

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4单元 数学工具

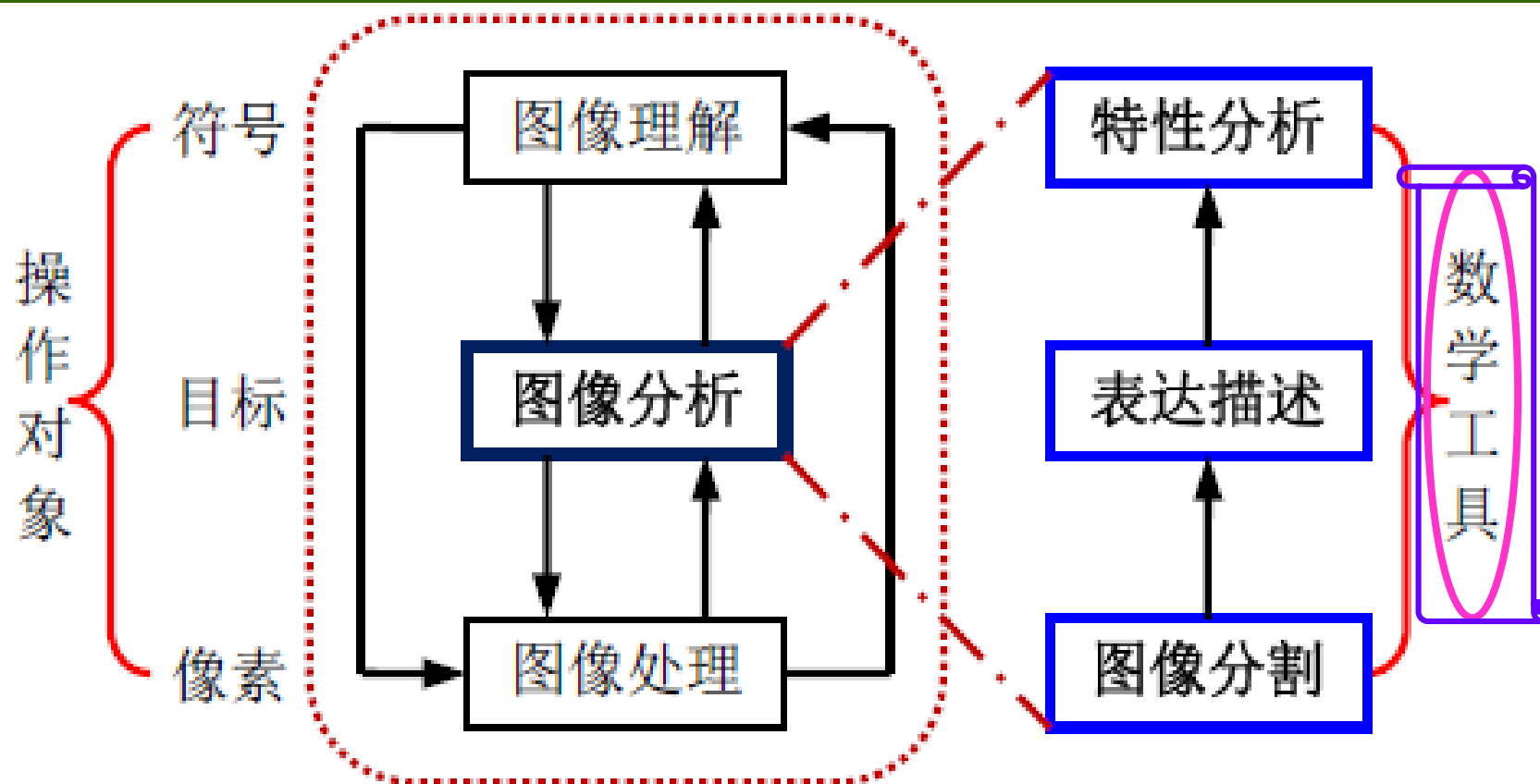
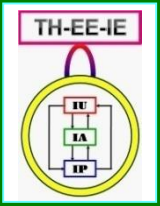


图 1.2.2 图像分析主要功能模块



第4单元 数学工具

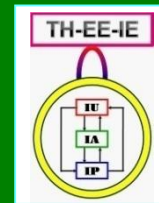
- 第13章 数学形态学：二值
- 第14章 数学形态学：灰度
- 第15章 图象识别

数学及其他学科理论和工具的支持

数学形态学表示以形态为基础对图象进行分析的数学工具

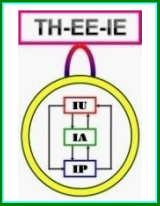
操作对象可以是二值或灰度图象

图象模式识别（简称图象识别）



第15章 图象识别

- 15.1 模式和分类
- 15.2 不变量交叉比
- 15.3 统计模式识别
- 15.4 感知机和支持向量机
- 15.5 结构模式识别

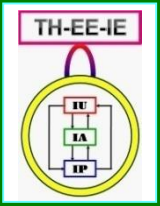


15.1 模式和分类

1. 模式和模式矢量

- 模式是一个广泛的概念，模式由一个或多个模式符（也可叫特征）组成
- 模式类由具有共同特性的模式组成
- 图象模式可定义为对图象中的目标或其它感兴趣部分定量或结构化的表达描述
- 矢量是一种定量描述模式的表达方式

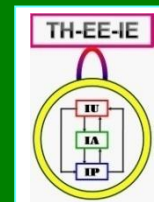
$$\mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_n]^T$$



15.1 模式和分类

2. 模式识别和分类

- (1) **验证** (identification/verification)，对一个事先见过的目标的识别
- (2) **推广** (generalization)，识别一个目标，尽管由于某些变换已使得它的外观发生了变化
- (3) **分类** (categorization)，将目标分到一组相似特性或属性的目标类中去
- (4) **类似/类比** (analogy)，发现不同目标变换后的相似之处



