



## Chapter 2

# Digital Image Fundamentals

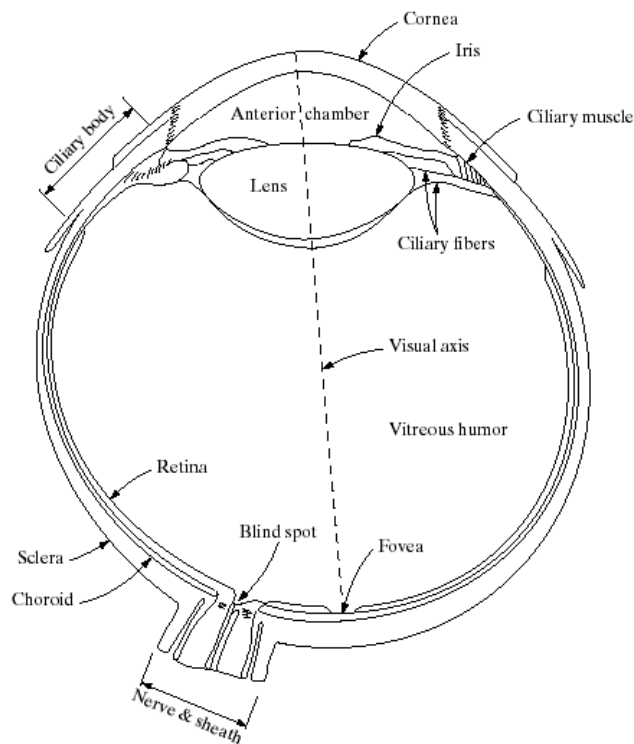
## 第二章：数字图像基础

### 2.1 视觉的一些基本要素

人的视觉是由眼睛中两部分光接收器（感觉细胞）组成的：锥状体和杆状体。

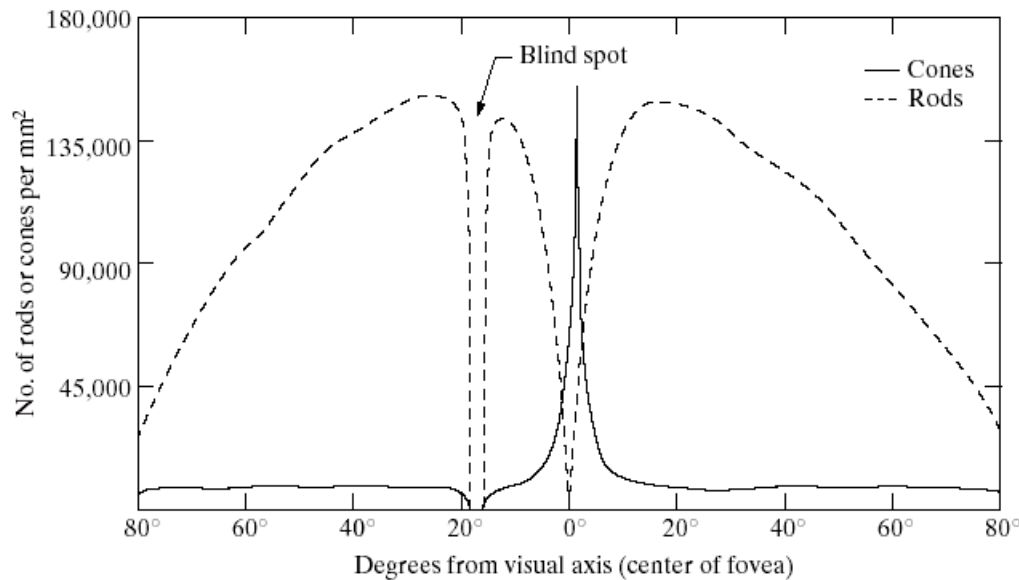
每只眼睛锥状体数目**6-7**百万，每个锥状体都连接到自己的神经末端，对色彩敏感，分布于视网膜中央凹部分，称为白昼视觉或亮光视觉。

杆状体**7500**万到**15000**万，分布在视网膜表面，几个杆状体连接到一个神经末端，不如锥状体灵敏。给出图像的总体轮廓，没有彩色感觉，在低照明度下对图像较敏感。称为夜视觉或暗视觉。



**FIGURE 2.1**  
Simplified  
diagram of a cross  
section of the  
human eye.

## 2.1 Elements of Visual Perception



**FIGURE 2.2**  
Distribution of  
rods and cones in  
the retina.

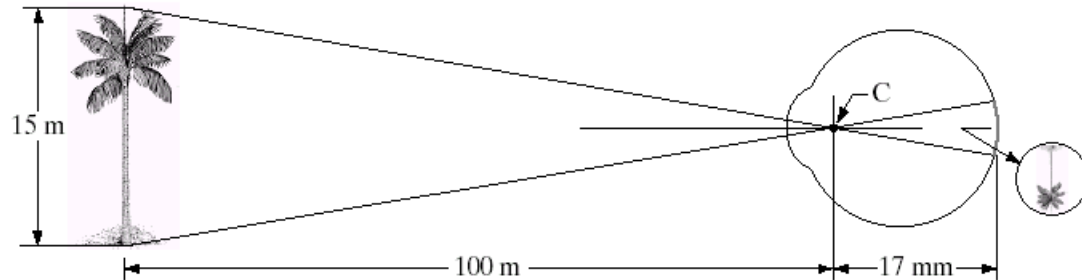
**分辨细节的基本能力：**每平方毫米150,000个像素，最高敏感区(中间凹)的接收阵列近似为 $1.5\text{mm} \times 1.5\text{mm}$ ，锥状体数量为337000个，相当于一个接收阵列不大于 $5\text{mm} \times 5\text{mm}$ 的中等分辨率的电耦合元件（CCD: Charge-coupled device）的成像芯片。

## 2.1 Elements of Visual Perception

**成像原理：**和光学透镜类似，但适应性强，是可自行调节的透镜。

**FIGURE 2.3**

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



看远处物体，肌肉会迫使晶状体变得扁平，晶状体的聚焦中心向前移动；物体离眼睛近时，肌肉使晶状体变厚，光心向视网膜成像区域靠近。光心到视网膜的距离在14mm到17mm之间变化。物体由远至近，焦距由17mm向14mm变化，晶状体的折射能力也由弱变强。当物体远于超过3米时，折射能力最弱，这也是为什么远处物体的细节难以分辨的原因之一。

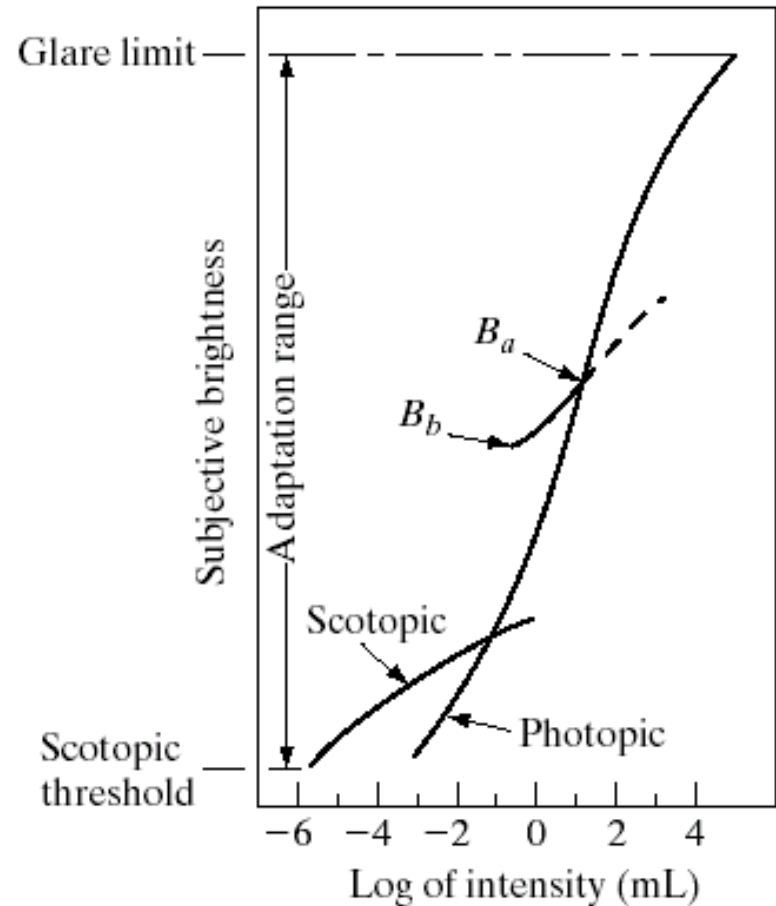
### 眼睛对亮度的适应和鉴别

#### 一些有意思的结论:

- 视觉系统能够适应的光强度级别很宽，从夜视域值到强闪光约有 $10^{10}$ 个量级(朗伯)，下一页图对应于对数坐标系横坐标范围-6~4；
- 但人的视觉绝对不可能同时感知整个范围，而是通过不断地改变其整个灵敏度来完成这一大变动的。与整个适应范围相比，能同时鉴别的光强度级别的总范围很小；
- 在低的照明级别，亮度辨别较差（杆状体起作用）。在背景照明增强时，亮度辨别得到明显的改善（锥状体起作用）。
- 背景照明保持恒定时，眼睛一般可辨别总共12到24级不同的强度。但这并不意味着仅用很小的强度值就可以表示一幅图。这是因为当眼睛扫视图像时，平均背景在变化，眼睛也会根据这种变化作调整，最后结果是眼睛能够辨别很宽的全部强度范围。

