

第二章 数字图像处理基础

- 人眼视觉基础
- 图像采样和量化
- 数字图像的基本格式
- 像素间的基本关系
- 纹理
- 图像插值方法



本章概要

2 学时授课

1 学时实验

人眼视觉基础

DIP主要是基于数学和概率等学科的理论体系发展起来的
人的视觉在选择某个DIP技术时扮演了中心的角色
一切DIP处理的效果可能都由人的视觉来衡量

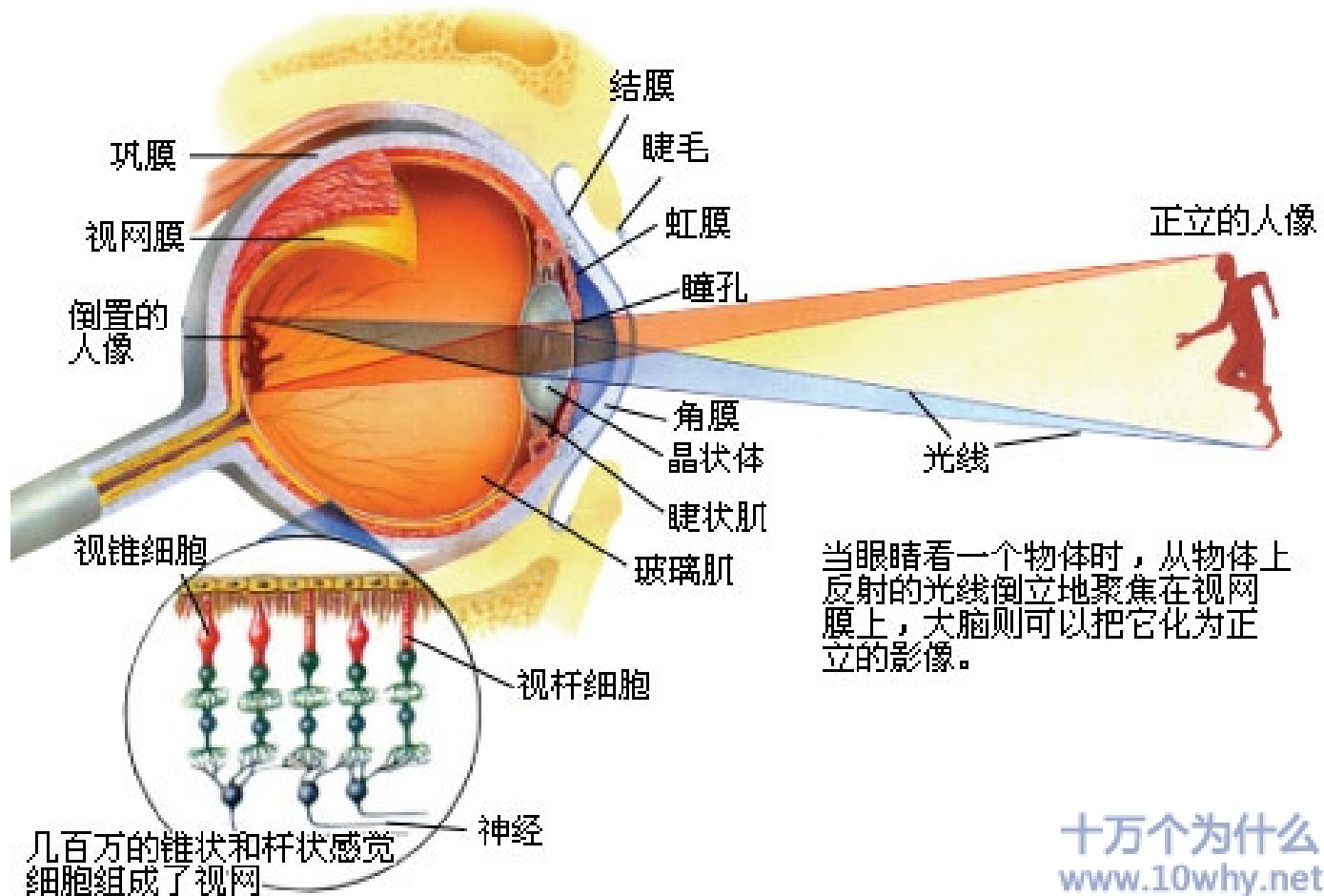


对人视觉的
基本了解有
必要

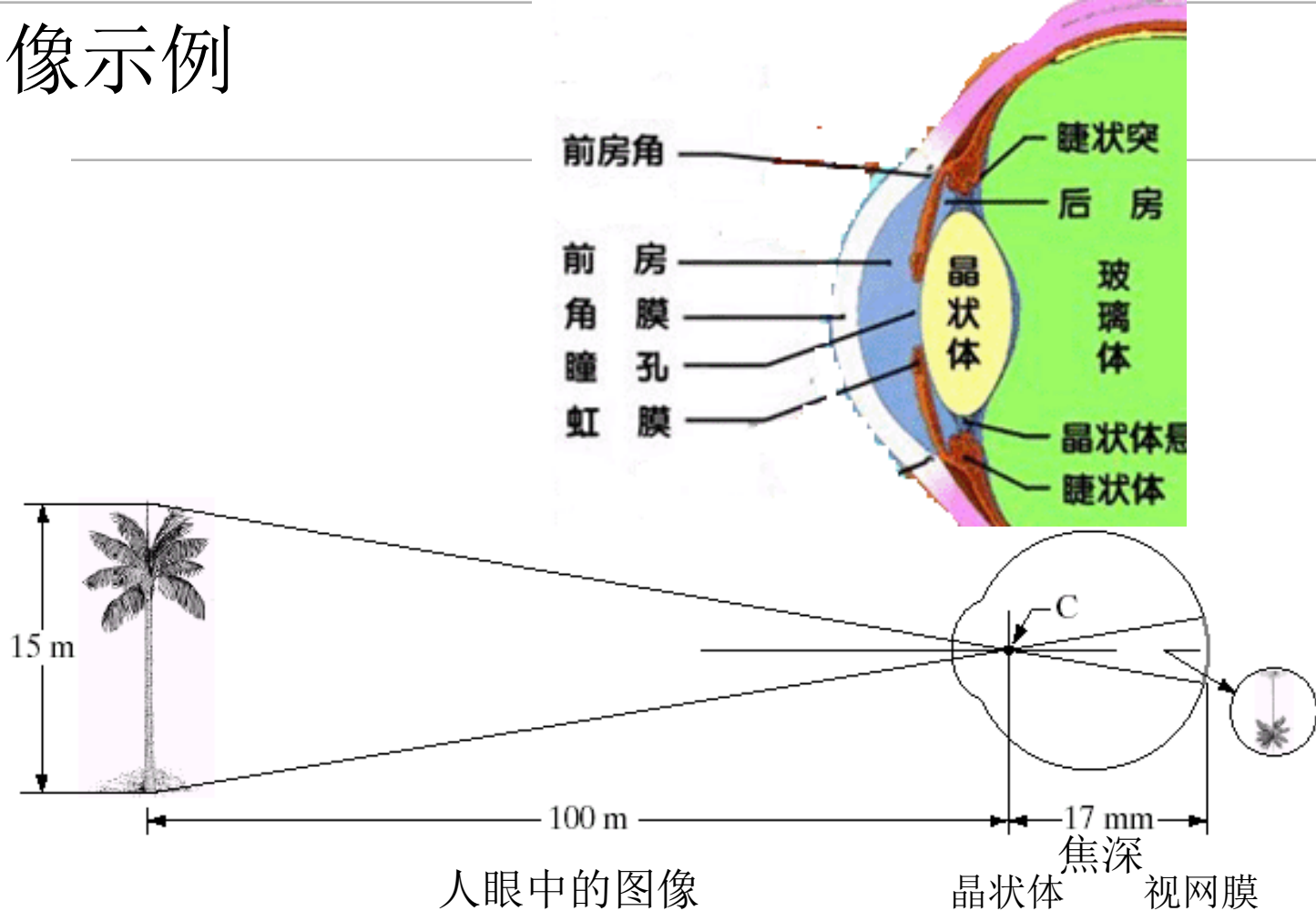
- 晶状体 ~ 适应性强的光学透镜、视网膜~光接受器、大脑解码得到图像
- 人感觉到的亮度 是 进入人眼光强的对数函数
- 人眼的马赫效应：亮度跃变时，感觉亮侧更亮、暗侧更暗
- 人眼感觉到的亮度：依赖于与背景的对比度，呈现正相关
- 人眼的幻觉：眼睛填充了不存在的信息（联想）
- 人眼的错觉：错误地感知物体的几何特点

