

## UNIT 1 Introduction

好了，我们正式开始。

既然这门课叫数字图像处理，那么你知道什么是数字图像吗？

你可能不太知道，告诉你，数字图像就是连续物体表面的离散数字表示。

那么何为离散？和连续的区别是啥？（可以联想一下我们概率论中涉及到的离散型随机变量和连续型随机变量）

---

数字图像的基本元素是啥？如何表示？

那么它的本质是啥？如果你不清楚，请利用python IDE(spyder、pycharm等)读取图片观察！

---

数字图像可分为彩色图、灰度图和二值图，那么它们有啥区别与联系？

二值图的“二”怎么理解？你要回答不出来那就坐实了“二”！

再者，为何大多数图像的预处理都要转为灰度图？

---

数字图像处理有哪三个level？

①**Low-level processing**: Perform various processing on the image to improve the effect of the image and prepare for image analysis

②**Intermediate processing(Image analysis)**: Detect and measure the ROI in the image

③**High-level processing(Image understanding)**: Study the nature and interrelationship of objects in the image

以上是图像处理的三大层次，如有翻译欠佳的地方，不吝赐教。

---

## 图像信息的数字化

数字化的本质是将连续的信号转为数字形式）

采样：图像空间坐标的数字化，说得再具体一点就是将一副图像分割成 $m*n$ 个网格，每一个网格都代表这一位置的亮度值，可以用 $f(x,y)$ 来表示。（连续→离散）

从视觉效果角度来说，采样越多，也就是分割成的网格数越多，图像分辨率越高。如果你还理解不了，想想，如果给你的自拍照用 $2 \times 2$ 的网络来表示和用 $2000 \times 2000$ 的网格来表示，是不是前者模糊得连你妈也认不出？

