

基于支持向量机的彩钢玻璃色差分离技术研究*

张梓然 张瑞峰 李 铨 南刚雷

(天津大学 电子信息工程学院 天津 300072)

摘要: 为了将彩钢玻璃中有色差的玻璃样本分离出来,提出一种利用分光测色仪检测,并与支持向量机相结合分离色差样本的方法。首先利用分光测色仪获取出样本玻璃的各采样点的 CIE 色度值等光学数据,建立数据样本集,然后选取特征参数及训练集并建立 SVM 模型,其次选取不同类型核函数及参数进行对比实验,使模型最优化,最后利用该模型对预测集进行处理,检测并分离出有色差的玻璃样本。实验结果表明该方法可以高效准确地检测出彩钢玻璃中有色差的样本。

关键词: 支持向量机;色差分离;CIE 色度值;彩钢玻璃

中图分类号: TP391.5 **文献标识码:** A **国家标准学科分类代码:** 510.40

Research of separating color discrepancy of glass based on support vector machine

Zhang Ziran Zhang Ruifeng Li Qiang Nan Ganglei

(School of Electronic information engineering, Tianjin University, Tianjin 300072, China)

Abstract: In order to distinguish the color discrepancy of the same color class, the method of detection based on spectrophotometer and support vector machine is proposed. First, spectrophotometer is utilized to obtain the CIE chromaticity coordinates of each point in the glasses and other optical data for establishing data sample set. Then selecting the feature parameters and establishing the training set are used for creating SVM model. Followed, the different types of kernel functions and parameters are compared to choose the best, so that they make the model optimizing. Finally, the model is used to work on the prediction set, which detecting and distinguishing the dissimilarity. The experimental results show that this method can be efficiently and accurately detected the discrepancy of color of the glass.

Keywords: support vector machine; CIE chromaticity coordinates; color discrepancy; color glass

1 引言

彩钢玻璃在目前市场上应用广泛,是多种电器外表封装中不可或缺的材料。随着客户对产品外观要求的日益提高,其生产过程中的品质监控急需重视。目前同一批次产品在生产中会出现色差,而这些色差会直接影响产品销量以及消费者对产品质量的可信度,因而需将有色差的不合格品筛选出来。目前在工业上,对其检测的方法多数采用人工手段,而在大批量生产过程中,既不方便也不精确。那么如何在工程上高效、精准地检测并分离出有色差的不合格品是生产厂家极为关注的问题。

针对上述这一问题,应首先将彩钢玻璃颜色数据化,其

次利用数据处理手段将有色差的数据分离。在颜色检测上,本研究采用测色仪器检测出大量样本玻璃关于颜色的光学数据,而面向大量数据的二分类,本文提出将支持向量机(support vector machine, SVM)的理论引入到其中,建立 SVM 模型后,分离出数据样本中有色差的元素,从而达到上述目的。

2 基本原理

2.1 CIE 色度值

本文采用的是 CIE1931 (国际照明委员会)标准色度系统^[1],在该系统中包括了一切物理上能实现的颜色。由式(1)可得到某颜色的三刺激值 X, Y, Z 。

收稿日期:2015-09

* 基金项目:国家自然科学基金(61471263)资助项目

$$\begin{cases} X = k \int_{\lambda} \phi(\lambda) \bar{x}(\lambda) d\lambda \\ Y = k \int_{\lambda} \phi(\lambda) \bar{y}(\lambda) d\lambda \\ Z = k \int_{\lambda} \phi(\lambda) \bar{z}(\lambda) d\lambda \end{cases} \quad (1)$$

式中： $\phi(\lambda)$ 称为颜色刺激函数，在式(1)中 $\bar{x}(\lambda)$ ， $\bar{y}(\lambda)$ ， $\bar{z}(\lambda)$ 是 CIE 规定的标准色度观察者的三刺激值，其中常数 k 被称之为归化系数。三刺激值 XYZ 与颜色的色品坐标 xyz 的关系为式(2)^[2]，因此可得出颜色的色品坐标，进而确定被测颜色在 CIE1931 色品图中的位置。

$$\begin{cases} x = \frac{X}{X+Y+Z} \\ y = \frac{Y}{X+Y+Z} \\ z = \frac{Z}{X+Y+Z} \end{cases} \quad (2)$$