**FIGURE 2.4**  
Range of  
subjective  
brightness  
sensations  
showing a  
particular  
adaptation level.

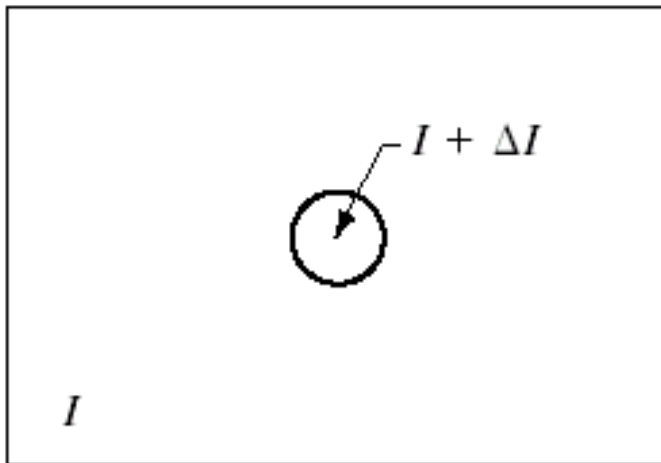
主观亮度（即由人的视觉系统感知的亮度）是进入人眼的光强的对数函数。长实线代表视觉系统能适应的光强范围。在亮视觉中，该范围大约是 $10^6$ 。由暗视觉逐渐过渡到亮视觉的近似范围约为  $0.001 \sim 0.1 \text{ mL}$ （郎伯）（在对数坐标中为  $-3 \sim -1 \text{ mL}$ ），图中画出了该范围内这一适应曲线的两个分支。



**韦伯比:** 用于度量人的眼睛特定的适应级别对亮度变化的辨别力。

由  $\frac{\Delta I_c}{I}$  决定。该比值较小意味着可辨别强度较小的百分比变化，

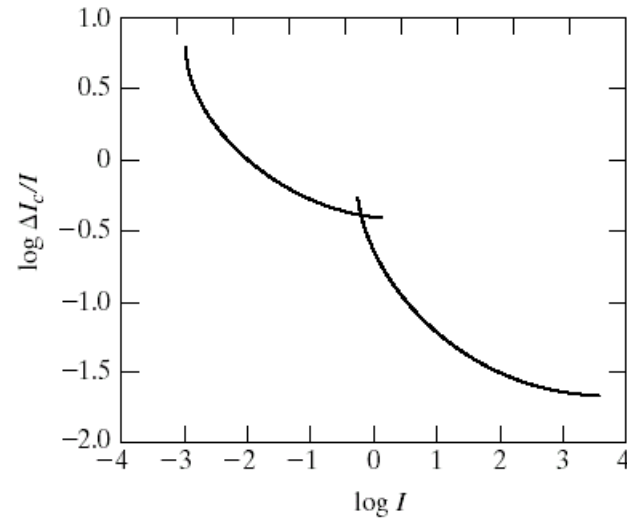
这表示亮度辨别能力“较好”，反之，表示亮度辨别能力“较差”。



**FIGURE 2.5** Basic experimental setup used to characterize brightness discrimination.

## 2.1 Elements of Visual Perception

**FIGURE 2.6**  
Typical Weber  
ratio as a function  
of intensity.

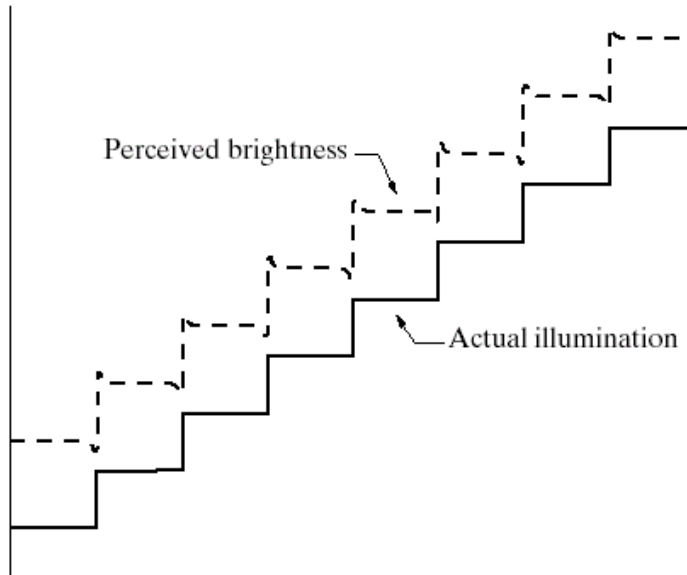
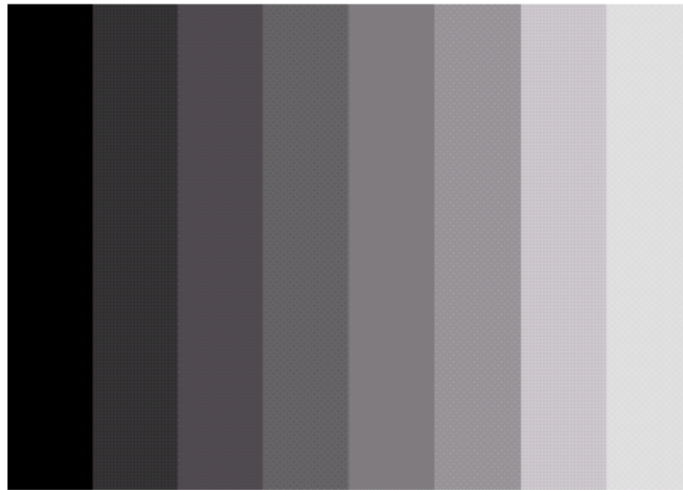


在低照明级别，亮度辨别力差；在高照明级别，亮度辨别力好。



### 两个现象说明了感觉亮度不是简单的强度函数

第一个现象是视觉系统倾向于不同强度区域边界周围的“欠调”（Undershoot）和“过调”（Overshoot）

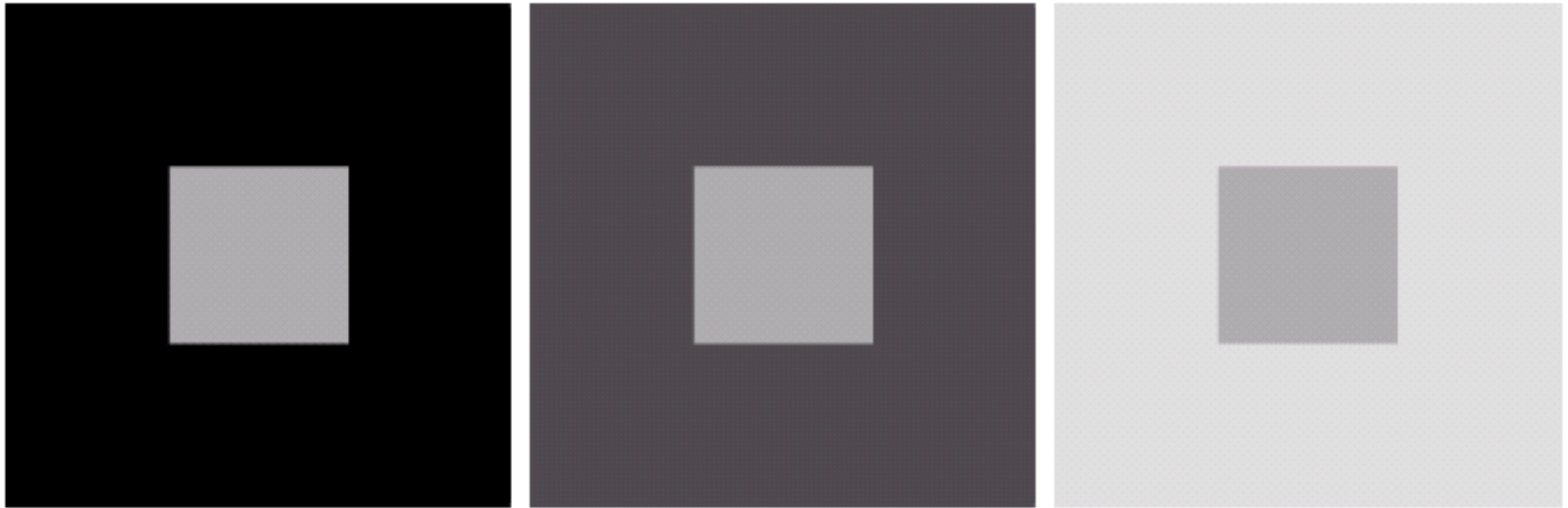


a  
b

**FIGURE 2.7**

(a) An example showing that perceived brightness is not a simple function of intensity. The relative vertical positions between the two profiles in (b) have no special significance; they were chosen for clarity.

