

机器视觉

题型分布：

分值 x 题数

单选 2x10

判断 1x10

简答 5x5

应用 10x3+15

注：应用题按步骤给分，尽量多写步骤，把步骤写清楚。

应用题拿分技巧： 如果不会可以写对于图像的预处理，比如去噪，对比度增强。（写清楚是怎么去噪怎么增强的）

绪论

机器视觉的概念： 机器视觉是用人的视觉感知和判断能力用摄像机和计算机来执行检测任务的替代品。它是图像的自动获取和分析，以获得控制或评估特定部分或活动所需的数据。

机器视觉的三个步骤（方面）：

图像处理：输入和输出都是图像（灰度变换）

图像分析：输入是图像输出是分析结果（图像特征），例如边缘，特征检测的特征

机器视觉：输入图像属性，输出对于机器的指令（动作指令）

图像获取

图像获取是从三维图像转换成二维图像的过程，在这个过程中会丢失**角度信息和距离信息**。

小孔成像：孔洞越小，图像越清晰

注意：小孔成像的孔并不是越小越好，当孔洞太小的时候会出现模糊

图像获取有关的信息：（摄影）

几个属性的概念

景深，焦距，光圈，f 指数，曝光，快门（之间的关系）

f 值和焦距不是一个东西，他代表的是焦距和有效通光孔径的比值。



■ **F-number:** focal length / aperture diameter

■ Aperture size is controlled by the F-number

f 值 = 焦距 (f) / 孔径直径 (D)（也就是光圈直径）

焦距和视场的关系，焦距越小，视场越大。视场是成像系统可以包含的场景范围。

景深，焦距，光圈：近大远小。 光圈直径、光圈大小、f 值

f 值和光圈成反比，光圈越大景深越浅（背景模糊），光圈越小景深越深（背景清晰）；

光圈与 f 值成反比；光圈越大，曝光度越大。

光圈越小，景深越大，进光量越小，成像时间越大，f 值越大，曝光度越小

光圈越大，景深越小，进光量越大，成像时间越小，f 值越小，曝光度越大

焦距越大，视角（视场）越来越小（窄），越远

焦距越小，视角（视场）越来越大（宽），越近 $F = \text{焦距} / \text{光圈直径}$

拍照近大远小的原理：（透视关系）

快门速度越快，曝光程度越强。



图像表示

图像数字化的过程：

抽样——决定图像的空间分辨率（空间分辨率和采样有关，依赖于光学传感器的采样点分布）是光信号转换成电信号的过程

量化——决定灰度分辨率，灰度级越多（量化级别越高，可以用于表示的灰度级越多）

图像原点：图像的左上角

灰度值：

我们把用来表示像素灰度取值范围的比例值称为动态显示范围（也就是灰度值的范围）；一幅图像的灰度值能够均匀的分布在其取值范围内（0~255），那么这幅图就成为高动态显示范围，说明这幅图的对比度高。

直方图均衡的目的就是让灰度值均匀地分布在 0-255 上，这种分布的图我们称之为高动态显示范围

$f(x, y)$ 表示的是像素点的坐标的灰度值。

N 比特图 $\rightarrow 2^N$

离散灰度级： $L=2$ 的 k 次方，2 的 8 次方有 256 灰度级，若 $k=8$ ，我们称之为 8 比特图。

存储空间： 一幅图像所能占的存储空间是和【它的量化的灰度级别】和【它抽样点数】有关，有多少个像素组成（宽*高/M * N），每颗像素（由于为 k 比特图）所占 k，所以为 $M * N * k$ (bits)，如果是彩色图像，要 x3，三个颜色通道

注意单位是比特

Digital image representation

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}$$

the range of values spanned by the gray scale is called the **dynamic range** of an image, and we refer to images whose gray levels span a significant portion of the gray scale as having a **high dynamic range**.

Discrete gray levels: $L = 2^k$

The number, b , of bits required to store a digitized image is

$$b = M \times N \times k$$

图像量化

灰度级越高，量化程度越高。

图像量化对低频区域更敏感，因为需要更多的灰度。

空间低频：低频指的是亮度或者颜色变化速度的快慢程度很低，变化缓慢，平滑。

图像保存格式

无损压缩（TIFF、BMP、GIF、PNG）

TIFF 文件大、浏览器不兼容

GIF 可变长编码，商标设计，透明通道，小的动画，支持透明度，支持动画模式。版本老、大可表示的颜色数较少，只有 256 种颜色；转换时会有噪声，用紫色点替代

BMP 位图 原图 无损压缩 1-24bit 全彩 windows 文件较大、浏览器不兼容

PNG 支持真彩色图像，24 位 RGB（24 位真彩色），支持透明通道；自然图像 通常不用来保存较大\细节较多的图像，比 jpg 大

*gif 不可以显示自然图像，png 不支持动画

有损压缩（JPG）

JPG（有损） 支持 24 位真彩色，自然图像 不适合保存线条，当保存图像为线性图像，在压缩时会导致图像失真（出现晕染现象）。 按照对象 8x8 进行压缩。

*如果压缩系数在变化，对于一张图压缩再打开，次数增多之后图像质量会越来越差。

*图像增强

灰度变换

灰度变换的目的：让图像的对比度更好，更清晰细节更明显。

灰度变换分为 3 种，图像翻转、对数(压缩)变换、指数变换(要分析 γ -gamma 的数值)。

对数变换和指数变换更常用。

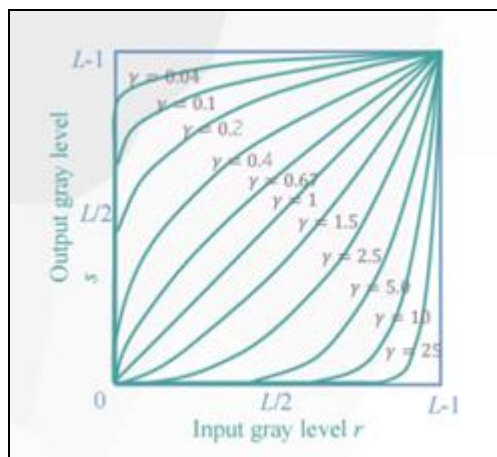
图像翻转：作用是获取亮区域中的细节。（如黑点）

对数变换：通常用来压缩灰度级跨度很高（几千、上万）的图像暗区域转细节

指数变换：对于常见的自然图像处理用的是指数变换，当 $\gamma < 1$ 时，暗区域拉伸，增强暗区域细节，压缩高灰度值（亮区域）； $\gamma > 1$ 时，亮区域拉伸，增强亮区域细节，压缩低灰度值（暗区域）。

