

图象工程（中）

# 图 象 分 析

（第4版）

章毓晋

清华大学电子工程系 100084 北京

# 第3单元 特性分析

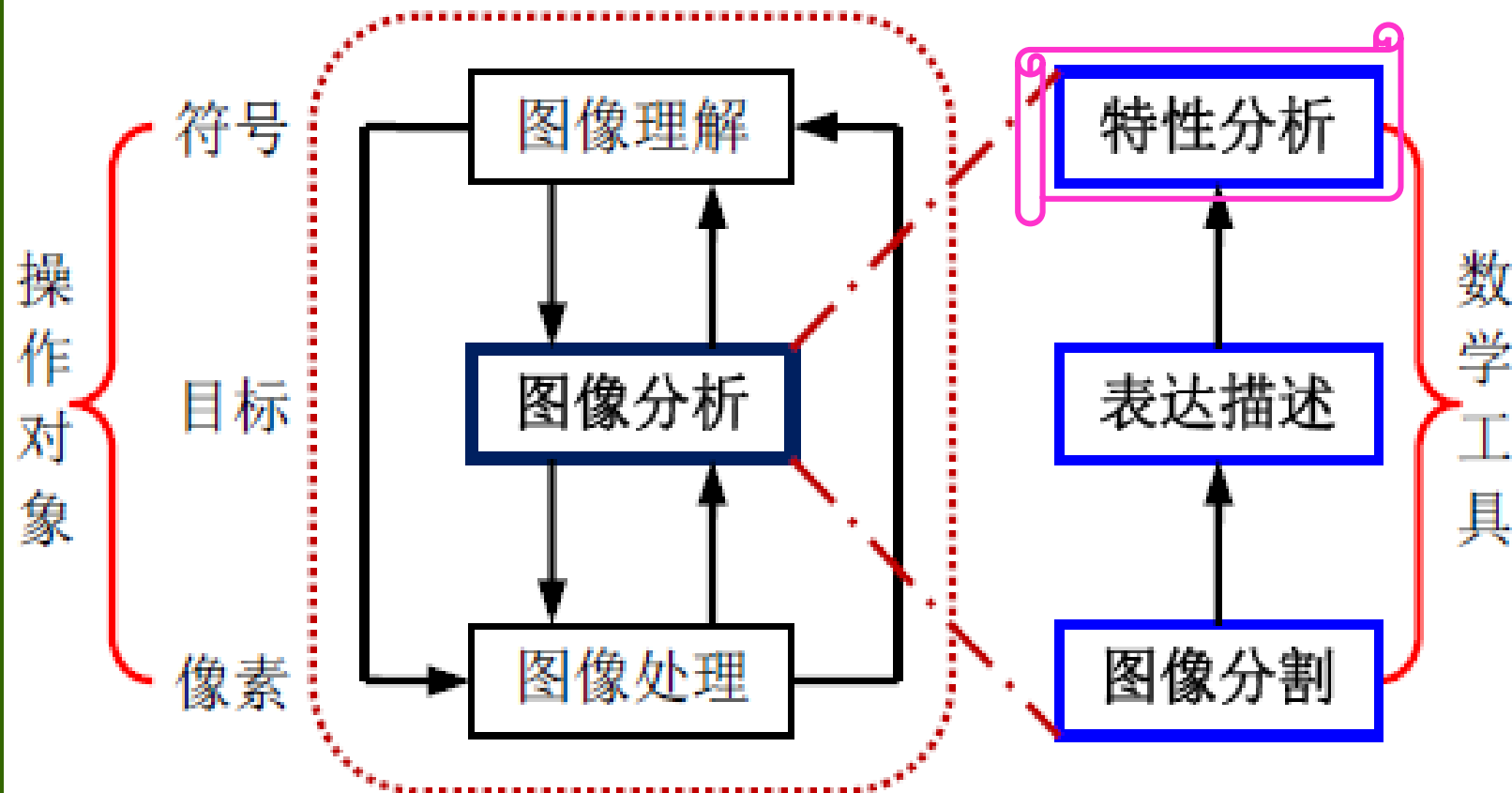
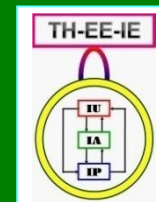
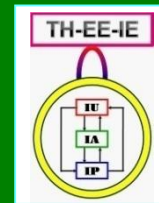


图 1.2.2 图像分析主要功能模块



# 第3单元 特性分析

- 第9章 纹理分析：对纹理的表达和描述，纹理图象分割，分类和合成
- 第10章 形状分析：特征的方法，形状变换的方法，基于关系的方法
- 第11章 运动分析：景的变化（外在照明改变、视场改变）和物的变化（目标位置朝向改变）
- 第12章 显著性和属性  
显著性：与主观感知相关联  
属性：人指定名称并可观察到的特性



