

2 数字图像基础

要想成功，就必须弄清楚基础问题。

——亚里士多德

本章主要内容

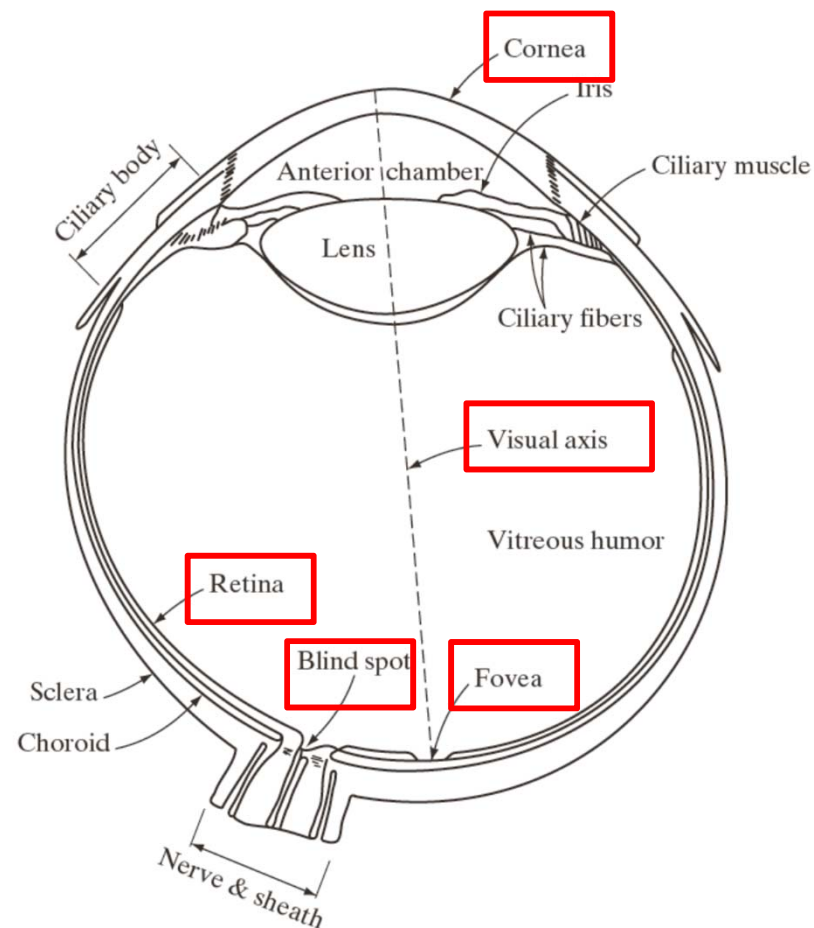
- 视觉感知要素
- 光和电磁波谱
- 图像感知和获取
- 图像采样和量化
- 像素间的基本关系
- 数字图像处理所用基本数学工具

为什么要先了解人的视觉感知要素？

- 人的**直觉**和**分析**在数字图像处理和计算机视觉等算法设计中起到重要作用；
- 基本了解人的视觉的形成与感知图像的机理和参数，便于和电子成像设备对比，明确努力方向；
- 实现并超越人类视觉的能力，这也是当今计算机视觉、人工智能努力的方向
- 从计算机视觉角度来说，借鉴人类视觉工作原理，开展类脑视觉计算，是一个重要的发展方向

人眼剖面结构

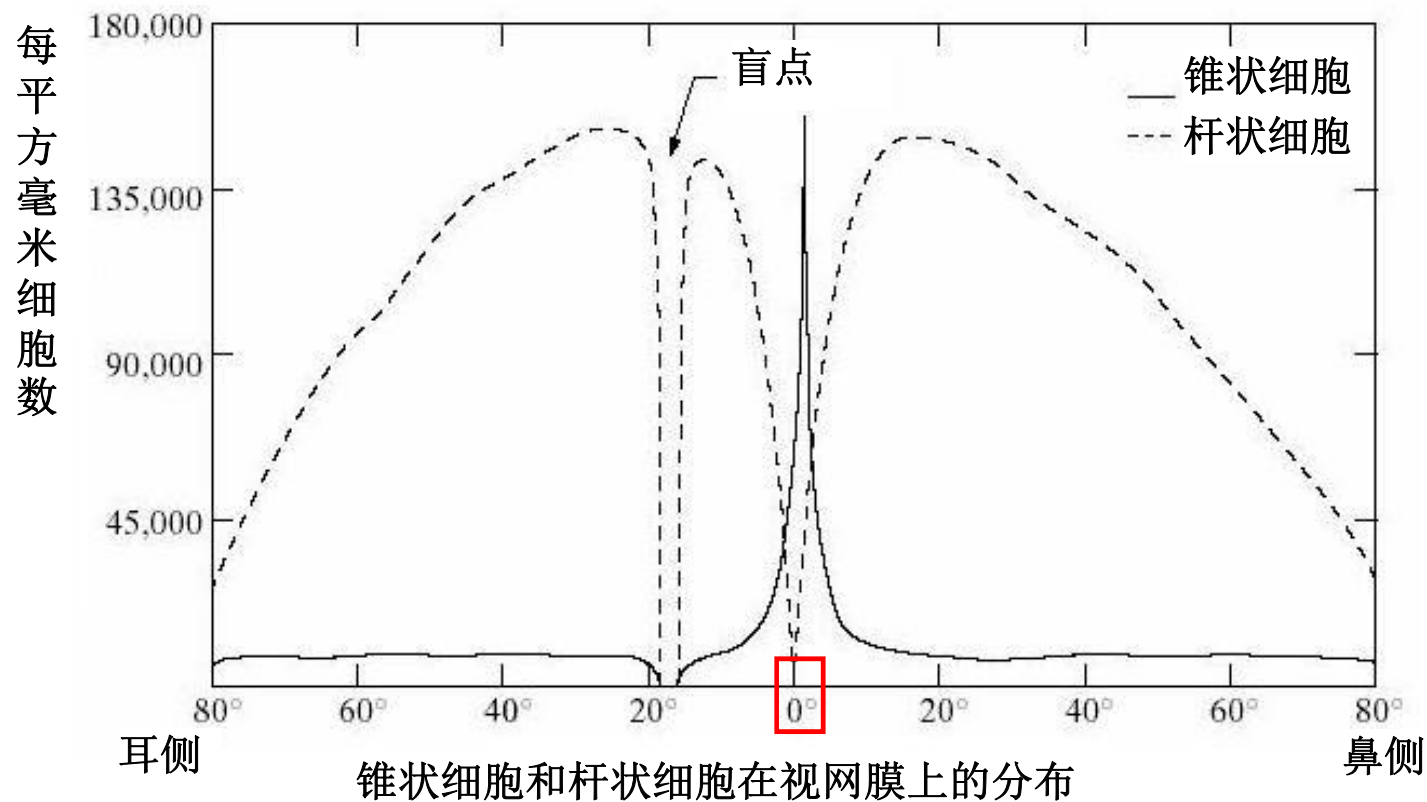
- 从外向内，有三层薄膜：
 - 角膜与巩膜外壳，脉络膜和视网膜
- 虹膜中间的开口为瞳孔
 - 直径2~8mm可变
- 中央凹
- 盲点



视觉感知要素 (1)

- 视网膜
 - 视网膜是眼睛的感觉神经层
 - 杆状体和锥状体是感光器件
 - 杆状体在低照度下提供视觉响应，且对形状敏感，暗视觉
 - 锥状体只在高照度下提供视觉，称为亮视觉

锥状细胞和杆状细胞在视网膜上的分布



中央凹，直径约1.5mm的圆形凹坑。锥状体数量近似为33万个，与640*480分辨率的相机靶面有些相当

眼睛中图像的形成

- 人眼焦距约为14~17mm

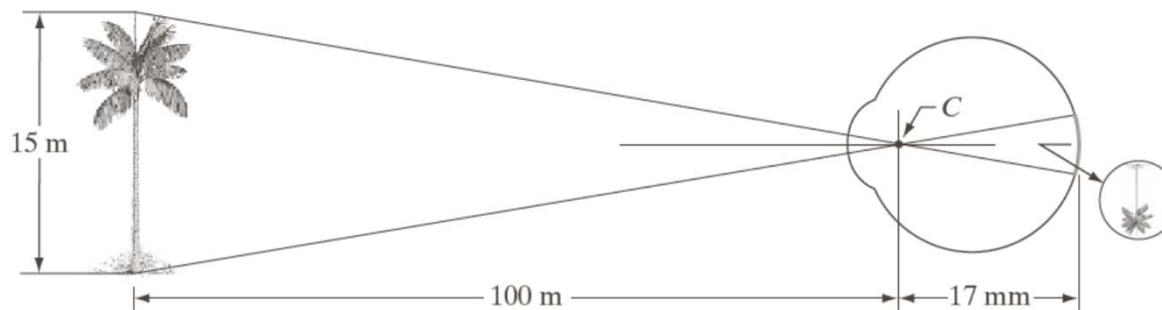
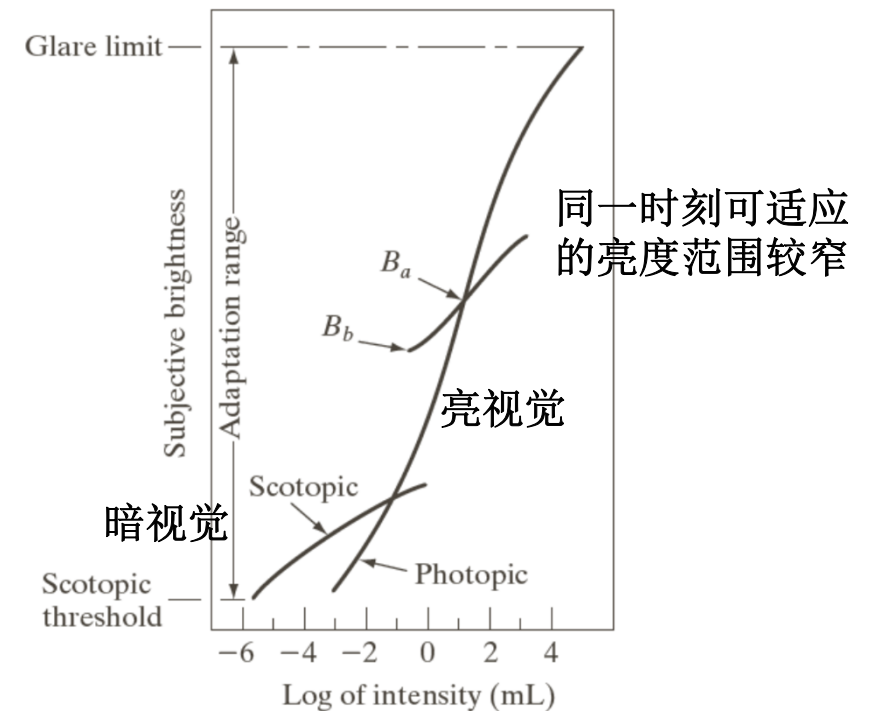


FIGURE 2.3
Graphical representation of the eye looking at a palm tree. Point C is the optical center of the lens.

亮度适应和辨别

- 人的视觉系统能够适应的光强度级别很宽，从暗到亮阈值约有 10^{10} 个量级。但同一时刻，可适应的亮度范围较窄。
- 主观亮度是进入人眼的光强的对数函数
- 视觉系统的当前灵敏度级别
 - 亮度适应级别，如图中 B_a



亮度适应和辨别

- 在任何特定的实验级别上，人眼辨别光强变化的能力。
- 韦伯比 $\Delta I_c / I$
- 低照度下，韦伯比较大，高照度下，韦伯比小，有更好辨别能力

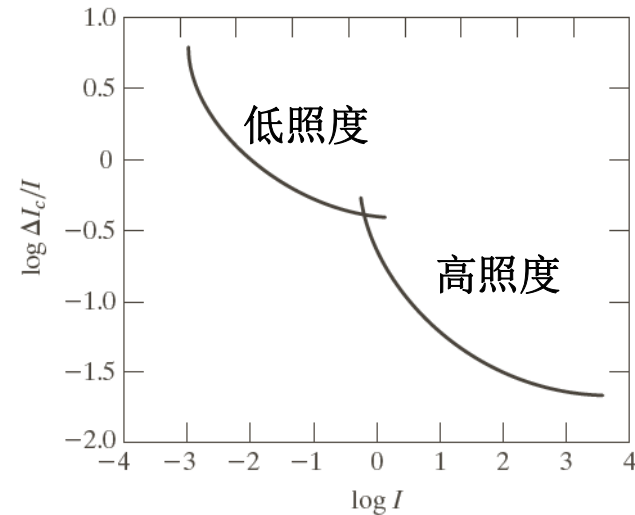
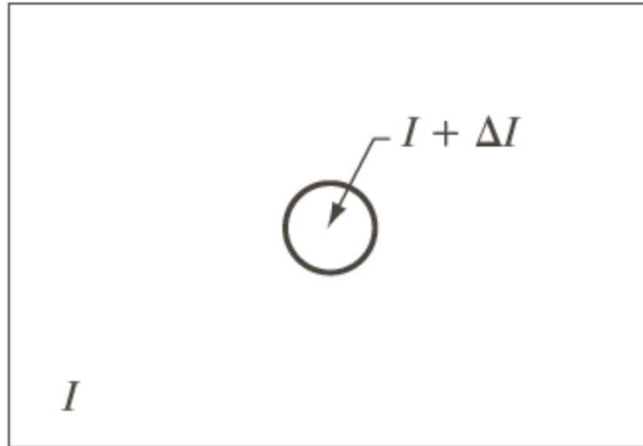


FIGURE 2.6
Typical Weber
ratio as a function
of intensity.

