

近红外光谱法检测乙醇柴油主要性能指标

欧阳爱国*, 唐天义, 王海阳, 刘燕德

(华东交通大学 光机电技术及应用研究所, 江西 南昌 330013)

摘要:乙醇柴油是柴油替代品的一种,它的使用越来越广泛,乙醇柴油品质由许多指标决定,采用传统方法检测这些指标不仅价格昂贵而且耗时长。近红外光谱技术是一种廉价、快速实时在线检测乙醇柴油品质的有效方法。本文采用近红外光谱技术结合最小二乘支持向量机检测了乙醇柴油的密度、粘度和乙醇含量,比较了线性和非线性校正技术(主成分回归、偏最小二乘回归和最小二乘支持向量机)对乙醇柴油品质的分析效果,同时也比较了不同预处理方法对预测模型能力的影响。实验结果表明,最小二乘支持向量机优于主成分回归和偏最小二乘回归模型,其对乙醇柴油密度、粘度、乙醇含量建模效果最优,相关系数分别是0.995 8、0.995 7和0.995 3;预测均方根误差分别为0.000 68、0.011 3和0.571 4。**关键词:**近红外光谱;乙醇柴油;密度;粘度;乙醇含量

中图分类号:O657.33; TE626.9 文献标识码:A doi:10.3788/CO.20171003.0363

Detection of key performance indicators of ethanol diesel by the infrared spectroscopy method

OUYANG Ai-guo*, TANG Tian-yi, WANG Hai-yang, LIU Yan-de

(*Institute of Optics-Mechanics-Electronics Technology and Application (OMETA),
East China Jiaotong University, Nanchang 330013, China*)

* Corresponding author, E-mail:ouyangaiguo1968711@163.com

Abstract: Ethanol diesel is one of the alternatives for the diesel, which is used more and more widely. Many indicators reflect the quality of diesel ethanol fuel. It is not only expensive but also time consuming to detect these indicators with traditional method. Near infrared spectroscopy method is an inexpensive, fast and real time online test for the quality of ethanol diesel oil. In this paper, the density, viscosity and the quality of the ethanol content of diesel are detected, and the analysis effect on the quality of ethanol diesel are compared by linear and non-linear calibration technology, including principal component regression, partial least squares regression and least squares support vector machines (LSSVM), and the effects of different pretreatment methods on the prediction model capabilities are also compared. Experimental results show that LSSVM is better than the principal component regression and partial least squares regression model, with the optimal modeling effect on the density of ethanol-diesel, viscosity, alcohol content. The correlation coefficients are 0.995 8, 0.995 7 and 0.995 3, and the root mean square error of prediction are 0.000 68, 0.011 3 and 0.571 4, re-

收稿日期:2017-01-25;修订日期:2017-03-22

基金项目:南方山地果园智能化管理技术与装备协同创新中心资助项目(No.赣教高字[2014]60号)

Supported by Foundation Project of Intelligent Management Technology and Equipment Collaborative Innovation Center of the Southern Mountain Orchard(No. Gan Jiao high [2014]60)

spectively.

Key words: NIRS; ethanol diesel; density; viscosity; ethanol content

1 引言

过去十几年中,汽车行业的快速发展,促使石油燃料消耗逐年增加,同时也带来了一系列如经济、社会、环境等问题。我国是石油进口大国,为了降低对石油进口的依赖程度和空气污染程度,乙醇柴油等一系列石油替代品被研制出来。乙醇柴油的物理性质和柴油相近,在一定浓度内不需要改变发动机的参数就可以直接使用,乙醇柴油代替柴油使用减少了温室气体、颗粒物和硫化物的排放,并作为一种商业品在世界各地推广使用。乙醇生产原料来自不同地方,生产工艺也有所差别,因此,最终乙醇柴油产品可能会有不同的属性,所以乙醇柴油的质量控制非常重要。目前我国还没有一种操作简单、成本低、实时在线检测方法去评价这些石油替代品的品质。