量化：灰度值的数字化，也就是说将每一个像素的灰度值转换为离散的整数值。量化级数可以是2，那么这幅图像就成了二值图；可以是256，也就是我们通常看到的数字图像。

从视觉效果角度来说，量化的越细致，也就是级数越大，图像的表现更加丰富（细节）

说了这么多，请你来说说我们的初心——数字化的原因？

---

## 图像文件大小

如何计算？（分辨率  $\times$  位数）

计算机存储信息的最小单位是啥？（学计算机的你肯定知道）

通常情况下，计算机用哪个单位表示占用空间大小？

那么两者之间有啥关系？

从理论上讲，各大字节，比如兆字节与千字节之间的比例关系是啥？如果应用到工业领域呢？

---

## 影响清晰度的因素

亮度、对比度、尺寸大小、饱和度、细微层次。

说说什么是对比度、饱和度。

尺寸大小和细微层次分别由啥决定？

---

## 邻域

如果某一个像素点用 $f(x,y)$ 表示的话，那么它的四邻域和八邻域如何表示？

你的理解是有问题的，而且你也没发现问题

---

## UNIT 2 灰度转换与空间域滤波

简述图像退化的原因，从对比度、噪声以及清晰度三个角度来说明。

所以为了处理图像，我们首先需要进行图像增强。

图像增强是为了使得这幅图像或者说是某一区域适用于特定应用。

它具有主观性，并不是客观意义上的增强图像，一切都依靠人的主观感觉来评估。

---

那么图像增强有哪些手段？

### （一）空间域处理

法一，灰度变换。

①可以直接调用函数（线性函数（图像反转）、对数函数(傅里叶)、幂律函数、非线性），也可以自己设计一个合适的函数来进行灰度转换。

②直方图处理（直方图均衡化啥的，过程是啥？计算会了吗？）

法二，空间滤波，也称之为邻域处理或空间卷积或模板处理。

①平滑、锐化滤波器

②自己设计一个模板来增强图像

---

那么问题来了，请接我夺命连环三问：

①平滑和锐化的区别是啥？从视觉上来看是咋样的？

②说说平滑滤波有哪些？锐化滤波呢？

③卷积是什么？怎么计算？卷积中的步长、大小是什么意思？如何设计卷积核或者说卷积核的特点是啥？

---

补充一下，你知道噪声到底是个啥玩意么？这里我们拿椒盐噪声来说明。

原图：

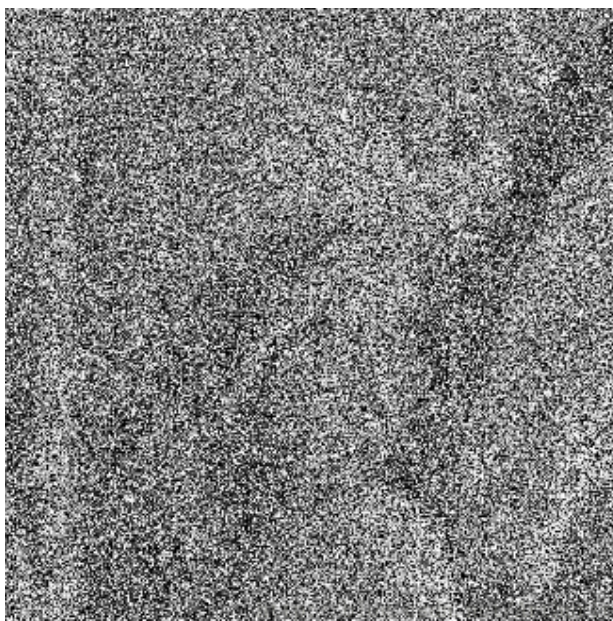


带有椒盐噪声的图：





如果你还看不出来，那么我要放大招了！！！！



所以啊，不难看出，噪声就是一个个小点点，这些点点呢会影响视觉美感，当然我不否定艺术家在创作时特意加入这些点点来体现个性。但是在图像处理领域，我们称之为影响美观的噪声。噪声其实是图像中灰度值突变的地方，你可以想想山峰的样子。那么图像增强就需要利用平滑滤波去除噪声，也就是将这些山峰给铲平，看上去光滑一点。

---

## （二）频率域处理

法一，傅里叶变换（转换），卷积定理

法二，自己设计合适模板

这一部分内容在下面我会详细解说原理以及方法

对了顺便问一下，傅里叶变换时为啥要使用对数变换？

因为傅里叶频谱的像素值动态范围很大，一些低值的细节容易丢失，所以需要取对数变换缩小范围。

---

## 图像反转

作用？

适用范围：当需要增强嵌入到图像暗色区域的白色或者灰色细节时可以进行反转。（暗色占主导地位）

一般医学领域应用比较多。

---

## 直方图

灰度直方图描述的是啥？（不同灰度级像素出现的次数）

如何计算均衡化直方图？（计算分析）

算法步骤：

- ①计算每一个灰度级像素出现的个数，然后归一化，即用频率表示；
- ②直方图的累积更新；
- ③根据原归一化直方图寻找最佳匹配的灰度级；
- ④根据映射关系获得新的灰度级像素个数
- ⑤计算新的直方图