15.1 模式和分类

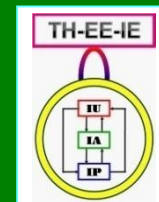
2. 模式识别和分类

- 对给定的 M 个模式类 s_1, s_2, \dots, s_M , 确定 M 个决策函数 $d_1(\mathbf{x}), d_2(\mathbf{x}), \dots, d_M(\mathbf{x})$ 。如果一个模式 \mathbf{x} 属于类 s_i , 则

$$d_i(\mathbf{x}) > d_j(\mathbf{x}) \quad j=1, 2, \dots, M; \quad j \neq i$$

- 如果对 \mathbf{x} 的值, 有 $d_i(\mathbf{x}) = d_j(\mathbf{x})$, 则得到将类 i 与类 j 分开的决策边界

$$d_{ij}(\mathbf{x}) = d_i(\mathbf{x}) - d_j(\mathbf{x}) = 0$$



15.2 不变量交叉比

15.2.1 交叉比

15.2.2 非共线点的不变量

15.2.3 对称的交叉比函数

15.2.4 交叉比应用示例

15.2.1 交叉比

交叉比

比率的比率，或比例的比例

不变量：不随某些变换而改变的量度

4个共线点 P_1, P_2, P_3 和 P_4 在成像变换后给出点 Q_1, Q_2, Q_3 和 Q_4

从任何投影角度来看，
由4个共线点得到的交叉比
都有相同的值

$$C(P_1, P_2, P_3, P_4) = \frac{(x_3 - x_1) / (x_2 - x_4)}{(x_2 - x_1) / (x_3 - x_4)} = R$$

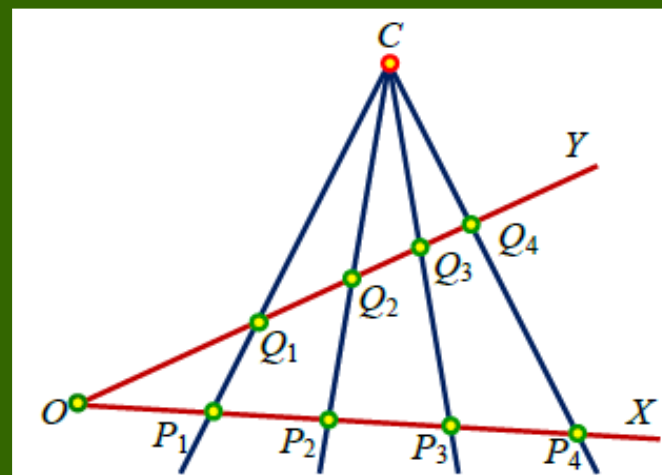
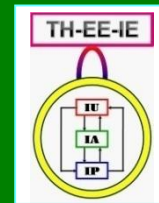


图 15.2.1 4个共线点的透视变换



15.2.1 交叉比

两种较少见的:

$$C(P_3, P_1, P_2, P_4) = 1 - C(P_1, P_3, P_2, P_4) = 1 - \frac{1}{R} = \frac{R-1}{R}$$

三种较多见的:

$$C(P_2, P_3, P_1, P_4) = \frac{1}{C(P_2, P_1, P_3, P_4)} = \frac{1}{1-R}$$

$$C(P_2, P_1, P_3, P_4) = \frac{(x_3 - x_2) / (x_1 - x_4)}{(x_1 - x_2) / (x_3 - x_4)} = C(P_1, P_2, P_4, P_3) = \frac{(x_4 - x_1) / (x_2 - x_3)}{(x_2 - x_1) / (x_4 - x_3)} = 1 - R$$

$$C(P_1, P_3, P_2, P_4) = \frac{(x_2 - x_1) / (x_3 - x_4)}{(x_3 - x_1) / (x_2 - x_4)} = C(P_4, P_2, P_3, P_1) = \frac{(x_3 - x_4) / (x_2 - x_1)}{(x_2 - x_4) / (x_3 - x_1)} = \frac{1}{R}$$

$$C(P_3, P_2, P_1, P_4) = \frac{(x_1 - x_3) / (x_2 - x_4)}{(x_2 - x_3) / (x_1 - x_4)} = C(P_1, P_4, P_3, P_2) = \frac{(x_3 - x_1) / (x_4 - x_2)}{(x_4 - x_1) / (x_3 - x_2)} = \frac{R}{R-1}$$