第二个现象称为同时对比现象



a b c

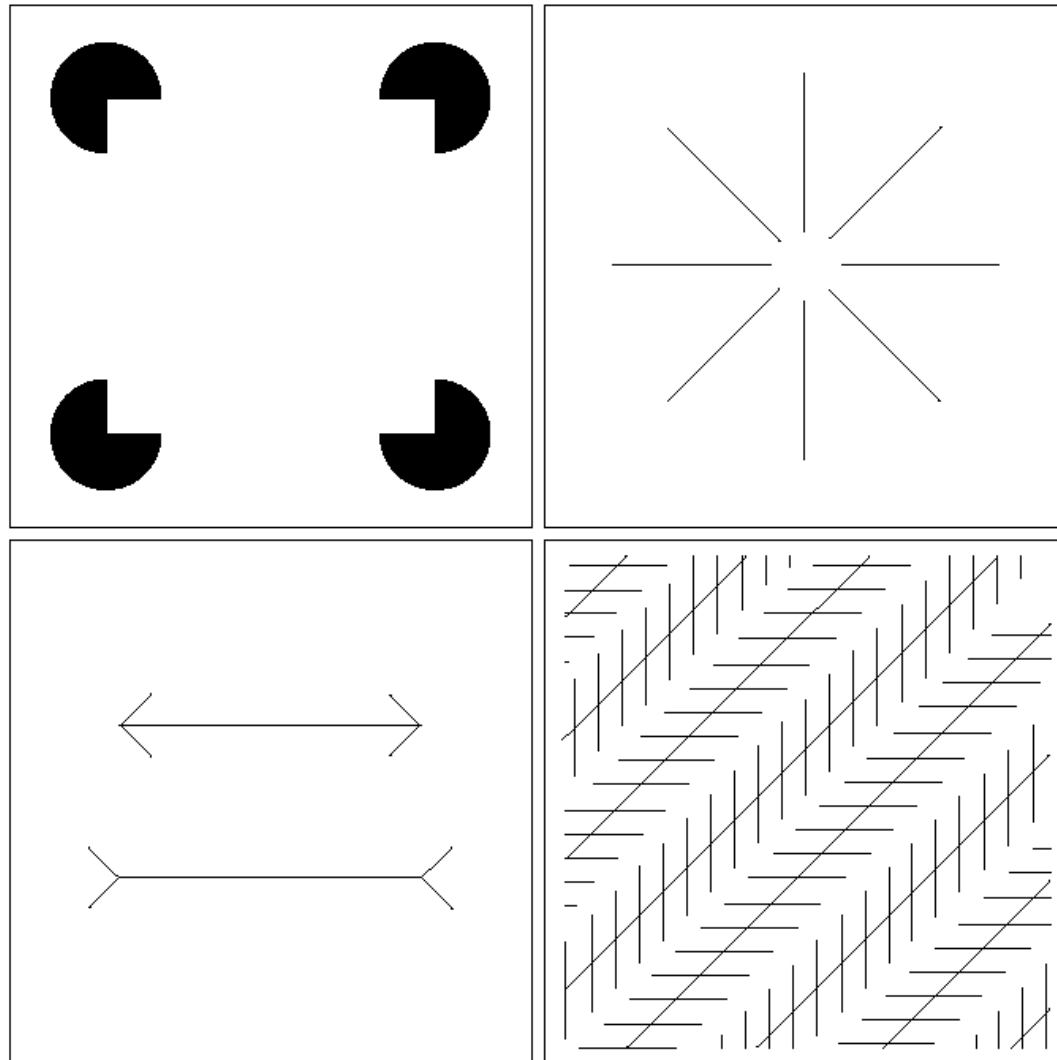
**FIGURE 2.8** Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

## 2.1 Elements of Visual Perception

视觉错觉：这是人类视觉系统的一个特性，尚未完全了解。

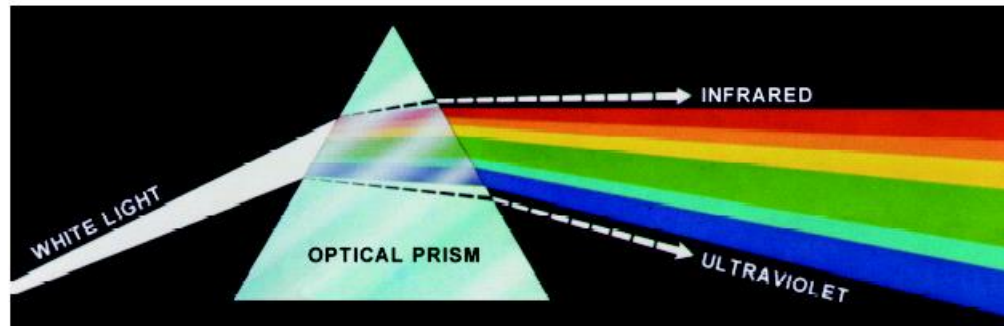
a b  
c d

**FIGURE 2.9** Some well-known optical illusions.



### 2.2光和电磁波

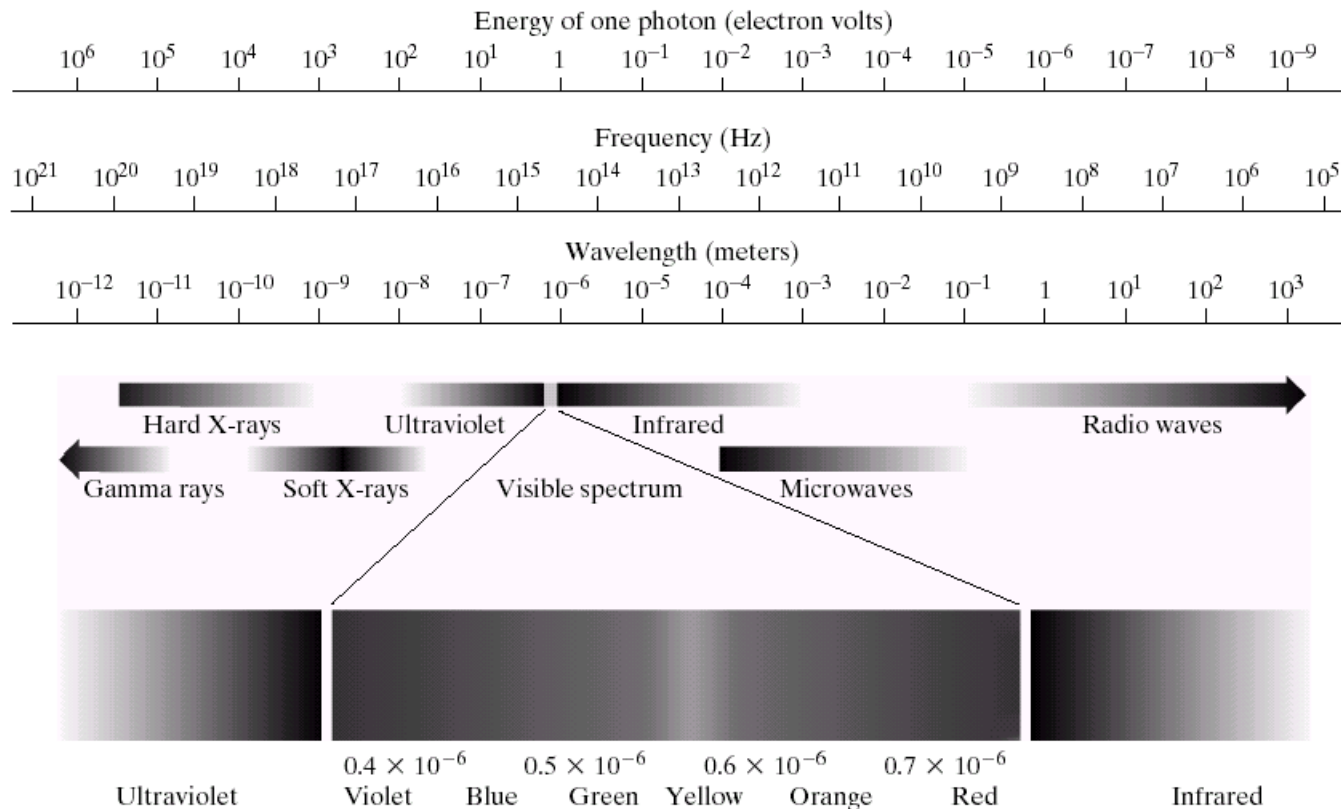
牛顿1666年发现了这样一个现象：当一束白光通过一个玻璃棱镜时，出来的光就不再是白光，而是一端是紫色另一端是红色的连续色谱。



**FIGURE 6.1** Color spectrum seen by passing white light through a prism. (Courtesy of the General Electric Co., Lamp Business Division.)

