

《数字图像处理》

第2讲： 图像处理基础

冯建江

清华大学，自动化系

2017-09-28

内容

1. 人眼
2. 采样与量化
3. 像素之间的关系
4. 图像处理中的数学工具

1. 人眼

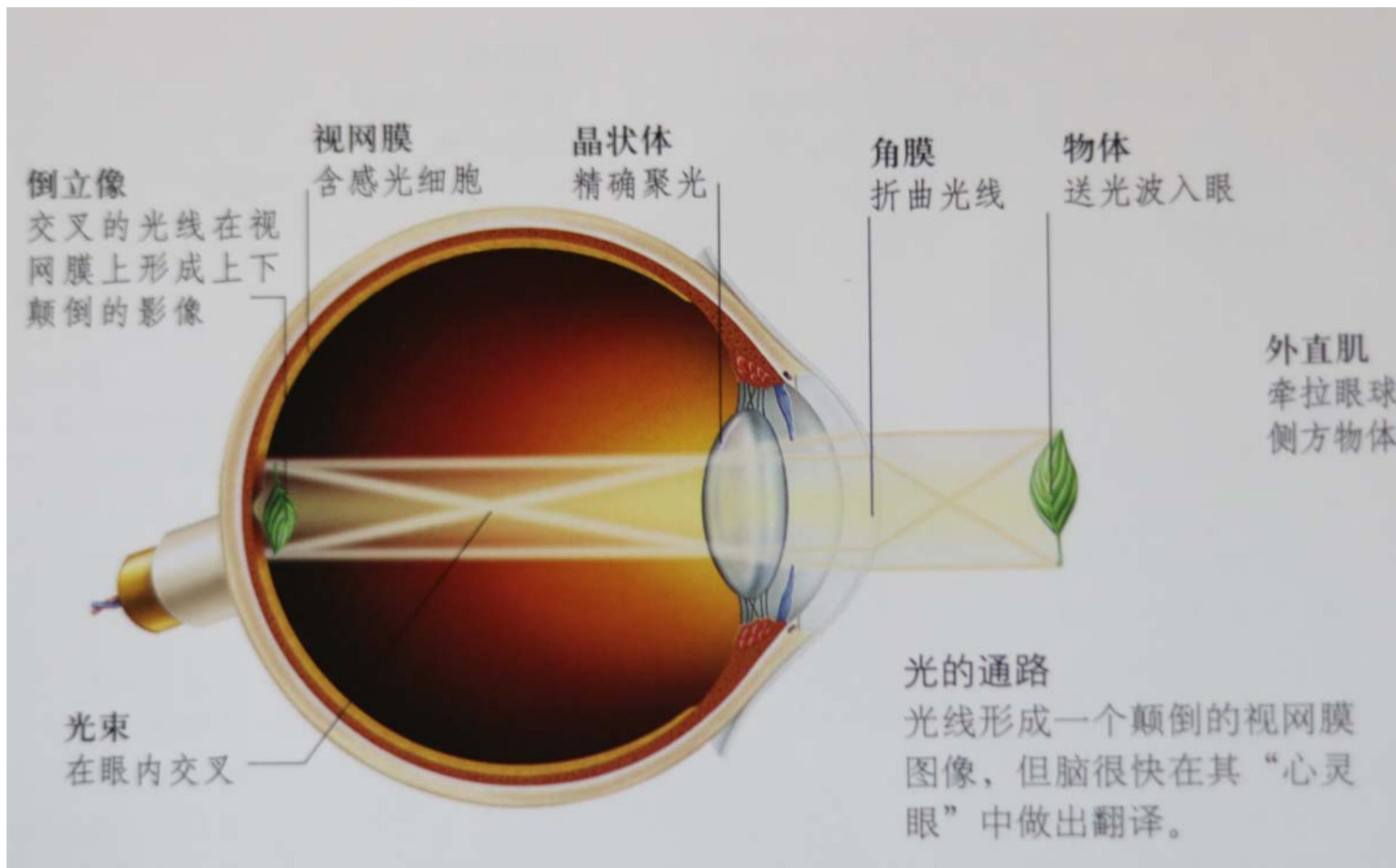
数字图像处理课为什么要讲人眼？

- 人眼有值得借鉴的地方
- 图像处理的一大用途是给人看， 需要考虑人眼的特点（例如电影的帧率24fps考虑了人眼的视觉暂留）

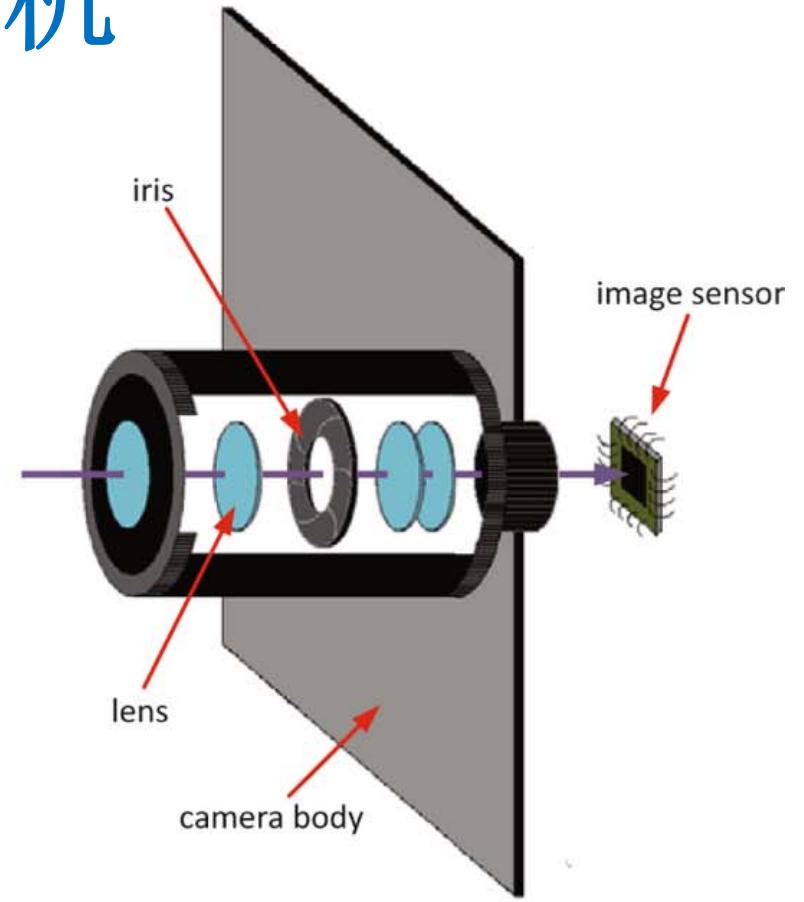
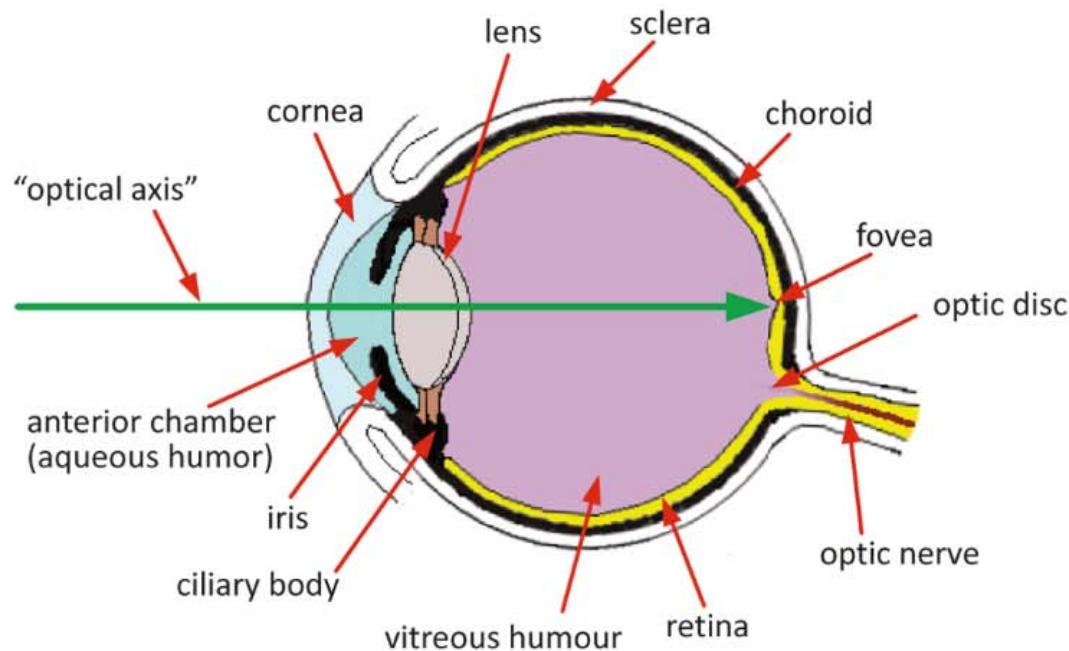
优酷



视觉是如何形成的



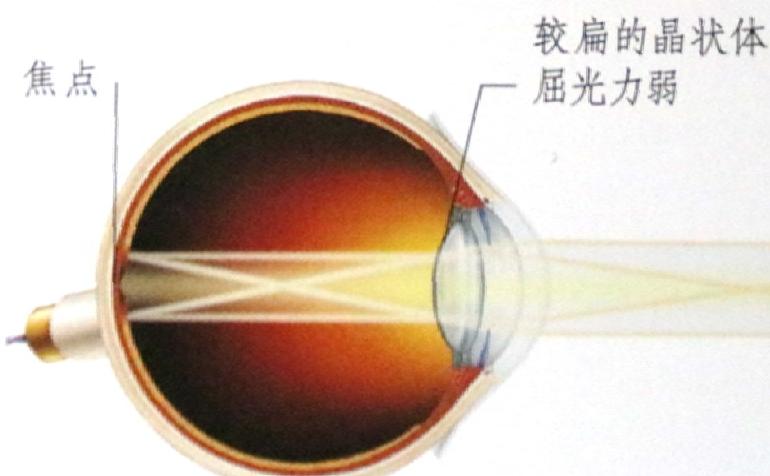
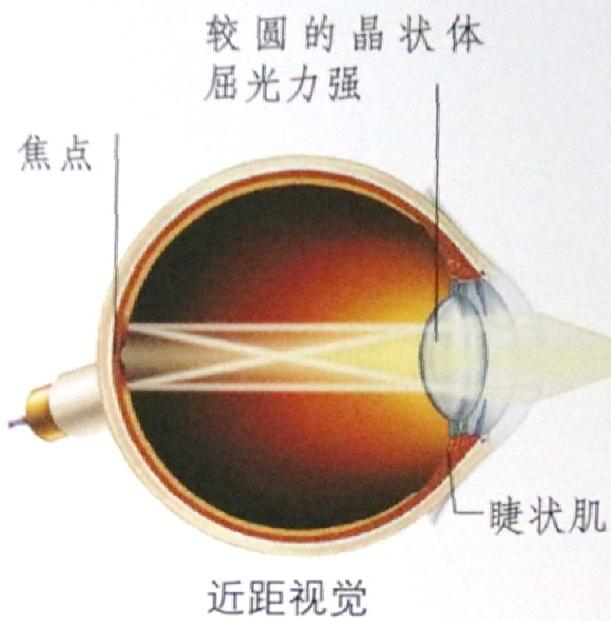
人眼与相机



- 晶状体—透镜：变焦
- 虹膜—光圈：控制进光量
- 视网膜—图像传感器：感光

人眼变焦

眼和视觉 — 神经系统



相机变焦

短焦

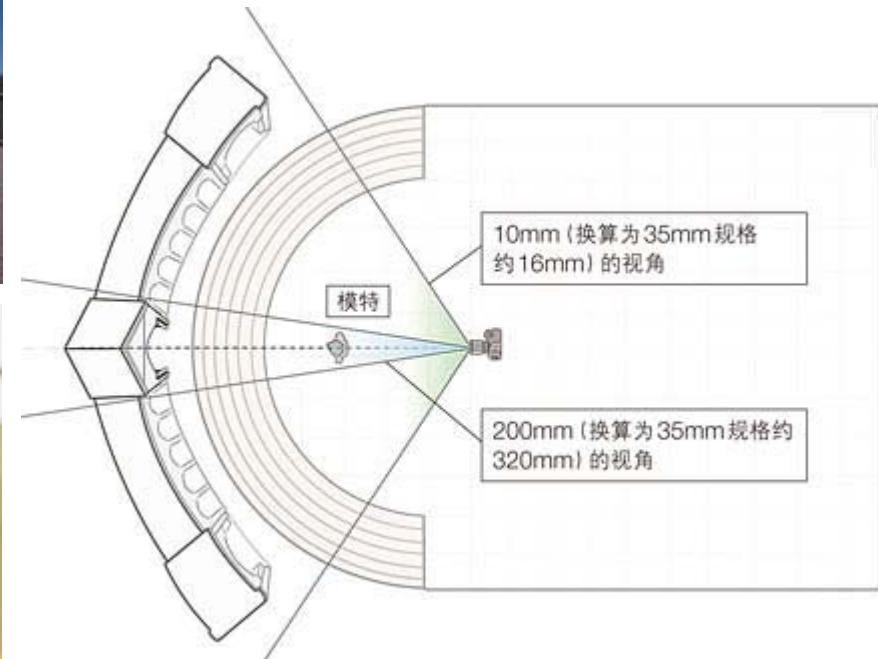


长焦

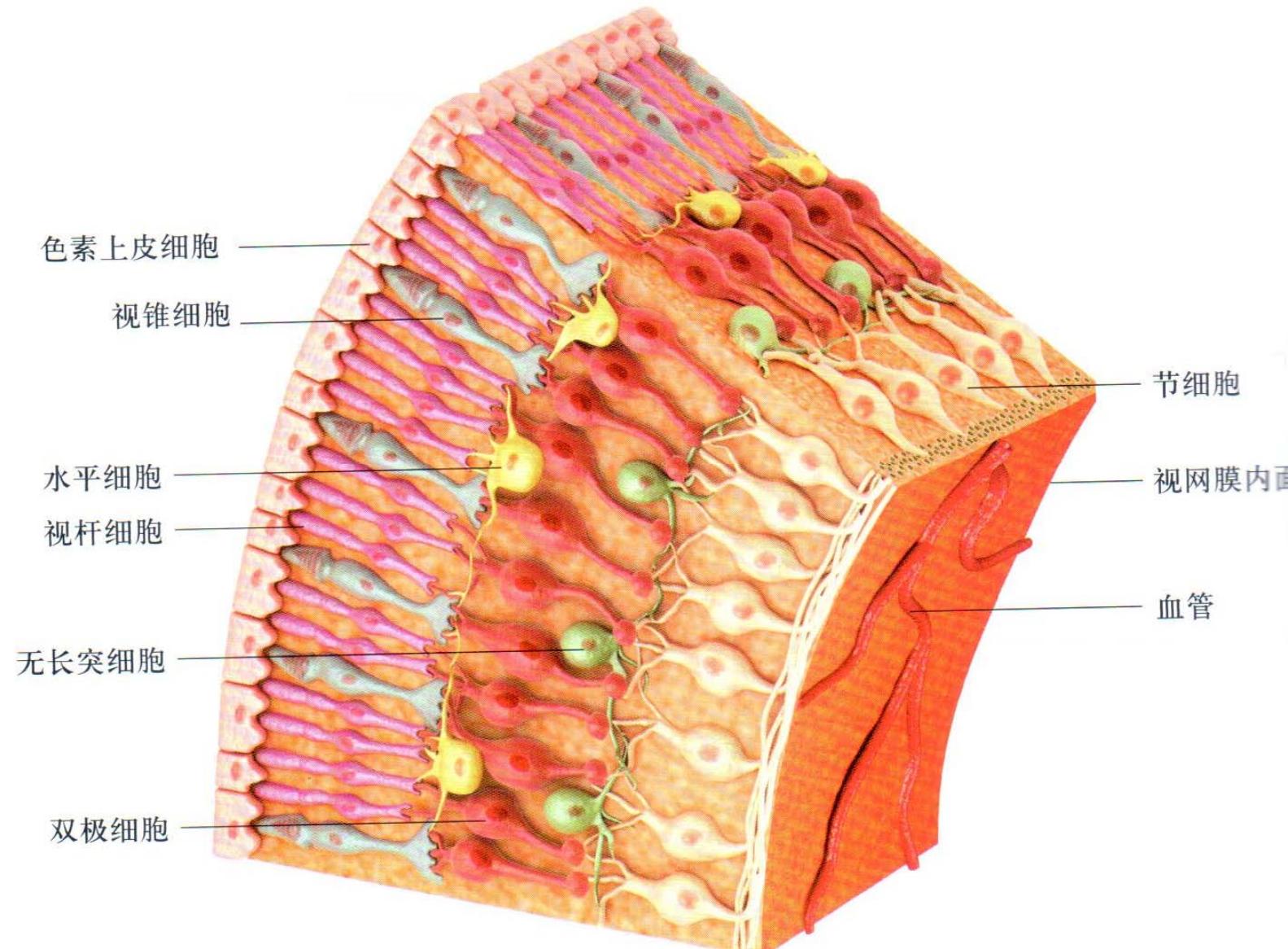


变焦镜头

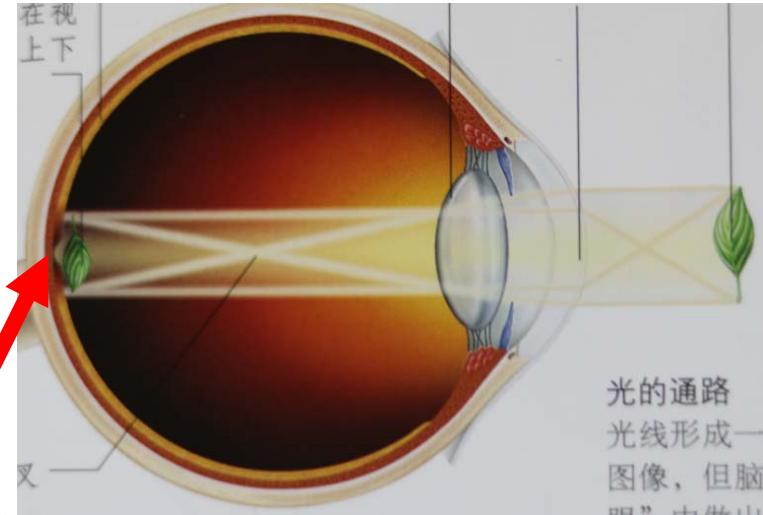
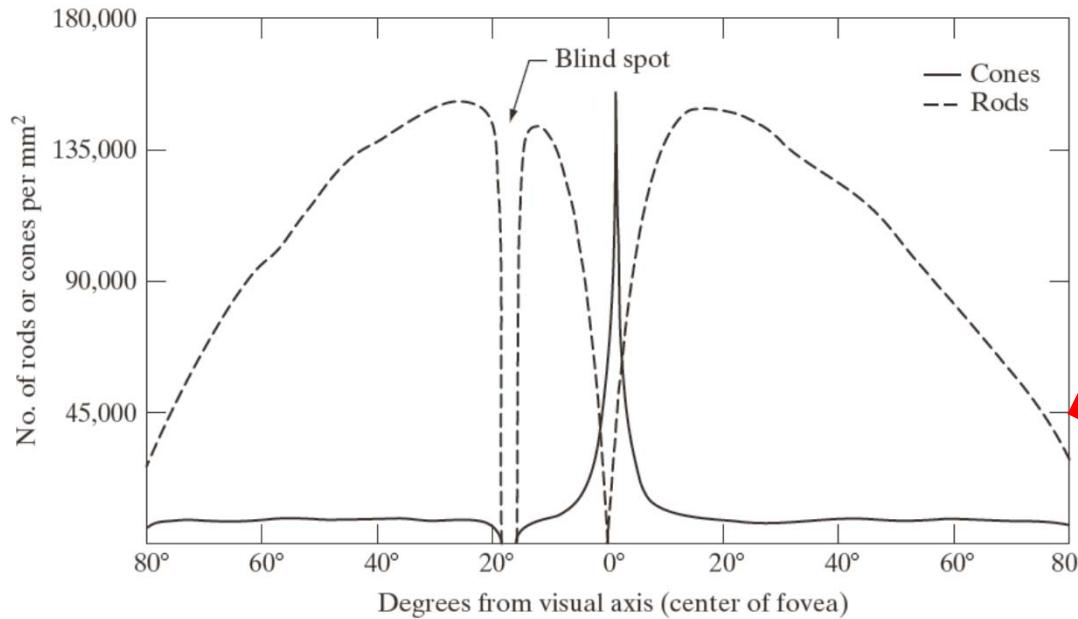
模特和摄影师都没有移动