光谱数据(一般由近红外光谱、质谱、核磁共振技术获得)与先进的统计方法相结合,促进光谱分析在分析检测方面的应用。其中,近红外相对其他分析技术的优势在于非破坏性、少量的样品制备、实时响应^[1]在很多领域都得到了应用^[2-5]。近红外光谱范围包括 780 ~ 2 500 nm (12 820 ~ 4 000 cm^{-1}) 区域内的电磁辐射吸收。目前,我国已有部分关于近红外与拉曼技术对甲/乙醇汽油中甲/乙醇含量检测^[6-9]。

本文,应用近红外光谱技术比较了线性(PCR、PLS)和非线性(LSSVM)校正技术对乙醇柴油主要性能指标预测。该模型检测乙醇柴油的三个重要性能指标分别为:密度(20 °C)、粘度(20 °C)、乙醇含量。我们还讨论了不同预处理方法(平滑,多元散射校正-平滑、标准正则变量变换-平滑、一阶微分-平滑、二阶微分-平滑等)对模型预测能力的影响,并将最小二乘支持向量机用于乙醇柴油主要性能指标的光谱检测。

2 实验与材料

2.1 样品制备

96 个不同浓度的乙醇柴油均在实验室配制,所用 0# 柴油购于南昌某加油站,乙醇由天津市大茂化学试剂厂生产,分析纯度 $\geq 99.7\%$,正丁醇来自西陇化工有限公司,分析纯度 $\geq 99.5\%$,电子天平和 50 mL 容量瓶购置于江西赣仪科技有限公司。配制乙醇柴油时,室温控制在 20 °C 左右,柴油、乙醇、正丁醇以 43:5:2 比例加入 50 mL 容量瓶内。实验用 50 mL 容量瓶用纯水清洗后烘干,记录电子天平测量空容量瓶质量,先通过移液枪将乙醇和正丁醇按 5:2 比例加入 50 mL 容量瓶中并充分振荡,形成变性乙醇,完成后加入柴油并滴定至 50 mL,振荡均匀,直至得到淡黄色透明液体,测量此时的重量,贴上带有浓度的标签静置一段

表 1 32 种浓度乙醇柴油

Tab. 1 Ethanol diesel with 32 kinds of concentration (% , v/v)

序号	浓度/%	序号	浓度/%	序号	浓度/%	序号	浓度/%
1	0.6	9	5.6	17	10.6	25	15.6
2	1.2	10	6.2	18	11.2	26	16.2
3	1.8	11	6.8	19	11.8	27	16.8
4	2.4	12	7.4	20	12.4	28	17.4
5	3.0	13	8.0	21	13.0	29	18.0
6	3.6	14	8.6	22	13.6	30	18.6
7	4.2	15	9.2	23	14.2	31	19.2
8	5.0	16	10.0	24	15.0	32	20.0

时间。同样的方法配制余下浓度的乙醇柴油,32种乙醇柴油浓度如表1所示。

2.2 乙醇柴油密度、粘度、乙醇含量测量

乙醇柴油的密度通过传统的方法测量,由配制前、后50 mL容量瓶的质量差和50 mL体积计算得出;乙醇柴油的粘度由上海精天仪器有限公司生产的NDJ-5S旋转粘度计测得,其原理是当转子在乙醇柴油中旋转时,由于乙醇柴油的粘滞性,转子就受到一个与粘度成正比的扭力,通过扭矩传感器测量这个扭力的大小,就可得到乙醇柴油的粘度。测量前先校准,以空气为参比,按下测量键,数字显示为零后按复位键。将乙醇柴油倒入直径不小于70 mm的烧杯中,准确地控制被测液体的温度。查阅乙醇及柴油的粘度选择相应规格的转子旋入连接螺杆。装卸转子时,避免碰坏转子。旋转升降旋钮,使仪器缓慢地下降,转子逐渐浸入乙醇柴油中,直至转子液面标志和液面相平为止,再次调整好仪器水平试样在测试温度下充分恒温(20℃左右),以保持示值稳定准确。当按下测量键后等待,当听到六声滴滴声后仪器屏幕显示的数值作为样本的粘度,重复测量五次取平均值,每种浓度的乙醇柴油都依据上述方法测量其密度和粘度。

