

数据挖掘与机器学习

潘斌

panbin@nankai.edu.cn

范孙楼227

1

上节回顾

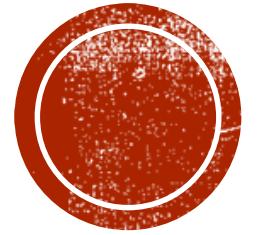
- 概念学习的求解方法
 - Find-S
 - 列表后消除算法
 - 候选消除算法
 - 奥卡姆剃刀原理和没有免费午餐原理



本节提要

- 分类的基本概念
- 特征





分 类

Classification

解决分类问题的一般方法

- 使用一种学习算法 (learning algorithm) 确定分类模型
 - 拟合输入数据中的类标号和属性集之间的联系
 - 正确预测未知样本的类标号
 - 决策树、基于规则的分类法、KNN、神经网络、支持向量机、贝叶斯分类.....
- 提取有效的样本特征



深入理解 MACHINE LEARNING

≈ LOOKING FOR A FUNCTION

- Speech Recognition

$$f(\text{[sound波形图]}) = \text{"How are you"}$$

- Image Recognition

$$f(\text{[猫的照片]}) = \text{"Cat"}$$



- Playing Go

$$f(\text{[围棋棋盘]}) = \text{"5-5"} \quad (\text{next move})$$



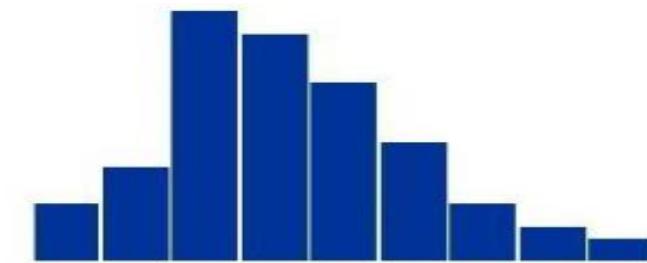
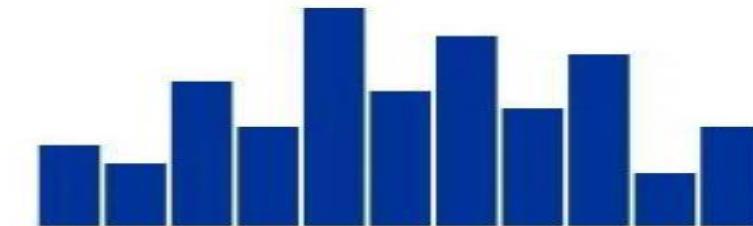
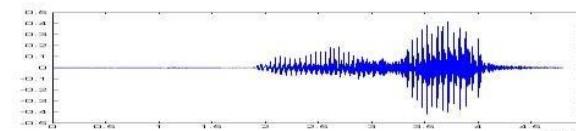
- Dialogue System

$$f(\text{"Hi"}) = \text{"Hello"}$$

(what the user said) (system response)



特征描述



“Hi”

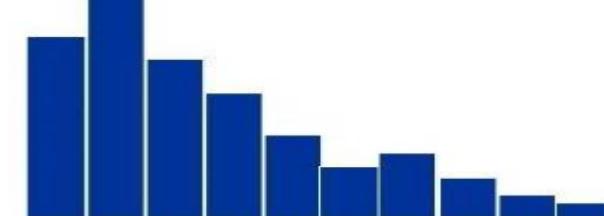
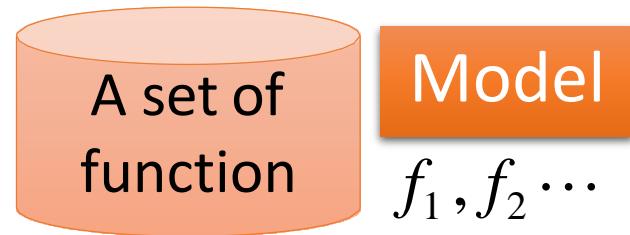


Image Recognition:

定义映射集合

$$f(\text{ }) = \text{ "cat"}$$

$$f_1(\text{ }) = \text{ "cat"}$$


$$f_2(\text{ }) = \text{ "money"}$$


$$f_1(\text{ }) = \text{ "dog"}$$

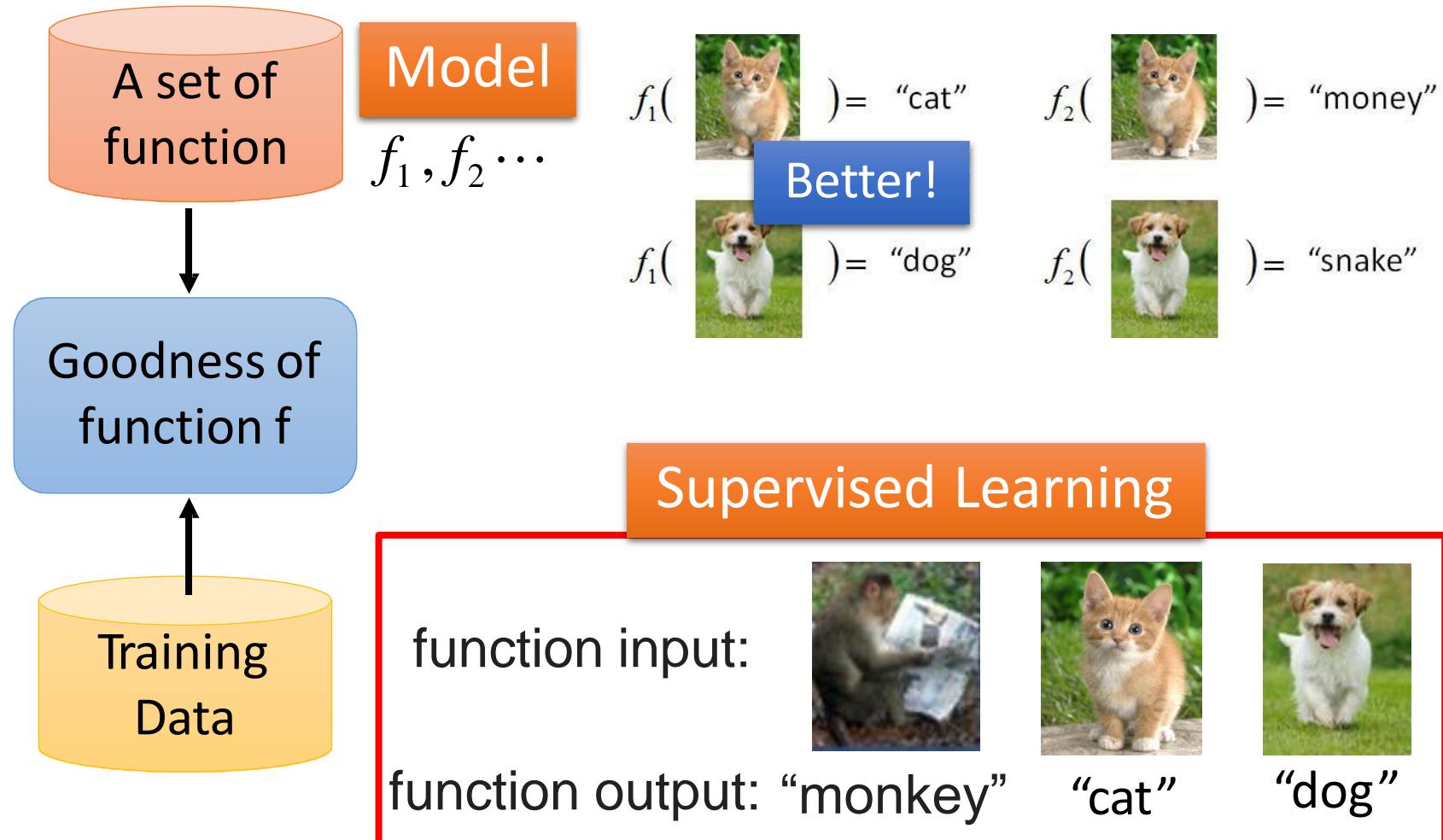

$$f_2(\text{ }) = \text{ "snake"}$$




确定最优准则

Image Recognition:

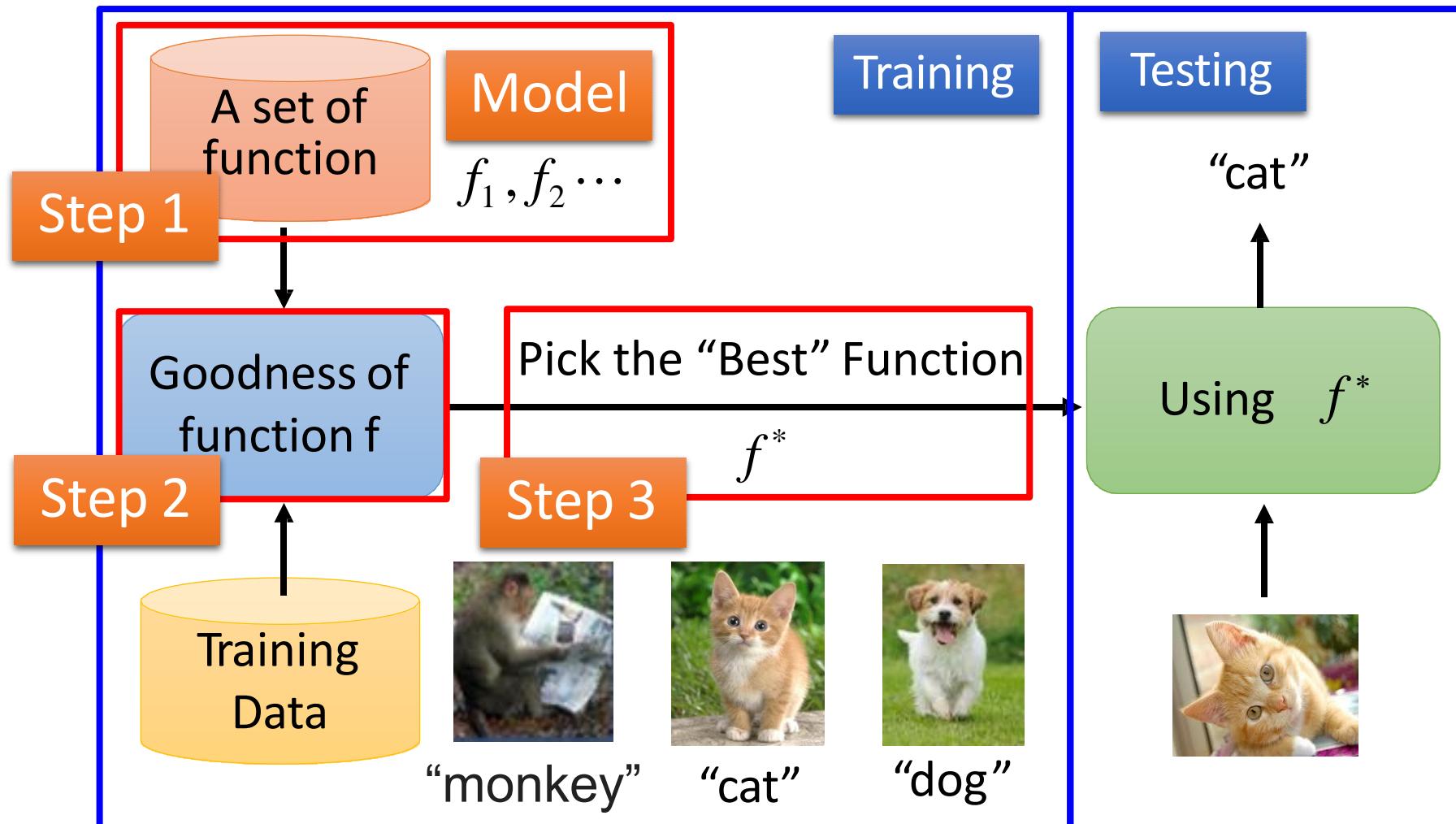
$$f(\text{cat}) = \text{"cat"}$$



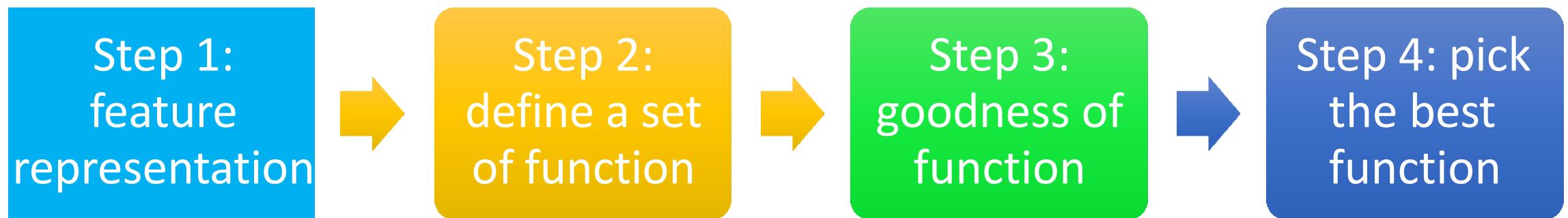
选择最优映射

Image Recognition:

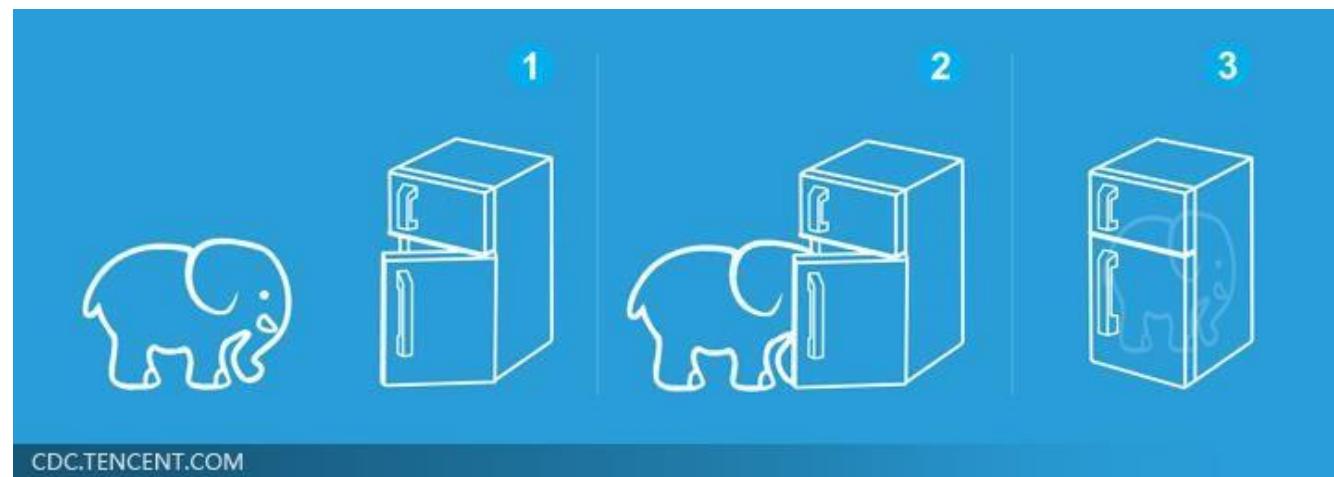
$$f(\text{cat}) = \text{"cat"}$$

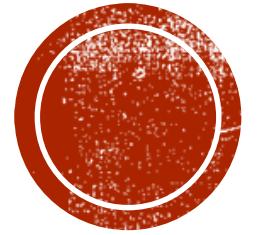


机器学习的4个步骤



Machine Learning is so simple





特征

Feature

§ 1.1 从图像到图像处理

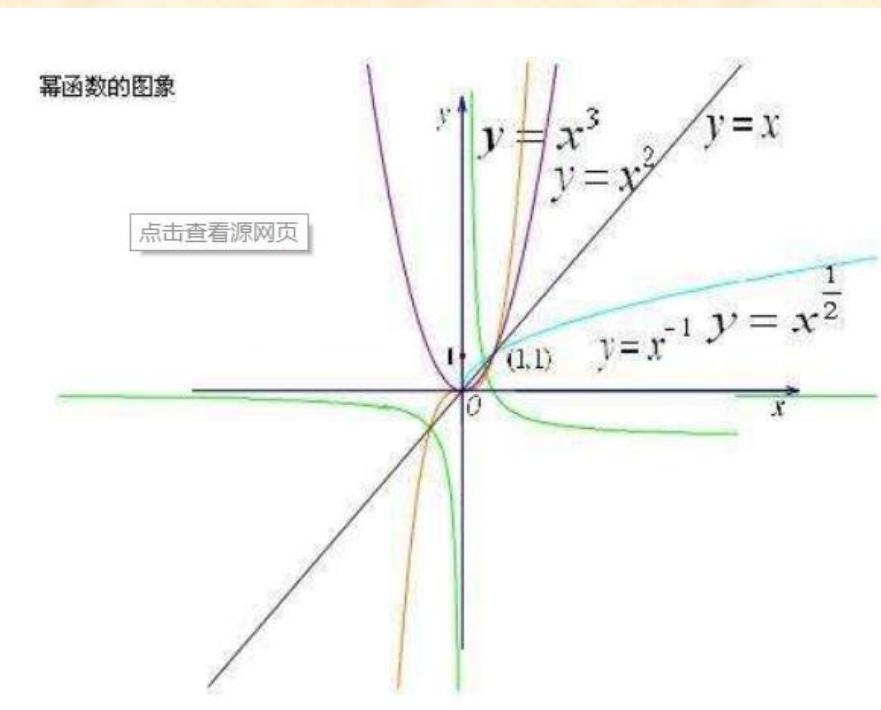
■ 图像和数字图像

■ 图象：

- 用各种观测系统以不同形式和手段观测客观世界而获得的，可以直接或间接作用于人眼并进而产生视觉的实体（照片）
- 人类从外界（客观世界）获得的信息约有75%来自视觉系统
- 图象（广义/抽象） ⊨ 图像（狭义/具体）

§ 1.1 从图像到图像处理

■ 图像和图象



仅图象



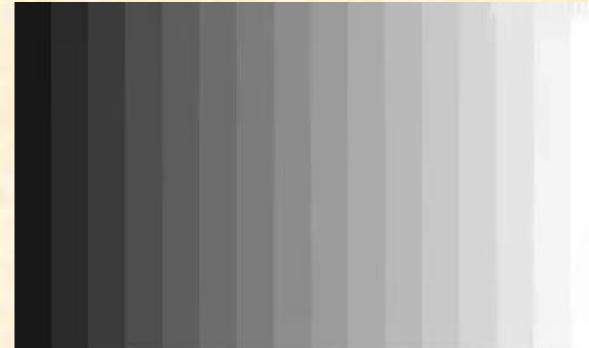
图像或图象

§ 1.1 从图像到图像处理

■ 数字图像概念：由像素组成的矩阵

- 像素：矩阵中的一个元素，像素的值代表该位置的亮度，被称为“灰度值”，**灰度值越高这个位置越亮**（接近白色），反之越暗（接近黑色）
- 1080P：图像的短边有1080个像素。常见的手机屏幕包含**1920*1080**个像素

0	1	2	4	8	16	32	128
1	2	8	16	32	64	128	192
2	8	8	32	32	100	150	208
4	16	32	32	100	150	175	240
8	32	32	100	128	175	200	248
16	64	100	150	175	200	224	252
32	128	150	175	200	224	250	254
128	192	208	240	248	252	254	255



§ 1.1 从图像到图像处理

■ 数字图像的数学表示: $f(x,y,\lambda,t)$

- x, y : 2-D空间中坐标点的位置
 - λ : 光线的波长
 - t : 时间
- 其中最基本的形式是静态图像, 视觉上就是黑白深浅的区别, 表示为 $f(x,y)$, 称为“灰度图”
 - 加入波长, 视觉上体现为颜色信息, 表示为 $f(x,y,\lambda)$, 称为“多光谱图”, 若仅考虑3个可见光波段, 就是我们常见的彩图
 - 进一步加入时间信息, 视觉上体现为“视频”



§ 1.1 从图像到图像处理

- 对于静态的彩色图像
 - 用红绿蓝三基色表示，图像是一个3个2-D数组 $f_R(x, y), f_G(x, y), f_B(x, y)$
 - 如果是多光谱图像，图像是一个n个2-D数组
- 对于动态的连续图像，可以是一个时间序列
 - 黑白电视信号 $f_1(x, y), f_2(x, y), \dots, f_k(x, y)$
 - 彩色电视信号 $f_{Rk}(x, y), f_{Gk}(x, y), f_{Bk}(x, y), k=1, 2, \dots$

§ 1.1 从图像到图像处理

■ 图像处理

■ 图像处理

□ 图像处理在广义上是各种与图像有关的技术的总称

□ 主要功能/作用包括：

- 对图像的各种加工（见下）
- 基于加工结果的判断决策和行为规划
- 为此进行的硬件设计及制作

■ 图像加工技术

□ 例如：图像的采集、获取、编码、存储和传输，合成和产生，显示和输出，变换、增强、恢复和重建，分割，目标的检测、表达和描述，特征的提取和测量，序列图象的校正，**3-D**景物的重建复原，图像数据库的建立、索引和抽取，图像的分类、表示和识别，图像模型的建立和匹配，图像和场景的解释和理解，.....

§ 1.1 从图像到图像处理

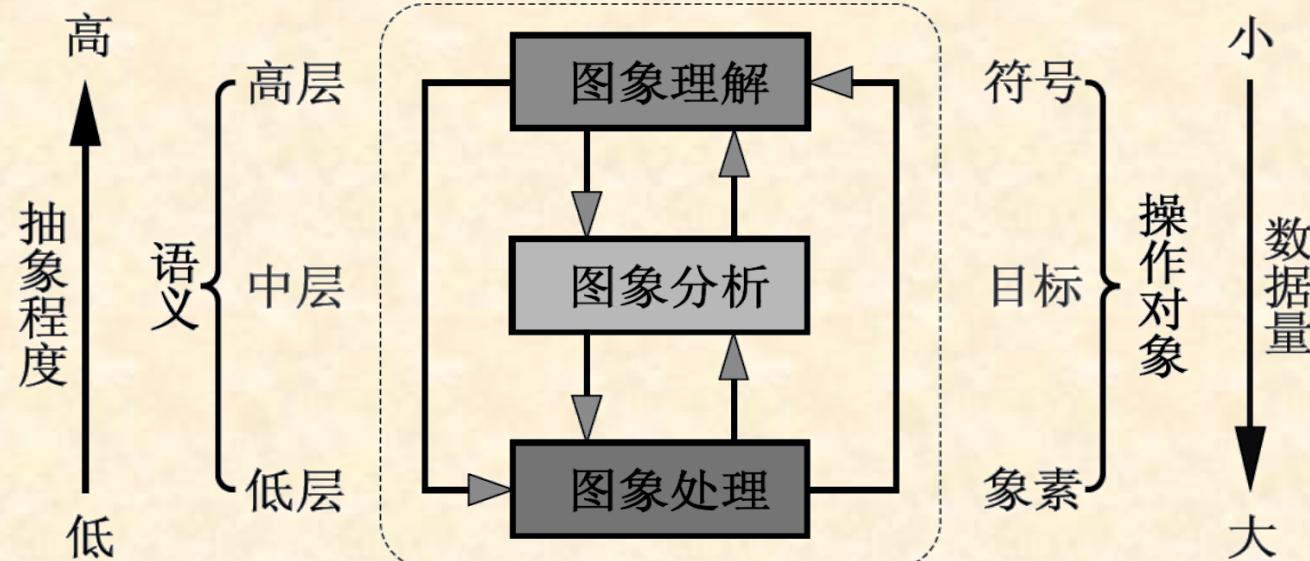
■ 图像处理应用领域示例

- 视频通信：可视电话，电视会议，按需电视，远程教育；
- 文字档案：文字识别，过期档案复原，邮件分拣，支票，签名辨伪，办公自动化；
- 生物医学：红白学球计数，染色体分析、**X光**、**CT**、**MRI**、**PET**图象分析，医学手术模拟规划，远程医疗；
- 遥感测绘：巡航导弹制导，无人驾驶飞机飞行，精确制导，航天侦查，资源探测，气象预报，自然灾害监测；
- 工业生产：工业检测，工业探伤，自动生产流水线监控，移动机器人，无损探测，金相分析，印刷板质量检验，精细印刷品缺陷检测；
- 军事公安：雷达图像分析、巡航导弹路径规划 / 制导，罪犯脸形合成、识别，指纹、印章的鉴定识别；
- 交通管理：太空探测、航天飞行、公路交通管理。

§ 1.1 从图像到图像处理

■ 图像处理三层次：

- 图像预处理（图像——>图像）
- 图像分析（图像——> 数据）
- 图像理解（图像——> 解释）



§ 1.1 从图像到图像处理

■ 图像处理三层次：

□ 图像预处理（狭义图像处理）



§ 1.1 从图像到图像处理

■ 图像处理三层次：

□ 图像分析（机器学习）



帽子

人脸

长发

§ 1.1 从图像到图像处理

■ 图像处理三层次：

□ 图像理解（机器学习）



帽子：时尚

人脸：容光焕发

长发：女性标志

图像理解：炯炯有神的美女模特

§ 1.2 图像处理的基本操作

■ 图像增强

- 改善图像视觉效果，增强图像的有用信息，消弱噪声的干扰



去雾



去模糊



§ 1.2 图像处理的基本操作

■ 图像恢复与重建

■ 恢复图像原本面貌，辐射校正、几何校正



几何校正

§ 1.2 图像处理的基本操作

■ 图像编码

■ 压缩数据，有效传输，节省空间



450KB



89KB

特征的分类：

根据人类视觉系统对场景的感知，特征分为**低层特征**和**高层语义特征**两个方面。

低层特征包括：外观特征、边缘特征、纹理特征、形状特征等。

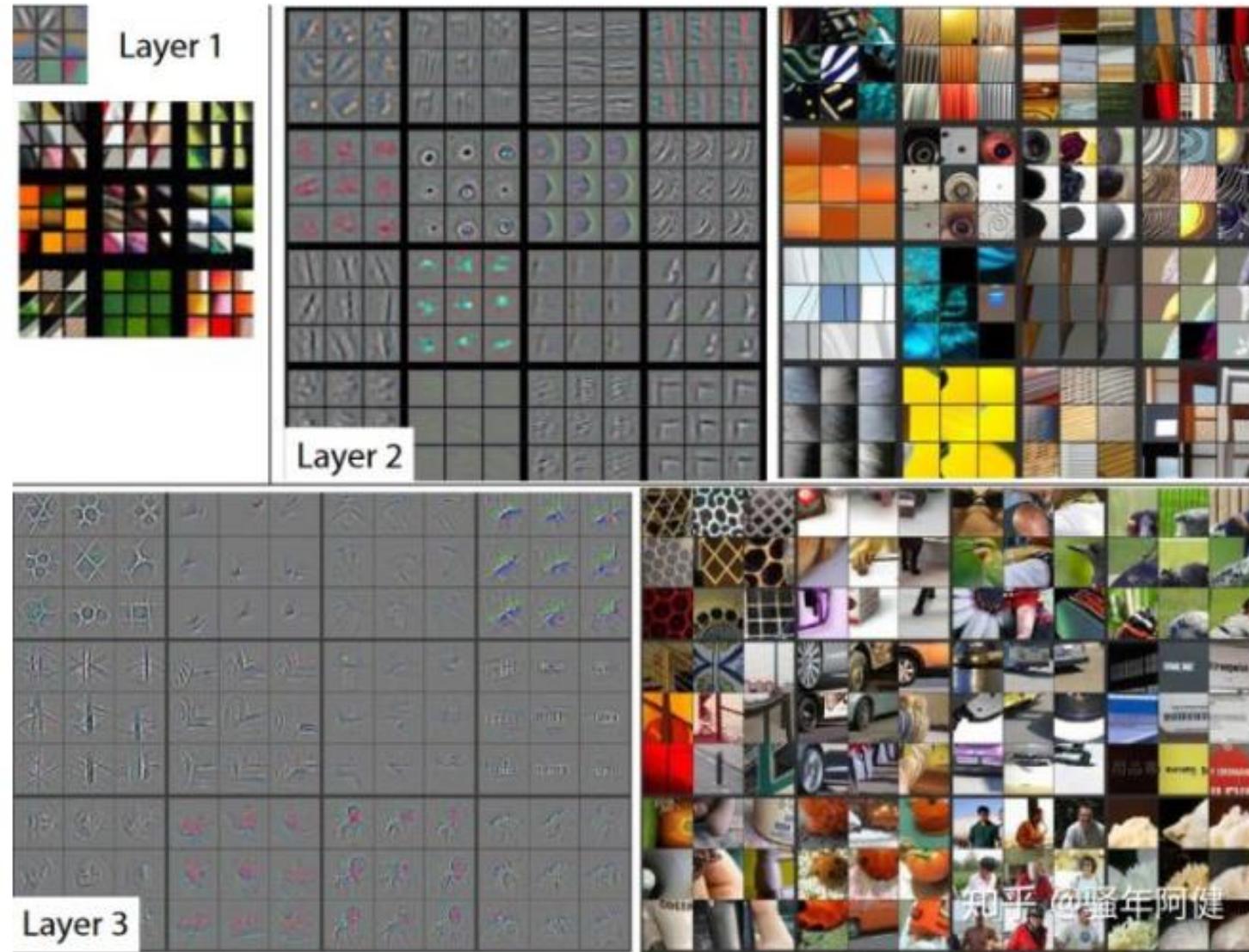
高层语义特征包括：场景语义、行为语义和情感语义等3个层次。

在底层图像特征与高层语义特征之间，往往存在着“语义鸿沟”，视觉物体的**中层语义表示**是连接低层与高层语义的纽带。

低层特征描述的是图像的视觉特征，中层语义是由视觉特征推导得到的中间语义特征，对应于**视觉词袋**和**中间语义（主题语义）**两种表示方法。



低层、中层和高层特征



知乎 @ 骚年阿健



有效的特征与特征描述举例：

- 统计特征（统计直方图、方差、均值、熵）
- 几何特征（包括面积、周长、位置、方向、距离）
- 边界特征（链码、傅立叶描述子）
- 形状描述（区域描述、图像矩）
- 灰度与颜色描述（梯度特征、灰度直方图特征、梯度方向直方图、特征直方图、颜色直方图和颜色直方图距离、颜色不变量特征）
- 角点特征（Harris、Susan、Fast、Brisk）
- 变换特征（傅里叶、小波）
- 纹理特征（灰度共生矩阵、LBP特征，结构相似性特征，Gabor特征）
- 局部不变特征描述子（SIFT、SURF、GLOH、WLD、D-nets） ...



