

# 计算机视觉

## 1 绪论

了解计算机视觉基本概念

### 1.1 视觉

视觉进一步可分为视感觉和视知觉

1. 视感觉主要与成像相关
2. 视知觉主要与成像后得到的图像或图像序列（视频）的处理、分析与理解相关

计算机视觉是指用计算机实现人类的视觉功能

- 研究方法

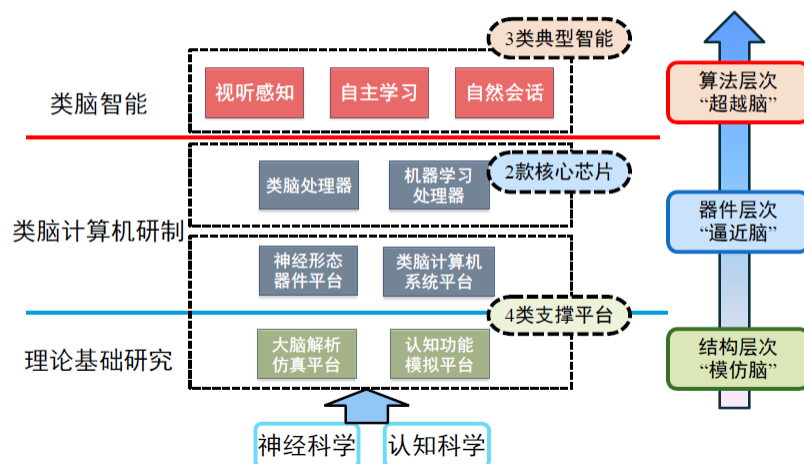
1. 仿生学的方法：参照人类视觉系统的结构原理
2. 工程的方法：实现系统的功能

- 研究目标

1. 建立计算机视觉系统来完成各种视觉任务
2. 加深对人脑视觉机理的掌握和理解

### 1.2 概述 ★

#### “脑认知与类脑计算”任务规划



### 1.3 应用实例

- 结冰测量

1. 打开第一帧图像
2. 人工标记迎风面
3. 调整迎风面
  - (a) 精确确定迎风面
  - (b) 边缘检测
4. 人工定标
5. 设置配准区域
  - (a) 克服部件运动的影响
  - (b) 相关匹配配准
6. 测量
  - (a) 冰层边缘检测

- 公交车客流统计

1. 选择正负样本训练集
2. 特征提取与分类训练
  - (a) HOG 特征
  - (b) 线性支撑向量机 SVM
3. 目标检测
  - (a) 多尺度分析
4. 目标跟踪
  - (a) 分析人头目标运动轨迹
  - (b) Camshift 相关匹配
5. 计数

## 1.4 图象

**图像：** 反射/辐射强度模式的空间分布

**图像表达函数：** 辐射能量在空间分布的函数

**通用图像表达函数：**  $T(x, y, z, t, l)$

**模拟图像：** 从连续的客观场景直接观察到

用一个 2D 数组  $f(xy)$  来表示， $f(xy)$  的值可以是任意实数

**数字图像：**把连续的模拟图像在坐标空间 XY 和性质空间 F 都离散化了的图像  
用  $f(xy)$  代表数字图像， $f(xy)$  都为整数

## 1.5 图像表达

- 矩阵表达

$$\mathbf{F} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1N} \\ f_{21} & f_{22} & \cdots & f_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ f_{M1} & f_{M2} & \cdots & f_{MN} \end{bmatrix}$$

- 矢量表达

$$\mathbf{F} = [\mathbf{f}_1 \quad \mathbf{f}_2 \quad \cdots \quad \mathbf{f}_N]$$

$$\mathbf{f}_i = [f_{1i} \quad f_{2i} \quad \cdots \quad f_{Mi}]^T \quad i = 1, 2, \cdots, N$$