2.3 乙醇柴油近红外光谱表征

乙醇柴油的近红外光谱由布鲁克公司生产的Tensor37型傅里叶变换红外光谱仪采集获得,采集前光谱仪先开机预热20 min左右,实验的环境温度调到25℃左右,湿度调到50%以下。先用胶头滴管吸取乙醇柴油慢慢滴入石英比色皿(光程5 mm)中,避免产生气泡,再将比色皿放入载物台上,关闭仪器的外盖。实验时,软件中光谱采集范围直接设为系统默认的 $12\ 500 \sim 4\ 000\ \text{cm}^{-1}$,扫描次数为32次,分辨率为 $8\ \text{cm}^{-1}$ 。每过45 min左右以空的比色皿采集背景通道光谱,为避免乙醇和柴油混合不匀或采集位置不同导致采集光谱的差异性,通过抬升载物台的高度使入射光通过比色皿上、中、下3个不同位置进行光谱采集,计算平均光谱作为样品的最终光谱。所有程序均在MATLABR2014a(The Math Work, Inc, USA)中运行。

2.4 乙醇柴油近红外光谱

乙醇柴油的近红外光谱图如图1所示,从图中可以看出在1 400 ~ 1 650 nm和2 000 ~ 2 150 nm两个光谱区域内,光谱谱带的吸光度与乙醇柴油浓度有关系。吸收特征峰大约在1 580 nm和2 100 nm处,这是由于乙醇分子中O—H伸缩振动的一级倍频在1 587 nm处,其组合频吸收谱带在2 173 nm附近,在1 667 ~ 2 000 nm范围内光谱数据与乙醇含量的线性相关性较差,这是由于柴油中烃类物质的化学键吸收峰重叠所致。

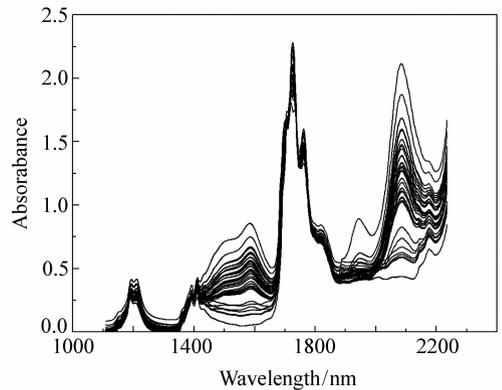


图1 乙醇柴油近红外光谱

Fig. 1 Near infrared spectra of ethanol diesel

2.5 光谱预处理

预处理均在Unscrambler(verion. 10.1; CAMO AS, Trondheim, Norway)化学计量学分析软件中进行,简单地说,在校正模型建立之前,常使用预处理技术对原始光谱进行处理。本文用到的预处理方法有以下5种:(1)SG:Savitzky-Golay;(2)MSC-SG:Mean Scattering correction followed by Savitzky-Golay;(3)SNV-SG:Standard Normal Variate scaling plus Savitzky-Golay;(4)SGD1-SG:First-order Savitzky-Golay Derivative followed by Savitzky-Golay;(5)SGD2-SG:Second-order Savitzky-Golay Derivative followed by Savitzky-Golay。

2.6 数据分析

乙醇柴油原始光谱进行数字化后,每条光谱为 1×2203 的向量,向量的长度是由光谱仪分辨率决定。所有程序均在MATLABR2014a(The MathWork, Inc, USA)中运行。本文比较了几种不

同的多元回归方法,从中找到最适合检测乙醇柴油指标的模型。由于校正模型好坏依赖其参数的选择,本文所使用模型的参数有:主成分回归(PCR):主成分(PC = 1 ~ 20);偏最小二乘回归(PLS):主因子数;最小二乘支持向量机(LSSVM):选择不同的核函数其参数也会有变化,视具体情况而定。