视网膜

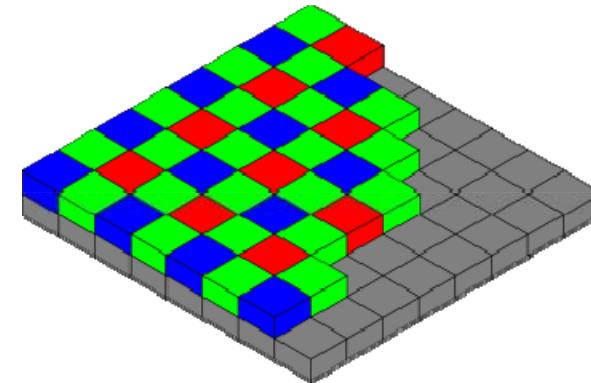
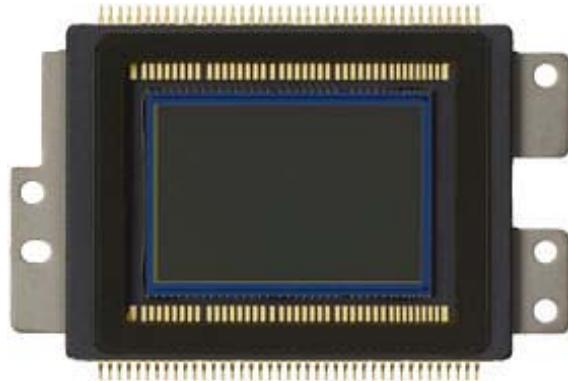


视锥细胞（Cones）和视杆细胞（Rods）



- 视锥细胞主要分布在中央凹（fovea），需要亮光，对颜色敏感，能分辨细节。
- 视杆细胞分布在整个视网膜，感受低亮度光，不能辨别颜色。

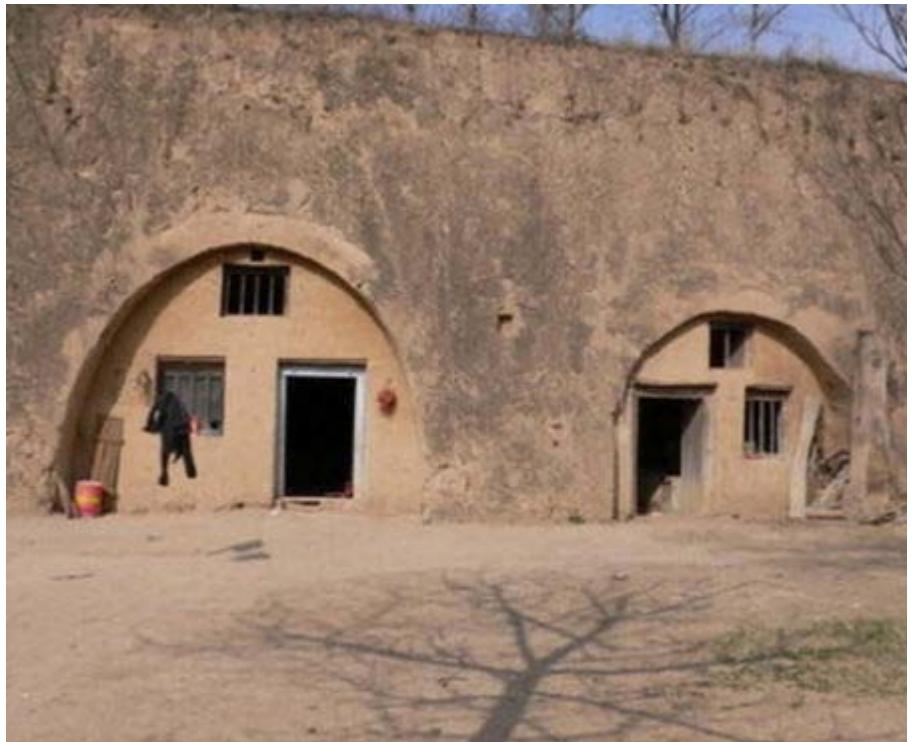
图像传感器 (Image sensor)



- 图像传感器类似于视网膜。
- 图像传感器中每个感光元件对应一个像素。
- 在接受光照后，感光元件产生对应的电流，电流大小与光强对应，电流是模拟的，经模数转换为数字。
- 由于感光元件只能感应光的强度，无法捕获色彩信息，因此必须在感光元件上方覆盖彩色滤光片。
- 常用做法是覆盖红绿蓝三色滤光片，以1:2:1的比例构成（因人眼对绿色敏感）。

https://en.wikipedia.org/wiki/Bayer_filter

http://www.canon.com.cn/specialsite/ds_abcbook/beginner01.html



白天，窑洞里看上去是黑的

白天窑洞里的光（来自太阳光）比夜间的光（来自日光灯）强？



瞳孔看上去是黑的？

透过瞳孔看到的是什么？

视网膜

视网膜是黑色的？



夜间闪光灯拍照有红眼

视网膜是橘红色

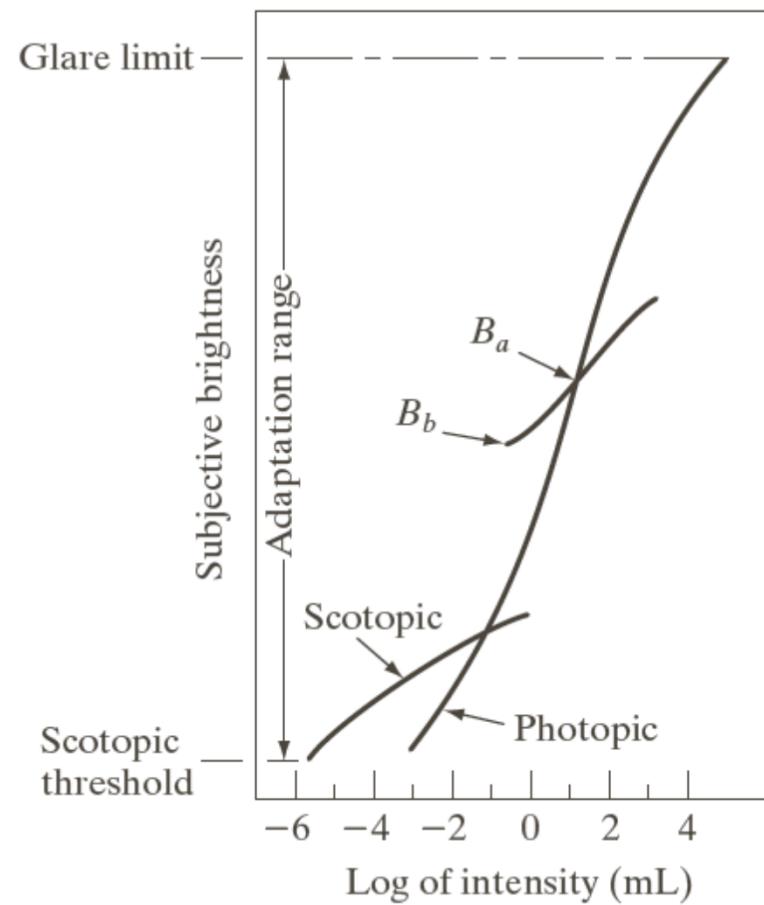
瞳孔黑和窑洞白天黑
是一个道理

红眼照片

<http://jandan.net/2014/07/21/causes-red-eyes.html>

人眼对光强的区分力

- 人眼能区分的光强范围非常广
($\frac{\text{最高值}}{\text{最低值}} \approx 10^{10}$)
- 主观亮度是光强的对数函数
- 亮视觉 (photopic vision) 和暗视觉 (scotopic vision)
- 人眼不能同时工作在整个区间
- 人眼的亮度适应能力 (brightness adaption)
- 亮度适应级 B_a , 低于 B_b 是黑
(白天窑洞是黑的)

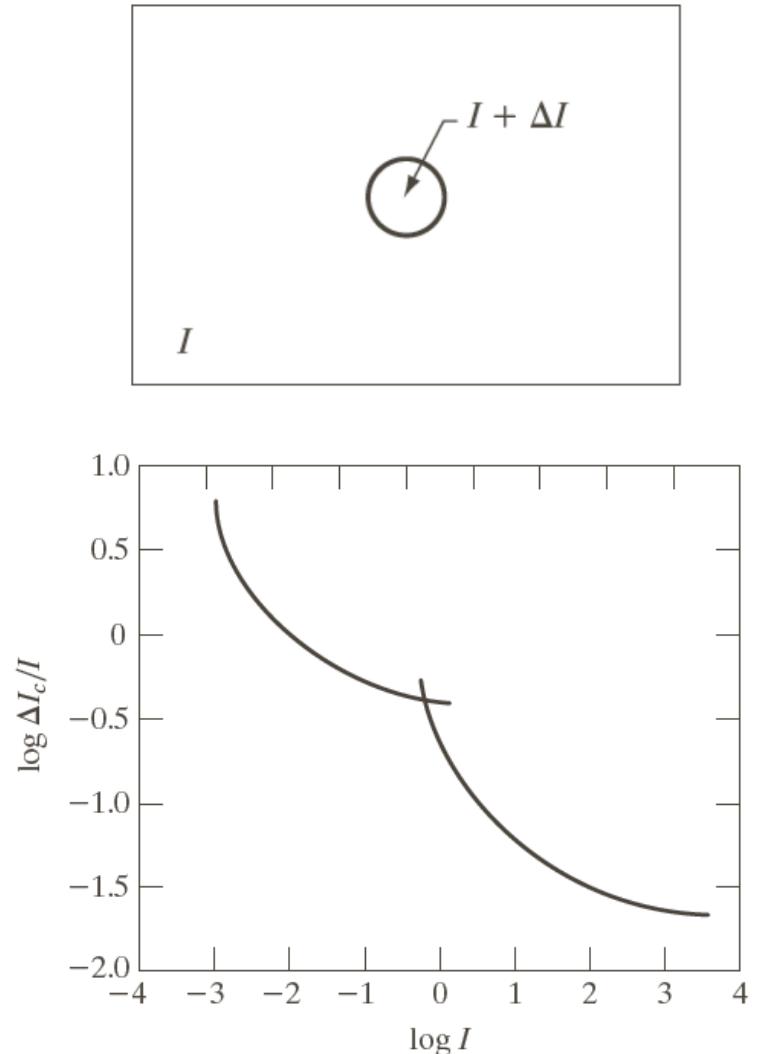


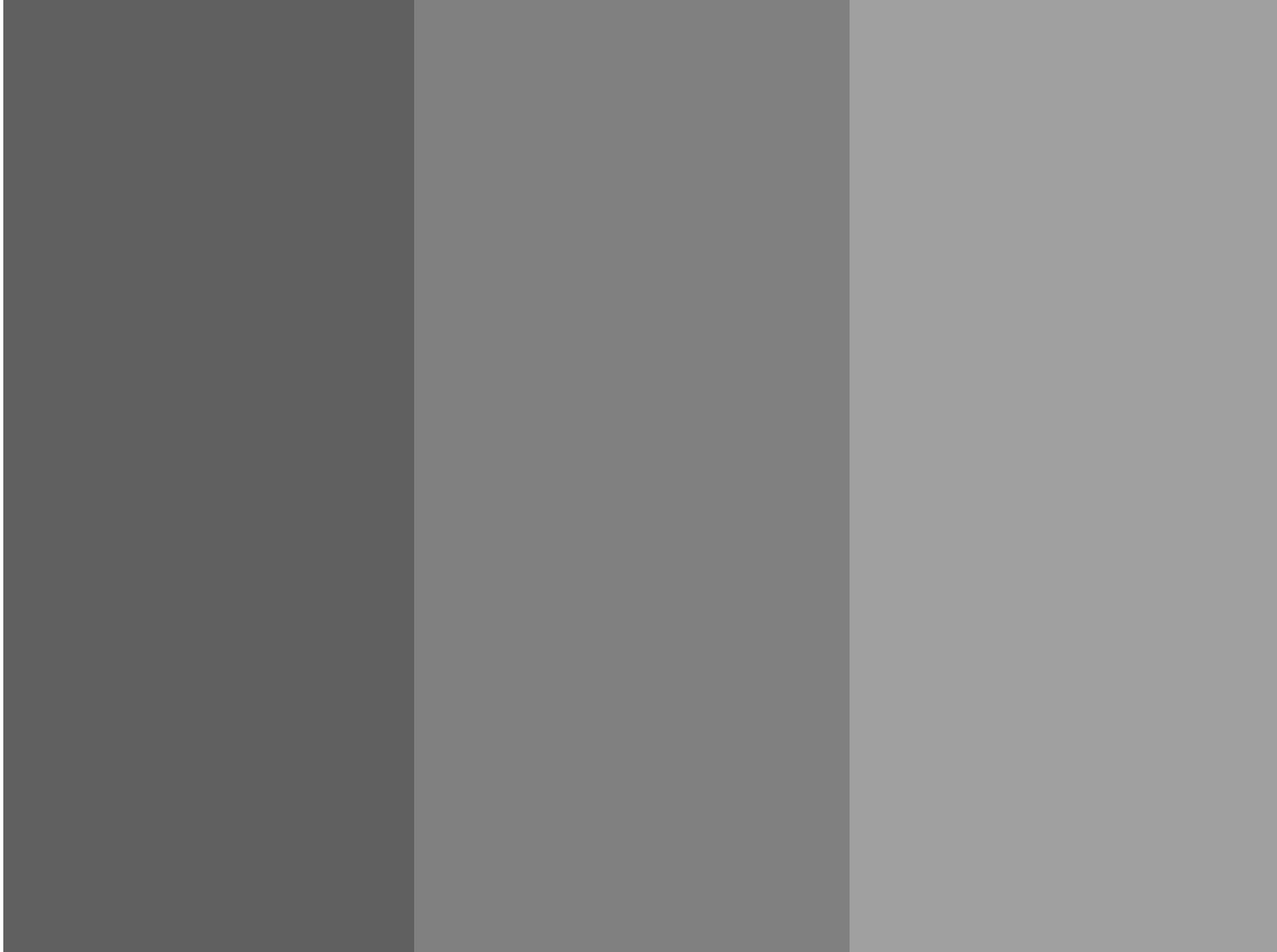
人眼对光强变化的区分力

实验方法：

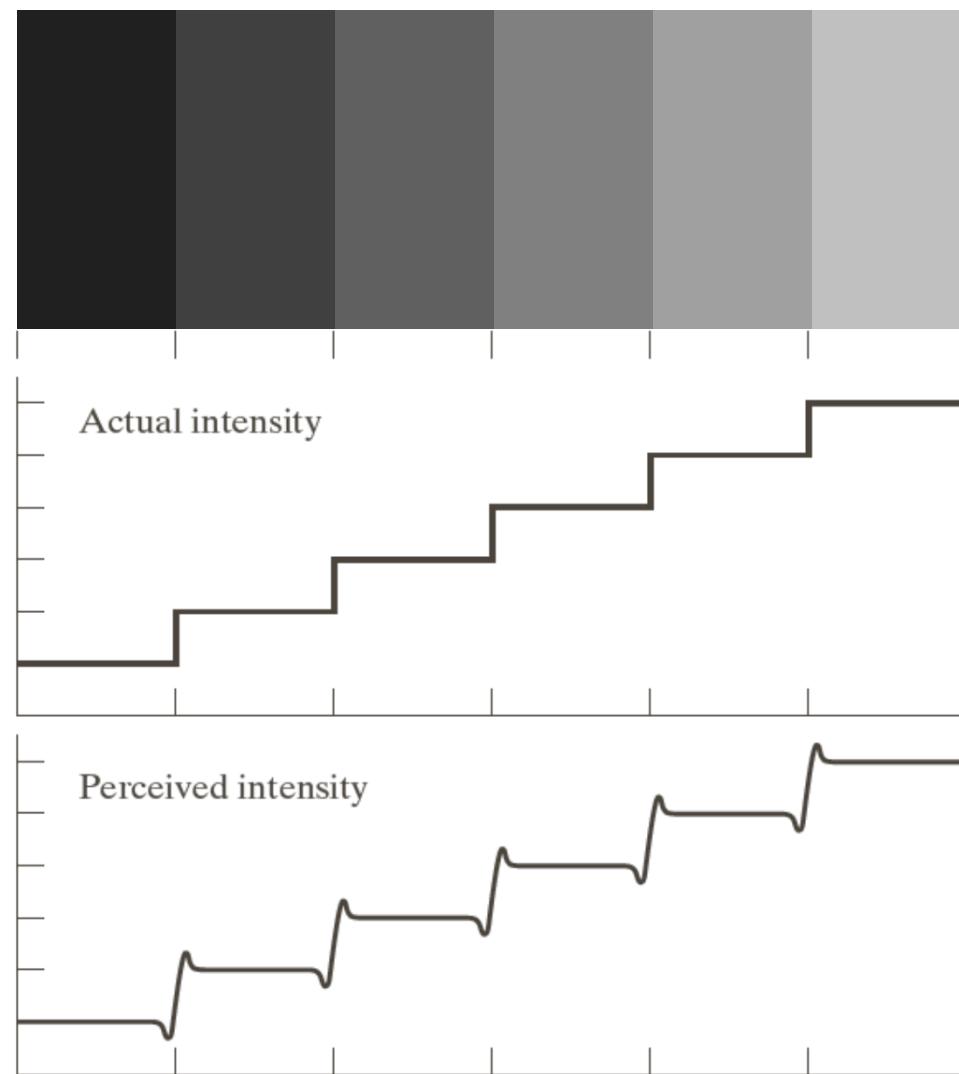
- 让人看一块足够大的均匀光照的平面（如光照 I 的点光源照射毛玻璃）
- 在中间圆形区域施加一个闪光(ΔI)
- 当 ΔI 很小时，实验者感觉不到；当 ΔI 很大时，实验者总能感觉到；
- 找到 ΔI_c (50%感觉到)
- Weber率： $\Delta I_c/I$