从视觉效果来看，直方图均衡化的增加了图像的对比度。

---

## 空间滤波

掩模运算的数学含义是啥？

### （一）平滑空间滤波

作用：模糊处理、减小噪声

分类：

①线性：均值滤波器

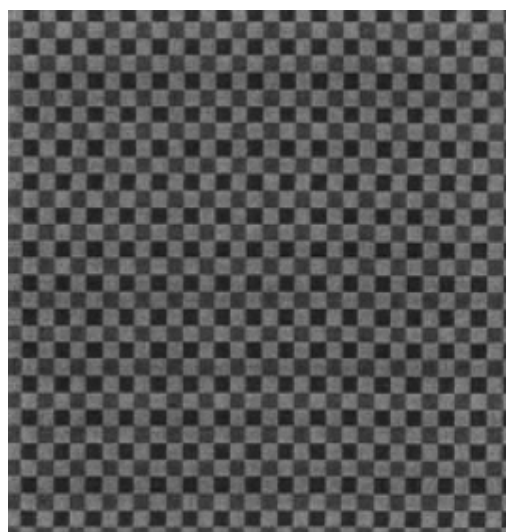
②非线性：最大值、最小值、中值滤波器

**cores:** 在邻域内按像素大小排序，由排序结果决定用某个值代替中心像素值，所以对于最大值滤波器而言，可以寻找图像中的最亮点；执行次数越多，图像越亮，但存在阈值，原因不必解释过多。

中值滤波的优点：

①可有效保留锐度和图像细节；

②有效去除脉冲噪声：黑白叠叠加混合而成的啦，别告诉我你有密集恐惧症。



## （二）锐化滤波

原理：空间微分（微分可以增强图像边缘，消除灰度变化缓慢的部分）

作用：突出图像中的细节，增强被模糊了的细节

分类：

### 1.一阶微分：梯度算子

哪一种算子对噪声较为敏感？哪一种算子作用效果正好与之相反？

### 2.二阶微分：拉普拉斯

应用范围：①图像检测或分割前的边缘提取；②可以改善超声探测图像的边缘模糊；③目标识别与定位

---

## UNIT3 频率域滤波

图像的空间域是啥？（二维平面坐标系）空间域上的幅值呢？

何为频率域？如何将空间域的东西转化到频率域中？（我上面已经提到了，如果你说不出来，你找我！）

请简述频率域研究图像的意义。

说了这么多频率域，那频率是啥？

又不知道了吧，听我说，频率就是对图像中灰度变换剧烈程度的一个衡量指标。所以说对于一些图像中突变的部分，比如上面讲的噪声，对应的频率值就比较高，而其他部分由于灰度变换比较平缓，所以频率值低。（所以对一副频谱图，低频和高频代表啥含义你心中也有答案了吧！如果没有的话，呜呜，难道是我解释得还不够透彻嘛，那就再把上面划线的部分看一遍。）

---



## 图像变换

(一) 目的:

①化繁为简;

②利于特征提取 (该部分等到设计卷积神经网络的时候我会和你们详细讲解)

(二) 图像变换通常是一种正交变换。有啥特点?

能量集中分布在低频率部分, 而边缘、线状信息分布在高频率部分, 利于图像处理。

可以理解吗? 不理解没关系。关于低频和高频到底是啥么个玩意我下面会详细讲解。

(三) 应用:

图像增强、恢复、特征提取、形状分析、图像压缩编码

直观的分析一个变换的方法就是计算图像的频谱。

---

关于傅里叶中的两个特殊名词有必要了解

幅度谱、频谱图:  $|F(u, v)| = [R^2(u, v) + I^2(u, v)]^{\frac{1}{2}}$

相位谱:  $\arctan[\frac{I(u, v)}{R(u, v)}]$

从梯度的角度来讲, 前者代表梯度幅值, 后者代表梯度的方向角。(梯度是啥? 说白了就是一阶导, 它具有两层含义, 一层是值的大小, 另一层是方向)

---

你以为了解了上面两个特殊名词就够了吗? 小了小了!

①周期性: 实际上对**DFT**来说, 不论其空间域还是频率域, 都是沿着X和Y方向无限延伸, 所以我们通过python调包显示的频谱图只是冰山一角啦。

②频谱的中心化: 经过变换后的图四周为低频, 中间为高频。所以通过某一个函数, python中是ffshift中心化, 将低频的部分移到中心 (至于为啥这么干, 可能是一种习惯吧 QAQ)

③高低频率的分布：中间最亮的是低频，属于直流(DC)分量，越往边走，频率越高。（这是对于中心化过后的频谱图哦，别搞混了！记得数形结合，我就不放图了！）

④能量分布：中心能量最多，依次递减

⑤纵横交错性：比如横向的变化会体现在频谱的纵向上

⑥平移和旋转：前者不会影响频谱图，但后者会影响，频谱图也会相应旋转对应的角度。（但平移是会改变相位图的）

---

## 如何从傅里叶的角度来理解低通滤波器？

低通滤波器，顾名思义，截止高频，只让低频的部分通过。在傅里叶变化中，对于一副频谱图，其低频区域对应着图像中变换不明显的部分，高频对应着图像中的一些边缘或者说之前提到的噪声部分。那么将这些频率都通过低通滤波器时，只留下了低频，也就是图像中变换不明显的区域，而高频部分（边缘和噪声）会被舍弃，那么从视觉上看就会变得模糊，当然还会出现振铃现象。（这一现象限于本人的知识很难给大家讲解，感兴趣的话自行查阅资料）

---

## 如何设计一个理想低通滤波器？

可以思考一下阈值化的过程，设置一个阈值  $T$ ，高于  $T$  的部分置为255，反之为0。那么设计该滤波器时就需要给定一个截止频率  $D$ ，用于区分“你所认为”的低频与高频。如果大于  $D$ ，就无法通过，反之则使其通过，最后留下的便是低频部分。

$D(u, v)$  是该  $(u, v)$  点到原点的距离。

那么，那么，那么天音波在滑行的途中呢，摸眼靠近他.....停停停！想啥呢？回来回来！那么如果我们设置的截止频率  $D_0$  较小的话，从视觉上来讲，图像有啥变化嘛？（如果你答不上来，那真的要好好体会我上面说的话了）