分段线性变换：把图像里某一个区间的对比度拉伸，会把分段线性函数写出来，三段直线方程。比如：给两个点，求三段的直线方程。

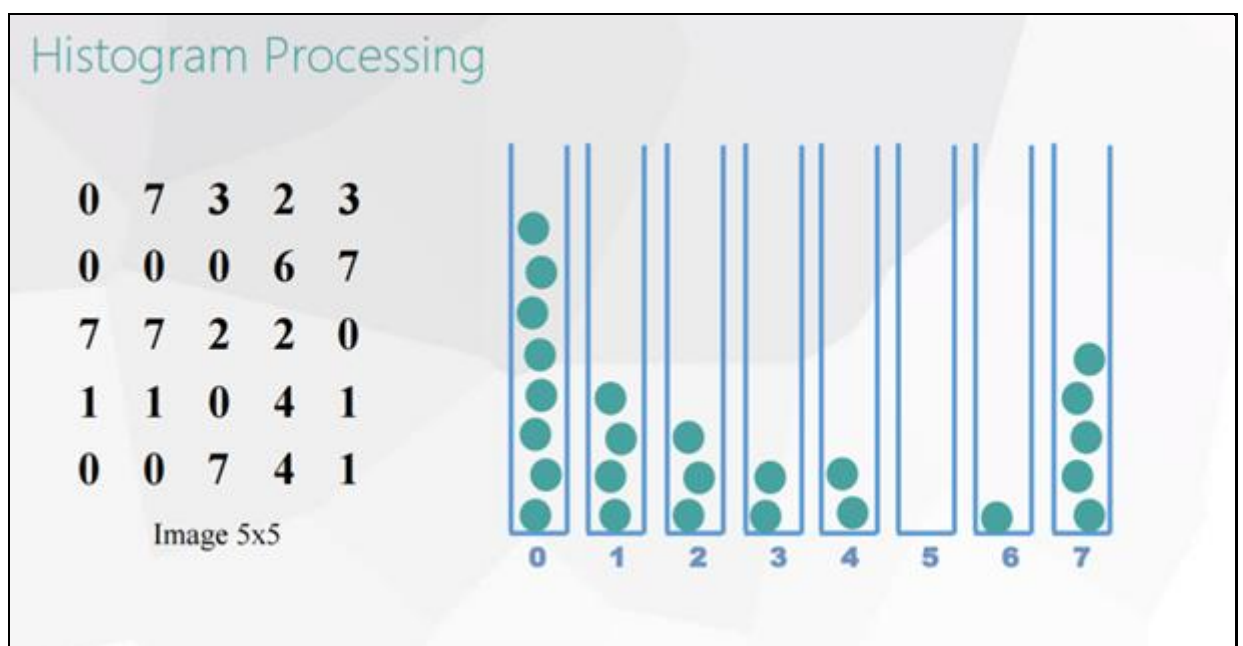
如果要在亮区域中找细节，可以用到的方法有：图像翻转，直方图均衡（自适应比全局更有效），同态滤波（有针对性）[同态滤波](#)。



直方图

如何画，把图像里面所有像素它的灰度值进行统计，某个灰度值的像素有多少颗

如果一幅图像过暗，它的直方图大概是靠在偏 0 这一侧；亮靠 255；灰度中等分布中间



直方图均衡

直方图均衡化：掌握 PPT 的例子，取整的时候四舍五入。（要学习计算）

全局均衡：可以使全局对比度最好（灰度级为 0~255），高动态显示范围。

自适应均衡：把图像分成很多小块，对于每一块进行全局均衡。怎么分？

注意计算：（计算均衡后的灰度值）

*对于灰度的范围是怎么确定的

■ Histogram equalization

• Example

Do histogram equalization on the 5x5 image with integer intensities in the range between one and eight

$$s_k = T(r_k) = (L-1) \sum_{j=0}^k p_r(r_j) = (L-1) \sum_{j=0}^k \frac{n_j}{n}$$

		PDF r	Intensity transformation function	Histogram value after HE			
0	7	3	2	3	$p(r_0) = 8/25 = 0.32$	$T(r_0) = 0.32*7$	$s_0 = 2.24 \rightarrow 2$
0	0	0	6	7	$p(r_1) = 4/25 = 0.16$	$T(r_1) = (0.32+0.16)*7=0.48*7$	$s_1 = 3.36 \rightarrow 3$
7	7	2	2	0	$p(r_2) = 3/25 = 0.12$	$T(r_2) = (0.32+0.16+0.12)*7=0.60*7$	$s_2 = 4.20 \rightarrow 4$
1	1	0	4	1	$p(r_3) = 2/25 = 0.08$	$T(r_3) = (0.32+0.16+0.12+0.08)*7=0.68*7$	$s_3 = 4.76 \rightarrow 5$
0	0	7	4	1	$p(r_4) = 2/25 = 0.08$	$T(r_4) = \dots = 0.76*7$	$s_4 = 5.32 \rightarrow 5$
					$p(r_5) = 0/25 = 0.00$	$T(r_5) = \dots = 0.76*7$	$s_5 = 5.32 \rightarrow 5$
					$p(r_6) = 1/25 = 0.04$	$T(r_6) = \dots = 0.80*7$	$s_6 = 5.60 \rightarrow 6$
					$p(r_7) = 5/25 = 0.20$	$T(r_7) = \dots = 1.00*7$	$s_7 = 7.00 \rightarrow 7$

杭州電子

概率密度函数 r，变换函数，直方图均衡之后

图像算数和逻辑

算数包括加减法，逻辑包括与或非。

更倾向于逻辑操作：因为在计算过程中位操作的速度快于加减法操作。

图像减法用原图减去一幅图，通常用在造影（凸显血管）、追踪检测（车、行人）、检测缺陷。图像减法的目的是什么，得到的结果是什么样的，起到了什么样的作用

图像帧平均算法

将很多在同一个地点连续拍摄的图像求平均，目的是去除 0 均值的高斯噪声。

通过平均 n 张图将方差变为原本 1/n （也是夜间拍摄的原理）

这里去除椒盐噪声会产生色差吗？是什么原因？

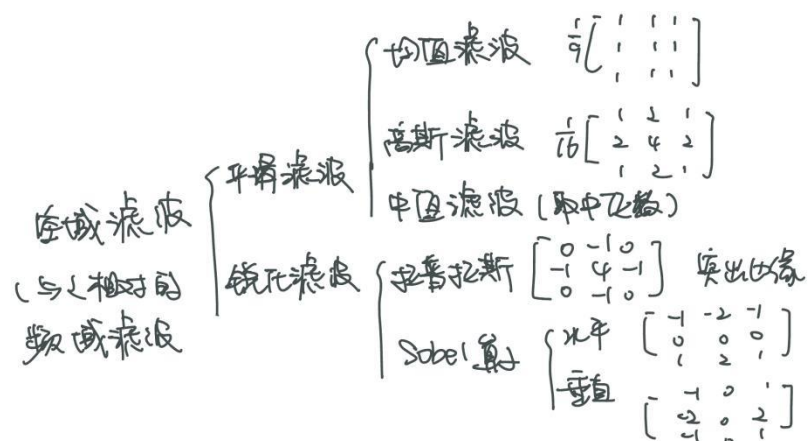
高斯噪声：均匀的雪花感，所有像素都受到影响。均值滤波、高斯滤波适用

椒盐噪声：黑白点，部分像素受到影响。中值滤波、自适应中值滤波、形态学滤波适用。

要说明为什么能去除？学会原理。

滤波

什么是空域滤波



空域滤波是图像处理中直接在像素空间域(即图像平面)上对像素邻域进行局部运算的技术。

滤波器

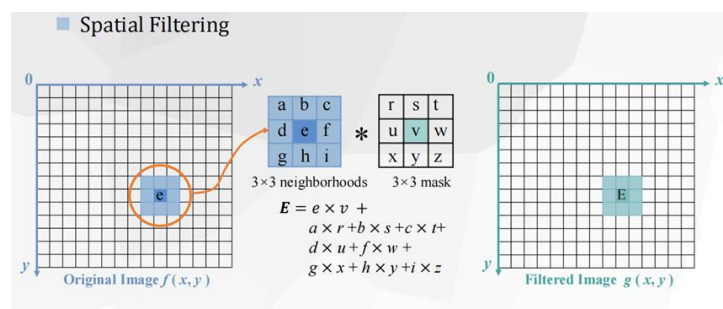
滤波器分两种：线性(均值)滤波器和非线性(基于统计的)中值滤波器。

均值滤波

均值滤波：

均值滤波和高斯滤波处理高斯噪声，中值滤波处理椒盐噪声效果更显著。不过用中值滤波处理彩色图像可能会出现图像颜色变浅的现象。(为什么)

模版覆盖，求均值或中值的计算过程

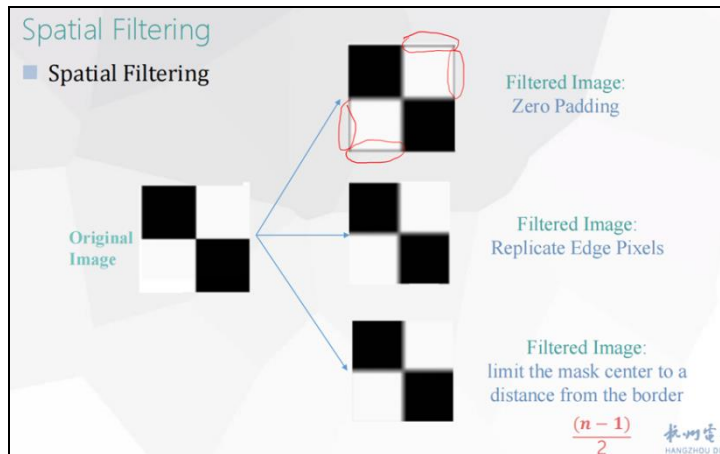


滤波时的图像边缘：

图像边缘补零、复制边缘像素、限制模板中心点到边缘的距离(边缘一圈像素不处理)

