

文章编号: 1001-0920(2013)04-0569-05

## 基于模糊度的计算几何分类器权重分配

张 涛<sup>a,b</sup>, 洪文学<sup>b</sup>

(燕山大学 a. 信息科学与工程学院, b. 电气工程学院, 河北 秦皇岛 066004)

**摘 要:** 为了充分发挥计算几何组合分类器的可视化特性, 从类空间类别分布模糊特性出发, 提出一种基于类空间模糊度的权重分配方法. 该方法首先将子分类器由空间的类别表示转变为类别的空间表示, 进而利用粗糙集理论分析各空间的模糊度, 并将其作为该子分类器的权重. 实验表明, 利用模糊度信息进行加权后的分类器不但能与可视化特性更好地吻合, 增强分类过程的可理解性, 而且在分类精度上得到了进一步的提升, 扩展了应用领域.

**关键词:** 模糊度; 粗糙集; 计算几何; 组合分类器; 可视化

**中图分类号:** TP391

**文献标志码:** A

## Weight calculation for computational geometry combining classifier using fuzzy of class space

ZHANG Tao<sup>a,b</sup>, HONG Wen-xue<sup>b</sup>

(a. College of Information Science and Engineering, b. College of Electrical Engineering, Yanshan University, Qinhuangdao 066004, China. Correspondent: ZHANG Tao, E-mail: zhtao@ysu.edu.cn)

**Abstract:** In order to take advantage of visualization characteristics of computational geometry combined classifiers, a weight calculation method based on space fuzzy is proposed according to the category distribution in class space. In this method, the space is turned from the category information in space to the space information in category. The space fuzzy is obtained from the later based rough set theory and taken as the weight for the sub classifier. The experiments show that the classifier weighted by using the fuzzy degree can enhance the visual performance and the classify performance of the classifier. It means that the comprehensibility of the classifier is enhanced and the application of the classifier is extended.

**Key words:** fuzzy degree; rough set theory; computational geometry; combining classifier; visualization

## 0 引 言

在模式识别领域中, 分类器的设计一直是核心研究问题之一<sup>[1-2]</sup>, 其最终目的是将不同类别数据在相应空间上找到可以将这些数据加以区分的分类界面, 从而完成数据的分类过程. 根据分类器方法的不同, 经典的分类器<sup>[3-5]</sup>可分为线性分类器、非线性分类器、非参数估计分类器以及源于 Cover 定理的支持向量机<sup>[6]</sup>等.

针对复杂分布数据, 文献[7]提出了采用可视化模式识别的方法进行分类器设计. 可视化模式识别将抽象数据变为视觉范围内的数据, 因此具有天然的物理意义简单、便于理解的特性. Zhang 等<sup>[8]</sup>利用可视化理论分别设计完成了数据表示、特征选择与分类器设计. 其中, 散点图分类器架构<sup>[9]</sup>作为典型的可视化

组合分类器, 在保留可视化分类良好的解释性的同时, 获得了较好的分类效果.

目前, 对于计算几何分类器的研究主要集中于其分类界面的形成过程. 作为典型的组合分类器, 其子分类器间的权重分配同样是该分类器的重要组成部分, 其分配效果将直接影响分类精度. 文献[9]提到了权重分配对分类结果的影响, 采用了经典的组合分类器权重分配方法. 这样的权重计算虽然与经典模式识别领域中的组合分类器理论相符, 但由于其未考虑可视化特色, 无法最大限度发挥可视化模式识别的优势.

鉴于此, 本文提出基于类空间模糊度的计算几何分类器权重分配方法. 该方法首先将子分类器由空间的类别表示转变为类别的空间表示, 进而利用空间中类别分布的概率特性计算模糊度, 从而获得当前类空

收稿日期: 2011-12-25; 修回日期: 2012-05-07.

基金项目: 国家自然科学基金项目(60904100, 61074195); 河北省自然科学基金项目(F2011203073).

作者简介: 张涛(1979—), 男, 讲师, 博士生, 从事模式识别、图像处理的研究; 洪文学(1953—), 男, 教授, 博士生导师, 从事可视化模式识别、信息融合等研究.

