

doi:10.11918/j.issn.0367-6234.2015.10.009

工业 CT 的高铁齿轮箱体材料缺陷识别

艾轶博¹, 王楠¹, 阙红波², 杨斌¹, 张卫冬¹

(1.北京科技大学 国家材料服役安全科学中心,100083 北京;
2.中国南车股份有限公司 南车戚墅堰机车车辆工艺研究所有限公司,213011 江苏 常州)

摘 要: 高铁齿轮箱是高速列车的重要部件,为保障高铁的安全、稳定运行,需要对高铁齿轮箱箱体出厂及检修时的铸件内部缺陷进行检验,并对箱体内部缺陷实现自动、准确的分类和识别.基于此利用三维工业 CT 技术,设计实验获取到高铁齿轮箱体材料的 4 种内部缺陷的三维体数据,根据齿轮箱体内部缺陷的物理背景知识,对三维体数据进行特征提取,设计 Adaboost_BTSVM 多分类算法,实现基于三维工业 CT 的箱体材料内部缺陷的自动分类识别,并使重点关注的收缩类缺陷的分类准确率达到 85% 以上、裂纹类缺陷的分类准确率达到 100%,为实现高铁齿轮箱箱体材料的缺陷自动识别提供技术保障.

关键词: 模式分类;支持向量机;三维特征提取;高铁齿轮箱体;铸造缺陷;工业 CT

中图分类号: TP391 **文献标志码:** A **文章编号:** 0367-6234(2015)10-0045-05

Material casting defect recognition of high-speed train gearbox shell based on industrial CT technology

AI Yibo¹, WANG Nan¹, QUE Hongbo², YANG Bin¹, ZHANG Weidong¹

(1.National Center for Materials Service Safety China, University of Science and Technology Beijing, 100083 Beijing, China;
2.Qishuyan Institute Co., Ltd., China South Locomotive & Rolling Stock Corporation limited, 213011 Changzhou, Jiangshu, China)

Abstract: High-speed train gearbox shell is an important component of high-speed train. In order to protect the operational safety of high-speed train gearbox shell, it is needed to detect the casting internal defect as product testing and maintenance inspection accurately and rapidly. In this paper, based on three-dimensional CT technology the test was developed to detect the casting defects of high-speed train gearbox shell; through the analysis of three-dimensional data of the four kinds detects, three-dimensional geometric features and characteristic values were obtained, and the Adaboost_BTSVM algorithm were used to achieve the automatic classification of casting defects of high-speed train gearbox shell. The according classification accuracy of shrinkage defects can be 85%, and the classification accuracy of crack defects can stand at 100%. These will provide an available automatic identification method for the defect of high-speed train gearbox shell.

Keywords: pattern classification; support vector machine; 3D feature extraction; high-speed train gearbox shell; casting defects; industrial CT technology

高铁齿轮箱箱体结构复杂,多采用铸造方法生产.由于铸造工序繁多,影响铸件质量的因素复杂且具有较大的相关性,难以综合控制,铸造缺陷难以避免^[1].同时,铸件的部分缺陷在服役过程中会不断演化,最终导致铸件失效,使高铁齿轮箱报废、缩短使

用寿命甚至造成高铁运行事故.因此,检测和识别高铁齿轮箱箱体的铸件内部缺陷^[2],对保障高铁齿轮箱的服役安全和高铁列车安全稳定运行具有重要意义,对高铁齿轮箱箱体铸件的内部缺陷检测和质量评估将成为高铁齿轮箱箱体安全评价的首要环节.

工程结构材料铸件中存在各种尺度和维数的缺陷,在服役过程中,微结构形态、缺陷分布和形貌、界面特性、孔隙率等都极大的影响着材料宏观整体韧性、强度、硬度等力学性能^[3].依据铸造铝合金规范,高铁齿轮箱体的铸造缺陷相当于晶粒级,与微结构强烈相互作用,超声、红外等无损检测手段并不适

收稿日期: 2014-07-10.

基金项目: 国家自然科学基金面上项目(61273205);教育部中央高校基本科研业务费项目(FRF-SD-12-028A);高等学校学科创新引智计划(B12012).