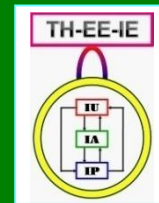
# 第12章 显著性和属性

## 中层分析

对纹理、形状、运动等的分析都涉及对多个像素组合而成的连通单元（片、目标、区域）的分析，类似层次：显著性、属性、超像素、部件等

显著性是一个与主观感知相关联的概念，视觉显著性常归因于场景区域在底层特性方面的变化或对比而导致的综合结果

属性是指可以由人指定名称并且能在图象中观察到的特性，它提供了有价值的新的语义线索



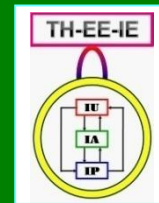
# 第12章 显著性和属性

## 中层分析

显著性与属性有一些共性，它们都在中间语义层次上表达了图象的特性，且都有一定的主观性，所以互相之间也有许多联系

图象中的显著极值通常与图象中的主要结构成分密切相关。而属性形态学分析就为图象显著极值的提取提供了一种有效的解决方案

提取图象的显著极值关键在于如何根据极值的区域形态属性对极值的显著性进行评估



# 第12章 显著性和属性

12.1 显著性概述

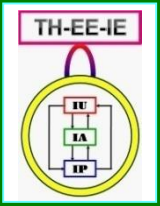
12.2 显著性检测

12.3 显著区域分割提取

12.4 属性概述

12.5 属性提取中的特征比较

12.6 属性应用



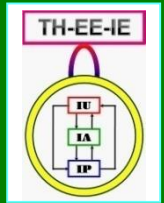
# 12.1 显著性概述

显著性指能使一个特征、图象点、图象区域或目标的鉴别性或相对于其环境更显眼的量/特性

## 1. 显著性的内涵

显著性与人对世界的关注或注意有关。关注或注意是一个心理学概念，是心理过程的一种具有共性的特征，属于认知过程的内容

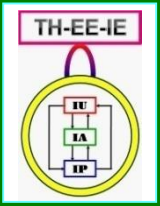
作为一个中层的语义线索，显著性可帮助填补低层特征和高层类别间的鸿沟



# 12.1 显著性概述

## 2. 显著区域的特点

- (1) 高层语义特征：人在观察中经常注意到的对象（如人脸、汽车等）经常对应图象中的显著区域，本身具有一定的认知语义含义
- (2) 认知稳定性：显著区域对场景亮度、对象位置、朝向、尺度、以及观察条件等比较鲁棒
- (3) 全局稀缺性：从全局范围来看，显著区域出现的频率比较低（局部、稀少），且不容易由图象中的其他区域复合而得到
- (4) 局部差异性：显著区域总是与周围区域具有明显的特性差异

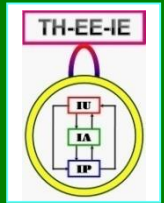


# 12.1 显著性概述

## 3. 显著图质量的评价

显著图反映了图象中各部分吸引人注意的程度  
显著图的质量是评价显著性检测算法好坏的重要标准，也与显著区域的特点密切关联

- (1) 能突出最为显著的物体
- (2) 能使整个显著物体各部分具有比较一致的突出程度
- (3) 能给出精确完整的显著物体边界
- (4) 能给出全分辨率的检测结果



# 12.1 显著性概述

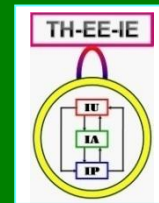
## 4. 视觉注意力机制和模型

视觉注意力机制有两个基本特征：指向性（表现为对出现在同一时段的多个刺激有选择性）和集中性（表现为对干扰性刺激的抑制能力）

注意力机制主要分为两大类：自底向上数据驱动的预注意机制和自顶向下任务驱动的后注意机制

用计算机来模拟人类视觉注意力机制的模型称为视觉注意力模型

一个视觉注意力模型对图象中的显著区域用视觉注意力（VA）图来表示



## 12.2 显著性检测

### 1. 显著性检测方法分类

- (1) 对图象信号的处理：在空域/变换域
- (2) 检测算法的流程：自底向上/自顶向下
- (3) 考虑计算对象：基于注视点/显著区域
- (4) 检测结果分辨率：像素级的方法和基于区域（包括超像素）的方法
- (5) 考虑辅助信息，将检测方法分为仅利用图象自身信息的内部方法和还利用图象“周边”信息的外部方法
- (6) 显著性与主观感知相关联，所以检测方法除可借助计算模型也可考虑仿生学的方法