间的权重. 该方法以视觉信息作为基础, 充分利用了计算几何分类器的可视化特色, 进一步完善了可视化模式识别方法.

## 1 计算几何组合分类器空间计算

基于计算几何原理的组合分类器原理如图1所示<sup>[8]</sup>.

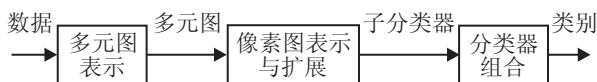


图1 计算几何分类器原理图

分类器的输入数据为传统模式识别的特征矢量, 多元图表示将抽象数据映射为直观的多元图表示. 在像素图表示与扩展模块中, 将传统多元图转换为可进行数据与类别概率表示的像素图, 从而达到可计算的要求, 并对像素图中的点区域进行扩展, 生成子分类器的分类界面. 分类器组合模块则是将像素图形成的子分类器集成, 形成一个组合分类器, 从而完成最终的分类. 从整个分类过程看, 数据在分类器中一直以图形或图像的形式表示, 因此称为可视化分类器; 而多维数据生成多个多元图, 又为组合分类器形式. 因此, 计算几何分类器是典型的可视化组合分类器.

设训练样本集合  $X = \{x_1, x_2, \dots, x_m\}$ , 共有  $m$  个训练样本, 每个样本可表示为  $x_i = \{x_{i1}, x_{i2}, \dots, x_{id}\} \in R^d$ . 训练样本类别标签集合为  $C = \{c_1, c_2, \dots, c_k\}$ , 显然  $k \leq m$ . 设表示训练样本  $x_i$  的类别为  $L(x_i)$ , 则必有  $L(x_i) \in C$ .

为了降低离群点对分类结果的影响, 可以进行量化. 设训练样本  $X$  中属性集合为  $X_j = \{x_{ij}|1, 2, \dots, m\}$ , 量化后的数据空间集合为

$$V = \{v_n|n = 1, 2, \dots, N\}. \quad (1)$$

区域  $v_n$  内样本为  $c_j$  类的概率可表示为

$$p(c_j|v_n) = \frac{\#\{i|x_i \in v_n, L(x_i) = c_j\}}{\sum_{p=1}^k \#\{i|x_i \in v_n, L(x_i) = c_p\}}. \quad (2)$$

为便于书写, 对符号简化, 令  $p_n(c_j) = p(c_j|v_n)$ ,  $p_n = \{p_n(c_j)|j = 1, 2, \dots, k\}$ .

利用计算几何原理, 可以通过对基点数据的分析获取非基点下可能的类别分布, 从而完成整个类空间的全息信息获取. 文献[10]给出其计算结果表达式

$$p_n = \frac{1}{2} |\text{sign}[p_{n-1}p_{n+1}] - 2|(p_{n-1} + p_{n+1})|. \quad (3)$$

其中:  $\text{sign}(\cdot)$  为符号函数,  $|\cdot|$  表示取绝对值运算.

由于主动生长获得的分类界面与传统分类界面不同, 其不受函数形式约束且允许多类别重叠区域存在. 该方法得到的分类界面不但符合数据本身分布特点, 而且不要求严格的样本完备性, 同时具有良好的

可视化特性和可解释特性.

## 2 子分类器权重计算

### 2.1 分类界面转换

主动生长过程为在高维空间的约束性生长, 在该过程中, 不同类别在相同空间下的扩展过程相互制约. 在每个子分类器的类空间中包含  $k$  个类的全部信息. 由主动生长算法可知, 生成后的类空间中,  $p(c_j|v_n)$  表示在特定空间上的类别分布. 但从权重计算的角度看, 应利用特定类别上的空间表示, 因此需要对该数据进行转化. 依全概率公式可知

$$p(v_n|c_j) = \frac{p(v_n)p(c_j|v_n)}{p(c_j)}, \quad (4)$$

其中  $p(v_n)$  表示当前位置点上出现的概率. 结合计算几何的计算特点可知

$$p(v_n) = 1/N, \quad (5)$$

$$p(c_j) = \frac{1}{N} \sum_{n=1}^N p(c_j|v_n). \quad (6)$$

将式(5)和(6)代入(4), 整理得

$$p(v_n|c_j) = p(c_j|v_n) / \sum_{n=1}^N p(c_j|v_n). \quad (7)$$

式(7)将分类界面由特定点类别信息转化为特定类别的分布信息. 从界面分解的角度看, 是将一个表示空间类别分布的分类界面分解为  $k$  个表示单个类别空间分布的界面, 这样便于后期的权重计算.