(从上到下的图示↓)

零填充会有一圈黑边(填充0为黑色)



均值滤波器， 3×3 ，系数和为 1，所以每个系数为 $1/9$ 。如果不强调，直说的 $n \times n$ ，直接按均值。

均值滤波器——图像平滑

滤波器大小对滤波的影响 均值、中值滤波，滤波模板越大，图像越模糊（中值滤波也会模糊吗）

均值滤波器和高斯滤波器只是将噪声扩散出去，进行了平均，没有去除噪声

均值滤波器模版越大，效果越模糊，模版越小越清晰。这个方法只是把强的噪声扩散，没有去除噪声。

高斯滤波则是按权重对噪声进行扩散，本质上和均值滤波器很像。而中值滤波器也可以减小噪声

非线性滤波器

最大，最小，模板怎么动，怎么算

差分滤波器：sobel 算子，不是线性滤波器。 sobel 算子的适用范围 (垂直方向和水平方向的梯度)

*梯度有水平梯度和垂直梯度，合在一起才是最后的梯度。

锐化：二阶导数——拉普拉斯算子；一阶导数——梯度算子 也就是把模糊的部分减去 拉普拉斯算子得到的是边缘图像，要进行锐化还需要+或者-去原图

滤波——图像处理

图像平滑，图像模糊——均值滤波器（高斯滤波器是和均值滤波器类似，更改了权重）

图像锐化：

1. 用拉普拉斯算子做锐化，但是拉普拉斯算子只得到了边缘轮廓，需要根据使用模版的中心系数在原图上+或-边缘轮廓的图像
2. 二阶导数——拉普拉斯算子；一阶导数——梯度算子 也就是把模糊的部分减去

统计滤波器（中值滤波器） 通过排序滤波，对噪声不敏感，主要通过排序实现，对椒盐噪声（黑白点噪声）最有效。

均值滤波器和拉普拉斯对噪声敏感，所以用拉普拉斯提取边缘需要提前进行预处理，进行图像平滑和去噪。

高增益滤波:

原图 \times 系数 k -原图作均值滤波的结果

(原图作均值滤波的结果也就是模糊, 原图-模糊=锐化)

钝化掩膜, 高频强调, 从原图里减去模糊的图像, 经过均值滤波, 用模糊的图像就可以得到清晰的图像, 但这样减会造成图像整体灰度值变暗, 所以我们用高频强调(对原图先乘 A 系数, 将原图的灰度值提高), 再减去模糊的图, 得到锐化后的图像

● High Frequency Emphasis

Using frequency domain terminology, the **unsharp masking** is a highpass filtered image

$$f_{hp}(x, y) = f(x, y) - f_{lp}(x, y)$$

and the **high-boost filtering** is

$$\begin{aligned} f_{hb}(x, y) &= Af(x, y) - f_{lp}(x, y) \\ &= (A-1)f(x, y) + f_{hp}(x, y) \end{aligned}$$

In the frequency domain, these become

$$H_{hp}(u, v) = 1 - H_{lp}(u, v)$$


and

$$H_{hb}(u, v) = (A-1) + H_{hp}(u, v)$$

Sometimes it becomes

$$H_{hfc}(u, v) = a + bH_{hp}(u, v)$$

where $a \geq 0$ (0.25 ~ 0.5) and $b > a$ (1.5 ~ 2.0).



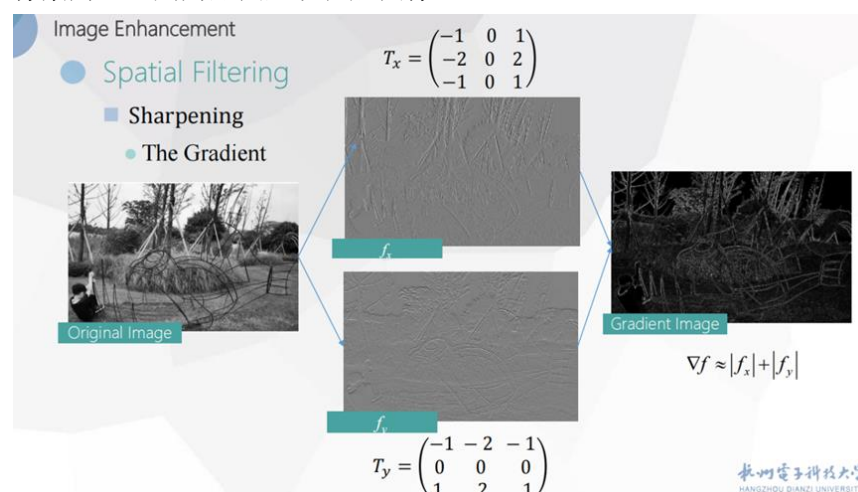
频域滤波

频域滤波——平滑、锐化、带通(阻)、同态滤波

频率增强, 梯度算子, 一阶二阶算子。若中心像素为负, 则为原图减锐化后的图像, 若中心像素为正, 则为原图加锐化后图像。

Image Enhancement

- Spatial Filtering
 - Sharpening
 - The Gradient



垂直边缘、水平边缘和算子的对应。如何看?

得到的两张图分别是频谱图和相位图。

频谱图中心点: 直流增量(能量占比最高的点)

相位图: 空间位置信息也就是结构特点(决定了图像的内容)

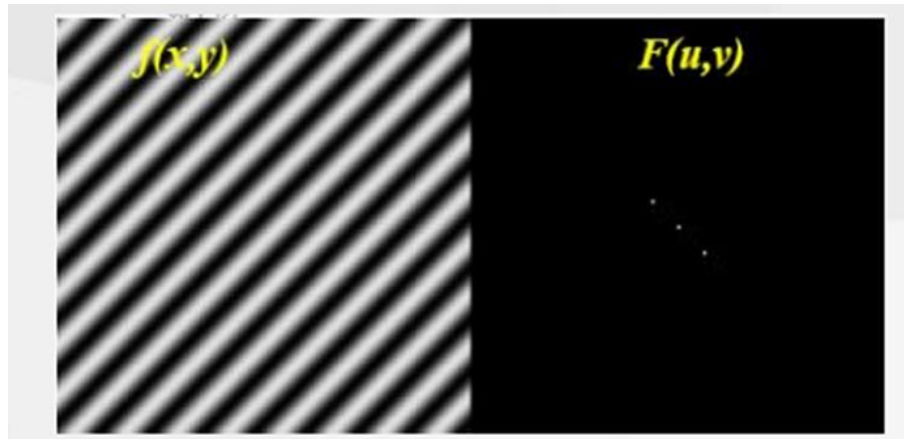
频谱图: 表示图像能量

什么是能量环, 怎么显示

能量环中心点占了 92.7%

频域滤波，将图像进行二维傅里叶变换转到频域再处理，转换的时候做了位移，把四个亮点放到中间，中心点称为直流分量，它占据的相当于是整幅图像的平均灰度值，中心点能量很高；如果图像水平方向的线较多，那么在频域会出现三个点，这三个点一定是垂直的（垂直关系）

