

类空间规整度的计算几何组合分类器权重分配

张涛^{1,2}, 洪文学²

¹(燕山大学 信息科学与工程学院, 河北 秦皇岛 066004)

²(燕山大学 电气工程学院, 河北 秦皇岛 066004)

E-mail: zhtao@ysu.com

摘要: 在计算几何组合分类器中, 子分类器的权重分配一直未能充分利用空间视觉信息, 使得分类器的可视化特性无法完全得到发挥. 本文从类空间类别分布特性出发, 提出基于类空间规整度的权重分配方法. 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, 进而利用共生原则分析不同类别在空间中的分布规整度. 由于分布规整度为类别分布信息的整体体现, 可以用于刻画类空间中不同类别样本的离散程度, 因此可以利用当前类空间的规整度信息作为该子分类器的权重. 实验表明, 利用规整度信息进行加权后的分类器不但与可视化特性更好的吻合, 增强了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: 规整度; 计算几何; 组合分类器; 可视化; 共生

中图分类号: TP391

文献标识码: A

文章编号: 1000-4220(2012)07-1572-05

Weight Calculation for Computational Geometry Combining Classifier Using Regularity of Class Space

ZHANG Tao^{1,2}, HONG Wen-xue²

¹(College of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, China)

²(College of Electrical Engineering, Yanshan University, Qinhuangdao 066004, China)

Abstract: In all the tissues about computational geometry combining classifier, the weight calculation for sub classifiers has not taken the advantage of visual information in the spaces, which retains the visual performance about classifier. According to the category distribution in class space, a weight calculation method based on space regulation is proposed. In this method, the space is turned from category information in space to space information in category. And the space regularity is obtained from the later based on co occur rules. As the regularity reflects the distribution of categories and describes the separation of the samples, which makes it as the weight for the sub classifier. The experiments show that the classifier weighted by the regularity not only enhance the visual performance, but also 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: regularity; computational geometry; combining classifier; visual; co occur

1 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^[1]. 其最终目的将不同类别数据在相应空间上找到可以将各类别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^[1-5]可分为线性分类器、非线性分类器、非参数估计分类器与源于 Cover 定理的支持向量机等.

根据待分类数据的分布特性, 不同的分类器表现出不同的分类性能. 其中, 线性分类器的原理是在高维空间中形成线性分类界面, 从而将不同类别数据划入不同区域. 该方法对于线性可分数据效果较好, 但是在实际应用中, 大量数据属于复杂的多元数据, 分布特点为非线性, 此时线性分类器难以胜任. 此时使用支持向量机等非线性分类器可以获得较好效果. 而对于非连续分布数据, 一般采用基于测量思想的分类器, 通

过计算待分类数据与训练数据间的关系判断类别, 属于 voronoi 图的分支应用之一^[6].

针对复杂分布数据, 文献[7]提出可采用可视化模式识别的方法进行分类器设计. 由于可视化模式识别将抽象数据变为视觉范围内, 因此具有天然的物理意义简单、便于理解的特性. Zhang 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^[8]. 其中, 散点图分类器架构^[9]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, 获得了良好的分类效果.

目前, 对于计算几何分类器的研究, 还主要集中在其分类界面的形成过程上. 而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, 采用的是经典的组合分类器权重分配方法. 这样的权重

收稿日期: 2011-05-23 收修改稿日期: 2011-06-24 基金项目: 国家自然科学基金项目(60904100, 61074195) 资助; 河北省自然科学基金项目(F2011203073) 资助. 作者简介: 张涛, 男, 1979年生, 博士研究生, 讲师, 研究方向为模式识别、图像处理; 洪文学, 男, 1953年生, 硕士, 教授, 博士生导师, 研究方向为可视化模式识别、信息融合.

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.文献[10]利用了简单的模糊区域比例方法计算分类器权重,在一定程度上利用了可视化的特性,但由于统计信息简单,对于界面复杂问题无能为力.

基于此,本文提出基于类空间规整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,进而利用共生原则分析不同类别在空间中的分布规整度,从而获得当前类空间的规整度信息作为该子分类器的权重.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.