将在后面具体说明

人眼的主要结构



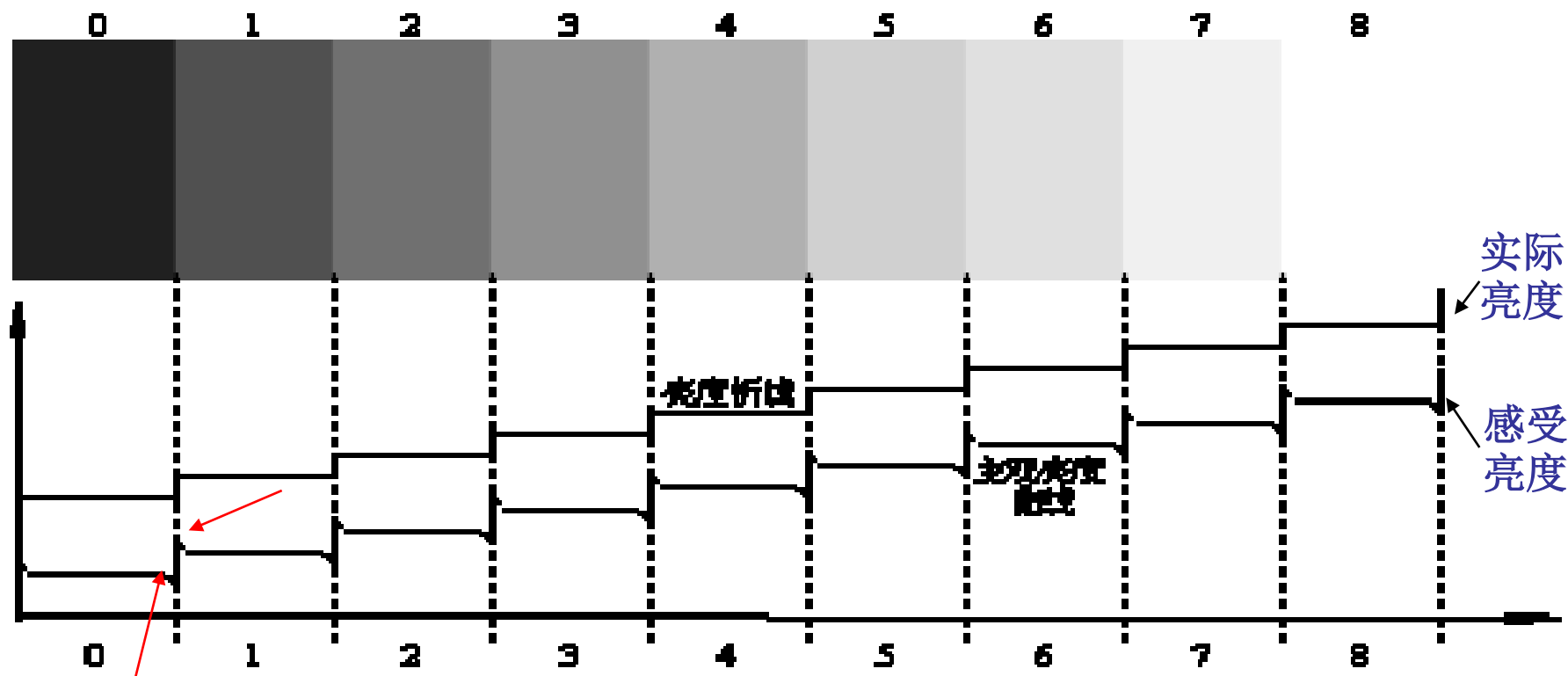
人眼成像示例



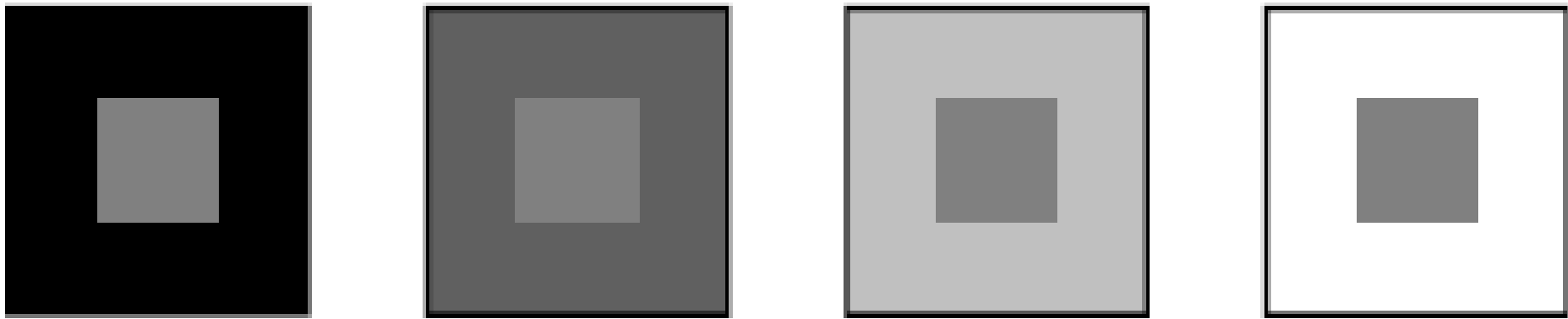
- 焦深：晶状体的中央到视网膜间的距离，大约17mm~14mm，其折射能力最小到最大。
- 虹膜收缩与扩张控制进入眼睛的光量
- 虹膜中间开口处为瞳孔，直径可变范围2~8mm
- 成像大小计算： $15/100=h/17$ ， $h=2.55\text{mm}$ 为物体在视网膜上图像的高。

马赫带效应 (Mach Bands Effect)

Mach效应（1865年由Ernst Mach所发现的现象）表明，人眼所感觉的亮度并不是强度的简单函数，人眼视觉系统有趋向于过高或过低估计不同亮度区域边界值的现象。在亮度跳跃附近，感觉亮的更亮，暗的更暗，如下图。



人眼感觉到的亮度依赖于背景亮度

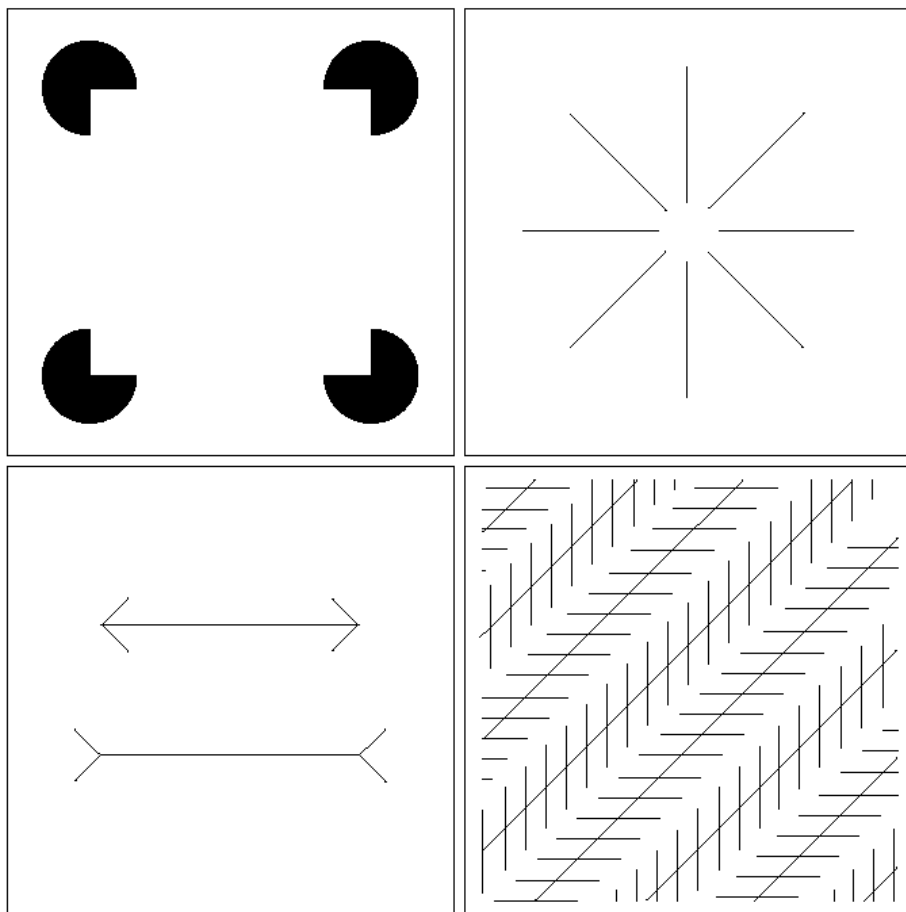


