



第二章：图像数字化

中国科学技术大学
电子工程与信息科学系

主讲教师：李厚强 (lihq@ustc.edu.cn)
周文罡 (zhwg@ustc.edu.cn)

助教：王 宁 (wn6149@mail.ustc.edu.cn)
邓家俊 (dengjj@mail.ustc.edu.cn)
陈 铮 (czczcz@mail.ustc.edu.cn)
胡鹤臻 (alexhu@mail.ustc.edu.cn)



图像数字化

□ 数字图像基础

- 成像过程
- 像素间联系
- 数字化
- 图像变换

□ 形态学

- 二值形态学
- 灰度形态学



图像数字化

□ 数字图像基础

■ 成像过程

- ✓ 视觉过程
- ✓ 成像变换
- ✓ 成像亮度
- ✓ 视觉系统
- ✓ 采样与量化

■ 像素间联系

■ 数字化

■ 图像变换

□ 形态学

视觉过程

□ 光学过程

- 视网膜成像为倒影
- 晶状体可以改变形状，调节聚焦焦距（范围：14~17mm），远离变扁，靠近加厚。

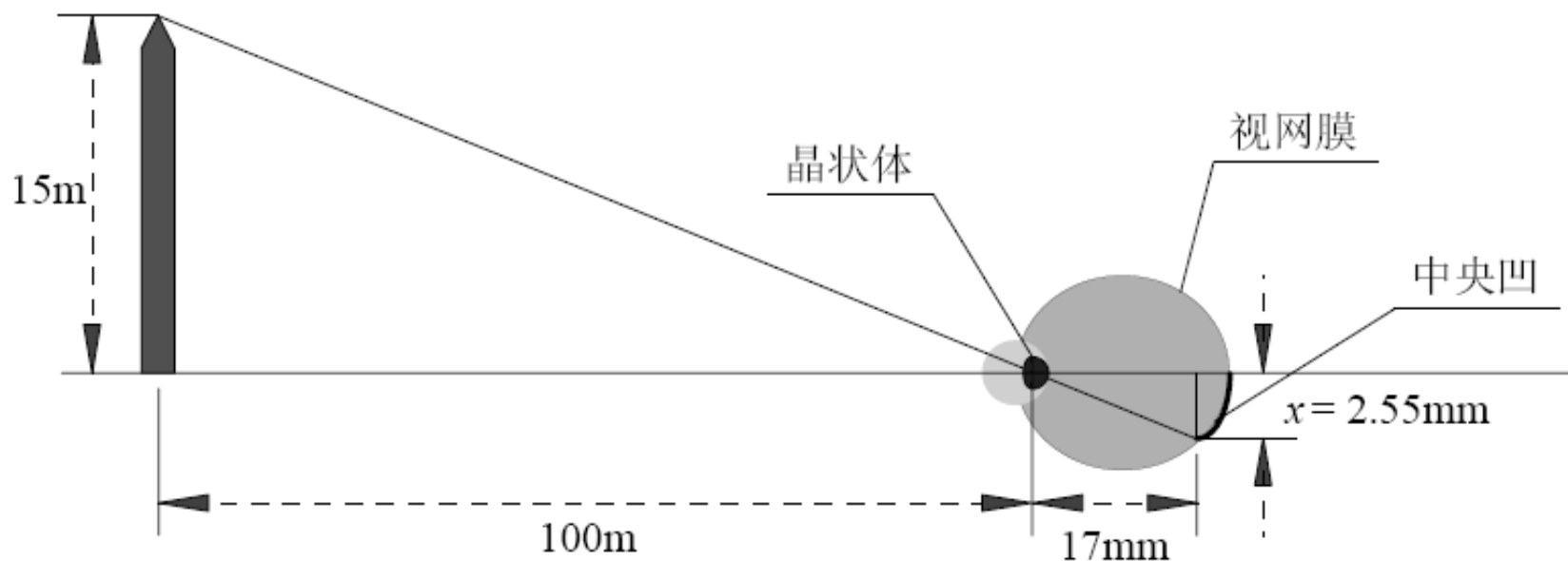


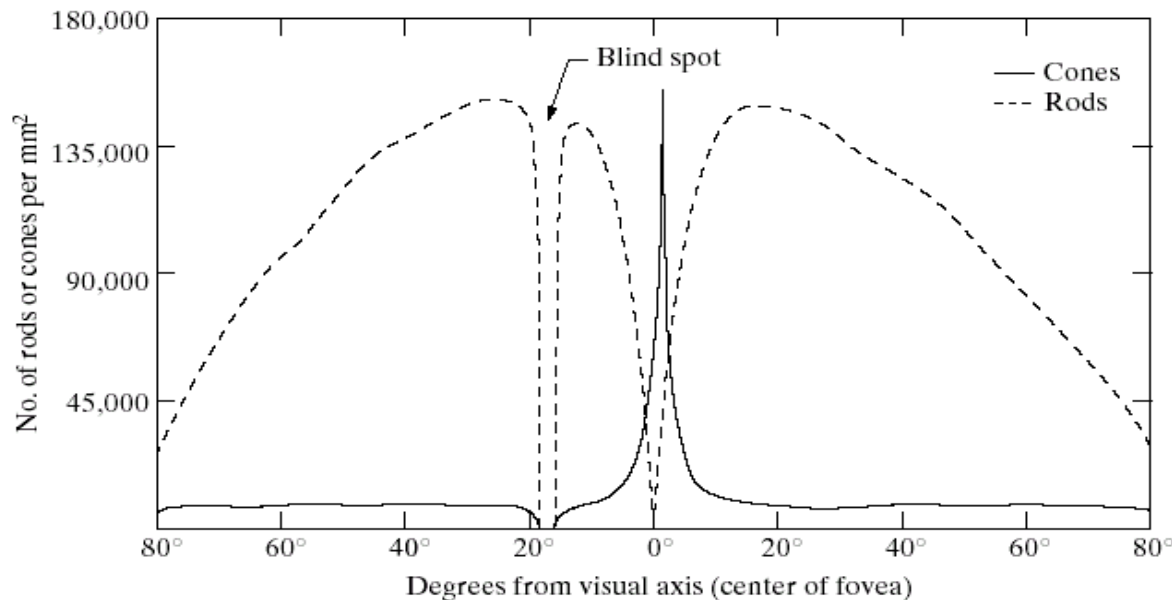
图 2.1.1 人眼水平横截面示意图

$$15 / 100 = 2.55 / 17$$

视觉过程

□ 化学过程

- 视网膜上布满两类光感受器：锥细胞和柱细胞
 - ✓ 锥细胞(cone): 数量少 (600~700万), 对颜色高度敏感
 - ✓ 锥细胞视觉: 明视觉或亮光视觉(photopic vision)
 - ✓ 柱细胞(rod): 数量多 (7500~15000), 分辨率比较低不感受颜色, 对低照度较敏感
 - ✓ 柱细胞视觉: 暗视觉或微光视觉(scotopic vision)

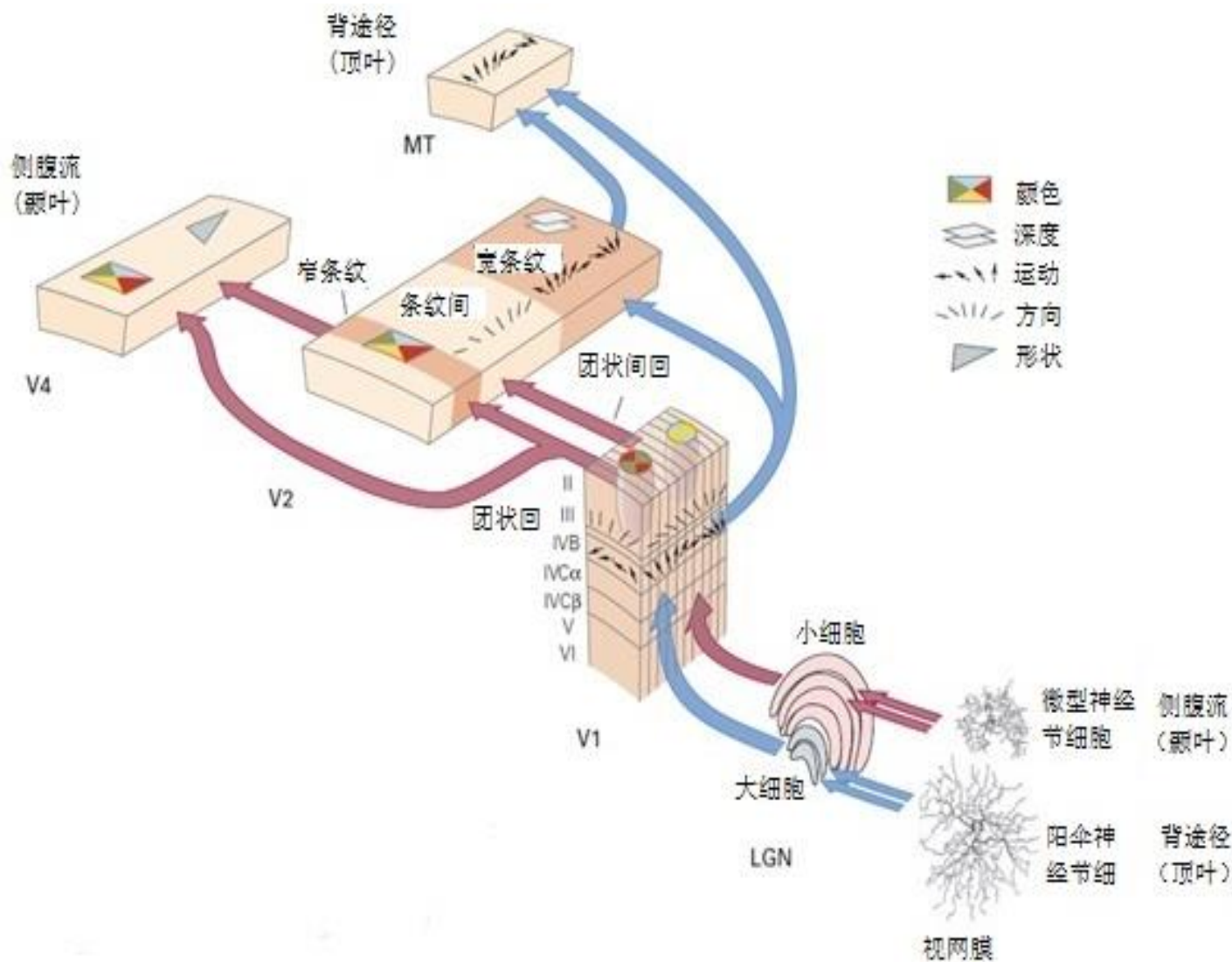


视觉过程

□ 神经处理过程

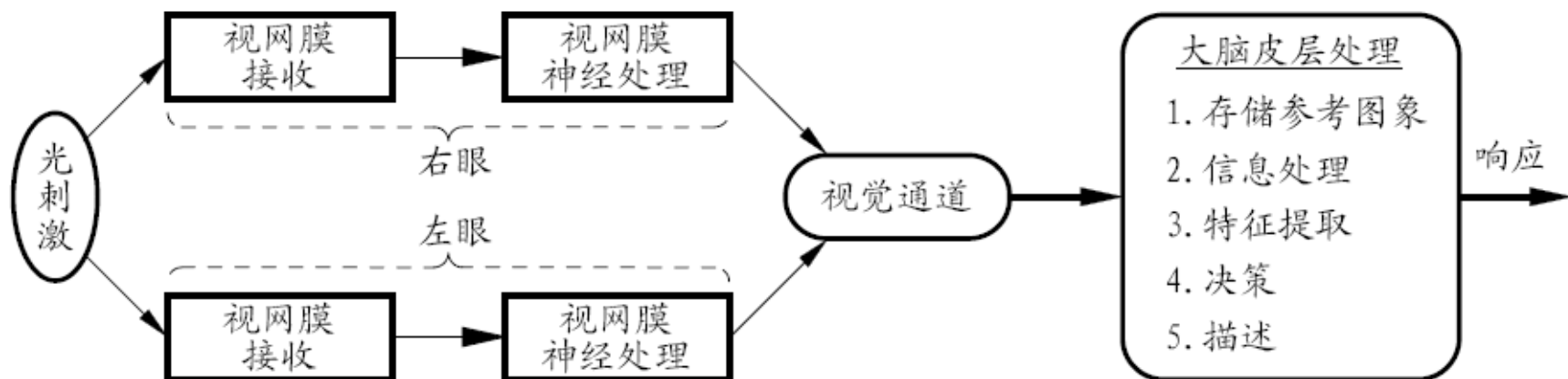
- 每个视网膜接收单元都与一个神经元细胞借助**突触**(synapse)相连
- 每个神经元细胞借助其它的**突触**与其它细胞连接, 从而构成光神经(optical nerve)网络
- 光神经进一步与大脑中的侧区域(side region of the brain)连接, 并到达大脑中的纹状皮层(striated cortex)
- 对光刺激产生的响应经过一系列处理最终形成关于场景的表像, 从而将对光的感觉转化为对景物的知觉

视觉过程：脑视觉信息处理通路



视觉过程

□ 整体视觉过程



视觉 = “视” + “觉”



成像变换

□ 投影成像

- 将3-D客观场景投影到2-D图像平面

□ 成像过程

- 三个坐标系统
 - ✓ 世界坐标系 XYZ
 - ✓ 摄像机坐标系 xyz
 - ✓ 图像平面 xy

从 XYZ 到 xyz , 从 xyz 到 xy

成像变换

□ 成像过程

■ 三维坐标系

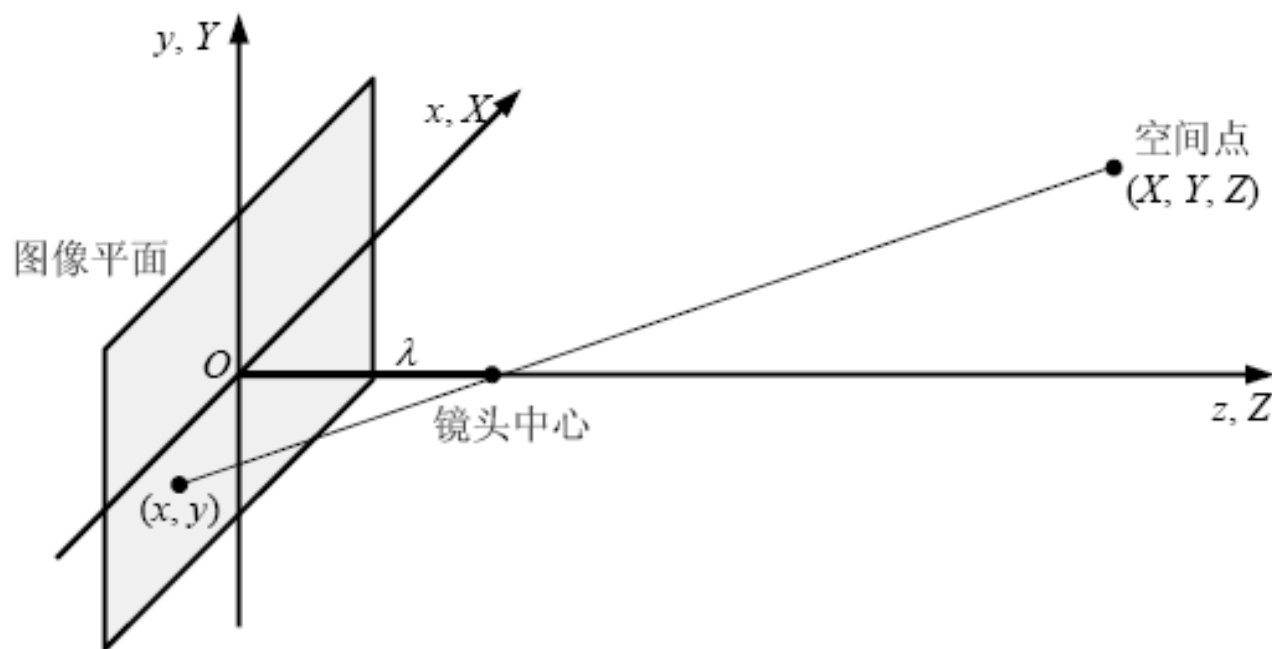


图 2.2.1 投影变换成像示意图

3-D点投影后的
图像平面坐标

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

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

非线性投影等式
(分母含变量Z)

成像变换

□ 齐次坐标

- 可用来将前述非线性（分母中含变量 Z ）等式表示成线性矩阵形式

- ✓ 笛卡尔坐标: $\mathbf{w} = [X \quad Y \quad Z]^T$

- ✓ 齐次坐标: $\mathbf{w}_h = [kX \quad kY \quad kZ \quad k]^T$ (k 为任意非零常数)

齐次坐标 \Rightarrow 笛卡尔坐标: 用第4个坐标量去除前3个坐标量

成像变换

□ 齐次坐标透视变换

$$\mathbf{c}_h = \mathbf{P}\mathbf{w}_h = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/\lambda & 1 \end{bmatrix} \begin{bmatrix} kX \\ kY \\ kZ \\ k \end{bmatrix} = \begin{bmatrix} kX \\ kY \\ kZ \\ -kZ/\lambda + k \end{bmatrix}$$

齐次形式的
摄像机坐标

用第4项分别去除前3个项

$$\mathbf{c} = \begin{bmatrix} x & y & z \end{bmatrix}^T = \begin{bmatrix} \frac{\lambda X}{\lambda - Z} & \frac{\lambda Y}{\lambda - Z} & \frac{\lambda Z}{\lambda - Z} \end{bmatrix}^T$$

