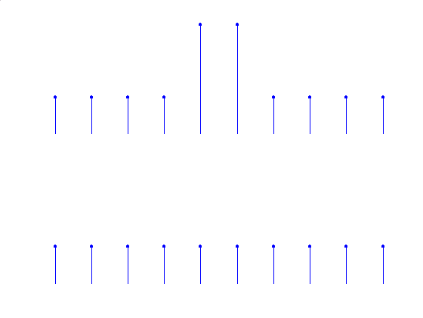
1. 傅里叶变换是一种线性的积分变换
2. 傅里叶变换的典型用途是将信号分解成\_\_\_\_\_分量和\_\_\_\_\_分量。
3. 一维傅里叶变换和傅里叶逆变换的定义
4. 傅里叶变换的复数和极坐标形式
5. 一维离散傅里叶变换与逆变换的定义
6. 一维离散傅里叶变换与逆变换中的1/N是什么作用
7. 二维离散傅里叶变换与逆变换的定义
8. 二维离散傅里叶变换的可分离性
9. 二维离散傅里叶变换的可平移性（频域和空域），如果时，有什么简化方法
10. 二维离散傅里叶变换的周期性、旋转性、线性性、尺度缩放、卷积与相关定理
11. F(0,0)与图像均值的关系
12. Matlab中与傅里叶变换相关的函数(六个)
13. 离散余弦变换的使用条件
14. 从一维离散傅里叶变换推导一维离散余弦变换(联想翻折)
15. 二维离散余弦变换的公式

# 五、图像增强

空域增强部分

1. 灰度线性变换的映射函数，如何实现对比度的变换？
2. 分段线性处理的映射函数，对数变换的映射函数，灰度切分的映射函数
3. 位面（对于8bit）：第0位将原图灰度级分成0-127和128-255这两个范围；第1位将灰度级分成0-63、128-191、64-127和192-255这四个范围。
4. 直方图均衡化是通过对原图像进行某种变换，使原图像的灰度直方图修正为\_\_\_\_\_\_\_\_的直方图的一种方法。
5. 直方图均衡化方法的基本思想是对在图像中像素个数多的灰度级进行\_\_\_\_\_，而对像素个数少的灰度级进行\_\_\_\_\_，从而达到清晰图像的目的。 所谓的平坦，其实是让最终保留下来的灰度级的分布尽可能均等。
6. 直方图均衡化的灰度变化函数（连续和离散）
7. 直方图均衡化步骤
8. 直方图均衡化的实质
9. 直方图均衡化缺点：增强效果不易控制
10. 直方图规定化动机：使直方图成为某个特定的形状
11. 直方图规定化(匹配)的计算过程
12. 空域滤波根据操作特点，分为\_\_\_\_滤波和\_\_\_\_滤波；根据滤波效果又分为\_\_\_\_\_滤波和\_\_\_\_\_滤波；空间滤波器是由两部分组成：\_\_\_\_\_和\_\_\_\_\_\_操作
13. 图像噪声的定义，可以借助什么来描述，有哪些特征
14. 噪声的分类：加性噪声，乘性噪声
15. 椒盐噪声：受噪声干扰的图像像素以50％的相同概率等于图像灰度的最大（表现为使相应的像素点变白）或最小（表现为使相应的像素点变黑）的可能取值。
16. 高斯噪声：噪声的取值分布呈高斯分布
17. 均值滤波，优点：算法简单、处理速度快；缺点：会使图像产生模糊，特别在边缘和细节处。而且邻域越大，在去噪能力增强的同时模糊程度越严重。
18. 典型的加权平均滤波器：



1. 中值滤波的一般计算过程
2. 对于离散的脉冲信号，当其连续出现的次数小于窗口尺寸的一半时，将被抑制掉，否则将不受影响。
3. 中值滤波保留图像中的相对平坦成份，或者比较模糊的边缘。优点：对椒盐噪声的抑制效果好，在抑制噪声的同时能有效地保护图像的边缘；缺点：随着窗口尺寸的增大，虽然能更好地滤除噪声，但图像的模糊程度加重；对点、线等细节较多的图像不太合适(原因如上一个所示)。
4. 对于椒盐噪声，为什么中值滤波效果比均值滤波效果好？
5. 对于高斯噪声，为什么均值滤波效果比中值滤波效果好？
6. 边界保持平滑滤波的动机
7. 超限像素平滑法的表达式
8. 超限像素平滑法的优点：对抑制椒盐噪声比较有效，对保护仅有微小灰度差的细节及纹理也有效；缺点：随着邻域增大，去噪能力增强，但模糊程度也变大。
9. 在超限像素平滑法的操作中，对窗口的大小及阈值𝑻的选择要慎重。𝑻太大，噪声消除不干净； 𝑻太小，易使图像模糊。
10. 线性锐化滤波器(高通滤波器)的动机：从原图像中减去它的低频分量，即原始图像经平滑处理后所得的图像



**平滑窗口**

**锐化模板**

1. 微分操作可以增强高频成分
2. 数字图像中的梯度(差分)：



1. Roberts算子，Prewitt算子，Sobel算子，三个算子如何不断改进，考虑了哪些因素
2. 梯度图像的几种输出方式
3. Laplacian算子：二阶微分算子，用二阶差分近似。
4. Laplacian算子获得的边界是比较细致的边界(边缘的边缘)，反映的边界信息包括了许多的细节信息，但是所反映的边界不是太清晰。
5. Laplacian增强算子的模板
6. 频域滤波分为\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_、\_\_\_\_\_\_
7. 振铃现象的出现是由于理想低通滤波器在时间域中采用 sinc 函数作为冲激响应，导致在信号或图像的边缘产生不必要的震荡。虽然理想滤波器在频域上能够精确处理信号，但在实际应用中，我们通常会使用更实际的滤波器(巴特沃斯滤波器)设计来减少或消除这种振铃效应。
8. 理想低通滤波所产生的“振铃”现象在2-D图像上表现为一系列同心圆环，且圆环半径反比于截断频率，截止频率越小，振铃效应越明显(低通滤波器)。
9. 高通滤波器，截止频率越低，保留下来的高频成分更加显著，高频震荡增加，振铃现象越明显。
10. 巴特沃斯滤波器优点：的通带与阻带之间没有明显的不连续性，因此在抑制噪声的同时，图像边缘的模糊程度大大减小，没有振铃效应产生；缺点：尾部含有较多的高频成分，对噪声的平滑效果不如理想低通滤波器。
11. 带阻、带通滤波器必须以关于原点对称的形式出现

1. 同态滤波的基本思想：减少入射分量，并同时增加反射分量来改善图像的显示效果



1. 在空间上变化缓慢，其频谱集中在低频段，反映图像的细节和边缘，其频谱集中在高频段。
2. 同态滤波步骤
3. 空域中的平滑滤波器在频域里对应低通滤波器

截止频率越高，空域越窄(可以理解为滤波器核大小)，平滑作用越弱

截止频率越低，空域越宽(可以理解为滤波器核大小)，模糊作用越强

1. 空域中的锐化滤波器在频域里对应高通滤波器

截止频率越高，空域越窄，检测的边缘越少

截止频率越低，空域越宽，检测的边缘越多

1. 频域技术和空域技术区别：

空域技术中无论使用点操作还是模板操作，每次都只是基于部分像素的性质，而频域技术每次都利用图像中所有像素的数据，具有全局性，有可能更好地体现图像的整体特性，如整体对比度和平均灰度值等。

1. 局部增强，多了一个选择局部区域的步骤





*M* — 图像***f(x,y)***的整体均值

压制灰度强变化的部分。变化越强，局部标准差越大，A越小，从而降低灰度差异值，使整体局部的灰度分布尽可能均衡。

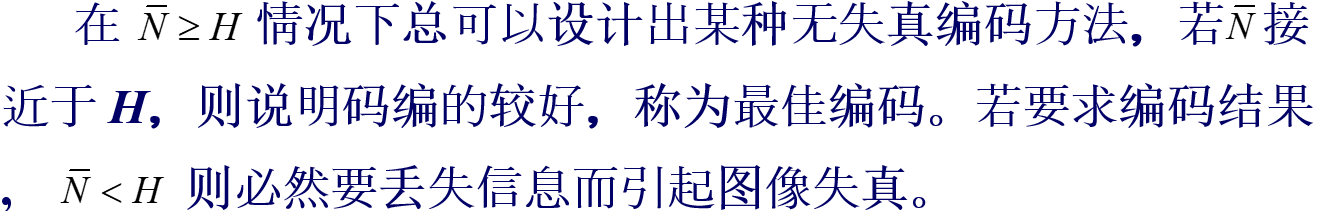
# 六、图像压缩

1. 图像数据本身固有的冗余性包含：\_\_\_\_\_\_、\_\_\_\_\_\_、\_\_\_\_\_\_。分别解释
2. 图像数据压缩的可能性：一般原始图像中存在很大的冗余度、用户通常允许图像失真、当信道的分辨率不及原始图像的分辨率时，降低输入的原始图像的分辨率对输出图像分辨率影响不大、用户对原始图像的信号不全都感兴趣，可以提取有用的信息。
3. 压缩比：原始图像每个像素的平均比特数n1和编码后每个像素的平均比特数n2的比值。
4. 无损压缩算法中删除的仅仅是图像数据中冗余的信息，因此在解压缩时能精确恢复原图像，无损压缩的压缩比很少能超过3：1,常用于要求高的场合
5. 有损压缩是通过牺牲图像的准确率以实现较大的压缩率，有损压缩在压缩比大于30：1时仍然可重构图像，而如果压缩比为10:1到20:1，则重构的图像与原图几乎没有差别。
6. 客观保真度准则：当信息损失的程度可以表示成初始图像或输入图像以及先被压缩而后被解压缩的输出图像的函数时,就说这个函数是基于客观保真度准则的。
7. 均方信噪比：



