

基于 WPA-SVM 的多分类故障混合诊断模型

冼广铭, 曾碧卿

XIAN Guang-ming, ZENG Bi-qing

华南师范大学 南海校区 计算机工程系, 广东 佛山 528225

Computer Engineering Department of Nanhai Campus, South China Normal University, Foshan, Guangdong 528225, China

XIAN Guang-ming, ZENG Bi-qing. Multiple classification fault diagnosis hybrid model based on WPA-SVM. Computer Engineering and Applications, 2008, 44(24): 242-245.

Abstract: Aiming at the difficulty in extracting eigenvectors and the difficulty of traditional SVM algorithm in diagnosis multi-classification, a novel hybrid model for machinery fault diagnosis combining wavelet packet analysis and multiple support vector machine classifier based on binary tree is put forward. According to the method, the energy of different frequency bands after wavelet packet decomposition, as eigenvectors, are input into the multiple support vector machine classifier depend on fault priority to find support vectors and construct hyperplane. And then testing samples are diagnosed according to the hyperplane. Experimental results show that the method proposed above is effective.

Key words: wavelet packet; Support Vector Machine; multiple classification

摘要: 针对目前机械故障诊断中难以进行特征提取和常规 SVM 算法诊断多类分类问题时存在困难等问题, 提出了结合了 WPA 理论和基于二叉树的多级 SVM 分类器的 WPA-SVM 多分类故障混合诊断模型。采用小波包分析对机械信号提取频域能量特征向量, 通过训练多个依赖故障优先级的基于二叉树的多级 SVM 分类器中, 找到样本中的支持向量, 并以此决定超平面。然后根据最优分类平面, 对测试集的样本进行故障诊断。通过对两种不同特征提取方法、三种不同 SVM 识别策略的实验比较结果可知, 该方法是有效的。

关键词: 小波包; 支持向量机; 多分类

DOI: 10.3778/j.issn.1002-8331.2008.24.073 文章编号: 1002-8331(2008)24-0242-04 文献标识码: A 中图分类号: TP18

1 引言

支持向量机方法是针对二分类问题提出的一种学习算法, 只能解决两类分类问题。但在现实应用中, 许多实际问题都是多类分类问题, 因此有必要采取一定的策略使支持向量机可以处理多类分类问题。

根据给定训练集

$$T = \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\} \in (X \times Y)^l$$

其中 $T = x_i \in X = R^n, y_i \in Y = \{1, 2, \dots, M\}, i = 1, \dots, l$, 寻找一个决策函数 $f(x): X = R^n \rightarrow Y$ 。

由此可见, 求解多分类问题, 实质上是找到一个把 R^n 上的点分成 M 部分的规则。

多级 SVM 分类器的构造方法主要有如下两种:

(1) 一种方法是 Weston^[1] 于 1998 年在经典 SVM 理论上构造新的分类模型并通过 SV 对其目标函数进行优化, 从而实现多级分类。该方法的缺点是目标函数过于复杂, 计算复杂度过高。

(2) 另一种方法^[2-5] 是通过组合多个二分类器来构造多级分类器。其构造策略有“one against all”、“one against one”和“one against others”三种。本文主要采用了这三种策略进行了

研究。

针对目前机械故障诊断^[6]中难以进行特征提取和常规 SVM 算法诊断多类分类问题时存在困难等问题, 本文提出了一种结合了 WPA 理论和基于二叉树的多级 SVM 分类器的 WPA-SVM 机械故障诊断方法。

如图 1 所示, 以小波包变换(或小波变换)作为 SVM 的前处理器以提取故障特征, 通过小波包(或小波)将其分解, 把分解系数输入支持向量机中进行训练, 找到样本中的支持向量, 并以此决定超平面; 第二步, 根据最优分类平面, 对测试集的样本进行故障诊断。这种方法称为基于小波包(小波)和支持向量机^[7-8]的故障诊断方法。

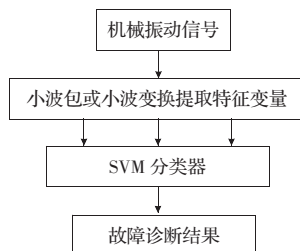
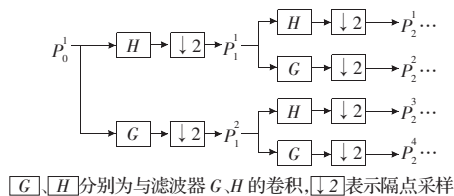