3 结果与讨论

3.1 线性校正模型:PCR 和 PLS

讨论了两种有效的多元线性方法对乙醇柴油

物理参数进行预测:PCR 和 PLS。相对于 MLR 来说,PCR 是一种更先进的光谱数据分析方法。在该方法中,通过减少变量的数目(从几百到几十)和去除噪声,以达到更好的预测精度。表 2 总结了应用 PCR 处理乙醇柴油的近红外光谱数据的结果。在 MSC-SG 预处理下,PCR 对乙醇柴油密度、粘度、乙醇含量的建模效果最好,而相对于粘度、乙醇含量这两个指标,PCR 对密度的建模效果最好,此时,模型的主成分数 PC 为 16,相关系数 R^2 为 0.909 4,预测均方根误差 RMSEP 为 0.001 7。

表 2 PCR 方法对乙醇柴油密度、粘度、乙醇含量的性能预测

Tab.2 Prediction results of principal component regression(PCR) approach for ethanol diesel properties including:density, viscosity, ethanol content

乙醇柴油性质	最优预处理方法	主成分 PC	R^2	RMSEP
密度/($g \cdot cm^{-3}$)	MSC-SG	16	0.909 4	0.001 7
粘度/($mPa \cdot s$)	MSC-SG	16	0.902 4	0.038 6
乙醇含量(体积比)(%)	MSC-SG	14	0.883 0	1.968 7

PLS 是一种与 PCR 相近的统计方法^[10],主要区别是,输入值 x 和输出数据 y 被投影到新的空间。在分析乙醇柴油时, x 和 y 分别是乙醇柴油的近红外光谱和物理参数(包括密度、粘度和乙醇含量)。PLS 操作简单和较低的计算要求,它被广泛用在非常不同的数据集的分析,是一种最受欢迎的用于建立校正模型化学计量学算法。此外,PLS 可以被视为一种标准的近红外光谱数据处理方法,是一种线性的数据分析方法。表 3 总

结了应用 PLS 的到乙醇柴油的近红外光谱数据的结果。PLS 预测乙醇柴油密度、粘度、乙醇含量在 SG 预处理下建模效果最好。相比 PCR 的结果,PLS 结果有明显提高,幅度提高最大的是乙醇含量的预测,其 R^2 提高了 7.7%,RMSEP 下降了 54.5%。PLS 是一种检测乙醇柴油性能的线性模型的选择方法,较少的输入值数量使得 PLS 模型比 PCR 更稳定,可以通过 SG 预处理技术结合 PLS 校正模型获得最佳的预测性能。

表 3 PLS 方法对乙醇柴油密度、粘度、乙醇含量的性能预测

Tab.3 Prediction results of partial least squares(PLS) regression approach for ethanol diesel properties including:density, viscosity, ethanol content

乙醇柴油性质	最优预处理方法	主因子数	R^2	RMSEP
密度/($g \cdot cm^{-3}$)	SG	8	0.955 3	0.001 5
粘度/($mPa \cdot s$)	SG	8	0.951 5	0.027 2
乙醇含量(体积比)(%)	SG	8	0.951 0	1.273 7

3.2 非线性校正模型:最小二乘支持向量机

支持向量机是机器学习领域的一种监督学习技术,适用于分类和回归分析^[11-18]。支持向量机

的主要缺点之一是必须解决大规模二次规划问题。最小二乘支持向量对传统的 SVM 做了修改,克服了这个缺点,将二次规划问题转化为成线性

问题,以减少优化过程的复杂性。表4总结了将LSSVM应用到乙醇柴油近红外光谱数据的结果。乙醇柴油的密度在SGD1-SG预处理下建立的LSSVM效果最好,此时, $\gamma = 131\ 969$ 、 $\text{sig}2 = 2\ 047$, R^2 为0.995 8, RMSEP为0.000 68。粘度和

乙醇含量分别在SNV-SG、SG预处理下建立的LSSVM模型效果最好,而乙醇柴油的密度和粘度结果较为接近,主要是因为密度与粘度之间有一定的联系,乙醇和柴油的比例影响乙醇柴油密度和粘度。