将点的排列顺序反过来不会改变交叉比的值

15.2.2 非共线点的不变量

非共线点的不变量

用5个点总可以获得具有不变性的交叉比

用5个共面的点可算出两个不同的交叉比，取两个参考点计算两个交叉比就可以唯一地确定所有剩下点之间的朝向关系

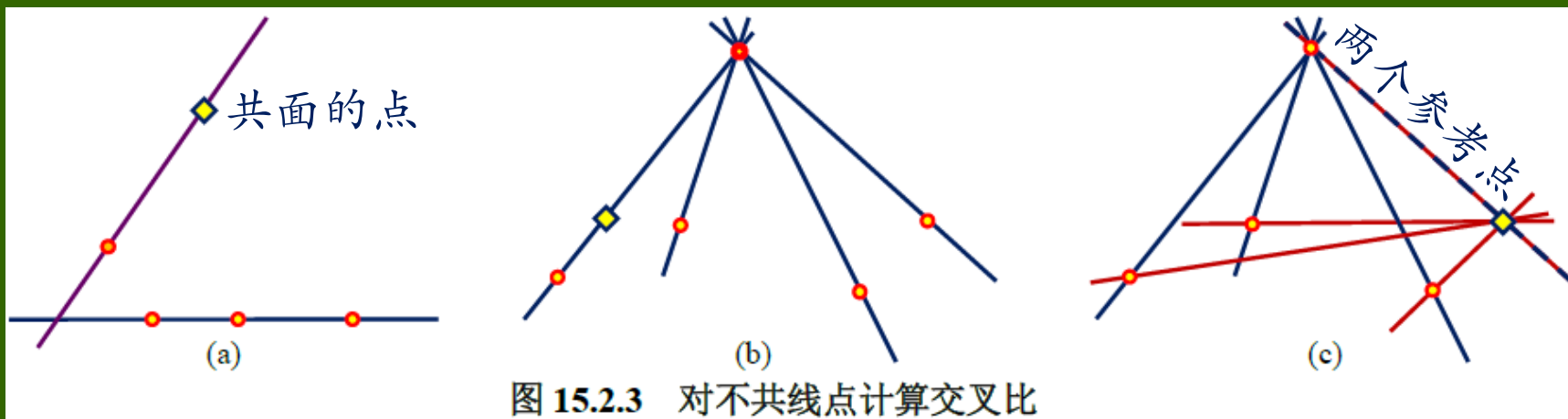
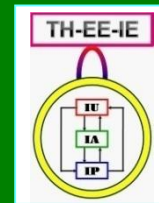


图 15.2.3 对不共线点计算交叉比



15.2.3 对称的交叉比函数

- 需要在各种可能排序情况下都有不变性
- 如果歧义源于分不清交叉比的值是 R 还是 $(1-R)$, 则可使用函数 $f(R) = R(1-R)$
- 如果歧义源于分不清交叉比的值是 R 还是 $1/R$, 则可使用函数 $g(R) = R+1/R$
- 如果源于分不清交叉比的值是 R 还是 $(1-R)$ 还是 $1/R$, 则需要能满足双重条件的函数 $h(R) = h(1-R) = h(1/R)$:

$$S(R) = \frac{(1 - R + R^2)^3}{R^2 (1 - R^2)^2}$$

15.2.4 交叉比应用示例

图(a)为4个共线点都在水平地面上的情况

图(b)为地面不平整交叉比不能保持常数的情况

图(c)为交叉比保持常数但4个共线点不都在水平地面上的情况，

图(d)为4个点不共面所以交叉比不为常数的情况

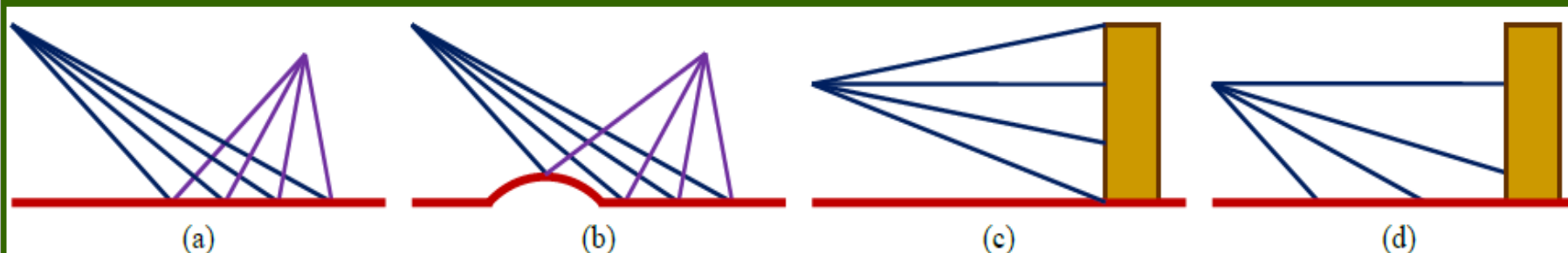
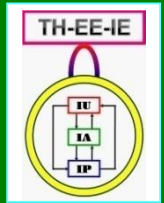


图 15.2.4 借助交叉比来确定地平面



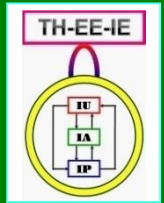
15.3 统计模式识别

- 根据模式统计特性确定决策函数
 - 选取特征表达描述模式
 - 设计分类器（根据特征）进行分类

15.3.1 最小距离分类器

15.3.2 最优统计分类器

15.3.3 自适应自举



15.3.1 最小距离分类器

- 每个模式类 s_j 用一个均值矢量表示

$$\mathbf{m}_j = \frac{1}{N_j} \sum_{\mathbf{x} \in s_j} \mathbf{x} \quad j=1, 2, \dots, M$$

- 利用欧氏距离来确定接近程度

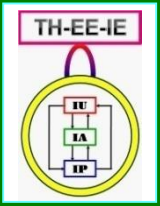
$$D_j(\mathbf{x}) = \|\mathbf{x} - \mathbf{m}_j\| \quad j=1, 2, \dots, M$$

如果 $D_j(\mathbf{x})$ 是最小的距离，则将 \mathbf{x} 赋给类 s_j

- 等价于计算决策函数

$$d_j(\mathbf{x}) = \mathbf{x}^T \mathbf{m}_j - \frac{1}{2} \mathbf{m}_j^T \mathbf{m}_j \quad j=1, 2, \dots, M$$

在 $d_j(\mathbf{x})$ 给出最大值时将 \mathbf{x} 赋给类 s_j



15.3.1 最小距离分类器

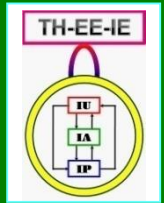
- 类 s_i 和 s_j 之间的决策边界

$$d_{ij}(\mathbf{x}) = d_i(\mathbf{x}) - d_j(\mathbf{x}) = \mathbf{x}^T (\mathbf{m}_i - \mathbf{m}_j) - \frac{1}{2} (\mathbf{m}_i - \mathbf{m}_j)^T (\mathbf{m}_i - \mathbf{m}_j) = 0$$