四个中间正方形具有相同的灰度，但人眼的感觉是由亮变暗，原因在于背景的亮度由暗变亮，二者呈有趣的关系：

对比度为正，感觉亮，亮度与对比度呈现正相关性；

对比度为负，感觉暗，亮度与对比度呈现正相关性。

人眼的幻觉与错觉



AB

CD

A光幻觉：内部正方形轮廓是幻觉效果

B光幻觉：线段之间的圆是幻觉效果

C光错觉：两条水平线段长度一样却感觉下面的长

D光错觉：45°斜线间平行等距，但人眼感觉不平行。

光幻觉与错觉是人视觉系统所特有的，迄今还没有清楚的解释。
由于以上各种特殊现象，在进行图像处理时，应该采取一些特殊的补偿措施。

图像采样和量化：模拟到数字

数字化处理的两个方面

□ 空间离散化

将坐标变量 (x,y) 离散化，也即将原来连续的二维图像，离散化为点阵，每个点称为像素(pixel)，这一过程一般称为采样；

□ 幅值离散化

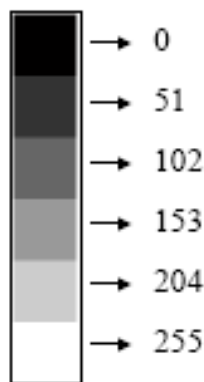
将灰度 f 离散化，将本来可以任意取值的灰度，取为有限的值，这一过程称为灰度级量化。