2.2 支持向量机

支持向量机是二类分类模型，该方法是在统计学理论基础发展起来的新一代学习算法是一种监督式学习算法^[3]。在线性分类时，其为特征空间上间隔最大的线性分类器；而在非线性分类时通过高维空间变换，将非线性问题转化为高维空间的线性分类问题后再进行分类^[4]。

设训练样本集是：

$$\{(x_1, y_1), \dots, (x_n, y_n), x_i \in \mathbf{R}^m, y_i \in \{-1, 1\}\}$$

式中： x 是特征向量， y 是类标志， n 是样本数， m 是输入维数^[5]。

分类器目标是在 m 维数据空间中找到一个超平面 (hyperplane) 将样本数据分为两类，在此超平面表示为：

$$\mathbf{w}^T \mathbf{x} + b = 0 \quad (3)$$

选取超平面的判定标准是最近的数据点距之间隔 (margin) 尽可能的大^[6]。即间隔值最大化：

$$\max_{\mathbf{w}, b} \frac{1}{\|\mathbf{w}\|} \quad (4)$$

$$y_i (\mathbf{w}^T \mathbf{x}_i + b) \geq 1, i = 1, \dots, n \quad (5)$$

式(5)为约束条件，此时分类间隔为 $2/\|\mathbf{w}\|$ ，其最大化，等价于最小化 $\omega^2/2$ ，这样上式转化为凸二次规划问题，由于给定了约束条件后求最优值，因此可将其转化成为带约束条件的优化问题，即求解拉格朗日乘子问题：

$$L(\mathbf{w}, b, \alpha) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^n \alpha_i (y_i (\mathbf{w}^T \mathbf{x}_i + b) - 1) \quad (6)$$

从而得到：

$$L(\mathbf{w}, b, \alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \quad (7)$$

对上式中 α 的极大值，即是关于问题的最优化，所以得下式：

$$\begin{cases} \max_{\alpha} \left[\sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \right] \\ s. t. C \geq \alpha_i \geq 0, \sum_{i=1}^n \alpha_i y_i = 0, i = 1, \dots, n \end{cases} \quad (8)$$

式中： C 为惩罚参数， C 越大表示对错误分类的惩罚越大^[7]， α_i 为拉格朗日乘子。

当遇到线性不可分时，SVM 的处理方法是引入核函数^[8]，通过将数据在低维度空间中运算，而实质的分类效果表现在高维度上，避免了高维度上计算复杂的问题。故可表示为如下：

$$\max_{\alpha} \left[\sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \right] \quad (9)$$

式中： $K(x_i, x_j)$ 为核函数。常见的核函数有多项式核函数，高斯核函数，线性核函数。

3 实验分析

为了分离有色差的玻璃，首先需要检测其颜色，在此利用分光测色仪器检测样品玻璃的颜色并用 CIE 色品坐标来表示，随后用 SVM 对数据进行分类处理。因为支持向量机作为二类分类器，有处理高维特征数据空间的能力^[9-10]，故适用于解决本文的色差分离问题。