## 2.2 Light and Electromagnetic Spectrum

光波等同于电磁波，反过来也成立。可见光的范围：电磁波约400~700nm（ $1\text{nm}=10^{-9}\text{m}$ ）的范围。波谱一端是无线电波，波长比可见光长几十亿倍；另一端是伽马射线，波长比可见光短几百万倍。



**FIGURE 2.10** The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

## 2.2 Light and Electromagnetic Spectrum

电磁波可以用波长、频率或能量来描述。波长( $\lambda$ )和频率( $\nu$ )关系式为:

$$\lambda = c / \nu$$

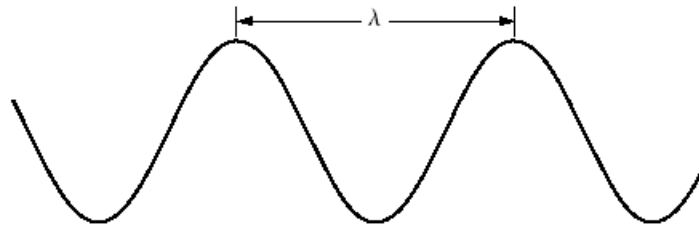
其中 $c$ 是光速 ( $2.998 \times 10^8 m/s$ )，电磁波谱的能量由下式给出

$$E = h\nu$$

$h$ 是普朗克常数。能量与频率成正比。

电磁波可以看成是以波长  $\lambda$  传播的正弦波。

**FIGURE 2.11**  
Graphical  
representation of  
one wavelength.



可见光是一种特殊的电磁波谱，只在电磁波谱中占很小的一部分。

眼睛从物体上感受到的颜色和物体发射光的性质有关。一个物体对所有可见光波长的反射是相对平衡的，则这个物体将呈现白色（灰色）。

单色光的属性是它的强度或大小。灰度级通常是用来描绘单色光的强度，它的范围从黑到灰，最后到白。

**通常有三个基本量用于描绘彩色光源的质量：**发光强度（从光源流出的能量）、光通量（观察者从光源感受的能量，例如：远红外光有实际的能量，但光通量为零）和亮度（亮度是描绘光感受的主观描绘，它实际上不能测量，包含无色的强度的概念，并且也是描述彩色感觉的参数之一）。

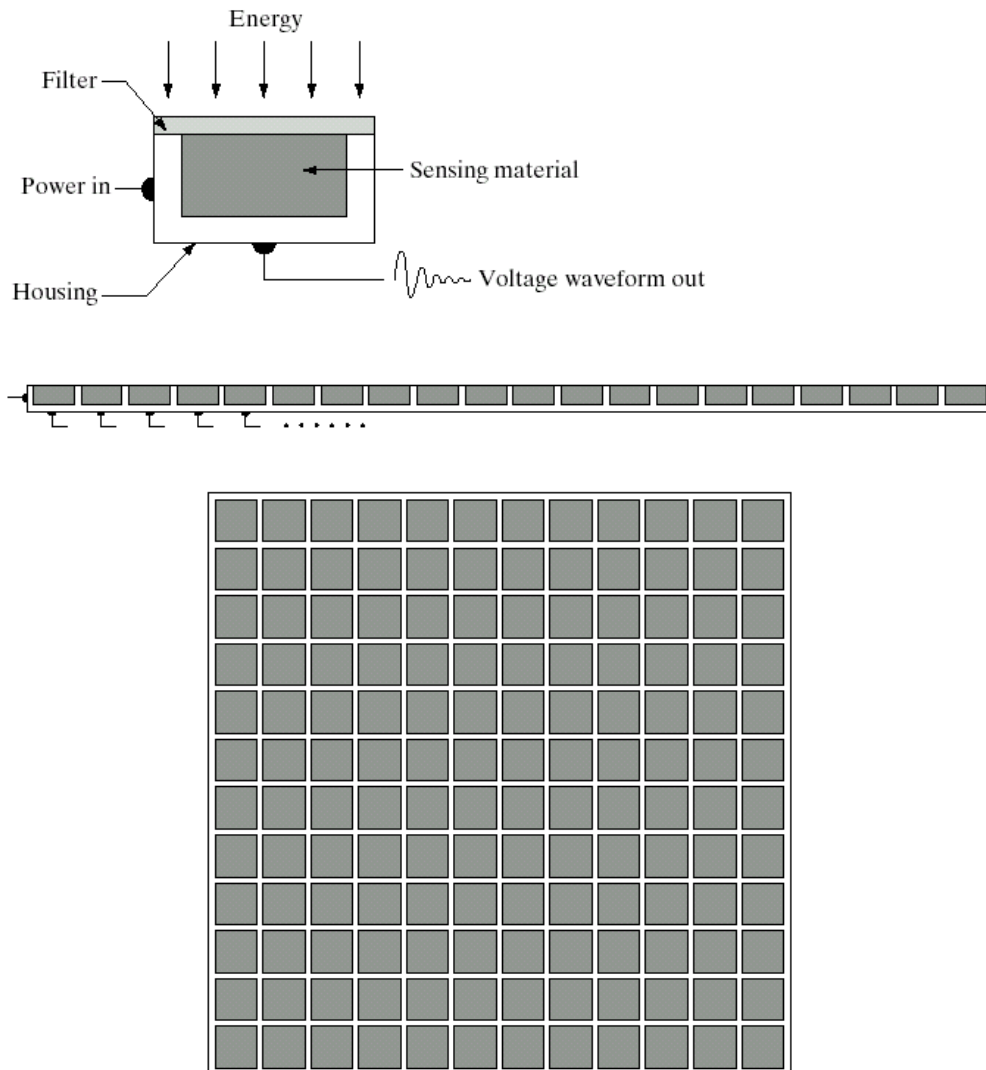
### 2.3 图像的感知和获取

一般来说，各类图像都是由“照射”源和形成图像的“场景”元素对光能的反射或吸收相结合而产生的。

a  
b  
c

**FIGURE 2.12**

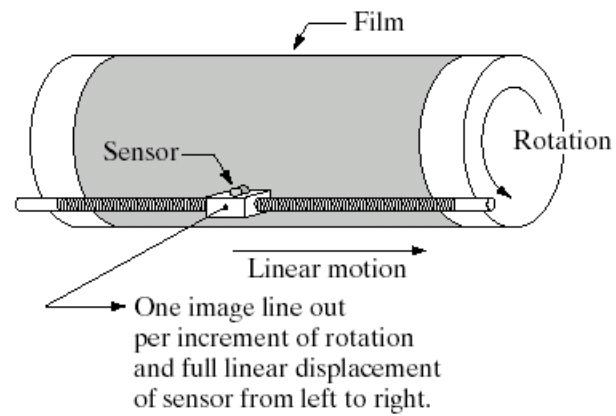
(a) Single imaging sensor.  
(b) Line sensor.  
(c) Array sensor.





