Chapter 2
Digital Image Fundamentals

第二章：数字图像基础

Notice

请各位同学在学者网的课程平台上注册，使用课程平台提交作业等。网站：

http://www.scholat.com/ 或

http://tel.scholat.com/

注册后到“课程”栏目里查找“数字图像处理”，申请加入，填写学号姓名等信息，选择班级，使用密码“dip2013”加入到相应班级中。

在最终的分析中，所有知识皆为历史。在抽象的意义下，所有科学皆为数学。在理性的世界里，所有判断皆为统计。

——黄文璋（国立高雄大学应用数学系教授）

登高而招，臂非加长也，而见者远；顺风而呼，声非加疾也，而闻者彰。假舆马者，非利足也，而致千里；假舟楫者，非能水也，而绝江河。

——荀子劝学

- **How to present your homework and project**
- **Resolution concepts (including spatial and gray level)**

第二章作业见课程网站：

http://gitl.sysu.edu.cn/dip/lib/exe/fetch.php/hw1.pdf

Preview

2.1 Visual Perception

2.2 Light and the Electromagnetic Spectrum

2.3 Image Sensing and Acquisition

2.4 Image Sampling and Quantization

2.5 Some Basic Relationships between Pixels

2.6 Mathematical Tools for Dip

2.1 Visual Perception

——Key points:

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（cones）和杆状体 (rods)。

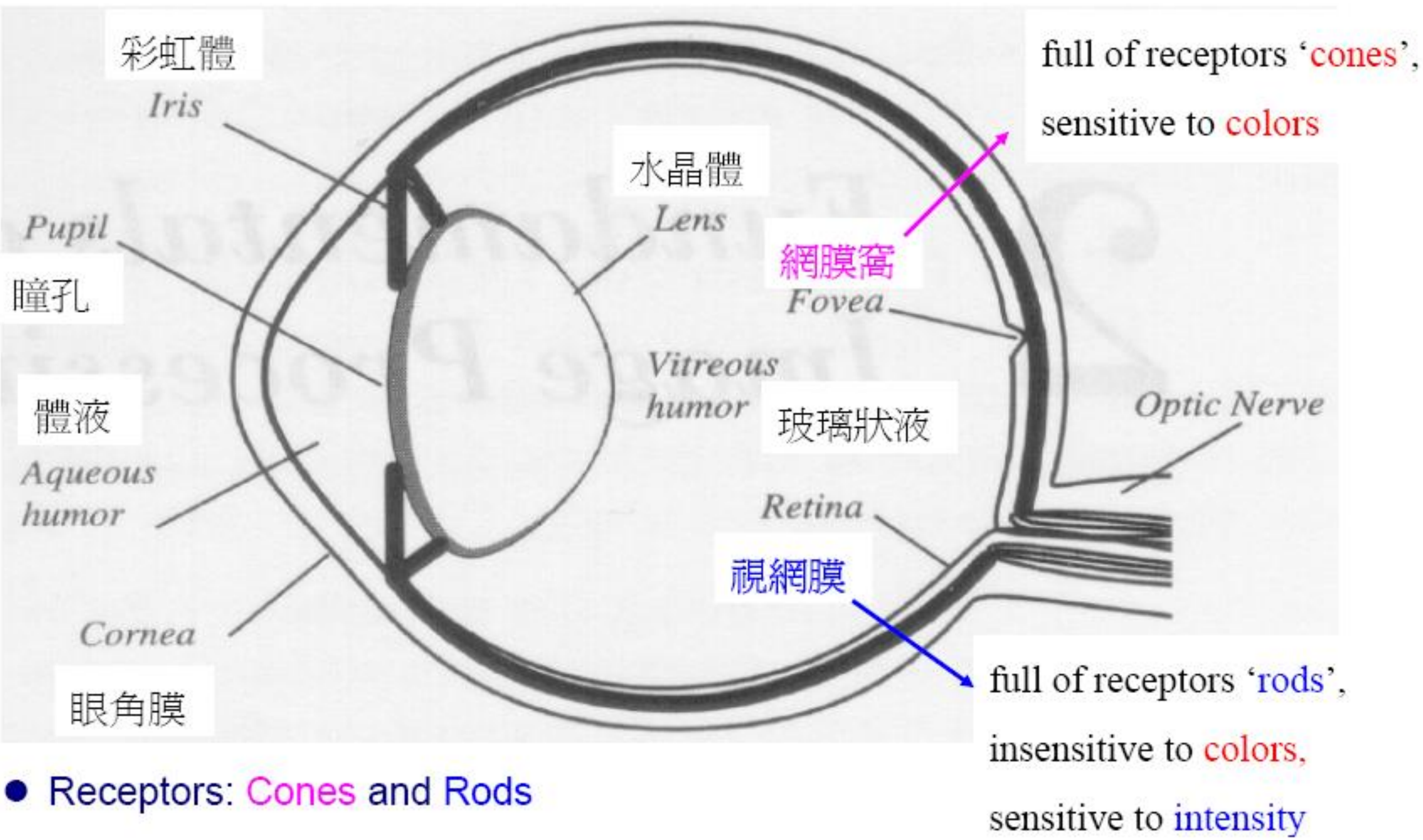
每只眼睛**锥状体**数目6-7百万，每个锥状体都连接到自己的神经末端，对色彩敏感，分布于视网膜中央凹部分，称为白昼视觉或亮光视觉。

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。

分辨细节的基本能力：每平方毫米150,000个像素，最高敏感区的接收阵列近似为1.5mm×1.5mm，相当于一个接收列阵不大于5mm×5mm的中等分辨率的电耦合元件（CCD: Charge-coupled device）的成像芯片。

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。

Structure of human eyes



- Receptors: Cones and Rods

Structure of human eyes

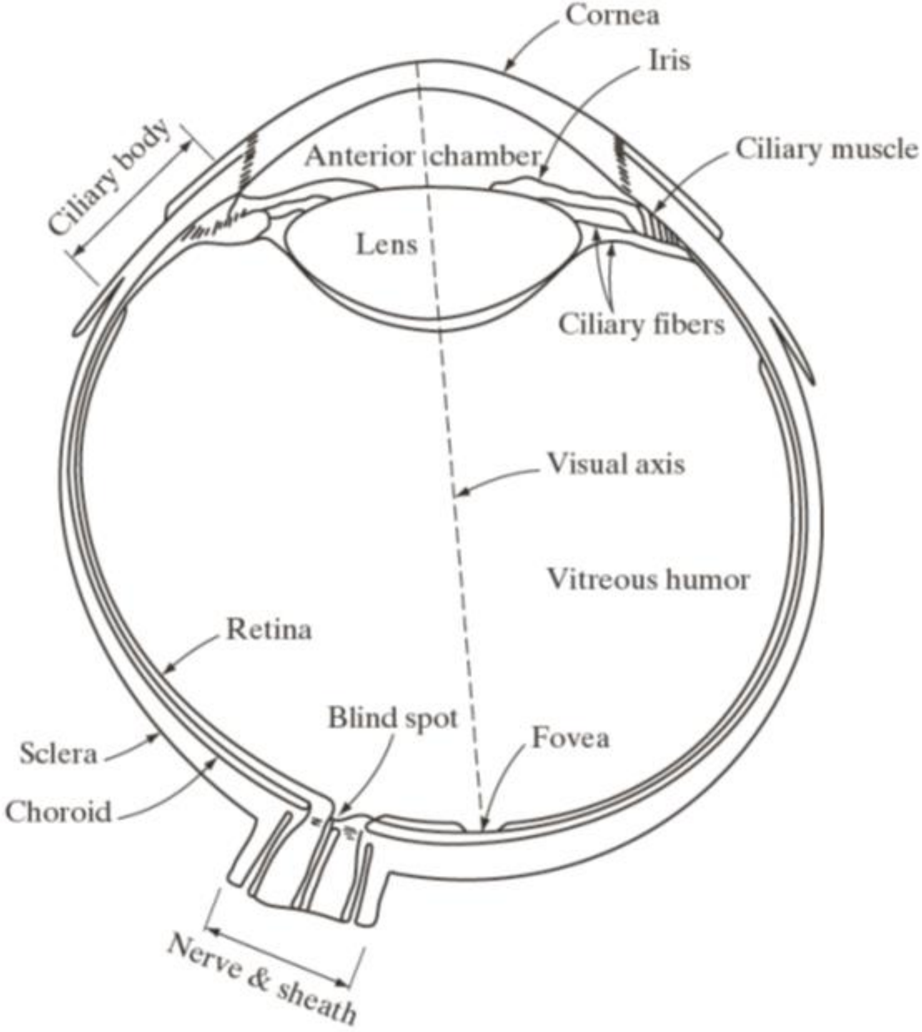
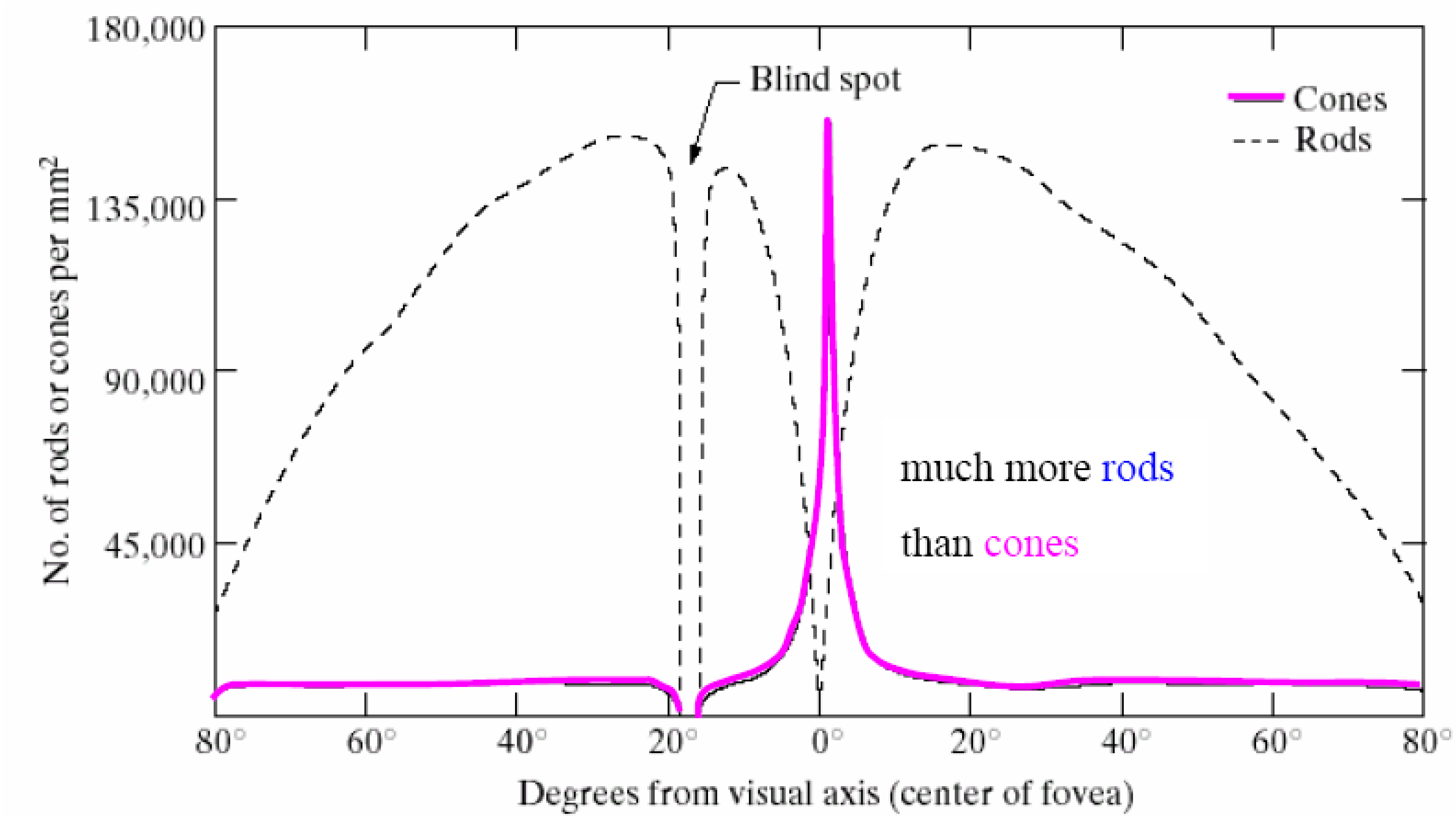
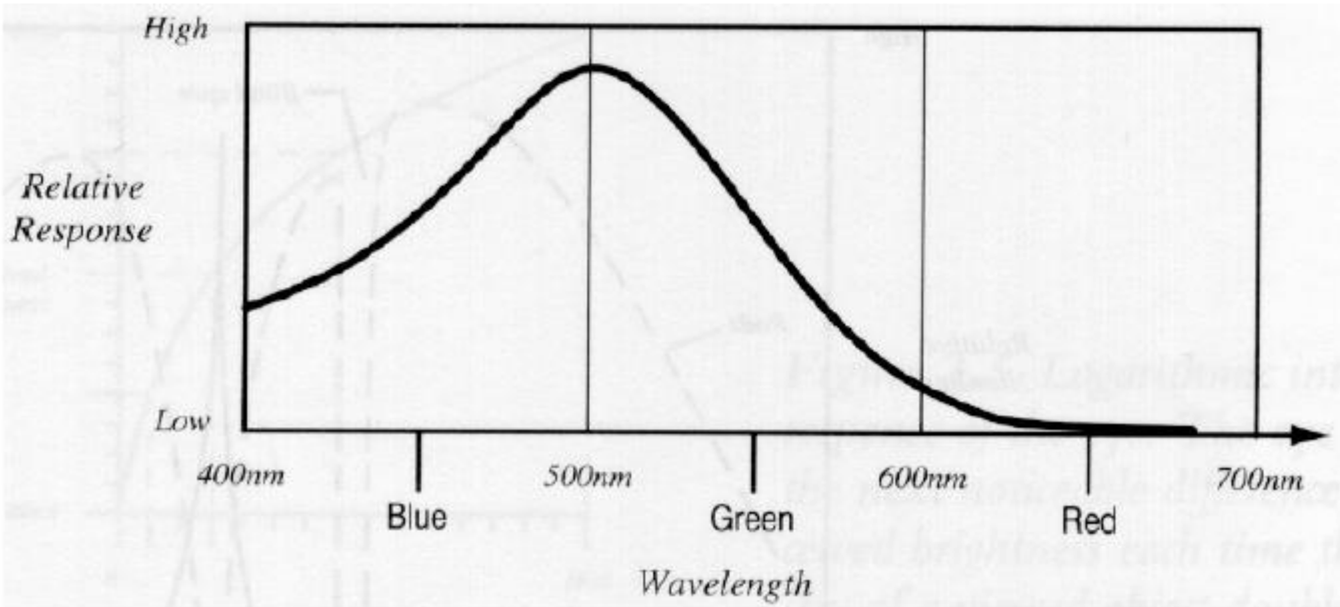


FIGURE 2.1 Simplified diagram of a cross section of the human eye.

- **Distribution of Cones and Rods on retina**



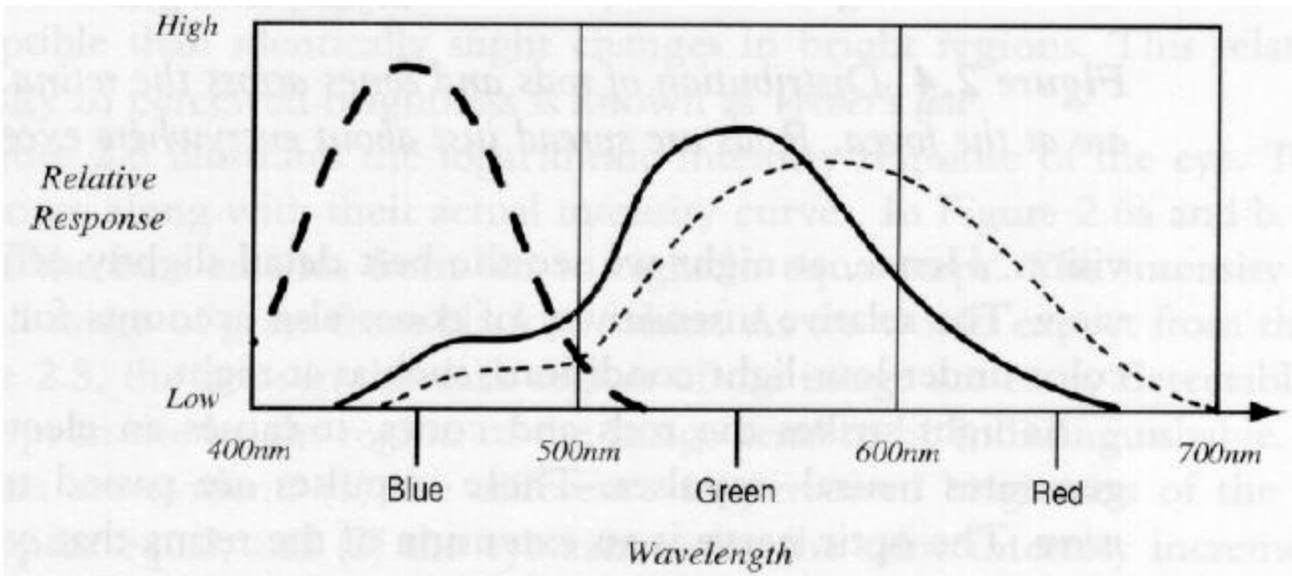
● Rods are highly responsive to light, but see only a single band of light, and therefore cannot discriminate color



Peripheral Vision: (rods)

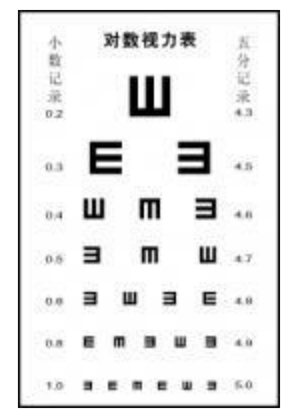
- Responsible for perception under low- intensity Light.
- Low spatial resolution
- Higher sensitivity to temporal variations
- Less sensitive to color

- **Cones** are less responsive, but **see three distinct color spectral bands** of light, enabling color vision.