作者简介: 艾轶博(1981—),女,博士研究生;
张卫冬(1974—),男,研究员,博士生导师.

通信作者: 张卫冬, zwd@ustb.edu.cn.

用,而工业 CT(computed tomography)是材料内部微观结构状况的最佳无损检测手段^[4-5].目前在齿轮箱体裂纹失效分析中,采集二维图像信息,很难将箱体材料疲劳短裂纹检测出来,其他缺陷,如气孔和渣孔等也同样难以识别,但三维立体结构数据能够有效、完整的记录材料中几乎所有的缺陷.因此本文以高铁齿轮箱箱体材料作为研究对象,重点研究高铁齿轮箱体材料的 4 种内部缺陷通过工业 CT 技术获取高铁齿轮箱箱体材料的三维体数据,提取铸件内部缺陷的三维几何特征作为缺陷的特征数据,并针对对待分类铸件内部缺陷特点设计齿轮箱体内部缺陷分类算法,应用 Adaboost 改进 BTSVM(二叉树 SVM)方法设计分类器,实现高铁齿轮箱体材料内部缺陷的有效检测和缺陷类型的识别.

1 高铁齿轮箱体铸件内部缺陷类别

高铁齿轮箱体材料的内部缺陷种类众多,根据国家分类标准中对不同种类铸造缺陷的定义,将铸件内部缺陷划分为 4 类:气孔类缺陷,收缩类缺陷,夹杂类缺陷和裂纹类缺陷^[6].

根据高铁齿轮箱体质检要求,在出厂检测和定期检修时,有裂纹类缺陷视为不合格,其他缺陷也要满足几何尺寸要求.裂纹类缺陷,会导致高铁齿轮箱在服役过程中,出现开裂、断裂等,严重影响高速列车的运行安全.虽然出厂检测合格的箱体被视为不存在裂纹类缺陷的,但高铁齿轮箱体在服役一段时间后,箱体材料中可能会出现裂纹类缺陷;此外,箱体材料中的其他缺陷在服役中也容易演化为裂纹缺陷,特别是收缩类缺陷,在服役过程中容易造成应力集中而使力学性能大大降低,演化为裂纹类缺陷.因此,在高铁齿轮箱体材料出厂检测和服役过程中定期检修时,需要特别注意裂纹缺陷和收缩类缺陷的识别.

2 高铁齿轮箱体材料内部缺陷数据采集

本文采用 GE v|tome|x m CT 全能型微焦点 X 射线工业 CT 检测系统,对高铁齿轮箱体材料试样进行内部缺陷检测,检测试样最小分辨率为 5 μm.

由铸件内部缺陷可知,合格的高铁齿轮箱体材料被视为不存在裂纹缺陷的,要得到裂纹类缺陷数据就需要对合格箱体的试样材料进行加速试验,模拟箱体材料的服役过程中内部裂纹类缺陷的产生.选取高频疲劳试验机对试样进行高频疲劳加载实验,依据箱体材料的屈服强度,确定疲劳加载载荷为 150 MPa,设置实验机频率下降 1 Hz 自动停止,观察是否出现裂纹,如此反复,直到试样出现肉眼可见时

裂纹停止实验.由于疲劳裂纹较小,设计了试样加持工具,使试样在进行 CT 扫描时,也能持续提供大约 1.5 kN 的拉力,提取清晰裂纹三维数据.

3 高铁齿轮箱体材料内部缺陷自动识别

3.1 箱体材料内部缺陷特征提取

对实验得到的数据进行处理,首先采用阈值分割的方法进行缺陷提取.通过灰度值,可以将夹杂类缺陷从背景材料中区分出来;裂纹类缺陷的三视图中可以发现试样内部条状缺陷,即为裂纹缺陷.如图 1 所示为夹杂类缺陷,图 2 所示为气孔类缺陷和收缩类缺陷,图 3 中箭头所指示的为裂纹缺陷.

