**Chap01**

**1.** 每个灰度图像可以看做是一个二维数组或者矩阵

灰度图像(grayscale image)一般每个像素（pixel）用一个字节（8bit）表示，此时最大能表示[0..255]共256个灰度级。

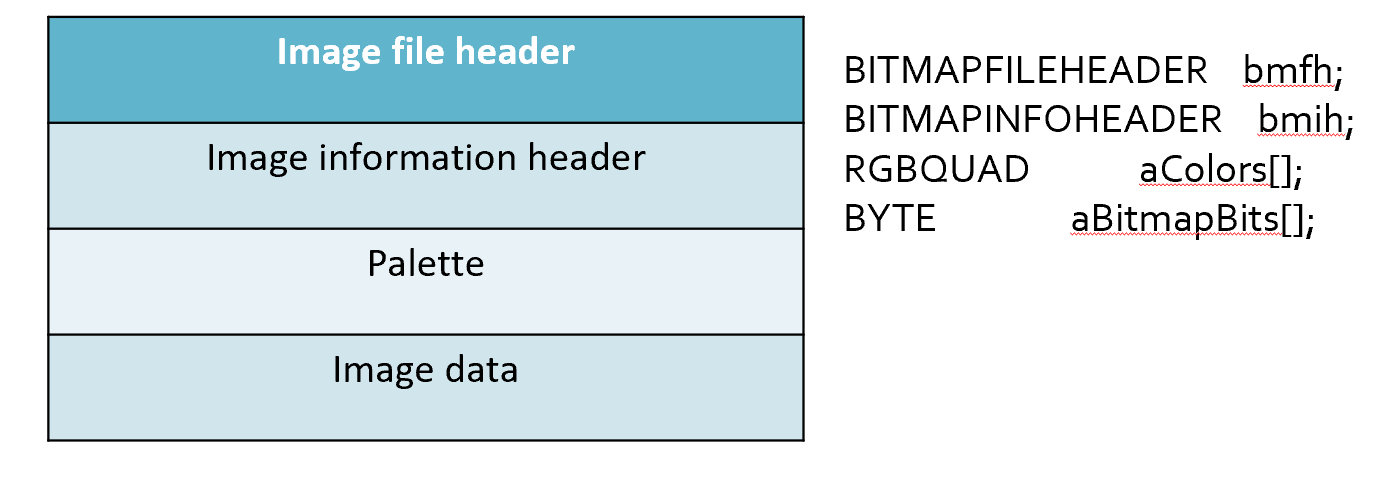
32 位多了一个阿尔法通道, 用来表示透明度.

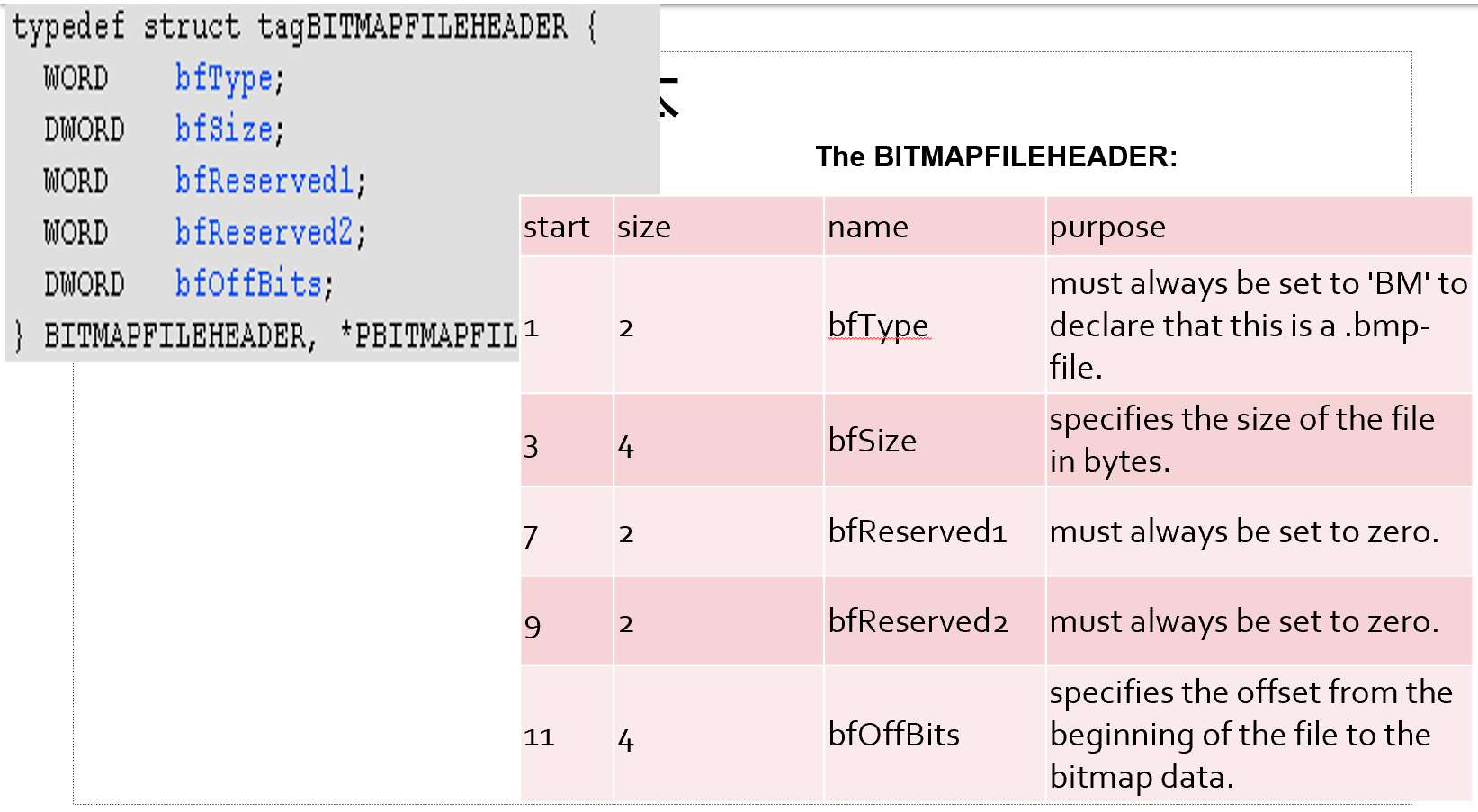
**2.BMP**

Windows操作系统的标准文件格式之一，有时候也存为.dib文件

大部分BMP文件是不压缩的形式

但它本身还是支持图像压缩的，如rle格式（行程长度编码，runlength encoding）压缩格式等





bfType

说明文件的类型，该值必需是0x4D42，也就是字符'BM'。

bfSize

说明该位图文件的大小，用字节为单位

bfReserved1

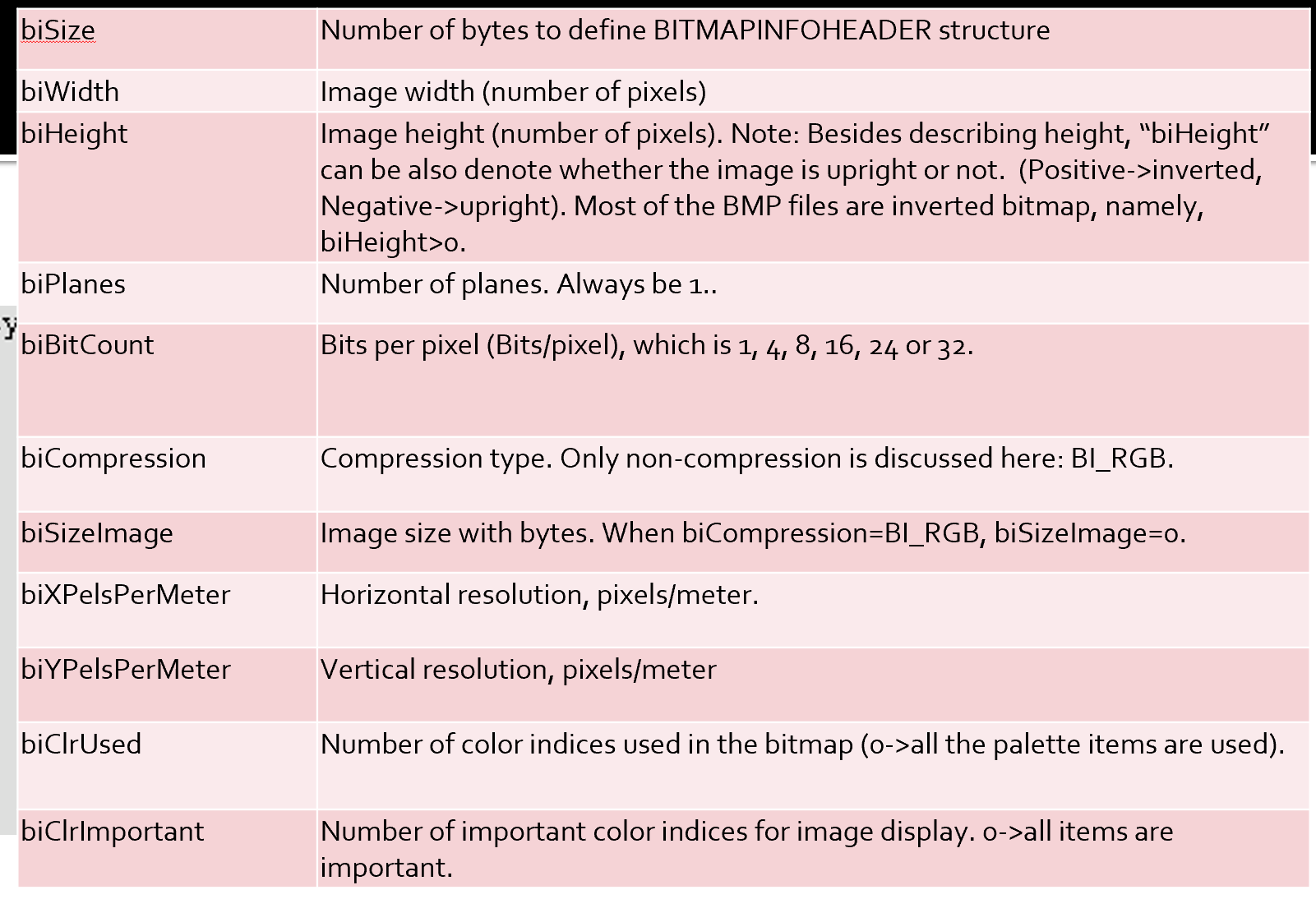
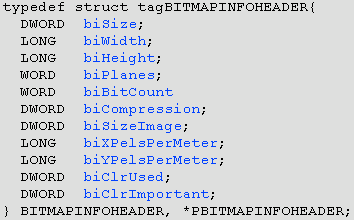
保留，必须设置为0

bfReserved2

保留，必须设置为0

bfOffBits

说明从文件头开始到实际的图象数据之间的字节的偏移量。这个参数是非常有用的，因为位图信息头和调色板的长度会根据不同情况而变化，所以你可以用这个偏移值迅速的从文件中读取到位数据。



biSize

How many bytes are needed to define 说明BITMAPINFOHEADER结构所需要的字数。

biWidth

说明图象的宽度，以象素为单位。

biHeight

说明图象的高度，以象素为单位。注：这个值除了用于描述图像的高度之外，它还有另一个用处，就是指明该图像是倒向的位图，还是正向的位图。如果该值是一个正数，说明图像是倒向的，如果该值是一个负数，则说明图像是正向的。大多数的BMP文件都是倒向的位图，也就是时，高度值是一个正数。

biPlanes

为目标设备说明位面数，其值将总是被设为1。

biBitCount

说明比特数/象素，其值为1、4、8、16、24、或32。但是由于我们平时用到的图像绝大部分是24位和32位的，所以我们讨论这两类图像。

biCompression

说明图象数据压缩的类型，我们只讨论没有压缩的类型：BI\_RGB。

biSizeImage

说明图象的大小，以字节为单位。当用BI\_RGB格式时，可设置为0。

biXPelsPerMeter

说明水平分辨率，用象素/米表示。

biYPelsPerMeter

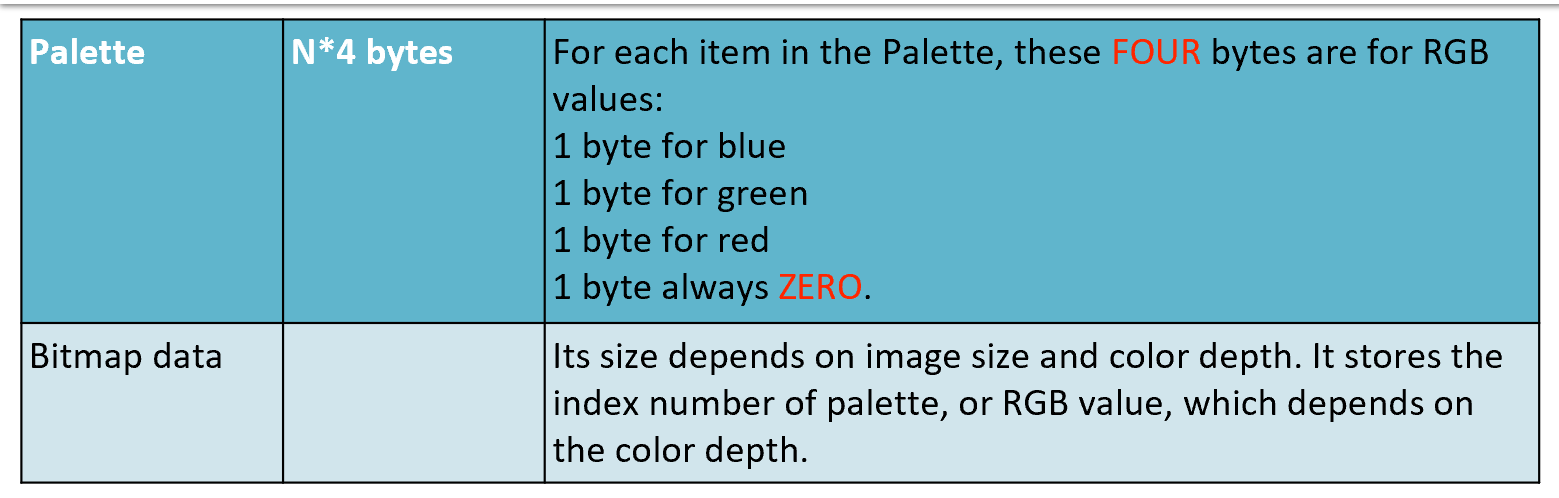
说明垂直分辨率，用象素/米表示。

biClrUsed

说明位图实际使用的彩色表中的颜色索引数（设为0的话，则说明使用所有调色板项）。

biClrImportant

说明对图象显示有重要影响的颜色索引的数目，如果是0，表示都重要。



1. 每一行的字节数必须是4的整倍数，如果不是，则需要补齐。
2. 一般来说，.bMP文件的数据从下到上，从左到右的。也就是说，从文件中最先读到的是图象最下面一行的左边第一个象素，然后是左边第二个象素……接下来是倒数第二行左边第一个象素，左边第二个象素……依次类推 ，最后得到的是最上面一行的最右一个象素。