左图：空域 右图：频域



频域滤波所用到的滤波器

平滑滤波：理想滤波器（低通）、巴特沃思滤波器（高通）、高斯滤波器（带通和带阻）

巴特沃斯滤波器：无阶数限制，阶数 n 越大，越接近理想滤波器，越容易出现振铃现象；

高斯滤波器：安全，不会出现振铃现象，效果相对较好

带通滤波器，带阻滤波器（处理摩尔噪声，有规律变化的噪声）

理想滤波器会出现振铃（涟漪）现象，原理：

理想滤波器在频域上是一个完美的阶跃，转到空域后是一个 sinc ，所以用理想滤波器对图像处理会出现振铃现象，频域相乘，空域卷积，图像的每个点都是脉冲，脉冲和 sinc 函数卷积，每个点都会有 sinc 函数，叠加后导致振铃现象

处理摩尔噪声（为什么会产生摩尔噪声）

具体过程：

1. 傅里叶变换
2. 平移四个点到中心
3. 带阻滤波
4. 对中心点平移再反变换

为什么要把四个点移到中心：这样中心点所对应的就是直流分量了，中心点包含大部分信息。

带通带阻滤波，对于加性噪声条纹或是有规律变化的摩尔噪声可以用带阻滤波找噪声点，去除后转回空域

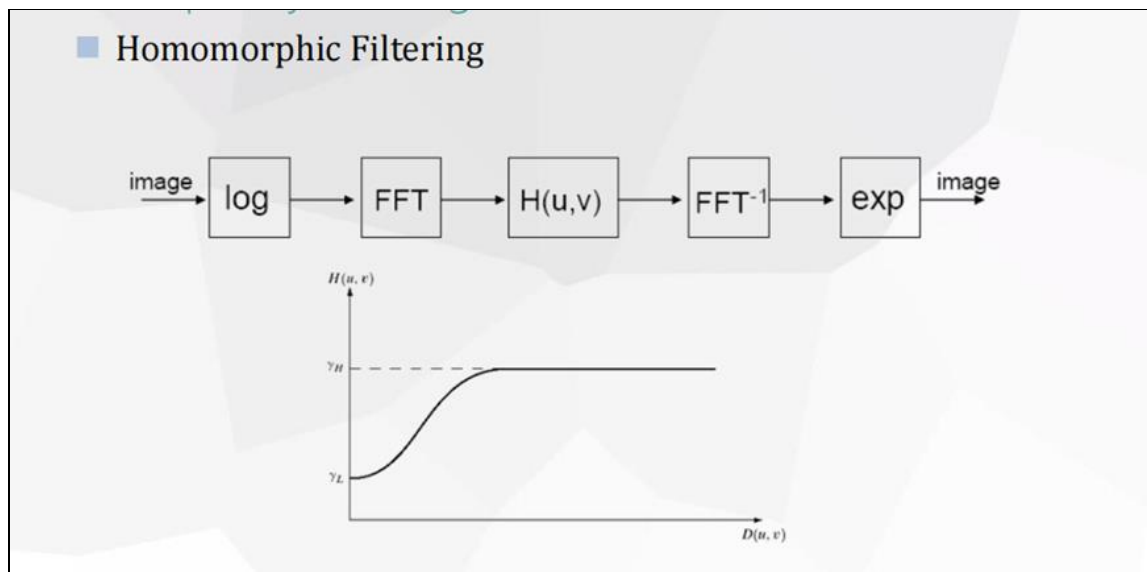
同态滤波

同态滤波，乘型滤波，先做对数变换，将乘性噪声变为加性噪声，再做傅里叶变换，利用高通滤波器（光照低频），再傅里叶反变换。再进行指数变换。

（原理）

为什么要进行同态滤波：

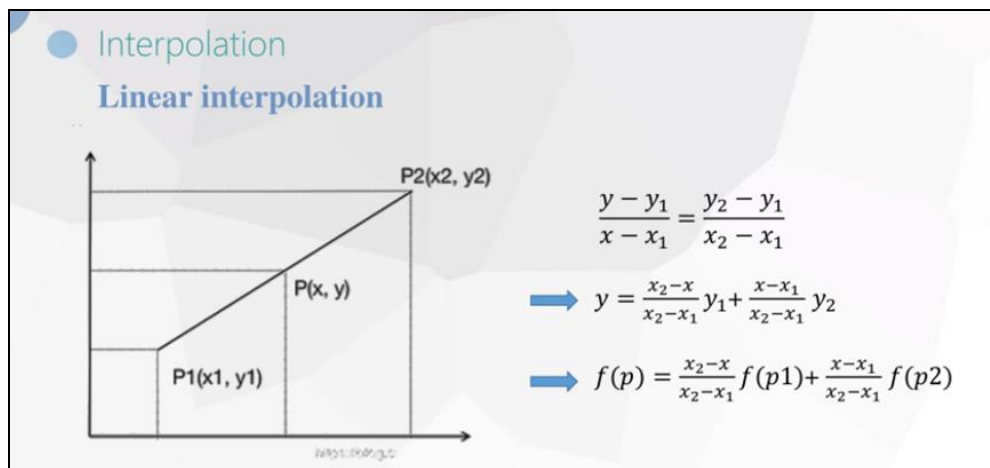
对于反射光去噪，要把乘积转换成加减，此过程需要进行对数变换

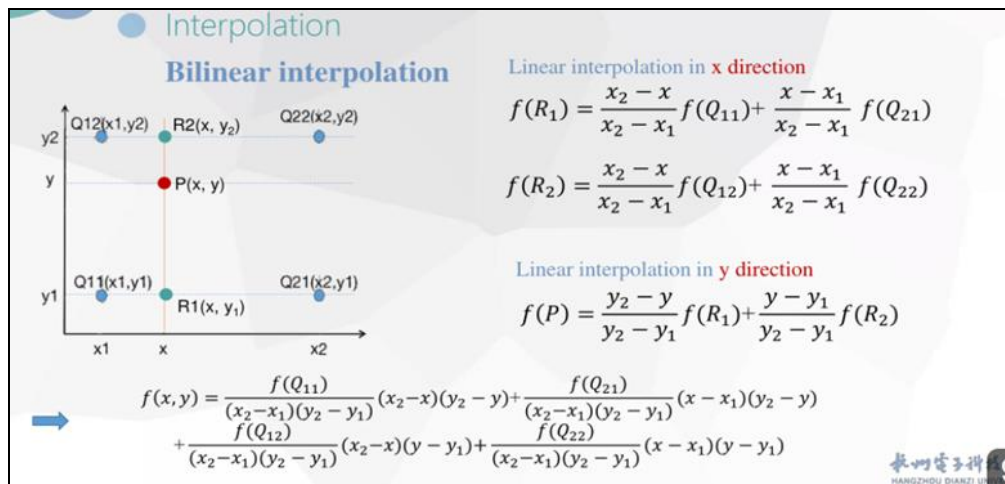


几何变换

插值——最近邻插值、线性插值、双线性插值

插值怎么实现。





最近邻插值：在最近邻插值中，对于要放大的新图像中的每个像素，都会在原始图像中找到最近的像素，并将其值用于新图像。具体来说，对于新图像的每个像素，找到其在原始图像中最接近的位置，然后将该位置的像素值赋给新图像中的对应位置。

Pros: 计算简单，速度快

Cons: 可能会导致放大后的图像出现锯齿状的边缘

双线性插值：双线性插值在最近邻插值的基础上更进一步，它考虑了相邻像素之间的权重关系。具体而言，对于新图像中的每个像素，首先确定它在原始图像中的四个相邻像素，然后根据距离和这四个像素之间的关系来计算新像素的值。这四个相邻像素的权重是根据距离和位置关系来确定的，**距离越近的像素权重越高，距离越远的像素权重越低。**