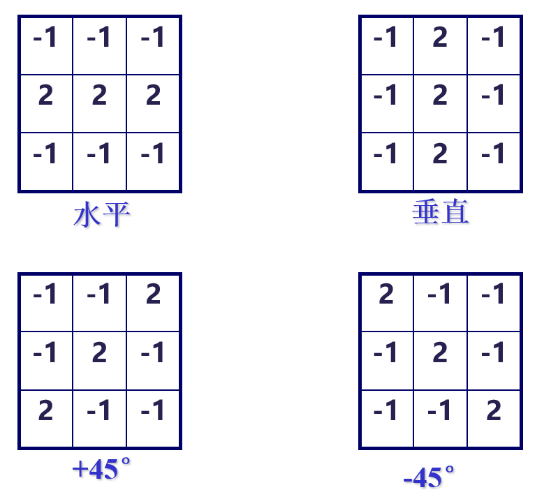


1. 图像的熵的定义，熵越大，信号的随机程度越大。
2. 图像的码字平均长度、编码效率和冗余度
3. 香农信源编码信息论：



1. 霍夫曼编码的过程
2. 霍夫曼编码的局限性：编出的码不是唯一的、码字长短不一，实时硬件实现很复杂、编码前必须有信源这方面的先验知识、只能用近似的整数位来表示单个符号，而不是理想的小数

# 七、灰度图像（边缘检测）

1. 图像分割的基本策略，基于灰度值的两个基本特性：区域之间的不连续性，区域内部的相似性
2. 不连续性检测，常用空域滤波实现，卷积结果大于阈值，即为被检测到，包含点检测，线检测

点检测模板

**-1**

**-1**

**-1**

**-1**

**8**

**-1**

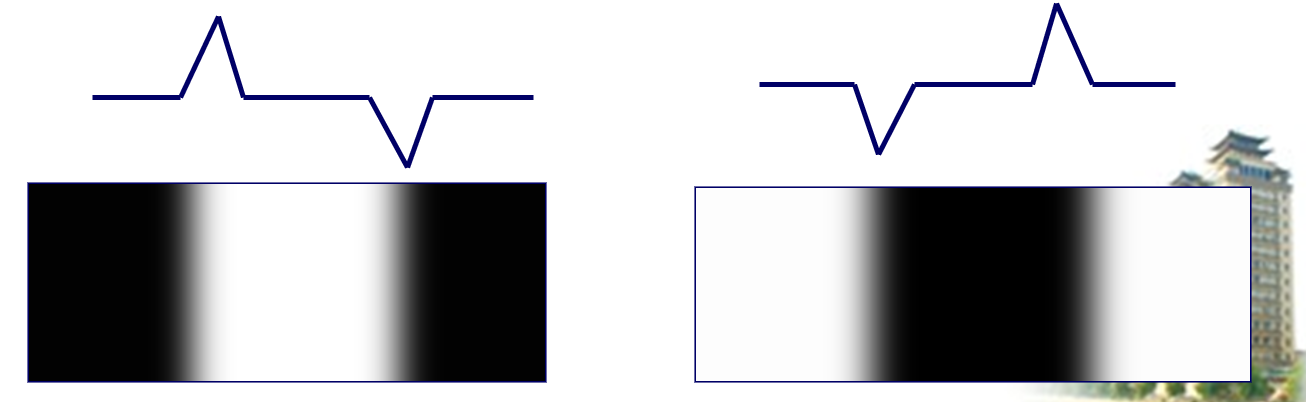
**-1**

**-1**

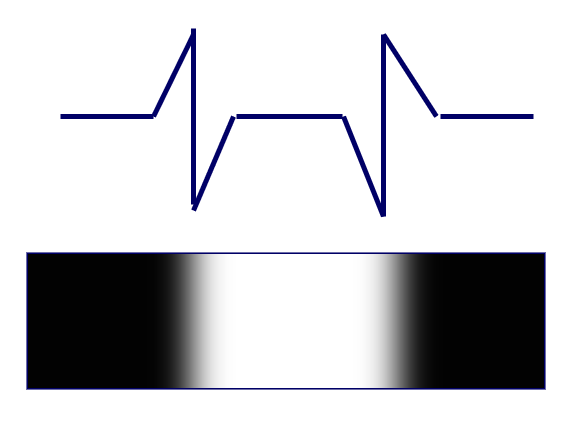
**-1**

线检测模板

1. 边缘检测：对于实际数字斜坡边缘，一阶导数的幅度峰值对应边缘位置；二阶导数的过零点对应边缘位置，二阶导数在过零点附近的符号确定边缘象素在图像边缘的暗区或明区。



一阶微分，黑到白，0-255，数量增加，所以一阶微分围成面积为+



二阶微分，在亮的一边是负的；在暗的一边是正的；因为暗对应0，亮对应255，暗只能增加，亮只能减少。

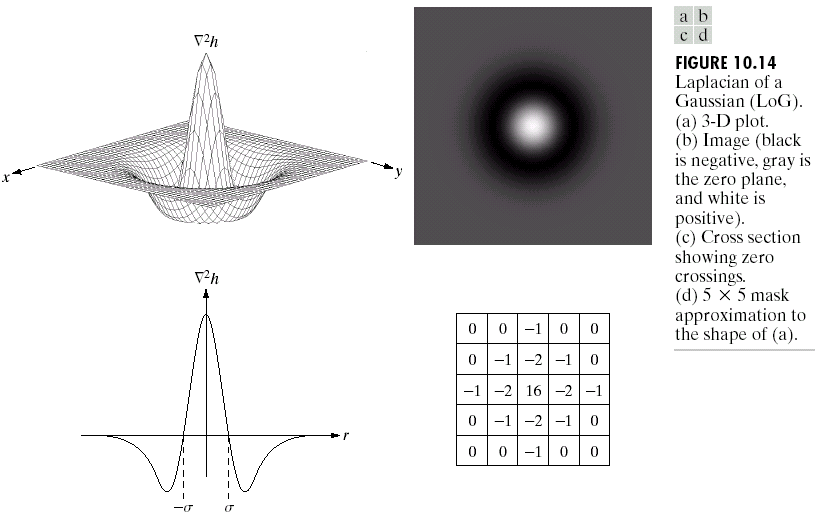
1. Sobel算子相较于Prewitt算子，保留了部分高频，即保留了更多的边缘，同时中间的像素值权重大，噪声抑制效果更好，是用于噪声较多的图像。但实际效果好像差不多(？)
2. 回顾拉普拉斯算子
3. 拉普拉斯算子提取更细的边缘，对噪声具有无法接受的敏感性，对于一个边缘，产生两个值，且不能检测边缘的方向
4. LOG算子(Laplacian of Gaussian)：先使用高斯平滑





再对平滑后的图像求二阶导数(Laplacian算子)

LOG算子可以通过任何一个方形核进行逼近，但要保证LOG核的所有元素的和或均值为0。



1. Canny算子的动机：对边缘的错误检测率要尽可能低、标记边缘要和图像上真正边缘的位置尽量接近、对同一边缘要有低的响应次数、克服噪声
2. Canny边缘检测算法：

* 用高斯滤波器平滑图像；
* 用一阶偏导有限差分计算梯度幅值和方向；
* 对梯度幅值应用非极大值抑制；极大值抑制的目的是找出图像梯度中的局部极大值点，把其它非局部极大值点置零以得到细化的边缘。
* 用双阈值算法检测和连接边缘。T1 = 2\*T2，得到两个边缘图，高阈值边缘图和低阈值边缘图，连接高阈值边缘图，通过低阈值边缘图补充高阈值边缘图(比如边缘的断裂处)。

1. Matlab中边缘检测的调用：



1. 几种算子的比较：

Roberts算子：边缘定位精度高，但容易丢失一部分边缘，且不具备能抑制噪声能力。

Sobel算子和Prewitt算子：对噪声具有一定的抑制能力，但不能完全排除检测结果中出现的虚假边缘。虽然这两个算子边缘定位效果不错，但检测出的边缘容易出现多像素宽度。

Laplacian算子：不依赖于边缘的方向，对阶跃型边缘定位准确，但对噪声非常敏感。所以，Laplacian算子容易丢失一部分边缘的方向信息，造成一些不连续的检测边缘。

LOG算子：在使用高斯函数对图像作平滑滤波处理来抑制噪声的同时也可能将原有的比较尖锐的边缘也平滑掉了，造成这些尖锐边缘无法检被测到。且对于不同图像应选择不同参数。

Canny算子：同样该算子也会将一些高频边缘平滑掉，造成边缘丢失。Canny算子其后所采用双阈值算法检测和连接边缘，采用的多尺度检测和方向性搜索较LOG算子要好。

# 八、灰度图像（阈值分割）

1. 阈值分割法的基本思想，可以分为全局阈值、局部阈值、动态阈值
2. 直方图得到阈值(要求直方图双峰明显)的改进：选取双峰的中间值而不是谷底、平滑以去除噪声
3. 迭代阈值选择，类似k-means。
4. 最佳阈值分割(最小错误率)：