- 给出一个连接 \mathbf{m}_i 和 \mathbf{m}_j 线段的垂直二分界
 - $M = 2$, 垂直二分界是一条线
 - $M = 3$, 垂直二分界是一个平面
 - $M > 3$, 则垂直二分界是一个超平面
- 仅由 $d_{ij}(\mathbf{x})$ 的符号就可判断模式属于哪个类

{例15.3.1}

{例15.3.2}



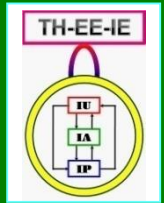
15.3.2 最优统计分类器

1. 最优统计分类原理

在平均意义上产生最小可能的分类误差

- 借助概率论：令 $p(s_i|\mathbf{x})$ 代表一个特定的模式 \mathbf{x} 属于类 s_i 的概率
- 如果模式分类器判决 \mathbf{x} 属于 s_j ，但事实上 \mathbf{x} 属于 s_i ，则认为分类器犯了一个误检错误，记为 L_{ij}
- 将 \mathbf{x} 赋给 s_j 产生的平均损失是（条件平均风险损失）

$$r_j(\mathbf{x}) = \sum_{k=1}^M L_{kj} p(s_k | \mathbf{x}) \quad \{\text{概率加权和}\}$$



15.3.2 最优统计分类器

1. 最优统计分类原理

- 平均风险的表达式（根据基本概率理论）

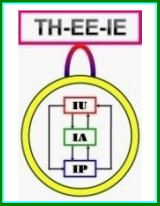
$$r_j(\mathbf{x}) = \frac{1}{p(\mathbf{x})} \sum_{k=1}^M L_{kj} p(\mathbf{x} | s_k) P(s_k)$$

- 贝叶斯分类器：

能最小化总平均损失的分类器

如果 $r_i(\mathbf{x}) < r_j(\mathbf{x})$ ，将 \mathbf{x} 赋给 s_i

$$\sum_{k=1}^M L_{ki} p(\mathbf{x} | s_k) P(s_k) < \sum_{l=1}^M L_{lj} p(\mathbf{x} | s_l) P(s_l)$$



15.3.2 最优统计分类器

1. 最优统计分类原理

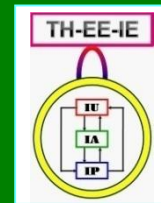
- 损失函数 $L_{ij}=1-\delta_{ij}$ {以项计}
- 贝叶斯分类器:

满足下面条件时将 \mathbf{x} 赋给类 s_i

$$p(\mathbf{x}|s_i)P(s_i) > p(\mathbf{x}|s_j)P(s_j) \quad j=1, 2, \dots, M; \quad j \neq i$$

- 贝叶斯分类器的决策函数

$$d_j(\mathbf{x}) = p(\mathbf{x}|s_j)P(s_j) \quad j=1, 2, \dots, M$$



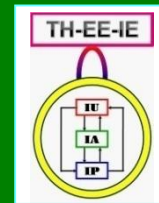
15.3.3 自适应自举

弱分类器： 分类效果在两类样本时仅略高于50%
需要将多个这样的独立分类器结合起来
以取得更好的效果

自举： 将这些分类器依次分别用于不同的训练
样本子集

自举算法将多个弱分类器结合成一个比
其中每个弱分类器都要好的新的强分类器

关键： (1) 如何选择输入弱分类器的训练样本子集
(2) 如何结合它们以构成一个强分类器

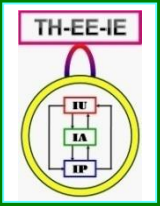


15.3.3 自适应自举

自适应自举 (Adaboost) :

- (1) 初始化 K , K 为需使用的弱分类器数量
- (2) 令 $k = 1$, 初始化权重 $W_1(i) = 1/m$
- (3) 对每个 k , 使用训练集合和一组权重 $W_k(i)$ 来训练弱分类器 C_k , 对每个模式 \mathbf{x}_i 赋一个实数, 即 $C_k: X \rightarrow \mathbf{R}$
- (4) 选择系数 $a_k > 0 \in \mathbf{R}$
- (5) 更新权重
相对增加错分
样本的权重

$$W_{k+1}(i) = \frac{W_k(i) \exp[-a_k c_i C_k(\mathbf{x}_i)]}{G_k}$$



15.3.3 自适应自举

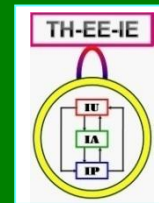
自适应自举:

- (6) 设置 $k = k + 1$
- (7) 如果 $k \leq K$, 回到步骤(3)
- (8) 最后的强分类器是

$$S(\mathbf{x}_i) = \text{sign} \left[\sum_{k=1}^K a_k C_k(\mathbf{x}_i) \right]$$