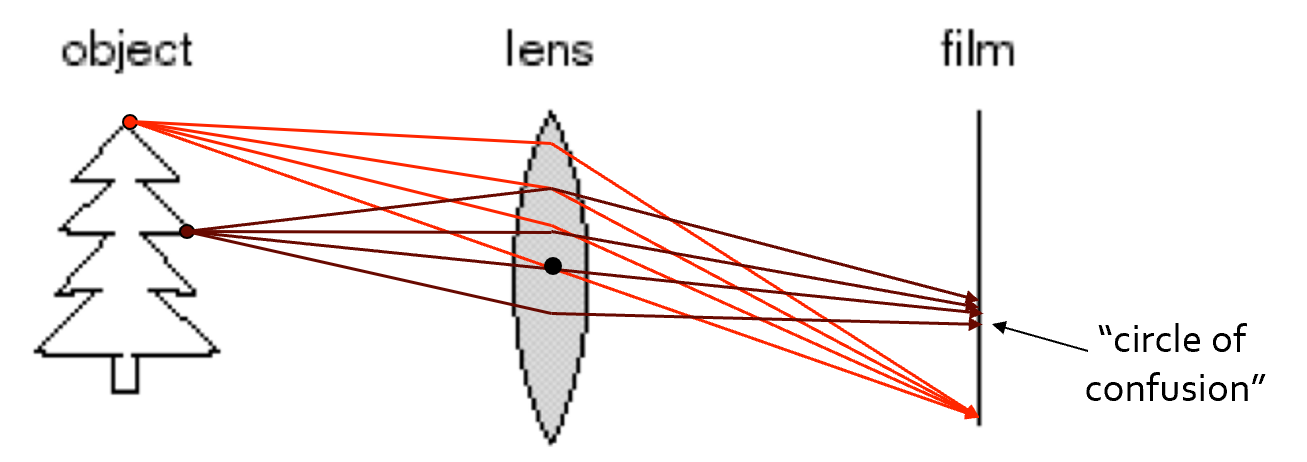
**3. 弥散圆**

如果孔径太小，可通过的光线就很少，导致光强太弱，同时当孔径小到一定程度时，会产生衍射现象。

通过使用透镜可以会聚更多的光线到一个成像点，只是只有在特殊的位置上才能达到这个目标，即大家以前就学习过的透镜聚焦。位于合适距离的物体点可以成清晰像，其他的点则会产生弥散圆。

通过改变镜头的形状可以改变聚焦时的物体点所需要的距离。

**弥散圆**(circle of confusion) ：在焦点前后，光线开始聚集和扩散，点的影像变模糊，形成一个扩大的圆。如果弥散圆的直径小于人眼的鉴别能力，在一定范围内实际影像产生的模糊是不能辨认的。不能辨认的直径叫做容许弥散圆。焦点前后各有一个容许弥散圆。



**4. 光圈**

**光圈对成像**：光圈孔径大的时候，会导致模糊；

同时，光圈孔径过小的时候，可通过的光线就很少，导致光强太弱，同时当孔径小到一定程度时，会产生衍射现象。

**光圈对景深**：

光圈的大小与景深有密切关系

焦点前后各有一个容许弥散圆，他们之间的距离叫做景深。景深随镜头的焦距、光圈值、拍摄距离而变化

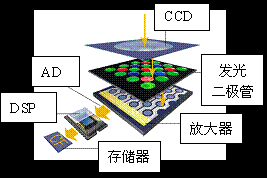
(1)、镜头光圈：  
光圈越大，景深越小；光圈越小，景深越大；  
(2)、镜头焦距  
镜头焦距越长，景深越小；焦距越短，景深越大；  
(3)、拍摄距离  
距离越远，景深越大；距离越近，景深越小

**单镜头反光相机原理**

所 谓“单镜头”是指摄影曝光光路和取景光路共用一个镜头，不像旁轴相机或者双反相机那样取景光路有独立镜头。“反光”是指相机内一块平面反光镜将两个光路分 开：取景时反光镜落下，将镜头的光线反射到五棱镜，再到取景窗；拍摄时反光镜快速抬起，光线可以照射到感光元件CMOS上。

**5. 成像过程**

感觉必考

(1)当使用数码相机拍

(2)当CCD曝光后，光电二极管受到光线的激发而释放出电荷，生成感光元件的电信号。摄景物时，景物反射的光线通过数码相机的镜头透射到CCD上。

(3) CCD控制芯片利用感光元件中的控制信号线路对发光二极管产生的电流进行控制，由电流传输电路输出，CCD会将一次成像产生的电信号收集起来，统一输出到放大器。

(4)经过放大和滤波后的电信号被传送到ADC，由ADC将电信号（模拟信号）转换为数字信号，数值的大小和电信号的强度与电压的高低成正比，这些数值其实也就是图像的数据。

(5)此时这些图像数据还不能直接生成图像，还要输出到DSP（数字信号处理器）中，在DSP中，将会对这些图像数据进行色彩校正、白平衡处理，并编码为数码相机所支持的图像格式、分辨率，然后才会被存储为图像文件。

(6)当完成上述步骤后，图像文件就会被保存到存储器上,我们就可以欣赏了。

**6. 彩色和消色**

色彩场景是由光与物体之间的交互过程产生的。如：反射、折射、散射、传播、吸收以及衍射等等。

色彩可以分为彩色（chromatic color）和消色（achromatic color）两大类。

彩色是指红、黄、蓝等单色以及它们的混合色。彩色物体对光谱各波长的反射具有选择性，所以它们在白光照射下呈现出不同的颜色。

消色，又称非彩色，也就是我们通常所说的灰度，是指白色，黑色以及各种深浅不同的灰色。消色物体对光谱各波长的反射没有选择性，它们是中性色。

**7. 色觉**

色觉是指不同波长的光线作用于视网膜而在大脑中引起的感觉。人眼可见光线的波长是390nm～780nm，一般可辨出包括紫、蓝、青、绿、黄、橙、红7种主要颜色在内的120～180种不同的颜色。

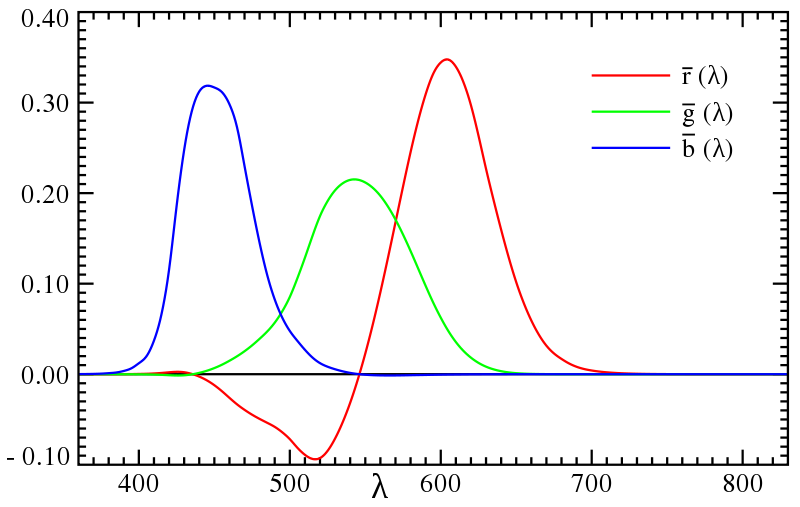
**视网膜**是人眼中最重要的组成部分，就像相机里的感光底片，专门负责感光成像。

视网膜上分布着两种视觉细胞，一种为杆状体，另一种为锥状体。杆状体细胞比较多，大约有上亿个，它对光极为灵敏，但没有区分色彩的能力。锥状体细胞则只有六、七百万个，它要在较强的照度下才能激发，它的存在使我们能够辨别各种不同的颜色。

**8.三原色**

三原色是目前最为代表性的色觉机制解释：

在视网膜上存在着三种分别对红、绿和蓝光线的波长特别敏感的视锥细胞或相应的感光色素，当不同波长的光线进入人眼时，与之相符或相近的视锥细胞发生不同程度的兴奋，于是在大脑产生相应的色觉；三种视锥细胞若受到同等程度的刺激，则产生消色。



700nm（R）、541.6nm（G）、435.8nm（B） 国际照明委员会-CIE

**8. 色觉的特点**

绝对色觉感知不准确，一般通过上下文判断。所以说有时候会出错，这个叫vision illusion

不同的人的感知能力也不同

感知到的对比度, 和背景的灰度成正比.

**感知的优先程度和敏感度**

优先程度：同等条件下，人们往往会注意到色调（Hue, H）的变化，然后是饱和度（Saturation, S），然后是亮度（Value, V）。

敏感度：人眼对于亮度的变化最为敏感，分辨能力最强。恰好与人眼的高动态能力相匹配。

**9.色彩空间**

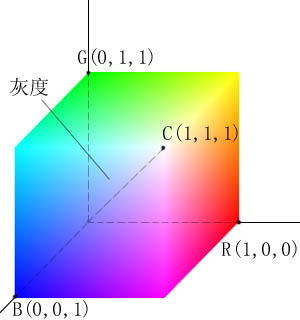
与设备有关的颜色模型:

RGB, CMY, HSV

与设备无关的颜色模型：

主要是国际照明委员会(CIE, Commission Internationale de L‘Eclairage/International Commission on Illumination)定义的CIE

CIE XYZ, CIE L\*a\*b和 CIE YUV

**（1）RGB**

RBG 颜色模型是三维直角坐标颜色系统中的一个单位正方体

在正方体的主对角线上，各原色的量相等，产生由暗到亮的白色，即灰度。（0，0，0）为黑，（1，1，1）为白，正方体的其他6个角点分别为红、黄、绿、青、蓝和品红。

RGB颜色模型构成的颜色空间是CIE原色空间的一个真子集。

RGB颜色模型通常用于彩色阴极射线管和彩色光栅图形显示器（计算机和电视机采用）。

**（2）CMY**

彩色印刷或者彩色打印的纸张不能发射光线，因而印刷机或打印机就只能使用一些能够吸收特定光波而反射其他光波的油墨或者颜色。

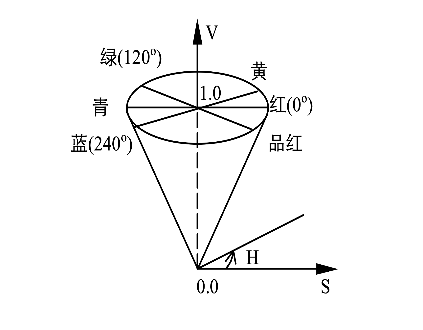
油墨或颜料的3种基色是以红、绿、蓝三色的补色青（Cyan）、品红（Magenta）、黄（Yellow）为基色。

用CMY模型产生的颜色称为相减色，是因为它减掉了为视觉系统识别颜色所需要的反射光

CMY空间与RGB空间互补，也就是用白色减去RGB空间中的某一颜色值就等于同样颜色在CMY空间中的值

没看懂 听听黎叔讲

**（3）HSV**

颜色空间是从人的视觉系统出发，用色调（Hue)、色饱和度（Saturation）和亮度（Intensity，或者Value）来

圆锥的顶面对应于V=1，它包含RGB模型中的R=1，G=1，B=1三个面，故所代表的颜色较亮。色度H由绕V轴的旋转角给定。红色对应于角度0o，绿色对应于角度120o，蓝色对应于角度240o。描述颜色。

在圆锥的顶点处，V=0，H和S无定义，代表黑色。

\* 符合人类视觉系统

\* 当采用RGB（或者CMY）颜色模型时，改变某一颜色的属性，比如改变色调就必须同时改变R、G、B（或者C、M、Y）三个坐标；而采用HSV颜色模型时只需改变H坐标。也就是说，HSV颜色模型中的三个坐标是独立的。合人眼对颜色的感觉。

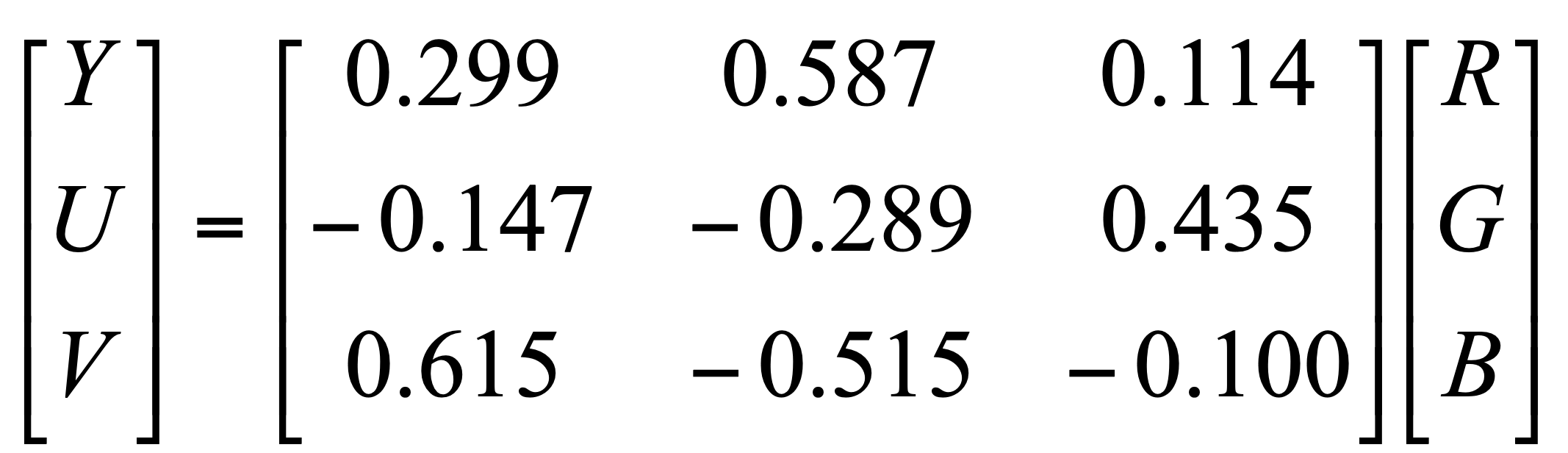
\* HSV颜色模型构成的是一个均匀的颜色空间，采用线性的标尺，彩色之间感觉上的距离与HSV颜色模型坐标上点的欧几里德距离成正比。

**（4）YUV**

在现代彩色电视系统中，通常采用三管彩色摄像机或彩色CCD（电荷耦合器件）摄像机，它把摄得的彩色图像信号，经分色，分别放大校正得到RGB，再经过矩阵变换电路得到亮度信号Y和两个色差信号R－Y、B－Y，最后发送端将亮度和色差三个信号分别进行编码，用同一信道发送出去。这就是我们常用的YUV颜色空间。

采用YUV颜色空间的重要性是它的亮度信号Y和色度信号U、V是分离的。如果只有Y信号分量而没有U、V分量，那么这样表示的图就是黑白灰度图。彩色电视采用YUV空间正是为了用亮度信号Y解决彩色电视机与黑白电视机的兼容问题，使黑白电视机也能接收彩色信号。

根据美国国家电视制式委员会，NTSC制式的标准，当白光的亮度用Y来表示时，它和红、绿、蓝三色光的关系可用如下式的方程描述：



**10. JPEG**

压缩策略：根据压缩比要求，从高频到低频逐步削减信息

好处

高频信息占用存储空间大，减少高频信息更容易获得高压缩比；

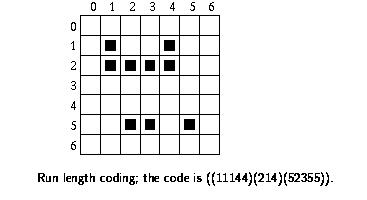
低频信息可以保留物体的基本轮廓和色彩分布，最大限度维持图像质量。

适合用于互联网

不适合用于线条画、文字、图标等，其有损压缩会导致这类对象的瑕疵严重

**11.RLE run length encoding**

* Each row of the image is described by a sublist, the first element of which is the row number.
* Subsequent terms are coordinate pairs; the first element of a pair is the beginning of a run and the second is the end.