- 1 基本概念
- 2 颜色特征描述
- 3 形状特征描述
- 4 图像的纹理分析技术
- 5 小结



◆ 目的

让计算机具有认识或者识别图像的能力，即图像识别。特征选择是图像识别中的一个关键问题。特征选择和提取的基本任务是如何从众多特征中找出最有效的特征。

◆ 特征形成

根据待识别的图像，通过计算产生一组原始特征，称之为特征形成。



◆ 特征提取

原始特征的数量很大，或者说原始样本处于一个高维空间中，通过映射或变换的方法可以将高维空间中的特征描述用低维空间的特征来描述，这个过程就叫特征提取。

◆ 特征选择

从一组特征中挑选出一些最有效的特征以达到降低特征空间维数的目的，这个过程就叫特征选择。

选取的特征应具有如下特点：

可区别性

可靠性

独立性好

数量少

8.2 颜色特征描述

8.2.1 颜色矩

8.2.2 颜色直方图



8.2.1 颜色矩

- ◆ 颜色矩是以数学方法为基础的，通过计算矩来描述颜色的分布。
- ◆ 颜色矩通常直接在RGB空间计算
- ◆ 颜色分布的前三阶矩表示为：

$$\mu_i = \frac{1}{N} \sum_{j=1}^N p_{i,j}$$

$$\sigma_i = \left(\frac{1}{N} \sum_{j=1}^N (p_{i,j} - \mu_i)^2 \right)^{\frac{1}{2}}$$

$$s_i = \left(\frac{1}{N} \sum_{j=1}^N (p_{i,j} - \mu_i)^3 \right)^{\frac{1}{3}}$$





8.2.2 颜色直方图

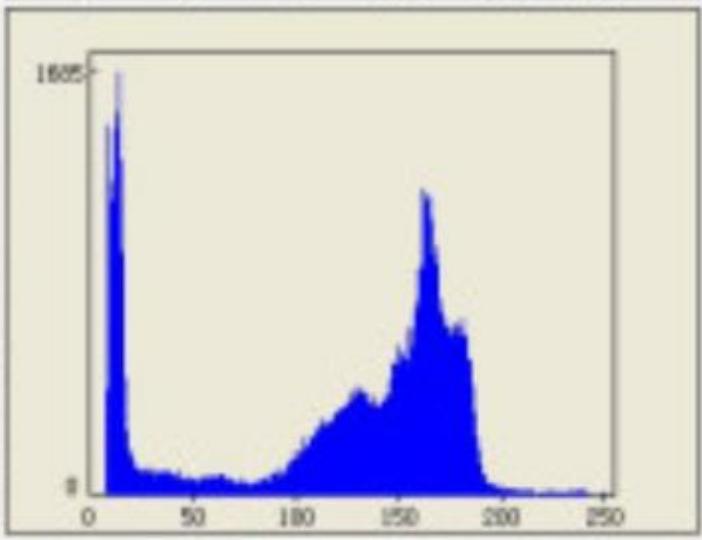
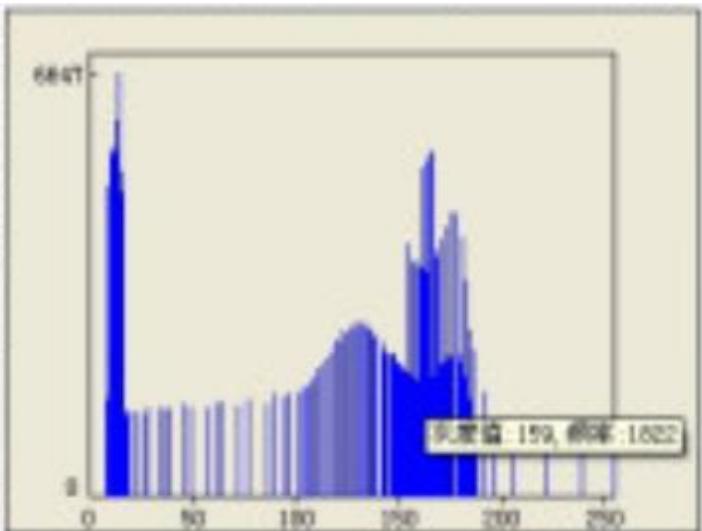
- ◆ 设一幅图像包含M个像素，图像的颜色空间被量化成N个不同颜色。颜色直方图H定义为：

$$p_i = h_i \quad (8-4)$$

h_i 为第i种颜色在整幅图像中具有的像素数。

- ◆ 归一化为：

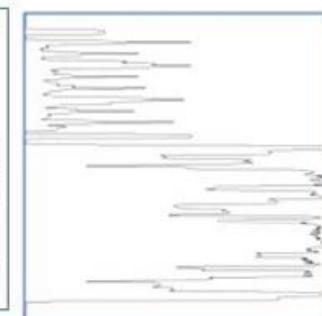
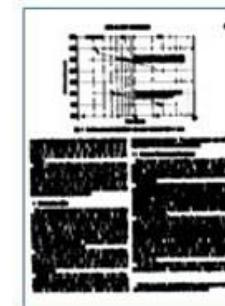
$$p_i = h_i / M \quad (8-5)$$



投影直方图

- 对于一个二值图像 $f(x, y)$, 其垂直和水平投影直方图可以分别记为 $p_v(y)$ 和 $p_h(x)$, 其中

- $p_v(y) = \sum_{x=0}^{w(f)} f(x, y), \quad p_h(x) = \sum_{y=0}^{h(f)} f(x, y),$
- $w(f)$ 和 $h(f)$ 分别为 $f(x, y)$ 的像素宽度和高度。



形状特征提取

- 基于图像内物体形状的检索是基于内容检索当中一个**最具挑战性**的问题之一，因为寻找符合人眼感知特性的形状特征不是一件简单的工作。
- 首要的**困难**是要将不同物体从图像中分割出来，这是计算机视觉的困难问题之一。
- 形状的描述也是困难的问题，常用的方法有**傅立叶描述子**、**矩不变量**、各种简单的**形状因子**（如**面积**、**圆度**、**偏心度**、**主轴方向**）等。
- 除了这些全局特征以外，有时也用一些局部特征（如**直线段**、**圆弧**、**角点**、**高曲率点**等），以解决遮挡问题。

福

寿



形状特征提取

Hu不变矩

Zernike不变矩

Hough变换

Radon变换

傅里叶描述子



HU不变矩

- 矩是一种完备的数学表示，其优点是能够直接用于包含感兴趣目标的区域而不需事先把目标分离出来。
- 区域的矩是用所有属于区域内的点计算出来的，因而不太受噪声等的影响。
- 由于低阶矩对噪声和量化误差不敏感，所以矩为描述区域提供了一种虽不完备但却有用的总体表示。
- 矩在图像检索中，特别是在商标图像检索中，得到了广泛的应用。
（背景简单）
- Hu不变矩具有良好的尺度、平移和旋转不变性



不变矩

◎ $f(x, y)$ 的 $j + k$ 阶矩定义为：

■ $m_{jk} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^j y^k f(x, y) dx dy,$

◎ $f(x, y)$ 的 $j + k$ 阶中心规格矩定义为：

■ $\mu_{jk} = \frac{1}{m_{00}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^j (y - \bar{y})^k f(x, y) dx dy,$

○ 其中 $\bar{x} = \frac{m_{10}}{m_{00}}$, $\bar{y} = \frac{m_{01}}{m_{00}}$, $j + k = 2, 3, 4, \dots$ 。

◎ 中心规格矩同时具有位置无关性和大小无关性。

◎ 使得二阶中心矩 μ_{11} 为最小的旋转角 θ 可以由如下公式得出

■ $\tan 2\theta = \frac{2\mu_{11}}{\mu_{20} - \mu_{02}}.$

○ 相对于主轴计算的中心规格矩，在物体缩放、平移和旋转时保持不变。

HU不变矩

- $\emptyset_1 = \mu_{20} + \mu_{02}$
- $\emptyset_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2$
- $\emptyset_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2$
- $\emptyset_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2$
- $\emptyset_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]$
- $\emptyset_6 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03})$
- $\emptyset_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2]$



HU不变矩



(a)



(b)



(c)



(d)

	\emptyset_1	\emptyset_2	\emptyset_3	\emptyset_4	\emptyset_5	\emptyset_6	\emptyset_7
(a)	0.001286	0.000000	0.000000	0.000000	-0.000000	0.000000	0.000000
(b)	0.411764	0.001129	0.000195	0.004179	0.000004	0.000120	-0.000002
(c)	0.001615	0.000000	0.000000	0.000000	0.000000	0.000000	-0.000000
(d)	0.310856	0.002382	0.005049	0.000931	-0.000001	-0.000009	0.000002



ZERNIKE不变矩

◎ Zernike不变矩是来自数学领域的概念，其定义是一组多项式序列，每个多项式在单位圆上是正交的。

◎ 存在奇数和偶数Zernike多项式，其中偶数多项式可以定义为

- $Z_n^m(\rho, \varphi) = R_n^m(\rho) \cos(m\varphi)$,

- 而奇数多项式则定义为

- $Z_n^{-m}(\rho, \varphi) = R_n^m(\rho) \sin(m\varphi)$,

- 其中， m 和 n 为非负整数，且 $n \geq m$ ， φ 是以弧度为单位的方位角， ρ 为规范化的径向距离，径向多项式 R_n^m 则定义为

- $$\odot R_n^m(\rho) = \sum_{k=0}^{(n-m)/2} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+m}{2} - k\right)! \left(\frac{n-m}{2} - k\right)!} \rho^{n-2k}.$$



ZERNIKE不变矩

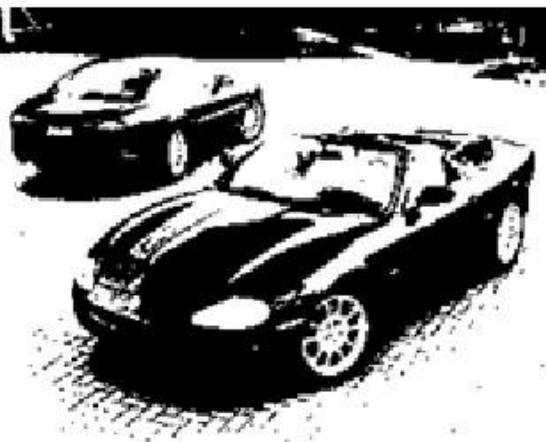
- ◎ 在Zernike不变矩的基础上，伪Zernike不变矩被提出，并具有更好的鲁棒性和对图像噪声的不敏感性。
 - 寻找目标像素所占据的外接矩形区域；
 - 根据所设定的阶数 n ，进行重复次数 l 的计算，从而进一步得到所需计算的不变矩系数的个数；
 - 针对每一对阶数 n 和重复次数 l ，进行如下四个子步骤的计算：
 - 将步骤1所得到的外接矩形区域归一化为一个单位圆；
 - 将所有图像像素从笛卡尔坐标系转换至以单位圆圆心为原点的极坐标系中；
 - 在极坐标系下计算每一个像素离开原点的径向距离；
 - 统计所有像素的径向距离得到在当前阶数 n 和重复次数 l 下的不变矩系数；
 - 将所有的不变矩系数组合起来得到一个系数序列，便是一幅二值图像所对应的伪Zernike不变矩的计算结果。



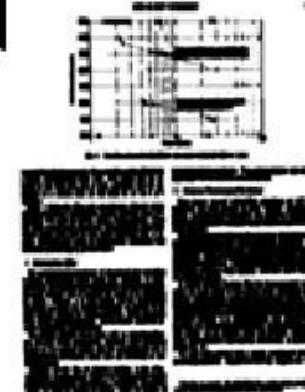
伪ZERNIKE不变矩计算



(a)



(b)



(c)



(d)

	A_{00}	A_{11}	A_{22}	A_{20}	A_{33}	A_{31}
(a)	0.318224	0.083217	0.289287	0.034964	0.163702	0.004639
(b)	0.318229	0.045778	0.423569	0.037042	0.074096	0.080350
(c)	0.318310	0.142310	0.385979	0.073231	0.185505	0.038667
(d)	0.318289	0.039561	0.186443	0.094189	0.019000	0.054490

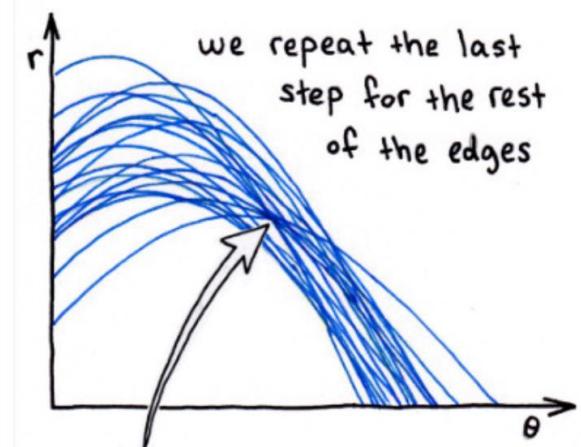
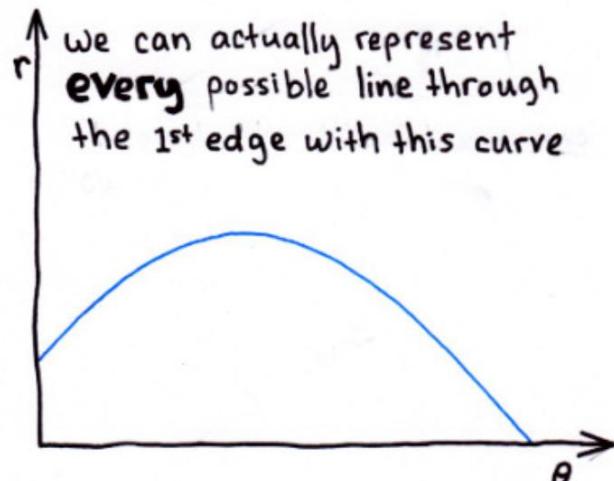
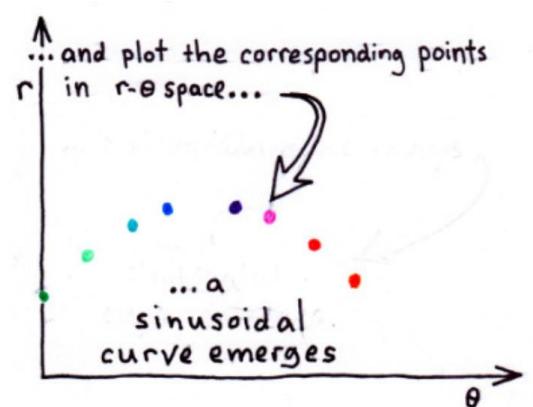
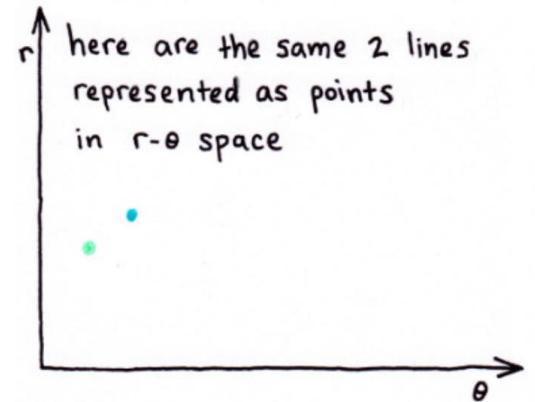
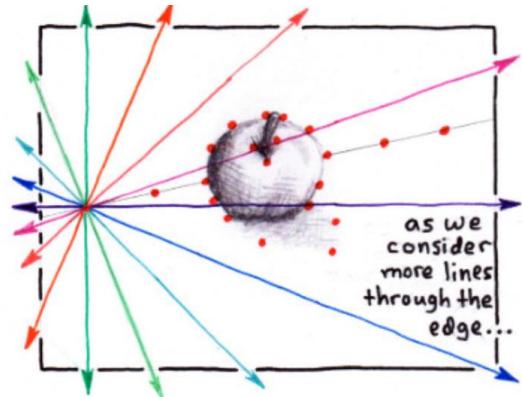
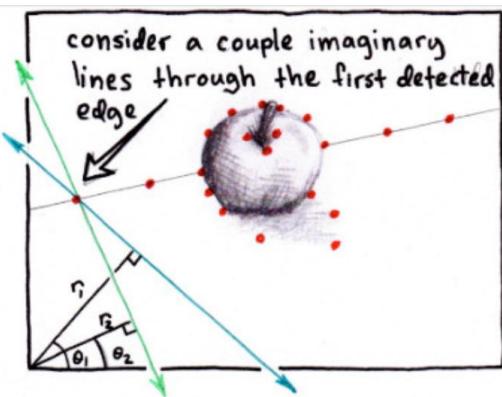
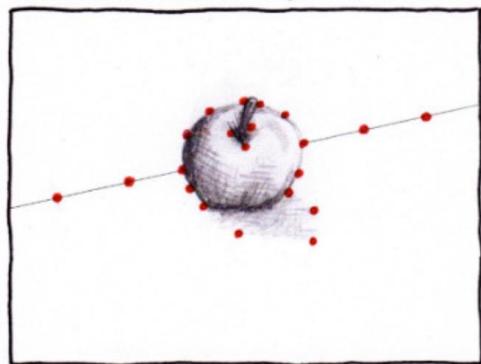