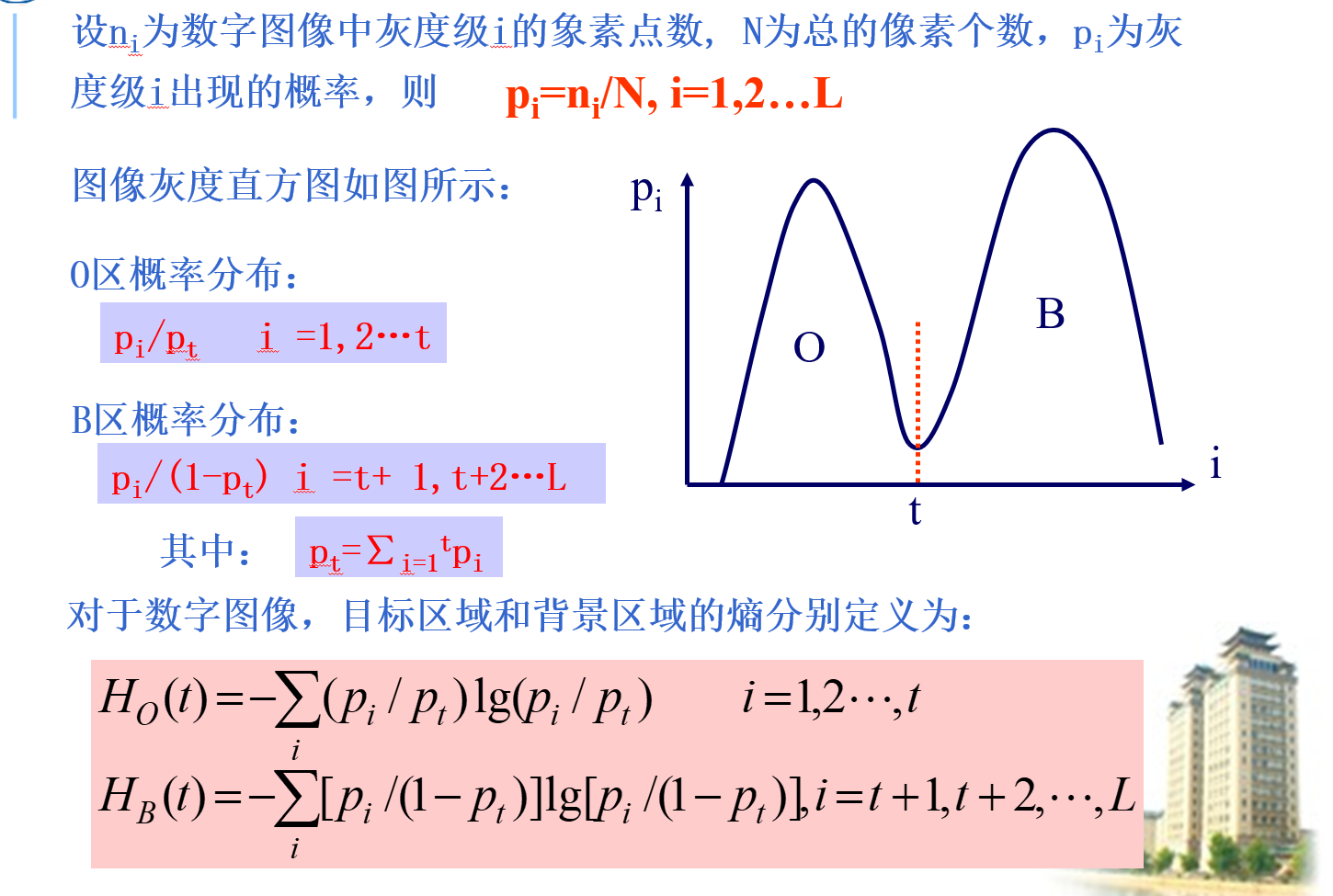
目标的灰度分布概率密度为*po(r)，*而背景的灰度分布概率密度为*pb(r)，*同时设目标占整个画面的百分比为θ，则背景占1-θ。若取阈值为t，有：



再带入分布进行计算得到t。

对于复杂图像，使用多阈值进行改进，结合邻域信息判断object还是background

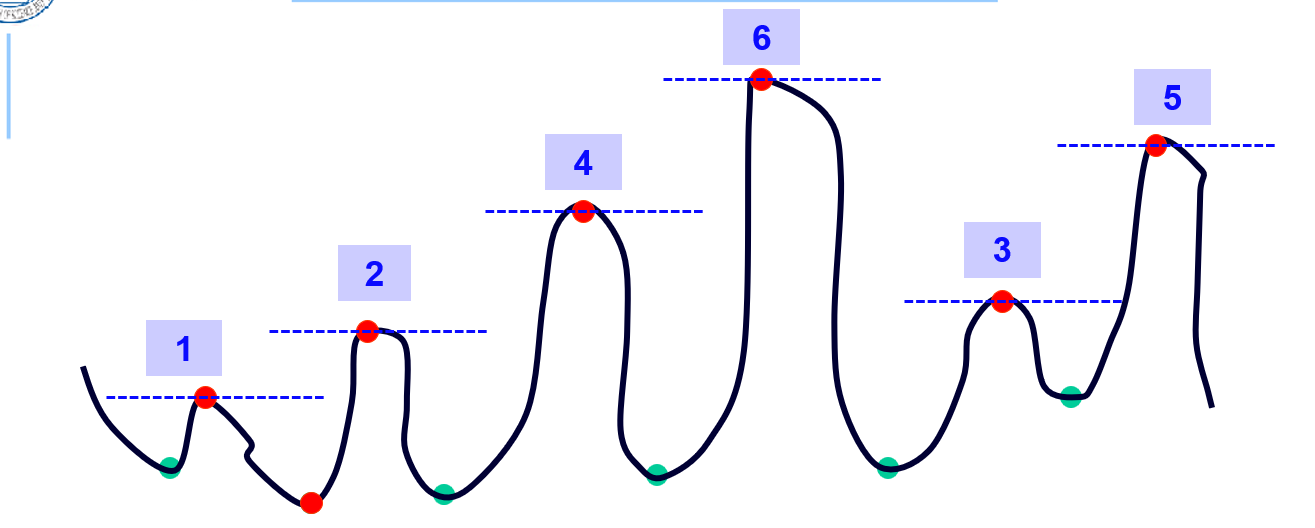
1. Otsu法阈值选择，思想是通过设定不同阈值，统计对应直方图，使类间方差最大。
2. 一维熵（描绘灰度直方图的一个测度，H越大，灰度分布越均匀）阈值分割：



使Ho+HB最大。

# 九、彩色图像分割（分水岭+MS）

1. 描述分水岭算法的递归过程



1. 分水岭方法应用在梯度图像，那么集水处对应灰度变化最小（平坦）的区域，而分水岭对应灰度变化相对最大（边缘）的区域。
2. 分水岭之前要进行平滑去噪处理，减少区域个数，防止过度分割
3. 标记约束分水岭算法：从先前已经定好的区域开始浸水，防止过度分割
4. 分水岭算法可以起平滑作用，一次分水岭，连通区域设为同一灰度值，再次进行分水岭，可以改善效果
5. 分水岭算法优点：快，效果较好；缺点：容易过分割，没有自上而下的信息（？）
6. Mean Shift 本质上是一种根据梯度上升进行自适应点搜索的方法，可用于对模型表面顶点的局部高度进行聚类分析
7. Mean Shift强大之处是可以选择任意的特征空间，包括颜色空间、尺度空间……在任意的特征空间中进行非参数密度的密度梯度估计，也可以融合多个特征空间。
8. 核函数用于计算当前点X与周围其他点Xi的相似度（Xi在X处的概率密度贡献）。
9. 用核密度估计（Kernel Density Estimation, KDE）来估算数据密度的梯度。这个梯度告诉我们在当前位置向哪里移动可以增加密度，即找到更密集的数据区域。



第一项为X处的概率密度，第二项为邻域质心-X，刻画偏移的方向与距离

1. Mean Shift的优缺点

优点：

* 应用无关：Mean Shift是一种通用工具，适用于各种数据类型和应用场景。
* 适合实际数据分析：适用于处理真实数据，能有效应对数据的复杂分布。
* 无需预设数据簇的形状：不需要假设数据簇的形状（如椭圆形等），能处理任意形状的聚类。
* 支持任意特征空间：能在不同类型的特征空间中工作，不局限于特定的数据特征。
* 仅需选择一个参数：算法只需一个参数，即窗口大小h，使得参数选择较为简单。
* 窗口大小h有物理意义：窗口大小h具有实际意义，即“局部邻域”，而不像K-Means中的聚类数那样抽象。窗口大小h决定了算法在当前位置周围多大范围内计算密度梯度。这个范围中的点对均值偏移方向的计算有贡献，而范围外的点则不计入。换句话说，h越大，纳入计算的邻域越大；h越小，算法只关注更靠近的数据点。

缺点：

* 窗口大小的选择较为复杂：选择合适的窗口大小（带宽）不容易。
* 不适合的窗口大小会导致聚类问题：窗口过大会导致簇被合并，窗口过小则可能生成额外的“浅层”簇。建议使用自适应窗口大小来应对不同密度的数据分布。

1. 最后一页要记住的：

分割的用途

* 提高效率
* 提取更好的特征
* 获得所需的分割对象

Mean-shift分割

* 一种通用的分割方法
* 适用于各种聚类和追踪任务

分水岭分割

* 适合层次分割
* 常与边界预测结合使用

# 十、灰度图像分割（数学形态学）

1. 数学形态学进行图像处理的基本思想是：用具有一定形态的结构元素探测目标图像，通过检验结构元素在图像目标中的可放性和填充方法的有效性，来获取有关图像形态结构的相关信息，进而达到对图像分析和识别的目的。
2. 二值形态学运算——腐蚀：