### 2.2 分类界面与粗糙集关系

空间转换将表示特定空间位置  $v_n$  上类别  $c_j$  分布的  $p(c_j|v_n)$  转换为特定类别  $c_j$  在位置  $v_n$  上出现的概率  $p(v_n|c_j)$  表示. 依据  $p(v_n|c_j)$  的分布可以对集合  $P_j = \{v_n|p(v_n|c_j), n \in [1, N]\}$  进行划分.

显然, 当  $p(v_n|c_j) = 0$  时, 表示当前位置对应的数据一定不属于  $c_j$  类; 当  $p(v_n|c_j) = 1$  时, 表示当前位置对应的数据一定属于  $c_j$  类; 而当  $p(v_n|c_j) \in (0, 1)$  时, 表示当前位置对应的数据可能属于  $c_j$  类. 因此, 该表示方法是对当前位置类别标签的模糊表示. 该过程与粗糙集中特定空间等价关系理论<sup>[11-12]</sup>相符, 因此可以采用粗糙集进行分析.

另外, 粗糙集无需提供问题所需处理数据集之外的任何先验信息, 所以对问题的不确定性的描述或处理更为客观, 这也符合多维筛分类器中权重计算的客观性要求. 因此, 粗糙集方法用于权重计算符合数据分布本身的要求.

### 2.3 特定类粗糙度计算

在特定空间中, 对于特定类别  $c_j$  而言, 其类别的位置分布可以表示为信息系统

$$S_{c_j} = \langle U_{c_j}, R_{c_j}, V_{c_j}, f_{c_j} \rangle. \quad (8)$$

其中:  $U_{c_j}$  为论域, 且  $U_{c_j} = \{v_n | p(v_n | c_j), n \in [1, N]\}$ ;  $R_{c_j}$  为属性集合, 此处属性仅有一个, 即概率;  $V_{c_j}$  为属性的值域, 有  $V_{c_j} = \bigcup_{n \in [1, N]} p(v_n | c_j)$ ;  $f_{c_j}$  为信息函数, 有  $f_{c_j} : U_{c_j} \times R_{c_j} \rightarrow V_{c_j}$ , 即  $f(v_n) = p(v_n | c_j)$ .

由粗糙集理论可知, 对于信息系统  $S_{c_j}$ , 其下近似集与上近似集分别为

$$R_{c_j}X = \{v_n | p(v_n | c_j) = 1, n \in [1, N]\}, \quad (9)$$

$$\bar{R}_{c_j}X = \{v_n | p(v_n | c_j) \in (0, 1], n \in [1, N]\}. \quad (10)$$

即: 信息系统  $S_{c_j}$  的下近似集为肯定属于该类的位置所构成的集合, 当待分类样本落入该集合中时, 表示该样本的类别为确定的  $c_j$ ; 而信息系统  $S_{c_j}$  的上近似集表示可能属于类别  $c_j$  的位置所对应的集合, 当待分类样本落入该集合中时, 表示该样本的类别可能为  $c_j$  且为  $c_j$  的概率是  $f_{c_j}(v_n) = p(v_n | c_j)$ .

对于信息系统  $S_{c_j}$ , 其精度为

$$d_{R_{c_j}}(X) = \frac{\#\{v_n | p(v_n | c_j) = 1, n \in [1, N]\}}{\#\{v_n | p(v_n | c_j) \in (0, 1], n \in [1, N]\}}, \quad (11)$$

其粗糙度为

$$r_{R_{c_j}}(X) = 1 - d_{R_{c_j}}(X) = \frac{\#\{v_n | p(v_n | c_j) \in (0, 1], n \in [1, N]\}}{\#\{v_n | p(v_n | c_j) \in (0, 1], n \in [1, N]\}}. \quad (12)$$

式(12)表明, 在分类空间中, 粗糙度  $r_{R_{c_j}}(X)$  表示  $c_j$  类在该分类空间中不确定分类位置与可能分类位置之比, 且有  $r_{R_{c_j}}(X) \in [0, 1]$ .