## 12.2 显著性检测

### 2. 基本检测流程

主要有5个模块

输入和输出可有不同的形式

其他3个模块也列出了一些常用的方法和技术

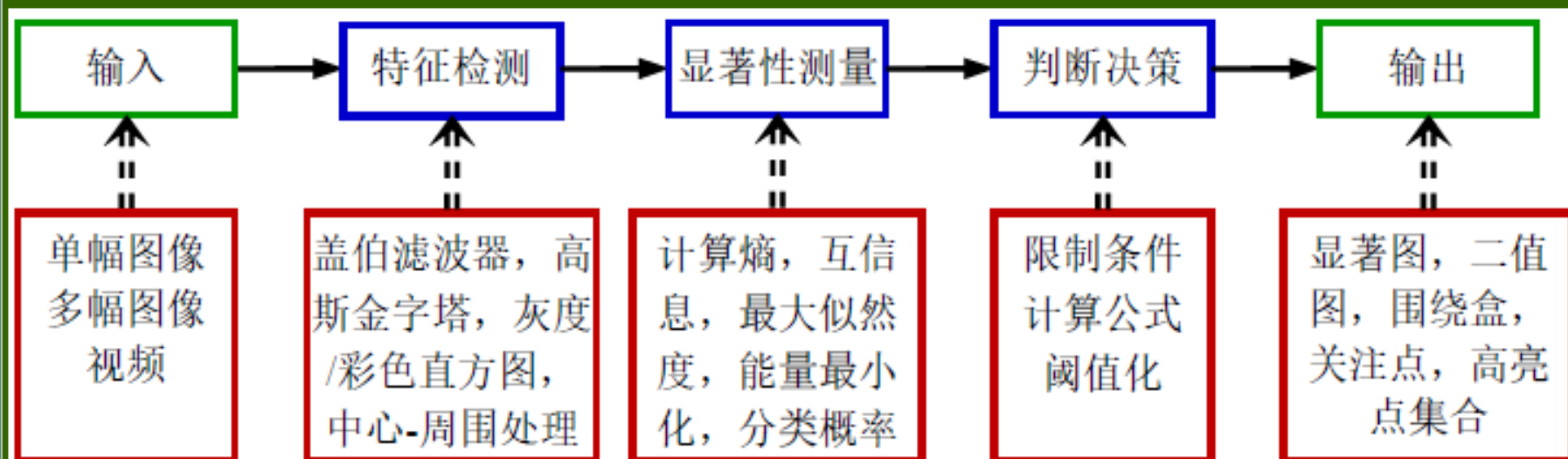
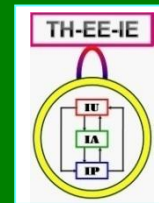


图 12.2.1 显著性检测流程



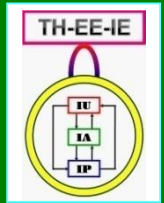
## 12.2 显著性检测

### 3. 对比度检测

现有的显著区域检测算法可分为两类：基于局部对比的算法和基于全局对比的算法

计算对比度的方法主要分为3大类：

- (1) 利用局部对比度先验知识  
将每个像素仅仅与图象局部中某些像素比较
- (2) 利用全局对比度先验知识  
将目标像素与图象中其余所有像素进行特征差异度计算
- (3) 利用背景先验知识  
背景先验：图象的四周对应背景



## 12.2 显著性检测

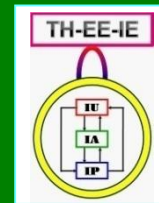
### 4. 显著区域提取

在对显著性进行检测的基础上，可进一步将显著区域提取出来：

- (1) 直接阈值分割
- (2) 基于交互图象分割
- (3) 结合矩形窗定位

显著区域的提取通常包括以下步骤：

- (1) 显著图计算
- (2) 初始显著区域定位
- (3) 精细显著区域提取



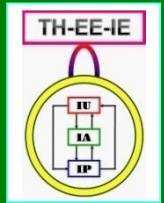
## 12.3 显著区域分割提取

12.3.1 基于对比度幅值

12.3.2 基于对比度分布

12.3.3 基于最小方向对比度

12.3.4 目标分割和评价



## 12.3.1 基于对比度幅值

### 1. 基于直方图对比度的算法

一个像素的显著性是该像素与图象中所有其他像素之间的颜色/灰度对比度之和

$$S(I_i) = \sum_{\forall I_j \in I} D_c(I_i, I_j)$$

$$S(I_i) = D_c(I_i, I_1) + D_c(I_i, I_2) + \cdots + D_c(I_i, I_N)$$

具有相同颜色/灰度的像素会具有相同的显著性值，将对应具有相同颜色/灰度值 $c_i$ 的像素 $I_i$ 分在同一组

$$S(I_i) = S(c_i) = \sum_{j=1}^C p_i D(c_i, c_j)$$

对应直方图表达，计算快捷

## 12.3.1 基于对比度幅值

### 2. 基于区域对比度的算法

将输入图象初步分割成多个区域，然后通过计算区域对比度（RC）获取每个区域的显著性值

$$S(R_i) = \sum_{R_i \neq R_j} W(R_i) D_c(R_i, R_j)$$

$D_c(R_i, R_j)$  度量区域之间在  $L^*a^*b^*$  空间中的颜色距离

$$D_c(R_i, R_j) = \sum_{k=1}^{C_i} \sum_{l=1}^{C_j} p(c_{i,k}) p(c_{j,l}) D(c_{i,k}, c_{j,l})$$

引入空间信息以加强邻近区域的权重并减少远离区域的权重

$$S(R_i) = W_s(R_i) \sum_{R_i \neq R_j} \exp \left[ -\frac{D_s(R_i, R_j)}{\sigma_s^2} \right] W(R_i) D_c(R_i, R_j)$$

## 12.3.2 基于对比度分布

### 整体流程

- (1) 对图象进行超像素分割
- (2) 计算超像素的最大环绕对比度
- (3) 计算相对对比度方差
- (4) 将该方差数值转换为显著性值

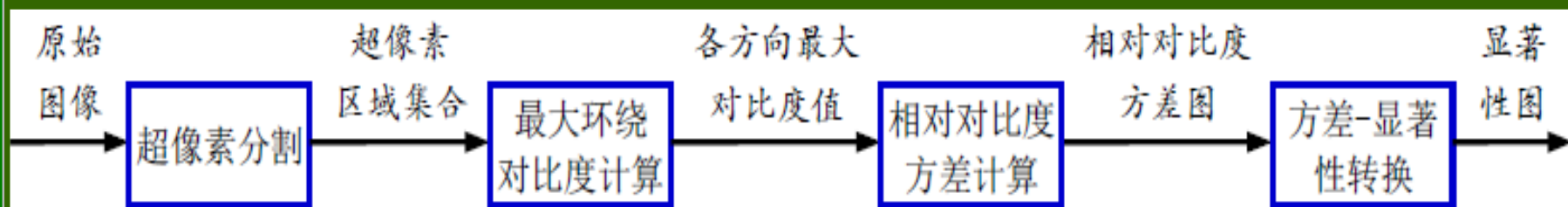
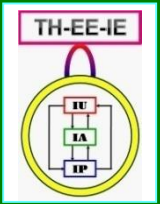


