

文章编号: 1001-0920(2005)04-0372-05

粗集理论中连续属性的广义离散化

张葛祥^{1,3}, 金炜东², 胡来招³

(1. 西南交通大学 计算机与通信工程学院, 四川 成都 610035; 2 西南交通大学 电气工程学院, 四川 成都 610031; 3 电子对抗国防科技重点实验室, 四川 成都 610036)

摘要: 提出一种处理区间属性值离散化问题的新方法, 其特点是在离散化过程中强调类别可分离性。首先拓展已有的基于断点划分的离散化定义, 给出基于类别可分离性的广义离散化定义; 然后描述了广义离散化算法并进行时间复杂性分析; 最后利用辐射源信号进行了仿真实验。结果表明, 该方法能有效离散区间属性决策系统, 简化分类器的设计和提高识别率, 从而拓展了粗集理论的应用范围。

关键词: 粗集理论; 连续属性; 离散化; 特征选择; 识别

中图分类号: TP18, TP301.6

文献标识码: A

Generalized discretization of continuous attributes in rough set theory

ZHANG Ge-xiang^{1,3}, JIN Wei-dong², HU Lai-zhao³

(1. School of Computer and Communication Engineering, Southwest Jiaotong University, Chengdu 610031, China; 2 School of Electrical Engineering, Southwest Jiaotong University, Chengdu 610031, China; 3 National EW Laboratory, Chengdu 610036, China Correspondent: ZHANG Ge-xiang, E-mail: dylan7237@sina.com)

Abstract: An approach is proposed to discretize continuous interval-valued attributes in rough set theory. The approach is characterized by emphasizing class-separability in discretization. The existing discretization definition based on cut-splitting is extended and the generalized discretization definition based on class-separability criterion is given. Then the detailed algorithm of generalized discretization is described and its time complexity is analyzed. Finally, simulation experiment is made on radar emitter signals. Experimental results show that the proposed approach discretizes the decision table with continuous interval-valued attributes efficiently and has good characteristics of simplifying classifier design and heightening recognition rate. So the approach expands the application scope of rough set theory.

Key words: rough set theory; continuous attribute; discretization; feature selection; recognition

1 引言

粗集理论(RST)是一种处理不完整不精确知识的新型数学工具,是当前倍受关注的一种软计算基础理论^[1]。RST 无需先验知识便能从大量数据或经验中获取知识并生成决策规则^[2,3],故 RST 自 Pawlak 提出以来得到了广泛应用^[3,4-8]。但 RST 只能处理离散的属性值,实际应用中大量存在的连续

属性值必须经离散化后才能利用 RST 方法进行处理,因此连续属性的离散化问题成为 RST 研究中的一个非常重要的方向^[2,6-8]。离散化问题的研究不仅可拓展 RST 的应用领域,而且可节省连续属性值的存储空间和处理时间,并可简化分类器的结构和设计过程,加快学习速度^[8]。

目前,虽然已提出多种离散化方法^[2,5-8],但无

收稿日期: 2004-06-10; 修回日期: 2004-08-09

基金项目: 国家自然科学基金项目(69574026); 电子对抗技术预研基金项目(NEWL51435QT220401); 教育部高等学校骨干教师计划项目(教技司[2000]65号)。

作者简介: 张葛祥(1974—),男,四川绵阳人,博士生,从事粗集理论、模式识别等研究; 金炜东(1959—),男,安徽淮南人,教授,博士生导师,从事智能信息处理、系统仿真等研究

论是硬离散化方法^[2,5,6], 还是软离散化方法^[7,8], 都是针对固定点上的连续属性值的. 在离散化前, 需要人为确定划分的维数或需要预先给定一些参数, 且这些离散化方法都是基于寻找最佳断点集. 在实际应用中, 尤其在模式识别中, 由于噪声等多种因素的影响, 特征值往往不是固定的, 而是在某一区间内变化的. 对于这类难以进行断点划分的离散化问题, 已有的离散化方法无法进行有效处理. 在模式识别中, 属性约简的重要目的是使找出的简化属性集获得最高的正确识别率, 至于是否需要进行断点划分以及是否找到最优断点集并不重要. 求解最佳断点集是一个 NP 难问题, 在实际应用中应该尽量避免.