**12 数字图像的重要性**

数字图像(或视频流)是信息呈现的主要形式。

对于人类来说，60%以上的信息是通过视觉获得的，这是感知和理解周围环境的基础。

计算机视觉和计算机图形学基础。

应用领域:计算机动画、视频监控、医学图像处理、远程图像处理、视觉效果增强等。

不同形式的成像

可见光，x射线，超声波，红外线

**Chap02**

1. **Binary Image**

二值图像（Binary Image）中像素（Pixel）的值只有[0,1]或者[0,255]，编程中一般用[0,255]来构造二值图像

优点：

更小的内存需求

运行速度更快

为二值图像开发的算法往往可以用于灰度级图像

更便宜

缺点：

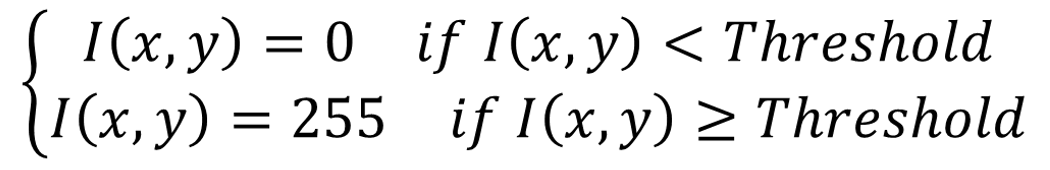
应用范围毕竟有限；

更无法推广到三维空间中

表现力欠缺，不能表现物体内部细节

无法控制对比度

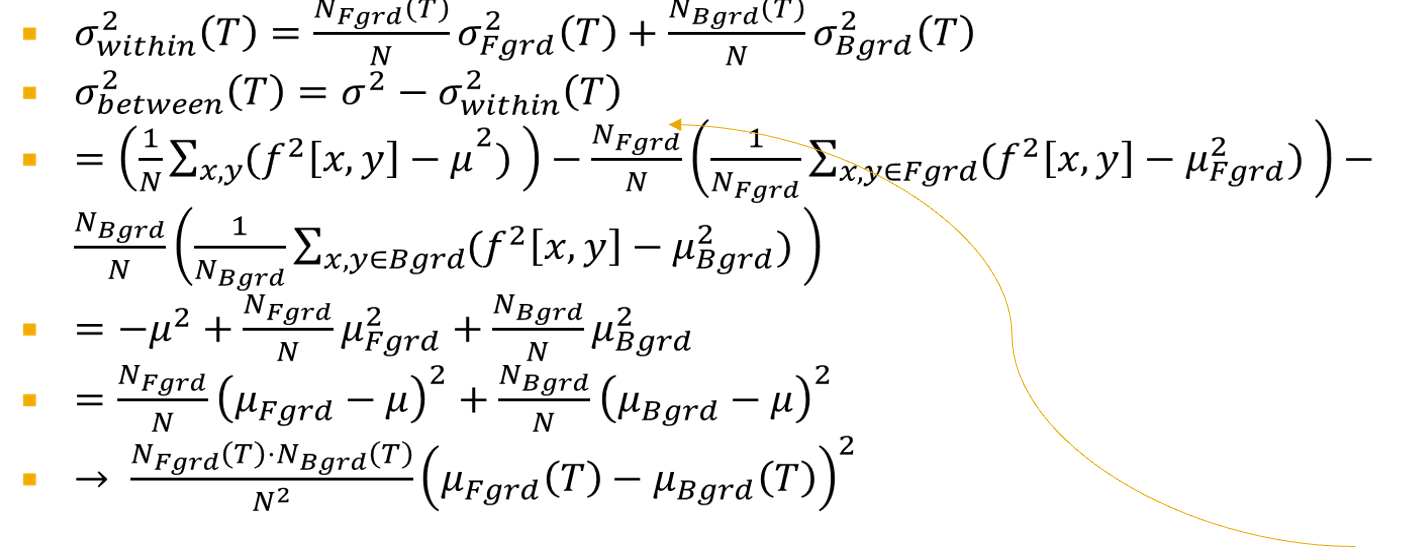
**2. binarization**



如何选取合适的threshold

基本思想：将二值化得到的二值图像视为两部分，一部分对应前景（Foreground），另一部分对应背景（Background）。尝试找到一个合适的threshold使得到的前景和背景的内部方差最小，而它们之间的方差则最大。

**推导过程！**





Step 1: 确定原始图像中像素的最大值和最小值；

Step 2: 最小值加1作为threshold对原始图像进行二值化操作；

Step 3: 根据对应关系确定前景和背景，分别计算当前threshold下的内部协方差和外部协方差；

Step 4: 回到Step 2直到达到像素最大值；

Step 5：找到最大外部和最小内部协方差对应的threshold.

**局部大津+滑动窗口！** (很有可能考)

局部自适应操作

设定一个局部窗口，在整个图像上滑动该窗口；

对于每一窗口位置，确定针对该窗口的threshold。

1. **Morphology 形态学**

**膨胀dilation、腐蚀erode、开操作和闭操作**

1. 膨胀

膨胀是将与物体**“接触”**的所有背景点合并到该物体中，使边界向外部扩张的过程。可以用来填补物体中的空洞。

1. 腐蚀

腐蚀是一种消除边界点，使边界向内部收缩的过程。可以用来消除小且无意义的物体。

Padding

1. 开操作

先腐蚀，后膨胀

用来消除小物体、在纤细点处分离物体、平滑较大物体的边界的同时并不明显改变其面积。

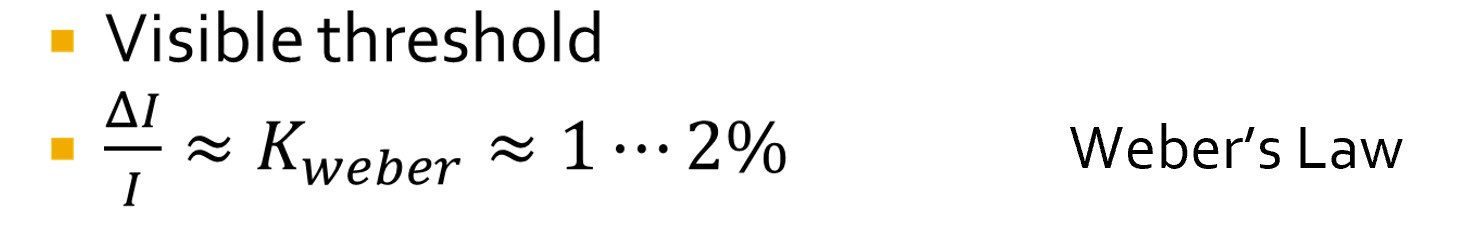
1. 闭操作

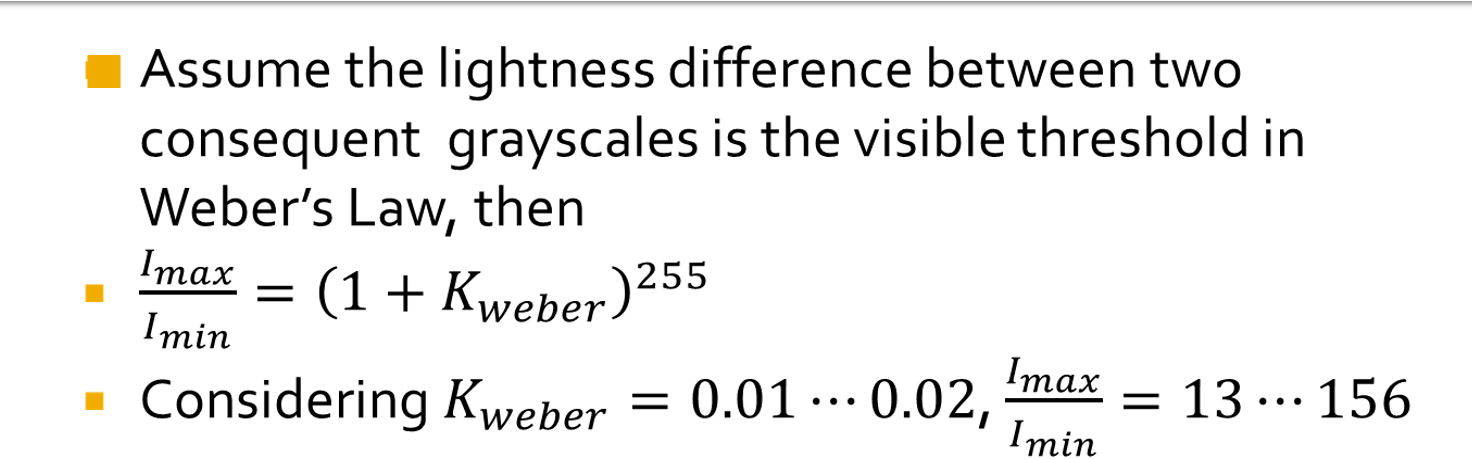
先膨胀，后腐蚀

用来填充物体内细小空洞、连接邻近物体、平滑其边界的同时并不明显改变其面积。

**Chap03 图像的灰度变换**

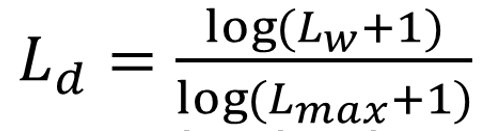
1. **灰度感知**





1. 可视化增强

为了增强图像的可视信息，对图像中的像素进行基于对数的操作

Ld 是显示亮度，Lw是真实亮度 Lmax是图像中的最大亮度 这个映射能够确保不管场景的动态范围是怎么样的，其最大值都能映射到1（白），其他的值能够比较平滑地变化。

3. 灰度图像及直方图

灰度直方图是一类统计图形，它表示一幅图像中各个灰度等级的像素个数在像素总数中所占的比重。

但是直方图把结构信息丢失，只知道颜色分布，不知道结构。

* 是空间域处理技术的基础。
* 反映图像灰度的分布规律，但不能体现图像中的细节变化情况。
* 对于一幅给定的图像，其直方图是唯一的。
* 不同的图像可以对应相同的直方图。

**4.直方图均衡化**

推导

直方图均衡化——寻找*T*（连续灰度变化）

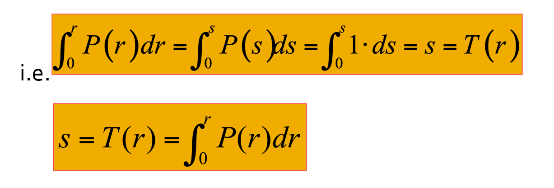
假设：

（1）令*r*和*s*分别代表变化前后图像的灰度级，并且 0≤*r*,*s* ≤1 。

（2）*P*(*r*)和*P*(*s*) 分别为变化前后各级灰度的概率密度函数（*r*和*s*值已归一化，最大灰度值为1）。

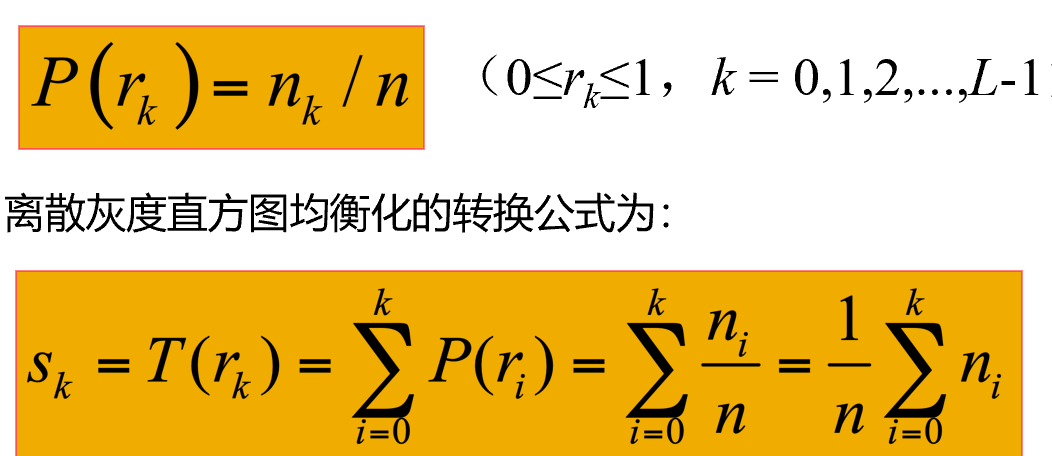
规定：

（1）在0≤*r* ≤1中，*T*(*r*)是单调递增函数，并且0≤*T*(*r*)≤1。

 （2）反变换*r* = *T*-1(*s*)也为单调递增函数。

考虑到灰度变换不影响像素的位置分布，也不会增减像素数目。所以有：

转换函数*T*在变量*r*处的函数值*s*，是原直方图中灰度等级为[0,*r*]以内的直方图曲线所覆盖的面积。



对于原直方图中的任意一个灰度级*rk*，只需将灰度级为[0,*rk*]以内的所有像素个数的和除以图像的像素总数，就可以得到转换之后的对应灰度级*sk*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *k* | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| *rk* | 0 | 1/7 | 2/7 | 3/7 | 4/7 | 5/7 | 6/7 | 1 |
| *nk* | 790 | 1023 | 850 | 656 | 329 | 245 | 122 | 81 |
| *P(rk)* | 0.19 | 0.25 | 0.21 | 0.16 | 0.08 | 0.06 | 0.03 | 0.02 |
| *sk* | 0.19 | 0.44 | 0.65 | 0.81 | 0.89 | 0.95 | 0.98 | 1.00 |
| *~sk* | 1/7 | 3/7 | 5/7 | 6/7 | 6/7 | 1 | 1 | 1 |
| *sk* | *s*1 | *s*3 | *s*5 | *s*6 | *s*6 | *s*7 | *s*7 | *s*7 |
| *nk* | 790 | 1023 | 850 | 985 |  | 448 |  |  |
| *P(sk)* | 0.19 | 0.25 | 0.21 | 0.24 |  | 0.11 |  |  |