本实验的主体流程框图如图 1 所示。

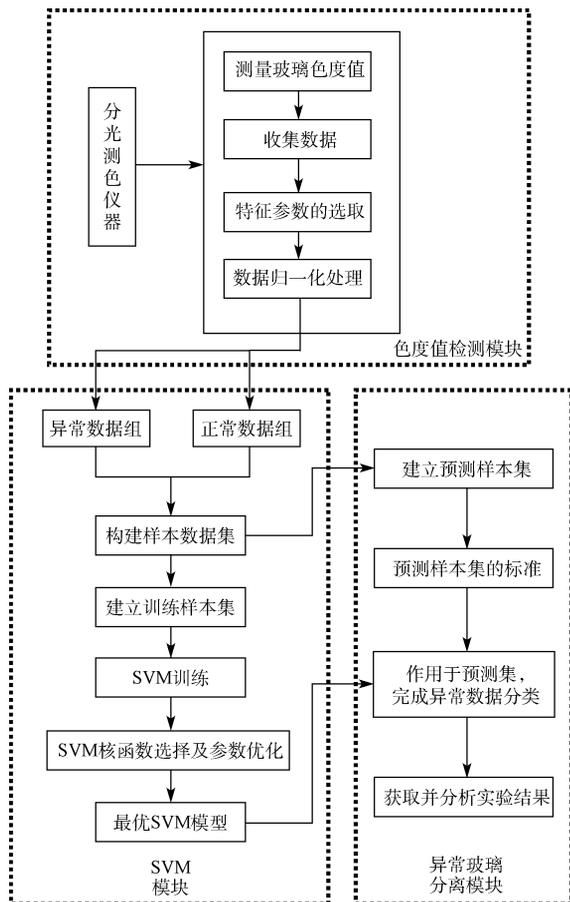


图 1 基于 SVM 的彩钢玻璃色差分离实验框图

3.1 彩钢玻璃颜色数据检测

本实验采用分光测色仪，其通过测量物体的光谱反射

或光谱投射特性后,利用 CIE 推荐的标准照明体,即指特定的光谱功率分布^[11],标准色度观察者的三刺激值及上述公式(1)得出被测物体的色品坐标。在实验中,视野观察角度为 10° ,标准照明体为 D65 照明体。

生产厂家提供 10 块样本玻璃,分别标记为 A, B, C, ..., H, I, J, 其中有颜色差别的不合格品为 E, F, G, 已标注完成。每次每块玻璃随机选取 5 个点,先后进行反复测量,共测得 740 组彩钢玻璃的光学数据,其中主要包括 CIE1931 色品图、CIE 色品坐标值、反射光谱图、主波长、色纯度、中心波长等。如图 2、3 所示,分别给出正常和异常采用点的 CIE 色品图,对比发现,正常与异常点位置相近,不易直接分辨。

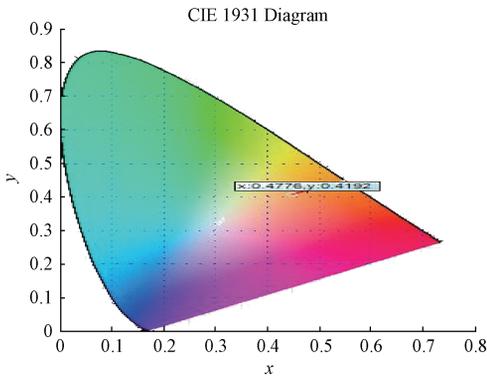


图 2 正常点 CIE 色品图

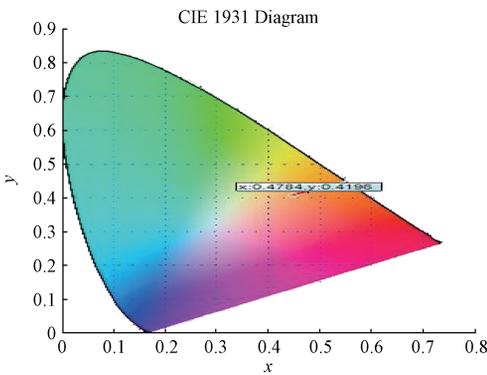


图 3 异常点 CIE 色品图

3.2 支持向量机的实验分析

将上述各采样点的光学数据作为支持向量机的样本集,并对样本集进行正负性标注。随后选取实验所需的特征参数,再从 740 组样本数据中抽取 200 组作为 SVM 的训练集,其余 540 组作为预测集。

SVM 把训练集在低维度空间完成计算后,映射到高维特征空间中,最终可在高维特征空间中构造分类器,完成分类^[12]。因此构建一个用于分离玻璃样本颜色差异的 SVM 模型是至关重要的,同时,最优核函数及参数的选取是构建 SVM 模型的关键。