Fovea Vision: (cones)

- Responsible for perception under high intensity light
- Very high spatial resolution (near 1 min. visual acuity)
- Less sensitive to temporal variations
- Highly sensitive to color



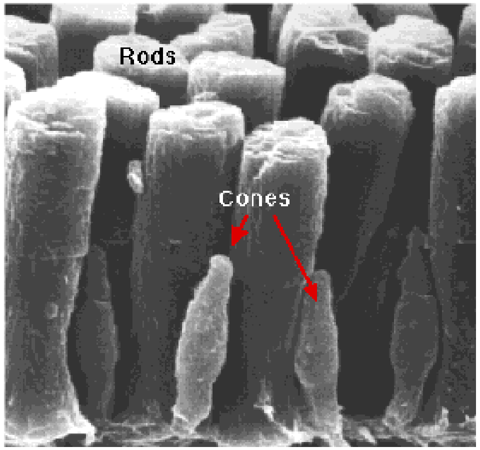
Light Detection: Rods and Cones

Rods:

-120 million rods in retina

-1000X more light sensitive than Cones

- Discriminate B/W brightness in low illumination
- Short wave-length sensitive



Cones:

- 6-7 million Cones in the retina
- Responsible for high-resolution vision
- Discriminate Colors
- Three types of color sensors (64% red, 32%, 2% blue)
- Sensitive to any combination of three colors

● Image formation in the eye

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。

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17mm到14mm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。当物体距离超过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。

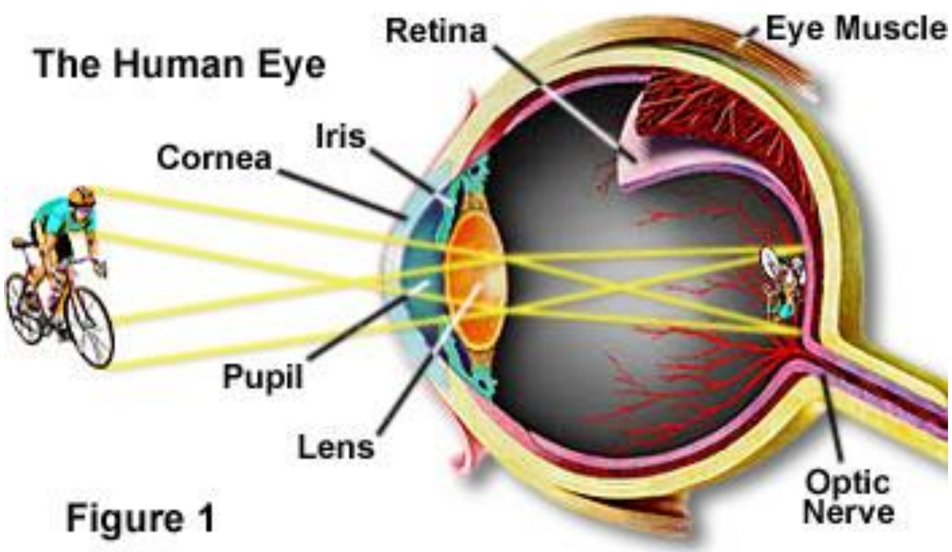
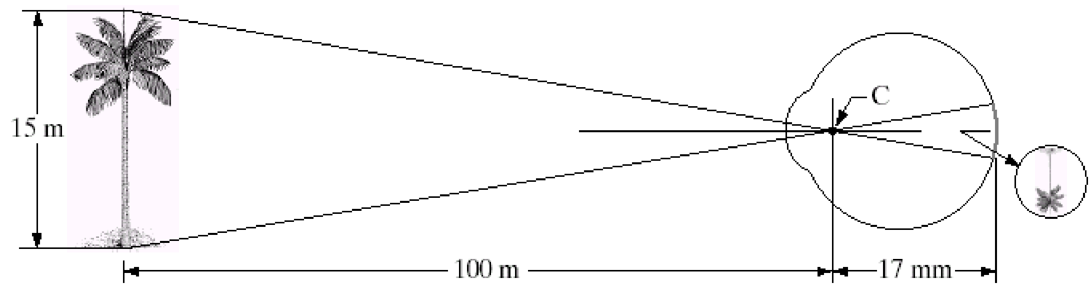


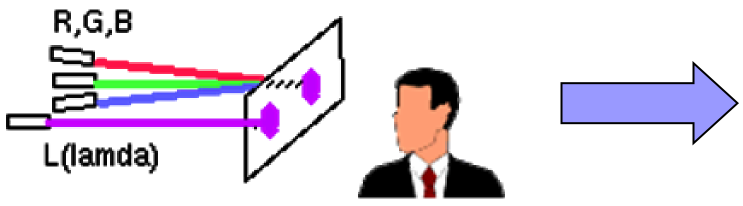
Figure 1

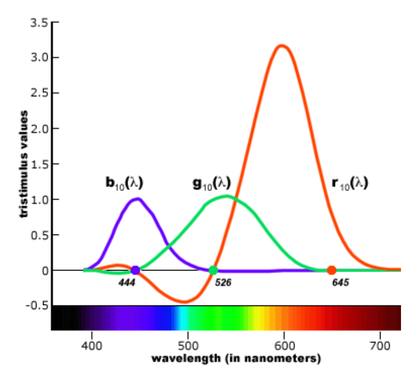
● Image formation in the eye

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



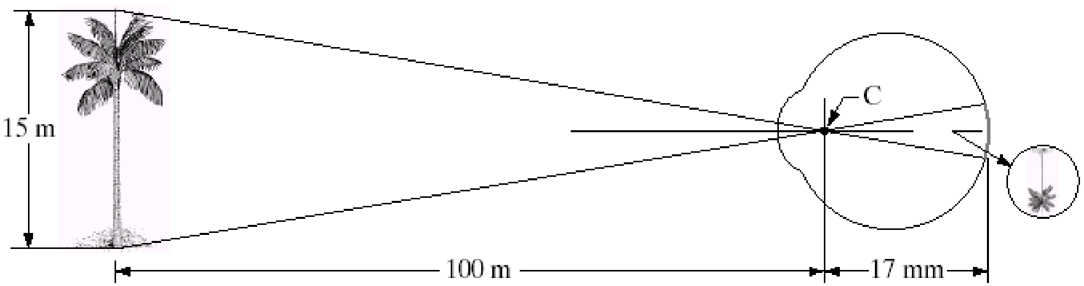
The focal length of the eye various from 14 mm to 17mm.



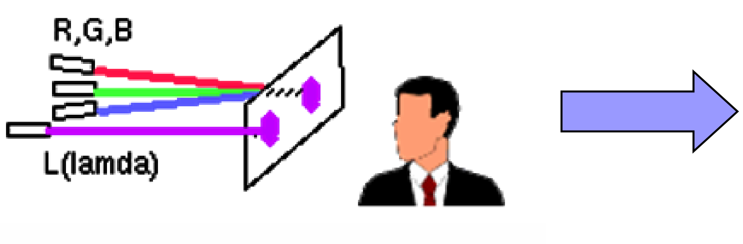


● Image formation in the eye

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



The focal length of the eye various from 14 mm to 17mm.



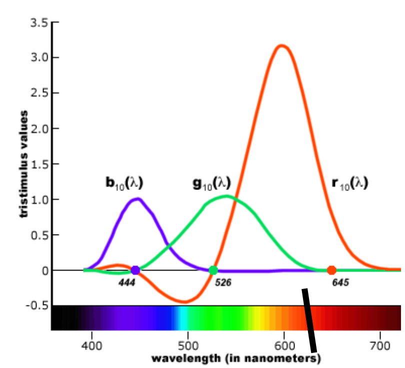
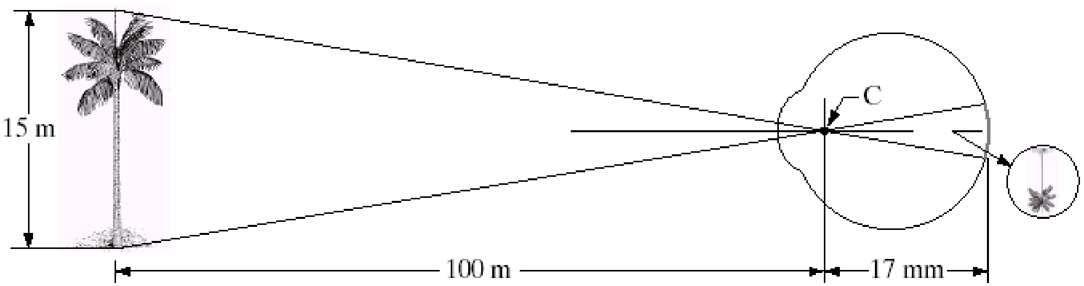
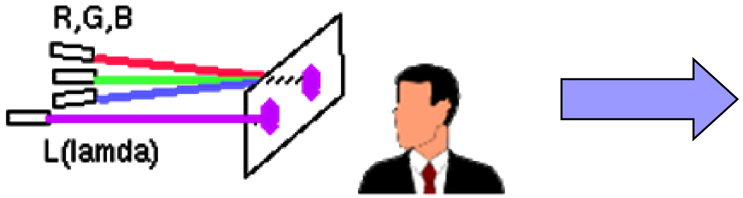


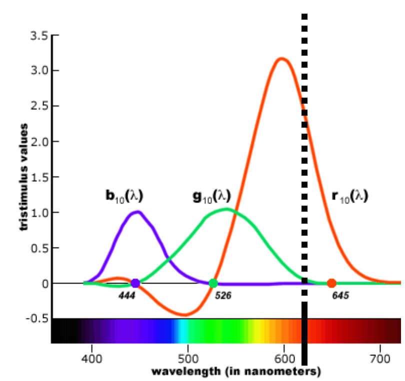
Image formation in the eye

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



The focal length of the eye various from 14 mm to 17mm.





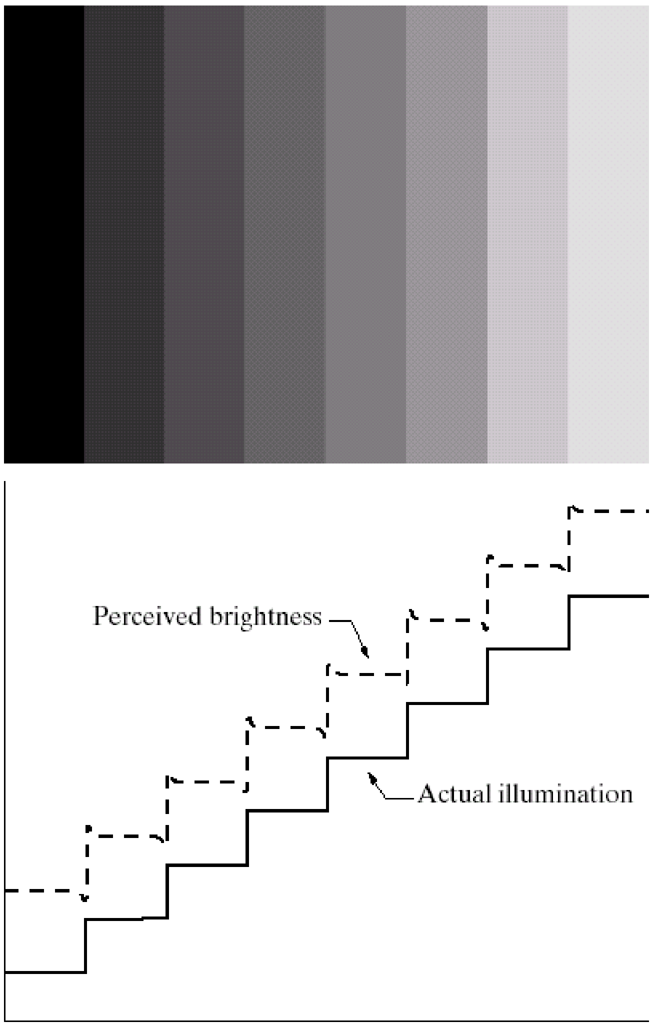
眼睛对亮度的适应和鉴别

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；
- 但人的视觉绝对不可能同时在一个范围内工作，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少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。

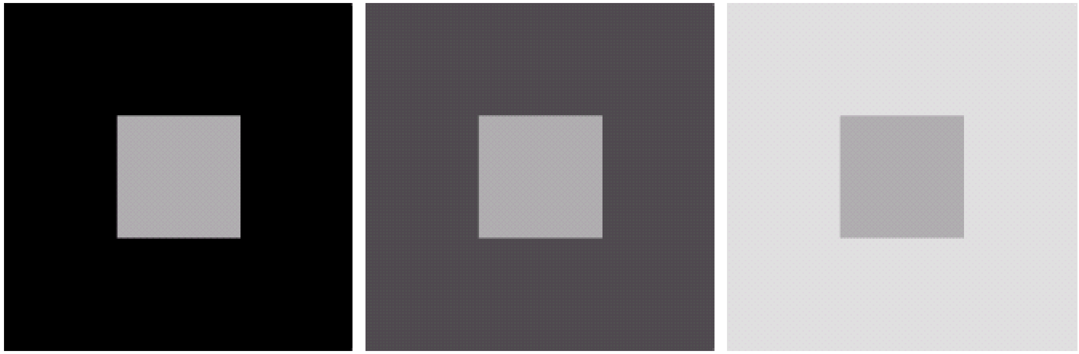
两个现象说明了亮度感觉不是简单的强度函数

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）



- The true change in intensity is a step function, as shown by solid line.
- Human eye tends to increase the contrast at the edge of two steps by undershooting and overshooting, as indicated by dash line.
- This phenomena is called the ‘Mach bands’(1865).

第二个现象称为同时对比现象(Simultaneous contrast)



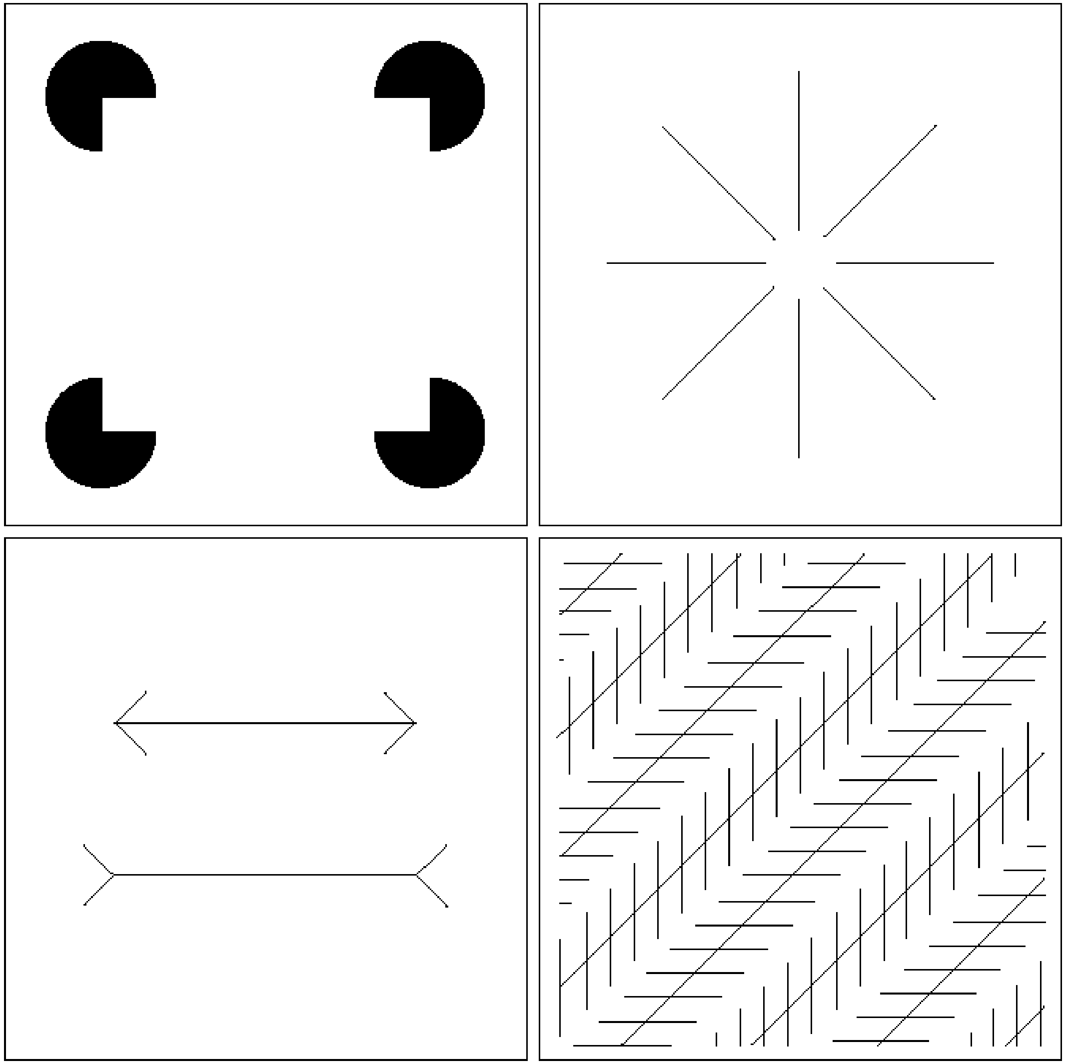
a b c

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.

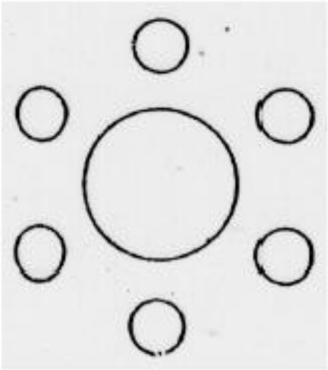
视觉错觉：这是人类视觉系统的一个特性，尚未完全了解。

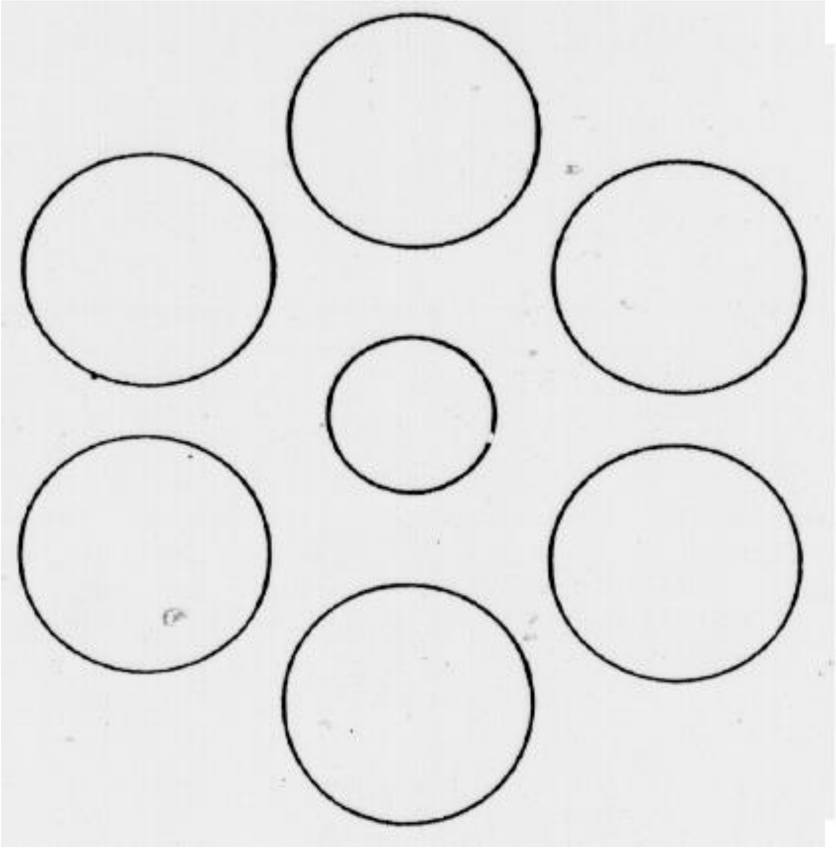
a b
c d

FIGURE 2.9 Some well-known optical illusions.



Ebbinghaus Illusion

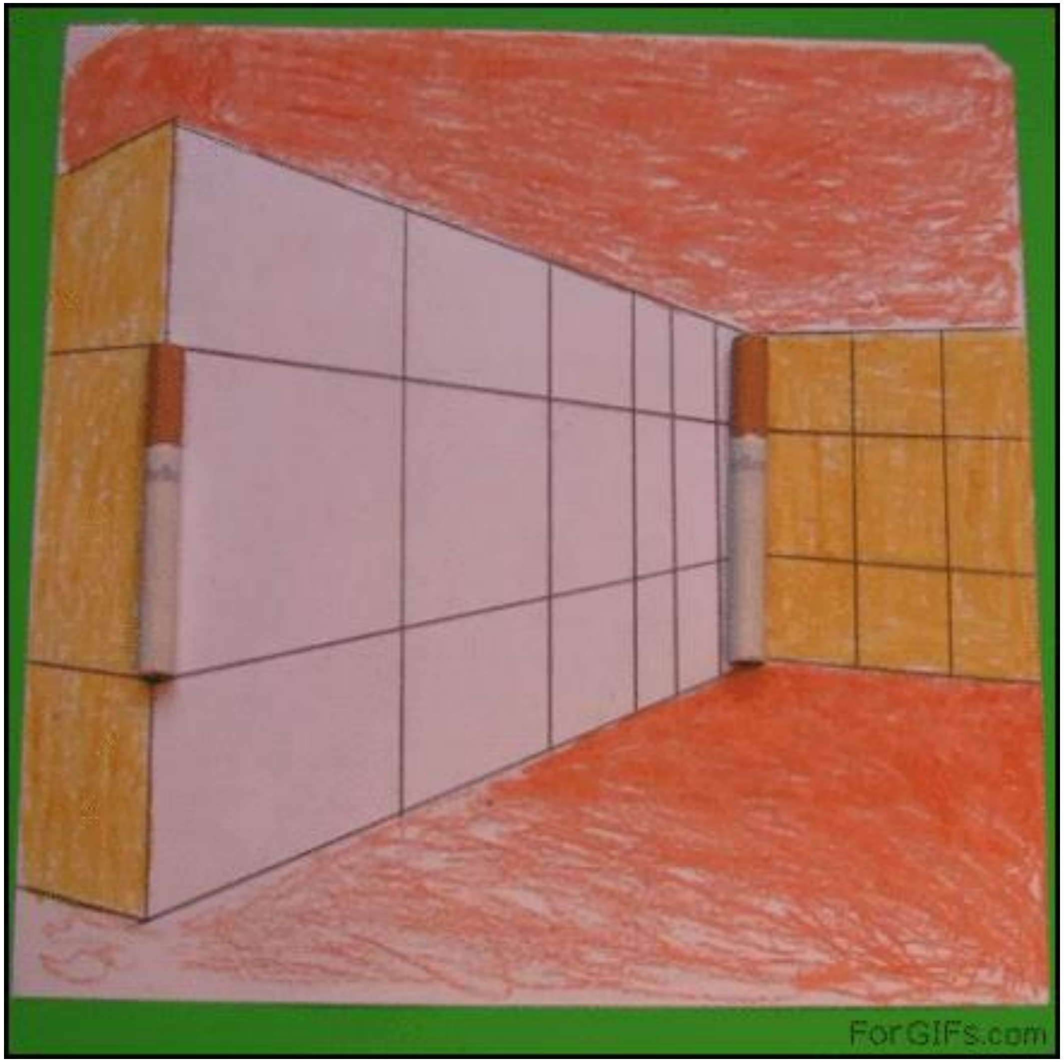




For fun


蘑菇街
mogujie.com
发现时尚，分享购物乐趣
请看列表并且说出颜色而不是单词
黄 蓝 橙
黑 红 绿
紫 黄 红
橙 绿 黑
蓝 红 紫
绿 蓝 橙
左右（脑）冲突
你的右脑尝试着说出颜色，但是
你的左脑坚持要阅读
下载蘑菇街客户端，逛街更方便
weibo.com/mogujie

For fun



Grand Teton Park

2.2 Light and Electromagnetic spectrum

Sir Isaac Newton 1666 discovered: when a beam of sunlight is passed through a glass prism, the emerging beam of light is not white again but contains a continuous spectrum of colors ranging from violet at one end to red at the other .

