

文章编号: 1001- 2486(2005) 02- 0092- 04

# 机械故障特征与分类器的联合优化\*

王新峰, 邱 静, 刘冠军

(国防科技大学 机电工程与自动化学院, 湖南 长沙 410073)

**摘要:** 在机械故障诊断中, 特征选择和分类器的参数优化都可以提高诊断精度。利用特征和分类器参数的依赖关系, 提出了特征选择和 SVM 参数的联合优化方法来提高诊断性能。联合优化方法采用支持向量机(SVM)作为故障分类器, SVM 半径-间距上界(RM 界)为目标计算诊断精度, 并应用遗传算法求解此优化问题。齿轮故障诊断试验结果表明, 联合优化的诊断精度要优于单独优化特征和 SVM 参数, 而且优化速度更快。因此在故障诊断中, 利用特征和分类器参数联合优化能够快速取得较好的诊断精度。

**关键词:** 特征选择; 支持向量机; 半径-间距上界; 遗传算法

中图分类号: TN06; TP391 文献标识码: A

## Joint Optimization of the Fault Feature and Classifier

WANG Xin-feng, QIU Jing, LIU Guan-jun

(College of Mechatronics Engineering and Automation, National Univ. of Defense Technology, Changsha 410073, China)

**Abstract:** Feature selection and parameters optimization of the fault classifier can enhance the fault diagnosis accuracy. Using the interdependent relationship between the feature selection and classifier parameter, a method of joint optimization of feature selection and classifier parameters is proposed to improve the diagnosis accuracy. By using the method we adopt the support vector machine (SVM) as a fault classifier, take into account of the radius-margin bounds for the accuracy evaluation of SVM classifier, and applies genetic algorithm (GA) to solve the joint optimization problem. In the gear fault diagnosis experiment, the joint optimization method guarantees better diagnosis accuracy and the optimization process has a higher rate than the single optimization of features or SVM parameters. So the joint optimization of fault features and classifier can fast achieve the better diagnosis accuracy in fault diagnosis.

**Key words:** feature selection; SVM; radius-margin bound; genetic algorithm

在机械故障诊断中, 由信号处理得到的故障特征集中常含有两类冗余特征, (1) 与分类目标无关的特征; (2) 与其他特征量有较高相关性的特征。这些特征会影响最后的诊断精度。所以在故障识别之前, 必须对故障特征集进行特征选择<sup>[1-3]</sup>, 找到优化特征集。另一方面, 诊断精度还与故障分类器有关, 分类器对诊断精度也有影响。在故障诊断过程中, 除了特征选择外, 还需要对故障分类器进行优化, 即必须实现特征选择和故障分类器的联合优化。这是因为特征选择和故障分类器的优化过程是互相依赖的, 优化其一需要有另一个才能测试。然而在一般应用中它们却常分开进行, 这样做存在一定的缺陷, 即无法确定先优化特征还是故障分类器。另外优化其中一个时需要随机地确定另一个, 即使轮流进行也不能保证两者同时达到最好。因此, 为了使诊断有更好的精度, 特征选择和故障分类器的优化应该联合进行。

本文采用支持向量机(support vector machine, SVM)作为故障分类器对故障进行诊断。SVM 是 20 世纪 90 年代出现的有效分类器, 已经应用到许多领域, 如模式识别、数据挖掘、特征选择等。支持向量机基于结构风险最小化原则, 它可以克服维数灾难问题, 在有限样本的基础上取得良好的分类效果, 目前得到越来越多的应用<sup>[3]</sup>。在使用 SVM 的过程中, SVM 的性能与 SVM 的核函数类型、核函数的参数有很大关系。在建立故障分类器时, 必须对这些参数进行寻优, 找到最合适的 SVM 故障分类器。

\* 收稿日期: 2004- 10- 22

基金项目: 国家部委基金资助项目(41319040202)

作者简介: 王新峰(1977-), 男, 博士生。

## 1 支持向量机简述

支持向量机是通过将低维的输入空间数据  $x \in R^n$  经非线性映射函数映射到高维属性空间  $H$ , 映射记为  $x \rightarrow \Phi(x) \in H$ , 在此高维空间找到最优分类超平面。这种映射是通过核函数实现的, 它定义了  $H$  空间的内积。在找到最优分类平面的过程中, 在高维空间实际上只需要进行内积运算, 这种内积运算可以用原空间中的函数实现, 甚至没有必要知道变换的具体形式。因此, 在最优分类面中采用适当的内积核函数  $K(x, y)$  就可以实现从低维空间向高维空间的映射, 从而实现某一非线性变换后的线性分类, 而计算的复杂度却没有增加<sup>[4]</sup>。