成像变换

□ 逆投影变换

- 根据2-D图像坐标来确定3-D客观景物的坐标

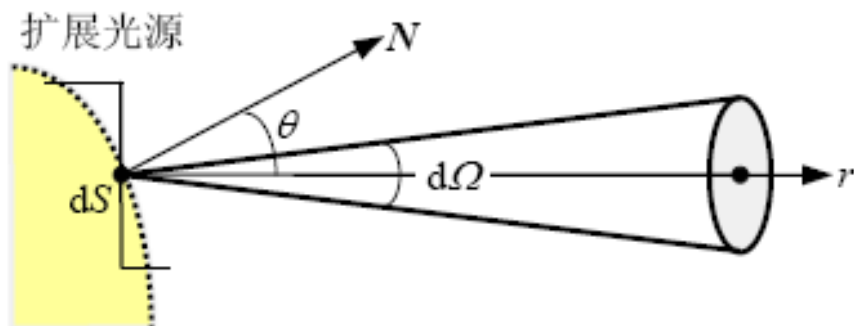
$$\mathbf{w}_h = \mathbf{P}^{-1} \mathbf{c}_h \quad \mathbf{P}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/\lambda & 1 \end{bmatrix}$$

- 从3D客观场景映射到图像平面的投影变换是一个一到多的变换
- 每一个2D图像平面上的坐标点对应过空间点 (X, Y, Z) 的一条直线

成像亮度

□ 景物所受到的光照度

- 光度学：研究光辐射强弱的学科
- 光通量：表示光辐射的功率或光辐射量
 - ✓ 单位：lm（流明）
- 发光强度：单位立体角内发出的光通量
 - ✓ 单位：cd（坎[德拉]）
 - ✓ $1 \text{ cd} = 1 \text{ lm/sr}$
- 亮度：单位投影面积的发光强度



$$B \equiv \frac{dI}{dS'} \equiv \frac{dI}{dS \cos \theta} \equiv \frac{d\Phi}{d\Omega dS \cos \theta}$$

图 2.2.2 扩展光源

成像亮度

□ 图像成像模型

■ 2-D亮度函数: $f(x, y)$

- ✓ 亮度是能量的量度, 一定不为零且为有限值

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

- 入射到可见场景上的光量
- 场景中目标对入射光反射的比率
- 与照度成分 $i(x, y)$ 和反射成分 $r(x, y)$ 相关

成像亮度

□ 景物表面对光的反射程度

- 景物的亮度既与入射光照度有关也与景物表面对光反射的性质有关

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

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

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

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

黑天鹅绒：0.01；

不锈钢：0.65；

白墙平面：0.80；

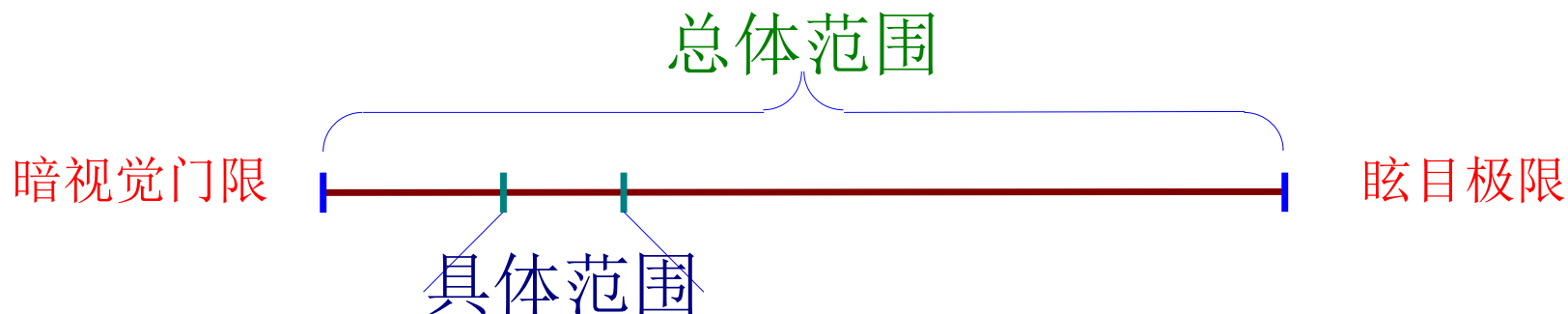
镀银器皿：0.90；

白雪：0.93

视觉系统对光的感知特点

□ 人眼适应的亮度范围

- 总体范围大：从暗视觉门限到眩目极限之间的范围在 10^{10} 量级
- 具体范围小：一般范围在 10^2 量级



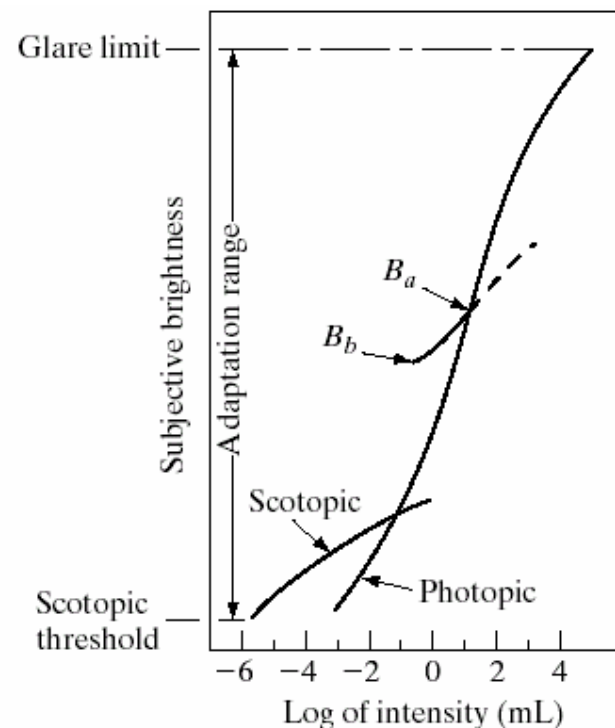
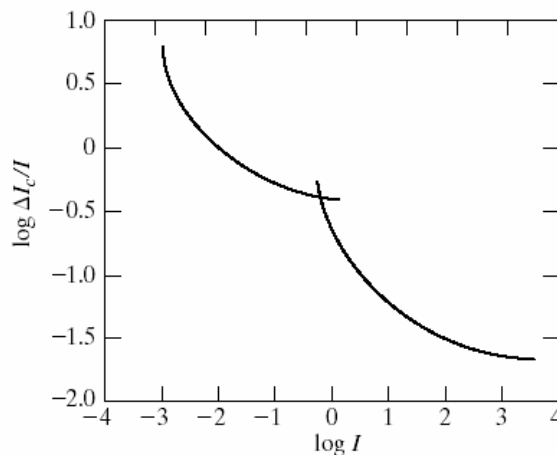
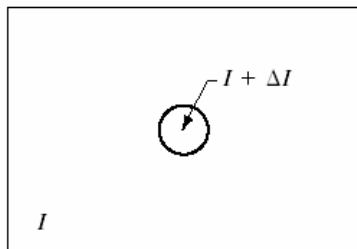
视觉系统对光的感知特点

□ 亮度变化的感知

- 人类视觉系统对**亮度变化**的感知比对亮度本身要敏感
- 人类视觉系统对光强度的响应不是线性的，而是**对数**形式的（对暗光时亮度的增加比对亮光时亮度的增加更敏感）

□ 亮度适应和分辨

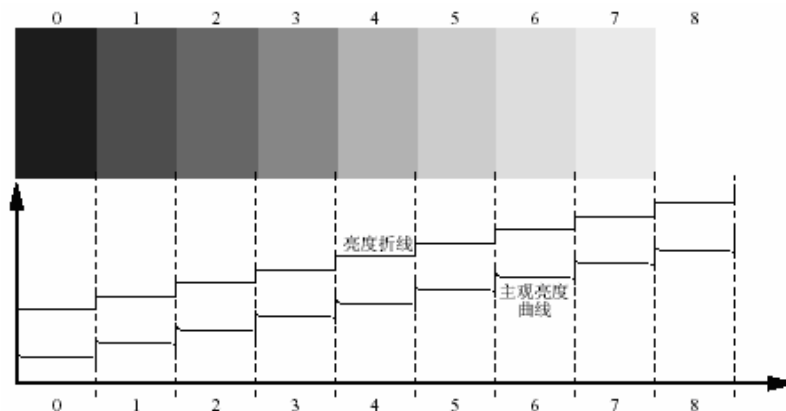
- 客观亮度、主观亮度、亮度适应级
- 亮度分辨率 (韦伯率 Weber ratio)



视觉系统对光的感知特点

□ 亮度错觉

- 马赫带：在亮度变化部分附近的亮区和暗区分别感知到一条更暗和更亮的条带



- 同时对比度

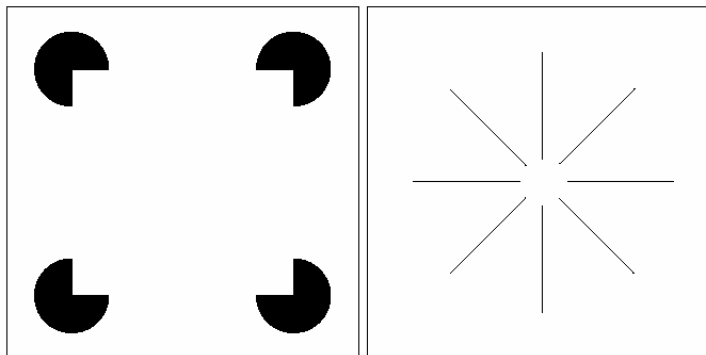


视觉系统对光的感知特点

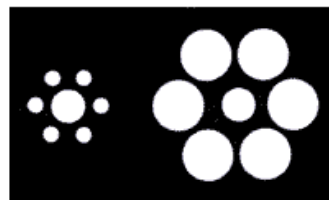
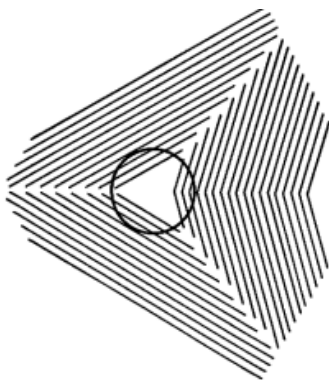
□ 视觉错觉

- 由视觉系统特性特性、心理、知识等方面有关，形成机制尚未完全了解

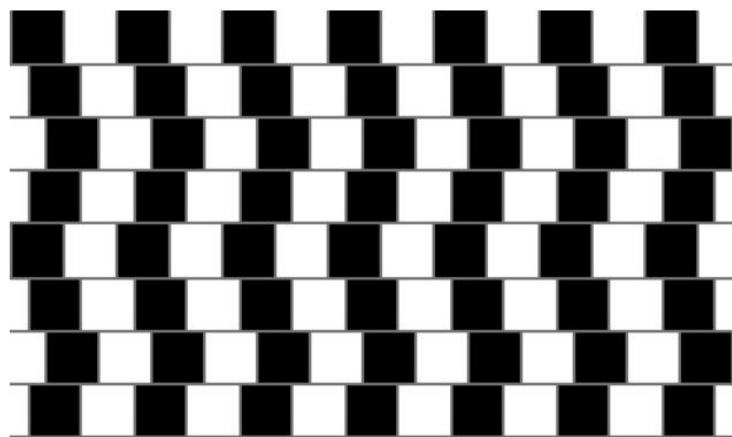
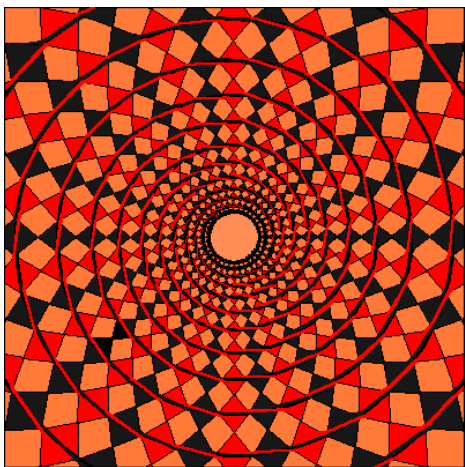
填充



几何特性



视觉系统对光的感知特点



Rubin Vase





采样与量化

□ 图像 \Rightarrow 数字图像

$$f(x, y) \Rightarrow I(r, c)$$

- 空间坐标的离散化叫做空间采样

$$x, y \Rightarrow r, c$$

- 幅度的离散化叫做幅度量化

$$f \Rightarrow I$$

采样与量化

□ 数字图像

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

图像（水平）尺寸 M : $M = 2^m$

图像（垂直）尺寸 N : $N = 2^n$

像素灰度级数 G (k -bit): $G = 2^k$

图像所需的位数 b : $b = M \times N \times k = N^2 k$

采样与量化

□ 数字图像

$$b = M \times N \times k = N^2 k$$

存储1幅 32×32 , 16个灰度级的图需要 4 096 bit

存储1幅 128×128 , 64个灰度级的图需要 98 304 bit

存储1幅 512×512 , 256个灰度级的图需要 2 097 152 bit

□ 常见的分辨率 (P: Progressive, 代表逐行扫描)

- 720P: 1280×720 (HD, 高清)
- 1080P: 1920×1080 (full HD, 全高清)
- 2K: 2560×1440 , 2048×1536 , 2560×1600
- 4K: 4096×2160 , 4096×3112 , 3656×2664 , 3840×2160

采样与量化

□ 图像空间分辨率变化所产生的效果 (515x512~16x16)

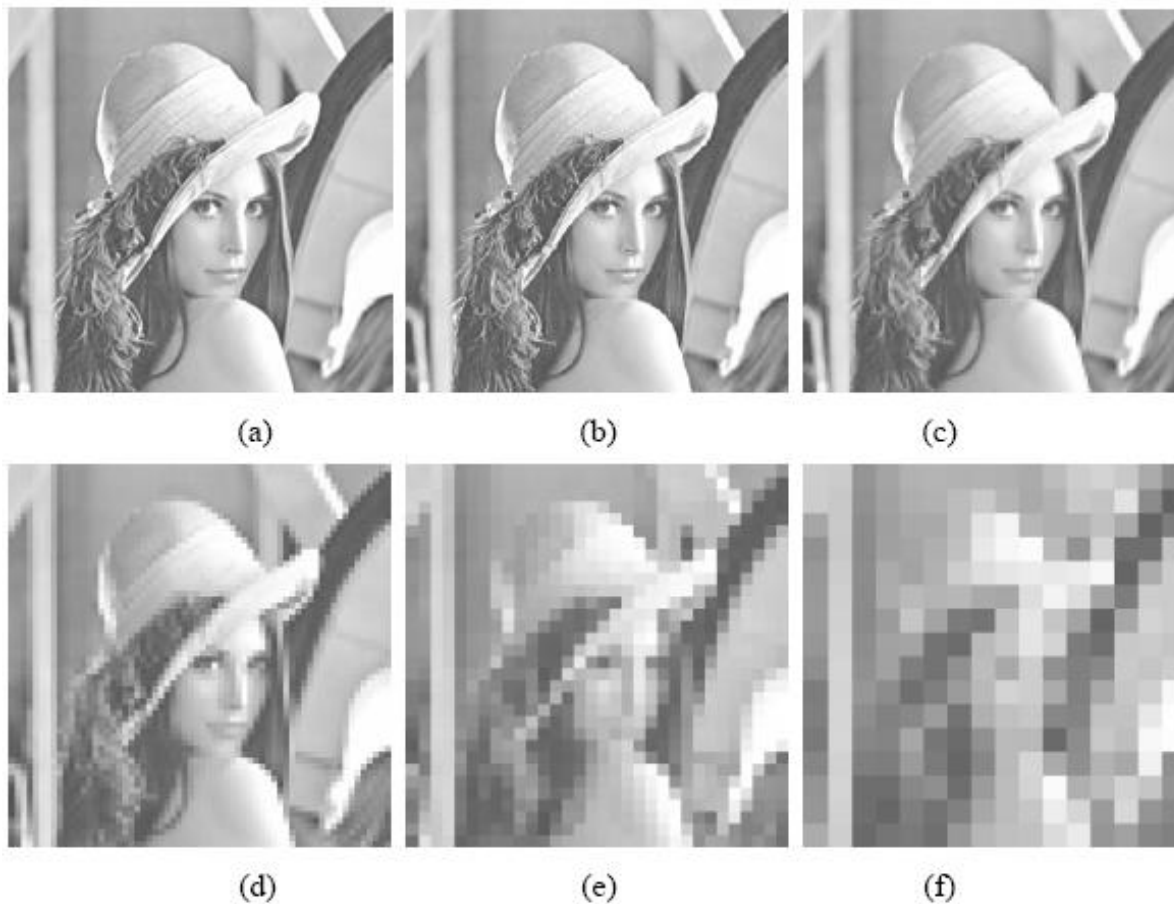


图 2.4.2 图像空间分辨率变化所产生的效果

采样与量化

□ 图像幅度分辨率变化所产生的效果 (256, 64, 16, 8, 4, 2)