综合以上分析, 对于集合  $P_j = \{v_n | p(v_n | c_j), n \in [1, N]\}$  所形成的信息系统  $S_{c_j} = \langle U_{c_j}, R_{c_j}, V_{c_j}, f_{c_j} \rangle$  而言, 其精度  $d_{R_{c_j}}(X)$  越高, 说明其对类别  $c_j$  进行分类获得精确结果的概率越高; 而粗糙度  $r_{R_{c_j}}(X)$  越高, 说明其对类别  $c_j$  进行分类获得模糊结果的概率越高. 将二者进行组合, 可定义当前信息系统的模糊度

$$\text{fuzzy}_{R_{c_j}}(X) = \frac{d_{R_{c_j}}(X)}{r_{R_{c_j}}(X)} = \frac{\#\{v_n | p(v_n | c_j) = 1, n \in [1, N]\}}{\#\{v_n | p(v_n | c_j) \in (0, 1], n \in [1, N]\}}. \quad (13)$$

模糊度  $\text{fuzzy}_{R_{c_j}}(X)$  表示当前空间下类别  $c_j$  的模糊程度, 同时可以表示对类别  $c_j$  进行分类的结果可信度. 该值越高, 说明分类结果可信度越高, 确定性越强; 反之,  $\text{fuzzy}_{R_{c_j}}(X)$  值越小, 说明该分类结果可信度越低, 模糊性越强.

## 2.4 子分类器空间权重计算

式(13)表示了特定分类空间中特定类别  $c_j$  的模糊度. 该特征表征了不确定空间点与可能空间点之比, 是对该类别不确定度的描述. 该值越高, 说明该类别的不确定性越高. 该精度表征了确定空间点与可能空间点之比, 是对该类别精确程度的描述. 该值

越高, 说明该类别的确定性越高. 对于组合分类器而言, 不但要关注特定空间下类别的分布, 更应关注各子分类器分类空间之间的权重关系. 例如, 图2表示了Iris数据形成的6个子分类器分类界面, 显然图2(b)、2(c)、2(d)、2(f)各类之间可分性更强, 应该获得较高的权重. 因此, 本节利用模糊度概念, 对子分类空间完成权重计算.

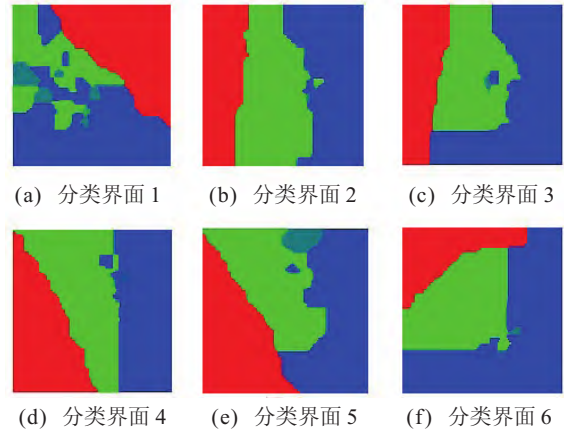


图2 Iris数据形成的子分类器分类界面

对于原始特征  $(u_{i_1}, u_{i_2})$  形成的特定分类空间而言, 其形成的知识系统为

$$S_{(i_1, i_2)} = \langle U_{(i_1, i_2)}, R_{(i_1, i_2)}, V_{(i_1, i_2)}, f_{(i_1, i_2)} \rangle. \quad (14)$$