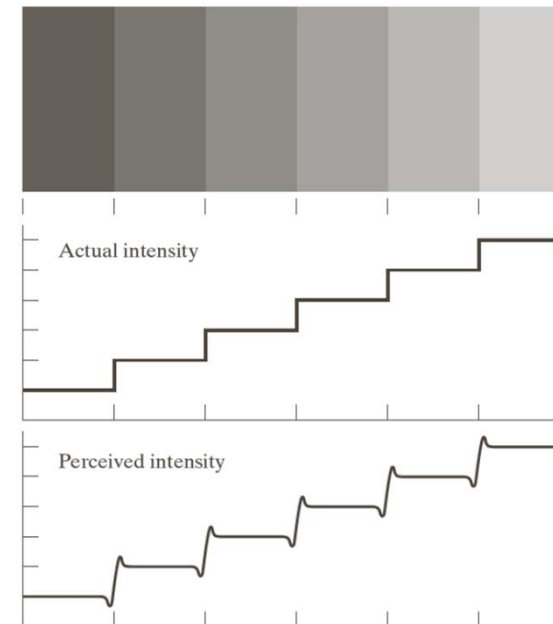
亮度适应和辨别

- 有两种现象表明，人眼感知亮度不是简单的强度的函数。
 - 1. 马赫带
 - 2. 同时对比



a b c

FIGURE 2.8 Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.



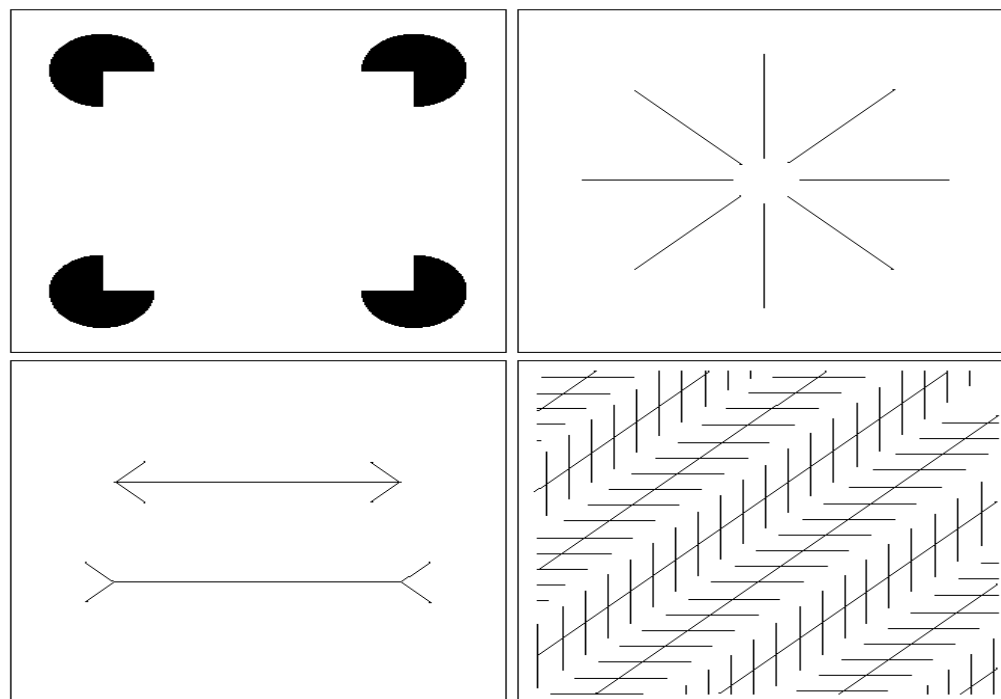
在观察均匀黑区域和均匀白区域形成的边界时，会感觉到边界附近暗区更暗，亮区更亮

人眼感知的错觉

- 人眼感知的另外一个现象——错觉

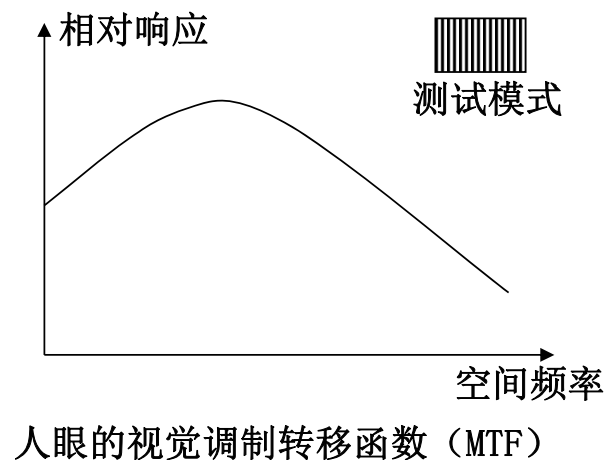
a b
c d

FIGURE 2.9 Some well-known optical illusions.

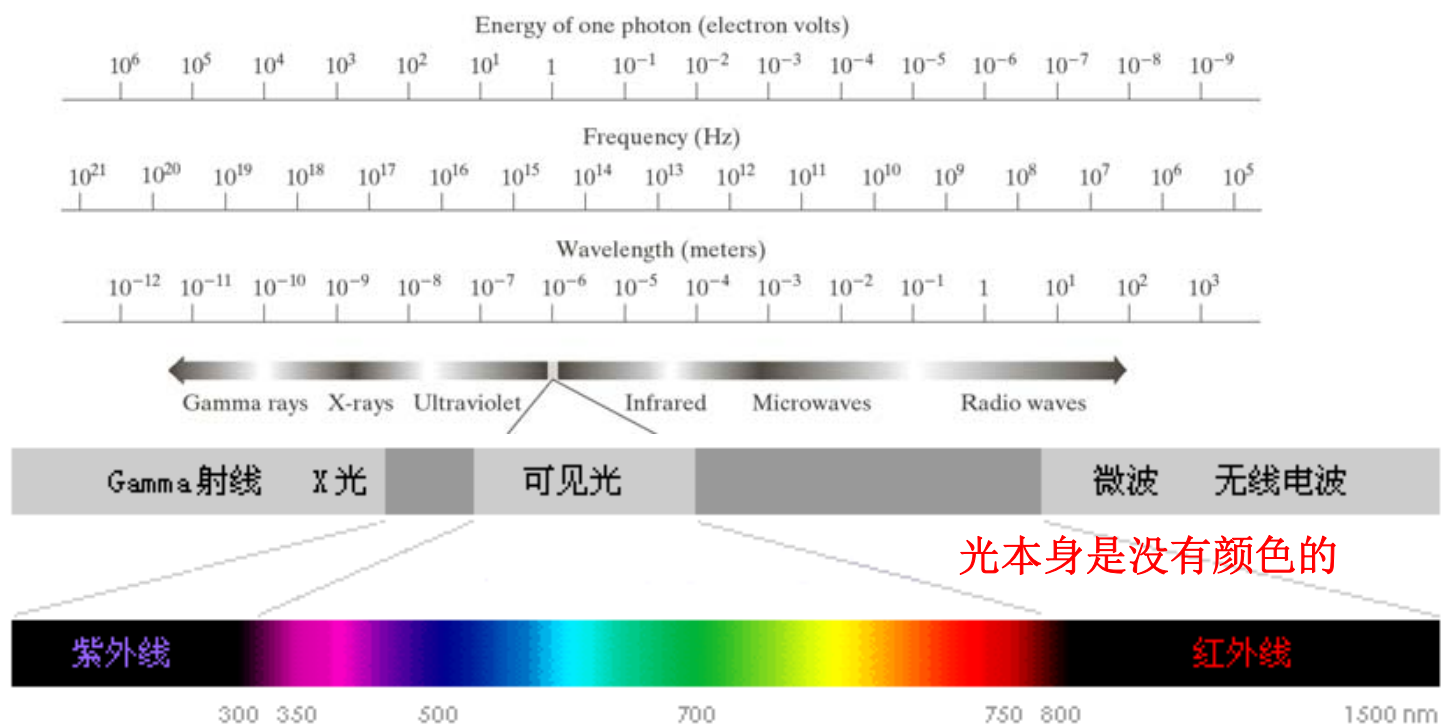


亮度适应和辨别

- 空间分辨率
 - 视野 :向左或右侧约 100° ，向上约 65° ，向下约 75°
 - 空间分辨率:**空间上**二相邻的视觉信号，人眼所刚能鉴别出二者存在的能力
 - 中央凹附近约 $1.5 \times 1.5 \text{mm}^2$ 的面积上，约有33万椎状体细胞，与中等分辨率CCD相当
 - 分辨率与照度有关



2. 2 光和电磁波谱



光本身是没有颜色的

可见光: 430nm~790nm

2. 2 光和电磁波谱

- 除了频率以外，有三个基本量用于描述彩色光源的质量：**发光强度**，**光通量**和**亮度**。
- **发光强度**是从光源流出的能量的总量，通常用**瓦特**来度量；
- **光通量**用**流明数**来度量（单位：**流明数**，lm），给出观察者从光源感受到的能量。如，远红外光光源发送出的光具有实际能量，但观察者很难感受到，光通量几乎是零。
- **亮度**是光感知的主观描绘子，实际上不能度量。
- 没有颜色的光称为单色光，唯一属性是**强度**或者灰度级。

- **照度**：照射在单位面积上的光通量：
 - 照度单位用lx(或lux,勒[克斯]), $1\text{lx}=1\text{lm}/\text{m}^2$ 。