按照均衡化的要求，在均衡化后的结果直方图中，各灰度级发生的概率应该是相同的，如右上图所示连续灰度级均衡化结果那样。但是，右下图中离散灰度级均衡化后，各灰度级出现的概率并不完全一样。

为什么？

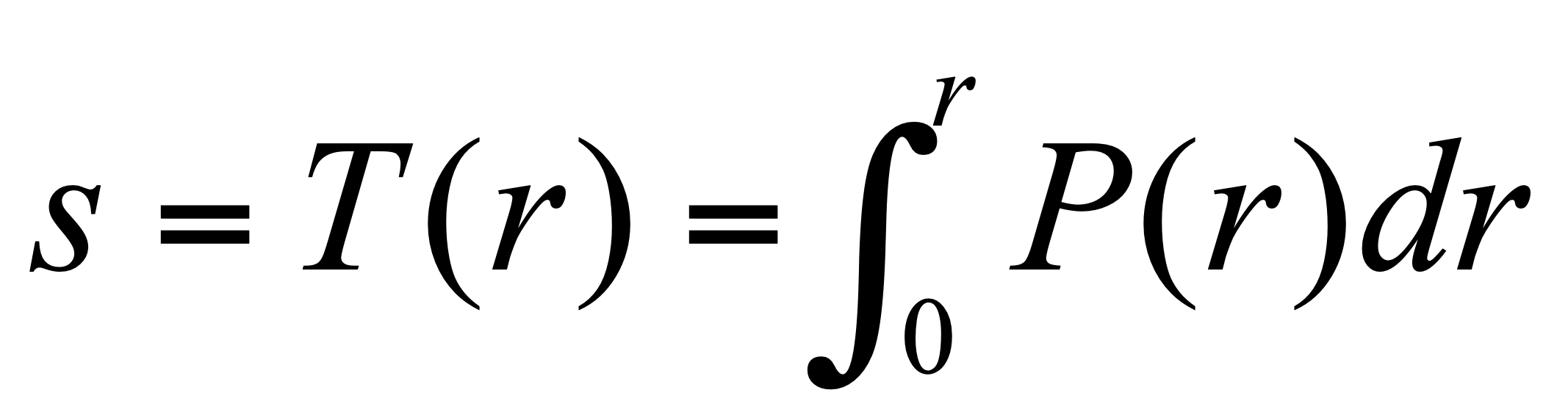
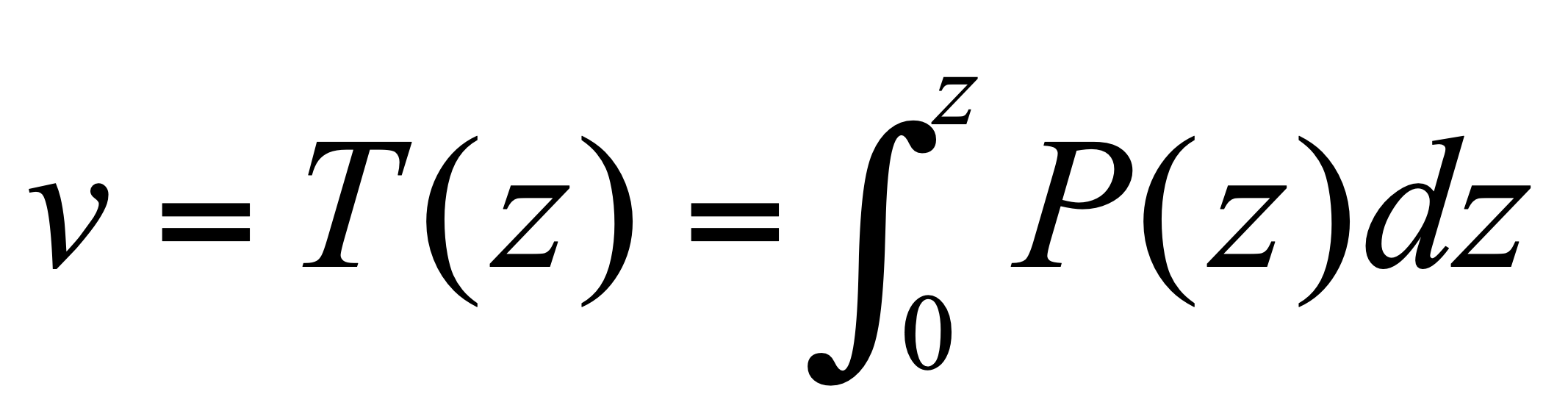
步骤2中，所得的sk不可能正好等于8级灰度值中的某一级，因此需要就近归入某一个灰度级中。这样，相邻的多个sk就可能落入同一个灰度级，需要在步骤3时将处于同一个灰度级的像素个数累加。因此，离散灰度直方图均衡化操作以后，每个灰度级处的概率密度（或像素个数）并不完全一样。

直方图均衡化实质上是减少图像的灰度级以换取对比度的加大。在均衡过程中，原来的直方图上出现概率较小的灰度级被归入很少几个甚至一个灰度级中，故得不到增强。若这些灰度级所构成的图象细节比较重要，则需采用局部区域直方图均衡化处理。

**5.直方图匹配**

* 所谓直方图匹配，就是修改一幅图像的直方图，使得它与另一幅图像的直方图匹配或具有一种预先规定的函数形状。
* 直方图匹配的目标，是突出我们感兴趣的灰度范围，使图像质量改善。
* 利用直方图均衡化操作，可以实现直方图匹配过程。

1. 基于公式将原始图像中的 r 映射为s



1. 基于公式将结果直方图中的z 映射到v
2. 步骤3：由*v* = *G*(*z*)得到*z* =*G*-1(*v*)。由于*s*和*v*有相同的分布，逐一取*v* = *s*，求出与*r*对应的*z* =*G*-1(*s*)。
3. 方法简述：

在步骤1和2中，分别计算获得两张表（参见直方图均衡化中的算例），从中选取一对*vk*、*sj*，使*vk* = *sj*，并从两张表中查出对应的*zk*、*rj*。这样，原始图像中灰度级为*rj*的所有像素都映射成灰度级*zk*，最终得到所期望的图像。

6.直方图变换

Ppt上例子 略

**Chap04 几何变换**

**1.**简单的几何变换

平移、镜像变换、旋转、比例放缩、错切

空洞问题

图像经过旋转变换以后，新图像中会出现许多空洞。用插值方法填补。

行插值——按顺序寻找每一行中的空洞像素，设置其像素值与同一行中前一个像素的像素值相同。

2.插值

(1)最邻近插值

输出像素的灰度值等于离它所映射到的位置最近的输入像素的灰度值。

为了计算几何变换后新图像中某一点P’处的像素值，可以首先计算该几何变换的逆变换，计算出P’所对应的原图像中的位置P。通常情况下，P的位置不可能正好处在原图像的某一个像素位置上（即P点的坐标通常都不会正好是整数）。寻找与P点最接近的像素Q，把Q点的像素值作为新图像中P’点的像素值。

**（2）双线性插值**

1. 定义双线性方程g(x,y)=ax+by+cxy+d 。
2. 分别将A、B、C、D四点的位置和灰度代入方程，得到方程组。
3. 解方程组，解出a、b、c、d四个系数。
4. 将P点的位置代入方程，得到P点的灰度。
5. morph

大小相同的两幅图的转换作静态变换。从一幅图a逐渐变化成第二幅图b。

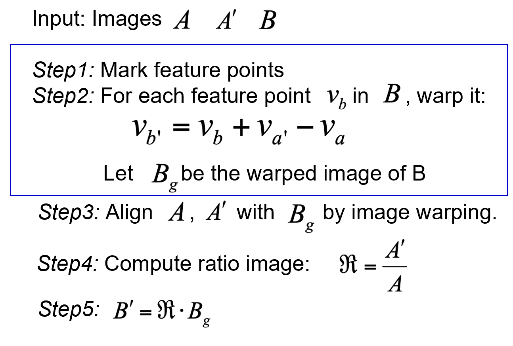
原理：让图a中每个像素的颜色，逐渐变成图b相同位置像素的颜色。

方法：根据变换的快慢，设置相应的步长，将图a每一点的RGB逐渐变成图b相同位置象素的RGB。可以选择等比或等差的方式，或其它方式让：ra-->rb。

对于灰度图像，可以直接用等比或等差级数定义步长，使颜色从原图变到目标图。

选择两幅结构相似、大小相同的画面作为起始和结束关键帧，这样才能比较容易地实现自然、连续的中间变形过程。

在起始和结束画面上确定和勾画出各部分（主要轮廓）的结构对应关系，也即从起始画面上的一个点变到结束画面上的另一个对应点的位置，这是变形运算所需要的参数。根据需要，对应点的位置可以任意移动。调整起始帧的对应点位置，可以模拟摄像中的镜头渐变效果。

