

·经验交流·

多级图像分类系统在硅钢冷轧 表面缺陷检测中的应用

甘胜丰¹, 孙林¹, 曹阳², 刘广益²

(1. 中国地质大学(武汉)机械与电子信息工程学院, 湖北 武汉 430074;
2. 武汉钢铁(集团)公司)

摘要:针对硅钢冷轧生产过程中带钢表面缺陷的特点,采用一种基于专家经验和决策树分类器的多级带钢表面缺陷图像分类系统,包含训练集与测试集的采集、特征提取、专家经验训练分类器、决策树分类器等。实际应用表明,该系统能较好地识别结疤、重皮、孔洞和边裂等严重缺陷图像,提高了武钢硅钢冷轧生产过程的自动化水平。

关键词:冷轧硅钢带钢;表面缺陷检测;图像分类;专家经验;决策树

中图分类号: TP181 **文献标志码:** A **文章编号:** 1000-7059(2009)02-0063-03

Application of multi-level image classification system in surface defect detection of cold-rolled silicon strip

GAN Sheng-feng¹, SUN Lin¹, CAO Yang², LU Guang-yi²

(1. Machinery and Electronic Information Engineering College, China University of Geosciences, Wuhan 430074, China;
2. Wuhan Iron and Steel Co., Ltd.)

Abstract: According to features of surface defect on cold-rolled silicon strip, a multi-level image classification system for strip surface defect based on expert knowledge and decision tree classifier was adopted. The system includes collection of training set and test image, extraction of image feature, expert knowledge training classifier and decision tree classifier. Practical applications show that the system can recognize scab, double skin, holes and edge crack etc. serious defects and automatic level of cold-rolled silicon strip production process is improved.

Key words: cold-rolled silicon strip; surface defect detection; image classification; expert knowledge; decision tree

0 引言

由于含硅量较高的带钢物理特性十分复杂,因此刚酸洗过的硅钢表面图像状况也非常复杂,除了酸洗不良引起的麻点、水渍等轻度表面缺陷外,还有数量庞大,种类与图像纹理变化复杂,无规律性的结疤、重皮、孔洞和边裂等严重缺陷。这些严重缺陷随机出现,且容易损坏轧辊或者造成

断带,严重影响生产效率。而一些轻度缺陷和严重缺陷的混合造成采集到的图像质量较差,类与类之间的重叠相当明显,即使是同一类缺陷图像也出现表面属性差别大等现象,使得现场在线检测难以达到自动控制所要求的识别率。

我们在提取几何特征、灰度特征、纹理特征、统计特征等低层特征^[1]的基础上,建立了一个基

收稿日期: 2008-06-26; 修改稿收到日期: 2008-10-10

基金项目: 湖北省自然科学基金资助项目(2006ABA346)

作者简介: 甘胜丰(1983-),男,湖北石首人,硕士研究生,主要研究方向为模式识别与设备自动控制。

于专家经验与决策树分类器的组合分类器,来实现具有一定学习能力、能够适应硅钢生产线相应变化的缺陷检测系统。通过研究和实验,该系统能够对结疤、重皮、孔洞和边裂等缺陷检出分级,参与武钢硅钢冷轧带钢自动化生产线的在线控制,从而避免轧制事故,提高生产效率、降低轧辊消耗。

1 组合式图像分类系统

根据硅钢带钢图像的特点提取和选择带钢缺陷特征,采用组合式图像分类系统,建立了基于专家经验训练和基于决策树 D3 算法的多级分类器。通过表面缺陷检测图像识别系统实现了训练集和测试集的收集、特征提取、专家经验训练分类、决策树分类器设计和分类决策等^[2]。图像识别流程图如图 1 所示。使用的训练集与测试集是从武钢冷轧硅钢生产线的带钢上采集得到,这些图像经过图像处理后即可进行特性提取等后续工作。