2 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图如图1所示^[8].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,该步骤为数据可视化重要过程.在像素图表示与扩展模块,将只进行数据表示的传统多元图转换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,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完成最终的分类.从整个的分类过程看,数据在分类器中一直以图形或图像进行表示,因此为可视化分类器;而多维数据生成多个多元图,最用为组合分类器形式.因此,计算几何分类器是典型的可视化组合分类器.



图1 计算几何分类器原理图

Fig.1 Theory graph for computational geometry combining classifier

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 $L(x_i) \in C$.

由于训练样本中存在离群点,这些离群点是过学习的主要成因.为了降低其对分类结果的影响,可以进行量化.设训练样本 X 中属性集合为 $X_j = \{x_{ij} | i = 1, 2, \dots, m\}$, 其分布范围为 $[\min X_j, \max X_j]$, 其中

$$\max X_j = \{x_{ij} | i = \arg \max_a x_{aj}\} \quad (1)$$

$$\min X_j = \{x_{ij} | i = \arg \min_a x_{aj}\} \quad (2)$$

本文采用线性量化完成量化过程,量化阶为 $\tau = \frac{\max X_j - \min X_j}{N}$. 量化后的数据空间集合为

$$V = \{v_n | n = 1, 2, \dots, N\} \quad (3)$$

其中, $v_n = \{x_{ij} | x_{ij} \in [(n-1)\tau, n\tau]\}$. 本文仅考虑1维的情况,因此 $j = 1$, 以下以 x_i 代替 x_{ij} . 通过对 v_n 的分析可以用于

表示样本在该量化区间内的概率分布,如区域 v_n 内样本为 c_j 类的概率可表示为:

$$p(c_j | v_n) = \frac{\#\{i | x_i \in v_n, L(x_i) = c_j\}}{\sum_{p=1}^k \#\{i | x_i \in v_n, L(x_i) = c_p\}} \quad (4)$$

为书写方便,令 $p_n(c_j) = p(c_j | v_n)$, $p_n = \{p_n(c_j) | j = 1, 2, \dots, k\}$. 所有 v_n 对应的类别概率分布集合为 $P = \{p_n | n = 1, 2, \dots, N\}$.

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:

$$p_n = \frac{1}{2} |\text{sign}(p_{n-1} \cdot p_{n+1}) - 2| \cdot (p_{n-1} + p_{n+1}) \quad (5)$$

其中 $\text{sign}(\cdot)$ 为符号函数, $|\cdot|$ 表示取绝对值运算.在二维条件下,其生成过程如图2所示.

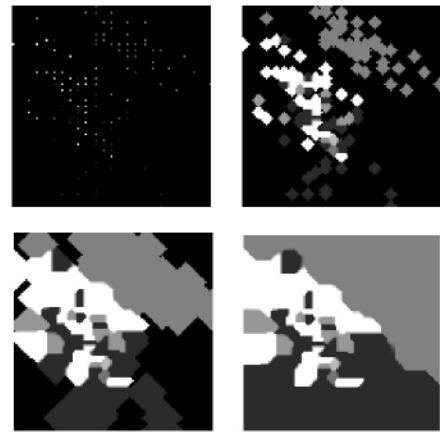


图2 像素图的生长过程

Fig.2 Expansion process of a pixel graph

由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,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.

3 子分类器权重计算

主动生长完成了由采样点集向类空间区域集合的转换.但由于散点图分类器采用了特征组合的方式,使得原始的 d 个属性的数据组成 $N_d = \binom{d}{2}$ 个特征对,从而形成了组合分类器形式.在组合分类器中,权重分配占有重要地位,其分配效果将直接影响最终的分类精度.计算几何分类器作为典型的组合分类器,其权重计算必须符合组合分类器中的分配结果.同时,作为可视化分类器,在权重分配阶段又必须保留可视化的特色.