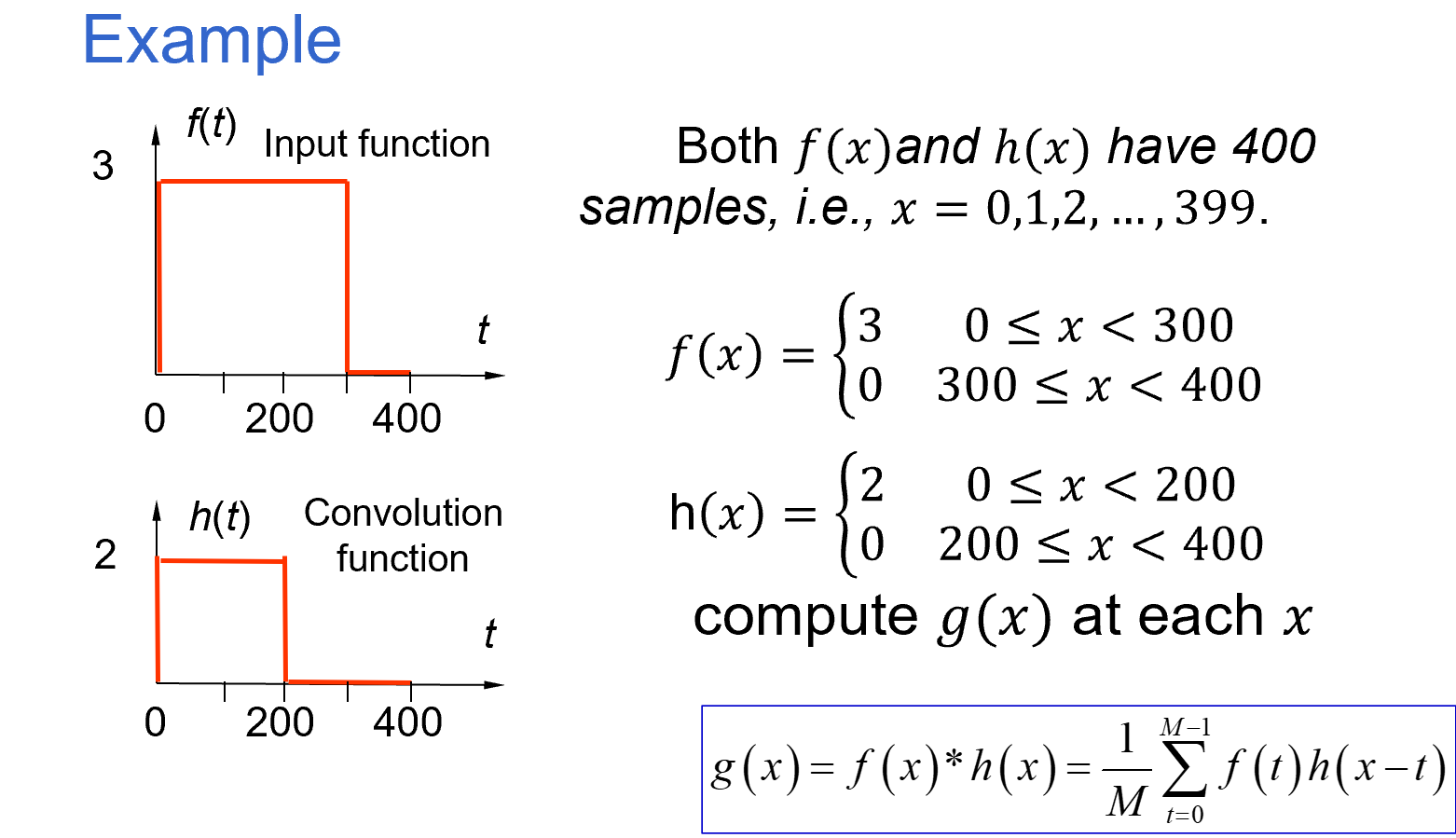
1. 表情变换

**Chap05 卷积和滤波**

1.卷积

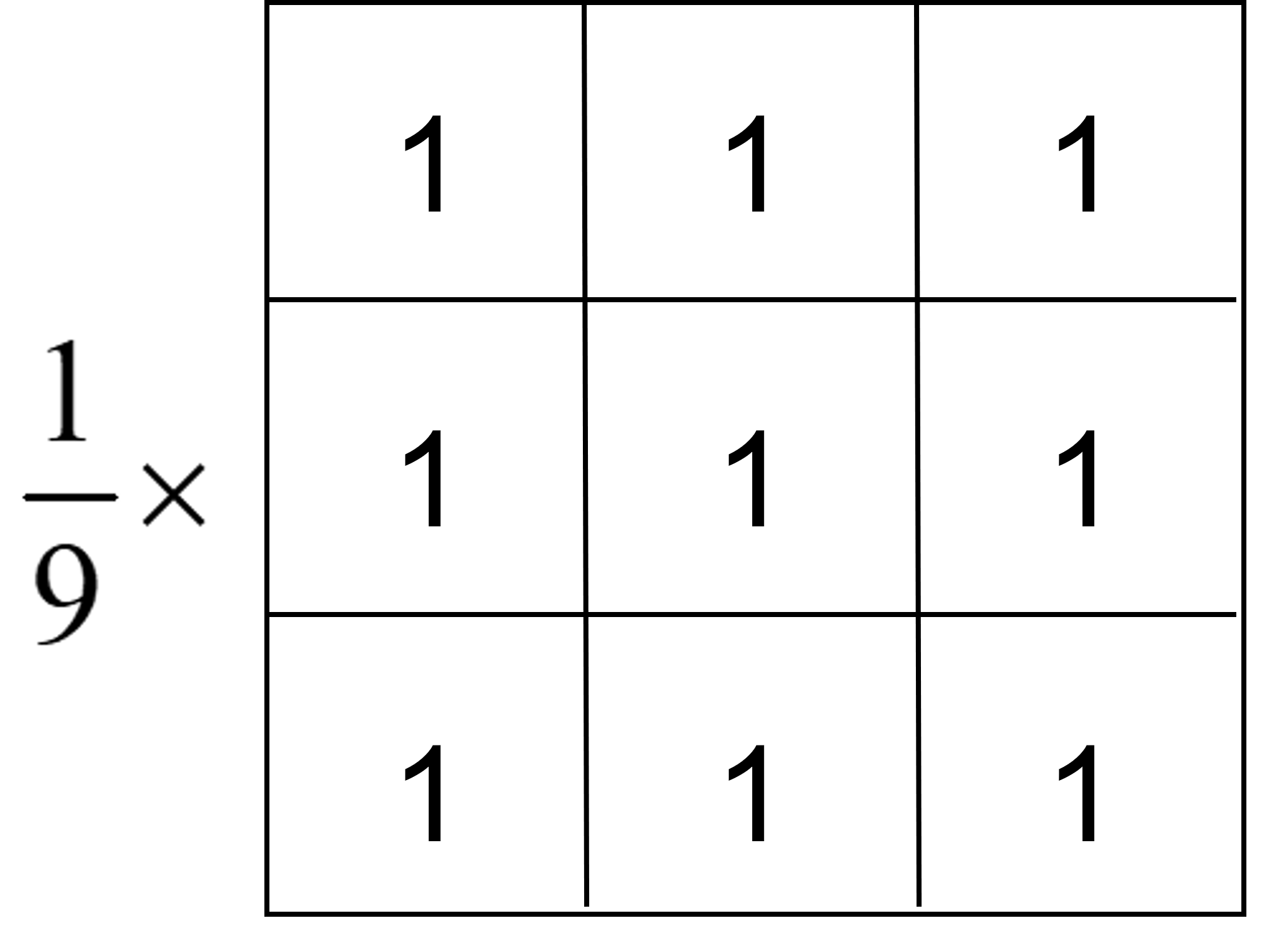
****

****



2.空间滤波

平滑空间滤波器用于模糊处理和减少噪声。模糊处理经常用于预处理，例如，在提取大的目标之前去除图像中一些琐碎的细节，桥接直线或曲线的缝隙。

1. 均值滤波

平滑线性空间滤波器的输出是包含在滤波掩模邻域内像素的简单平均值。因此，这些滤波器也称为均值滤波器。

均值滤波器的主要应用是去除图像中的不相干细节，即那些与滤波掩模尺寸相比更小的像素区域。

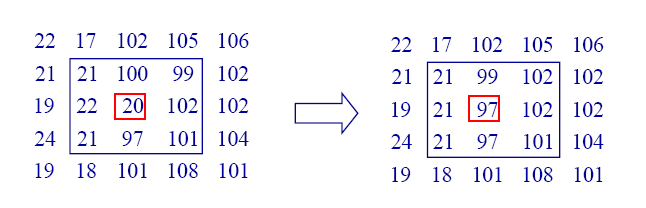
滤波掩模的大小与图像的平滑效果有直接的关系。当掩模比较小时，可以观察到在整幅图像中有轻微的模糊，当掩模大小增加，模糊程度也随之增加。

为了对感兴趣物体得到一个粗略的描述而模糊一幅图像，这样，那些较小物体的强度与背景混合在一起了，较大物体变得像“斑点”而易于检测。掩模的大小由那些即将融入背景中去的物件尺寸来决定。

1. 中值滤波

统计滤波器是一种非线性的空间滤波器，它的响应是基于窗口内图像区域中像素值的排序，由统计排序结果决定的值代替中心像素的值。

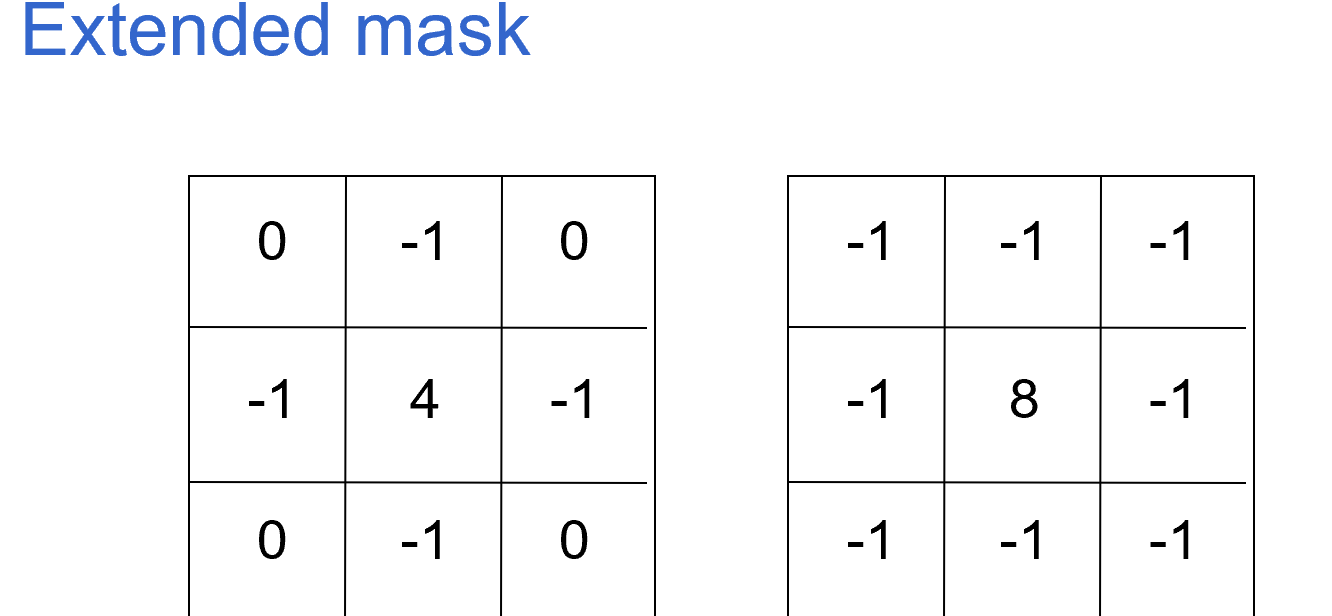
统计滤波器中最常见的例子就是中值滤波器。

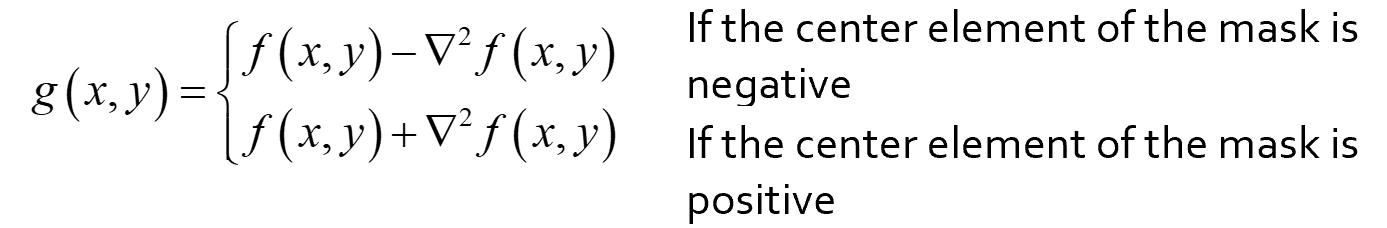
* 用像素邻域内灰度的中值代替该像素的值。
* 提供了优秀的去噪能力，比小尺寸的线性平滑滤波器的模糊程度明显要低。
* 对处理脉冲噪声（也称为椒盐噪声）非常有效，因为这种噪声是以黑白点叠加在图像上的。

突出图像中的细节或者增强被模糊了的细节。

（3）拉普拉斯算子





优点是让图像更加sharp,缺点是让图像噪声更加明显

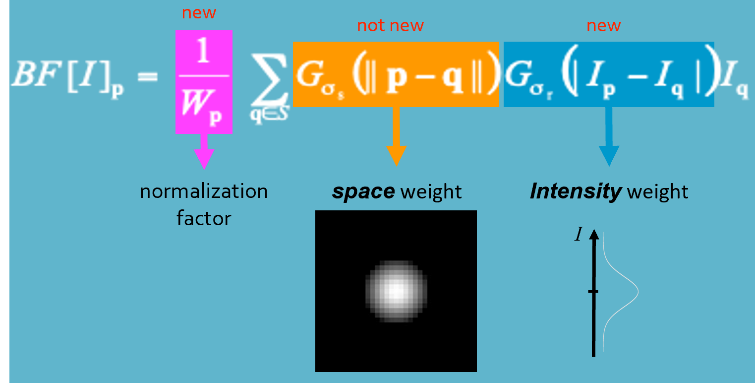
将原始图像和拉普拉斯图像叠加在一起的简单方法可以保护拉普拉斯锐化处理的效果，同时又能复原背景信息。