光线暗时，人眼对光强变化的区分力弱

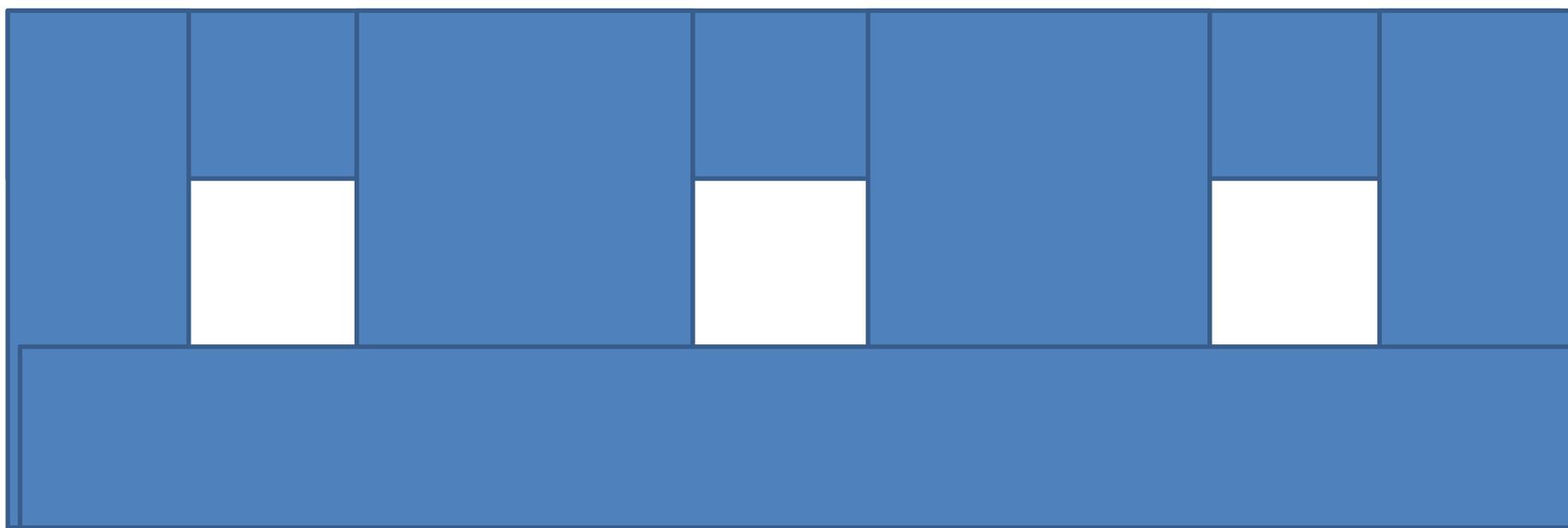
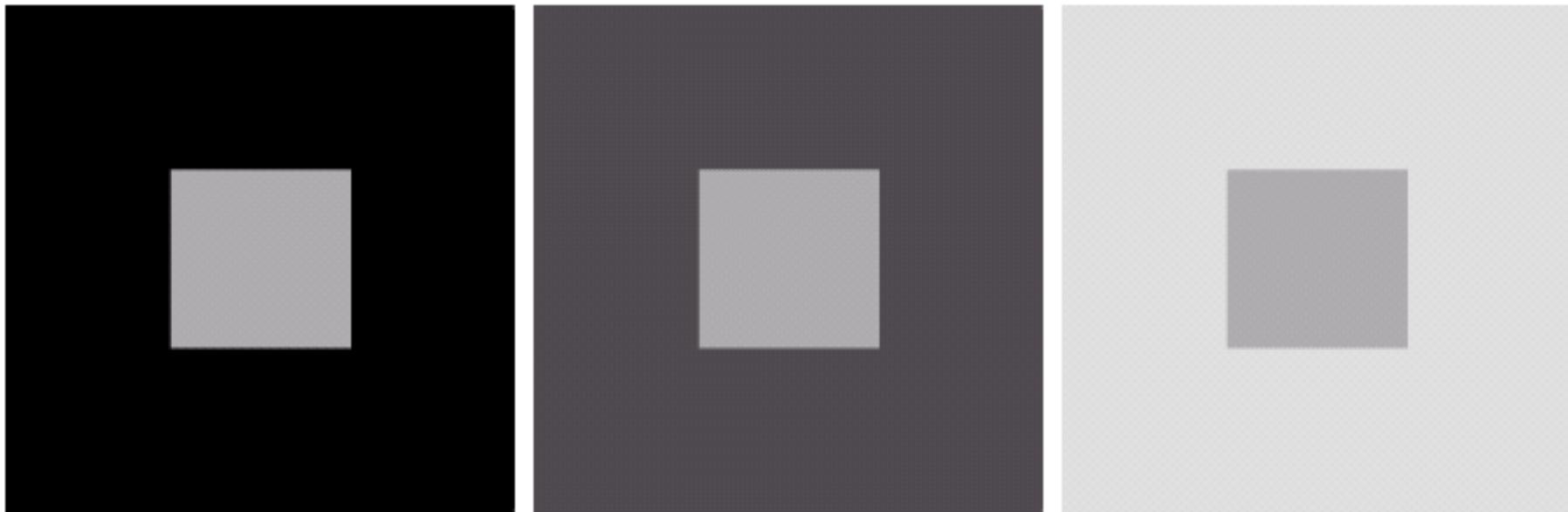