HOUGH变换

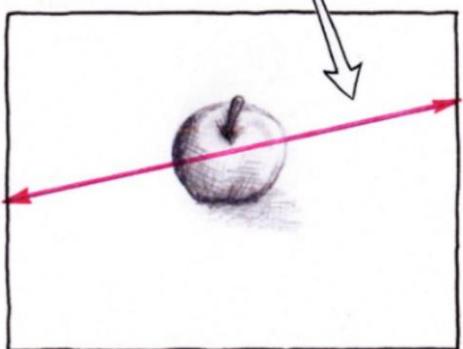
- ◎ 经典的Hough变换主要涉及图像中的直线检测，但是后来Hough变换得到了扩展，被用于任意形状位置的检测，其中最常用的是圆形或者椭圆。
 - Hough变换最简单的实例就是用于直线检测的线性变换。
- ◎ 使用参数空间 (r, θ) ，直线公式可以重写为
 - $y = \left(-\frac{\cos\theta}{\sin\theta}\right)x + \left(\frac{r}{\sin\theta}\right)$,
 - 这也就等价于公式
 - $r = x \cos\theta + y \sin\theta$ 。

直角坐标系中的一条直线对应于极坐标系下的一个点。

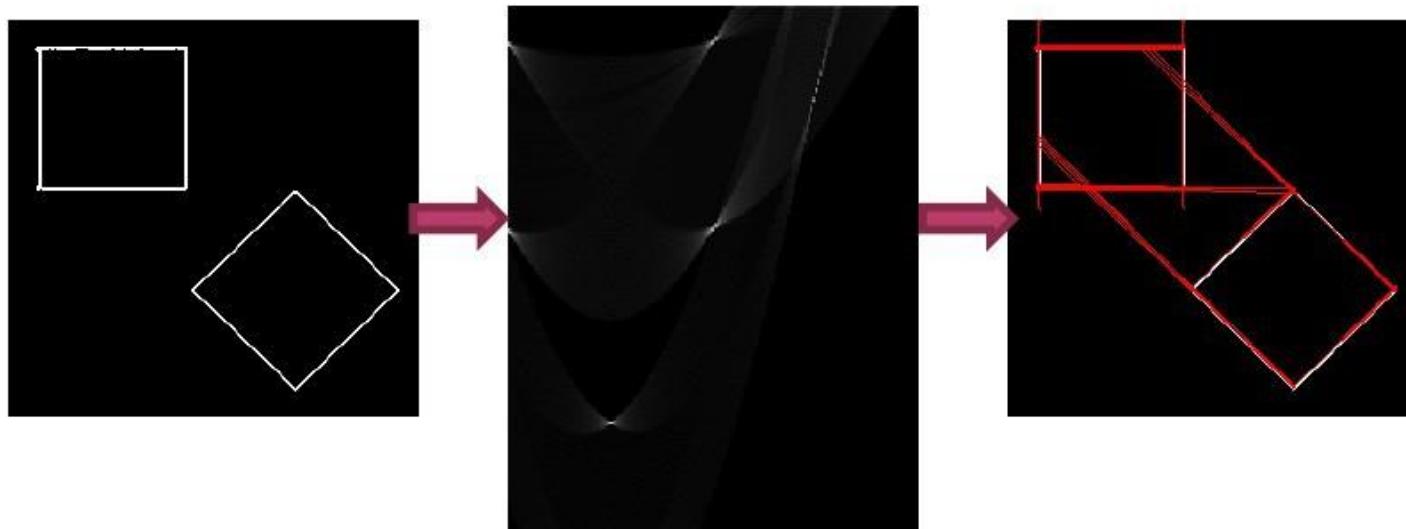
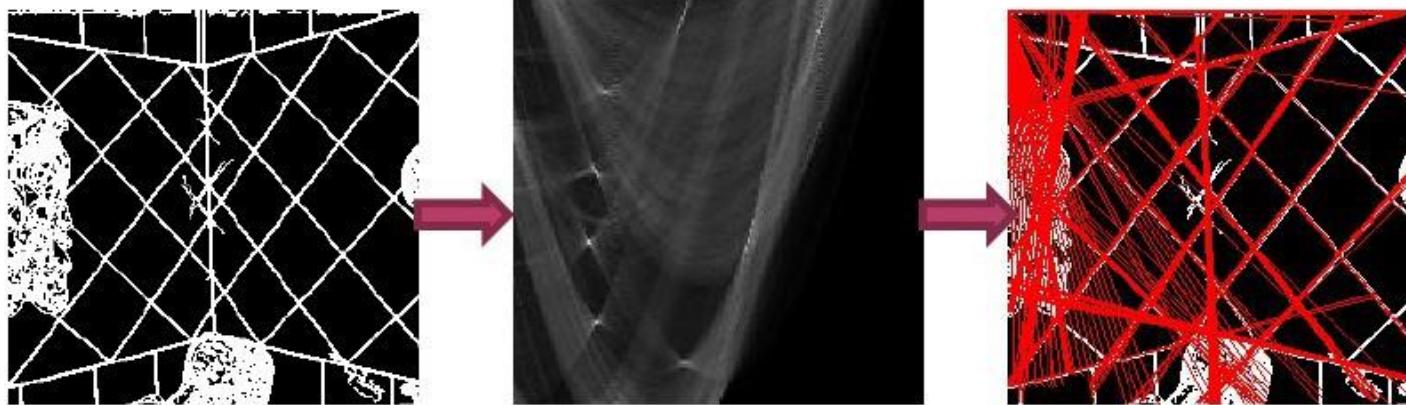
detected edges



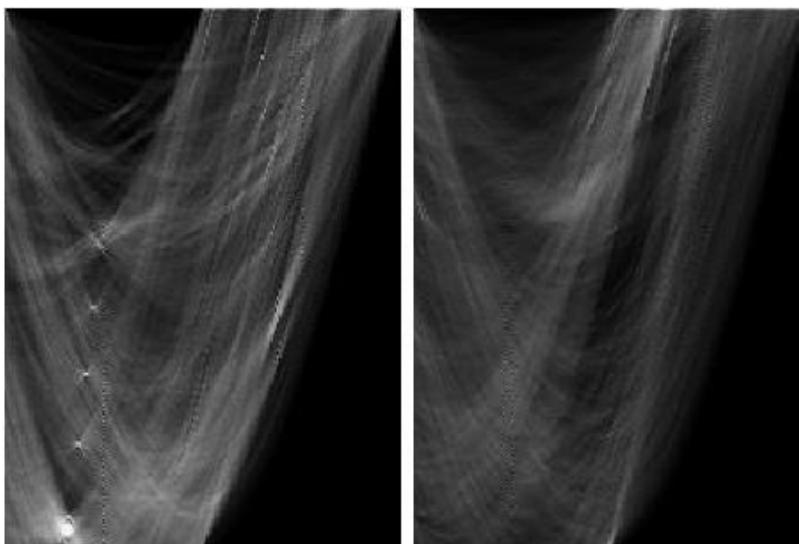
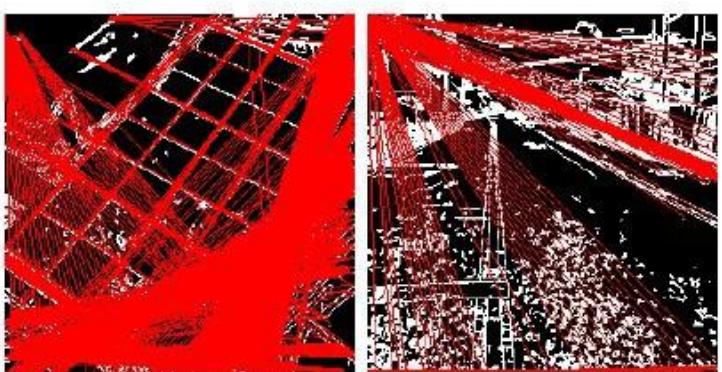
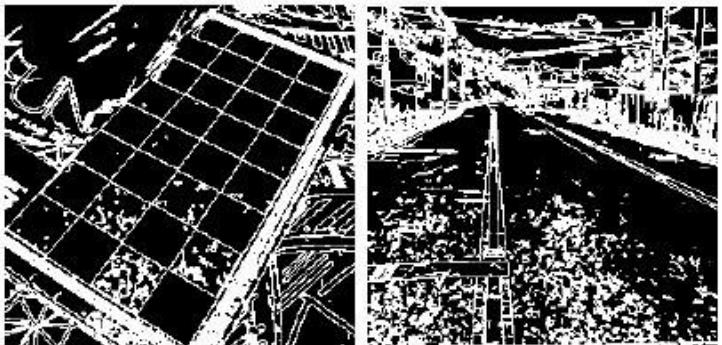
The point with the most curve-crossings identifies an imaginary line in the original image that passes through the most edges



HOUGH变换



HOUGH变换



RADON变换

- ◎ 二维Radon变换属于一种积分变换，是由众多直线函数上的积分组成。
- ◎ Radon变换被广泛的应用于断层摄影术。
 - 如果 f 表示一个未知的密度函数，那么Radon变换表示一种发散数据，可用于断层摄影术的输出。
 - 因此，Radon逆变换可以用于从发散数据中重构原始密度函数，从而形成了断层摄影术重构技术的数学基础。
- ◎ Radon变换数据通常被称为正弦图，因为Dirac三角函数的Radon变换结果看上去是呈一个正弦波的分布。
 - 许多小物体的Radon变换在显示上都会呈现出模糊正弦波的样子，而这些正弦波具有不同的振幅和相位。

RADON变换

◎ Radon变换的数学公式可以描述为

- $R(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(\rho - x\cos\theta - y\sin\theta) dx dy,$
- 其中 $\rho = x\cos\theta + y\sin\theta$ 表示一条直线， $\delta(\cdot)$ 为 Dirac 三角函数。

◎ Dirac 三角函数大致可以认为是一种在原点无穷大而其余地方为零的函数，即

- $\delta(x) = \begin{cases} +\infty & x = 0 \\ 0 & x \neq 0 \end{cases},$

- 并且满足如下条件

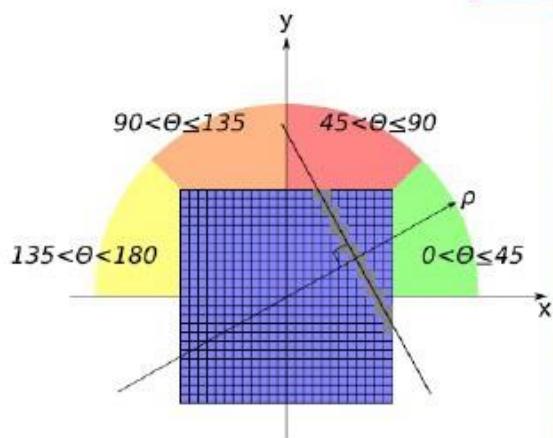
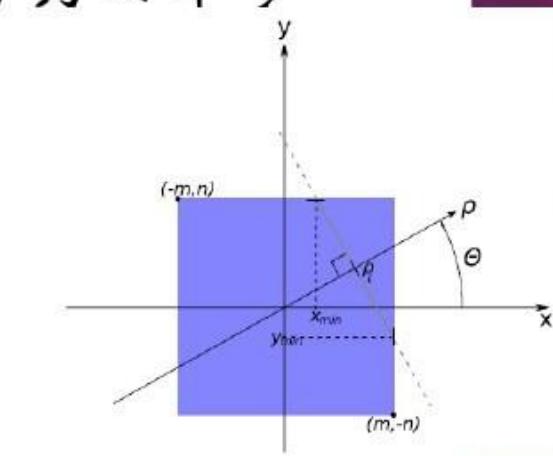
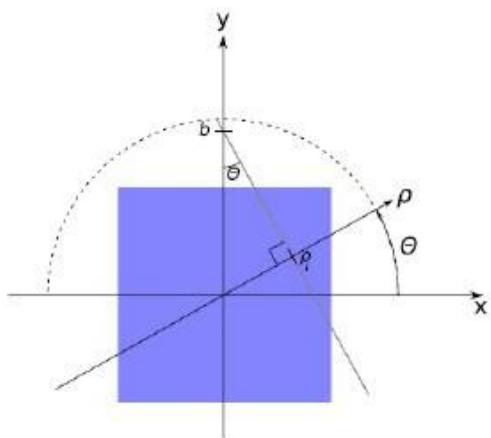
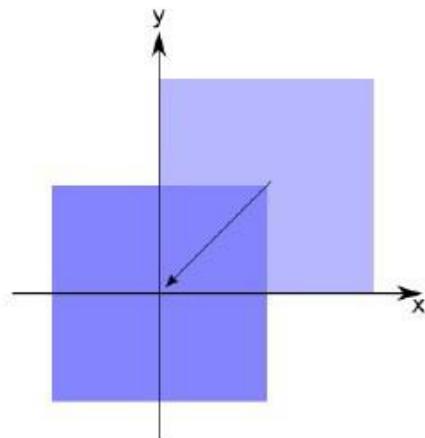
- $\int_{-\infty}^{\infty} \delta(x) dx = 1.$