其中:  $U_{(i_1, i_2)}$  为论域, 且  $U_{(i_1, i_2)} = \{v_n, n \in [1, N]\}$ ;  $R_{(i_1, i_2)}$  为属性集合, 即概率;  $V_{(i_1, i_2)}$  为属性的值域, 有  $V_{(i_1, i_2)} = \bigcup_{n \in [1, N]} \bigcup_{j \in [1, k]} p(v_n | c_j)$ ;  $f_{(i_1, i_2)}$  为信息函数, 有  $f_{(i_1, i_2)} : U_{(i_1, i_2)} \times R_{(i_1, i_2)} \rightarrow V_{(i_1, i_2)}$ , 即  $f_{(i_1, i_2)}(v_n) = p(v_n | c_j)$ . 根据2.3节的推导过程可知, 其模糊度为

$$\text{fuzzy}_{(i_1, i_2)}(X) = \frac{d_{R_{(i_1, i_2)}}(X)}{r_{R_{(i_1, i_2)}}(X)} = \frac{\#\{v_n | p(v_n | c_j) = 1, n \in [1, N], j \in [1, k]\}}{N^2 - \#\{v_n | p(v_n | c_j) = 1, n \in [1, N], j \in [1, k]\}}. \quad (15)$$

空间的模糊度与类别模糊度类似, 都是分类结果可信度的表示. 与类别模糊度不同的是, 空间模糊度表示空间中多个类别分类结果的总可信度. 因此, 空间模糊度可理解为在特定空间中, 所有类别模糊度的融合表示.

最后, 对模糊度做归一化, 完成模糊度到权重的映射, 即

$$\alpha_{(i_1, i_2)} = \frac{\text{fuzzy}_{(i_1, i_2)}(X)}{\sum_{j_1 \in [1, d], j_2 \in [1, d]} \text{fuzzy}_{(j_1, j_2)}(X)}. \quad (16)$$

图2对应的各子分类器分类界面模糊度与权重如表1所示. 可以看到, 权重的分布与视觉上空间的类别模糊程度正相关, 因此该权重计算方法符合可视化特性.

表 1 图 2 中各子分类器分类界面模糊度与权重

子分类器	模糊度	权重
(a)	0.448 2	0.158 089 661 7
(b)	0.490 7	0.173 080 314 6
(c)	0.482 2	0.170 082 184 0
(d)	0.486 9	0.171 739 973 9
(e)	0.452 8	0.159 712 179 5
(f)	0.474 3	0.167 295 686 2

表 3 不同加权方法下的分类精度对比 %

数据集	平均加权	模糊度加权
diabetes	71.88	75.13
liver	59.42	63.41
Iris	89.67	96.67
wine	76.97	98.02
ionosphere	83.71	90.32
breast cancer	90.09	95.43

### 3 实验与分析

#### 3.1 实验数据集与测试方法

为了全面测试多维筛分类器的性能, 本文利用 UCI 机器学习数据库中的多个数据集对分类器进行测试. 其中: diabetes 和 liver 是简单的二分类问题; Iris 和 wine 是对分类器进行性能测试中最常用的数据集, 可用于测试针对中等分类复杂度下的多类别数据的分类性能; 而 ionosphere 和 breast cancer 数据维数均超过 30 维, 数据维数相对较高, 用于评价高维情况下的分类情况.

表 2 实验用到的各数据集属性

数据集	特征数	类别数	样本数
diabetes	8	2	768
liver	6	2	345
Iris	4	3	150
wine	13	3	178
ionosphere	34	2	351
breast cancer	30	2	569

实验中用到的数据集来自物理科学与生命科学领域, 具体的应用包括了疾病诊断、产品分类、物种识别、气象分析等, 且均为实际测量的实验数据, 含有一定的测量误差, 因此在一定程度上代表了分类器在实际应用中的分类性能<sup>[13-14]</sup>.

为了获得分类性能的直观认识, 本文在基本参数相同的情况下, 分别利用等加权方式和模糊度加权方式对表 2 中各数据集进行分类性能测试. 该对比测试的目的是为了在更大的数据范围内获取数据精度信息, 进而分析本文提出的模糊度加权在分类精度上的改进.