针对上述问题, 本文首先对目前 RST 研究中普遍采用的离散化定义进行分析, 并对其进行拓展, 给出广义离散化问题的定义, 然后提出一种离散化新方法, 对属性值在一定范围内连续变化的决策表进行处理. 实验结果证明了该方法的有效性和实用性.

2 广义离散化问题描述

为了分析方便, 首先将已有的离散化定义^[2]重新描述如下:

决策表 $S = \langle U, R, V, f \rangle$, $R = C \cup \{d\}$ 为属性集合, 子集 C 和 $\{d\}$ 分别称为条件属性集和决策属性集, 论域 $U = \{x_1, x_2, \dots, x_n\}$ 为有限的对象集合. 设决策种类的个数为 $r(d)$; 属性 a 的值域 V_a 上的一个断点可记为 (a, c) , $a \in R, c$ 为实数集; 在值域 $V_a = [l_a, r_a]$ 上的任意一个断点集合 $\{(a, c_1^a), (a, c_2^a), \dots, (a, c_{k_a}^a)\}$ 定义了 V_a 上的一个分类 P_a , 即

$$P_a = \{[c_0^a, c_1^a], [c_1^a, c_2^a], \dots, [c_{k_a}^a, c_{k_a+1}^a]\}, \quad (1)$$

$$l_a = c_0^a < c_1^a < \dots < c_{k_a}^a < c_{k_a+1}^a = r_a, \quad (2)$$

$$V_a = [c_0^a, c_1^a) \cup [c_1^a, c_2^a) \cup \dots \cup [c_{k_a}^a, c_{k_a+1}^a] \quad (3)$$

因此, 任意的 $P = \{P_a\}_{a \in R}$ 定义了一个新的决策表 $S^P = \langle U, R, V^P, f^P \rangle$, $f^P(x_a) = i \Leftrightarrow (x_a \in [c_i^a, c_{i+1}^a))$, 对于 $x \in U, i \in \{0, 1, \dots, K_a\}$, 即经离散化后, 一个新的信息系统代替原信息系统.

上述离散化定义的本质可归结为利用选取的断点划分由条件属性构成的空间. 由于条件属性是确定的实数值, 通过确定断点就可划分条件属性空间, 而已有离散化方法^[2,5-8]恰好致力于求解最佳断点集, 且已证明它是一个 NP 难问题.

在模式识别或机器学习中, 决策表中的条件属性是从对象中提取的各个特征, 于是条件属性空间的划分问题转化为特征空间的划分问题. 利用寻找断点的方法划分特征空间需要人为确定划分的子空间数, 它未考虑特征的可分性问题, 因而该离散化方法在模式识别和机器学习中是不合理的; 而且由于

噪声等多种因素的影响, 所提取出的特征常常不是一个固定值, 而是一个区间值或近似服从某种分布的函数值, 这样, 断点划分法就无法对特征值进行离散化. 基于上述原因, 本文给出下列广义离散化问题的描述:

$S = \langle U, R, V, f \rangle$, $R = C \cup \{d\}$ 为属性集合, 其中: $C = \{c_1, c_2, \dots, c_m\}$ 和 $\{d\}$ 分别称为条件属性集和决策属性集, $U = \{x_1, x_2, \dots, x_n\}$ 为有限的对象集合, 即论域; $c_i (c_i \in C, i = 1, 2, \dots, m)$ 的值域为

$$V_{c_i} = \{[v_{c_i}^{x_1, \min}, v_{c_i}^{x_1, \max}], [v_{c_i}^{x_2, \min}, v_{c_i}^{x_2, \max}], \dots, [v_{c_i}^{x_n, \min}, v_{c_i}^{x_n, \max}]\}. \quad (4)$$

其中: $v_{c_i}^{x_j, \min}, v_{c_i}^{x_j, \max} \in R (j = 1, 2, \dots, n)$ 采用类别可分离性准则函数 $J(V_{c_i})$ 对属性 $c_i (c_i \in C)$ 进行划分, 得到论域 U 的一个等价关系 (论域 U 的一个分类) R_{c_i} . 这样, 对于属性集 C , 可以得到由 m 个等价关系 $R_{c_1}, R_{c_2}, \dots, R_{c_m}$ 构成的等价关系族

$$P = \{R_{c_1}, R_{c_2}, \dots, R_{c_m}\}.$$

P 定义了一个新的决策表

$$S^P = \langle U, R, V^P, f^P \rangle,$$

$$f^P(x) = k, x \in U, k = \{0, 1, \dots\}.$$

经离散化后, 原决策表被一个新的决策表代替.