一些典型场所的光照情况：

室内：

仓库 **20—75 Lux**

紧急通道 **30—75 Lux**

楼梯走廊 **75—200 Lux**

商店 **75—300 Lux**

办公室及接待室 **300—500 Lux**

银行 **200—1000 Lux**

会议厅 **300—1000 Lux**

室外：

晴天 **10000—1000000 Lux**

阴天 **100—10000 Lux**

黎明黄昏 **1—10 Lux**

满月之夜 **0.1—1 Lux**

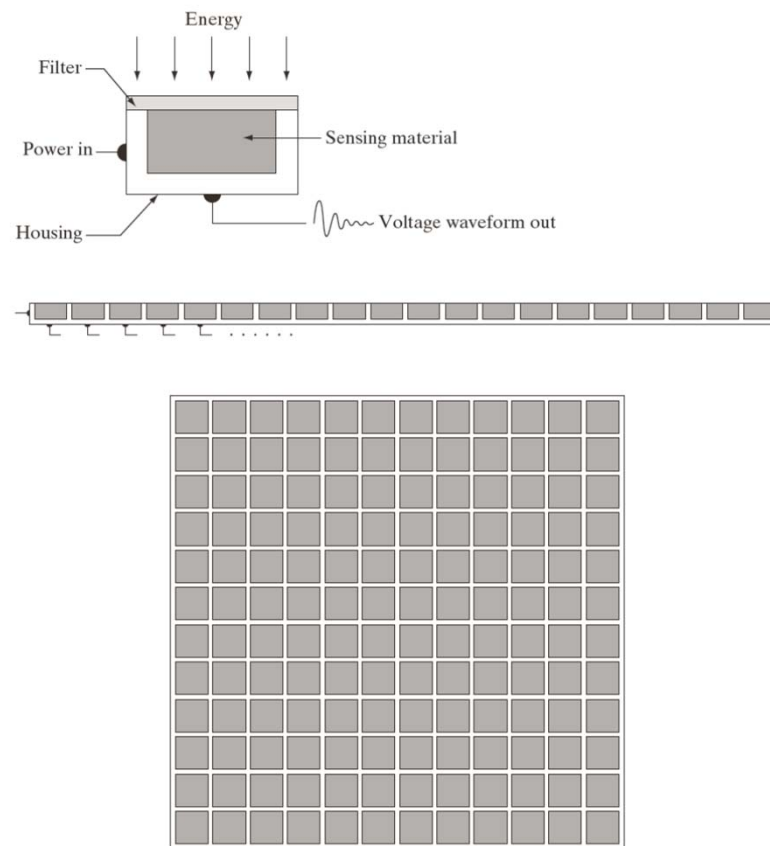
多云之夜 **0.01—0.1 Lux**

晴天时星光下 **0.001—0.01 Lux**

阴天时星光下 **0.0001—0.001 Lux**

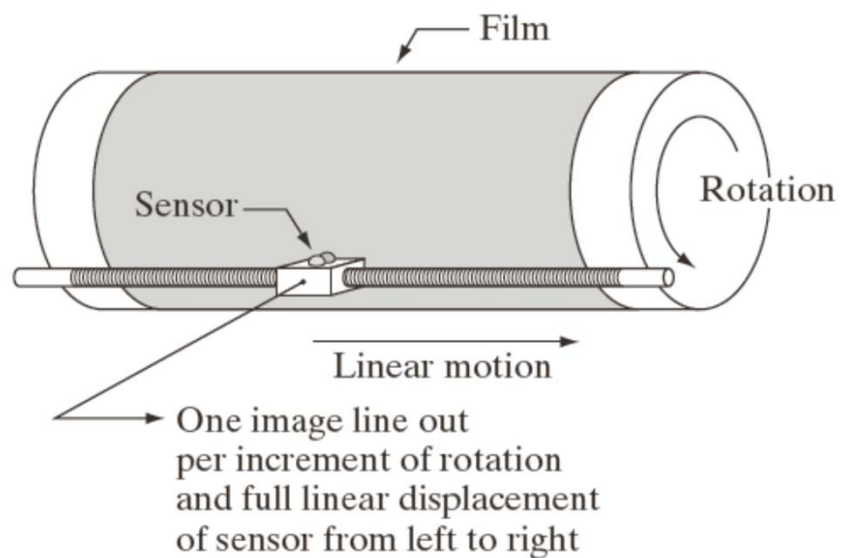
2.3 图像感知和获取

- 多数图像都是有“照射”源和形成图像的“场景”对光能的反射或者吸收产生的。
- 成像传感器配置。



图像感知和获取

- 使用单个传感器获取图像

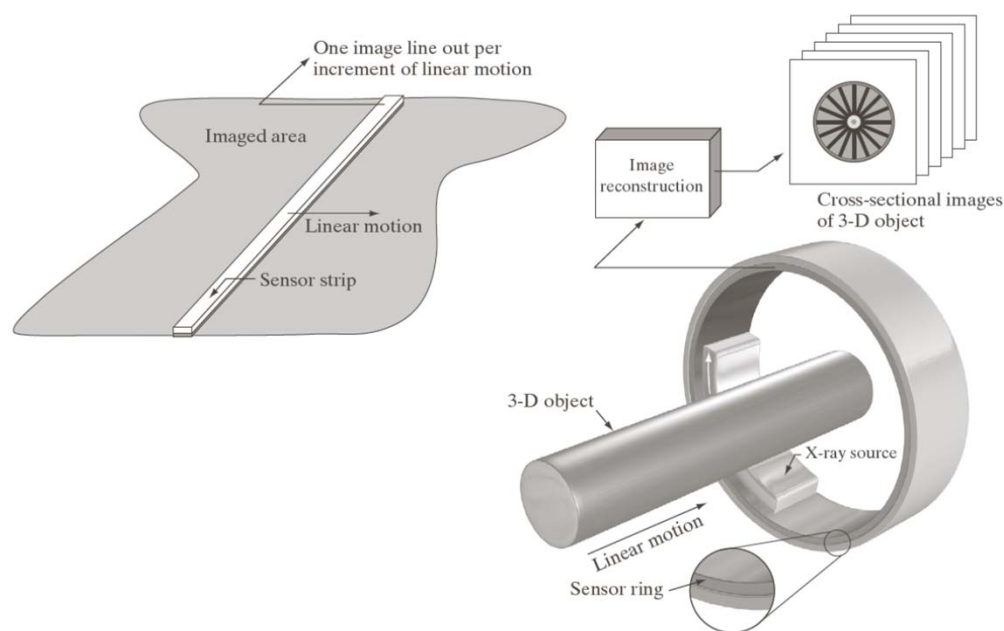


- 另一个例子：单线激光雷达



图像感知和获取

- 使用条带传感器获取图像
- 例子：平板扫描仪，磁共振成像（圆环扫描），多线激光雷达



图像感知和获取

- 使用传感器阵列获取图像（光被物体反射或者透射成像，不需要运动）

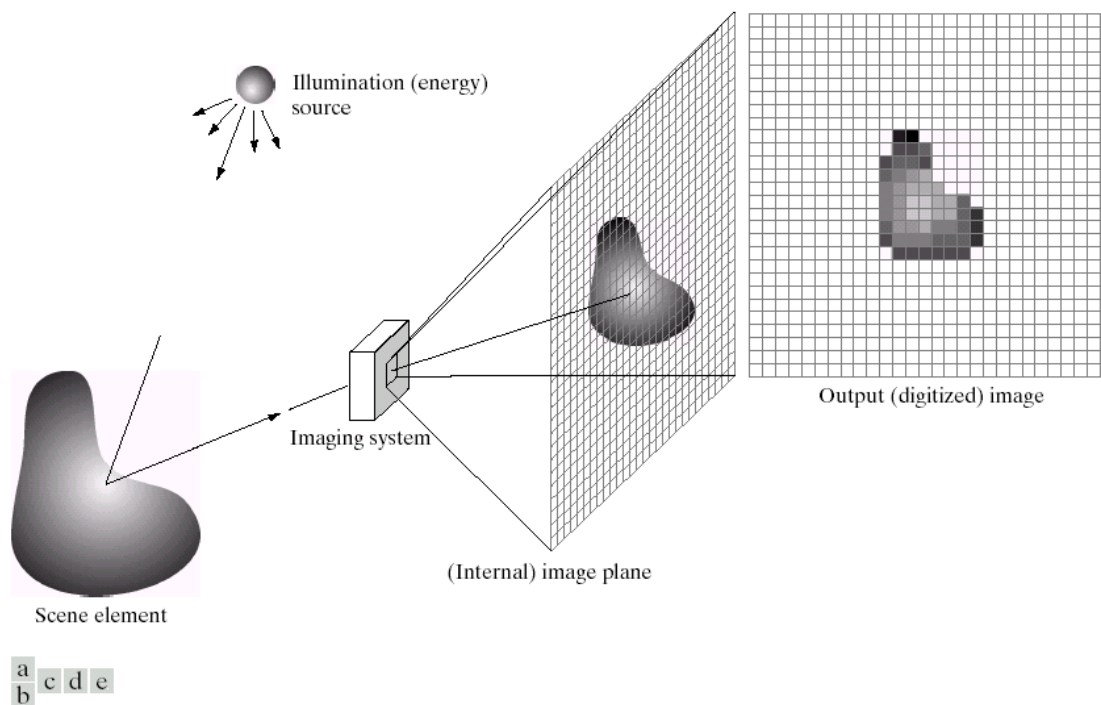


FIGURE 2.15 An example of the digital image acquisition process. (a) Energy (“illumination”) source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

2.3 图像感知和获取

- 简单的图像形成模型
- 用一个二维亮度函数 $f(x,y)$ 表示图像，则 $0 < f(x,y) < \infty$
- 对自身不发光的景物： $f(x,y) = i(x,y)r(x,y)$
 - $i(x,y)$ 是入射分量或者照度， $0 < i(x,y) < \infty$
 - $r(x,y)$ 为反射系数， $0 < r(x,y) < 1$
 - $r(x,y)$ 的典型值：黑天鹅绒0.01，不锈钢0.65，白墙0.8，雪0.93.

三维成像几何

- 三维空间中的一点 (X, Y, Z) ，在像平面上成像为点 (x, y)

则：

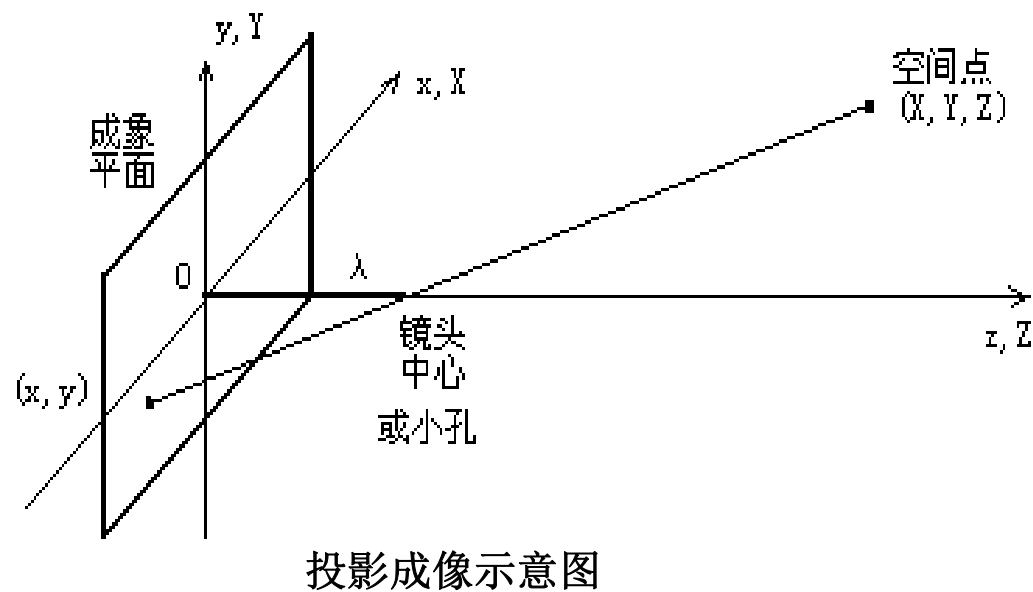
$$\frac{-x}{X} = \frac{\lambda}{Z - \lambda}$$

$$\frac{-y}{Y} = \frac{\lambda}{Z - \lambda}$$

故

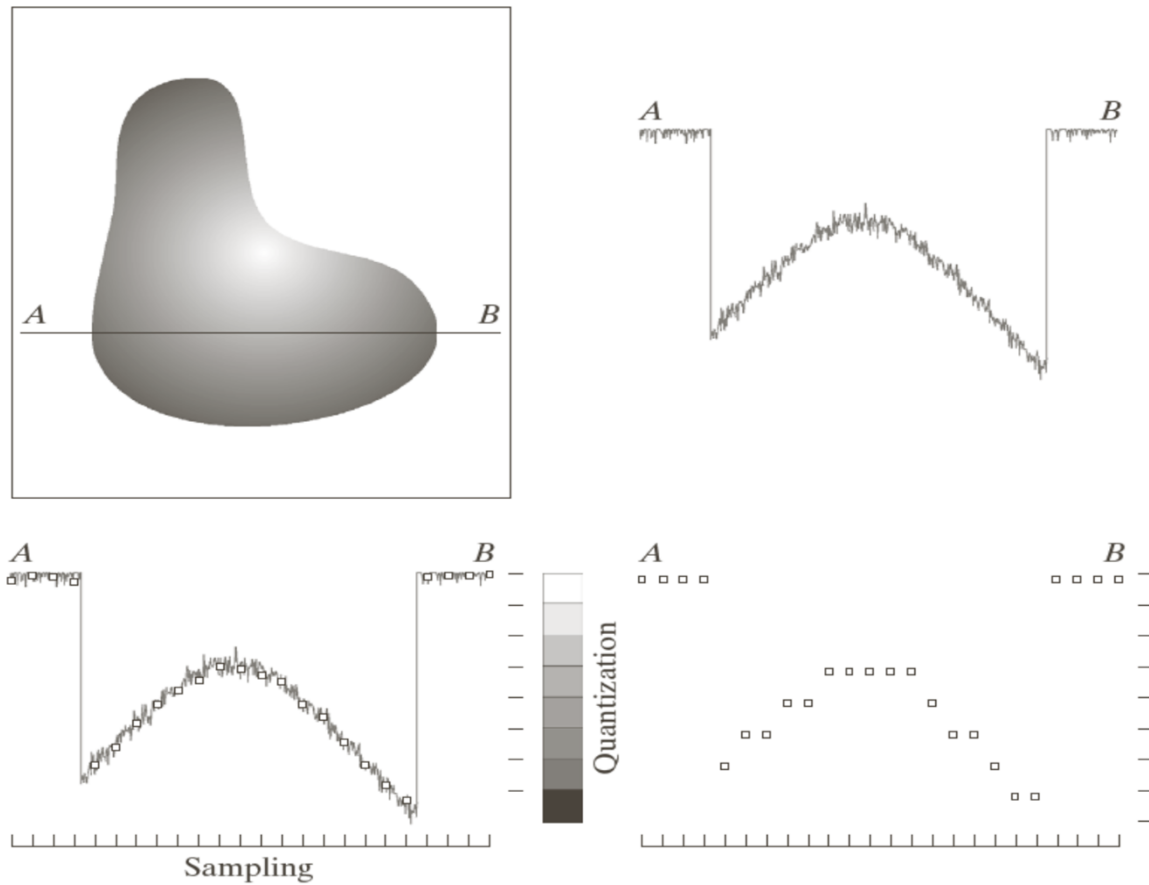
$$x = \frac{\lambda X}{\lambda - Z}$$

$$y = \frac{\lambda Y}{\lambda - Z}$$



2.4 图像取样和量化

- 基本概念



a b
c d

FIGURE 2.16
Generating a digital image.
(a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

2.4 图像取样和量化

- 数字图像的质量在很大程度上取决于取样和量化中所用的**样本数**和**灰度级**。

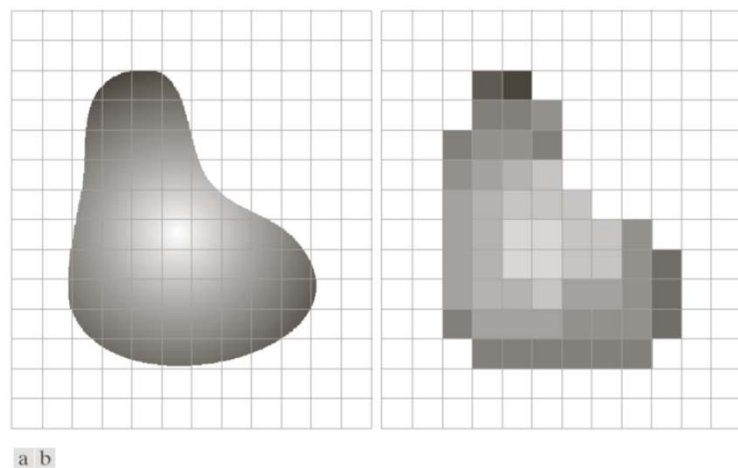
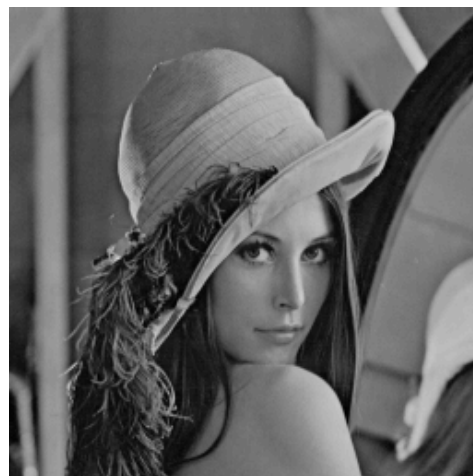


FIGURE 2.17 (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.