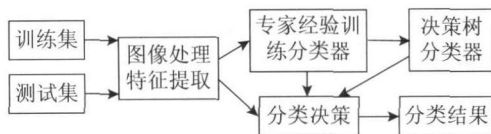


图 1 图像识别流程

Fig. 1 Flow diagram of image recognition

1.1 特征提取

根据硅钢带钢表面缺陷的特点,我们提取了 55 个缺陷图像的低层属性特征,可概括为位置特征、外形特征、段特征、灰度特征等。特征数学描述越完备信息丢失也就越少,但是特征越多效果不一定会越好,有时会降低分类效果。根据图像属性的划分,抽取的部分特征列表如表 1 所示。其中位置特征缺陷是所在带钢表面的空间位置特征;外形特征是缺陷的形状特征;段特征是单个行扫描上的一系列相邻(横向连续)阈值限定像素上的特征,段宽是横向连续阈值限定像素的长度,段宽比是介于 0.0~1.0 之间的比率,表示缺陷是否具有恒定的长宽比;灰度特征是将视图分为灰度级库,其中四个用于暗像素(暗度库 1, 2, 3 和 4),四个用于亮像素(亮度库 1, 2, 3 和 4)。暗度库 4 为最大极限(最暗),亮度库 4 为最大极限(最亮), q 和 r 分别为灰度级库各等级的代替符。“%亮度库 $q+r$ ”是缺陷像素中属于亮度库 q 和 r 的缺陷像素所占的百分比。

表 1 部分表面缺陷图像特征列表

Table 1 Part surface defects image feature

位置特征	外形特征	段特征	灰度特征
中心距	长宽比	段宽比	亮度库 q 面积
左(右)边距	缺陷面积	最大段宽	暗度库 $q+r$ 面积
最近边距离	缺陷周长	平均段宽	%亮度库 $q+r$

1.2 专家经验训练分类器

由于许多缺陷是前工序或固定工艺所造成的,形成原因较为固定,或者在形成该缺陷的设备中也有相应的输出控制,并且许多缺陷有固定的位置、纹理、灰度等特征,所以容易根据专家经验通过简单的决策来识别。第 1 级分类器是用 C++ 语言编写的一组应用程序,它集成实际生产中的带钢特征参数和缺陷类别作为可配置选项。依据硅钢工艺专家的经验,设置第 1 级分类器中的可配置选项来确定工艺中的固定缺陷的分类决策。例如:面积小于 15 mm^2 的缺陷可以不做特征提取和分类;由工艺原因产生的大规模氧化皮缺陷可以不做特征提取;距边界 15 mm 以内的缺陷均可以定义为边界缺陷;表面结疤的“%亮度库 3+4”属性应该高于 15% 等等。因此根据专家经验建立相对独立的决策来分离这些缺陷,可以简化后续决策树的节点,提高决策树分类器的速度,减少分类器的误识率,提高系统的有效缺陷运算能力,增加冷轧硅钢带钢表面缺陷的在线检测效率。

1.3 决策树分类器^[3]

采用决策树的分类方法设计第 2 级分类器。决策树是用于分类的一种树结构,树的每个内部节点代表对某个属性的一次测试,每条边代表一个测试结果,叶节点代表某个类或者类的分布,最上面的节点是根节点。与其他分类算法相比,决策树分类法有如下优点:(1)速度快,计算量相对较小,且容易转化成分类规则;(2)挖掘出的分类规则准确性高,便于理解,可以清晰地显示哪些字段比较重要。

本系统中带钢缺陷图像识别所采用的是决策树 D3 算法,其应用过程如下^[4]。

步骤 1,节点的设计。计算训练集中所有带钢属性的信息增益,选择信息增益最大的属性作为根节点;测试集中至少有 2 组叶节点,即测试集含有两组以上的已分类数据。

步骤 2,由根节点属性的不同建立多组决策树分支。在训练集数据比较多的情况下,使用一组

决策树会使得决策结果变得很庞大,不易于理解,也会在决策测试集时增加判断时间,不宜于在线检测,所以我们可以根据根节点属性的不同选择组建多组决策树。同时为了保证每组决策树形成的决策结果不会过度拟合,可选择一定的决策树生长长度。这样就保证了训练集全部参与训练且得到较好的决策结果。应用中采用的最大生长长度为 20 层,决策组数为 10 组。

步骤 3,采用递归的方法,对各分支的子集仍然选择信息增益最大的属性作为子节点直到所有的子集包含同一类别的数据为止。为保证分类准确和防止过度拟合,设计叶节点容错率为 25%。