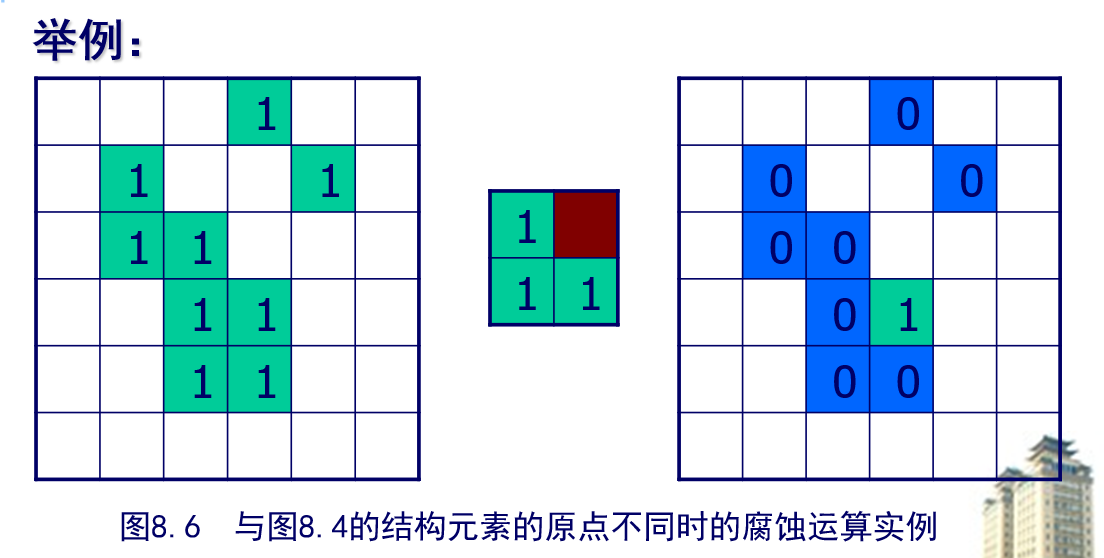
A为图像，B为结构元素



1. 腐蚀运算的过程

把结构元素B看作为一个卷积模板，每当结构元素平移到其原点位置与目标图像A中那些像素值为“1”的位置重合时，就判断被结构元素覆盖的子图像的其它像素的值是否都与结构元素相应位置的像素值相同；只有当其都相同时，就将结果图像中的那个与原点位置对应的像素位置的值置为“1”，否则置为0。

**注意：当结构元素在目标图像上平移时，结构元素中的任何元素不能超出目标图像的范围。**



1. 二值形态学运算——膨胀：

A为图像，B为结构元素



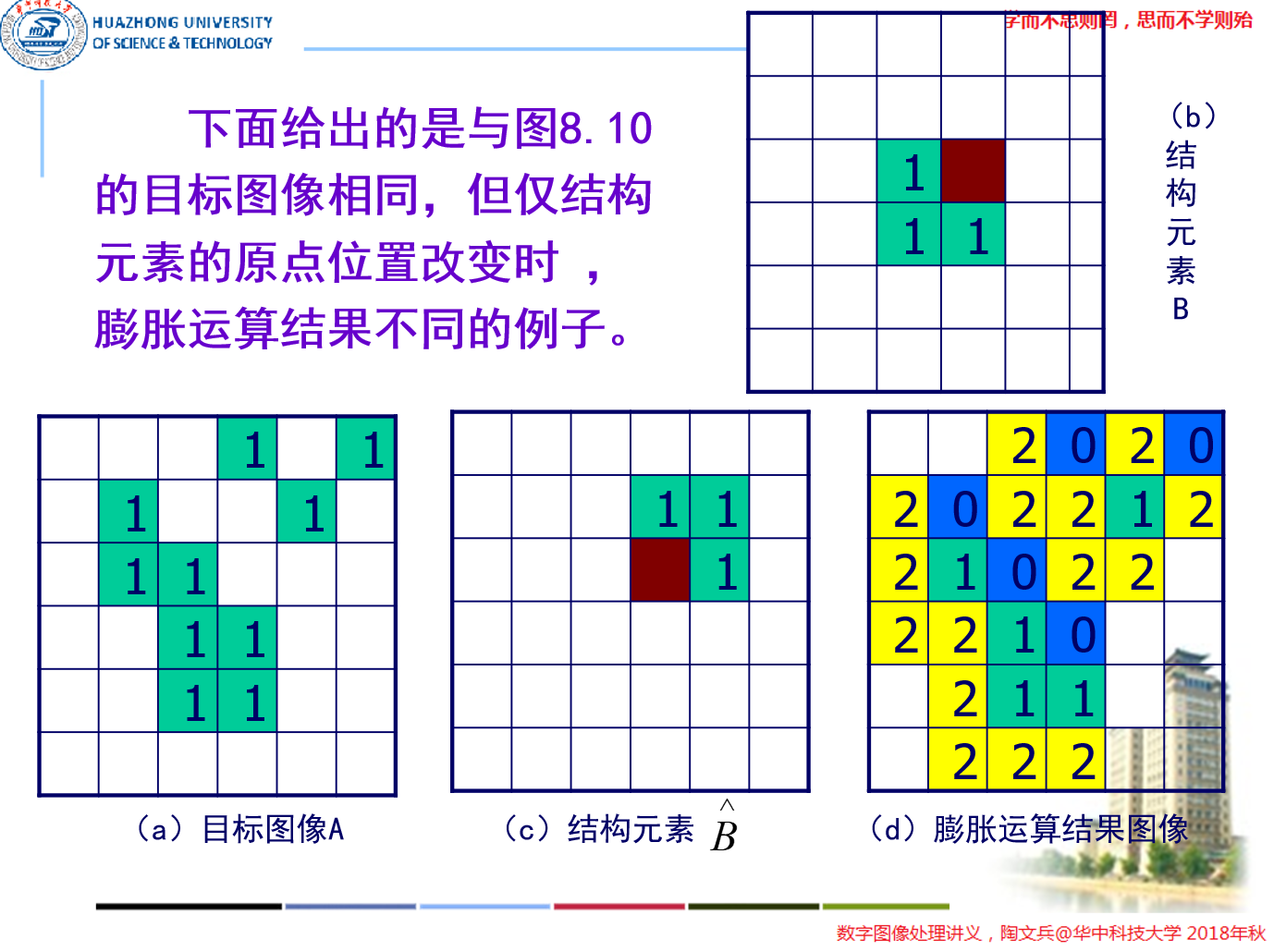
1. 膨胀运算的过程：

（1）求结构元素B关于其原点的反射集合；

（2）每当结构元素B在目标图像A上平移后，结构元素B与其覆盖的子图像中至少有一个元素相交时，就将目标图像中与结构元素B的原点对应的那个位置的像素值置为“1”，否则置为0。

**注意：（1）当结构元素中原点位置的值是0时，仍把它看作是0；而不再把它看作是1。**

**（2）当结构元素在目标图像上平移时，允许结构元素中的非原点像素超出目标图像范围。**

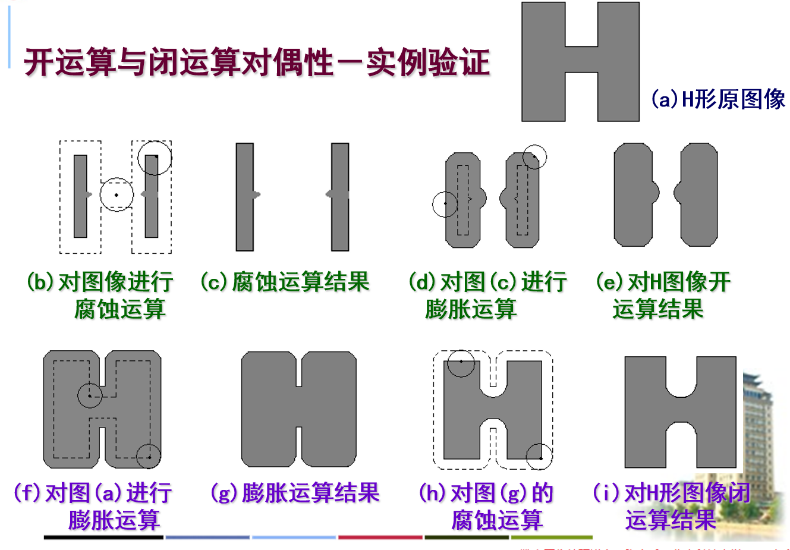


1. 对偶性：对目标图像（置为1）的膨胀运算，相当于对图像背景（置为1）的腐蚀运算操作（结构元素B要反射）；对目标图像的腐蚀运算，相当于对图像背景的膨胀运算操作（结构元素B要反射然后再加上膨胀的反射）。
2. 开运算（主调为腐蚀）：使用同一个结构元素对目标图像先进行腐蚀运算,然后再进行膨胀运算称为开运算。



1. 闭运算（主调为膨胀）：使用同一个结构元素对目标图像先进行膨胀运算,然后再进行腐蚀运算称为闭运算。



1. 开闭运算互为对偶，闭运算具有磨光物体内边界（内部0->1）的作用，而开运算具有磨光图像外边界（边界1->0）的作用。

# 十一、彩色图像分割 （Ncuts分割）

1. Ncuts分割的思想：图像分割问题就是根据像素点之间的相似程度来确定不同的区域，从而分割图像，则定义相邻像素点之间的相似程度为边，相似程度越大，边的权值越大。而且很显而易见的一点是:不同区域之间的差异是非常大的，因此相邻区域之间连通的边的权值，一定都是较小的值(不同区域之间相似程度小，相邻边的权值就很小)。所以我们可以这样想: 找到了权值之和比较小的边集，就可能相当于找到了区域的边界。
2. 边的权值的定义：