设给定的训练集为  $\{(x_i, y_i)\}_{1 \leq i \leq l}$ , 其中样本  $x_i \in R^n$ , 它所属的类别  $y_i \in \{-1, 1\}$ , 采用的核函数为  $K(x, y)$ , 支持向量机分类函数为:

$$f(x) = \sum_i \alpha_i^* y_i K(x_i, x) + b \quad (1)$$

系数  $\alpha_i^*$  可以通过求下式的最大值得到:

$$Q^2(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (2)$$

约束式为  $\sum_{i=1}^l \alpha_i y_i = 0, \alpha_i \geq 0, i = 1, 2, \dots, l$ 。

系数  $\alpha_i^*$  定义了在高维空间的最大分离间隙超平面。对于不可分的情况, 所谓的广义最优分类面问题, 即在追求最大化分类间隔的同时最小化错分样本的数目。通过改变核的形式  $K \leftarrow K + (1/C)I$  来作为错分样本的惩罚。其中  $C$  为自定义的惩罚系数,  $I$  为辨识矩阵, 代表训练样本关于广义分离超平面的偏差。

## 2 特征选择和 SVM 参数的联合优化

### 2.1 数学模型

特征选择和 SVM 参数联合优化的数学描述如下:

$$\max_{S, P} T(S, P) \quad (3)$$

其中,  $T$  为 SVM 诊断精度,  $S$  为特征参数,  $P$  为 SVM 参数。式(3)中的  $(S, P)$  就是混合向量, 它同时描述特征和 SVM 参数。该问题可以用遗传算法求解。其中, 个体为混合向量  $(S, P)$ , 其适应度为 SVM 性能  $T$ 。在计算个体的适应度时, 先根据混合向量  $(S, P)$  重建特征数据集和 SVM 参数, 再通过 SVM 的一次训练计算出精度  $T$ 。这样,  $S$  和  $P$  的优化不存在先后关系而是互相关联。

特征选择和 SVM 参数的联合优化是组合优化问题。遗传算法具有强大的隐并行模式空间搜索能力和随机、自适应、稳健的解空间搜索特点, 是解决优化问题的有力工具<sup>[1]</sup>。它曾经被用来特征选择和优化 SVM 参数<sup>[5,6]</sup>, 其种群中的个体分别为特征向量和 SVM 参数向量。两者的适应度都是分类器性能。因此, 考虑合并特征选择向量和 SVM 参数向量作为种群中的个体, 这样就可以用遗传算法进行联合优化。遗传算法的实施过程基本类似, 不同的是针对问题的具体编码方式和适合度函数的实现过程。

### 2.2 SVM 性能计算

计算 SVM 诊断精度作为遗传算法的适应度。一般计算 SVM 诊断精度时, 应该将样本数据分为训练和测试数据两个集合, 对多次训练和计算的结果取均值, 如 Leave-One-Out 和 Cross-Validation 方法。对于 SVM, 这些方法的计算量很大。本文采用 SVM 的一种错误率的上界半径-间隙界(Radius-Margin, 简称 RM 界)作为 SVM 诊断精度的评价, 因为计算 RM 界只需要训练一次 SVM, 它的计算效率要高于 Leave-One-Out 和 Cross-Validation 方法估计错误率的方式。

定理: 设  $\rho$  为最大的分类间隙, 采用内积核函数  $K(x, y)$  实现训练特征样本从低维空间向高维空间的映射, 所有映射都包含在一个最小半径为  $R$  的超球内。假设  $\alpha^*$  为式(2)所求得的  $\alpha$  值, 如果高维

映射在  $R$  中以间隙  $\rho$  为可分的, 则期望错误率  $EP_{err}$  存在上界  $EP_{RM}$  为:

$$EP_{err} \leq EP_{RM} = \frac{1}{l} E \left[ \frac{R^2}{\sigma^2} \right] = \frac{1}{l} E \{ R^2 Q^2(\alpha^*) \} \quad (4)$$

$R^2$  可通过下式计算得到<sup>[7]</sup>:

$$R^2 = \max_{\beta} \sum_{i=1}^l \beta_i K(x_i, x_j) - \sum_{i,j=1}^l \beta_i \beta_j K(x_i, x_j) \quad (5)$$

约束式为  $\sum_{i=1}^l \beta_i y_i = 0, \beta_i \geq 0, i = 1, 2, \dots, l$ 。

在多故障情况下, 可以计算每对故障类别间的错误率上界, 求其和作为总的上界。故 SVM 诊断精度  $T$  可以用  $EP_{RM}$  来表示, 即  $T = 1 - EP_{RM}$ 。