表4 LSSVM对乙醇柴油密度、粘度、乙醇含量的性能预测

Tab.4 Prediction results of least squares support vector machine (LSSVM) regression approach for ethanol diesel properties including: density, viscosity, ethanol content

乙醇柴油性质	最优预处理方法	[γ , sig2]	R^2	RMSEP
密度/($\text{g} \cdot \text{cm}^{-3}$)	SGD1-SG	131 969, 2 047	0.995 8	0.000 68
粘度/($\text{mPa} \cdot \text{s}$)	SNV-SG	111 607, 5 690	0.995 7	0.011 3
乙醇含量(体积比)(%)	SG	28 707, 5 963	0.995 3	0.571 4

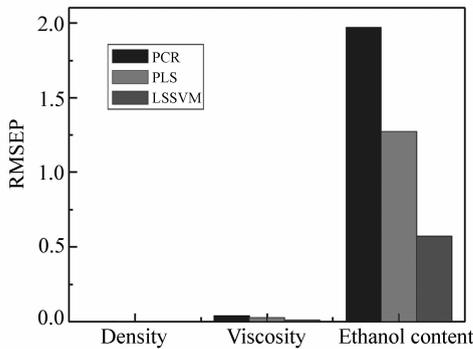


图2 乙醇柴油密度、粘度和乙醇含量在PCR, PLS, LSSVM建模下的RMSEP

Fig.2 RMSEP of ethanol diesel oil density, viscosity and ethanol content under PCR, PLS and LSSVM modeling RMSEP

图2是乙醇柴油密度、粘度和乙醇含量在PCR, PLS和LSSVM建模下的RMSEP变化, LSSVM在检测乙醇柴油密度、粘度和乙醇含量得到的RMSEP是3种模型中最小的,分别是0.000 68、0.011 3和0.571 4,且检测乙醇含量的RMSEP下降最明显,降幅达到54.5%。PCR, PLS和LSSVM建模效果顺序为: PCR < PLS < LSSVM,建模消耗时间为PCR \approx PLS < LSSVM,操作简易性为PCR \approx PLS < LSSVM。LSSVM模型参数的优化比较复杂,消耗时间较长,而PCR和PLS均借助于Unscrambler软件自带的算法建模,操作简单,消耗时间较少^[16]。以最小二乘支持向量机方法建立的乙醇柴油密度、粘度及乙醇含量的模型预测结果如图3(a)、3(b)、3(c)所示。

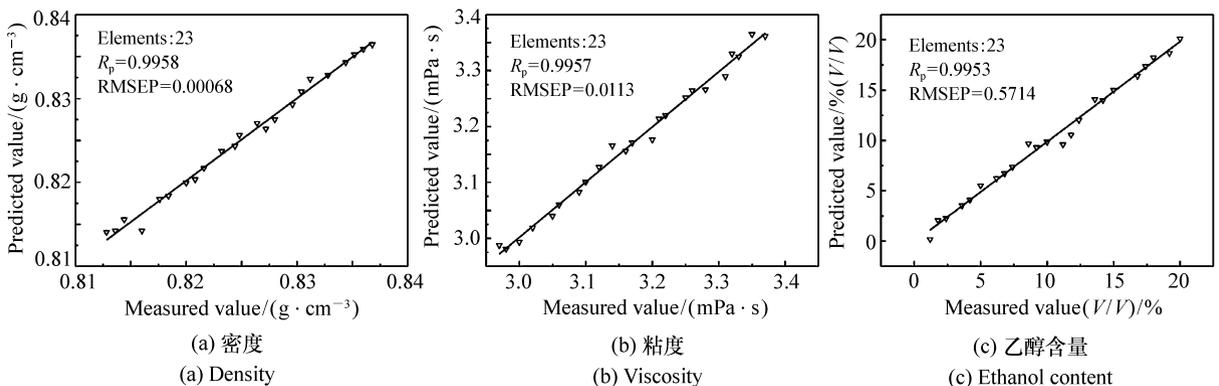


图3 最小二乘支持向量机最佳模型预测结果

Fig.3 Predicted results of optimized LSSVM model

4 结 论

本文建立了基于近红外光谱数据的乙醇柴油性能预测模型。最小二乘支持向量机优于 PCR 和 PLS,其对乙醇柴油密度、粘度、乙醇含量建模最好,相关系数分别是 0.995 8、0.995 7 和