图 12.3.2 基于对比度分布算法的流程图



## 12.3.2 基于对比度分布

### 显著性定义

一个超像素区域的显著性与该区域与其他区域的彩色对比度成比例，且与该区域与其他区域的空间距离也成比例

$$S(R_i, R_j) = \exp\left[\frac{D_s(R_i, R_j)}{-\sigma_s^2}\right] D_c(R_i, R_j)$$

$D_c(R_i, R_j)$ 为区域 $R_i$ 和 $R_j$ 之间的颜色距离， $D_s(R_i, R_j)$ 为区域 $R_i$ 和 $R_j$ 之间的空间距离， $\sigma_s$ 控制空间距离加权的强度（其值越大则远离区域 $R_i$ 的其他区域对其的加权强度越大）

## 12.3.2 基于对比度分布

### 对比度分布计算

对属于内部的超像素区域，其各个方向的最大对比度值都比较大；而对属于外部的超像素区域，其各个方向的最大对比度值有比较大的差别

在显著物体内部

在显著物体外部

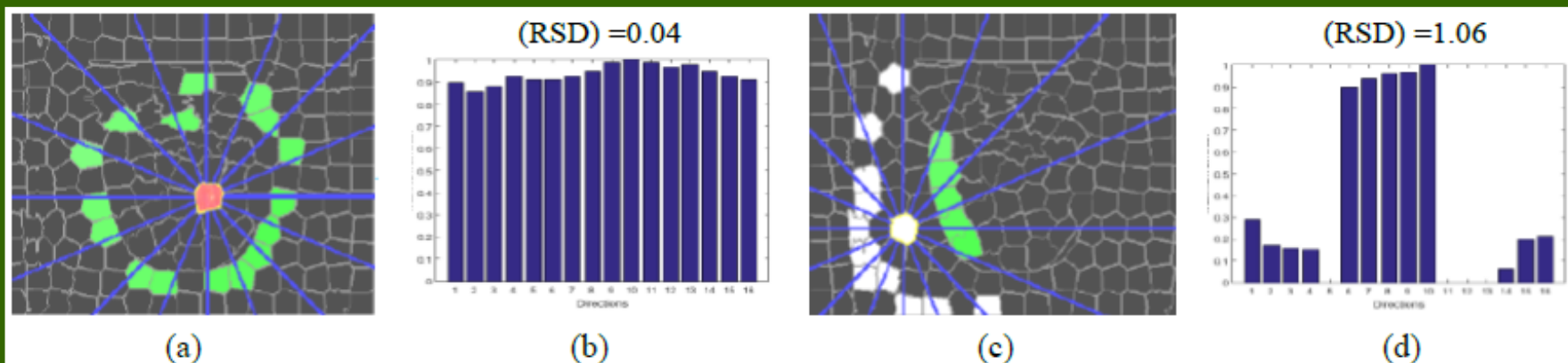


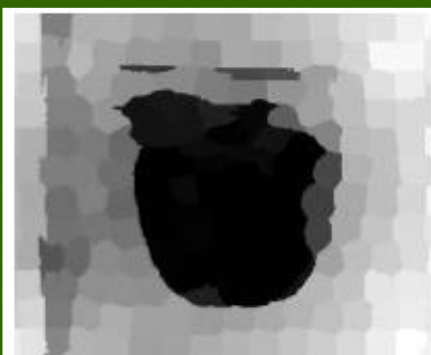
图 12.3.3 对比度分布的计算示意图

## 12.3.2 基于对比度分布

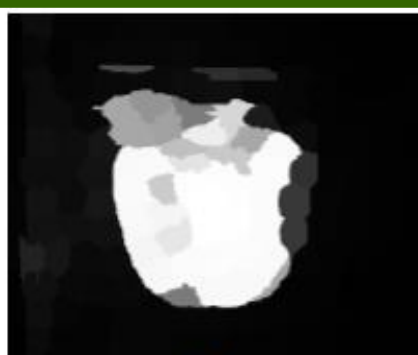
### 显著性图

属于显著物体内部的超像素区域的相对标准方差（RSD）会比属于显著物体外部的超像素区域的相对标准方差小

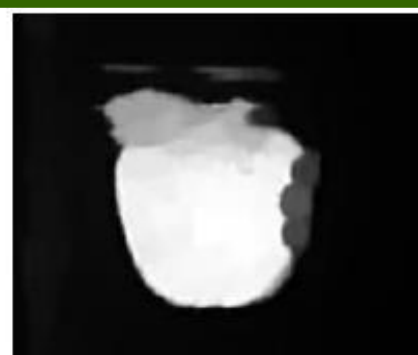
相对标准方差      显著性图      后处理结果      真值



(a)