灰度级

□ 表示某像素位置上亮暗程度的整数称为灰度。

□ 目前，使用的灰度级一般是64(6bit)~256(8bit)，而人眼所能分辨的灰度级大约是40级。
以256阶灰度为例，一般以0表示全黑，而255表示全白。

□ 实际上，人眼并不能分辨256灰阶中相邻的两个灰度。



采样要满足香农采样定理：采样频率如果低于图像的两倍截止频率（欠采样），则会发生混叠现象。

实际情况中，对于非周期函数的采样，采样定律总是难以满足，混叠效应不可避免。

减少混叠效应的一个基本方法是在采样前通过使图像模糊来减少高频分量。

图像空间和灰度分辨率

空间分辨率：一幅图像中可分辨的最小细节，或单位距离内可分辨的最大数目的线对数。

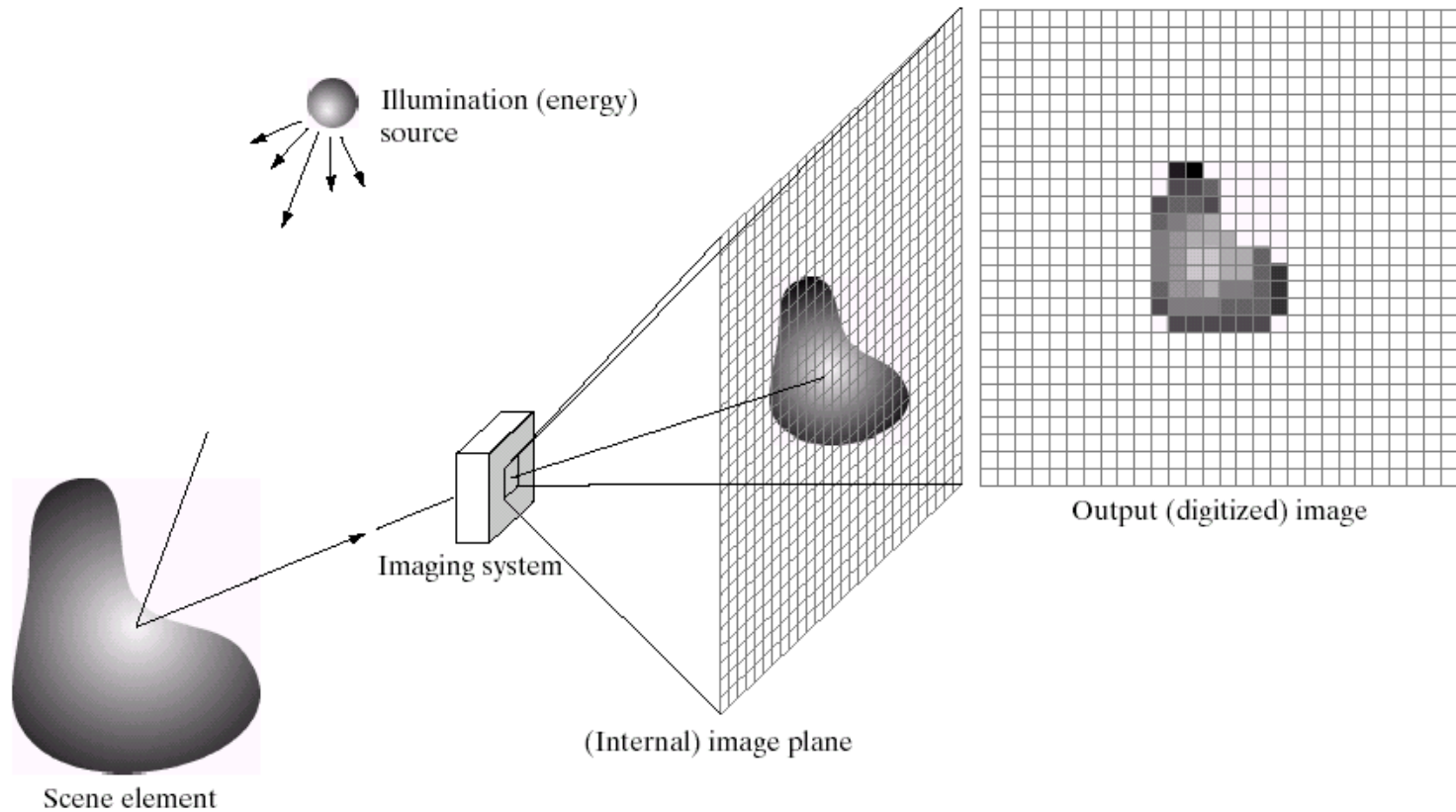
灰度分辨率：可察觉的最小灰度变化（主观的）

图像的空间分辨率主要由采样决定

灰度分辨率主要由量化所决定

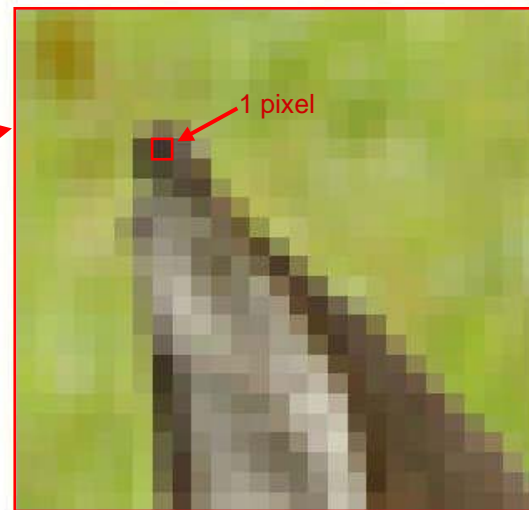
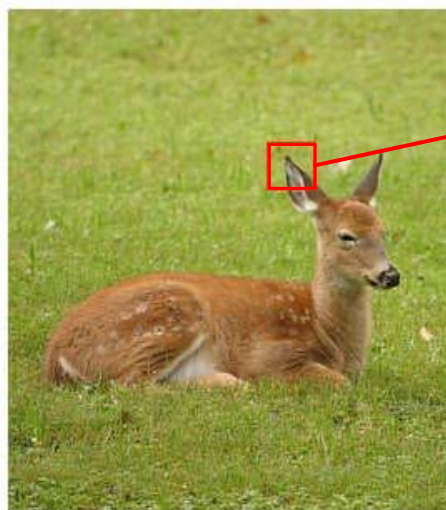
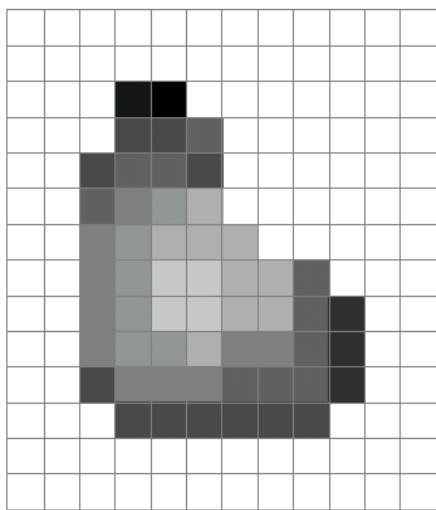
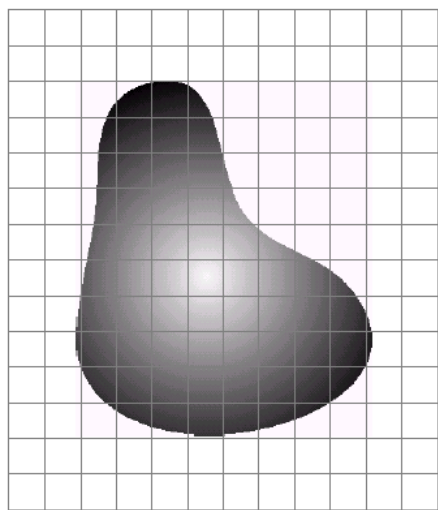
一般都将这些量取作2的整数幂