第一项：亮度、颜色或纹理相似度，这一项衡量的是像素u和v在亮度或颜色上的差异，通常称为“强度相似度”。

第二项：空间距离相似度，这一项衡量的是像素u和v在空间位置上的距离差异，称为“空间相似度”。

分母均为归一化参数

只考虑空间距离小于r的像素对之间的连接：

**限制的原因**

1. **减少计算量**：在图像中每个像素与所有其他像素都计算权值会导致大量计算。通过限定距离r，只考虑距离较近的像素对，可以显著减少计算量。
2. **局部相关性**：图像中的相似性通常是局部的，即距离较近的像素更可能属于同一物体或区域。限制距离可以更好地捕捉图像中的局部结构（抓边缘），减少错误连接远距离像素的可能性。
3. **提高分割效果**：过远的像素通常不具有直接的相关性，不应相互影响。在图像分割中，只考虑距离较近的像素，有助于更精确地识别图像中的边界。

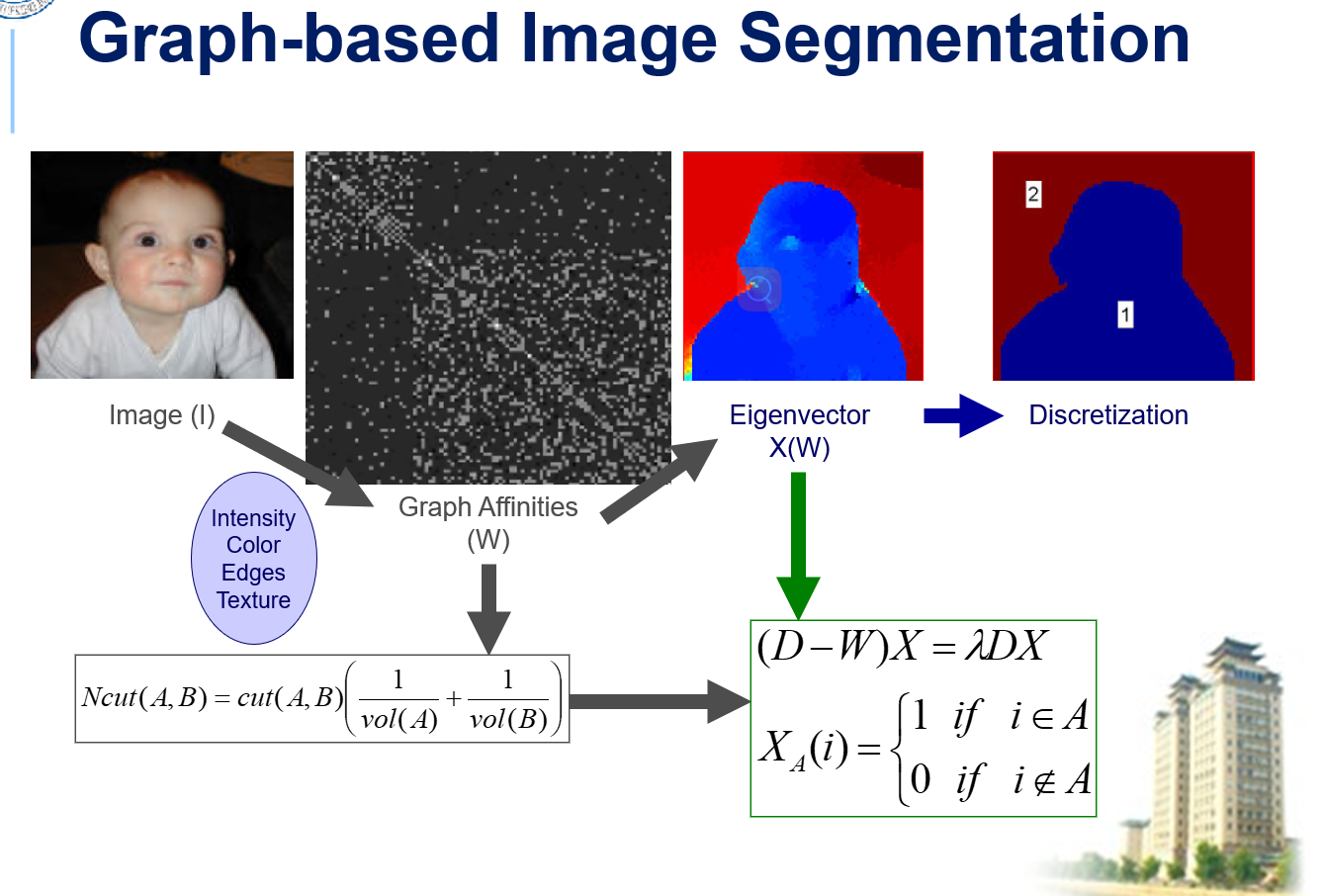


1. 如何解决分割出孤立点的情况：

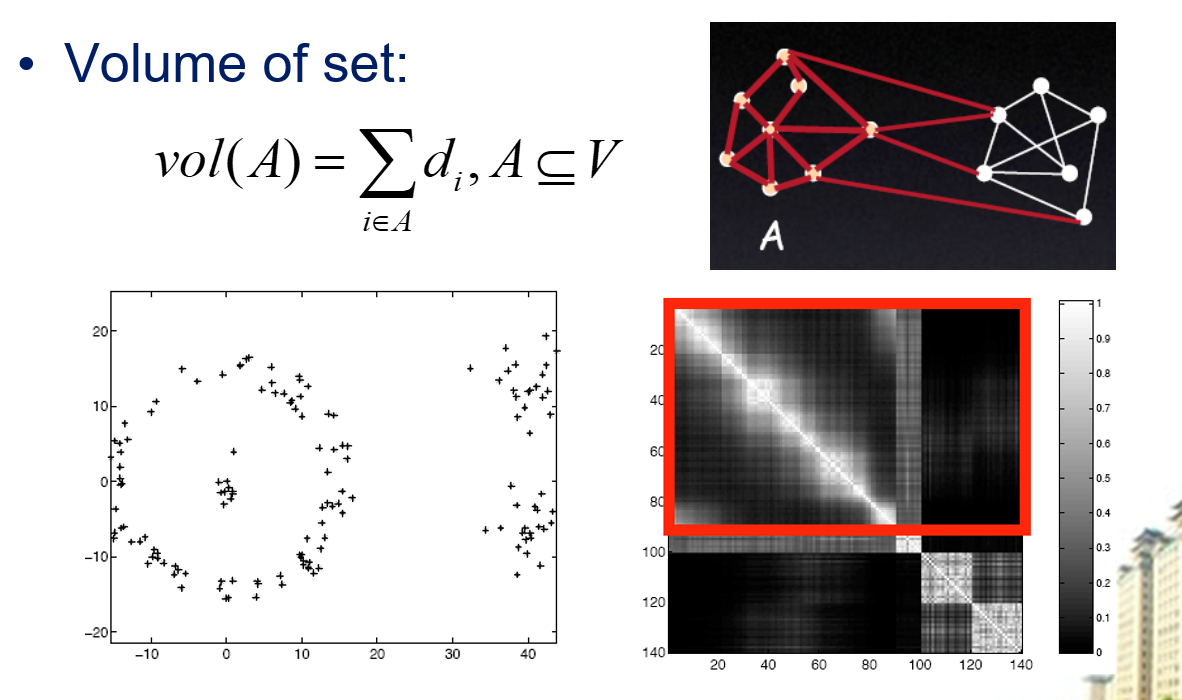
其中

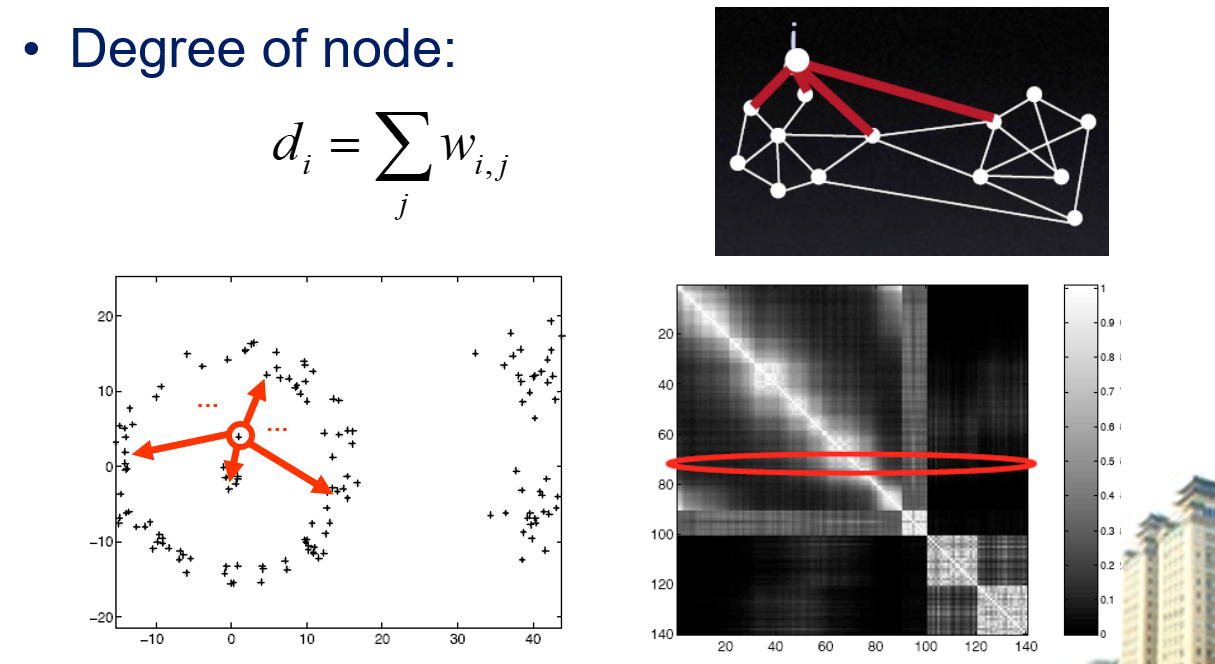
A,B均为点集，V为整个图像的总像素点集。ANB = cut(A,B)就是分割A，B两集合的割集的权值和，assoc(A,V)表示A中所有点与图中所有点 (即点集V) 相连的边的权值和。引入assoc这个函数后，就能比较好地避免分割出孤立点的情况。