步骤 4,测试训练集。将训练好的分类决策法用于测试训练集,人工判断其结果,一方面便于将人工分类错误加以纠正;另一方面可以判断相关参数设置的优劣,便于找到最佳分类参数。

2 现场应用

首先取 340 个严重缺陷图像作为训练集,再取现场 10 组带钢缺陷图像(共 172 个)作为测试集。将图像分为 4 类:重皮、边裂、结疤、孔洞,并将其作为决策树的叶节点。将提取的 55 个缺陷图像低层属性特征作为决策树的内部节点来构建相应的决策树。

由 D3 算法可以将分类能力最好的特性作为根节点,然后通过信息最大化的网络修剪^[4],构造出一颗分类为结疤、重皮、孔洞和边裂的决策树,部分决策方式如图 2 所示。

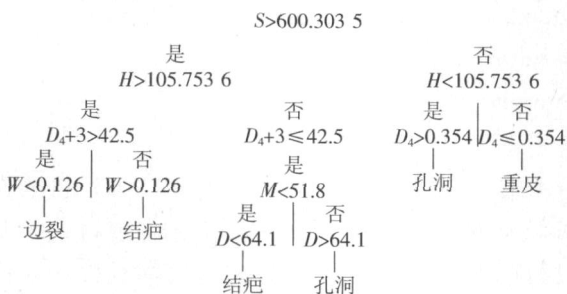


图 2 缺陷分类决策树结构

Fig. 2 Structure of defects classification decision tree

S—暗域面积, mm²; H—高度, mm; D₄+3—暗度库 4+3 面积, mm²; D₄—暗度库 4 面积, mm²; W—宽度比, %; M—最大宽度, mm; D—灰度比, %

用专家经验训练和决策树训练得到的分类决策对测试集进行分类,可以得到如表 2 所示的测

试结果。

表 2 缺陷识别率

Table 2 Rate of defect recognition

%			
结疤	重皮	孔洞	边裂
88	76	94	86

由上述结果可以看出,专家经验训练和决策树组成的多级分类器对于严重表面缺陷有较高的识别率。因此使用多级分类器进行在线检测能够检测到较准确的严重缺陷,并可以给控制系统提供比较正确的控制信号。但由上述结果还可看出在图像比较复杂的重皮等缺陷方面,还需要加以改进,以提高相应的识别率和整体识别率。

3 结论

采用提取带钢图像低层属性特征作为特征属性,使用决策树和专家经验分类的方法对硅钢带钢表面缺陷进行识别,该分类方法于 2007 年 11 月应用于武钢冷轧硅钢片厂冷轧生产线上,缺陷识别状况良好,可以较好地识别结疤、重皮、孔洞和边裂等重要缺陷,严重缺陷的在线整体识别率达到了 76% 以上,有效地提高了硅钢冷轧生产过程自动化水平。

参考文献:

[1] 李 炜, 黄心汉, 王 敏, 等. 基于机器视觉的带钢表面缺陷检测系统 [J]. 华中科技大学学报(自然科学版), 2003, 31(2): 72-74.
LI Wei, HUANG Xin-han, WANG Min, et al. The steel surface inspection system based on computer vision [J]. Journal of Huazhong University of Science and Technology (Nature Science Edition), 2003, 31(2): 72-74.

[2] 王陈飞, 肖诗斌. 基于 SVM 的图像分类研究 [J]. 计算机与数字工程, 2006, 34(8): 74-75.
WANG Chen-fei, XIAO Shi-bin. Study of image classification based on SVM [J]. Computer & Digital Engineering, 2006, 34(8): 74-75.

[3] Quinlan J R. Introduction of decision trees [J]. Machine Learning, 1986(1): 84-100.

[4] 黄 文. 决策树的经典算法: D3 与 C4. 5 [J]. 四川文理学院学报, 2007, 17(5): 16-18.
HUANG Wen. The algorithm of decision trees: D3 and C4. 5 [J]. Sichuan University of Arts and Science Journal, 2007, 17(5): 16-18.

[编辑:沈黎颖]