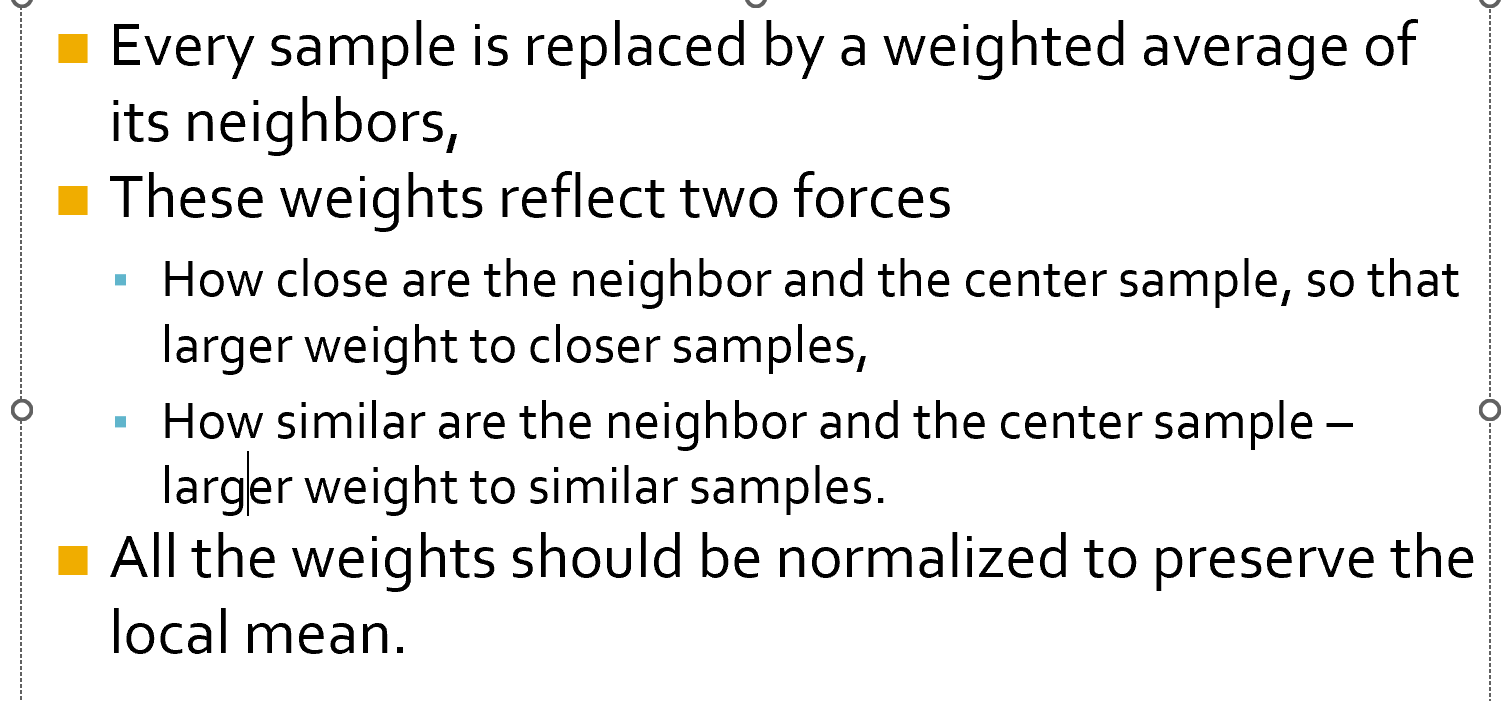
4.双边滤波

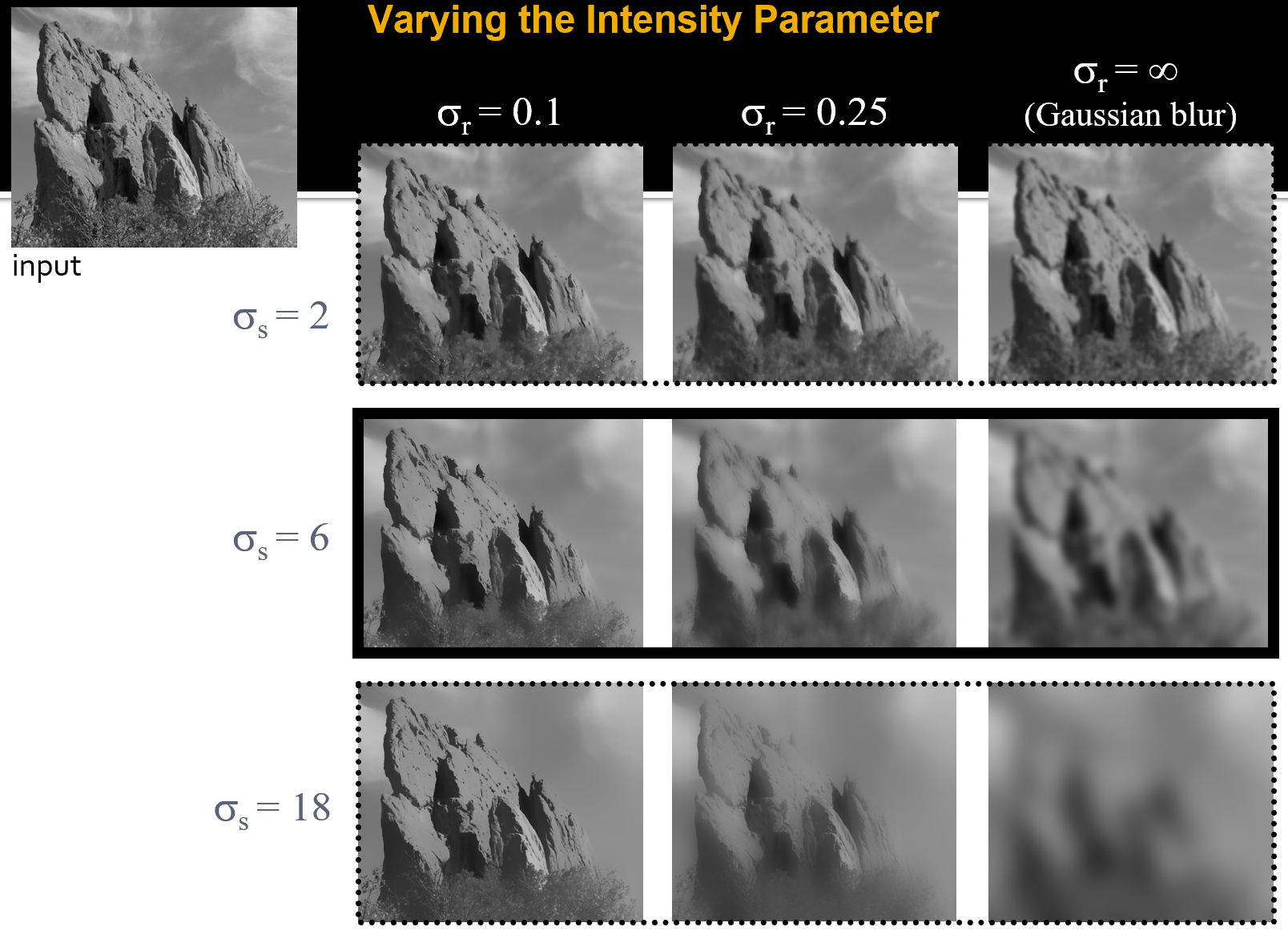
一幅图像有两个主要特征

空间域S，它是一幅图像中可能位置的集合。这与分辨率有关，也就是图像中的行数和列数。

强度域R，是可能像素值的集合。用来表示像素值的位数可能会变化。常见的像素表示形式是无符号字节(0到255)和浮点数。

高斯滤波只考虑了前一个域。





1. **引导滤波**

****

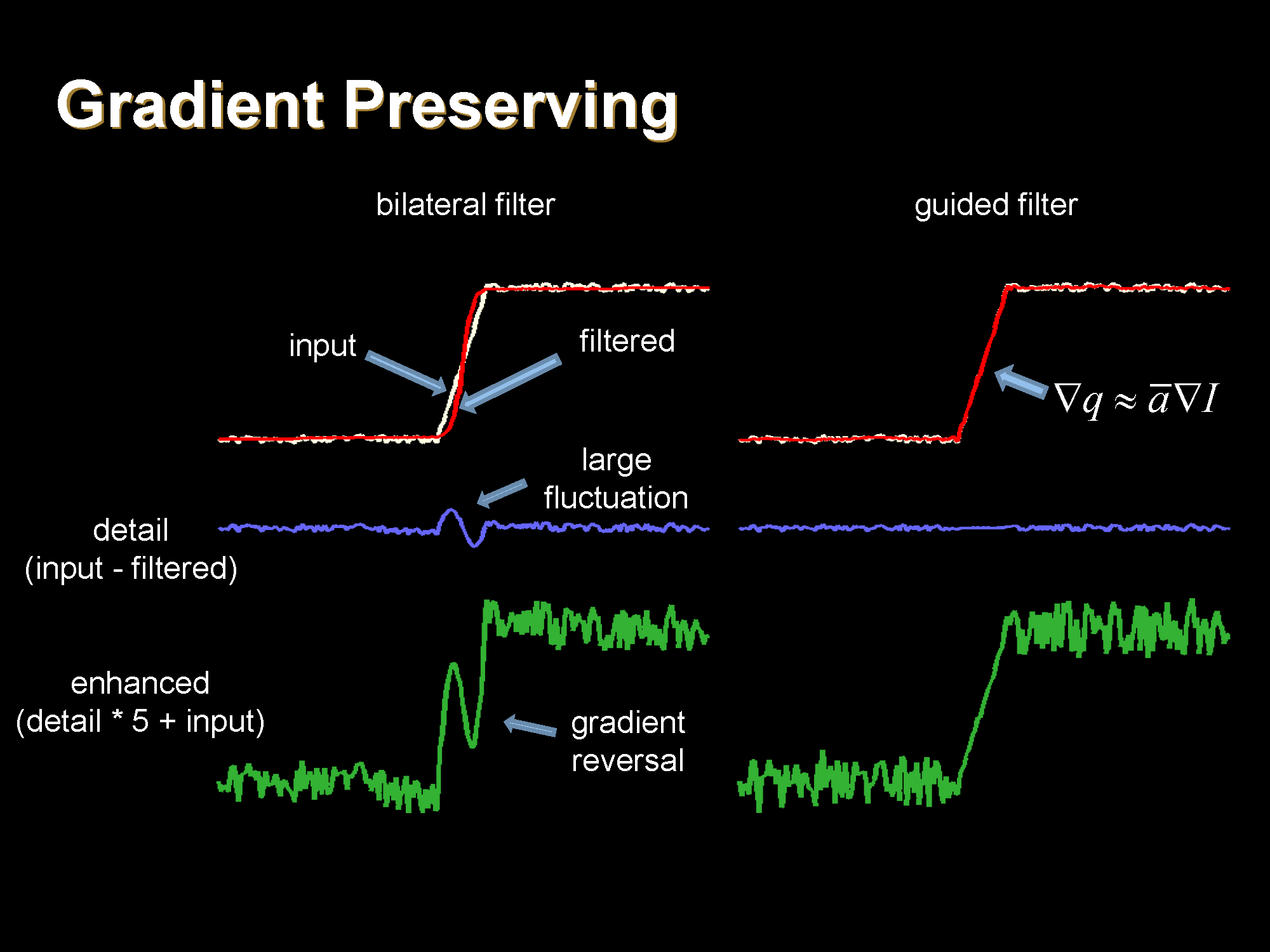
**Guided Filter 的优点**

保边（保梯度就一定能保边，反之不一定）

非迭代

O1的时间，快且不需要通过近似的方法

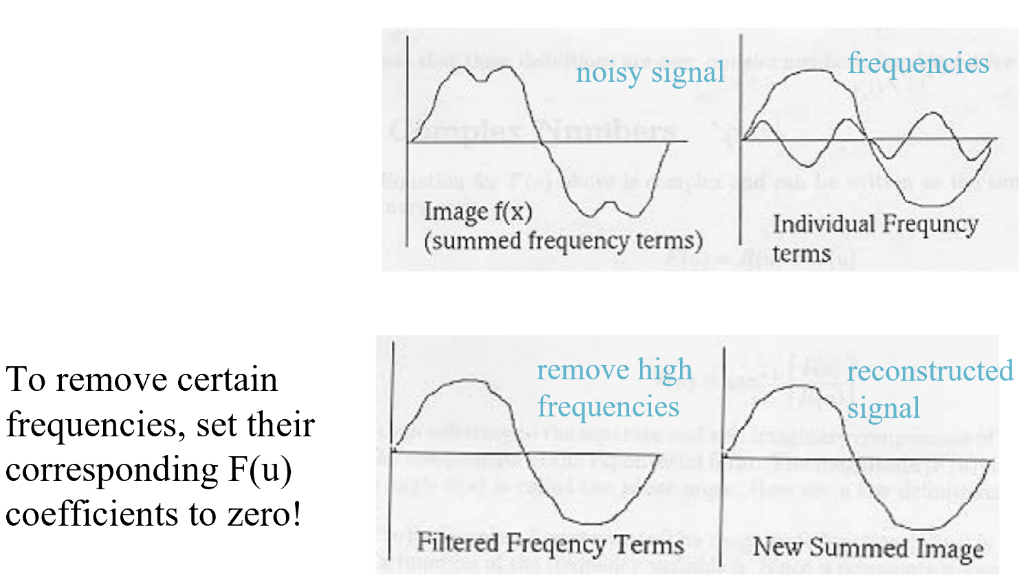
不存在梯度逆转的问题

****

**GF>BF**

因为Guided filter考虑了保梯度，不会造成梯度的逆转，但是双边滤波会产生梯度的逆转，因为双边保的是difference的绝对值，而不是方向。

1. Sparse Norm Filter

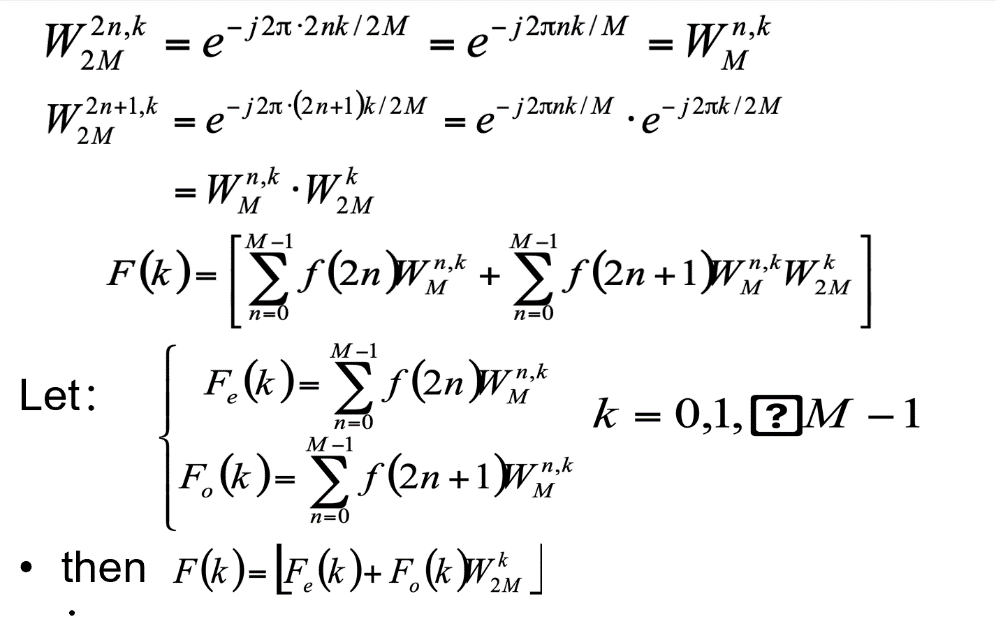
**Chap06**

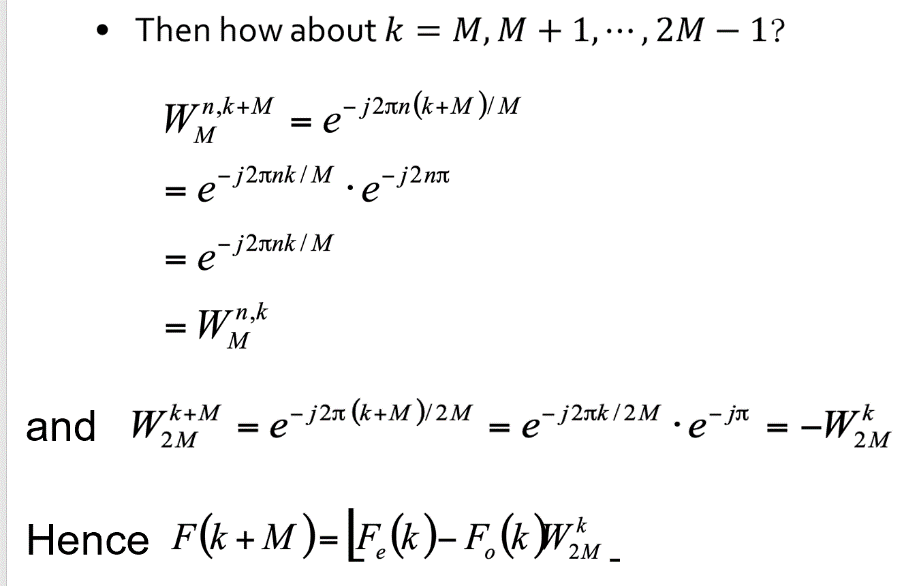
1. 傅里叶变换

用处：去除噪声

1. FFT的推导



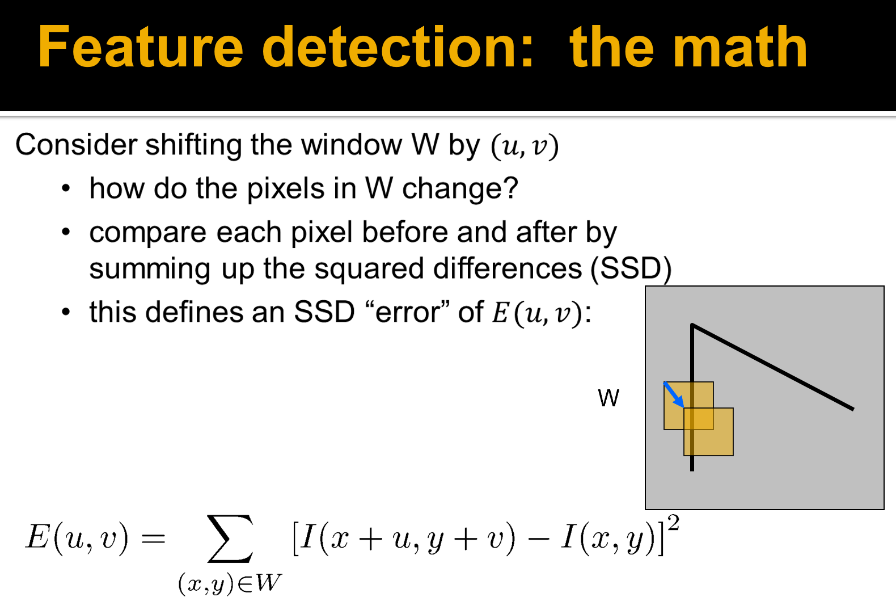
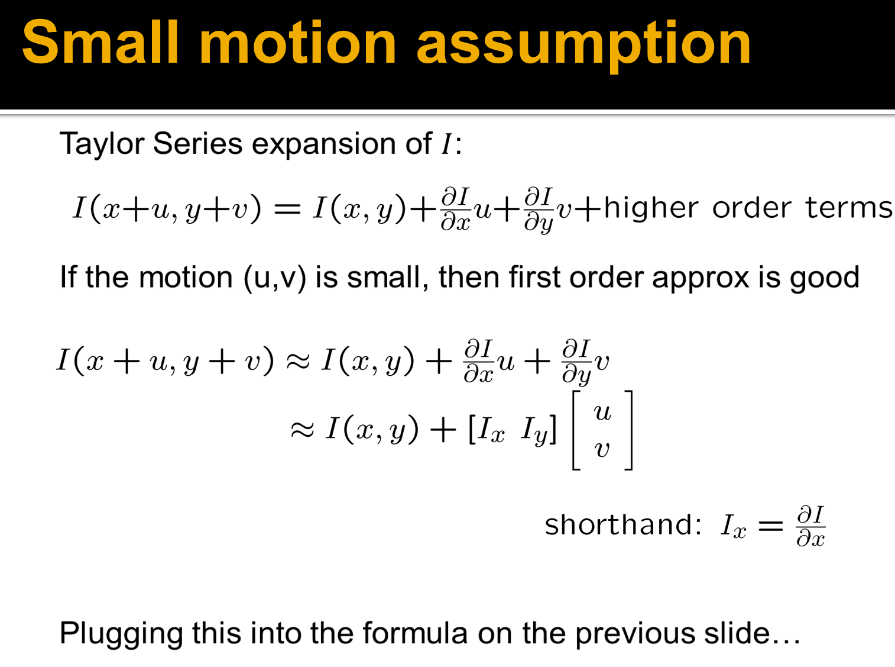