数字图像获取实例

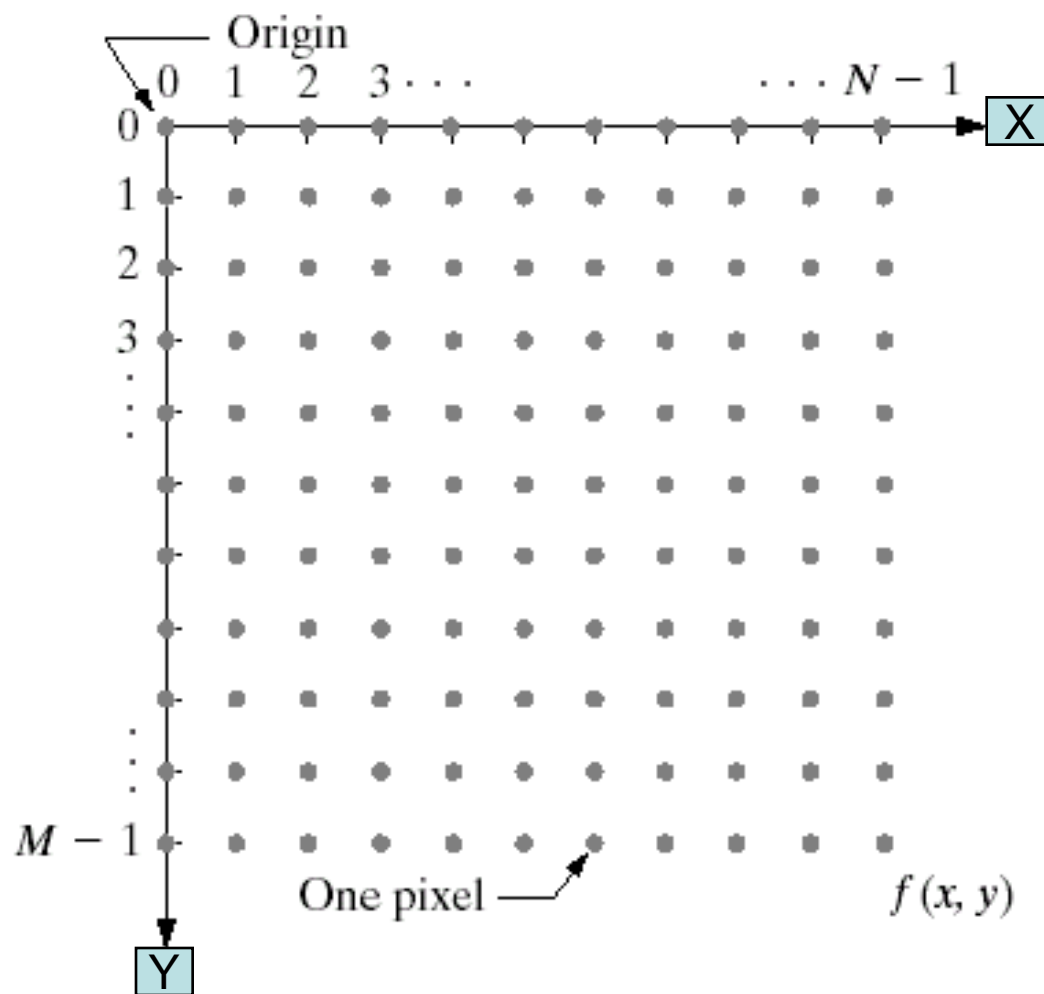


数字图像：对连续图像的 空间量化+幅值量化

- 注意：“digitization”表示数字图像是现实中图像的近似值（空间及灰度的近似）。
- 在图像处理软件中，使用放大工具，对图像的局部进行观察时能更清楚地了解量化



数字图像坐标系



二值图像

- ❑ 二值图像是一种黑白图像，用逻辑值1表示黑色，用0表示白色，则每个像素只需要1bit。

例如，以灰度180为阈值，对彩色图片作二值化：

取值0/1

显示0/255



数字图像的矩阵表示

对连续图像 $f(x, y)$ 按照等间隔采样，得到 $M \times N$ 阵列：

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

如果是灰度图，则 f 为灰度；如果是彩色图像，一般用红(R)、绿(G)、蓝(B)三个矩阵来表示（各8bit，共用24bit）。

数字图像的下采样示例

下图给出1组空间分辨率变化所产生效果的例子，其中图(a)为1幅 512×512 ，256级灰度的图象，其余各图的空间分辨率依次为 256×256 ， 128×128 ， 64×64 ， 32×32 ， 16×16 。



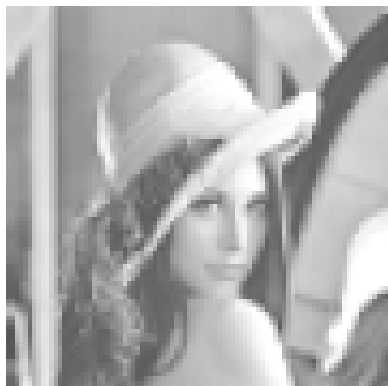
(a)



(b)



(c)



(d)



(e)



(f)

空间细节
丧失

数字图像的灰度级降低示例

下图给出1组幅度分辨率变化所产生效果的例子，其中图(a)为幅 512×512 ，256级灰度图。其余各图依次为保持空间分辨率不变将灰度级数逐次减小为64，16，8，4，2所得到的结果。