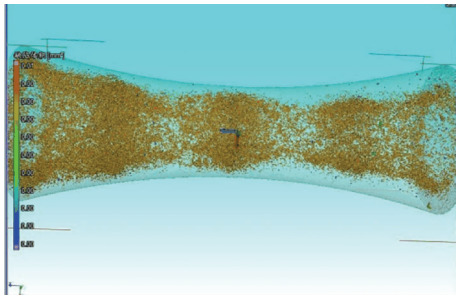


图 1 夹杂类缺陷

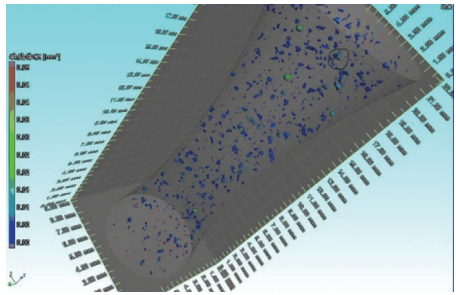


图 2 气孔类缺陷和收缩类缺陷

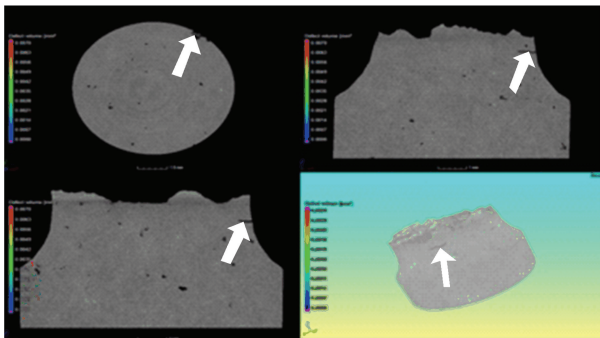


图 3 裂纹(箭头所指)类缺陷

在区分缩孔类缺陷和气孔的时候,可以借助缺陷的断层截面图和 3D 图进行区分.根据气孔与缩孔的定义,气孔为圆形孔洞且内壁光滑,而缩孔形状不规则且孔壁表面粗糙,在图 4 和图 5 中可明显区分出 2 种缺陷.