◎ Radon 变换实际上是从笛卡尔直角坐标 (x, y) 到极坐标 (ρ, θ) 的映射

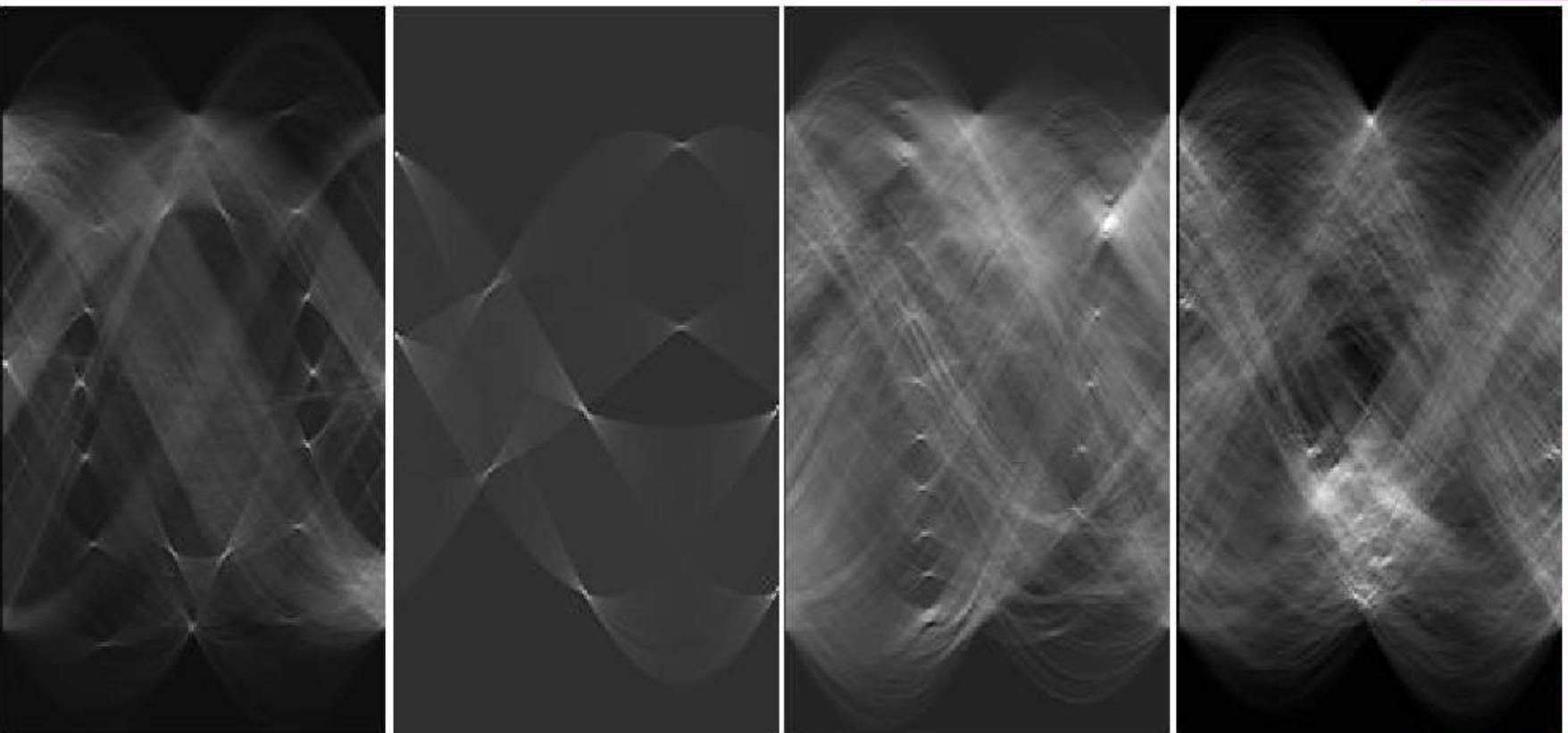
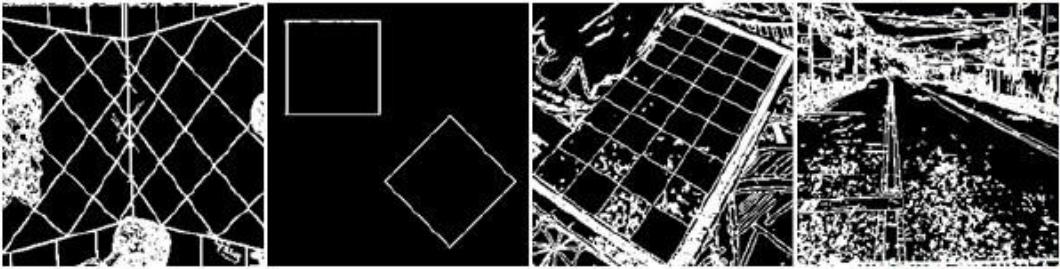
RADON变换

◎ Radon变换的具体实现，大致可以分为四个步骤

- 将坐标原点移至图像中心
- 直线参数计算
- 直线坐标的最大最小值计算
- 根据角度进行分区讨论



RADON变换



傅里叶描述子

- ① 傅里叶描述子，就是对封闭曲线进行傅里叶变换，将变换后得到的傅里叶系数作为表示图像形状的特征。
- ② 可以将形状表示为具有 N 个顶点的集合 $\{z(i): i, \dots, N\}$ ，那么傅里叶描述子 $\{c(k): k = -\frac{N}{2} + 1, \dots, \frac{N}{2}\}$ 就是集合 Z 的傅里叶变换系数集合，即

- $z_i = \sum_{k=-\frac{N}{2}+1}^{\frac{N}{2}} c_k e^{2\pi j \frac{ki}{N}}$ ，
- 而 $c(k)$ 与 $z(i)$ 之间的逆变换为
 - $c_k = \frac{1}{N} \sum_{i=1}^N z_i e^{-2\pi j \frac{ik}{N}}$ 。

- ③ 为了进行傅里叶描述子的计算，就需要两步操作：
 - 一是从二值图像中得到一维复数数组，用以作为一维傅里叶变换的输入
 - 二是进行一维傅里叶正变换。

傅里叶描述子

- 为了验证所得到的傅里叶描述子是否正确，可以对傅里叶描述子进行傅里叶逆变换，然后从所得到的一维复数数组中恢复二值图像，通过比较新旧图像就可以看出所得到的傅里叶描述子是否正确。



■ 区域形状的描述

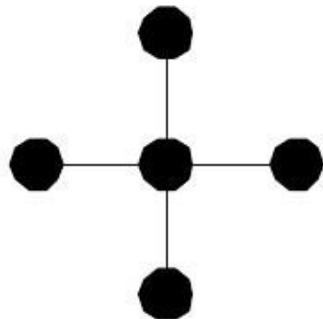
- 常针对二值图
- 属于手工特征



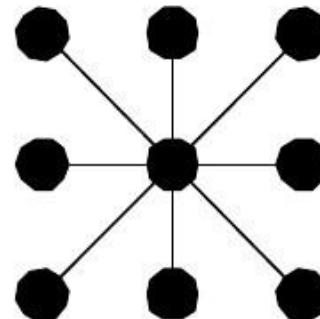
8.3.1几个基本概念

◆ 邻域与邻接

- ▶ 对于任意像素 $(i, j), (s, t)$ 是一对适当的整数，则把像素的集合 $\{(i+s, j+t)\}$ 叫做像素 (i, j) 的邻域.
- ▶ 直观上看，这是像素 (i, j) 附近的像素形成的区域.
- ▶ 最经常采用的是4-邻域和8-邻域



(a)



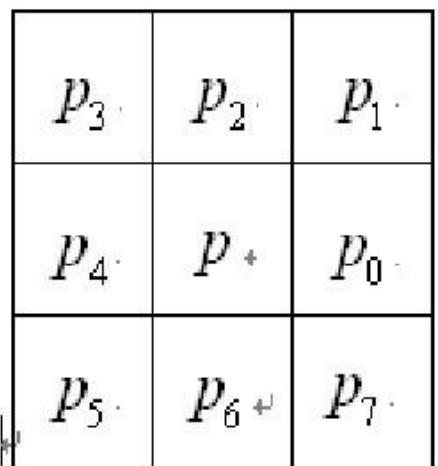
(b)

4-邻域和8-邻域



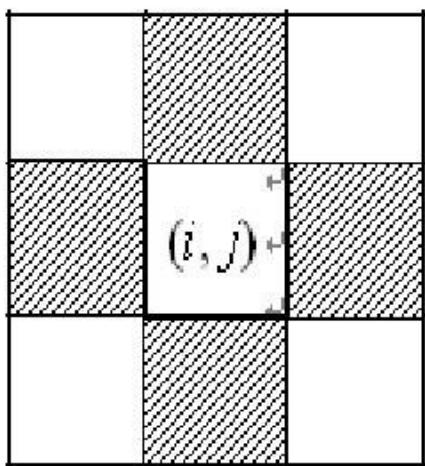
◆ 邻域与邻接

- ▶ 互为4-邻域的两像素叫4-邻接。
- ▶ 互为8-邻域的两像素叫8-邻接。



(a)

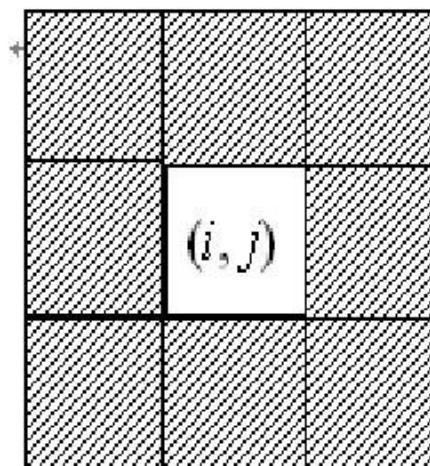
(a) 像素的编号



(b)

图 8-1 邻接像素的种类

(b) 4-邻接

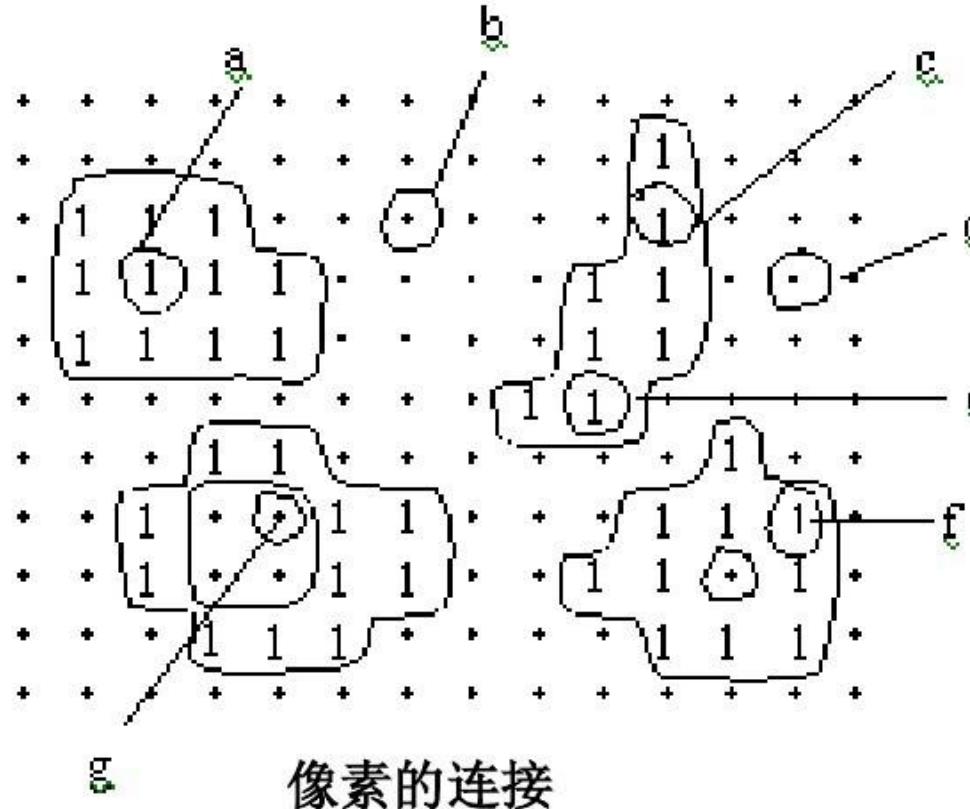


(c)

(c) 8-邻接

◆ 像素的连接

对于图像中具有相同值的两个像素A和B，如果所有和A、B具有相同值的像素序列 $L_0 (= A), L_1, L_2, \dots, L_{n-1}, L_n (= B)$ 存在，并且 L_{i-1} 和 L_i 互为4-邻接或8-邻接，那么像素和叫做4-连接或8-连接，以上的像素序列叫4-路径或8-路径。



像素的连接

◆ 连接成分

- ▶ 在图像中，把互相连接的像素的集合汇集为一组，于是具有若干个0值的像素和具有若干个1值的像素的组就产生了。把这些组叫做连接成分，也称作连通成分。
- ▶ 在研究一个图像连接成分的场合，若1像素的连接成分用4-连接或8-连接，而0像素连接成分不用相反的8-连接或4-连接就会产生矛盾。
- ▶ 假设各个1像素用8-连接，则其中的0像素就被包围起来。如果对0像素也用8-连接，这就会与左下的0像素连接起来，从而产生矛盾。因此0像素和1像素应采用互反的连接形式，即如果1像素采用8-连接，则0像素必须采用4-连接。

0	0	0	0	0
0	1	1	1	0
0	1	0	1	0
0	0	1	1	0
0	0	0	0	0

连接性矛盾示意图



► 在0-像素的连接成分中，如果存在和图像外围的1行或1列的0-像素不相连接的成分，则称之为孔。不包含有孔的1像素连接成分叫做单连接成分。含有孔的1像素连接成分叫做多重连接成分。

0	0	0	0	0
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
0	0	0	0	0

(a)孤立点

0	0	0	0	0
0	1	1	0	0
0	1	1	1	0
0	1	1	1	0
0	0	0	0	0

(b)

⊕

孔

1	1	1	1	1
1	0	1	1	1
1	1	1	0	1
1	0	0	0	1
1	1	1	1	1

(c)

图 8-4 连接成分

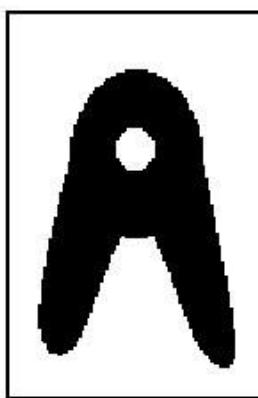
(a)孤立点 (b)单连接成分(c)多重连接成分

8.3.2 区域内部空间域分析

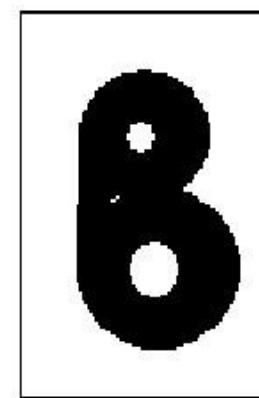
- ◆ 区域内部空间域分析是不经过变换而直接在图像的空间域，对区域内提取形状特征。