图 1 小波包(或小波)结合支持向量机的故障诊断模型

2 小波包快速算法

小波包分解快速算法也可以用图 2 所示。



[G], [H] 分别为与滤波器 G, H 的卷积, [↓2] 表示隔点采样

图 2 小波包分解快速算法

小波包分解过程可以用图 3 来表示。

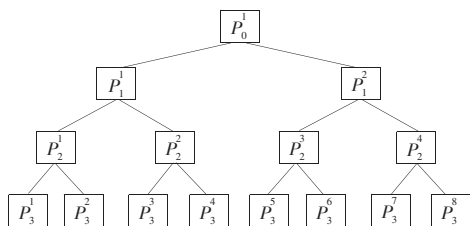
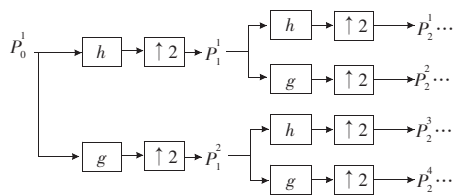


图 3 小波包分解二叉树

小波包重构快速算法(如图 4 所示)的核心包括三个关键的运算:

- (1)与小波滤波器的卷积。
- (2)隔点采样。
- (3)隔点插零。



[g], [h] 分别为与滤波器 h 和 g 的卷积, [↑2] 表示隔点插零

图 4 小波包重构快速算法

3 多类分类的支持向量机

3.1 一对多区分法

支持向量机多类分类方法最早使用的算法就是一对多方法^[9-10](one-against-all)。要得到多类分类机,通常的方法是构造一系列两类分类机,其中的每一个分类机都把其中的一类同余下的各类分开。然后据此推断某个输入 x 的归属。一对多方法对于 k 类问题需要构造 k 个支持向量机子分类器。在构造第 i 个支持向量机子分类器时,将属于第 i 类别的样本数据标记为正类,不属于 i 类别的样本数据标记为负类。测试时,对测试数据分别计算各个子分类器的决策函数值,并选取函数值最大所对应的类别为测试数据的类别。第 i 个支持向量机需要解决下面的最优化问题:

$$\min_{w^i, b^i, \xi^i} \frac{1}{2} (w^i)^T w^i + C \sum_{j=1}^l \xi_j^i \quad (1)$$

$$\text{s.t. } (w^i)^T \phi(x_j) + b^i \geq 1 - \xi_j^i, \text{ if } y_j = i \quad (2)$$

$$(w^i)^T \phi(x_j) + b^i \leq -1 + \xi_j^i, \text{ if } y_j \neq i \quad (3)$$

$$\xi_j^i \geq 0, j=1, \dots, l \quad (4)$$

解式(1)~(4)的最优化问题后,就可以得到 k 个决策函数:

$$(w^i)^T \phi(x) + b^i$$

$$\vdots$$

$$(w^k)^T \phi(x) + b^k$$

对于待测样本 x ,将其输入这 k 个决策函数中,得到 k 个值,取得最大值的函数对应的类别即为该样本所属分类。

通过以上叙述的一对多方法的算法过程,可以看到这种方法的一个明显优点是,只需要训练 k 个两类分类支持向量机,故其所得到的分类函数的个数(k 个)较少,其分类速度相对较快。

这种方法的缺点是,每个分类器的训练都是将全部的样本作为训练样本,这样需要求解 k 个 n 个变量的二次规划问题,因为每个支持向量机的训练速度随着训练样本的数量的增加急剧减慢,因此,这种方法训练时间较长。

3.2 一对一区分法

一对一方法^[9-11](one-against-one)是分别选取两个不同类别构成一个 SVM 子分类器,这样共有 $k(k-1)/2$ 个 SVM 子分类器。在构造类别 i 和类别 j 的 SVM 子分类器时,在样本数据集选取属于类别 i 、类别 j 的样本数据作为训练样本数据,并将属于类别 i 的数据标记为正,将属于类别 j 的数据标记为负。一对一方法需要解决如下的最优化问题:

$$\min_{w^{ij}, b^{ij}, \xi^{ij}} \frac{1}{2} (w^{ij})^T w^{ij} + C \sum_i \xi_i^{ij} \quad (5)$$

$$\text{s.t. } (w^{ij})^T \phi(x_i) + b^{ij} \geq 1 - \xi_i^{ij}, \text{ if } y_i = i \quad (6)$$