(b)

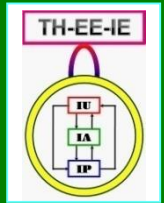


(c)



(d)

图 12.3.4 根据对比度分布得到的显著性图



## 12.3.3 基于最小方向对比度

### 1. 最小方向对比度

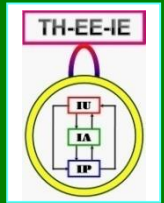
来自各个区域的方向对比度 (DC)

$$DC_{i,H_l} = \sqrt{\sum_{j \in H_l} \sum_{k=1}^K (f_{i,k} - f_{j,k})^2} \quad l=1,2,3,4$$

前景与背景像素的方向对比度分布很不一样

用最小方向对比度 (MDC)，即所有方向中对比度最小的DC值，作为初始的显著性的度量值：

$$S_{\min}(i) = \min_{H_l} DC_{i,H_l} = \sqrt{\min_{H_l} \sum_{j \in H_l} \sum_{k=1}^K (f_{i,k} - f_{j,k})^2} \quad l=1,2,3,4$$



## 12.3.3 基于最小方向对比度

### 2. 降低计算复杂度

对每个像素计算MDC，需要计算4个方向的对  
比度。如果直接计算，则计算复杂度将是 $O(N)$ ，  
其中 $N$ 是整幅图的像素个数

$$\sum_{j \in H} \sum_{k=1}^K (f_{i,k} - f_{j,k})^2 = \sum_{j \in H} \sum_{k=1}^K f_{i,k}^2 - 2 \sum_{k=1}^K \left\{ \sum_{j \in H} f_{j,k} \right\} f_{i,k} + |H| \sum_{k=1}^K f_{i,k}^2$$

上式中等号右边第1项以及第2项里花括号中的部  
分都可借助积分图象来计算，这样一来，对每个  
像素的计算复杂度将可以减到 $O(1)$

## 12.3.3 基于最小方向对比度

### 3. 显著性平滑

为了降低计算复杂度，对颜色进行量化

简单的颜色量化，可能会导致两种相似的颜色量化到两种取值，进而得到的显著性在视觉上不连续。借助边界连通性的平滑来消除这种问题



## 12.3.3 基于最小方向对比度

### 4. 显著性增强

增加前景和背景区域之间的对比度  
使用基于标记的分水岭算法增强显著性  
前景像素为红色、背景像素为绿色

平滑结果      初始标记区域      最终标记区域      真值



(a)



(b)

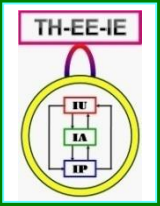


(c)



(d)

图 12.3.7 显著性检测后处理

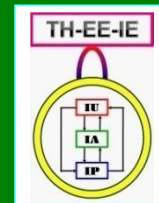


## 12.3.4 显著目标分割和评价

### 1. 目标分割和提取

对图割方法改进，以分割显著性图中的目标

- (1) 用一个固定的阈值对显著性图进行二值化
- (2) 将显著性值大于阈值的最大连通区域作为显著目标的初始候选区域
- (3) 将这个候选区域标记为未知，而把其他区域都标记为背景
- (4) 利用标记为未知的候选区域训练前景颜色以帮助算法确定前景像素
- (5) 用能给出高查全率（召回率）的潜在前景区域来初始化GrabCut算法，并迭代优化以提高查准率（精确度）
- (6) 迭代执行GrabCut算法，并使用膨胀和腐蚀（见下章）



## 12.4 属性描述概况

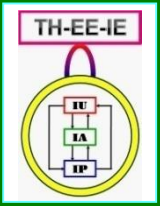
### 1. 属性的类型

从描述对象看，可将属性分成3种类型：

- (1) 二进制属性：如有轮子/没有轮子
- (2) 相对属性：如脸更宽，腿更长
- (3) 语言属性：二进制和相对属性的组合

从描述功能看，可将属性分成3种类型：

- (1) 语义属性：用于标注某类物体
- (2) 鉴别性属性：某类物体区分性的性质
- (3) 可命名属性：可明确被语言表达的属性



## 12.4 属性描述概况

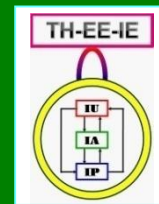
### 2. 属性的层次

属性往往看作是一种含有语义信息的中层特征。它联结了底层的基本特征与高层的实物模型

属性是底层特征的抽象和总结，是对象的一般化高层描述

属性比底层特征含有更丰富的语义信息，可被人类感知、解释，因而更方便与先验知识结合

属性与高层抽象的文本标签有相容但有不同的关系，不是所有文本标签都可被称作属性



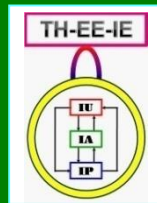
## 12.4 属性描述概况

### 3. 属性学习结构/学习框架

属性学习的过程是建立底层特征 $\Leftrightarrow$ 属性、属性 $\Leftrightarrow$ 高层语义的关系

在自然场景中，可以把视觉属性当作一组坐标基，针对每一种属性，训练该属性对应的预测器，然后把图象中视觉属性的预测分数当作该图象的坐标，从而得到图象在视觉属性基下的坐标

