

基于支持向量机集成算法的煤矿顶板状态检测

付家才,张铁山,任 众

(黑龙江科技学院,黑龙江 哈尔滨 150027)

摘要: 针对小样本、多类煤矿顶板的状态检测问题,提出了基于支持向量机集成的算法。采用深度优先搜索对支持向量机集成参数进行优化,并结合煤矿顶板敲击声信号的人耳听觉功率谱特征,对煤矿顶板中浮石、剥层等现象进行分类识别。实验表明,该算法能够对多类顶板状态进行有效分类,且识别率较高,能够作为安全开采的保障手段。

关键词: 支持向量机集成;人耳听觉功率谱;顶板状态检测;深度优先搜索

中图分类号: TP277

文献标识码: A

文章编号: 1674-7720(2012)17-0015-03

Detection of coal mine roof based on support vector machine ensemble

Fu Jiakai, Zhang Tieshan, Ren Zhong

(Heilongjiang Institute of Science & Technology, Harbin 150027, China)

Abstract: In this paper, a method based on support vector machine ensemble about the detection of Many types of coal mine roof is proposed. In optimization process, the depth-first search is used. The classification and recognition of security risk is according to power spectral characteristics of human ear of mine roof percussion signal. The experiments show that the algorithm can classify the security risk of coal mine roof effectively.

Key words: SVM; power spectral characteristics of human ear; detection of coal mine roof; depth-first search

我国是一个煤炭大国,煤炭资源丰富,且煤炭消费量逐年增加。为保证国民经济的持续健康发展,需要有计划地大规模开采煤炭资源。然而,我国在煤炭开采中的安全事故经常发生。其中,冒顶事故是最常见的安全事故之一,给国家和人民带来了巨大的损失。而要避免和减少冒顶事故的发生,关键就是要对顶板的安全性能进行及时、有效地检测,以便工程技术人员及时排除安全隐患。

敲帮问顶(wall tapping and roof sounding)是依靠人耳极其灵敏的听觉系统和长期的实践经验来判断顶板的安全稳定性的,是目前被广泛沿用的对顶板进行安全检测的重要途径。但这种方法对工人自身技能要求较高,且对工人的人身安全有一定的安全隐患。因此,研究新的检测技术就显得非常迫切。其中,基于模式识别的检测是目前研究的重要方向之一。在应用现代模式识别技术对顶板进行状态监测和故障诊断中,功率特征提取和分类器模型设计是两个至关重要的因素。本文采用基于听觉模型和支持向量机集成算法。基于听觉模型的声

信号特征提取、分类、识别的研究,近年来得到了国内外学者的高度重视,特别在语音信号处理方面。但是针对煤矿顶板敲击声音信号方面重视较少。支持向量机(SVM(Support Vector Machines)是Vapnik等人根据统计学习理论中结构风险最小化原则提出的。而支持向量机集成SVM(Support Vector Machine Ensemble)可以提高支持向量机的分类性能^[1-2]。

本文正是基于以上原因,针对小样本、多类煤矿顶板,提出了基于支持向量机集成算法对煤矿顶板安全性能进行检测的方法。

1 人耳听觉功率谱特征提取

人们通过对生理声学、心理声学和信号处理的研究和分析^[3-4],建立了人耳听觉模型,并且利用此模型来对声音信号进行听觉谱特征提取。该听觉谱特征^[5]提取过程如图1所示。