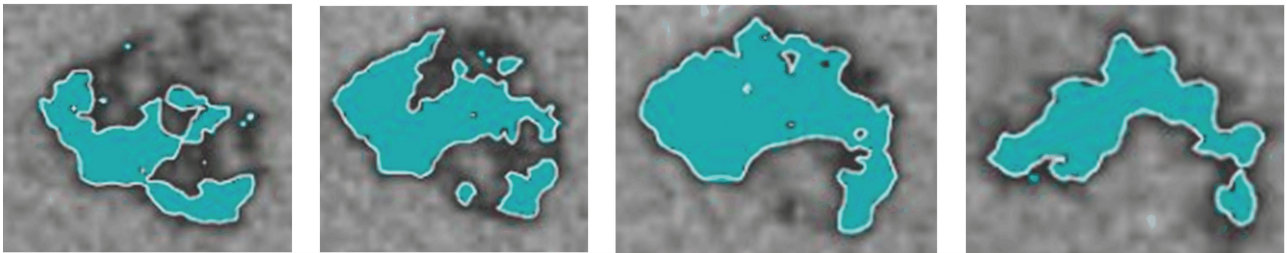


图 4 收缩类缺陷断层截面

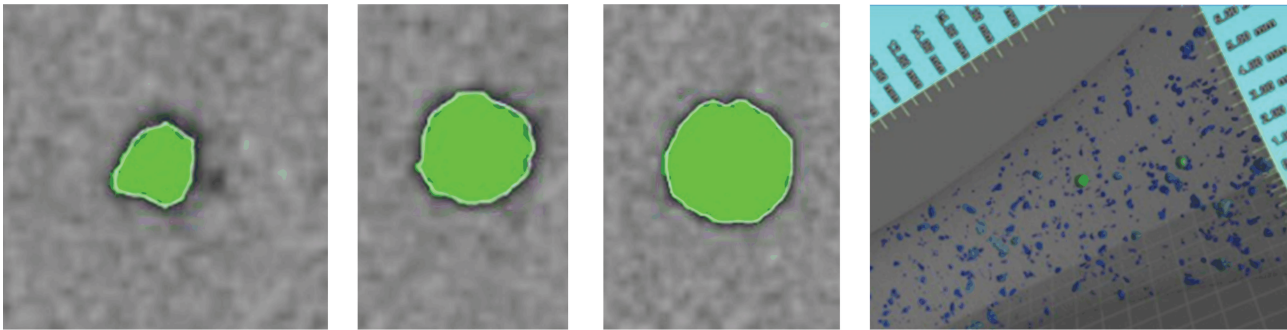


图 5 气孔断层截面图和气孔 3D 图

因此,利用工业 CT 三维数据进行缺陷类别的识别需要考虑缺陷灰度、缺陷内壁光滑程度、缺陷的基本形状、缺陷几何大小等等.4 类铸造缺陷的三维数据类别较多,最终选取如下参量作为分类的特征值.

- 1)半径、表面积、体素等从整体上描述缺陷的大小.
- 2)紧密度.缺陷表面及其体积之间的比值.对于体积差不多的气孔类缺陷和收缩类缺陷,由于气孔类缺陷内壁光滑,收缩类缺陷内壁凹凸不平,紧密度差距很大.
- 3)球度.与缺陷体积相同的球体表面与缺陷表面之间的比率.对于裂纹类缺陷、气孔类缺陷以及收缩类缺陷的区分很明显.
- 4)尺寸(X 、 Y 、 Z)、 PX 、 PY 、 PZ (封闭缺陷的边界框的尺寸和缺陷投影形状).对于气孔类缺陷、收缩类缺陷和裂纹类缺陷的区别非常有效.

5)平均灰度值、灰度偏差(缺陷内的平均灰度值和灰度值偏差).用于区分夹杂类缺陷,并对所有缺陷的灰度平均值和灰度偏差进行归一化后作为分类特征值.

虽然依据缺陷的几何特征,可以对缺陷类型进行人工标记,但铸件检测尤其在服役检修时,内部缺陷较多,对于缺陷的分类识别还需要开展自动识别方法研究.

3.2 多分类器设计

支持向量机(SVM)是一种机器学习算法^[7-8],通过构造不同的超平面,引入非负松弛变量集合,在

错误最小的情况下尽量将样本准确分离^[9].SVM 算法在解决小样本^[10]、非线性、数据不平衡^[11]及高维模式^[12]识别中表现出许多特有的优势.SVM 本质是二分类器^[13],通过对 K 个类别设定 I 个分类器实现多分类功能,二叉树方法可以与 SVM 结合,实现多分类算法.

本研究中,由于缺陷数据中,夹杂类缺陷较多,气孔、缩孔类缺陷次之,裂纹缺陷较少,夹杂类缺陷差别区分度大、裂纹类缺陷次之,且各类缺陷数据严重不平衡,考虑选用 SVM 进行作为二叉树多分类器的基础分类器^[11],即 BTSVM 分类器.

依据各类铸件内部缺陷的不同特点,对缺陷进行分类,分类器设计主要考虑如下:

- 1)夹杂类缺陷.空隙的灰度值低于周围材料的灰度值,而夹杂的灰度值高于周围材料的灰度值,由此设计第 1 层分类器用以区分夹杂类缺陷.
- 2)裂纹类缺陷.呈现条状或者片状(平面),与气孔收缩类缺陷有明显不同.
- 3)考虑裂纹缺陷在产品出厂检测时的重要性,在进行多分类时,第 2 层分类把裂纹看做一类,气孔和缩孔看做一类,设计第 2 层分类器.在第 2 层分类时就将裂纹缺陷区分出来,可以有效避免裂纹缺陷遭受 2 次分类错误.如果将裂纹缺陷放在第 2 层分类器区分,第 1 层区分时会对一部分裂纹缺陷进行错误区分,第 2 层区分时又会对裂纹缺陷进行一次错误区分,裂纹的准确率就会降低.
- 4)第 3 层分类器,将气孔类缺陷和收缩类缺陷分开.

高铁齿轮箱箱体铸造缺陷多分类器设计如图 6 所示.