256级



64级



16级



8级



4级
第二章



2级

量化误差峰值信噪比

- 信噪比：信号峰峰值平方与误差的均方值之比
- 对于量化误差 Δ ，如果量化间隔足够小，可近似认为量化误差在 $[-0.5\Delta, +0.5\Delta]$ 之间均匀分布，则量化噪声的方差为

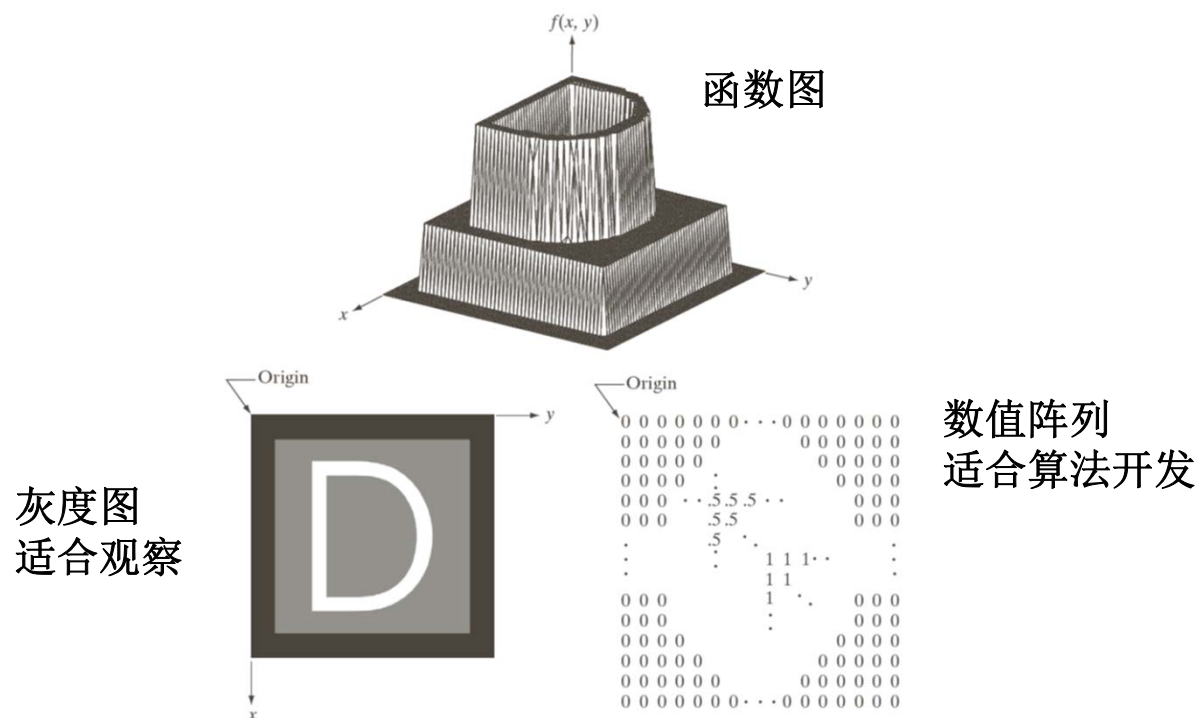
$$\sigma_n^2 = \frac{\Delta^2}{12} = \frac{2^{-2n} A^2}{12},$$

$$SNR = 10 \log_{10} \left(\frac{A^2}{\sigma_n^2} \right) = 10 \log_{10} (12 \times 2^{2n}) = 6.02n + 10.8 \text{ dB}$$

量化位数每增加1位，信噪比提升6db。

数字图像表示

- 数字图像的三种表示形式，后两种最有用。



数字图像表示

■ 矩阵表示

$$f(x, y) = \begin{bmatrix} f_{00} & f_{01} & \cdots & f_{0,N-1} \\ f_{10} & f_{11} & \cdots & f_{1,N-1} \\ \cdots & \cdots & \cdots & \cdots \\ f_{M-1,0} & f_{M-1,1} & \cdots & f_{M-1,N-1} \end{bmatrix}$$

f_{ik} 是图像的像元（或叫像素，像点），数字化图像量化为一个坐标和幅值都是整数的二维函数。

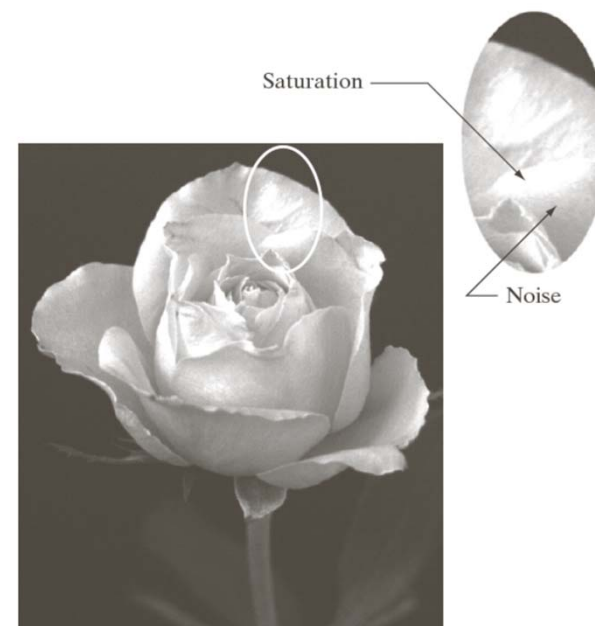
■ 像素值有非负性和有界性

■ 图像的能量也是有界的：

$$E = \sum_{j=1}^M \sum_{k=1}^N f_{ik}^2 < \text{上限值}$$

数字图像表示

- 灰度级一般取为2的整数次幂， $L=2^k$
- 图像系统的**动态范围**定义为系统中最大可度量灰度与最小可检测灰度之比。
- 实际中，上限取决于饱和度，下限取决于噪声。
- 图像的对比度：
 - 最高和最低灰度级之间的灰度差



- 不同分辨率和量化技术对应的存储图像的比特数。
若方形图像， 比特数 $b=N^2k$

TABLE 2.1

Number of storage bits for various values of N and k .

N/k	1 ($L = 2$)	2 ($L = 4$)	3 ($L = 8$)	4 ($L = 16$)	5 ($L = 32$)	6 ($L = 64$)	7 ($L = 128$)	8 ($L = 256$)
32	1,024	2,048	3,072	4,096	5,120	6,144	7,168	8,192
64	4,096	8,192	12,288	16,384	20,480	24,576	28,672	32,768
128	16,384	32,768	49,152	65,536	81,920	98,304	114,688	131,072
256	65,536	131,072	196,608	262,144	327,680	393,216	458,752	524,288
512	262,144	524,288	786,432	1,048,576	1,310,720	1,572,864	1,835,008	2,097,152
1024	1,048,576	2,097,152	3,145,728	4,194,304	5,242,880	6,291,456	7,340,032	8,388,608
2048	4,194,304	8,388,608	12,582,912	16,777,216	20,971,520	25,165,824	29,369,128	33,554,432
4096	16,777,216	33,554,432	50,331,648	67,108,864	83,886,080	100,663,296	117,440,512	134,217,728
8192	67,108,864	134,217,728	201,326,592	268,435,456	335,544,320	402,653,184	469,762,048	536,870,912

图像空间和灰度分辨率

- 空间分辨率
 - 每单位距离线对数
 - 每单位距离点数（印刷业用dpi）
 - 必须针对空间单位来规定才有意义，比如印刷业。
- 右图例子中，分辨率分别为1250dpi, 300dpi, 150dpi, 72dpi.



图像空间和灰度分辨率

- 灰度分辨率，一般用8比特，过高视觉上无法分辨。
- 例：从256级分别降到128, 64, 32级灰度

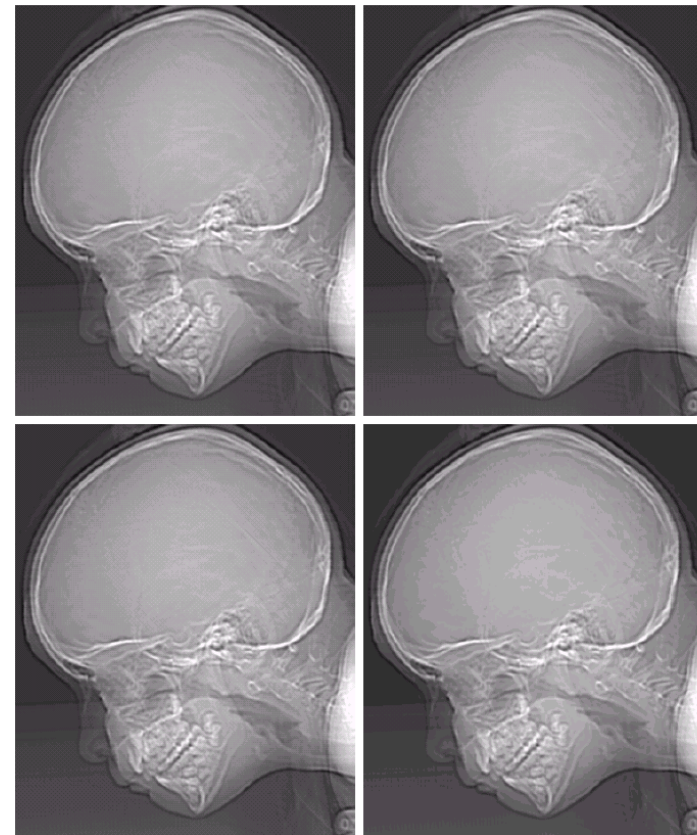
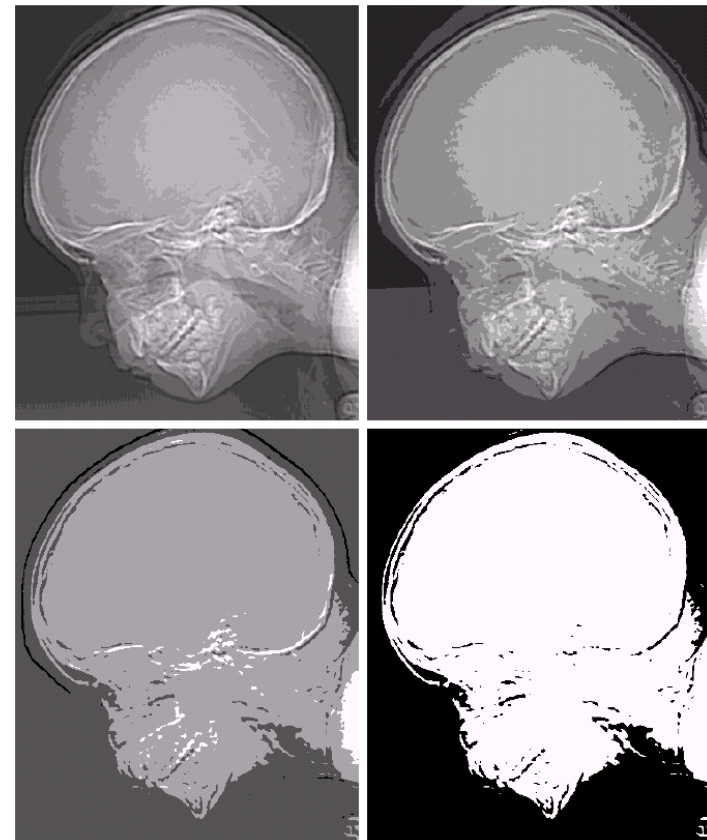


FIGURE 2.21
(a) 452×374 ,
256-level image.
(b)–(d) Image
displayed in 128,
64, and 32 gray
levels, while
keeping the
spatial resolution
constant.