## 1.6 图像显示

1. 离散点集
2. 覆盖区域
3. 矩阵表达

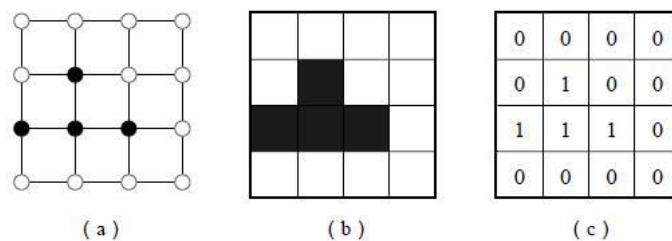


图 1.2.1 3 种表达同一幅  $4 \times 4$  的二值图像的方式

## 1.7 像素邻域

- 4 - 邻域:  $N_4(p)$
- 对角邻域:  $N_D(p)$
- 8 - 邻域:  $N_8(p)$

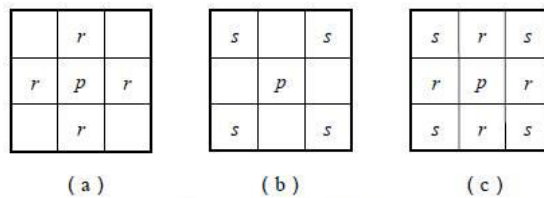


图 1.3.1 像素的邻域

### 三种像素位置

#### 邻接:

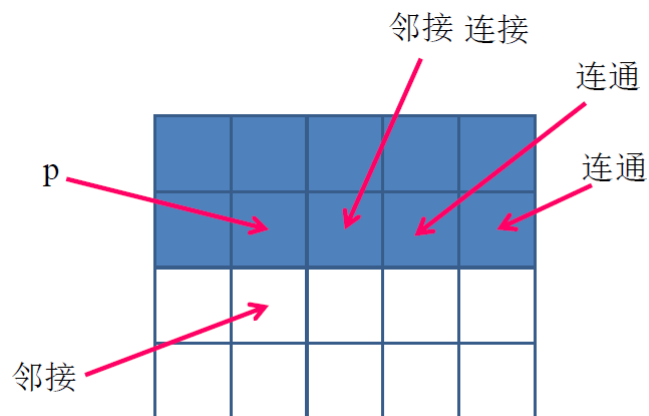
对两个像素  $p$  和  $q$  来说, 如果  $q$  在  $p$  的邻域中, 则称  $p$  和  $q$  满足邻接关系 (空间)

#### 连接:

$p$  和  $q$  邻接且灰度值均满足某个特定的相似准则 (空间+灰度)

#### 连通:

不 (直接) 邻接, 但均在另一个像素的相同邻域中, 且这 3 个像素的灰度值均满足某个特定的相似准则 (空间+灰度+空间传递)



二值图像, 4-邻域

## 1.8 像素间距离

### 距离与范数

$$D_w(p, q) = \left[ |x - s|^w + |y - t|^w \right]^{1/w}$$

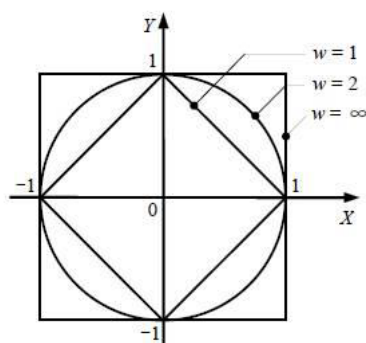


图 1.3.4 3 种范数和 3 种距离

欧氏距离:  $D_E(p, q) = [(x-s)^2 + (y-t)^2]^{1/2}$

城区距离:  $D_4(p, q) = |x-s| + |y-t|$

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

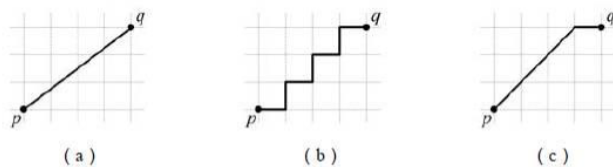
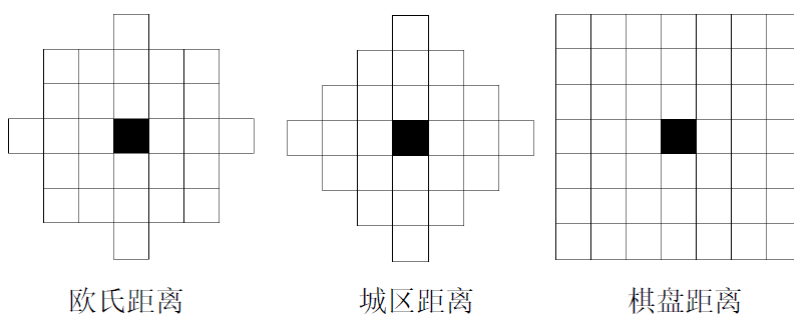