广义离散化定义将点属性值离散化问题转化为区间属性值离散化问题, 其实质是将连续属性的离散化过程视为一个函数映射关系, 拓展了离散化过程的含义. 定义中还将这种函数关系称为类别可分离性准则函数, 以说明离散化过程强调的是类别的可分性. 该定义为属性值是区间值的离散化问题提供较好的求解途径, 同时将粗集理论中连续属性的离散化问题与模式识别、机器学习中的特征选择问题直接联系起来. 在离散过程中考虑类别可分性问题, 有助于简化分类器的设计和提高分类识别率.

如果属性值 $v_{c_j}^{x_j, \min} = v_{c_j}^{x_j, \max}, j = 1, 2, \dots, n$, 则区间属性值离散化问题就转化为点属性值离散化问题. 此时, 如果类别可分离性准则函数是一个特殊的函数, 它将属性的值域划分为若干个区间, 即在属性的值域上确定了一个断点集, 则广义离散化定义即变为一般离散化定义.

3 离散化方法

从对广义离散化问题的描述中可看出, 实现区间属性值离散化的关键是类别可分离性准则函数的选取. 在描述实现连续属性离散化的具体算法之前, 首先给出一种类别可分离性准则函数.

当属性值在某一区间上取值时, 一般而言, 属性值总会呈现出一定的规律性. 虽然由于实际应用领域的复杂性, 这种规律将呈现出多样性, 但通过测

量而得到的特征值遵循一定的统计规律。本文只讨论属性值有一定规律性的离散化问题,可将这种规律性看成近似地遵循某种概率分布。设 $f(x)$ 和 $g(x)$ 均为一维连续非负实函数,分别表示决策表中属性 c_i ($c_i \in C, i = 1, 2, \dots, m$)对应于论域 U 中任意两个对象的属性值概率分布函数,于是可引入如下的类别可分离性准则函数:

$$J = 1 - \frac{\int f(x)g(x)dx}{\sqrt{\int f^2(x)dx} \sqrt{\int g^2(x)dx}} \quad (5)$$

函数 J 满足文献[9]中的类别可分离性判据的3个条件:

- 1) 准则函数值为非负;
- 2) 当两类分布函数完全不交叠时,准则函数值取得最大值;
- 3) 当两类分布函数完全相同时,准则函数值应为0

由于 $f(x)$ 和 $g(x)$ 均为非负实函数,可由Cauchy-Schwartz不等式得到

$$0 \leq \int f(x)g(x)dx \leq \sqrt{\int f^2(x)dx} \sqrt{\int g^2(x)dx}, \quad (6)$$

即式(5)中准则函数 J 的值域为 $[0, 1]$ 。

J 可用于判断两类别是否可分,现说明如下:若函数 $f(x)$ 和 $g(x)$ 分别表示论域 U 中两个对象 A 和 B 的属性值概率分布函数,则 A 类和 B 类可分离的几种情况如图1所示。

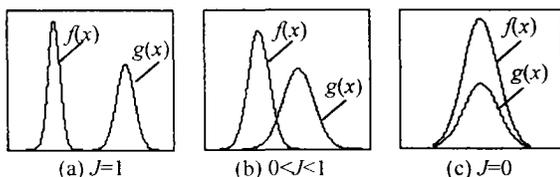


图1 函数 $f(x)$ 和 $g(x)$ 可分离的3种情况

1) A 与 B 完全可分。若对所有使 $f(x) > 0$ 的点都有 $g(x) = 0$,或对所有使 $g(x) > 0$ 的点都有 $f(x) = 0$,则 A 与 B 两类是完全可分的,此时准则函数 J 取得最大值1,如图1(a)所示。

2) A 与 B 部分可分。若对某些使 $f(x) > 0$ 的点都有 $g(x) > 0$,则 A 与 B 两类是部分可分的,此时准则函数 $0 < J < 1$,如图1(b)所示。

3) A 与 B 完全不可分。若对所有 x 都有 $f(x) = kg(x)$, $k \in \mathbb{R}^+$,则 A 与 B 两类是完全不可分的,准则函数取得最小值 $J = 0$,如图1(c)所示,图中 $k =$