---

## 频域增强原理

通过改变不同频率的分量实现

方法：

①图像转换，比如傅里叶变换

②在频域空间增强图像，将其与频率滤波器相乘

③图像逆转换，从频域空间转换到图像空间，比如傅里叶逆变换

频率滤波：低通、高通、带通和带阻滤波，同态滤波

所以说到这里，低通滤波、高通滤波、平滑以及锐化滤波有啥区别和联系？

空域和频域密不可分哦！

---

## UNIT 4 图像退化与复原

### 图像退化

含义：图像质量下降

形式：模糊、失真、有噪声（说白了不就图像不清晰了嘛）

原因：①相机未聚焦；②传感器噪声；③自然因素

模型：

$$g(x, y) = f(x, y) * h(x, y) + n(x, y)$$

其中 $n(x, y)$ 代表噪声

模型灰常简单，那么问题来了，你知道这是哪一个域的表达嘛？

---

### 图像复原

也称之为图像恢复，不用纠结。

含义：利用先验知识来重建或恢复退化的图像

本质：可以理解为图像退化的逆过程，对退化过程估计并建立数学模型，然后反过来补偿图像。

## 图像恢复和增强的区别

同为改善图像质量，前者是客观过程，后者基于人为主观判断。前者使用反卷积，后者使用卷积。（同为空间或频域滤波）

---

## 图像恢复存在的技术难题

- ①无法得知退化的过程，因此存在盲目复原的现象（盲复原）
  - ②退化后的图像由于存在一些噪声与模糊导致复原的不确定性与challenge
- 

## 图像复原的方法

### （一）逆滤波

简单说一下，这种滤波呢在没有噪声的情况下效果较好，如果有噪声，那就可以选择下一种复原方法。

### （二）维纳滤波

- ①使用条件：图像信号可看成平稳随机过程
- ②原理：最小均方误差
- ③描述过程：假设现在有一个噪声图像 $y=x+\delta$ ，我们需要设计一个滤波器 $h$ ，使输出的 $y=s$ 的这幅图新尽可能接近 $x$ ，这只是定性分析。那么从定量分析来看，就是使 $s$ （估计图像）与 $x$ （原噪声图）的均方误差最小：

$$\min : e^2 = E[x - s]^2$$

因此，对于维纳滤波，我们也可以称之为最小均方差滤波、最小二乘或最小平放滤波器。

---

## UNIT 5 彩色图像处理

---

说说有哪些可见光？它们的频率和波长有啥关系？

如果你连这都想不起来，那你可以和高中物理老师叙叙旧了。

---

### 三原色

我想要的是RGB三原色，也就是光学领域的，你不会不知道吧。

反正我不管.....