通过对模仿人的听觉感知机理进行分析,听觉谱特征提取算法主要有三个方面的处理:(1)临界频带段分析处理;(2)等响度级预处理;(3)等响度转换。利用临

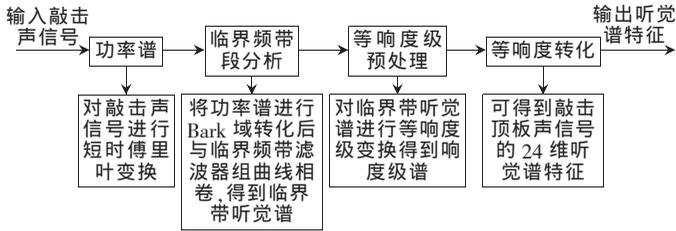


图1 听觉谱特征提取原理

界频段分析,计算模型考虑了耳蜗的分频特性,由此可以反映人耳的掩蔽效应。

2 支持向量机集成算法

顶板状态模型为利用已知类别的顶板状态样本建立支持向量机分类器模型。首先给定 n 个独立样本,每个样本 x_i 有 d 个特征,样本 x 分别属于 m 类顶板状态。

$$X = \{(x_i, y_i)\} | \{x_i \in R^d, y_i \in \{1, 2, \dots, m\}, i=1, 2, \dots, n\}$$

通过对给定的样本集进行训练得到一个分类器,利用此分类器对测试样本集进行分类的算法称为机器学习算法。

为了提高支持向量机的分类性能和分类精度,可以把多个分类器集成起来。针对小样本、多类顶板状态诊断问题,本文提出的支持向量机集成(SVME)算法,结合了敲击声音信号的人耳听觉谱特征,对顶板状态进行诊断。

SVME 算法流程如图 2 所示。



图2 支持向量机集成(SVME)算法流程

如图 2 所示,对于 SVME 算法而言,首先应给定训练样本集 X 和训练样本的初始加权系数 $w_1(i) = \frac{1}{n}, i=1, 2, \dots, n$; 然后,进行一个循环函数以训练 T 个个体 SVM。

在循环体中,首先从给出的训练样本集 X 中按照一定的概率抽取样本并获得新的训练样本 X_i ; 然后,依据新获得的训练样本 X_i 来优化选择个体 SVM 核函数参数,并训练 SVM 以获得分类超平面 $f_i(X_i)$; 再次,计算出该分类超平面的加权分类错误率;最后,做出一个判断:即判断计算出的加权分类错误率是否 ≥ 0.5 或 $= 1$ 。如果条件成立,则结束该循环,并运用加权多数投票法集成

个体 SVM; 反之,则改变加权系数,进行下一轮循环,直至条件成立,退出循环。

要想获得一个高性能和高精度的支持向量机集成模型,就必需对优化不敏感损失函数 ϵ 、惩罚系数 c 和核参数 γ 等参数进行优化处理,而这些参数之间相互影响。实际上,选择过程是一个优化搜索的过程,由于深度优先算法具有全局搜索能力强、并行性和启发式等优点,所以本文采用深度优先算法对 SVM 参数进行寻优。深度优先算法^[6]是一种基于图论的优化算法,其基本思想:为了求得问题的解,先选择某一种可能情况向前(子结点)探索,在探索过程中,一旦发现原来的选择不符合要求,则回溯至父亲结点重新选择另一结点,继续向前探索。如此反复进行,直至求得最优解。深度优先搜索的实现方式可以采用递归或者栈来实现。

采用深度优先搜索对 SVM 参数 c 、 γ 和 ϵ 进行优化的具体步骤如下:

(1) 对 SVM 中的参数 c 、 γ 和 ϵ 进行初始范围的确定,按照步长对参数 c 、 γ 和 ϵ 进行等量划分,以此获得离散数组。

(2) 首先利用 LIBSVM 软件采用交叉验证方法对参数 c 、 γ 和 ϵ 进行建立模型和预测; 然后对每组均方根误差进行比较; 最后选择最小的均方根误差所对应的参数为本次的 SVM 最佳参数组合。