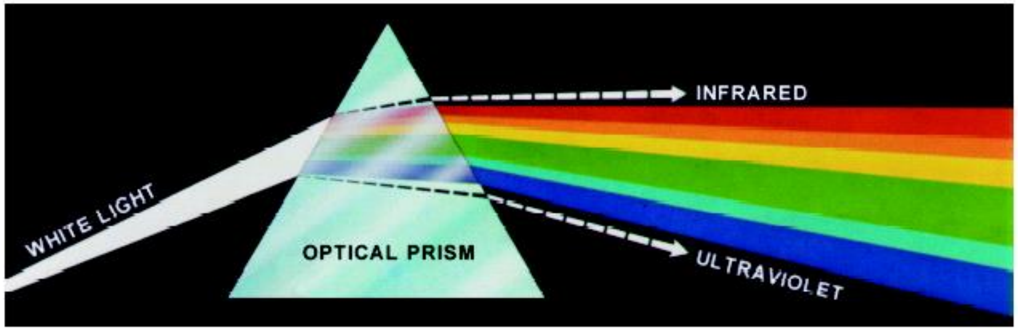


FIGURE 6.1 Color spectrum seen by passing white light through a prism. (Courtesy of the General Electric Co., Lamp Business Division.)

光波等同于电磁波，反过来也成立。可见光的范围：电磁波400~700nm（1nm=10^{-9}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。

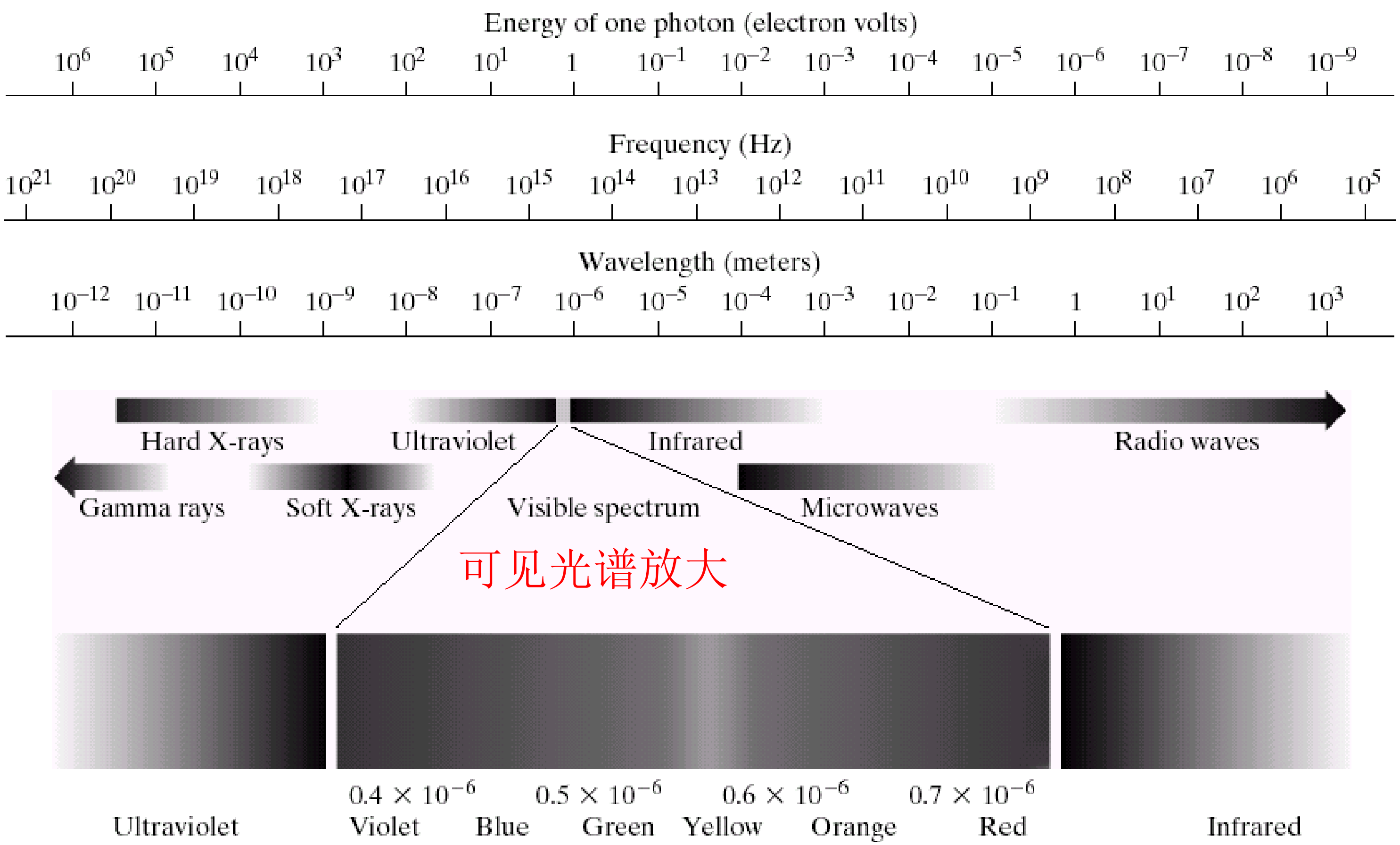


FIGURE 2.10 The electromagnetic spectrum. The visible spectrum is shown zoomed to facilitate explanation, but note that the visible spectrum is a rather narrow portion of the EM spectrum.

2.2 Light and Electromagnetic Spectrum

　　电磁波可以用波长、频率或能量来描述。波长（λ）和频率（v）关系式为：

$$\lambda = v / c$$

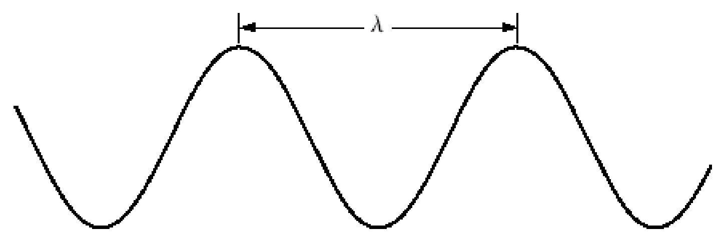
其中c是光速（2.998x10^8 m/s），电磁波谱(的分量由下式给出

$$E=hv$$

h是普普克朗常数。能量与频率成正比。

　　电磁波可以看成是以波长 λ 传播的正弦波。

FIGURE 2.11 Graphical representation of one wavelength.



关于可见光的一些基本事实：

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）。

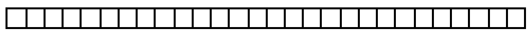
单色光的属性是它的强度或大小。灰度级通常是用来描绘单色光的强度，它的范围从黑到灰，最后到白。

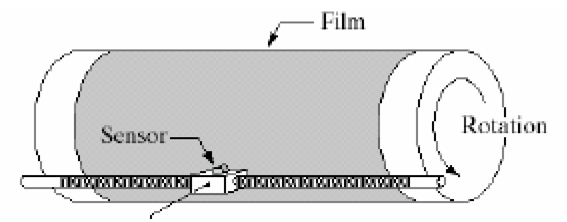
通常有三个基本量用于描绘彩色光源的质量：发光强度（从光源流出的能量）、光通量（观察者从光源感受的能量，例如：远红外光有实际的能量，但光通量为零）和亮度（亮度是描绘光感受的主观描绘，它实际上不能测量，包含无色的强度的概念，并且也是描述彩色感觉的参数之一。

波长和被“观察”物体的大小之间的关系？

2.3 Image sensing and Acquisition

- **Line camera(不要求掌握）**





➢ Line camera is used when the object moves (rotate) along the other direction to complete a 2-D scanning.

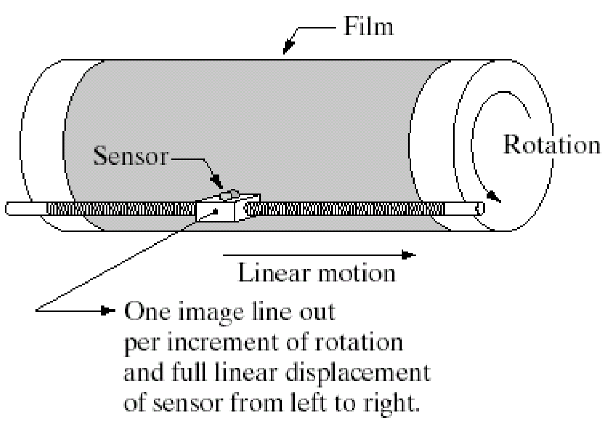
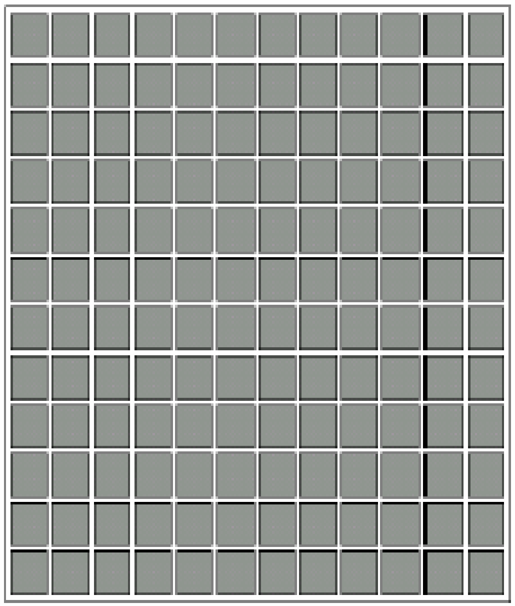


FIGURE 2.13 Combining a single sensor with motion to generate a 2-D image.

不要求掌握

- Field camera



- Field camera is usually carried by an X-Y table to complete a 2-D scanning.
- The scene that a field camera capture in each shoot is called a 'field of view (FOV)'.

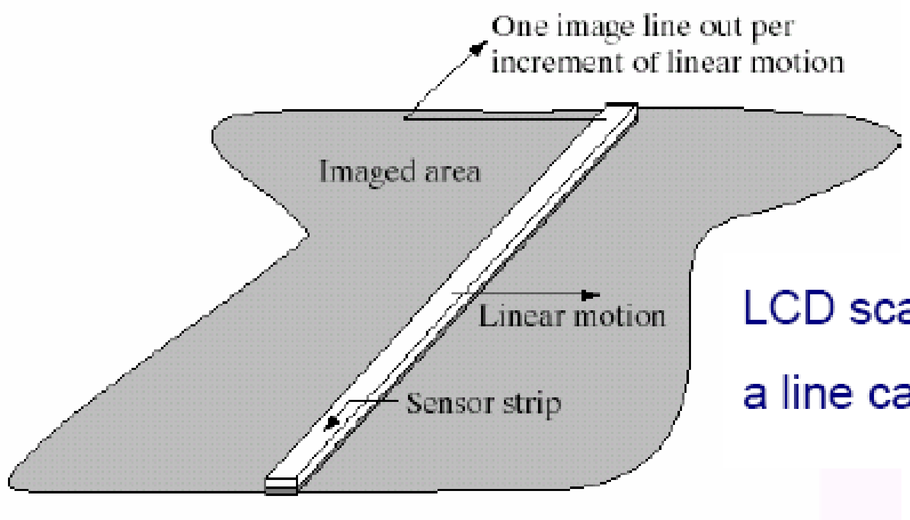
- Comparison

The motion companying a line camera is one-dimensional, hence allow faster scanning and smoother motion (less blurring).

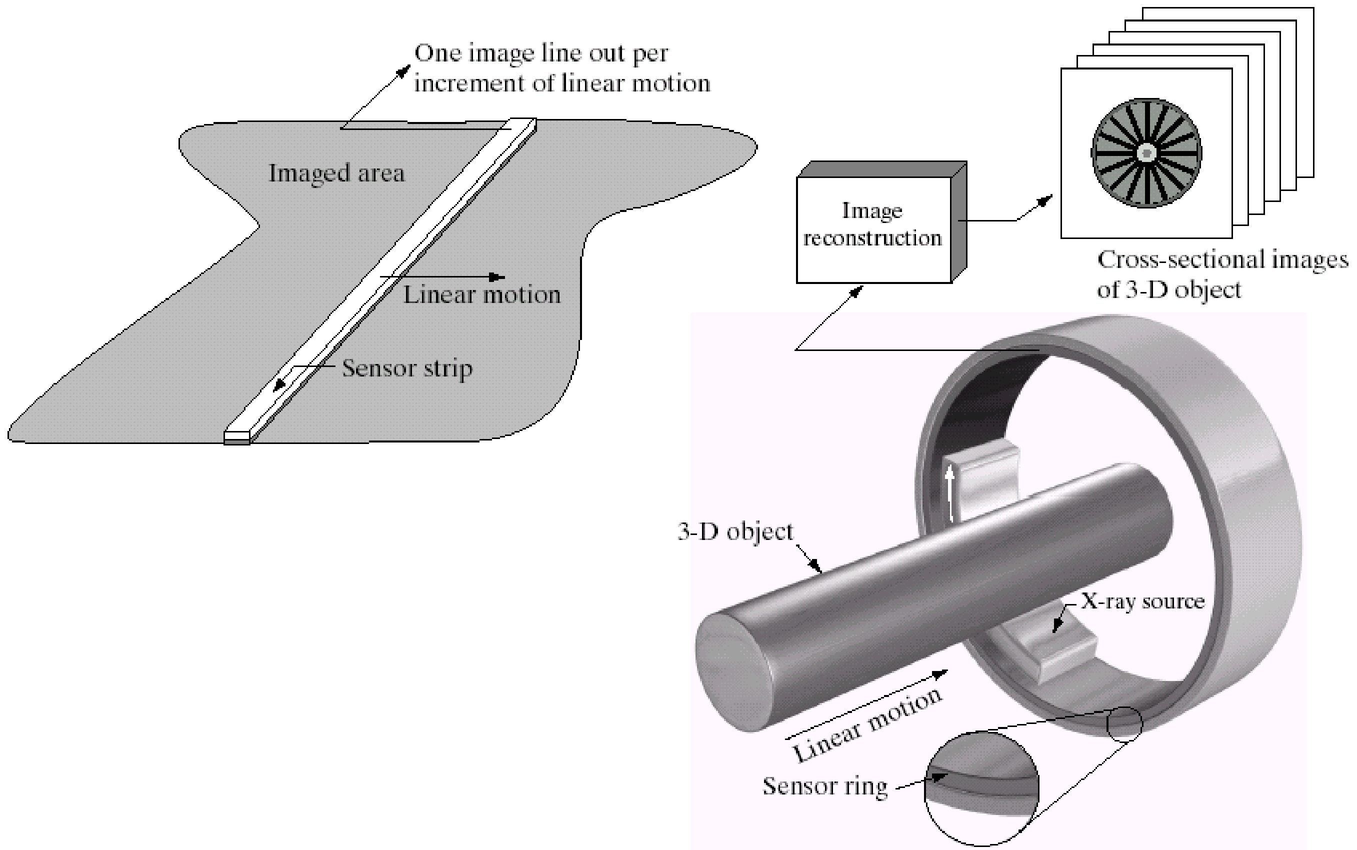
不要求掌握

However, a line camera adopts very limited illumination arrangement (mainly illuminate from a specific angle, typically from top).

Automatic inspection (AOI) in SMT requires complicated illumination arrangements, hence uses field camera. AOI in LCD industry requires very fast scanning and good image quality and can adapt to fixed lighting, hence uses line camera.



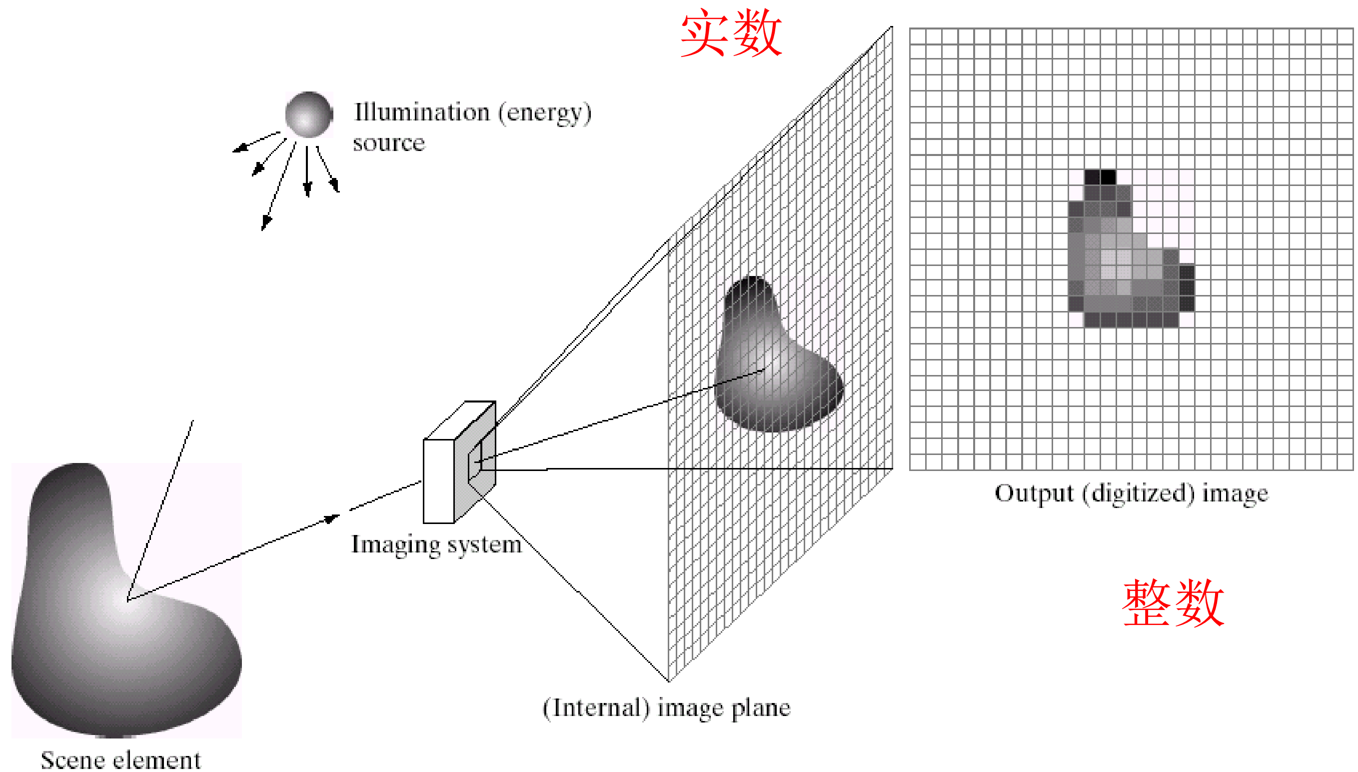
LCD scanning using a line camera



a b

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

- **Schematic diagram showing image capturing process**



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（全反射）和1（全吸收）之间。

2.4 Image Sampling and Quantization

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。这就包含了两种数字化处理，取样（时空域）和量化（光色强度等）。

- **Concepts of Image Sampling and Quantization**

这两个概念和数值计算方法中的网格划分、有限维逼近（如分片常数逼近）和量化等概念非常类似(learned before）。对应于空间分辨率和灰度分辨率。

关键词：分辨率（resolution)