1. 欧拉数 欧拉数=连接成分数-孔数

- ✓ 图像的欧拉数是图像的拓扑特性之一，它表明了图像的连通性。下图 (a) 的图形有一个连接成分和一个孔，所以它的欧拉数为0，而下图 (b) 有一个连接成分和两个孔，所以它的欧拉数为-1。
- ✓ 可见通过欧拉数可用于目标识别。



(a)



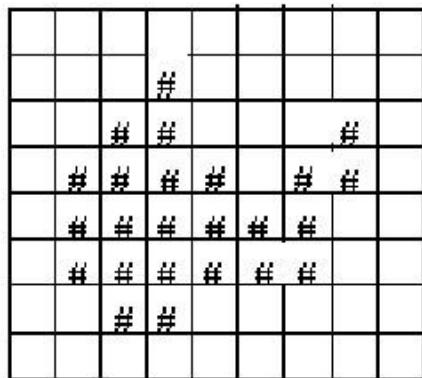
(b)

具有欧拉数为0和-1的图形

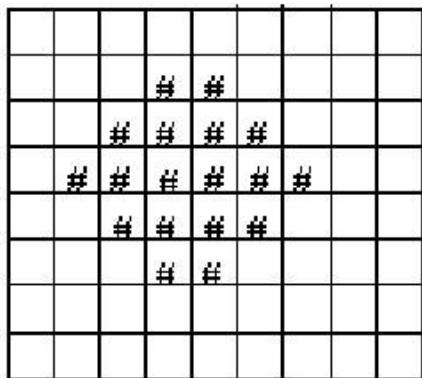


2.凹凸性

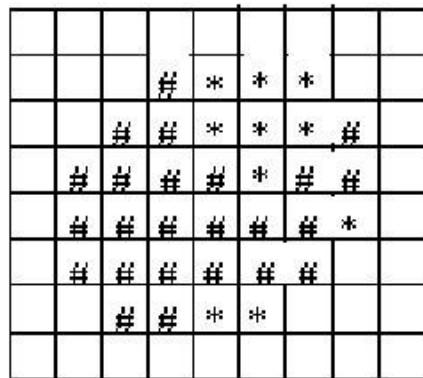
- ✓ 凹凸性是区域的基本特征之一，区域凹凸性可通过以下方法进行判别：区域内任意两像素间的连线穿过区域外的像素，则此区域为凹形。相反，连接图形内任意两个像素的线段，如果不通过这个图形以外的像素，则这个图形称为是凸的。任何一个图形，把包含它的最小的凸图形叫这个图形的凸闭包。
- ✓ 凸图形的凸闭包就是它本身。从凸闭包除去原始图形的部分后，所产生的图形的位置和形状将成为形状特征分析的重要线索。凹形面积可将凸封闭包减去凹形得到。



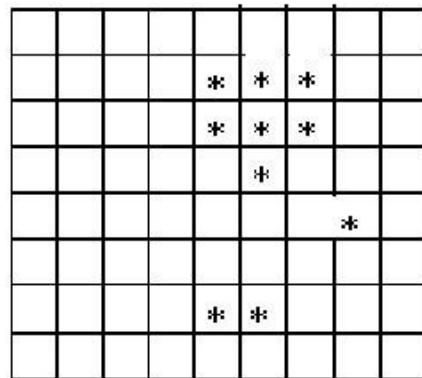
(a) 凹形



(b) 凸形



(c)(a) 中凹形的凸封闭包



(d) 凹形面积

区域的凹凸性



3. 距离

- ✓ 距离在实际图像处理过程中往往作为一个特征量出现，因此对其精度的要求并不是很高。所以对于给定图像中三点A, B, C，当函数 $D(A, B)$ 满足下式的条件时，把 $D(A, B)$ 叫做A和B的距离，也称为距离函数。

$$\begin{cases} D(A, B) \geq 0 \\ D(A, B) = D(B, A) \\ D(A, C) \leq D(A, B) + D(B, C) \end{cases}$$

- ✓ 第一个式子表示距离具有非负性，并且当A和B重合时，等号成立；
- ✓ 第二个式子表示距离具有对称性
- ✓ 第三个式子表示距离的三角不等式。



✓ 计算点(i, j)和(h, k)间距离常采用的几种方法：

(1) 欧氏距离，用 d_e 来表示。

$$d_e[(i, j), (h, k)] = [(i - h)^2 + (j - k)^2]^{\frac{1}{2}}$$

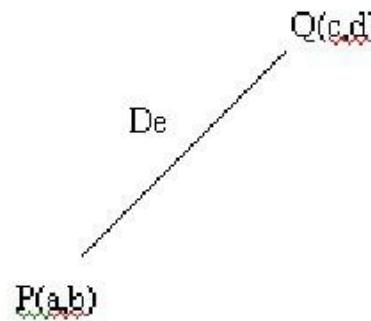
(2) 4-邻域距离，也称为街区距离。

$$d_s[(i, j), (h, k)] = [(i - h)^2 + (j - k)^2]^{\frac{1}{2}}$$

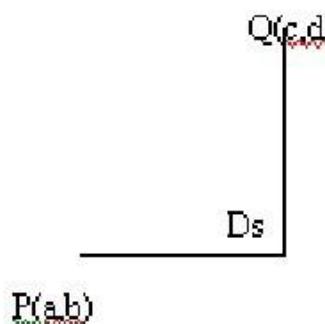
(3) 8-邻域距离，也称为棋盘距离。

$$d_g[(i, j), (h, k)] = \max(|i - h|, |j - k|)$$

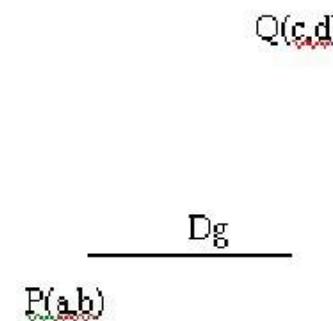
这三种距离之间的关系： $d_g \leq d_s \leq d_e$ ，如图所示。街区距离和棋盘距离都是欧式距离的一种近似。



(a) 欧式距离



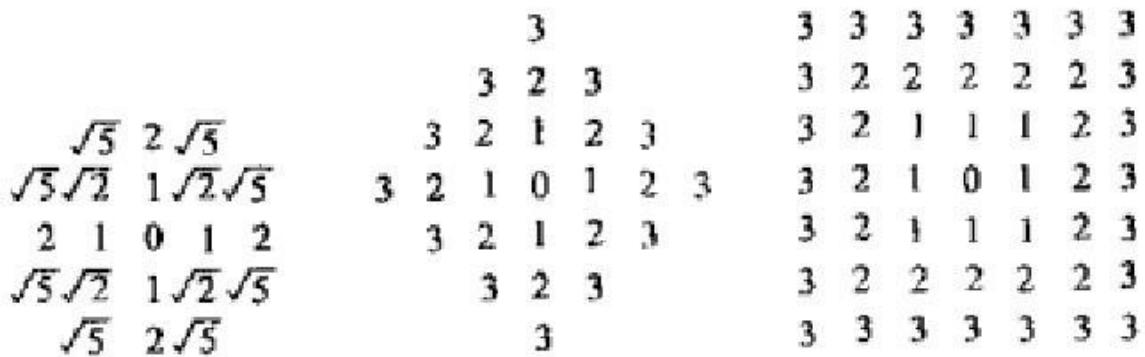
(b) 街区距离



(c) 棋盘距离

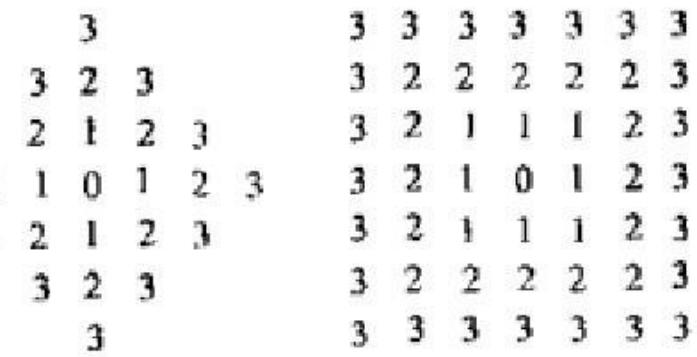


- ✓ 下图中表示了以中心像素为原点的各像素的距离。从离开一个像素的等距离线可以看出，在欧氏距离中大致呈圆形，在棋盘距离中呈方形，在街区距离中呈倾斜45度的正方形。街区距离是图像中两点间最短的4-连通的长度，而棋盘距离则是两点间最短的8-连通的长度。
- ✓ 此外，把4-邻域距离和8-邻域距离组合起来而得到的八角形距离有时也被采用，它的等距线呈八角形。



(a)

(a) 欧几里德距离



(b)

(b) 4-邻域距离

(c)

离开单个像素的距离

(c) 8-邻域距离

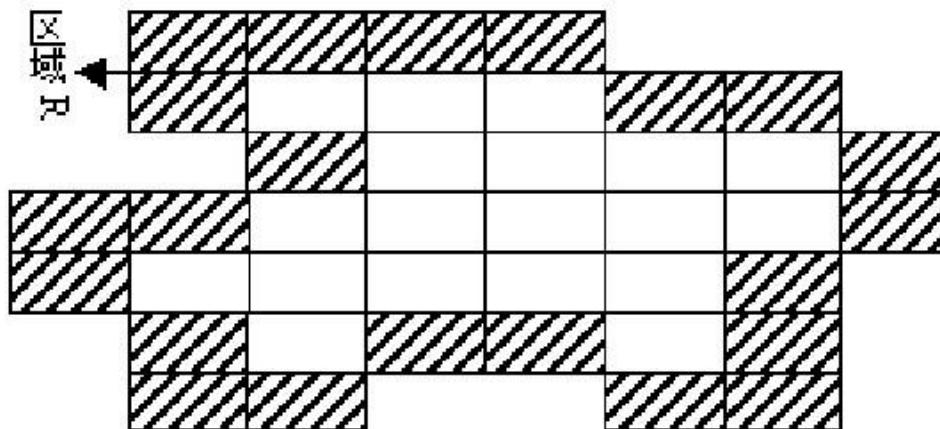


4.区域的测量

- ✓ 区域的大小及形状表示方法主要包括以下几种：

(1) 面积S：图像中的区域面积S可以用同一标记的区域内像素的个数总和来表示。

福



王

区域的面积和周长

按上述表示法区域R的面积 $S=41$ 。区域面积可以通过扫描图像，累加同一标记像素得到，或者是直接在加标记处理时计数得到。



(2) 周长L: 区域周长L是用区域中相邻边缘点间距离之和来表示。采用不同的距离公式, 关于周长L的计算有很多方法。常用的有两种:

一种计算方法是采用欧式距离, 在区域的边界像素中, 设某像素与其水平或垂直方向上相邻边缘像素间的距离为1, 与倾斜方向上相邻边缘像素间的距离为 $\sqrt{2}$ 。周长就是这些像素间距离的总和。这种方法计算的周长与实际周长相符, 因而计算精度比较高。

另一种计算方法是采用8邻域距离, 将边界的像素个数总和作为周长。也就是说, 只要累加边缘点数即可得到周长, 比较方便, 但是, 它与实际周长间有差异。

(3) 圆形度 R_0 : 圆形度 R_0 用来描述景物形状接近圆形的程度，它是测量区域形状常用的量。其计算公式为：

$$R_0 = 4\pi S / L^2$$

式中为S区域面积；L为区域周长 R_0 值的范围为 $0 < R_0 \leq 1$ ， R_0 值的大小反映了被测量边界的复杂程度，越复杂的形状取值越小。 R_0 值越大，则区域越接近圆形。



(4) 形状复杂性e: 形状复杂性常用离散指数表示，其计算公式为：

$$e = L^2 / S$$

该式描述了区域单位面积的周长大小，e值越大，表明单位面积的周长大，即区域离散，则为复杂形状；反之，则为简单形状。e值最小的区域为圆形。

典型连续区域的计算结果为：圆形e=12.6；正方形e=16.0；正三角形e=20.8。

此外，常用的特征量还有区域的幅宽、占有率和直径等。

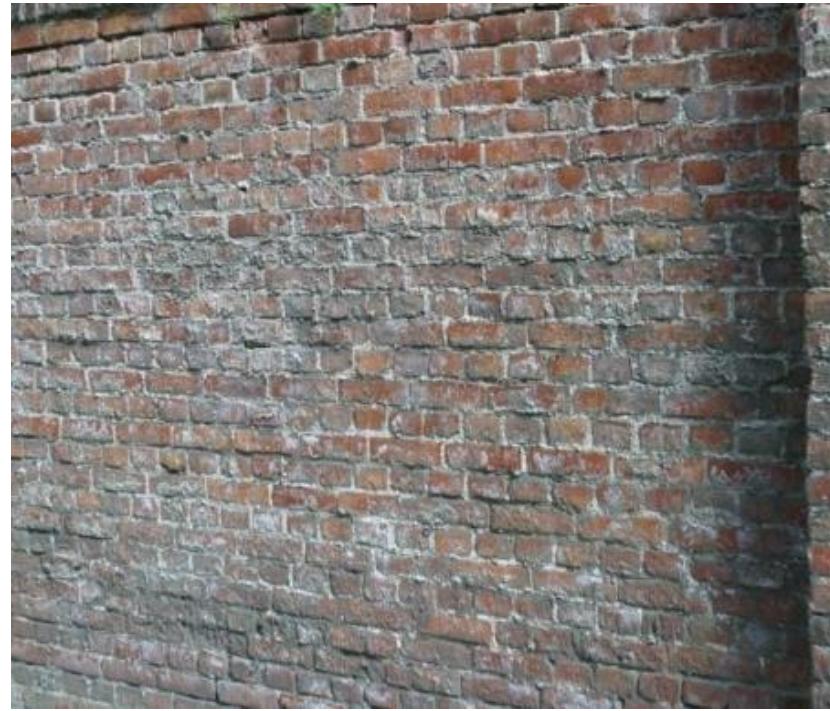
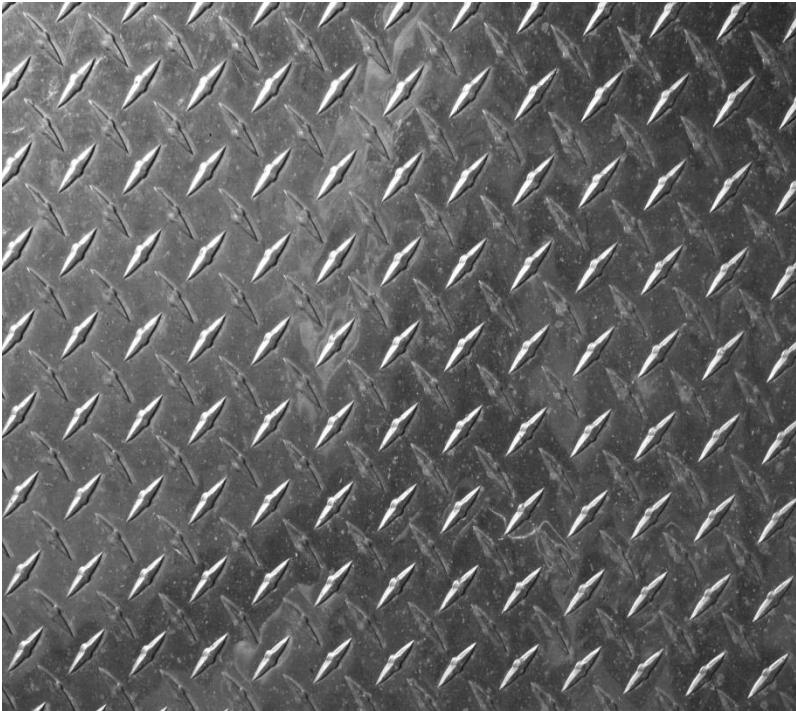