3.1 分类界面转换

通过对文献[10]的分析可知,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扩展过程相互制约.在每个界面中包含 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. 但作为空间规整度信息的权重计算, 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:

$$p(v_n | c_j) = \frac{p(v_n) p(c_j | v_n)}{p(c_j)} \quad (6)$$

其中, $p(v_n)$ 表示当前位置点上出现的概率. 从界面生成过程可知, 主动生长的终止条件为所有位置点均为基点, 即表示类别信息. 因此 $p(v_n)$ 为均匀分布, 有

$$p(v_n) = \frac{1}{N} \quad (7)$$

$p(c_j)$ 表示当前空间中类别 c_j 出现的概率, 即主动生长后的类别总概率. 因此有

$$\begin{aligned} p(c_j) &= \frac{\sum_{n=1}^N p(c_j | v_n)}{\sum_{i=1}^k \sum_{n=1}^N p(c_i | v_n)} = \\ &= \frac{\sum_{n=1}^N p(c_j | v_n)}{\sum_{n=1}^N \sum_{i=1}^k p(c_i | v_n)} = \\ &= \frac{\sum_{n=1}^N p(c_j | v_n)}{\sum_{n=1}^N \sum_{i=1}^k p(c_i | v_n)} \end{aligned} \quad (8)$$

考虑到主动生长后所有区域均可进行类别的概率表示, 显然

$$\sum_{n=1}^N \sum_{i=1}^k p(c_i | v_n) = N \quad (9)$$

将式(9)代入式(8)得

$$p(c_j) = \frac{1}{N} \sum_{n=1}^N p(c_j | v_n) \quad (10)$$

由此可得, 在特定的特征组合类空间中, 位置为 v_n 的位置上出现类别 c_j 的概率为

$$p(v_n | c_j) = \frac{p(v_n) p(c_j | v_n)}{p(c_j)} = \frac{p(c_j | v_n)}{\sum_{i=1}^k p(c_i | v_n)} \quad (11)$$

式(11)将分类界面由特定点类别信息转化为特定类别的



图3 子分类器分类空间

Fig.3 Category space of sub classifier

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, 便于后期的权重计算. 对于图3的转化结果如图4所示. 其中, 图3为 Iris 数据集的一个子分类器分类空间. 由于 Iris 数据具有 3 个类别, 因此分解为 3 个独立的类别分布空间, 如图4所示. 为了清楚的表示类别的概率分布情况, 图4采用伪彩色表示, 其色彩对应的概率值如各子图右侧标注所示.

3.2 分类界面中共生关系计算

由第3.1节可知, 对特定的组合下, $P_j = \{v_n | p(v_n | c_j), n \in [1, N]\}$ 表示在当前界面下, 类别 c_j 的分布情况. 由经典的模式识别理论可知, "类内紧, 类间松" 是分类器设计的重要

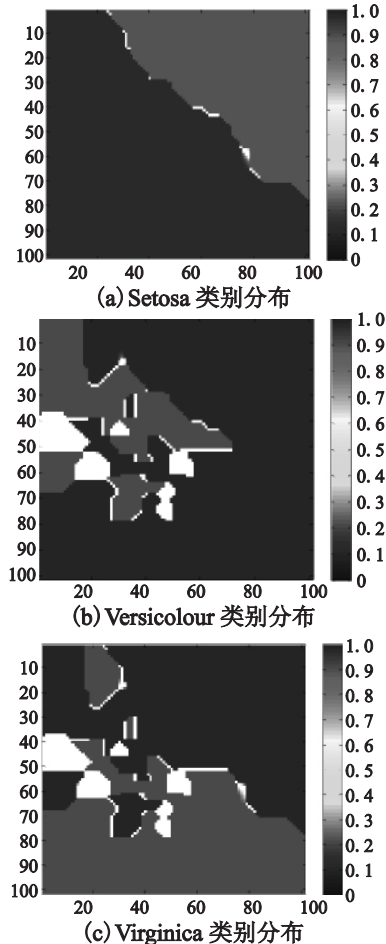


图4 经过空间转换后的类别分布界面

Fig.4 Category probability distribution space turned from Fig.3

原则. 在可视化分类中, 虽然通过主动生长的方式不再强调表示相同类别的区域必须联通, 但由于过多的分类块将加大模糊区域的面积, 导致分类数据落入不确定区域的概率加大, 从而可能导致误判. 因此, 对于该类分类器, 仍遵从聚集优先的原则. 下页图5表示了 Iris 数据形成的 6 个子分类器分类界面, 显然图(f)各类之间可分性更强, 因此更为合理.