在每个步骤, 要确定弱分类器 C_k 以使其性能与权分布 $W_k(i)$ 相适应。在二分类情况下, 弱分类器训练要最小化的目标函数为

$$e_k = \sum_{i=1}^m P_{i \sim W_k(i)} [C_k(\mathbf{x}_i) \neq c_i]$$



15.4 感知机和支持向量机

- 直接通过训练或学习得到所需的决策函数
- 神经网络方法
- 最优设计方法

15.4.1 感知机

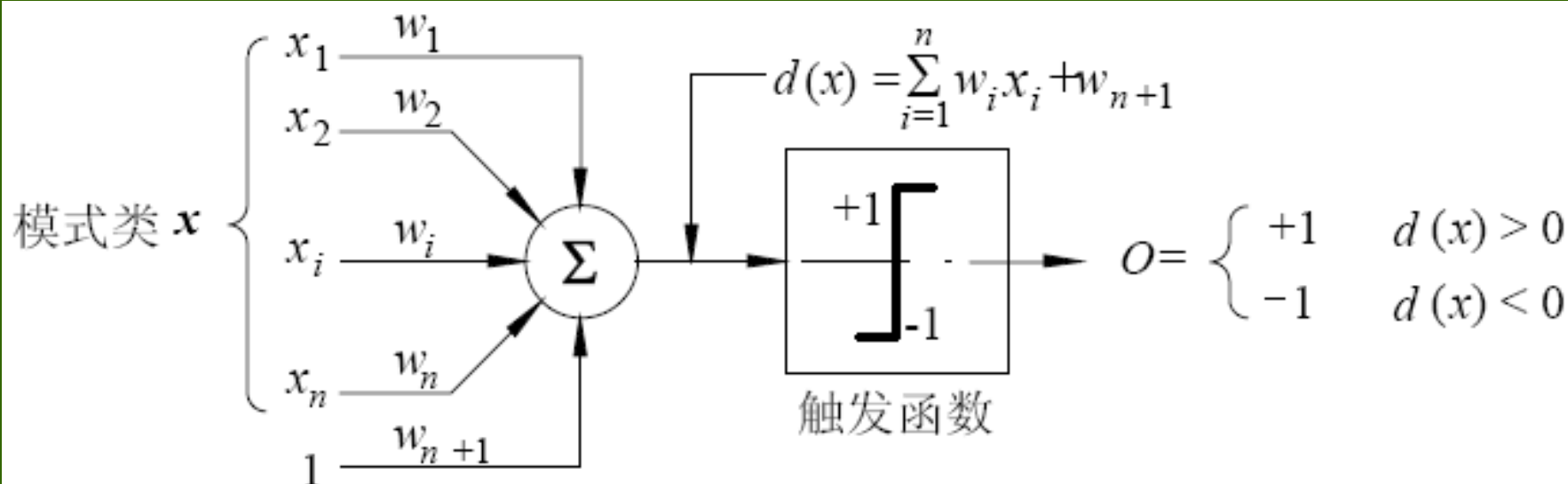
15.4.2 支持向量机

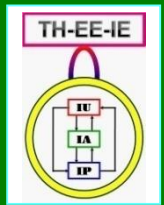
15.4.1 感知机

➤ 最基本的感知机建立能将两个线性可分训练集分开的线性决策函数

➤ 决策边界

$$d(\mathbf{x}) = \sum_{i=1}^n w_i x_i + w_{n+1} = 0$$





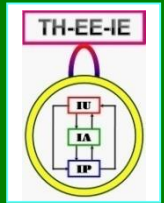
15.4.1 感知机

- **变型：**对模式矢量增加第 $n + 1$ 个元素
- 构建一个扩充模式矢量 \mathbf{y} ，让 $y_i = x_i$ ， $i = 1, 2, \dots, n$ ，且后面加一个元素 $y_{n+1} = 1$

$$d(\mathbf{x}) = \sum_{i=1}^{n+1} w_i y_i = \mathbf{w}^T \mathbf{y}$$

其中 $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_n \ 1]^T$ 是一个扩充模式矢量， $\mathbf{w} = [w_1 \ w_2 \ \dots \ w_n \ w_{n+1}]^T$ 是一个权矢量

- 关键问题：用模式矢量的给定训练集确定 \mathbf{w}



15.4.1 感知机

1. 线性可分类

由两个线性可分训练集获取权矢量

令 $\mathbf{w}(1)$ 代表一个任意选定的初始权矢量

如果 $\mathbf{y}(k) \in s_1$, $\mathbf{w}^T(k)\mathbf{y}(k) \leq 0$

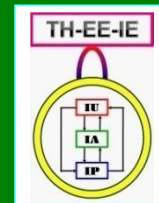
$$\mathbf{w}(k+1) = \mathbf{w}(k) + c\mathbf{y}(k)$$

固定增量

如果 $\mathbf{y}(k) \in s_2$, $\mathbf{w}^T(k)\mathbf{y}(k) \geq 0$

$$\mathbf{w}(k+1) = \mathbf{w}(k) - c\mathbf{y}(k)$$

否则: $\mathbf{w}(k+1) = \mathbf{w}(k)$ {正确分类的奖励就是不给惩罚}



15.4.1 感知机

2. 线性不可分类

最小化实际响应和希望响应间的误差

$$J(\mathbf{w}) = \frac{1}{2}(\mathbf{r} - \mathbf{w}^T \mathbf{y})^2$$

沿 $J(\mathbf{w})$ 负梯度的方向逐步增加 \mathbf{w} 以寻找上述函数的最小值。最小值应在 $\mathbf{r} = \mathbf{w}^T \mathbf{y}$ 时出现

通用的梯度下降算法可写成：

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \alpha \left[\frac{\partial J(\mathbf{w})}{\partial \mathbf{w}} \right]_{\mathbf{w}=\mathbf{w}(k)}$$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \alpha \left[\mathbf{r}(k) - \mathbf{w}^T(k) \mathbf{y}(k) \right] \mathbf{y}(k)$$