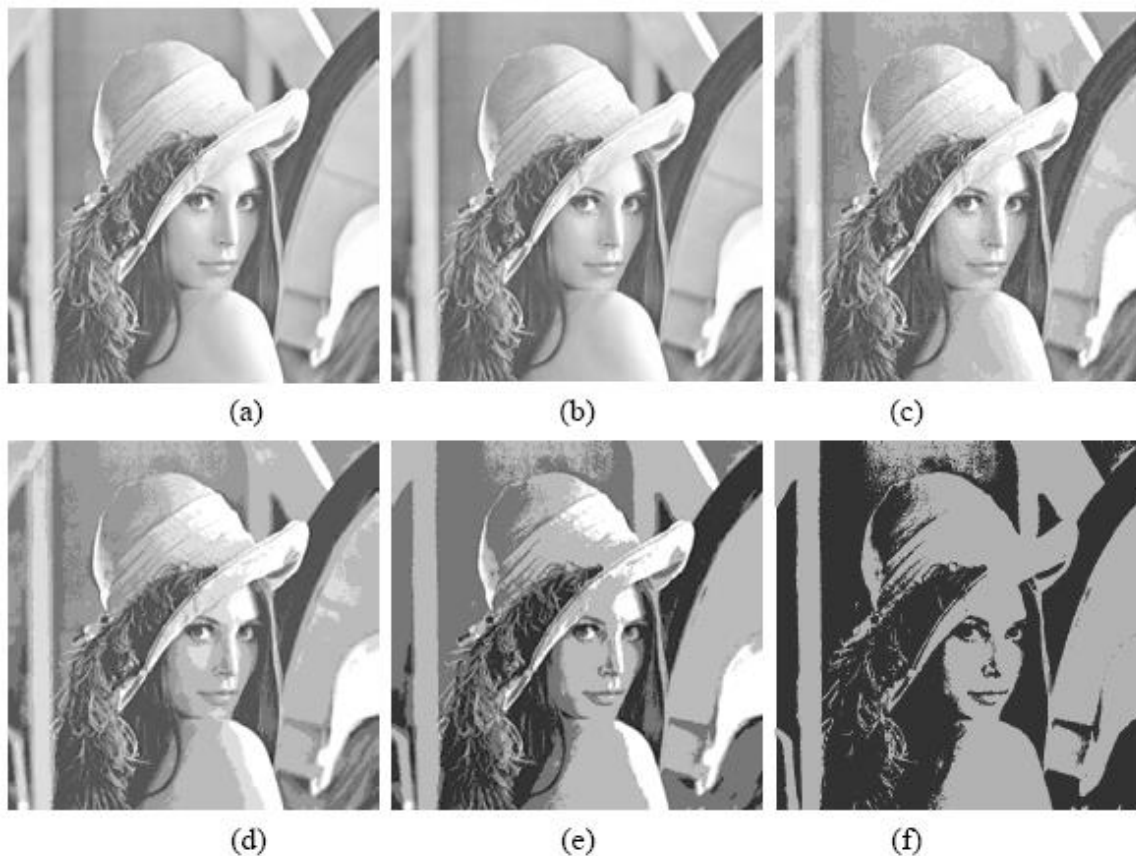


图 2.4.3 图像幅度分辨率变化所产生的效果

采样与量化

□ 空间和幅度分辨率同时变化所产生的效果



256x256,128;
181x181,64;
128x128,32;
90x90,16;
64x64,8;
45x45,4

图 2.4.4 图像空间和幅度分辨率同时变化所产生的效果



图像数字化

□ 数字图像基础

■ 成像过程

■ 像素间联系

✓ 像素的邻域

✓ 像素间的邻接，连接和连通

✓ 像素间的距离

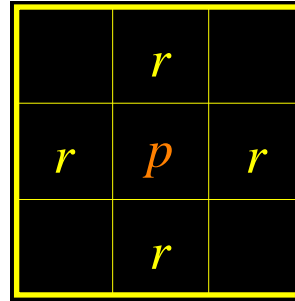
■ 数字化

■ 图像变换

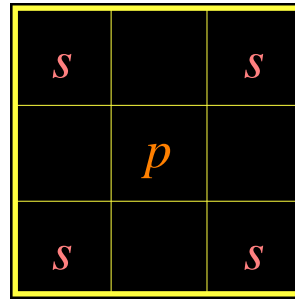
□ 形态学

像素的邻域

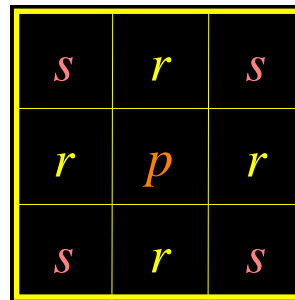
□ 4-邻域—— $N_4(p)$



□ 对角邻域—— $N_D(p)$



□ 8-邻域—— $N_8(p)$





像素间的邻接，连接和连通

□ 连接和连通

(adjacency, 邻接) **vs.** (connectivity, 连接)

邻接仅考虑像素间的空间关系

两个像素是否连接：

- (1) 是否接触（邻接）
- (2) 灰度值是否满足某个特定的相似准则（同在一个灰度值集合中取值）

像素间的邻接，连接和连通

□ 3种连接

■ 4-连接

✓ 2个像素 p 和 r 在 V 中取值且 r 在 $N_4(p)$ 中

■ 8-连接

✓ 2个像素 p 和 r 在 V 中取值且 r 在 $N_8(p)$ 中

0	1	0
0	1	0
0	0	1

0	1	0
0	1	0
0	0	1

像素间的邻接，连接和连通

□ 3种连接

- m-连接（混合连接）：2个像素 p 和 r 在 V 中取值且满足下列条件之一
 - ✓ r 在 $N_4(p)$ 中
 - ✓ r 在 $N_D(p)$ 中且集合 $N_4(p) \cap N_4(r)$ 不包含 V 中取值的像素

	a		
b	p	c	
	d	r	e
		f	

(a)

	1		
1	1	0	
	0	1	0
		1	

(b)

	1		
1	1	1	
	0	1	0
		1	

(c)

像素间的邻接，连接和连通

□ 混合连接的应用：消除8-连接可能产生的歧义性

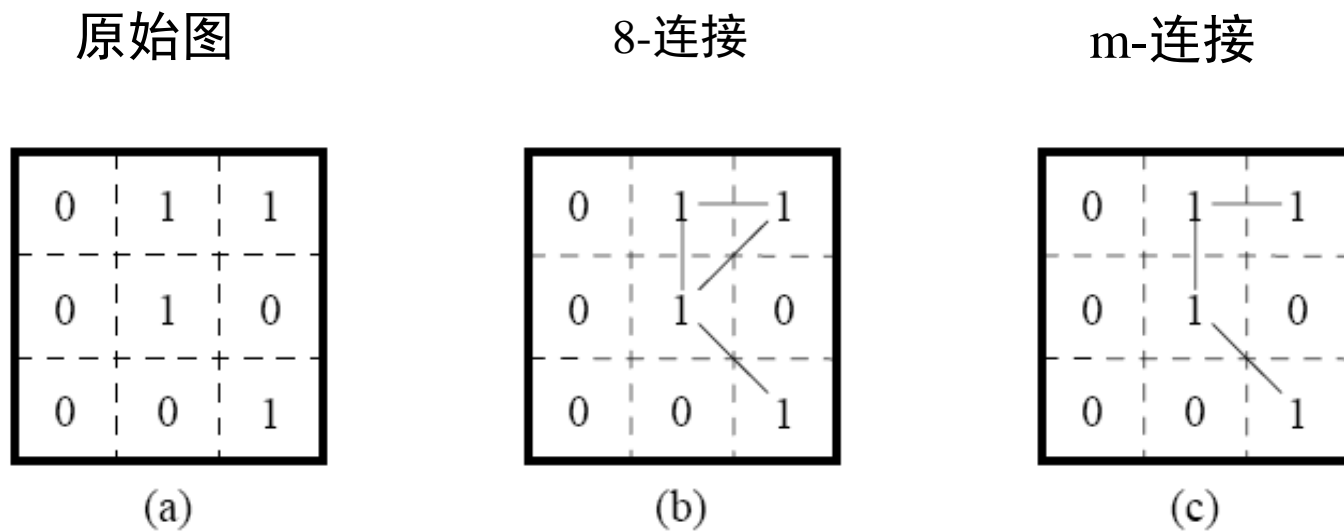


图 3.1.3 像素间的混合连接



像素间的邻接，连接和连通

□ 连通

- 连接是连通的一种特例

□ 通路

- 由一系列依次连接的像素组成

从具有坐标 (x, y) 的像素 p 到具有坐标 (s, t) 的像素 q 的一条通路由一系列具有坐标 (x_0, y_0) , (x_1, y_1) , ..., (x_n, y_n) 的独立像素组成。

这里 $(x_0, y_0) = (x, y)$, $(x_n, y_n) = (s, t)$, 且 (x_i, y_i) 与 (x_{i-1}, y_{i-1}) 连接, 其中 $1 \leq i \leq n$, n 为通路长度

4-连通, 8-连通 \Rightarrow 4-通路, 8-通路



像素间的邻接，连接和连通

□ 像素集合的连接和连通

- 对2个图像子集 S 和 T 来说，如果 S 中的一个或一些像素与 T 中的一个或一些像素连接，则可以说2个图像子集 S 和 T 是连接的
- 完全在一个图像子集中的像素组成的通路上的像素集合构成该图像子集中的一个**连通组元**
- 如果 S 中只有1个连通组元，即 S 中所有像素都互相连通，则称 S 是一个**连通集**

像素间的距离

□ 距离度量

■ 距离度量函数的定义

✓ 给定3个像素 \mathbf{x} , \mathbf{y} , \mathbf{z} , 坐标分别为 (x_1, x_2) , (y_1, y_2) , (z_1, z_2) , 如果下列条件满足, 则 D 是距离度量函数:

① $D(\mathbf{x}, \mathbf{y}) \geq 0$ ($D(\mathbf{x}, \mathbf{y}) = 0$, 当且仅当 $\mathbf{x} = \mathbf{y}$)

② $D(\mathbf{x}, \mathbf{y}) = D(\mathbf{y}, \mathbf{x})$

③ $D(\mathbf{x}, \mathbf{z}) \leq D(\mathbf{x}, \mathbf{y}) + D(\mathbf{y}, \mathbf{z})$

■ L_p 范数 (norm) 及其诱导出的距离度量: 对于向量 $\mathbf{u} = [u_1, u_2, \dots, u_k]^T$

✓ L_p 范数: $\|\mathbf{u}\|_p = (\sum_{i=1}^k |u_i|^p)^{\frac{1}{p}}$

✓ L_p 范数诱导的距离度量: $D(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|_p = (\sum_{i=1}^k |u_i - v_i|^p)^{\frac{1}{p}}$

像素间的距离

□ 距离量度函数

■ 欧氏 (Euclidean) 距离

✓ 等价于 L_2 距离

$$D_e(\mathbf{x}, \mathbf{y}) = [(x_1 - y_1)^2 + (x_2 - y_2)^2]^{\frac{1}{2}}$$

■ 城区 (city-block) 距离

✓ 等价于 L_1 距离

$$D_4(\mathbf{x}, \mathbf{y}) = |x_1 - y_1| + |x_2 - y_2|$$

■ 棋盘 (chessboard) 距离

✓ 等价于 L_∞ 距离 (请证明该结论)

$$D_8(\mathbf{x}, \mathbf{y}) = \max(|x_1 - y_1|, |x_2 - y_2|)$$

像素间的距离

□ 距离量度函数

等距离轮廓图案

			3			
	2.8	2.2	2	2.2	2.8	
	2.2	1.4	1	1.4	2.2	
3	2	1	0	1	2	3
	2.2	1.4	1	1.4	2.2	
	2.8	2.2	2	2.2	2.8	
			3			

(a)

D_E

			3			
		3	2	3		
	3	2	1	2	3	
3	2	1	0	1	2	3
	3	2	1	2	3	
		3	2	3		
			3			

(b)

D_4

3	3	3	3	3	3	3
3	2	2	2	2	2	3
3	2	1	1	1	2	3
3	2	1	0	1	2	3
3	2	1	1	1	2	3
3	2	2	2	2	2	3
3	3	3	3	3	3	3

(c)

D_8



像素间的距离

□ 用距离定义邻域

- 考虑在空间点 (x_p, y_p) 的像素 p

- ✓ 4-邻域—— $N_4(p)$

$$N_4(p) = \{r \mid D_4(p, r) = 1\}$$

- ✓ 8-邻域—— $N_8(p)$

$$N_8(p) = \{r \mid D_8(p, r) = 1\}$$



图像数字化

□ 数字图像基础

- 成像过程
- 像素间联系
- 数字化
 - ✓ 图像采集网格
 - ✓ 数字化模型
 - ✓ 离散直线性
 - ✓ 距离变换
 - ✓ 3-D图像的连通与拓扑
- 图像变换

□ 形态学

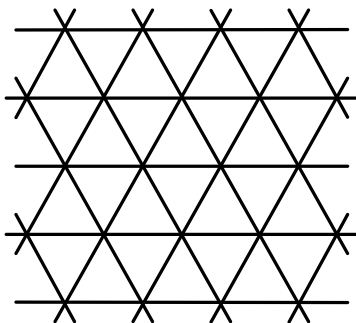
图象采集网格

□ 三种不同的采样模式

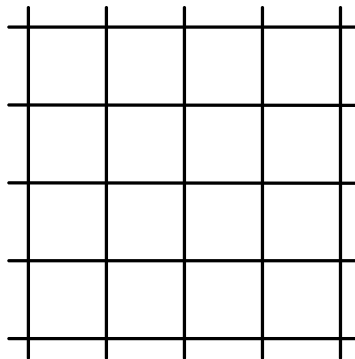
- 图象采集：用一个离散的模式采样
- 三种规则的形式：

$$m \cdot \frac{(n-2)\pi}{n} = 2\pi$$

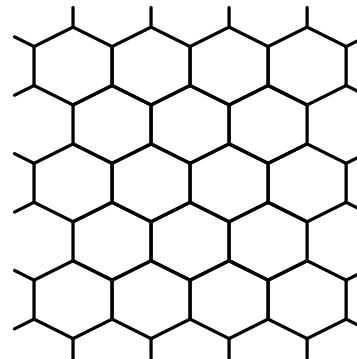
三角形



正方形



六边形



网格：将图象平面分解成小单元的集合

采样时，采集在各多边形的中央点的数据

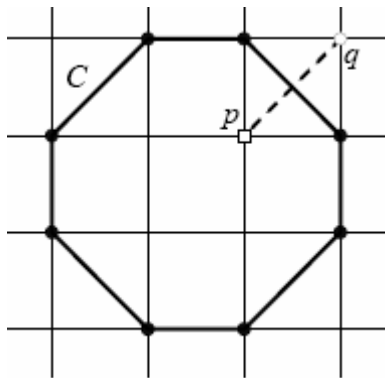
- 图象网格与采样模式互补
 - ✓ 三角形模式 \Leftrightarrow 六边形网格
 - ✓ 正方形模式 \Leftrightarrow 正方形网格
 - ✓ 六边形模式 \Leftrightarrow 三角形网格

图象采集网格

□ 三种不同的采样模式

■ 图象网格与采样模式互补

✓ 正方形网格



广泛使用：直观，无边界问题

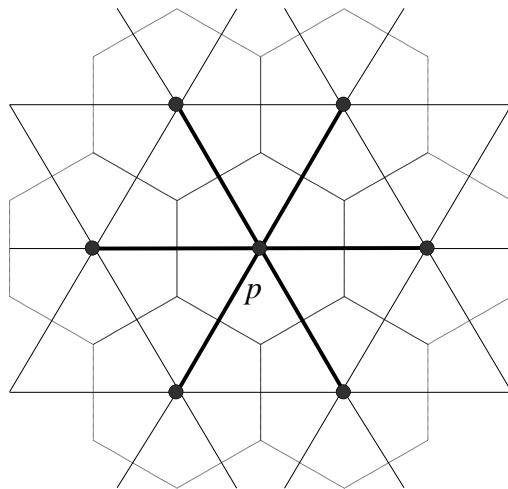
结构问题：“连通悖论”

图象采集网格

□ 三种不同的采样模式

■ 图象网格与采样模式互补

✓ 三角形网格



相邻像素：有共同边
粗实线连接相邻像素
细线表示三角形网格
点线对应采样模式

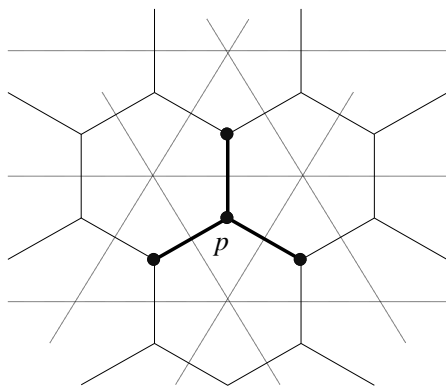
对像素 p ，它的6-邻域记为 $N_6(p)$

图象采集网格

□ 三种不同的采样模式

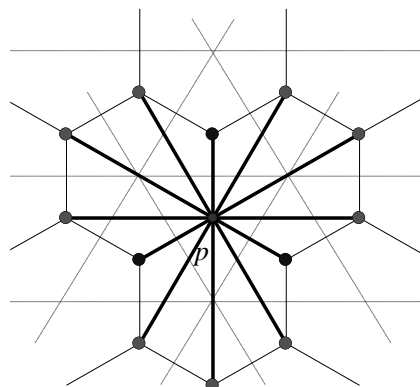
■ 图象网格与采样模式互补

✓ 六边形网格



$N_3(p)$: 有公共边

邻域过于稀疏



$N_{12}(p)$: 有公共顶点

图象采集网格

□ 三种图象采样效率比较

■ 采样效率定义为：单位圆面积与覆盖该单位圆的网格面积之比

✓ 三角形采样效率： $\pi / 3\sqrt{3} = (0.60)$

✓ 四边形采样效率： $\pi / 4 = (0.79)$

✓ 六边形采样效率： $\pi / 2\sqrt{3} = (0.91)$



数字化模型

- 与图象采集密切相关
 - 数字化模型基础
 - 方盒量化
 - 网格相交量化



数字化模型基础

□ 两个定义

■ 预图象 (pre-image)

- ✓ 给定一个离散点集合 P ，一个其数字化为 P 的连续点集合 S 称为 P 的预图象

■ 域 (domain)