图 1.3.2 像素间距离的计算

欧氏距离

城区距离

棋盘距离



欧氏距离

城区距离

棋盘距离

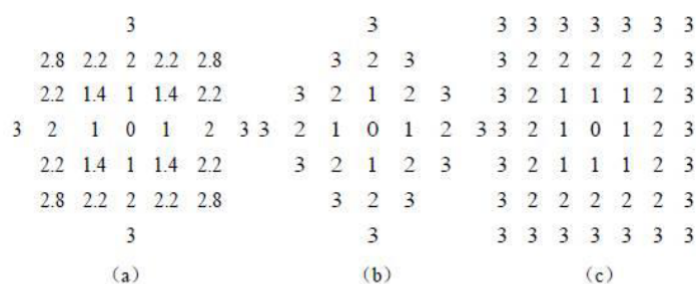


图 1.3.3 等距离轮廓示例

欧氏距离

城区距离

棋盘距离

## 距离和邻域

4-邻域:  $D_4 = 1$  的像素集合

8-邻域:  $D_8 = 1$  的像素集合

**等距离轮廓:** 与中心像素的某种距离等于某个值的像素集合

$\Delta_i(r)$ : 与中心像素的距离小于或等于  $r$  的等距离轮廓

$\# \Delta_i(r)$ : 除中心像素外  $\Delta_i(r)$  所包含像素个数

$$\#[\Delta_4(r)] = 4 \sum_{j=1}^r j$$

$$\#[\Delta_8(r)] = 8 \sum_{j=1}^r j$$

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