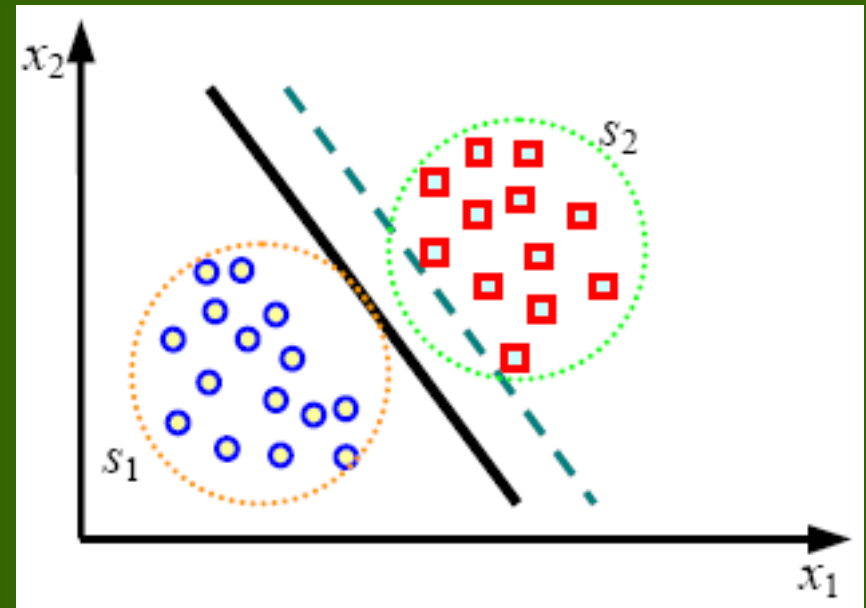
15.4.2 支持向量机

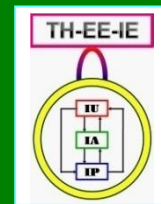
1. 线性可分类

线性分类器的设计目的：要设计一个超平面

$$g(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + w_0 = 0$$

- 满足条件的超平面一般不惟一
- 用离开两个类都比较远的超平面来分类的结果会更好些，可能的错误率也会更小一些





15.4.2 支持向量机

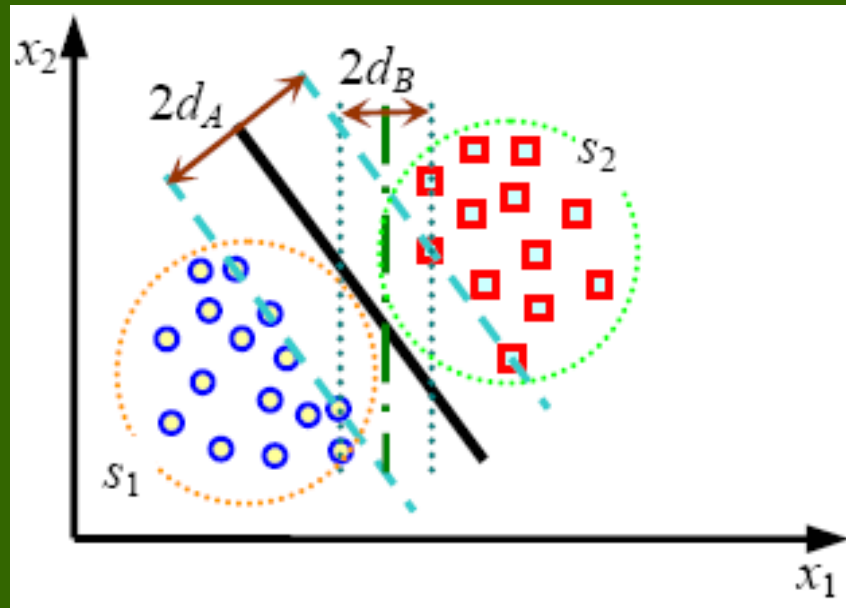
1. 线性可分类

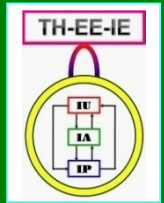
对每个朝向，与两个类距离相等的超平面应该是与两个类都有最大距离的超平面

确定能给出类距离最大之朝向的超平面

从一个点到超平面的距离：

$$d = |g(x)| / \|w\|$$





15.4.2 支持向量机

1. 线性可分类

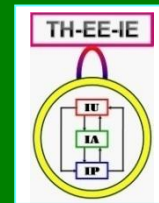
对每个类 s_i ，记其标号为 t_i ，其中 $t_1 = 1$ ， $t_2 = -1$

- 问题变为：计算超平面的 \mathbf{w} 和 w_0 ，在满足约束条件（线性不等式）
$$t_i(\mathbf{w}^T \mathbf{x}_i + w_0) \geq 1 \quad i=1, 2, \dots, N$$

的情况下最小化（非线性代价函数）
$$C(\mathbf{w}) \equiv \frac{1}{2} \|\mathbf{w}\|^2$$

- 用拉格朗日乘数法来解：

$$L(\mathbf{w}, w_0, \boldsymbol{\lambda}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} - \sum_{i=1}^N \lambda_i [t_i(\mathbf{w}^T \mathbf{x}_i + w_0) - 1]$$



15.4.2 支持向量机

1. 线性可分类

结果为

$$\mathbf{w} = \sum_{i=1}^N \lambda_i t_i \mathbf{x}_i$$

$$\sum_{i=1}^N \lambda_i t_i = 0$$

拉格朗日乘数
可取正值或零

最优解的向量参数 \mathbf{w} 是 N_s 个 ($N_s \leq N$) 与 $\lambda_i \neq 0$ 相关的特征向量的线性组合 (支持向量)