- ✓ 由所有可能的预图象 S 的并集所定义的区域称为 P 的域

数字化模型基础

- 示例：将一个正方形图象网格覆盖到连续的目标 S 上，一个象素用一个正方形网格上的交点 p 表示，该象素当且仅当 $p \in S$ 时属于 S 的数字化结果

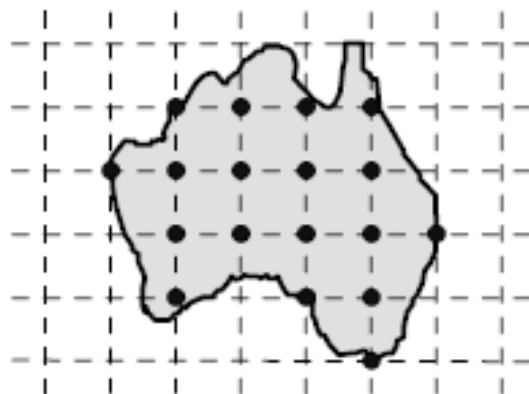


图 3.2.1 一个简单的数字化模型示例

- S 在图中用阴影部分表示，黑色圆点代表属于 S 的象素 p ，所有 p 组成集合 P

数字化模型基础

□ 采用不同采样步长数字化连续集合效果

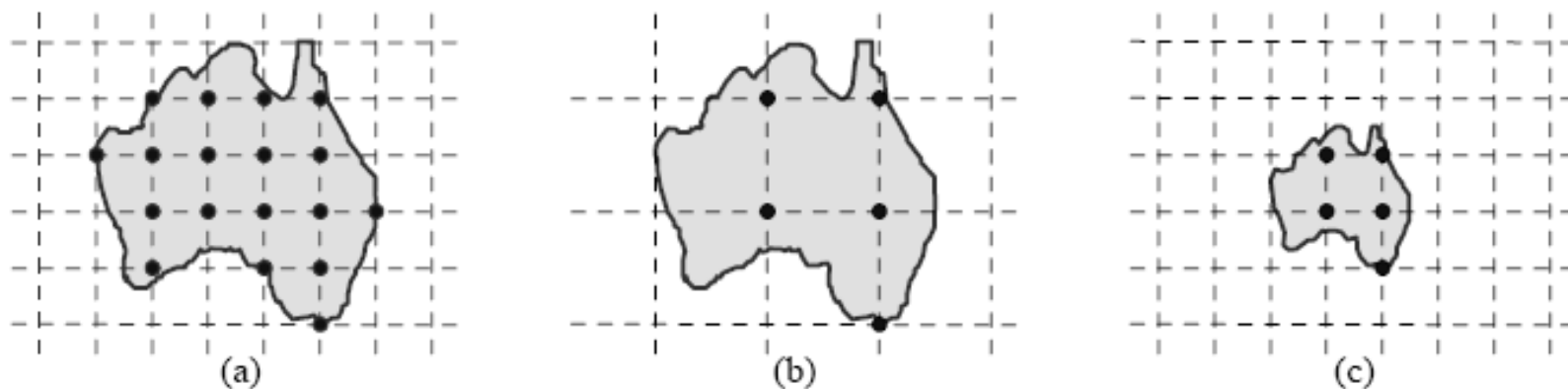


图 3.2.2 用不同的采样步长数字化连续集合的效果

数字化模型基础

□ 不一致性

- 一个非空集合 S 有可能映射到一个空的数字化集合中
- 该数字化模型不是平移不变(称为混叠: aliasing)
- 给定一个数字化集合 P , 并不能保证精确地刻画它的预图象 S

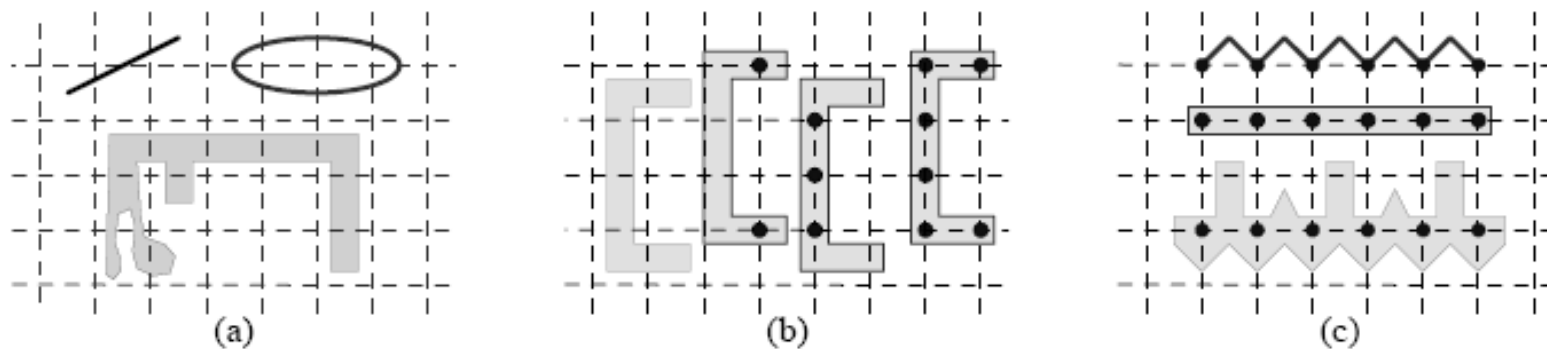


图 3.2.3 数字化模型的不一致性示例

数字化模型基础

□ 合适的数字化模型应具有的特征

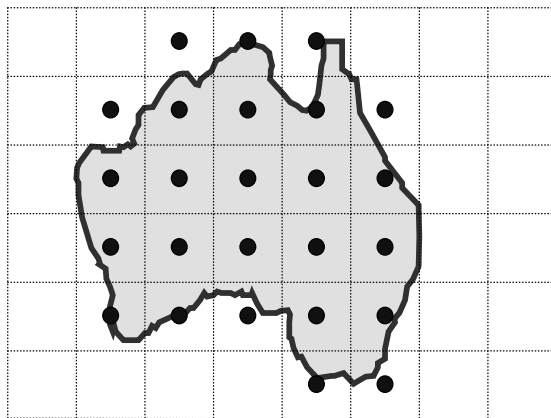
- 对一个非空的连续集合的数字化结果应该是非空的
- 数字化模型应该尽可能平移不变(即混叠效应尽可能小)
- 给定一个数字化集合 P ，其各个预图象应在一定准则下相似。
更严格说， P 的域应该有限且越小越好

□ 常用的数字化模型

- 方盒量化
- 网格量化
- 目标量化

方盒量化

- 一种数字化模型：对任何象素 $p_i = (x_i, y_i)$ ，都有一个对应的数字化盒 $B_i = (x_i - 1/2, x_i + 1/2) \times (y_i - 1/2, y_i + 1/2)$
- 数字化盒等价于中心为象素位置的分割多边形。
- 一个象素 p_i 当且仅当 $B_i \cap S \neq \emptyset$ 时（即它对应的数字化盒 B_i 与 S 相交）处在 S 的数字化集合 P 中



方盒量化

□ 方盒量化特性

- 数字化盒 $B_i = [x_i - 1/2, x_i + 1/2) \times [y_i - 1/2, y_i + 1/2)$, B_i 的闭包 $[B_i] = [x_i - 1/2, x_i + 1/2] \times [y_i - 1/2, y_i + 1/2]$
- 一个像素 p_i 当且仅当 $B_i \cap S \neq \emptyset$ 时, 则其处在 S 的数字化集合 P 中 (半开片量化)



方盒量化

□ 方盒量化特性

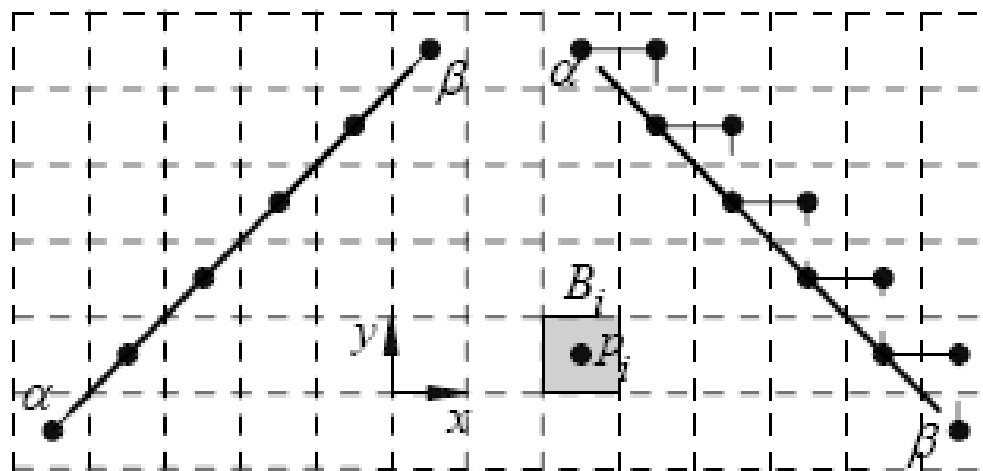
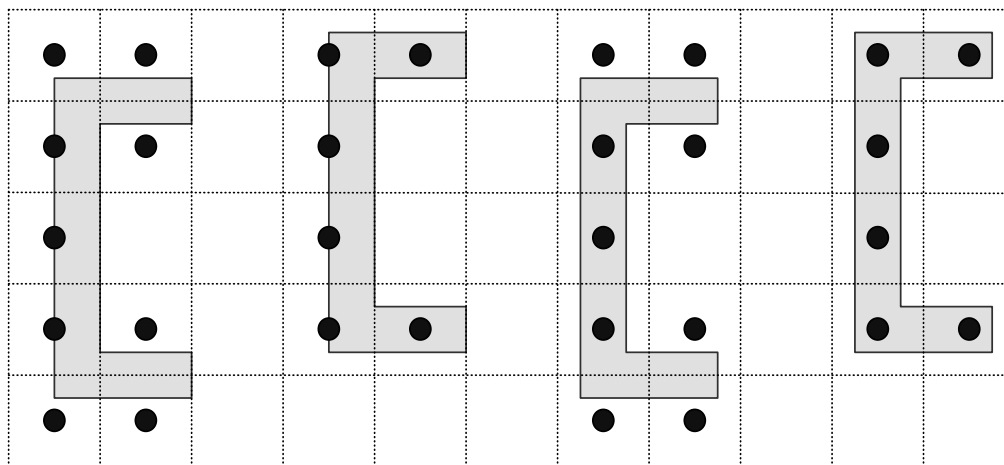


图 3.2.5 对连续直线段的半开片量化

方盒量化

□ 方盒量化特性

- 对一个连续直线段的方盒量化 (SBQ) 的结果是一个4-数字弧
- 方盒量化的定义保证了非空集合S会被映射到非空离散集合P
- 方盒量化不保证完全的平移不变性



网格相交量化

- 给定一个连续的细目标 C ，它与网格线的交点定义一个实点 $t = (x_t, y_t)$ ，该点视 C 与垂直网格线相交或与水平网格线相交分别满足 $x_t \in I$ 或 $y_t \in I$ （ I 代表整数集合）。这个点 $t \in C$ 将被映射到一个网格点 $p_i = (x_i, y_i)$ ，这里 $t \in (x_i - 1/2, x_i + 1/2) \times (y_i - 1/2, y_i + 1/2)$
- 在特殊情况（如 $x_t = x_i + 1/2$ 或 $y_t = y_i + 1/2$ ）下，取落在左边或下边的点 p_i 属于离散集合 P

网格相交量化

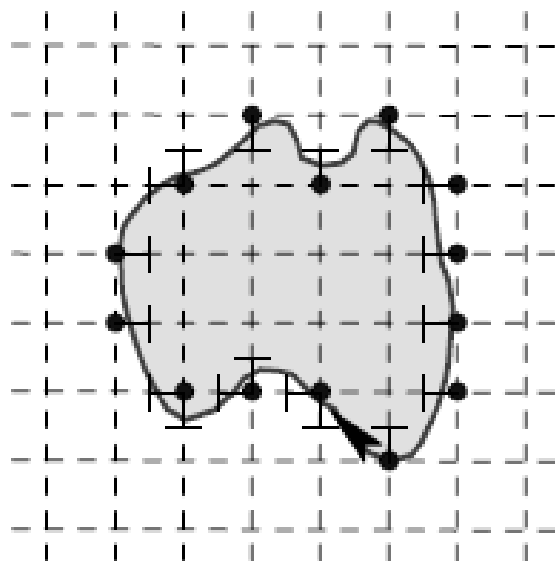


图 3.2.8 对连续曲线 C 的网格相交量化示例

网格相交量化

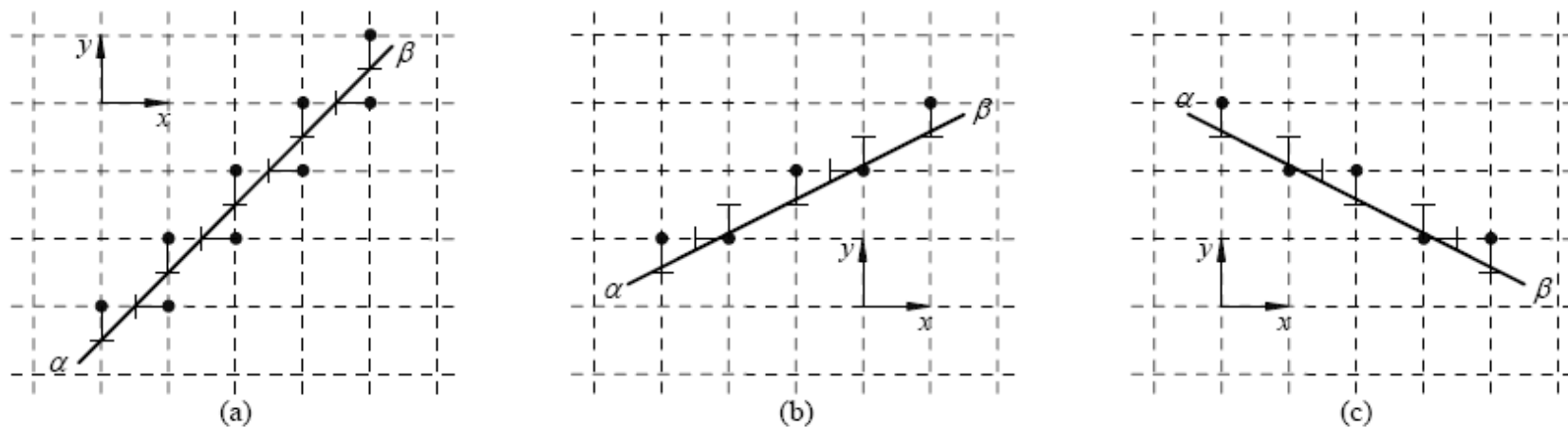


图 3.2.9 对连续直线段的网格相交量化示例

网格相交量化

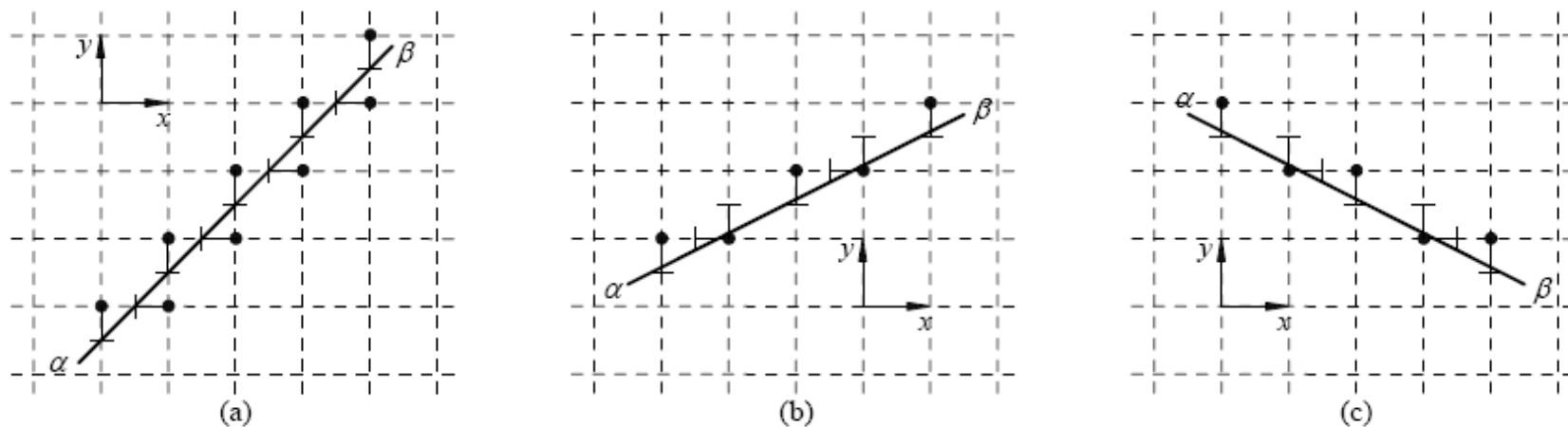


图 3.2.9 对连续直线段的网格相交量化示例



离散直线性

- 有关直线性的定理和性质可以用来判断一个数字弧是否是一条数字直线段（弦）
 - 弦和弧
 - 直线性

弦和弧

□ 弧

- 非自相交曲线上介于两点间的部分

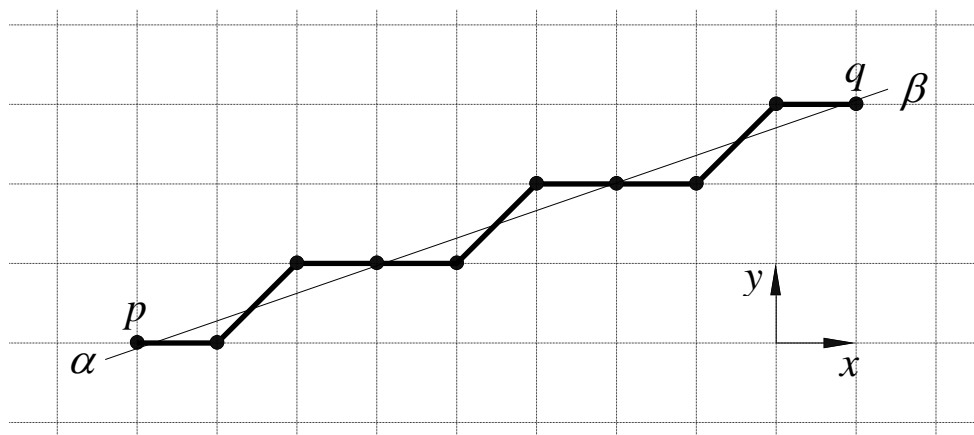
□ 数字弧

- 从点 p 到点 q 的数字弧 P_{pq} 定义为满足下列条件的弧 $P_{pq} = \{p_i, i = 0, 1, \dots, n\}$
 - ✓ $p_0 = p, p_n = q$
 - ✓ $\forall i = 1, \dots, n-1$, 点 p_i 在弧 P_{pq} 中正好有两个相邻点: p_{i-1} 和 p_{i+1}
 - ✓ 端点 p_0 / p_n 在弧 P_{pq} 中正好有一个相邻点: p_1 / p_{n-1}