### 2.3 特征选择和 SVM 参数遗传编码

给定特征集  $F = \{f_1, f_2, \dots, f_N\}$ , 其中  $N$  是特征集大小, 特征选择可以用一个二进制向量来表示,  $S = \{s_1, s_2, \dots, s_N\}$ ,  $s_i \in \{0, 1\}, i = 1, \dots, N$ 。如果  $s$  的第  $i$  位为 1, 则此特征被选中; 如果为 0, 则此特征未被选中。SVM 在训练之前必须确定一些参数, 其中最重要的就是核函数。常用的核函数有:

(1) 多项式(Poly)内积函数

$$K_{poly}(x, y) = [(x \cdot y) + 1]^q \quad (6)$$

(2) 径向基(RBF)内积函数

$$K_{rf}(x, y) = \exp \left[ - \frac{|x - y|^2}{\sigma^2} \right] \quad (7)$$

其中,  $x, y$  为输入空间的两个向量,  $p$  和  $\sigma$  分别为多项式和 RBF 核函数的参数。Smits 指出每一种核函数对提高 SVM 性能的能力有限, 并引用一个参数构造了混合模型<sup>[7]</sup>:

$$K = \lambda K_{rf} + (1 - \lambda) K_{poly}, \quad 0 \leq \lambda \leq 1 \quad (8)$$

Smits 未指出如何选择  $\lambda$  以及核函数的参数  $p$  和  $\sigma$ 。结合核函数, 式(3)中 SVM 参数向量集  $P$  为:

$$P = \{ \lambda, \sigma, p \}, \quad 0 \leq \lambda \leq 1 \quad (9)$$

考虑算法的快速性和实用性, 对于 SVM 参数不需要找到准确值, 只考虑它的取值范围即可。对于多项式核函数, 当指数  $p$  的值太大时, 会使核函数的值很大, 计算量增大, 优化效率降低, 且对提高分类器性能的作用也不大, 所以  $p$  的取值不宜太大。基于上述情况, 把  $p$  的优化范围选定为 1~ 8, 用染色体  $(b_1, b_2, b_3)$ ,  $b_i \in \{0, 1\}$  表示, 即  $p = (b_1, b_2, b_3) + 1$  来实现。对于径向基核函数, 宽度系数  $\sigma$  的取值不能太大或太小, 一般取值为 0.1 ~ 5, 选择为 0.1, 0.5, 1, 1.5, 2, 3, 4, 5;  $\lambda$  取值为 0, 0.15, 0.3, 0.45, 0.6, 0.75, 0.9, 1, 同样可以用三字节  $(b_1, b_2, b_3)$  表示各自的值。

## 3 实例验证

### 3.1 实验数据

实验对象为某直升机中减速器中的齿轮, 数据在减速器试车台测得。齿轮故障类型包括两种最为常见的局部性故障——点蚀和裂纹。故障部位为输入齿轮, 齿数为 29, 额定啮合频率为 1 689Hz, 选定采样频率为 10kHz。根据直升机的工作状况, 对 5 种工作(负载)状况下减速器进行了试验, 每种工况 10 次重复采样, 总共有三种状况, 即正常、点蚀和裂纹, 得到 150 个样本, 每种情况 50 个样本。

由于减速器工作时, 引起其振动的因素很多, 通过安装在减速器上的振动传感器所拾取的振动信号是非常复杂的宽带信号。所以本文采用小波包分解的方法进行信号预处理。对信号经过小波包 5 层分解后, 可以得到 32 个特征频带的小波包系数。各频带带宽为 156.25Hz, 由低到高分布, 如频带 1 的范围为 0Hz~ 156.25Hz, 频带 2 的范围为 156.25Hz~ 312.5Hz, 依次类推。然后利用小波包能量谱的方法得到原始特征。在小波包能量谱中, 选取各个频带内信号的平方和作为能量的标志。对于频带  $\omega_i$ , 小波包变换结果用序列  $\{ \omega_i(k), k = 1, 2, \dots, M \}$  表示, 其中  $M$  为该频带的样本长度。由于小波包分解后的各频带宽度相同, 因此利用小波包分解的结果作为能量谱的输入, 则各频带的能量为  $G_i = \sum_{k=1}^M | \omega_i(k) |^2$ , 令

$G = \sum_{i=1}^{32} G_i$ , 可以计算出各频带的相对能量比值, 即  $G_i = G_i / G$ , 以此作为故障样本特征, 可得到 32 个特征。