因此,利用式(5)的准则函数判断两类的可分离程度是合理的。需要说明的是,式(5)中的两函数 $f(x)$ 和 $g(x)$ 为满足条件的任意分布函数。

根据类别可分离性准则函数,给出如下区间属性值的离散化算法:

Step 1: 初始化。确定论域 U 中对象的数目 n 和属性集中的元素数目 m 。

Step 2: 构造决策表。将所有属性值排列成以对象为行以属性为列的二维决策表,表中的属性值均表示为区间值形式。

Step 3: 选择类别可分离性阈值 T_h 。从式(5)和图1可知,该阈值的含义表示离散化过程中允许两对象重叠的比例为 $1 - T_h$ 。在选取过程中,根据不同应用对象而定,一般取接近于1的正数。

Step 4: 将属性 c_i (开始时 $i = 1$)中的所有属性值按中心点值从小到大排序,排序后结果为 v_1, v_2, \dots, v_n 。

Step 5: 将属性 c_i 中的最小属性值 v_1 编码为Code = 0,作为离散过程的初始值。

Step 6: 从属性 c_i 中的最小属性值 v_1 开始,按排序后的顺序依次计算 v_k 和 v_{k+1} ($k = 1, 2, \dots, n - 1$)的可分离性准则函数值 J_k ,如果 $J_k \geq T_h$,表示这两类对象完全可分,则使属性值 v_{k+1} 的对应位置编码Code加1,否则, $J_k < T_h$,表示该两类对象不完全可分,属性值 v_{k+1} 对应位置的编码Code保持不变。

Step 7: 重复执行Step 6,直到属性 c_i 中的所有属性值均被离散化为止。

Step 8: 如果 $i = m$,表明还有属性没有进行离散化,则 $i = i + 1$,算法转到Step 4,直到 $i > m$ 为止,表明决策表中所有属性值均被离散化。

Step 9: 采用离散化后的决策表代替原决策表,用于属性化简。

该算法的时间复杂度主要由Step 4和Step 6~Step 8决定。在Step 4中采用的排序算法不同,相应的时间复杂度也不同。简单排序算法、快速排序算法和线性时间排序算法的时间复杂度分别为 $O(n^2)$ 、 $O(n \log(n))$ 和 $O(n)^{[10]}$ 。Step 6和Step 7的时间复杂度为 $O(n)$,Step 8的时间复杂度为 $O(m)$ 。当采用简单排序算法、快速排序算法和线性时间排序算法时,离散化算法的时间复杂度分别为 $O((n^2 + n)m)$ 、 $O((n \log(n) + n)m)$ 和 $O((n + n)m)$,即 $O(mn^2)$ 、 $O(mn \log(n))$ 和 $O(mn)$ 。

4 应用实例

在雷达辐射源信号识别中,一般采用经验式或启发式的方法进行特征提取。提取出的特征往往带有主观性和猜测性,因而在分类识别前需要提取多

表 1 离散处理前的属性表

属性	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
x_1	[0 690, 0 707]	[1 402, 1 404]	[0 122, 0 125]	[0 271, 0 272]	[0 435, 0 437]	[1 962, 2 301]	[0 952, 1 021]	[0 420, 0 433]
x_2	[0 671, 0 682]	[1 430, 1 576]	[0 115, 0 162]	[0 214, 0 277]	[0 362, 0 428]	[1 088, 2 991]	[1 115, 1 261]	[0 406, 0 455]
x_3	[0 546, 0 598]	[1 582, 1 586]	[0 292, 0 304]	[0 225, 0 226]	[0 374, 0 376]	[0 726, 1 111]	[0 691, 0 755]	[0 534, 0 550]
x_4	[0 893, 0 896]	[1 437, 1 490]	[0 374, 0 652]	[0 990, 1 000]	[0 856, 0 867]	[11 738, 20 222]	[2 157, 2 960]	[0 199, 0 201]
x_5	[0 897, 0 909]	[1 561, 1 768]	[0 200, 0 223]	[0 725, 0 727]	[0 330, 0 337]	[6 817, 7 748]	[2 899, 3 130]	[0 199, 0 202]
x_6	[0 419, 0 453]	[1 219, 1 224]	[0 562, 0 586]	[0 141, 0 142]	[0 243, 0 244]	[2 618, 2 967]	[0 195, 0 310]	[0 772, 0 787]
x_7	[0 396, 0 413]	[1 281, 1 284]	[0 080, 0 082]	[0 040, 0 040]	[0 000, 0 001]	[6 404, 6 726]	[0 257, 0 326]	[0 676, 0 687]
x_8	[0 657, 0 684]	[1 479, 1 483]	[0 066, 0 067]	[0 182, 0 191]	[0 115, 0 129]	[2 147, 3 314]	[1 423, 1 664]	[0 393, 0 413]
x_9	[0 521, 0 562]	[1 394, 1 400]	[0 078, 0 081]	[0 707, 0 729]	[0 584, 0 612]	[17 351, 19 069]	[0 922, 0 977]	[0 504, 0 524]
x_{10}	[0 905, 0 908]	[1 384, 1 459]	[0 332, 0 348]	[0 625, 0 630]	[0 556, 0 561]	[0 345, 4 668]	[3 114, 3 737]	[0 202, 0 211]