Pros: 通常比较更平滑，更接近真实图像。

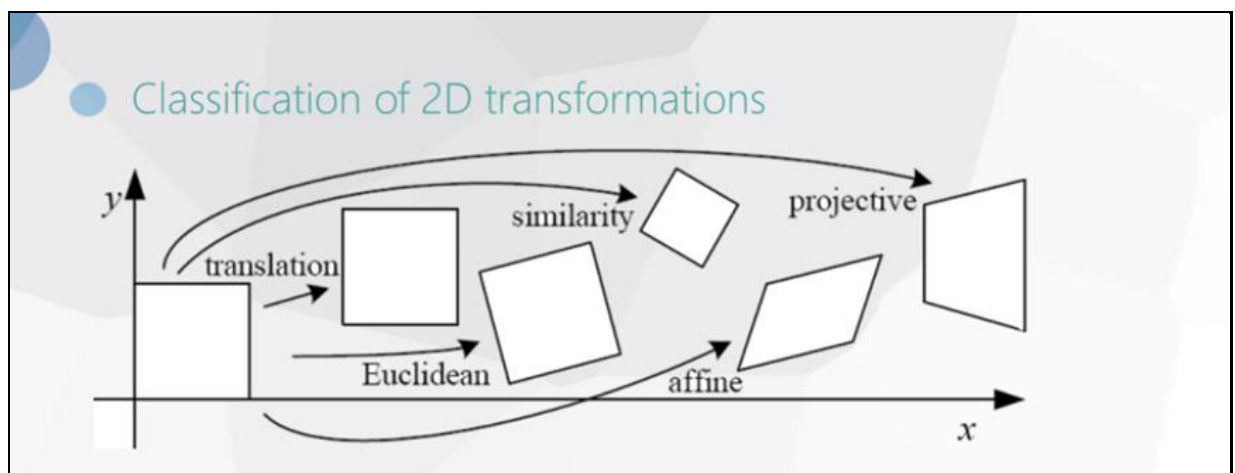
Cons: 计算量稍大。

变换——平移、旋转、缩放、仿射变换、投影变换

平移、旋转变换（欧几里得、刚体变换）、缩放（相似性变换）；仿射变换；投影变换

仿射变换包括平移，旋转，缩放这些基本几何变换

投影涉及畸变



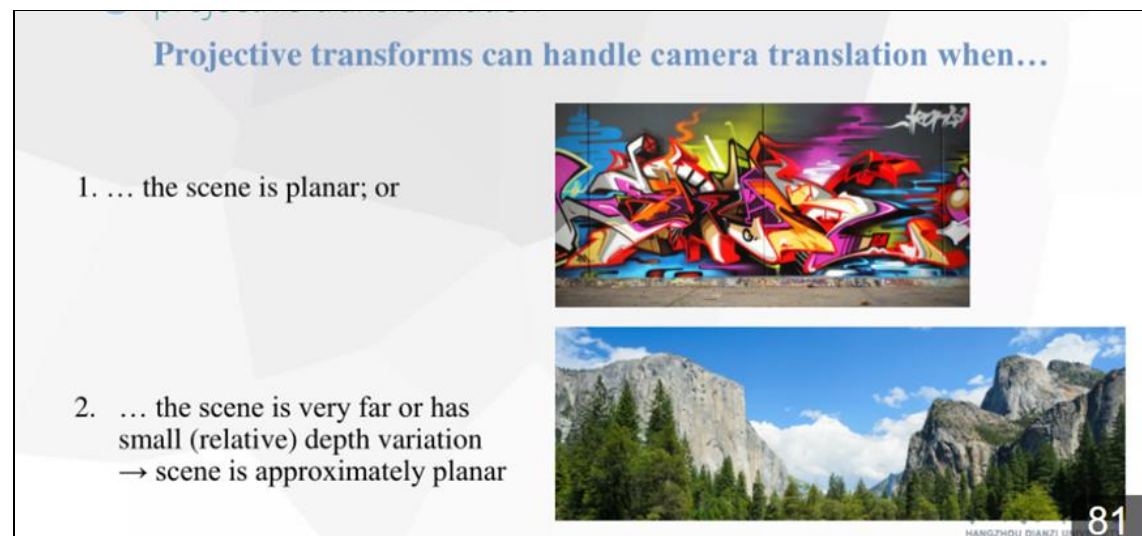
仿射：除了平移，都可以用 2×2 的矩阵表示 任何几何变换都可以用矩阵 x 的形式表示

平移可以通过 3×3 的矩阵表示，所以为了方便会用 3×3 的矩阵表示仿射变换
投影：实现投影变换需要四组点（每组两个点）（变形前和变形后）
怎么用矩阵表示平移，缩放，旋转

投影变换在笛卡尔坐标系二维的基础上加了第三位 z ，用来控制物距，加了 w 项 (x/w y/w)
在做投影变换时要将两幅图片接成一幅图片，要将两幅图片投影到同一个平面上，要求拍摄场景要么是一个平面，或者这个场景可能是凹凸不平的，但是凹凸度远远小于我们的物距，就是摄像头到场景的距离，这个情况下也能用投影变换。

投影和仿射变换的区别 仿射：除了平移，都可以用 2×2 的矩阵表示 任何几何变换都可以用矩阵 x 的形式表示

插值：谁离得近谁用处大（用距离作为系数）



彩色图像

彩色基础

视锥体：比较少，感受颜色和细节，晚上不工作，叫昼视觉

视感体：多，感受大的轮廓，晚上工作的，叫夜视觉

颜色特性：

Hue 色调：组成这种颜色的所有光波里占主要地位的光波决定了色调 0-255

Saturation 饱和度：向纯色里面加入白光的稀释程度叫饱和度；

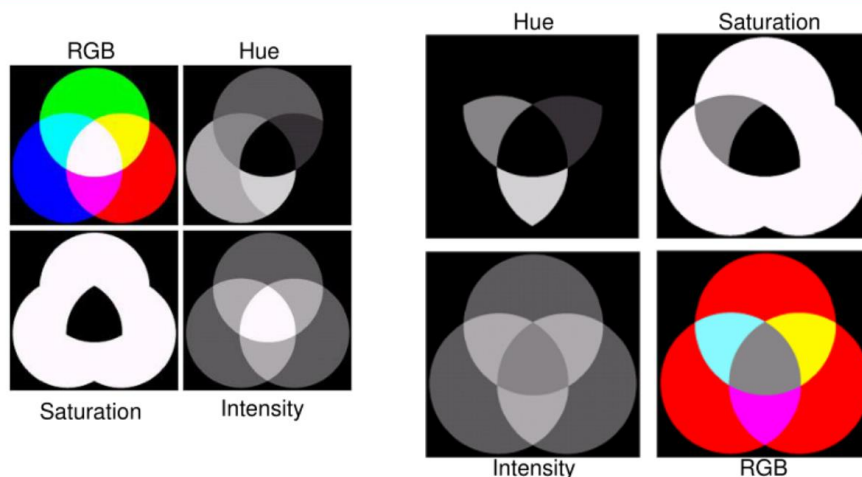
亮度：明暗程度

色调和饱和度合在一起称为 Intensity 色度，色度决定物体的颜色。

Intensity 强度、亮度，明暗程度。0-1

RGB, HSI, CMY 之间的区别，ppt 图片。

Manipulating Images In The HSI Model

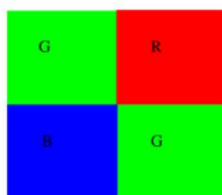


These result illustrate clearly the power of the HSI color model in allowing *independent* control over hue, saturation, and intensity, quantities with which we are quite familiar when describing colors.