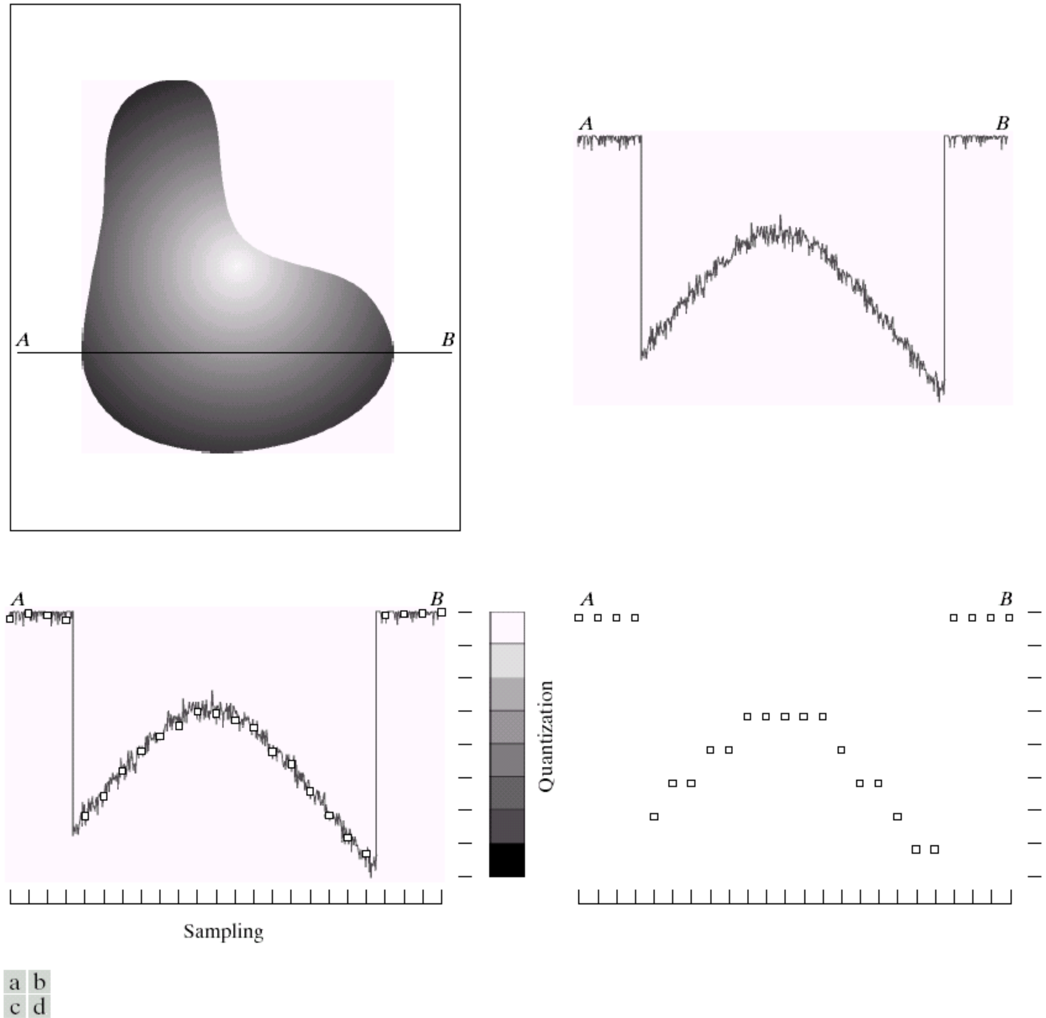
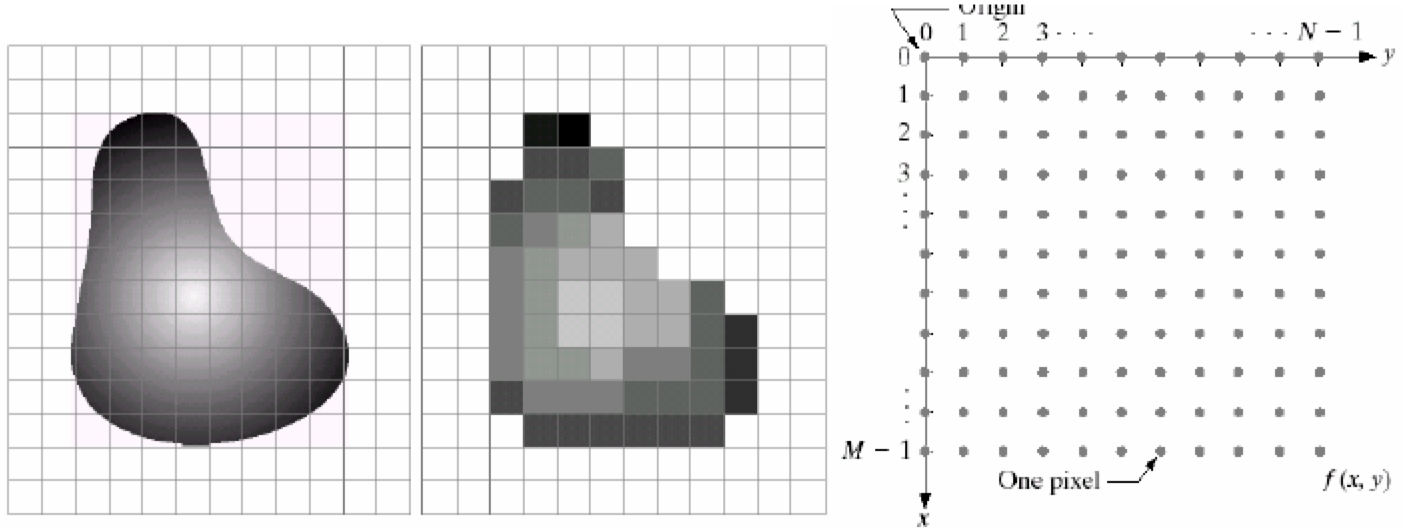


FIGURE 2.16 Generating a digital image. (a) Continuous image. (b) A scan line from *A* to *B* in the continuous image, used to illustrate the concepts of sampling and quantization. (c) Sampling and quantization. (d) Digital scan line.

- Sampling and quantized using 8-bits grey level



The image is sampled (with fixed spatial resolution/frequency) and quantized into 8-bits grey level, then the grey level (of the image) is recorded as a matrix with the coordinate system shown on the upper right.

- Sampled image can be represented by a two-D matrix

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

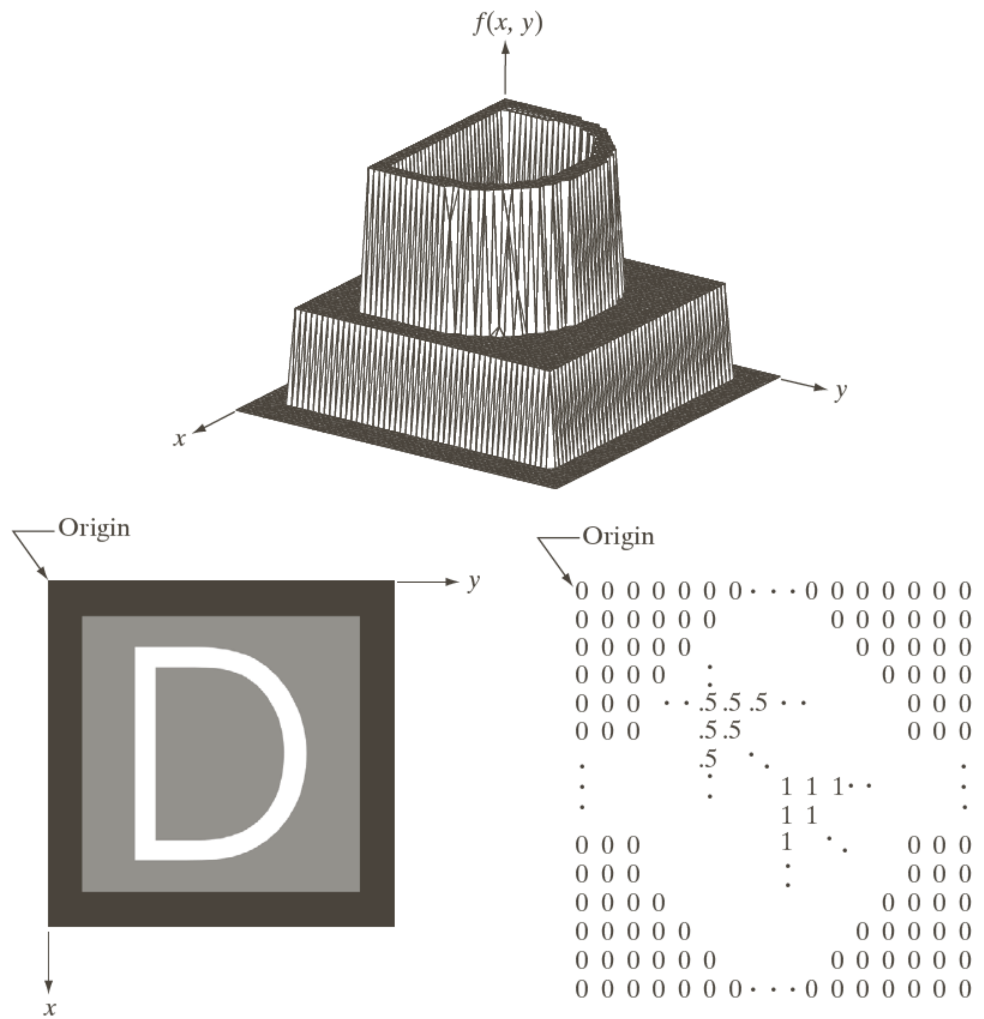
or

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

二维矩阵是表示数字图像的重要数学形式。一幅M×N的图像可以表示为矩阵：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”，“值”是这一“位置”像素的颜色或者强度。

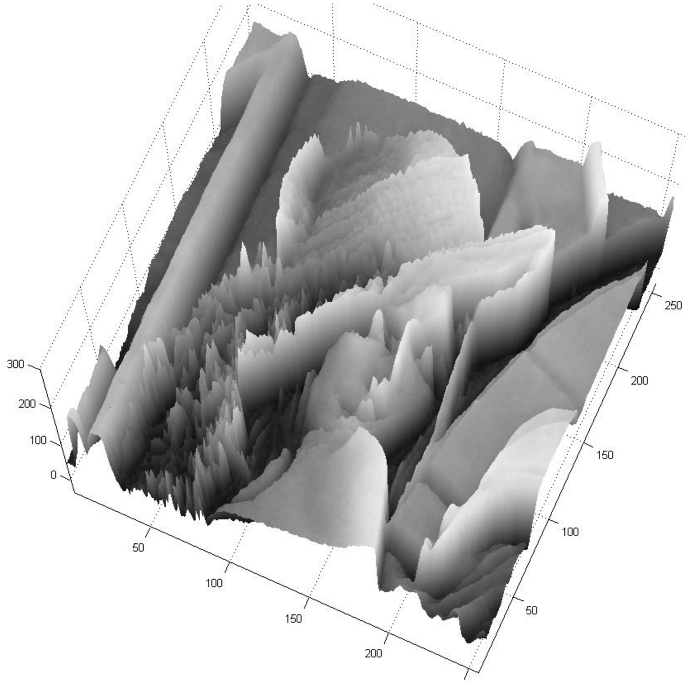
2.4 Image Sampling and Quantization

- Sampled image can be represented by a two-D matrix



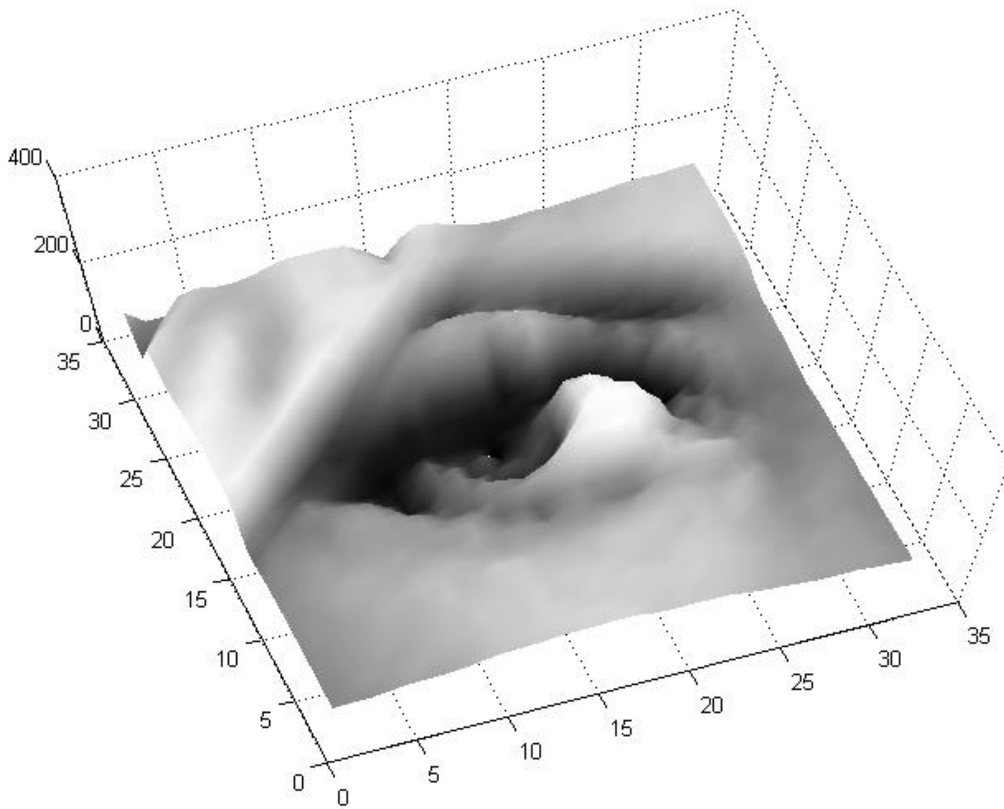
a
b c

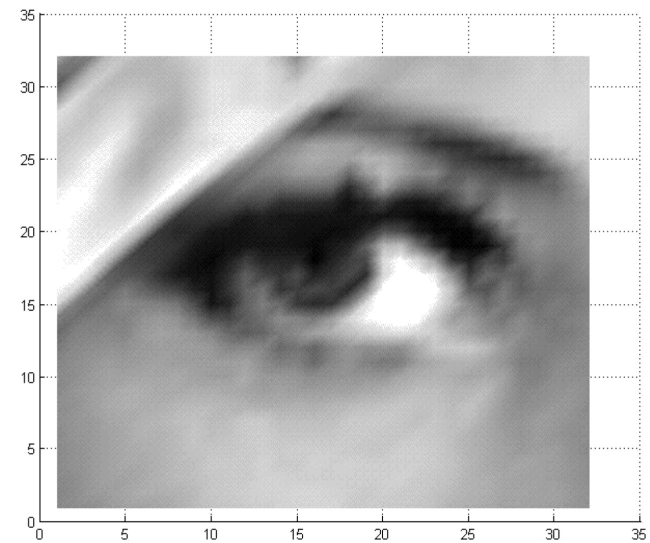
FIGURE 2.18
(a) Image plotted as a surface. (b) Image displayed as a visual intensity array. (c) Image shown as a 2-D numerical array (0, .5, and 1 represent black, gray, and white, respectively).

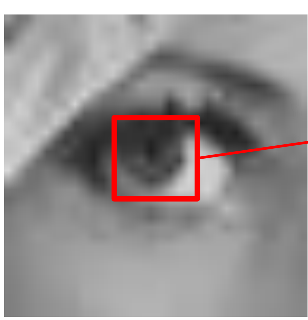


8x8

57	54	52	50	49	50	51	50
52	53	48	46	45	47	62	90
50	49	48	43	54	60	98	163
60	62	60	35	58	77	85	183
81	72	89	58	70	85	67	182
114	71	77	86	80	63	121	205
120	108	74	63	73	117	187	207
114	132	126	118	139	179	198	203







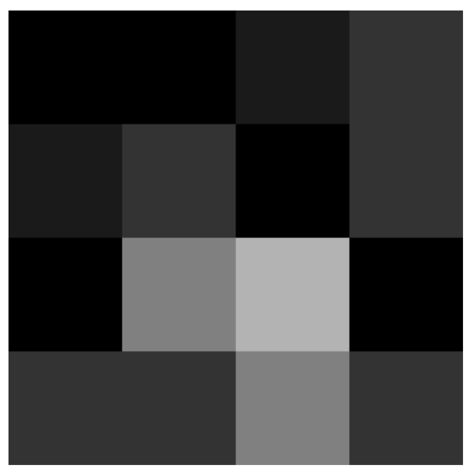
8x8

57	54	52	50	49	50	51	50
52	53	48	46	45	47	62	90
50	49	48	43	54	60	98	163
60	62	60	35	58	77	85	183
81	72	89	58	70	85	67	182
114	71	77	86	80	63	121	205
120	108	74	63	73	117	187	207
114	132	126	118	139	179	198	203

- Sampled image can be represented by a two-D matrix

Representation of gray value

$$\begin{bmatrix} 0 & 0 & 1 & 2 \\ 1 & 2 & 0 & 2 \\ 0 & 5 & 7 & 0 \\ 2 & 2 & 5 & 2 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 0 & 0 & 1/7 & 2/7 \\ 1/7 & 2/7 & 0 & 2/7 \\ 0 & 5/7 & 1 & 0 \\ 2/7 & 2/7 & 5/7 & 2/7 \end{bmatrix}$$



存储问题：

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×N×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

- **Spatial and Gray-Level Resolution**

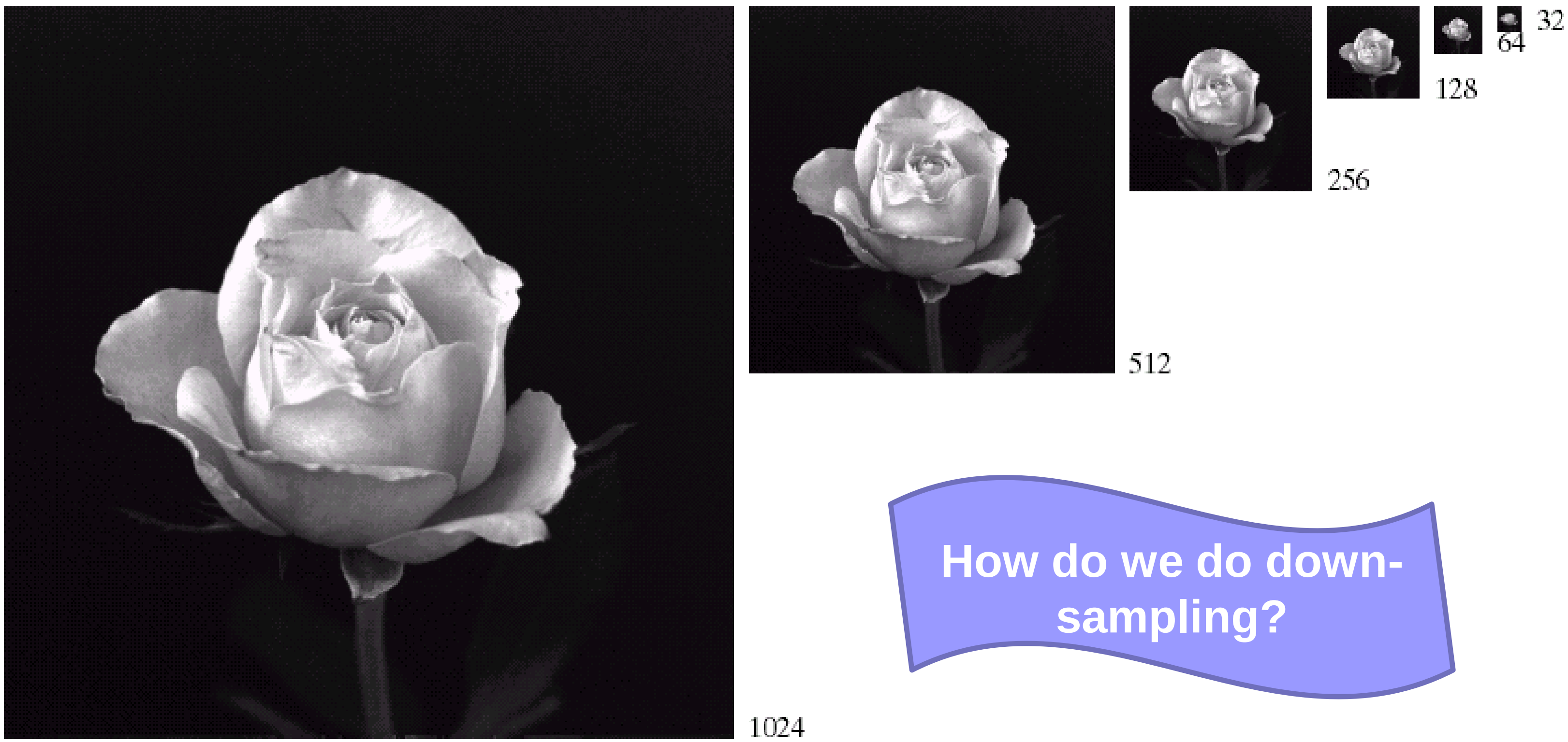
Different Types of Sampling

- Sampling： continuous image to digital (or discrete) image, so called digitalization;
- Down Sampling: high resolution image to low resolution image;
- Up Sampling: low resolution to “high resolution”.