表 2 离散处理后的属性表

属性	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
x_1	2	2	0	0	1	0	2	1
x_2	2	2	0	0	1	0	3	1
x_3	1	2	1	0	1	0	1	2
x_4	3	2	1	1	2	3	5	0
x_5	3	2	0	1	1	2	5	0
x_6	0	0	1	0	1	0	0	4
x_7	0	1	0	0	0	1	0	3
x_8	2	2	0	0	1	0	4	1
x_9	1	2	0	1	1	3	2	2
x_{10}	3	2	1	0	1	0	5	0

种特征, 并从中挑选出最有效的特征, 以消除特征提取的主观性, 简化分类器设计和提高正确识别率。由于受噪声等多种因素的影响, 雷达辐射源信号的特征值会在一定范围内连续变化, 当采用 RST 对各特征进行筛选时, 需要首先进行离散化处理。

本文选取 10 种典型雷达辐射源信号进行仿真实验。10 种信号分别表示为 x_1, x_2, \dots, x_{10} , 它们构成论域 U , 属性集由 8 种特征组成, 分别表示为 a_1, a_2, \dots, a_8 。在 5 ~ 20 dB 的信噪比范围内对 10 种雷达辐射源信号进行 8 种特征提取, 得到如表 1 所示的属性表。表中所有属性值均为区间值, 对于每个属性值, 表中给出的是其最小值和最大值。

采用上述方法对表 1 中的属性值进行离散化处理, 得到如表 2 所示的新决策表, 在离散化过程中, 参数 $n = 10, m = 8, T_h = 0.99$, 各属性值均视为以数学期望和方差为参数的高斯分布函数。

将连续属性进行离散化后, 即可采用粗集理论方法对表 2 中的属性进行约简。本文采用基于可辨识矩阵和逻辑运算的属性约简方法^[2,3]对表 2 进行约简, 属性约简的算法如下:

Step 1: 计算决策表的可辨识矩阵 C_D 。

Step 2: 对于可辨识矩阵中的所有取值为非空集合的元素 $C_{ij} (C_{ij} \neq 0, C_{ij} \neq \emptyset)$, 建立相应的析取逻辑表达式

$$L_{ij} = \bigvee_{a_i \in C_{ij}} a_i \quad (7)$$

Step 3: 将所有的析取逻辑表达式 L_{ij} 进行合取运算, 得到一个合取范式

$$L = \bigwedge_{C_{ij} \neq 0, C_{ij} \neq \emptyset} L_{ij} \quad (8)$$

Step 4: 将合取范式 L 转化为析取范式, 得

$$L = \bigvee_i L_i \quad (9)$$

Step 5: 输出属性约简结果。析取范式中的每个合取项对应一个属性约简的结果, 每个合取项中所包含的属性组成约简后的条件属性集合。

对表 2 进行属性约简后得到如下结果:

$$L = a_7(a_6 + a_3a_4 + a_5(a_4 + a_1(a_2 + a_3) + a_2a_8)) \quad (10)$$

采用粗集理论方法可得到所有可能的约简结果, 往往导致简化后的属性子集存在多解问题, 因此本文引入特征提取时间复杂性对多个约简结果进行再次挑选。由式(10)可知, 表 2 经约简后共有 6 种简化结果, 通过对 6 种结果进行复杂性计算后, 得到具有最低复杂性的结果 $a_4a_5a_7$ 。