### 3.2 优化方法比较

对故障数据进行特征和 SVM 参数的联合优化的同时, 为进行比较, 分别对特征和 SVM 参数进行单独优化。遗传算法的编码如前所示, 遗传算法的初始设置相同, 种群数为 30, 杂交率为 0.9, 变异概率为 0.09, 迭代次数为 30。优化结果如图 1 所示。

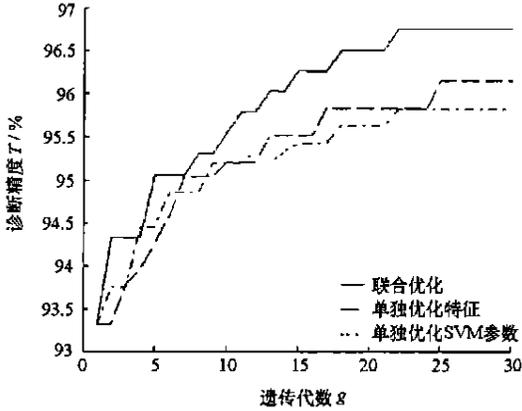


图 1 联合优化和单独优化比较

Fig. 1 Comparing diagnosis accuracy of joint optimization with single optimization

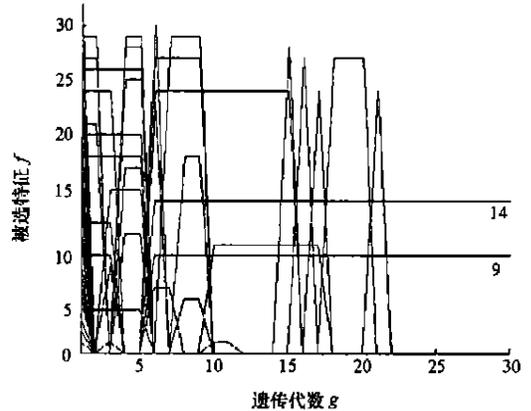


图 2 特征随寻优次数的变化趋势

Fig. 2 Presence of features selected with genetic generations

可以看出, 无论是特征选择和 SVM 的单独优化, 还是联合优化, 都能提高诊断精度  $T$ , 其中联合优化的精度提高最多, 最后的诊断性能为 96.75%, 比单独优化提高了。而单独优化特征要比优化 SVM 参数精度要高。可见, 联合优化性能最好, 特征选择的可优化性要比 SVM 的可优化性稍微高一点。从改善速度的角度看, 联合优化比单独优化要好得多。在第 12 代种群的时候, 联合优化方法已达到 SVM 单独优化的最终结果, 而在 15 代种群的时候, 联合优化方法已达到特征选择单独优化的最终结果。

图 2 为联合优化特征随遗传代数的发展情况。可以看出经过 22 次迭代, 特征子集找到了优化值。最优特征为第 9 和 14 个频带能量特征。

## 4 结论

提出了特征选择和故障分类器 SVM 的联合优化方法, 并与单独优化方法比较。减速器齿轮故障特征数据的优化实验结果表明, 与分别优化特征和 SVM 参数相比, 联合优化可以达到最佳诊断精度, 而且联合优化的速度更快, 能够以更短的时间搜索出最优特征和故障分类器, 提高诊断精度。

## 参考文献:

- [1] Kohavi R, John G. Wrappers for Feature Subset Selection [J]. Artificial Intelligence, 1997, (12): 273-324.
- [2] Jain A, Zongker D. Feature Selection: Evaluation, Application, and Small Sample Performance [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(2): 153-158.
- [3] 谢涛, 张育林. 基于遗传算法与最大最小原理的故障模式特征选择[J]. 国防科技大学学报, 1998, (2): 17-21.
- [4] Vapnik V. The Nature of Statistical Learning Theory [M]. New York: Springer-Verlag, 1995.
- [5] Vapnik V, Chapelle O. Bounds on Error Expectation for SVM [A]. Advances in Large Margin Classifiers[M], MIT Press, Cambridge, MA, 2000: 261-280.
- [6] Siedlecki W, Sklarsky J. A Note on Genetic Algorithms for Large Scale Feature Selection [J]. Pattern Recognition Letters, 1989, 10: 335-347.
- [7] Fröhlich H, Chapelle O, Schölkopf B. Feature Selection for Support Vector Machines by Means of Genetic Algorithms [A]. Proceedings of the 15th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'03) [C]. IEEE, 2003.
- [8] Smits G F, Jordaen E M. Improved SVM Regression Using Mixtures of Kernels [A]. Proc of the 2002 Int Joint Conf on Neural Networks [C], Honolulu, HI, USA: IEEE, 2002, 3: 2785-2790.