## 2.3 Image Sensing and Acquisition

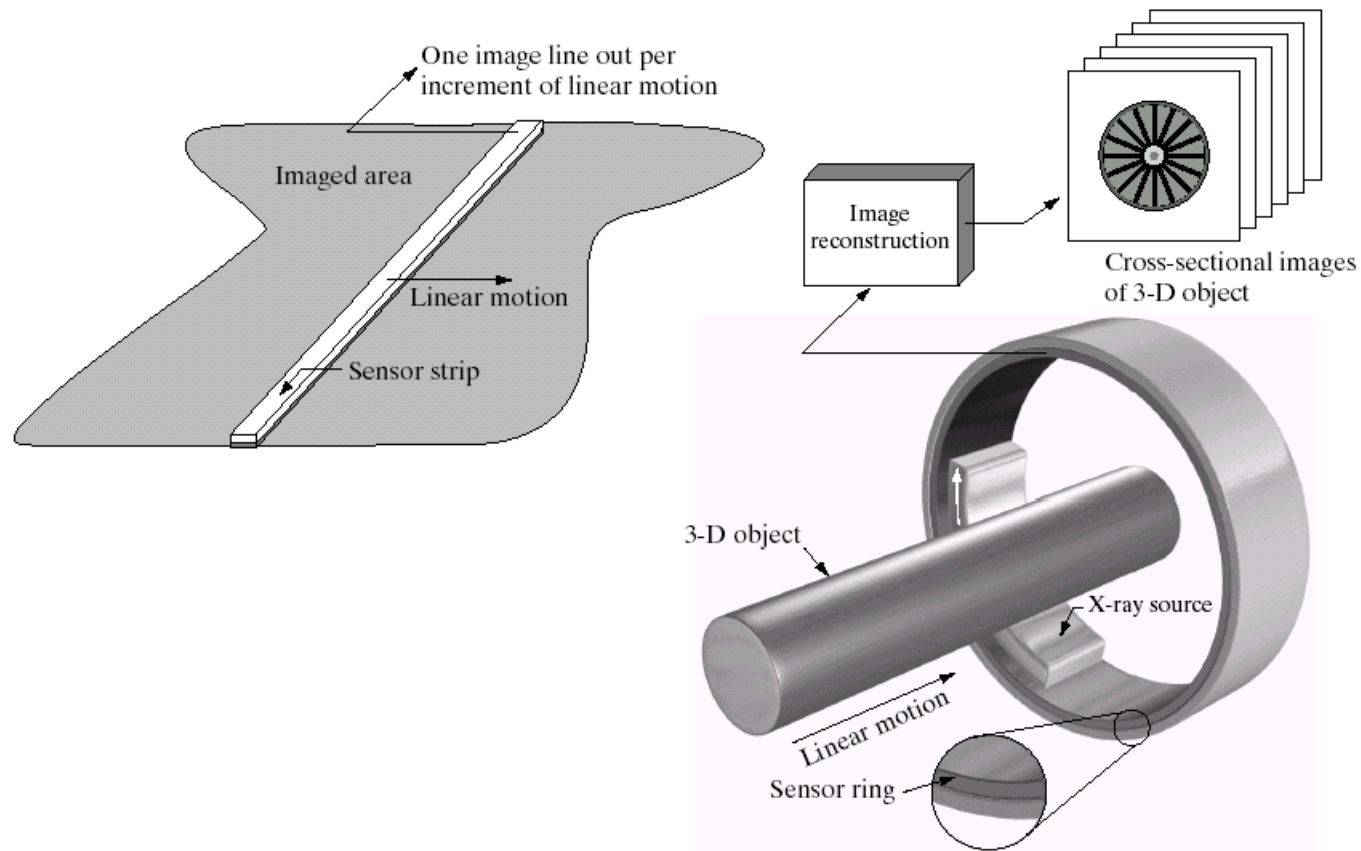
使用单个传感器获取图像



**FIGURE 2.13** Combining a single sensor with motion to generate a 2-D image.

## 2.3 Image Sensing and Acquisition

### 使用条带传感器获取图像

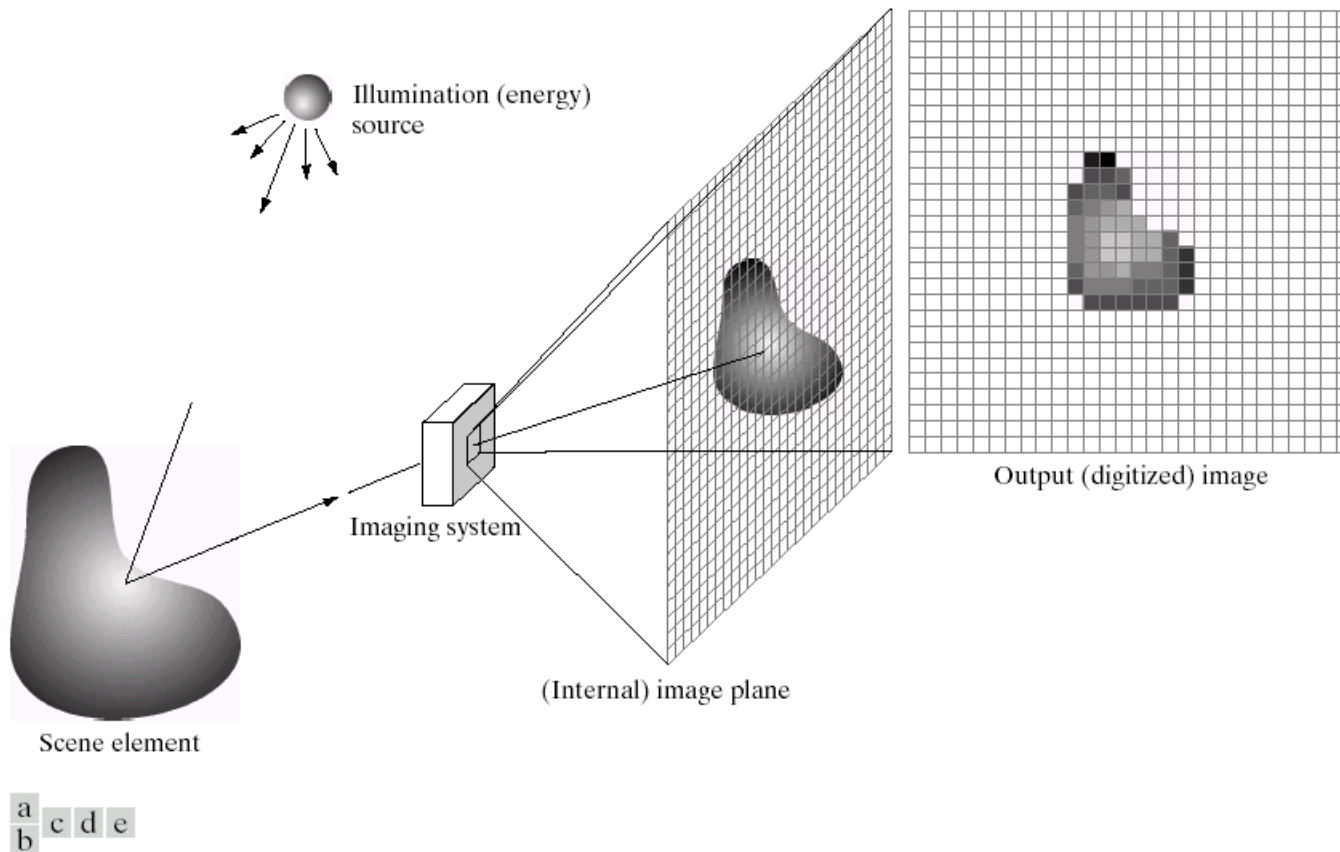


a b

**FIGURE 2.14** (a) Image acquisition using a linear sensor strip. (b) Image acquisition using a circular sensor strip.

## 2.3 Image Sensing and Acquisition

### 使用传感器阵列获取图像



**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.

### 简单的图像模型

用二维函数的形式表示一幅单色图像。当一幅图像从物理过程产生时，它的值正比于物理源的辐射能量（如电磁波）。故一定有

$$0 < f(x, y) < \infty$$

另外，函数 $f(x, y)$ 有两个分量来表征：

- （1）入射到观察场景的光源总量和
- （2）场景中物体反射光的总量。

称为入射分量和反射分量，并分别用 $i(x, y)$ 和 $r(x, y)$ 表示。两个函数合并形成图像函数 $f(x, y)$ ：

$$f(x, y) = i(x, y) r(x, y)$$

其中  $0 < i(x, y) < \infty$

$$0 < r(x, y) < 1$$

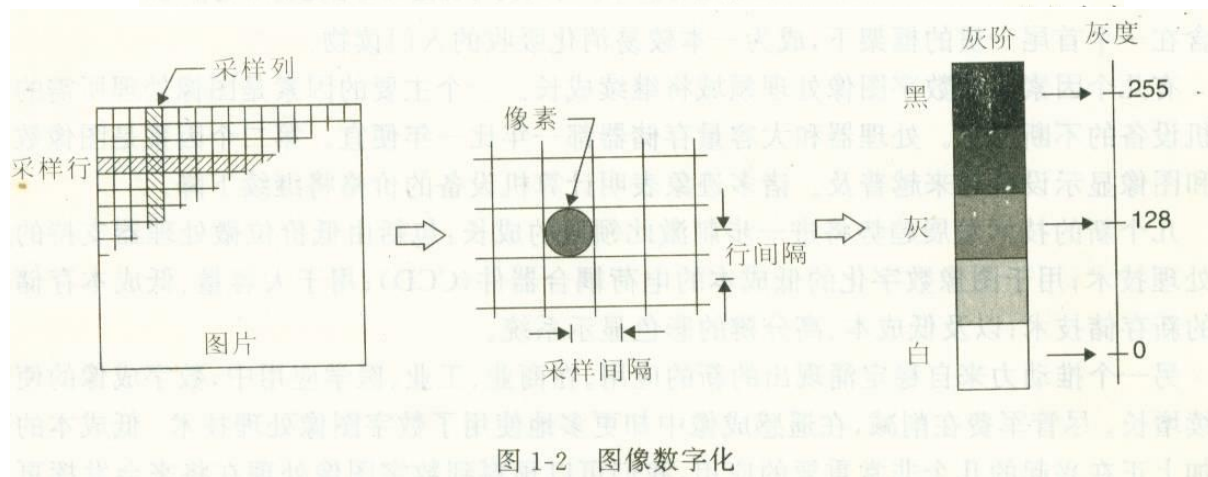
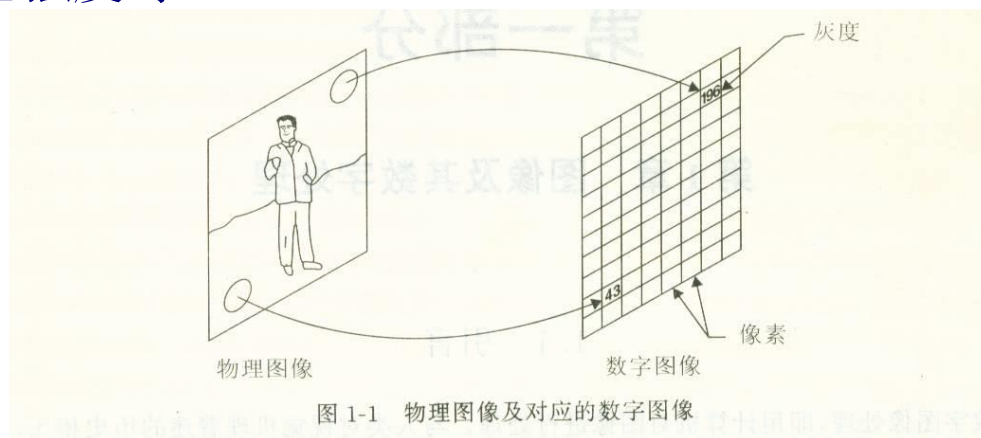
单色图像上任一点的强度就是图像在那一点的灰度级。反射分量限制在0（全吸射）和1（全反射）之间。