- 例：灰度级
继续分别降到16, 8, 4, 2
个灰度级。注意伪轮
廓的出现。

e f
g h
FIGURE 2.21
(Continued)
(e)–(h) Image
displayed in 16, 8,
4, and 2 gray
levels. (Original
courtesy of
Dr. David
R. Pickens,
Department of
Radiology &
Radiological
Sciences,
Vanderbilt
University
Medical Center.)



分辨率和灰度级共同对视觉效果的影响

- 对具有不同细节数量的图像进行的测试表明，对于有大量细节的图像，只需要较少的灰度级。此时获得的图像主观感受质量是近似的。



a b c

FIGURE 2.22 (a) Image with a low level of detail. (b) Image with a medium level of detail. (c) Image with a relatively large amount of detail. (Image (b) courtesy of the Massachusetts Institute of Technology.)

三幅图细节数逐步由少变多

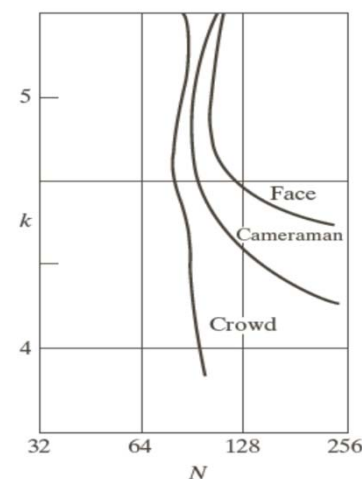
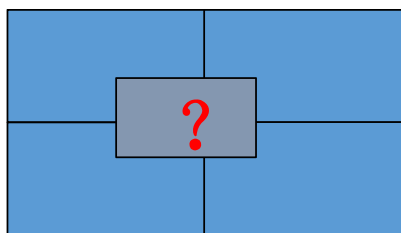


FIGURE 2.23
Typical
isopreference
curves for the
three types of
images in
Fig. 2.22.

N-K平面中的等偏爱曲线

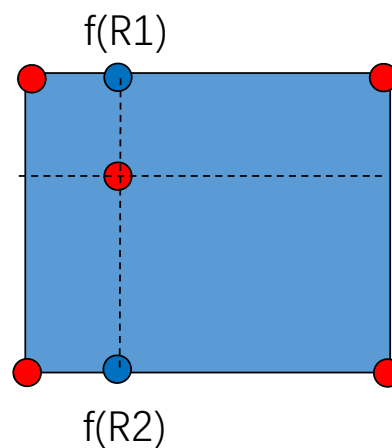
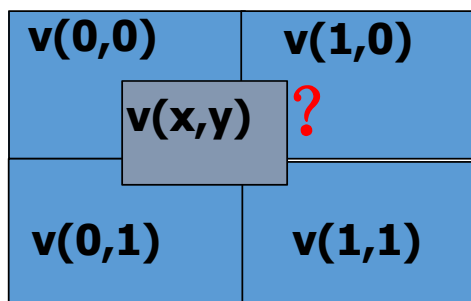
图像内插

- 内插是图像放大、收缩、旋转和几何校正等任务中广泛使用的基本工具。是一种基本的图像重取样方法。
- 是用已知数据来估计未知位置的数值的处理



图像内插

- 最近邻内插
 - 把原图像中最近邻的灰度赋给每个内插的新位置
- 双线性内插(用4个最近邻)
 - $V(x,y)=ax+by+cxy+d$

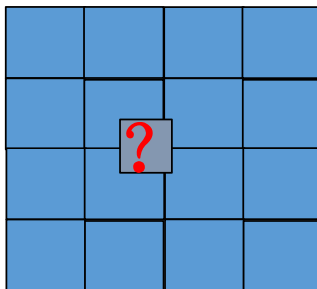


注意：实质上是一种非线性内插方法

图像内插

- 双三次内插（用16个最近邻，保持细节更好）

$$v(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j$$



图像内插

(1) **1250dpi**缩到**72dpi**后重新放大，分别用最近邻、双线性内插和双三次方内插

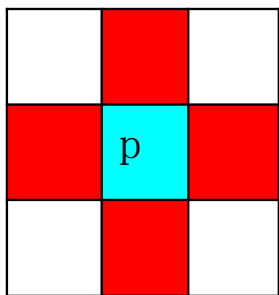
(2) **1250dpi**缩到**150dpi**后重新放大，分别用最近邻、双线性内插和双三次方内插



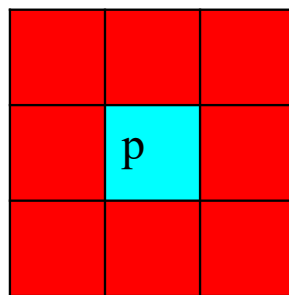
FIGURE 2.24 (a) Image reduced to 72 dpi and zoomed back to its original size (3692×2812 pixels) using nearest neighbor interpolation. This figure is the same as Fig. 2.20(d). (b) Image shrunk and zoomed using bilinear interpolation. (c) Same as (b) but using bicubic interpolation. (d)–(f) Same sequence, but shrinking down to 150 dpi instead of 72 dpi [Fig. 2.24(d) is the same as Fig. 2.20(c)]. Compare Figs. 2.24(e) and (f), especially the latter, with the original image in Fig. 2.20(a).

2.5 像素间的基本关系

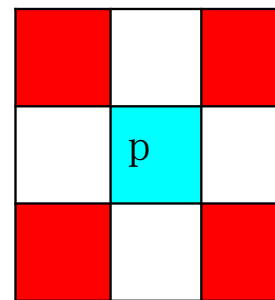
- 一 . 邻域



p 的4-邻域 $N_4(p)$



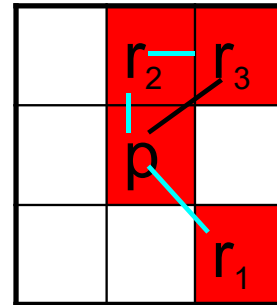
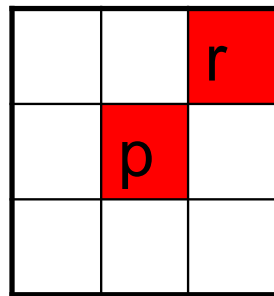
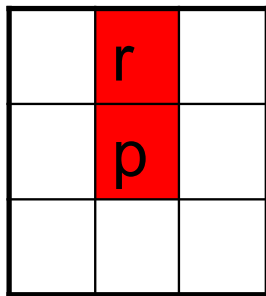
p 的8-邻域 $N_8(p)$



**p 的4个对角近邻
像素记为 $N_D(p)$**

邻接性

- 二. 邻接性
 - 相同性质的像素相邻 —— 称它们是邻接的
 - 设像素p和r有相同的性质，定义如下3种连接：
 - (1) 4-连接：r在 $N_4(p)$ 中，则p和r为4-邻接；
 - (2) 8-连接：r在 $N_8(p)$ 中，则p和r为8-邻接；
 - (3) m-连接（混合连接）：若r和p满足
 - r在 $N_4(p)$ 中
 - 或r在 $N_D(p)$ 中，且p与r没有共同的4连接像素则称r和p是m-邻接的。



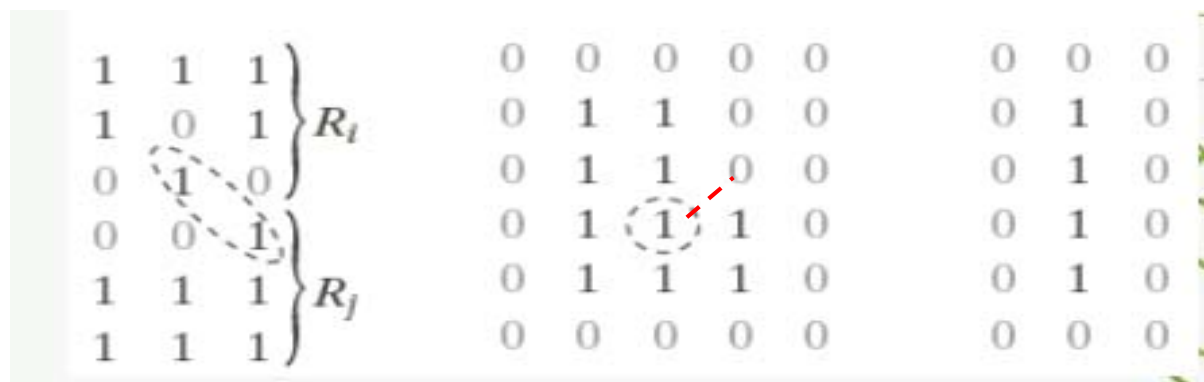
混合邻接，即：
要么4邻接，要么对角邻接，都有的话，只算4邻接通路，为了使计算通路长度时有唯一性。

连通性

- 若两个具有相同性质的像素能找到两两邻接的一条通路, 则称它们是连通的
- 连通关系是一种等价关系
 - 对称性;
 - 传递性。
 - 集合中的每个像素都与自身连通

像素间的基本关系

- 通路 与 闭合通路
- 如果图像中的一个像素子集S的全部像素之间存在着一个通路，则称S是连通的，构成一个连通集，或者称为一个区域。
- 连通区域与边界（边界定义为与区域R的补集中的点邻接）



- **外边界** 对应于背景边界
 - 很多算法为了保证形成一个闭合通路，通常采用外边界

采用**8**邻接，则圆圈中的“**1**”为前景区域边界的一部分

“**1**”为区域边界
 没有内部闭合通路
 （**内边界**）

边缘与边界

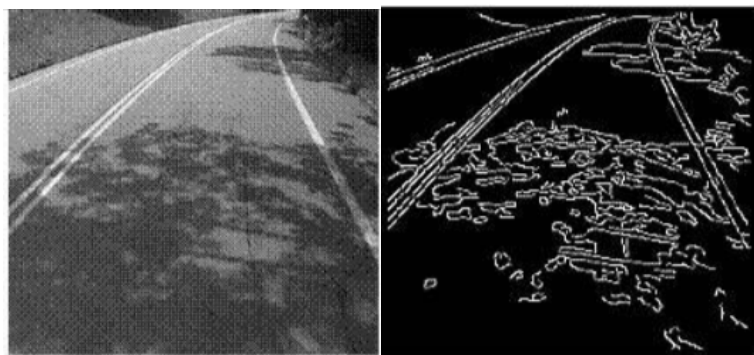
- 边缘

- 边缘由某些具有超过预先设定阈值的导数值的像素形成，是局部概念
- 灰度不连续的像素连接而成

- 边界

- 一个（分割后）有限区域的边界形成一条闭合通路，是整体的概念。

边缘



边界



像素间距离

- 像素间距离

- 距离函数须满足三个条件:

- $D(p,q) \geq 0$, ($D(p,q)=0$ iff $p=q$) 非负性;
 - $D(p,q)=D(q,p)$ 对称性;
 - $D(p,r) \leq D(p,q)+D(q,r)$ 相当于三角形的二边之和大于第三边。

- 常用的欧氏 (Euclidean) 距离

$$D_e(p, q) = \left[(x-s)^2 + (y-t)^2 \right]^{1/2}$$

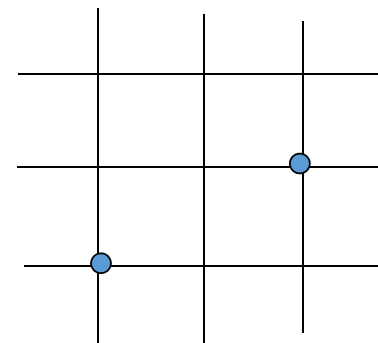
- 也可用其它形式的距离

- 街区 (city-block) 距离:

$$D_4(p,q)=|x-s|+|y-t|$$

- D8 (棋盘, chessboard) 距离:

$$D_8(p,q) = \max(|x-s|, |y-t|)$$



2.6 相关数学工具介绍

- 阵列与矩阵操作

阵列操作 $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}$

矩阵操作 $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$

线性与非线性操作

- 线性与非线性操作——图像处理方法最主要的分类之一
 - 线性满足加性和同质性
- 线性操作具有很大应用价值，但也存在很多非线性操作，如最大算子，中值滤波等。

$$H[f(x, y)] = g(x, y)$$

$$\begin{aligned} H[a_i f_i(x, y) + a_j f_j(x, y)] &= a_i H[f_i(x, y)] + a_j H[f_j(x, y)] \\ &= a_i g_i(x, y) + a_j g_j(x, y) \end{aligned}$$

线性与非线性操作

- 非线性操作例子：考虑如下两个图像，求最大算子，设 $a_1=1$, $a_2=-1$

$$f_1 = \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} \text{ 和 } f_2 = \begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix}$$