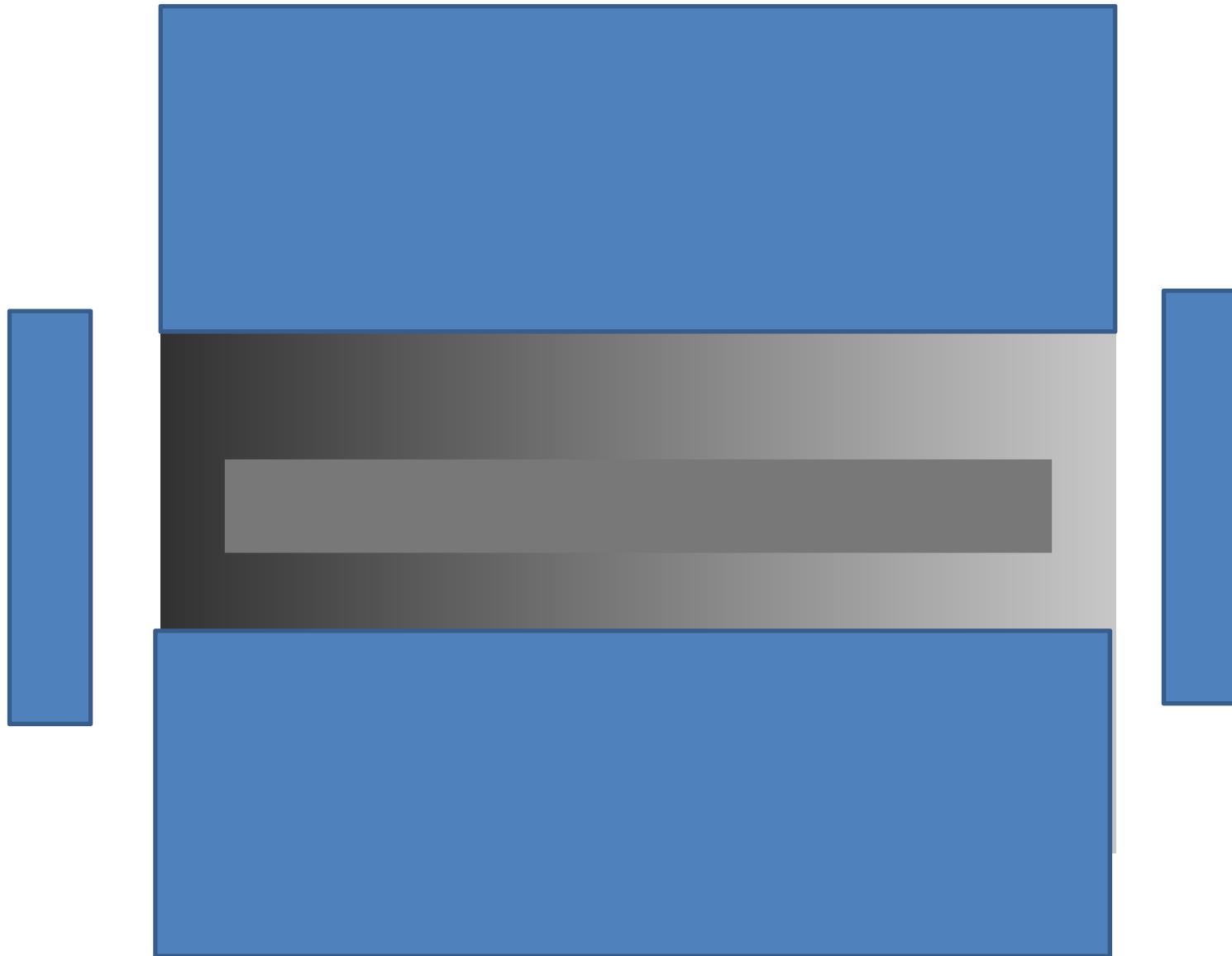
主观亮度不是光强的简单函数



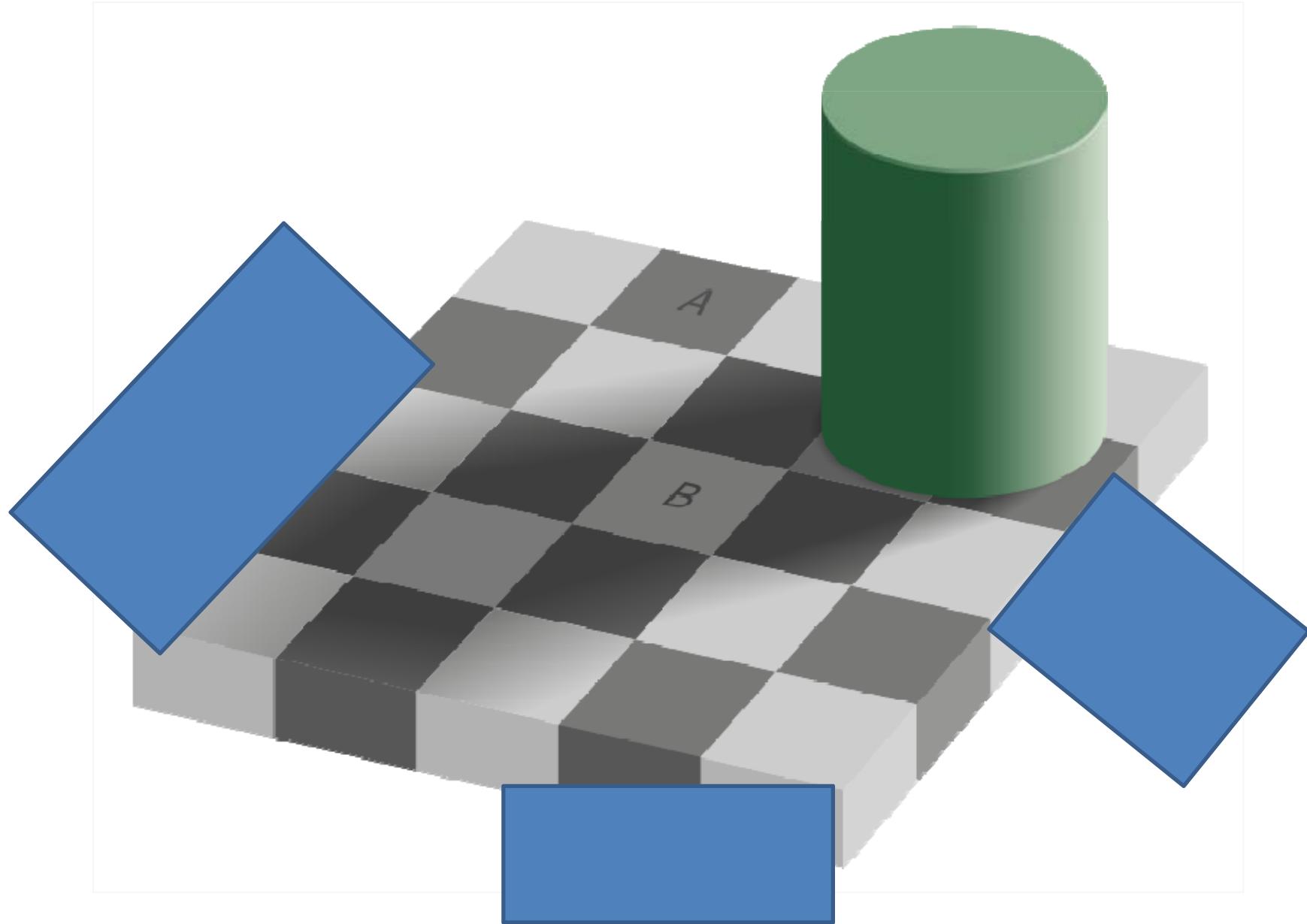
Simultaneous Contrast Illusion



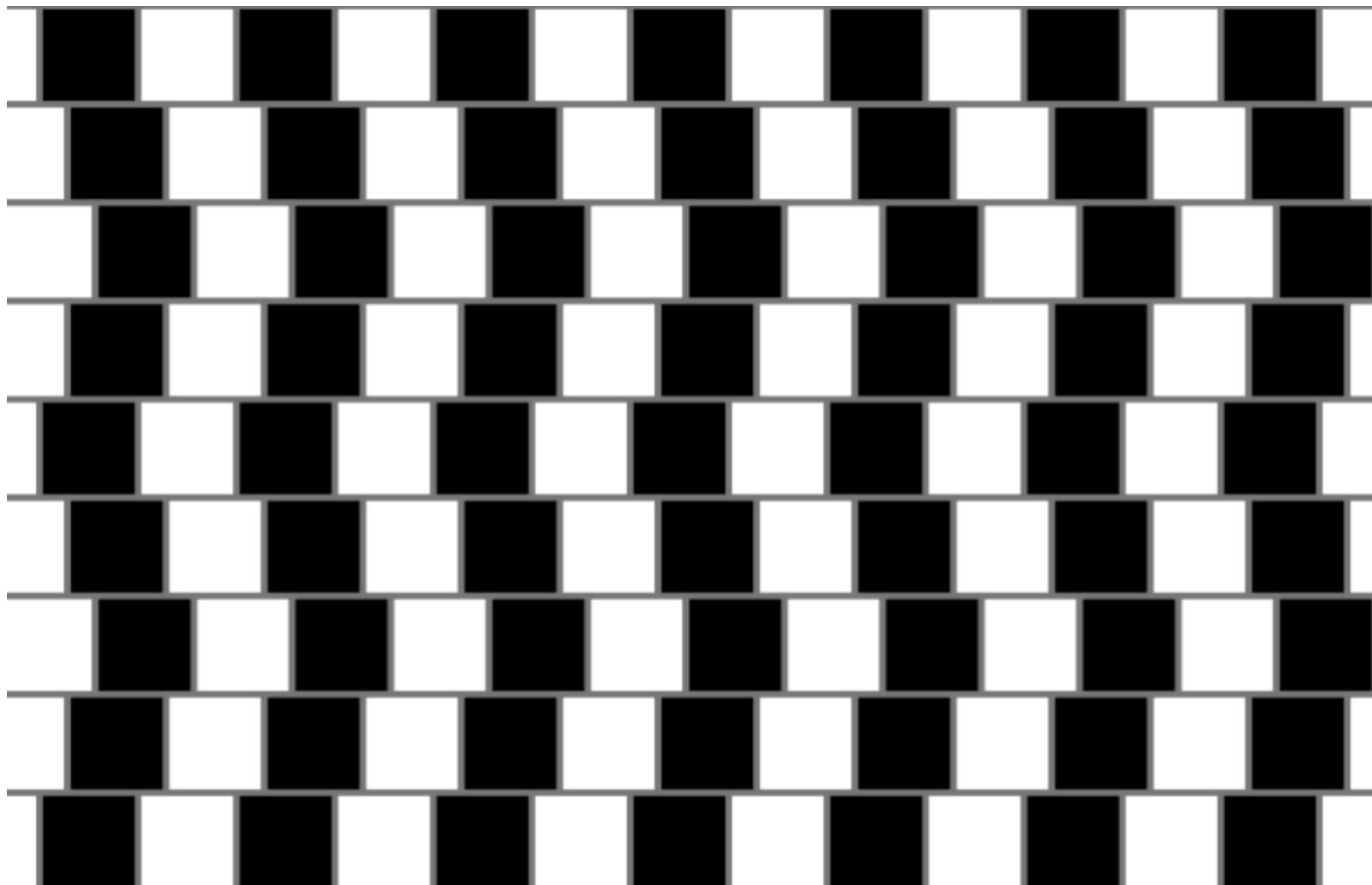
Simultaneous Contrast Illusion



Simultaneous Contrast Illusion

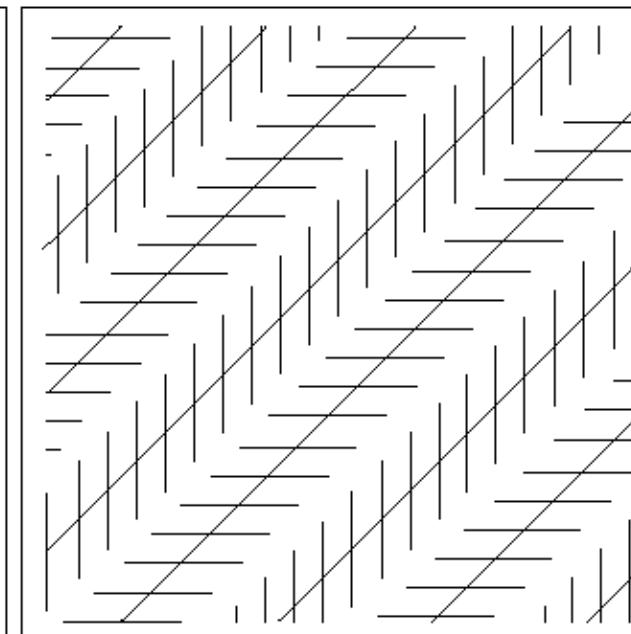
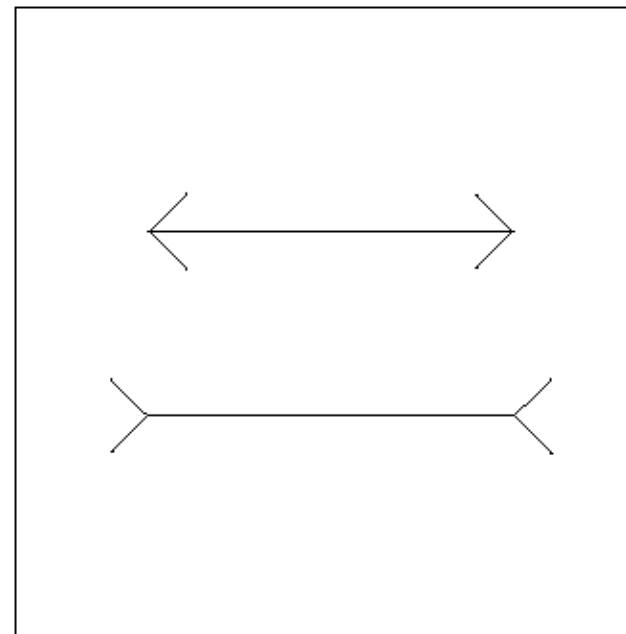
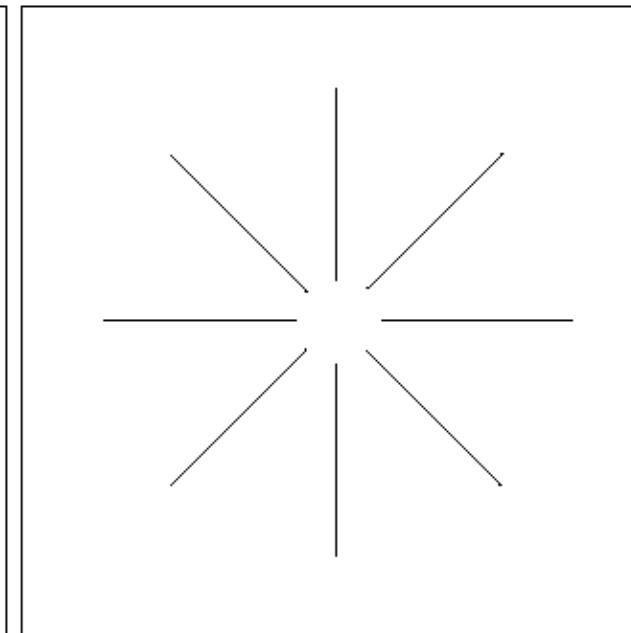
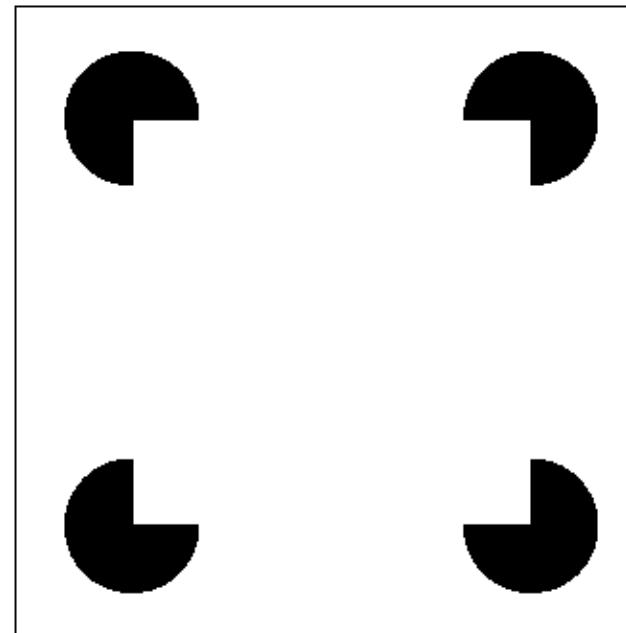


Optical Illusion



https://en.wikipedia.org/wiki/Optical_illusion

Optical Illusion



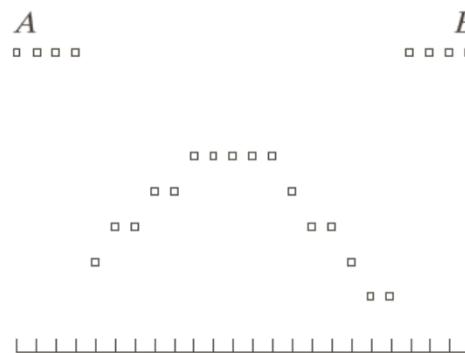
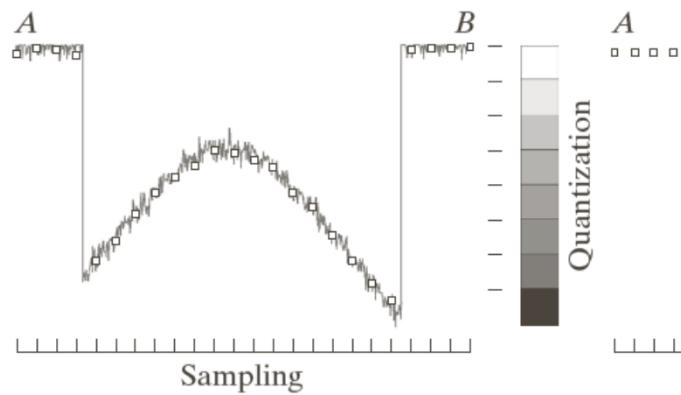
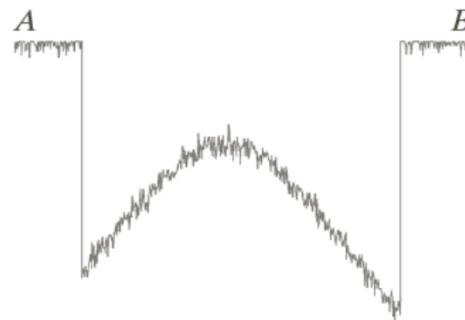
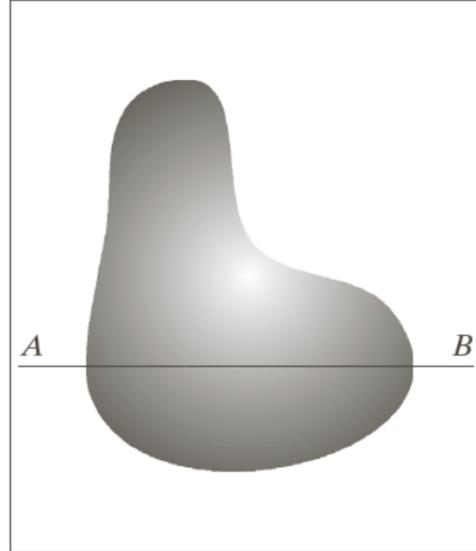
内容

1. 人眼
2. 采样与量化
3. 像素之间的关系
4. 图像处理中的数学工具

2. 图像采样与量化

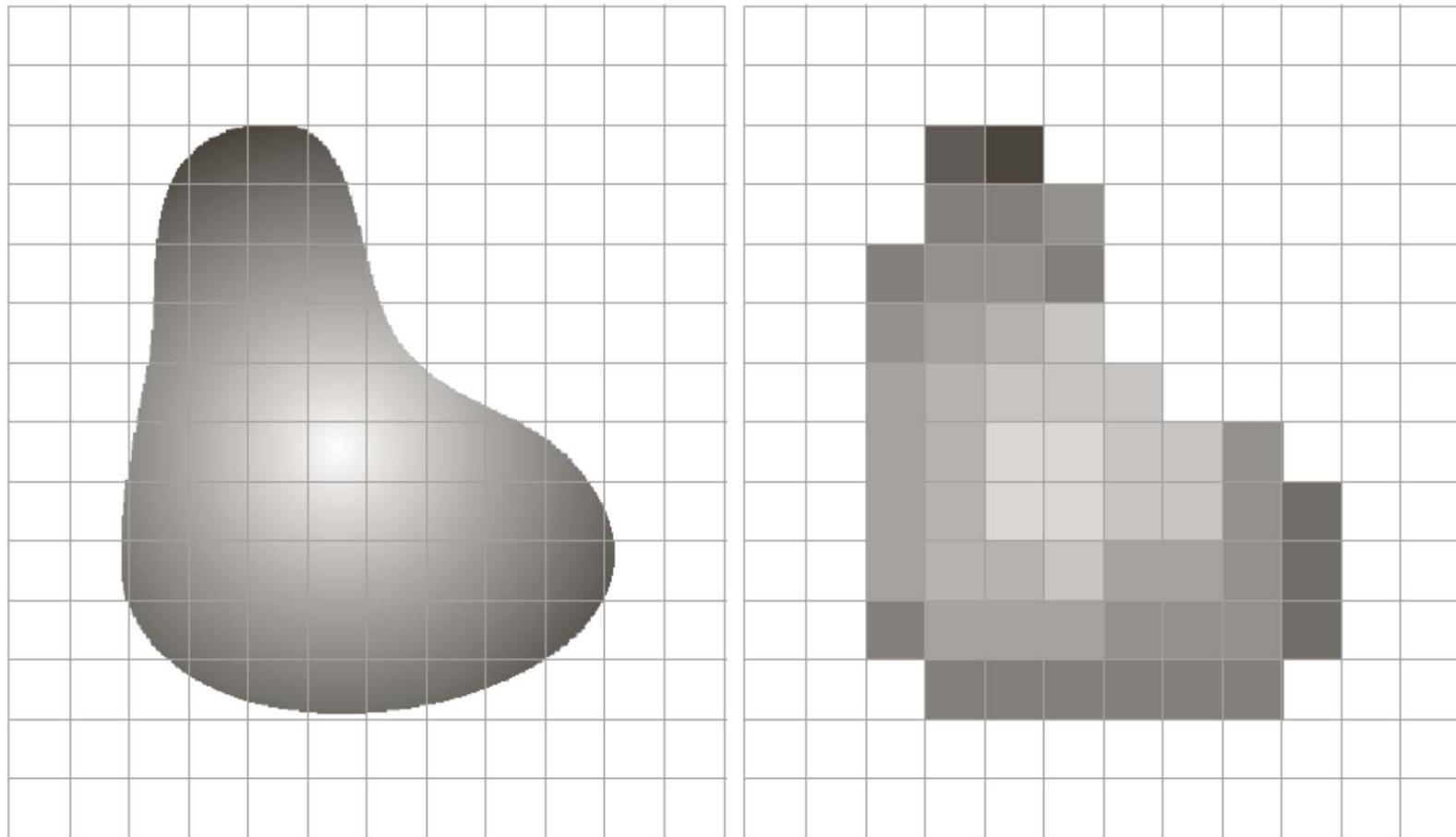
- 采样与量化
- 数字图像的表示
- 空间和亮度的分辨率
- 图像插值

采样 (sampling) 和量化 (quantization)



- 实际的物理世界，图像是二元函数，函数的2个自变量（ xy 坐标）和函数值（亮度）都是连续的。
- 为了在计算机上进行处理，需要数字化。
 - **采样**：将连续的坐标数字化
 - **量化**：将连续的亮度数字化

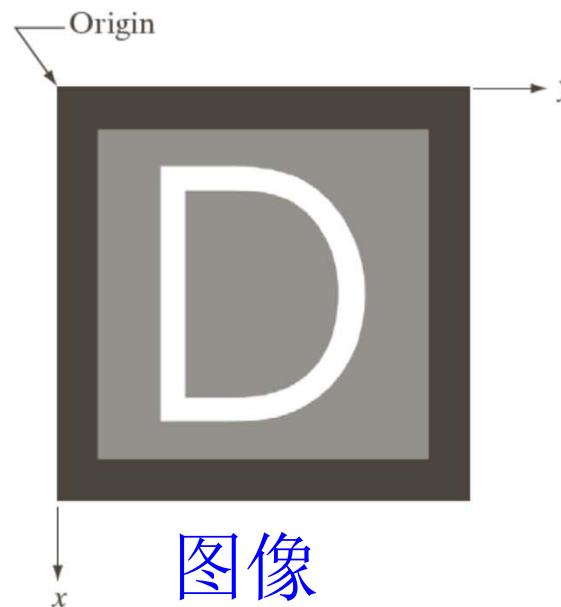
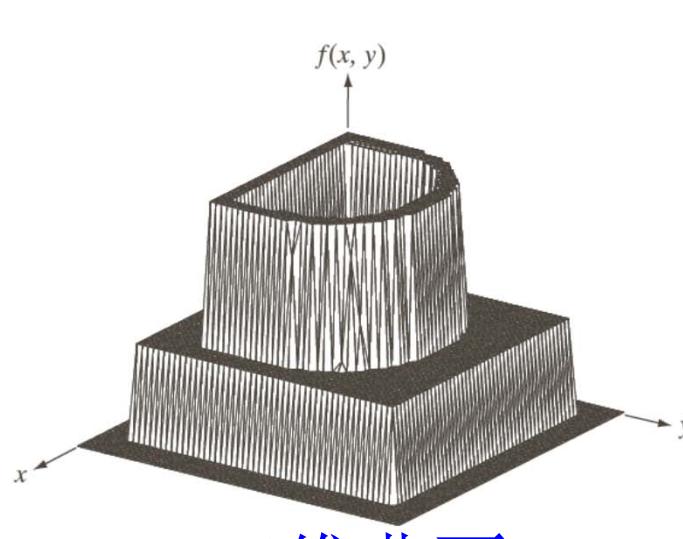
采样和量化



数字图像的表示

- 数字图像 $f(x, y)$ 是一个二维数组
- $x = 0, 1, 2, \dots, M - 1$
- $y = 0, 1, 2, \dots, N - 1$
- $$f(x, y) = \begin{bmatrix} f(0, 0) & \cdots & f(0, N - 1) \\ \vdots & \ddots & \vdots \\ f(M - 1, 0) & \cdots & f(M - 1, N - 1) \end{bmatrix}$$
- $f(x, y)$ 的2种含义：整幅图像或 (x, y) 处的图像灰度
- $f(x, y) \in [0, L - 1]$, $L = 2^k$ 为灰度阶数（级数）
- 最常见的灰度图像： $k = 8, L = 256$

数字图像的3种显示方式



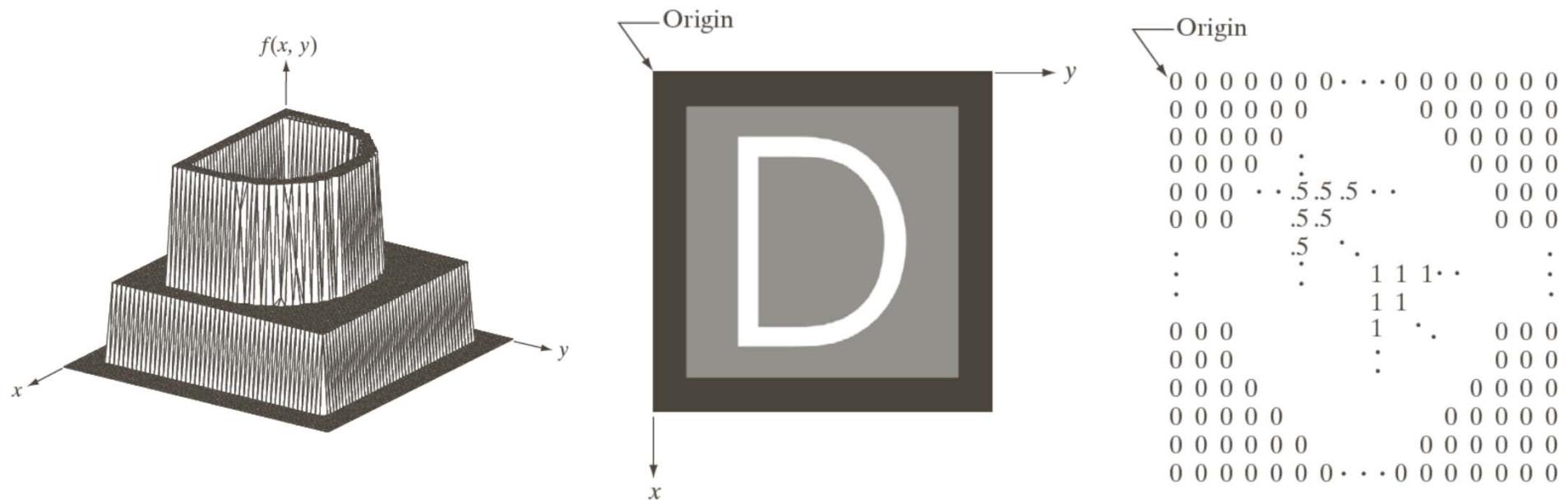
Origin

0	0	0	0	0	0	0	...	0	0	0	0	0	0	0
0	0	0	0	0	0	0		0	0	0	0	0	0	0
0	0	0	0	0	0	0	.	0	0	0	0	0	0	0
0	0	0	0	0	0	0	.	0	0	0	0	0	0	0
0	0	0	0	0	0	0	..	5	5	5	..	0	0	0
0	0	0	0	0	0	0	5	5	5	5	5	0	0	0
.	5	5	5	5	5	0	0
.	1	1	1	1
.	1	1	1	1	1
0	0	0	0	0	0	0	1	1	1	1	1	..	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