为验证属性约简结果的性能, 本文采用 3 层 BP 神经网络分类器对 10 种雷达辐射源信号进行分类识别。首先对每一种雷达辐射源信号在 5 ~ 20 dB 的信噪比(只考虑脉内噪声)范围内每隔 5 dB 产生 150 个样本, 这样每种雷达辐射源信号总共有 600 个样本, 其中 200 个样本用于神经网络分类器训练, 其余 400 个样本用于分类识别的测试实验。因属性约简的结果为 $a_4a_5a_7$, 故 BP 神经网络分类器的结构选为 3-15-10。即输入层有 3 个神经元, 分别作为 3 种特征的输入; 隐含层的神经元数目选为 15, 神经元

表3 属性约简前后的平均正确识别率

信号	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	平均
约简前	96.61	90.85	95.32	96.67	98.35	98.56	100.0	98.76	95.82	89.62	96.05
约简后	97.05	87.10	98.71	100.0	99.5	97.65	100.0	100.0	99.59	92.71	97.23

传递函数采用 Tansig 函数; 输出层有 10 个神经元, 对应 10 种雷达辐射源信号, 输出层传递函数采用 Logsig 函数

BP 神经网络分类器的训练算法采用弹性反向传播算法^[11], 监督信号取为 0-1 型信号, 输出容差为 0.05, 训练误差取为 0.001. 经 20 次测试实验, 所得 10 种雷达辐射源信号的平均正确识别率如表 3 所示. 为了进行对比, 本文将属性约简前的 8 种特征也用于 10 种雷达辐射源信号的分类识别, 分类器仍采用 BP 神经网络, 结构为 8-20-10, 其余参数或条件与前面相同, 分类识别结果如表 3 所示. 表 3 显示, 经属性约简, 不仅简化了分类器结构, 而且获得了 97.23% 的平均正确识别率, 比属性约简前提高 1.18%.

5 结 语

在模式识别、机器学习等领域, 当采用粗集理论方法进行特征选择时, 常常存在一类属性值为区间的连续属性离散化问题, 已有的离散化方法难以处理这类问题. 本文针对该问题, 在基于对寻找最佳断点集的离散化定义进行推广的基础上, 提出一种基于类别可分离性准则的离散化方法, 可有效解决区间属性值的离散化问题. 雷达辐射源信号特征选择和分类识别的仿真实验结果表明, 本文方法能对区间属性值信息系统进行合理的离散化, 避免了寻找最佳断点集的难题. 此方法对类似的连续属性离散化问题的研究具有一定的参考价值.

参考文献(References)

[1] Lin T Y. Introduction to the special issue on rough sets

[J]. *Int J of Approximate Reasoning*, 1996, 15(4): 287-289.

[2] 王国胤. *Rough 理论与知识获取*[M]. 西安: 西安交通大学出版社, 2001. 92-143.

[3] Walczak B, Massart D L. Rough sets theory [J]. *Chemometrics and Intelligent Laboratory Systems*, 1999, 47(1): 1-16.

[4] Pawlak Z. Rough sets [J]. *J of Information and Computer Science*, 1982, 11(5): 341-356.

[5] Shen L X, Tay F E H, Qu L S, et al. Fault diagnosis using rough set theory [J]. *Computers in Industry*, 2000, 43(1): 61-72.

[6] Dai J H, Li Y X. Study on discretization based on rough set theory [A]. *Proc of 1st Int Conf on Machine Learning and Cybernetics*[C]. Piscataway: IEEE Press, 2002: 1371-1373.

[7] Roy A, Pal S K. Fuzzy discretization of feature space for a rough set classifier [J]. *Pattern Recognition Letter*, 2003, 24(6): 895-902.

[8] Sumaga R. Analyzing discretizations of continuous attributes given a monotonic discrimination function [J]. *Intelligent Data Analysis*, 1997, 1(3): 157-179.

[9] 边肇祺, 张学工. *模式识别*[M]. 第 2 版. 北京: 清华大学出版社, 2000: 176-210.

[10] 傅清祥, 王晓东. *算法与数据结构*[M]. 北京: 电子工业出版社, 1998: 253-275.

[11] Riedmiller M, Braun H. A direct adaptive method for faster back propagation learning: The RPROP algorithm [A]. *Proc of the IEEE Int Conf on Neural Networks*[C]. Piscataway: IEEE Press, 1993: 586-591.