8.4 图像的纹理分析技术

8.4.1 纹理分析概念

- ◆ 指的是图像像素灰度级或颜色的某种变化，主要研究如何获得图像纹理特征和结构的定量描述和解释，以便于图像分析、分割和理解。
- ◆ 一般来说，可以认为纹理由许多相互接近、相互编织的元素构成，并常富有周期性。
- ◆ 纹理的定义大体可以从三个方面来描述：
 - ▶ 具有某种局部的序列性，并在该序列更大的区域内不断重复；
 - ▶ 序列由基本部分非随机排列组成；
 - ▶ 各个部分大致都是均匀的统一体。





图像的纹理特征



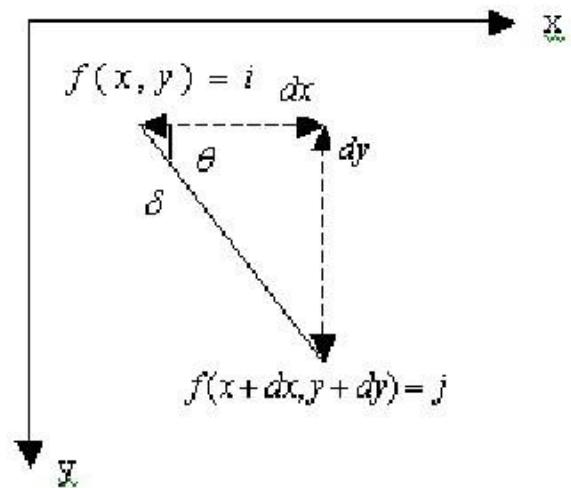
- ◆ 纹理分析是指通过一定的图像处理技术抽取出纹理特征，从而获得纹理的定量或定性描述的处理过程。
- ◆ 纹理特征是从图像中计算出来的一个值，它对区域内部灰度级变化的特征进行量化。
- ◆ 纹理分析基本过程是从像素出发，在纹理图像中提取出一些辨识力比较强的特征，作为检测出的纹理基元，并找出纹理基元排列的信息，建立纹理基元模型，然后再利用此纹理基元模型对纹理图像进一步分割、分类或是辨识等处理。



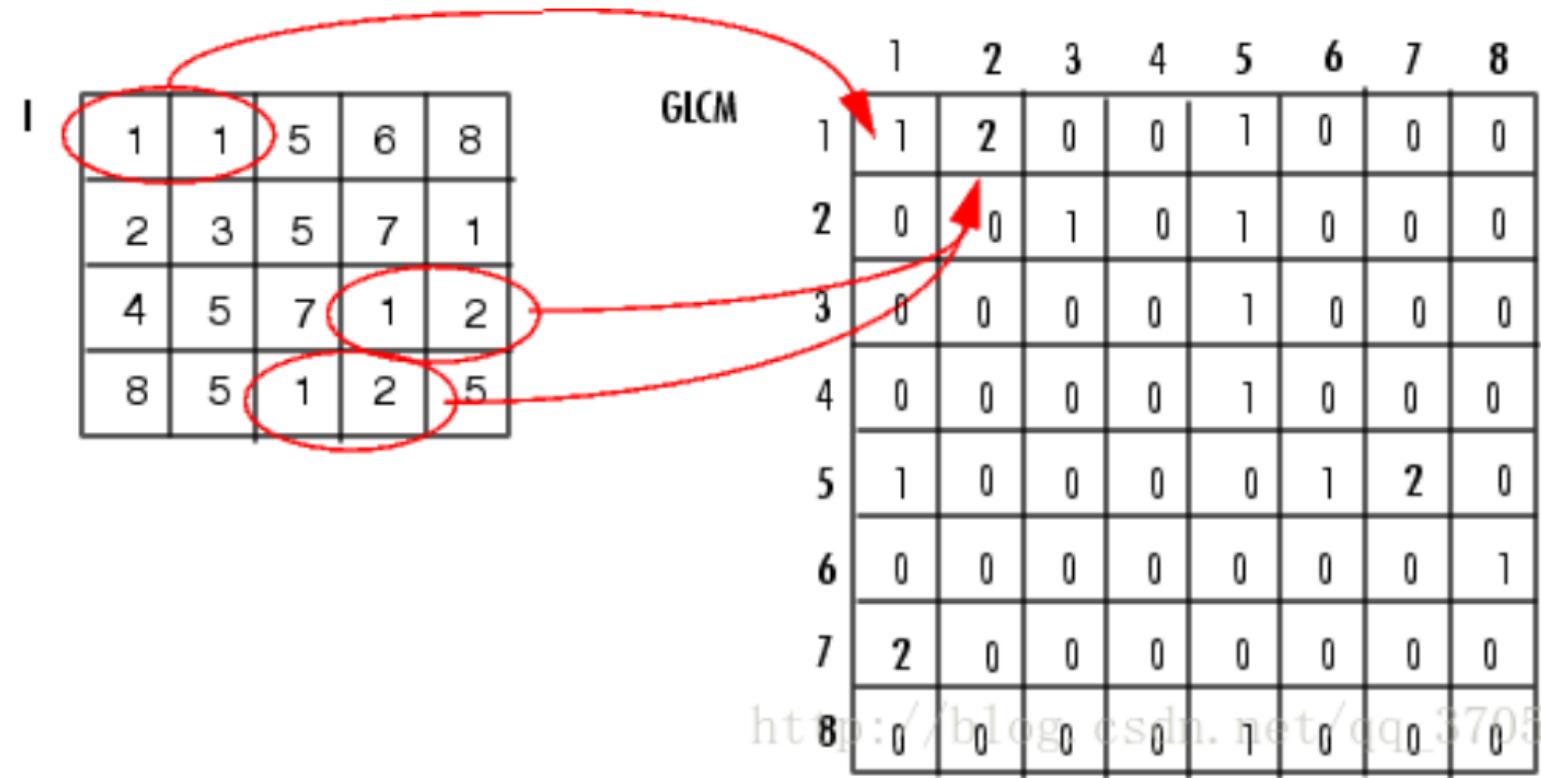
8.4.2 空间灰度共生矩阵

- ✓ 灰度共生矩阵就是从 $N \times N$ 的图像 $f(x, y)$ 的灰度为 i 的像素出发，统计与距离为 $\delta = (\alpha x^2 + \alpha y^2)^{1/2}$ ，灰度为 j 的像素同时出现的概率 $P(i, j, \delta, \theta)$ 。用数学表达式则为：

$$P(i, j, \delta, \theta) = \{(x, y), (x + \alpha x, y + \alpha y) \mid f(x, y) = i, f(x + \alpha x, y + \alpha y) = j\}$$



灰度共生矩阵的像素对



- 分 方 向
- 构 造 矩 阵
- 统 计 个 数
- 形 成 特 征 向 量

http://blog.csdn.net/cqq_37059



1. 0° 方向灰度共生矩阵

当 $\theta = 0^\circ$, $\alpha_x = 1, \alpha_y = 0$ 由于所给图像中只有4个灰度级, 因此所求得的灰度共生矩阵的大小为 4×4 。

可设置有向或无向

↑

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

一幅数字灰度图像

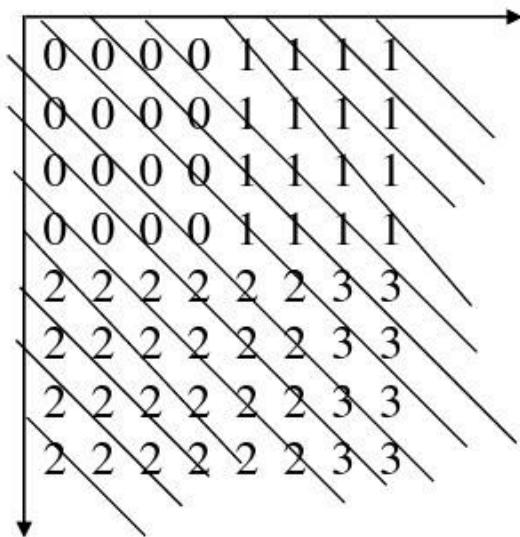
↓

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

0° 方向灰度共生矩阵计算示意图

2. 45° 方向灰度共生矩阵

当 $\theta = 45^\circ$ 时, $\alpha_x = 1, \alpha_y = -1$ 。



45° 方向灰度共生矩阵计算示意图



3. 90° 方向灰度共生矩阵

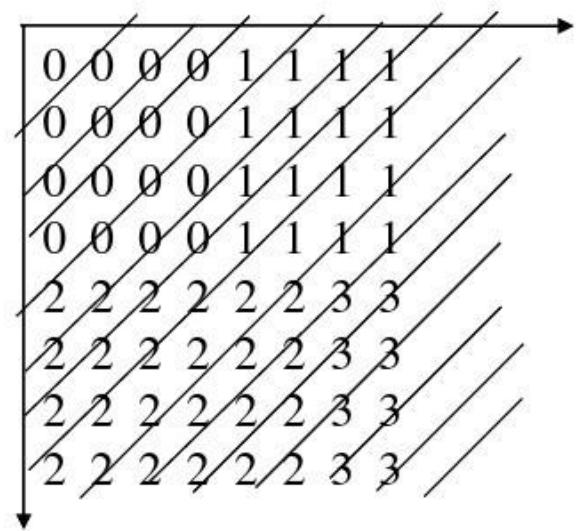
当 $\theta = 90^\circ$ 时, $d_x = 0, d_y = -1$ 。

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

90° 方向灰度共生矩阵计算示意图

4. 135° 方向灰度共生矩阵

当 $\theta = 135^\circ$ 时, $\alpha_x = -1, \alpha_y = -1$ 。



135° 方向灰度共生矩阵计算示意图

灰度共生矩阵计算结果

$$P(0^\circ) = \begin{bmatrix} 12 & 4 & 0 & 0 \\ 4 & 12 & 0 & 0 \\ 0 & 0 & 20 & 4 \\ 0 & 0 & 4 & 4 \end{bmatrix}$$

(a)

$$P(45^\circ) = \begin{bmatrix} 9 & 3 & 4 & 0 \\ 3 & 9 & 1 & 2 \\ 4 & 1 & 15 & 3 \\ 0 & 2 & 3 & 3 \end{bmatrix}$$

(b)

$$P(90^\circ) = \begin{bmatrix} 12 & 0 & 4 & 0 \\ 0 & 12 & 2 & 2 \\ 4 & 2 & 18 & 0 \\ 0 & 2 & 0 & 6 \end{bmatrix}$$

(c)

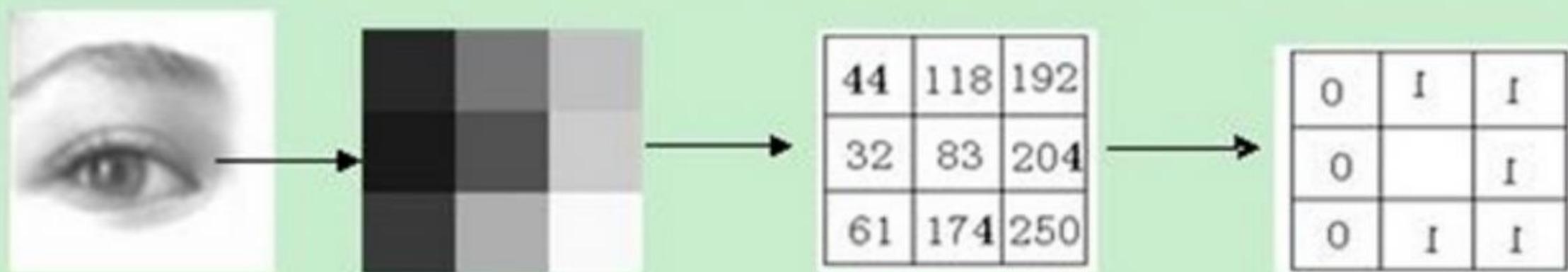
$$P(135^\circ) = \begin{bmatrix} 9 & 3 & 3 & 0 \\ 3 & 9 & 3 & 1 \\ 3 & 3 & 15 & 3 \\ 0 & 1 & 3 & 3 \end{bmatrix}$$

(d)

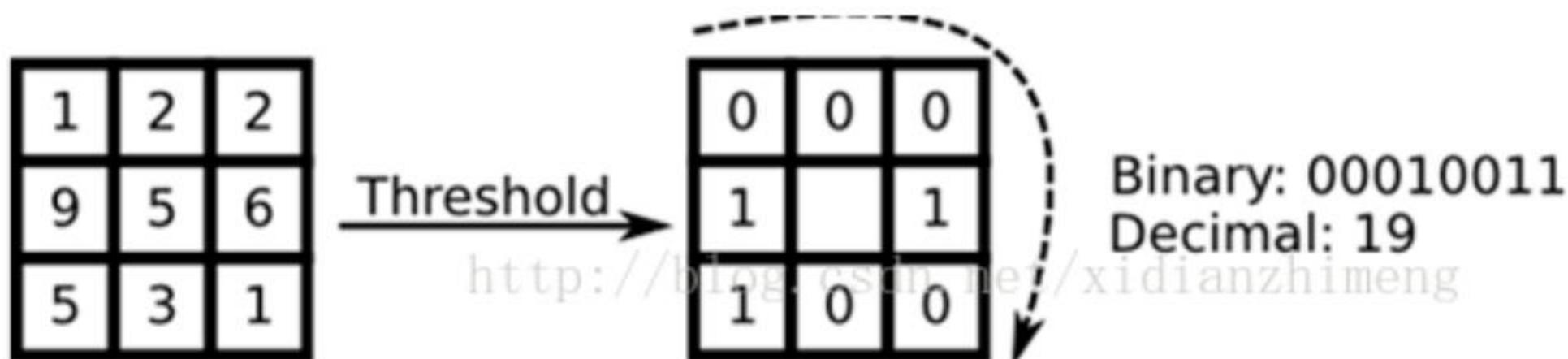
局部二值模式特征 (Local Binary Pattern, LBP)

- LBP算子定义为在 $3*3$ 的窗口内，以窗口中心像素为阈值，将相邻的8个像素的灰度值与其进行比较，若周围像素值大于中心像素值，则该像素点的位置被标记为1，否则为0。这样， $3*3$ 邻域内的8个点经比较可产生8位二进制数（通常转换为十进制数即LBP码，共256种），即得到该窗口中心像素点的LBP值，并用这个值来反映该区域的纹理信息（例如亮点和暗点）
- 最终得到LBP图，然后计算直方图





$$(01111100)_{10} = 124$$



视觉词袋 (Bag-of-words, BOW)