引入属性概念把在目标识别时的关注焦点从辨识独特的物体，向描述物体的属性发生转移

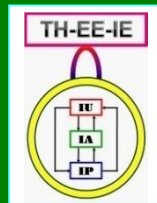


## 12.5 属性提取中的特征比较

对图象提取其属性  $\Rightarrow$  属性学习或属性分类  
一般涉及3个步骤的工作：① 特征提取；② 特征选择；③ 分类器训练和预测

### 1. 特征和特征组合

- (1) 彩色特征：彩色直方图
- (2) 纹理特征：纹理基元、局部二值模式
- (3) 形状特征：边缘方向直方图、SIFT
- (4) 混合特征：彩色尺度不变特征变换SIFT
- (5) 深度特征：卷积神经网络（CNN）特征



## 12.5 属性提取中的特征比较

### 1. 特征和特征组合

对特征进行组合或融合。常可分3种方式

#### (1) 早期融合

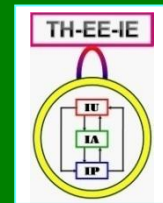
连接低层特征矢量，训练分类器

#### (2) 晚期融合

构建对应的分类器，将特征分类器结合起来计算预测分数

#### (3) 加权晚期融合

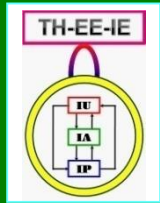
对每个分类器的分数给予不同的权重



## 12.5 属性提取中的特征比较

### 2. 实验数据库

- (1) a-Pascal: 对Pascal VOC 2008数据库增加属性标记而得到, 20个目标类别, 64个属性
- (2) a-Yahoo: 对Yahoo数据库增加属性标记而得到, 12个目标类别, 64个属性
- (3) Shoes: 源于一个属性发现数据库, 4种商品类别(仅取鞋), 10个属性
- (4) People: 是一个用于人的属性识别的数据库, 9个属性



## 12.5 属性提取中的特征比较

### 3. 实验结果和分析

#### (1) 单独特征的性能

总体上CNN特征最好，HOG和CSIFT特征性能也与属性类别有关

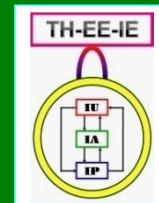
同一类特征的性能与数据库有关

#### (2) 低层特征组合的性能（表12.5.1 →）

#### (3) 低层特征与CNN特征组合的性能

晚期融合比早期融合的效果好

加权晚期融合比晚期融合还要略好一些



## 12.6 属性应用

12.6.1 跨类目标分类

12.6.2 属性学习和目标识别

12.6.3 基于局部动作属性的  
动作分类

# 12.6.1 跨类目标分类

考虑用不同的训练集和测试集进行学习并实现分类的问题

## 1. 传统分类器

对于训练集中出现的类别  $x_1, \dots, x_M$ ，可以通过训练出参数向量  $s_1, \dots, s_M$ ，并构建分类器来完成分类。然而对于类别  $z_1, \dots, z_N$ ，由于它们没有出现在训练集中，无法得到相应的参数向量来构建分类器

训练集和测试集中都存在      训练集中不存在而测试集中可能存在

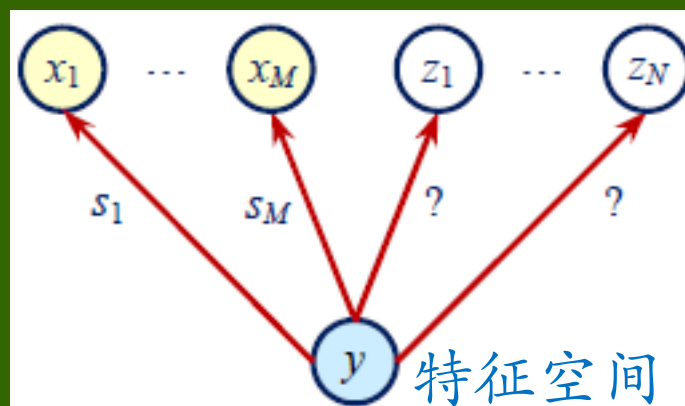
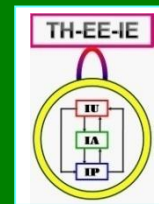