简单说一下颜色的混合效果， $R+B=M$ ， $G+B=C$ ， $R+G=Y$ （ $R+G+B=W$ ）

（如果理解起来有些抽象，我可以用ps给你演示，你愿意听吗？）

---

### 颜色的三特性

色调、饱和度、亮度

说明一下，色调和波长有关。

彩色=色调+饱和度

---

### 图像矩阵的取值范围

这和图像的类型有关，一般都是八位图，即uint8，所以取值范围在 $[0, 2^8]$ 之间。

像素的深度：即每一个像素的比特数，如果是八位图，那就是8bit。

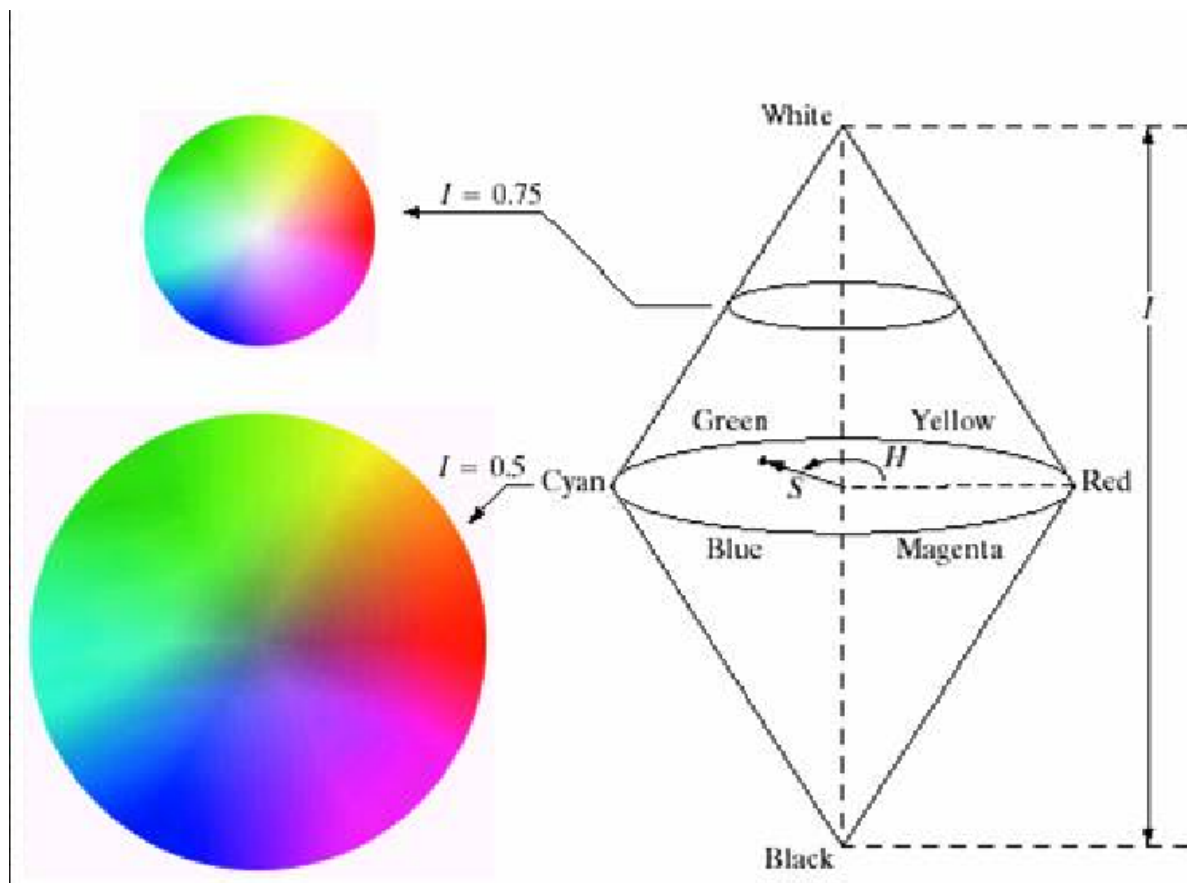
---

### HSI

三个参数分别代表色调（Hue）、饱和度（Saturation）和亮度（intensity）强度。

从模型上来看，色调是点S与原点连线的向量与红轴夹角；饱和度是向量的长度；而强度，或者说是亮度由平面在垂直强度轴的位置给出。





一些特殊情况，H的角度范围是 $[0, 2\pi]$ ，其中红色角度为0，绿色为 $2\pi/3$ .蓝色为 $4\pi/3$

其中色调和饱和度是彩色图的主要成分。

色调的专业术语是颜色的频率，不过我们就理解为是不同的颜色就好啦。

如果我让你描述一下啥是**HSI**模型，你咋办？

CSDN呀，不过不用这么麻烦，我来告诉你：

**HSI**模型就是根据人的视觉感知，从视觉系统出发，用色调、饱和度和亮度来描述色彩的一种模型

那么问题来了，为啥**HSI**模型在图像处理领域被广泛关注呢？

想想每一个字母的含义，其一，因为I分量代表着能量与强度，和彩色信息压根没有关系。也就是说，处理图像的时候只需要处理I分量即可，这样可以保证原彩色图像的种类。其二，因为**HSI**模型完全反映了人类感知颜色的基本属性，H和S与我们感受颜色的方式息息相关呀。

啊，冬天到了，大家都穿着深色的衣服，结果你穿一件橘色的衣服，走在路上大家眼里都是你。这就是色调，啥是饱和度？你们同样穿着红色的衣服，但你是洋红，别人是深红，那你走在路上还是最靓的仔，耶比~

---

从RGB模型到HSI模型的转变式线性变换吗？反过来呢？

---

## 全彩色图像处理

彩色图可以分解成三个单通道的灰度图，而灰度图又是由 $m \times n$ 个像素点组成的。所以我们可以将全彩色图像分为两大类：其一，对每一分量进行处理，然后再合成；其二，直接对彩色像素进行处理。