### 2.4 图像的取样和量化

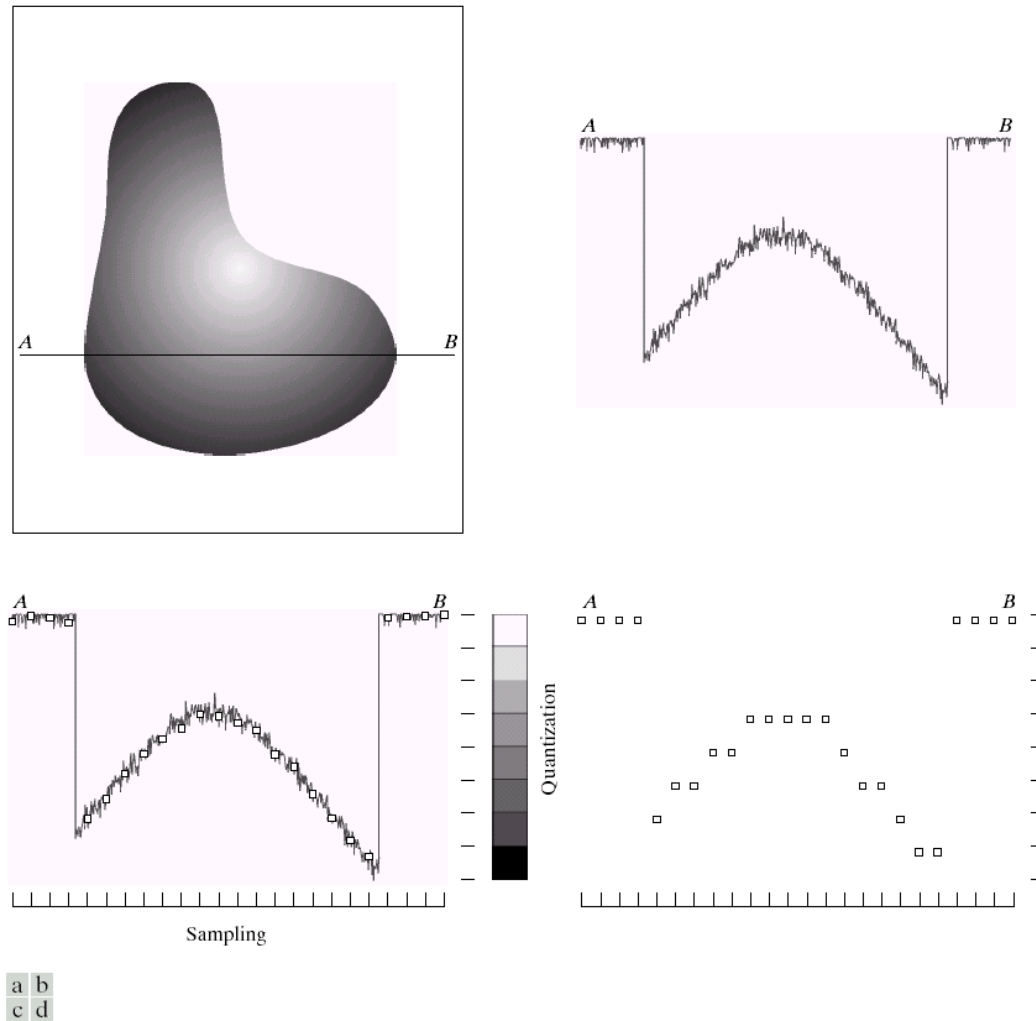
大多数传感器的输出是连续的电压波形（图像），为了产生一幅数字图像，需要把连续的感知数据转换为数字形式。这就包含了两种处理，取样（时空域）和量化（光色强度等）。

#### 2.4.1 取样和量化的基本概念

图像数字化 = 取样 + 量化

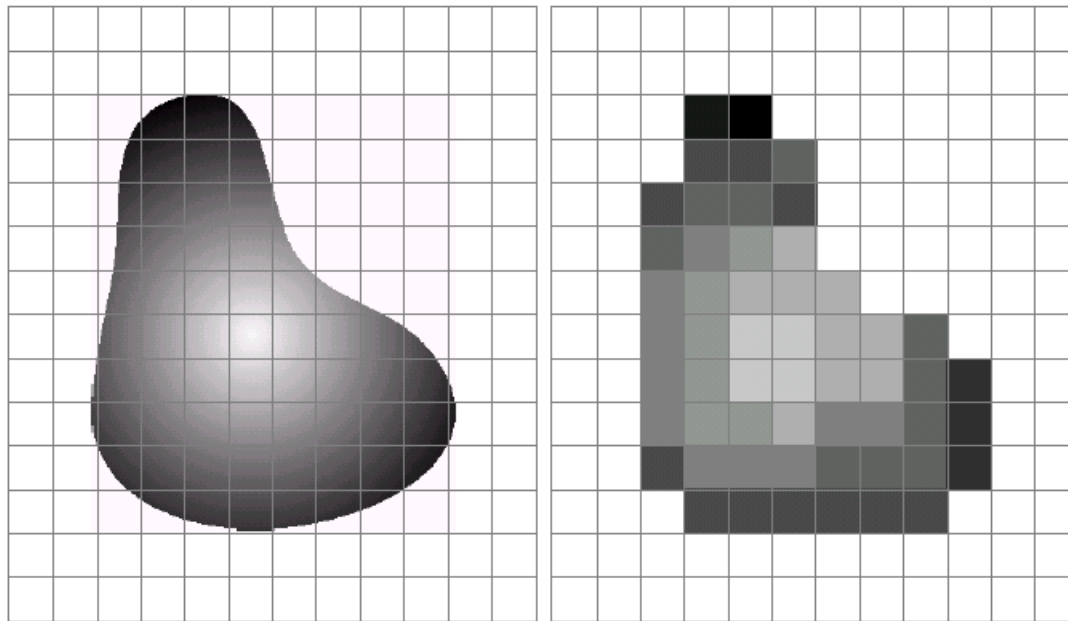


## 2.4 Image Sampling and Quantization



**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 Image Sampling and Quantization



a b

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

---

### 2.4.2 数字图像表示

二维矩阵是表示数字图像的重要数学形式。一幅 $M \times N$ 的图像可以表示为：

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

或者

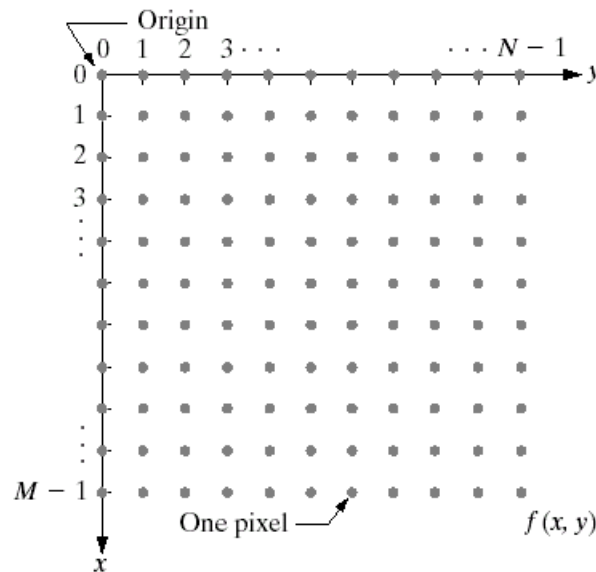
$$A = \begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,N-1} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,N-1} \\ \vdots & & & \\ a_{M-1,0} & a_{M-1,1} & \cdots & a_{M-1,N-1} \end{bmatrix}$$

矩阵中的每个元素称为图像的“像素”。每个像素都有它自己的“位置”和“值”。



## 2.4 Image Sampling and Quantization

取样和量化的正规数学描述：令 $\mathbf{Z}$ 和 $\mathbf{R}$ 分别表示整数集、实数集。取样过程可以看作是把 $xy$ 平面分为网格，每一网格中心的坐标是笛卡尔坐标 $\mathbf{Z}^2$ 的一对元素， $\mathbf{Z}^2$ 是所有元素对 $(z_i, z_j)$ 的集合， $z_i$ 和 $z_j$ 是 $\mathbf{Z}$ 中的整数。因此，如果 $(x, y)$ 是 $\mathbf{Z}^2$ 中的元素，并且 $f$ 是把灰度级值（即实数集 $\mathbf{R}$ 中的一个实数）赋予特定坐标，则 $f(x, y)$ 就表示一幅数字图像。这个赋值过程就是前面描述的量化过程。



**FIGURE 2.18**

Coordinate convention used in this book to represent digital images.