为了对同类分布进行度量, 本文提出空间规整度特征. 对于集合 P_j , 由于其表示类别 c_j 在当前空间下的概率分布, 因此该集合下所有信息同类别, 即

$$L(\{v_n | p(v_n | c_j), n \in [1, N]\}) = c_j \quad (12)$$

因此该集合中所有元素对与类别 c_j 具有共生关系. 当前类别下出现概率为 p_1, p_2 的联合概率密度为:

$$\begin{aligned} S_{c_j}(p_1, p_2) &= \\ &= \frac{\#\{(v_i, v_k) | p(v_i | c_j) = p_1, p(v_k | c_j) = p_2, |i - k| \leq dis\}}{N} \end{aligned} \quad (13)$$

式中, $\#(\cdot)$ 表示集合的元素个数, dis 表示在计算联合概率

密度时所需考虑的分布直径. 显然, 对于量化级为 N 的空间, 理论上 $dis \leq N$. 但在实际使用中, 由于将更多的关注类别概率的递变特性, 因此取 $dis = 1$.

由式 (13) 可知, 矩阵 S_{c_j} 表示在当前类别空间中 c_j 类的共生特性. 描述的是在直径为 dis 的范围内 p_1, p_2 共同出现的联合概率. 该联合概率以概率对的形式对类别分布信息进行刻画, 也是对概率对 (p_1, p_2) 出现相对频率的描述. 通过 S_{c_j} , 可以直观的分析集合 P_j 中同类别概率对的分布. 对于聚类良好的空间, 由于其概率为渐变, 因此 S_{c_j} 的对角线附近将出现较高的能量; 而概率突变点或明确的边界大多是由于生硬划分造成, 其必然在矩阵的左下角和右上角造成高频度分布.



图 5 Iris 数据形成的子分类器分类界面

Fig. 5 Category spaces of sub classifiers from Iris dataset

由以上分析可知, S_{c_j} 描述了该类概率在邻近空间中的分布特性. 对其进行统计量分析可以获得界面的分布信息. 对于权重分析而言, 主要目的是分析相似概率的分布状况, 因此可采用 S_{c_j} 概率差的一阶矩:

$$g_{ij} = \int_0^1 \int_0^1 |s_1 - s_2| S_{c_j}(s_1, s_2) d_{s_1} d_{s_2} \quad (14)$$

其中, g_{ij} 表示第 i 中特征组合空间下的第 j 类空间分布规整度. 将式 (13) 带入式 (14) 有,

$$\begin{aligned} g_{ij} &= \frac{1}{N} \int_0^1 \int_0^1 |p_1 - p_2| \# \{ (v_{n1}, v_{n2}) | p(v_{n1} | c_j) = \\ & p_1, p(v_{n2} | c_j) = p_2, |n1 - n2| \leq dis \} d_{p_1} d_{p_2} = \\ & \frac{1}{N} \int_0^1 \int_0^1 DiffP \# \{ (v_{n1}, v_{n2}) | p(v_{n1} | c_j) - p(v_{n2} | c_j) | = \\ & DiffP, |n1 - n2| \leq dis \} d_{DiffP} = \\ & \frac{1}{N} \int_0^1 \int_0^1 DiffP \left| \int_{v_{n1}-dis}^{v_{n1}+dis} sign(|p(v_{n1} | c_j) - p(v_{n2} | c_j)|) = \right. \\ & \left. DiffP \& |n1 - n2| \leq dis \} d_{v_{n2}} d_{v_{n1}} d_{DiffP} \quad (15) \end{aligned}$$

由式 (15) 可知, g_{ij} 表现的是当前空间下类别以该概率差的出现概率为权重的概率差加权和. 该值可以反映整体的概率分布情况, 即特定类别在当前类空间下的规整度. 该数值越小, 则规整度越强, 该界面越适合用于进行分类. 图 4 中各类别分类界面规整度如表 1 所示. 由此可以明显看出, Setosa 类别分布最为理想, Versicolour 类与 Virginica 类分布规整度相当, 但都不如 Setosa 类分布规则性强.