矩阵

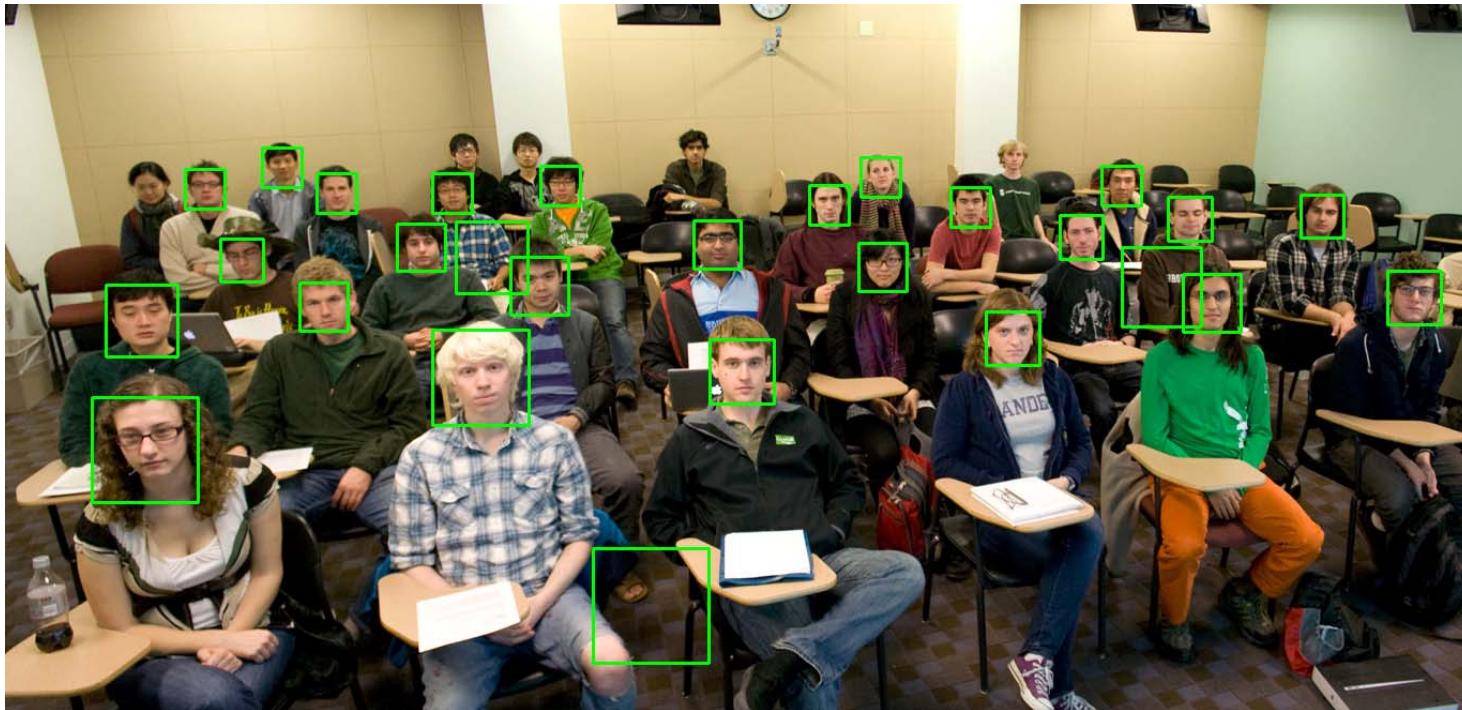
- 图像坐标系与二维矩阵的行列排序是一致的
- 图像坐标系不同于常见的坐标系
- 图像坐标系逆时针转90度，就是常见的坐标系
- 图像坐标系也是右手坐标系

数字图像的3种显示方式



- 图像: 最常见, 适合人眼来检查算法结果
- 矩阵: 程序调试时检查数值
- 三维曲面: 图像值是深度、与深度相关、峰值具有意义

外行容易低估图像识别的困难



- 外行很难理解图像识别的困难
- 不但外行不理解，早期的大科学家也曾经没有意识到
- 人工智能先驱明斯基（Marvin Minsky）1966年布置的暑假作业：在计算机接上摄像头，写个程序识别出摄像头看到了什么。
- 深入浅出地解释专业问题是极其重要的能力

向外行解释图像识别的困难

	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	
49	185	186	185	186	188	188	189	191	189	190	190	191	191	192	193	193
50	184	184	184	184	187	188	187	189	191	190	190	191	191	193	193	193
51	183	184	184	187	188	186	188	191	189	189	190	190	190	191	192	192
52	184	186	185	186	186	186	188	189	188	188	189	190	190	191	191	192
53	183	185	184	184	186	186	186	187	186	187	190	190	190	190	190	192
54	182	184	183	184	186	187	186	188	187	188	190	189	191	190	191	191
55	184	185	185	185	186	186	186	186	186	189	189	188	190	190	190	190
56	183	183	184	184	185	185	185	184	185	188	187	187	189	191	191	190
57	182	183	183	181	184	187	185	184	185	187	188	188	190	190	190	191
58	183	182	183	183	185	187	185	184	185	187	190	188	189	189	190	190
59	184	183	184	184	185	186	185	184	186	187	188	186	188	190	190	190
60	182	182	184	183	183	183	183	184	185	186	185	187	188	189	189	188
61	183	181	183	184	183	183	183	185	186	185	184	187	189	188	186	186
62	182	183	183	182	183	183	182	184	185	184	184	186	187	187	187	187
63	180	182	182	182	183	182	182	183	184	184	184	185	185	185	185	187
64	180	181	181	182	182	182	183	183	184	185	185	185	185	186	186	186
65	182	180	180	179	181	182	184	184	183	184	184	185	186	187	185	185
66	182	181	181	180	181	183	183	184	182	182	184	185	185	186	186	185
67	180	181	181	180	181	182	182	183	183	182	183	184	184	186	186	186
68	180	180	179	179	180	180	182	183	182	182	183	184	184	185	185	185
69	180	181	179	179	180	180	181	182	180	182	183	183	183	183	183	183
70	179	170	180	170	180	180	181	181	180	180	181	181	182	181	181	180

- 计算机只会加减乘除和逻辑运算
- 人脸检测：用一系列基本运算来从一组数组中找出人脸

空间分辨率 (spatial resolution)



产品型号	Nokia/诺基亚216DS
屏幕尺寸	2.4英寸
分辨率	320×240
适用网络	移动2G/联通2G
运行内存	16MB
存储容量	1.5MB 可扩展32GB
操作系统	Series 30+
摄像头	前置30万像素+后置30万像素
电池容量	1020mAh
键盘类型	12键标准键盘
特色功能	双卡双待 支持蓝牙