## 2 视觉与视知觉

**重点:** 能够理解视觉过程和特点, 以及三种视觉感知, 解释日常生活中的相关现象

### 2.1 形状知觉

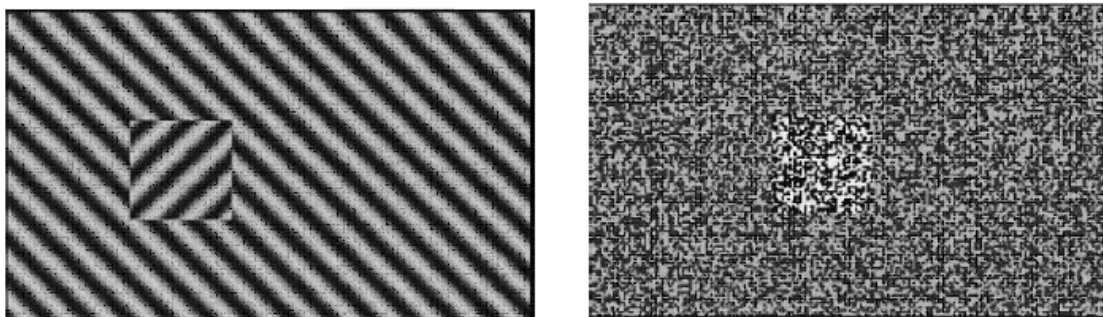
#### 2.1.1 形状的感知

##### 1. 视觉边缘和目标

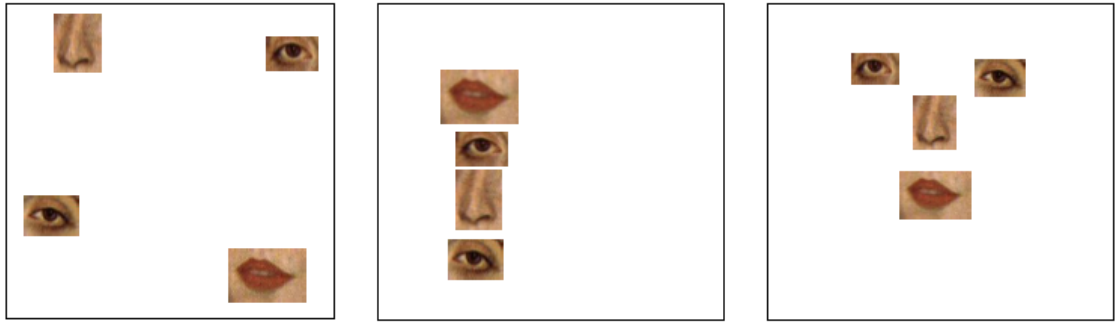
(a) 外部刺激导致的视觉边缘提供形状感知所需信息



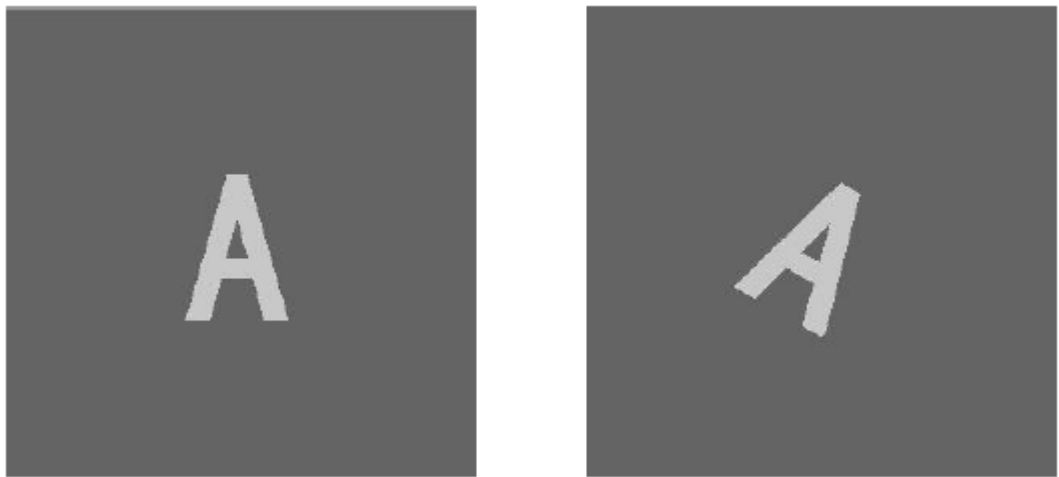
(b) 边缘可以为亮度、颜色、纹理等变化造成



(c) 与目标整体有关



(d) 图形形状不变性：位置、大小、方位、组成单元

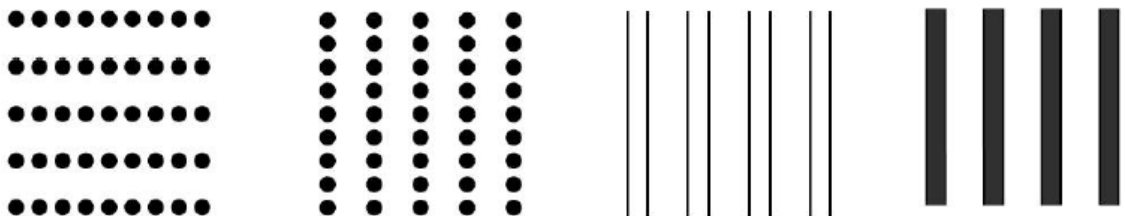


## 2. 前景 (figure) 和背景 (background) 的分离

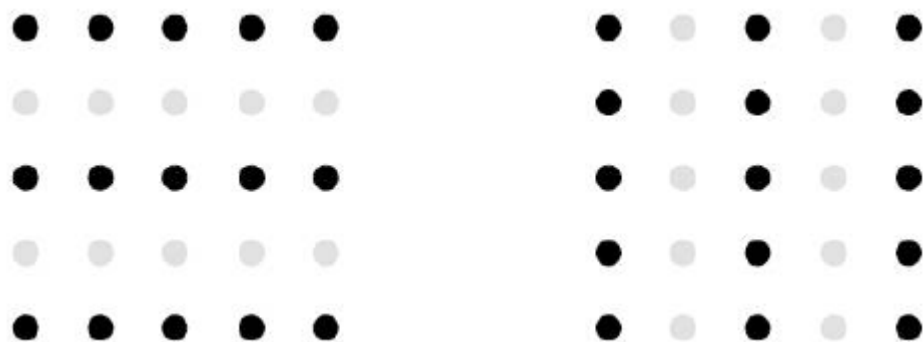
- (a) 前景位于轮廓内部，它具有某种形状
- (b) 背景一般位于轮廓外部，并不具有特定的形状