(3) 将本次最优参数和上一次最优参数的均方根误差进行比较,如果均方根误差小于上一次均方根误差,则跳转到步骤(4),否则跳转到步骤(5)。

(4) 根据上一次的最优参数,并且采用启发式方式在该参数附近进行参数范围修改,从而加快参数搜索速度; 然后跳转到步骤(1),重新进行递归运算。

(5) 参数优化结束后,退出运算,取此时的参数 c 、 γ 和 ϵ 为 SVM 的最优参数。

深度优先搜索参数优化流程图如图 3 所示。

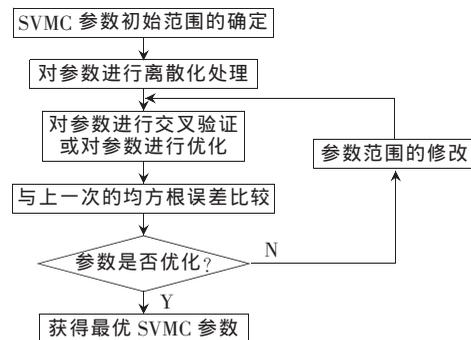


图3 深度优先搜索参数优化流程

3 实验及结果

3.1 实验数据集

本文实验数据是从我国东北某大型煤矿井下顶板上采集的,共有 4 类目标:一类是安全顶板,另 3 类是 3 种危险顶板(浮石、剥层和断裂)。信号的采样频率是

20 000 Hz, 每个样本长度为 4 096 个点。安全顶板、浮石顶板、剥层顶板和断裂顶板的样本数目分别为 2 000、200、500、150。由于受到实验数据采样频率(20 000 Hz)的限制, 本文只在 20~10 000 Hz 频率范围内(即前 22 个临界频带内)提取了 4 类顶板的敲击声信号的听觉谱特征。

3.2 实验

在分类实验中, 训练样本数目和测试样本数目的比值是 1:4, 提取样本的听觉谱特征后, 分别用 SVM 算法和单个 SVMC 对 4 类目标进行顶板状态诊断实验。为了避免样本选取的随机性对实验结果带来的误差, 实验分别进行了 20 次, 最终结果为 20 次实验结果的平均值。检测结果如表 1 所示。

从表 1 可以看出, 支持向量机集成算法可以有效识别安全顶板和危险顶板, 并进行状态诊断, 而且其识别 4 类目标的正确识别率都比单个 SVMC 要高。特别是本文所提出的算法对危险顶板的正确识别率比单个 SVMC 提高了 2.57%~3.22%。实验结果证明, 本文提出的支持向量机集成(SVME)算法是有效的, 可以应用于煤矿顶板状态识别和故障诊断中。

本文提出了支持向量机集成(SVME)算法, 采用了深度优先搜索对支持向量机集成参数进行优化, 并结合敲击声信号的人耳听觉谱特征用于顶板状态的分类识别。实验证明, 本文所提出的 SVM 算法识别率较高, 可以满足现场检测的要求, 为煤矿顶板状态诊断提供了新的方法。

参考文献

表 1 四类顶板检测结果

	样本数目/个		正确识别率/%	
	训练	测试	SVME	SVM
安全顶板	300	1 200	99.68	98.12
浮石顶板	30	120	94.15	91.02
剥层顶板	80	320	94.58	91.36
断裂顶板	25	100	97.58	95.01

- [1] VAPNIK V. The nature of statistical learning theory [M]. New York: Springer-Verlag, 1999.
- [2] 杨宏晖, 孙进才. 基于支持向量机集成的水下目标自动识别系统[J]. 测控技术, 2006, 25(12).
- [3] ZWICKER E, FASTL H. Psychoacoustics facts and models [M]. New York: Springer-Verlag, 1999.
- [4] HERMANSK H. Perceptual linear predictive (PLP) analysis of speech [J]. Journal of the Acoustical Society of America, 1990, 87(4): 1738-1751.
- [5] 杨宏晖, 侯宏, 曾向阳. 基于声信号人耳听觉谱特征的风机故障诊断[J]. 仪器仪表学报, 2009, 30(1).
- [6] 李立红, 许元飞. 深度优先搜索的支持向量机参数优化算法[J]. 计算机仿真, 2011, 28(7): 216-219.

(收稿日期: 2012-05-06)

作者简介:

付家才, 男, 1954 年生, 教授, 主要研究方向: 现代矿山电气控制。