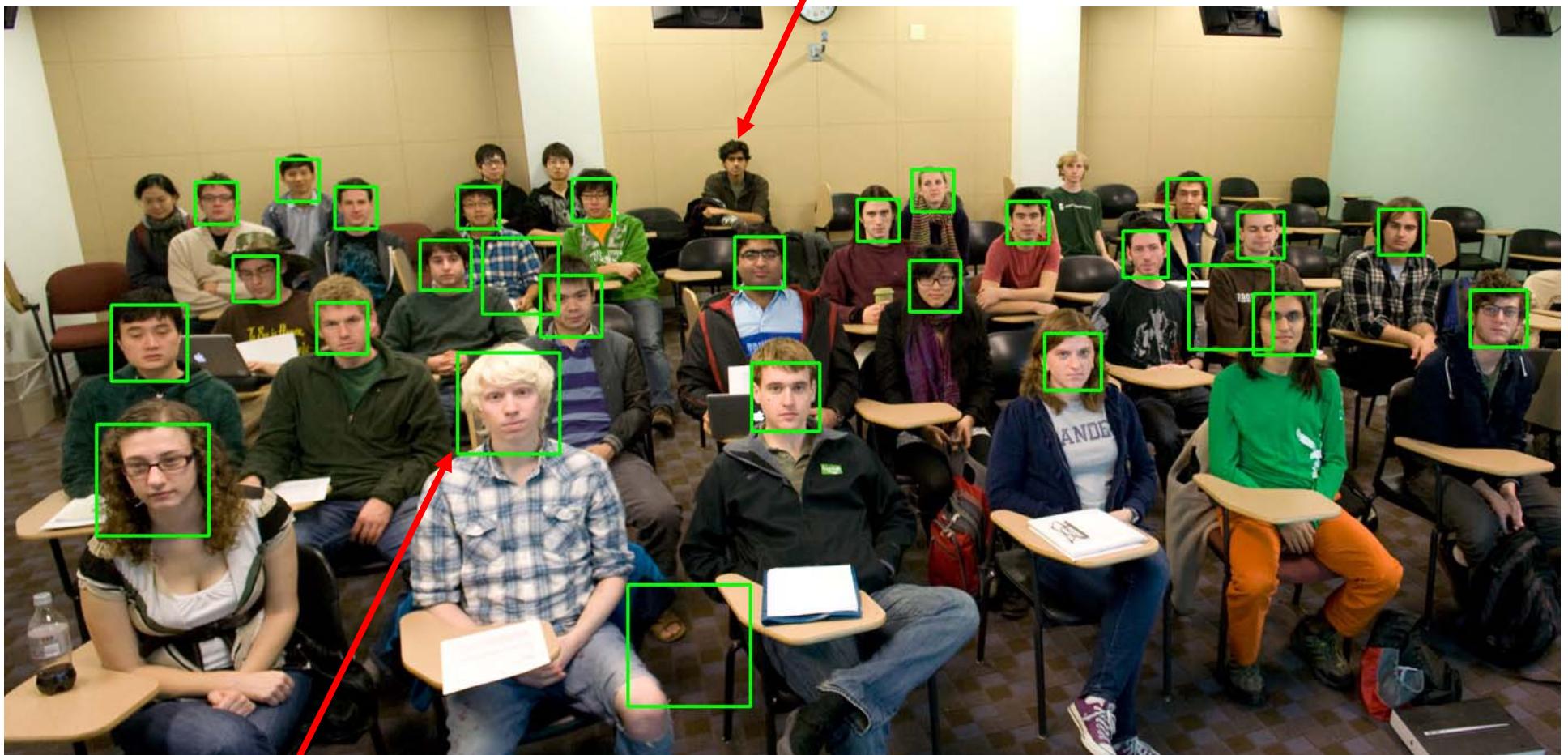
R11 Plus
前后 2000 万，拍照更清晰。



- 手机的一个重要指标是摄像头分辨率
- 分辨率越高， 价格越高

脱离拍摄对象的物理尺寸，图像分辨率意义不大

该图像的分辨率为200万像素



20x20像素

100x100像素

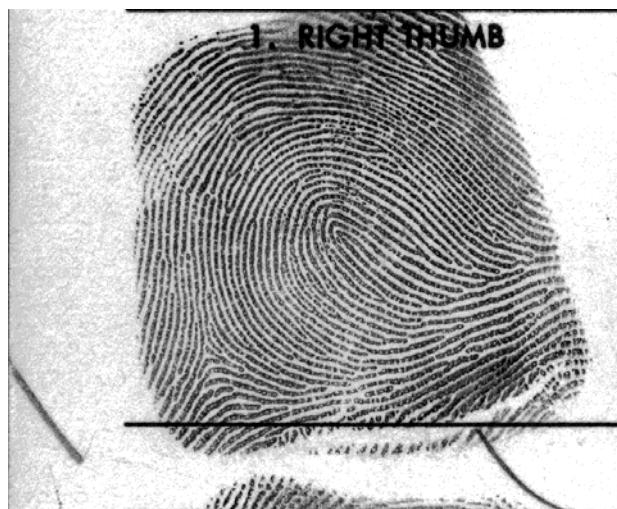
空间分辨率

有些图像中（指纹、CT图像、扫描文档），分辨率有物理意义

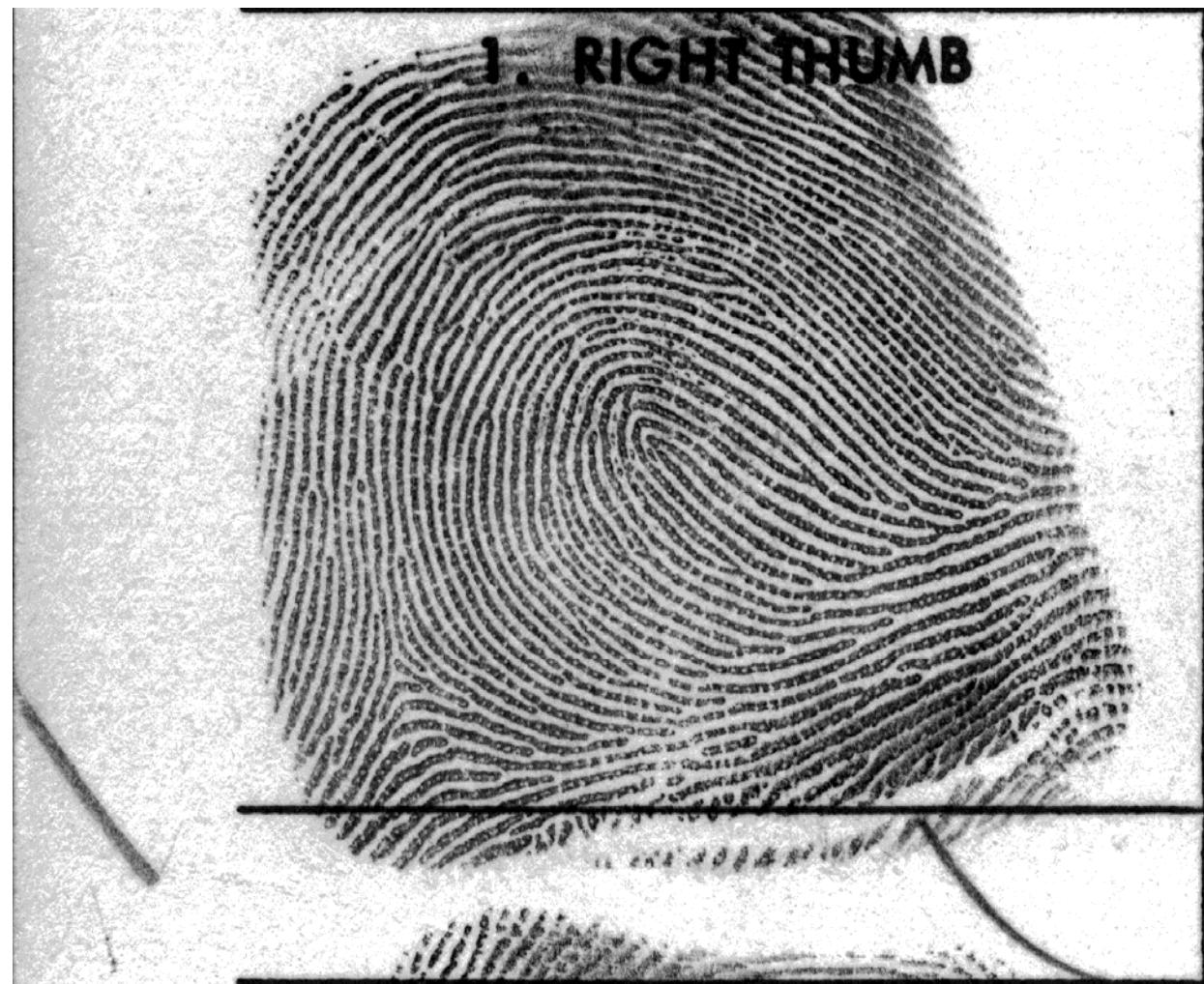
成年男性指纹脊线平均间距为0.5毫米

身份证指纹的标准分辨率是500 DPI

犯罪前科人员的指纹分辨率是1000 DPI

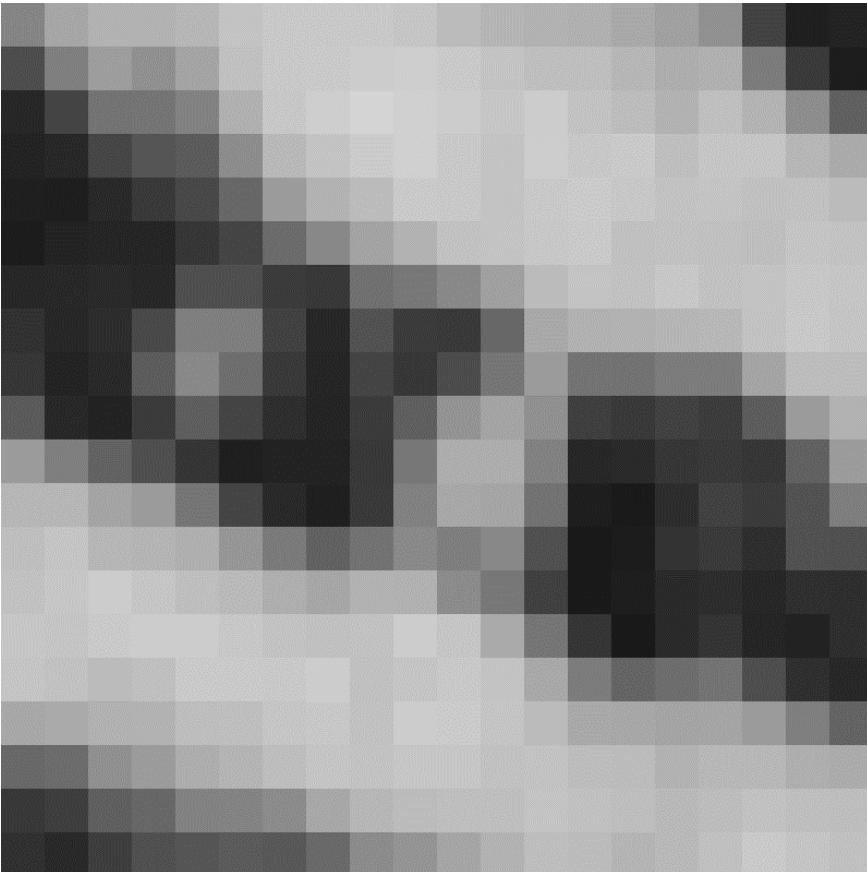


500 DPI

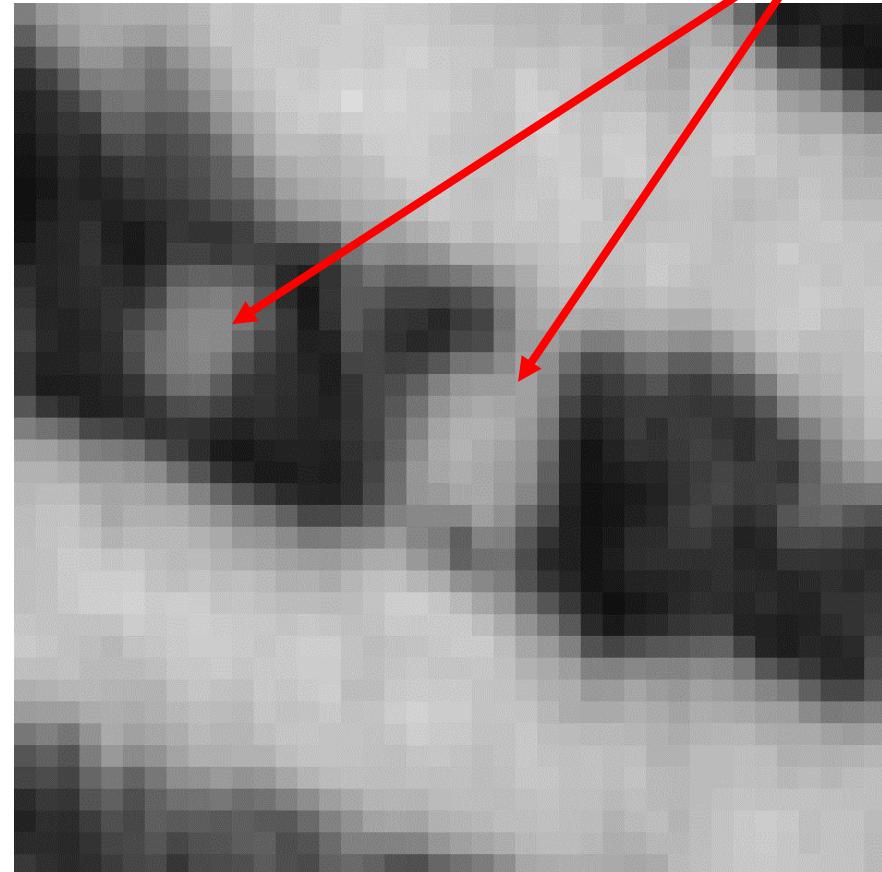


1000 DPI

不同分辨率的指纹局部



500 DPI



1000 DPI

放大到相同尺寸下的指纹局部图像

不同空间分辨率



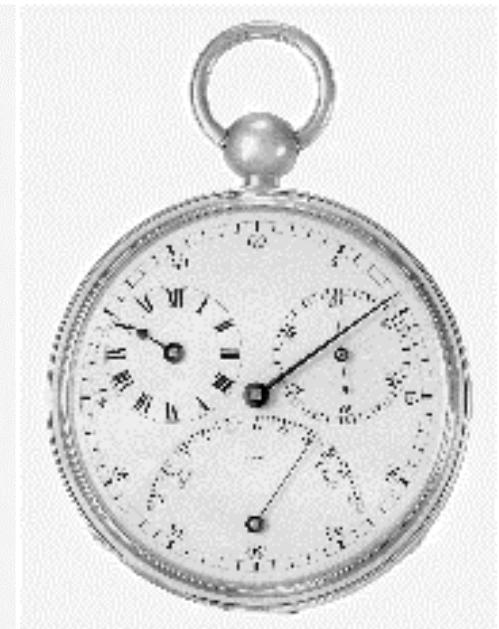
3692 × 2812



923 × 703



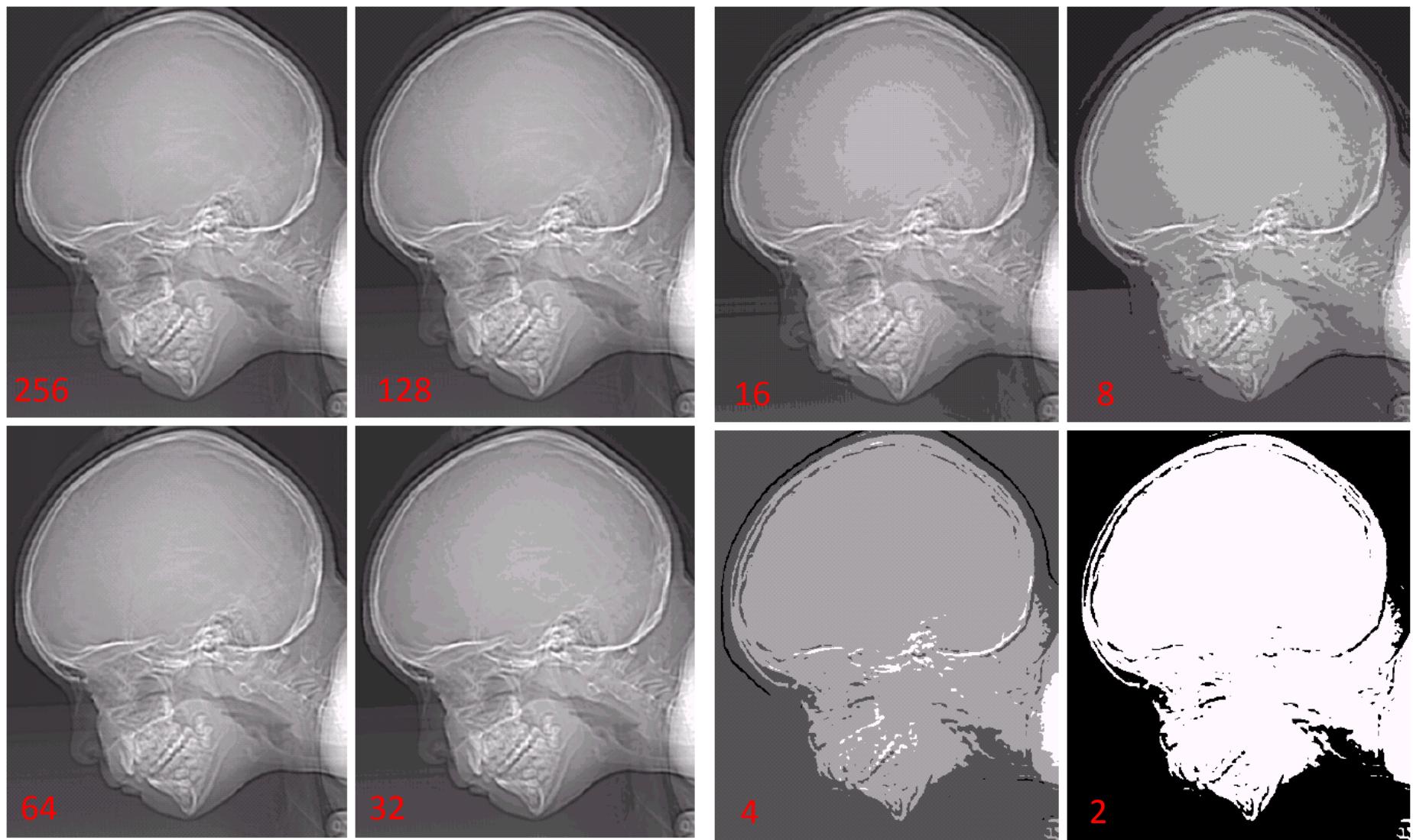
462 × 352



231 × 176

灰度分辨率 (intensity resolution)

即像素的灰度级数

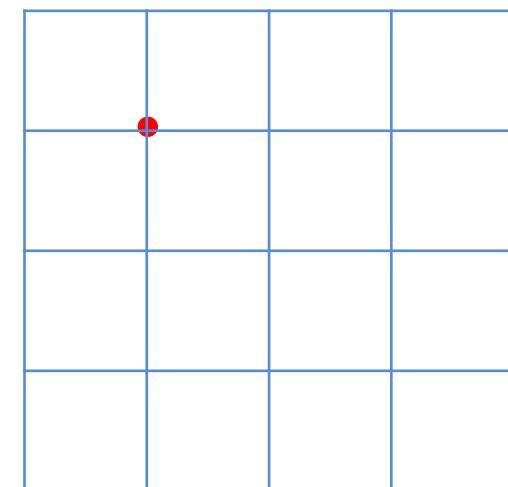
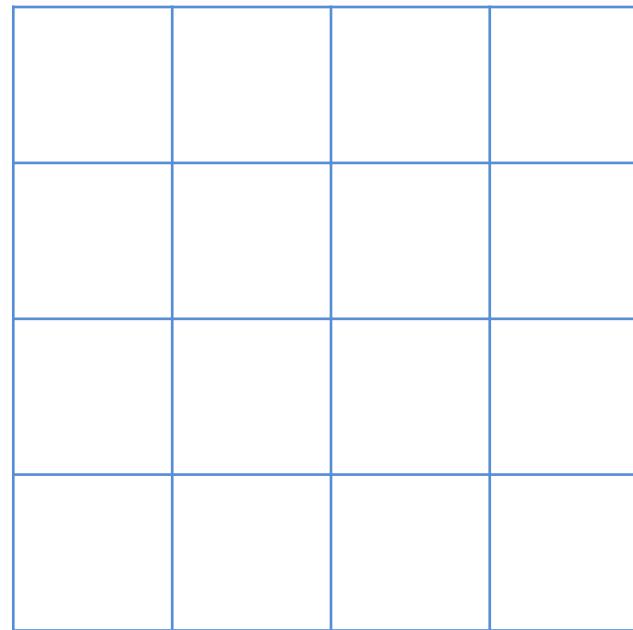
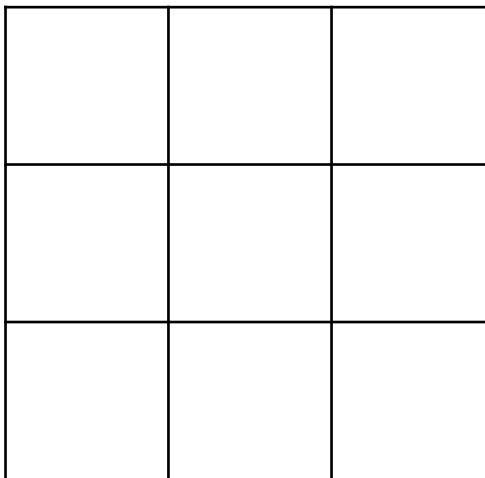


图像插值 (interpolation)

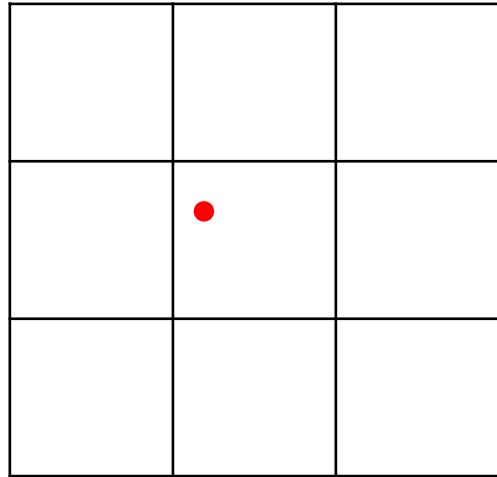
- 数字图像是对连续图像进行采样和量化的结果。
- 当对数字图像进行缩小、放大时，需要做重采样 (resampling)。
- 新采样点往往不在原始像素上，取值未知，因此需要从已知数字图像估算新采样点处的灰度值。
- 图像的各种几何变换（缩放、旋转、非线性的变换）都需要用到插值。
- 下面介绍3种最常见的插值方法。

最近邻插值 (nearest neighbor interpolation)

- 4×4 图像放大到 5×5 图像（线是采样网格，顶点是像素）
- 这样理解放大过程：生成 5×5 采样网格，将其尺寸缩小为原始图像，对于每个采样点取最近邻像素

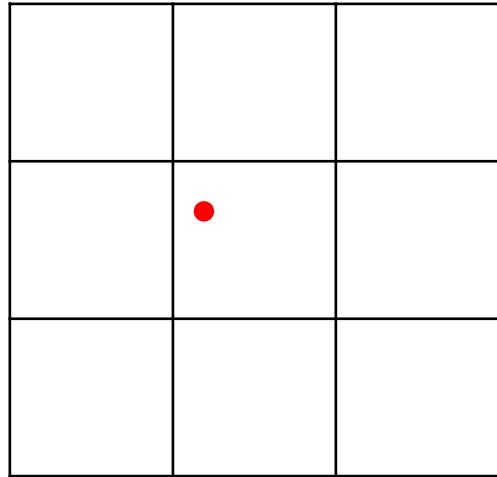


双线性插值 (bilinear interpolation)



采样点 (x, y) 处的插值公式: $v(x, y) = ax + by + cxy + d$
4个系数 (a, b, c, d) 由所在方格4个顶点对应的4个等式计算得到

双三次插值 (bicubic interpolation)



$$\text{插值公式: } v(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j$$

16个系数由16个顶点对应的16个等式计算得到

这是Photoshop图像变换默认使用的插值方法

高级的图像超分辨率技术

- 对于图像缩小，基本的插值方法够用，没有可研究的。
- 对于图像放大（也叫图像超分辨率），基本的插值方法不理想，是研究热点。
- 放大需要增加信息，这些信息从哪里来？
- 目前高性能的超分辨率技术利用机器学习

内容

1. 人眼
2. 采样与量化
3. 像素之间的关系
4. 图像处理中的数学工具

3. 像素之间的关系

- 像素之间的关系
- 区域之间的关系
- 像素之间的距离

像素之间的关系

- 坐标为 (x, y) 的像素 p 的近邻
 - 2个水平近邻: $(x, y - 1), (x, y + 1)$
 - 2个垂直近邻: $(x - 1, y), (x + 1, y)$
 - 4个对角近邻: $(x - 1, y - 1), (x - 1, y + 1), (x + 1, y - 1), (x + 1, y + 1)$
 - 4邻域 $N_4(p)$
 - 8邻域 $N_8(p)$

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

邻接、区域

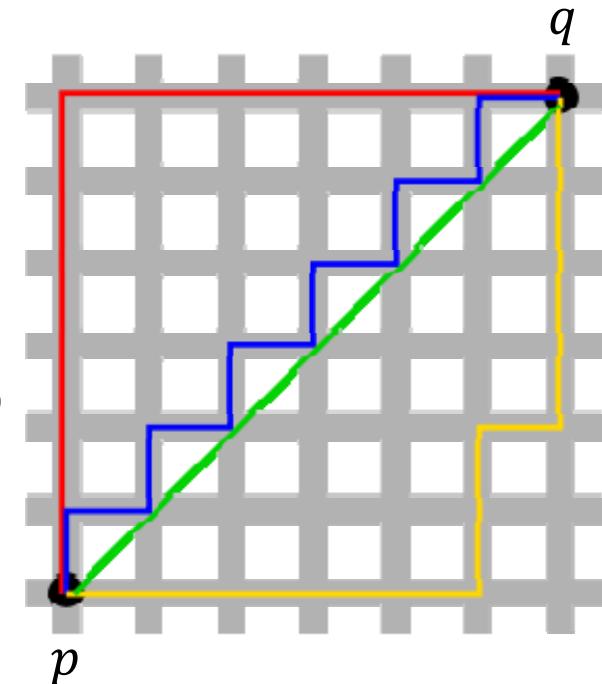
- 二值图像中，1为前景，0为背景
- 4邻接、8邻接：像素值都为1的4近邻、8近邻
- 路径：2个像素之间的一序列邻接像素
- S 是一幅图像中一些像素的集合；
- 如果像素 p 和像素 q 在 S 中存在路径， p 和 q 是连通的。
- 集合 S 中，与像素 p 连通的像素构成连通成分；如果只有1个连通成分，也称连通集或者区域。
- 如果2个区域能构成连通集，它们是邻接的。

$$\begin{matrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix} \left. \begin{array}{l} \\ \\ \{ \\ \\ \} \\ \end{array} \right\} R_i$$
$$\begin{matrix} \\ \\ \\ \\ \\ \end{matrix} \left. \begin{array}{l} \\ \\ \{ \\ \\ \} \\ \end{array} \right\} R_j$$

像素之间的距离

像素 p 和像素 q 之间的距离 D , p 坐标为 (x, y) , q 坐标为 (s, t)

- Euclidean distance (欧氏距离) D_e
- City-block distance (城市街区距离)
$$D_4 = |x - s| + |y - t|$$
- Chessboard distance (棋盘格距离)
$$D_8 = \max(|x - s|, |y - t|)$$



$D_4 \leq 2$ 的像素

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

$D_8 \leq 2$ 的像素

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

内容

1. 人眼
2. 采样与量化
3. 像素之间的关系
4. 图像处理中的数学工具

4. 数学工具

- Array vs. Matrix Operations
- Linear vs. nonlinear operations
- Arithmetic operations
- Set and logical operations
- Spatial operations
- Spatial transformations
- Image transforms
- Probabilistic methods

Array vs. Matrix Operations

- Two 2×2 arrays:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ and } \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

- Array product of these two images:

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

- Matrix product of these two images:

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

Linear vs. Nonlinear Operations

- 系统 H 对于输入图像 $f(x, y)$ 的响应（输出图像）为 $g(x, y)$

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

- 线性系统（线性运算）：输入为几幅图像的加权和，输出是各自响应的加权和

$$H[af_1(x, y) + bf_2(x, y)] = aH[f_1(x, y)] + bH[f_2(x, y)]$$

- 线性系统例子：

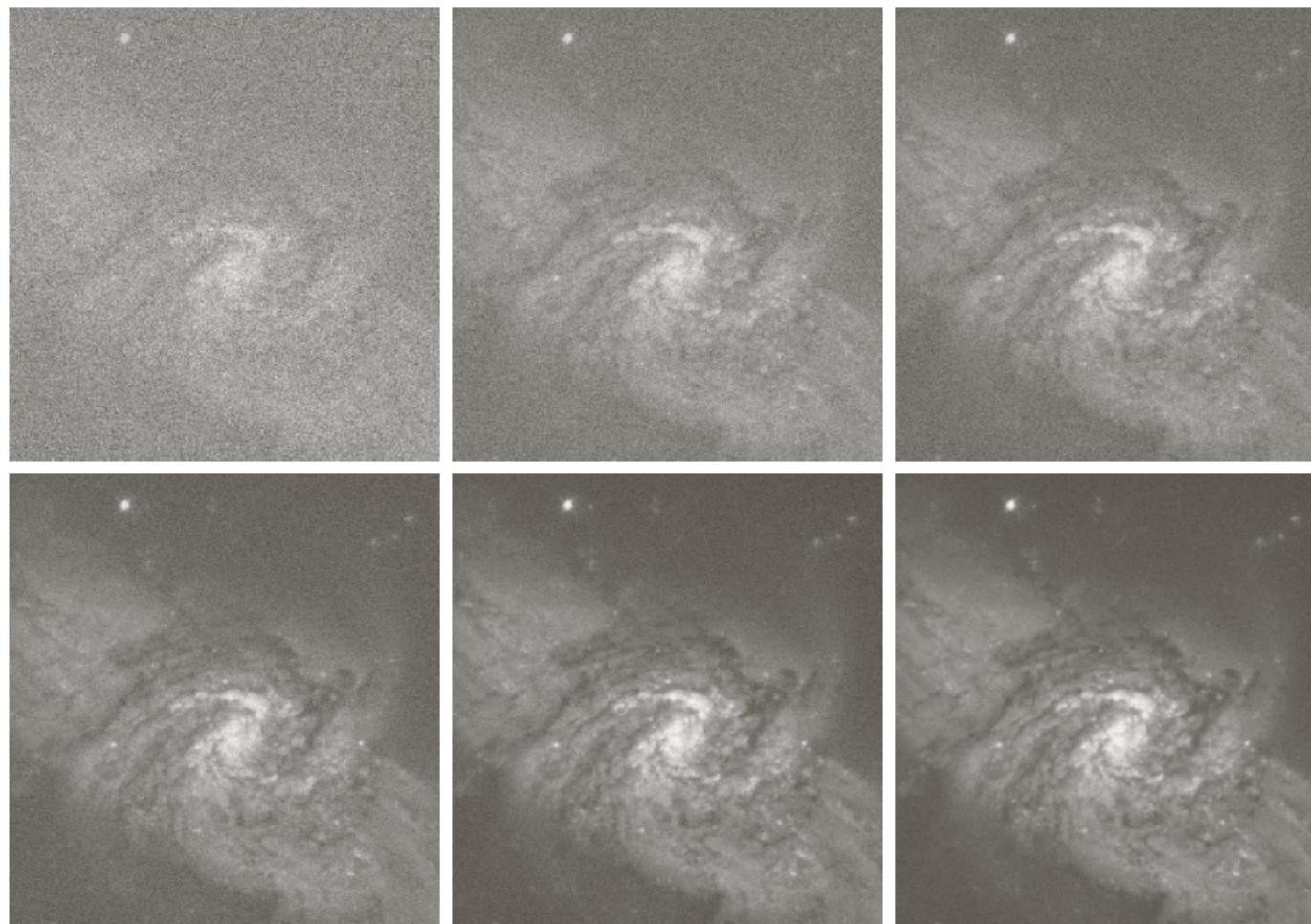
- 差分 $g(x, y) = f(x, y) - f(x - 1, y)$

- 均值滤波 $g(x, y) = \frac{1}{9} \sum_{i=-1}^1 \sum_{j=-1}^1 f(x - i, y - j)$

- 非线性系统例子： $g(x, y) = f^2(x, y)$

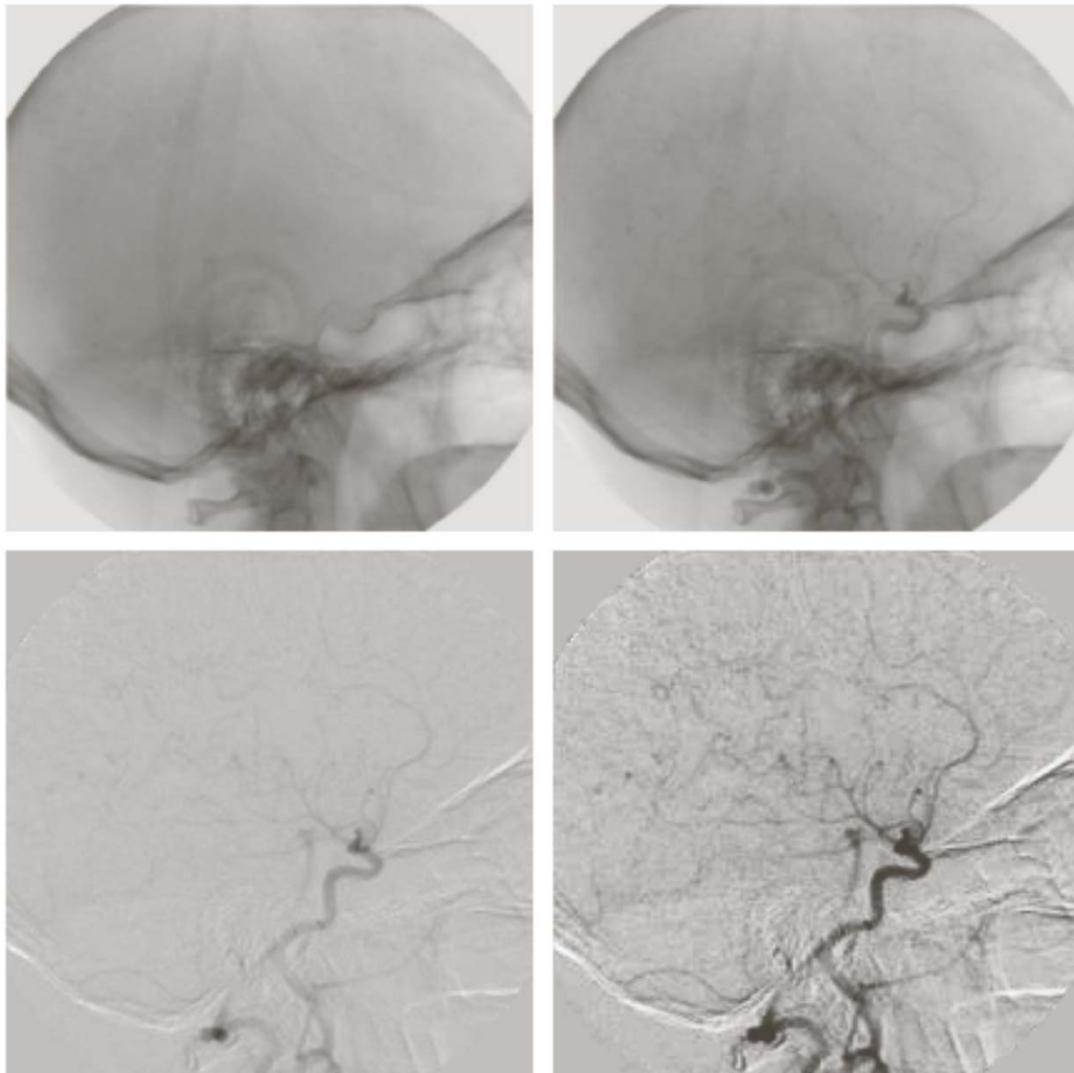
Arithmetic operations

- 图像加 (Array operation) , 图像大小相等
- 典型用途：去噪



Arithmetic operations

- 图像减，典型用途：突出差异



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.)

Arithmetic operations

- 图像乘，典型用途：阴影校正

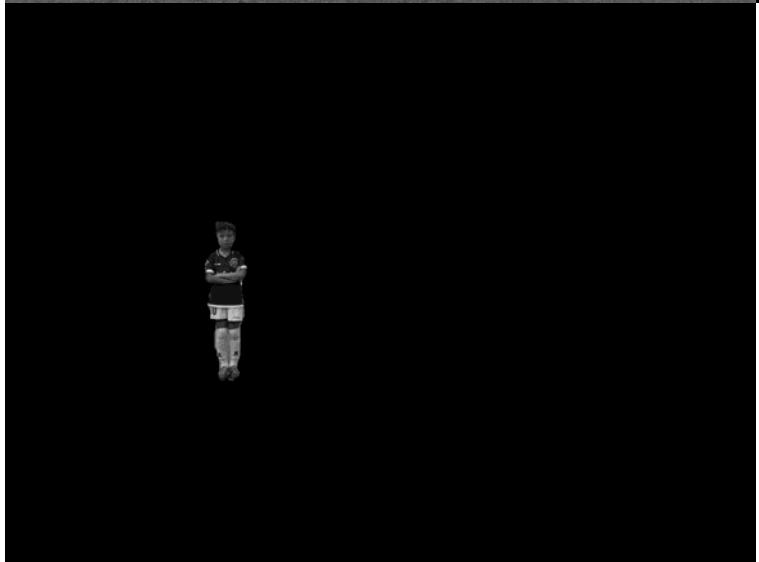
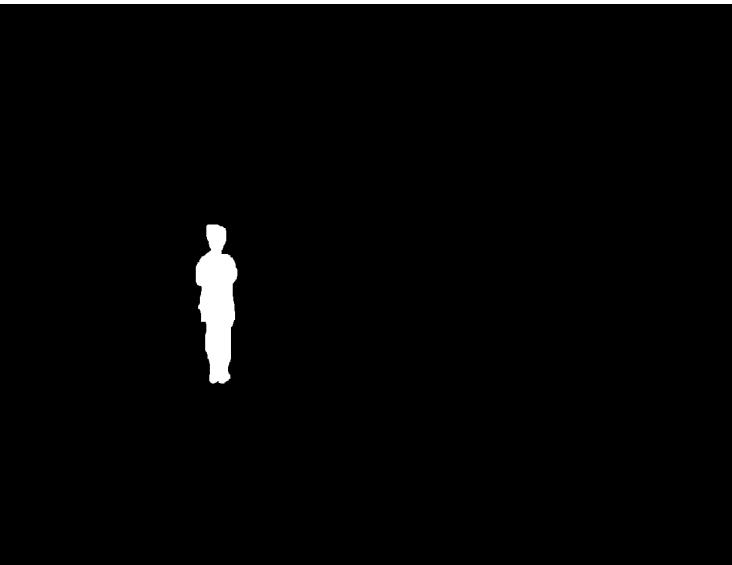


a b c

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.)

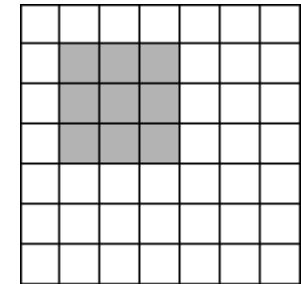
Arithmetic operations

- 图像乘，典型用途：提取感兴趣区域（ROI）

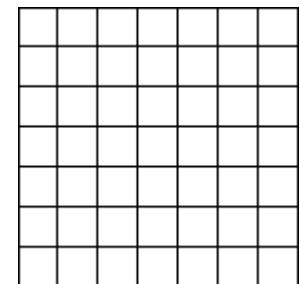


集合的概念

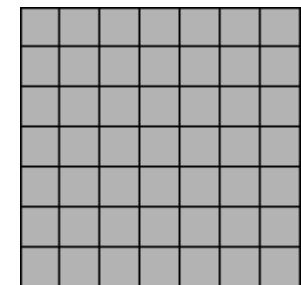
- 形态学图像处理主要关注像素坐标的集合
- 定义集合的方法：罗列元素或者用表达式
$$A = \{(2,2), (2,3), (2,4), (3,2), (3,3), (3,4), (4,2), (4,3), (4,4)\}$$
$$A = \{(x,y) | 2 \leq x \leq 4, 2 \leq y \leq 4\}$$
- 两个特殊的集合
 - 空集 \emptyset (empty set) : 没有元素的集合
 - 全集 U (universe set) : 全部像素坐标的集合



集合 A



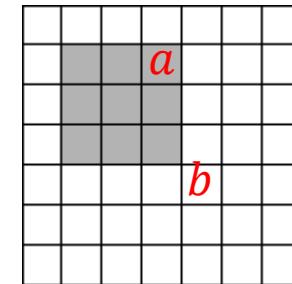
空集



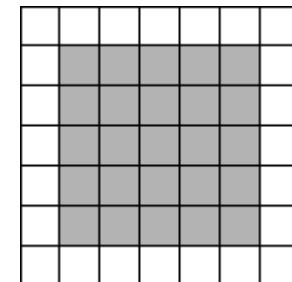
全集

集合的概念

- 如果 $a = (x, y)$ 是 A 的一个元素，称 a 属于 A ，记 $a \in A$ 。例如 $(4,2)$
- 如果 b 不是 A 的元素，称 b 不属于 A ，记 $b \notin A$ 。例如 $(5,5)$
- 子集：如果集合 A 的所有元素也属于集合 B ，则称 A 是 B 的子集，记为 $A \subseteq B$



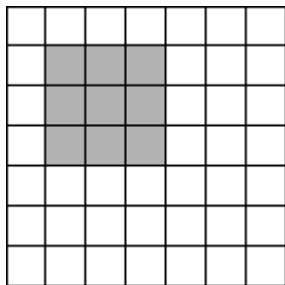
集合 A



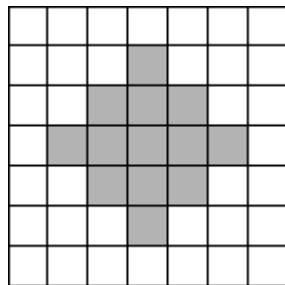
集合 B

集合的运算

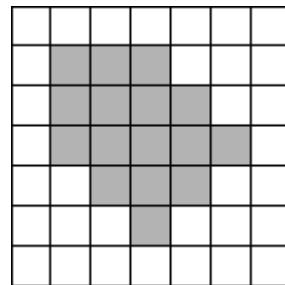
- A 和 B 的并集 (union) : $C = A \cup B$
- A 和 B 的交集 (intersection) : $D = A \cap B$
- 不相连 (互斥) : $A \cap B = \emptyset$
- A 和 B 的差集: $A - B = \{w | w \in A, w \notin B\} = A \cap B^c$
- A 的补集: $A^c = \{w | w \notin A\} = U - A$



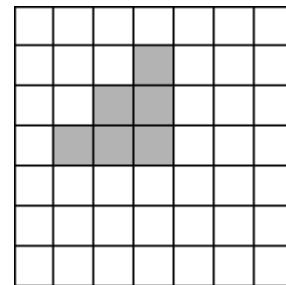
集合 A



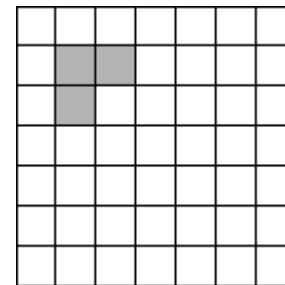
集合 B



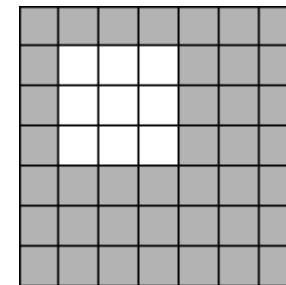
$A \cup B$



$A \cap B$



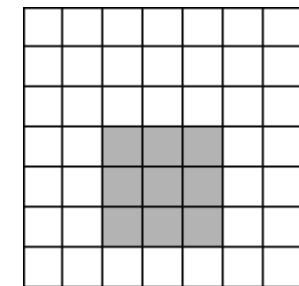
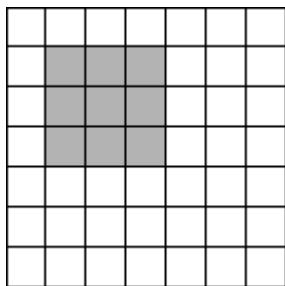
$A - B$



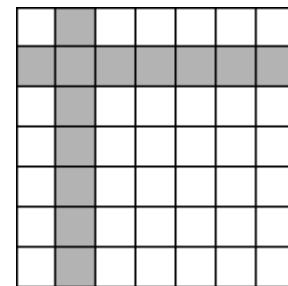
A^c

集合的运算

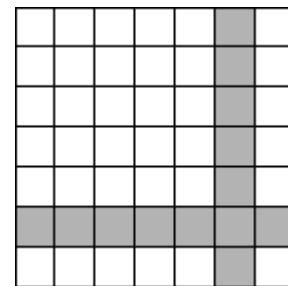
- 形态学运算需要将小图（结构元素）在另一幅大图上移动。集合的平移： $(A)_z$
- 有的形态学运算还需要将结构元素旋转180度（类似空域卷积）。集合的反射（旋转180度）： \hat{B}



平移



反射

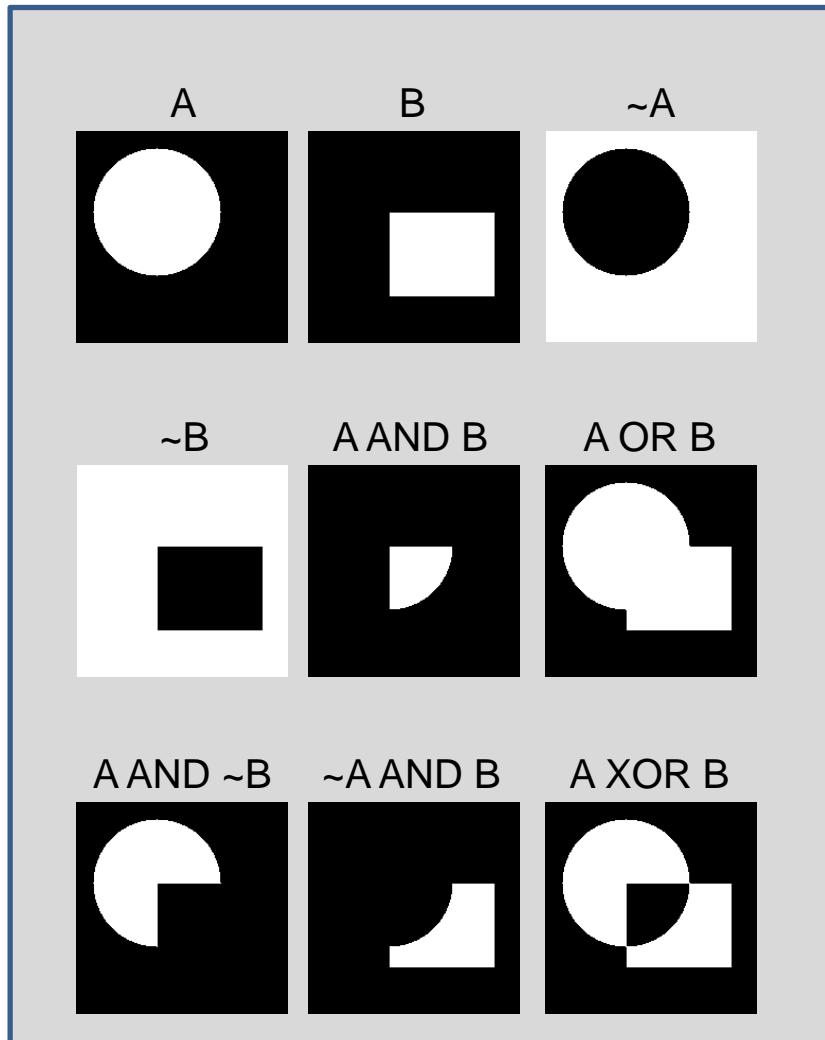


逻辑运算

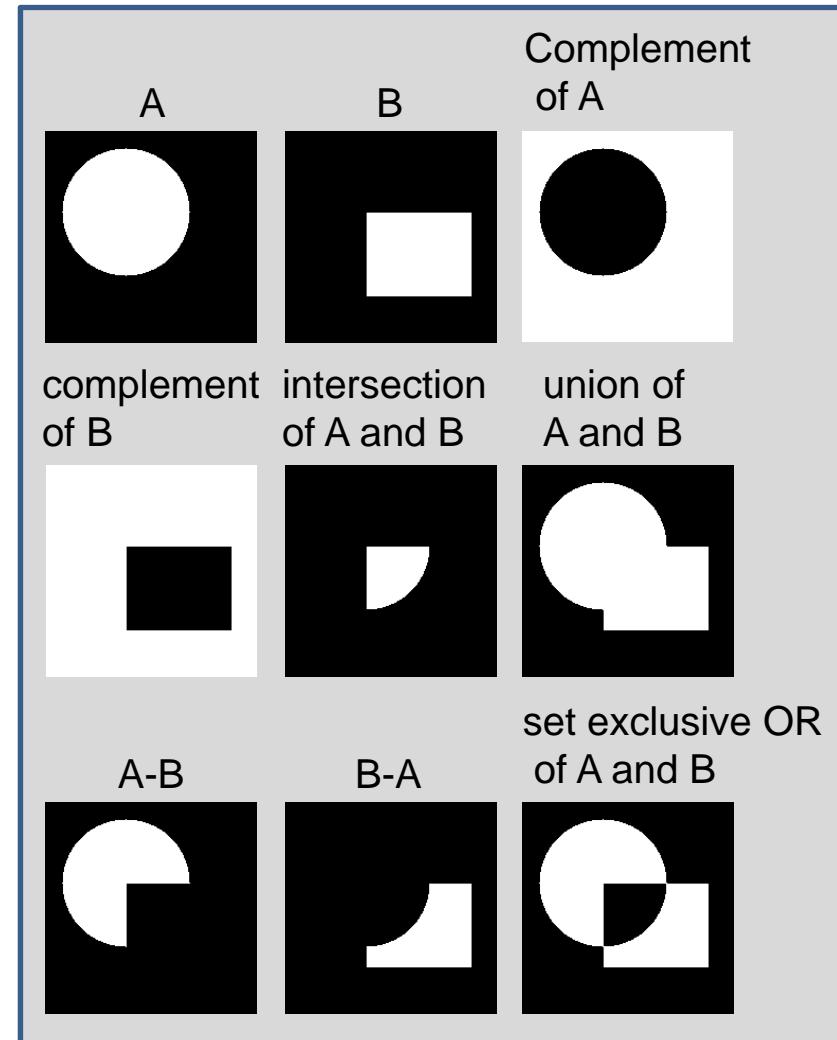
形态学主要是针对二值图像
三种最基本的逻辑运算：与、或、非（补）

p	q	$p \text{ AND } q$ (also $p \cdot q$)	$p \text{ OR } q$ (also $p + q$)	$\text{NOT } (p)$ (also \bar{p})
0	0	0	0	1
0	1	0	1	1
1	0	0	1	0
1	1	1	1	0