(a)



(b)



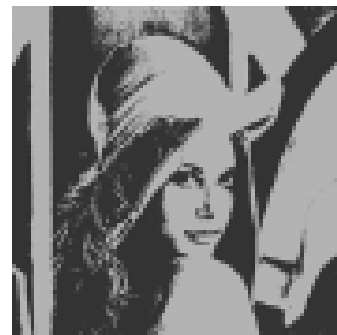
(c)



(d)



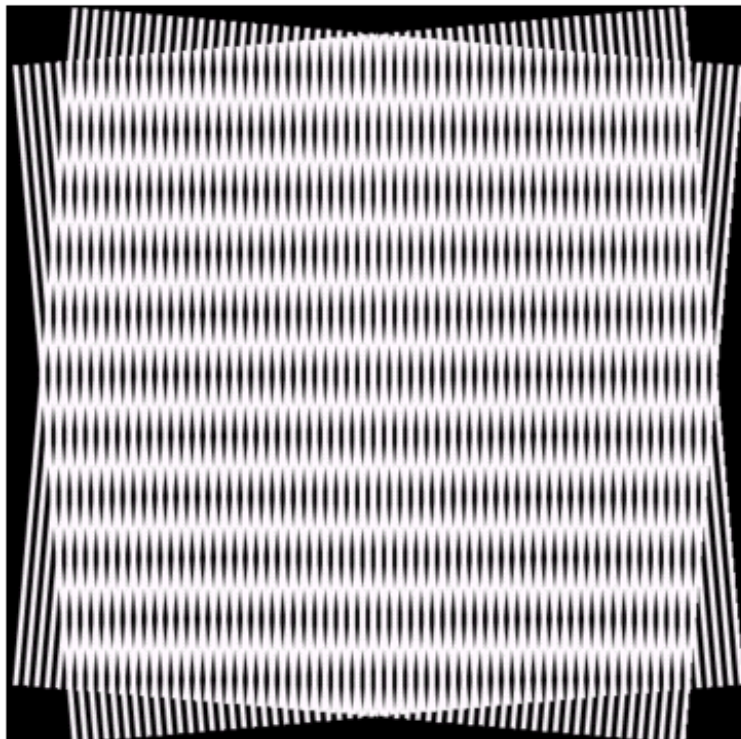
(e)



(f)

灰度细节
丧失

莫尔效应 (Moire Effect)



莫尔效应指的是：两组线阵条纹相互重叠产生的效果,新线段不同于原来的线段

当注视一组线或点与另一组线或点的叠层时的一种视觉效果。

莫尔效应能产生有趣和美丽的几何图样；降低了图画的质量，例如计算机显示的图像（有点阵）被通过照相的方式复制，然后以点阵格式还原时也会产生这种情况。

图像文件存储格式

- ❑ 图像数据文件的格式基本上有两种形式，一种是矢量形式，另一种是光栅形式。在矢量形式中，图像是用一系列线段或线段的组合体来表示的(Flash)。矢量文件主要用于图形数据文件。图像数据文件主要使用光栅形式，该形式与人对图像的理解一致（一幅图像是许多图像点的集合），比较适合色彩、阴影或形状变化复杂的图像。

- ❑ 四种应用比较广泛的格式为：

BMP（BitMaP）格式

GIF（Graphics Interchange Format）格式

TIFF（Tagged Image Format File）格式

JPEG（Joint Photographic Experts Group）格式

格式不同，其对图像数据的压缩编码方式不同。

图像文件存储格式注解

- **BMP格式**: Windows环境下的标准、点阵位图、多数非压缩但支持压缩、灰度与彩色
- **JPEG** 静止图像压缩、有损编码、变换编码, 对高频、低频信号区别对待、灰度与彩色
- **GIF格式**: 灰度与彩色、压缩 (比例1-3)、可存放多幅图像
- **TIFF格式**: 独立于操作系统和文件系统的格式、可存放多幅图像、可压缩 (多种压缩模式)、灰度与彩色
- 最简单的头的数据格式是灰度图像**PGM (Portable GreyMap)**、彩色图像**PPM (Portable PixMap)**。无压缩、头只包含图像宽度、图像高度、最大灰度值或颜色值。

PGM格式

P5

宽度

高度

255

裸数据(8位)

PPM格式

P6

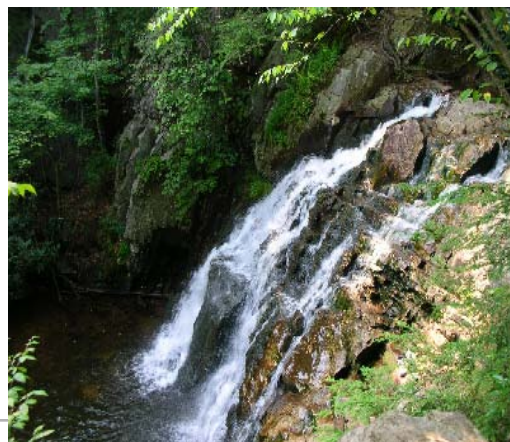
宽度

高度

255 255 255

裸数据(RGB) (3字节)

- 一般的，图像裸数据可以是如下几种情况：
 - 每像素1个字节 (二值或灰度图像)
 - 每像素3个字节 (RGB或其他彩色图像格式)
 - 每像素4个字节(彩色图像+透明度)



图像像素间的基本关系

像素间空间基本关系的描述是基于像素邻域操作的算法的基础

- 像素邻域 (neighbors): 像素 p 的4邻域 $N_4(p)$ 与8邻域 $N_8(p)$
- 像素间的连通性: 空间相邻, 灰度相似(主要是二值或标号图像)
- 相邻像素的邻接: 4邻接、8邻接、混合邻接
- 两个图像子集相邻: 至少存在一对像素来自这两个子集, 且邻接
- 连通路径: 一串连通的像素序列
- 连通区域: 区域内的任意两像素存在连通路径
- 像素间距离: 常用的有欧式、街区、棋盘

像素邻域

像素邻域：像素 p 的4邻域 $N_4(p)$ 与8邻域 $N_8(p)$

位于坐标 (x, y) 的一个像素 p 有 4 个水平和垂直的相邻像素,其坐标由下式给出:

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

这些像素称 p 的4邻域, 表示为 $N_4(p)$ (*像素).