$$(w^{ij})^T \phi(x_j) + b^{ij} \leq -1 + \xi_j^{ij}, \text{ if } y_j = j \quad (7)$$

$$\xi_i^{ij} \geq 0 \quad (8)$$

解这一最优化问题后,也即用训练样本进行训练后就可以得到 $k(k-1)/2$ 个 SVM 子分类器。

测试时,将测试数据对 $k(k-1)/2$ 个 SVM 子分类器分别进行测试,并累计各类别的得分,选择得分最高者所对应的类别为测试数据的类别。在这种方法中,需要许多两类问题的分类机。实际上对 k 类问题,就有 $k(k-1)/2$ 个两类分类机。

这个数目常常要比上面的一对多方法得到的分类机的数目大很多。例如 $k=10$,需要得到 45 个分类机,而一种方法只需要 10 个。但是尽管如此,一对一方法的每个分类问题的规模却小了许多,要学习的问题也比较简单。如果 k 太大, $k(k-1)/2$ 就会非常大,这时这个方法就会慢许多。

总的来说,这种方法的优点是训练速度较一对多方法快,缺点是分类器的数目 $k(k-1)/2$ 随类数 k 急剧增加,导致在决策时速度很慢,且存在不可分区域。

3.3 基于二叉树的多级 SVM 分类器

该方法适合故障诊断等多分类决策的二叉树^[12]分类,由于采用了“one against others”的分类器构造策略,故具有直观、简单和重复训练样本少的特点。

给定一个 k 类分类问题,学习样本为 $(x_1, y_1), \dots, (x_l, y_l)$, $x_i \in R^n, j=1, \dots, l, y_j \in \{1, \dots, k\}$,该分类算法是一个四元组:

$$\langle F, P, SVM, CS \rangle$$

$F = \{f_1, \dots, f_i, \dots, f_k\}$, 是二叉树的终止节点集合,由待识别系统的 k 个可能的模式构成。

$P = \{p_1, \dots, p_i, \dots, p_k\}$ 表示各模式类发生的优先级,由各种状态发生的频率的高低排序确定。最可能优先级定为 p_1 ,发生可能性最低的状态定为最后一级 p_k 。 P 决定了分类器中多个支持向量机的安排方式。

$$SVM=\{SVM_{p_1}, \dots, SVM_{p_i}, \dots, SVM_{p_{k-1}}\}$$

是由所设计的 $k-1$ 个支持向量机组成的二叉树的全部非终点的集合, 对一个 k 类分类问题, 需要构造 $k-1$ 个 SVM 机, 其中第 i 个 SVM 决定的模式的优先级为 p_i 。

$SC=\{SC_1, \dots, SC_i, \dots, SC_k\}$ 为属于 k 个模式类别的全部学习样本集合, 其中 $SC_i=\{(x_1, y_1), \dots, (x_i, y_i), \dots, (x_k, y_k)\}$, 表示第 i 类的样本组成, $x_j \in R^n, y \in \{1, \dots, k\}, \sum_{SC} l_i=1$ 构成全部学习样本。

第 i 级支持向量 SVM_{p_i} 的训练样本 S_{p_i} , 按下述原则确定:

$$\begin{cases} S_{p_i} = \text{All Samples} \\ S_{p_i} = \overline{SC_{i-1}} & i=1, \dots, k-1 \\ S_{p_i} = \overline{SC} \ominus \sum_{j=1}^{i-1} SC_j \end{cases} \quad (9)$$

\overline{SC} 表示不属于第 $1, \dots, i$ 类模式的样本全体。可见随着优先级的降低, 训练样本数逐渐减少, 第 i 个 SVM 解决下列问题:

$$\min_{\omega_i, b_i, \xi} \frac{1}{2} (\omega_i)^T \omega_i + C \sum_{j=1}^l \xi_j^i \quad (10)$$

如果, $y_j=i$, 则 $(\omega_i)^T H(x_j) + b_i \geq 1 - \xi_j^i$ 。

如果, $y_j \neq i$, 则 $(\omega_i)^T H(x_j) + b_i \leq 1 - \xi_j^i$ 。

$\xi_j^i \geq 0, j=1, \dots, l, H(x_j)$ 是训练样本 x_j 在高维特征空间中的映射, C 是不可分情况下的惩罚因子, 以降低训练错误数目。