- 左侧

$$\max \left\{ (1) \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} + (-1) \begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix} \right\} = \max \left\{ \begin{bmatrix} -6 & -3 \\ -2 & -4 \end{bmatrix} \right\} = -2$$

- 右侧

$$(1) \max \left\{ \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix} \right\} + (-1) \max \left\{ \begin{bmatrix} 6 & 5 \\ 4 & 7 \end{bmatrix} \right\} = 3 + (-1)7 = -4$$

算术操作

- 是阵列操作，涉及对应像素的加，减，乘，除操作。
- 不改变图像的大小

$$s(x, y) = f(x, y) + g(x, y)$$

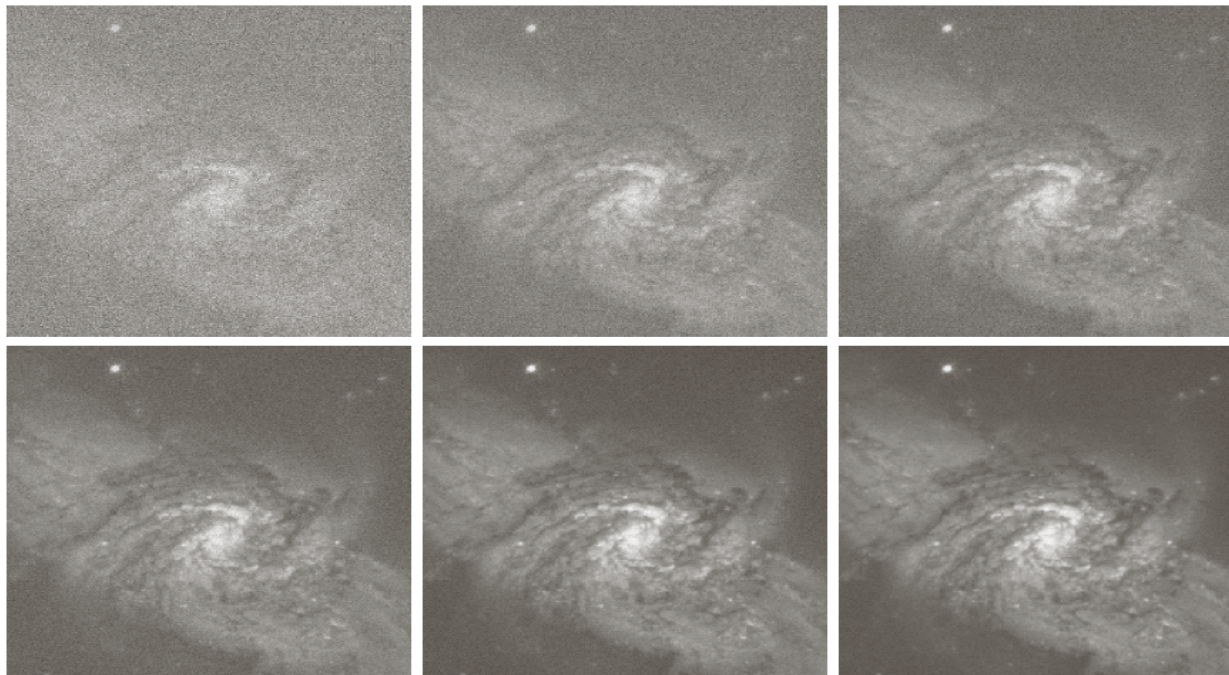
$$d(x, y) = f(x, y) - g(x, y)$$

$$p(x, y) = f(x, y) \times g(x, y)$$

$$v(x, y) = f(x, y) \div g(x, y)$$

算术操作

- 针对降噪的带噪图像相加平均



a	b	c
d	e	f

FIGURE 2.26 (a) Image of Galaxy Pair NGC 3314 corrupted by additive Gaussian noise. (b)–(f) Results of averaging 5, 10, 20, 50, and 100 noisy images, respectively. (Original image courtesy of NASA.)

- 针对降噪的带噪图像相加平均原理:

$$g(x, y) = f(x, y) + \eta(x, y)$$

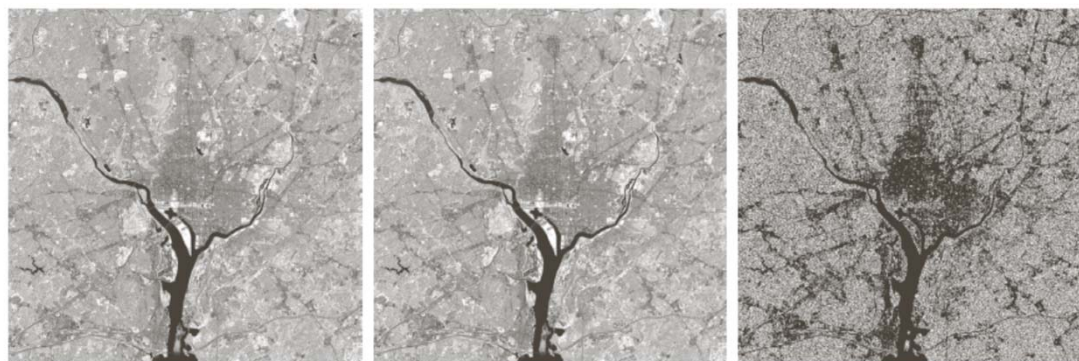
$$\bar{g}(x, y) = \frac{1}{K} \sum_{i=1}^K g_i(x, y)$$

$$E\{\bar{g}(x, y)\} = f(x, y)$$

$$\sigma_{\bar{g}(x, y)}^2 = \frac{1}{K} \sigma_{\eta(x, y)}^2$$

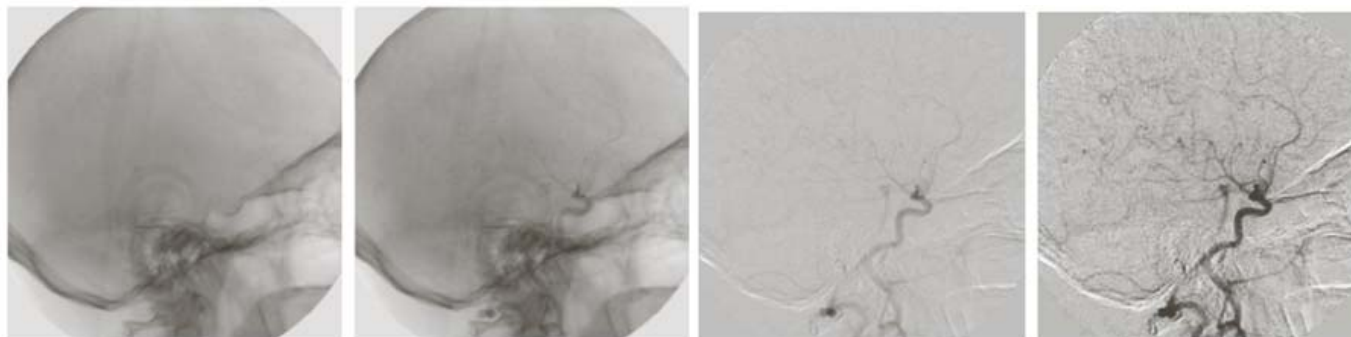
$$\sigma_{\bar{g}(x, y)} = \frac{1}{\sqrt{K}} \sigma_{\eta(x, y)}$$

图像相减用于增强差别



a b c

FIGURE 2.27 (a) Infrared image of the Washington, D.C. area. (b) Image obtained by setting to zero the least significant bit of every pixel in (a). (c) Difference of the two images, scaled to the range $[0, 255]$ for clarity.



模版

活体

差值

增强后的差值

使用图像相乘或相除来校正阴影

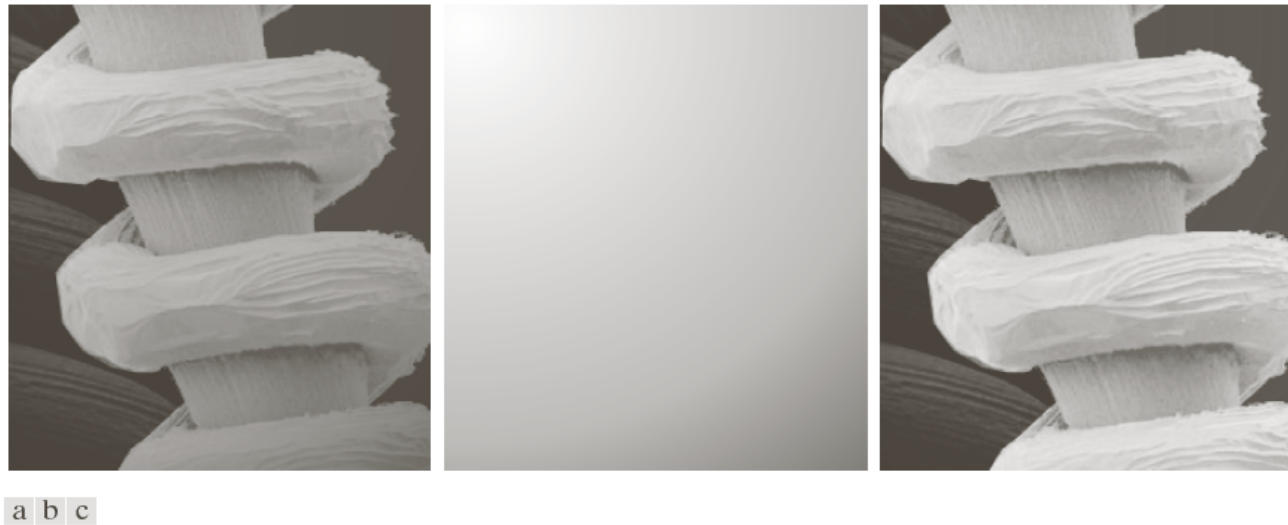


FIGURE 2.29 Shading correction. (a) Shaded SEM image of a tungsten filament and support, magnified approximately 130 times. (b) The shading pattern. (c) Product of (a) by the reciprocal of (b). (Original image courtesy of Mr. Michael Shaffer, Department of Geological Sciences, University of Oregon, Eugene.)

原图、阴影模式、相除结果

$$g(x, y) = f(x, y)h(x, y)$$

- 利用模版相乘做ROI提取操作



FIGURE 2.30 (a) Digital dental X-ray image. (b) ROI mask for isolating teeth with fillings (white corresponds to 1 and black corresponds to 0). (c) Product of (a) and (b).