1. Ncuts的过程：



Vol：





D为对角矩阵，对角线上记录了每个像素点与其相邻像素点的总相似度之和

最终求解得特征向量，用于分割

1. Ncut的改进方法：

### 原来的 Ncut 方法

在标准的 Ncut 方法中，图像的每个像素都被视作一个节点，计算每个节点之间的连接权重。如果一幅图像有 NNN 个像素，那么理论上需要计算每对像素之间的连接权重，这样计算量就是 **N×(N−1)/2**，计算复杂度是 O(N^2)。对于高分辨率图像，这个计算量是非常庞大的，因为图像的像素数量通常是数百万甚至更多。

### 改进后的方法：基于灰度级集合

改进方法将像素按灰度级分组。灰度级通常是 256 级，即从 0 到 255。如果我们将图像分成256个灰度级集合V0,V1,…,V255​，那么我们只需要计算每对灰度级集合之间的连接权值，而不是每对像素之间的连接权值。这样计算量从对每对像素的计算简化为对**灰度级集合之间**的计算。

假设我们只需计算每个灰度级集合之间的权值，那总共需要的计算量是 **256×255/2=32640**次，而不再取决于图像的像素数目。这对大图像来说，计算量减少得非常显著。

### 举个例子：班级交流网络

假设我们要统计一个班级中同学之间的社交联系。如果班级里有 100 名同学，按照传统方式，我们需要统计每对同学之间是否有联系，总共就是 **100×99/2=4950**次计算。

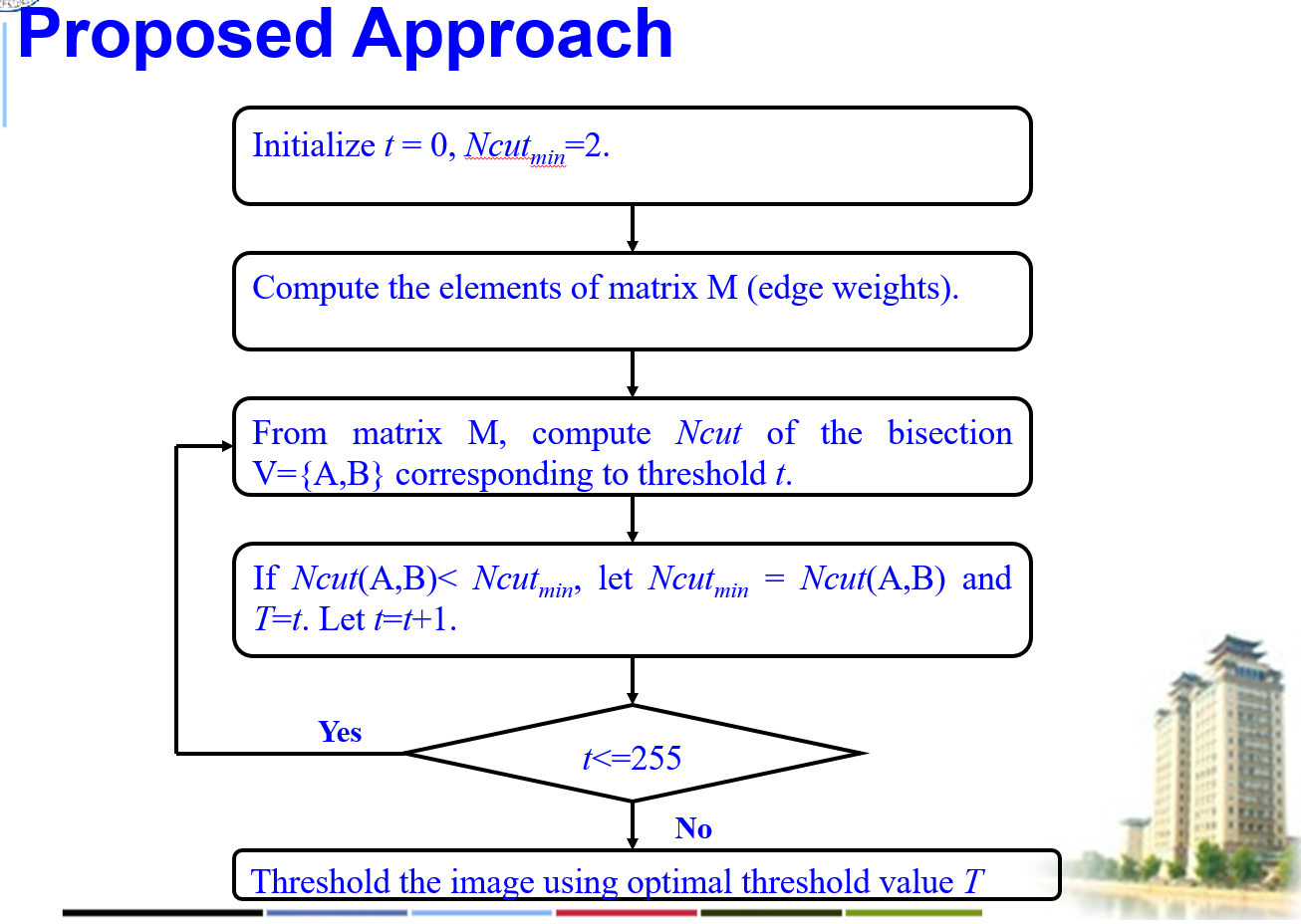
但是我们发现，这些同学可以按不同的兴趣分成若干小组，比如：喜欢篮球的、喜欢音乐的、喜欢编程的等。每个兴趣小组有几个或几十个同学。于是我们可以简化统计过程，不去关注每个同学的具体联系，而只统计不同兴趣小组之间是否有交流。比如，如果篮球组和音乐组之间有联系，我们就认为这两个组之间有交流，而不具体区分每个同学。

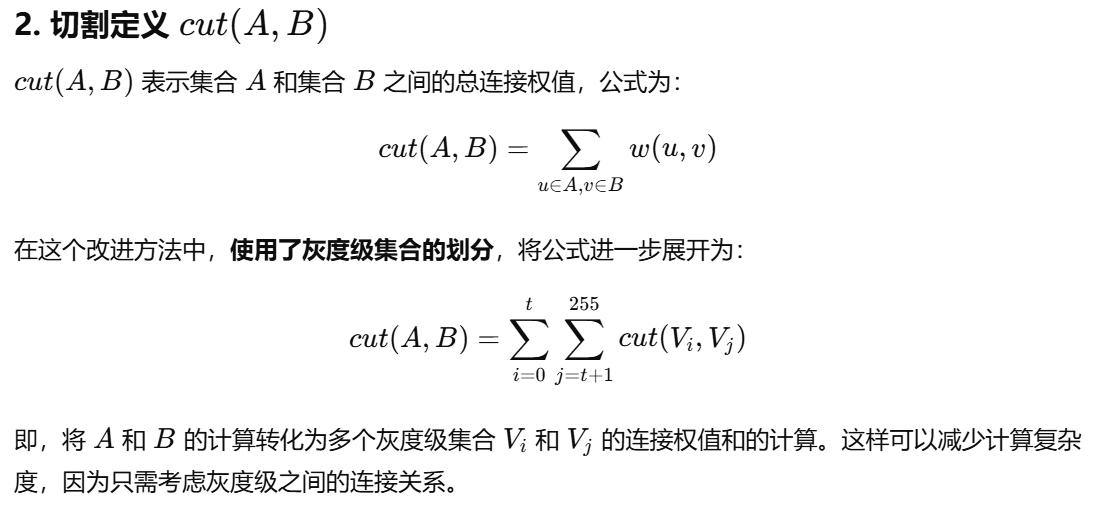
这样一来，计算量从“所有同学对所有同学”变成了“所有小组对所有小组”。假设有 10 个兴趣小组，那么只需要 **10×9/2=45**次计算，大幅减少了工作量。

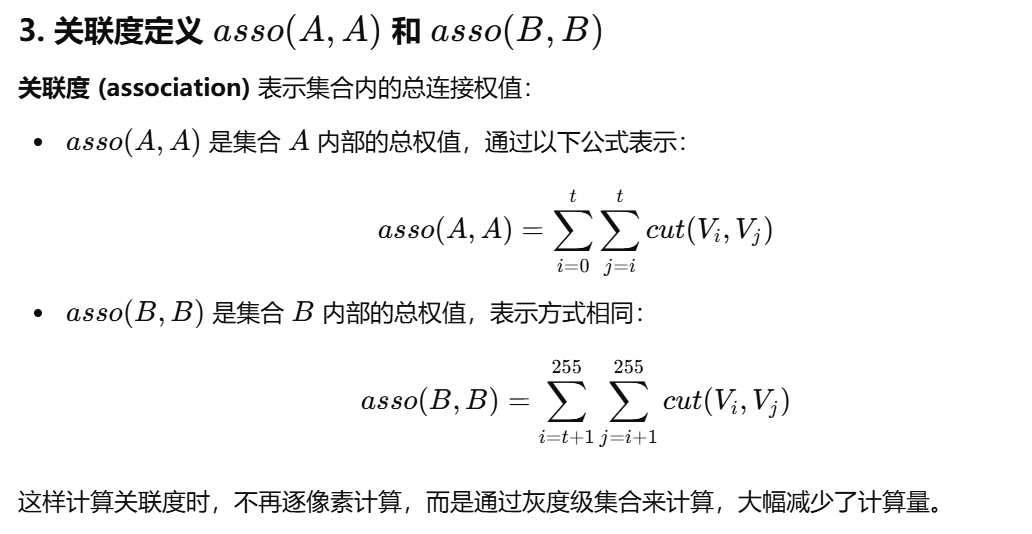
### 回到图像处理的例子

在图像处理中，按灰度级分组相当于将所有灰度值相同的像素视为一个集合，不再区分每个像素的具体位置。然后只需要计算每个灰度级集合之间的关系，而不是每对像素的关系。这样不仅减少了计算量，还保持了图像中相似灰度像素的连贯性。