0.995 3; RMSEP 分别为 0.000 68、0.011 3 和 0.571 4。本文配制的乙醇柴油浓度涵盖了目前市面上销售的乙醇柴油,能够满足乙醇柴油几个主要指标的检测,近红外光谱检测乙醇柴油所得到的结果可以拓展到其他燃料如生物柴油性能快速、准确的分析。

参考文献:

- [1] BAPTISTA P, FELIZARDO P, MENEZES J C, *et al.*. Multivariate near infrared spectroscopy models for predicting the methyl esters content in biodiesel[J]. *Anal. Chim. Acta*, 2008, 607(2): 153-159.
- [2] 张明祥, 闵顺耕, 李宁, 等. 乙醇混合燃料近红外定量分析[J]. 分析测试学报, 2003, 6: 15-18.
ZHANG M X, MIN S G, LI N, *et al.*. Ethanol blends near-infrared quantitative analysis[J]. *Instrumental Analysis*, 2003, 6: 15-18. (in Chinese)
- [3] 蔡剑华, 王先春, 胡惟文. 基于形态小波的烟草尼古丁含量近红外光谱检测[J]. 农业工程学报, 2012(15): 281-286.
CAI J H, W X CH, HU W W. Near Infrared Spectroscopy based on morphological wavelet nicotine content of tobacco[J]. *J. Agricultural Engineering*, 2012(15): 281-286. (in Chinese)
- [4] 孙通, 应义斌, 刘魁武, 等. 梨可溶性固形物含量的在线近红外光谱检测[J]. 光谱学与光谱分析, 2008(11): 2536-2539.
SUN T, YING Y B, LIU K W, *et al.*. Pear-line Near Infrared Spectroscopy of soluble solids[J]. *Spectroscopy and Spectral Analysis*, 2008(11): 2536-2539. (in Chinese)
- [5] 孙旭东, 郝勇, 高荣杰, 等. 脐橙糖度近红外光谱在线检测数学模型优化研究[J]. 光谱学与光谱分析, 2011, 31(5): 1230-1235.
SUN X D, HAO Y, GAO R J. Orange sugar content line Near Infrared spectroscopy detection mathematical model optimization[J]. *Spectroscopy and Spectral Analysis*, 2011, 31(5): 1230-1235. (in Chinese)
- [6] 史永刚, 粟斌, 李华峰, 等. 拉曼光谱及其在石油产品分析中的应用[J]. 现代仪器, 2010, 16(9106): 9-13.
SHI Y G, SU B, LI H F, *et al.*. Merlin Raman spectroscopy and its application in petroleum product analysis[J]. *Modern Instruments*, 2010, 16(9106): 9-13. (in Chinese)
- [7] 李津蓉, 俞柯. 一种基于拉曼光谱技术的甲醇汽油定量分析方法[J]. 科技创新与应用, 2016(15305): 59.
LI J R, YU K. Methanol gasoline quantitative analysis method based on Raman spectroscopy[J]. *Science and Technology Innovation and Application*, 2016(15305): 59. (in Chinese)
- [8] 欧阳爱国, 刘军. 基于近红外与拉曼光谱的甲醇汽油定量分析[J]. 西南师范大学学报(自然科学版), 2012, 37(18609): 98-101.
OUYANG A G, LIU J. Methanol-based gasoline Quantitative Near Infrared and Raman Spectra of[J]. *Southwest China Normal University(Natural Science)*, 2012, 37(18609): 98-101. (in Chinese)
- [9] 淡图南, 戴连奎. 近红外光谱的甲醇汽油定量分析[J]. 计算机与应用化学, 2011, 3: 329-332.
DAN T N, DAI L K. Quantitative methanol gasoline near infrared spectroscopy[J]. *Computer and Applied Chemistry*, 2011, 3: 329-332. (in Chinese)
- [10] 姚捷, 戴连奎, 林艺玲. 基于拉曼特征峰的甲醇汽油甲醇含量测定[J]. 光散射学报, 2013, 1: 59-65.
YAO J, DAI L K, LIN Y L. Determination of methanol content of methanol-based gasoline Raman characteristic peak[J]. *J. of Light Scattering*, 2013, 1: 59-65. (in Chinese)
- [11] BALABIN R M, SAFIEVA R Z, LOMAKINA E I. *Chemometr Intell Lab*, 2007: 88: 183.
- [12] 刘燕德, 周延睿, 潘圆媛. 基于最小二乘支持向量机的辣椒可溶性固形物和维生素 C 含量近红外光谱检测[J]. 光学精密工程, 2014(2): 281-288.
LIU Y D, ZHOU Y R, PAN Y Y. Based on least squares support chili soluble solids and vitamin C content of vector machine near infrared spectroscopy[J]. *Opt. Precision Eng.*, 2014(2): 281-288. (in Chinese)