弦和弧

□ 数字化集合

■ 网格相交 (grid-intersect) 量化模型



在 $[\alpha, \beta]$ 之间与网格线相交的点都映射到它们最接近的整数点
(相等时取 $[a, b]$ 左边的)

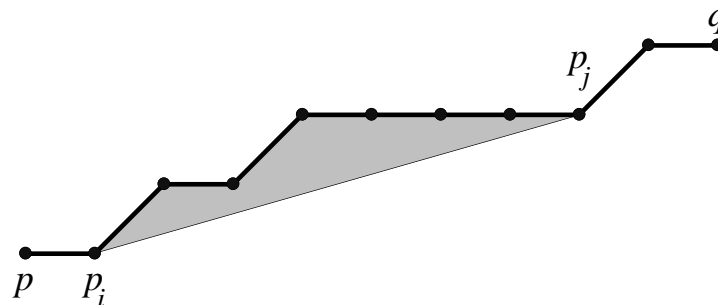
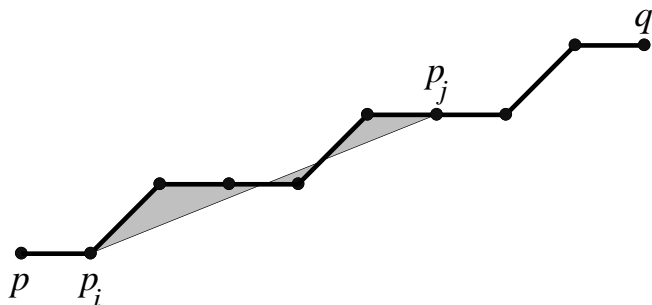
弦和弧

□ 弦

- 连接圆锥曲线上任意两点间的直线段

□ 数字弦（思路）

- 给定一条从 $p = p_0$ 到 $q = p_n$ 的数字弧 $P_{pq} = \{p_i\} i = 0, \dots, n$, 连续线段 $[p_i, p_j]$ 和各段之和 $\sum [p_i, p_{i+1}]$ 间的距离可用离散距离函数来测量, 且不应该超过一定的阈值

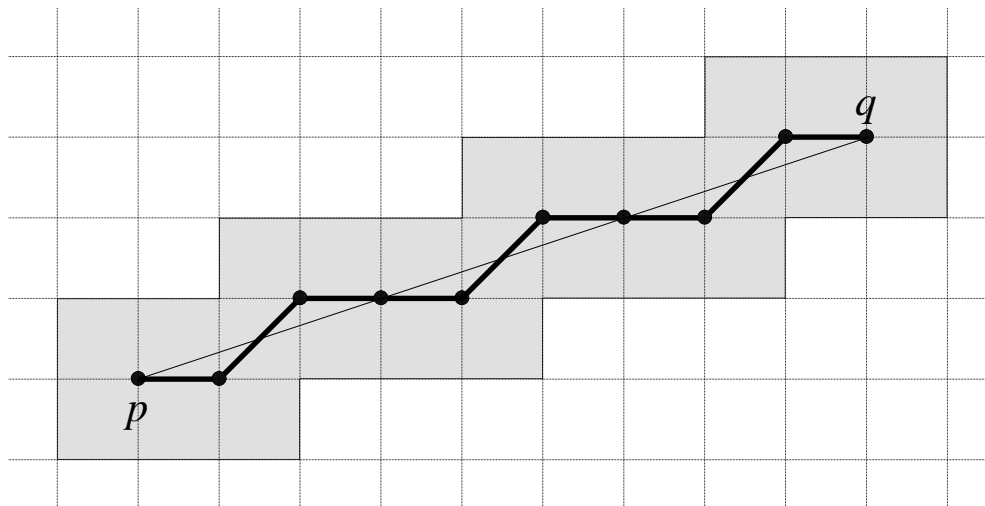


弦和弧

□ 数字弦的性质（8-数字弧）

- 一条8-数字弧 $P_{pq} = \{p_i\}_{i=0, \dots, n}$ 满足弦的性质，如果当且仅当对 P_{pq} 中的任意两个离散点 p_i 和 p_j 以及连续线段 $[p_i, p_j]$ 中的任意实点 ρ ，存在一个点 $p_k \in P_{pq}$ 使得 $d_8(\rho, p_k) < 1$

阴影多边形给出
点 $\rho \in \mathbf{R}^2$ 的集合，
可以看出总存在
一个点 $p_k \in P_{pq}$
使得 $d_8(\rho, p_k) < 1$



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

弦和弧

□ 数字弦的性质（8-数字弧）

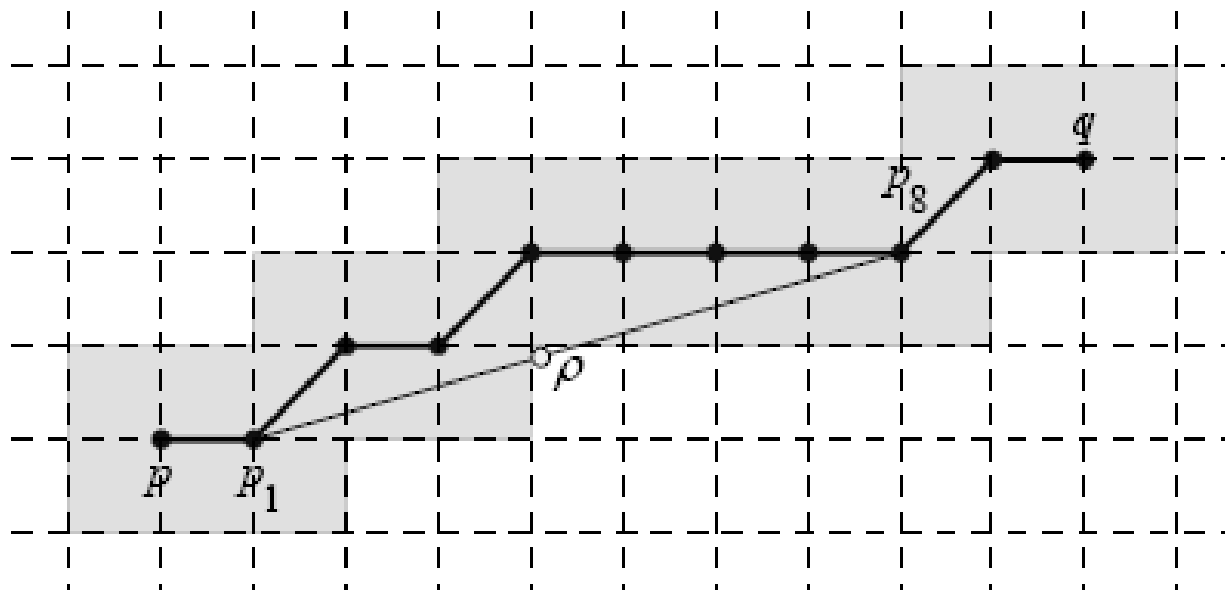


图 3.3.4 弦性质不成立的例子



弦和弧

□ 数字弦的性质（8-数字弧）

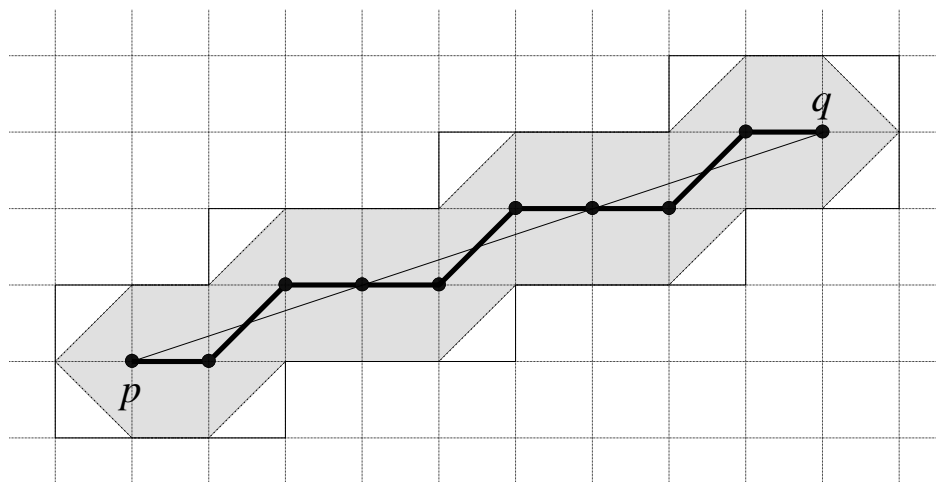
- 在8数字空间中，任一直线段的数字化结果，必是一条数字弦
- 任一条数字弦，也必存在一条直线段，其数字化结果与该弦同

弦和弧

□ 紧致弦性质

- 一条8-数字弧 $P_{pq} = \{p_i\} i = 0, \dots, n$ 满足紧致 (compact) 弦性质：
如果当且仅当对 P_{pq} 中的任意两个不同的离散点 p_i 和 p_j 以及连续
线段 $[p_i, p_j]$ 中的任意实点 a ，在各段之和 $\bigcup_i [p_i, p_{i+1}]$ 中存在一个
实点 $b \in \mathbf{R}^2$ 使得 $d_4(a, b) < 1$

紧致弦可见多边形包含
在弦可见多边形之中

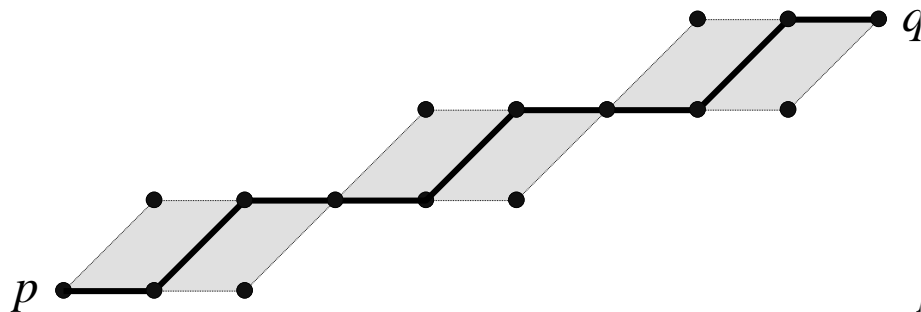


$$D_4(p, q) = |x - s| + |y - t|$$

直线性

□ 8-数字直线段的上下限

- 数字直线段可表示成一系列特定线段的组合



链码为 $\{c_i\}_{i=1, \dots, n} = \{0, 1, 0, 0, 1, 0, 0, 1, 0\}$

平移 $n-1$ 次可产生 $n-1$ 个平移的链码

对应从 p 到 q 的不同的数字直线段

上限



平移	链码	平移	链码	平移	链码
0	{0, 1, 0, 0, 1, 0, 0, 1, 0}	3	{0, 1, 0, 0, 1, 0, 0, 1, 0}	6	{0, 1, 0, 0, 1, 0, 0, 1, 0}
1	{1, 0, 0, 1, 0, 0, 1, 0, 0}	4	{1, 0, 0, 1, 0, 0, 1, 0, 0}	7	{1, 0, 0, 1, 0, 0, 1, 0, 0}
2	{0, 0, 1, 0, 0, 1, 0, 0, 1}	5	{0, 0, 1, 0, 0, 1, 0, 0, 1}	8	{0, 0, 1, 0, 0, 1, 0, 0, 1}

距离变换

- 距离变换基于对距离的计算，其本身是一个全局概念，但可以借助对局部距离的计算而化整为零地进行
 - 定义和性质
 - 局部距离的计算
 - 离散距离变换的实现
 - 3-D距离变换

定义和性质

- 距离变换计算区域中的每个点与最近的区域外的点之间距离，把二值图象变换为灰度图象
- 给定一个点集 P 、一个子集 B 以及满足测度条件的距离函数 $d(., .)$ ，对 P 的距离变换中赋予点 $p \in P$ 的值为：

$$DT(p) = \min_{q \in B} \{d(p, q)\}$$

- 距离图 (map) 可用矩阵 $[DT(p)]$ 来表示

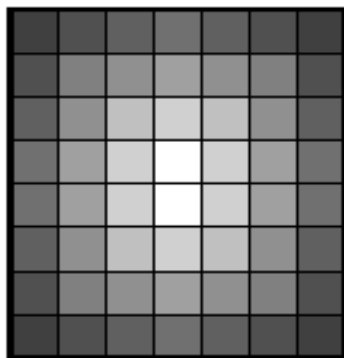


图 3.4.1 用灰度图表示的距离图

局部距离的计算

□ 全局的操作，所以计算量会很大

性质： 给定一个离散集合 P 和它的一个子集 B ，用 d 表示计算距离图的离散距离函数。那么，对任何点 $p \in P^\circ$ (即 $p \in P-B$)，存在 p 的一个邻域点 q (即 $q \in N(p)$)，使得在 p 的离散距离变换值 $DT(p)$ 满足 $DT(p) = DT(q) + d(p, q)$

进一步，因为 p 和 q 互为邻接点，从 p 移动到 q 的长度为 $l(p, q) = d(p, q)$ 。这样，对任意点 $p \notin B$ ，离散距离变换可由 $DT(p) = \min\{DT(q')\} + l(p, q')$ ， $q' \in N(p)$ 来刻画

局部距离的计算

□ 用于局部距离扩展的模板

- 模板基于4-邻域定义且被用来扩展 d_4 距离
- 模板基于8-邻域且被用来扩展 d_8 距离或 $d_{a,b}$ 距离 ($a = 1, b = 1$)
- 模板基于16-邻域且被用来扩展 $d_{a,b,c}$ 距离

∞	a	∞
a	0	a
∞	a	∞

(a)

b	a	b
a	0	a
b	a	b

(b)

∞	c	∞	c	∞
c	b	a	b	c
∞	a	0	a	∞
c	b	a	b	c
∞	c	∞	c	∞

(c)

图 3.4.2 用于计算距离变换的模板

局部距离的计算

□ 初始化距离图

$$DT^{(0)}(p) = \begin{cases} 0 & \text{当 } p \in B \\ \infty & \text{当 } p \notin B \end{cases}$$

□ 用下面规则将距离值从像素 $q = (x_p + k, y_p + l)$ 传播到 p

$$DT^{(t)}(p) = \min_{k,l} \{DT^{(t-1)}(q) + M(k, l); q = (x_p + k, y_p + l)\}$$

□ 更新过程持续进行到距离图不再变化而停止

离散距离变换的实现

□ 串行实现

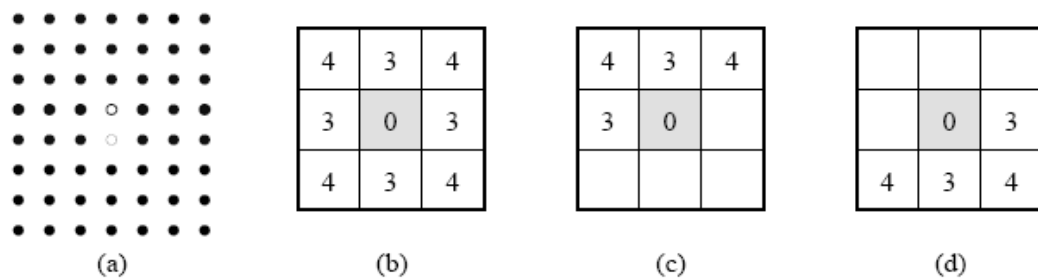


图 3.4.3 串行计算距离变换

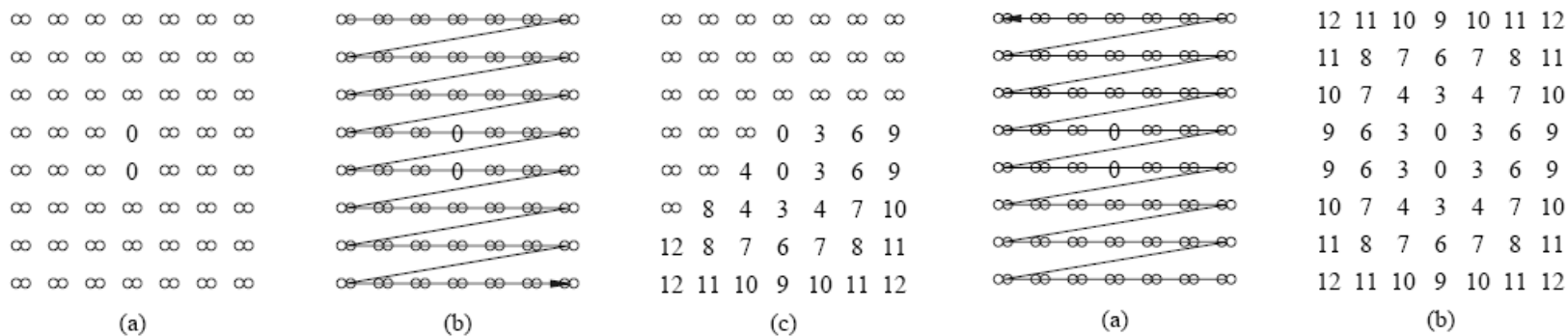


图 3.4.4 前向扫描

图 3.4.5 反向扫描

离散距离变换的实现

□ 并行实现

$$DT^{(t)}(p) = \min_{k,l} \{DT^{(t-1)}(q) + M(k,l); q = (x_p + k, y_p + l)\}$$

∞	∞	∞	∞	∞	∞	∞
∞	∞	∞	∞	∞	∞	∞
∞	∞	4	3	4	∞	∞
∞	∞	3	0	3	∞	∞
∞	∞	3	0	3	∞	∞
∞	∞	4	3	4	∞	∞
∞	∞	∞	∞	∞	∞	∞
∞	∞	∞	∞	∞	∞	∞