图 12.6.1 传统分类器图示



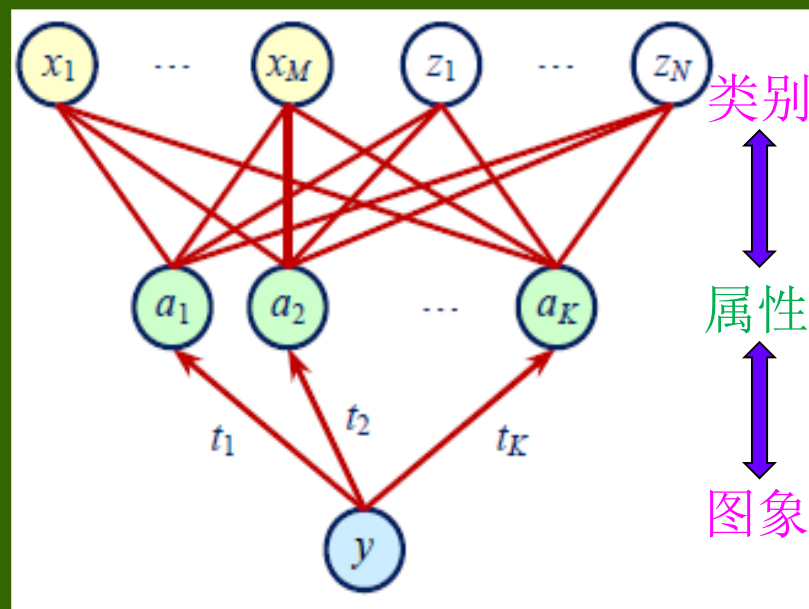
## 12.6.1 跨类目标分类

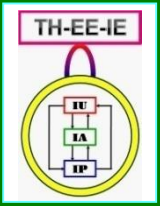
### 2. 利用属性跨类学习

#### (1) 直接属性预测模型

使用属性变量之间的层以将图象层与标签层耦合。在测试时对每个测试样本的属性值进行预测，并推断测试类的标签

这里测试的类别可以不同于用于训练类别，只要属性层不需要训练





## 12.6.1 跨类目标分类

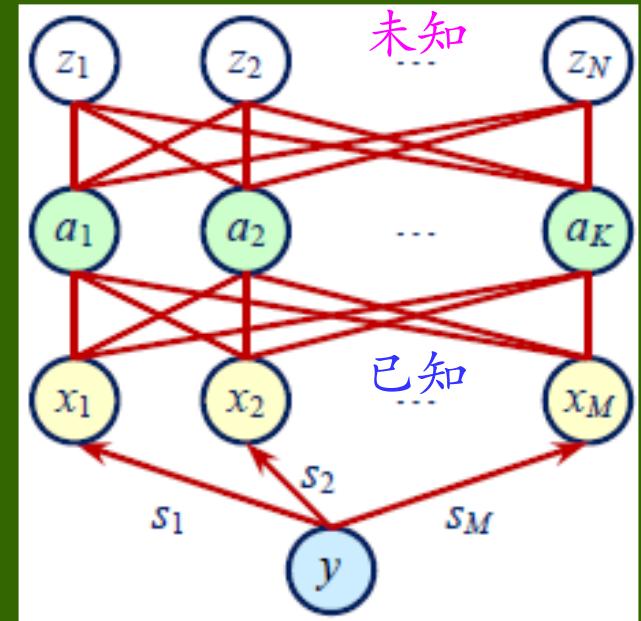
### 2. 利用属性跨类学习

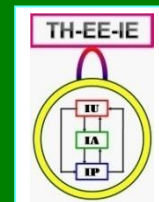
#### (2) 间接属性预测模型

属性构成两层标签之间的连接层，这两个层一个是训练时已知的而另一个则是训练时未知的

它也使用属性来传递类别间的知识

训练阶段是普通的多类分类，测试时推断测试类的标记





## 12.6.2 属性学习和目标识别

### 1. 基于属性的目标识别框架

从特征出发借助支持向量机获得语义属性和鉴别性属性，从而计算出针对任意一幅测试图象的属性值，获得关于测试图象的中层“特征”，根据属性与目标类别间的联系最终得到高层的目标分类结果