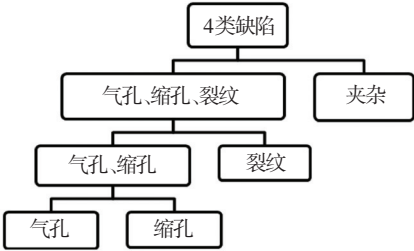


图 6 高铁齿轮箱箱体内部缺陷多分类器设计

由于铸件各类内部缺陷的数量差别很大,待分类的样本数量不均衡,此外气孔和缩孔两类缺陷在形貌上相似度高、错分率较高,特别是对裂纹缺陷的识别希望尽可能达到 100%,需要设计有效的分类器来实现铸件内部缺陷的分类识别.

3.3 Adaboost_BTSVM 多分类算法

目前,基于 Boosting 方法中 Adaboost“adaptive boosting”(自适应增强)方法应用较多,它能够自适应的调整弱学习算法的错误率,使其在弱学习算法学习的过程中经过若干次迭代以后整体错误率能够达到期望值^[14].Adaboost 方法能够调整增加易分错样本的权值,增加易错分样本训练次数,并通过不断地加入新的“弱分类器”,直到联合生成一个强分类器,以达到某个预定的足够小的误差率,这正与高铁齿轮箱体材料铸造缺陷分类识别的数据特点相吻合,因此采用 Adaboost 算法与 BTSVM 分类器相结合来进行铸造缺陷分类研究.

Adaboost 算法的具体流程如图 7 所示,总的训练样本集为 D ,有 n 个样本,每个样本都有自己的权值 $W_k(i)$,初始权值为 $1/n$,在选取每个分类器的样本时,都是由权值从大到小选择. C_k 为多个弱分类器,每次由当前权值(根据上一次训练结果,将易分错的样本权值增加,分对的样本的权值减小)分布选取不同的样本放到 C_k 中进行训练(按权值大小将对样本进行排队,权值大的排在前面,由前往后选取样本放到 C_k 中进行训练),以达到对容易分错的样本进行多次训练的目的,从而提高分类准确度.

对于 k 个弱分类器,也都对应一个权值 a_k ,表示这个分类器的准确程度.最终的强分类器 $g(x)$ 由 k 个弱分类器组成,分类准确的弱分类器相应的权值较大,分类不准确的分类器,权值较小, $g(x) = [\sum_{k=1}^{k_{\max}} a_k h_k(x)]$.测试时,将样本代入每个分类器进行检查,根据每个分类器的权值,计算 $g(x)$,得到样本的类别.

本文选择 SVM 作为二叉树节点位置的分类器,

组成二叉树 SVM 多分类器.在每个节点处,用 Adaboost 改进 SVM 算法,即 SVM 为 Adaboost 的弱分类器.这就是 Adaboost_BTSVM 多分类算法.

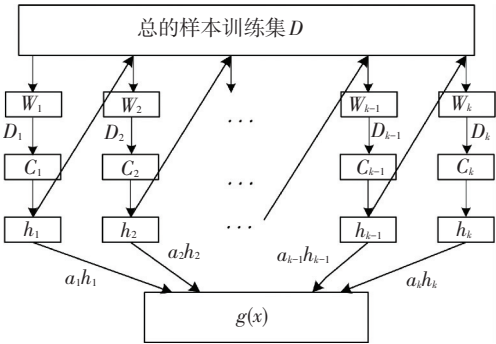


图 7 Adaboost 算法流程

4 结果分析

依据上述分类算法,选取了 320 个缺陷进行训练和测试,其中气孔类缺陷有 110 个,收缩类缺陷有 170 个,裂纹类缺陷有 40 个.分别用一对多 SVM,二叉树 SVM 和用 Adaboost 改进二叉树 SVM 方法进行多分类学习.利用数据的 60%作为训练样本,40%作为测试样本,分类准确率比较如表 1 所示.