(a)

∞	∞	∞	∞	∞	∞	∞
∞	8	7	6	7	8	∞
∞	7	4	3	4	7	∞
∞	6	3	0	3	6	∞
∞	6	3	0	3	6	∞
∞	7	4	3	4	7	∞
∞	8	7	6	7	8	∞
∞	∞	∞	∞	∞	∞	∞

(b)

12	11	10	9	10	11	12
11	8	7	6	7	8	11
10	7	4	3	4	7	10
9	6	3	0	3	6	9
9	6	3	0	3	6	9
10	7	4	3	4	7	10
11	8	7	6	7	8	11
12	11	10	9	10	11	12

(c)

图 3.4.6 并行计算过程中的各个步骤所得到的距离图

3-D距离变换

□ 3-D距离

$$d_E(p, q) = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2 + (z_p - z_q)^2}$$

$$d_4(p, q) = |x_p - x_q| + |y_p - y_q| + |z_p - z_q|$$

$$d_8(p, q) = \max\{|x_p - x_q|, |y_p - y_q|, |z_p - z_q|\}$$

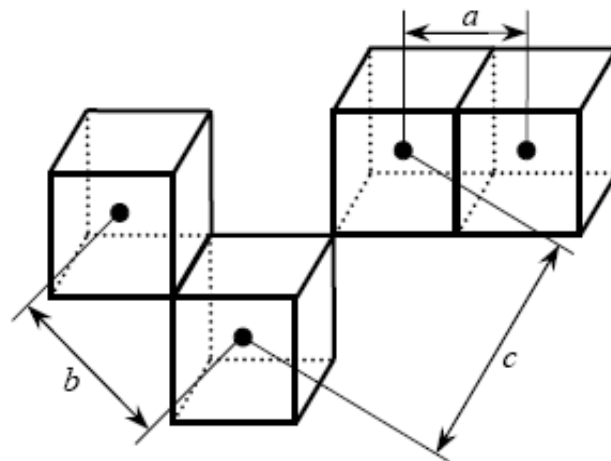


图 3.4.8 3-D 体素间的 3 种距离

3-D距离变换

□ 3-D距离变换的模板

- 前向扫描：从图象第一层的左上角向最后一层的右下角进行

$$\begin{array}{ccc} c & b & c \\ b & a & b \\ c & b & c \end{array} \quad \begin{array}{ccc} b & a & b \\ a & 0 & 0 \\ 0 & 0 & 0 \end{array} \quad \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}$$

图 3.4.9 3-D 前向扫描模板

- 反向扫描：从最后一层的右下角向第一层左上角进行

3-D图象中的连通和拓扑

□ 邻域和连通

■ 邻域的通用定义

设 $\mathbf{x} = (x_0, \dots, x_n)$ 为图象网格上的一个单元 \mathbf{x} 的 V_1^r 邻域定义为:

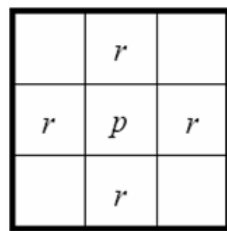
$$V_1^r = \{\mathbf{y} \mid D_1(\mathbf{x}, \mathbf{y}) \leq r\}$$

\mathbf{x} 的 V_∞^r 邻域定义为: $V_\infty^r = \{\mathbf{y} \mid D_\infty(\mathbf{x}, \mathbf{y}) \leq r\}$

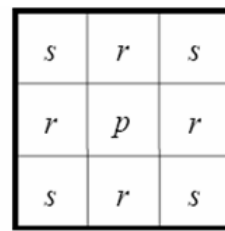
$$N_4(\mathbf{x}) = V_1^1(\mathbf{x}) \quad N_8(\mathbf{x}) = V_\infty^1(\mathbf{x})$$

3-D图象中的连通和拓扑

□ 2-D像素邻域



(a)

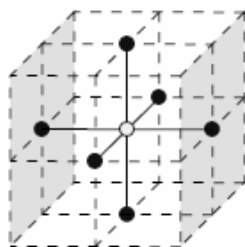


(b)

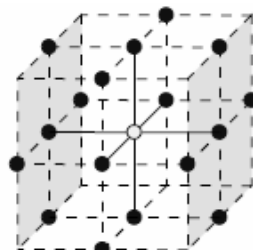
$$N_4(\mathbf{x}) = V_1^1(\mathbf{x})$$

$$N_8(\mathbf{x}) = V_\infty^1(\mathbf{x})$$

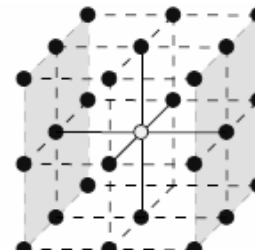
□ 3-D体素邻域



(a)



(b)



(c)

$$N_6(\mathbf{x}) = V_1^1(\mathbf{x})$$

$$N_{18}(\mathbf{x}) = V_\infty^1(\mathbf{x}) \cap V_1^2(\mathbf{x})$$

$$N_{26}(\mathbf{x}) = V_\infty^1(\mathbf{x})$$

3-D图象中的连通和拓扑

□ 像素点连通

- 存在一系列依次连接的像素组成的通路

□ 连通组元

- 完全在一个图象子集中的像素组成的通路上的像素集合构成该图象子集中的一个连通组元

□ 连通集

- 如果 S 中只有1个连通组元，即 S 中所有像素都互相连通，则称 S 是一个连通集



图像数字化

□ 数字图像基础

- 成像过程
- 像素间联系
- 数字化
- 图像变换
 - ✓ 可分离和正交图像变换
 - ✓ 离散傅立叶变换 (DFT)
 - ✓ 离散余弦变换 (DCT)
 - ✓ Karhunen-Loeve变换 (KLT)
 - ✓ 小波变换 (DWT)

□ 形态学

可分离和正交图象变换

□ 1-D变换

■ 正变换

$$T(u) = \sum_{x=0}^{N-1} f(x)h(x,u) \quad u = 0, 1, \Lambda, N-1$$

正向变换核

■ 反变换

$$f(x) = \sum_{u=0}^{N-1} T(u)k(x,u) \quad x = 0, 1, \Lambda, N-1$$

反向变换核

可分离和正交图象变换

□ 2-D可分离变换

(例子：傅里叶变换)

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

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} T(u, v) k(x, y, u, v) \quad (2)$$

正向变换核

变换核与
原始函数及
变换后函数无关

反向变换核



可分离和正交图象变换

□ 可分离

$$h(x, y, u, v) = h_1(x, u)h_2(y, v)$$

1个2-D变换分成2个1-D变换

$$T(x, v) = \sum_{y=0}^{N-1} f(x, y)h_2(y, v) \quad T(u, v) = \sum_{x=0}^{N-1} T(x, v)h_1(x, u)$$

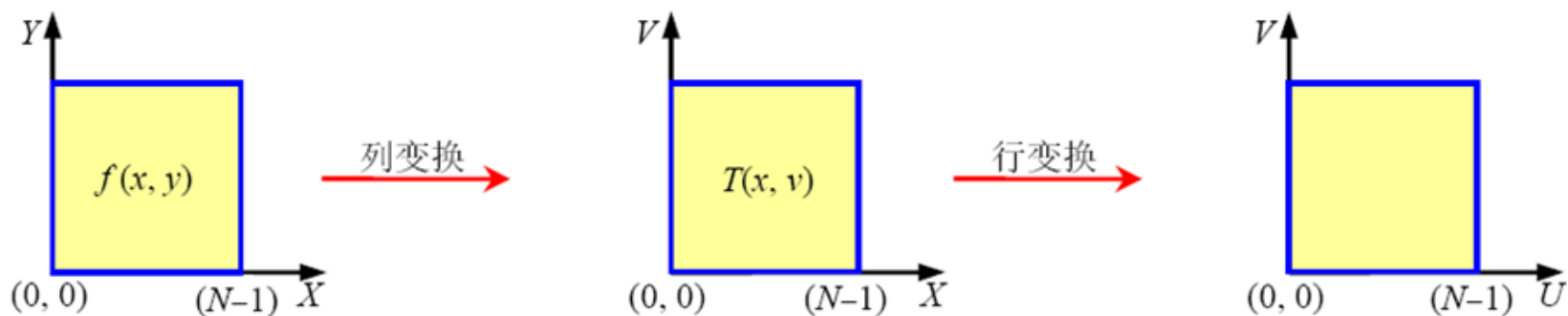
□ 对称

$$h(x, y, u, v) = h_1(x, u)h_1(y, v)$$

(h_1 与 h_2 的函数形式一样)

可分离和正交图象变换

- 具有可分离变换核的2-D变换可以分成两个步骤计算，每个步骤用一个1-D变换



由 2 步 1-D 变换计算 2-D 变换

可分离和正交图象变换

□ 可分离且对称

变换结果

$$\mathbf{T} = \mathbf{AFA}$$

对称变换矩阵

图象矩阵

反变换矩阵

$$\mathbf{BTB} = \mathbf{BAFAB}$$

$$\mathbf{F} = \mathbf{BTB} \quad \mathbf{B} = \mathbf{A}^{-1}$$

$$\hat{\mathbf{F}} = \mathbf{BAFAB} \quad \mathbf{B} \neq \mathbf{A}^{-1}$$



可分离和正交图象变换

□ 正交

考虑变换矩阵： $B = A^{-1}$ $F = BTB$

酉矩阵（*代表共轭）： $A^{-1} = A^{*T}$

如果A为实矩阵，且： $A^{-1} = A^T$

则A为正交矩阵，构成正交变换对



离散傅立叶变换 (DFT)

□ 二维离散傅立叶变换式

■ 对于 $N \times N$ 的二维矩阵（方阵），二维离散傅立叶变换对为：

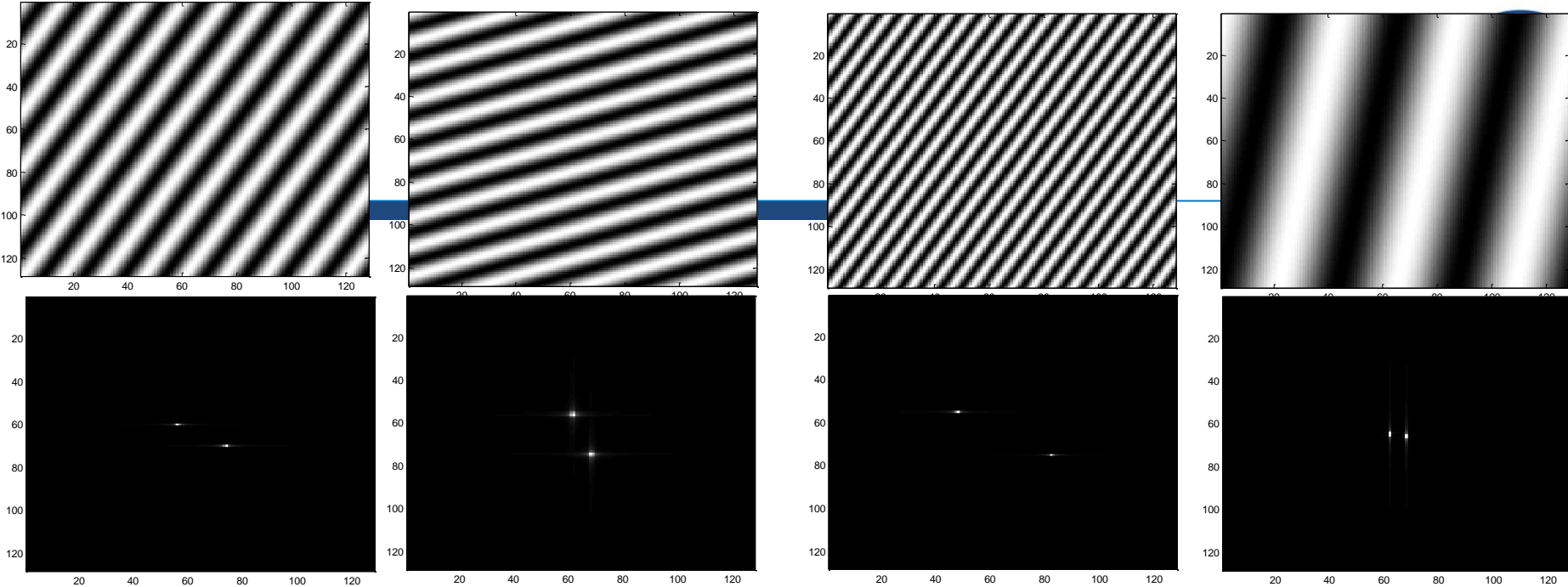
$$F(u, v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \exp\left\{\frac{-2\pi j(ux + vy)}{N}\right\}$$

$$f(x, y) = \frac{1}{N} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) \exp\left\{\frac{+2\pi j(ux + vy)}{N}\right\}$$

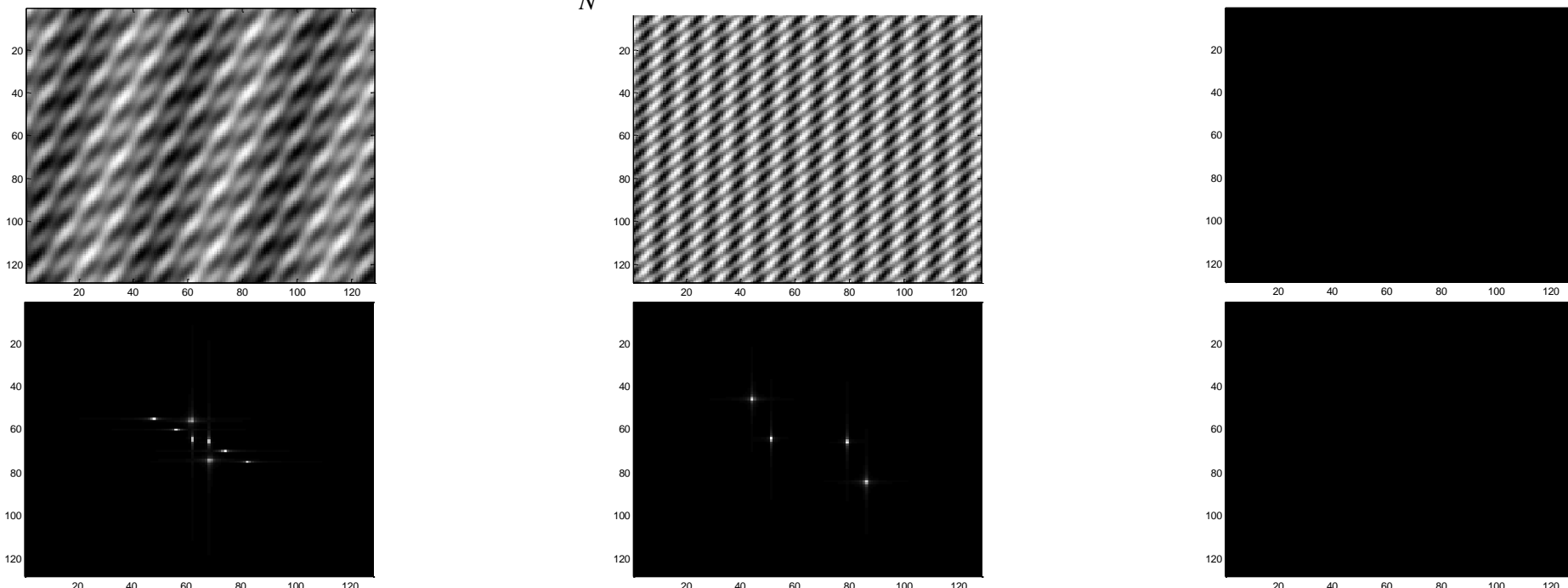


二维DFT的性质

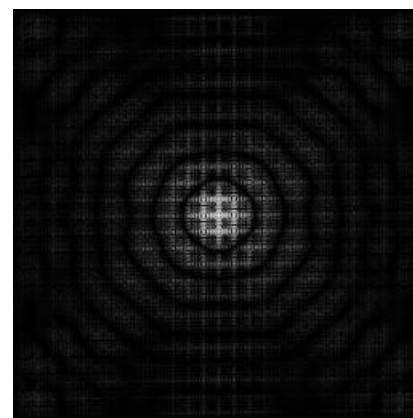
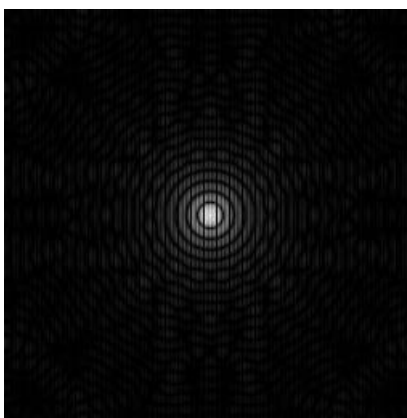
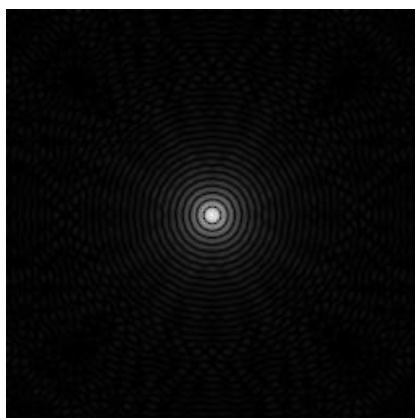
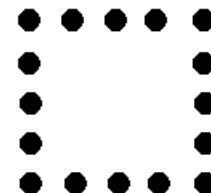
- 线性 $f_1(x, y) + f_2(x, y) \quad F_1(u, v) + F_2(u, v)$
- 比例 $f(ax, by) \quad \frac{1}{ab} F\left(\frac{u}{a}, \frac{v}{b}\right)$
- 平移 $f(x-a, y-b) \quad e^{-j2\pi(au+bv)} F(u, v)$
 $e^{j2\pi(cx+dy)} f(x, y) \quad F(u-c, v-d)$
- 卷积 $f_1(x, y) * f_2(x, y) \quad F_1(u, v) F_2(u, v)$
 $f_1(x, y) f_2(x, y) \quad F_1(u, v) * F_2(u, v)$
- 旋转 $f(x \cos \theta + y \sin \theta, -x \sin \theta + y \cos \theta)$
 $F(u \cos \theta + v \sin \theta, -u \sin \theta + v \cos \theta)$