对二值图像，逻辑运算与集合运算等价

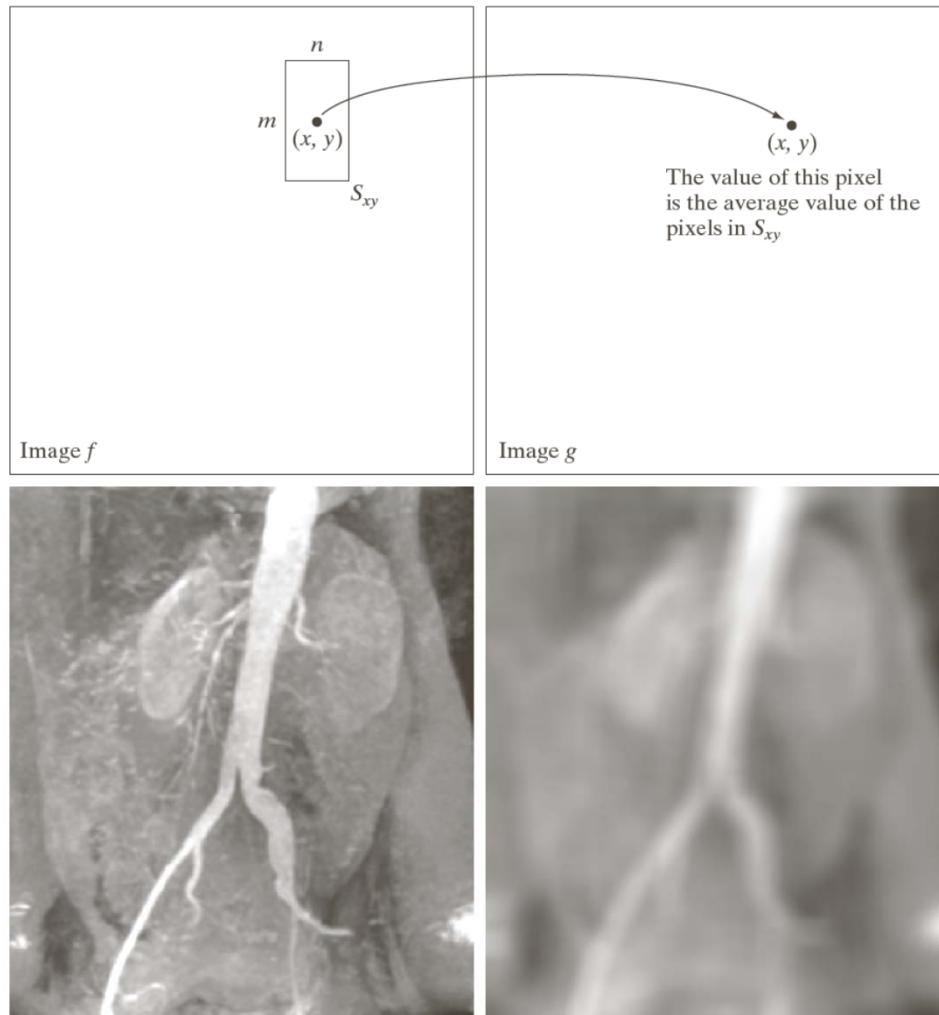
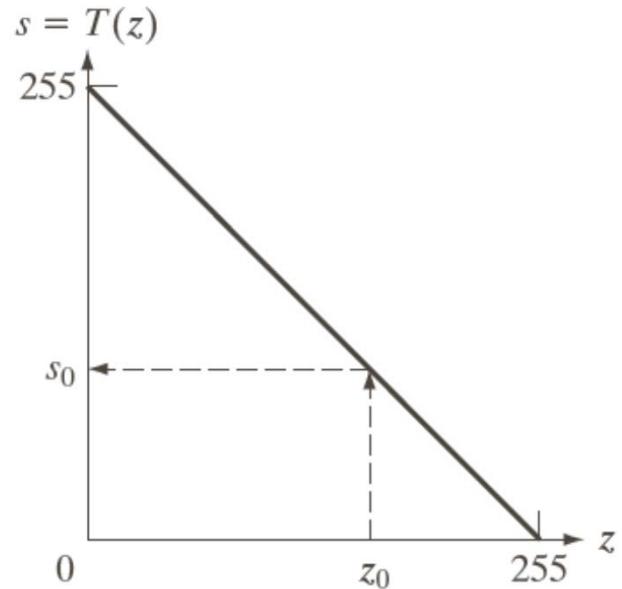


逻辑运算



集合运算

Spatial operation (空域运算)



单像素运算（如反色）

邻域运算（如局部平均）

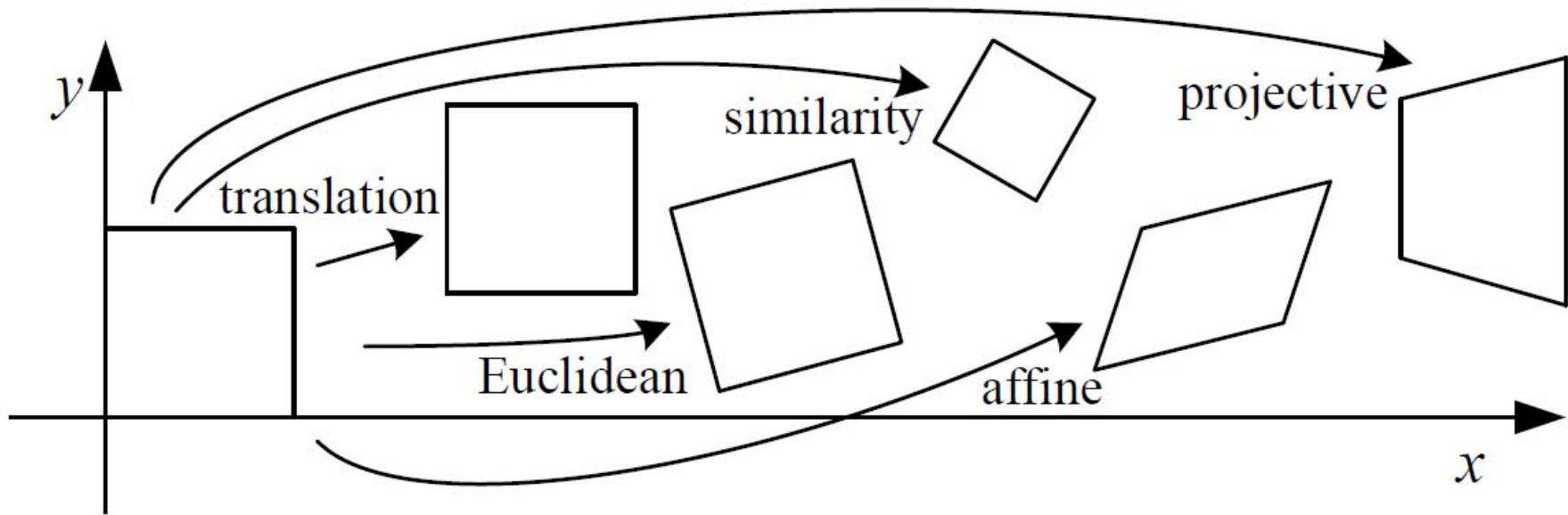
Spatial transformation

- 把照片转正、立体效果
- 矩阵运算，线性变换





Spatial transformations



刚体变换 (rigid transformation)

$$(x, y) = T\{(v, w)\}$$

变化后坐标 原始坐标

$$\begin{aligned}[x & \quad y & \quad 1] &= [v & \quad w & \quad 1]T \\ &= [v & \quad w & \quad 1] \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{bmatrix}\end{aligned}$$

Spatial transformations

- 空间变换除计算像素的新坐标外，还需做图像插值
- 变换函数是前向。而在计算变换后的图像时，用后向变换。

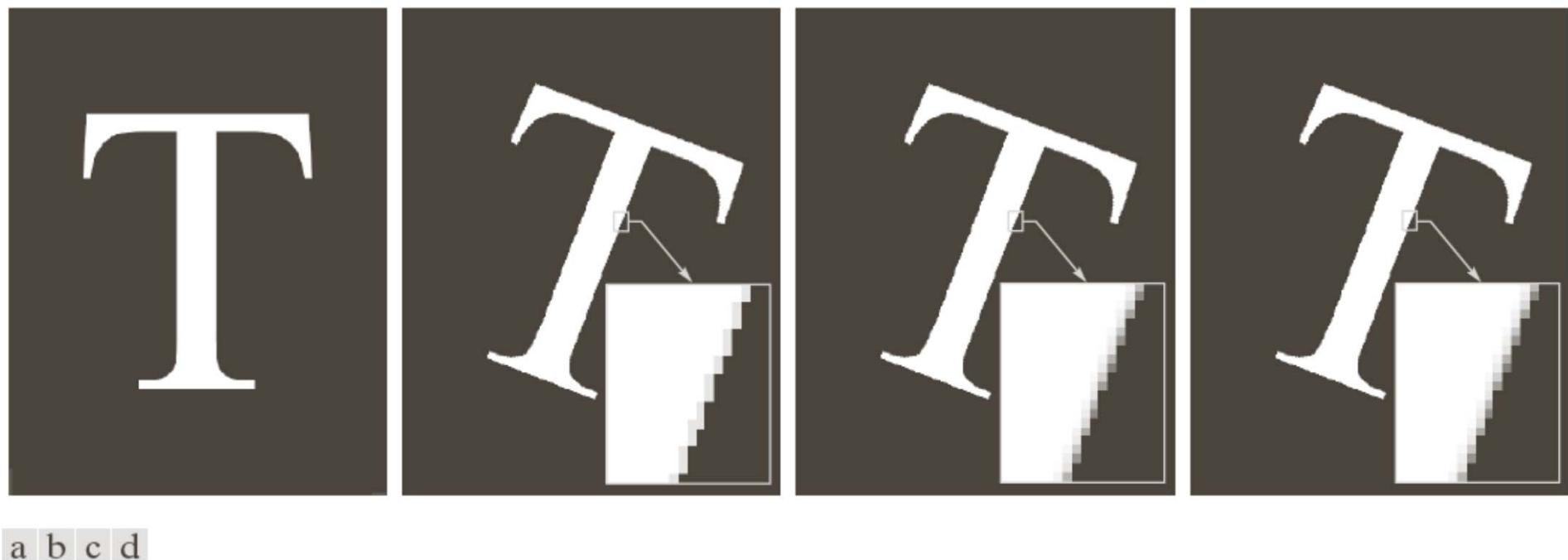
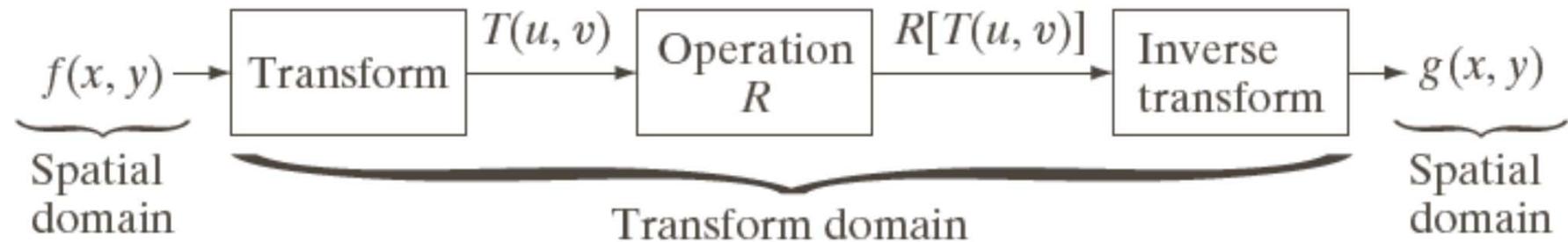


FIGURE 2.36 (a) A 300 dpi image of the letter T. (b) Image rotated 21° clockwise using nearest neighbor interpolation to assign intensity values to the spatially transformed pixels. (c) Image rotated 21° using bilinear interpolation. (d) Image rotated 21° using bicubic interpolation. The enlarged sections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

Image transforms (图像变换)

变换域图像处理很重要



图像的变换和逆变换（通用形式）

$f(x, y)$ 的 前向变换核
正变换 (forward transformation kernel)

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

u, v 为变换域变量

$T(u, v)$ 的
逆变换

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

Image transforms (图像变换)

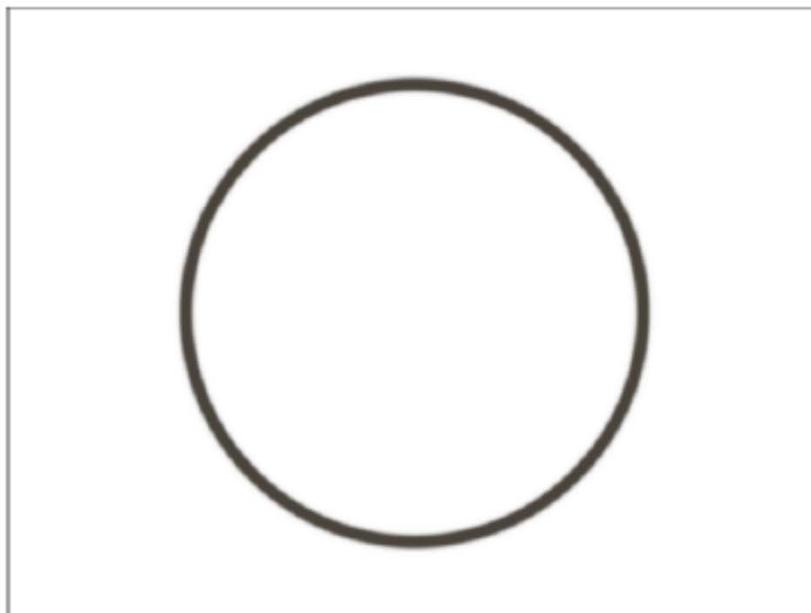
2D离散傅里叶变换 (DFT)

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

具有可分离性（分为 x 和 y 两个变换核）、对称性（两个变换核的形式一样）

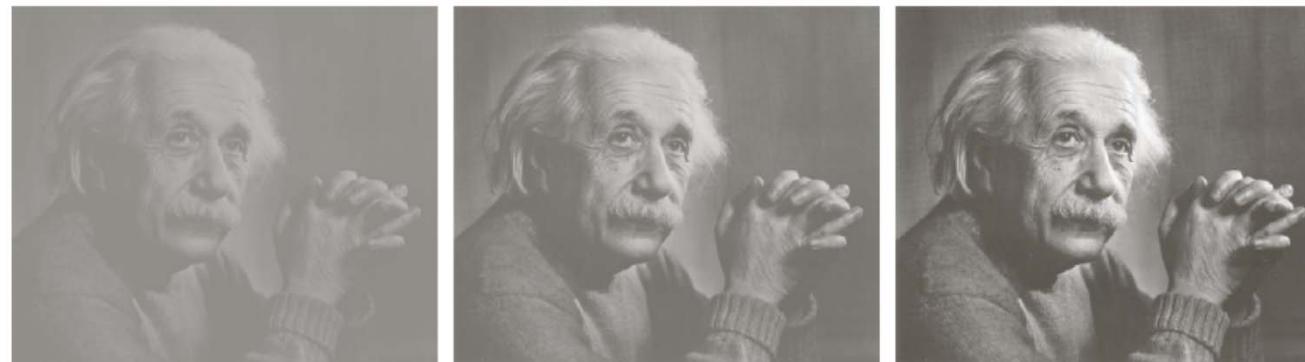
可分离性对于实际应用意味着什么？

2D DFT应用



Probabilistic methods

- 概率方法在图像处理中被广泛应用
- 简单例子：灰度值视为随机变量
- 一幅 $M \times N$ 的图像所有可能的灰度值 $z_i, i = 0, 1, 2, \dots, L - 1$
- $p(z_k) = \frac{n_k}{MN}$, n_k 是灰度值为 z_k 的像素数
- 一阶矩、二阶矩...



标准偏差由低到高

概率方法的应用

- 直方图均衡
- 图像恢复
- 图像分割
- 图像压缩
- 图像识别
- 图像合成
- ...

下次课

- 第3周放假
- 第4周介绍Matlab及图像处理工具箱