对于特定的子分类器, 由于界面由 k 个类别共同组成. 因此特定特征组合下的子分类空间的平均规整度为

$$g_i = \frac{1}{k} \sum_{j=1}^k g_{ij} \quad (16)$$

由于规整度数值越小表示规整度越高, 进而表示该子分类器所占权重应该越大. 因此规整度与权重数值上应为反比

表 1 图 4 中各类别规整度

Table 1 Regularity of Fig. 4

类 别	规整度
Setosa 类	0.3106
Versicolour 类	0.6817
Virginica 类	0.6778

例关系. 考虑到对于实际数据 $g_i \neq 0$ 且计算几何分类器共生 N_d 个子分类器, 因此第 i 个子分类器的权重为

$$w_i = \frac{(g_i)^{-1}}{\sum_{l=1}^{N_d} (g_l)^{-1}} \quad (17)$$

由此可知, 图 5 对应各子分类器分类界面规整度与权重如表 2 所示. 可以看到, 权重的分布与视觉上的类别聚合程度正相关, 因此该权重计算方法符合可视化特性.

表 2 图 5 中各子分类器分类界面规整度与权重

Table 2 Regularity and weight of Fig. 5

子分类器	规整度	权重
(a)	0.5567	0.1500
(b)	0.6644	0.1257
(c)	0.5290	0.1578
(d)	0.5991	0.1394
(e)	0.4434	0.1883
(f)	0.3496	0.2388

在分类过程中, 设某数据在第 i 个子分类器中对应类别 c_j 的概率为 $p(c_j | i)$, 则组合分类器对该类别的概率判断为

$$p(c_j) = \sum_{i=1}^{N_d} w_i p(c_j | i) \quad (18)$$

4 实验设计与结果分析

4.1 分类实验

为了测试该加权算法在计算几何分类器中的分类性能, 本文利用 UCI 数据库中的几个典型数据集进行分类性能的测试与分析. 根据综合数据的分布特点与采集来源领域的不同, 本文所选择的进行测试的数据集特性列表见下页表 3.

从数据分布特点上看, Pima-Indians-diabetes 数据集与 liver-disorder 数据集类别数目较少且大多为线性分布; Iris 数据集, wine 数据集和 glass 数据集是分类器分类性能测试中最常用的几个数据. 不但因为其类别为多类数据, 而且数据分布在不同类别间也表现为明显的线性与非线性交叉分布; 而 ionosphere 数据集与 breast cancer 数据集虽然分类类别较少, 但特征数目较多, 长用对高维数据分类器的性能分析.

从数据来源看, 本文所选用的 7 个数据集内容涉及到产品分类、物种识别、疾病诊断、气象分析等多个应用, 隶属于生命科学领域和物理科学领域, 对其进行测试不但具有理论意

义,而且具有现实意义^[10,11]。

在评价方法上,本文采用留一法交叉验证(leave one out cross validation, LOOCV)下的分类器分类精度进行分类性能表示。原因在于留一法可以避免训练集与测试集之间的依赖问题,其测试结果是无偏的。

表3 实验用到的各数据集属性

Table 3 Attributes of the data sets in our experimentations

数据集	所属领域	特征数	类别数	样本数
Pima-Indians-diabetes	生命科学	8	2	768
liver-disorder	生命科学	6	2	345
Iris	生命科学	4	3	150
wine	物理科学	13	3	178
glass	物理科学	9	6	214
ionosphere	物理科学	34	2	351
breast cancer	生命科学	30	2	569

在测试结果的表现上,本文在参数相同情况下分别利用等加权方式、模糊加权方式^[10]与本文提出的规整度加权方式对表3中各数据集进行分类性能测试。在不同加权方法下,其分类精度如表4所示。