本实验中构建分离玻璃颜色差异的 SVM 模型框图如图 4 所示。

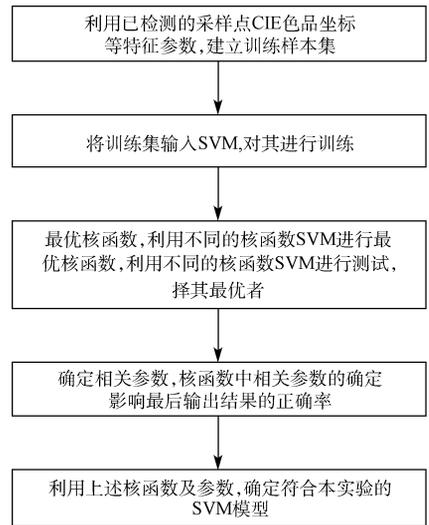


图 4 分离玻璃色差的 SVM 模型建立框图

4 结果和分析

4.1 特征参数的选取

在每组样本数据中,选取 CIE1931 色品坐标值、主波长、色纯度、CIE 白度值、黄变指数、中心波长、半高宽(FWHM)作为样本的特征参数。将各点 CIE 色品坐标值绘制于三维图中,形成如图 5 所示的样本集 CIE 坐标点云图,或称之为各点红色聚类^[13],此外,为了将其余各维度可视化,故绘制成二维图,如图 6 所示。

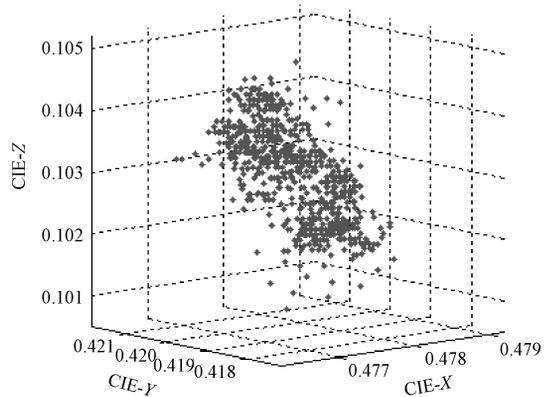


图 5 样本集的 CIE 色品坐标点云图

4.2 特征参数处理及样本的建立

由于参数的范围不同,首先对特征参数整体进行归一化处理,将数据归至 $[-1, 1]$,然后建立样本集,具体步骤如下。

1) 本实验将所测 740 组数据作为样本集,其中 400 组为正常玻璃颜色数据,340 组为异常颜色数据;

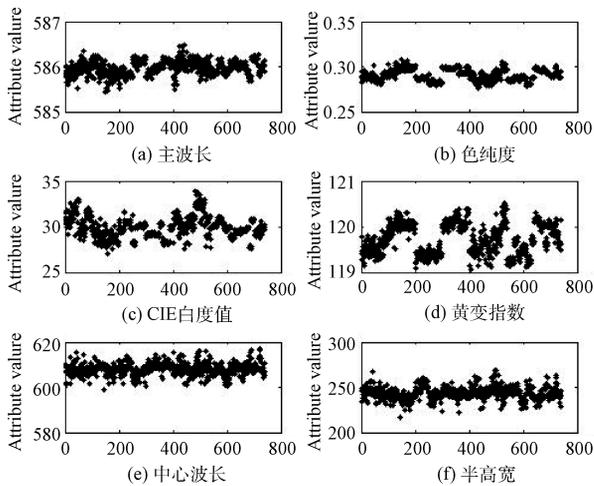


图 6 特征参数的各维可视化图

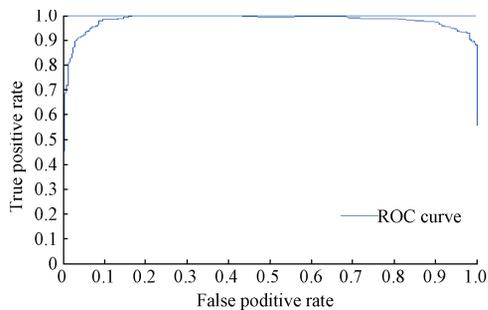


图 7 ROC 曲线

表 2 系统评估参数表

指标	数据
AUC	0.983382
Accuracy	0.948148
Precision ratio	0.927672
Recall ratio	0.983333