说明一下，这里的彩色像素是向量，因为全彩色图像至少有三个分量。

条件：普遍性和独立性

①保证处理对向量和标量都可用；

②对向量每一分量的操作对其他分量必须是独立的；

---

## 彩色图像平滑

两种方法：其一，对RGB每一分量平滑；其二，对HSI图像的强度分量I平滑，保持H和S分量不变。（为何不变，我前面在讲理论基础的时候已经涉及，先不往前面翻，回忆一下，为啥图像处理领域HSI模型被广泛关注咧？相信你可以的！）

那么，那么彩色图像锐化和平滑一样的处理方法。

---

## 彩色图像直方图处理

均匀得扩展彩色强度，保留彩色本身。害，说来说去都是那么一个HSI模型！！

这是一种方法，还有一种呢是独立地进行彩色图像分量的直方图均衡。

不过你想，人怎么能脱离这个社会独立存在，又不是山里的鸟兽，所以第一种方法或许更符合逻辑，自然效果应该更好一点叭~

---

## UINT6 图像压缩与编码

---

Frames Per Second，简称FPS啦，男孩子都懂，中文是每秒传输帧数。那这个帧数又是啥玩意呢？说白了就是图片啦，因为图片是视频的组成元素，就不多说啦。

那么FR呢？FR是帧率，指每秒重新绘制图像的次数。

bps：比特率，即单位时间内传输的比特数。

1Kbps = 1024Mbps

1Gbps = 1024Kbps

1Bps = 1bps

✓ 一张原始图像(1920x1080)，如果每个像素32bit表示 (RGBA)

1920x1080x4 = 8294400 Byte，大约8M

✓ 视频是1920x1080，30fps，1小时，大约8Mx30x60\*60 = 864000M

---

## 图像压缩

原因：数据冗余

那么数据冗余的原因呢？刨根问底！

①编码冗余

②时间和空间冗余

③心理视觉上的冗余，存在不相干的信息

方法：

①有损压缩

②无损压缩

（”损“是对信息是否确实的度量）

---

## 信息的度量

（一）信息公理

信息由不确定性程度进行度量，可以回忆一下AI数学涉及的惊奇程度，通常我们以负对数的概率来表示

不确定性程度越高，信息量越大。

（二）DMS的信息度量

何为信息熵，惊奇程度的期望，没啥好说的，AI数学都教给你们了。

通常情况下单位是比特/字符，以底为2的对数来计算。

（三）平均码字长

$$\vec{N} = \sum_{i=1}^m \beta_i p(d_i)$$

比特\字符

（四）编码效率

$$\eta = \frac{H}{\vec{N}}$$

---

## 霍夫曼编码

该编码是否会对信息造成损害呢？

如何计算，自己领悟！如果你DS学好了，你可以利用霍夫曼树来推导

---

## UINT7 图像分割

可以用一句话来描述你对图像分割的理解吗？ -----EXTRACT ROI

---

请简述你对OTSU的理解。

一种自适应的阈值分割算法。

按照灰度特性可以将图像分为前景和背景两部分，所以呢我们需要找一个阈值T，将图像分为两部分，使得这两部分的类间方差最大（分离性最大）。如果能这样的话，就能使得错分的概率最小化。（所以OTSU也可以称之为最大类间方差法）

---

## OTSU算法步骤描述

①计算归一化直方图

②计算灰度均值

$$\mu = \sum_{i=1}^M ip(i)$$

其中M为灰度级数，如果是八位图那么M为8,  $p(i)$ 就是归一化的产物，即每一个灰度级灰度出现的概率

③计算直方图的零阶w[i]和一级距u[i]（说实话这里没理解）

④计算找到最大类间方差对应此最大方差的灰度值，作为我们阈值化的阈值。

⑤阈值化

---

## 形态学运算

开运算用于消除小的亮的细节，闭运算消除小的暗的细节。

那么开运算的操作顺序是什么？（先膨胀后腐蚀，还是倒过来？）

分别会引起什么视觉变化？

---



## UINT8 表示与描述

链码：是一种边界的编码表表示法，用边界的方向作为编码的依据。

如何表示4链码和8链码？

何为原链码？归一化链码呢？

原链码存在哪些问题，如何解决？

何为形状数？

链码如何转为一阶差分链码？（注意逆时针）

好了，下课！