Spatial and Gray-Level Resolution

Down Sampling

Sampling down from 1024×1024 to 32×32



How do we do down-sampling?

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

a b
c d

FIGURE 2.20 Typical effects of reducing spatial resolution. Images shown at: (a) 1250 dpi, (b) 300 dpi, (c) 150 dpi, and (d) 72 dpi. The thin black borders were added for clarity. They are not part of the data.

Spatial and Gray-Level Resolution

Down Sampling

$$
\begin{bmatrix}
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2}
\end{bmatrix}
\begin{bmatrix}
0 & 0 & 1 & 1 & 2 & 2 & 3 & 3 \\
0 & 0 & 1 & 1 & 2 & 2 & 3 & 3 \\
4 & 4 & 5 & 5 & 6 & 6 & 7 & 7 \\
4 & 4 & 5 & 5 & 6 & 6 & 7 & 7 \\
8 & 8 & 9 & 9 & 10 & 10 & 11 & 11 \\
8 & 8 & 9 & 9 & 10 & 10 & 11 & 11 \\
12 & 12 & 13 & 13 & 14 & 14 & 15 & 15 \\
12 & 12 & 13 & 13 & 14 & 14 & 15 & 15
\end{bmatrix}
$$

$$
\Rightarrow
\begin{bmatrix}
0 & 0 & 1 & 1 & 2 & 2 & 3 & 3 \\
4 & 4 & 5 & 5 & 6 & 6 & 7 & 7 \\
8 & 8 & 9 & 9 & 10 & 10 & 11 & 11 \\
12 & 12 & 13 & 13 & 14 & 14 & 15 & 15
\end{bmatrix}
$$

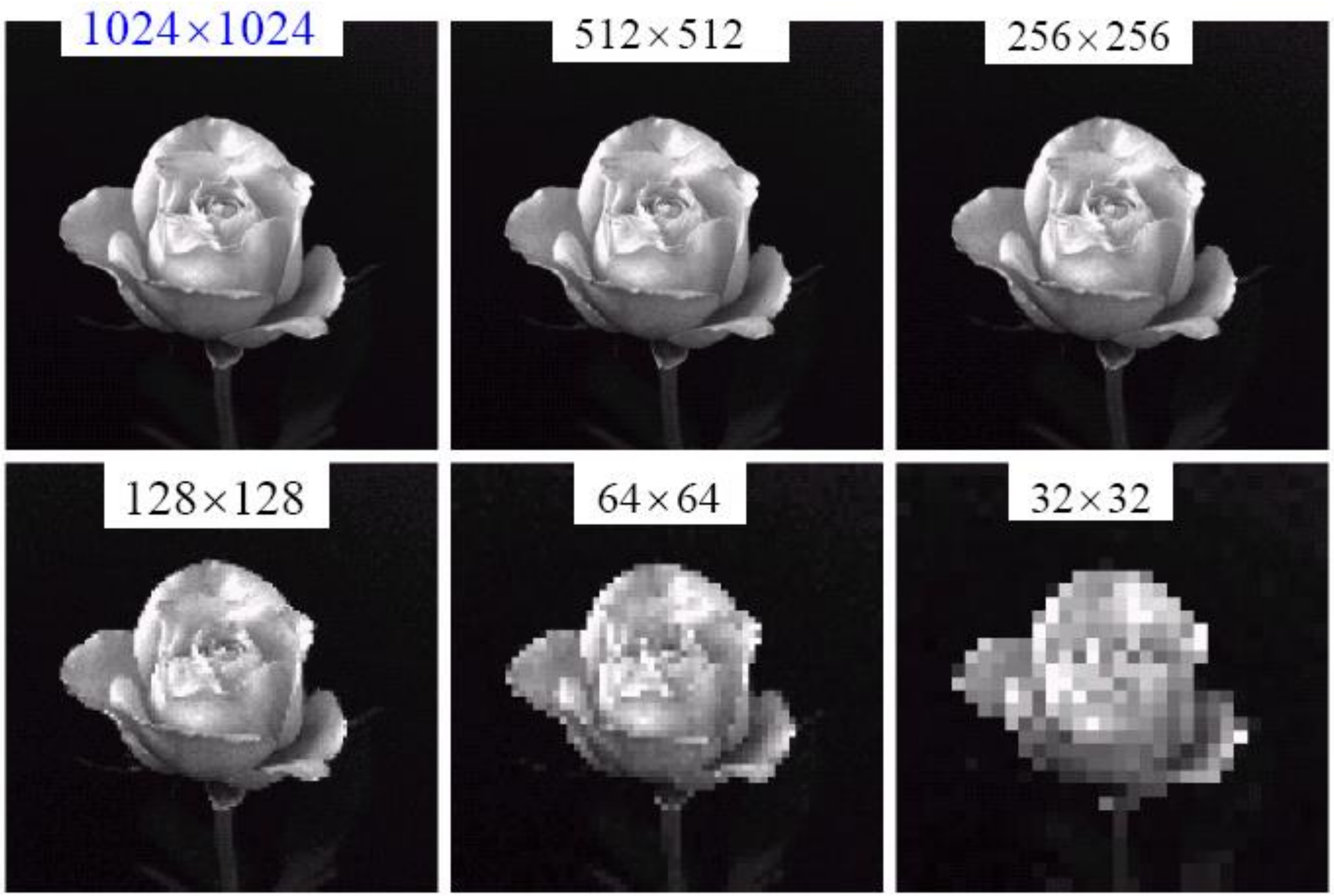
Spatial and Gray-Level Resolution

Down Sampling

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 2 & 2 & 3 & 3 \\ 4 & 4 & 5 & 5 & 6 & 6 & 7 & 7 \\ 8 & 8 & 9 & 9 & 10 & 10 & 11 & 11 \\ 12 & 12 & 13 & 13 & 14 & 14 & 15 & 15 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

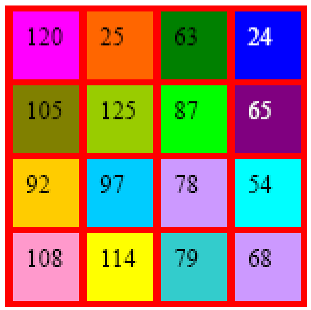
$$\Rightarrow \begin{bmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \\ 8 & 9 & 10 & 11 \\ 12 & 13 & 14 & 15 \end{bmatrix}$$

- Up sampling

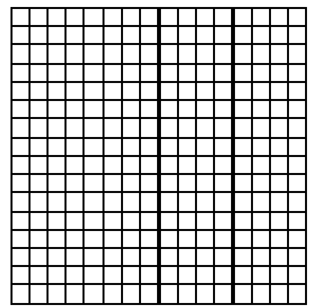


Zooming (up-sampling) a digital image

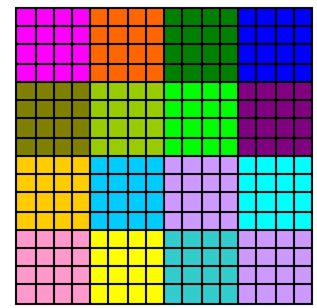
Zooming (up-sampling) by nearest neighbor interpolation



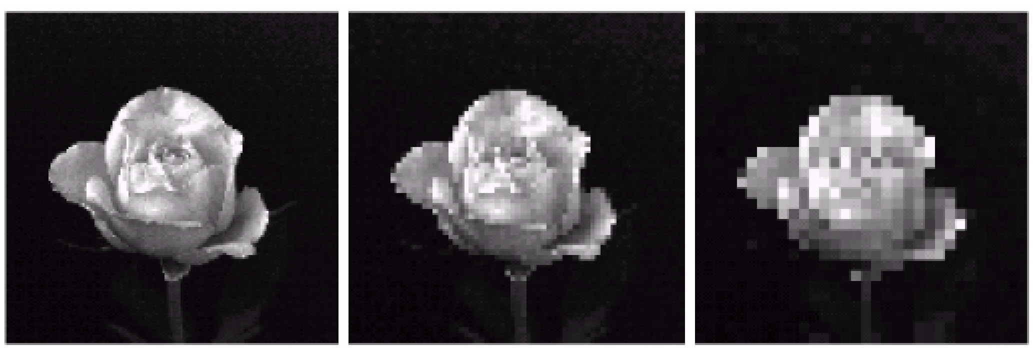
low resolution image



high resolution grids (up-sample)

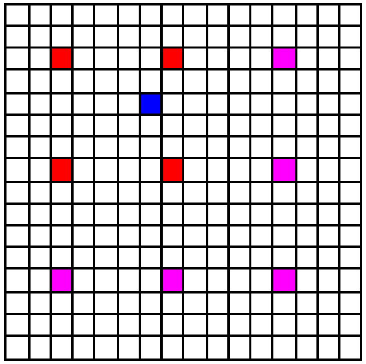


high resolution image



Up-sampled into 1024x1024 from 128x128, 64x64, and 32x32 by nearest neighbor

➢ Zooming (up-sampling) by bilinear interpolation



Let the coordinate of the blue pixel be (x, y), and its output intensity be assigned $v(x, y)$, then the bilinear interpolation model is given by

$$v(x, y) = ax + by + cxy + d$$

Intensity of the four original pixels nearest to the blue pixel can be used to solve for four unknown a, b, c, d before solving v.

Up-sampled into 1024x1024 from 128x128, 64x64, and 32x32 by bilinear neighbor

a b c
d e f

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).

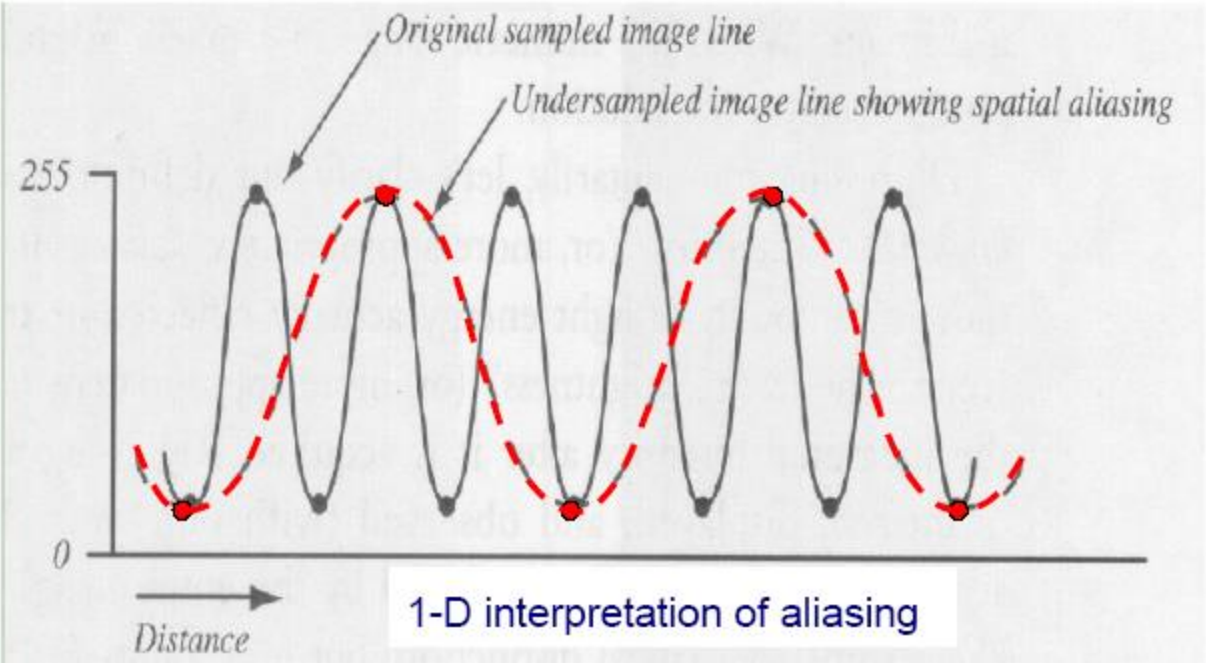
Question

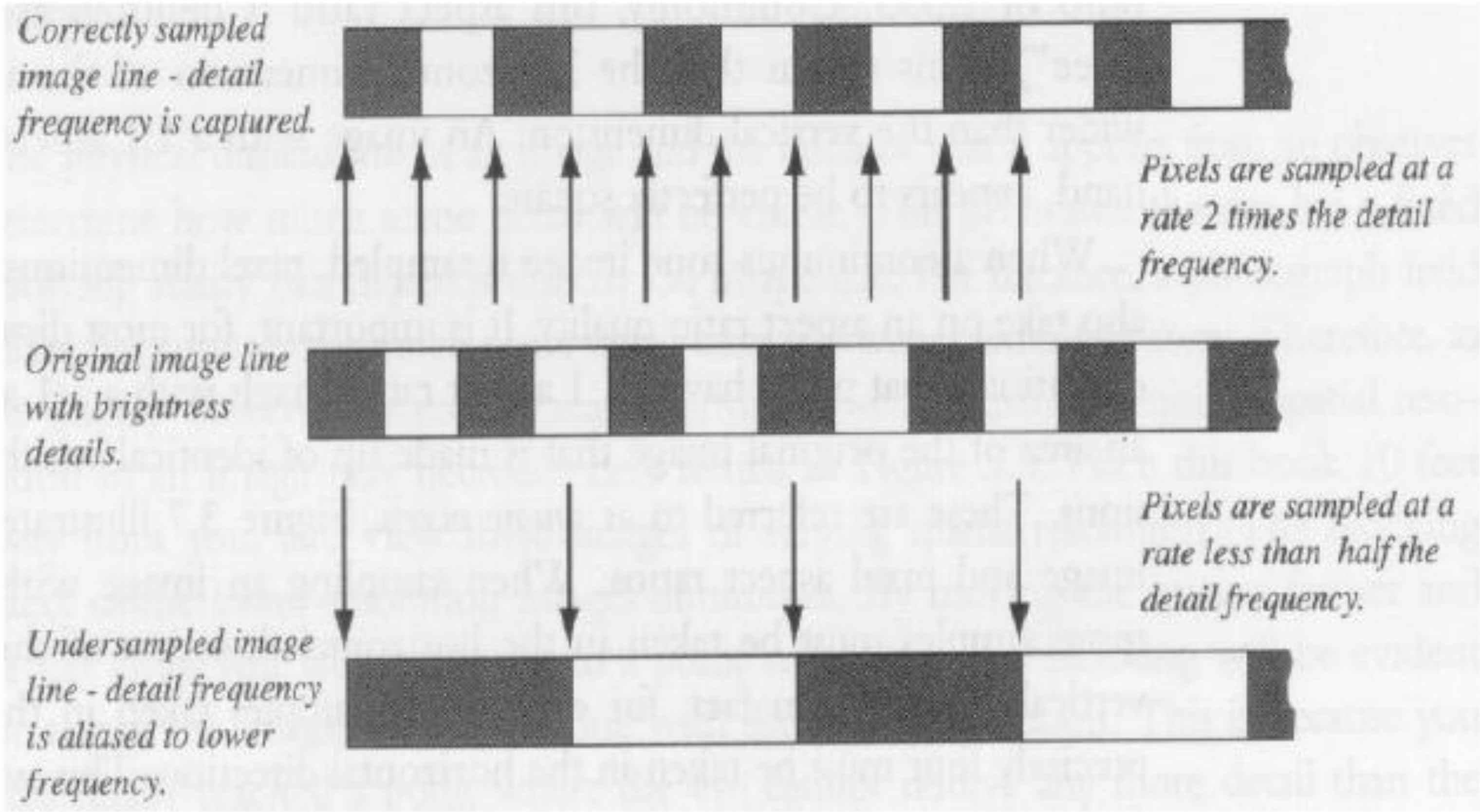
Can a down-sampled image be recovered as it is by

up-sampling?

下采样的图像能否通过上采样恢复原图?

- **Sampling, Aliasing and Moiré pattern**
  - ➢ Aliasing: high frequency pattern mis-interpreted as low frequency pattern, it occur when the image is sampled at too low a frequency.

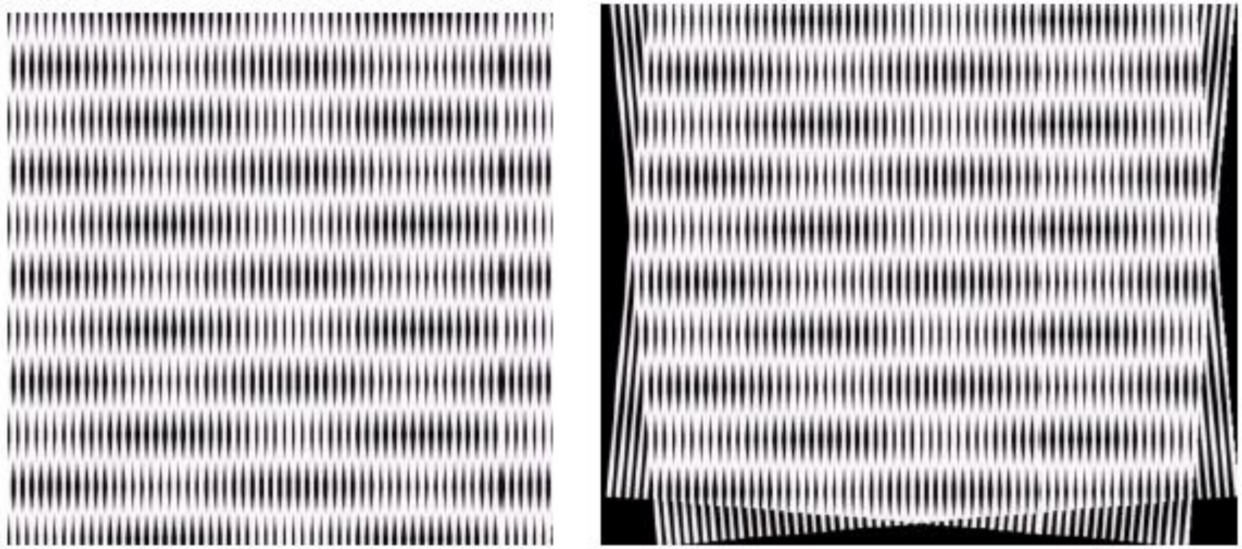




To avoid aliasing, the Nyquist sampling theorem has to be obeyed:

sampling frequency $\Omega_S \geq 2\Omega_P$ maximum frequency in the pattern

➢ Moiré pattern occur when two identical periodic pattern are overlapped with an intersection angle as shown below



One set of periodic lines can be perceived as the sampling pulses train to modulate the other pattern of periodic lines.

$\Omega_S = \Omega_P < 2\Omega_P$, aliasing occur, a pattern at a lower frequency (periodic horizontal lines) occurs called Moiré pattern.