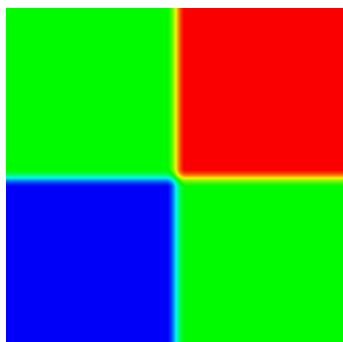


题：考察下面的一幅500x500的彩色图像，其中的方块部分分别为纯的红、绿和蓝色。

1. 如果将此图像转换到HSI空间，对H分量图像用一个25x25的算术平均掩模进行处理，再转换回到RGB空间，得到的结果将是怎么样的？
2. 采取上面同样的步骤，只是这次处理的是S分量，结果又会怎样？



对 H 分量：



对 S 分量：

没有变化

原理：H 分量调节的是图像的色调，对色调进行平滑导致了混色。

彩色模型——HSI

RGB 模型主要用在显示器和摄影机上

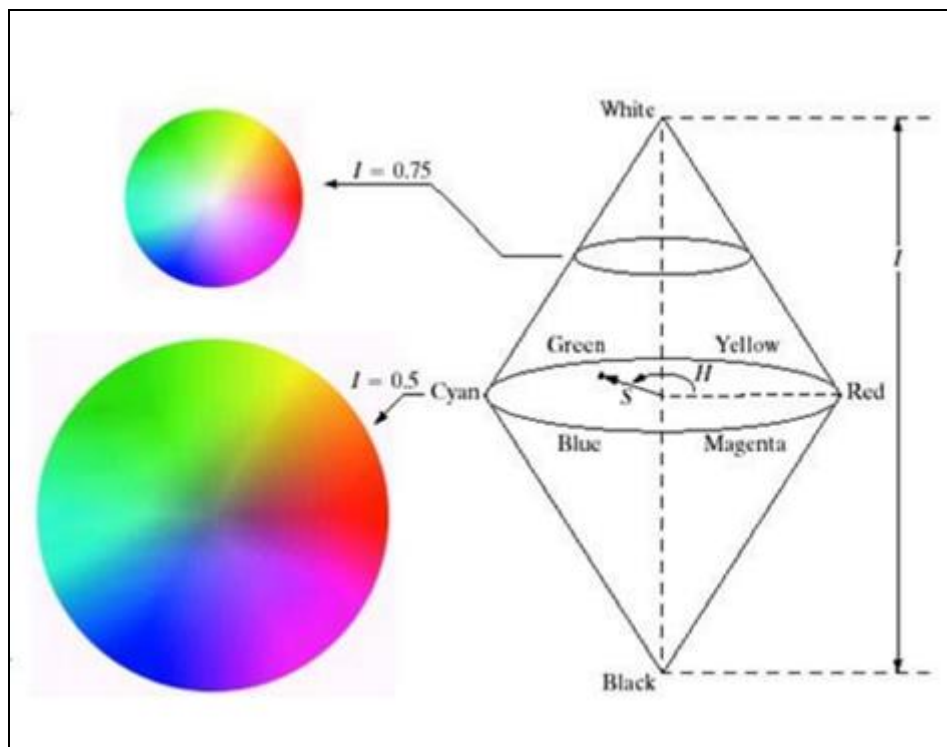
CMY 用在打印机和扫描仪；

HSI 主要是用来描述色彩的；

补色：红 青；绿 品红；蓝 黄

HSI 模型用锥形来表示，锥形的中心轴是 I 强度，底部是黑，顶部是白，0-360，表示灰度值；锥体截面圆的半径表示了饱和度，越靠近中心饱和度越低，接近 0，靠近边缘接近 1；色调从 0 度开始，旋转 360 度再回来，**0 度和 360 度都是红色**

因此 RGB 和 HSI 模式不是一一对应的，所以他们之间的转换属于不可逆转换，但是可以直接通过相应公式进行转换。



灰度(I 强度)=(R+B+G)/3

hsi 三个分量是什么：

h：色调（主要波长）

s：饱和度（加入白光的程度）

i：强度（明暗变化）（前两个决定色彩）

调节亮度，HSI 只需调节 I，RGB 里对三个分量调节，调节要一致，否则颜色会发生改变，因为调节的比例不一样。

伪色彩：对于灰度图像根据灰度值进行分层

彩色分割

取一个区域分别求 RGB 的均值和方差，用来实现分割的球体或是立方体，取一个中心点，看其距离

色彩分割步骤：

1. 提取目标区域，对目标区域的色彩作均值和方差
2. 以均值为中心点，方差为半径/边长，确定分割
3. 球体内的色彩分割为目标，球体外的为背景

色彩变换：（怎么调整亮度）

hsi: 调节 i 分量

rgb: 三个通道同时调整

噪声在 rgb，转换成 hsi 之后

除了 i 里的噪声会减小，hs 中的噪声会增大

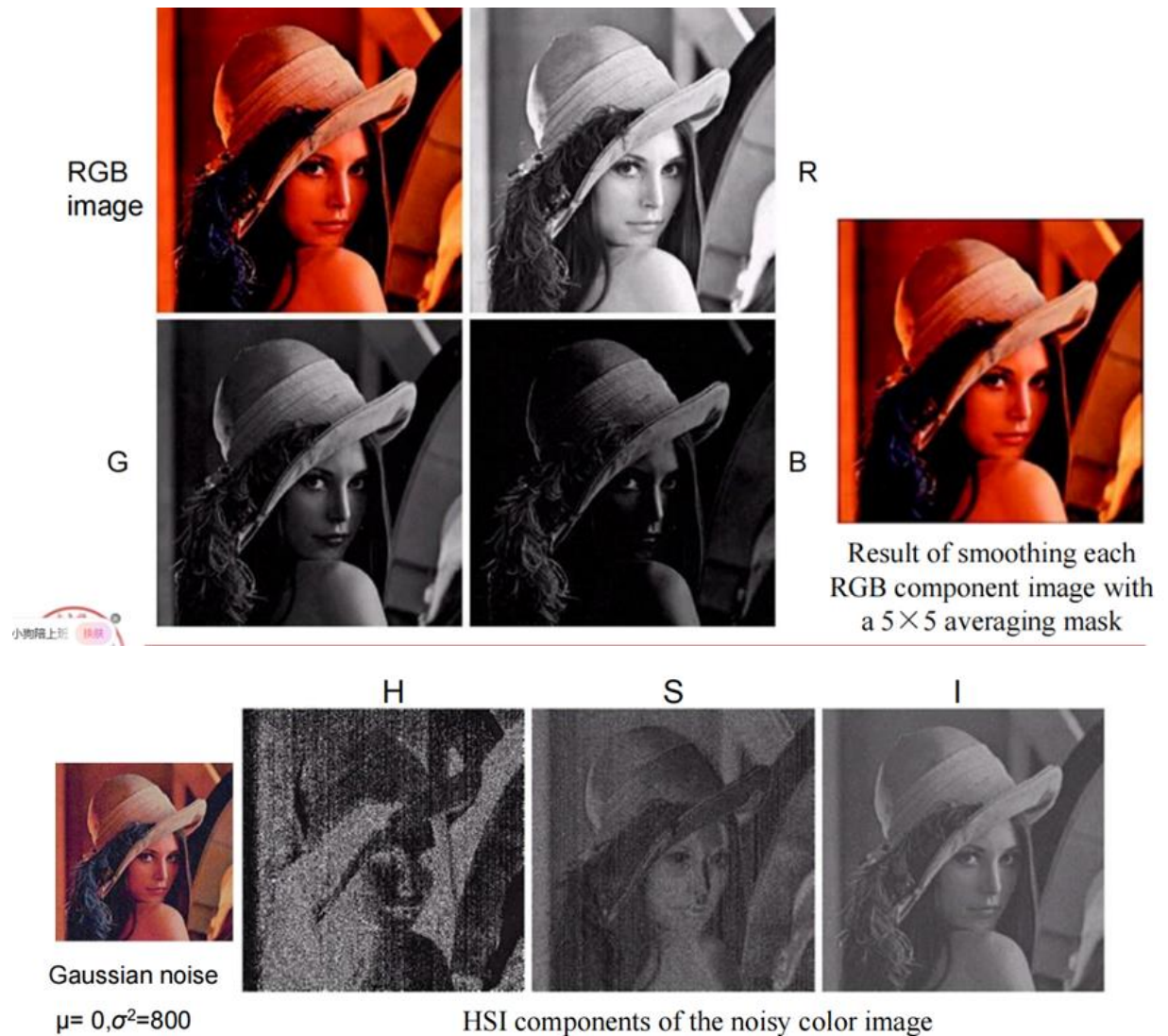
为什么有噪声的 RGB 图像转变成 HSI 图像会出现这样的问题

HSI 的转换设计非线性数学运算，其中对于 HS（色调和饱和度）的计算更明显，其中使用的是一种三角函数计算，会导致放大噪声扰动。而 I 分量进行的是线性平均操作，平均可以减少噪声的方差。

1.2 彩色平滑

我们在做彩色平滑时，这个噪声转换到 RGB 三个通道均存在，然后我要去噪声，对三个通道均作（高斯），再转回来作 RGB 显示，但是这个时候处理后的图颜色会变，将有噪声的 RGB 转到 HSI，噪声怎么变，如果要去噪（对 HSI 三个通道均去噪）再转回 RGB 显示，图像会怎么样，为什么会出现这样的结果。

1. 对于图像在每个通道单独平滑，破坏了 RGB 的比例关系，原始色比和滤波后色比不同。红色物体偏粉红，绿色物体偏黄绿
2. 仅对 I 滤波，颜色保留完好，亮度进行平滑。如果对三个通道滤波，会导致色差和饱和度失真。
3. HSI 滤波的原理影响：H 通道滤波导致颜色混合错误，扭曲光谱分布。S 通道会导致色彩浑浊，高饱和度边缘被模糊。



- ❑ The hue and saturation image are significantly degraded by noise due to the nonlinearity of **cos** and **min** operations.
- ❑ The intensity component is smoother than any of the three noisy RGB component images because it is the average of the three components.