在实验过程中, 为了确保分类性能更加客观, 并避免训练集和测试集的依赖, 分类器精度的估计采用留一法交叉验证 (LOOCV). 留一法是指设数据集共有  $N$  个样本, 使用  $(N-1)$  个样本设计分类器, 并估计剩余的一个样本; 对于训练集重复  $N$  次. 这种估计虽然计算量大, 但却是无偏的.

#### 3.2 实验结果与分析

针对表 2 所给出的 6 个测试数据集, 在多维筛框架下分别采用平均加权和模糊度加权两种方法进行加权, 其分类精度如表 3 所示.

实验结果表明, 由于模糊度加权考虑了视觉特性, 其分类性能高于未考虑分布特性的平均加权算法. 该结论与经典模式识别理论中关于组合分类器的加权计算理论相符, 说明不同的子分类器对于分类性能的贡献度不同.

对于低维度二分类问题, 模糊度加权分类性能提升有限. 其主要原因在于: 低维度数据形成的子分类器个数较少, 且每个类空间中数据分布相对简单, 对于组合分类贡献相当. 因此, 在简单数据分布下, 空间的模糊度加权对分类精度的提高不大.

对于中等复杂度的分类问题, 模糊度加权的分类精度远远超过了平均加权算法. 主要原因在于: 随着数据类别数的增加, 各子分类器的分布特性变得更加复杂, 各类别均需多个域进行数据分布表示, 此时不同的子分类器在分类精度上出现了较大差异, 模糊度加权考虑了各子分类空间的模糊特性, 表现出了对界面特征的描述能力, 进而获得了更好的分类精度.

对于高维分类问题, 模糊度加权同样具有远高于平均加权方法的分类精度. 其主要原因是: 随着数据维数的增加, 带来了子分类器个数的增加, 从而使各子分类器间的关系变得复杂, 模糊度加权从各子分类空间的模糊特性出发对各子分类器进行不确定性的分析, 从而使分类精度获得了较大提升.

### 4 结 论

本文从分类界面模糊性的角度出发, 利用粗糙集方法对各子分类界面进行分析, 提取模糊度作为衡量各子分类器权重的标准. 由于模糊度与分类界面具有良好的视觉对应性, 符合可视化模式识别中对权重计算的要求; 同时, 由于考虑了不同分类空间下模糊特性, 分类精度得到了有效提高.

从数据分布角度看, 目前的模糊度方法是从分类界面的统计角度进行权重分配. 如何综合利用分类界面的统计与结构信息, 更为全面地提取子分类器的权重信息将是下一步研究的方向.

#### 参考文献(References)

- [1] Duda R O, Hart P E, Stork D G. Pattern classification[M]. New York: Wiley, 2000: 8-20.
- [2] McLachlan G J. Discriminant analysis and statistical