- [13] 刘燕德,周延睿. 基于 GA-LSSVM 的苹果糖度近红外光谱检测[J]. 西北农林科技大学学报(自然科学版),2013,7:229-234.
LIU Y D,ZHOU Y R. Near infrared spectroscopy based GA-LSSVM apple brix[J]. *J. Northwest Sci-Tech University of Agriculture and Forestry*(Natural Science Edition),2013,7:229-234. (in Chinese)
- [14] 彭秀辉,黄常毅,刘飞,等. LSSVM 模型下的近红外光谱联合区间波长筛选方法[J]. 光谱学与光谱分析,2014,3:668-672.
PENG X H,HUANG C Y,LIU F,*et al.*. Near infrared spectroscopy synergy interval wavelength selection method using the LSSVM model[J]. *Spectroscopy and Spectral Analysis*,2014,3:668-672. (in Chinese)
- [15] 伍铁斌,朱红求,孙备,等. PLS-LSSVM 模型在锌净化中的应用[J]. 计算机工程,2012,10:212-214.
WU T B,ZHU H Q,SUN B,*et al.*. Application of PLS-LSSVM model in zinc purification[J]. *Computer Engineering*,2012,10:212-214. (in Chinese)
- [16] 刘燕德,邓清. 基于高光谱成像技术的脐橙叶片的叶绿素含量及其分布测量[J]. 发光学报,2015,8:957-961.
LIU Y D,DENG Q. Measurement of chlorophyll distribution in navel orange leaves based on hyper-spectral imaging technique[J]. *J. Luminescence*,2015,8:957-961. (in Chinese)
- [17] 鲍一丹,陈纳,何勇,等. 近红外高光谱成像技术快速鉴别国产咖啡豆品种[J]. 光学精密工程,2015,2:349-355.
BAO Y D,CHEN N,HE Y,*et al.*. Near infrared hyper-spectral imaging technology for rapid identification of domestic coffee bean varieties[J]. *Opt. Precision Eng.*,2015,2:349-355. (in Chinese)
- [18] 郭文川,董金磊. 高光谱成像结合人工神经网络无损检测桃的硬度[J]. 光学精密工程,2015,6:1530-1537.
GUO W CH,DONG J L. High spectral imaging combined with artificial neural network for nondestructive detection of peach hardness[J]. *Opt. Precision Eng.*,2015,6:1530-1537. (in Chinese)
- [19] YANG H,GRIFFITHS P R,TATE J D. Comparison of partial least squares regression and multi-layer neural networks for quantification of nonlinear systems and application to gas phase Fourier transform infrared spectra[J]. *Anal Chim Acta*,2003,489:125-136.

作者简介:



欧阳爱国(1968—),男,江西常德人,学士,教授,主要从事车辆性能检测方面的研究。E-mail:ouyang-
gaiguo1968711@163.com