翻译:

由于 **cos** 和最小操作的非线性，色调和饱和图像显著降低。H、S 会有较大的变化。强度分量比三个噪声 RGB 分量图像中的任何一个都更平滑，因为它是三个分量的平均值。

H 色调变化失真、S 饱和度变化——颜色失真

像素深度：在 RGB 空间里用来表示每一个像素所用到的比特数，全彩色是 24

全彩色图像：有 24 位，RGB 各有 8 位的，0 到 255 取值。

图像分割

图像分割，检测水平边缘，垂直边缘应该用什么算子，通常用梯度算子检测出来的线是否就是边缘(不是，只是一些离散的点，我们还需要边缘连接，判断梯度大小和方向，把梯度值

接近和梯度方向接近的点连接在一起才是边缘)

边缘检测

用 sobel 算子、canny 算子，得到完整边缘，用霍夫变换用来检测直线和圆，二次自动阈值分割，基于最大类间方差，把背景和前景分为两类，求其最大方差，取到的这个值就是最佳阈值

sobel 算子

不仅可以检测水平和垂直上的边缘，还可以检测 45 度的边缘

因为 sobel 和拉普拉斯算子对噪声敏感的原因，需要先对图像作平滑操作来减小噪声造成的影响。

霍夫变换（圆和直线）

霍夫变换不能直接使用在原图上，需要先对原图提取点，或者进行二值化。[为什么会导致计算不准确以及非常大程度影响运算和检测效率](#)

检测圆需要已知半径，通过调整阈值，只有部分圆也可以检测。

阈值分割（原理，对于背景和目标求方差）

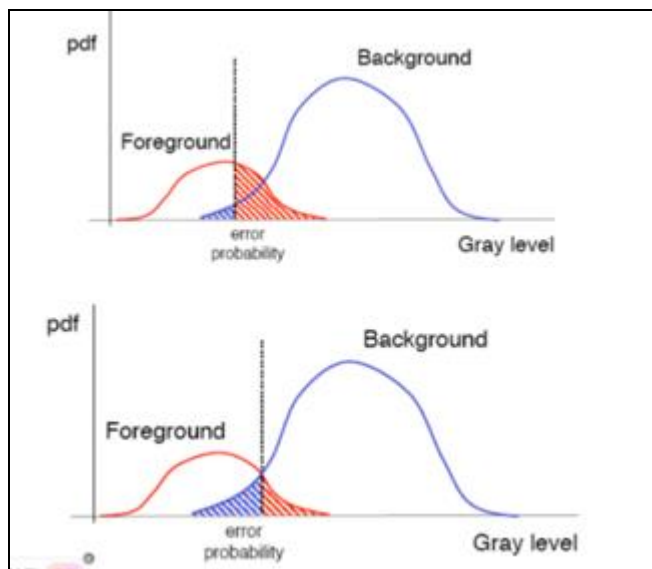
奥克斯自动阈值分割

一种自适应阈值分割，把图片对象分割成小块，对于灰度值进行遍历，再进行阈值分割（[怎么操作的，算法原理是什么？](#)）

迭代法：

先给一个阈值，不断去算更多阈值

高斯滤波（高斯模糊）+ 梯度计算 + 非极大值抑制 + 双阈值检测 + 边缘连接



形态学

1.1.1 开闭操作（腐蚀膨胀）

膨胀腐蚀原理（分别是什么）

膨胀算子怎么用，怎么写，结果怎么画

腐蚀：去除小的区域想，狭长链接断开

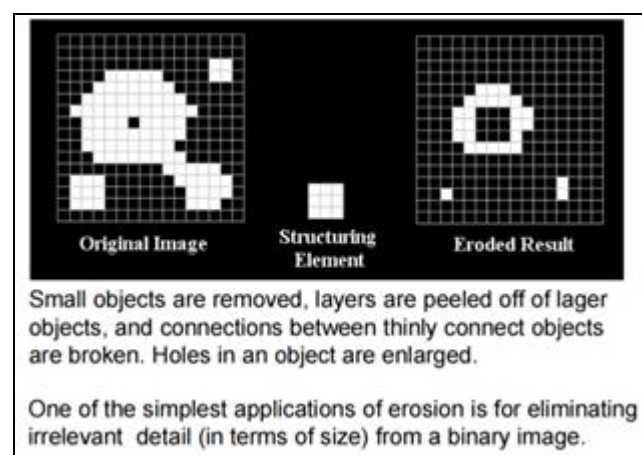
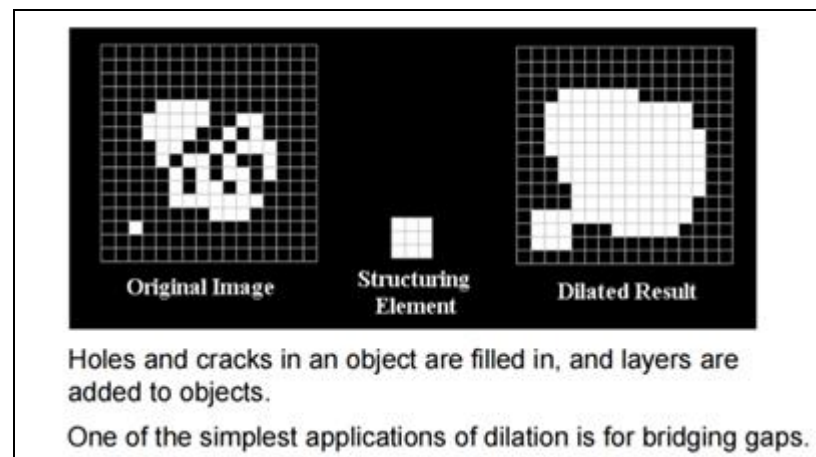
膨胀：填充空洞，进行连接

用开闭操作实现梯度求取

二值图的腐蚀膨胀的过程，基于腐蚀膨胀的开闭操作；

开操作：先腐蚀后膨胀，断开狭长的连接，去掉比算子小的物体；

闭操作：先膨胀后腐蚀，连接比较窄的断开的地方，填充孔洞，膨胀腐蚀



灰度图的腐蚀膨胀：

膨胀：算子和图像和的最大值

腐蚀：算子和图像差的最小值

对灰度图膨胀，整幅图的灰度值提高，图像变亮；

对灰度图腐蚀，整幅图的灰度值减小，图像变暗；

膨胀去除小的暗细节，腐蚀去除小的亮细节，开闭操作对应

（开操作先腐蚀后膨胀，所以去除亮细节）开=腐蚀+膨胀（变暗+亮），除亮细节

（闭操作先膨胀后腐蚀，所以去除暗细节）闭=膨胀+腐蚀（变亮+暗），除暗细节

开闭操作在二值图上，无论你对一幅图像做多少次开操作，图像不变

形态学滤波，灰度求取

1.1.2 形态学应用

边缘提取、区域填充、求凸集、细化、粗化、连通区域

图像分割

特征提取

分为，自然特征，人为特征

自然特征：眼睛看到的，包括颜色、纹理、形状、边缘；

人为特征：计算获取，矩、直方图、SIFT、HOG

一个好的图像特征，应该具有：（旋转，缩放，平移）几何不变性和光度不变性。SIFT 有这样的特性。

特征特性：好的特征具有局部性，抗遮挡，几何不变光照不变曝光不变性

局部二值模式（LBP 特征）：对单调灰度变换具有不变性，但是对噪声比较敏感，受噪声影响，同时也受光照影响。

1.1 HOG 算子

计算过程，求梯度值、梯度方向

● HOG (Histogram of Oriented Gradient, 方向梯度直方图)