图像算术操作中的归一化

- 算术操作后会出现很多负值以及超出原图灰度范围的灰度值，需要归一化

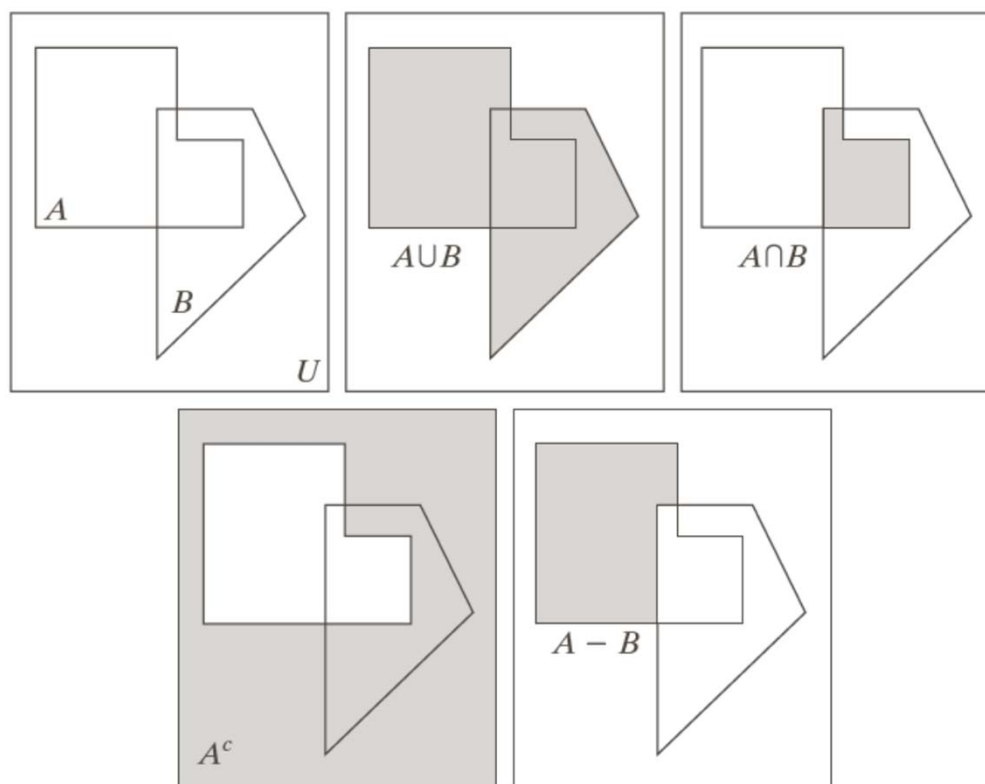
$$f_m = f - \min(f)$$

$$f_s = K [f_m / \max(f_m)]$$

当处理**8bit**图像时，**K=255**

集合和逻辑操作

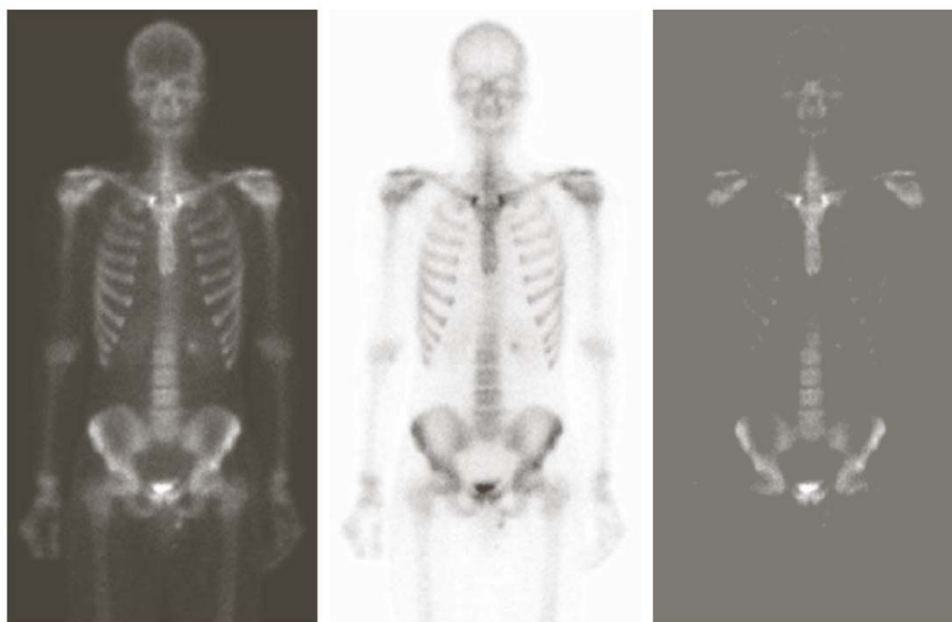
- 基本集合操作（图像二值，集合元素为坐标）



灰度图像的集合操作

- 由空间相应元素对间的最大灰度形成的阵列

$$A \cup B = \left\{ \max_z(a, b) | a \in A, b \in B \right\}$$



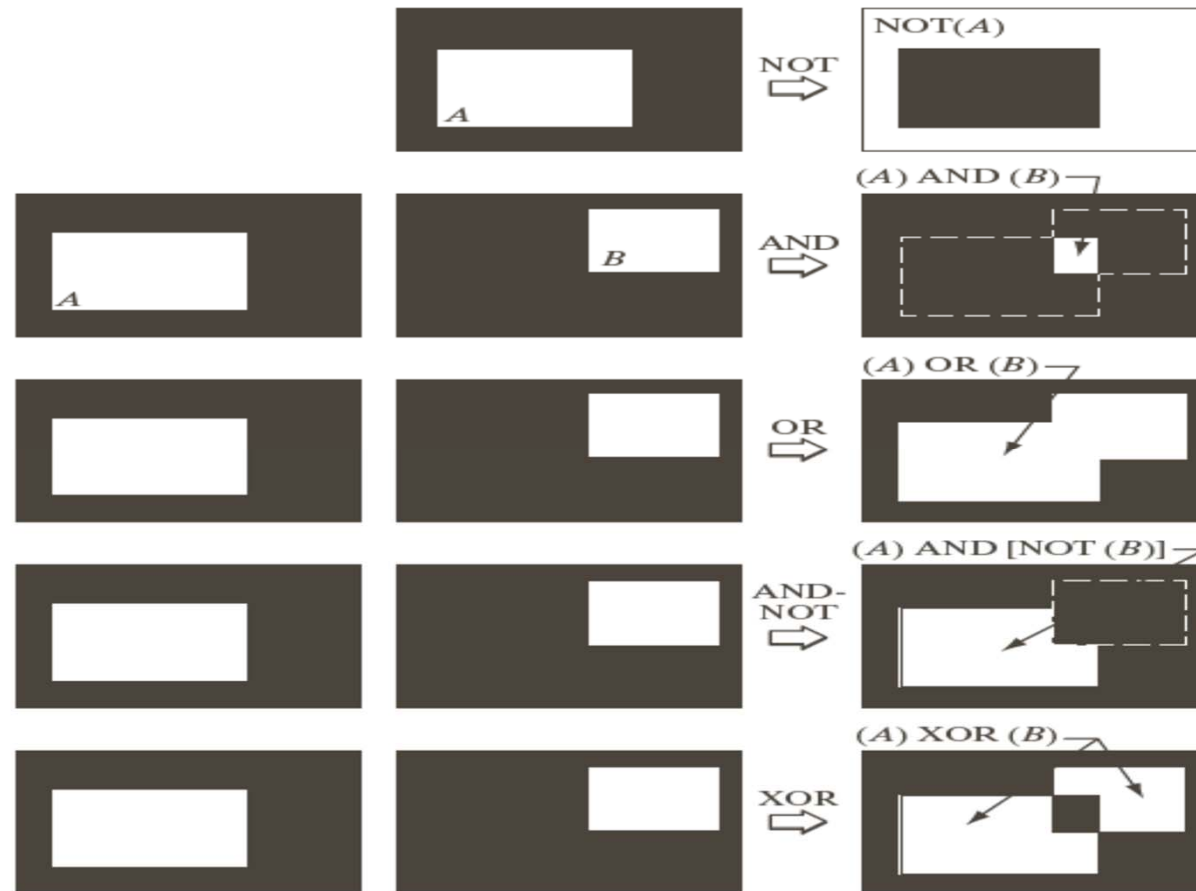
原图

补集

原图与常数图像的并

逻辑操作

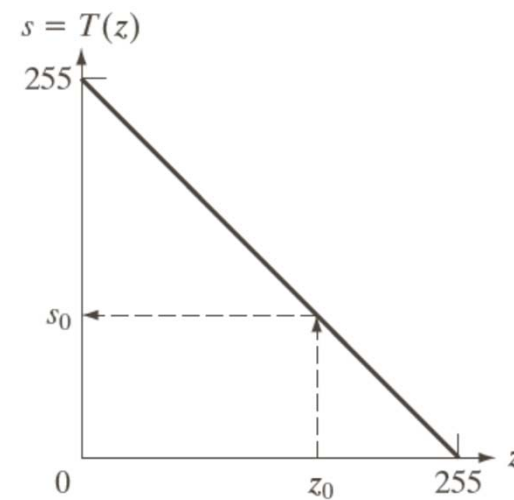
- 逻辑操作



空间操作

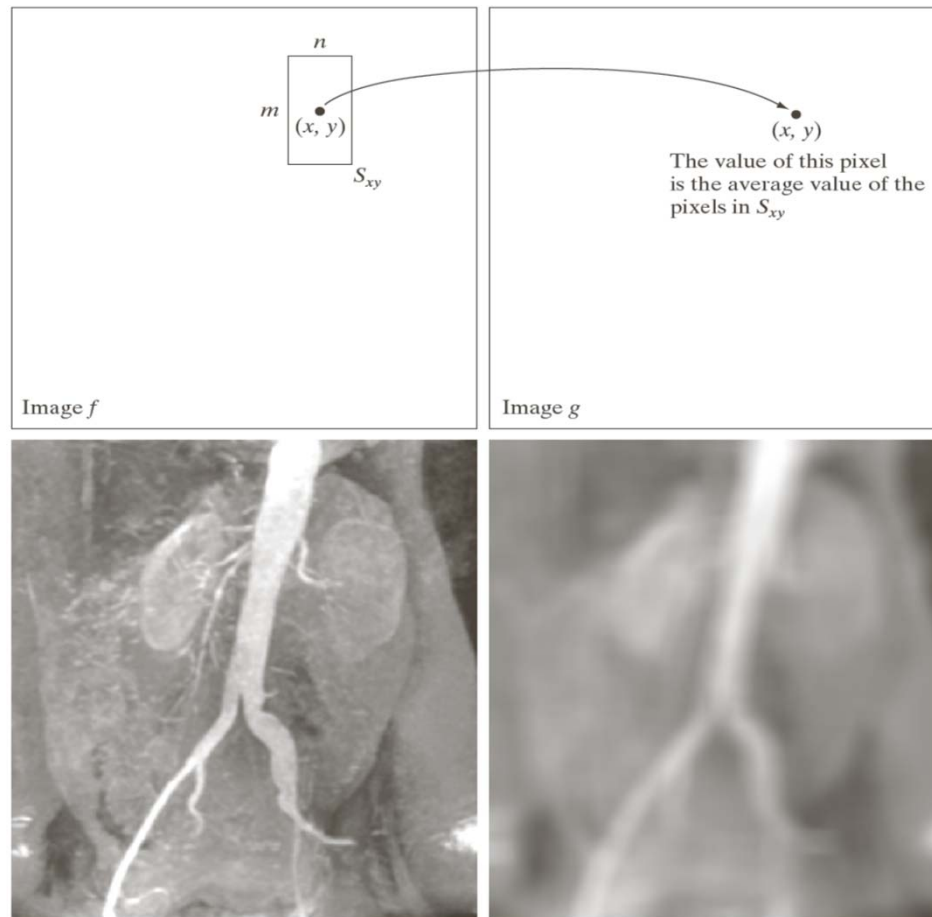
- 包含三类：单像素操作、邻域操作、集合空间变换
- 单像素操作

$$s = T(z)$$



- 邻域操作

$$g(x, y) = \frac{1}{mn} \sum_{(r, c) \in S_{xy}} f(r, c)$$



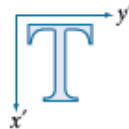
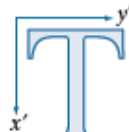
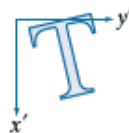
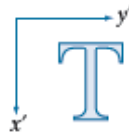
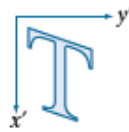

图像配准

- 几何空间变换和图像配准
 - 几何变换由两个基本操作组成
 - 坐标的空间变换
 - 灰度内插
 - 常用的几何空间变换——仿射变换

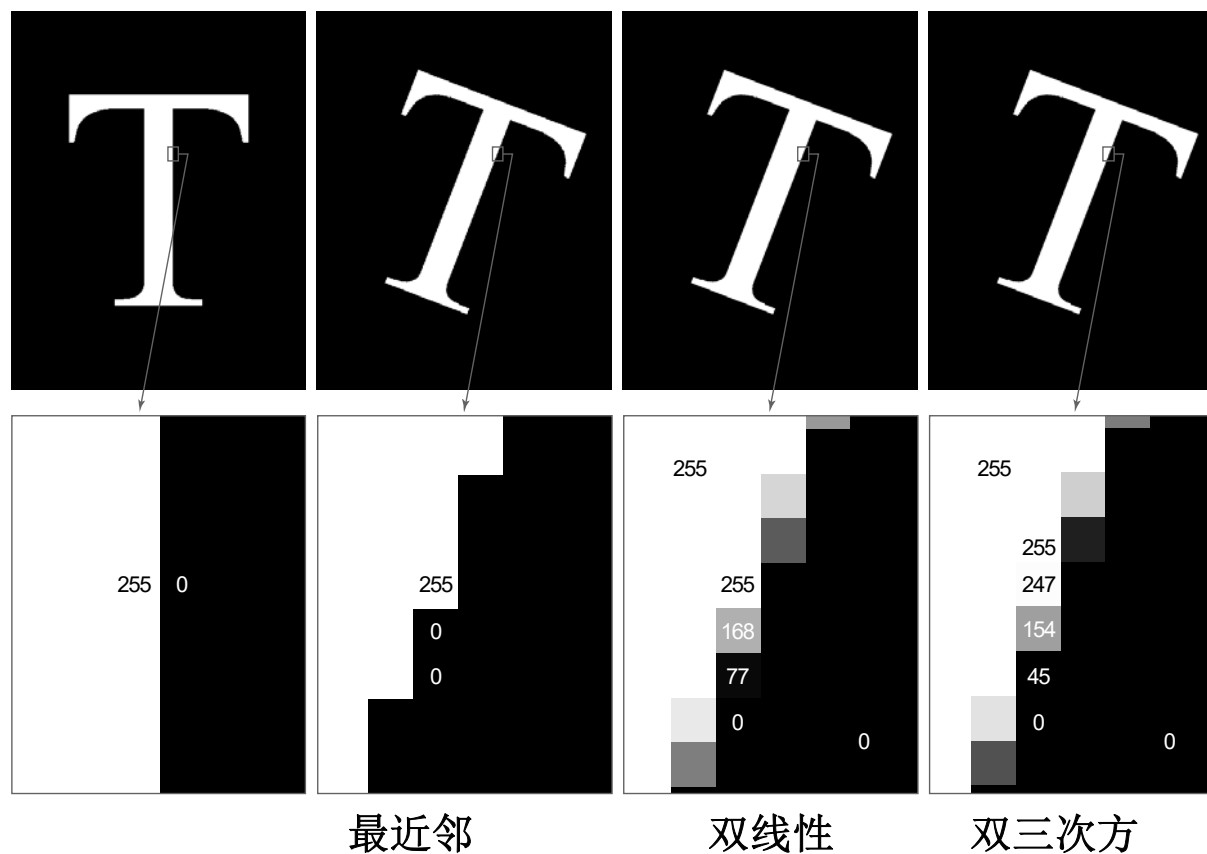
$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = T \begin{bmatrix} v \\ w \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ w \\ 1 \end{bmatrix}$$

仿射变换

TABLE 2.3
Affine
transformations
based on
Eq. (2-45).