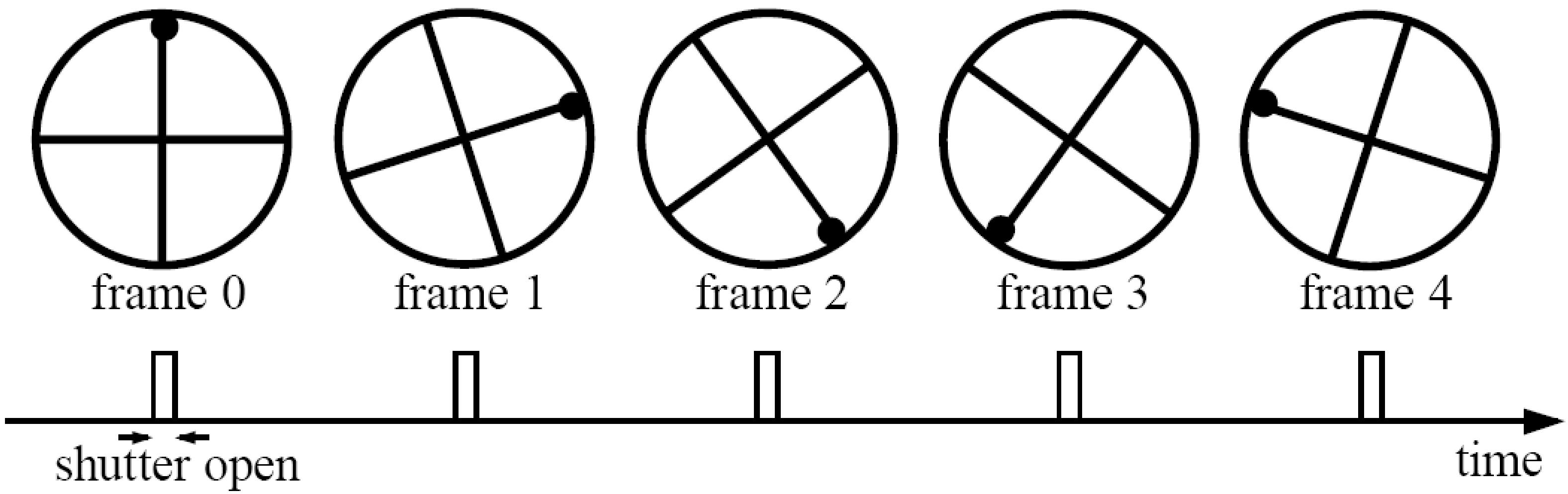
Properly sampled image of brick wall

Spatial aliasing in the form of a Moiré pattern

Really Bad in Video

Imagine a spoked wheel moving to the right (rotating clockwise). Mark wheel with dot so we can see what's happening.

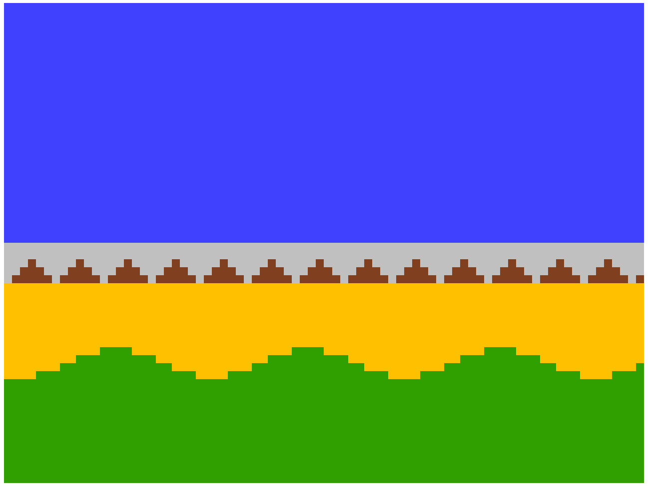
If camera shutter is only open for a fraction of a frame time (frame time = 1/30 sec. for video, 1/24 sec. for film):



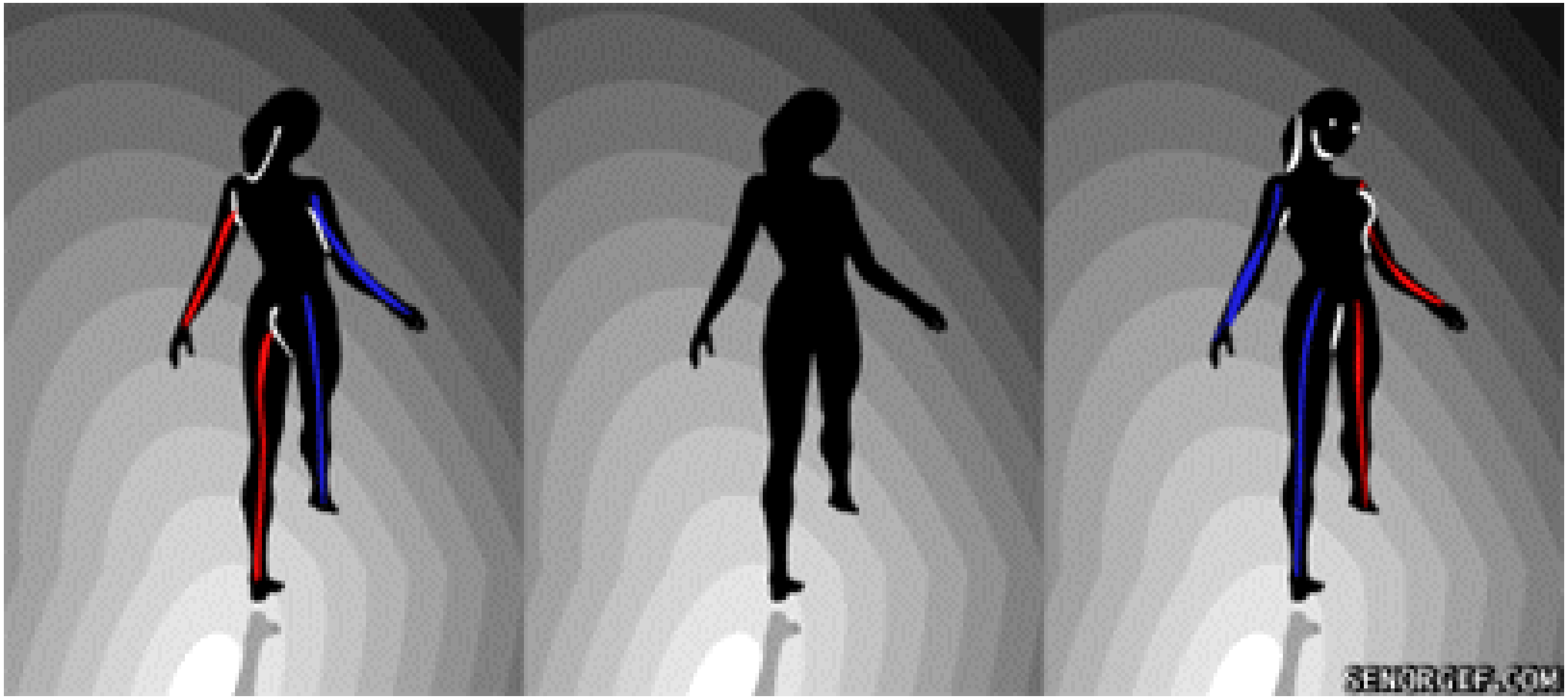
Without dot, wheel appears to be rotating slowly backwards! (counterclockwise)

Click here

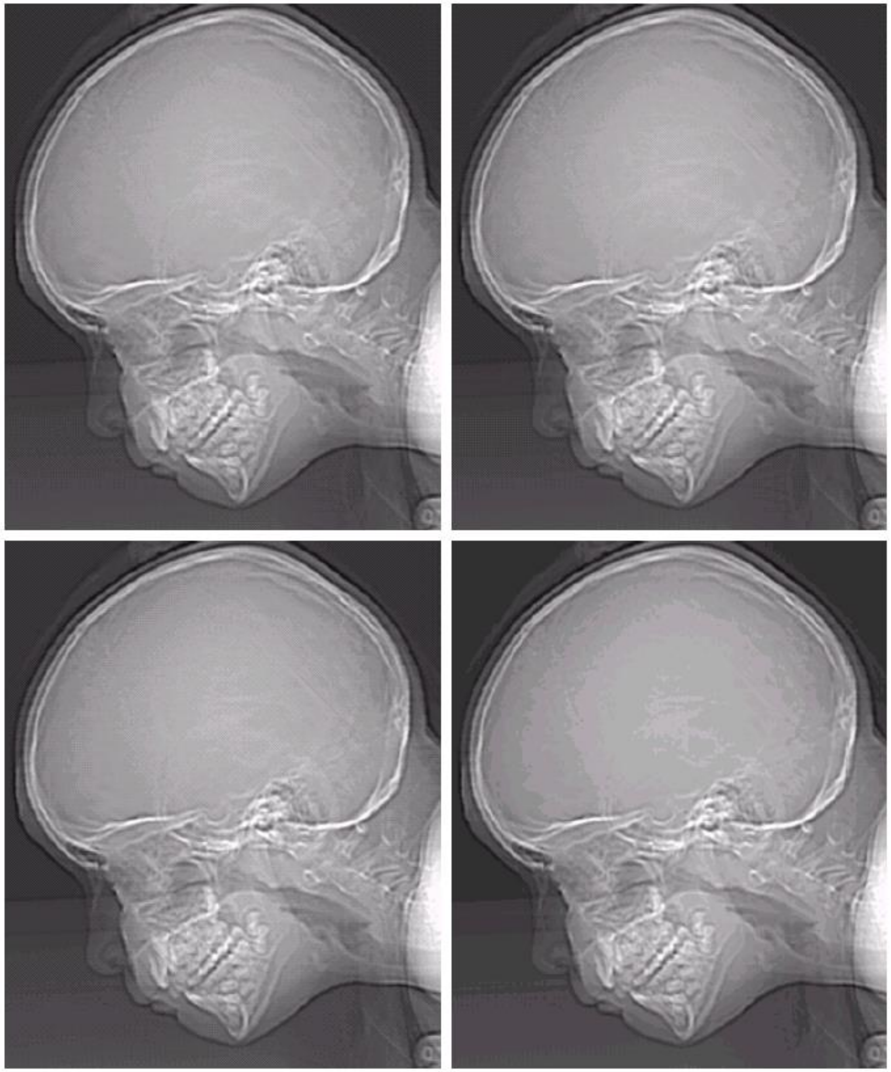
Really Bad in Video



Really Bad in Video

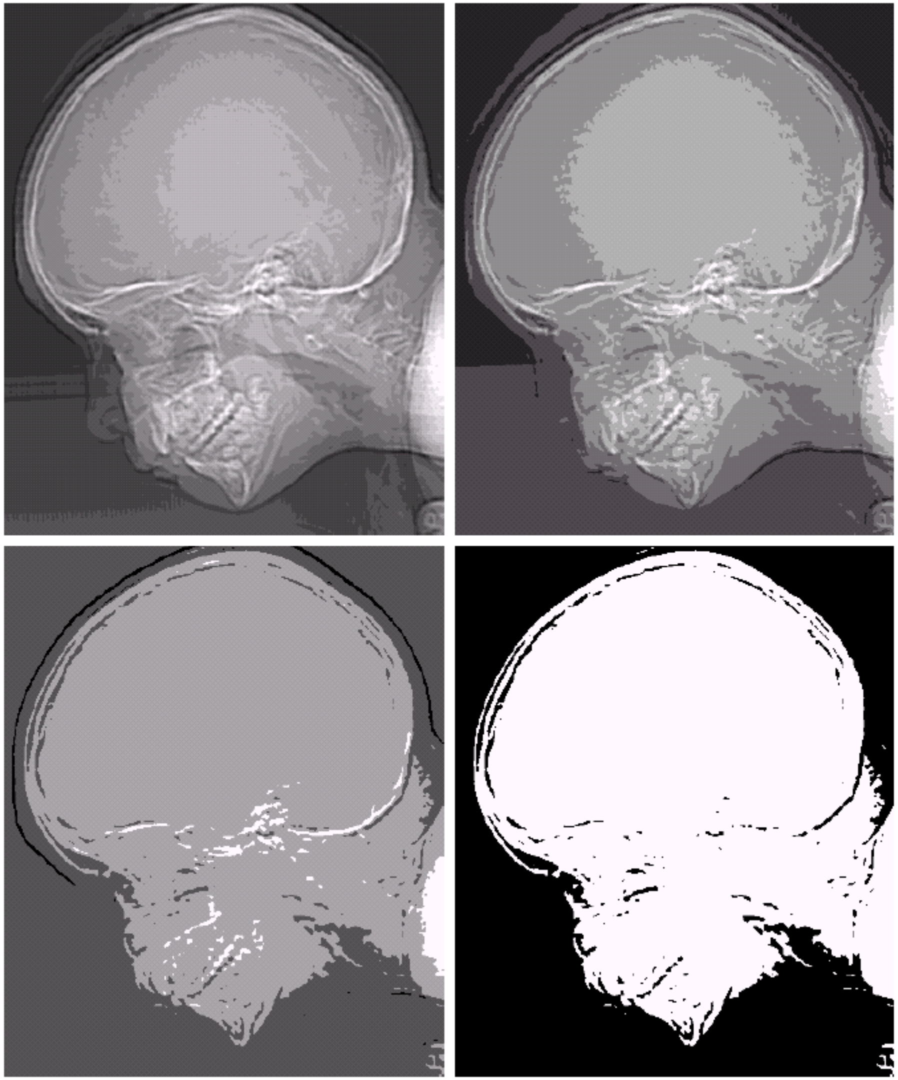


Intensity (grey level) resolution (强(灰)度分辨率)



a b
c d

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.



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 levels

4 levels

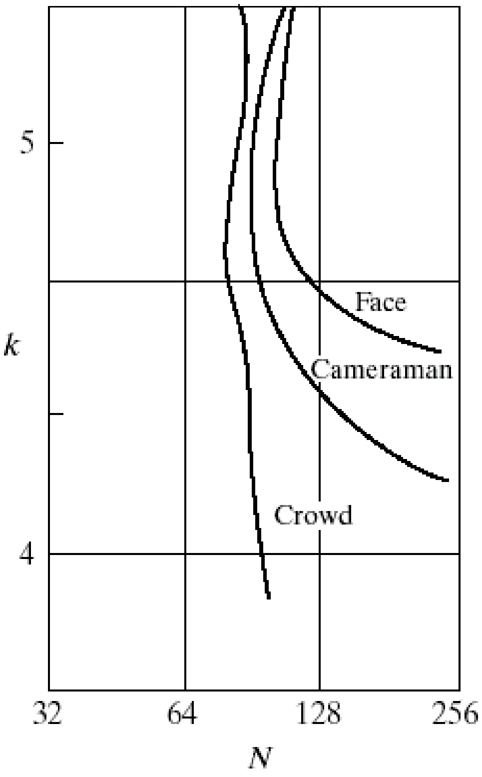
256 levels

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.



- **Spatial and Gray-Level Resolution**

Facts about Gray-level Resolution

- ✓ Gray-level resolution of a common genetic monitor usually is 256

Different Types of Quantization

- ➢ Continuous image to digital (or discrete) image, so called digitalization;
- ➢ High dynamic range (HDR) image to low dynamic range (LDR) image;

PS处理的图像

我们的方法一

高动态图像在不同曝光度下的显示效果对比

低曝光

中曝光

中高曝光

高曝光

经过使用某种图像处理技术，获得下述图像

用途: 改善数码相机在这种环境下的输出图像的质量? 现代数码相机已经具备在高动态范围(**High Dynamic Range**)成像的能力, 但在转化为24位数码图像需要量化时, 通常会出现类似前面的情况.

存在的问题:

2.5 Basic spatial relationships between pixels

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。

*	•	*
•	(x, y)	•
*	•	*

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：

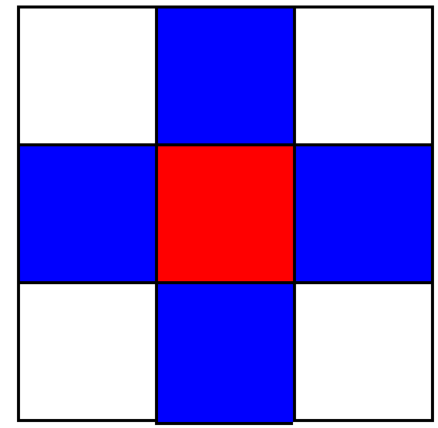
(x-1, y-1) , (x+1, y-1), (x-1, y+1), (x+1, y+1)

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

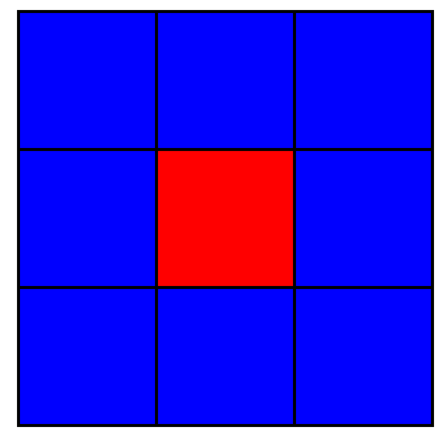
•	*	•
*	(x, y)	*
•	*	•

- Neighbors（相邻点）：主要有两种定义：4-neighbors 和8-neighbors

4-neighbors 定义较严：一个点（紅色）周边有四个相邻点（蓝色）

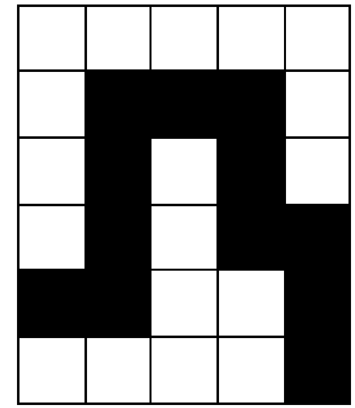


8-neighbors 定义较松：一个点（紅色）周边有八个相邻点（蓝色）



4-neighbors 一定是8-neighbors，反之不一定。

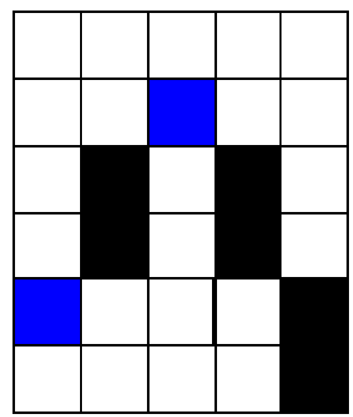
- Path（路径）：也有4-path 和 8-path

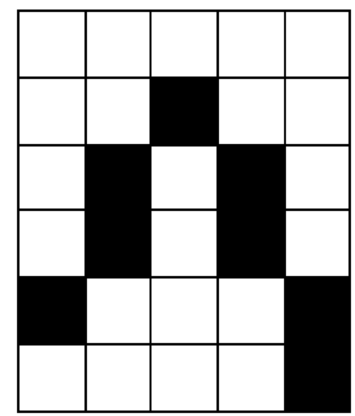


4-path 定义较严：一个路径上所有点互为 4-neighbors

8-path 定义较松：一个路径上所有点互为 8-neighbors

在4-path 定义下，蓝色点不再属于4-neighbor，路径因此断裂。





原本在8-path 定义下为连续的路径，在4-path 定义下可能会断裂。譬如上右图在8-path 定义下为连续的路径，在4-path 定义下变成左上图

- 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中的任一个子集。

一般我们考虑三种邻接性：

（a）4邻接：如果点q在$N_4(P)$中，数值在V中，则q和p是4邻接的；

（b）8邻接：如果点q在$N_8(P)$中，数值在V中，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。

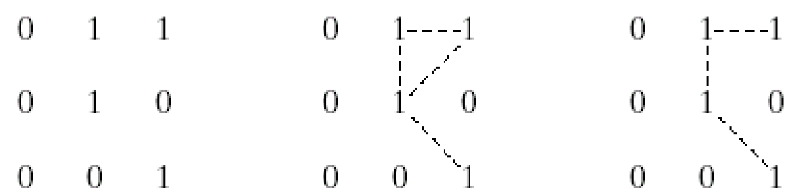
（i）q在中$N_4(P)$

（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。

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S¹中的某些像素和S²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）

- Distance measure

For pixels p,q, and z with coordinates (x,y), (s,t) and (u,v) respectively, D is distance function or metric if

(a) $D(p,q) \geqq 0$ 『$D(p,q)=0$ if and only if $p=q$』

(b) $D(p,q)=D(q,p)$

(c) $D(p,z) \leqq D(p,q)+D(q,z)$

- Euclidean distance

$$D_e(p,q) = \sqrt{(x-s)^2 + (y-t)^2}$$

- The D4 distance (also called the 'city block distance')

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

		2		
	2	1	2	
2	1	0	1	2
	2	1	2	
		2		

- The D8 distance (also called the 'chessboard distance')

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

2	2	2	2	2
2	1	1	1	2
2	1	0	1	2
2	1	1	1	2
2	2	2	2	2

2.6 An introduction to the Mathematical Tools Used in DIP

2.6.1. Array versus Matrix Operation

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Array product

$$\mathbf{AB} = \begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \end{bmatrix}$$

Matrix product

$$\mathbf{AB} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{bmatrix}$$

2.6.2. Linear versus Nonlinear Operation

H—— an operator, both *f* and *g* are images. A Linear operator has to satisfy the following rules (*a* and *b* are scalar number).