## 2.4 Image Sampling and Quantization

出于处理、存储和硬件的考虑，灰度级别通常是2的整数幂

$$L = 2^k$$

$L$ 是最大的灰度级别。这时，图像中所有像素的灰度是区间  $[0, L-1]$  的整数。一幅数字图像占用的空间： $M \times N \times k$ 。

**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

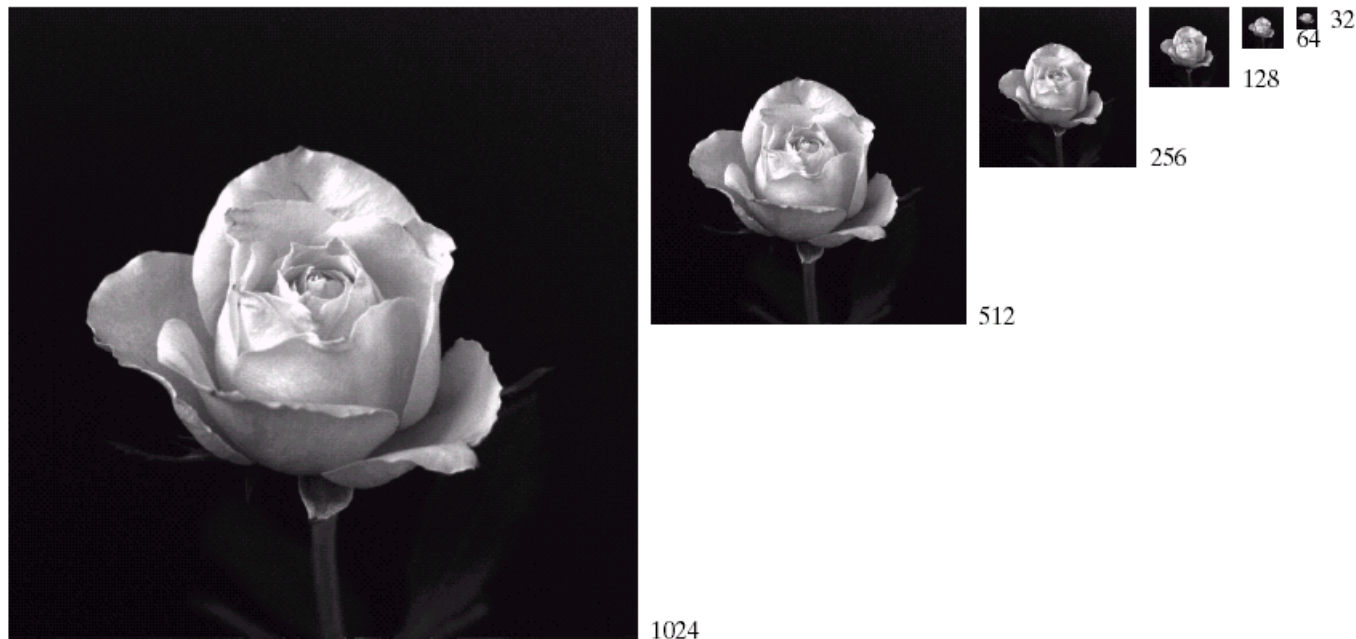
$M = N$

### 2.4.3 空间和灰度的分辨率

#### 空间分辨率

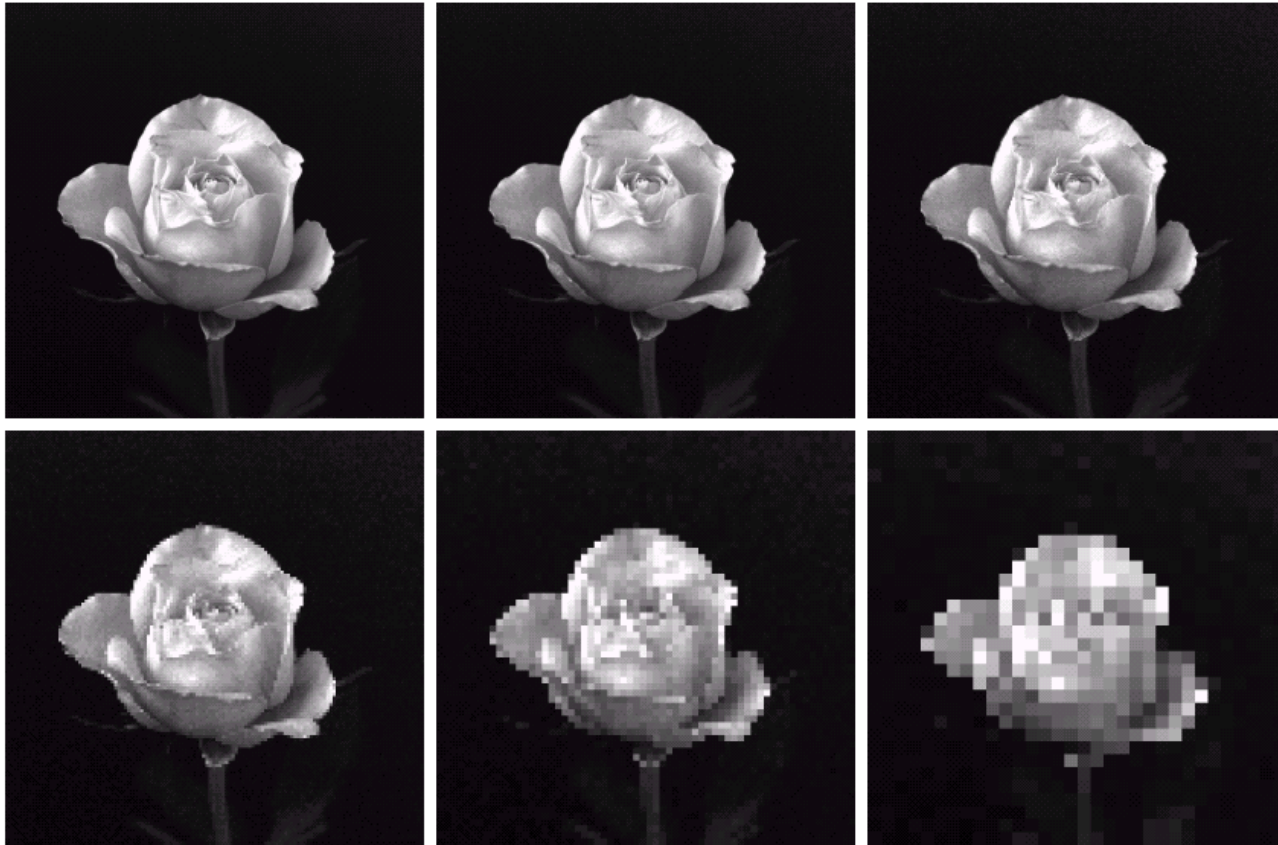
空间分辨率是图像中可分辨的最小细节。广泛使用的分辨率的意义是在每单位距离可以分辨的最小线对数目。

通常，空间分辨率即图像大小（最大行数×每行最大像素数）。



**FIGURE 2.19** A  $1024 \times 1024$ , 8-bit image subsampled down to size  $32 \times 32$  pixels. The number of allowable gray levels was kept at 256.