#	*	#
*	p	*
#	*	#

p 的 4 个对角邻像素有如下坐标:

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

表示为 $N_D(p)$ (#像素). 并与 $N_4(p)$ 一起称为 p 的8邻域, 表示为 $N_8(p)$ (*#像素) .

三维图像中, 邻域有**6邻域**($d=1$)、**18邻域**($d \leq 2$)、**26邻域**($d \leq 3$)

$$d(x, y, z; x', y', z') = |x - x'| + |y - y'| + |z - z'|, |x - x'| \leq 1, |y - y'| \leq 1, |z - z'| \leq 1$$

相邻像素间的连通性

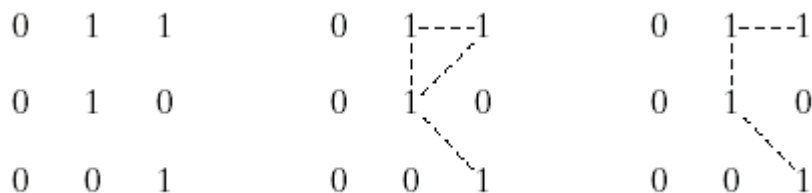
空间连通性分析主要用于二值图像或标号图像（像素具有相同的标号）

像素 $p(x, y)$ 与像素 $q(x', y')$ 之间的连通性定义

- 4邻接: q 在 $N_4(p)$ 中, 即像素 q 是像素 p 的4邻域像素
- 8邻接: q 在 $N_8(p)$ 中, 即像素 q 是像素 p 的8邻域像素
- 混合邻接: q 在 $N_4(p)$ 或 q 在 $N_D(p)$ 中、 $N_4(p)$ 与 $N_4(q)$ 无交集

混合邻接可避免8邻接带来的二义性。

像素 p 与 q 是4邻接, 或在它们间无4邻接像素的情况下是8邻接



左: 9像素; 中: 与中心像素的8邻接像素; 右: 与中心像素的混合邻接像素

像素间的距离

像素 $p(x, y)$ 与像素 $q(x', y')$ 间的几种常见距离：满足距离的三要素

欧式距离

$$D_E(p, q) = \sqrt{(x - x')^2 + (y - y')^2}$$

街区距离

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

棋盘距离

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

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

中心像素的棋盘距离

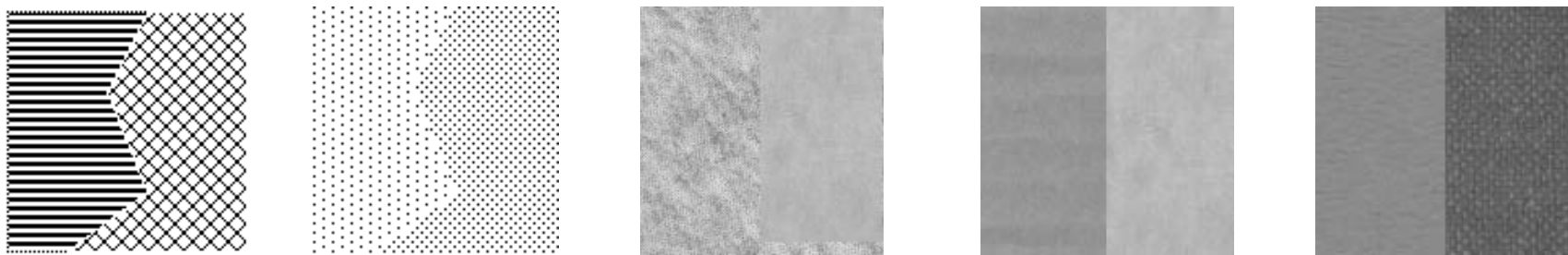
			2	
		2	1	2
2	1	0	1	2
		2	1	2
			2	

中心像素的街区距离

图像纹理Texture

图像纹理是一种普遍存在的视觉现象，尚无统一认可的定义

- 定义1：按照一定规则对基元进行排列所形成的重复模式,该基元称为纹理素(Texels)。
- 定义2：如果图像函数的一组局部属性是恒定、缓变或近似周期性的，则图像中的对应区域具有恒定的纹理。
- 纹理是区域属性，当区域小于纹理素时，就不存在纹理
- 最经典、最常用的纹理描述方法是基于灰度共生矩阵GLCM的描述子



灰度共生矩阵

灰度共生矩阵(grey level co-occurrence matrix, GLCM)

初始像素灰度为 i , 位移
 Δx , Δy 后的像素灰度为 j .

$$P(i, j, \Delta x, \Delta y) = \# \{ [f(x_1, y_1) = i, f(x_1 + \Delta x, y_1 + \Delta y) = j] \} / \#S$$

S 为所有满足空间间隔为 $(\Delta x, \Delta y)$ 的像素对的集合, $\#$ 表示集合的元素个数

实际应用中, 常固定间隔的距离, 而改变角度

常见的角度为 0° 、 45° 、 90° 、 135°

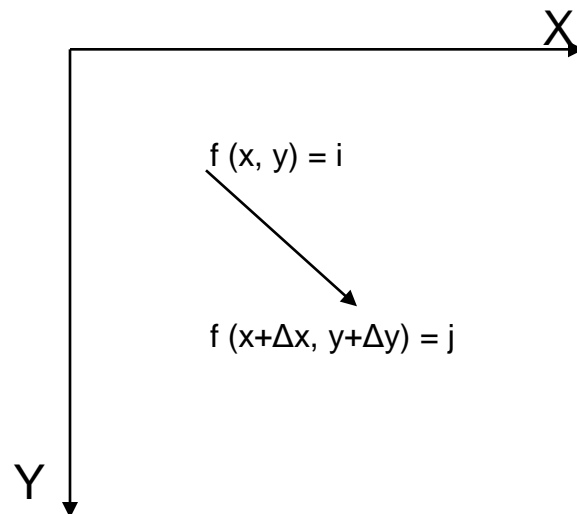
给定 $(\Delta x, \Delta y)$, 常计算

能量或角度二阶矩、熵、对比度、逆差矩、相关

若要得到具有旋转不变性的特征, 简单的方法是对同一特征在四个角度下的量求平均值、与标准差

通常为减少计算量, 先把256灰阶转化为16级灰阶

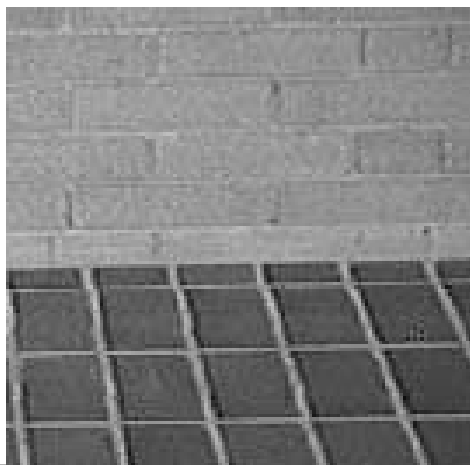
还可通过计算直方图间接计算GLCM以加速(见Parker p 159)