$$\mathbf{w} = \sum_{i=1}^{N_s} \lambda_i t_i \mathbf{x}_i$$

支持向量机：最优的超平面分类器

15.4.2 支持向量机

2. 线性不可分类

训练特征向量可以分成以下三类：

- (1) 向量落在分类带之外且被正确地分了类
 (2) 向量落在分类带之内且被正确地分了类

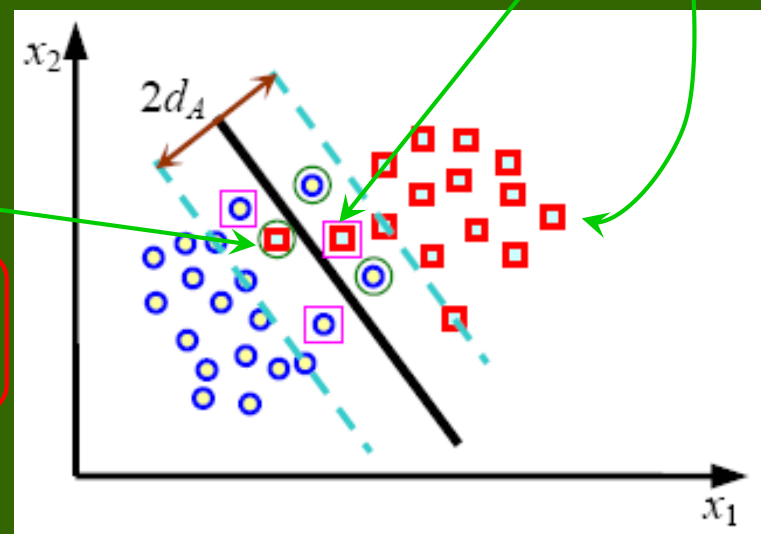
$$0 \leq t_i(\mathbf{w}^T \mathbf{x} + w_0) < 1$$

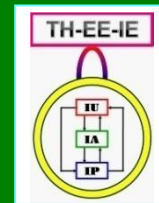
- (3) 向量被错误地分了类

$$t_i(\mathbf{w}^T \mathbf{x} + w_0) < 0$$

松弛
变量

统一： $t_i(\mathbf{w}^T \mathbf{x} + w_0) \geq 1 - r_i$





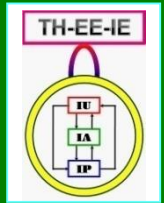
15.5 结构模式识别

句法模式识别

一组模式基元，一组确定这些基元相互作用的规则（句法）和一个识别器（称为自动机）

15.5.1 字符串结构识别

15.5.2 树结构识别



15.5.1 字符串结构识别

1. 字符串文法

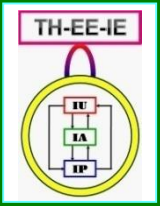
文法（语法）：一组句法规则，控制字符集中符号产生句子的过程

语言：由一个文法 G 所产生的一组句子，记为 $L(G)$

句子：符号的串，代表模式，语言对应模式类

文法 G_1 中的句法规则只允许产生对应类 s_1 中模式的句子，文法 G_2 中的句法规则只允许产生对应类 s_2 中模式的句子，.....

识别：对一个表示未知模式的句子，决定在哪种语言中该模式表示了一个可成立的句子



15.5.1 字符串结构识别

1. 字符串文法

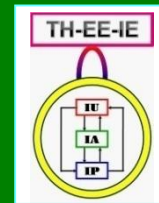
定义一个四元组: $G = (N, T, P, S)$

N 为一个有限的变量集, 称为非终结符号集
用大写字母 A, B, \dots, S, \dots 代表非终结符号

T 为一个有限的常量集, 称为终结符号集
用小写字母 a, b, c, \dots 表示终结符号

P 是一组称为产生式的重写规则集

S 在 N 中, 称为起始符号



15.5.1 字符串结构识别

1. 字符串文法

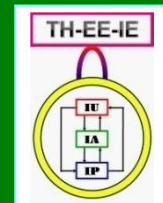
字符串文法的特点由产生式规则的形式决定：

- 规则文法

产生式规则： $A \rightarrow aB$ 或 $A \rightarrow a$

- 前后文无关文法

产生式规则：它只包含 $A \rightarrow \alpha$ 的形式， α 在集合 $(N \cup T) - \lambda$ 中，即 α 可以是除了空集 $\{\lambda\}$ 以外的、由终结符号和非终结符号组成的任何字符串



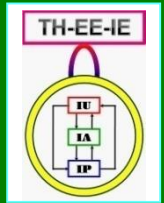
15.5.1 字符串结构识别

2. 语义应用

句法确定了目标的结构，而语义（规则）主要与其（内容/含义）正确性有关

产生式	语义信息
$S \rightarrow aA$	与 a 的连接只在圆点处， a 的方向（用 θ 表示）与两个线段端点间连线正交。每个线段长度为 2。
$A \rightarrow bA$	与 b 的连接只在圆点处，不允许多重连接。 a 和 b 的方向必须相同， b 的长度是 1。这个规则不能用 5 次以上。
$A \rightarrow bB$	a 和 b 的方向必须相同。连接为简单连接且只在圆点处发生。
$B \rightarrow c$	a 和 c 的方向必须相同。连接为简单连接且只在圆点处发生。

通过使用语义信息，可用较少的句法规则以描述较广泛（但是有限）的模式类



15.5.1 字符串结构识别

3. 用自动机作为字符串识别器

自动机能识别一个输入模式字符串是否属于与自动机关联的语言

有限自动机：由规则文法产生的语言识别器

$$A_f = (Q, T, \delta, q_0, F)$$

Q 是一个有限非空状态集； T 是一个有限输入字符集； δ 是一个从 $Q \times T$ （即由 Q 和 T 的元素组成的排序对集合）到所有 Q 子集的映射； q_0 是初始状态； F （ Q 的一个子集）是一个最终或可接受状态集合