1. 改进算法：







1. 优点：低计算成本、适合实时视觉处理、鲁棒性强

# 十二、SIFT（尺度不变特征变换匹配算法）

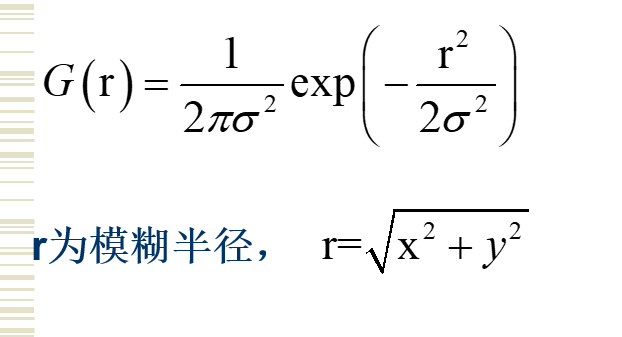
1. SIFT的动机：急需提出一种鲁棒性强、在不同光照、不同位姿等情况下能够有效识别目标的方法。
2. SIFT简介：一幅图像映射（变换）为一个局部特征向量集；特征向量具有平移、缩放、旋转不变性，同时对光照变化、仿射及投影变换也有一定不变性。

通俗易懂的理解：假设你用地图定位自己的家。你记住了家周围的四个地标：超市、银行、公园和餐馆。如果你以后迷路了，只要能看到这几个地标，你就可以通过方向和距离再次找到家。SIFT中的描述子也是类似地通过周围几个方向的梯度信息“记住”了这个关键点的位置，使得在图片匹配中即使有轻微变化（如角度或亮度变化），它也能找到“家”——即该关键点的位置。

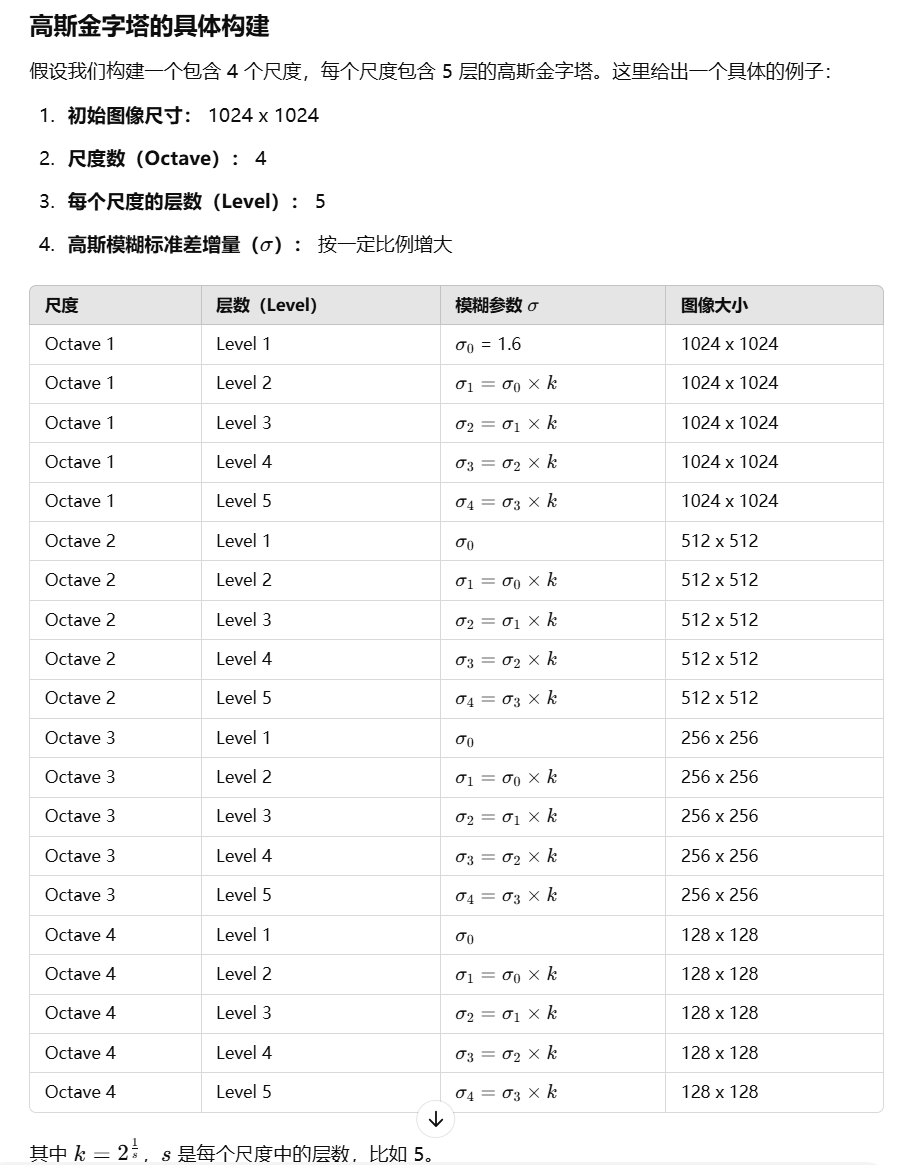
1. SIFT算法的实质：在不同尺度空间上查找特征点（关键点）的问题。



1. 关键点（特征点）：在不同尺度空间的图像下检测出的具有方向信息的局部极值点。比如角点、边缘点、暗区域的亮点以及亮区域的暗点。它们具有三个特征：尺度、方向、大小
2. 尺度空间：通过对原始图像进行尺度变换，获得图像多尺度下的尺度空间表示序列，实现边缘、角点检测和不同分辨率上的特征提取等。
3. 尺度空间中各尺度图像的模糊程度逐渐变大，能够模拟人在距离目标由近到远时目标在视网膜上的形成过程。尺度越大图像越模糊。
4. 采样之前要对图像进行高斯模糊处理，保证在采样图像中不会出现虚假的高频信息



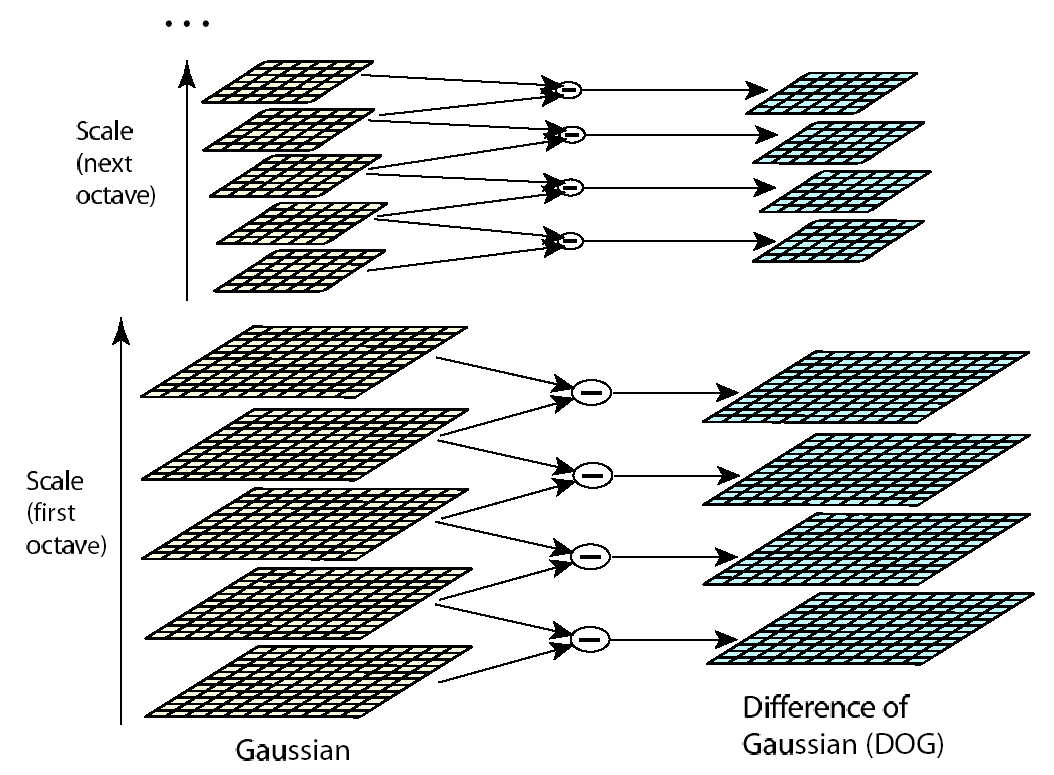
1. 高斯核大小：3σ之外可忽略，故大小为
2. 高斯模糊具有线性可分的性质，也可以在二维图像上对两个独立的一维空间分别进行计算。这样可以大大减少了运算的次数。
3. 使用半径分别为 6 和 8 的两次高斯模糊变换得到的效果等同于一次半径为 10 的高斯模糊效果。根据这个关系，使用多个连续较小的高斯模糊处理不会比单个高斯较大处理时间要少。
4. 高斯金字塔的例子（chat给出）：