## 2.4 Image Sampling and Quantization



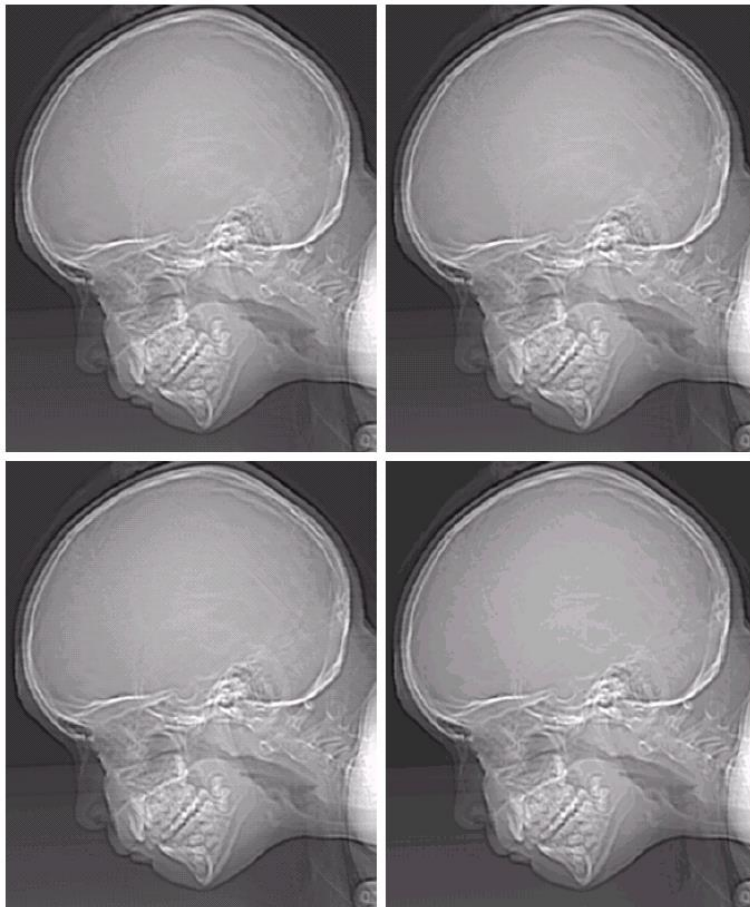
a	b	c
d	e	f

**FIGURE 2.20** (a)  $1024 \times 1024$ , 8-bit image. (b)  $512 \times 512$  image resampled into  $1024 \times 1024$  pixels by row and column duplication. (c) through (f)  $256 \times 256$ ,  $128 \times 128$ ,  $64 \times 64$ , and  $32 \times 32$  images resampled into  $1024 \times 1024$  pixels.



### 灰度分辨率

**灰度分辨率：**一个像素值单位幅度上包含的灰度级数。灰度级数通常是2的整数幂级数，如：用一个byte存一个像素值，则256级。  
用一个4bit存一个像素值，则16级

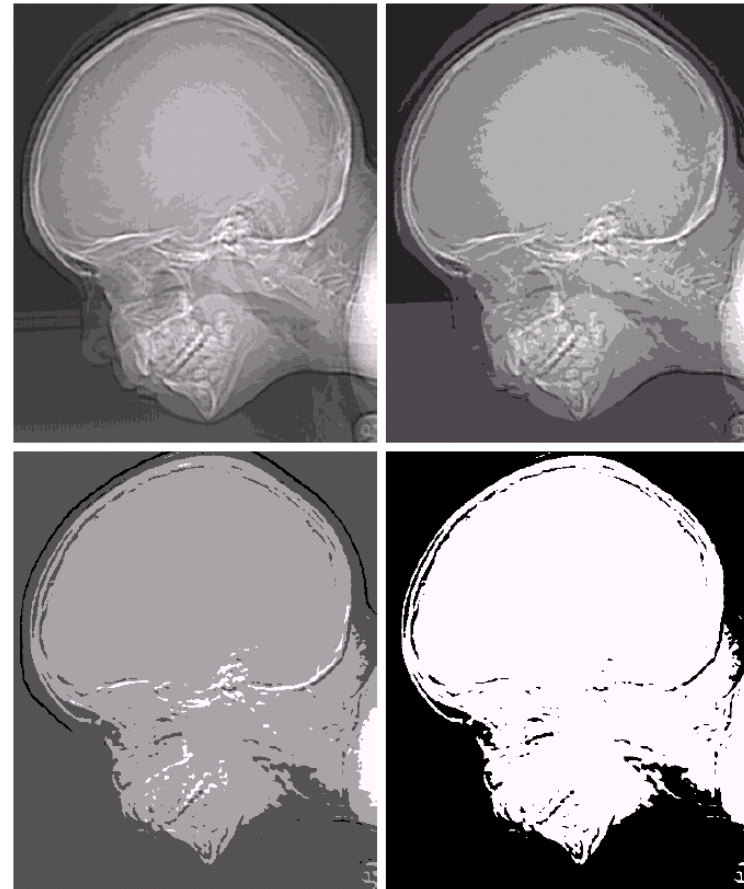


a b  
c d

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

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.)



## 2.4 Image Sampling and Quantization

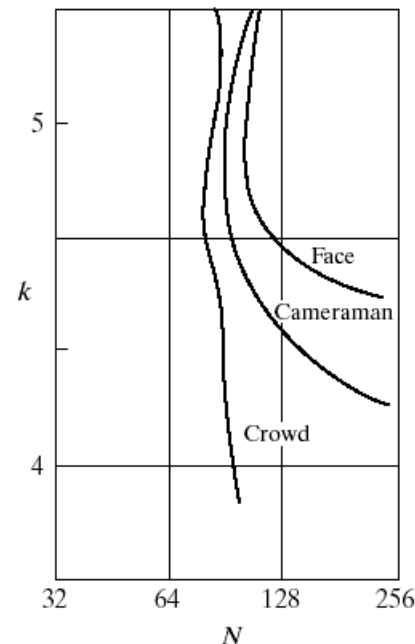
人眼对灰度分辨率的敏感程度和图像内容的复杂程度相关



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.)

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

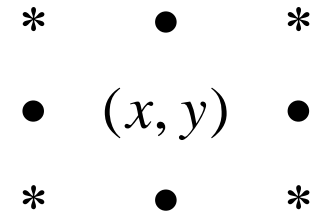


偏爱曲线

### 2.5 像素间的一些基本关系

#### 相邻像素

位于坐标 $(x, y)$ 的像素 $p$ 有四个水平和垂直的相邻像素，每个像素距 $(x, y)$ 一个单位距离。

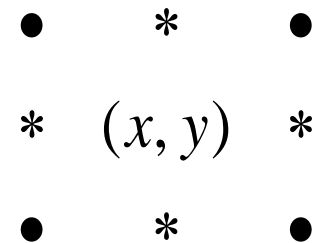


坐标分别为:  $(x-1, y)$ ,  $(x+1, y)$ ,  $(x, y-1)$ ,  $(x, y+1)$ 。此像素集合定义为像素 $p$ 的4邻域，用 $N_4(p)$ 表示。

另外， $p$ 有4个对角相邻像素，坐标为:

$(x-1, y-1)$ ,  $(x+1, y-1)$ ,  $(x-1, y+1)$ ,  $(x+1, y+1)$

用 $N_D(p)$ 表示。



$N_D(p)$ 和 $N_4(p)$ 一起称为 $p$ 的8邻域，用 $N_8(p)$ 表示

### 邻接性、连通性、区域和边界

邻接性有两个要素：一个是灰度值的邻接性（值域 $V$ ）、一个是物理位置的邻接性（邻域，如 $N_4(p)$ 等）。例如，二值图象中，像素值都为1（或都为0）的像素才有可能被称为是邻接的。在一般图像中，可定义一个值域 $V$ ， $V$ 是0到255中的一个子集。

一般我们考虑三种邻接性：

(a) 4邻接：如果点 $q$ 在 $N_4(p)$ 中，并 $q$ 和 $p$ 具有 $V$ 中的数值，则 $q$ 和 $p$ 是4邻接的；

(b) 8邻接：如果点 $q$ 在 $N_8(p)$ 中，并 $q$ 和 $p$ 具有 $V$ 中的数值，则 $q$ 和 $p$ 是8邻接的；

(c)  $m$ 邻接（混合邻接）：满足下列条件的任一个，则具有 $V$ 中数值的 $p$ 和 $q$ 是 $m$ 连接的。

(i)  $q$ 在中 $N_4(p)$

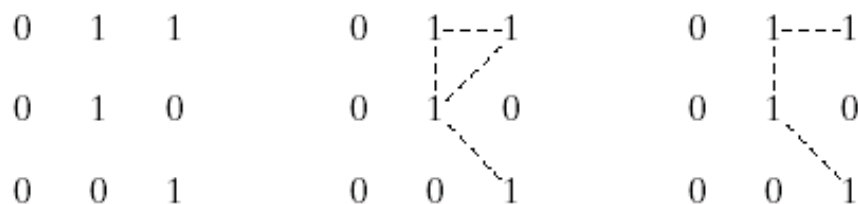
(ii)  $q$ 在 $N_D(p)$ 中，且集合 $N_4(p) \cap N_4(q)$ 中没有 $V$ 值的像素。



## 2.5 Some Basic Relationships Between Pixels

注意：混合邻接是8邻接的改进，为了消除8邻接的二义性。例如图2.26。

两个集合邻接的概念：如果集合 $S^1$ 中的某些像素和 $S^2$ 中的某些像素邻接，则称这两个集合是邻接的。这里说的邻接指的是4、8或者 $m$ 邻接。连通性等概念暂时略过。



a b c

**FIGURE 2.26** (a) Arrangement of pixels; (b) pixels that are 8-adjacent (shown dashed) to the center pixel; (c)  $m$ -adjacency.

距离度量（见书本）

### 2.6 线性和非线性操作

$$H(af + bg) = aH(f) + bH(g)$$

其中， $f$ 、 $g$ 代表图像， $H$ 代表操作。

作业：2.11、2.15