这样可以得到 $k-1$ 个决策函数:

$$\begin{aligned} & (\omega_1)^T H(x) + b_1 \\ & \vdots \end{aligned}$$

$$(\omega_{k-1})^T H(x) + b_{k-1}$$

R_{others} 表示为重复训练样本数和重复次数的乘积, 其公式为

$$R_{others} = \sum_{i=2}^{k-1} l_i \cdot (i-1) + l_k \cdot (k-2), k \geq 3 \quad (11)$$

对于一个由 l 个训练样本组成的 k 个模式的分类问题, “one against all” 和 “one against one” 策略来说, “one against others” 大大降低了样本的重复训练量。

“one against others” 策略属于多级 SVM 分类器构造方法中的第二种。对于多类多峰分布问题, 通过优先级的确定可以迅速建立一个简单的二叉树分类器, 它不是企图用一种算法、一个决策规则去把多个类别一次分开, 而是采用分级的形式, 把一个复杂的多分类问题转化若干个简单的二分类问题, 使分类问题得到逐步的解决。这种策略不存在 “one against all” 和 “one against one” 策略的无法分类阴影问题, 因此分类精度较高。

在机械故障诊断模型中, 首先设计出第一个 SVM, 称为故障检测 SVM, 完成系统正常样本和故障样本的分类, 即检测出系统是否发生故障, 如果诊断出系统正常则无须进入下一级 SVM 检测。第二个 SVM 只在故障样本中进行, 完成第一类故障和其它故障类型的分类, 往下依次类推, 直到第 $k-1$ 个 SVM 输出分类结果为止。

4 实验结果和分析

采用 WPA-SVM 或 WT-SVM 对故障信号进行诊断的策略

是: 首先, 对某振动设备进行检测, 获得其状态样本数据。样本数据的状态分为三种: 正常状态样本数据 ($-2 < |A| < 2$)、第一类振动故障 ($-1.5 < |A| < 2.5$) 和第二类振动故障的状态样本数据 ($-2.5 < |A| < 2.5$), 分别对应类别 1、2 和 3。提取振动信号在不同频带的特征参量, 建立学习样本, 然后 SVM 机按照不同的策略进行故障诊断。

在此基础上, 建立的两个支持向量机分类器分别为: SVM_1 对应正常状态; SVM_2 对应第一类振动故障。

(1) 特征提取方法的比较

选择 210 个样本, 使用小波 (WT) 和小波包 (WPA) 进行特征提取 (“db5” 小波和 “shannon” 熵), 然后采用 SVM 的 “one against one” 策略对其进行识别。图 3 表示第一类振动故障的小波分解结果, 图 4 表示第一类振动故障的小波包分解结果, 表 1 为部分样本的小波包特征参量。当采用小波作为特征提取方法时, SVM 的识别率为 84.8%; 当采用小波包作为特征提取

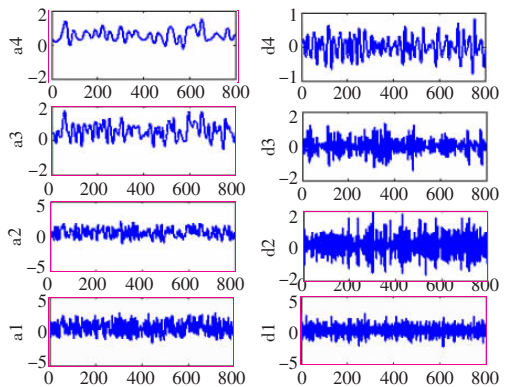


图 3 第一类振动故障的小波分解结果 (“db5” 小波和 “shannon” 熵)

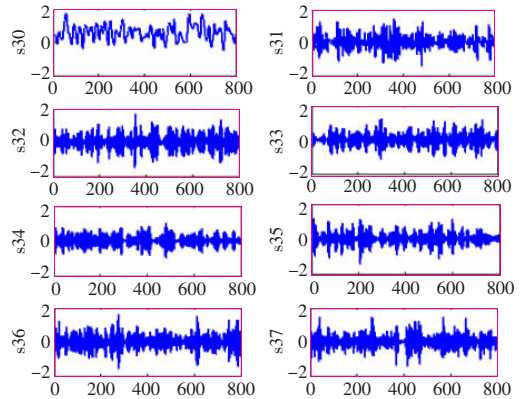


图 4 第一类振动故障的小波包分解结果 (“db5” 小波和 “shannon” 熵)