15.5.1 字符串结构识别

3. 用自动机作为字符串识别器

有限自动机状态图

$$Q = \{q_0, q_1, q_2\}$$

$$T = \{a, b\}$$

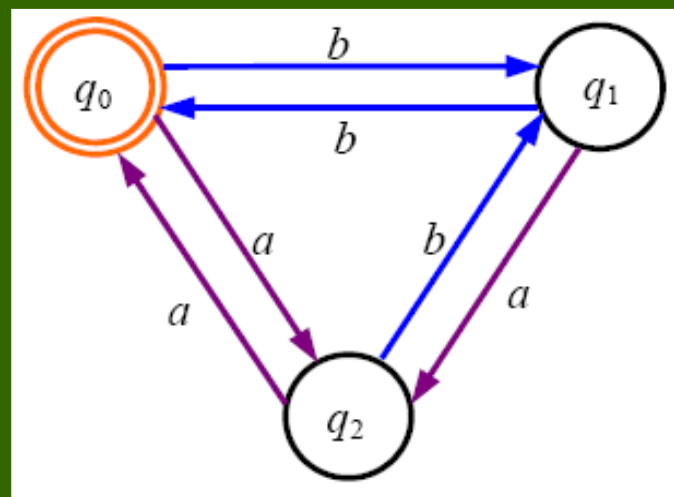
$$F = \{q_0\}$$

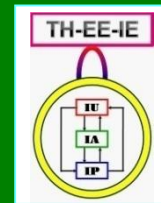
映射规则:

$$\delta(q_0, a) = \{q_2\}, \quad \delta(q_0, b) = \{q_1\}, \quad \delta(q_1, a) = \{q_2\}$$

$$\delta(q_1, b) = \{q_0\}, \quad \delta(q_2, a) = \{q_0\}, \quad \delta(q_2, b) = \{q_1\}$$

能识别字符串 *abbabb*, 但不能识别字符串 *aabab*





15.5.2 树结构识别

1. 树文法

$$G = (N, T, P, r, S)$$

N : 非终结符号集

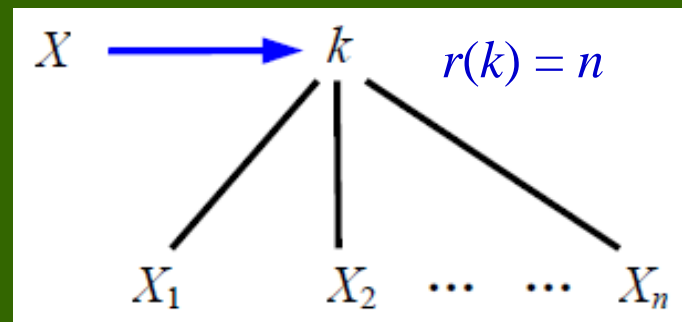
T : 终结符号集

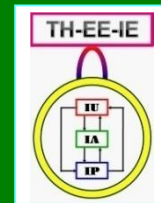
S : 一个包含在 N 中的起始符号，它一般是一棵树

P : 一组产生式规则，其一般形式为 $T_i \rightarrow T_j$ ，其中 T_i 和 T_j 都是树

r : 排序函数，它记录了一个其标号是文法中终结符号结点的直接后裔数目

扩展树文法的产生式规则





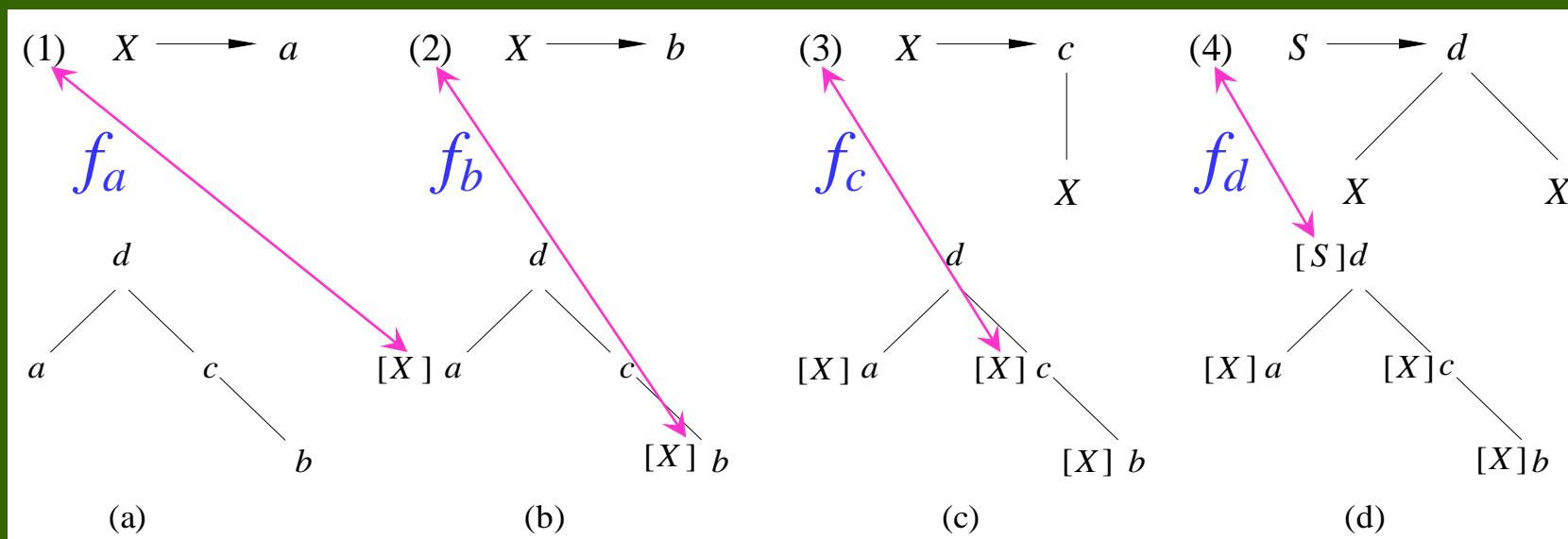
15.5.2 树结构识别

2. 树自动机

一个从树叶向树根扫描的自动机

$$A_t = (Q, F, \{f_k \mid k \in T\})$$

示例: $Q = \{S, X\}$, $F = \{S\}$, $\{f_k \mid k \in S\} = \{f_a, f_b, f_c, f_d\}$





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
- ☞ 个人主页：oa.ee.tsinghua.edu.cn/~zhangyujin/