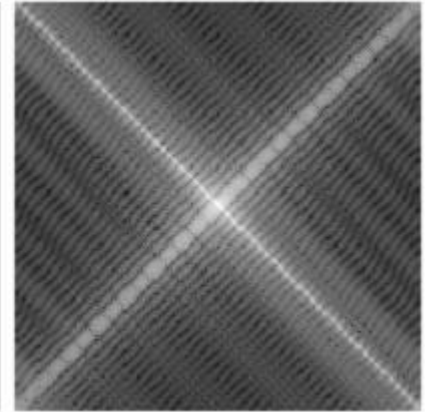
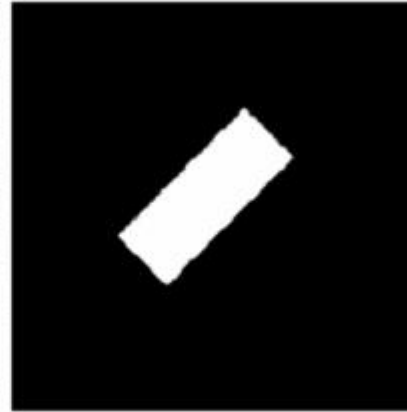
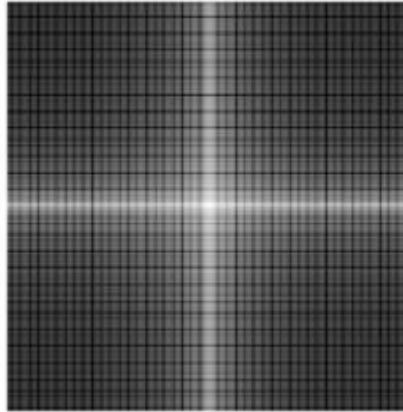
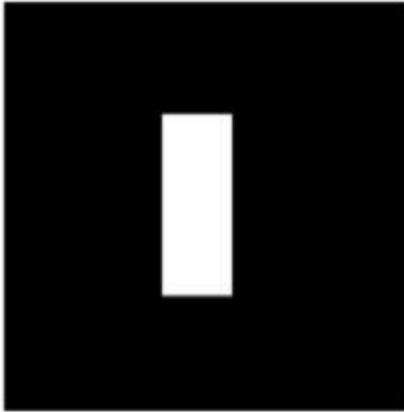
$$\cos\left[\frac{\pi}{N}(u \cdot x + v \cdot y)\right] \quad u = R \cos \theta, v = R \sin \theta$$



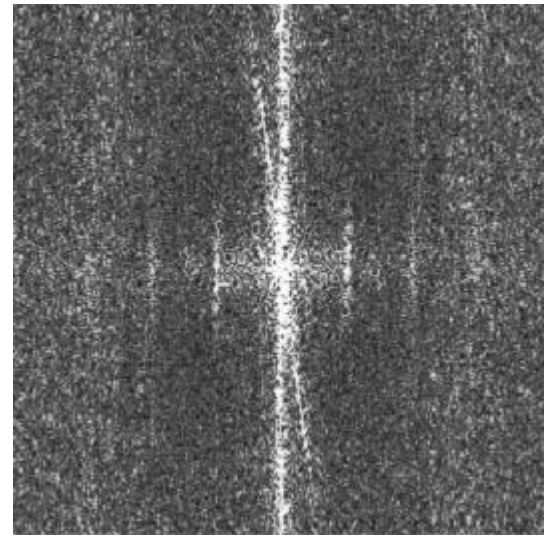
线性叠加及尺度变化



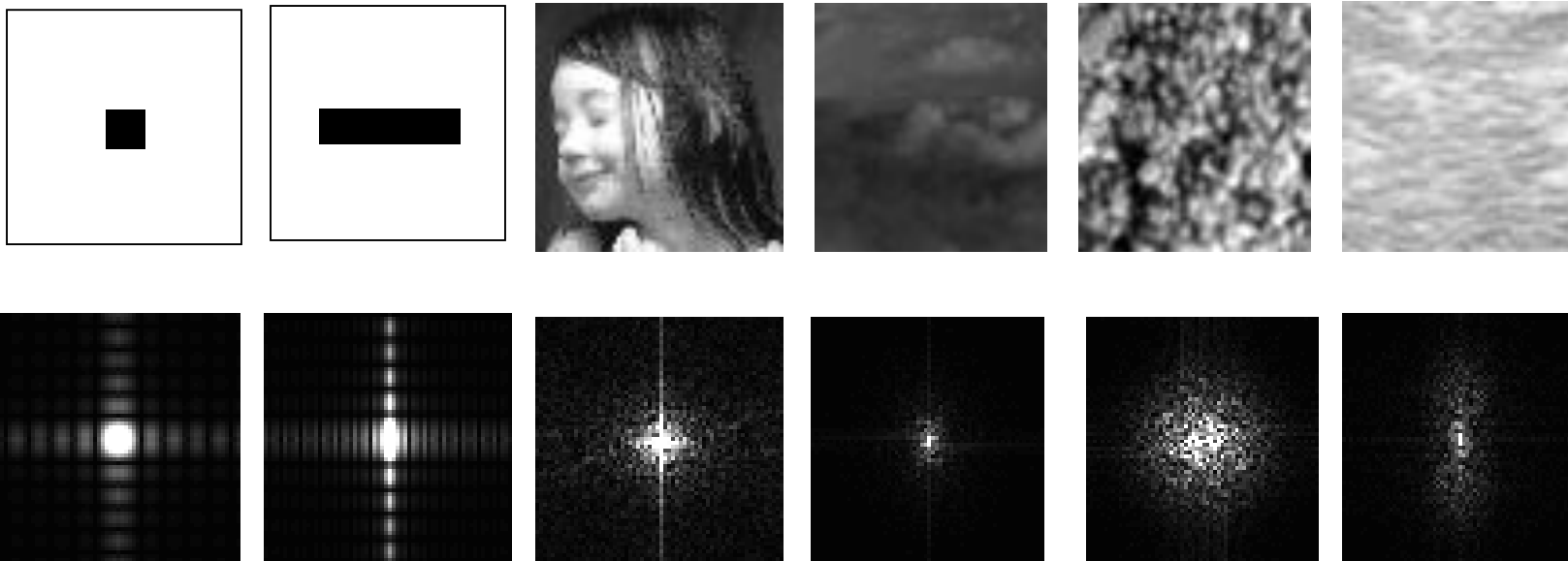
旋转性



实例



典型图象的频谱



离散余弦变换 (DCT)

- 一种可分离、正交、对称的变换
- 1-D离散余弦变换

$$C(u) = a(u) \sum_{x=0}^{N-1} f(x) \cos \left[\frac{(2x+1)u\pi}{2N} \right] \quad u = 0, 1, \Lambda, N-1$$

$$f(x) = \sum_{u=0}^{N-1} a(u) C(u) \cos \left[\frac{(2x+1)u\pi}{2N} \right] \quad x = 0, 1, \Lambda, N-1$$

$$a(u) = \begin{cases} \sqrt{1/N} & \text{当 } u = 0 \\ \sqrt{2/N} & \text{当 } u = 1, 2, \Lambda, N-1 \end{cases}$$



离散余弦变换 (DCT)

□ 2-D离散余弦变换

$$C(u, v) = a(u)a(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos \left[\frac{(2x+1)u\pi}{2N} \right] \cos \left[\frac{(2y+1)v\pi}{2N} \right]$$

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} a(u)a(v) C(u, v) \cos \left[\frac{(2x+1)u\pi}{2N} \right] \cos \left[\frac{(2y+1)v\pi}{2N} \right]$$

讨论可分离性和对称性

$$h(x, y, u, v) = h_1(x, u)h_2(y, v) \quad h(x, y, u, v) = h_1(x, u)h_1(y, v)$$



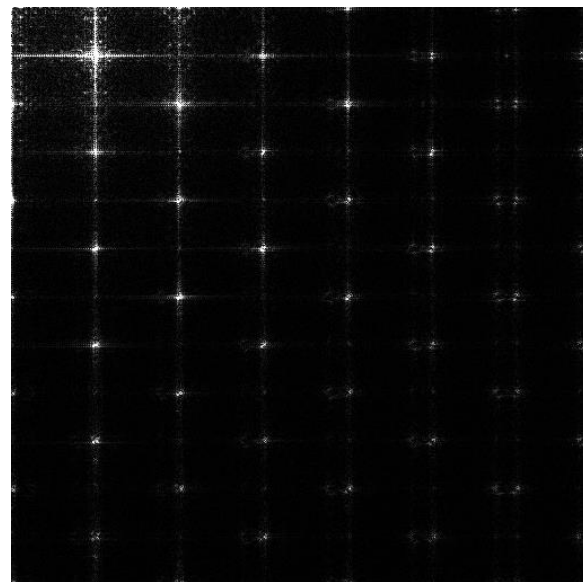
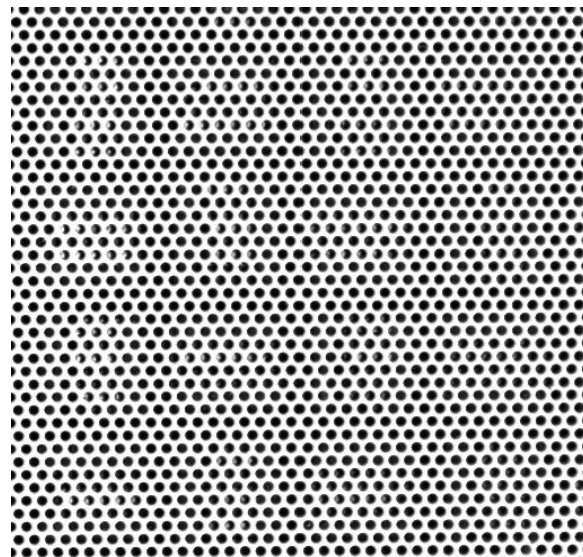
DCT的性质

- 实规范正交基
- 与DFT的关系
 - DCT对应实偶函数的DFT
- 有快速算法
- 能量压缩
 - 应用于JPEG压缩编码

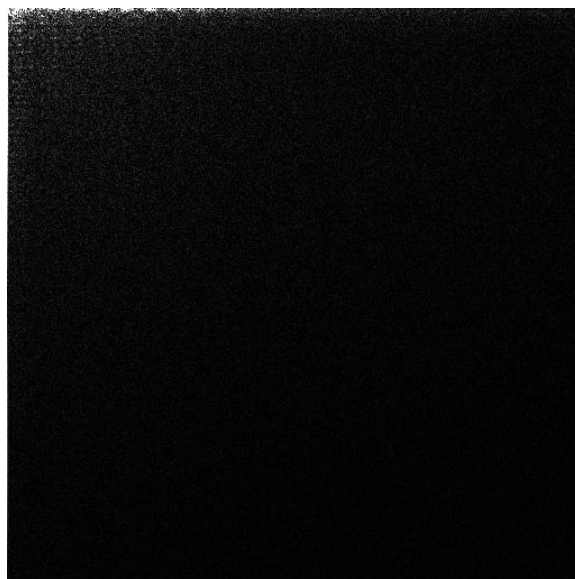
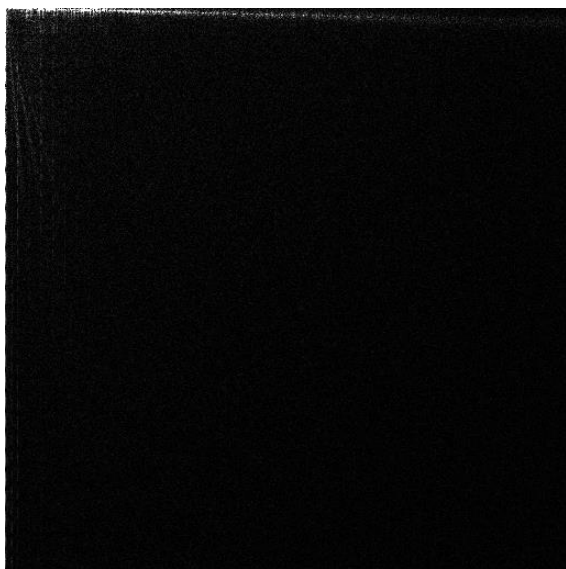
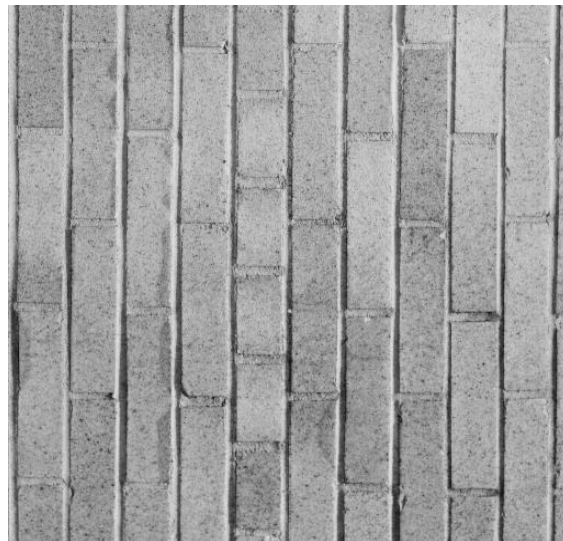
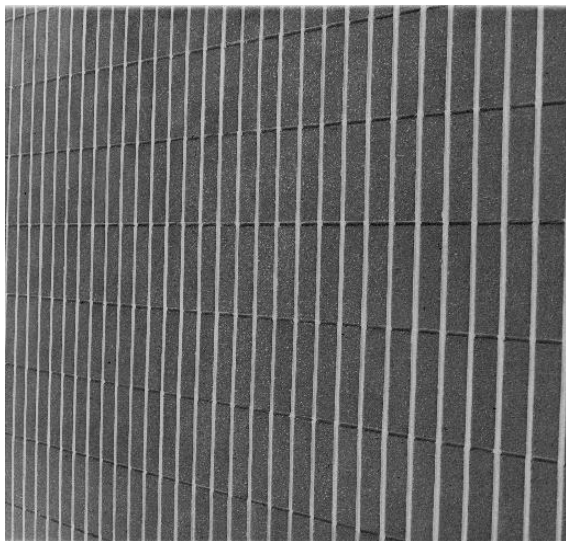
DCT变换结果示例



DCT变换结果示例



DCT变换结果示例





Karhunen-Loeve变换 (PCA)

Principle Component Analysis

- 对于零均值的 $f(x, y)$ ，对其自相关函数 R_{ff} ，可以找到一组正交变换基 ϕ_{mn} ， ϕ_{mn} 满足积分方程：

$$\iint R_{ff}(x, y, u, v) \phi_{mn}(u, v) du dv = \gamma_{mn} \phi_{mn}(x, y)$$

其中 $R_{ff}(x, y, u, v) = \text{cov}(f(x, y), f(u, v))$

- 用这样的正交基系列进行变换，可使变换后**完全去相关**
 - 该变换的**变换核**要根据原信号的统计性质求出
- KL变换是建立在统计特性基础上的一种变换
 - 一种特征提取方法
 - 最小均方误差意义下的最优正交变换
 - 在消除模式特征之间的相关性、突出差异性方面有最优的效果



协方差矩阵及优化目标

- 我们希望原始数据在新的基上的坐标尽量分散，这样可以使得数据便于区分。即：数据在每个基上的坐标分布的方差尽可能的大
- 我们希望原始数据的新的表达在每个维度上不存在（线性）相关性，因为相关性意味着数据的不同维度间不完全独立，就必然存在重复表示的信息。即：数据的不同维度的协方差为0

我们希望由新的基所得到的数据表达的协方差矩阵中，除对角线上的方差元素外，其余所有的协方差元素全部为0（矩阵对角化）。

协方差矩阵及优化目标

□ 设原始数据为M个N维向量，首先将数据每个维度减去各自维度的均值，使每个维度的均值都变为0，记为矩阵X（每一列对应一个样本向量）

□ 基变换矩阵记为矩阵P，则基变换后的数据可以记为：

$$Y = PX$$

□ 显然，Y每个维度的均值也为0。因此Y的协方差矩阵为：

$$\begin{aligned} D_Y &= \frac{1}{M} Y Y^T \\ &= \frac{1}{M} (PX)(PX)^T \\ &= \frac{1}{M} P X X^T P^T \\ &= P \left(\frac{1}{M} X X^T \right) P^T \\ &= P D_X P^T \end{aligned}$$

目标变换矩阵P：
能让原始数据协
方差矩阵对角化



协方差矩阵及优化目标

□ 我们知道：

- 协方差矩阵 D_X 是一个实对称矩阵
- 实对称矩阵不同特征值对应的特征向量必然正交

□ P是协方差矩阵 D_X 的特征向量单位化后按行排列出的矩阵，其中每一行都是 D_X 的一个特征向量

□ 如果P按照特征值从大到小，将特征向量从上到下排列，则用P的前k行组成的矩阵乘以原始数据矩阵X，就得到了我们需要的降维后的数据矩阵Y

算法步骤

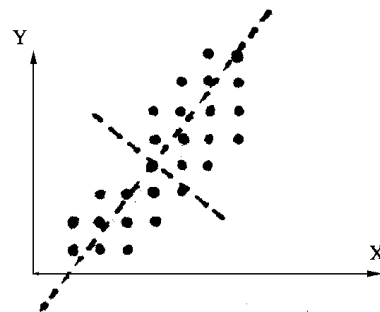
原始数据为M个N维向量：

1. 将原始数据按列组成N行M列的矩阵X；
2. 将X的每一行（每个维度）进行零均值化；
3. 求出协方差矩阵 $D_X = \frac{1}{M}XX^T$ ；
4. 求出 D_X 的特征值及对应的特征向量；
5. 将特征向量按对应特征值大小从上到下按行排列成矩阵，取前k行组成矩阵P；
6. $Y = PX$ 即为降维到k维后的数据。