表 1 部分样本的小波包特征参量

| 序号 | 频带 1 | 频带 2 | 频带 3 | 频带 4 | 频带 5 | 频带 6 | 频带 7 | 频带 8 | 类别 |
|----|----------|----------|----------|----------|----------|----------|----------|----------|----|
| 1 | 12.561 3 | 11.201 5 | 10.496 7 | 11.716 9 | 10.775 1 | 10.522 6 | 11.497 4 | 10.237 2 | 1 |
| 2 | 11.963 3 | 11.926 0 | 11.921 1 | 10.782 5 | 12.201 9 | 11.896 5 | 9.846 5 | 11.742 8 | 1 |
| 3 | 12.101 8 | 12.601 9 | 11.715 5 | 12.102 0 | 12.435 1 | 11.377 2 | 10.886 1 | 11.365 1 | 1 |
| 4 | 12.561 3 | 11.201 5 | 10.496 7 | 11.716 9 | 10.775 1 | 10.522 6 | 11.497 4 | 10.237 2 | 1 |
| 5 | 17.356 3 | 11.783 1 | 12.622 9 | 12.596 8 | 11.304 7 | 12.067 4 | 10.650 8 | 12.124 0 | 2 |
| 6 | 17.625 9 | 11.456 4 | 11.725 8 | 11.391 8 | 11.529 2 | 11.025 9 | 12.446 1 | 11.593 3 | 2 |
| 7 | 18.650 7 | 10.527 3 | 12.068 1 | 11.378 6 | 12.555 6 | 11.614 3 | 9.990 7 | 11.469 1 | 2 |
| 8 | 18.422 5 | 12.022 6 | 12.473 1 | 11.172 4 | 10.836 4 | 12.666 5 | 11.651 1 | 12.945 0 | 2 |
| 9 | 19.424 6 | 12.858 3 | 11.347 3 | 10.060 6 | 11.557 6 | 12.508 3 | 13.836 1 | 11.456 4 | 3 |
| 10 | 21.120 1 | 11.256 1 | 12.647 6 | 10.539 8 | 11.339 9 | 12.126 7 | 10.969 1 | 11.602 2 | 3 |
| 11 | 20.109 5 | 10.950 7 | 11.281 5 | 10.694 9 | 12.148 4 | 11.349 8 | 12.238 1 | 11.243 5 | 3 |
| 12 | 19.192 1 | 11.802 9 | 11.302 6 | 12.539 4 | 12.218 6 | 10.305 4 | 12.565 6 | 9.949 3 | 3 |

方法时,SVM 的识别率为 88.0%(如图 5 所示)。

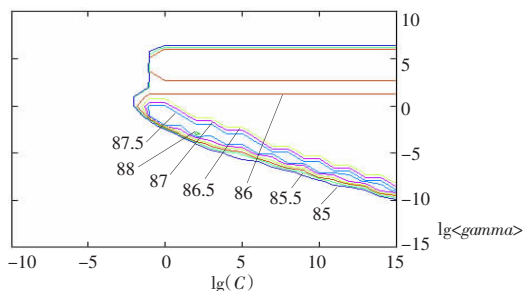


图 5 WPA-SVM 的识别结果(最佳参数选择: $C=4, \gamma=0.125$)

可见,作为特征提取方法,小波包比小波要好。

(2)SVM 三种策略的比较

图 6 反映了三种策略的识别性能。在本次的测试实验中,“one against all”的识别准确率可以达到 87.1%，“one against one”识别准确率可以达到 88.6%,而“one against others”的识别准确率可以达到 92.1%。可见“one against one”的识别效果略优于“one against all”,“one against others”的识别效果比“one against one”和“one against all”都要好,特别是在训练样本数目较大(大于 120)的时候,其优势更加明显。

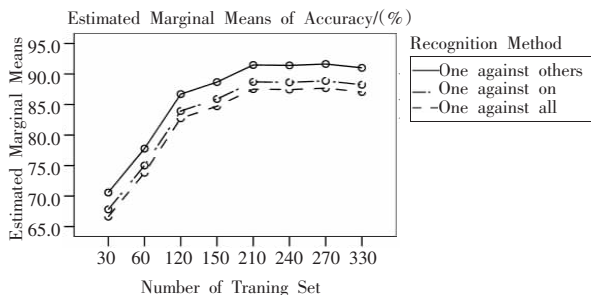


图 6 三种策略的识别性能比较

5 结论

针对目前机械故障诊断中难以进行特征提取和常规 SVM 算法在解决故障诊断这种典型的多类分类问题时存在困难等问题,提出了一种小波包分析与基于 SVM 的二叉树多级分类器结合起来的机械故障诊断方法。采用小波包分析对机械信号提取频域能量特征向量,通过建立多 SVM 分类器完成故障诊