- pattern recognition[M]. New York: Wiley Interscience, 2004: 12-15.
- [3] Vapnik V. Statistical learning theory[M]. New York: Wiley Interscience, 1998: 34-36.
- [4] Anil K Jain, Robert P W Duin, Mao Jianchang. Statistical pattern recognition: A review[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22(1): 4-37.
- [5] Tominaga Y. Comparative study of class data analysis with PCA-LDA, SIMCA, PLS, ANNs, and  $k$ -NN[J]. Chemometrics and Intelligent Laboratory Systems, 1999, 49(1): 105-115.
- [6] Maurizio Filippone, Francesco Camastra, Francesco Masulli, et al. A survey of kernel and spectral methods for clustering[J]. Pattern Recognition, 2008, 41(1): 176-190.
- [7] Xu Yonghong, Hong Wenxue, Chen Na, et al. Parallel filter: A visual classifier based on parallel coordinates and multivariate data analysis[C]. The 3rd Int Conf on Intelligent Computing. Qingdao, 2007, 4682: 1172-1183.
- [8] Zhang Tao, Hong Wenxue. A novel visual combining classifier based on a two-dimensional graphical representation of the attribute data[C]. The 6th Int Conf on Fuzzy Systems and Knowledge Discovery. Tianjin, 2009, 1: 71-75.
- [9] 张涛, 洪文学. 多元图表示原理的散点图分类器设计[J]. 小型微型计算机系统, 2010, 31(7): 1433-1438.
- (Zhang T, Hong W X. Scatter classifier design based on multivariate graphical representation theory[J]. J of Chinese Computer Systems, 2010, 31(7): 1433-1438.)
- [10] 张涛, 洪文学. 基于计算几何的非线性可视化分类器设计[J]. 电子学报, 2011, 39(1): 53-58.
- (Zhang T, Hong W X. A nonlinear visual classifier based on computational geometry[J]. Acta Electronica Sinica, 2011, 39(1): 53-58.)
- [11] Chen Degang, Zhao Suyun. Local reduction of decision system with fuzzy rough sets[J]. Fuzzy Sets and Systems, 2010, 161(13): 1871-1883.
- [12] Hao Jing, Li Qingguo. The relationship between  $L$ -fuzzy rough set and  $L$ -topology[J]. Fuzzy Sets and Systems, 2011, 178(1): 74-83.
- [13] 李洁, 邓一鸣, 沈士团. 基于模糊区域分布的分类规则提取及推理算法[J]. 计算机学报, 2008, 31(6): 934-941.
- (Li J, Deng Y M, Shen S T. Classification rule extraction based on fuzzy area distribution and classification reasoning algorithm[J]. Chinese J of Computers, 2008, 31(6): 934-941.)
- [14] Zhou Enwang, Khotanzad Alireza. Fuzzy classifier design using genetic algorithms[J]. Pattern Recognition, 2007, 40(12): 3401-3414.

(上接第568页)

- [5] Hu X H, Cercone N. Learning in relational satabases: A rough set approach[J]. Computational Intelligence, 1995, 11(2): 323-337.
- [6] 张文修, 吴伟志, 梁吉业, 等. 粗糙集理论与方法[M]. 北京: 科学出版社, 2001: 22-25.
- (Zhang W X, Wu W Z, Liang J Y, et al. Theory and method of rough set[M]. Beijing: Science Press, 2001: 22-25.)
- [7] 刘文军, 谷云东, 冯艳宾, 等. 基于可辨别矩阵和逻辑运算的属性约简算法[J]. 模式识别与人工智能, 2004, 17(1): 119-123.
- (Liu W J, Gu Y D, Feng Y B, et al. An improved attribute reduction algorithm of decision table[J]. Pattern Recognition and Artificial Intelligence, 2004, 17(1): 119-123.)
- [8] 支天云, 苗夺谦. 二进制可辨矩阵的变换及高效属性约简算法的构造[J]. 计算机科学, 2002, 29(2): 140-142.
- (Zhi T Y, Miao D Q. The binary discernibility matrix's transformation and high efficiency attributes reduction algorithm's conformation[J]. Computer Science, 2002, 29(2): 140-142.)
- [9] 徐章艳, 杨炳儒, 宋威. 基于简化的二进制差别矩阵的快速属性约简算法[J]. 计算机科学, 2006, 33(4): 155-158.
- (Xu Z Y, Yang B R, Song W. Quick attribution reduction algorithm based on simple binary discernibility matrix[J]. Computer Science, 2006, 33(4): 155-158.)
- [10] 杨萍, 李济生, 黄永宣. 一种基于二进制区分矩阵的属性约简算法[J]. 信息与控制, 2009, 38(1): 70-74.
- (Yang P, Li J S, Huang Y X. An attribute reduction algorithm based on binary discernibility matrix[J]. Information and Control, 2009, 38(1): 70-74.)
- [11] 蒙祖强, 史忠植. 一种新的启发式知识约简算法[J]. 小型微型计算机系统, 2009, 30(7): 1249-1255.
- (Meng Z Q, Shi Z Z. Novel heuristic algorithm for knowledge reduction[J]. J of Chinese Computer Systems, 2009, 30(7): 1249-1255.)
- [12] 葛浩, 李龙澍, 杨传健. 新的可分辨矩阵及其约简方法[J]. 控制与决策, 2010, 25(12): 1891-1895.
- (Ge H, Li L S, Yang C J. New discernibility matrix and attribute reduction method[J]. Control and Decision, 2010, 25(12): 1891-1895.)