表4 不同分类器分类性能对比

Table 4 Comparison of performance from different classifiers

数据集	等加权	模糊加权	规整度加权
Pima-Indians-diabetes	71.88%	75.13%	75.13%
liver-disorder	59.42%	68.41%	70.15%
Iris	89.67%	96.67%	97.33%
wine	76.97%	98.31%	99.44%
glass	64.49%	76.17%	82.49%
ionosphere	83.71%	90.32%	88.50%
breast cancer	92.09%	95.43%	96.13%

4.2 结果分析与讨论

以上实验结果表明,由于模糊加权与本文提出的规整度加权均考虑了视觉特性,因此其分类性能均高于未考虑分布特性的平均加权算法。对于低维度二分类问题,本文构造的规整度加权方式与模糊加权差别相比优势并不明显。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。同时,模糊加权与规整度加权均考虑了数据的分布特性,因此权重计算结果相仿。

对于中等复杂度的分类问题,规整度加权方式的分类精度则全面超过了模糊加权的分类精度。主要原因在于随着数据的维数或类别数的增加,各子分类器的分布特性变得复杂。此时简单的模糊加权已经无法对分类界面的规则进行细致描述。而由于规整度方法基于共生关系,因此表现出了更好界面特征描述能力。

而对于高维分类问题,规整度加权方式的分类精度则与模糊加权的分类器精度性能相当。其主要原因在于虽然维数的增加带来了子分类器个数的增加,但类别数据较少使得在单个子分类界面中,数据分布相对简单,因此二者分析结果相近。

从对7个UCI数据的综合实验结果来看,规整度加权方式在liver-disorder、Iris、glass等5个数据集的分类精度优于

现有的计算几何分类器水平,而其他几个数据集的分类精度也与模糊加权方式分类器相当。因此,该实验充分证明用规整度信息进行加权后的分类器不但与分类器的可视化特性更好的吻合,而且在分类精度上得到了进一步的提升。

5 结论

针对计算几何可视化组合分类器的加权问题,本文提出一种基于规整度的子分类器权重计算方法。该方法通过对子分类器中相同类别数据的概率分布分布其共生特性,并以共生概率差的一阶矩作为规整度特性进行分析。通过UCI数据集的实验分析表明,利用该方法的加权不但与分类器的可视化特性更好的吻合,而且在分类精度上得到了进一步的提升。

但目前该加权方法存在这一定问题,比如为了获取分类界面规整度必须完成复杂的主动生长过程,如何在主动生长之前对规整度进行估算,从而简化整个计算几何分类器的计算复杂度是今后研究中需要重点解决的问题。

References:

- [1] Duda R O, Hart P E, Stork D G. Pattern classification[M]. New York: Wiley, 2000.
- [2] McLachlan G J. Discriminant analysis and statistical pattern recognition[M]. New York: Wiley Interscience, 2004.
- [3] Vapnik V. Statistical learning theory[M]. New York: Wiley Interscience, 1998.
- [4] Anil K Jain, Robert P W Duin, Mao Jian-chang. Statistical pattern recognition: a review[J]. IEEE Transaction 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]. Chemo Intel. Lab. Syst, 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 Yong-hong, Hong Wen-xue, Chen Na, et al. Parallel filter: a visual classifier based on parallel coordinates and multivariate data analysis[C]. Third International Conference on Intelligent Computing, 2007, 4682: 1172-1183.
- [8] Zhang Tao, Hong Wen-xue. A novel visual combining classifier based on a two-dimensional graphical representation of the attribute data[C]. Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 2009, 1: 71-75.
- [9] Zhang Tao, Hong Wen-xue. Scatter classifier design based on multivariate graphical representation theory[J]. Journal of Chinese Computer Systems, 2010, 31(7): 1433-1438.
- [10] Zhang Tao, Hong Wen-xue. A nonlinear visual classifier based on computational geometry[J]. Acta Electronica Sinica, 2011, 39(1): 53-58.
- [11] Li Jie, Deng Yi-ming, Shen Shi-tuan. Classification rule extraction based on fuzzy area distribution and classification reasoning algorithm[J]. Chinese Journal of Computers, 2008, 31(6): 934-941.

附中文参考文献:

- [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.
- [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.
- [11] 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.