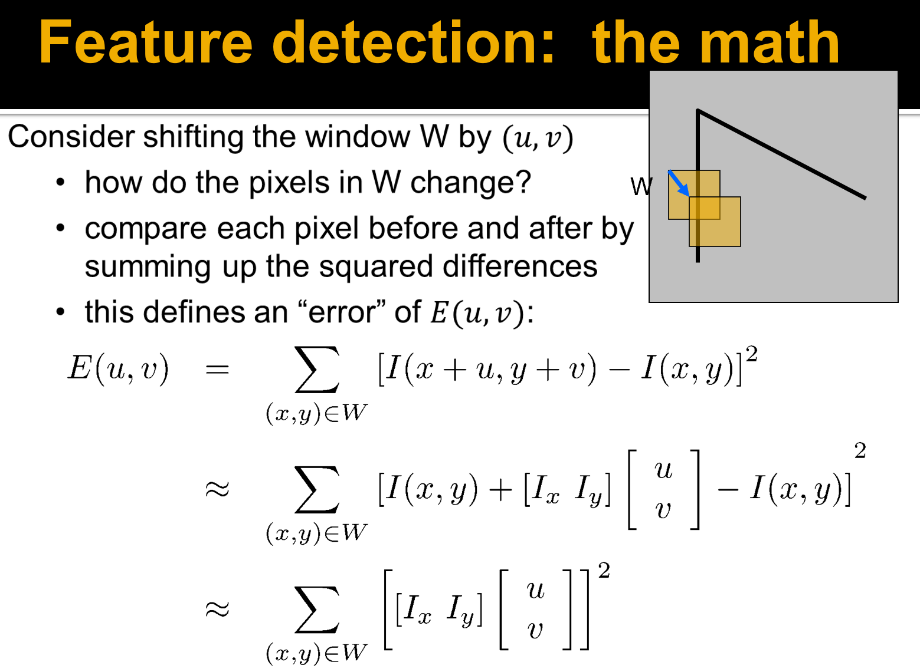
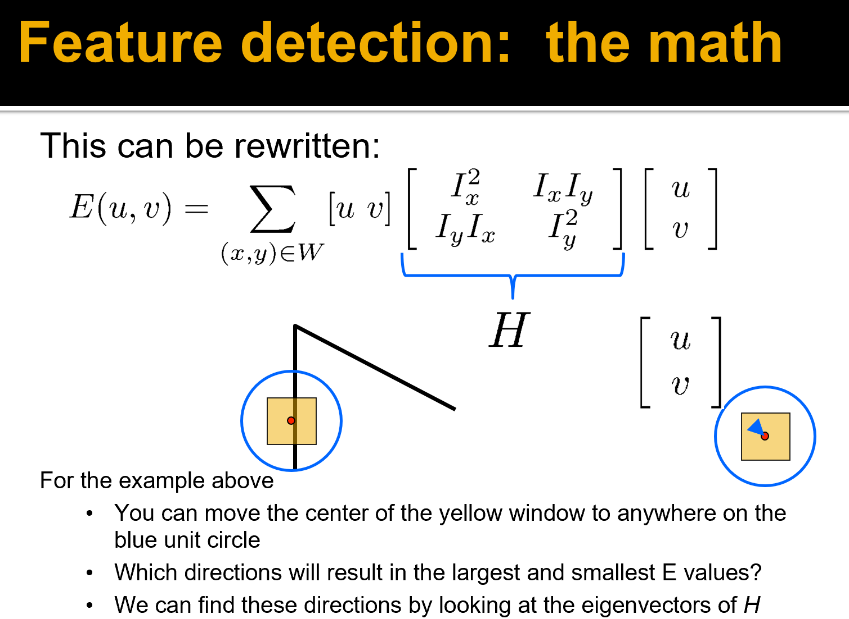


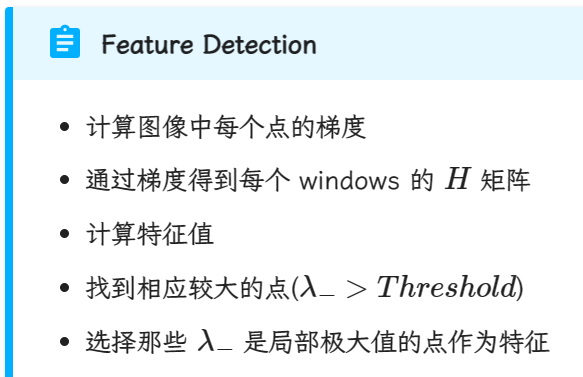
3. DFT中幅值和相位哪个更重要

相位可以用来重建图像但是幅值不可以

**Chap07**

1. **特征检测**

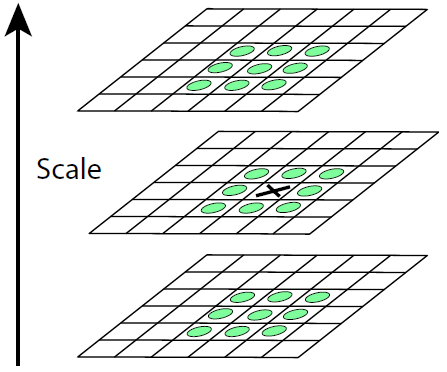




旋转不变性，尺度不具有不变性！

1. Harris Laplacian

初始化:多尺度Harris角点检测；基于拉普拉斯函数的尺度选择

1. SIFT
2. Detector

Difference of Gaussians in space and scale

（2）descriptor

SIFT为什么能做到旋转不变性

能够根据主要的梯度方向来旋转小块，使其置于规范方向。

SIFT为什么能做到尺度不变性

Basic idea

1. 在检测特征附近创建一个16\*16的正方形窗口

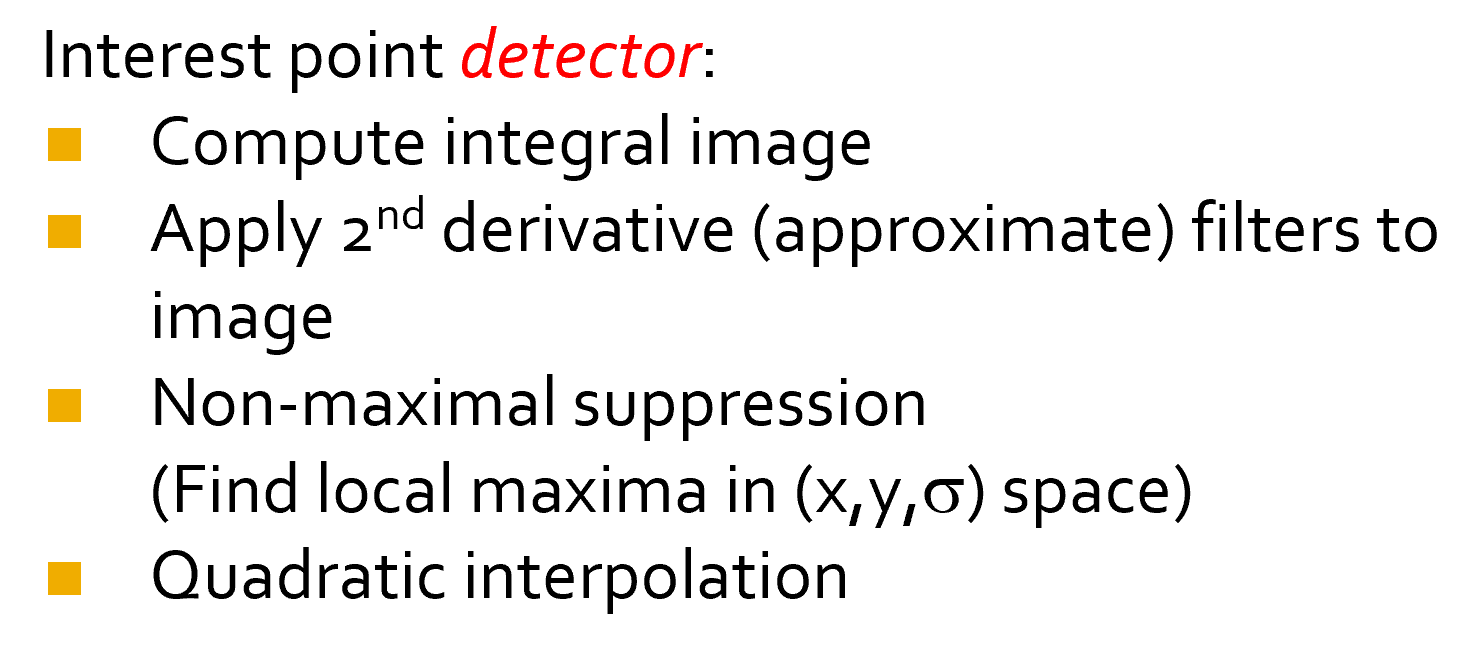
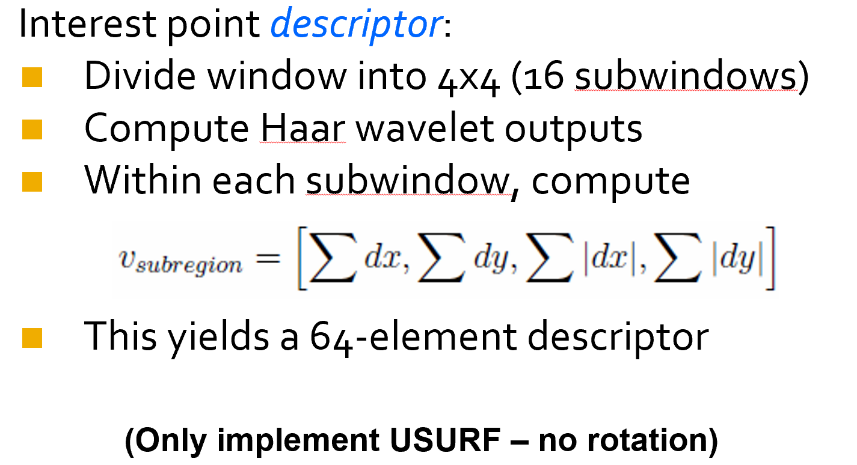
2. 对于每个像素计算其边缘方向

3. 除去弱的边(小于阈值梯度幅度)