- HOG特征通过计算和统计图像局部区域的梯度方向直方图来构成特征。HOG特征结合SVM分类器已经被广泛应用于图像识别中，尤其在行人检测中获得了极大的成功。

- **主要思想：**在一幅图像中，局部目标的表象和形状能够被梯度或边缘的方向密度分布很好地描述。（本质：梯度的统计信息，而梯度主要存在于边缘的地方）。
- **实现方法：**首先将图像分成小的连通区域，我们把它叫细胞单元（cell）。然后采集细胞单元中各像素点的梯度的或边缘的方向直方图。最后把这些直方图组合起来就可以构成特征描述器。



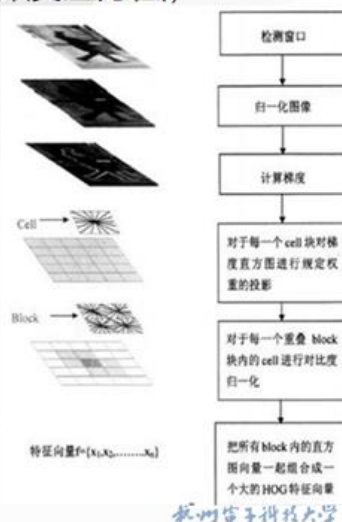
杭州电子科技大学
HANGZHOU DIANZI UNIVERSITY

● HOG (Histogram of Oriented Gradient, 方向梯度直方图)

■ HOG特征提取实现过程

HOG特征提取方法就是将一个图片（要检测的目标或者扫描窗口）：

- 1) 灰度化（将图像看做一个x, y, z（灰度）的三维图像），归一化；
- 2) 计算图像每个像素的梯度（包括大小和方向）；
- 3) 将图像划分成小cells（例如6x6像素/cell）；
- 4) 统计每个cell的梯度直方图（不同梯度的个数），即可形成每个cell的描述特征；
- 5) 将每几个cell组成一个block（例如3x3个cell/block），一个block内所有cell的特征串联起来便得到该block的HOG特征。
- 6) 将图像内的所有block的HOG特征串联起来就可以得到该image（要检测的目标）的HOG特征了。这个就是最终的可供分类使用的特征向量了。



杭州电子科技大学

- 1、图像预处理：灰度化+归一化
- 2、计算梯度：计算每个像素的梯度（大小和方向）
- 3、将图像划分成小的 cells（如 6*6 像素/cell）
- 4、统计每个 cell 的方向梯度直方图（不同梯度个数），形成描述特征。
- 5、将每几个 cell 组成一个 block（如 33cell/block），串联 cell 特征形成 block 的 HOG 特征。
- 6、将所有 block 的 HOG 特征串联起来得到 image（要检测的目标）的 HOG 特征，即最终的可供分类使用的特征向量。

1.2 SIFT

高斯分解，分解成不同的分辨率的图像，再求高斯差，求大致过程

● SIFT (Scale-invariant feature transform) 描述子

■ 在SIFT 算法的步骤

- (1) 构造高斯差分尺度空间
- (2) 检测尺度空间极值点
- (3) 精确定位极值点，去除不好的特征点
- (4) 给特征点赋值一个128维的方向参数。每个关键点都包含三个信息：位置、尺度和方向。
- (5) 关键点描述子的生成

- 1、构造高斯差分尺度空间
- 2、检测尺度空间极值点
- 3、精确定位极值点，去除不好的特征点
- 4、给特征点赋值一个 128 维的方向参数，每个关键点包含：位置、尺度、方向
- 5、关键点描述子的生成

附加

立体视觉

深度与视差的关系：反比——深度越深视差越小

模版匹配

匹配方式： 模版匹配 直方图匹配（求直方图，看直方图相似度） 特征匹配（计算特征求特征之间的距离） HOG、SIFT 模板匹配（模板和图像卷积或相关，相关度最高）

习题部分

伽马变换

伽马变换通常用于进行对比度增强、丰富画面细节。

是通过对图像的像素值进行幂次运算实现。

具体过程如下： 1.灰度级归一化：首先，将图像的灰度级从[0, 255]的范围归一化到[0, 1]之间，这是因为在进行 Gamma 变换时，通常要对图像进行归一化处理。

2.Gamma 变换：对归一化后的像素值进行 Gamma 变换。其中需要设定 gamma 的数值， $\text{gamma}>1$ ，整体变暗， $\text{gamma}<1$ ，整体变亮。

3.像素值重新缩放（直方图均衡）：将变换后的像素值重新缩放到[0, 255]的范围内，以便于显示和处理。

需要确定一个合适的 gamma 值，中间需要调试。

去噪处理（高斯噪声）

1. 对于拥有方差，均值为 0 的高斯噪声进行去噪处理，需要多幅图像，用到了高斯滤波器或者均值滤波器进行去噪。

2. 对于每个图像该位置的像素进行算术平均

3. 得到目标图像，导出

去噪处理（椒盐噪声）

1. 确定使用中值滤波器进行去噪处理，设定滤波器大小

2. 对目标图像进行滤波操作（使用复制边界值的方式进行边界处理）

直方图均衡

对于直方图计算其累计分布函数，然后将其归一化到 0-255 的灰度范围，在原始图像和新的图像进行映射，得到的新的图像

锐化处理

1. 对图像使用拉普拉斯算子提取边界

2. 根据使用的算子和得到的边界图像，与原图进行相加或者相减

3. 对于处理得到的图进行灰度变换，直方图均衡，增强对比度

去除摩尔噪声（频域滤波）

1. 对原图作傅里叶变换

2. 将四周的点移到中心点（直流分量）

3. 对于频域图作阻滤波

4. 把中心点移回原位，进行反傅里叶变换得到原图

同态滤波

1. 对原图作傅里叶变换，并移动点（同上）
2. 对频域图作高通滤波，去除低频反射光
3. 把中心点移回原位
4. 得到原图

彩色图像分割

1. 对于目标区域的颜色进行方差和均值的计算
2. 根据均值和方差确定中心点和半径，以此作圆确定范围
3. 对于目标范围进行均值滤波或者高斯滤波

霍夫变换

1. 对图像进行平滑去噪处理，降低噪声对算子的影响
2. 使用 sobel 算子或者拉普拉斯算子对于原图进行边缘提取
3. 转为二值图，或者取点
4. 使用霍夫变换提取目标直线或者圆
5. 调整直线或者圆的相对参数

击中击不中问题

1. 对原图进行锐化处理（如上）
2. 设置目标结构
3. 设置需要排除的目标结构（目标结构外轮廓）
4. 当满足目标结构且不存在需要排除的外轮廓时才判定击中

插值问题

双线性插值结果：

算法原理：在最近邻插值的基础上考虑相邻像素的权重关系，非单一寻找最接近的像素，而是首先确定其中（在原始图像）中的四个相邻像素，然后根据距离和四个相邻像素来计算新像素的数值。这四个像素的数值是由距离关系和位置关系来确定的。

在最近邻插值中，对于目标图像的每一个像素，都会在原始图像找到最接近的像素，并将其运用于新图像。新图像的每个像素都会去找到原始图像中的最接近位置的像素。计算简单速度较快。但是可能会出现边缘锯齿。

灰度形态学

特征提取匹配——HOG 和 SIFT