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

Examples: sum operator, Σ; max operation……

2.6.3. Arithmetic Operation

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

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

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

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

- Averaging K images to remove zero-mean noise

$f(x,y)$——原始图像

$\eta(x,y)$——随机噪音，在各个坐标点（x,y）上的噪音不相关，且均值为零

$g(x,y)=f(x,y)+\eta(x,y)$——带噪音的图像

对 K 幅带噪音的图像取平均

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

- Averaging K images to remove zero-mean noise

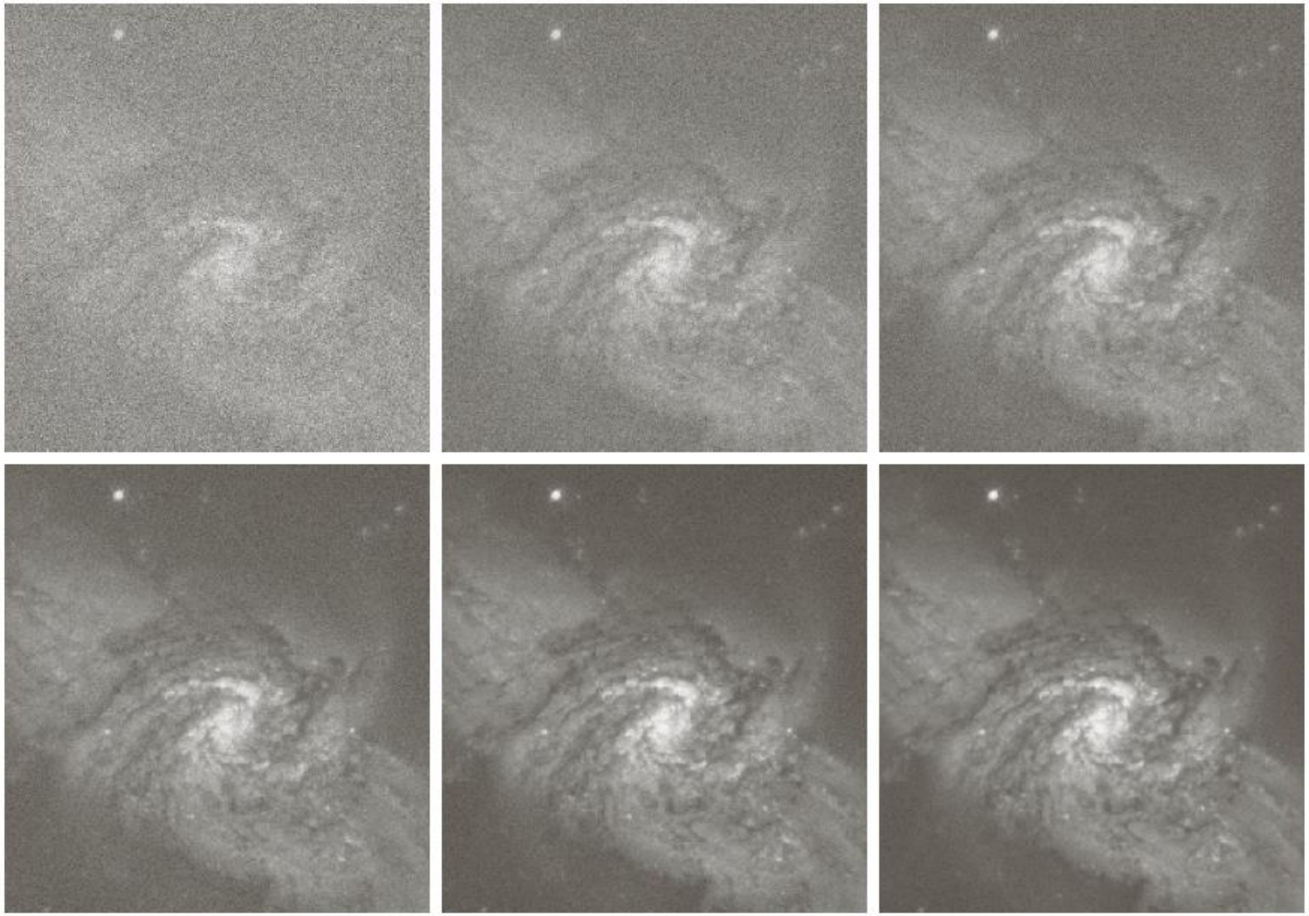
则有

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

$$\sigma^2_{\overline{g}(x,y)} = \frac{1}{K}\sigma^2_{\eta(x,y)}$$

其中，$E\{\overline{g}(x,y)\}$是 $\overline{g}$ 的期望值，$\sigma^2_{\overline{g}(x,y)}$ 和 $\sigma^2_{\eta(x,y)}$ 分别是 $\overline{g}$ 和 η 的方差。在平均图像中，任何一点的标准差为：

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



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.)

✓ 希腊夜景

- ✓ 希腊夜景

关于位图

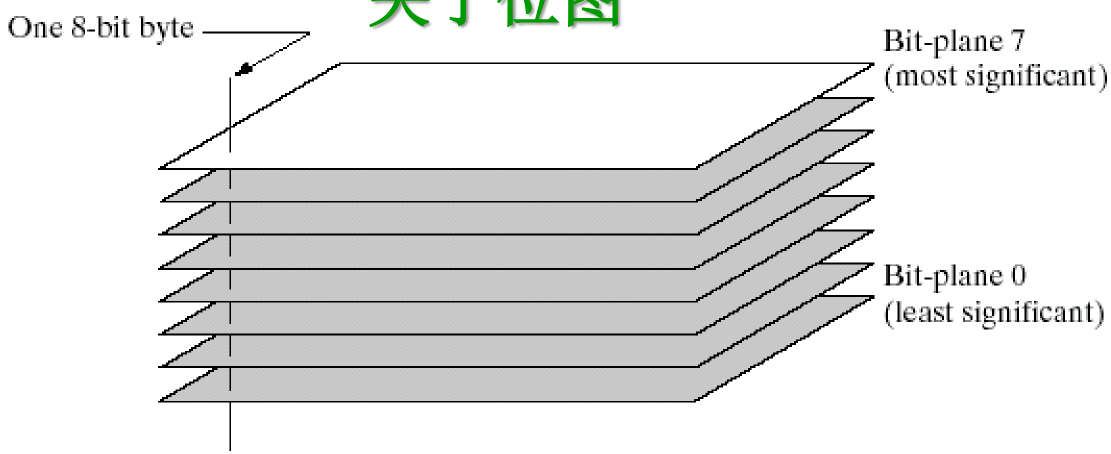


FIGURE 3.12 Bit-plane representation of an 8-bit image.

$$\begin{bmatrix} 255 & 130 \\ 80 & 24 \end{bmatrix} \xrightarrow{\text{Bit-plane}}$$

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

255	130	80	24
1	1	0	0
1	0	1	0
1	0	0	0
1	0	1	1
1	0	0	1
1	0	0	0
1	1	0	0
1	0	0	0

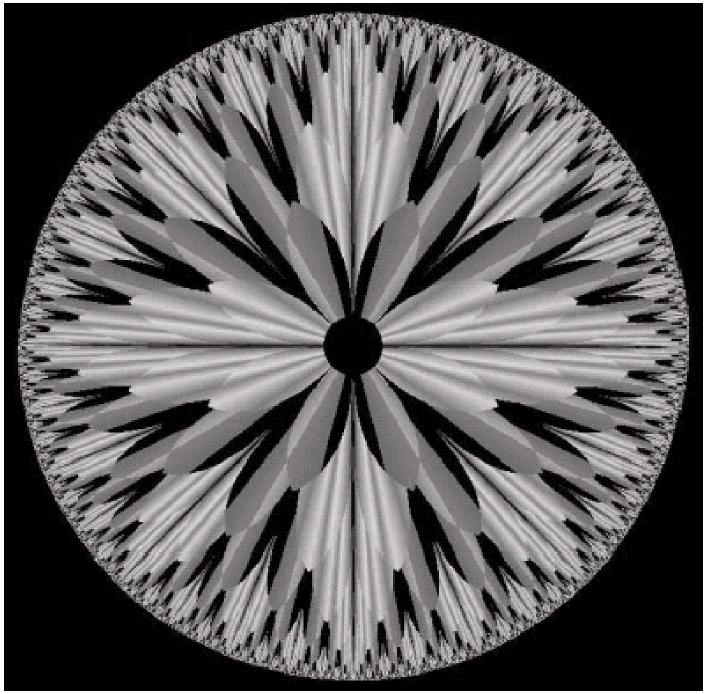


FIGURE 3.13 An 8-bit fractal image. (A fractal is an image generated from mathematical expressions). (Courtesy of Ms. Melissa D. Binde, Swarthmore College, Swarthmore, PA.)

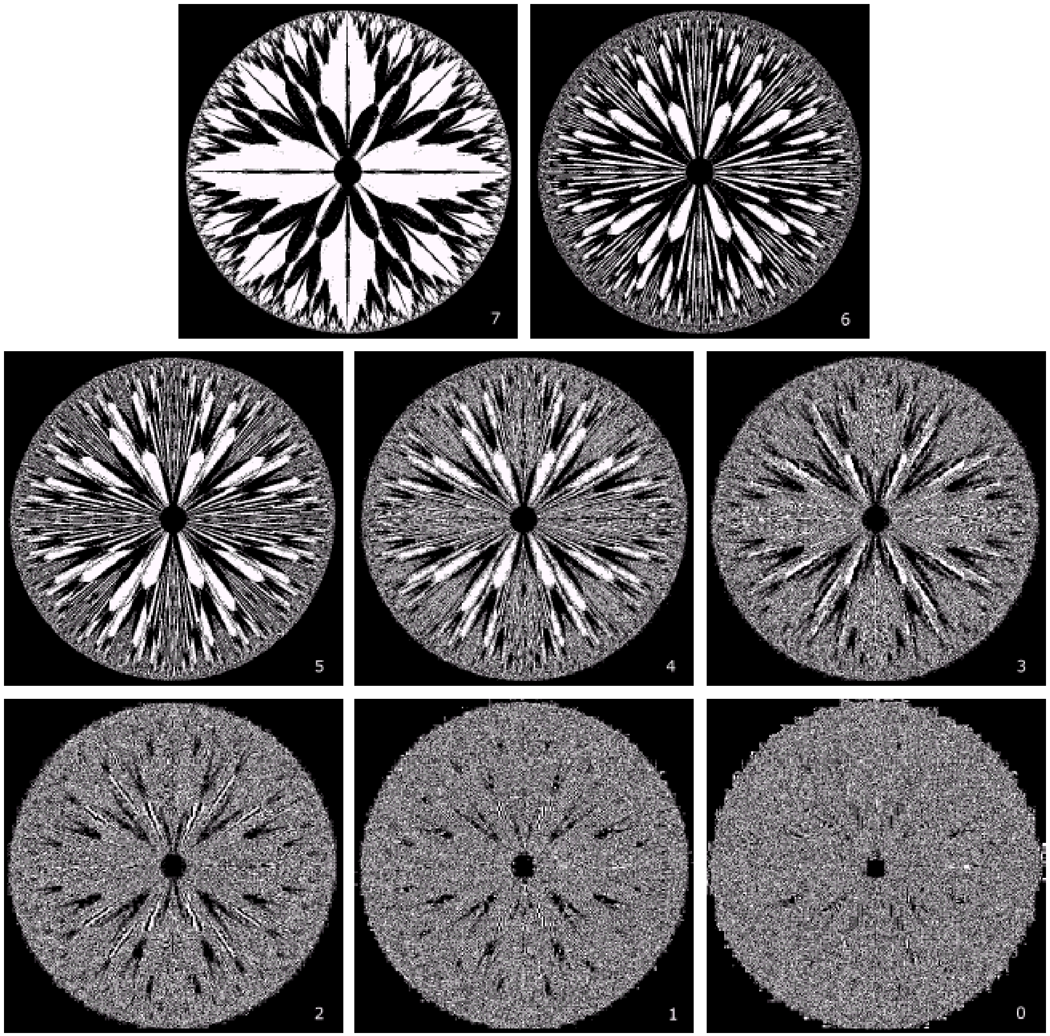
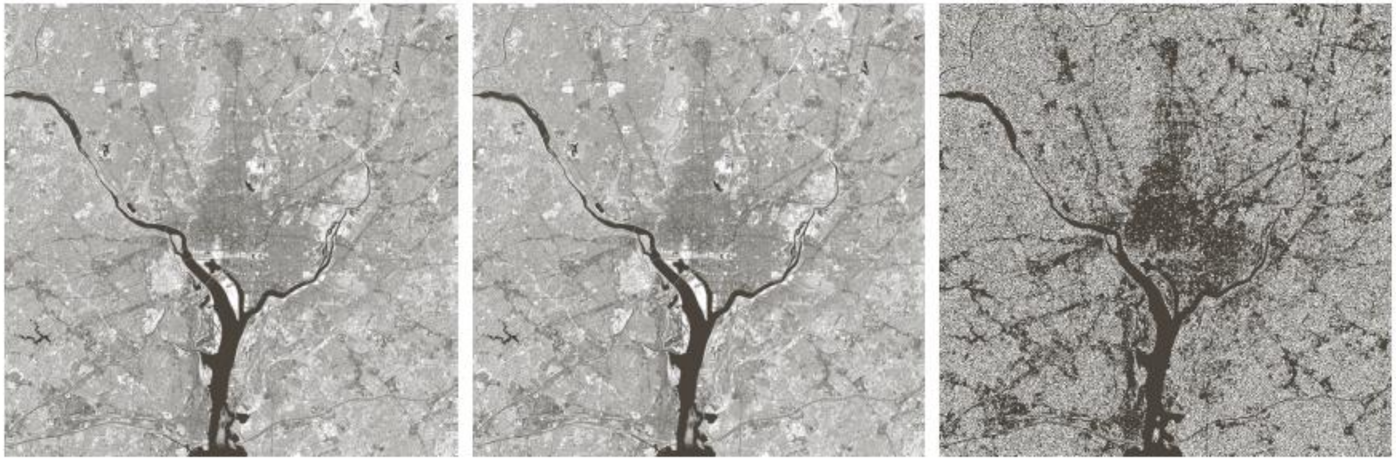


FIGURE 3.14 The eight bit planes of the image in Fig. 3.13. The number at the bottom, right of each image identifies the bit plane.

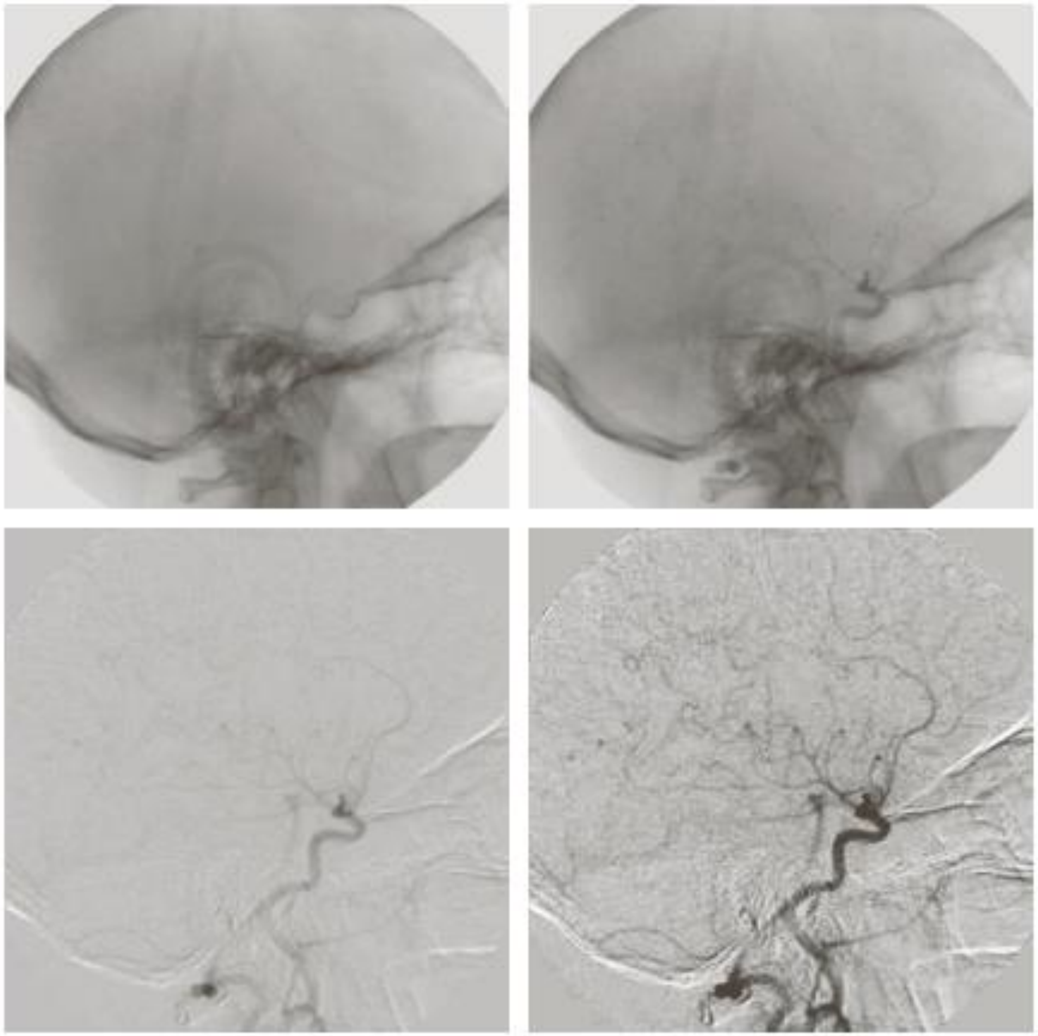
Image subtraction



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.

Image subtraction



a b
c d

FIGURE 2.28 Digital subtraction angiography. (a) Mask image. (b) A live image. (c) Difference between (a) and (b). (d) Enhanced difference image. (Figures (a) and (b) courtesy of The Image Sciences Institute, University Medical Center, Utrecht, The Netherlands.)

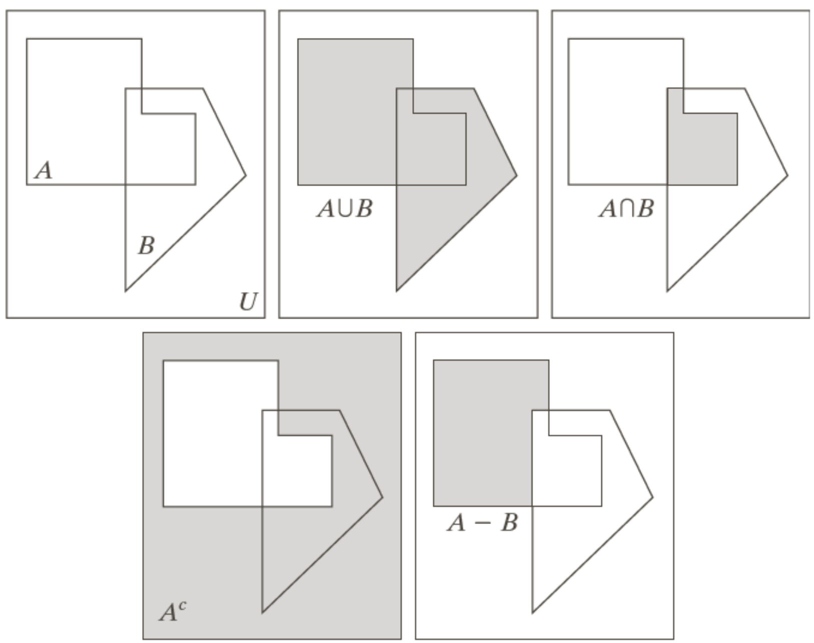
Image multiplication (and division)

a b c

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.)