断,获得了令人满意的效果。实验表明,在特征提取方法中,小波包(WPA)比小波(WT)的效果更好;而在 SVM 策略中,采用“one against others”策略不存在“one against all”和“one against one”策略的无法分类阴影问题,因此分类精度比它们更高,特别是在训练样本数目较大的时候,其优势更加明显。

参考文献:

- [1] Weston J,Watkins C.Multi-class Support Vector Machines[D].Royal Holloway College,1998.
- [2] KreBel U.Pairwise classification and support vector machines[M].Cambridge MA:MIT Press,1999:255-256.
- [3] Motton L,Cortes C.Comparison of classifier methods:a case study in handwriting digit recognition[C]//ICPR94,1994:77-87.
- [4] Hsu C W,Lin C J.A comparison of methods for multi-class support vector machine,2001.
- [5] 韩家新,何家灿.SVM 分类器及其在文本分类中的应用研究[J].计算机应用研究,2004(1):23-24.
- [6] 何学文,赵海鸣.支持向量机及其在机械故障诊断中的应用[J].中南大学学报:自然科学版,2005,36(1):97-101.
- [7] Hubert D,Stephan W.Detection of cochlear hearing loss applying wavelet packets and support vector machines[C]//Conference Record of the Thirty-Eighth Asilomar Conference on Signals,Systems and Computers,2004,2:1575-1579.
- [8] Jaime G,Ignacio M,Juan S.Wavelet time shift properties integration with support vector machines[C]//Lecture Notes in Artificial Intelligence:Modeling Decisions for Artificial Intelligence,2004,3131:49-59.
- [9] Hsu Chih-wei,Lin Chih-jen.A comparison of methods for multi-class support vector machines[J].IEEE Transactions on Neural Networks,2002,13(2):415-425.
- [10] Blanz V,Schuolkopf B,Bulthoii H.Comparison of view-based object recognition algorithms using realistic 3D models[C]//Artificial Neural Networks-ICANN,1996:375-381.
- [11] Krelel U.Pairwise classification and support vector machines[C]//Schuolkopf B,Burges C J C,Smola A J.Advances in Kernel Methods:Support Vector Learning,Cambridge,MA,1999:255-268.
- [12] 马笑潇,黄席樾.基于 SVM 的二叉树多分类算法及其在故障诊断中的应用[J].控制与决策,2003,18(3):272-276.

(上接 241 页)

- [22] 赵杰,姜健.基于强化学习的未知环境多机器人协作搜索[J].计算机工程与应用,2007,43(10):19-21.
- [23] 张文志,吕恬生.强化学习理论在机器人应用中的几个关键问题探讨[J].计算机工程与应用,2007,43(4):69-72.
- [24] Prandini M,Hespanha J,Pappas G J.Greedy control for hybrid pursuit games[C]//ECC Conf,Porto,Portugal,Sept 2001.
- [25] Bopardikar S D,Bullo F,Hespanha J P.A cooperative Homicidal Chauffeur game[C]//IEEE Conf on Decision and Control,New Orleans,LA,December 2007:1-9.
- [26] Choi Han-Lim,Tahk Min-Jea,Bang Hyo-Choong.Neural network guidance based on pursuit-evasion games with enhanced performance[C]//15th IFAC World Congress 2003:101-106.

- [27] Littman M L,Moore A W.Reinforcement learning:a survey [J].JAIR,1996(4):237-285.
- [28] Mitchell T M.Machine learning[M].USA:McGraw-Hill Companies Inc,1997:367-387.
- [29] Pack D J,Mullins B E.Toward finding a universal search algorithm for swarm robots[C]//IEEE/RSJ International Conference on Intelligent Robots and Systems(IROS2003),2003,2:1945-1950.
- [30] Sheng Weihua,Yang Qingyan,Tan Jindong,et al.Distributed multi-robot coordination in area exploration[J].Robotics and Autonomous Systems,2006,54:945-955.
- [31] 高阳,周如益,王皓,等.平均奖赏强化学习算法研究[J].计算机学报,2007,30(8):1372-1378.