相邻两组的同一层尺度为2倍的关系

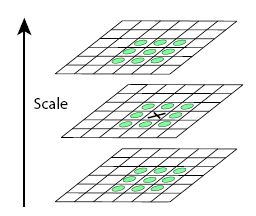
上一组图像的底层是由前一组图像的倒数第二层图像隔点采样生成的。这样可以保持尺度的连续性。

1. DOG算子（高斯差分算子）：



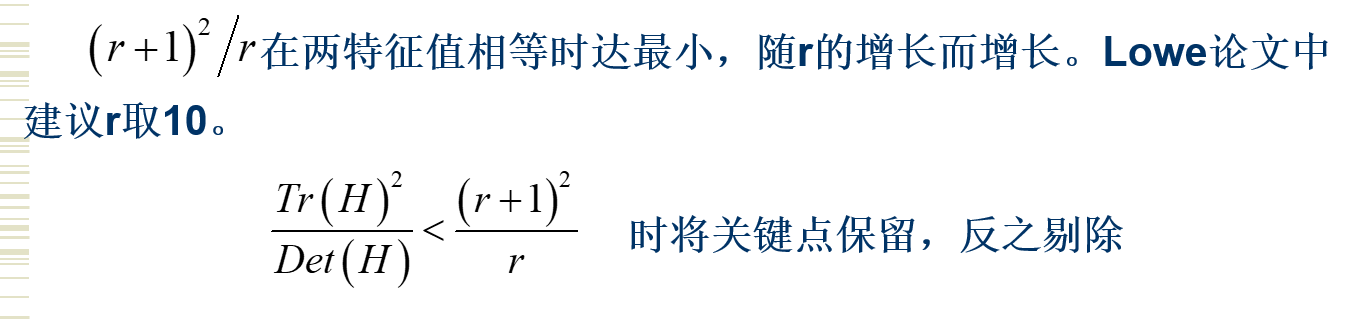
高斯差分算子提取强大边缘，既能度过下采样，个头又要大。

1. 关键点是由DOG空间的局部极值点组成的。为了寻找DoG函数的极值点，每一个像素点要和它所有的相邻点比较，看其是否比它的图像域和尺度域的相邻点大或者小。

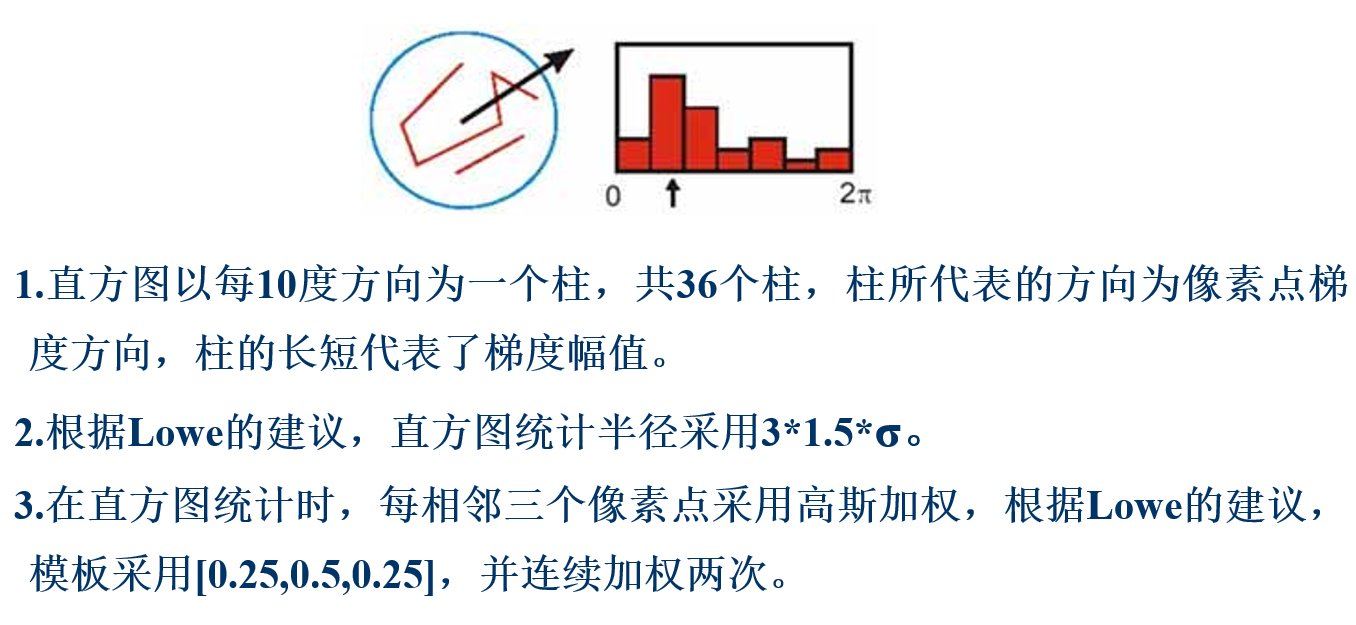


1. 关键点精确定位：去除对比度低的不稳定极值点
2. 去除边缘响应：





1. 关键点方向匹配：采用梯度直方图统计法，统计以关键点为原点，一定区域内的图像像素点对关键点方向生成所作的贡献。



关键点主方向：极值点周围区域梯度直方图的主峰值也是特征点方向

关键点辅方向：在梯度方向直方图中，当存在另一个相当于主峰值 80%能量的峰值时，则将这个方向认为是该关键点的辅方向。

1. 关键点描述：描述的目的是在关键点计算后，用一组向量将这个关键点描述出来，这个描述子不但包括关键点，也包括关键点周围对其有贡献的像素点。用来作为目标匹配的依据，也可使关键点具有更多的不变特性，如光照变化、3D视点变化等。
2. 关键点描述思路：通过对关键点周围图像区域分块，计算块内梯度直方图，生成具有独特性的向量，这个向量是该区域图像信息的一种抽象，具有唯一性。
3. 关键点描述过程：

统计领域内点的梯度，分块统计梯度直方图，采用4×4×8＝128维向量表征，综合效果最优（下图右为2×2×8＝32维）。



1. 描述子生成步骤：

* 确定区域，计算区域半径



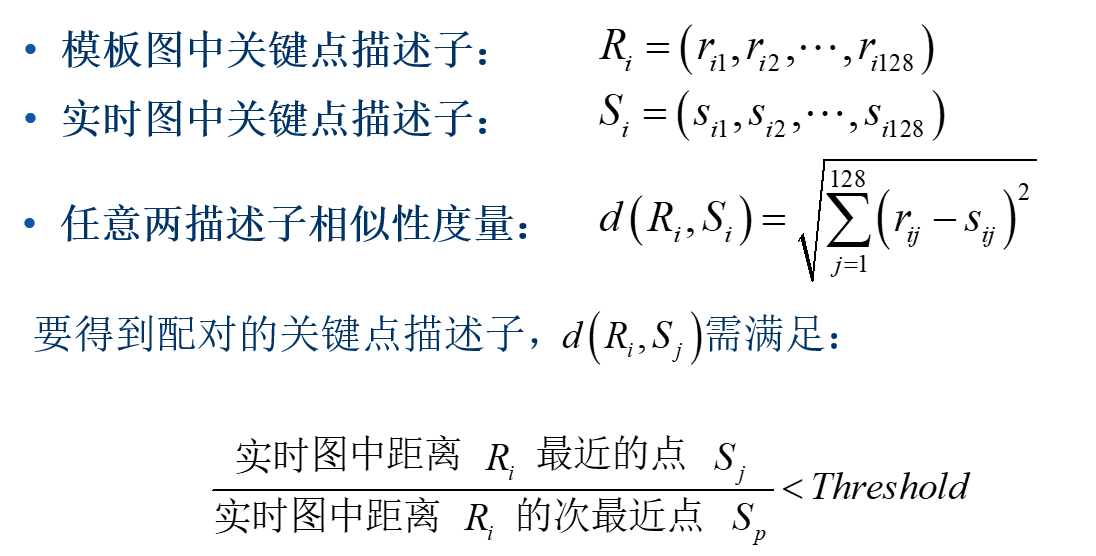
* 旋转坐标轴到关键点主方向上
* 对区域内每个像素点求其梯度幅值和方向，然后对每个梯度幅值乘以高斯权重参数，生成方向直方图。



* 在窗口宽度为2X2的区域内计算8个方向的梯度方向直方图，绘制每个梯度方向的累加值，即可形成一个种子点。然后再在下一个2X2的区域内进行直方图统计，形成下一个种子点，共生成16个种子点。
* 方向直方图每个方向上梯度幅值限制在一定门限值以下（门限一般取0.2），规范化（类似归一化）得到128维向量。



1. 规范化的目的：去除光照影响
2. 每一对关键点匹配过程（关键点之间进行两两匹配）：



1. 穷举太多？使用kd树
2. 消除错匹点：RANSAC，实质上就是一个反复测试、不断迭代的过程。设计目标函数，不断抽取，划分内点外点，使目标函数达到极值
3. SIFT应用：

物体识别

机器人定位与导航

图像拼接

三维建模

手势识别

视频跟踪

笔记鉴定

指纹与人脸识别

犯罪现场特征提取

1. SIFT缺点：实时性不高，慢，有时特征点较少，对边缘模糊的目标无法准确提取。
2. 改进：PCA-SIFT（降维，化简128维），CSIFT（针对彩色图像），SURF（计算量小，运算速度快），ASIFT（抗反射变换）。