2) 将上述数据进行归一化处理,并对原始数据进行正负性标注;

3) 于样本集中提取正常及异常数据各 100 组,共 200 组作为训练集;

4) 将其余 540 组数据作为预测集,其中正常颜色数据 300 组,异常颜色数据 240 组。

4.3 最优核函数选取及参数确定

建立最合适的 SVM 模型是实验的重要步骤,而建立模型的关键是寻找最佳核函数。我们常用的核函数有多项式核函数,线性核函数,高斯核函数(径向基核函数 RBF)^[14-15]。本实验分别用上述核函数对训练集进行分类并比较分类效果,如表 1 所示,其中线性核函数及高斯核函数都有着较高准确度,而后者使用效果最佳,故实验选用高斯核函数(RBF)为 SVM 模型的核函数。

表 1 不同类型核函数的正确率表

核函数类型	正确率/(%)
Linear	93.7037
Polynomial	56.1111
RBF	94.4444

高斯核函数中的参数 g 与惩罚因子 C 的确定至关重要,其直接影响系统输出结果的正确率,因此参数的最优可实现正确率的最大化。经过循环测试,比较得出 $C=4.00, g=0.08$ 是本文实验的最优参数。

4.4 实验 SVM 模型评估

利用上述优化好的 SVM 模型对预测集进行分类,准确度可达到 94.8148%,比例为 512/540。采用 ROC 曲线对整个 SVM 模型进行评估,曲线图如图 7 所示,同时统计出评估系统输出结果的参数,包括正例准确度(precision ratio)、整体准确度(accuracy)、曲线下面积 AUC、查全率(recall ratio),如表 2 所示。

图 7 中横坐标是 FPR(false positive rate)值,代表特异性,纵坐标是 TPR(true positive rate)值,代表敏感性,AUC 反应了模型的准确性。AUC 在 0.5~0.7 时有较低准确性,AUC 在 0.7~0.9 时有一定准确性,AUC 在 0.9 以上时有较高准确性。本实验的 AUC 为 0.9833382,表明了优化后的 SVM 模型有较高准确性,同时查全率及正例正确度都在 90%以上,说明该模型拥有较强的色差分离能力且精度高。

5 结论

本文先用分光测色仪器检测出彩钢玻璃各点的 CIE 色度值及多种光学数据,在之后数据分离处理时,利用支持向量机的二类分类器这一本质特点,将 SVM 引入该方法。从最终的结果上看,本实验的 SVM 模型具有较高的预判准确度。因此得出本文提出的方法可应用于彩钢玻璃色差判别,为解决这一工业生产问题提供了新思路,有一定的工程应用价值。

参考文献

- [1] 袁琨,严惠民,金尚忠. 基于光泽分析的分光测色仪 SCI 误差修正模型[J]. 光电工程, 2014, 41(2): 12-18.
- [2] 李温温. 瓷砖色差在线检测系统的研究[D]. 天津: 天津大学, 2009.
- [3] 刘松松,张辉,毛征,等. 基于 HRM 特征提取和 SVM 的目标检测方法[J]. 国外电子测量技术, 2014, 33(10):38-41.
- [4] 杨海. SVM 核参数优化研究与应用[D]. 浙江:浙江 大学,2014.
- [5] 程淑红,刘洁,朱丹丹. 基于计算机视觉与 SVM 的水