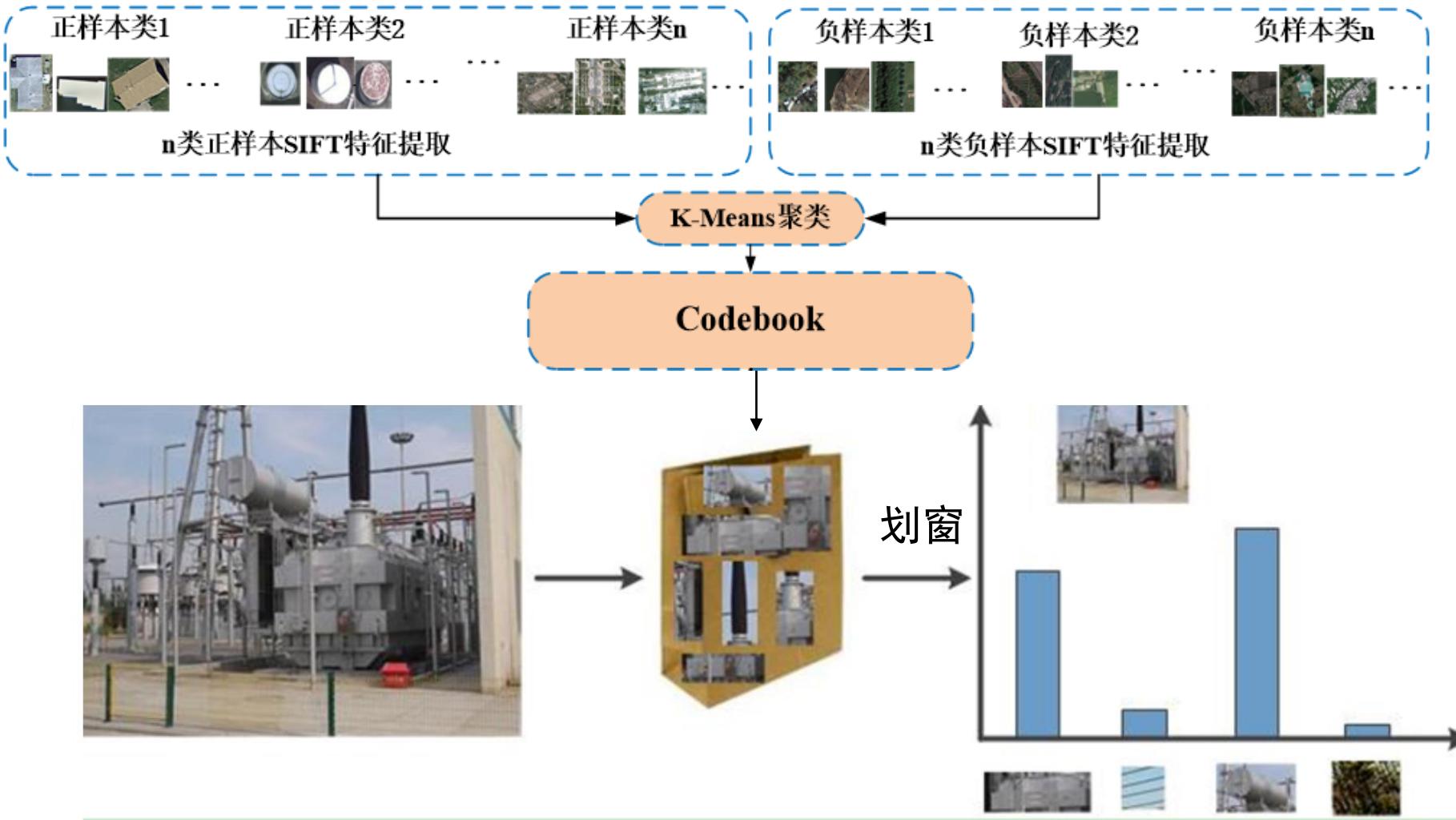
词袋是一种统计关键词出现频率的文档表示方法，2004年被引入计算机视觉领域。词袋模型将图像局部特征空间进行量化，得到一组离散的具有代表性的特征向量，成为视觉单词，视觉单词的集合称为视觉词典，通常采用的视觉词典的构建方法是利用K-means算法对局部特征进行聚类，每个聚类中心对应于一个视觉词汇，视觉词袋方法将图像中的局部特征匹配到距离最近的一个视觉词汇，用视觉词汇出现频率的直方图表示图像。

特征袋 (Bag-of-Features, BoF) :

该模型将多幅图像作为文档集合，分别提取出文档集合中单个文档的局部特征，然后采用聚类算法生成码书 (codebook)，每个聚类的中心即代表一个视觉词汇，码书则表示为视觉词汇的集合，通过量化提取的输入图像特征来计算属于码书中视觉词汇的统计直方图，即为图像的特征袋模型。



基于特征袋模型的分类器



图像

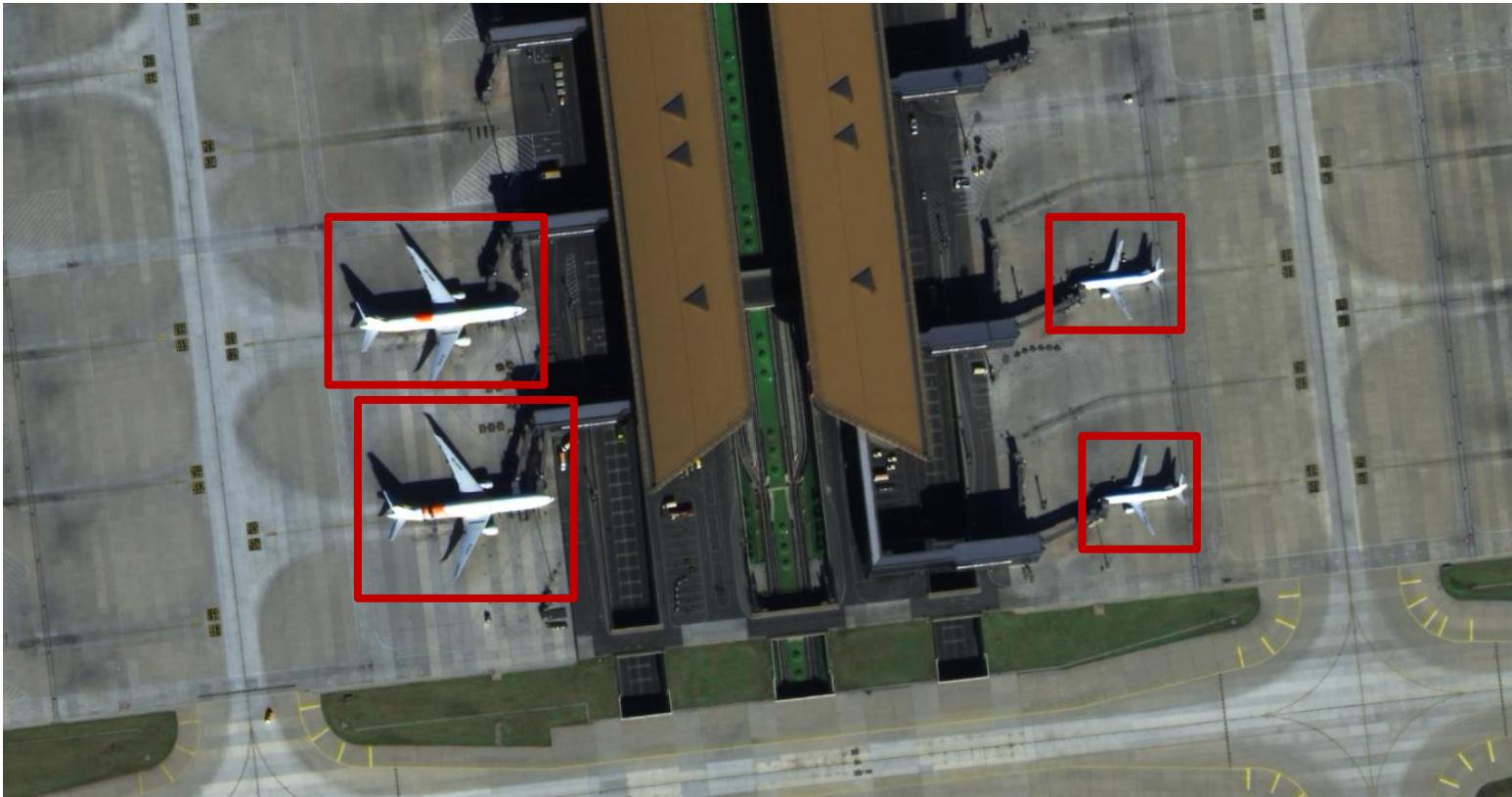
视觉词汇字典

直方图



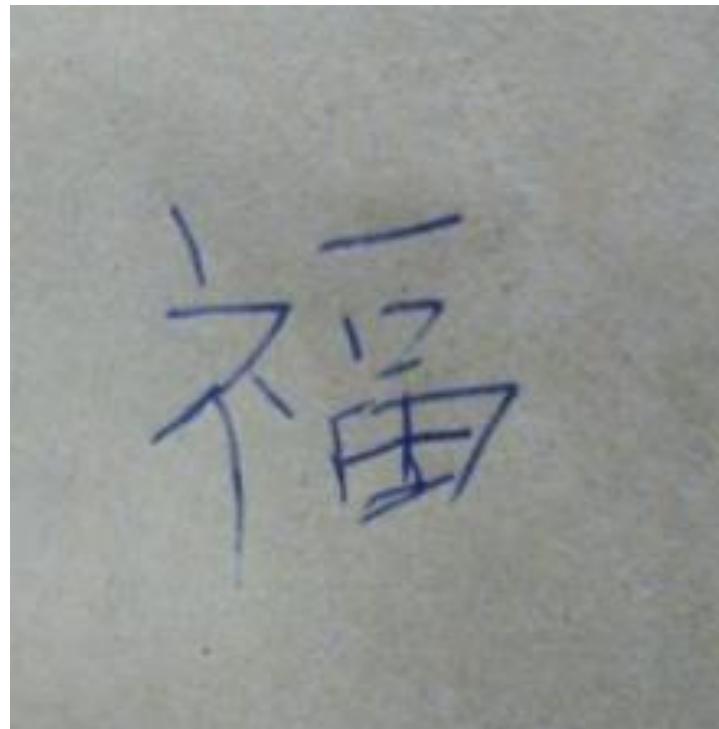
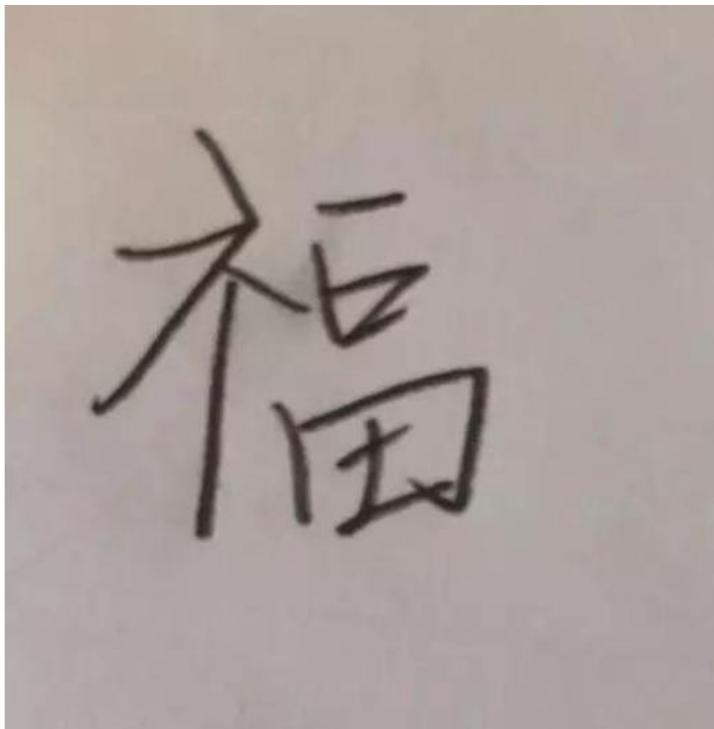
课程大作业（四选一）

■任务1：遥感图像飞机检测



课程大作业（四选一）

- 任务2：“福”字识别-解决类别不平衡问题



课程大作业（四选一）

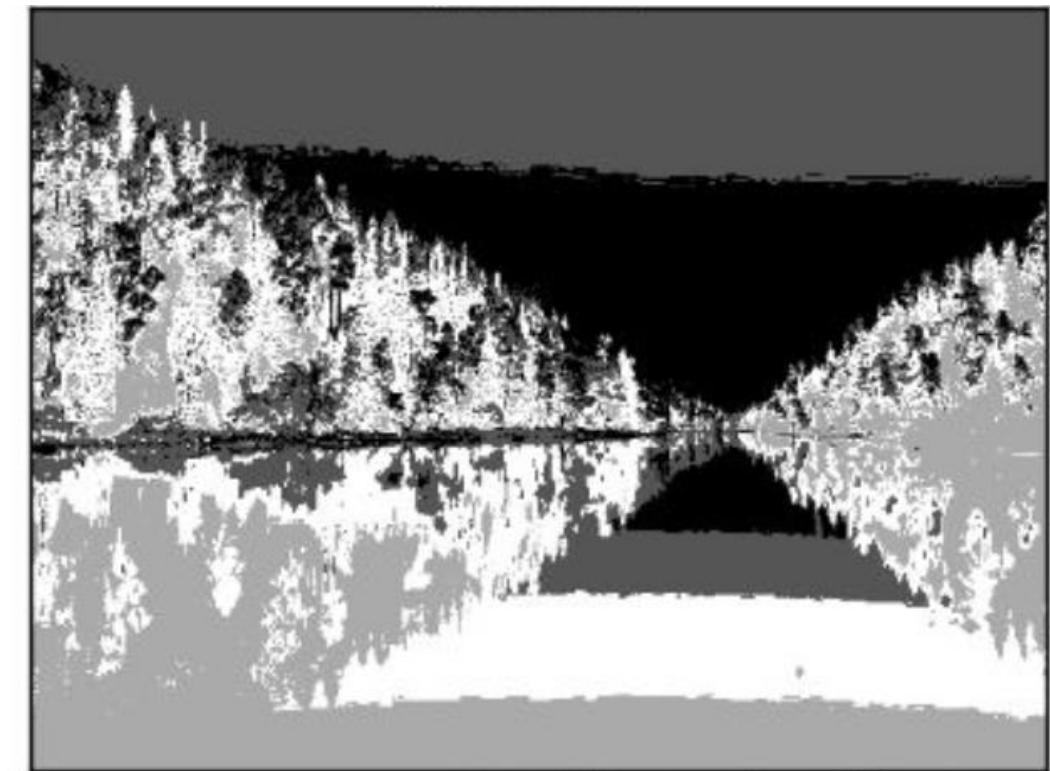
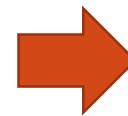
■任务3：台风预报



66666 1815	21	0018	1815	0	6	LEEPI
2018081012	1	174	1448	1006	13	
2018081018	1	179	1444	1004	15	
2018081100	1	185	1441	1004	15	
2018081106	1	192	1438	1004	15	
2018081112	2	200	1435	1000	18	
2018081118	2	208	1432	995	20	
2018081200	2	217	1427	990	23	
2018081206	2	227	1420	990	23	
2018081212	3	235	1415	982	28	
2018081218	3	245	1407	982	28	
2018081300	3	253	1398	982	28	

课程大作业（四选一）

- 任务4：图像区域分割提取-如何保持空间相关性？



大作业1：福字识别



提供矩阵形式





THE END!