2.6.4. Set and Logical Operations

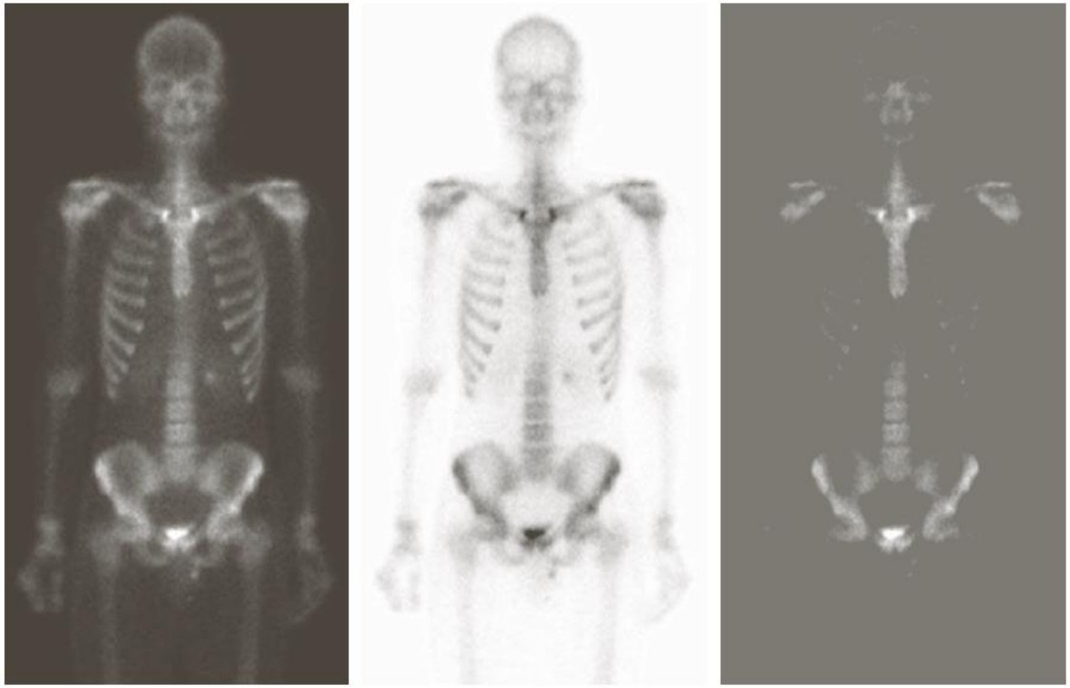
- Basic set operations
  1. Union
  2. Intersection
  3. Complement
  4. Difference



a b c
d e

FIGURE 2.31 (a) Two sets of coordinates, A and B, in 2-D space. (b) The union of A and B. (c) The intersection of A and B. (d) The complement of A. (e) The difference between A and B. In (b)–(e) the shaded areas represent the member of the set operation indicated.

2.6.4. Set and Logical Operations



a b c

FIGURE 2.32 Set operations involving gray-scale images. (a) Original image. (b) Image negative obtained using set complementation. (c) The union of (a) and a constant image. (Original image courtesy of G.E. Medical Systems.)

Example 2.8: A gray image **A** can be represented by **(x, y, z)**, where ***z*** denotes intensity. $\mathbf{A}^c=\{(x, y, K-z)\}$, K is a constant (such as $K=2^8-1$).

- logical operations

Foreground=1, background=0

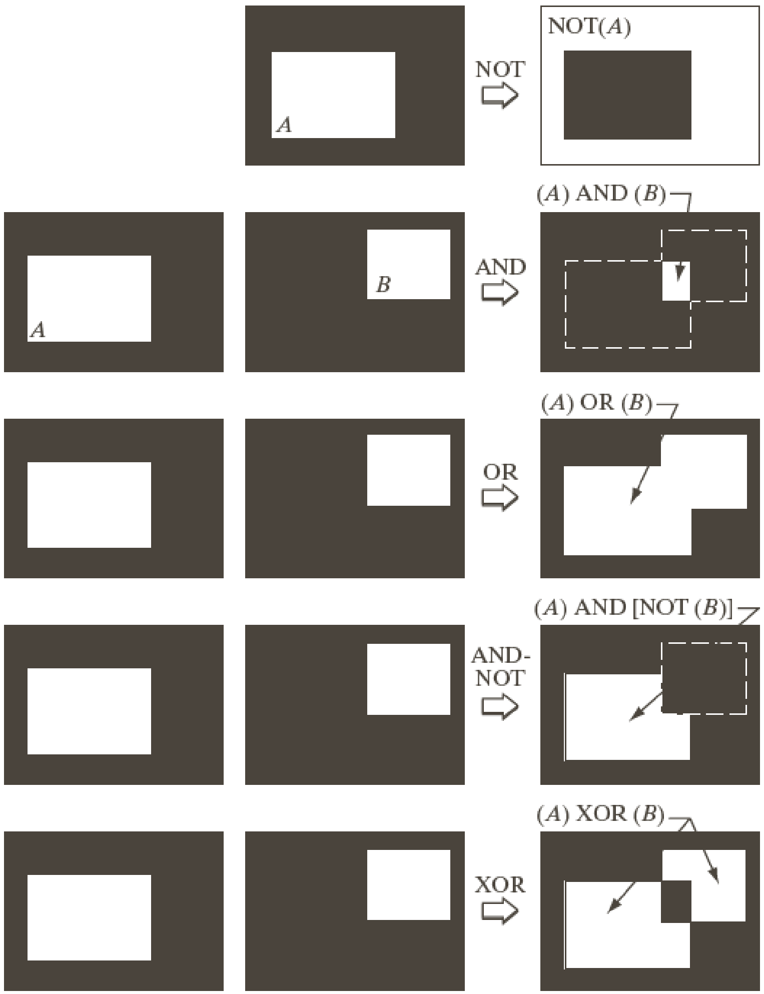
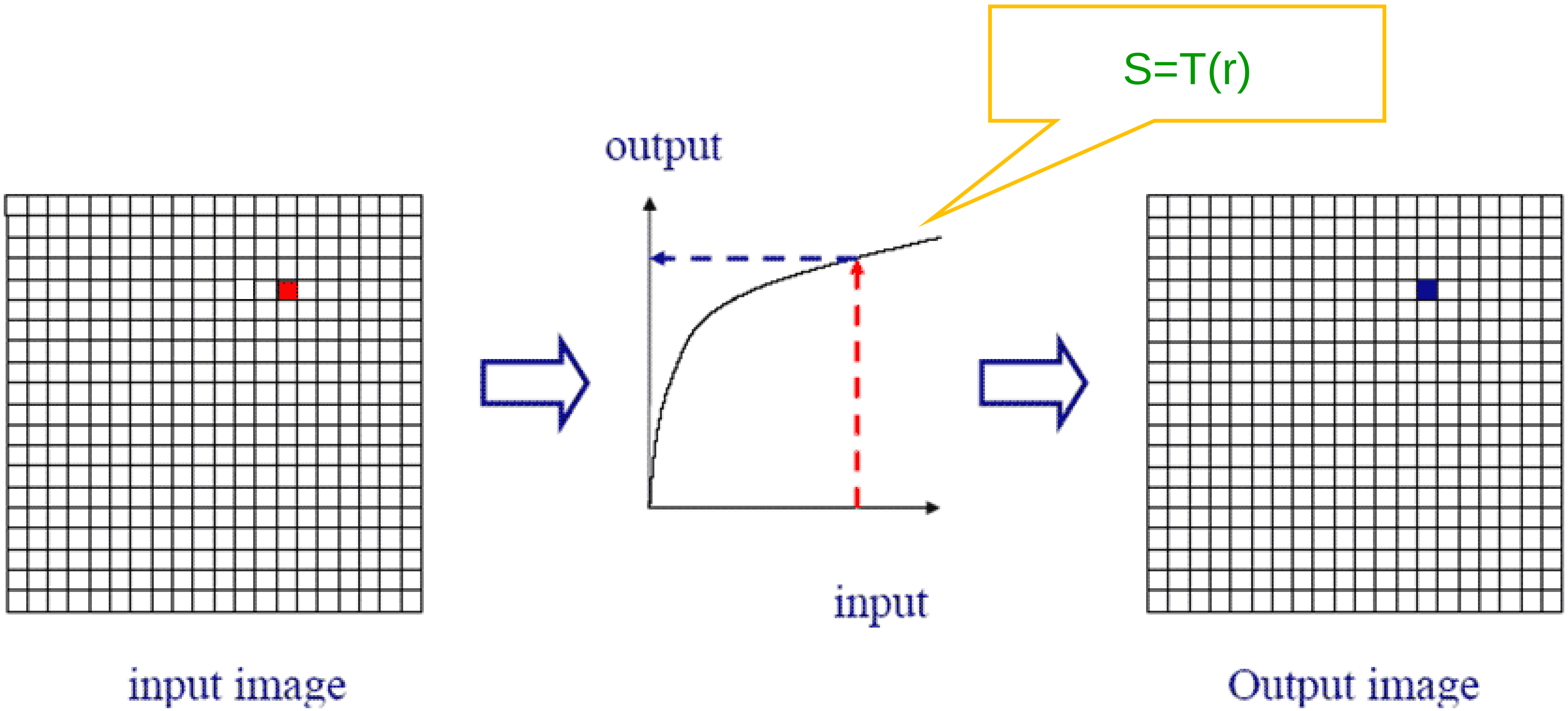


FIGURE 2.33 Illustration of logical operations involving foreground (white) pixels. Black represents binary 0s and white binary 1s. The dashed lines are shown for reference only. They are not part of the result.

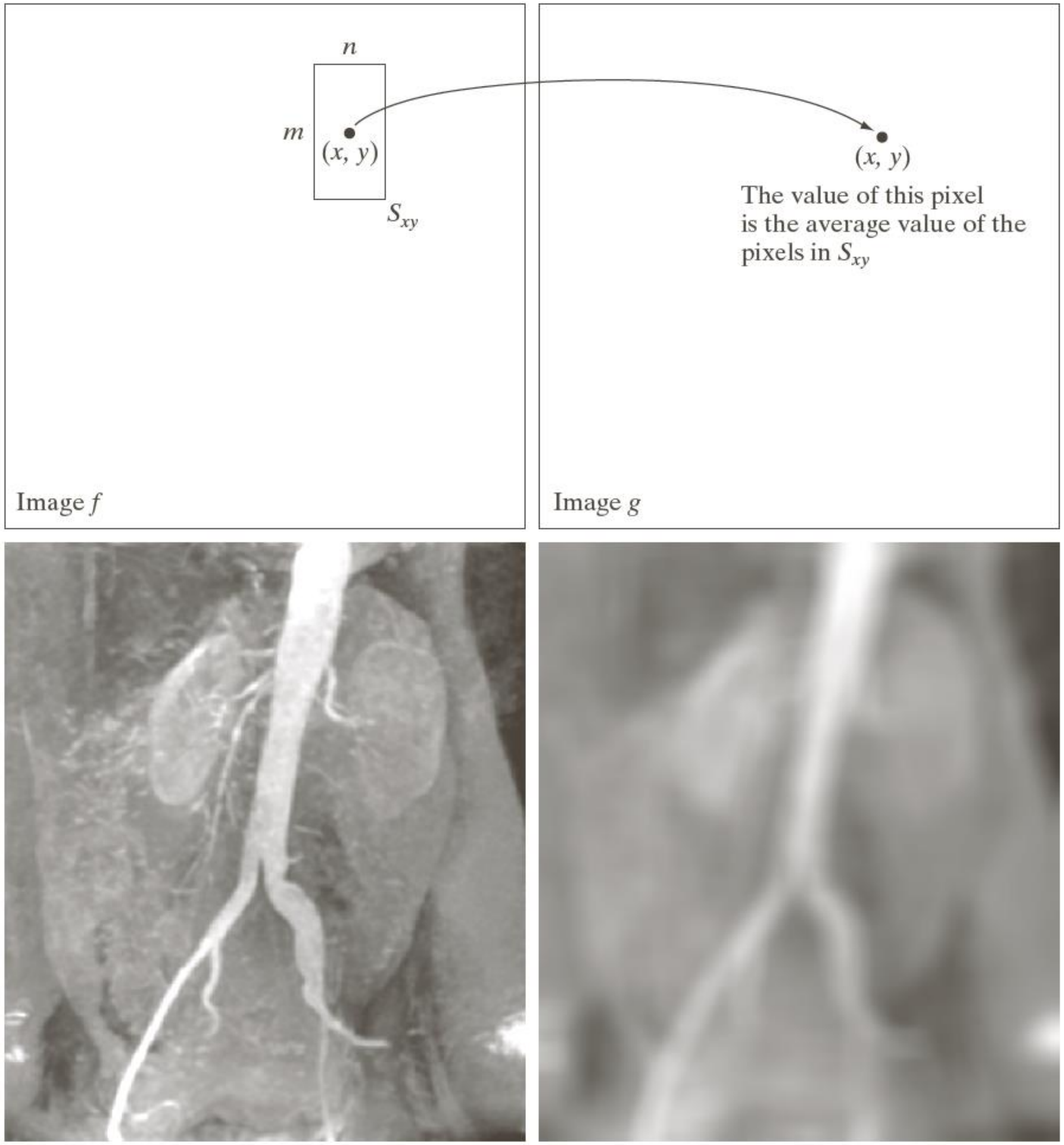
- Fuzzy set (模糊集)： examples

- **Spatial Operation**

  - ✓ **Single-pixel operations:** one pixel in the original image to produce one pixel in the output image **s=T(z)**



✓ Neighborhood operations:



a b
c d

FIGURE 2.35 Local averaging using neighborhood processing. The procedure is illustrated in (a) and (b) for a rectangular neighborhood. (c) The aortic angiogram discussed in Section 1.3.2. (d) The result of using Eq. (2.6-21) with $m = n = 41$. The images are of size 790×686 pixels.

✓ Geometric spatial transformations and image registration:

Geometric transformations modify the spatial relationship between pixels. They are often called rubber-sheet transformations:

- analogous to “printing” an image on a sheet of rubber
- and then stretching the sheet according to a predefined set of rules

Basic formula

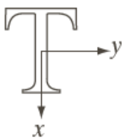
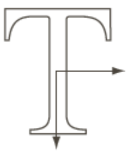
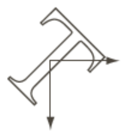
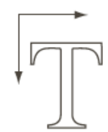
$$(x, y) = \boldsymbol{T}\{(v, w)\} \tag{2.6-22}$$

By using 3D homogeneous coordinates, we have

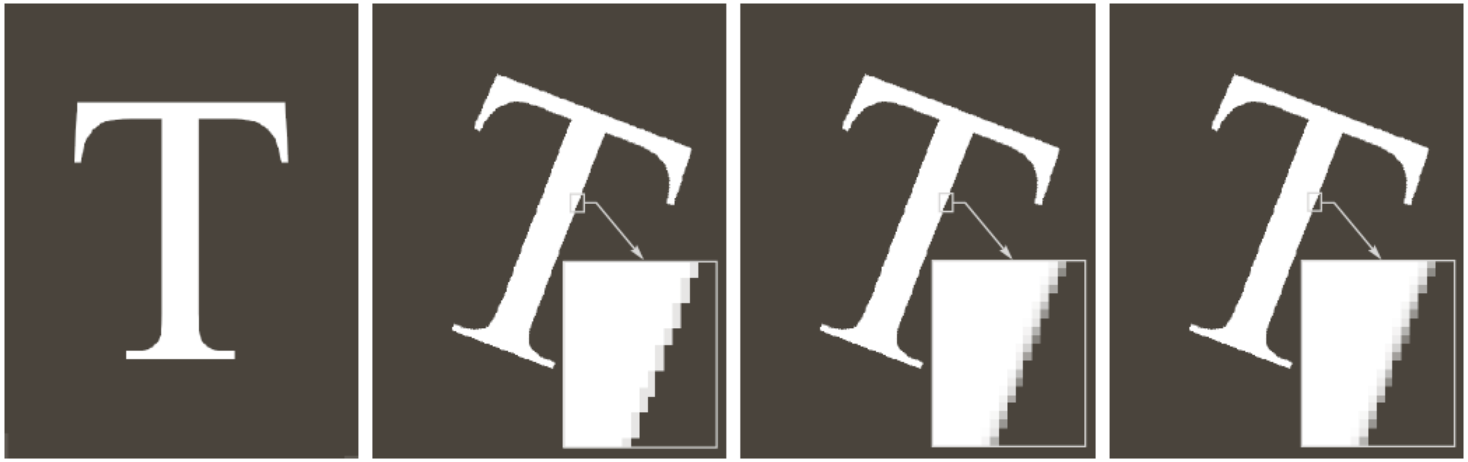
$$\begin{bmatrix} x & y & 1 \end{bmatrix} = \begin{bmatrix} v & w & 1 \end{bmatrix} \mathrm{T} = \begin{bmatrix} v & w & 1 \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{bmatrix} \tag{2.6-23}$$

✓ Geometric spatial transformations and image registration:

TABLE 2.2
Affine transformations based on Eq. (2.6.–23).

Transformation Name	Affine Matrix, T	Coordinate Equations	Example
Identity	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v$ $y = w$	
Scaling	$\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = c_x v$ $y = c_y w$	
Rotation	$\begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v\cos\theta - w\sin\theta$ $y = v\cos\theta + w\sin\theta$	
Translation	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix}$	$x = v + t_x$ $y = w + t_y$	
Shear (vertical)	$\begin{bmatrix} 1 & 0 & 0 \\ s_v & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v + s_v w$ $y = w$	
Shear (horizontal)	$\begin{bmatrix} 1 & s_h & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$x = v$ $y = s_h v + w$	

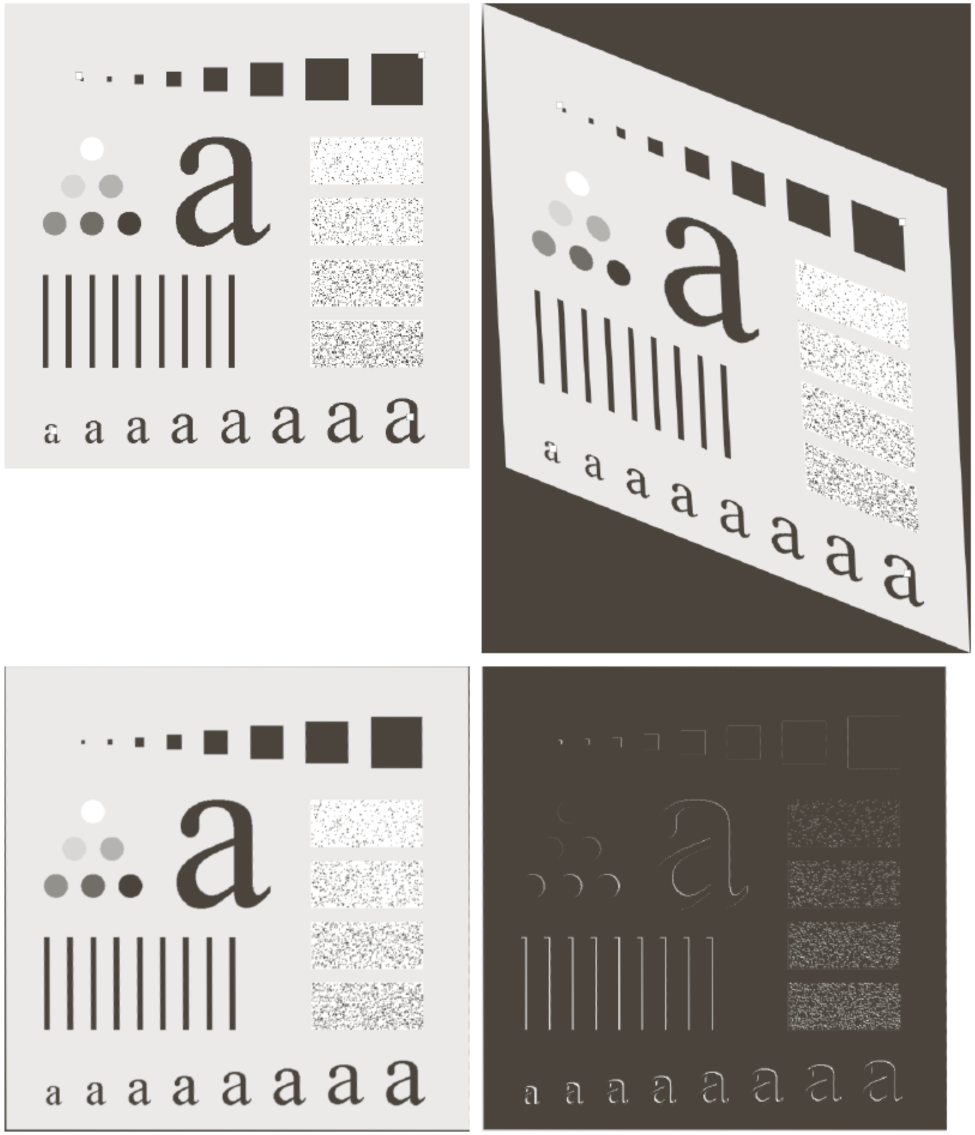
✓ Geometric spatial transformations and image registration:



a b c d

FIGURE 2.36 (a) A 300 dpi image of the letter T. (b) Image rotated 21° clockwise using nearest neighbor interpolation to assign intensity values to the spatially transformed pixels. (c) Image rotated 21° using bilinear interpolation. (d) Image rotated 21° using bicubic interpolation. The enlarged sections show edge detail for the three interpolation approaches.

✓ **Geometric spatial transformations and image registration:**



a b
c d

FIGURE 2.37 Image registration. (a) Reference image. (b) Input (geometrically distorted image). Corresponding tie points are shown as small white squares near the corners. (c) Registered image (note the errors in the borders). (d) Difference between (a) and (c), showing more registration errors.