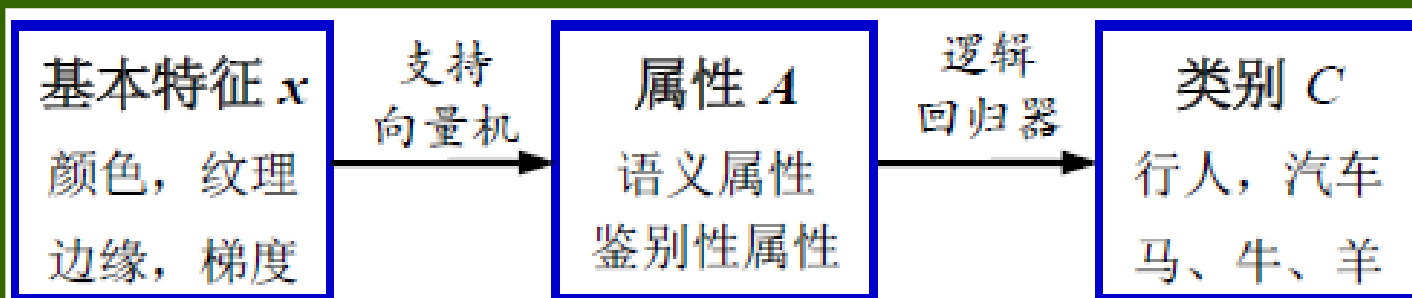


图 12.6.3 基于属性的目标识别框架和主要步骤

## 12.6.2 属性学习和目标识别

### 2. 基于属性描述的属性学习框架

使用属性分类器，除了进行目标识别外，还可以用属性描述未知的目标类别，指出非典型的属性，并从很少的样本中学习新的类别，甚至实现（仅用文本描述的）零样本学习

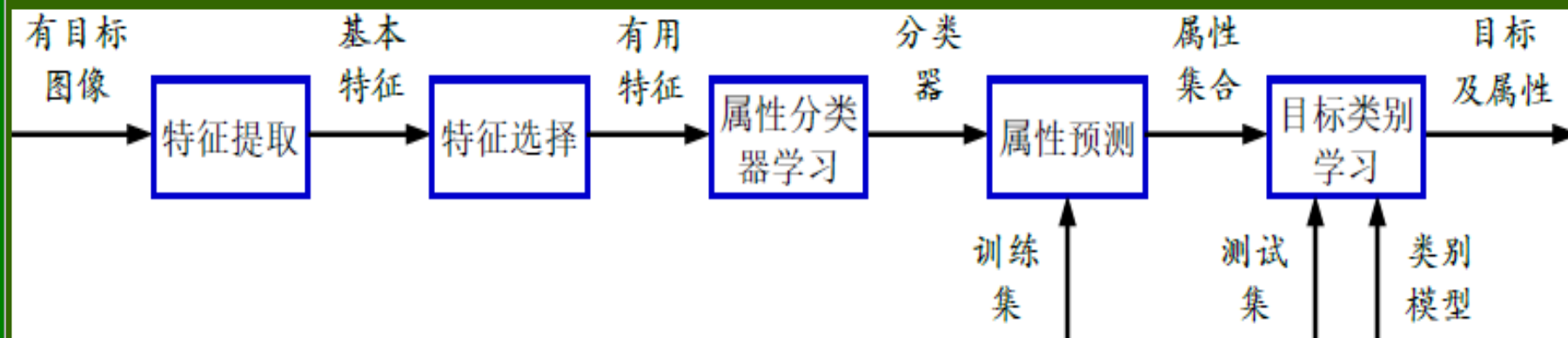


图 12.6.4 一种属性学习框架和主要步骤

## 12.6.2 属性学习和目标识别

### 3. 鉴别性可命名属性模型

给定一组 $N$ 幅图象以及它们的表达和相关的类别标记。现在希望确定对 $M$ 个属性分类器的中间表达 $A = [a_1, \dots, a_M]$ 。  $A$ 是可命名的，即它有一个语义词与之关联。而且， $A$ 的输出合起来是一个具有鉴别性的分类器

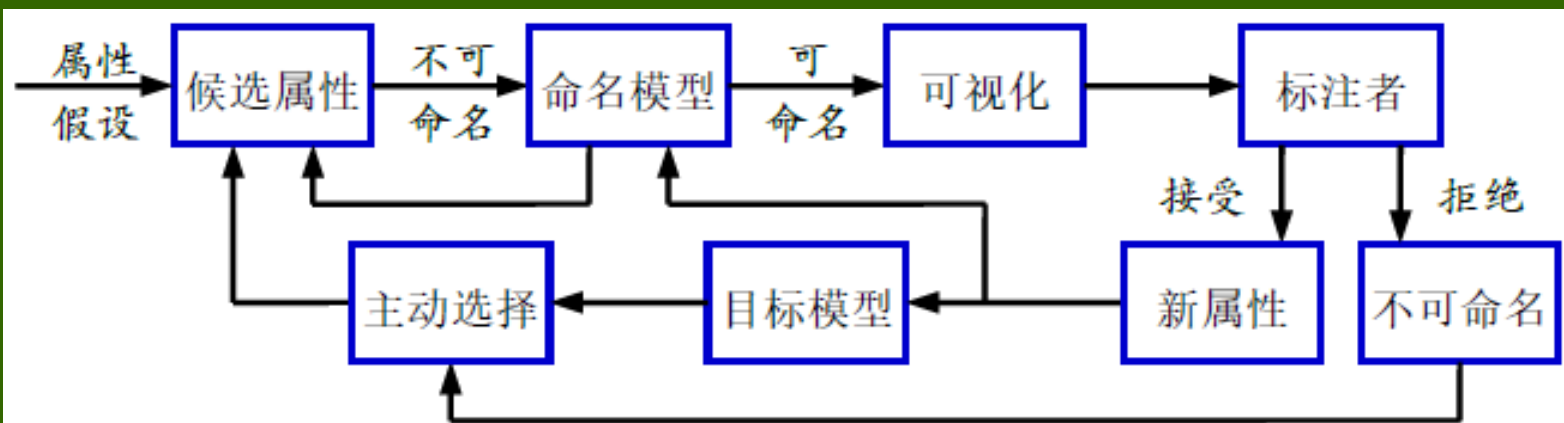
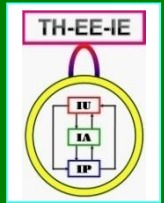


图 12.6.5 一种交互学习属性的框架和主要步骤



## 12.6.3 基于局部动作属性的动作分类

部件是一种像素和目标之间的单元，与属性可以结合使用

先将人体分为若干相互关联的部件，再利用部件的局部动作属性来描述人类行为，就能够构建用于人类行为识别的描述性模型

可以称局部动作属性为“局部部件动作属性”

为了将人体结构信息嵌入到局部动作属性中去，要实现基于部件的低层动作表示来构建属性描述符



# 联系信息

- ☞ 通信地址：北京清华大学电子工程系
- ☞ 邮政编码：100084
- ☞ 办公地址：清华大学，罗姆楼，6层305室
- ☞ 办公电话：(010) 62798540
- ☞ 传真号码：(010) 62770317
- ☞ 电子邮件：[zhang-yj@tsinghua.edu.cn](mailto:zhang-yj@tsinghua.edu.cn)
- ☞ 个人主页：[oa.ee.tsinghua.edu.cn/~zhangyujin/](http://oa.ee.tsinghua.edu.cn/~zhangyujin/)