- 质异常监测方法.[J]. 光电工程, 2014, 41(5): 28-33.
- [6] VORA A, PAUNWALA C N, PAUNWALA M. Statistical analysis of various kernel parameters on SVM based multimodal fusion[C]. India Conference (INDICON), 2014 Annual IEEE. IEEE, 2014: 1-5.
- [7] 王道明,鲁昌华,蒋薇薇,等. 基于粒子群算法的决策树 SVM 多分类方法研究[J]. 电子测量与仪器学报, 2015, 29(4): 611-615.
- [8] 姚亚夫,张星. 基于瞬时能量熵和 SVM 的滚动轴承故障诊断[J]. 电子测量与仪器学报, 2013, 27(10): 957-962.
- [9] LIN K P, CHEN M S. On the design and analysis of the privacy-preserving SVM classifier[J]. Knowledge and Data Engineering, IEEE. 2011, 23 (11): 1704- 1717.
- [10] ZAIDI N A, SQUIRE D M G. Local adaptive SVM for object recognition [C]. 2010 International Conference on Digital Image Computing: Techniques and Applications (DICTA). IEEE, 2010: 196-201.
- [11] 陈硕,沈模卫,水仁德. CIE 标准照明体和 Munsell 颜色体计算机仿真[J], 浙江大学学报:工学版, 2003, 37(4): 401-405.
- [12] 周绍磊,廖剑,史贤俊. RBF-SVM 的核参数选择方法及其在故障诊断中的应用[J]. 电子测量与仪器学报, 2014, 28(3): 240-246.
- [13] 常发亮,黄翠,刘成云,等. 基于高斯颜色模型和 SVM 的交通标志检测[J]. 仪器仪表学报, 2014, 35(1): 43-49.
- [14] 陈强,田杰,黄海宁,等. 基于统计和纹理特征的 SRS 图像 SVM 分割研究[J]. 仪器仪表学报, 2013, 34(6): 1413-1420.
- [15] 刘萌萌,刘亚岚,孙国庆,等. 结合纹理特征的 SVM 样本分层土地覆盖分类[J]. 遥感技术与应用, 2014, 29(2): 315-323.

作者简介

张梓然(通讯作者),1990 年出生,天津大学在读硕士,主要研究工作为机器视觉。

E-mail: cottage314@163. com

张瑞峰,1974 年出生,天津大学副教授,主要研究领域为机器视觉、硬件开发。

E-mail: zhangruifeng@tju. edu. cn

李锐,1974 年出生,天津大学教授,中国电子学会高级会员。主要研究领域为智能信息处理(红外图像处理与应用、音乐信号处理、医学信号处理、自适应与盲信号处理)、滤波器设计等。

E-mail: liqiang@tju. edu. cn

南刚雷,1991 年出生,天津大学在读硕士,主要研究工作为机器视觉。

E-mail: Ganglei_Nan@tju. edu. cn

是德科技与中国移动研究院在 5G 关键技术领域展开深度合作

2016 年 5 月 19 日,中国北京——是德科技有限公司(NYSE:KEYS)今天宣布和世界上最大的移动运营商,无线通信网络的领导者中国移动在 5G 关键技术领域 5G 信道测量建模、大规模天线系统的空口测试方面开展深度合作,进一步拓展与中国移动研究院从 2014 年起在 5G 方面的长期合作,共同推进 5G 发展。

作为业内领先的电子设计与测试解决方案公司,是德科技高度重视 5G 无线通信关键技术的研发,与业界技术领导者和学术研究团队开展了紧密合作,以提供独到的 5G 设计及测试解决方案。与中国移动研究院在 5G 关键技术上的合作是其中一个重要的例子。

5G 的信道测量与建模对于 5G 技术研发和标准化至关重要。是德科技有业界领先的 5G MIMO 信道测量解决方案,能够解决 5G 信道测量中,超大频率覆盖范围、超大带宽以及多通道的测量挑战。基于是德科技的 5G 信道测

量解决方案,是德科技实验室与中国移动研究院开展了 5G 信道测量与建模的合作,有效支持了中国移动研究院在 6 GHz~100 GHz 频率范围内 5G 信道模型的研究工作。

5G massive MIMO 系统的 OTA 测试是 5G 系统研发与验证的一个关键。基于中国移动研究院独有的 Massive MIMO 天线技术: SmartTile 和是德科技独有的 massive MIMO 测试技术,是德科技实验室和中国移动研究院在 5G Massive MIMO 系统的 OTA 测试方面开展了深入的合作。在 Massive MIMO 天线测试方面已经取得了初步成果。

在 2016 年的 MWC 展会上,中国移动宣布成立 5G 联合创新中心,是德科技成为中国移动 5G 联合创新中心宣布的首批 11 家合作伙伴之一。双方在 5G 信道测量与建模,Massive MIMO OTA 方面的合作是基于 5G 联合创新中心框架下的合作的进一步深化。