## 3. 形状构造的规律

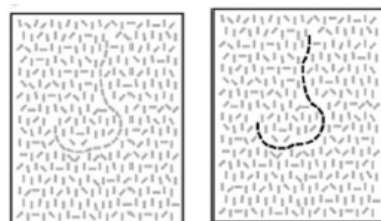
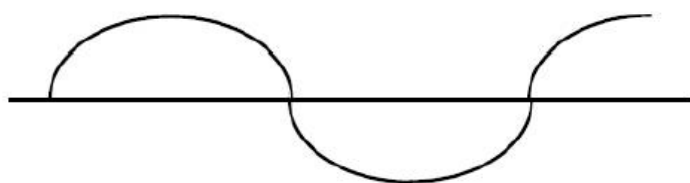
- (a) 接近规则：空间中相接近的元素比相分离的元素更容易被感知为属于共同的形状



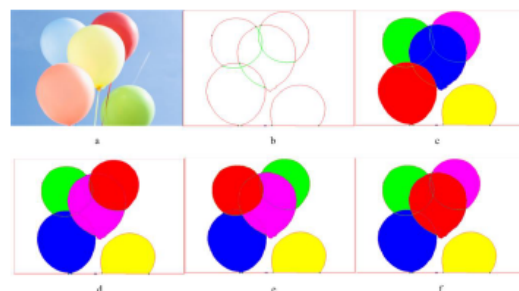
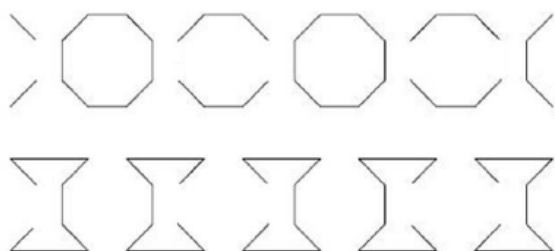
- (b) 相似规则：类似形状或尺寸的元素更容易被感知为属于相似的集体形状



- (c) 连续规则：不完整的形状有被感知为完整形状的趋势（如果一个图形的某些部分可以被看作是连接在一起的，如光滑连接，那么这些部分就相对容易被我们知觉为一个整体。）



- (d) 封闭规则：移动一个形状时，同时移动的元素被感知为属于同一个形状整体（没有闭合的残缺形状有被感知为闭合形状的倾向）



#### 4. 形状和信息

- (a) 前景中的信息冗余：亮度均匀/结构对称  
(b) 最小原理：感知最简单形状

##### 2.1.2 轮廓

**主观轮廓：**在没有直接刺激作用下产生的轮廓知觉。主观轮廓的形成是在定感觉信息的基础上进行知觉假设的结果

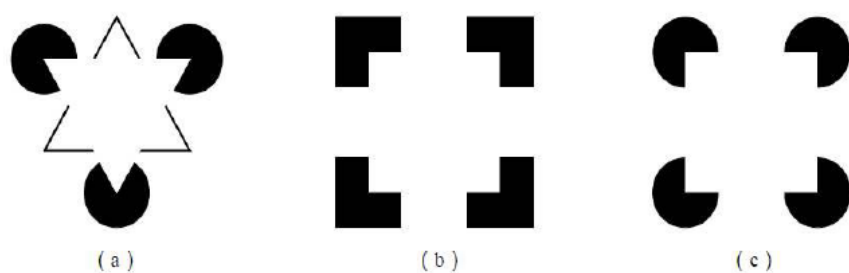


图 2.2.2 主观轮廓示例



## 形成图形和背景的影响因素

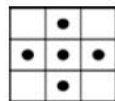
1. 视野中的距离：格式塔原理中的接近规则
2. 相同或相似：格式塔原理中的相似规则
3. 将接近规则和相似规则进行综合

$$w_{ij} = \exp(-\|F_i - F_j\|_2^2 / \sigma_I^2) * \begin{cases} \exp(-\|X_i - X_j\|_2^2 / \sigma_X^2) & \text{if } \|X_i - X_j\|_2 < r \\ 0 & \text{其他} \end{cases}$$