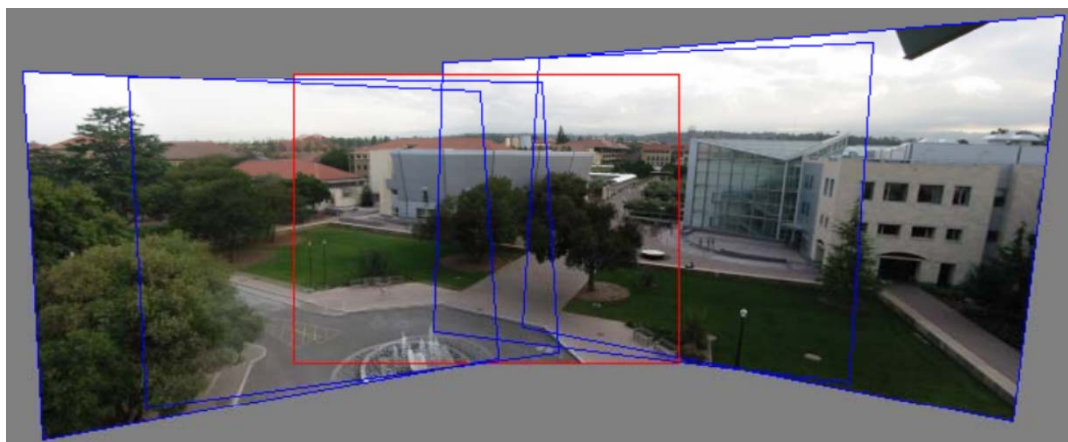
Transformation Name	Affine Matrix, A	Coordinate Equations	Example
Identity	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = y$	
Scaling/Reflection (For reflection, set one scaling factor to -1 and the other to 0)	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = c_x x$ $y' = c_y y$	
Rotation (about the origin)	$\begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x \cos \theta - y \sin \theta$ $y' = x \sin \theta + y \cos \theta$	
Translation	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + t_x$ $y' = y + t_y$	
Shear (vertical)	$\begin{bmatrix} 1 & s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x + s_v y$ $y' = y$	
Shear (horizontal)	$\begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x' = x$ $y' = s_h x + y$	

图像旋转与灰度内插



图像配准

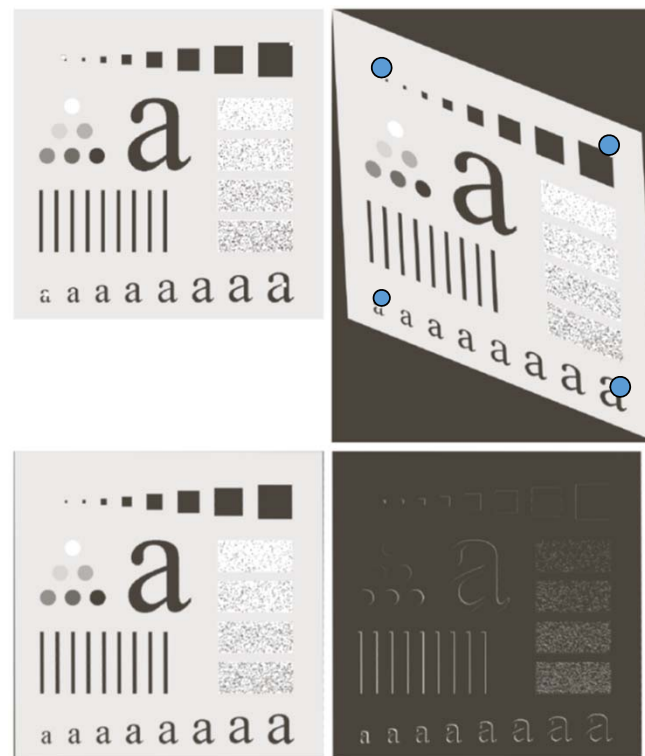
- 用于对其两幅或者多幅相同或者相似场景的图像，是图像处理的一种重要应用；
- 主要问题是需要估计两幅图像之间的几何变换函数；
- 通常采用约束点（或者特征点）对应的方法求取。



图像配准

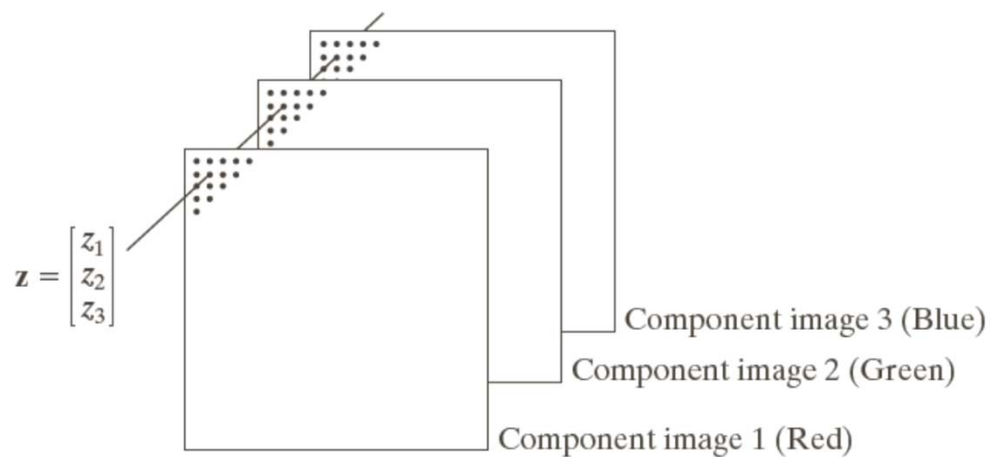
- 图像配准

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = T \begin{bmatrix} v \\ w \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ w \\ 1 \end{bmatrix}$$



向量与矩阵操作

- 1 彩色图像可以看做是三维向量，多光谱图像则维数更高
- 2. 整幅图像可以看做矩阵或者向量来处理。把尺寸为 $M \times N$ 的图像描述为一个 $MN \times 1$ 维向量来处理，如线性处理： $\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}$



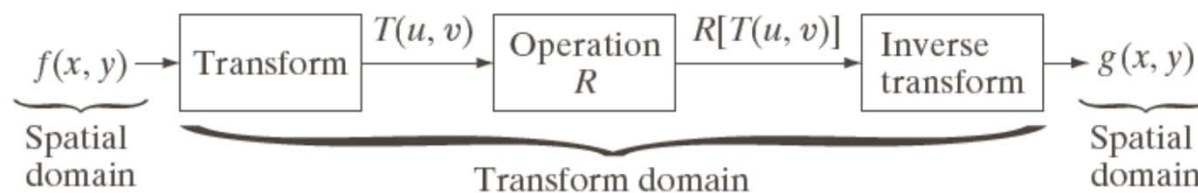
图像变换

- 很多时候，除了空间域，在图像的变换域做处理会更好。
- 正变换：

$$T(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)r(x, y, u, v)$$

- 反变换：

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} T(u, v)s(x, y, u, v)$$



变换举例

- 变换核可分

$$r(x, y, u, v) = r_1(x, u)r_2(y, v)$$

- 变换核对称

$$r(x, y, u, v) = r_1(x, u)r_1(y, v)$$

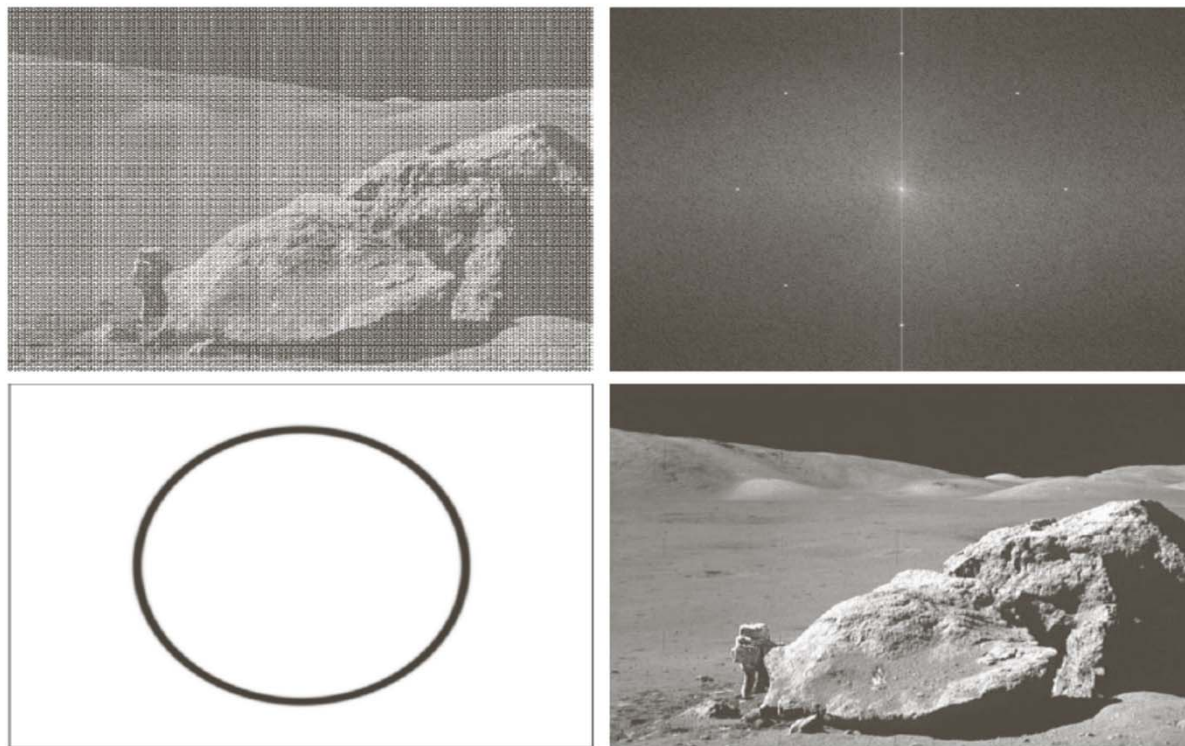
- 傅里叶变换

$$T(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M+vy/N)}$$

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} T(u, v) e^{j2\pi(ux/M+vy/N)}$$

变换域图像处理例子

- 被正弦噪声干扰的图像复原



概率方法

- 概率方法在图像处理中应用广泛。最基本的方式是以随机量来处理灰度值。
- 将图像看成为是一个随机场，可以获得很多相关的统计量。如：

给定灰度级的概率：

$$p(z_k) = \frac{n_k}{MN}$$

图像灰度均值：

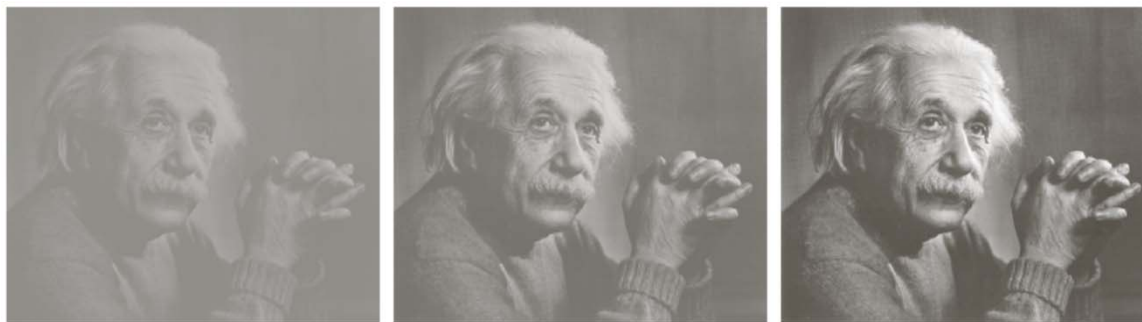
$$m = \sum_{k=0}^{L-1} z_k p(z_k)$$

图像灰度方差：

$$\sigma^2 = \sum_{k=0}^{L-1} (z_k - m)^2 p(z_k)$$

统计量与视觉特性

- 图像的均值和方差对于图像的视觉特性有明显的直接关系。
- 图像灰度标准差



标准差分别为**14.3, 31.6, 49.2**

本章作业

- 书面作业（第4版教材）（下周二课前提交书面作业本）：
 - 习题2.8, 2.14, 2.16, 2.18, 2.22, 2.26, 2.28, 2.37.
- 上机作业（两周内，学在浙大网上提交，实验报告和源代码）：
 - 将Lena图像降采样后，分别利用最近邻、双线性和双三次方插值恢复至原分辨率，并比较效果差异。要求除了图像输入和输出显示外，核心处理函数自己编写，不能直接调用opencv或者matlab里的相关函数。