K-L变换一例(1)

图象点序列:

(1, 1), (1, 2),
(2, 1), (2, 2), (2, 3),
(3, 1), (3, 2), (3, 3),
(4, 2), (4, 3), (4, 4), (4, 5), (4, 6),
(5, 3), (5, 4), (5, 5), (5, 6), (5, 7),
(6, 4), (6, 5), (6, 6), (6, 7), (6, 8),
(7, 5), (7, 6), (7, 7), (7, 8)



均值
$$\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix} = E \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4.4 \\ 4.3 \end{bmatrix}$$

协方差
$$\Sigma_x = E \left\{ \begin{bmatrix} x - \bar{x} \\ y - \bar{y} \end{bmatrix} \begin{bmatrix} x - \bar{x} & y - \bar{y} \end{bmatrix} \right\} = \begin{bmatrix} 3.286 & 3.099 \\ 3.099 & 4.579 \end{bmatrix}$$

K-L变换一例(2)

- 在维数小时，由本征多项式为零求协方差矩阵的本征值：

$$\begin{bmatrix} 3.268 - \lambda & 3.099 \\ 3.099 & 4.579 - \lambda \end{bmatrix} = \lambda^2 - 7.865\lambda + 5.443 = 0$$

$$\lambda_1 = 7.098 \quad \lambda_2 = 0.768$$

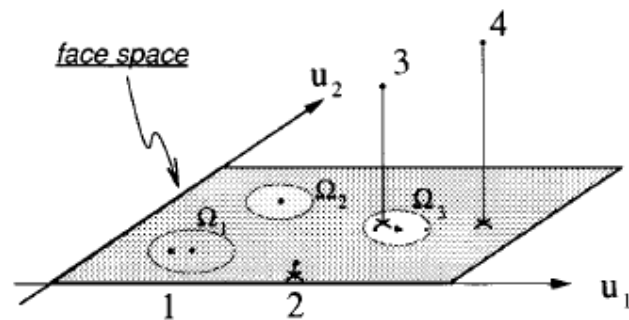
- 再把本征值代入 $\sum_X \phi_i^{\omega} = \lambda_i \phi_i^{\omega}$ ，求出特征矢量：

$$\phi_1^{\rho} = \begin{bmatrix} 1 \\ 1.23 \end{bmatrix} \quad \phi_2^{\rho} = \begin{bmatrix} -1.23 \\ 1 \end{bmatrix}$$

- 把相互垂直的二特征矢量作为新的坐标，新坐标的主轴方向为所变换数据方差最大的方向

KL变换应用实例 —— 人脸识别

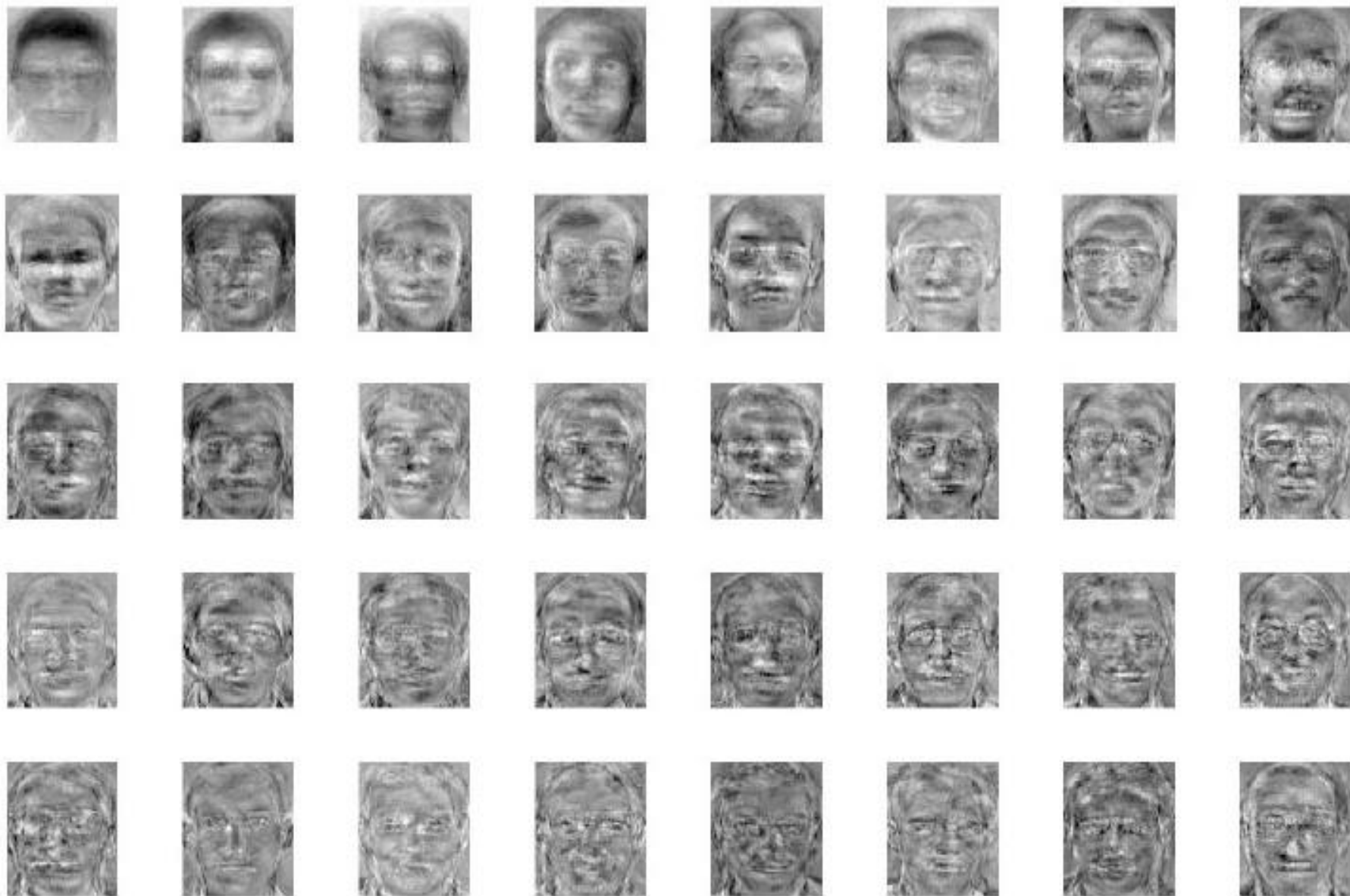
1. 把每一幅人脸列化，视为随机向量 F 的不同实现。
2. 估计 F 协方差矩阵 C ，并计算其特征值特征向量。
(C 是半正定矩阵，维数不大于图像数，对应不同特征值的特征向量正交，特征脸)
3. 选择对应少量最大特征值的特征向量组成特征脸空间
4. 每张脸映射为特征脸空间的点，以其坐标作为特征向量。
5. 采用模式识别方法，进行分类识别。
(如欧式距离)



人脸库



特征脸



特征脸空间 (top 8)





小波变换 (DWT)

□ 傅立叶和小波

- 傅立叶变换在分析瞬态信号时仍有不足
- 小波是一组衰减的波

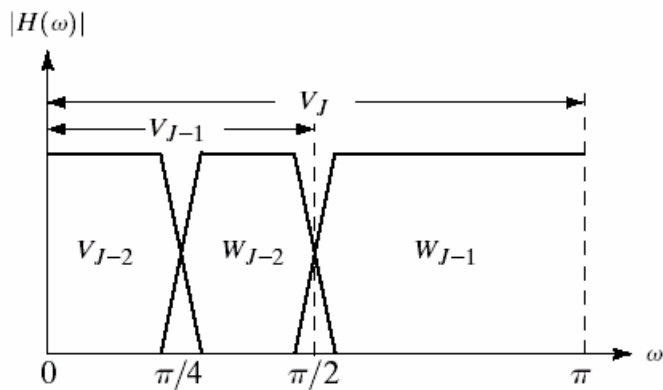
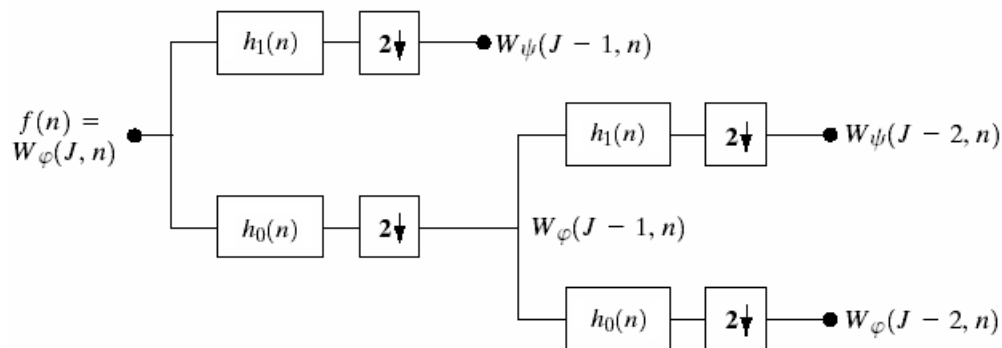
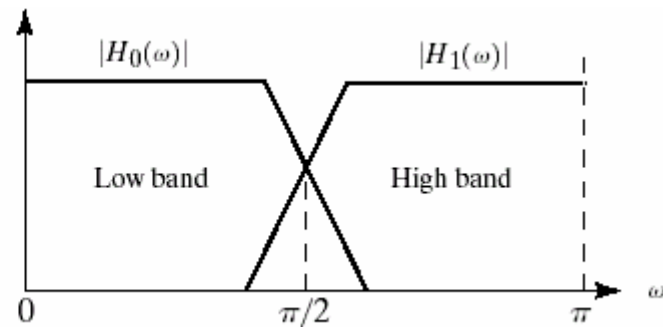
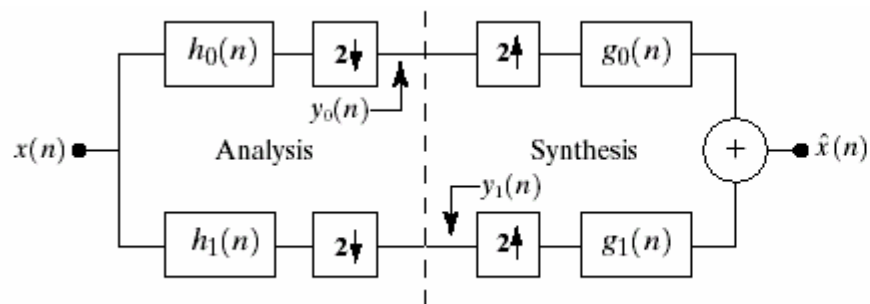
□ 时频域分析

- 小波变换在二维时频空间分析信号

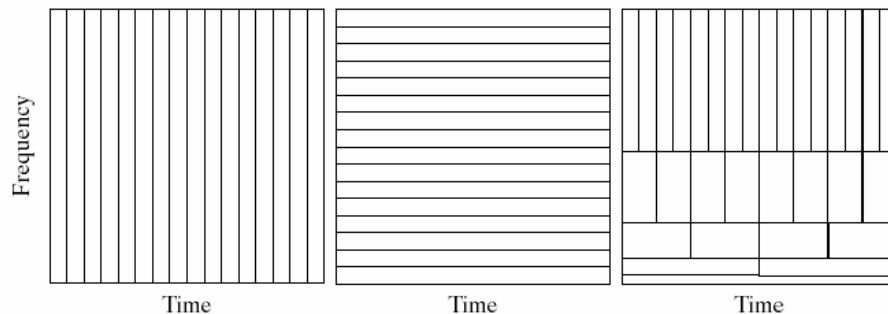
□ 变换

- 理想的基本小波是过程很短的振荡函数
- 如同傅立叶变换有连续、离散的变换，小波也有连续、离散的变换

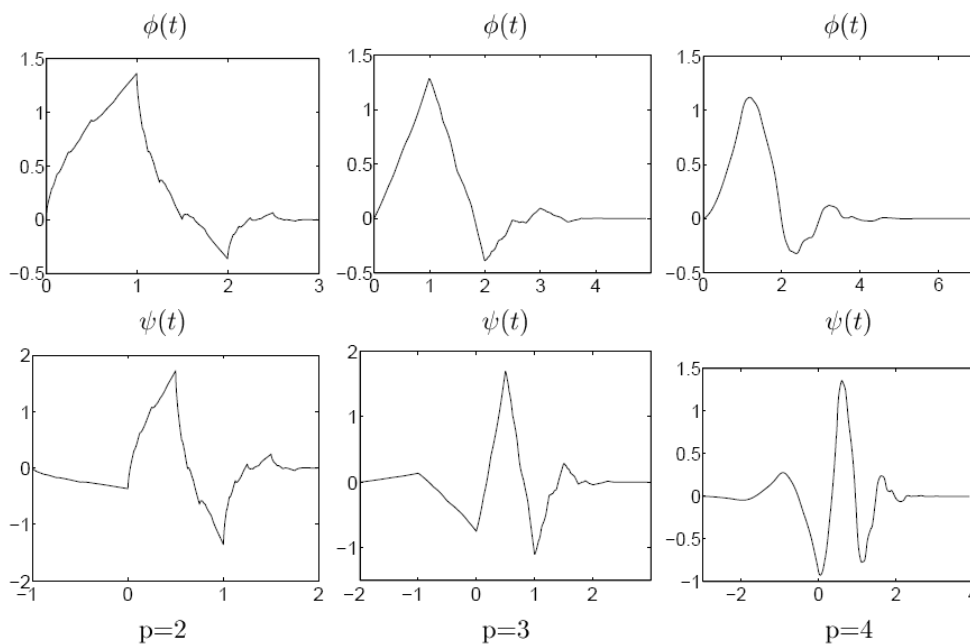
1维离散小波变换



1维离散小波变换

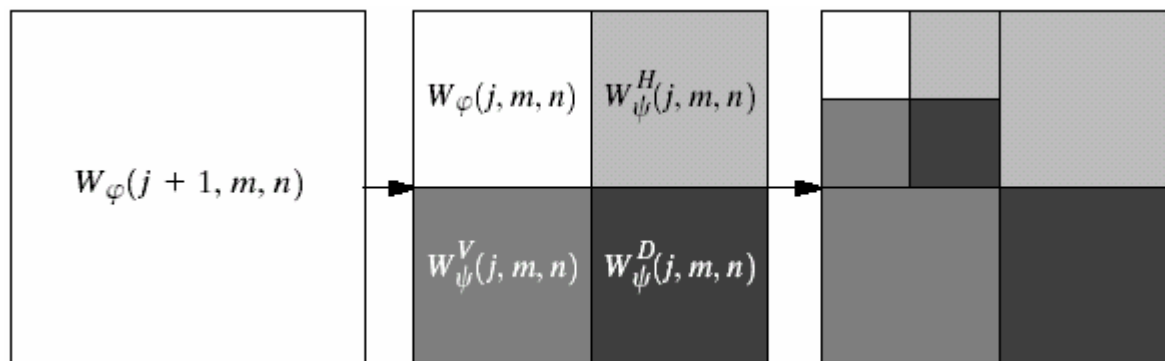
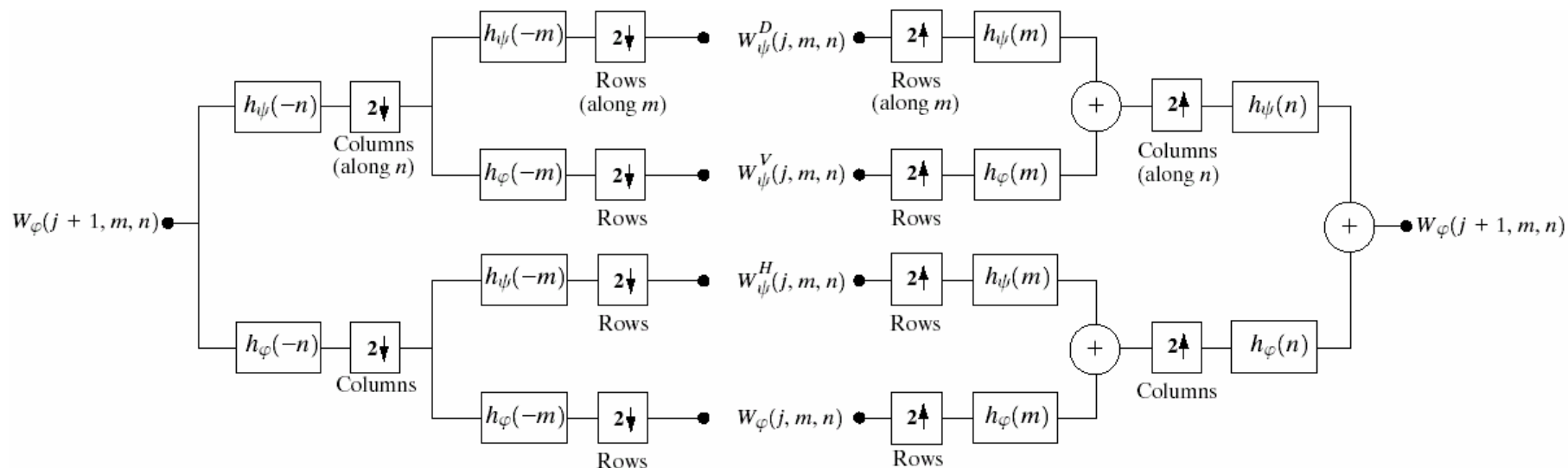


时频铺叠（从左到右：Dirac、Fourier、wavelet）



Daubechies 小波及尺度函数。（ p 为消失矩 Vanishing Moment）

2维离散小波变换



实例