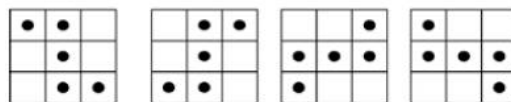
$F$  : 像素灰度值       $X$  : 像素空间位置

### 4. 良好图形

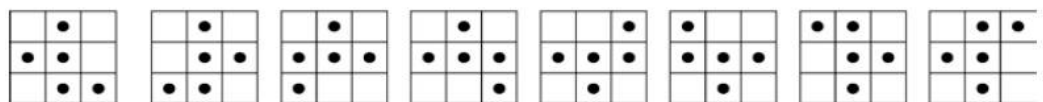
- (a) 有一定的意义
- (b) 刺激性因素
  - i. 封闭（格式塔原理规则之一）
  - ii. 连续（格式塔原理规则之一）
  - iii. 对称
- (c) 非刺激性因素（主观因素）
- (d) “良好性”的度量
  - i. 信息论：冗余图形，可从局部预测整体
  - ii. 可变图形的数量



(a)



(b)



(c)

## 2.2 空间知觉

### 2.2.1 双目深度线索

# Geometry

The diagram illustrates the geometry of stereo vision. A point  $p$  in the scene is projected onto two image planes. The left image plane is at a distance  $f$  from the left camera center  $O_l$ . The right image plane is at a distance  $f$  from the right camera center  $O_r$ . The horizontal distance between the camera centers is  $T$ . The projections of  $p$  on the left and right image planes are  $p_l$  and  $p_r$ , respectively. The vertical distance from the baseline to the point  $p$  is  $Z$ . The coordinates of the projections are  $x_l$  and  $x_r$ .

<http://www.cse.psu.edu/~zyin/Demo/Stereo%20geometry.jpg>

### 随机点立体图

1	0	1	0	0	1	0	1	0	1	0
0	1	0	1	1	0	1	0	1	0	1
1	0	0	1	0	1	0	0	1	0	1
0	0	1	Y	A	A	B	B	0	1	0
0	1	0	X	A	B	A	B	0	0	1
1	0	1	Y	B	A	B	A	1	0	0
0	1	0	X	B	B	A	A	0	1	0
1	0	0	1	0	1	0	0	1	0	1
1	0	1	0	0	0	1	1	0	0	0
0	1	0	1	1	0	0	1	0	1	0

1	0	1	0	0	1	0	1	0	1	0
0	1	0	1	1	0	1	0	1	0	1
1	0	0	1	0	1	0	0	1	0	1
0	0	1	A	A	B	B	X	0	1	0
0	1	0	A	B	A	B	Y	0	0	1
1	0	1	B	A	B	A	X	1	0	0
0	1	0	B	B	A	A	Y	0	1	0
1	0	0	1	0	1	0	0	1	0	1
1	0	1	0	0	0	1	1	0	0	0
0	1	0	1	1	0	0	1	0	1	0

图 2.3.3 随机点立体图示例

### 2.2.2 单目深度线索

### 知觉深度和距离的线索(举例说明)

1. 大小和距离
2. 照明的变化
3. 线性透视
4. 纹理梯度
5. 物体的遮挡

## 2.3 运动知觉★

运动速度的上下限是以下变量的函数

1. 物体的尺寸
2. 亮度和反差
3. 环境，运动的感知有一定的相对性

**表观运动：**在一定条件下，当实际中没有景物运动时也可能感知到运动

**眼睛的运动：**感知在一定程度上依赖这种运动，包括急速运动、跟踪运动、补偿运动、漂移运动等

**运动感知理论：**

- 1、一阶运动系统主要感知亮度的运动，即所谓的一阶运动
- 2、二阶运动系统主要感知由对比度、空间频率、闪烁频率等特征定义的二阶运动
- 3、三阶运动系统主要感知由显著性特征（根据输入景象的显著性图计算）定义的三阶运动