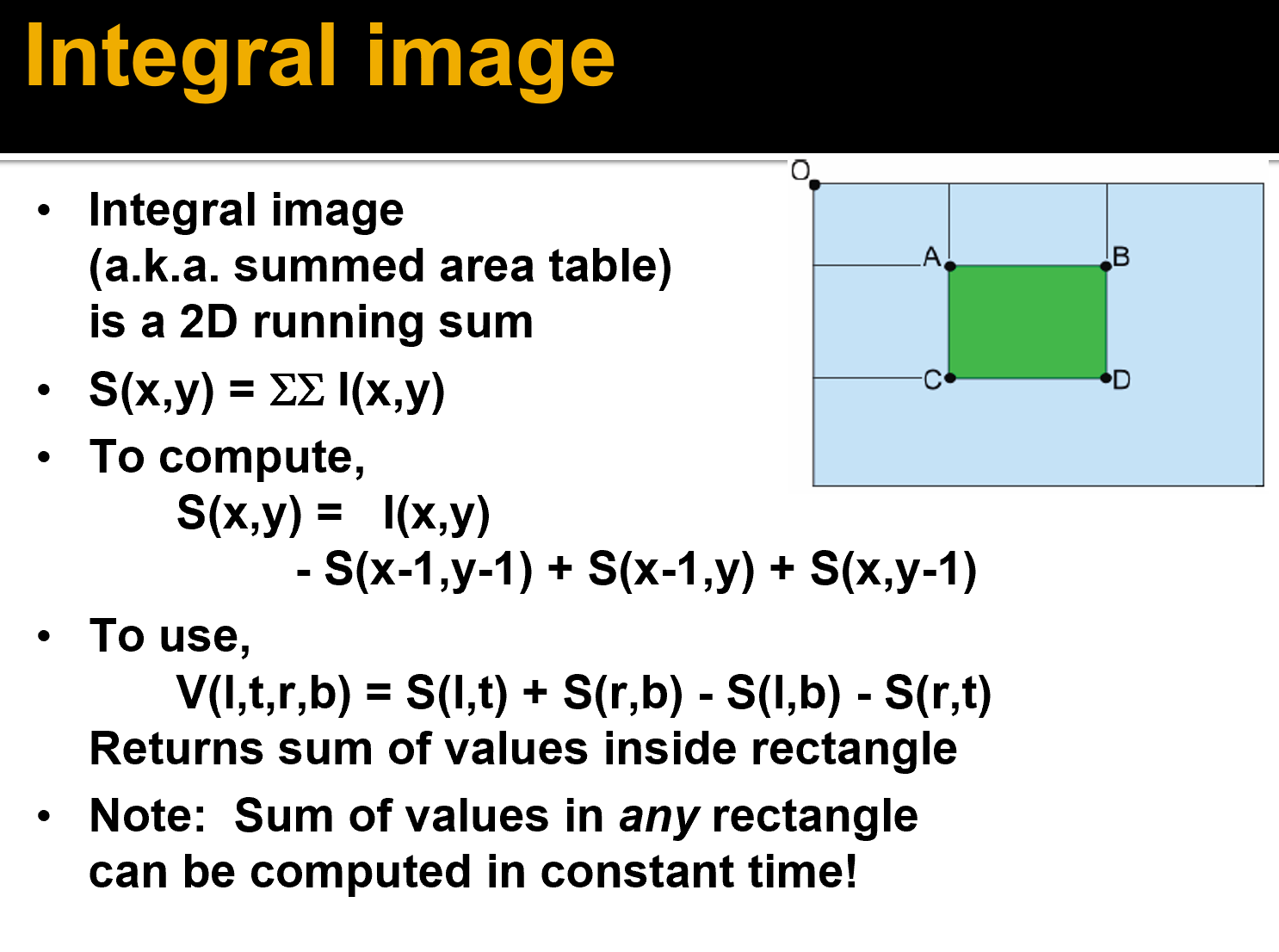
4. 对剩下的边创建一个直方图

Full version

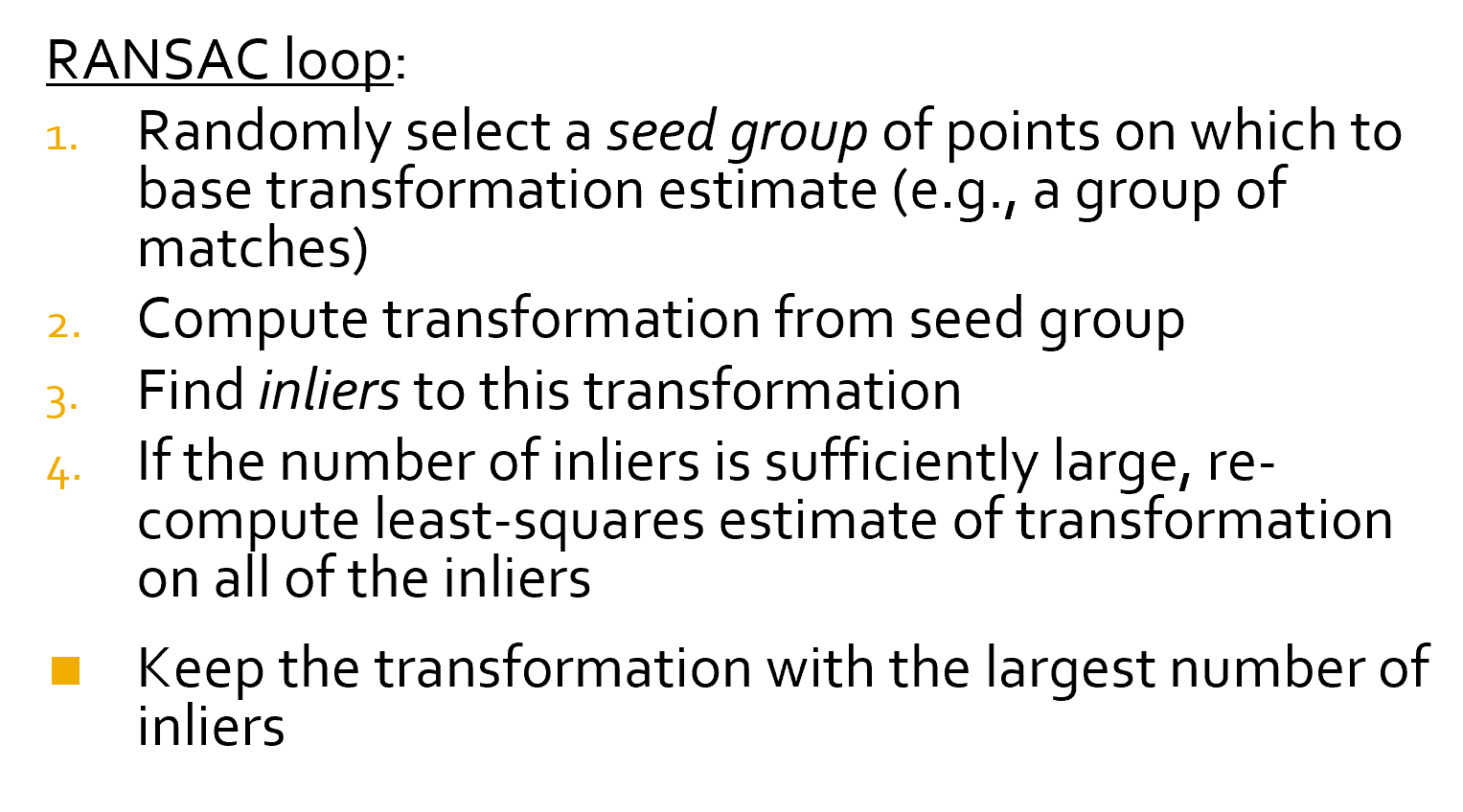
* 将 16×16 的窗口 划分为 4×4 的网格
* 对每个网格计算其方向直方图
* 16 cells \* 8 orientations = 128 dimensional descriptor(128 维向量)

1. SURF

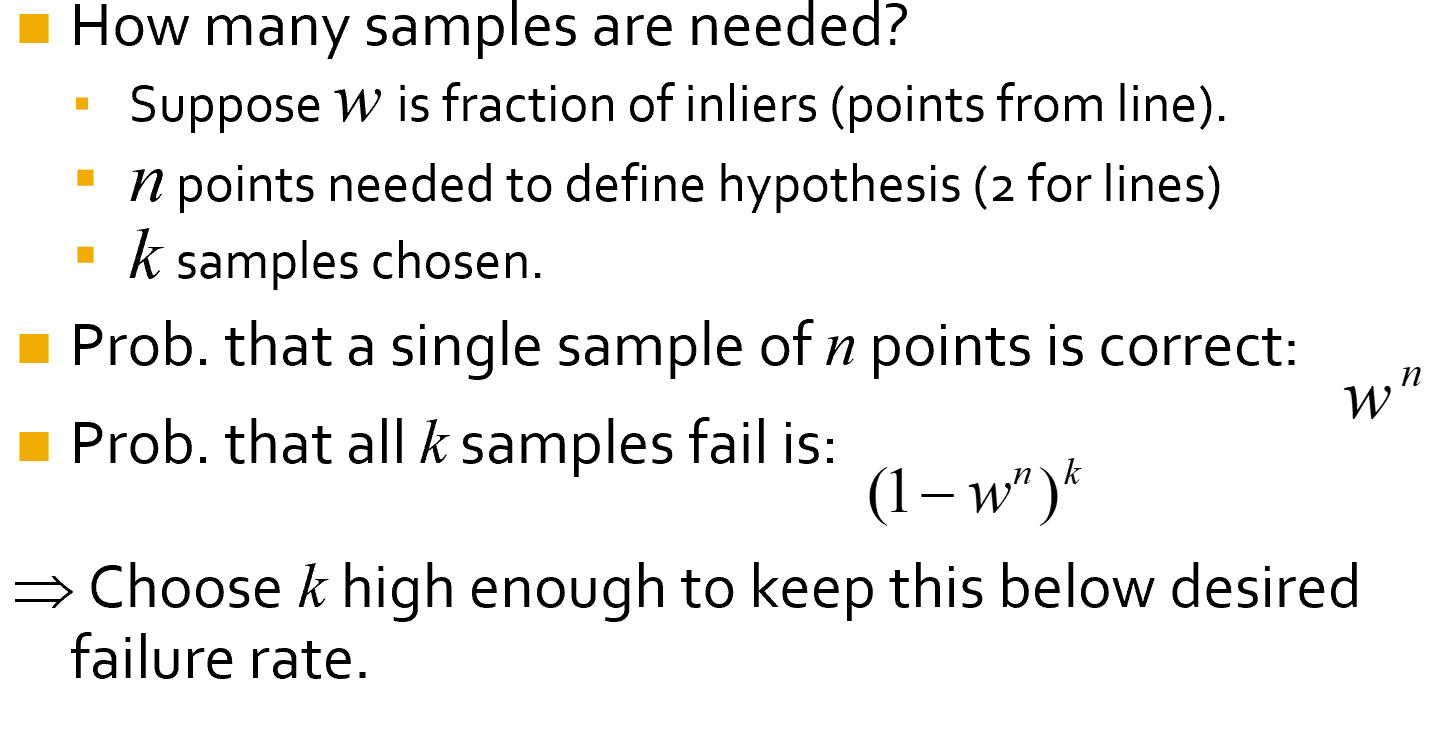
积分图



1. RANSAC



需要多少轮迭代？



1. 图像拼接的过程

Detect key points

Build the SIFT descriptors

Match SIFT descriptors

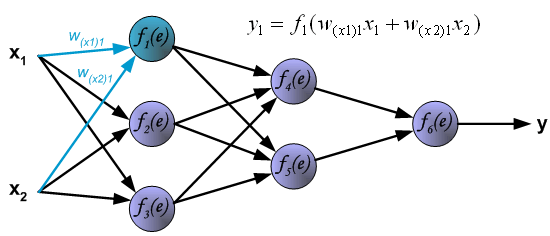
Fitting the transformation

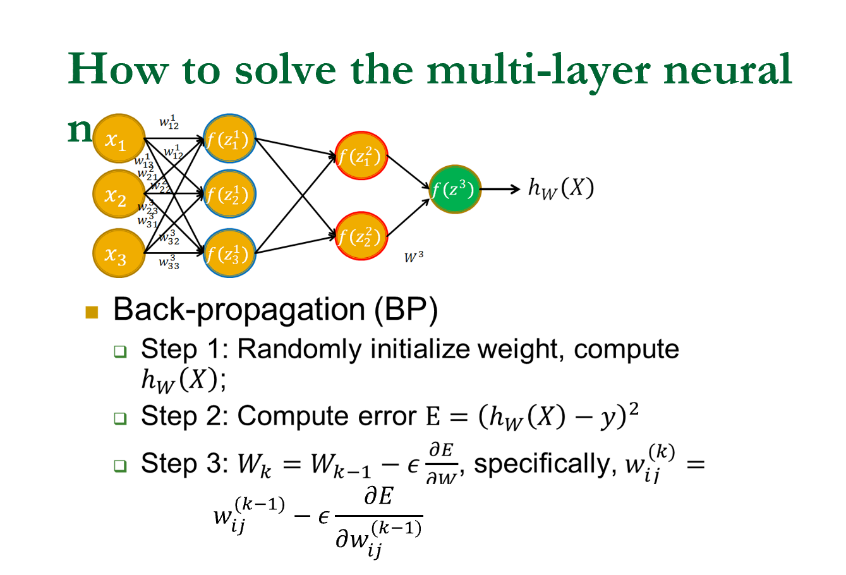
RANSAC

Image Blending

**Chap08 DL**

要会算多层神经网络







如何计算梯度

用一个batch B

学习率低，收敛慢，容易掉到坑里陷入局部最优；学习率高，可能使得收敛过程不稳定，来回震荡，一直不收敛

**PPT1**

数字图像的重要性 那一页 其他的不用仔细看

**PPT2**

图像表达二维数组/矩阵 每个像素用字节表示（灰1彩3）

图像格式 （常见的 BMP JPEG TIFF GIF PNG）编码方式：有损or无损，简单介绍

BMP格式 文件结构，具体size不用记，但要知道各部分及其功能 bitCount(1 4 8 16 24 32分别对应 1二值 8灰度 16伪彩色 真彩色 带阿尔法通道的彩色) 分辨率 物理尺寸上的像素

图像数据 倒立存储，每一行字节要是 4 的倍数

光圈大小的影响 以及光圈对景深的影响（为什么光圈越大，景深越浅？）

数码相机成像原理过程 背一遍

色彩分为彩色（RGB）和消色（灰度）

三原色 对应波长（R最长 B最短）

强调：感知的HSV 色调>饱和度>亮度（这里是优先级 优先级不等于感知能力、敏感度）

颜色空间（设备有关/无关） RGB CMY HSV; CIE XYZ,LAB,YUV 式子不用记

加色和减色,RGB加 CMY减 要知道字母代表什么

JPEG格式 了解压缩的基本策略及其优点 JPEG的缺点 具体编码方式不管

链式图像表达方式 Run Length Encoding 好好理解

**PPT3**

重点：如何得到二值图像（低中高级） 局部自适应+大津

大津算法推导、算法过程

形态学：四种操作（以膨胀为例，膨胀的物理意义，能模拟例子结果）

**PPT4**

重点：（基于对数）可视增强、直方图（均衡化）

Weber’s Law

直方图均衡化 必须掌握 推导过程、具体例子计算（可以用计算器）、特点带来的好处/坏处

直方图匹配三步、直方图变换

**PPT5**

几何变换举例子，简单的计算（公式不用记，变换矩阵会给）

插值 重点掌握最近邻插值、双线性插值

人脸Morph 表情映射的过程 以及建模（光照模型）

**PPT6**

一维卷积（例子，计算）

均值滤波（平滑滤波）、中值滤波（！） 拉普拉斯算子

**PPT7**

双边滤波，掌握思想，不用记式子（给结果 猜参数大小）

**PPT8**

快速双边滤波 了解 不考

引导滤波 式子怎么来的

保梯度（给图区分双边滤波和引导滤波）

优点

Sparse Norm Filter

**PPT9**

如何使用傅里叶变换（图像去噪 p89）

FFT 推导

幅值 相位哪个更重要？

特征检测：数学推导（一直到得到 H 矩阵）分析； 有几个不变性

Harris Laplace 怎么做的（各尺度Harris corner+拉普拉斯做尺度选择）

**PPT10**

SIFT 特征怎么提取的(basic idea, full version)

利用 bob 特征 理解一下就好

**PPT11**

SURF特征 提取方法（作为detector descriptor） 积分图

图像拼接完整步骤

RANSAC 算法过程 需要多少轮迭代

**PPT12**

DL 重点掌握多层神经网络的 BP （可能有计算）

卷积里 Pooling操作包含哪两方面

学习率高低的弊端

如何计算梯度更稳定（引入 batch B）