表 1 3 种分类器分类准确率比较				%
分类方法	总准确率	裂纹类	气孔类	缩孔类
一对多 SVM	76.38	43.33	81.82	77.94
二叉树 SVM	77.17	33.33	81.82	83.82
Adaboost_BTSVM	84.42	100.00	80.00	85.29

通过表 1 比较学习结果,可见:

- 1) 用 Adaboost 改进二叉树 SVM 方法,取得了较好的分类结果,这是由于应用 Adaboost 算法,通过反复迭代与权值调整,减少了由于裂纹数据训练样本少而导致错误分类结果.
- 2) Adaboost_BTSVM 算法的裂纹识别准确率最高能达到 100%,收缩类缺陷的准确率达到 85.29%,显著提高了收缩类缺陷和裂纹缺陷的分类准确率,对于有效识别裂纹缺陷保障高铁齿轮箱的安全运行具有重要意义.
- 3) 结果中气孔和缩孔类缺陷识别率较低,是由于而且物理几何形态相似,且训练样本数据有限,存在一定的错分情况.

5 结 论

- 1) 本研究针对高铁齿轮箱体铸件的 4 种缺陷的特征,对缺陷工业 CT 三维数据进行处理,选取缺陷的半径、表面积、紧密度、球度、体素、平均灰度值、灰度偏差等几何特征作为缺陷分类的特征值.

2) 利用 Adaboost 改进二叉树 SVM 算法进行高铁齿轮箱体内部缺陷自动识别后,较 SVM 方法与二叉树 SVM 方法准确率有较大提高,收缩类缺陷和裂纹类缺陷分类准确率明显提高,总准确率达到 84.42%,其中裂纹类缺陷的分类准确率达到 100%,收缩类缺陷的准确率达到 85.29%.

参考文献

- [1] 曹庆峰,王立志,李琪,等.齿轮箱箱体裂纹失效分析[J].热加工工艺,2008,37(17):134-136.
- [2] MERY D, JAEGER T, FILBERT D. A review of methods for automated recognition of casting defects[J]. Insight, 2002, 44(7): 428-436.
- [3] AMANAT N, TSAFNAT N, JONES A S, et al. An investigation into compression properties and microstructure using X-ray microtomography[J]. Scripta Materialia, 2009, 60(2): 92-95.
- [4] WITTE F, FISCHER J, BECKMANN F, et al. Three-dimensional microstructural analysis of Mg-Al-Zn alloys by synchrotron-radiation-based microtomography[J]. Scripta Materialia, 2008, 58(6): 453-456.
- [5] 敖波,张小海,李春阳,等.疲劳裂纹扩展高分辨率 CT 分析新进展[J].CT 理论与应用研究,2010,19(2): 35-41.
- [6] 范金辉,华勤.铸造工程基础[M].北京:北京大学出版社,2009.
- [7] 张宇航,张晔.SVM 和 RVM 对高光谱图像分类的应用潜能分析[J].哈尔滨工业大学学报,2012,44(3): 34-39.
- [8] DEMETGUL M. Fault diagnosis on production systems with support vector machine and decision trees algorithms[J]. The International Journal of Advanced Manufacturing Technology, 2013, 67(9/10/11/12): 2183-2194.
- [9] SCHÖLKOPF B, PLATT J C, SHAWE-TAYLOR J, et al. Estimating the support of a high-dimensional distribution[J]. Neural Computation, 2001, 13(7): 1443-1471.
- [10] WASIKOWSKI M, XUEWEN C. Combating the small sample class imbalance problem using feature selection[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1388-1400.
- [11] 杨智明,彭喜元. μ SVM——一种适用于解决不平衡数据分类问题的新方法[J].仪器仪表学报,2008,4(29): 795-798.
- [12] BRAUN A C, WEIDNER U, HINZ S. Classification in high-dimensional feature spaces-Assessment using SVM, IVM and RVM with focus on simulated EnMAP data[J]. IEEE Selected Topics in Applied Earth Observations and Remote Sensing, 2012, 5(2): 436-443.
- [13] ZHANG Shuning, WANG Fuli, HE Dakuo, et al. Online quality prediction for cobalt oxalate synthesis process using least squares support vector regression approach with dual updating[J]. Control Engineering Practice, 2013, 21(10): 1267-1276.
- [14] MEIR R, RAETSCH G. An introduction to boosting and leveraging[J]. Advanced Lectures on Machine Learning, 2003, 2600: 118-183.

(编辑 张 红)