图像灰度插值方法

图像灰度插值要解决的是：非整数点的灰度确定问题

- 最近邻插值(nearest neighbor interpolation): 最简单, 误差最大
- 双线性插值(bilinear interpolation): 有低通滤波性质
- 三次插值: 多项式、B样条等, 精度高但计算量大
- 作业1: 读图(Chapter2_1.pgm)、顺时针方向旋转 15° (利用上述几种插值方法)、输出图像 (测试图像见下图), 可用C++或Matlab编程



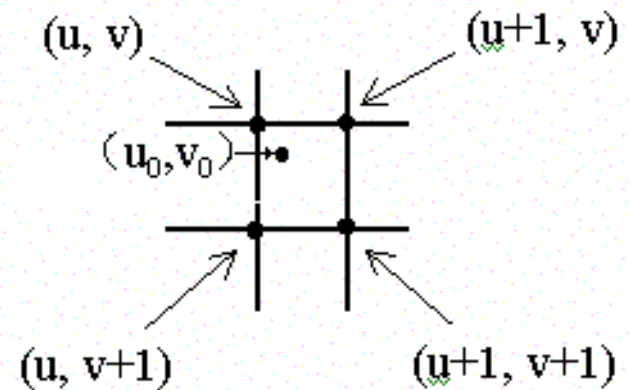
对结果进行讨论:用量化指标支持你的讨论

灰度最近邻插值

(Zero-order or Nearest-neighbor Interpolation)

用四个相邻格点中与 (u_0, v_0) 点最近的点的灰度值作为该点灰度值。假设，图中整数坐标 (u, v) 点与 (u_0, v_0) 点距离最近，则有

$$f(u_0, v_0) = f(u, v)$$



特点：只用到距离及一个点的灰度值，简单、快速。

但当像素间灰度差值大时，该方法的误差也较大。

双线性插值法 (Bilinear Interpolation)

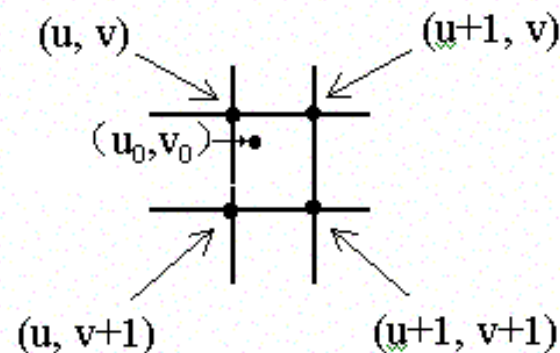
用 $[S]$ 表示不超过 S 最大整数,

则 $u = [u_0]$ $v = [v_0]$

$$\alpha = u_0 - [u_0] \quad \beta = v_0 - [v_0]$$

根据 (u_0, v_0) 4个邻点灰值, 插值计算 $f(u_0, v_0)$,

$$F(u_0, v_0) = (1-\alpha)(1-\beta)f(u, v) + \alpha(1-\beta)f(u+1, v) \\ + (1-\alpha)\beta f(u, v+1) + \alpha\beta f(u+1, v+1)$$



相邻4像素的线性组合, 权重反比于在X 及Y方向上的坐标差的乘积

特点: 一般能够得到满意结果

但有低通滤波性质, 使图像的高频分量受损失。

三次多项式插值 (Cubic Polynomial Interpolation)

基于 $c(x) = \sin(\pi x)/(\pi x)$ 的三次多项式逼近

$$u = [u_0] \quad v = [v_0]$$

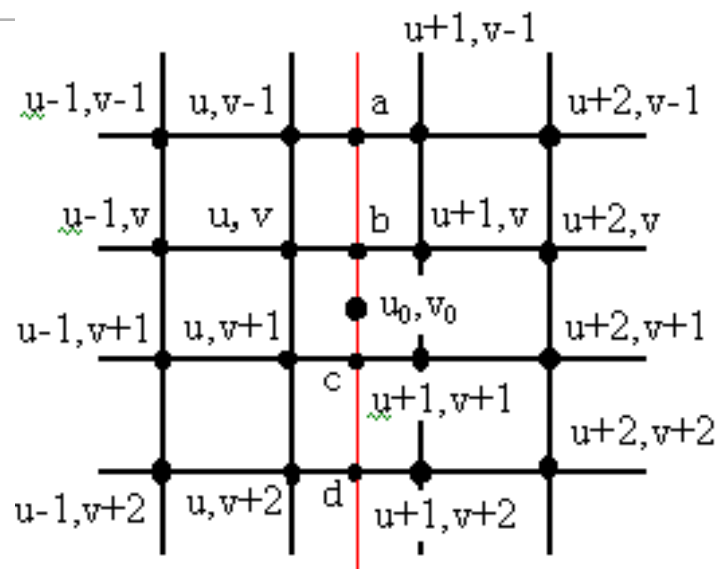
$$\alpha = u_0 - [u_0] \quad \beta = v_0 - [v_0]$$

$$f(u_0, v_0) = A * B * C$$

$$A = [c(1+\alpha), c(\alpha), c(1-\alpha), c(2-\alpha)]$$

$$C = [c(1+\beta), c(\beta), c(1-\beta), c(2-\beta)]^T$$

$$B = \begin{bmatrix} f(u-1, v-1) & f(u-1, v) & f(u-1, v+1) & f(u-1, v+2) \\ f(u, v-1) & f(u, v) & f(u, v+1) & f(u, v+2) \\ f(u+1, v-1) & f(u+1, v) & f(u+1, v+1) & f(u+1, v+2) \\ f(u+2, v-1) & f(u+2, v) & f(u+2, v+1) & f(u+2, v+2) \end{bmatrix}$$



要用到邻近的16个像素的灰度值(见罗述谦, 周果宏 编著 医学图像处理与分析 第二版 p34)
插值精度高, 但计算量大

还可使用三阶B样条插值! (感兴趣的同学可推导, 提示: 借助于非线性配准文献)